

the following is a

combination of all first pages of most of the papers
that i was reading in 2017

ok thanks

kanzure@gmail.com

2017-10-22

h/t nsh

<https://twitter.com/kanzure/status/922164370288898048>

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/259566905>

Where is mTOR and what is it doing there?

Article in *The Journal of Cell Biology* · November 2013

DOI: 10.1083/jcb.201306041 · Source: PubMed

CITATIONS

151

READS

534

2 authors, including:



Charles Betz

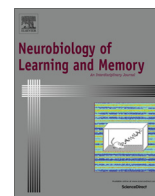
University of Basel

9 PUBLICATIONS 569 CITATIONS

SEE PROFILE

All content following this page was uploaded by [Charles Betz](#) on 19 December 2013.

The user has requested enhancement of the downloaded file.



Review

Prefrontal–hippocampal pathways underlying inhibitory control over memory

Michael C. Anderson^{a,*}, Jamie G. Bunce^b, Helen Barbas^b^a MRC Cognition & Brain Sciences Unit, 15 Chaucer Road, Cambridge, England CB2 7EF, United Kingdom^b Neural Systems Laboratory, Boston University, 635 Commonwealth Ave., Boston, MA 02215, USA

ARTICLE INFO

Article history:

Received 17 August 2015

Revised 6 November 2015

Accepted 17 November 2015

Available online 28 November 2015

Keywords:

Retrieval suppression

Inhibitory control

Forgetting

Hippocampus

Anterior cingulate

Nucleus reuniens

ABSTRACT

A key function of the prefrontal cortex is to support inhibitory control over behavior. It is widely believed that this function extends to stopping cognitive processes as well. Consistent with this, mounting evidence establishes the role of the right lateral prefrontal cortex in a clear case of cognitive control: retrieval suppression. Retrieval suppression refers to the ability to intentionally stop the retrieval process that arises when a reminder to a memory appears. Functional imaging data indicate that retrieval suppression involves top-down modulation of hippocampal activity by the dorsolateral prefrontal cortex, but the anatomical pathways supporting this inhibitory modulation remain unclear. Here we bridge this gap by integrating key findings about retrieval suppression observed through functional imaging with a detailed consideration of relevant anatomical pathways observed in non-human primates. Focusing selectively on the potential role of the anterior cingulate cortex, we develop two hypotheses about the pathways mediating interactions between lateral prefrontal cortex and the medial temporal lobes during suppression, and their cellular targets: the entorhinal gating hypothesis, and thalamo-hippocampal modulation via the nucleus reuniens. We hypothesize that whereas entorhinal gating is well situated to stop retrieval proactively, thalamo-hippocampal modulation may interrupt an ongoing act of retrieval reactively. Isolating the pathways that underlie retrieval suppression holds the potential to advance our understanding of a range of psychiatric disorders characterized by persistent intrusive thoughts. More broadly, an anatomical account of retrieval suppression would provide a key model system for understanding inhibitory control over cognition.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Memories, like physical actions, sometimes need to be controlled. For example, although good memory for the past typically is welcomed, this feature poses a problem when memories are unpleasant and intrusive. When people encounter an unwelcome reminder, they strive to limit awareness of the unwanted memory by stopping its retrieval. This retrieval stopping process, known as retrieval suppression, is mediated by an inhibitory control mechanism that suppresses unwanted traces, rendering them less likely to be retrieved in the future (Anderson & Green, 2001; see Anderson & Hanslmayr, 2014; Anderson & Huddleston, 2011 for reviews). Over the last decade, evidence has grown showing that the brain systems underlying retrieval suppression exhibit important similarities and differences to other putative forms of inhibitory control, such as motor response stopping. Like motor

stopping, retrieval suppression engages the right lateral prefrontal cortex; but, instead of modulating motor cortical regions, the prefrontal cortex suppresses hippocampal activity that supports retrieval (Anderson et al., 2004; Benoit & Anderson, 2012; Depue, Curran, & Banich, 2007; Depue, Orr, Smolker, Naaz, & Banich, 2015; Gagnepain, Henson, & Anderson, 2014; Levy & Anderson, 2012; Paz-Alonso, Bunge, Anderson, & Ghetti, 2013). These findings suggest that mnemonic functions of the hippocampus are subject to inhibitory control by the prefrontal cortex. If so, retrieval suppression may provide an important model system for studying inhibitory control over thought that complements and generalizes models of inhibitory control based on stopping action.

Whereas the anatomical pathways underlying action stopping are increasingly well characterized (e.g., see, e.g., Schmidt, Leventhal, Mallet, Chen, & Berke, 2013; for a review, see Aron, Robbins, & Poldrack, 2014), little is known about how the lateral prefrontal cortex modulates hippocampal activity to suppress retrieval. In this article, we begin to close this gap. In particular, we review anatomical findings observed with non-human

* Corresponding author.

E-mail address: michael.anderson@mrc-cbu.cam.ac.uk (M.C. Anderson).

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/2595891>

Two Types Of Mechanical Reversible Logic

Article · February 1970

Source: CiteSeer

CITATIONS

31

READS

176

1 author:



[Ralph C. Merkle](#)

Institute for Molecular Manufacturing

73 PUBLICATIONS 6,285 CITATIONS

SEE PROFILE

BC^2

Jeremy Rubin

Jerome Siegel

Brain mechanisms that control sleep and waking

Published online: 2 July 2004
© Springer-Verlag 2004

Abstract This review paper presents a brief historical survey of the technological and early research that laid the groundwork for recent advances in sleep–waking research. A major advance in this field occurred shortly after the end of World War II with the discovery of the ascending reticular activating system (ARAS) as the neural source in the brain stem of the waking state. Subsequent research showed that the brain stem activating system produced cortical arousal via two pathways: a dorsal route through the thalamus and a ventral route through the hypothalamus and basal forebrain. The nuclei, pathways, and neurotransmitters that comprise the multiple components of these arousal systems are described. Sleep is now recognized as being composed of two very different states: rapid eye movements (REMs) sleep and non-REM sleep. The major findings on the neural mechanisms that control these two sleep states are presented. This review ends with a discussion of two current views on the function of sleep: to maintain the integrity of the immune system and to enhance memory consolidation.

Historical background

Three advances in technology and a clinical research finding during the first half of the twentieth century were critical for laying a foundation for the scientific advances that were generated in the latter half of the twentieth century.

This article is based in part on material in the book “The neural control of sleep and waking” (J. Siegel, 2002).

J. Siegel (✉)
University of Delaware,
Newark, DE 19716, USA
e-mail: jsiegel@udel.edu

J. Siegel
Warsaw School of Social Psychology,
Chodakowska 19/31, 03-815 Warsaw, Poland

Electroencephalography

The single most important tool for the scientific study of sleep and waking has been the electroencephalograph. This instrument records electrical activity of the brain and provides an objective measure of the state of the brain and the states of consciousness. Hans Berger in Jena, Germany, is credited with developing the first instrument that recorded electrical activity generated by the human brain. Berger, in 1925, had access to one of the first vacuum tube amplifiers manufactured by the Siemens Company and used this sensitive galvanometer to record the electrical activity from his young son’s brain. This landmark electroencephalographic (EEG) recording was published by Berger in 1929 (Berger 1929). This important advance was built upon earlier research with animals that demonstrated the electrical nature of neural activity (Brazier 1961).

Berger’s early work showed a clear EEG difference between sleep and waking. Research conducted 24 years later (Aserinsky and Kleitman 1953), showed that the EEG during sleep could be differentiated into at least two categories. One type of sleep was found to be associated with the occurrence of dreams and the other with non-dream sleep. Since dream sleep was found to be accompanied by episodes of rapid eye movement (REM), this state is often called REM sleep. Non-dream sleep is also called non-REM sleep. Figure 1a shows examples of EEG tracings during waking and during different stages of sleep.

An awake record is characterized by EEG waves that are low voltage (5–50 μ V) and high frequency (20–40 Hz) relative to voltages and frequencies seen during non-REM sleep. When a person falls asleep, a brief transitional state between waking and sleep is characterized by an EEG that resembles waking. During this period, recordings from electrodes placed on the skin around the eyes show the presence of sporadic slow eye movements. This transitional state is referred to as stage 1 sleep. As the individual progresses toward deeper sleep, stages 2, 3, and 4, the EEG shows higher amplitudes and lower frequencies.

A theory of working memory

without consciousness or sustained activity

Darinka Trübutschek^{1, 2, 3, *}, Sébastien Marti³, Andrés Ojeda⁴, Jean-Rémi King^{5, 6}, Yuanyuan Mi⁷, Misha Tsodyks^{8, 9}, and Stanislas Dehaene^{3, 10}

1 - Ecole des Neurosciences de Paris Ile-de-France, 15 rue de l'école de médecine, 75006 Paris, France

2 - Ecole Doctorale Cerveau-Cognition-Comportement, Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France

3 - Cognitive Neuroimaging Unit, CEA DSV/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin center, 91191 Gif/Yvette, France

4 - Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom

5 - Department of Psychology, New York University, 4 Washington Place, New York, USA

6 - Frankfurt Institute for Advanced Studies, Frankfurt, Germany

7 - Brain Science Center, Institute of Basic Medical Sciences, 100850 Beijing, China

8 - Department of Neurobiology, Weizmann Institute of Science, 76100 Rehovot, Israel

9 - Department of Neuroscience, Columbia University, 10032 New York, USA

10 - Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France

*Correspondence: darinkat87@gmail.com

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7169144>

Gene expression correlates of unexplained fatigue

Article in *Pharmacogenomics* · May 2006

DOI: 10.2217/14622416.7.3.395 · Source: PubMed

CITATIONS

32

READS

74

6 authors, including:



Toni Whistler

Centers for Disease Control and Prevention

84 PUBLICATIONS 972 CITATIONS

SEE PROFILE



Cameron Craddock

Child Mind Institute

106 PUBLICATIONS 3,811 CITATIONS

SEE PROFILE



Gordon Broderick

Rochester General Hospital

112 PUBLICATIONS 1,692 CITATIONS

SEE PROFILE



Nancy G Klimas

Nova Southeastern University

254 PUBLICATIONS 10,252 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Testing the Model: A Phase I/II Randomized Double Blind Placebo Control Trial of Therapeutics: Liposomal Glutathione and Curcumin [View project](#)



Molecular Patterns of Persistent Immune Activation in a Post-infectious Adolescent Cohort [View project](#)

All content following this page was uploaded by [Toni Whistler](#) on 05 June 2014.

The user has requested enhancement of the downloaded file.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

**THE BROAD INSTITUTE, INC., MASSACHUSETTS INSTITUTE
OF TECHNOLOGY, and PRESIDENT AND
FELLOWS OF HARVARD COLLEGE,**
(Patents 8,697,359; 8,771,945; 8,795,965; 8,865,406; 8,871,445; 8,889,356;
8,895,308; 8,906,616; 8,932,814; 8,945,839; 8,993,233; 8,999,641
and Application 14/704,551),

Junior Party,

v.

**THE REGENTS OF THE UNIVERSITY OF CALIFORNIA,
UNIVERSITY OF VIENNA, and EMMANUELLE CHARPENTIER**
(Application 13/842,859),

Senior Party.

Patent Interference No. 106,048 (DK)

DECISION ON MOTIONS

37 C.F.R. § 41.125(a)

Before RICHARD E. SCHAFER, SALLY GARDNER LANE, and
DEBORAH KATZ, *Administrative Patent Judges.*

Per curiam.

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

**THE BROAD INSTITUTE, INC., MASSACHUSETTS INSTITUTE
OF TECHNOLOGY, and PRESIDENT AND FELLOWS
OF HARVARD COLLEGE,**
(Patents 8,697,359; 8,771,945; 8,795,965; 8,865,406; 8,871,445; 8,889,356;
8,895,308; 8,906,616; 8,932,814; 8,945,839; 8,993,233; 8,999,641
and Application 14/704,551),

Junior Party,

v.

**THE REGENTS OF THE UNIVERSITY OF CALIFORNIA, UNIVERSITY
OF VIENNA, and EMMANUELLE CHARPENTIER**
(Application 13/842,859),

Senior Party.

Patent Interference No. 106,048 (DK)

JUDGMENT
37 C.F.R. § 41.127(a)

Before RICHARD E. SCHAFER, SALLY GARDNER LANE, and
DEBORAH KATZ, *Administrative Patent Judges.*

Per curiam.

SpeechJammer: A System Utilizing Artificial Speech Disturbance with Delayed Auditory Feedback

Kazutaka Kurihara

National Institute of Advanced
Industrial Science and Technology
Chuo Dai 2, 1-1-1 Umezono,
Tsukuba, Ibaraki, Japan
k-kurihara@aist.go.jp

Koji Tsukada

Ochanomizu University / JST
PRESTO
2-1-1 Otsuka Bunkyo-ku
Tokyo, Japan
tsuka@acm.com

Abstract

In this paper we report on a system, "SpeechJammer", which can be used to disturb people's speech. In general, human speech is jammed by giving back to the speakers their own utterances at a delay of a few hundred milliseconds. This effect can disturb people without any physical discomfort, and disappears immediately by stop speaking. Furthermore, this effect does not involve anyone but the speaker. We utilize this phenomenon and implemented two prototype versions by combining a direction-sensitive microphone and a direction-sensitive speaker, enabling the speech of a specific person to be disturbed. We discuss practical application scenarios of the system, such as facilitating and controlling discussions. Finally, we argue what system parameters should be examined in detail in future formal studies based on the lessons learned from our preliminary study.

Keywords

DAF, artificial speech jamming

ACM Classification Keywords

H.5.m. [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage

Alexis C. Komor^{1,2}, Yongjoo B. Kim^{1,2}, Michael S. Packer^{1,2}, John A. Zuris^{1,2} & David R. Liu^{1,2}

Current genome-editing technologies introduce double-stranded (ds) DNA breaks at a target locus as the first step to gene correction^{1,2}. Although most genetic diseases arise from point mutations, current approaches to point mutation correction are inefficient and typically induce an abundance of random insertions and deletions (indels) at the target locus resulting from the cellular response to dsDNA breaks^{1,2}. Here we report the development of ‘base editing’, a new approach to genome editing that enables the direct, irreversible conversion of one target DNA base into another in a programmable manner, without requiring dsDNA backbone cleavage or a donor template. We engineered fusions of CRISPR/Cas9 and a cytidine deaminase enzyme that retain the ability to be programmed with a guide RNA, do not induce dsDNA breaks, and mediate the direct conversion of cytidine to uridine, thereby effecting a C→T (or G→A) substitution. The resulting ‘base editors’ convert cytidines within a window of approximately five nucleotides, and can efficiently correct a variety of point mutations relevant to human disease. In four transformed human and murine cell lines, second- and third-generation base editors that fuse uracil glycosylase inhibitor, and that use a Cas9 nickase targeting the non-edited strand, manipulate the cellular DNA repair response to favour desired base-editing outcomes, resulting in permanent correction of ~15–75% of total cellular DNA with minimal (typically ≤1%) indel formation. Base editing expands the scope and efficiency of genome editing of point mutations.

The clustered regularly interspaced short palindromic repeat (CRISPR) system has been widely used to mediate genome editing in a variety of organisms and cell lines³. CRISPR/Cas9 protein–RNA complexes localize to a target DNA sequence through base pairing with a guide RNA, and natively create a dsDNA break (DSB) at the locus specified by the guide RNA. In response to DSBs, cellular DNA repair processes mostly result in random insertions or deletions (indels) at the site of DNA cleavage through non-homologous end joining (NHEJ). In the presence of a homologous DNA template, the DNA surrounding the cleavage site can be replaced through homology-directed repair (HDR). HDR competes with NHEJ during the resolution of DSBs, and indels are generally more abundant outcomes than gene replacement. For most known genetic diseases, however, correction of a point mutation in the target locus, rather than stochastic disruption of the gene, is needed to study or address the underlying cause of the disease⁴.

Motivated by this need, researchers have sought to increase the efficiency of HDR and suppress NHEJ. Despite recent progress (see Supplementary Information for a detailed discussion), current strategies to correct point mutations using HDR under therapeutically relevant conditions remain inefficient (typically ~0.1 to 5%)^{5,6}, especially in unmodified, non-dividing cells. These observations highlight the need to develop alternative approaches to correct point mutations in genomic DNA that do not require DSBs.

We envisioned that direct conversion of one DNA base to another at a programmable target locus without requiring DSBs could increase

the efficiency of gene correction relative to HDR without introducing an excess of random indels. Catalytically dead Cas9 (dCas9), which contains Asp10Ala and His840Ala mutations that inactivate its nuclease activity, retains its ability to bind DNA in a guide RNA-programmed manner, but does not cleave the DNA backbone⁷. In principle, conjugation of dCas9 with an enzymatic or chemical catalyst that mediates the direct conversion of one base to another could enable RNA-programmed DNA base editing.

The deamination of cytosine (C) is catalysed by cytidine deaminases⁸ and results in uracil (U), which has the base-pairing properties of thymine (T). Most known cytidine deaminases operate on RNA, and the few examples that are known to accept DNA require single-stranded (ss) DNA⁹. Recent studies on the dCas9–target DNA complex reveal that at least nine nucleotides (nt) of the displaced DNA strand are unpaired upon formation of the Cas9–guide RNA–DNA ‘R-loop’ complex¹⁰. Indeed, in the structure of the Cas9 R-loop complex, the first 11 nt of the protospacer on the displaced DNA strand are disordered, suggesting that their movement is not highly restricted¹¹. It has also been speculated that Cas9 nickase-induced mutations at cytosines in the non-template strand might arise from their accessibility by cellular cytosine deaminase enzymes¹². We reasoned that a subset of this stretch of ssDNA in the R-loop might serve as an efficient substrate for a dCas9–tethered cytidine deaminase to effect direct, programmable conversion of C to U in DNA (Fig. 1a).

Four different cytidine deaminase enzymes (human AID, human APOBEC3G, rat APOBEC1, and lamprey CDA1) were evaluated for ssDNA deamination. Of the four enzymes, rat APOBEC1 showed the highest deaminase activity under the conditions tested (Extended Data Fig. 1a). Fusing rat APOBEC1 to the amino terminus, but not the carboxy terminus, of dCas9 preserves deaminase activity (Extended Data Fig. 1a). We expressed and purified four rat APOBEC1–dCas9 fusions with linkers of different length and composition (Extended Data Fig. 1b), and evaluated each fusion for single guide RNA (sgRNA)–programmed dsDNA deamination *in vitro* (Fig. 1b and Extended Data Fig. 1c–f).

We observed efficient, sequence-specific, sgRNA-dependent C to U conversion *in vitro* (Fig. 1c). Conversion efficiency was greatest using rat APOBEC1–dCas9 linkers over nine amino acids in length. The number of positions susceptible to deamination (the ‘activity window’) increases from approximately 3 to 6 nt as the linker length was extended from 3 to 21 amino acids (Extended Data Fig. 1c–f). The 16-residue XTEN linker¹³ offered a promising balance between these two characteristics, with an efficient deamination window of approximately 5 nt, typically from positions 4 to 8 within the protospacer, counting the end distal to the protospacer-adjacent motif (PAM) as position 1. The rat APOBEC1–XTEN–dCas9 protein served as the first-generation base editor (BE1).

We assessed the ability of BE1 *in vitro* to correct seven T→C mutations relevant to human disease (Extended Data Fig. 2). BE1 yielded products consistent with efficient editing of the target C, or of at least

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ²Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA.

A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells

Brian Chaikind^{1,2}, Jeffrey L. Bessen^{1,2}, David B. Thompson^{1,3}, Johnny H. Hu^{1,3} and David R. Liu^{1,2,*}

¹Department of Chemistry & Chemical Biology, Harvard University, Cambridge, MA 02138, USA, ²Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138 USA and ³Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA 02138, USA

Received June 24, 2016; Revised August 01, 2016; Accepted August 02, 2016

ABSTRACT

We describe the development of 'recCas9', an RNA-programmed small serine recombinase that functions in mammalian cells. We fused a catalytically inactive dCas9 to the catalytic domain of Gin recombinase using an optimized fusion architecture. The resulting recCas9 system recombines DNA sites containing a minimal recombinase core site flanked by guide RNA-specified sequences. We show that these recombinases can operate on DNA sites in mammalian cells identical to genomic loci naturally found in the human genome in a manner that is dependent on the guide RNA sequences. DNA sequencing reveals that recCas9 catalyzes guide RNA-dependent recombination in human cells with an efficiency as high as 32% on plasmid substrates. Finally, we demonstrate that recCas9 expressed in human cells can catalyze *in situ* deletion between two genomic sites. Because recCas9 directly catalyzes recombination, it generates virtually no detectable indels or other stochastic DNA modification products. This work represents a step toward programmable, scarless genome editing in unmodified cells that is independent of endogenous cellular machinery or cell state. Current and future generations of recCas9 may facilitate targeted agricultural breeding, or the study and treatment of human genetic diseases.

INTRODUCTION

Efficient, programmable and site-specific homologous recombination remains a longstanding goal of genetics and genome editing (1). An enzyme that catalyzes recombination at sites specified by the researcher would be a valuable tool for studying the phenotypic effects of genetic

alterations, enabling gene integration or gene deletion-based therapeutic strategies or exchanging genes in a locus-specific manner during agricultural breeding. Early attempts at directing recombination to loci of interest relied on the transfection of donor DNA with long flanking sequences that are homologous to a target locus (2,3). This strategy was hampered by very low efficiency and thus, the need for a stringent selection to identify integrants. More recent efforts have exploited the ability of double-stranded DNA breaks (DSBs) to induce homology-directed repair (HDR). Homing endonucleases (4) and later programmable endonucleases such as zinc finger nucleases (5), TALE nucleases (6,7), Cas9 (8,9) and fCas9 (10,11) have been used to introduce targeted DSBs and induce HDR in the presence of donor DNA. In most post-mitotic cells, however, DSB-induced HDR is strongly down regulated and generally inefficient (12). Moreover, repair of DSBs by error-prone repair pathways such as non-homologous end-joining (NHEJ) or single-strand annealing (SSA) causes random insertions or deletions (indels) of nucleotides at the DSB site (12–14) at a higher frequency than HDR. The efficiency of HDR can be increased if cells are subjected to conditions forcing cell-cycle synchronization or if the enzymes involved in NHEJ are inhibited (15–17). However, such conditions can be highly perturbative to cells, limiting potential applications.

As a complementary approach, site-specific recombinases directly catalyze the cleavage, strand exchange and religation of two double-stranded DNA sequences. Recombination can result in the insertion, deletion or inversion of sequences of interest, depending on the relative orientation of the substrate sequences. In contrast with DNA nucleases, direct catalysis by recombinases typically does not provoke error-prone DNA repair processes that result in indel formation, is not dependent on endogenous cellular DNA repair machinery, and therefore leads primarily to a single defined product. Recombinase-mediated genome modifica-

*To whom correspondence should be addressed. Tel: +1 617 496 1067; Fax: +1 617 496 5688; Email: liu@chemistry.harvard.edu

Surgically disconnected temporal pole exhibits resting functional connectivity with remote brain regions

DAVID E. WARREN^{*1,2}, MATTHEW J. SUTTERER^{*2}, JOEL BRUSS², TAYLOR J. ABEL^{2,3}, ANDREW JONES², HIROTO KAWASAKI³, MICHELLE VOSS⁴, MARTIN CASSELL⁵, MATTHEW A. HOWARD³, and DANIEL TRANEL^{†2,4}

¹Department of Neurological Sciences, University of Nebraska Medical Center

²Department of Neurology, University of Iowa

³Department of Neurosurgery, University of Iowa

⁴Department of Psychological and Brain Sciences, University of Iowa

⁵Department of Anatomy and Cell Biology, University of Iowa

April 14, 2017

Abstract

Functional connectivity, as measured by resting-state fMRI, has proven a powerful method for studying brain systems in the context of behavior, development, and disease states. However, the relationship of functional connectivity to structural connectivity remains unclear. If functional connectivity relies on structural connectivity, then anatomical isolation of a brain region should eliminate functional connectivity with other brain regions. We tested this by measuring functional connectivity of the surgically disconnected temporal pole in resection patients (N=5; mean age 37; 2F, 3M). Functional connectivity was evaluated based on coactivation of whole-brain fMRI data with the average low-frequency BOLD signal from disconnected tissue in each patient. In sharp contrast to our prediction, we observed significant functional connectivity between the disconnected temporal pole and remote brain regions in each disconnection case. These findings raise important questions about the neural bases of functional connectivity measures derived from the fMRI BOLD signal.

1. INTRODUCTION

THE current consensus understanding of functional connectivity measured with resting-state functional magnetic resonance imaging (fMRI) is that measures of functional connectivity reflect — in part — temporal coactivation between neuronal populations. This follows from two premises. First, temporally correlated fMRI signals between brain regions reflect coordinated neural activity. Converging evidence from intracranial recordings, EEG, and fMRI supports this premise (David et al., 2008; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). Second, a necessary condition of functional connectivity that reflects communication between neuronal populations is structural connectivity, whether monosynaptic or polysynaptic (e.g., Buckner, Krienen, & Yeo, 2013). This premise

yields the straightforward prediction that a brain region with no structural connections to other brain regions would be functionally isolated. Some empirical support for this prediction exists (Gillebert & Mantini, 2013; Grekkes & Fink, 2011; Lu et al., 2011; Putnam, Wig, Grafton, Kelley, & Gazzaniga, 2008; Shen et al., 2015; Tovar-Moll et al., 2014; Tyszk, Kennedy, Adolphs, & Paul, 2011), but the strongest test of this prediction has not been conducted. Specifically, can brain tissue that has been fully, surgically disconnected from the rest of the brain show functional connectivity with other brain regions?

Prior efforts that addressed the complex relationship between structural and functional connectivity have provided some insight, but the implications of their findings are sometimes contradictory (Damoiseaux & Greicius, 2009; Greicius, Supekar, Menon, & Dougherty, 2009; Shen et al., 2015). Structural connectivity does pre-

^{*}These authors contributed equally.

[†]Corresponding author: daniel-tranel@uiowa.edu

© 2017 by the authors. All rights reserved.

Classification: Biological Sciences; Evolution, Genetics.

Title:
Refactoring the Genetic Code for Increased Evolvability

Gur Pines^{a,b,1,2}, James D. Winkler^{a,b,1,3}, Assaf Pines and Ryan T. Gill^{a,b}

Affiliations:

^a Renewable and Sustainable Energy Institute (RASEI), University of Colorado Boulder, Boulder, Colorado 80303, United States

^b Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States

¹ G.P. and J.D.W contributed equally to this work.

² Corresponding author: Gur Pines, Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States, gur.pines@colorado.edu

³ Current address: Shell Biodomain, 3333 Texas 6, Houston, TX, USA.

Author contributions: G.P, J.D.W and R.T.G designed research; G.P and J.D.W performed research; G.P, J.D.W and A.P analyzed data; G.P, J.D.W and R.T.G wrote the paper.

The authors declare no conflict of interest.

Development of New Inkjet Head Applying MEMS Technology and Thin Film Actuator

Kenji MAWATARI *, Koich SAMESHIMA *, Mitsuyoshi MIYAI *, Shinya MATSUDA *

Abstract

We developed a new inkjet head by applying MEMS technology and thin film piezo actuator. Jetting properties of inkjet heads were calculated by the simulation method of the equivalent circuit model generated from actuator properties and ink flow channels of the inkjet head.

We manufactured a test piece to investigate the jetting properties and oscillation forms of the actuator. As a result, our test piece was driven at maximum 70 kHz and ejected 3 pl droplet with an ink which viscosity was 10 mPa·s. We found that the experimented jetting properties and vibration forms agreed very well with the simulation.

1 Introduction

Konica Minolta, Inc. introduced newly developed inkjet printer “KM-1” at drupa 2012. Figure 1 shows an image of KM-1. This printer can print on up to B2 paper size with 1200×1200 dpi resolution at the speed of maximum 3300 sheets per hour simplex [1]. Recently, demands for high-speed and high-resolution printers are increasing more and more; this machine is one of the answer for these demands.



Fig. 1 KONICAMINOLTA KM-1.

Developing a small-sized, high-speed and high nozzle density inkjet head is demanded for achieving the high printing performance. Reducing the printing head cost is also needed because the number of print heads for one printer will increase more in the future. To meet these requirements for inkjet heads, we are trying to develop a new inkjet head by applying MEMS technology and thin film actuator. Inkjet head using thin film actuator is becoming popular recently [2]. We can arrange high density nozzles into the inkjet head with lower cost by using thin film actuator because they are manufactured by patterning a thin piezo film coated on a Si substrate.

We had reported the development of inkjet head utilizing MEMS technology for printed electronics applications at NIP 28 [3]. That head consisted of silicon (Si) actuator plate attached a piezo ceramic element (bulk actuator) and Si nozzle plate.

* 2nd R&D Division, Inkjet Business Unit, Konica Minolta, Inc.; Takatsuki, Osaka, Japan.

The Night Watch

JAMES MICKENS



James Mickens is a researcher in the Distributed Systems group at Microsoft's Redmond lab. His current research

focuses on web applications, with an emphasis on the design of JavaScript frameworks that allow developers to diagnose and fix bugs in widely deployed web applications. James also works on fast, scalable storage systems for datacenters. James received his PhD in computer science from the University of Michigan, and a bachelor's degree in computer science from Georgia Tech. mickens@microsoft.com

As a highly trained academic researcher, I spend a lot of time trying to advance the frontiers of human knowledge. However, as someone who was born in the South, I secretly believe that true progress is a fantasy, and that I need to prepare for the end times, and for the chickens coming home to roost, and fast zombies, and slow zombies, and the polite zombies who say “sir” and “ma’am” but then try to eat your brain to acquire your skills. When the revolution comes, I need to be prepared; thus, in the quiet moments, when I’m not producing incredible scientific breakthroughs, I think about what I’ll do when the weather forecast inevitably becomes RIVERS OF BLOOD ALL DAY EVERY DAY. The main thing that I ponder is who will be in my gang, because the likelihood of post-apocalyptic survival is directly related to the size and quality of your rag-tag group of associates. There are some obvious people who I’ll need to recruit: a locksmith (to open doors); a demolitions expert (for when the locksmith has run out of ideas); and a person who can procure, train, and then throw snakes at my enemies (because, in a world without hope, snake throwing is a reasonable way to resolve disputes). All of these people will play a role in my ultimate success as a dystopian warlord philosopher. However, the most important person in my gang will be a systems programmer. A person who can debug a device driver or a distributed system is a person who can be trusted in a Hobbesian nightmare of breathtaking scope; a systems programmer has seen the terrors of the world and understood the intrinsic horror of existence. The systems programmer has written drivers for buggy devices whose firmware was implemented by a drunken child or a sober goldfish. The systems programmer has traced a network problem across eight machines, three time zones, and a brief diversion into Amish country, where the problem was transmitted in the front left hoof of a mule named Deliverance. The systems programmer has read the kernel source, to better understand the deep ways of the universe, and the systems programmer has seen the comment in the scheduler that says “DOES THIS WORK LOL,” and the systems programmer has wept instead of LOLed, and the systems programmer has submitted a kernel patch to restore balance to The Force and fix the priority inversion that was causing MySQL to hang. A systems programmer will know what to do when society breaks down, because the systems programmer already lives in a world without law.

Self-Deployed Extremely Large Low Mass Space Structures

Devon G. Crowe, Prakash Joshi,
Ed Rietman, and Kophu Chang*

Physical Sciences, Inc.

March 7, 2007

*With a little help from our friends W. Cash and A. Labeyrie

THE FAILURE RESOLUTION OF LEHMAN BROTHERS

- The experience of resolving Lehman in the bankruptcy courts has led to an active debate about the effectiveness of U.S. Chapter 11 proceedings for complex financial institutions.
- Lehman’s poor pre-bankruptcy planning may have substantially reduced the value of Lehman’s estate and contributed to many ensuing disputes with creditors.
- For over-the-counter (OTC) derivatives transactions, where much of the complexity of Lehman’s bankruptcy resolution was rooted, creditors’ recovery rate was below historical averages for failed firms comparable to Lehman.
- The settlement of OTC derivatives was a long and complex process, occurring on different tracks for different groups of derivatives creditors.
- Some of the losses borne by Lehman investors stemmed from the manner in which Lehman failed and could have been avoided in a more orderly process.

Michael J. Fleming is a vice president and Asani Sarkar an assistant vice president at the Federal Reserve Bank of New York.
michael.fleming@ny.frb.org; asani.sarkar@ny.frb.org

1. INTRODUCTION

Lehman Brothers Holdings Inc. (LBHI) filed for Chapter 11 bankruptcy on September 15, 2008, while its subsidiaries did so over the subsequent months (see Exhibit 1 for Lehman’s organizational structure).¹ With 209 registered subsidiaries in twenty-one countries, Lehman’s Chapter 11 filing was one of the largest and most complex in history. Creditors filed about \$1.2 trillion of claims against the Lehman estate (LBHI, “The State of the Estate,” September 22, 2010), which was party to more than 900,000 derivatives contracts at the time of bankruptcy.

Several bodies of law applied to Lehman’s various corporate entities (Exhibit 2):

- The U.S. Bankruptcy Code applied to LBHI and its subsidiaries.
- The Securities Investor Protection Act (SIPA) regime applied to the insolvent broker-dealer, Lehman Brothers Inc. (LBI).
- More than eighty jurisdictions’ insolvency laws applied to the non-U.S. Lehman Brothers entities, such as Lehman’s U.K.-based broker-dealer Lehman Brothers International (Europe) (LBIE).

¹ When referring to LBHI and all its subsidiaries as an ensemble, we use “Lehman.” Otherwise, when referring to the holding company (subsidiary), we use “LBHI” (the subsidiary name). Appendix A lists the acronyms and initialisms used in the article.

The authors thank Tobias Adrian, Wilson Ervin, Sahil Godiwala, Anna Kovner, Lisa Kraidin, Antoine Martin, James McAndrews, Hamid Mehran, João Santos, Joseph Sommer, and Emily Warren for helpful discussions and/or comments on earlier drafts as well as Samuel Antill, Weiling Liu, and Parinitha Sastry for excellent research assistance. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

A simple approach to the numerical simulation with trimmed CAD surfaces

G. Beer^{a,b}, B. Marussig and J. Zechner^a

^a*Institute for Structural Analysis, Graz University of Technology, Lessingstrasse 25, Graz, Austria, e-mail: gernot.beer@tugraz.at*

^b*Centre for Geotechnical and Materials Modelling, The University of Newcastle, Australia*

Abstract

In this work a novel method for the analysis with trimmed CAD surfaces is presented. The method involves an additional mapping step and the attraction stems from its simplicity and ease of implementation into existing Finite Element (FEM) or Boundary Element (BEM) software. The method is first verified with classical test examples in structural mechanics. Then two practical applications are presented one using the FEM, the other the BEM, that show the applicability of the method.

Keywords: trimming, CAD, isogeometric analysis.

1. Introduction

The original idea of isogeometric analysis was to use data from CAD programs directly for describing the geometry of the problem without the need to generate a mesh. Geometry data are provided by CAD programs in a standard ASCII format (IGES standard¹) and can be read by the analysis program. CAD programs describe surfaces using NURBS functions. For complex surfaces, especially involving a union of surfaces, trimming is used. The data provided in the IGES file contain NURBS information of surfaces (control points, knot vectors and weights), as well as trimming information, if the surface is trimmed. The trimming information is supplied in the global and local coordinate system of the surface to be trimmed. Figure 1 shows an example of an IGES output.

The basic idea of our approach is to use information provided in the IGES file directly for the definition of the geometry and other derived values such as the computation of the Jacobian and first and second derivatives. We are aware of only three papers that have addressed the problem of using trimmed CAD data for analysis. The first two papers ([1] and [2]) propose to use a regular grid of elements, that are defined by knot spans. A searching algorithm is employed, that allows to determine how the elements are transected by trimming curves. For the numerical integration trimmed elements are subdivided into triangular subregions.

¹available from http://diyhlpl.us/bryan/papers2/IGES5-3_forDownload.pdf

Towards Storytelling from Visual Lifelogging: An Overview

Marc Bolaños*, Mariella Dimiccoli*, and Petia Radeva

Abstract—Visual lifelogging consists of acquiring images that capture the daily experiences of the user by wearing a camera over a long period of time. The pictures taken offer considerable potential for knowledge mining concerning how people live their lives, hence, they open up new opportunities for many potential applications in fields including healthcare, security, leisure and the quantified self. However, automatically building a story from a huge collection of unstructured egocentric data presents major challenges. This paper provides a thorough review of advances made so far in egocentric data analysis, and in view of the current state of the art, indicates new lines of research to move us towards storytelling from visual lifelogging.

Index Terms—visual lifelogging, egocentric vision, storytelling

I. INTRODUCTION

LIFELOGGING consists of a user continuously recording their everyday experiences, typically via wearable sensors including accelerometers and cameras, among others. When the visual signal is the only one recorded, typically by a wearable camera, it is referred to as *visual lifelogging*. This is a trend that is rapidly increasing thanks to advances in wearable technologies over recent years. Nowadays, wearable cameras are very small devices that can be worn all-day long and automatically record the everyday activities of the wearer in a passive fashion, from a first-person point of view. As an example, Fig. 1 shows pictures taken by a person walking down a street while wearing such a camera.

Most wearable cameras on the market like GoPro, MeCam, Looxcie or Google Glass (see Fig. 2 (a) and (c)) are video cameras, which have relatively High Temporal Resolution (HTR) (e.g. from 25 up to 60 frames per second) and are more suitable to record specific moments, such as cooking or doing sports. A limited number of wearable cameras, such as Narrative Clip and SenseCam (see Fig. 2 (b) and (d)) are photographic cameras, which have Low Temporal Resolution (LTR) (2-3 frames per minute), and hence are more suitable for acquiring data over long periods of time. On the one hand, data recorded at specific moments with video cameras offer potential for in-depth analysis of daily or special activities, allowing to capture even *how* something happened. On the other hand, data acquired over long periods of time, commonly called *visual lifelogs*, offer considerable potential for inferring knowledge about e.g. behaviour patterns, and hence enable many applications that would not be possible with HTR cameras. As shown by Doherty et al. [32], visual lifelogs captured through a SenseCam, which as opposed to video cameras can capture



Fig. 1. Example of a sequence acquired by the Narrative Clip wearable camera while the user is walking down a street. The temporal leaps between neighbouring pictures produced by photographic cameras are common in dynamic environments and make the extraction of information from closely spaced images very difficult.



Fig. 2. Examples of wearable cameras on the market: (a) GoPro (2002). (b) SenseCam (2005). (c) Looxcie (2011). (d) Narrative Clip (2013).

the whole day, could be used to prevent non-communicable diseases associated with unhealthy trends and risky profiles (such as obesity or depression, among others). Additionally, they could also help prevent cognitive and functional decline in elderly people [29], [44], [57]. However, visual lifelogs present a significant challenge for automatic visual analysis. Indeed, due to the free motion of the camera and to its LTR, abrupt changes in lighting conditions and image content are very frequent (see Fig. 1). In such situations, computer vision techniques based on temporal coherence and motion estimation become unreliable. Recognition algorithms have to cope with the huge variety of objects that appear. In addition, due to the non-intentional nature of the pictures captured, they generally contain severely occluded objects, artefacts such as blurring or light saturation [89] and a large number of non-informative images that capture non-meaningful information such as walls, the sky, parts of objects, etc. Furthermore, the sheer number of data that a visual lifelog consists of and the rate at which they increase (up to 2,000 images per day or around 800,000 images every year) imposes a need for efficient methods to extract and locate relevant content concerning the wearer from the photo stream. Regarding HTR cameras, if they were employed for a lifelog analysis, the problem of the amount of data would be even more acute, and would additionally imply the need of huge computational resources.

In response to the challenges and opportunities introduced by analysis of visual lifelogs, and more generally, by wearable cameras, computer vision scientists have rapidly become more

M. Bolaños, marc.bolanos@ub.edu, M. Dimiccoli, mariella.dimiccoli@cvc.uab.es, P. Radeva, petia.ivanova@ub.edu - Universitat de Barcelona, Barcelona, Spain and Computer Vision Center, Bellaterra, Spain. *The first two authors contributed equally to this work.

DEEP COMPRESSION: COMPRESSING DEEP NEURAL NETWORKS WITH PRUNING, TRAINED QUANTIZATION AND HUFFMAN CODING

Song Han

Stanford University, Stanford, CA 94305, USA
songhan@stanford.edu

Huizi Mao

Tsinghua University, Beijing, 100084, China
mhzi12@mails.tsinghua.edu.cn

William J. Dally

Stanford University, Stanford, CA 94305, USA
NVIDIA, Santa Clara, CA 95050, USA
dally@stanford.edu

ABSTRACT

Neural networks are both computationally intensive and memory intensive, making them difficult to deploy on embedded systems with limited hardware resources. To address this limitation, we introduce “deep compression”, a three stage pipeline: pruning, trained quantization and Huffman coding, that work together to reduce the storage requirement of neural networks by $35\times$ to $49\times$ without affecting their accuracy. Our method first prunes the network by learning only the important connections. Next, we quantize the weights to enforce weight sharing, finally, we apply Huffman coding. After the first two steps we retrain the network to fine tune the remaining connections and the quantized centroids. Pruning, reduces the number of connections by $9\times$ to $13\times$; Quantization then reduces the number of bits that represent each connection from 32 to 5. On the ImageNet dataset, our method reduced the storage required by AlexNet by $35\times$, from 240MB to 6.9MB, without loss of accuracy. Our method reduced the size of VGG-16 by $49\times$ from 552MB to 11.3MB, again with no loss of accuracy. This allows fitting the model into on-chip SRAM cache rather than off-chip DRAM memory. Our compression method also facilitates the use of complex neural networks in mobile applications where application size and download bandwidth are constrained. Benchmarked on CPU, GPU and mobile GPU, compressed network has $3\times$ to $4\times$ layerwise speedup and $3\times$ to $7\times$ better energy efficiency.

1 INTRODUCTION

Deep neural networks have evolved to the state-of-the-art technique for computer vision tasks (Krizhevsky et al., 2012)(Simonyan & Zisserman, 2014). Though these neural networks are very powerful, the large number of weights consumes considerable storage and memory bandwidth. For example, the AlexNet Caffemodel is over 200MB, and the VGG-16 Caffemodel is over 500MB (BVLC). This makes it difficult to deploy deep neural networks on mobile system.

First, for many mobile-first companies such as Baidu and Facebook, various apps are updated via different app stores, and they are very sensitive to the size of the binary files. For example, App Store has the restriction “apps above 100 MB will not download until you connect to Wi-Fi”. As a result, a feature that increases the binary size by 100MB will receive much more scrutiny than one that increases it by 10MB. Although having deep neural networks running on mobile has many great

SQUEEZE NET: ALEX NET-LEVEL ACCURACY WITH 50X FEWER PARAMETERS AND <0.5 MB MODEL SIZE

Forrest N. Iandola¹, Song Han², Matthew W. Moskewicz¹, Khalid Ashraf¹,
William J. Dally², Kurt Keutzer¹

¹DeepScale* & UC Berkeley ²Stanford University

{forresti, moskewicz, kashraf, keutzer}@eecs.berkeley.edu

{songhan, dally}@stanford.edu

ABSTRACT

Recent research on deep convolutional neural networks (CNNs) has focused primarily on improving accuracy. For a given accuracy level, it is typically possible to identify multiple CNN architectures that achieve that accuracy level. With equivalent accuracy, smaller CNN architectures offer at least three advantages: (1) Smaller CNNs require less communication across servers during distributed training. (2) Smaller CNNs require less bandwidth to export a new model from the cloud to an autonomous car. (3) Smaller CNNs are more feasible to deploy on FPGAs and other hardware with limited memory. To provide all of these advantages, we propose a small CNN architecture called SqueezeNet. SqueezeNet achieves AlexNet-level accuracy on ImageNet with 50x fewer parameters. Additionally, with model compression techniques, we are able to compress SqueezeNet to less than 0.5MB (510 \times smaller than AlexNet).

The SqueezeNet architecture is available for download here: <https://github.com/DeepScale/SqueezeNet>

1 INTRODUCTION AND MOTIVATION

Much of the recent research on deep convolutional neural networks (CNNs) has focused on increasing accuracy on computer vision datasets. For a given accuracy level, there typically exist multiple CNN architectures that achieve that accuracy level. Given equivalent accuracy, a CNN architecture with fewer parameters has several advantages:

- **More efficient distributed training.** Communication among servers is the limiting factor to the scalability of distributed CNN training. For distributed data-parallel training, communication overhead is directly proportional to the number of parameters in the model (Iandola et al., 2016). In short, small models train faster due to requiring less communication.
- **Less overhead when exporting new models to clients.** For autonomous driving, companies such as Tesla periodically copy new models from their servers to customers' cars. This practice is often referred to as an *over-the-air* update. Consumer Reports has found that the safety of Tesla's *Autopilot* semi-autonomous driving functionality has incrementally improved with recent over-the-air updates (Consumer Reports, 2016). However, over-the-air updates of today's typical CNN/DNN models can require large data transfers. With AlexNet, this would require 240MB of communication from the server to the car. Smaller models require less communication, making frequent updates more feasible.
- **Feasible FPGA and embedded deployment.** FPGAs often have less than 10MB¹ of on-chip memory and no off-chip memory or storage. For inference, a sufficiently small model could be stored directly on the FPGA instead of being bottlenecked by memory bandwidth (Qiu et al., 2016), while video frames stream through the FPGA in real time. Further, when deploying CNNs on Application-Specific Integrated Circuits (ASICs), a sufficiently small model could be stored directly on-chip, and smaller models may enable the ASIC to fit on a smaller die.

*<http://deepscale.ai>

¹For example, the Xilinx Vertex-7 FPGA has a maximum of 8.5 MBytes (i.e. 68 Mbits) of on-chip memory and does not provide off-chip memory.

Generative Adversarial Text to Image Synthesis

Scott Reed, Zeynep Akata, Xinchun Yan, Lajanugen Logeswaran
Bernt Schiele, Honglak Lee

REEDSCOT¹, AKATA², XCYAN¹, LLAJAN¹
SCHIELE², HONGLAK¹

¹ University of Michigan, Ann Arbor, MI, USA (UMICH.EDU)

² Max Planck Institute for Informatics, Saarbrücken, Germany (MPI-INF.MPG.DE)

Abstract

Automatic synthesis of realistic images from text would be interesting and useful, but current AI systems are still far from this goal. However, in recent years generic and powerful recurrent neural network architectures have been developed to learn discriminative text feature representations. Meanwhile, deep convolutional generative adversarial networks (GANs) have begun to generate highly compelling images of specific categories, such as faces, album covers, and room interiors. In this work, we develop a novel deep architecture and GAN formulation to effectively bridge these advances in text and image modeling, translating visual concepts from characters to pixels. We demonstrate the capability of our model to generate plausible images of birds and flowers from detailed text descriptions.

1. Introduction

In this work we are interested in translating text in the form of single-sentence human-written descriptions directly into image pixels. For example, “this small bird has a short, pointy orange beak and white belly” or “the petals of this flower are pink and the anther are yellow”. The problem of generating images from visual descriptions gained interest in the research community, but it is far from being solved.

Traditionally this type of detailed visual information about an object has been captured in attribute representations - distinguishing characteristics the object category encoded into a vector (Farhadi et al., 2009; Kumar et al., 2009; Parikh & Grauman, 2011; Lampert et al., 2014), in particular to enable zero-shot visual recognition (Fu et al., 2014; Akata et al., 2015), and recently for conditional image generation (Yan et al., 2015).

While the discriminative power and strong generalization

Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 2016. JMLR: W&CP volume 48. Copyright 2016 by the author(s).

this small bird has a pink breast and crown, and black primaries and secondaries.



the flower has petals that are bright pinkish purple with white stigma



this magnificent fellow is almost all black with a red crest, and white cheek patch.



this white and yellow flower have thin white petals and a round yellow stamen



Figure 1. Examples of generated images from text descriptions. Left: captions are from zero-shot (held out) categories, unseen text. Right: captions are from the training set.

properties of attribute representations are attractive, attributes are also cumbersome to obtain as they may require domain-specific knowledge. In comparison, natural language offers a general and flexible interface for describing objects in any space of visual categories. Ideally, we could have the generality of text descriptions with the discriminative power of attributes.

Recently, deep convolutional and recurrent networks for text have yielded highly discriminative and generalizable (in the zero-shot learning sense) text representations learned automatically from words and characters (Reed et al., 2016). These approaches exceed the previous state-of-the-art using attributes for zero-shot visual recognition on the Caltech-UCSD birds database (Wah et al., 2011), and also are capable of zero-shot caption-based retrieval. Motivated by these works, we aim to learn a mapping directly from words and characters to image pixels.

To solve this challenging problem requires solving two sub-problems: first, learn a text feature representation that captures the important visual details; and second, use these fea-

ZONEOUT: REGULARIZING RNNs BY RANDOMLY PRESERVING HIDDEN ACTIVATIONS

David Krueger^{1,*}, Tegan Maharaj^{2,*}, János Kramár²

Mohammad Pezeshki¹, Nicolas Ballas¹, Nan Rosemary Ke², Anirudh Goyal¹

Yoshua Bengio^{1†}, Aaron Courville^{1‡}, Christopher Pal²

¹ MILA, Université de Montréal, `firstname.lastname@umontreal.ca`.

² École Polytechnique de Montréal, `firstname.lastname@polymtl.ca`.

* Equal contributions. † CIFAR Senior Fellow. ‡ CIFAR Fellow.

ABSTRACT

We propose zoneout, a novel method for regularizing RNNs. At each timestep, zoneout stochastically forces some hidden units to maintain their previous values. Like dropout, zoneout uses random noise to train a pseudo-ensemble, improving generalization. But by preserving instead of dropping hidden units, gradient information and state information are more readily propagated through time, as in feedforward stochastic depth networks. We perform an empirical investigation of various RNN regularizers, and find that zoneout gives significant performance improvements across tasks. We achieve competitive results with relatively simple models in character- and word-level language modelling on the Penn Treebank and Text8 datasets, and combining with recurrent batch normalization (Cooijmans et al., 2016) yields state-of-the-art results on permuted sequential MNIST.

1 INTRODUCTION

Regularizing neural nets can significantly improve performance, as indicated by the widespread use of early stopping, and success of regularization methods such as dropout and its recurrent variants (Hinton et al., 2012; Srivastava et al., 2014; Zaremba et al., 2014; Gal, 2015). In this paper, we address the issue of regularization in recurrent neural networks (RNNs) with a novel method called **zoneout**.

RNNs sequentially construct fixed-length representations of arbitrary-length sequences by folding new observations into their hidden state using an input-dependent transition operator. The repeated application of the same transition operator at the different time steps of the sequence, however, can make the dynamics of an RNN sensitive to minor perturbations in the hidden state; the transition dynamics can magnify components of these perturbations exponentially. Zoneout aims to improve RNNs’ robustness to perturbations in the hidden state in order to regularize transition dynamics.

Like dropout, zoneout injects noise during training. But instead of setting some units’ activations to 0 as in dropout, zoneout randomly replaces some units’ activations with their activations from the previous timestep. As in dropout, we use the expectation of the random noise at test time. This results in a simple regularization approach which can be applied through time for any RNN architecture, and can be conceptually extended to any model whose state varies over time.

Compared with dropout, zoneout is appealing because it preserves information flow forwards and backwards through the network. This helps combat the vanishing gradient problem (Hochreiter, 1991; Bengio et al., 1994), as we observe experimentally.

We also empirically evaluate zoneout on classification using the permuted sequential MNIST dataset, and on language modelling using the Penn Treebank and Text8 datasets, demonstrating competitive or state-of-the-art performance across tasks. In particular, we show that zoneout performs competitively with other proposed regularization methods for RNNs, including recently-proposed dropout variants. Code for replicating all experiments can be found at: <http://github.com/teganmaharaj/zoneout>

Egalitarian computing

Alex Biryukov
University of Luxembourg

Dmitry Khovratovich
University of Luxembourg

Abstract

In this paper we explore several contexts where an adversary has an upper hand over the defender by using special hardware in an attack. These include password processing, hard-drive protection, cryptocurrency mining, resource sharing, code obfuscation, etc.

We suggest memory-hard computing as a generic paradigm, where every task is amalgamated with a certain procedure requiring intensive access to RAM both in terms of size and (very importantly) bandwidth, so that transferring the computation to GPU, FPGA, and even ASIC brings little or no cost reduction. Cryptographic schemes that run in this framework become *egalitarian* in the sense that both users and attackers are equal in the price-performance ratio conditions.

Based on existing schemes like Argon2 and the recent generalized-birthday proof-of-work, we suggest a generic framework and two new schemes:

- MTP, a memory-hard Proof-of-Work based on the memory-hard function with fast verification and short proofs. It can be also used for memory-hard time-lock puzzles.
- MHE, the concept of memory-hard encryption, which utilizes available RAM to strengthen the encryption for the low-entropy keys (allowing to bring back 6 letter passwords).

Keywords: MTP, MHE, Argon2, memory-hard, asymmetric, proof-of-work, botnets, encryption, time-lock puzzles.

I. INTRODUCTION

A. Motivation

Historically attackers have had more resources than defenders, which is still mostly true. Whether it is secret key recovery or document forgery, the attackers are ready to spend tremendous amount of computing power to achieve the goal. In some settings it is possible to make most attacks infeasible by simply setting the key length to 128 bits and higher. In other settings the secret is limited and the best the defender can do is to increase the time needed for the attack, but not to render the attack impossible.

Passwords, typically stored in a hashed form, are a classical example. As people tend to choose passwords of very low entropy, the security designers added unique salts and then increased the number of hash iterations. In response the attackers switched to dedicated hardware for password cracking, so that the price of single password recovery dropped dramatically, sometimes by a few orders of magnitude. A similar situation occurred in other contexts. The Bitcoin cryptocurrency relies on continuous preimage search for the SHA-256 hash function,

which is much cheaper on custom ASICs, consuming up to 30,000 times less energy per solution than most efficient x86 laptops [2]. Eventually, the original concept of an egalitarian cryptocurrency [25] vanished with the emergence of huge and centralized mining pools. Related problems include password-based key derivation for hard-drive encryption, where the data confidentiality directly depends on the password entropy, and where offline attack is exceptionally easy once the drive is stolen. Similar situation arise in the resource sharing and spam countermeasures. In the latter it is proposed that every user presents a certain proof (often called proof-of-work), which should be too expensive for spammers to generate on a large scale. Yet another setting is that of code obfuscation, in which powerful reverse-engineering/de-compilation tools can be used in order to lift the proprietary code or secrets embedded in the software.

B. Egalitarian computing

Our idea is to remedy the disparity between ordinary users and adversaries/cheaters, where latter could use botnets, GPU, FPGA, ASICs to get an advantage and run a cheaper attack. We call it *egalitarian computing* as it should establish the same price for a single computation unit on all platforms, so that the defender's hardware is optimal both for attack and defence. Equipped with egalitarian crypto schemes, defenders may hope to become to be on par with the most powerful attackers.

The key element of our approach is large (in size) and intensive (in bandwidth) use of RAM as a widely available and rather cheap unit for most defenders. In turn, RAM is rather expensive on FPGA and ASIC¹, and slow on GPU, at least compared to memoryless computing tasks. All our schemes use a lot of memory and a lot of bandwidth — almost as much as possible.

We suggest a single framework for this concept and concrete schemes with an unique combination of features.

In the future, adoption of our concept could allow a homogenization of computing resources, a simplified security analysis, and relaxed security requirements. When all attackers use the same hardware as defenders, automated large-scale attacks are no longer possible. Shorter keys, shorter passwords, faster and more transparent schemes may come back to use.

¹The memory effect on ASICs can be illustrated as follows. A compact 50-nm DRAM implementation [17] takes 500 mm² per GB, which is equivalent to about 15000 10 MHz SHA-256 cores in the best Bitcoin 40-nm ASICs [1] and is comparable to a CPU size. Therefore, an algorithm requiring 1 GB for 1 minute would have the same AT cost as an algorithm requiring 2⁴² hash function calls, whereas the latter can not finish on a PC even in 1 day. In other words, the use of memory can increase the AT cost by a factor of 1000 and more at the same time cost for the desktop user.

Towards an integration of deep learning and neuroscience

Adam H. Marblestone

MIT Media Lab

Cambridge, MA 02139, USA

AMARBLES@MEDIA.MIT.EDU

Greg Wayne

Google Deepmind

London, EC4A 3TW, UK

GREGWAYNE@GOOGLE.COM

Konrad P. Kording

Rehabilitation Institute of Chicago

Northwestern University

Chicago, IL 60611, USA

KOERDING@GMAIL.COM

Editor: TBN

Abstract

Neuroscience has focused on the detailed implementation of computation, studying neural codes, dynamics and circuits. In machine learning, however, artificial neural networks tend to eschew precisely designed codes, dynamics or circuits in favor of brute force optimization of a cost function, often using simple and relatively uniform initial architectures. Two recent developments have emerged within machine learning that create an opportunity to connect these seemingly divergent perspectives. First, structured architectures are used, including dedicated systems for attention, recursion and various forms of short- and long-term memory storage. Second, cost functions and training procedures have become more complex and are varied across layers and over time. Here we think about the brain in terms of these ideas. We hypothesize that (1) the brain optimizes cost functions, (2) these cost functions are diverse and differ across brain locations and over development, and (3) optimization operates within a pre-structured architecture matched to the computational problems posed by behavior. Such a heterogeneously optimized system, enabled by a series of interacting cost functions, serves to make learning data-efficient and precisely targeted to the needs of the organism. We suggest directions by which neuroscience could seek to refine and test these hypotheses.

Keywords: Cost Functions, Neural Networks, Neuroscience, Cognitive Architecture

Contents

1	Introduction	3
1.1	<i>Hypothesis 1</i> – The brain optimizes cost functions.	3
1.2	<i>Hypothesis 2</i> – Cost functions are diverse across areas and change over development.	4
1.3	<i>Hypothesis 3</i> – Specialized systems allow efficiently solving key computational problems.	4
2	The brain can optimize cost functions	5
2.1	Local self-organization and optimization without multi-layer credit assignment	7
2.2	Biological implementation of optimization	7
2.2.1	The need for efficient gradient descent in multi-layer networks	7
2.2.2	Biologically plausible approximations of gradient descent	8
2.3	Alternative mechanisms for learning	12
2.3.1	Exploiting biological neural mechanisms	12
2.3.2	Learning in the cortical sheet	13

Rationalizing Neural Predictions

Tao Lei, Regina Barzilay and Tommi Jaakkola

Computer Science and Artificial Intelligence Laboratory

Massachusetts Institute of Technology

{taolei, regina, tommi}@csail.mit.edu

Abstract

Prediction without justification has limited applicability. As a remedy, we learn to extract pieces of input text as justifications – rationales – that are tailored to be short and coherent, yet sufficient for making the same prediction. Our approach combines two modular components, generator and encoder, which are trained to operate well together. The generator specifies a distribution over text fragments as candidate rationales and these are passed through the encoder for prediction. Rationales are never given during training. Instead, the model is regularized by desiderata for rationales. We evaluate the approach on multi-aspect sentiment analysis against manually annotated test cases. Our approach outperforms attention-based baseline by a significant margin. We also successfully illustrate the method on the question retrieval task.¹

1 Introduction

Many recent advances in NLP problems have come from formulating and training expressive and elaborate neural models. This includes models for sentiment classification, parsing, and machine translation among many others. The gains in accuracy have, however, come at the cost of interpretability since complex neural models offer little transparency concerning their inner workings. In many applications, such as medicine, predictions are used to drive critical decisions, including treatment options. It is necessary in such cases to be able to verify and under-

¹Our code and data are available at <https://github.com/taolei87/rcnn>.

<i>Review</i>		
the beer was n't what i expected, and i'm not sure it's "true to style", but i thought it was delicious. a very pleasant ruby red-amber color with a relatively brilliant finish, but a limited amount of carbonation, from the look of it. aroma is what i think an amber ale should be - a nice blend of caramel and happiness bound together.		
<i>Ratings</i>		
	<i>Look: 5 stars</i>	<i>Smell: 4 stars</i>

Figure 1: An example of a review with ranking in two categories. The rationale for Look prediction is shown in bold.

stand the underlying basis for the decisions. Ideally, complex neural models would not only yield improved performance but would also offer interpretable justifications – rationales – for their predictions.

In this paper, we propose a novel approach to incorporating rationale generation as an integral part of the overall learning problem. We limit ourselves to extractive (as opposed to abstractive) rationales. From this perspective, our rationales are simply subsets of the words from the input text that satisfy two key properties. First, the selected words represent short and coherent pieces of text (e.g., phrases) and, second, the selected words must alone suffice for prediction as a substitute of the original text. More concretely, consider the task of multi-aspect sentiment analysis. Figure 1 illustrates a product review along with user rating in terms of two categories or aspects. If the model in this case predicts five star rating for color, it should also identify the phrase “*a very pleasant ruby red-amber color*” as the rationale underlying this decision.

In most practical applications, rationale genera-

Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks

Hendrik Strobelt, Sebastian Gehrmann, Bernd Huber, Hanspeter Pfister, and Alexander M. Rush
*Harvard School of Engineering and Applied Sciences*¹
 {hstrobelt, gehrmann, huber, pfister, rush}@seas.harvard.edu



Fig. 1. The LSTMVis user interface. The user interactively *selects* a range of text specifying a hypothesis about the model. This range is then used to *match* similar hidden state patterns in the dataset which are displayed below. The selection is made by specifying a start-stop range in the text (gray border (b) and blue highlight (c)) and an activation threshold (red dashed line). The selection is visualized by (a) the hidden states selected, (d) the number of active states and (e) the activation ranges for each hidden state. The tool can then *match* this selection with similar hidden state patterns in the data set of varying lengths (f), providing insight into the representations learned by the model. The match view additionally includes user-defined meta-data (such as part-of-speech tags) (g) which allows the user to further refine or confirm the selection hypothesis.

Abstract— Recurrent neural networks, and in particular long short-term memory networks (LSTMs), are a remarkably effective tool for sequence modeling that learn a dense black-box hidden representation of their sequential input. Researchers interested in better understanding these models have studied the changes in hidden state representations over time and noticed some interpretable patterns but also significant noise. In this work, we present LSTMVis a visual analysis tool for recurrent neural networks with a focus on understanding these hidden state dynamics. The tool allows a user to select a hypothesis input range to focus on local state changes, to match these states changes to similar patterns in a large data set, and to align these results with domain specific structural annotations. We further show several use cases of the tool for analyzing specific hidden state properties on data sets containing nesting, phrase structure, and chord progressions, and demonstrate how the tool can be used to isolate patterns for further statistical analysis.

1 INTRODUCTION

Recurrent neural networks (RNNs) [6] have proven to be a very effective general-purpose model for capturing long-term dependencies in textual applications. Recent strong empirical results indicate that internal representations learned by RNNs capture complex relationships between the words within a sentence or document. These improved representation have led directly to end applications in machine translation [12, 22], speech recognition [2], music generation [4], and text classification [5], among a variety of other applications.

While RNNs have shown clear improvements for sequence modeling, the models themselves are black boxes, and it remains unclear

exactly *how* a particular model is representing long-distance relationships within a sequence. Typically, RNNs contain millions of parameters and utilize repeated non-linear transformations of large hidden representations under time-varying conditions. These factors make the model inter-dependencies challenging to interpret without sophisticated mathematical tools. How do we enable users to explore complex network interactions in an RNN and directly connect these abstract representations to human understandable inputs?

In this paper, we focus on visual analysis to allow experimenters to explore and form hypotheses about RNN hidden state dynamics in their

Designing bistable [2]rotaxanes for molecular electronic devices

BY WILLIAM R. DICHTEL^{1,2}, JAMES R. HEATH^{2,*}
AND J. FRASER STODDART^{1,*}

¹*California NanoSystems Institute and Department of Chemistry and Biochemistry, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095, USA*

²*Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA*

The development of molecular electronic components has been accelerated by the promise of increased circuit densities and reduced power consumption. Bistable rotaxanes have been assembled into nanowire crossbar devices, where they may be switched between low- and high-conductivity states, forming the basis for a molecular memory. These memory devices have been scaled to densities of 10^{11} bits cm^{-2} , the 2020 node for memory of the International Technology Roadmap for Semiconductors. Investigations of the kinetics and thermodynamics associated with the electromechanical switching processes of several bistable [2]rotaxane derivatives in solution, self-assembled monolayers on gold, polymer electrolyte gels and in molecular switch tunnel junction devices are consistent with a single, universal switching mechanism whose speed is dependent largely on the environment, as well as on the structure of the switching molecule. X-ray reflectometry studies of the bistable rotaxanes assembled into Langmuir monolayers also lend support to an oxidatively driven mechanical switching process. Structural information obtained from Fourier transform reflection absorption infrared spectroscopy of rotaxane monolayers taken before and after evaporation of a Ti top electrode confirmed that the functionality responsible for switching is not affected by the metal deposition process. All the considerable experimental data, taken together with detailed computational work, support the hypothesis that the tunnelling current hysteresis, which forms the basis of memory operation, is a direct result of the electromechanical switching of the bistable rotaxanes.

Keywords: supramolecular chemistry; molecular electronics; bistable rotaxanes; switching; molecular memory; nanofabrication

1. Interlocked molecules in molecular electronic devices

‘Bottom-up’ approaches to the fabrication of molecular electronic devices (MEDs) are likely to become increasingly important as a means of overcoming the fundamental limitations of ‘top-down’ lithographic techniques. In the context

* Authors and address for correspondence: California Nanosystems Institute and Department of Chemistry and Biochemistry, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095-1569, USA (heath@chem.ucla.edu; stoddart@chem.ucla.edu).

One contribution of 12 to a Discussion Meeting Issue ‘Supramolecular nanotechnology for organic electronics’.

Adaptive Neuron Apoptosis for Accelerating Deep Learning on Large Scale Systems

Charles Siegel, Jeff Daily, Abhinav Vishnu

Pacific Northwest National Laboratory

Richland, WA 99352

charles.siegel@pnnl.gov, jeff.daily@pnnl.gov, abhinav.vishnu.pnnl.gov

Abstract—We present novel techniques to accelerate the convergence of Deep Learning algorithms by conducting low overhead removal of redundant neurons – *apoptosis* of neurons – which do not contribute to model learning, during the training phase itself. We provide in-depth theoretical underpinnings of our heuristics (bounding accuracy loss and handling apoptosis of several neuron types), and present the methods to conduct adaptive neuron apoptosis. Specifically, we are able to improve the training time for several datasets by 2-3x, while reducing the number of parameters by up to 30x (4-5x on average) on datasets such as ImageNet classification. For the Higgs Boson dataset, our implementation improves the accuracy (measured by Area Under Curve (AUC)) for classification from 0.88/1 to 0.94/1, while reducing the number of parameters by 3x in comparison to existing literature. The proposed methods achieve a 2.44x speedup in comparison to the default (no apoptosis) algorithm.

I. INTRODUCTION

Deep Learning algorithms emulate computation structure of a brain by learning models using *neurons* and their interconnections (*synapses*, also known as *parameters/weights*) [1]. Using a cascade of neurons, Deep Learning algorithms are known to learn complex non-linear functions. These functions can be applied to both *supervised* (input dataset with ground truth labels) and *unsupervised* (input data with no labels) problems. Naturally, Deep Learning algorithms are being applied to several domains including Computer Vision [2], Speech Recognition [3], and High Energy Physics [4].

An important aspect of Deep Learning algorithms is the topology of a *Neural Network* (used interchangeably with Deep Learning with rest of the paper). A candidate topology may have a single input and an output layer, with possibly several *hidden layers*. Convolutional Neural Networks (CNN) – a class of Deep Learning algorithms – may have several *convolutional layers*, followed by several *fully-connected layers*. In practice, the neurons and synapses are implemented by using matrices, where each row/column represents a neuron and each element represents the strength (weight) of a synapse. The output of a neural network is the weight matrices, which may be used for Machine Learning tasks such as classification, or clustering.

Usually, a neural network topology is user-specified, which includes the number of hidden layers and number of neurons in each layer (an example is shown in Figure 1 (a)). *Deeper neural networks* (with more layers) are used for model generation from increasingly complex datasets, possibly learning

complex non-linear functions. Bigger networks – which have larger number of neurons per layer – may also be used in addition to deeper networks for this purpose.

However, deeper and/or bigger networks do not necessarily provide better models. With increasing number of network *parameters* (the overall number of synapses), the training time per epoch increases significantly [5], [6]. Deeper networks tend to suffer from problems such as vanishing gradients [7], where the weights of network parameters change slowly. In addition, deeper networks are known to cause *overfitting* – a scenario in which the model has learned very well from the training set, but does not generalize well on new samples. This scenario generally occurs when a neural network has learned a significantly more complex function than implied by the training set. A few possible solutions such as Dropout [8] exist – but they do not reduce the overall time to solution. Lastly, larger number of parameters have prohibitive storage and computational requirements during the testing phase (when the model is applied on new data) – which is problematic for deployment on power/memory constrained devices.

A possible solution to this problem is to prune the neural network. Usually, this is conducted after the training is completed by removing *unnecessary* weights (and possibly) neurons [9], [10], [11]. After pruning, the network is re-trained to stabilize the parameters [10] (The scenario is shown on the right of Figure 1(b)). Usually pruning implies that unnecessary parameters (and in a few cases neurons) may be removed, without accuracy loss.

However, there are several shortcomings of existing approaches: 1) Re-training is a time-consuming process. As an example, Han *et al.* report a slowdown of up to 2.5x when re-training after pruning [10]. This is especially problematic for large datasets, where increasing the training time is unattractive. 2) Another problem with the current approaches is the lack of theoretical underpinnings for network pruning. Usually, this is not required for state of the art networks, because they use the final network as starting point for re-training. After each sub-step of pruning and re-training, the accuracy is compared against the reference network – to ensure no empirical loss of accuracy. However, this approach is not always efficient, since this exploration is expensive in time due to several re-training steps.

Our objective in this paper is to remove unnecessary neurons (and hence synapses) for Deep Learning algorithms by

Quantum state readout of individual quantum dots by electrostatic force detection

Yoichi Miyahara, Antoine Roy-Gobeil and Peter Grutter

Department of Physics, McGill University, 3600 rue University, Montreal, H3A 2T8, Quebec, Canada

E-mail: yoichi.miyahara@mcgill.ca

Abstract. Electric charge detection by atomic force microscopy (AFM) with single-electron resolution (*e*-EFM) is a promising way to investigate the electronic level structure of individual quantum dots (QD). The oscillating AFM tip modulates the energy of the QDs, causing single electrons to tunnel between QDs and an electrode. The resulting oscillating electrostatic force changes the resonant frequency and damping of the AFM cantilever, enabling electrometry with a single-electron sensitivity. Quantitative electronic level spectroscopy is possible by sweeping the bias voltage. Charge stability diagram can be obtained by scanning the AFM tip around the QD. *e*-EFM technique enables to investigate individual colloidal nanoparticles and self-assembled QDs without nanoscale electrodes. *e*-EFM is a quantum electromechanical system where the back-action of a tunneling electron is detected by AFM; it can also be considered as a mechanical analog of admittance spectroscopy with a radio frequency resonator, which is emerging as a promising tool for quantum state readout for quantum computing. In combination with the topography imaging capability of the AFM, *e*-EFM is a powerful tool for investigating new nanoscale material systems which can be used as quantum bits.

1. Introduction

Quantum dots (QD) are one of the most interesting and important entities in nanoscience and nanotechnology. QDs are often called artificial atoms as their electronic states become discrete just like those in atoms because of quantum confinement in three dimensions. The emergence of atom-like discrete energy levels leads to particular optical and electronic properties of QDs. Unlike real atoms, one can engineer the size and shape of the QDs, leading to their tunable optical and electronic properties. Therefore, understanding the energy level structure and its relation to the shape and size of the QD is a key to understand the properties of QDs, which remains to be elucidated. For this end, spectroscopic measurement on individual QDs is essential. QDs have been attracting much attention in the field of quantum information processing as their atom-like discrete energy levels can host charge or spin qubit and their quantum state can be read out by probing the charge state of the QD [1]. Many studies have already been reported [2, 3, 4].

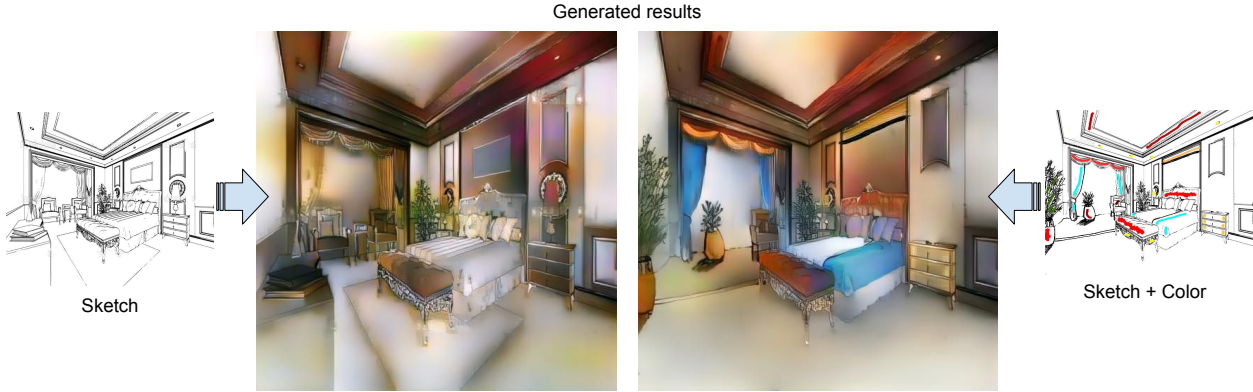


Figure 1. A user can sketch and scribble colors to control deep image synthesis. On the left is an image generated from a hand drawn sketch. On the right several objects have been deleted from the sketch, a vase has been added, and the color of various scene elements has been constrained by sparse color strokes. For best resolution and additional results, see <http://scribbler.eye.gatech.edu>

Scribbler: Controlling Deep Image Synthesis with Sketch and Color

Patsorn Sangkloy¹, Jingwan Lu², Chen Fang², Fisher Yu³, James Hays¹

¹Georgia Institute of Technology ²Adobe Research ³Princeton University

Abstract

Recently, there have been several promising methods to generate realistic imagery from deep convolutional networks. These methods sidestep the traditional computer graphics rendering pipeline and instead generate imagery at the pixel level by learning from large collections of photos (e.g. faces or bedrooms). However, these methods are of limited utility because it is difficult for a user to control what the network produces. In this paper, we propose a deep adversarial image synthesis architecture that is conditioned on sketched boundaries and sparse color strokes to generate realistic cars, bedrooms, or faces. We demonstrate a sketch based image synthesis system which allows users to scribble over the sketch to indicate preferred color for objects. Our network can then generate convincing images that satisfy both the color and the sketch constraints of user. The network is feed-forward which allows users to see the effect of their edits in real time. We compare to recent work on sketch to image synthesis and show that our approach can generate more realistic, more diverse, and more controllable outputs. The architecture is also effective at user-guided colorization of grayscale images.

1. Introduction

Recently, numerous image synthesis methods built on neural networks have emerged [40, 24, 12, 36, 21, 13]. These methods can generate detailed and diverse (if not quite photorealistic) images in many domains. However, it is still unclear how to *control* these powerful new tools. How can we enable everyday users (non-artists) to harness the power of deep image synthesis methods and produce realistic imagery? Several recent methods have explored controllable deep synthesis [8, 49, 54, 14, 55, 18, 45] and we focus on two complementary forms of control – sketches and color strokes. Sketches are a compelling form of control because anyone can draw (potentially very badly) and because it is easy to edit sketches, e.g. to remove or add objects, whereas the equivalent operations in the image domain require artistic expertise. Color is a compelling form of control because many sketches or grayscale scenes are fundamentally ambiguous with respect to color [53], but it is easy for a user to intervene, e.g. to scribble that drapes should be blue and the valance should be red (Figure 1). Both forms of control are relatively sparse and require a deep network to synthesize image detail beyond what is contained in the input. The deep network must also implicitly learn a significant amount of high-level image understanding, e.g. what colors are allowable for particular objects, the boundaries

StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks

Han Zhang¹, Tao Xu², Hongsheng Li³,
Shaoting Zhang⁴, Xiaogang Wang³, Xiaolei Huang², Dimitris Metaxas¹

¹Rutgers University ²Lehigh University ³The Chinese University of Hong Kong ⁴Baidu Research
{han.zhang, dnm}@cs.rutgers.edu, {tax313, xih206}@lehigh.edu
{hsli, xgwang}@ee.cuhk.edu.hk, zhangshaoting@baidu.com

Abstract

Synthesizing high-quality images from text descriptions is a challenging problem in computer vision and has many practical applications. Samples generated by existing text-to-image approaches can roughly reflect the meaning of the given descriptions, but they fail to contain necessary details and vivid object parts. In this paper, we propose Stacked Generative Adversarial Networks (StackGAN) to generate 256×256 photo-realistic images conditioned on text descriptions. We decompose the hard problem into more manageable sub-problems through a sketch-refinement process. The Stage-I GAN sketches the primitive shape and colors of the object based on the given text description, yielding Stage-I low-resolution images. The Stage-II GAN takes Stage-I results and text descriptions as inputs, and generates high-resolution images with photo-realistic details. It is able to rectify defects in Stage-I results and add compelling details with the refinement process. To improve the diversity of the synthesized images and stabilize the training of the conditional-GAN, we introduce a novel Conditioning Augmentation technique that encourages smoothness in the latent conditioning manifold. Extensive experiments and comparisons with state-of-the-arts on benchmark datasets demonstrate that the proposed method achieves significant improvements on generating photo-realistic images conditioned on text descriptions.

1. Introduction

Generating photo-realistic images from text is an important problem and has tremendous applications, including photo-editing, computer-aided design, etc. Recently, Generative Adversarial Networks (GAN) [8, 5, 23] have shown promising results in synthesizing real-world images. Conditioned on given text descriptions, conditional-

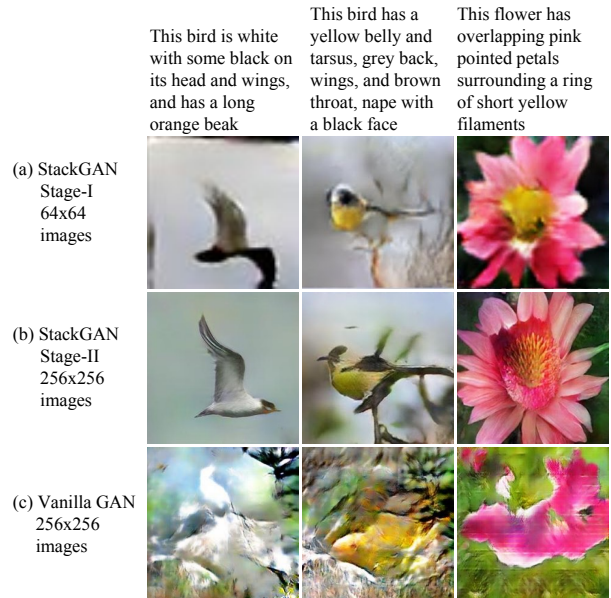


Figure 1. Comparison of the proposed StackGAN and a vanilla one-stage GAN for generating 256×256 images. (a) Given text descriptions, Stage-I of StackGAN sketches rough shapes and basic colors of objects, yielding low-resolution images. (b) Stage-II of StackGAN takes Stage-I results and text descriptions as inputs, and generates high-resolution images with photo-realistic details. (c) Results by a vanilla 256×256 GAN which simply adds more upsampling layers to state-of-the-art GAN-INT-CLS [26]. It is unable to generate any plausible images of 256×256 resolution.

GANs [26, 24] are able to generate images that are highly related to the text meanings.

However, it is very difficult to train GAN to generate high-resolution photo-realistic images from text descriptions. Simply adding more upsampling layers in state-of-the-art GAN models for generating high-resolution (e.g., 256×256) images generally results in training instability

Imaging action potential in single mammalian neurons by tracking the accompanying sub-nanometer mechanical motion

Yunze Yang¹, Xian-Wei Liu^{1,2,*}, Hui Yu¹, Yan Guan^{1,4}, Shaopeng Wang¹, and Nongjian Tao^{1,3,*}

¹Center for Biosensors and Bioelectronics, Biodesign Institute, Arizona State University, Tempe, Arizona 85287, USA.

²CAS Key Laboratory of Urban Pollutant Conversion, School of Chemistry and Materials Science, University of Science & Technology of China, Hefei 230026, China.

³School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona 85287, USA.

*Corresponding authors: xianweiliu@ustc.edu.cn (XW Liu); njtao@asu.edu (NJ Tao)

Abstract

Action potentials in neurons have been studied traditionally by the patch clamp and more recently by the fluorescence detection methods. Here we describe a label-free optical imaging method that can measure mechanical motion in single cells with sub-nanometer detection limit and sub-millisecond temporal resolution. Using the method, we have observed sub-nanometer mechanical motion accompanying the action potential in single mammalian neurons. The shape and width of the transient displacement are similar to those of the electrically recorded action potential, but the amplitude varies from neuron to neuron, and from one region of a neuron to another, ranging from 0.2 - 0.4 nm. The work indicates that action potentials may be studied non-invasively in single mammalian neurons by label-free imaging of the accompanying sub-nanometer mechanical motion.

Keywords: action potential, sub-nanometer, mechanical motion, label-free imaging, tracking

Software Engineering at Google

31 Jan 2017

Fergus Henderson

<fergus@google.com> (work) or
<fergus.henderson@gmail.com> (personal)

Abstract

We catalog and describe Google's key software engineering practices.

Biography

Fergus Henderson has been a software engineer at Google for over 10 years. He started programming as a kid in 1979, and went on to academic research in programming language design and implementation. With his PhD supervisor, he co-founded a research group at the University of Melbourne that developed the programming language Mercury. He has been a program committee member for eight international conferences, and has released over 500,000 lines of open-source code. He was a former moderator of the Usenet newsgroup comp.std.c++ and was an officially accredited "Technical Expert" to the ISO C and C++ committees. He has over 15 years of commercial software industry experience. At Google, he was one of the original developers of Blaze, a build tool now used across Google, and worked on the server-side software behind speech recognition and voice actions (before Siri!) and speech synthesis. He currently manages Google's text-to-speech engineering team, but still writes and reviews plenty of code. Software that he has written is installed on over a billion devices, and gets used over a billion times per day.

Crossing Nets: Combining GANs and VAEs with a Shared Latent Space for Hand Pose Estimation

Chengde Wan¹, Thomas Probst¹, Luc Van Gool^{1,3}, and Angela Yao²

¹ETH Zürich ²University of Bonn ³KU Leuven

Abstract

State-of-the-art methods for 3D hand pose estimation from depth images require large amounts of annotated training data. We propose to model the statistical relationships of 3D hand poses and corresponding depth images using two deep generative models with a shared latent space. By design, our architecture allows for learning from unlabeled image data in a semi-supervised manner. Assuming a one-to-one mapping between a pose and a depth map, any given point in the shared latent space can be projected into both a hand pose and a corresponding depth map. Regressing the hand pose can then be done by learning a discriminator to estimate the posterior of the latent pose given some depth maps. To improve generalization and to better exploit unlabeled depth maps, we jointly train a generator and a discriminator. At each iteration, the generator is updated with the back-propagated gradient from the discriminator to synthesize realistic depth maps of the articulated hand, while the discriminator benefits from an augmented training set of synthesized and unlabeled samples. The proposed discriminator network architecture is highly efficient and runs at 90FPS on the CPU with accuracies comparable or better than state-of-art on 3 publicly available benchmarks.

1. Introduction

We address the problem of estimating 3D hand pose from single depth images. Accurate estimation of the 3D pose in real-time has many challenges, including the presence of local self-similarity and self-occlusions. Since the availability of low-cost depth sensors, the progress made in developing fast and accurate hand trackers have relied heavily on having a large corpus of depth images annotated with hand joints. This is especially true for the recent success of deep learning-based methods [48, 22, 23, 34, 9, 53, 47, 50] which are all fully-supervised.

Accurately annotating 3D hand joints on a depth map is



Figure 1. **Random walk in the learned shared latent space.** A set of points is sampled on the connecting line between two points in the shared latent space. The pose and corresponding depth map are then reconstructed through our network. Our method generates meaningful and realistic interpolations in both pose and appearance space.

both difficult and time-consuming. While it is possible to synthesize data with physical renderers, there are usually discrepancies between the real and synthesized data. Generated hand poses are not always natural nor reflective of poses seen in real-life applications. More importantly, it is very difficult to accurately model and render depth sensor noise in a realistic way.

On the other “hand”, it is very simple to collect an unlabeled dataset of real human hands with standard consumer depth cameras. This begs the question: how can one use these unlabeled samples for training? To date, there has been virtually no work presented on semi-supervised learning for hand pose estimation. The one notable exception [43] is a discriminative approach using transductive random forests and largely ignores high-order pixel correlations of unlabeled depth maps.

Previous works from neuroscience [31], robotics [1] and hand motion capture [16] have demonstrated that hand movements exhibit strong correlations between joints. We therefore conclude that the space of realistic hand poses can be represented by a manifold in a lower-dimensional subspace. We further intuit that depth maps of the hand can be similarly encoded in a low-dimensional manifold, and be

One-Step Time-Dependent Future Video Frame Prediction with a Convolutional Encoder-Decoder Neural Network

Vedran Vukotić^{1,2,3}, Silvia-Laura Pintea², Christian Raymond^{1,3},
Guillaume Gravier^{1,4}, and Jan C. van Gemert²

¹ INRIA/IRISA Rennes, Rennes, France

² TUDelft, Delft, The Netherlands

³ INSA Rennes, Rennes, France

⁴ CNRS, France

{vedran.vukotic,christian.raymond,guillaume.gravier}@irisa.fr
{s.l.pintea,j.c.vangemert}@tudelft.nl

Abstract. There is an inherent need for autonomous cars, drones, and other robots to have a notion of how their environment behaves and to anticipate changes in the near future. In this work, we focus on anticipating future appearance given the current frame of a video. Existing work focuses on either predicting the future appearance as the next frame of a video, or predicting future motion as optical flow or motion trajectories starting from a single video frame. This work stretches the ability of CNNs (Convolutional Neural Networks) to predict an anticipation of appearance at an arbitrarily given future time, not necessarily the next video frame. We condition our predicted future appearance on a continuous time variable that allows us to anticipate future frames at a given temporal distance, directly from the input video frame. We show that CNNs can learn an intrinsic representation of typical appearance changes over time and successfully generate realistic predictions at a deliberate time difference in the near future.

Keywords: action forecasting, future video frame prediction, appearance prediction, scene understanding, generative models, CNNs.

1 Introduction

For machines to successfully interact in the real world, anticipating actions and events and planning accordingly, is essential. This is a difficult task, despite the recent advances in deep and reinforcement learning, due to the demand of large annotated datasets. If we limit our task to anticipating future appearance, annotations are not needed anymore. Therefore, machines have a slight advantage, as they can employ the vast collection of unlabeled videos available, which is perfectly suited for unsupervised learning methods. To anticipate future appearance based on current visual information, a machine needs to successfully be able to

Neural Episodic Control

Alexander Pritzel
 Benigno Uria
 Sriram Srinivasan
 Adrià Puigdomènech
 Oriol Vinyals
 Demis Hassabis
 Daan Wierstra
 Charles Blundell
 DeepMind, London UK

APRITZEL@GOOGLE.COM
 BURIA@GOOGLE.COM
 SRSRINIVASAN@GOOGLE.COM
 ADRIAP@GOOGLE.COM
 VINYALS@GOOGLE.COM
 DEMISHASSABIS@GOOGLE.COM
 WIERSTRA@GOOGLE.COM
 CBLUNDELL@GOOGLE.COM

Abstract

Deep reinforcement learning methods attain super-human performance in a wide range of environments. Such methods are grossly inefficient, often taking orders of magnitudes more data than humans to achieve reasonable performance. We propose Neural Episodic Control: a deep reinforcement learning agent that is able to rapidly assimilate new experiences and act upon them. Our agent uses a semi-tabular representation of the value function: a buffer of past experience containing slowly changing state representations and rapidly updated estimates of the value function. We show across a wide range of environments that our agent learns significantly faster than other state-of-the-art, general purpose deep reinforcement learning agents.

1. Introduction

Deep reinforcement learning agents have achieved state-of-the-art results in a variety of complex environments (Mnih et al., 2015; 2016), often surpassing human performance (Silver et al., 2016). Although the final performance of these agents is impressive, these techniques usually require several orders of magnitude more interactions with their environment than a human in order to reach an equivalent level of expected performance. For example, in the Atari 2600 set of environments (Bellemare et al., 2013), deep Q-networks (Mnih et al., 2016) require more than 200 hours of gameplay in order to achieve scores similar to those a human player achieves after two hours (Lake et al., 2016).

The glacial learning speed of deep reinforcement learning has several plausible explanations and in this work we focus on addressing these:

1. Stochastic gradient descent optimisation requires the use of small learning rates. Due to the global approximation nature of neural networks, high learning rates cause catastrophic interference (McCloskey & Cohen, 1989). Low learning rates mean that experience can only be incorporated into a neural network slowly.
2. Environments with a sparse reward signal can be difficult for a neural network to model as there may be very few instances where the reward is non-zero. This can be viewed as a form of class imbalance where low-reward samples outnumber high-reward samples by an unknown number. Consequently, the neural network disproportionately underperforms at predicting larger rewards, making it difficult for an agent to take the most rewarding actions.
3. Reward signal propagation by value-bootstrapping techniques, such as Q-learning, results in reward information being propagated one step at a time through the history of previous interactions with the environment. This can be fairly efficient if updates happen in reverse order in which the transitions occur. However, in order to train on uncorrelated minibatches DQN-style, algorithms train on randomly selected transitions, and, in order to further stabilise training, require the use of a slowly updating *target network* further slowing down reward propagation.

In this work we shall focus on addressing the three concerns listed above; we must note, however, that other recent advances in exploration (Osband et al., 2016), hierarchical reinforcement learning (Vezhnevets et al., 2016) and transfer learning (Rusu et al., 2016; Fernando et al., 2017) also make substantial contributions to improving data efficiency in deep reinforcement learning over baseline agents.

In this paper we propose Neural Episodic Control (NEC), a method which tackles the limitations of deep reinforcement learning listed above and demonstrates dramatic im-

End-to-End Prediction of Buffer Overruns from Raw Source Code via Neural Memory Networks

Min-je Choi and Sehun Jeong and Hakjoo Oh and Jaegul Choo

Department of Computer Science and Engineering

Korea University, Seoul

{devnote5676, gifaranga, hakjoo_oh, jchoo}@korea.ac.kr

Abstract

Detecting buffer overruns from a source code is one of the most common and yet challenging tasks in program analysis. Current approaches based on rigid rules and handcrafted features are limited in terms of flexible applicability and robustness due to diverse bug patterns and characteristics existing in sophisticated real-world software programs. In this paper, we propose a novel, data-driven approach that is completely end-to-end without requiring any hand-crafted features, thus free from any program language-specific structural limitations. In particular, our approach leverages a recently proposed neural network model called memory networks that have shown the state-of-the-art performances mainly in question-answering tasks. Our experimental results using source code samples demonstrate that our proposed model is capable of accurately detecting different types of buffer overruns. We also present in-depth analyses on how a memory network can learn to understand the semantics in programming languages solely from raw source codes, such as tracing variables of interest, identifying numerical values, and performing their quantitative comparisons.

1 Introduction

Detecting potential bugs in software programs has long been a challenge ever since computers were first introduced. To tackle this problem, researchers in the domain of programming languages developed various techniques called static analysis, which tries to find potential bugs in source codes without having to execute them based on a solid mathematical framework [Cousot and Cousot, 1977]. However, designing a static analyzer is tightly coupled with a particular programming language, and it is mainly based on a rigid set of rules designed by a few experts, considering numerous types of possible program states and bug cases. Thus, even with its slight syntax changes frequently found in real-world settings, e.g., several variants of ANSI C languages, a significant amount of engineering effort is required to make a previously designed analyzer applicable to the other similar languages.

To overcome these limitations, one can suggest data-driven, machine learning-based approaches as the rapid growth of deep neural networks in natural language processing has proved its effectiveness in solving similar problems such as defect predictions. Studies show that deep convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are capable of learning patterns or structures within text corpora such as source codes, so they can be applied to programming language tasks such as bug localization [Lam *et al.*, 2016], syntax error correction [Bhatia and Singh, 2016; Pu *et al.*, 2016], and code suggestion [White *et al.*, 2015].

Despite their impressive performances at detecting syntax-level bugs and code patterns, deep neural networks have shown less success at understanding how data values are transferred and used within source codes. This semantic level of understanding requires not only knowledge on the overall structure but also the capability to track the data values stored in different variables and methods. Although the aforementioned deep learning models may learn patterns and structures, they cannot keep track of how values are changed. This restriction greatly limits their usefulness in program analysis since run-time bugs and errors are usually much more difficult to detect and thus are often treated with greater importance.

In response, we introduce a new deep learning model with the potential of overcoming such difficulties: memory networks [Weston *et al.*, 2015b; Sukhbaatar *et al.*, 2015]. Memory networks are best described as neural networks with external memory ‘slots’ to store previously introduced information for future uses. Given a question, it accesses relevant memory slots via an attention mechanism and combines the values of the accessed slots to reach an answer. While long short-term memories (LSTMs) and earlier models also have external memories, theirs tend to evolve as longer sequences of information are fed in to the network, thus failing to fully preserve and represent information introduced at earlier stages. Memory networks on the other hand can preserve the given information even during long sequences.

This unique aspect of memory networks makes it and its variant models [Kumar *et al.*, 2016; Henaff *et al.*, 2016] perform exceptionally well at question answering tasks, e.g., the Facebook bAbI task [Weston *et al.*, 2015a], a widely-used QA benchmark set. The structure of these tasks comprises a story, a query, and an answer, from which a model has to predict the correct answer to the task mentioned in the query

I2T2I: LEARNING TEXT TO IMAGE SYNTHESIS WITH TEXTUAL DATA AUGMENTATION

Hao Dong, Jingqing Zhang, Douglas McIlwraith, Yike Guo

Data Science Institute, Imperial College London

ABSTRACT

Translating information between text and image is a fundamental problem in artificial intelligence that connects natural language processing and computer vision. In the past few years, performance in image caption generation has seen significant improvement through the adoption of recurrent neural networks (RNN). Meanwhile, text-to-image generation begun to generate plausible images using datasets of specific categories like birds and flowers. We’ve even seen image generation from multi-category datasets such as the Microsoft Common Objects in Context (MSCOCO) through the use of generative adversarial networks (GANs). Synthesizing objects with a complex shape, however, is still challenging. For example, animals and humans have many degrees of freedom, which means that they can take on many complex shapes. We propose a new training method called Image-Text-Image (I2T2I) which integrates text-to-image and image-to-text (image captioning) synthesis to improve the performance of text-to-image synthesis. We demonstrate that I2T2I can generate better multi-categories images using MSCOCO than the state-of-the-art. We also demonstrate that I2T2I can achieve transfer learning by using a pre-trained image captioning module to generate human images on the MPII Human Pose dataset (MHP) without using sentence annotation.

Index Terms— Deep learning, GAN, Image Synthesis

1. INTRODUCTION

As an image can be described in different ways and a caption can also be translated into different images, there is no bijection between images and captions. For example, in the sentence “a group of giraffes standing next to each other”, the number of giraffes, the background and many other details are uncertain. It is an open issue of text-to-image mapping that the distribution of images conditioned on a sentence is highly multi-modal. In the past few years, we’ve witnessed a breakthrough in the application of recurrent neural networks (RNN) to generating textual descriptions conditioned on images [1, 2], with Xu *et al.* showing that the multi-modality problem can be decomposed sequentially [3]. However, the lack of datasets with diversity descriptions of images limits the performance of text-to-image synthesis on multi-categories dataset like MSCOCO [4]. Therefore, the problem

of text-to-image synthesis is still far from being solved.



Fig. 1: Comparing text-to-image synthesis with and without textual data augmentation. I2T2I synthesizes the human pose and bus color better.

Recently, both fractionally-strided convolutional and convolutional neural networks (CNN) have shown promise for image generation. PixelCNN [5] learns a conditional image generator which is able to generate realistic scenes representing different objects and portraits of people. Generative adversarial networks (GAN) [6], specifically DCGAN [7] have been applied to a variety of datasets and shown plausible results. Odena *et al.* [8] has shown the capability of GANs to generate high-quality images conditioned on classes. To solve the multi-modality problem for text-to-image synthesis, Reed *et al.* [9] proposed GAN-CLS which bridges advances of DCGAN and an RNN encoder to generate images from an latent variables and embedding of image descriptions. Besides these, recent studies learned to generate images by conditions [10, 11]. However, GAN-CLS fails to generate plausible images on more complicated and changeable realistic scenes such as those illustrating human activities.

In this paper, we illustrate that sentence embedding should be able to capture details from various descriptions (one image can be described by plenty of sentences). This determines the robustness of understanding image detail. The main contribution of our work is a new training method called Image-Text-Image (I2T2I), which applies textual data augmentation to help text-to-image synthesis learn more descriptions of each image. I2T2I is composed of three modules: 1) the

A Proposal for an Electron-Transfer Mechanism of Avian Magnetoreception

Shao-Qing Zhang^{1,2†}

¹ Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6396.

² Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California at San Francisco, San Francisco, CA 94158-9001.

[†] zhangsh@sas.upenn.edu

Abstract

In spite of many years of research, the mechanism of avian magnetoreception remains a mystery due to its seemingly insurmountable intricacies. Recently Xie and colleagues proposed that IscA1 can act as a protein biocompass due to the measured intrinsic ferromagnetism, and thus named it MagR. However, Meister's calculations showed that the interaction energy of the magnetic moment of IscA1 with Earth's magnetic field is five magnitudes smaller than thermal fluctuation at room temperature. The other long-proposed compass protein is cryptochrome (Cry) with a mechanism of forming singlet-triplet radical pairs. However, this sensory mechanism still has no inferable information transmission routes. We propose a magnetoreception mechanism involving both the Cry and IscA1 proteins, through which photoinduced electrons are transported to redox-regulated ion channels to provoke neuronal responses. The structural features of the Cry-IscA1 complex that make it suitable for long-range electron transfer are discussed and how the magnetic effect leads to neuronal activity is described.

Keywords: magnetic compass, avian magnetoreception, quantum biology, long-range electron transfer, coherence, electron tunnelling, radical pairs

Auto-painter: Cartoon Image Generation from Sketch by Using Conditional Generative Adversarial Networks

Yifan Liu^{1,2}, Zengchang Qin¹, Zhenbo Luo², and Hua Wang²

¹Intelligent Computing and Machine Learning Lab, School of ASEE
Beihang University, Beijing 100191, China

² Samsung R&D Institute China Beijing
18F TaiTangGong Plaza, Beijing, 100028, China
{yifan_liu, zcqin}@buaa.edu.cn, {zb.luo, hua00.wang}@samsung.com

Abstract. Recently, realistic image generation using deep neural networks has become a hot topic in machine learning and computer vision. Images can be generated at the pixel level by learning from a large collection of images. Learning to generate colorful cartoon images from black-and-white sketches is not only an interesting research problem, but also a potential application in digital entertainment. In this paper, we investigate the sketch-to-image synthesis problem by using conditional generative adversarial networks (cGAN). We propose the auto-painter model which can automatically generate compatible colors for a sketch. The new model is not only capable of painting hand-draw sketch with proper colors, but also allowing users to indicate preferred colors. Experimental results on two sketch datasets show that the auto-painter performs better than existing image-to-image methods.

1 Introduction

Human beings possess a great cognitive capability of comprehend black-and-white cartoon sketches. Our mind can create realistic colorful images when see black-and-white cartoons. It needs a great artistic talent to choose appropriate colors, also with proper changes in light and shade to create vivid cartoon images. It is not easy for ordinary people to do so. How to automatically paint the sketch into colorful cartoon images is a useful application for digital entertainment. In this work, we are interested in solving this problem by employing deep neural networks with designed constraints to transfer the line draft into specific cartoon style. Practically, the new model can make up for ordinary people's artistic talent and even inspire the artists to create new images. Ideally, readers may have the freedom to choose to generate different style of cartoons based on their own tastes of colors.

Cartoon image generation from sketches can be regarded as an image synthesis problem. Previously, many non-parametric models were proposed [8,1,6] by matching the sketch to a database of existing image fragments. Recently, numerous image synthesis methods based on deep neural networks have emerged

Real-Time User-Guided Image Colorization with Learned Deep Priors

RICHARD ZHANG*, University of California, Berkeley

JUN-YAN ZHU*, University of California, Berkeley

PHILLIP ISOLA, University of California, Berkeley

XINYANG GENG, University of California, Berkeley

ANGELA S. LIN, University of California, Berkeley

TIANHE YU, University of California, Berkeley

ALEXEI A. EFROS, University of California, Berkeley



Fig. 1. Our proposed method colorizes a grayscale image (left), guided by sparse user inputs (second), in real-time, providing the capability for quickly generating multiple plausible colorizations (middle to right). Photograph of *Migrant Mother* by Dorothea Lange, 1936 (Public Domain).

We propose a deep learning approach for user-guided image colorization. The system directly maps a grayscale image, along with sparse, local user “hints” to an output colorization with a Convolutional Neural Network (CNN). Rather than using hand-defined rules, the network propagates user edits by fusing low-level cues along with high-level semantic information, *learned from large-scale data*. We train on a million images, with simulated user inputs. To guide the user towards efficient input selection, the system recommends likely colors based on the input image and current user inputs. The colorization is performed in a single feed-forward pass, enabling real-time use. Even with randomly simulated user inputs, we show that the proposed system helps novice users quickly create realistic colorizations, and offers large improvements in colorization quality with just a minute of use. In addition, we demonstrate that the framework can incorporate other user “hints” to the desired colorization, showing an application to color histogram transfer. Our code and models are available at <https://richzhang.github.io/ideepcolor>.

CCS Concepts: • **Computing methodologies** → **Image manipulation**; *Computational photography*; *Neural networks*;

Additional Key Words and Phrases: Colorization, Edit propagation, Interactive colorization, Deep learning, Vision for graphics

1 INTRODUCTION

There is something uniquely and powerfully satisfying about the simple act of adding color to black and white imagery. Whether as a way of rekindling old, dormant memories or expressing artistic creativity, people continue to be fascinated by colorization. From remastering classic black and white films, to the enduring popularity

of coloring books for all ages, to the surprising enthusiasm for various (often not very good) automatic colorization bots online¹, this topic continues to fascinate the public.

In computer graphics, two broad approaches to image colorization exist: user-guided edit propagation and data-driven automatic colorization. In the first paradigm, popularized by the seminal work of Levin et al. (2004), a user draws colored strokes over a grayscale image. An optimization procedure then generates a colorized image that matches the user’s scribbles, while also adhering to hand-defined image priors, such as piecewise smoothness. These methods can achieve impressive results but often require intensive user interaction (sometimes over fifty strokes), as each differently colored image region must be explicitly indicated by the user. Because the system purely relies on user inputs for colors, even regions with little color uncertainty, such as green vegetation, need to be specified. Less obviously, even if a user knows what general color an object should take on, it can be surprisingly difficult to select the exact desired natural chrominance.

To address these limitations, researchers have also explored more data-driven colorization methods. These methods colorize a grayscale photo in one of two ways: either by matching it to an exemplar color image in a database and non-parametrically “stealing” colors from that photo, an idea going back to Image Analogies (Hertzmann et al. 2001), or by learning parametric mappings from grayscale to color from large-scale image data. The most recent methods in this paradigm proposed by Iizuka et al. (2016), Larsson et al. (2016),

2017. The definitive Version of Record was published in *ACM Transactions on Graphics*, <https://doi.org/http://dx.doi.org/10.1145/3072959.3073703>.

¹e.g., <http://demos.algorithmia.com/colorize-photos/>

* indicates equal contribution

When Will AI Exceed Human Performance?

Evidence from AI Experts

Katja Grace^{1,2}, John Salvatier², Allan Dafoe^{1,3}, Baobao Zhang³, and Owain Evans¹

¹*Future of Humanity Institute, Oxford University*

²*AI Impacts*

³*Department of Political Science, Yale University*

Abstract

Advances in artificial intelligence (AI) will transform modern life by reshaping transportation, health, science, finance, and the military [1, 2, 3]. To adapt public policy, we need to better anticipate these advances [4, 5]. Here we report the results from a large survey of machine learning researchers on their beliefs about progress in AI. Researchers predict AI will outperform humans in many activities in the next ten years, such as translating languages (by 2024), writing high-school essays (by 2026), driving a truck (by 2027), working in retail (by 2031), writing a bestselling book (by 2049), and working as a surgeon (by 2053). Researchers believe there is a 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years, with Asian respondents expecting these dates much sooner than North Americans. These results will inform discussion amongst researchers and policymakers about anticipating and managing trends in AI.

Introduction

Advances in artificial intelligence (AI) will have massive social consequences. Self-driving technology might replace millions of driving jobs over the coming decade. In addition to possible unemployment, the transition will bring new challenges, such as rebuilding infrastructure, protecting vehicle cyber-security, and adapting laws and regulations [5]. New challenges, both for AI developers and policy-makers, will also arise from applications in law enforcement, military technology, and marketing [6]. To prepare for these challenges, accurate forecasting of transformative AI would be invaluable.

Several sources provide objective evidence about future AI advances: trends in computing hardware [7], task performance [8], and the automation of labor [9]. The predictions of AI experts provide crucial additional information. We survey a larger and more representative sample of AI experts than any study to date [10, 11]. Our questions cover the timing of AI advances (including both practical applications of AI and the automation of various human jobs), as well as the social and ethical impacts of AI.

Survey Method

Our survey population was all researchers who published at the 2015 NIPS and ICML conferences (two of the premier venues for peer-reviewed research in machine learning). A total of 352 researchers responded to our survey invitation (21% of the 1634 authors we contacted). Our questions concerned the timing of specific AI capabilities (e.g. folding laundry, language translation), superiority at specific occupations (e.g. truck driver, surgeon), superiority over humans at all tasks, and the social impacts of advanced AI. See [Survey Content](#) for details.

Time Until Machines Outperform Humans

AI would have profound social consequences if all tasks were more cost effectively accomplished by machines. Our survey used the following definition:

“High-level machine intelligence” (HLMI) is achieved when unaided machines can accomplish every task better and more cheaply than human workers.

cGAN-based Manga Colorization Using a Single Training Image

Paulina Hensman*
Dept. Information and
Communication Eng.
University of Tokyo
phensman@kth.se

Kiyoharu Aizawa
Dept. Information and
Communication Eng.
University of Tokyo
aizawa@hal.t.u-tokyo.ac.jp

ABSTRACT

The Japanese comic format known as Manga is popular all over the world. It is traditionally produced in black and white, and colorization is time consuming and costly. Automatic colorization methods generally rely on greyscale values, which are not present in manga. Furthermore, due to copyright protection, colorized manga available for training is scarce.

We propose a manga colorization method based on conditional Generative Adversarial Networks (cGAN). Unlike previous cGAN approaches that use many hundreds or thousands of training images, our method requires only a single colorized reference image for training, avoiding the need of a large dataset.

Colorizing manga using cGANs can produce blurry results with artifacts, and the resolution is limited. We therefore also propose a method of segmentation and color-correction to mitigate these issues. The final results are sharp, clear, and in high resolution, and stay true to the character's original color scheme.

CCS CONCEPTS

• **Computing methodologies** → **Computer vision tasks; Adversarial learning; Image manipulation; Image segmentation; Neural networks; Interest point and salient region detections;**

KEYWORDS

Manga, colorization, segmentation, GAN, cGAN

1 INTRODUCTION

Manga is the main comic style used in Japan, and it is enjoyed by all ages across the world. It is produced at a rapid pace, and is traditionally drawn in black and white. Instead of classic shading, screentones are used to show light and texture.

A colorized manga is often more visually appealing, and with digital distribution methods color printing costs are no longer an issue. That being said, the actual colorization of the image still takes time and requires a skilled artist. Any way to automate this process could potentially be very helpful in production of colorized manga.

While a fully automatic method for manga colorization would be optimal, some user input will always be required for a correct output. Certain parts of a manga image such as skin color and background elements could possibly be correctly estimated and colorized by a



(a) Training image

(b) Our results

Figure 1: On the left, a colorized training image. On the right, images colorized by our method. ©Whomor

good method. However, hair color and other attributes that make a character unique can not possibly be inferred from a monochrome image. We therefore propose a method using a colorized image as reference to colorize other images of the same character.

The same characters tend to recur many times in a manga, especially if the manga goes on for several volumes. Additionally, uniforms such as school uniforms, team uniforms, and military uniforms are common. Requiring only a single reference image for each character or each uniform could thus save a lot of time when colorizing a manga. An example of a recurring character is shown in Fig. 1.

Comics are protected by copyright as works of art, making it difficult to gather large datasets of manga to use for training. The largest available dataset is Manga109[10][1], which unfortunately contains only monochrome images. This is a constraint for manga colorization methods, as methods that rely heavily on large amounts of training data can not be used.

We propose a solution for the task of manga colorization using conditional Generative Adversarial Networks (cGANs). Unlike previous cGAN literature, we demonstrate convincing results using a single training image of the character to be colorized. This eliminates the need for a large dataset of colorized manga unprotected

*Exchange from Royal Institute of Technology, Sweden

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
© 2017 Copyright held by the owner/author(s).

Exoplanet Transits as the Foundation of an Interstellar Communications Network

Duncan H. Forgan^{1,2}

July 13, 2017

¹SUPA, School of Physics and Astronomy, University of St Andrews

²St Andrews Centre for Exoplanet Science

Word Count: 5,366

Direct Correspondence to:

D.H. Forgan

Email: dhf3@st-andrews.ac.uk

Rocket Launching: A Universal and Efficient Framework for Training Well-performing Light Net

Guorui Zhou , Ying Fan , Runpeng Cui , Weijie Bian
Xiaoqiang Zhu and Kun Gai

Alibaba Inc.

Beijing, China

{guorui.xgr, fanying.fy, weijie.bwj, xiaoqiang.zxq}@alibaba-inc.com

cuirunpeng@gmail.com

jingshi.gk@taobao.com

Abstract

Models applied on real time response task, like click-through rate (CTR) prediction model, require high accuracy and rigorous response time. Therefore, top-performing deep models of high depth and complexity are not well suited for these applications with the limitations on the inference time. In order to get neural networks of better performance given the time limitations, we propose an unified framework that exploits a booster net to help train the lightweight net for prediction. We dub the whole process rocket launching, where the booster net is used to guide the learning of our light net throughout the whole training process. We analyze different loss functions aiming at pushing the light net to behave similarly to the booster net, and adopt the loss with best performance in our experiments. We use one technique called gradient block to improve the performance of light net and booster net further. Experiments on benchmark datasets and real-life industrial advertisement data show the effectiveness of our proposed method.

Introduction

Deep networks have achieved state-of-the-art results in many areas, such as computer vision (Huang et al. 2016) and nature language processing (Bahdanau, Cho, and Bengio 2014). From AlexNet (Krizhevsky, Sutskever, and Hinton 2012) to recently proposed DenseNet (Huang et al. 2016), better performances are accompanied with deeper and wider networks and more complex and adaptable structures. A more complex structure of neural networks means longer inference time, which is not tolerated in the real-life industrial environment. Networks mentioned above only consider the evaluation criterion of accuracy, while neglect the necessity of real-time response in industrial applications.

At the same time, some nets like DIN (Zhou et al. 2017) and wide & deep model (Cheng et al. 2016) get more and more attention. These net shares some characteristics: nets are shallow, layers are very simple and with less computation cost. In real-life industrial applications, e.g. online advertising systems, models have to make prediction of hundreds of advertisements for one user in several milliseconds, which restricts the complexity of model. Only simple and shallow structure meets the stringent response time requirements in real-life industry.

Accuracy and latency are the two points that we pay attention to. In general, there are two solutions to reduce runtime complexities while keeping a decent performance. Some works factorize or compress to directly simplify the computation, such as matrix SVD (Denton et al. 2014), MobileNet (Howard et al. 2017), and ShuffleNet (Zhang et al. 2017). Other approaches adopt the teacher-student strategy. They use light networks with fewer layers and parameters to decrease the inference time, while the light nets are trained helped by a complicated teacher network that trained in advance, like knowledge distillation (Hinton, Vinyals, and Dean 2015) and FitNet (Romero et al. 2014). These teacher-student methods decrease the runtime complexities, and can be further combined with approaches of the first category. In this work, we propose a novel unified framework to train decent small networks, motivated by the potential of teacher-student methods.

In this work, we develop a novel network training process dubbed rocket launching. The light net is the target network for inference, the booster relates to the deeper and more complex network from the architecture. Both the light and the booster net compose the architecture of rocket network. At the training stage, the light and booster networks are trained simultaneously on the same task. Besides, the light net also keeps getting knowledge learned by the booster through the optimization of the hint loss, which is included in the objective function to make both nets have similar behaviour during training. The booster guides the optimization of the target light network all along the training process. At the inference stage, only the trained light network is used. Different from previous teacher-student methods (Hinton, Vinyals, and Dean 2015; Romero et al. 2014), we make the light model share some lower layers with the cumbersome one and train them simultaneously.

In this paper, we propose a universal method aiming to obtain a well-behaved light net considering limitations on inference time. Our method is suitable to many different network structures. In brief, our contributions can be summarized as follows:

- We propose a novel universal training process called rocket launching, which makes use of the booster net to supervise the learning of the light network through whole training process. We show that a light model can be trained to perform close to deeper and more complex

DolphinAttack: Inaudible Voice Commands

Guoming Zhang*
Zhejiang University
realzgm@zju.edu.cn

Chen Yan*
Zhejiang University
yanchen@zju.edu.cn

Xiaoyu Ji†
Zhejiang University
xji@zju.edu.cn

Tianchen Zhang
Zhejiang University
tianchen-zhang@zju.edu.cn

Taimin Zhang
Zhejiang University
ztm1992fly@zju.edu.cn

Wenyuan Xu†
Zhejiang University
xuwenyuan@zju.edu.cn

ABSTRACT

Speech recognition (SR) systems such as Siri or Google Now have become an increasingly popular human-computer interaction method, and have turned various systems into voice controllable systems (VCS). Prior work on attacking VCS shows that the hidden voice commands that are incomprehensible to people can control the systems. Hidden voice commands, though ‘hidden’, are nonetheless audible. In this work, we design a completely inaudible attack, DolphinAttack, that modulates voice commands on ultrasonic carriers (e.g., $f > 20$ kHz) to achieve inaudibility. By leveraging the nonlinearity of the microphone circuits, the modulated low-frequency audio commands can be successfully demodulated, recovered, and more importantly interpreted by the speech recognition systems. We validate DolphinAttack on popular speech recognition systems, including Siri, Google Now, Samsung S Voice, Huawei HiVoice, Cortana and Alexa. By injecting a sequence of inaudible voice commands, we show a few proof-of-concept attacks, which include activating Siri to initiate a FaceTime call on iPhone, activating Google Now to switch the phone to the airplane mode, and even manipulating the navigation system in an Audi automobile. We propose hardware and software defense solutions. We validate that it is feasible to detect DolphinAttack by classifying the audios using supported vector machine (SVM), and suggest to re-design voice controllable systems to be resilient to inaudible voice command attacks.

KEYWORDS

Voice Controllable Systems, Speech Recognition, MEMS Microphones, Security Analysis, Defense

1 INTRODUCTION

Speech recognition (SR) technologies allow machines or programs to identify spoken words and convert them into machine-readable

formats. It has become an increasingly popular human-computer interaction mechanism because of its accessibility, efficiency, and recent advances in recognition accuracy. As a result, speech recognition systems have turned a wide variety of systems into voice controllable systems (VCS): Apple Siri [5] and Google Now [21] allow users to initiate phone calls by voices; Alexa [4] has enabled users to instruct an Amazon Echo to order takeouts, schedule a Uber ride, etc. As researchers devote much of their effort into improving the performance of SR systems, what is less well understood is how speech recognition and the voice controllable systems behave under intentional and sneaky attacks.

Prior work [10, 61] has shown that obfuscated voice commands which are incomprehensible to human can be understood by SR systems, and thus may control the systems without being detected. Such voice commands, though ‘hidden’, are nonetheless audible and remain conspicuous. This paper aims at examining the feasibility of the attacks that are difficult to detect, and the paper is driven by the following key questions: *Can voice commands be **inaudible** to human while still being audible to devices and intelligible to speech recognition systems?* *Can injecting a sequence of inaudible voice commands lead to unnoticed security breaches to the voice controllable systems?* To answer these questions, we designed DolphinAttack, an approach to inject inaudible voice commands at VCS by exploiting the ultrasound channel (i.e., $f > 20$ kHz) and the vulnerability of the underlying audio hardware.

Inaudible voice commands may appear to be unfeasible with the following doubts. (a) *How can inaudible sounds be audible to devices?* The upper bound frequency of human voices and human hearing is 20 kHz. Thus, most audio-capable devices (e.g., phones) adopt audio sampling rates lower than 44 kHz, and apply low-pass filters to eliminate signals above 20 kHz [32]. Previous work [61] considers it impossible to receive voices above 20 kHz. (b) *How can inaudible sounds be intelligible to SR systems?* Even if the ultrasound is received and correctly sampled by hardware, SR systems will not recognize signals that do not match human tonal features, and therefore unable to interpret commands. (c) *How can inaudible sounds cause unnoticed security breach to VCS?* The first step towards controlling VCSs is to activate them. Many VCSs (e.g., smartphones and smart home devices) implement the always-on feature that allows them to be activated by speaker-dependent wake words, i.e., such systems utilize voice recognition to authenticate a user. A random voice command will not pass the voice recognition. We solved all these problems, and we show that the DolphinAttack voice commands, though totally inaudible and therefore imperceptible

*Guoming and Chen are co-first authors.

†Corresponding faculty authors.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CCS '17, October 30–November 3, 2017, Dallas, TX, USA

© 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-4946-8/17/10...\$15.00

<https://doi.org/10.1145/3133956.3134052>

Acetaminophen attenuates error evaluation in cortex

Daniel Randles,¹ Julia W.Y. Kam,² Steven J. Heine,³ Michael Inzlicht,¹ and Todd C. Handy³

¹Department of Psychology, The University of Toronto, Toronto, ON, Canada, ²Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA, and ³Department of Psychology, The University of British Columbia, Vancouver, BC, Canada

Correspondence should be addressed to Daniel Randles, University of Toronto, 1265 Military Trail, Toronto, ON M1C 1A4, Canada. E-mail: dan.randles@utoronto.ca.

Abstract

Acetaminophen has recently been recognized as having impacts that extend into the affective domain. In particular, double blind placebo controlled trials have revealed that acetaminophen reduces the magnitude of reactivity to social rejection, frustration, dissonance and to both negatively and positively valenced attitude objects. Given this diversity of consequences, it has been proposed that the psychological effects of acetaminophen may reflect a widespread blunting of evaluative processing. We tested this hypothesis using event-related potentials (ERPs). Sixty-two participants received acetaminophen or a placebo in a double-blind protocol and completed the Go/NoGo task. Participants' ERPs were observed following errors on the Go/NoGo task, in particular the error-related negativity (ERN; measured at FCz) and error-related positivity (Pe; measured at Pz and CPz). Results show that acetaminophen inhibits the Pe, but not the ERN, and the magnitude of an individual's Pe correlates positively with omission errors, partially mediating the effects of acetaminophen on the error rate. These results suggest that recently documented affective blunting caused by acetaminophen may best be described as an inhibition of evaluative processing. They also contribute to the growing work suggesting that the Pe is more strongly associated with conscious awareness of errors relative to the ERN.

Key words: acetaminophen; ERN; Pe; error

Introduction

Long used as an analgesic and antipyretic, acetaminophen (paracetamol)—or Tylenol, as its brand name is more commonly known—has recently been recognized as having impacts that extend into the affective domain. In particular, double-blind placebo controlled trials have revealed that acetaminophen reduces the magnitude of reactivity to social pain and frustration (Dewall *et al.*, 2010), dissonance and existential anxiety triggered by unsettling experiences (Randles *et al.*, 2013; Dewall *et al.*, 2015), and to both negatively and positively valenced attitude objects (Durso *et al.*, 2015). Given this diversity of consequences, it has recently been proposed that the psychological effects of acetaminophen are not specific to the affective domain, but instead reflect a more widespread blunting of evaluative processing in general (Durso *et al.*, 2015). If this hypothesis is correct, acetaminophen should be able to inhibit or interfere with cognitive signals associated with evaluative processing. We test this by focusing on two event-related potentials (ERPs) that emerge in cortex in

quick temporal succession following a choice response error: the error-related negativity (ERN) and positivity (Pe).

Converging evidence suggests that a general conflict detection mechanism in the brain supports cognitive control (Botvinick *et al.*, 2004). Conflict itself is defined as any disagreement or discrepancy between mental representations, response tendencies or actual behavior (Holroyd and Coles, 2002; Botvinick *et al.*, 2004). This system scrutinizes the moment-to-moment representations of action tendencies for potential conflicts, so that inhibitory mechanisms may be engaged to override the unwanted tendency and promote effective goal pursuit. The ERN/Pe ERP complex appears to signal activity of this conflict detection system. The ERN emerges 0–100 ms post-error and is typically maximal at frontal-central midline electrode sites (Falkenstein *et al.*, 1990; Gehring *et al.*, 1993), with the Pe following at 180–350 ms post-response, emerging more clearly along parietal-central midline sites (Hajcak *et al.*, 2003; Santesso *et al.*, 2005; Steinhauser and Young, 2010).

Received: 13 October 2015; Revised: 17 December 2015; Accepted: 11 February 2016

© The Author (2016). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

Is genetic drift a force?

Charles H. Pence¹

Received: 23 August 2013 / Accepted: 20 January 2016
© Springer Science+Business Media Dordrecht 2016

Abstract One hotly debated philosophical question in the analysis of evolutionary theory concerns whether or not evolution and the various factors which constitute it (selection, drift, mutation, and so on) may profitably be considered as analogous to “forces” in the traditional, Newtonian sense. Several compelling arguments assert that the force picture is incoherent, due to the peculiar nature of genetic drift. I consider two of those arguments here—that drift lacks a predictable direction, and that drift is constitutive of evolutionary systems—and show that they both fail to demonstrate that a view of genetic drift as a force is untenable. I go on to diagnose the reasons for the stubborn persistence of this problem, considering two open philosophical issues and offering some preliminary arguments in support of the force metaphor.

Keywords Evolutionary theory · Genetic drift · Force · Causation · Brownian motion

1 Introduction

Evolutionary theory can, as we all know, be decomposed into multiple components or factors. It characterizes change in populations over time, that is, as the result of natural selection, genetic drift, mutation (Mani and Clarke 1990; Merlin 2010), migration (Shpak and Proulx 2007), linkage disequilibrium (Lewontin and Kojima 1960; Lewontin 1964), meiotic drive (Lytle 1993), extinction (Lynch et al. 1995; Jablonski 2005), increase in complexity (McShea 1996, 2005; Carroll et al. 2001), and so on.

✉ Charles H. Pence
charles@charlespence.net

¹ Department of Philosophy and Religious Studies, Louisiana State University, 102 Coates Hall, Baton Rouge, LA 70803, USA



Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Genetic analysis of hyperemesis gravidarum reveals association with intracellular calcium release channel (RYR2)

Marlena Schoenberg Fejzo ^{a,b,*}, Ronny Myhre ^c, Lucía Colodro-Conde ^d, Kimber W. MacGibbon ^e, Janet S. Sinsheimer ^f, M.V. Prasad Linga Reddy ^g, Päivi Pajukanta ^g, Dale R. Nyholt ^h, Margaret J. Wright ⁱ, Nicholas G. Martin ^j, Stephanie M. Engel ^k, Sarah E. Medland ^d, Per Magnus ^c, Patrick M. Mullin ^b

^a Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA^b Department of Maternal-Fetal Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA^c Norwegian Institute of Public Health, Oslo Norway^d Psychiatric Genetics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia^e Hyperemesis Education and Research Foundation, Damascus, OR, USA^f Departments of Biostatistics, Biomathematics, & Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA^g Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA^h Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australiaⁱ Queensland Brain Institute and Centre for Advanced Imaging, University of Queensland, Brisbane, Australia^j Genetic Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia^k Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ARTICLE INFO

Article history:

Received 9 July 2016

Received in revised form

19 August 2016

Accepted 19 September 2016

Available online xxx

Keywords:

Hyperemesis gravidarum

Nausea

Vomiting

Pregnancy

RYR2

ABSTRACT

Hyperemesis Gravidarum (HG), severe nausea/vomiting in pregnancy (NVP), can cause poor maternal/fetal outcomes. Genetic predisposition suggests the genetic component is essential in discovering an etiology. We performed whole-exome sequencing of 5 families followed by analysis of variants in 584 cases/431 controls. Variants in *RYR2* segregated with disease in 2 families. The novel variant L3277R was not found in any case/control. The rare variant, G1886S was more common in cases ($p = 0.046$) and extreme cases ($p = 0.023$). Replication of G1886S using Norwegian/Australian data was supportive. Common variants rs790899 and rs1891246 were significantly associated with HG and weight loss. Copy-number analysis revealed a deletion in a patient. *RYR2* encodes an intracellular calcium release channel involved in vomiting, cyclic-vomiting syndrome, and is a thyroid hormone target gene. Additionally, *RYR2* is a downstream drug target of Inderal, used to treat HG and CVS. Thus, herein we provide genetic evidence for a pathway and therapy for HG.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Nausea and vomiting of pregnancy is a common symptom affecting 70% of pregnant women (Goodwin, 1998). Clinical intervention is necessary in the severest form, Hyperemesis Gravidarum (HG), which affects up to 2% of pregnancies (Christodoulou-Smith et al., 2011). HG leads to significant weight loss, dehydration, electrolyte imbalance, and ketonuria (Fairweather, 1968; Goodwin et al., 1992a,b; Goodwin, 1998). Although maternal mortality is

rare, 6 deaths due to HG have been reported recently (MacGibbon et al., 2015), as well as morbidity including Wernicke's encephalopathy (Chiossi et al., 2006), acute renal failure (Hill et al., 2002), liver function abnormalities (Ahmed et al., 2013), splenic avulsion (Nguyen et al., 1995), esophageal rupture (Woolford et al., 1993), pneumothorax (Schwartz and Rosoff, 1994), and post-traumatic stress symptoms (Christodoulou-Smith et al., 2011). HG is also associated with poor fetal/child outcomes including a 4-fold increased risk of preterm birth and a 3-fold increased risk of neurodevelopmental delay in children (Fejzo et al., 2013, 2015).

A variety of potential causative factors have been investigated, but the etiology remains unknown. Evidence for a genetic predisposition is provided by classic twin studies of Norwegian, Spanish,

* Corresponding author. 5535 MRL Bldg. Charles E Young Dr. S LA, CA 90095, USA.
E-mail address: mfejzo@mednet.ucla.edu (M.S. Fejzo).

A braindump about some speculative directed evolution projects

Bryan Bishop <kanzure@gmail.com>

0E4C A12B E16B E691 56F5 40C9 984F 10CC 7716 9FD2

2017-02-03

<http://diyhpl.us/~bryan/irc/2017-02-03-beacon.pdf>

A braindump about some speculative directed evolution projects

Bryan Bishop <kanzure@gmail.com>

2017-02-03

A braindump about some speculative directed evolution projects

Bryan Bishop <kanzure@gmail.com>

2017-02-03

Ordered Multistep Synthesis in a Single Solution Directed by DNA Templates**

Thomas M. Snyder and David R. Liu*

The use of DNA templates to direct chemical reactions enables the principles of biological evolution to be applied to the discovery of synthetic molecules and new chemical reactions.^[1] In addition, the use of DNA hybridization to modulate the effective molarity of DNA-linked reactants allows synthetic molecules to be manipulated in ways that cannot be accessed using traditional synthetic methods.^[2] For example, many oligomeric natural products, including proteins, nonribosomal peptides, and polyketides, are biosynthesized in a strictly ordered manner even though all of their constituent building blocks are simultaneously present in the cellular milieu.^[3] Nature achieves single-solution ordered multistep synthesis by increasing the effective molarity of specific sets of reactants at precise moments during biosynthesis. Compared to the strategy most frequently used by chemists to execute ordered multistep synthesis—dividing the construction of a molecule into a sequence of isolated reaction steps—nature's single-solution approach is not only remarkably elegant and efficient but also sufficiently selective to obviate the need for protecting groups.

In the absence of enzymes, ordered multistep synthesis in a single solution has proven to be a challenge. The ordered oligomerization of nucleotides on nucleic acid templates^[1a,4] has been achieved, but these methods have not allowed the synthesis of non-nucleic acid structures. Tamura and Schimmel^[5] have used RNA-templated synthesis to direct peptide-bond formation in an order determined by intrinsic differences in substrate reactivity. Relying on substrate reactivity differences, however, imposes significant constraints on the order of building blocks within the possible products. Even with precisely tuned reactivities, multistep syntheses generally still require sequential additions of reactants to form ordered products. Herein, we report the ordered multistep syntheses of both a triolefin and a tripeptide using

DNA-linked substrates of comparable intrinsic reactivity that are simultaneously present in one solution. In both cases, temperature-sensitive variations in DNA secondary structure orchestrate a series of effective molarity changes among four reactants to primarily generate one ordered product out of many possibilities. These results represent two strategies for ordered synthesis in a single solution without the structural constraints imposed by enzymes and can significantly enhance the efficiency and selectivity of multistep DNA-templated synthesis.^[6]

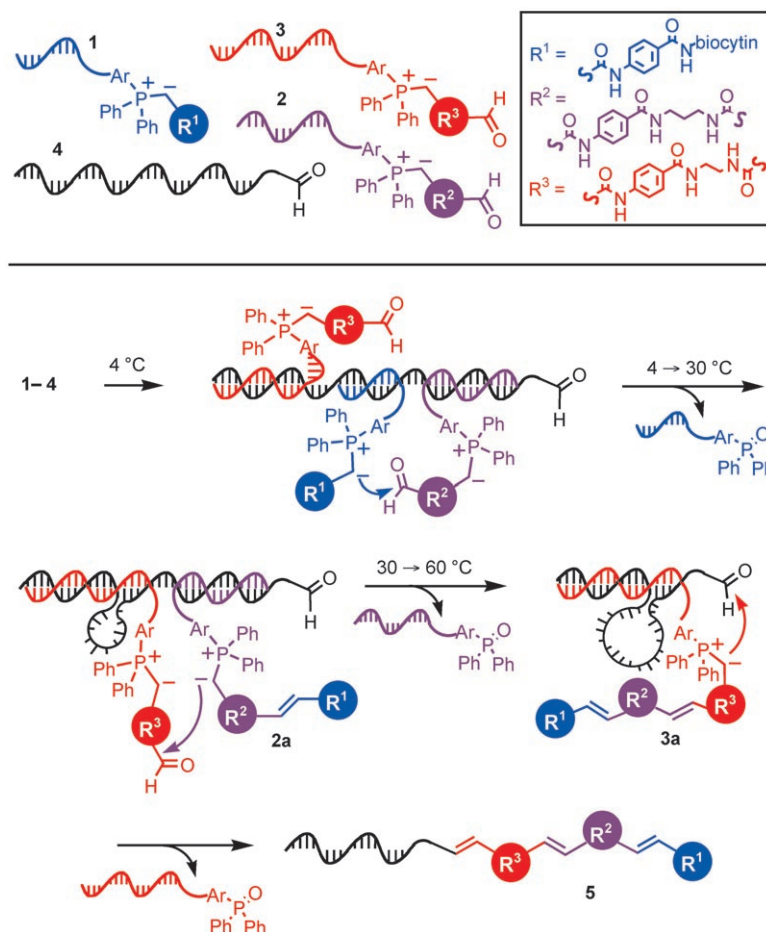


Figure 1. Strategy for the single-solution synthesis of an ordered triolefin. Building blocks are transferred sequentially among phosphorane reagents 1–3 before addition to an aldehyde-linked template 4. The rigidity of double-stranded DNA enforces Wittig olefination regioselectivity. As the reaction temperature is elevated, the DNA secondary structure undergoes sequence-programmed changes that enables the desired Wittig olefination to take place selectively.

[*] T. M. Snyder, Prof. D. R. Liu
Harvard University
Department of Chemistry and Chemical Biology
12 Oxford Street, Cambridge, MA 02138 (USA)
Fax (+1) 617-496-5688
E-mail: drliu@fas.harvard.edu

[**] Funding was generously provided by the NIH/NIGMS (R01 M065865), the Office of Naval Research (N00014-03-1-0749), and an NSF Graduate Research Fellowship to T.M.S.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

The first strategy (Figure 1) passes synthetic intermediates from site to site on a template in a manner controlled by changes in the DNA secondary structure. Three Wittig olefination substrates (1–3, with 1 containing a biotin group),^[7] each linked to DNA oligonucleotides of varying melting temperature (T_m), were prepared and hybridized to an aldehyde-terminated DNA template 4 at 4 °C. If these substrates were combined at the high concentrations (mM–M) typical of organic synthesis, a complex mixture of many

Changes in Sleep-Wakefulness after Kainic Acid Lesion of the Preoptic Area in Rats

Joshi JOHN, Velayudhan MOHAN KUMAR, Gomathy GOPINATH*,
Vijay RAMESH, and Hrudananda MALLICK

*Department of Physiology and *Department of Anatomy, All India
Institute of Medical Sciences, New Delhi, 110 029 India*

Abstract The role of the preoptic area (POA) neurons in the regulation of sleep-wakefulness (S-W) has been investigated in this study. The cell-specific neurotoxin, kainic acid (KA), was injected ($0.8 \mu\text{g}$ in $0.2 \mu\text{l}$) intracerebrally for lesioning of the POA. S-W was assessed (on the basis of EEG, EMG, and EOG recordings) for a day before bilateral lesion of the POA, and for 3 weeks after the lesion. There was an increase in wakefulness, and a decrease in all the stages of sleep after KA lesion of the POA. The reduction in deep slow wave sleep (S2) and REM sleep (PS) were more marked than light slow wave sleep (S1), and these had not shown any recovery even after 3 weeks of lesion. Two days after the lesion, the reduction in sleep was much more marked during the daytime than at night. There was an increase in locomotor activity, especially during the daytime, though it was only statistically significant on the 6th and the 10th day after the lesion. This study shows that the POA neurons are involved in the induction and maintenance of sleep. The lesion did not have a long lasting effect on the circadian distribution of sleep but the changes in locomotor activity seem to persist for a longer period.

Key words: preoptic area, kainic acid, sleep, wakefulness, locomotor activity.

The preoptic area (POA) is known to regulate slow wave sleep and REM sleep (PS) [1–3]. Nauta described it as the area for the “capacity of sleeping” on the basis of surgical destruction studies in rats [2]. Severe suppression of sleep was also reported after electrolytic lesion of the POA in rats [1] and cats [3]. Recently, radiofrequency lesions of the POA have shown that this area is important not only for sleep, but also for the regulation of the sleep-wakefulness (S-W) cycle [4]. The electrolytic and radiofrequency lesions destroy both the cells and fibers of passage. It could be argued that the changes in sleep after these lesions, might have resulted from the destruction of fibres as the POA has cells and fibers of passage [5]. This

Received on September 16, 1993; Accepted on April 18, 1994

SECURITIES AND EXCHANGE COMMISSION
(Release No. 34-80206; File No. SR-BatsBZX-2016-30)

March 10, 2017

Self-Regulatory Organizations; Bats BZX Exchange, Inc.; Order Disapproving a Proposed Rule Change, as Modified by Amendments No. 1 and 2, to BZX Rule 14.11(e)(4), Commodity-Based Trust Shares, to List and Trade Shares Issued by the Winklevoss Bitcoin Trust

Bats BZX Exchange (“Exchange” or “BZX”) has filed a proposed rule change to list and trade shares of the Winklevoss Bitcoin Trust.¹ When an exchange makes such a filing,² the Commission must determine whether the proposed rule change is consistent with the statutory provisions, and the rules and regulations, that apply to national securities exchanges.³ The Commission must approve the filing if it finds that the proposed rule change is consistent with these legal requirements, and it must disapprove the filing if it does not make such a finding.⁴

¹ The Exchange filed notice of the proposed rule change on June 30, 2016, and the Commission published the notice in the Federal Register on July 14, 2016. See Exchange Act Release No. 78262 (July 8, 2016), 81 FR 45554 (July 14, 2016) (“Notice”). On August 23, 2016, the Commission designated a longer period within which to act on the proposed rule change. See Exchange Act Release No. 78653 (Aug. 23, 2016), 81 FR 59256 (Aug. 29, 2016). On October 12, 2016, the Commission instituted proceedings under Section 19(b)(2)(B) of the Securities Exchange Act of 1934 (“Exchange Act”), 15 U.S.C. 78s(b)(2)(B), to determine whether to approve or disapprove the proposed rule change. See Exchange Act Release No. 79084 (Oct. 12, 2016), 81 FR 71778 (Oct. 18, 2016). On October 20, 2016, the Exchange filed Amendment No. 1 to the proposed rule change, replacing the original filing in its entirety, and Amendment No. 1 was published for comment in the Federal Register on November 3, 2016. See Exchange Act Release No. 79183 (Oct. 28, 2016), 81 FR 76650 (Nov. 3, 2016) (“Amendment No. 1”). On January 4, 2017, the Commission designated a longer period for Commission action on the proposed rule change. See Exchange Act Release No. 79725 (Jan. 4, 2017), 82 FR 2425 (Jan. 9, 2017) (designating March 11, 2017, as the date by which the Commission must either approve or disapprove the proposed rule change). On February 22, 2017, the Exchange filed Amendment No. 2 to the proposed rule change (“Amendment No. 2”). Amendment No. 2 further modified the Exchange’s proposal by (a) changing the size of a creation and redemption basket from 10,000 shares to 100,000 shares, (b) changing the bitcoin value of a share from 0.1 bitcoin to 0.01 bitcoin, and (c) changing the Exchange’s representation about the number of shares outstanding at the commencement of trading from 100,000 shares to 500,000 shares. Because Amendment No. 2 does not materially alter the substance of the proposed rule change, Amendment No. 2 is not subject to notice and comment. Amendment No. 2 is available on the Commission’s website at <https://www.sec.gov/comments/sr-batsbzx-2016-30/batsbzx201630-1594698-132357.pdf>.

² Such filings are made under Section 19(b)(1) of the Exchange Act, 15 U.S.C. 78s(b)(1), and Exchange Act Rule 19b-4, 17 CFR 240.19b-4.

³ See Exchange Act Section 19(b)(2)(C), 15 U.S.C. 78s(b)(2)(C).

⁴ See id.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/264634582>

CA3 size predicts the precision of memory recall

Article in *Proceedings of the National Academy of Sciences* · July 2014

DOI: 10.1073/pnas.1319641111 · Source: PubMed

CITATIONS

21

READS

57

3 authors, including:



Martin Chadwick

Google DeepMind

25 PUBLICATIONS 682 CITATIONS

[SEE PROFILE](#)



Heidi M Bonnici

University of East Anglia

18 PUBLICATIONS 382 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Hypoxia ischaemia, hippocampal pathology leading to developmental amnesia [View project](#)

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/264710474>

Exome sequencing of three cases of familial exceptional longevity

Article in *Aging cell* · August 2014

DOI: 10.1111/accel.12261

CITATIONS

3

READS

87

14 authors, including:



Leocadio Rodríguez-Mañas

Hospital Universitario de Getafe

263 PUBLICATIONS 4,143 CITATIONS

[SEE PROFILE](#)



Catalina Santiago

European University of Madrid

126 PUBLICATIONS 1,667 CITATIONS

[SEE PROFILE](#)



Nuria Garatachea

University of Zaragoza

173 PUBLICATIONS 1,821 CITATIONS

[SEE PROFILE](#)



Anna Gonzalez-Neira

Centro Nacional de Investigaciones Oncológi...

245 PUBLICATIONS 8,438 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



VIVIFRIL [View project](#)



Breast Cancer [View project](#)

All content following this page was uploaded by **Nuria Garatachea** on 17 November 2014.

The user has requested enhancement of the downloaded file.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/271443248>

Aptamer Nanostructures as Signaling Molecular Switches in Electrochemical Biosensing

Chapter · January 2013

CITATIONS

0

READS

208

2 authors:



Tatsuro Goda

Tokyo Medical and Dental University

62 PUBLICATIONS 811 CITATIONS

SEE PROFILE



Yuji Miyahara

Tokyo Medical and Dental University

180 PUBLICATIONS 2,738 CITATIONS

SEE PROFILE

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/291419723>

Nausea and Vomiting During Pregnancy is Highly Heritable

Article in *Behavior Genetics* · January 2016

DOI: 10.1007/s10519-016-9781-7

CITATIONS

3

READS

124

8 authors, including:



Lucía Colodro Conde

Queensland Institute of Medical Research

89 PUBLICATIONS 145 CITATIONS

[SEE PROFILE](#)



Patrick Jern

Åbo Akademi University

81 PUBLICATIONS 1,293 CITATIONS

[SEE PROFILE](#)



Juan Francisco Sánchez

University of Murcia

35 PUBLICATIONS 110 CITATIONS

[SEE PROFILE](#)



Juan R Ordoñana

University of Murcia

102 PUBLICATIONS 406 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Molecular bases in hypersexual disorder and gender dysphoria [View project](#)



Diagnosis, etiology and psychobehavioral treatment of premature ejaculation [View project](#)

All content following this page was uploaded by **Lucía Colodro Conde** on 02 February 2016.

The user has requested enhancement of the downloaded file.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/302589015>

CLARITY-compatible lipophilic dyes for electrode marking and neuronal tracing

Article in *Scientific Reports* · September 2016

DOI: 10.1038/srep32674

CITATIONS

2

READS

460

2 authors:



Kristian H. Reveles Jensen

University of Copenhagen

8 PUBLICATIONS 25 CITATIONS

[SEE PROFILE](#)



Rune W Berg

University of Copenhagen

39 PUBLICATIONS 1,133 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Collision of Action Potentials in Invertebrates [View project](#)



Model-based learning, dopamine and stress [View project](#)

All content following this page was uploaded by [Rune W Berg](#) on 10 May 2016.

The user has requested enhancement of the downloaded file.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/304250067>

The emerging role of Notch pathway in ageing: Focus on the related mechanisms in age- related diseases

Article in *Ageing research reviews* · June 2016

DOI: 10.1016/j.arr.2016.06.004

CITATIONS

10

READS

145

4 authors, including:



Carmela Rita Balistreri

Università degli Studi di Palermo

148 PUBLICATIONS 2,506 CITATIONS

[SEE PROFILE](#)



Rosalinda Madonna

Università degli Studi G. d'Annunzio Chieti e ...

147 PUBLICATIONS 1,599 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Endothelial progenitor cells in age-related diseases [View project](#)

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/280094272>

Preparation of Information-Containing Macromolecules by Ligation of Dyad-Encoded Oligomers

Article in *Chemistry - A European Journal* · July 2015

DOI: 10.1002/chem.201502414

CITATIONS

25

READS

69

5 authors, including:



Tam Thanh Trinh

University of Strasbourg

7 PUBLICATIONS 103 CITATIONS

[SEE PROFILE](#)



Delphine Chan-Seng

French National Centre for Scientific Research

30 PUBLICATIONS 625 CITATIONS

[SEE PROFILE](#)



Laurence Charles

Aix-Marseille Université

177 PUBLICATIONS 1,635 CITATIONS

[SEE PROFILE](#)



Jean-François Lutz

French National Centre for Scientific Research

202 PUBLICATIONS 11,735 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Distinction of Sugar Isomers Using the Kinetic Method [View project](#)



Mass Spectrometry of Plasma Polymers [View project](#)

All content following this page was uploaded by **Tam Thanh Trinh** on 07 July 2016.

The user has requested enhancement of the downloaded file.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/304778629>

Smile and Laughter Elicited by Electrical Stimulation of the Frontal Operculum

Article in *Neuropsychologia* · July 2016

DOI: 10.1016/j.neuropsychologia.2016.07.001

CITATIONS

5

READS

203

5 authors, including:



Fausto Caruana

Università degli studi di Parma

52 PUBLICATIONS 852 CITATIONS

[SEE PROFILE](#)



Veronica Pelliccia

Epilepsy Surgery Center "Claudio Munari" Ni...

28 PUBLICATIONS 119 CITATIONS

[SEE PROFILE](#)



Massimo Cossu

Azienda Ospedaliera Niguarda Ca' Granda

124 PUBLICATIONS 2,644 CITATIONS

[SEE PROFILE](#)



Pietro Avanzini

Italian National Research Council

41 PUBLICATIONS 272 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Basic Emotions: Still Necessary After All These Years? [View project](#)



Understanding Tool Behavior [View project](#)

All content following this page was uploaded by **Fausto Caruana** on 14 July 2016.

The user has requested enhancement of the downloaded file.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/305807508>

Restoring Latent Visual Working Memory Representations in Human Cortex

Article in *Neuron* · August 2016

DOI: 10.1016/j.neuron.2016.07.006

CITATIONS

15

READS

566

3 authors:



Thomas C Sprague

New York University

13 PUBLICATIONS 169 CITATIONS

[SEE PROFILE](#)



Edward F Ester

Florida Atlantic University

41 PUBLICATIONS 828 CITATIONS

[SEE PROFILE](#)



John T Serences

University of California, San Diego

95 PUBLICATIONS 5,199 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Mechanisms of distraction in visual short-term memory [View project](#)

Workshop Results

By David Markowitz

This document summarizes the gating research challenges and paths to the development of practically useful DNA storage technology that were identified by participants in a recent IARPA/SRC workshop on molecular information storage. Workshop results were highlighted in Nature and IEEE Spectrum.^{1,2}

1. Introduction

The goal of the envisioned DNA Storage program is to develop a fundamentally new storage paradigm that encodes bits using molecular scale polymers (e.g. DNA, RNA or synthetic polymers), which can be densely packed to achieve extraordinarily high volumetric information density. To achieve this goal, methods and tools need to be developed for synthesizing, storing and reading information from synthetic polymers with low latency and high bandwidth. These tools will then be integrated to produce an exabytescale archival storage technology that reduces the power, cost and physical footprint requirements of today's "cold storage" systems by three orders of magnitude.

A number of recent studies have demonstrated the feasibility of molecular information storage^{1,2}. Furthermore, a large and highly collaborative research community has developed around this problem since 2012, with contributors from academia (10+ universities in the US and Europe), the biotechnology industry (Illumina, Pacific Biosciences, Twist Bioscience, Gen9), the semiconductor industry (Intel, Micron), and the software industry (Microsoft, Autodesk). Three publicly disclosed academic-industry partnerships now exist, sponsored by Micron, Microsoft, and Semiconductor Research Corporation (SRC)^{3,4}.

2. Technical Discussion Summary

To guide discussion during the DNA Storage workshop, the workshop participants were asked to consider how one might build a molecular "hard drive" with the following performance characteristics:

- Total storage capacity: one exabyte,
- Write bandwidth: one petabyte/day, at an operating cost of \$1k/petabyte,
- Read bandwidth: ten petabytes/day (1% of total archive), at a cost of \$1k / 10 petabytes,
- Infrastructure cost, power and physical footprint requirements: Three orders of magnitude smaller than the requirements of today's exascale cold storage data centers:

Parameter	Current mainstream technology	Target
Infrastructure Cost	\$1B / exabyte	\$1M / exabyte
Power	200 MW	200 kW
Footprint	1,000,000 sq. ft / exabyte	1,000 sq. ft. / exabyte

The consensus by workshop attendees was that practical, exabytescale molecular information storage is achievable in 5 years. Evidence to support this view is provided below for three key technical areas: polymer synthesis, sequencing, and operation of a molecular storage system.

Tempo and mode of antibat ultrasound production and sonar jamming in the diverse hawkmoth radiation

Akito Y. Kawahara^{a,1,2} and Jesse R. Barber^{b,1,2}

^aFlorida Museum of Natural History, University of Florida, Gainesville, FL 32611; and ^bDepartment of Biological Sciences, Boise State University, Boise, ID 83725

Edited by May R. Berenbaum, University of Illinois at Urbana–Champaign, Urbana, IL, and approved January 5, 2015 (received for review August 29, 2014)

The bat–moth arms race has existed for over 60 million y, with moths evolving ultrasonically sensitive ears and ultrasound-producing organs to combat bat predation. The evolution of these defenses has never been thoroughly examined because of limitations in simultaneously conducting behavioral and phylogenetic analyses across an entire group. Hawkmoths include >1,500 species worldwide, some of which produce ultrasound using genital stridulatory structures. However, the function and evolution of this behavior remain largely unknown. We built a comprehensive behavioral dataset of hawkmoth hearing and ultrasonic reply to sonar attack using high-throughput field assays. Nearly half of the species tested (57 of 124 species) produced ultrasound to tactile stimulation or playback of bat echolocation attack. To test the function of ultrasound, we pitted big brown bats (*Eptesicus fuscus*) against hawkmoths over multiple nights and show that hawkmoths jam bat sonar. Ultrasound production was immediately and consistently effective at thwarting attack and bats regularly performed catching behavior without capturing moths. We also constructed a fossil-calibrated, multigene phylogeny to study the evolutionary history and divergence times of these antibat strategies across the entire family. We show that ultrasound production arose in multiple groups, starting in the late Oligocene (~26 Ma) after the emergence of insectivorous bats. Sonar jamming and bat-detecting ears arose twice, independently, in the Miocene (18–14 Ma) either from earless hawkmoths that produced ultrasound in response to physical contact only, or from species that did not respond to touch or bat echolocation attack.

acoustic | antipredator defense | bat–moth interactions | evolution | Sphingidae

Bats and moths have dominated the night sky for more than 60 million y (1, 2). The shared evolutionary history between these diverse groups has armed bats with effective attack behaviors and moths with powerful counter measures (3). Of the nearly 140,000 moth species (4), more than half may possess ultrasonically sensitive bat-detecting ears (5). Moth ears are connected to neuronal circuits that steer the animals away from bats at low sonar call intensities, trigger aerobatic evasive behaviors—such as directed turns, loops, spirals, and power dives (3)—and elicit ultrasound production at high sonar intensities (6–8). Ultrasonic ears are known to have independently evolved at least 18 times in seven insect orders (9) and possibly 10 or more times in Lepidoptera (5).

Recent work on bat–moth aerial encounters has focused on tiger moths (Arctiinae; reviewed in ref. 3, but also see ref. 10). Tiger moths can hear bat echolocation and have paired metathoracic tymbals that beam ultrasonic clicks back at bats (6, 7). The sounds that these moths make can startle bats (11), warn them of bad taste (12, 13), and jam bat sonar (14–16). Jamming likely functions to confuse the bat because it interferes with the bat's perception of its own echoes that are reflecting off of its prey (16, 17). The evolution of moth antibat ultrasound production is thought to have originated from simple warning sounds and developed into complex signals that jam bat sonar (14). However, moth ultrasound production, one of the most

sophisticated responses to bat attack (3), has never been examined in a phylogenetic context.

Hawkmoths (family Sphingidae) are among the most conspicuous nocturnal Lepidoptera (18); they are generally long-lived, strong flying pollinators (19, 20). Because of their large size and lack of known adult toxins (21, 22), sphingids are important prey for bats (23–28). Some hawkmoths possess specialized sound-receiving structures on the head, comprised of the labral pilifers and palps that are sensitive to the ultrasonic frequencies of bat sonar (29, 30). One group of hawkmoths, the Choerocampina, has been shown to produce ultrasound by stridulating modified genital valves when stimulated by playback of bat sonar attack (6). However, the function and taxonomic breadth of ultrasound production in hawkmoths is unknown.

To determine the function of ultrasound production, we conducted bat–moth interaction experiments in the laboratory. We predicted that the pattern over time of a naïve bat's success in catching stridulating hawkmoths would differentiate the three hypotheses for their function (12). If hawkmoths are jamming bat sonar, the moth sounds should be immediately and consistently effective at thwarting attack. Alternatively, if the moth sounds startle bats, the predators should initially be deterred but habituate over time. If moths are warning bats of distasteful toxins, bats should initially capture and drop moths, then learn to avoid them (12, 13). We paired these laboratory experiments with high-throughput behavioral field assays and used genetic

Significance

Ultrasound production is one of the most sophisticated antibat strategies in nocturnal insects, yet it has never been thoroughly studied in a phylogenetic framework. We conducted high-throughput field assays using playback of echolocation attack sequences, laboratory bat–moth interaction experiments, and fossil-calibrated phylogenetic analyses to provide the first evidence that multiple unrelated hawkmoth species produce ultrasound and jam bat echolocation. Our robust tree demonstrates that sonar jamming evolved twice during the Miocene after the radiation of insectivorous bats. We provide an example of the power behind collaborative science for revealing the function and historic pattern of behavior, and predict that ultrasound production is a widespread antibat strategy in the extraordinary diversity of nocturnal insects.

Author contributions: A.Y.K. and J.R.B. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KP719983–KP720300). Datasets and accompanying files are available from the Dryad Data Repository (www.datadryad.org; accession no. 10.5061/dryad.3450r).

¹A.Y.K. and J.R.B. contributed equally to this work.

²To whom correspondence may be addressed. Email: kawahara@flmnh.ufl.edu or jessebarber@boisestate.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1416679112/-DCSupplemental.

A Powerful Generative Model Using Random Weights for the Deep Image Representation

Kun He^{*} Yan Wang[†]

Department of Computer Science and Technology
Huazhong University of Science and Technology, Wuhan 430074, China
brooklet60@hust.edu.cn, yanwang@hust.edu.cn

John Hopcroft

Department of Computer Science
Cornell University, Ithaca 14850, NY, USA
jeh@cs.cornell.edu

Abstract

To what extent is the success of deep visualization due to the training? Could we do deep visualization using untrained, random weight networks? To address this issue, we explore new and powerful generative models for three popular deep visualization tasks using untrained, random weight convolutional neural networks. First we invert representations in feature spaces and reconstruct images from white noise inputs. The reconstruction quality is statistically higher than that of the same method applied on well trained networks with the same architecture. Next we synthesize textures using scaled correlations of representations in multiple layers and our results are almost indistinguishable with the original natural texture and the synthesized textures based on the trained network. Third, by recasting the content of an image in the style of various artworks, we create artistic images with high perceptual quality, highly competitive to the prior work of Gatys et al. on pretrained networks. To our knowledge this is the first demonstration of image representations using untrained deep neural networks. Our work provides a new and fascinating tool to study the representation of deep network architecture and sheds light on new understandings on deep visualization. It may possibly lead to a way to compare network architectures without training.

1 Introduction

In recent years, Deep Neural Networks (DNNs), especially Convolutional Neural Networks (CNNs), have demonstrated highly competitive results on object recognition and image classification [1, 2, 3, 4]. With advances in training, there is a growing trend towards understanding the inner working of these deep networks. By training on a very large image data set, DNNs develop a representation of images that makes object information increasingly explicit at various levels of the hierarchical architecture. Significant visualization techniques have been developed to understand the deep image representations on trained networks [5, 6, 7, 8, 9, 10, 11].

Inversion techniques have been developed to create synthetic images with feature representations similar to the representations of an original image in one or several layers of the network. Feature representations are a function Φ of the source image x_0 . An approximate inverse Φ^{-1} is used to

^{*}The three authors contributing equally.

[†]Corresponding author.



October 2017

BIODEFENSE

Federal Efforts to Develop Biological Threat Awareness

A Sequential Strand-Displacement Strategy Enables Efficient Six-Step DNA-Templated Synthesis

Yu He and David R. Liu*

Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, United States.

S Supporting Information

ABSTRACT: We developed a sequential strand-displacement strategy for multistep DNA-templated synthesis (DTS) and used it to mediate an efficient six-step DTS that proceeded in 35% overall yield (83% average yield per step). The efficiency of this approach and the fact that the final product remains linked to a DNA sequence that fully encodes its reaction history suggests its utility for the translation of DNA sequences into high-complexity synthetic libraries suitable for in vitro selection.

DNA-templated synthesis (DTS), the use of DNA hybridization to dramatically increase the effective molarity of reactants linked to oligonucleotides, is a powerful strategy to control chemical reactivity in a DNA sequence-programmed manner.^{1–4} Because the reaction products of DTS are encoded by the sequences of the associated DNA templates, they can be subjected to in vitro selection followed by PCR amplification and DNA sequence analysis to enable the discovery of functional small molecules,^{3,5–7} synthetic polymers,^{8–10} or novel chemical reactions.^{11–15} We recently reported the three-step DNA-templated synthesis of a 13,824-membered small-molecule macrocycle library.¹⁶ The library was subjected to in vitro selection for binding affinity to a variety of proteins of biomedical interest, ultimately yielding a new class of macrocyclic kinase inhibitors.⁶ Other complementary approaches to generating DNA-encoded libraries have led to the discovery of bioactive small molecules,^{17–29} including a number of examples in the pharmaceutical industry.^{5,30,31}

Generating DNA-encoded small molecules of significant structural complexity requires multistep DNA-programmed or DNA-tagged synthesis.^{3,5,23,30,32–37} A number of strategies have been developed to enable multistep DTS. The simplest uses a DNA template strand containing several codons and reagents linked to complementary anticodon oligonucleotides that are added successively.^{3,32–34} While this approach is conceptually straightforward, it requires several manipulations after each step that increase required time and effort and can decrease overall yields. The relative geometry between reactants on the template and reagent strands also changes after each step in this approach, potentially altering reaction efficiencies.³⁸ More complex self-assembled DNA structures and devices can also mediate multistep DTS. For example, a DNA three-way junction that contains multiple reagents at the junction has been developed for the construction of DNA-encoded peptides.⁵ We developed a DNA mechanical device that moves along a DNA track and mediates autonomous multistep organic synthesis in a single isothermal solution.³⁵ McKee et al. recently used a DNA strand-exchange strategy to achieve a three-step DTS in which products are

swapped between new and old DNA strands with the assistance of a “remover strand” that displaces expended reagent oligonucleotides.³⁶

Despite these significant advances and the diversity of approaches to generating multistep DTS products, all multistep DNA-templated small-molecules syntheses reported to date have used only three or fewer DNA-templated steps, and overall yields are generally low (typically <10%). Here we present a new strand-displacement strategy for multistep DTS and its use to mediate a six-step synthesis with an overall yield of 35% (average yield of 83%). By providing products of six-step DNA-programmed reaction sequences in good overall yield, the approach presented here may provide access to high-complexity DNA-templated small-molecule libraries.

Our strategy exploits “toehold displacement,” the known ability of a single-stranded DNA oligonucleotide (AB) to invade an asymmetric DNA duplex (A'B':B) that contains a single-stranded hybridization site (A') for the invading strand.³⁹ Once A:A' hybridization takes place, base pairing with the invading strand continues, ultimately resulting in strand displacement of the shorter, and therefore less-favorably hybridized, B strand. Displacement results in the formation of a new Watson–Crick complex suitable for DTS. We hypothesized that this approach could represent a highly efficient and very simple way to access products of several consecutive DNA-programmed reactions while preserving the correspondence between DNA sequence and reaction product structure that is required for in vitro selection.

The application of this “toehold displacement” strategy to multistep DTS is summarized in Figure 1. A single-stranded DNA template (T) consists of a 16-base initial reaction site (black) followed by five consecutive 8-base coding segments (colored) that also serve as toeholds to initiate sequential DNA strand displacement. In the first step, substrate DNA S1, which is tethered to the first reactant, hybridizes to T, initiating DTS. If self-cleaving reagents are used, the first reactant group is transferred from S1 to T as a natural consequence of the DNA-templated reaction. After sufficient time to react has elapsed, the second substrate strand S2 is added without requiring product purification or any other manipulation. S2 partially hybridizes with T through the red toehold region and displaces S1 by forming a more stable DNA duplex due to the fact that the oligonucleotide in S2 is eight nucleotides longer than that in S1. As a result, S1 spontaneously leaves the Watson–Crick complex and the second reactant's group is transferred to the T-linked intermediate through a DNA-templated reaction. Thereafter, an

Received: February 13, 2011

Published: June 09, 2011

**STATEMENT OF
TIM HUGHES
SENIOR VICE PRESIDENT FOR GLOBAL BUSINESS & GOVERNMENT AFFAIRS
SPACE EXPLORATION TECHNOLOGIES CORP. (SPACEX)**

**BEFORE THE
SUBCOMMITTEE ON SPACE, SCIENCE & TECHNOLOGY
COMMITTEE ON COMMERCE, SCIENCE & TECHNOLOGY
UNITED STATES SENATE**

July 13, 2017

Mr. Chairman, Ranking Member Markey, and Members of the Committee,

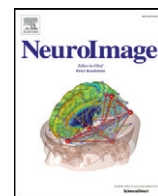
Thank you for the opportunity to participate in this important hearing on “Reopening the American Frontier.” SpaceX is a firm believer that public-private partnerships between U.S. commercial space entities and the Government are the optimal vehicles to rapidly, safely, and cost-effectively advance space exploration and settlement of the solar system.

Under your leadership, the Committee recently has reviewed an array of matters, including regulatory reform to enable commercial space to thrive and revisions to the Outer Space Treaty, which are critical to ensuring the Nation’s continued leadership in space exploration. Today’s hearing provides a timely opportunity to discuss the nature of NASA’s recent successful partnerships with private industry and to review how the United States can leverage such innovative approaches in its deep space endeavors going forward. SpaceX’s direct and significant experience working under unique, innovative public-private partnerships with NASA should help to shape the contours of this dialogue. In addition to existing programs at NASA focused on deep space exploration transportation and architectures, NASA again should pursue a parallel track that leverages non-traditional, public-private partnership approaches to increase the likelihood of success for the Nation’s space exploration objectives.

From its beginning, SpaceX has leveraged American innovation, technical savvy, and an iterative culture to yield the most advanced space launch vehicle and spacecraft systems in history. We are grateful for NASA’s ongoing support, which has been critical SpaceX’s success. We are proud to provide a dependable and affordable ride to space for NASA, the Department of Defense, and the world’s most sophisticated commercial satellite manufacturers and operators. Today, we regularly conduct critical un-crewed cargo resupply missions to and from the International Space Station (ISS) with our Dragon spacecraft—which was developed in partnership with NASA—and next year, we will begin launching American astronauts on American rockets for the first time since the Space Shuttle was retired in 2011. Commercially, SpaceX has restored the U.S. as a leader in global commercial satellite launch, taking back a market that had been wholly ceded to Russia and France for over a decade. As we look to the future, SpaceX is committed to continuing to support America’s space program and to contribute to our national exploration objectives through reliable, innovative, and affordable access to space.

To begin, it bears noting that the National Aeronautics and Space Act of 1958 identifies one of NASA’s core mission areas as follows: “[t]o seek and encourage, to the maximum extent possible, the fullest commercial use of space.”¹ Additionally, the National Space Transportation Policy expressly directs federal agencies to “[p]romote and maintain a dynamic, healthy, and efficient domestic space transportation industrial base,” and to do so, in part, by cultivating “increased technological innovation and

¹ Pub. L. 115–10, title III, §305(b), title IV, §443(b), Mar. 21, 2017, 131 Stat. 32, 47, added items 20148 and 20149.



Brains of verbal memory specialists show anatomical differences in language, memory and visual systems



James F. Hartzell^{a,*}, Ben Davis^a, David Melcher^a, Gabriele Miceli^a, Jorge Jovicich^a, Tanmay Nath^b, Nandini Chatterjee Singh^b, Uri Hasson^a

^a Center for Mind/Brain Sciences (CIMEC), University of Trento, 38060, Italy

^b National Brain Research Centre, Manesar, Gurgaon Dist., Haryana 122 050, India

ARTICLE INFO

Article history:

Accepted 8 July 2015

Available online 15 July 2015

Keywords:

Cortical thickness

Gray matter density

Diffusion tensor imaging

Language

Memory

Plasticity

Hippocampus

ABSTRACT

We studied a group of verbal memory specialists to determine whether intensive oral text memory is associated with structural features of hippocampal and lateral-temporal regions implicated in language processing. Professional Vedic Sanskrit Pandits in India train from childhood for around 10 years in an ancient, formalized tradition of oral Sanskrit text memorization and recitation, mastering the exact pronunciation and invariant content of multiple 40,000–100,000 word oral texts. We conducted structural analysis of gray matter density, cortical thickness, local gyrification, and white matter structure, relative to matched controls. We found massive gray matter density and cortical thickness increases in Pandit brains in language, memory and visual systems, including *i*) bilateral lateral temporal cortices and *ii*) the anterior cingulate cortex and the hippocampus, regions associated with long and short-term memory. Differences in hippocampal morphometry matched those previously documented for expert spatial navigators and individuals with good verbal working memory. The findings provide unique insight into the brain organization implementing formalized oral knowledge systems.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

A large body of research has established that acquisition of certain long-term skill sets or knowledge is linked to plasticity in both grey matter (GM) and white matter (WM) in multiple cortical and subcortical regions (May, 2011; Zatorre et al., 2004). As reviewed by May (2011, see references within), various expert groups such as sportsmen, mathematicians, ballet dancers, and professional board-game players all show particular morphological features that may be related to learning and plasticity.

Our goal in the current work was to examine the potential impact of extensive memorization and verbal recital practice on brain plasticity, as identifying brain regions implicated in these functions can elucidate the functional capacities of both lateral and medial temporal regions, as detailed below. To investigate the potential impact of extensive memorization and verbal recital practice on brain plasticity we recruited a sample group of traditional Sanskrit learners—Yajurveda Sanskrit Pandits—who memorize and recite one set of the most ancient Sanskrit

texts, the Vedas and their subsidiary texts (*Vedāṅgas*). The Sanskrit Vedas are late bronze/early iron-age oral texts passed down for over 3000 years in an unbroken tradition in India. They form the core of the ancient Sanskrit knowledge system, which developed extensive oral and later written literature in a wide range of traditional subjects still taught in India's Sanskrit institutions using traditional oral memorization and recitation methods (Rashtriya Sanskrit Sansthan, 2014). Professional Vedic Pandits undergo rigorous training in exact pronunciation and invariant content of these oral texts for 7 or more years, with 8–10 h of daily practice (totaling ~10,080 h over the course of the initial training), starting in their childhood, and mastering multiple 40,000 to 100,000 word oral texts (compared to ~38,000 in the book of Genesis). The training methods strongly emphasize traditional face-to-face oral learning, and the Yajurveda recitation practice includes right hand and arm gestures to mark prosodic elements. After graduation from training, professional Yajurveda Pandits work as teachers or Vedic priests, with daily recitation reduced to ~3 h.

We note that while the ability of Yajurveda Pandits to perform large-scale, precise oral memorization and recitation of Vedic Sanskrit texts may, *prima facie*, appear extraordinary or bordering on impossible, textual memorization and recitation are in fact standard practice in traditional Sanskrit education in India (Rashtriya Sanskrit

* Corresponding author at: Center for Mind/Brain Sciences (CIMEC), Via delle Regole 101, Mattarello, TN, Italy.

E-mail address: James.Hartzell@unitn.it (J.F. Hartzell).

AAAI-17 Accepted Papers

Main Technical Track

18: Axiomatic Characterization of Game-Theoretic Network Centralities
Oskar Skibski, Tomasz Michalak and Talal Rahwan

19: On Predictive Patent Valuation: Forecasting Patent Citations and Their Types
Xin Liu, Junchi Yan, Shuai Xiao, Xiangfeng Wang, Hongyuan Zha and Stephen Chu

23: Algorithms for Max-Min Share Fair Allocation of Indivisible Chores
Haris Aziz, Gerhard Rauchecker, Guido Schryen and Toby Walsh

24: Complexity of Manipulating Sequential Allocation
Haris Aziz, Sylvain Bouveret, Jérôme Lang and Simon Mackenzie

25: Unsupervised Learning for Lexicon-Based Classification
Jacob Eisenstein

38: Universum Prescription: Regularization using Unlabeled Data
Xiang Zhang and Yann Lecun

43: Unsupervised deep learning for optical flow estimation
Zhe Ren, Junchi Yan, Bingbing Ni, Bin Liu, Xiaokang Yang and Hongyuan Zha

51: Multiple Kernel k-means with Incomplete Kernels
Xinwang Liu, Miaomiao Li, Lei Wang, Yong Dou, Jianping Yin and En Zhu

60: A Multiview-based Parameter Free Framework for Group Detection
Xuelong Li, Mulin Chen, Feiping Nie and Qi Wang

61: Quantifying and Detecting Collective Motion by Manifold Learning
Xuelong Li, Mulin Chen and Qi Wang

67: Regularized Diffusion Process for Visual Retrieval
Song Bai, Xiang Bai, Qi Tian and Longin Jan Latecki

72: SSP: Semantic Space Projection for Knowledge Graph Embedding with Text Descriptions
Han Xiao, Minlie Huang and Xiaoyan Zhu

82: Incremental Growth of Semantic Branches on CNNs via Multi-Shot Learning
Quanshi Zhang, Ruiming Cao, Ying Nian Wu and Song-Chun Zhu

Distributed Patterns of Reactivation Predict Vividness of Recollection

Marie St-Laurent¹, Hervé Abdi², and Bradley R. Buchsbaum¹

Abstract

■ According to the principle of reactivation, memory retrieval evokes patterns of brain activity that resemble those instantiated when an event was first experienced. Intuitively, one would expect neural reactivation to contribute to recollection (i.e., the vivid impression of reliving past events), but evidence of a direct relationship between the subjective quality of recollection and multiregional reactivation of item-specific neural patterns is lacking. The current study assessed this relationship using fMRI to measure brain activity as participants viewed and mentally replayed a set of short videos. We used multivoxel pattern analysis to train a classifier to identify individual videos based on brain ac-

tivity evoked during perception and tested how accurately the classifier could distinguish among videos during mental replay. Classification accuracy correlated positively with memory vividness, indicating that the specificity of multivariate brain patterns observed during memory retrieval was related to the subjective quality of a memory. In addition, we identified a set of brain regions whose univariate activity during retrieval predicted both memory vividness and the strength of the classifier's prediction irrespective of the particular video that was retrieved. Our results establish distributed patterns of neural reactivation as a valid and objective marker of the quality of recollection. ■

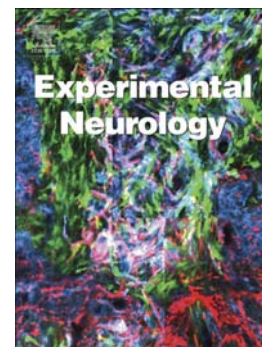
INTRODUCTION

Recollection allows people to mentally travel back in time to relive past events. The recollective experience is personal and subjective, and thus historically, it has been difficult to capture with objective measures. However, the recent application of multivoxel pattern analysis (MVPA; Haxby, 2012; Mahmoudi, Takerkart, Regragui, Boussaoud, & Brovelli, 2012; Tong & Pratte, 2012; Haynes & Rees, 2006; Norman, Polyn, Detre, & Haxby, 2006) to functional brain imaging data provides a way of examining recollection from an objective standpoint. MVPA can quantify cortical reinstatement or reactivation (Rissman & Wagner, 2012; Danker & Anderson, 2010; Rugg, Johnson, Park, & Uncapher, 2008), the phenomenon by which stimulus-specific patterns of brain activation elicited at perception are reactivated during subsequent memory retrieval. Although reactivation by itself does not guarantee the phenomenological experience of recollection (Johnson, McDuff, Rugg, & Norman, 2009), we propose that reactivation is a prerequisite for a faithful and vivid recollective experience because it reflects the specificity with which stimuli are represented in memory. With the current study, we tested this proposition by examining the correspondence between distributed patterns of stimulus-specific neural reactivation and the subjective quality of the recollective experience.

A few studies have reported a link between MVPA measures of cortical reinstatement and the subjective experience of recollection (Gordon, Rissman, Kiani, & Wagner, 2013; Ritchey, Wing, Labar, & Cabeza, 2013; Staresina, Henson, Kriegeskorte, & Alink, 2012; Kuhl, Rissman, Chun, & Wagner, 2011; Johnson et al., 2009; McDuff, Frankel, & Norman, 2009). However, most of these studies have assessed recollection as an all-or-none phenomenon that is either present or absent (e.g., with a remember/know paradigm; Tulving, 1985). For example, Johnson et al. (2009) reported significantly greater reactivation when an event is “remembered” (recollected) than when it is “known” (recognized with a sense of familiarity but not recollected). Although these results are compelling, very few studies have assessed whether reactivation also reflects the graded nature of conscious memory retrieval. Although the presence and absence of recollection at retrieval is thought to reflect the engagement of qualitatively different processes (Yonelinas, Aly, Wang, & Koen, 2010; Yonelinas, 2002), recollection also results in a complex range of experiences that vary in vividness and level of detail (e.g., Rubin, Schrauf, & Greenberg, 2003). So far, only a few reactivation studies have sampled this range of “above-threshold” recollective experience. For example, Leiker and Johnson (2014) have shown that the magnitude of task-specific pattern reactivation reflects the amount of information recollected. Johnson, Kuhl, Mitchell, Ankudowich, and Durbin (2015) have shown that aging influences correlations between classification accuracy and vividness of recall across brain ROIs. Wing, Ritchey,

¹Rotman Research Institute at Baycrest, Toronto, Ontario, Canada, ²The University of Texas at Dallas

Accepted Manuscript



A \rightarrow [C] \rightarrow cognitive \rightarrow [P] \rightarrow prostheses for \rightarrow [M] \rightarrow memory \rightarrow [F] \rightarrow facilitation by \rightarrow [C] \rightarrow closed-loop \rightarrow [F] \rightarrow functional \rightarrow [E] \rightarrow ensemble \rightarrow [S] \rightarrow stimulation of \rightarrow [H] \rightarrow hippocampal \rightarrow [N] \rightarrow neurons in \rightarrow [P] \rightarrow piriform \rightarrow [B] \rightarrow brain

Sam A. Deadwyler, Robert E. Hampson, Dong Song, Ioan Opris, Greg A. Gerhardt, Vasilis Z. Marmarelis, Theodore W. Berger

PII: S0014-4886(16)30149-2
DOI: doi: [10.1016/j.expneurol.2016.05.031](https://doi.org/10.1016/j.expneurol.2016.05.031)
Reference: YEXNR 12308

To appear in: *Experimental Neurology*

Received date: 1 December 2015
Revised date: 21 May 2016
Accepted date: 23 May 2016

Please cite this article as: Deadwyler, Sam A., Hampson, Robert E., Song, Dong, Opris, Ioan, Gerhardt, Greg A., Marmarelis, Vasilis Z., Berger, Theodore W., A \rightarrow [C] \rightarrow cognitive \rightarrow [P] \rightarrow prostheses for \rightarrow [M] \rightarrow memory \rightarrow [F] \rightarrow facilitation by \rightarrow [C] \rightarrow closed-loop \rightarrow [F] \rightarrow functional \rightarrow [E] \rightarrow ensemble \rightarrow [S] \rightarrow stimulation of \rightarrow [H] \rightarrow hippocampal \rightarrow [N] \rightarrow neurons in \rightarrow [P] \rightarrow piriform \rightarrow [B] \rightarrow brain, *Experimental Neurology* (2016), doi: [10.1016/j.expneurol.2016.05.031](https://doi.org/10.1016/j.expneurol.2016.05.031)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Derivatives Safe Harbors in Bankruptcy and Dodd-Frank: A Structural Analysis

The Harvard community has made this article openly available.
[Please share](#) how this access benefits you. Your story matters.

Citation	Stephen Adams, Derivatives Safe Harbors in Bankruptcy and Dodd-Frank: A Structural Analysis (April 30, 2013).
Accessed	October 22, 2017 1:52:37 PM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:10985175
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/318373380>

Advances in design of protein folds and assemblies

Article in *Current opinion in chemical biology* · October 2017

DOI: 10.1016/j.cbpa.2017.06.020

CITATIONS

0

READS

61

3 authors, including:



[Ajasja Ljubetič](#)

National Institute of Chemistry

20 PUBLICATIONS 40 CITATIONS

[SEE PROFILE](#)



[Roman Jerala](#)

National Institute of Chemistry

224 PUBLICATIONS 5,120 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Innate immunity [View project](#)



Protein origami [View project](#)

A HIGHLY DENSE NANONEEDLE ARRAY FOR INTRACELLULAR GENE DELIVERY

Seung-Joon Paik^{1*}, Seonhee Park², Vladimir Zarnitsyn²,
Seungkeun Choi¹, Xin Dong Guo², Mark R. Prausnitz², and Mark G. Allen^{1,2}

¹School of Electrical and Computer Engineering

²School of Chemical and Biomolecular Engineering

Georgia Institute of Technology, Atlanta, Georgia, USA

ABSTRACT

We present a dense needle array with sharp nano-tips for intracellular gene and drug delivery, with the ultimate goal of gene correction in a high throughput manner. The nanoneedles are fabricated using isotropic dry etching of silicon, followed by anisotropic dry etching and thermal-oxidation-based tip sharpening. The nanoneedle density is as large as 10^6 needles per square centimeter. *In-vitro* intracellular delivery tests demonstrated that the nanoneedle array can effectively deliver molecular probes into cells without causing significant cell death from cell membrane penetration. Uptake of molecular probes by up to 34% of cells after nanoneedle treatment was observed.

INTRODUCTION

Investigation into cellular functions requires the ability to apply specific and controlled treatment to cells. Such treatment includes the delivery of biological effectors across cell membranes using various approaches such as chemical, mechanical or electrical perturbations to cells [1]. In addition to investigation, gene correction represents an active treatment in which genetic diseases and cancer may be treated by correcting the causative mutations in the genetic code. Sickle cell disease is the most common inherited blood disorder caused by a mutation of a single nucleotide of a gene. The disease causes hemoglobin-containing red blood cells to tend to deform, clump and break apart, resulting in clogged blood vessels and causing severe pain, serious infection and organ damage. Bone marrow transplantation, the only permanent cure for the disease, has the undesirable limitation of donor compatibility. Research has shown that transforming defective blood-forming cells into normal ones by inserting corrective genes can alleviate the disease [2].

Several techniques currently exist for gene correction, including biological methods or physical injection of therapeutic agents into cells. Physical methods such as conventional microinjection typically use glass microcapillaries; however, these

techniques have relatively low throughput due to the time-consuming operation of targeting individual cell nuclei [1]. A micron-scale microcapillary array, made of SiO_2 , was previously demonstrated for fluorescent dye injection [3] and DNA injection [4] to plant cells. In this paper we propose a silicon nanoneedle array with needle tip diameters in tens of nanometers, enabling pinpoint injection of individual cells in a high throughput manner without causing cell death by puncture (Fig. 1).

FABRICATION

Fabrication proceeds by patterning a positive photoresist mask (SC1813, Shipley) on a silicon wafer bearing 0.7 micron of thermal oxide. After oxide etching, isotropic dry etching of Si with an SF_6 plasma is performed in order to achieve sharp tips utilizing under-etching of silicon under the etch mask (Figs 2a and 2b). Arrays of circular patterns of $7\mu\text{m}$ -diameter remain attached to the silicon substrate after the isotropic etching, so that the circular pattern can act as an etch mask during the subsequent anisotropic etching process (Fig. 2c). Due to the directivity from a DC bias during the isotropic etch step [5], a silicon neck region is formed several microns below the mask. After removing residual masks and cleaning, nano-scale silicon tips are formed using an oxidation sharpening process [6] (Fig. 2e). The nanoneedle array is separated into individual $5\text{mm} \times 5\text{mm}$ die and the remaining SiO_2 is removed using HF (Fig. 2f). Figure 3 shows formation of silicon nano-tips after growth of $0.8\mu\text{m}$ -thick thermal oxide. Figures 3(a) and 3(b) show SEM micrographs of structures during the fabrication process corresponding to the fabrication steps shown in Figures 2(d) and 2(f), respectively.

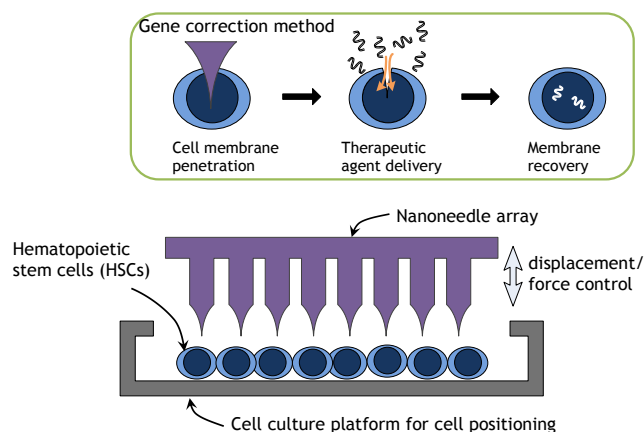


Figure 1: Schematic diagram of intracellular delivery system for insertion of therapeutic genes.

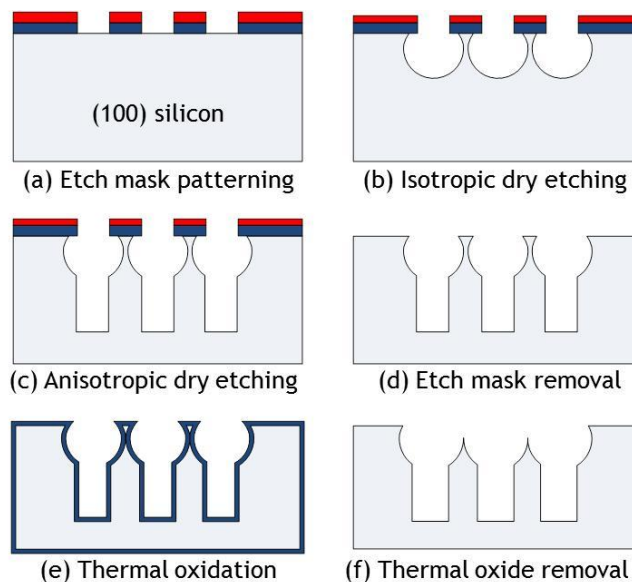


Figure 2: Fabrication process for highly dense silicon nanoneedles (photoresist in red and SiO_2 in blue).

Bacteria-mediated delivery of nanoparticles and cargo into cells

DEMIR AKIN^{1,2,3*}, JENNIFER STURGIS^{2,4}, KATHY RAGHEB^{2,4}, DEBBY SHERMAN⁵,
KRISTIN BURKHOLDER⁶, J. PAUL. ROBINSON^{2,3,4}, ARUN K. BHUNIA⁶, SULMA MOHAMMED⁷
AND RASHID BASHIR^{1,2,3,8*}

¹Birck Nanotechnology Center, Purdue University, 1205 W State Street, West Lafayette, Indiana 47907, USA

²Bindley Biosciences Center, Purdue University, West Lafayette, Indiana 47907, USA

³Weldon School of Biomedical Engineering, Purdue University, 206 S Intramural Drive, West Lafayette, Indiana 47907, USA

⁴Department of Basic Medical Sciences, Purdue University, West Lafayette, Indiana 47907, USA

⁵Department of Biology, Purdue University, West Lafayette, Indiana 47907, USA

⁶Molecular Food Microbiology Laboratory, Department of Food Science, 745 Agriculture Mall Drive, West Lafayette, Indiana 47907, USA

⁷Department of Comparative Pathobiology, 725 Harrison Street, West Lafayette, Indiana 47907, USA

⁸School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA

*e-mail: akin@purdue.edu; bashir@purdue.edu

Published online: 10 June 2007; doi:10.1038/nnano.2007.149

Nanoparticles and bacteria can be used, independently, to deliver genes and proteins into mammalian cells for monitoring or altering gene expression and protein production. Here, we show the simultaneous use of nanoparticles and bacteria to deliver DNA-based model drug molecules *in vivo* and *in vitro*. In our approach, cargo (in this case, a fluorescent or a bioluminescent gene) is loaded onto the nanoparticles, which are carried on the bacteria surface. When incubated with cells, the cargo-carrying bacteria ('microbots') were internalized by the cells, and the genes released from the nanoparticles were expressed in the cells. Mice injected with microbots also successfully expressed the genes as seen by the luminescence in different organs. This new approach may be used to deliver different types of cargo into live animals and a variety of cells in culture without the need for complicated genetic manipulations.

One of the most significant challenges facing the treatment of diseases is early intervention to deliver specific therapeutic cargo efficiently into cells to alter gene expression and subsequent protein production. Recent advances in nanotechnology have been used to deliver such cargoes into single cells through the use of nanoparticles for imaging^{1–3}, diagnostics^{4,5} and therapeutics^{6–8}. Although significant advances have been made, many difficulties remain in delivering the nanoparticles to the tumour sites, mainly because of the physical barriers encountered in solid tumours, such as malformed blood supplies, elevated interstitial pressure, and large transport distances in the tumour interstitium^{9,10}.

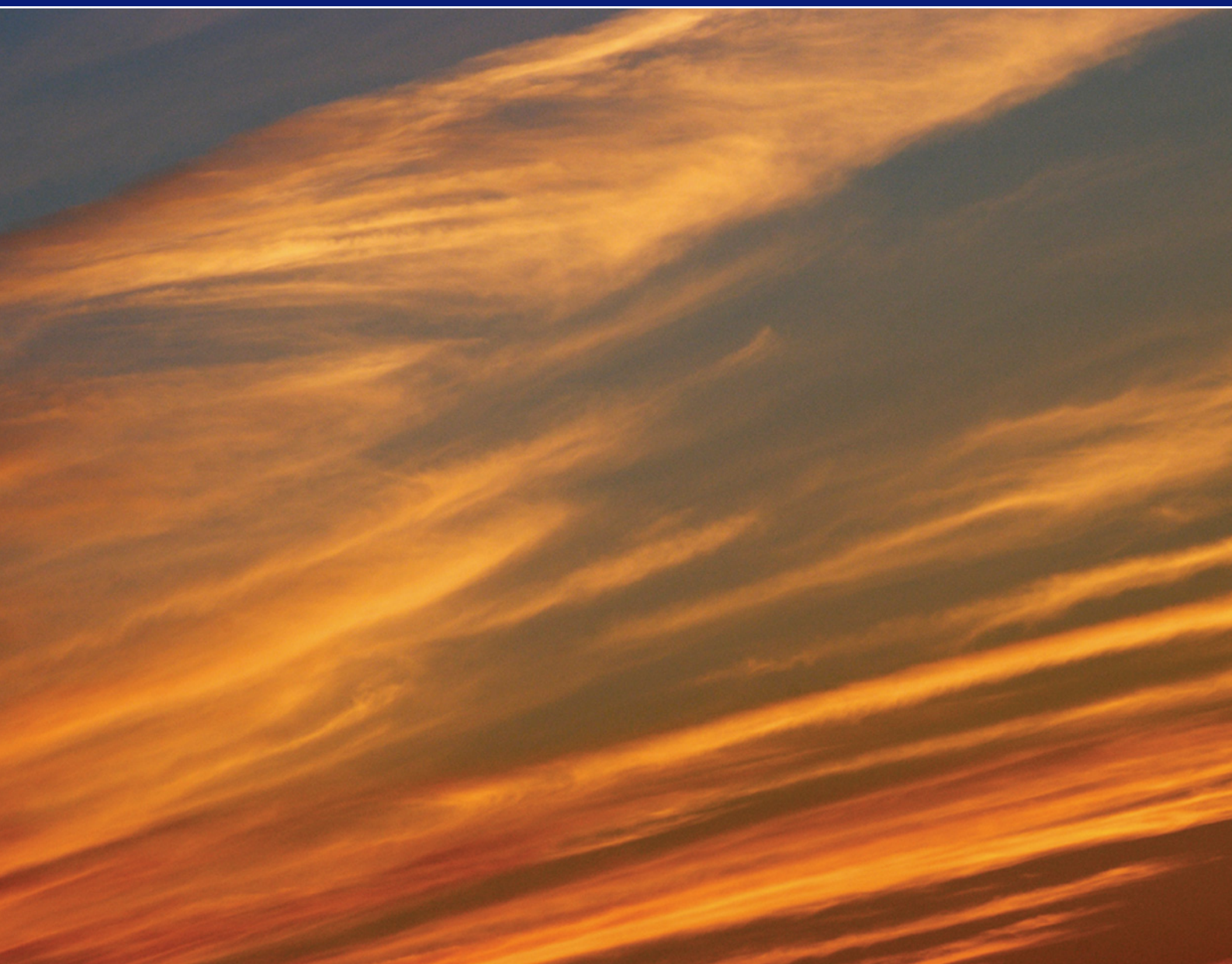
Bacteria have been used as a non-viral means to transfer plasmid DNA into mammalian cells through a process called 'bactofection' (reviewed in ref. 11). Several intracellular bacteria, including *Listeria monocytogenes*, which is responsible for food-borne infections in humans and animals¹², can penetrate mammalian cells that are normally non-phagocytic. These bacteria need specific surface molecules that interact with host-cell receptors for this invasion step^{13–15} once inside the cells, the bacteria carriers are disrupted—by treatment with antibiotics—and the DNA is released. *L. monocytogenes*-based bactofection systems have shown efficient transfer of genetic material inside the cells^{16,17}. Other earlier reports include use of attenuated (reduced infectivity) bacteria such as *Shigella*¹⁸ and *Salmonella typhimurium*^{19,20} for the delivery of DNA-based vaccines. Bacteria

themselves have additional advantages as delivery systems. For example, attenuated strains of *Escherichia coli*, *S. typhimurium*, *Vibrio cholerae* and *L. monocytogenes* have been shown to be capable of multiplying selectively in tumours²¹, and in the case of *Clostridium* and *Bifidobacterium* spp., they even inhibit tumour growth^{20,22}. Some of the unique properties of attenuated *Listeria* strains make them an ideal non-viral gene delivery vehicle^{23–25}. It should also be noted that antibiotics can control bacterial replication in the body or activate gene-based therapeutic molecules, as in the case with tetracycline-regulated control of gene expression²⁶.

Here, we report a novel technique for delivery of nanoparticles into cells, which takes advantage of the invasive properties of bacteria. The gene or cargo is not carried inside the bacteria, but rather remains on the surface conjugated to nanoparticles. Hence, our approach does not require bacterial disruption for delivery, or any genetic engineering of the bacteria for different cargo. Although more than one gene can be delivered by means of bactofection, many more copies of a target cargo can be carried with one bacterium using the method described here. We also show that nucleic acid-based model drugs (plasmid DNAs coding for green fluorescence protein (GFP), luciferase and secreted alkaline phosphatase (SEAP)) loaded on the nanoparticles can be released from the carriers and eventually find their way into the nucleus, with subsequent transcription and translation of their respective proteins, for both *in vitro* and

An Analysis of Climate Engineering as a Response to Climate Change

J Eric Bickel & Lee Lane



ASICs and Decentralization FAQ

1 About this document.

1. What is the thesis of this document?

The main thrust of this document is that proof-of-work schemes ought to be as simple and dependent on raw computational power as possible. That is, a proof-of-work should tend toward the thermodynamic limit (see a later section for this term) as quickly and directly as possible.

2. Is Proof of Work interesting?

Not really ☹. Replacing or tweaking the proof-of-work algorithm is one of the most popular changes to Bitcoin done by copycat “alt” currencies, but it does not enable any new use cases or features for users of the currency. It is in effect a bikeshed painting¹ change, though as we will see it is not without its dangers.

3. Why did you write this document?

For two reasons:

- (a) to organize and lay out some folklore about proof-of-work which has not been written down in one place;
- (b) to answer some common questions and suggestions regarding Bitcoin’s proof-of-work and those of other currencies (which as established above, is not interesting and therefore no fun to answer repeatedly).

2 What is Proof of Work?

1. What is Proof of Work?

As it applies to Bitcoin, a *proof of work* is a computational proof that some scarce resource was consumed. Such a proof is possible because it appears that computational resources are physically bounded by available time, space and energy²

Further, a proof of work commits to some data, effectively “signing” it with some consumed resource.

2. How and why is it used?

¹See <http://bikeshed.com/> for a history of this phrase.

²See, for example, <http://arxiv.org/abs/quant-ph/0502072> for a discussion on the emergence of fundamental computational limits from physics.



Review in Advance first posted online
on May 10, 2017. (Changes may
still occur before final publication
online and in print.)

A Single-Molecule View of Genome Editing Proteins: Biophysical Mechanisms for TALEs and CRISPR/Cas9

Luke Cuculis¹ and Charles M. Schroeder^{1,2}

¹Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801;
email: cms@illinois.edu

²Department of Chemical & Biomolecular Engineering, University of Illinois at
Urbana-Champaign, Urbana, Illinois 61801

Annu. Rev. Chem. Biomol. Eng. 2017.
8:25.1–25.21

The *Annual Review of Chemical and Biomolecular
Engineering* is online at chembioeng.annualreviews.org

<https://doi.org/10.1146/annurev-chembioeng-060816-101603>

Copyright © 2017 by Annual Reviews.
All rights reserved

Keywords

genome editing, TALEN, CRISPR/Cas9, single molecule, DNA search

Abstract

Exciting new advances in genome engineering have unlocked the potential to radically alter the treatment of human disease. In this review, we discuss the application of single-molecule techniques to uncover the mechanisms behind two premier classes of genome editing proteins: transcription activator-like effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated system (Cas). These technologies have facilitated a striking number of gene editing applications in a variety of organisms; however, we are only beginning to understand the molecular mechanisms governing the DNA editing properties of these systems. Here, we discuss the DNA search and recognition process for TALEs and Cas9 that have been revealed by recent single-molecule experiments.

pombe that will complement the *spt15-21* mutation in FW1259, our approach of functional cloning using this comprehensive library will enable us to search exhaustively for any genes encoding minor forms of TFIID-like proteins that may be expressed at low levels. □

Received 5 April; accepted 15 May 1990.

1. Forsburg, S. & Guarente, L. *A. Rev. Cell. Biol.* **5**, 153-180 (1989).
2. Struhl, K. *Cell* **49**, 295-297 (1987).
3. Johnson, P. & McKnight, S. A. *Rev. Biochem.* **58**, 799-839 (1989).
4. Mitchell, P. J. & Tjian, R. *Science* **228**, 371-378 (1989).
5. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. *Cell* **56**, 549-561 (1989).
6. Hawley, D. & Roeder, R. *J. Biol. Chem.* **262**, 3452-3461 (1987).
7. Van Dyke, M. W., Roeder, R. G. & Sawadogo, M. *Science* **241**, 1335-1338 (1988).
8. Buratowski, S., Hahn, S., Sharp, P. & Guarente, L. *Nature* **334**, 37-42 (1988).
9. Cavallini, B. *et al. Nature* **334**, 77-80 (1988).
10. Horikoshi, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 4843-4847 (1989).
11. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. *Cell* **58**, 1173-1181 (1989).
12. Eisenmann, D. M., Dollard, C. & Winston, F. *Cell* **58**, 1183-1191 (1989).
13. Horikoshi, M. *et al. Nature* **341**, 299-303 (1989).

14. Schmidt, M. C., Kao, C. C., Pei, R. & Berk, A. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7785-7789 (1989).
15. Cavallini, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9803-9807 (1989).
16. Ammerer, G. *Meth. Enzym.* **101**, 192-201 (1983).
17. Winston, F., Chaleff, D., Valent, B. & Fink, G. *Genetics* **107**, 179-197 (1984).
18. Matsumoto, D. & Yanagida, M. *EMBO J.* **4**, 3531-3538 (1985).
19. Russell, P. & Hall, B. D. *J. Biol. Chem.* **258**, 143-149 (1983).
20. Russell, P. R. *Gene* **40**, 125-130 (1985).
21. Hoeijmakers, J. J. *Nature* **343**, 417-418 (1990).
22. Nagai, K. *Nature* **343**, 418 (1990).
23. Sherman, F., Fink, G. R. & Hicks, J. B. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, New York, 1986).
24. Rothstein, R. J. *Meth. Enzym.* **101**, 202-211 (1983).
25. Henikoff, S. *Gene* **28**, 351-359 (1984).
26. Del Sal, G., Manfioletti, G. & Schneider, D. *Biotechniques* **7**, 514-519 (1989).
27. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4767-4771 (1987).
28. Sheen, J.-Y. & Seed, B. *Biotechniques* **6**, 942-944 (1988).

ACKNOWLEDGEMENTS. This work represents equivalent efforts by J.D.F. and D.M.B., who should be considered equally as first author. We thank Biorad for loan of an electroporator, Drs David Eide and Gerry Fink for strains, and Drs Eide and Olivia Bermingham-McDonogh for comments on the manuscript. D.M.B. is indebted to Dr A. Lee Osterman (Hospital of the University of Pennsylvania) for his expertise. This work was supported by grants from the NIH (L.G. and F.W.) and postdoctoral fellowships from the American Cancer Society (J.D.F. and D.M.B.). The nucleotide sequence reported here has been deposited in the EMBL, GenBank and DDBJ databases under accession number X53383.

Atomic-scale imaging of DNA using scanning tunnelling microscopy

Robert J. Driscoll, Michael G. Youngquist
& John D. Baldeschwieler*

Division of Chemistry and Chemical Engineering,
Arthur Amos Noyes Laboratory of Chemical Physics, Mail Code 127-72,
California Institute of Technology, Pasadena, California 91125, USA

THE scanning tunnelling microscope (STM) has been used to visualize DNA¹ under water², under oil³ and in air⁴⁻⁶. Images of single-stranded DNA have shown that submolecular resolution is possible⁷. Here we describe atomic-resolution imaging of duplex DNA. Topographic STM images of uncoated duplex DNA on a graphite substrate obtained in ultra-high vacuum are presented that show double-helical structure, base pairs, and atomic-scale substructure. Experimental STM profiles show excellent correlation with atomic contours of the van der Waals surface of A-form DNA derived from X-ray crystallography. A comparison of variations in the barrier to quantum mechanical tunnelling (barrier-height) with atomic-scale topography shows correlation over the phosphate-sugar backbone but anticorrelation over the base pairs. This relationship may be due to the different chemical characteristics of parts of the molecule. Further investigation of this phenomenon should lead to a better understanding of the physics of imaging adsorbates with the STM and may prove useful in sequencing DNA. The improved resolution compared with previously published STM images of DNA may be attributable to ultra-high vacuum, high data-pixel density, slow scan rate, a fortuitously clean and sharp tip and/or a relatively dilute and extremely clean sample solution. This work demonstrates the potential of the STM for characterization of large biomolecular structures, but additional development will be required to make such high resolution imaging of DNA and other large molecules routine.

We have used a Caltech-constructed ultra-high vacuum (UHV) STM system similar to the IBM 'pocket-size' STM⁸. On gallium arsenide our vertical resolution has been as high as ~1 pm (unpublished data). The system incorporates *in vacuo* tip and sample transfer and has a base pressure of ~3 × 10⁻¹¹ torr.

Although UHV provides a clean experimental environment it may cause some DNA denaturation; intact DNA is probably

TABLE 1 Comparison of DNA dimensions derived from the STM and X-ray crystallography

	STM	X-ray
Helix pitch	29 Å	28.5 Å
Minor groove width	10 Å	11.0 Å
Major groove width	3 Å	2.7 Å
Molecular width	23 Å	23 Å
Phosphate backbone width	10 Å	11.6 Å
Axial nucleotide rise	2.6 Å	2.59 Å
Base-pair angle	+18°	+19°
Helix symmetry	11	11.0

Comparison of the dimensions of general features in the STM image and dimensions derived from X-ray crystallography for 'random-sequence' A-DNA. The STM dimensions are average approximate values based on Fig. 1a. STM differs from X-ray crystallography in its ability to study local dimension and orientation variations on an individual molecule. Crystallography data reflect the average dimensions of many molecules in a crystal and therefore cannot be used to study local structural variations.

in an A-type conformation⁹. The conventional A form, determined by X-ray crystallography, is characterized by a width of ~23 Å, helical symmetry of 11 base pairs per turn, a +19° base-pair tilt to the helix axis, a 28.5 Å pitch and an axial nucleotide rise of 2.59 Å (ref. 9). Subtracting the van der Waals width (11.6 Å) of the backbone phosphate groups reveals that the major groove is narrow and deep (2.7 Å and 13.5 Å, respectively) and the minor groove is wide and shallow (11.0 Å and 2.8 Å).

Figure 1a shows an image of DNA from our UHV STM. The double helix and the major-minor groove alternation are apparent, as are parallel features spanning the minor grooves at a +18° ± 3° angle to the helix axis. Their orientation and spacing are consistent with base pairs bridging the thick phosphate backbones (~10 Å wide; instrument calibrated using the graphite atomic lattice). Roughly four turns of the helix are shown; the helix symmetry estimated from the bases visible in the minor grooves is ~11 base pairs per turn, consistent with A-DNA. The bottom portion of Fig. 1a shows a wide (~10 Å) minor groove with considerable substructure leading into a narrow (~3 Å), less-resolved major groove, consistent with the structure of A-DNA. The major groove is too narrow for the STM tip to enter. During a scan line in the second minor groove, a tip instability occurred. After this event the DNA molecule looks slightly different; the major groove appears as a notch on the right side of the molecule in the transition between one minor groove to the next. This may be explained by the tip having moved the upper half of the DNA slightly. But this same change in orientation is evident in previous images and probably

* To whom correspondence should be addressed.



An Epigenetics-Inspired DNA-Based Data Storage System

Clemens Mayer[†], Gordon R. McInroy[†], Pierre Murat[†], Pieter Van Delft, and Shankar Balasubramanian*

Abstract: Biopolymers are an attractive alternative to store and circulate information. DNA, for example, combines remarkable longevity with high data storage densities and has been demonstrated as a means for preserving digital information. Inspired by the dynamic, biological regulation of (epi)genetic information, we herein present how binary data can undergo controlled changes when encoded in synthetic DNA strands. By exploiting differential kinetics of hydrolytic deamination reactions of cytosine and its naturally occurring derivatives, we demonstrate how multiple layers of information can be stored in a single DNA template. Moreover, we show that controlled redox reactions allow for interconversion of these DNA-encoded layers of information. Overall, such interlacing of multiple messages on synthetic DNA libraries showcases the potential of chemical reactions to manipulate digital information on (bio)polymers.

Means to access, circulate, and preserve information have shaped human society by increasing knowledge, stimulating the economy and enriching culture. In this respect, the development of optical and magnetic storage devices has facilitated an unprecedented increase of accessible information, but their limited shelf lives and storage densities have prompted a search for alternative data carriers.^[1] Current lines of research focus on further increasing storage densities by compacting information into single atoms,^[2] supramolecular systems,^[3] or biopolymers.^[4] Nucleic acids, for example, are remarkably compact and long-lived, and have been proposed for storing digital information. The advent of high-throughput oligonucleotide synthesis^[5] and DNA sequencing^[6] has allowed DNA-based data storage to rapidly

progress from proof-of-concept studies toward systems that can rival established storage media.^[7] While such systems have enabled writing and reading of non-trivial amounts of information with synthetic DNA templates, the “one template, one information layer” coding scheme employed (Figure 1 A) is in stark contrast to nature’s dynamic control over the primary information encoded in genomes. In order to produce a complex organism from a single genetic makeup, cells regulate access to different layers of information by modifying histone proteins and DNA nucleobases (Figure 1 B).^[8] This epigenetic regulation orchestrates processes such as gene expression and ultimately drives cell differentiation. Herein, we apply principles from biological regulation toward DNA data storage, through the controlled chemical transformations of nucleobases^[9] and their associated binary value. As a result, we were able to (reversibly) recover multiple layers of binary data from a single DNA template (Figure 1 C).

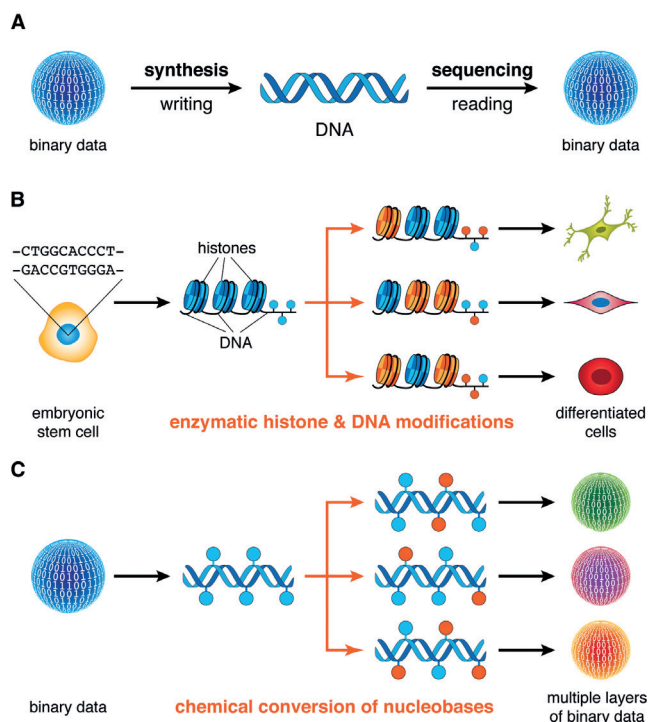


Figure 1. Emulating nature’s control over biopolymers for information storage. A) In DNA-based data storage binary data is written onto oligonucleotides by synthesis and read by sequencing on demand. B) The epigenome (histone and DNA modifications) of an embryonic stem cell undergoes controlled changes upon cell differentiation leading to the development of different phenotypes from the same genetic information. C) The use of selective chemical reactions facilitates alteration of the primary sequence of a synthetic DNA template, and thus the retrieval of multiple layers of information from it.

[*] Dr. C. Mayer,^[‡] Dr. G. R. McInroy,^[‡] Dr. P. Murat,^[‡] Dr. P. Van Delft, Prof. S. Balasubramanian
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge, CB2 1EW (UK)
E-mail: sb10031@cam.ac.uk
Prof. S. Balasubramanian
Cancer Research, UK, Cambridge Institute, Li Ka Shing Centre
University of Cambridge
Robinson Way, Cambridge, CB2 0RE (UK)
and
School of Clinical Medicine, University of Cambridge
Cambridge, CB2 0SP (UK)

[‡] These authors contributed equally to this work.

Supporting information for this article can be found under:
<http://dx.doi.org/10.1002/anie.201605531>.

© 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Bacterial morphology: why have different shapes?

Kevin D Young

The fact that bacteria have different shapes is not surprising; after all, we teach the concept early and often and use it in identification and classification. However, why bacteria should have a particular shape is a question that receives much less attention. The answer is that morphology is just another way microorganisms cope with their environment, another tool for gaining a competitive advantage. Recent work has established that bacterial morphology has an evolutionary history and has highlighted the survival value of different shapes for accessing nutrients, moving from one place to another, and escaping predators. Shape may be so important in some of these endeavors that an organism may change its morphology to fit the circumstances. In short, if a bacterium needs to eat, divide or survive, or if it needs to attach, move or differentiate, then it can benefit from adopting an appropriate shape.

Address

Department of Microbiology and Immunology, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND 58202-9037, USA

Corresponding author: Young, Kevin D (kyoung@medicine.nodak.edu)

Current Opinion in Microbiology 2007, **10**:596–600

This review comes from a themed issue on
Prokaryotes
Edited by Martin Dworkin

Available online 5th November 2007

1369-5274/\$ – see front matter
© 2007 Elsevier Ltd. All rights reserved.

DOI [10.1016/j.mib.2007.09.009](https://doi.org/10.1016/j.mib.2007.09.009)

Introduction

The discussion of bacterial morphology has been dominated by questions about how a cell manages to create a rod shape, which, of course, is but one example of the more general question of how a cell constructs *any* shape. The expectation is that by answering this (deceptively) simple question we may acquire knowledge that will point us to a universal mechanism of shape control. This emphasis is understandable because we are both more familiar with and more comfortable with answering *how*-type questions. And, indeed, this approach has produced exciting new information, highlighted by other articles in this issue.

What has not been as well explored is *why* bacteria find it advantageous to exhibit such a prodigious number of different shapes; and so the purpose of this article is to examine some of the reasons that lie behind this variety. I will highlight a few research areas that bear on why bacteria have certain morphologies, but only in a brief

and qualitative way. More depth, more examples, and a bit more quantitative treatment can be found in a recent review and the references therein [1^{••}]. Portions of this topic have also been discussed by Beveridge [2], Dusenbery [3[•]], Koch [4], and Mitchell [5^{••}].

Shape has selective value

The first issue to get settled is that the shape of a bacterium has biological relevance. One argument favoring this assertion is that even though bacteria have a wide variety of shapes, any one genus typically exhibits a limited subset of morphologies, hinting that, with a universe of shapes to choose from, individual bacteria adopt only those that are adaptive. Another clue is that some bacteria can modify their morphology in response to environmental cues or during the course of pathogenesis (e.g. reference [6[•]]), suggesting that shape is important enough to merit regulation.

Two evolutionary arguments also support the utility of bacterial shape. Firstly, shape has a vector through evolutionary time—rod-like organisms having arisen first and coccoid forms being derivatives at the ends of evolutionary lines [7–11]. Progressive development of a trait implies that selective forces are operating. Secondly, prokaryotes with different genealogies may converge morphologically, indicating that a similar shape may confer advantages in certain environments. So, for example, although they have a non-peptidoglycan-based cell wall, the Archaea exhibit a range of morphological forms similar to that of the bacteria [12]. The simplest conclusion is that morphological adaptation serves an important biological function.

How, then, might morphology contribute to natural selection? Simply put, bacteria with different shapes present different physical features to the outside world, and these features help cells cope with and adapt to external conditions. Even a 0.01% increase in the growth rate of *E. coli* can impart a fitness advantage of ~10% compared to its unaltered competitors [5^{••}], so improvements need not be dramatic to be useful. Consistent with these expectations, shape contributes a measure of survival value in the face of three ‘Primary’ selective pressures: (1) nutrient acquisition, (2) cell division, and (3) predators; and in optimizing five ‘Secondary’ mechanisms: (4) attachment to surfaces, (5) passive dispersal, (6) active motility, and (7) internal or (8) external differentiation [1^{••}] (Table 1). The first three are Primary in that they represent fundamental conditions that determine whether cells live or die because cells must grow and multiply and resist being killed. The last five are Secondary in that they represent a suite of

Beyond editing to writing large genomes

Raj Chari and George M. Church

Abstract | Recent exponential advances in genome sequencing and engineering technologies have enabled an unprecedented level of interrogation into the impact of DNA variation (genotype) on cellular function (phenotype). Furthermore, these advances have also prompted realistic discussion of writing and radically re-writing complex genomes. In this Perspective, we detail the motivation for large-scale engineering, discuss the progress made from such projects in bacteria and yeast and describe how various genome-engineering technologies will contribute to this effort. Finally, we describe the features of an ideal platform and provide a roadmap to facilitate the efficient writing of large genomes.

DNA reading (sequencing) and writing (synthesis) are intimately connected: meaningful use of either requires the other. For example, the most widely used sequencing platforms read DNA during the synthesis of a complementary strand^{1–3}. Targeted DNA reading, such as exome sequencing, depends on vast libraries of synthetic oligonucleotides for target capture, and the annotation and understanding of genome reads benefit greatly from genome-scale synthesis and testing. Furthermore, meaningful use of genome engineering generally requires whole-genome sequencing to design strategies for avoiding off-target writing and to verify the quality of the final genome product. Genome editing and writing have benefited from diverse nanomachines harvested by reading the biosphere. Since 1975, reading and writing platforms have exhibited increases in throughput of three-million-fold (for reading; see <https://www.genome.gov/sequencingcostsdata/>) and one-billion-fold (for writing of raw oligonucleotides)⁴, with a steep increase at an exponential rate since 2004. The technical progress of both reading and writing has depended not only on mere automation of human protocols but also on rethinking the whole process to permit miniaturization and multiplexing (in which millions of reactions

can be performed in the same space, at the same cost and with the same effort as a single reaction).

Although much of the attention has been on the progress in DNA reading, namely, advances in DNA sequencing technologies^{2,3}, the improvements in writing, which can largely be attributed to harnessing and stimulating natural cellular processes such as homologous recombination (HR) to enable genome modification, should not be overlooked. From the initial work of describing HR at a very low level in higher-order eukaryotic cells to current work, where we can achieve reasonably high efficiencies of HR in stem cells^{4–7} (FIG. 1), our ability to direct targeted changes in genomes has opened a large number of avenues that previously may not have seemed possible.

One of the main areas that can be realistically pursued is the ability to perform large-scale editing of genomes, even multiplexing libraries of billions of changes⁸. From an industrial perspective to a health care perspective and to researchers aiming for fundamental understanding of how biological systems work, the motivations to pursue ambitious genome-engineering projects originate from a diverse group of stakeholders, and the results of these projects could have considerable positive impact on multiple aspects of society. Combining

the technological advancements in genome engineering with those in both DNA synthesis and DNA sequencing, the key components that are necessary to approach large-scale genome engineering have fallen into place.

In this Opinion article, we describe the motivation for and current state of large-scale genome engineering and compare strategies for how it can be achieved. We provide an overview of key genome-engineering technologies and how they can be utilized for large-scale editing. Finally, we detail the key features that will be necessary for a robust genome-writing platform and provide a roadmap to engineer a large genome. As seen for whole-genome reading, the cost and quality of methods for genome writing are improving by factors of millions, and rigorous, comprehensive tests for the accuracy of genome writing will shift from being a luxury to becoming essential and routine.

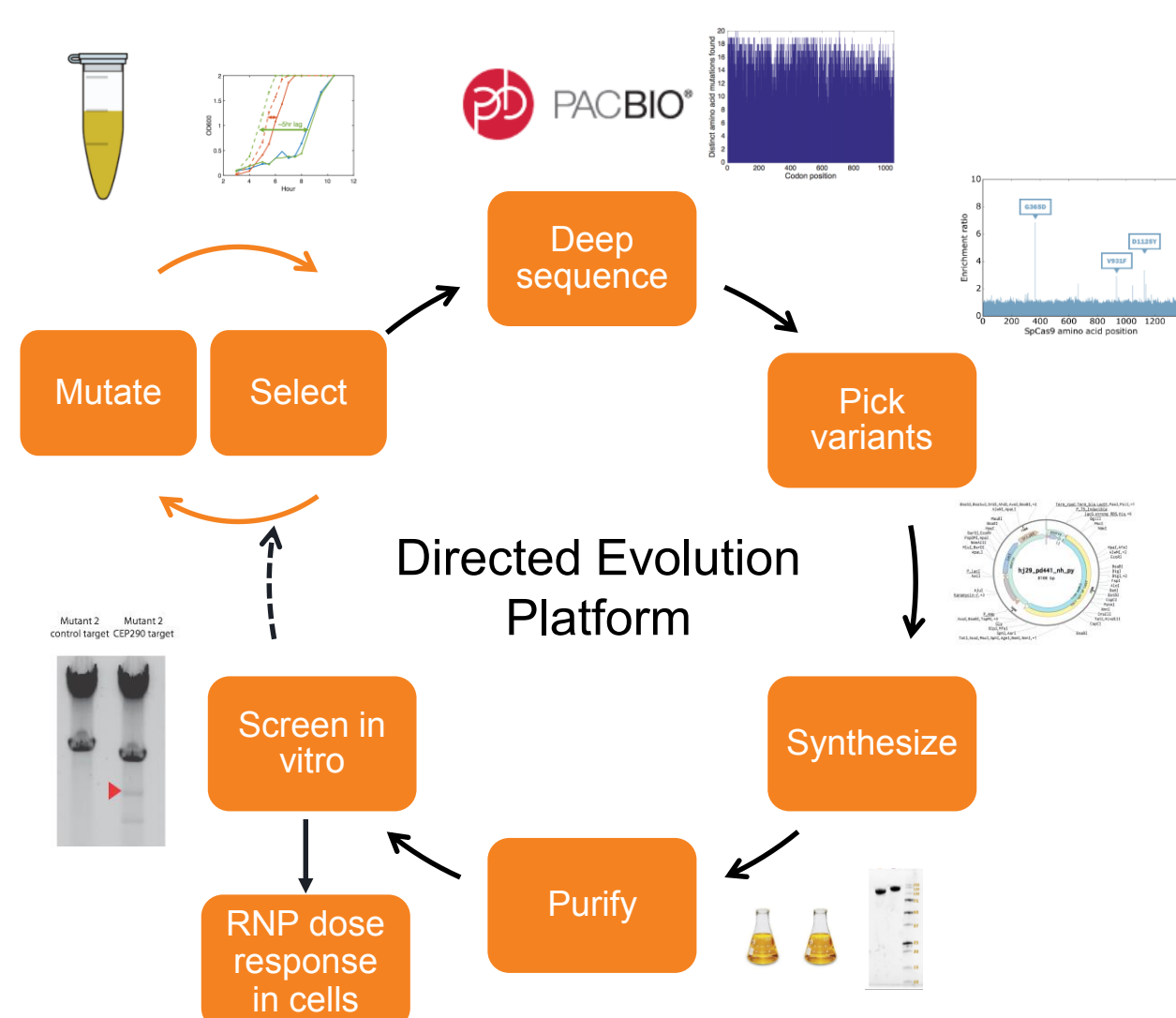
Why should we engineer genomes?

When we refer to genome engineering and genome editing, our use of ‘genome’ specifically means large-scale engineering or alteration of several loci across the genome. We feel that such terminology is unhelpfully misapplied when referring to single-locus applications, such as using genome engineering to refer to editing that is done directly in the genome rather than in a plasmid or transgenic setting or using the term genome editing when the changes are neither precise editing nor genome scale. The question for this section is why should we design and generate whole (or nearly whole) genomes? An initiative known as *Genome Project-write* has recently advocated large-scale genome engineering⁹. The motivation and goals are well aligned for both small-scale and large-scale genome engineering. With respect to non-human organisms, bacteria are common chassis for various bioproduction applications, from uses in bioremediation to pharmaceuticals to bioenergy^{10–12} (FIG. 2). From a purely functional perspective, re-engineering the genomes of particular strains could render the bacteria less susceptible to viruses and/or more efficient at producing a biomolecule of interest^{13,14}. Similarly, yeast strains, plants

Abstract

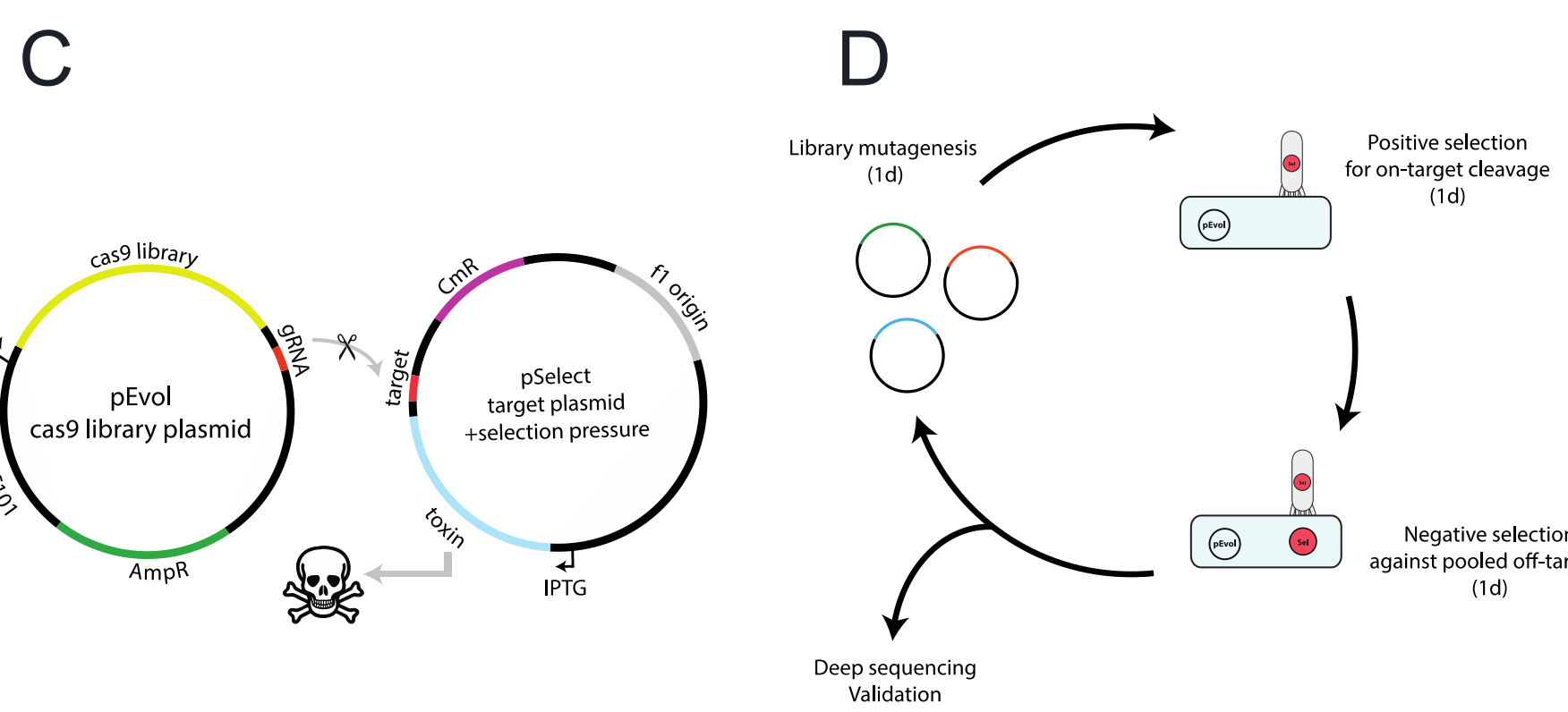
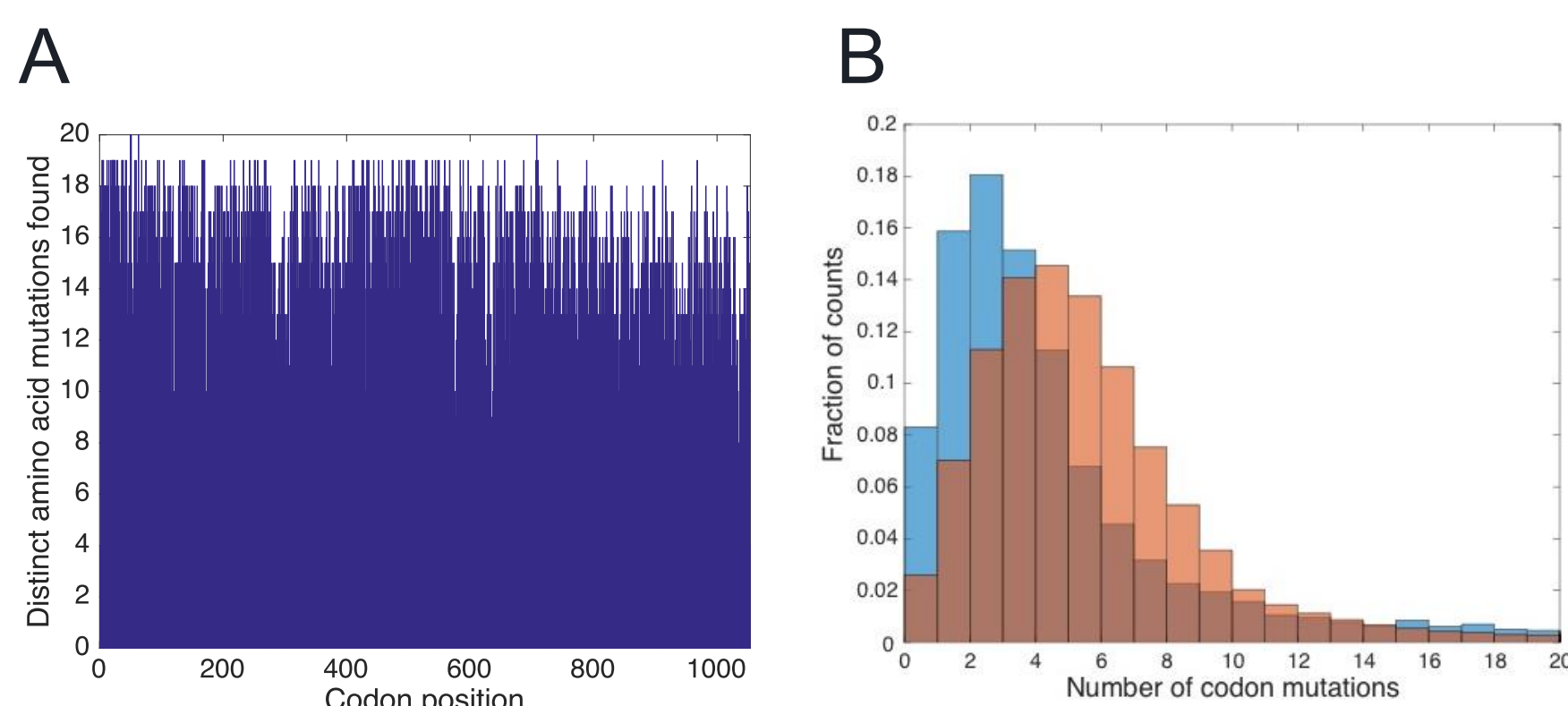
CRISPR/Cas9 has wide reaching scientific and therapeutic applications, but off-target cleavage can potentially represent a safety risk in therapeutic use. Several cas9 variants (Hi-Fi Cas9, e-Cas9) have been engineered to generally increase specificity; however, off-targets may still remain, especially when limited in guide choice. Further these variants are known to have reduced efficacy at some on target sites. In these cases it may be more beneficial to selectively mitigate specific off-targets while maintaining on-target activity. We identified three potential off-targets via GUIDE-Seq when treating T-cells with a given cas9-guide complex. In order to quickly engineer new nuclease variants with decreased cleavage of these off-targets, we developed a directed evolution platform to positively select for DNA cleavage in bacteria at on-targets and negatively select against undesired target cleavage. Our system allows rapid mutation of all amino acids throughout the protein, expanding potential diversity of functional mutants. These mutant libraries are challenged with phage containing target sites in a competitive pool. Using this method, we develop *S. pyogenes* cas9 variants which have maintained on-target cleavage efficiency but have reduced cutting at off-target loci, providing methods to potentially rescue promiscuous guides for therapeutic use.

Directed Evolution Overview



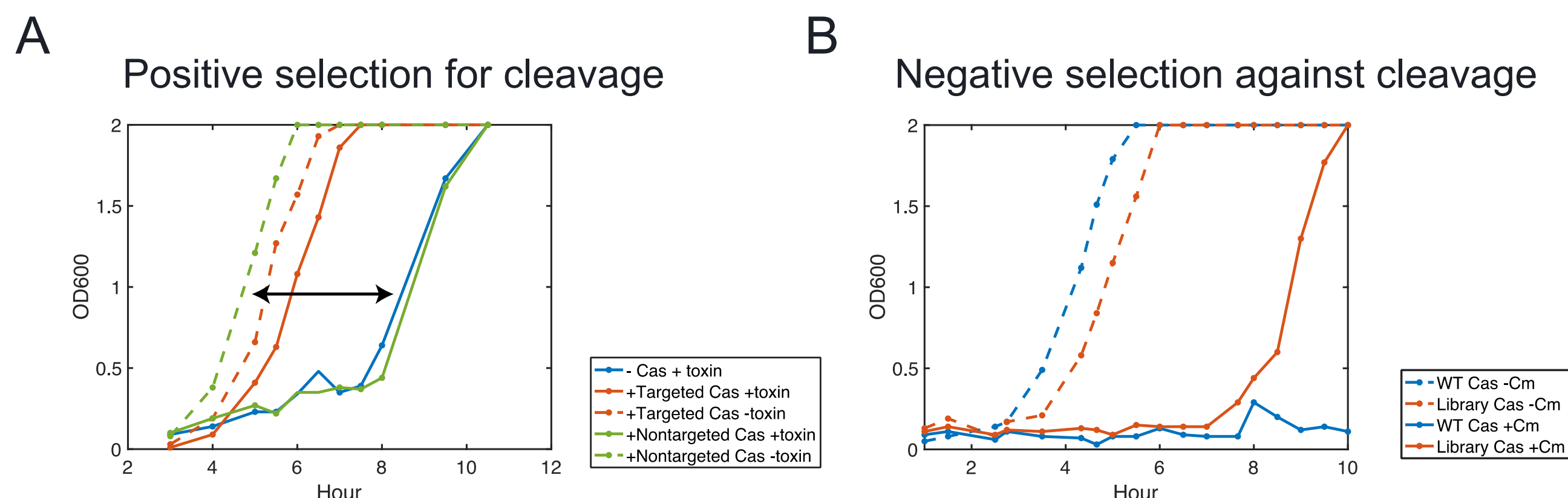
Our directed evolution platform begins with generating SpCas9 libraries and subjecting them to several rounds of selection. After deep sequencing of selected libraries, we synthesize and purify mutants of interest for further downstream characterization.

SMART library generation comprehensively mutates every amino acid



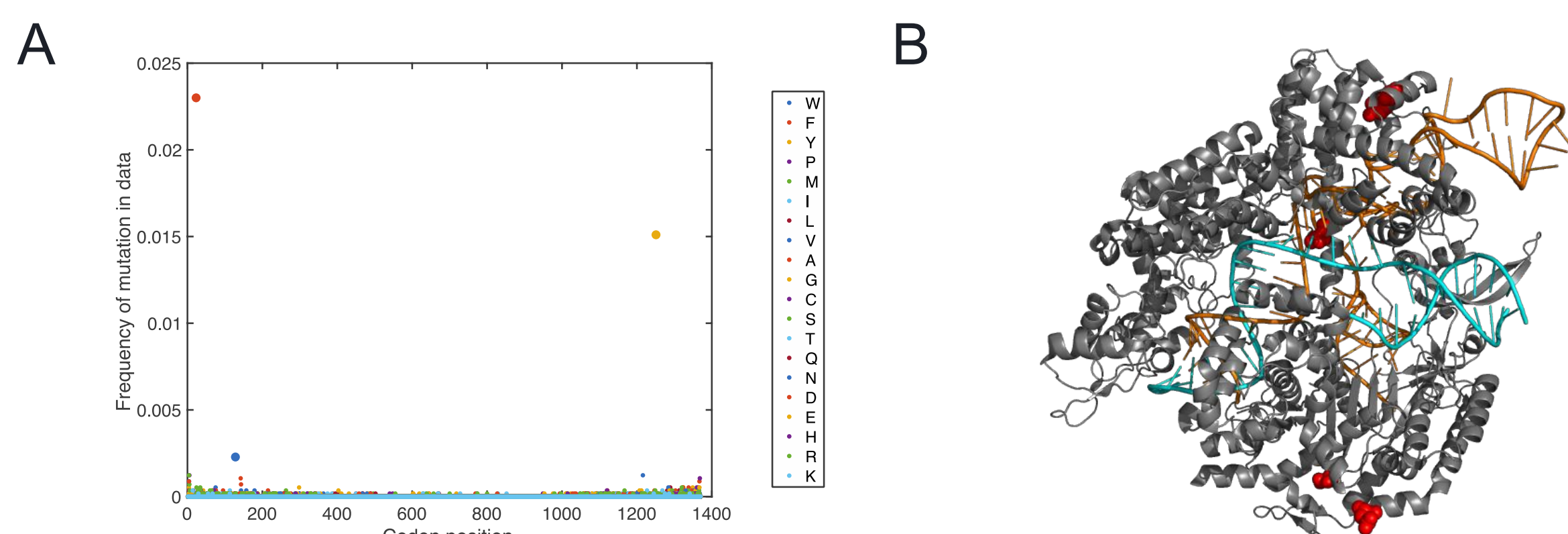
Mutant SpCas9 libraries were engineered using scanning mutagenesis at random targets (SMART). The process begins with a modified PCR step, in which forward oligos, containing codon mismatches, introduce degeneracy at every amino acid position across the protein (A). Undesired reaction products were removed from the library with a cocktail of nucleases. The number of mutations can be adjusted depending on reaction conditions (B). Libraries were then transformed into *E. coli* and challenged with phage for three rounds of both positive and negative selection (C, D).

Selection against off-target cleavage is effective in liquid culture



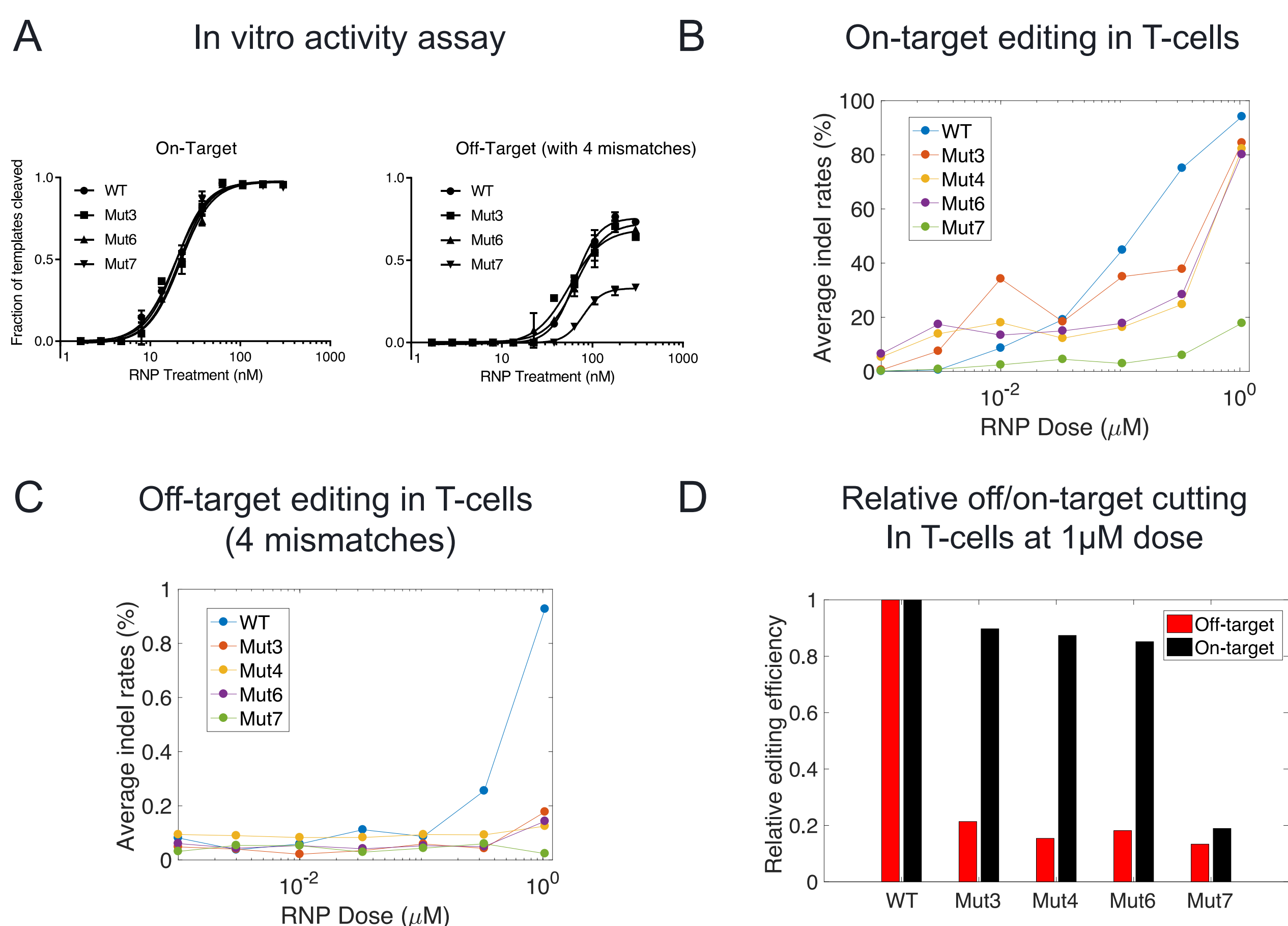
We challenge SpCas9 library plasmids to cleave phage DNA containing a given target. Positive or negative selection can be induced depending on desired activity of cleavage of each target. Growth curves were measured under varying conditions to assess the selection stringency of our system. The toxin greatly reduces the growth rate of cells in liquid media, as indicated by the solid green and blue lines (A). Expression of Cas and successful Cas9 cleavage of the target rescues growth and provides a competitive advantage, as shown by the black arrow. Phage can also be used for negative selection by selecting for cells retaining chloramphenicol resistance (B). Because our platform allows for competitive selection in liquid culture, we can challenge variants against several selective pressures in a single round. Each round also functions as a true selection rather than a screen for a single cleavage event, as seen in other platforms.

Deep sequencing identifies mutants of interest



After three rounds of directed evolution, libraries were subjected to PacBio next-generation sequencing. Several residues were enriched throughout the protein (as shown in red spheres). Combinations of these mutations were synthesized and tested *in vitro*.

Mutants have reduced off-target editing in T-cells while maintaining similar on-target activity



We synthesized seven mutant proteins to test *in vitro* (A) and in T-cells (B-D). Cells were treated with each variant in a dose-response followed by amplification at on- and off-targets and next-generation sequencing to determine indel rates at each locus. All synthesized variants showed decreased relative off-target editing at an off-target site relative to the wildtype enzyme (WT), while many variants maintained comparable levels of editing efficiency at the on-target site. *In vitro* specific activity indicates that genomic context is important for relative function of these nucleases pointing to the merits of a cellular selection for positive and negative selection for on and off-target editing.

BioSpaun: A large-scale behaving brain model with complex neurons

Chris Eliasmith, Jan Gosmann, Xuan Choo

Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, Ontario, Canada

Abstract

We describe a large-scale functional brain model that includes detailed, conductance-based, compartmental models of individual neurons. We call the model BioSpaun, to indicate the increased biological plausibility of these neurons, and because it is a direct extension of the Spaun model [1]. We demonstrate that including these detailed compartmental models does not adversely affect performance across a variety of tasks, including digit recognition, serial working memory, and counting. We then explore the effects of applying TTX, a sodium channel blocking drug, to the model. We characterize the behavioral changes that result from this molecular level intervention. We believe this is the first demonstration of a large-scale brain model that clearly links low-level molecular interventions and high-level behavior.

Keywords: Spaun, Neural Engineering Framework, Semantic Pointer Architecture, conductance neurons, biological cognition

1. Introduction

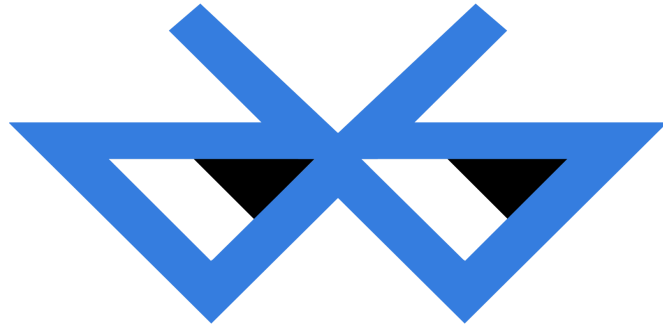
Recently, several large-scale brain models have been described. These include a biophysically detailed model from Markram’s group in the Human Brain Project (HBP) [2], which includes about 31,000 compartmental neurons and 37 million synapses, modelled with many equations per cell. This model is large-scale because of the amount of computation required to simulate its behavior at this level of biological detail. Another model reported earlier by the Synapse project has simulated 500 billion neurons – more than 5x the number in the human brain – although each neuron is much simpler than those in the HBP model, and the connectivity is far more limited [3, 4].



BioXpTM 3200 SYSTEM

ACCELERATING GENOMIC DISCOVERY





BlueBorne

The dangers of Bluetooth implementations: Unveiling zero day vulnerabilities and security flaws in modern Bluetooth stacks.

Ben Seri & Gregory Vishnepolsky



Challenges for Brain Emulation: Why is Building a Brain so Difficult?

Rick Cattell
SynapticLink.org
rick@cattell.net

Alice Parker
Department of Electrical Engineering
University of Southern California
parker@eve.usc.edu

February 5, 2012

Abstract – In recent years, half a dozen major research groups have simulated or constructed sizeable networks of artificial neurons, with the ultimate goal to emulate the entire human brain. At this point, these projects are a long way from that goal: they typically simulate thousands of mammalian neurons, versus tens of billions in the human cortex, with less dense connectivity as well as less-complex neurons. While the outputs of the simulations demonstrate some features of biological neural networks, it is not clear how exact the artificial neurons and networks need to be to invoke system behavior identical to biological networks and it is not even clear how to prove that artificial neural network behavior is identical in any way to biological behavior. However, enough progress has been made to draw some conclusions and make comparisons between the leading projects. Some approaches are more scalable, some are more practical with current technologies, and some are more accurate in their emulation of biological neurons. In this paper, we examine the pros and cons of each approach and make some predictions about the future of artificial neural networks and the prospects for whole brain emulation.

Keywords – biomimetic, neuromorphic, electronic, artificial brain, neuron, intelligence

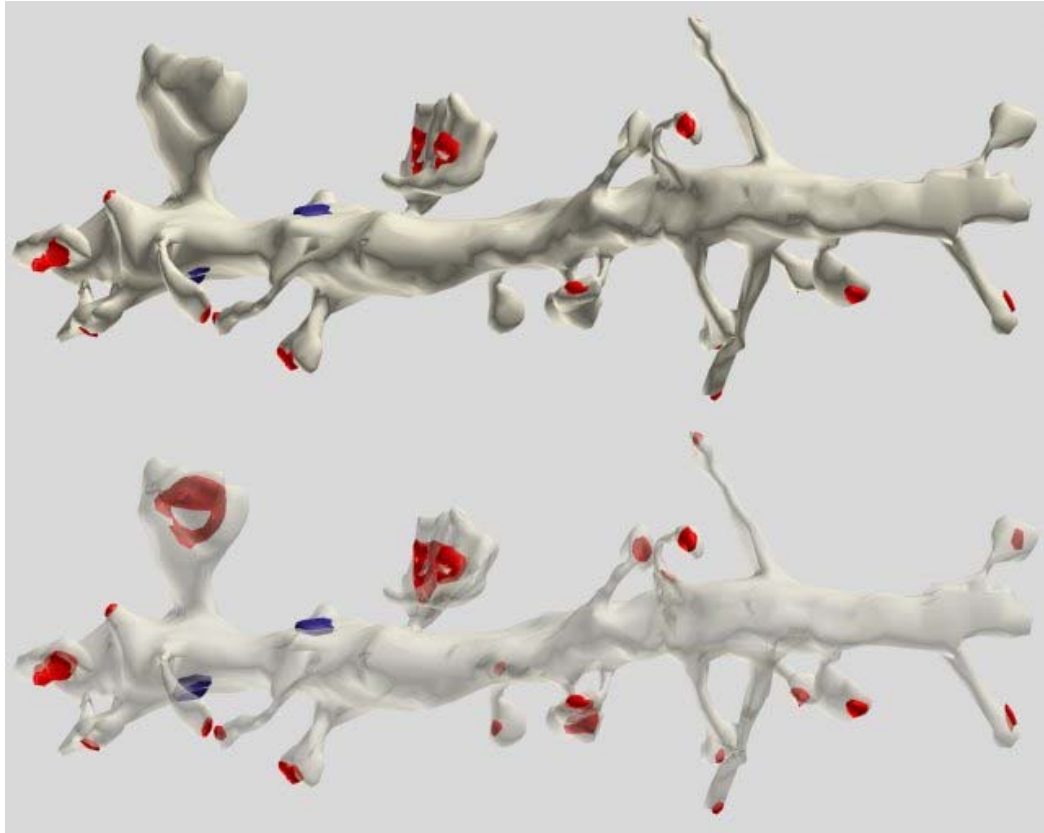
1. Introduction

Reverse-engineering the brain is one of the Grand Challenges posed by the United States National Academy of Engineering [1]. In this paper, we assess current status in approaching this difficult goal of brain emulation. We contrast competing approaches, and examine the major obstacles.

Artificial neurons and neural networks were proposed as far back as 1943, when Warren McColluch and Walter Pitts [2] proposed a “Threshold Logic Unit” with multiple weighted binary inputs combined to produce a binary output based on a threshold value. More sophisticated neural models were subsequently developed, including Rosenblatt’s popular “perceptron” model [3] and others we examine in this article. In 1952, Hodgkin and Huxley [4] published a model of ionic currents that provided the first basis for mathematical modeling and simulation of biological neurons and their action potentials, with the help of Wilfred Rall’s [5] theory of spatiotemporal integration, non-linear summation, and conductance of synaptic signals. These models have likewise been enhanced over the years by researchers examining synaptic transmission, integration, and plasticity.

Whole Brain Emulation

A Roadmap



(2008) *Technical Report #2008-3*

Anders Sandberg*
Nick Bostrom

Future of Humanity Institute
Faculty of Philosophy & James Martin 21st Century School
Oxford University

CITE: Sandberg, A. & Bostrom, N. (2008): *Whole Brain Emulation: A Roadmap*, Technical Report #2008-3, Future of
Humanity Institute, Oxford University
URL: www.fhi.ox.ac.uk/reports/2008-3.pdf

(*) Corresponding author: anders.sandberg@philosophy.ox.ac.uk

THE EPISTEMOLOGY OF COGNITIVE ENHANCEMENT

J. Adam Carter & Duncan Pritchard

University of Edinburgh

ABSTRACT. A common epistemological assumption in contemporary bioethics held by both proponents and critics of non-traditional forms of cognitive enhancement is that cognitive enhancement aims at the facilitation of the accumulation of human knowledge. This paper does three central things. First, drawing from recent work in epistemology, a rival account of cognitive enhancement, framed in terms of the notion of *cognitive achievement* rather than knowledge, is proposed. Second, we outline and respond to an axiological objection to our proposal that draws from recent work by Leon Kass (2004), Michael Sandel (2009), and John Harris (2011) to the effect that ‘enhanced’ cognitive achievements are (by effectively removing obstacles to success) not worthy of pursuit, or are otherwise ‘trivial’. Third, we show how the cognitive achievement account of cognitive enhancement proposed here fits snugly with recent active externalist approaches (e.g., extended cognition) in the philosophy of mind and cognitive science.

Keywords: *cognitive enhancement; bioethics; epistemic value; cognitive achievement; knowledge*

1. COGNITIVE ENHANCEMENT AND KNOWLEDGE

One of the most provocative and pressing topics in recent bioethics concerns human *enhancement*.¹ The latest science and medicine makes it increasingly possible to improve human functioning along several different dimensions: physical, cognitive and (arguably) even moral.² On the assumption that such improvements are available, there is a *pro tanto* instrumental reason to pursue them. After all, to the extent that we are able to function more effectively by being enhanced, we can better achieve our adopted ends. But as critics of human enhancement suggest, the overall balance of reasons may nonetheless speak against human enhancement. That is, all things considered, it may be that we’re better off *not* improving certain dimensions of our functioning, at least in certain ways.

The variety of improvements to human functioning that will be of interest in what follows will be specifically *cognitive* improvements—*viz.*, improvements in aspects of human intelligence, such as memory, reasoning, computation and decision making,³ as opposed to (for example) improvements in the affective (i.e., emotional) or conative (volitional) states or

Future, Opportunities and Challenges of Inkjet Technologies

J.R. Castrejón-Pita¹, W.R.S. Baxter², J. Morgan³, S. Temple⁴, G.D. Martin¹, and I.M. Hutchings¹

¹*Inkjet Research Centre, University of Cambridge,*

17 Charles Babbage Road, Cambridge, CB3 0FS, United Kingdom

²*Inca Digital Printers Ltd, 515 Coldhams Lane, Cambridge, CB1 3JS, United Kingdom*

³*Domino Printing Sciences Plc, Bar Hill, Cambridge, CB23 8TU, United Kingdom and*

⁴*Templetech Ltd, 10 Cambridge Road, Impington, CB24 9NU, United Kingdom*

Inkjet printing relies on the formation of small liquid droplets to deliver precise amounts of material to a substrate under digital control. Inkjet technology is becoming relatively mature and is of great industrial interest thanks to its flexibility for graphical printing and its potential use in less conventional applications such as additive manufacturing and the production of printed electronics and other functional devices. Its advantages over traditional methods of printing include the following: it produces little or no waste, it is versatile because several different methods exist, it is non-contact and does not require a master template so that printed patterns can be readily modified on demand. However, the technology is in need of further development to become mainstream in emerging applications such as additive manufacturing (3D printing). This review contains a description of conventional and less common inkjet methods and surveys the current applications of inkjet in industry. This is followed by specific examples of the barriers, limitations and challenges faced by inkjet in both graphical printing and manufacturing.

Accepted for publication in *Atomization and Sprays* (2013).

Available at: <https://www.repository.cam.ac.uk/handle/1810/238990>

Keywords: inkjet, drop on demand, continuous jet, droplet.

I. INTRODUCTION

Inkjet printing can be considered an umbrella term covering a range of technologies all of which involve the ejection of droplets of ink from a printhead on to a substrate. Since the 1970s when inkjet printing was first commercialised, it has provided a method for delivering and positioning precise small volumes of liquid at high repetition rates under digital control. A key feature of all embodiments of this printing method is that it is non-contact. Whilst inkjet technology has great value for printing purposes, it also has the potential to act as an element of manufacturing processes, whether in the form of accurate dispensing, as can be used for pharmaceutical, bio-chemical or chemical applications, or for building three-dimensional objects from engineering materials (so-called 3D printing, an example of additive manufacturing) so long as they can be supplied in liquid form. The future of inkjet in its traditional niche of printing is still very promising, as the market for graphical printing is very valuable. In principle, inkjet can replace a variety of conventional methods of printing in a large number of commercial applications. Regardless of these extensive potential uses, however, the adoption of inkjet has so far been concentrated in just a few such applications. Several inkjet methods exist that operate under very different principles and utilize diverse inks for various applications. This variety of methods in part reflects the fact that inkjet printing has some fundamental limitations and that these limitations have a different impact on each of the various methods employed to eject droplets. The barriers to wider adoption of inkjet in

non-traditional applications involve both technical and commercial issues. A brief survey of current inkjet processes is presented below and followed by some specific examples of obstacles to the adoption of inkjet.

II. EXAMPLES OF INKJET TECHNOLOGIES

Many different techniques fall under the generic title “inkjet”. The most basic common feature is that some kind of liquid (possibly also containing particles) is transported from a printhead, across a gap, to a receiving substrate. The liquid then, either as it is, or after subsequent modification (e.g. by drying, curing, absorption etc.), performs some function (e.g. forms an image on paper or a conductive structure on an insulating substrate). Another feature, shared by all inkjet systems, is an ability to control the flow of the liquid in a precise way so that a known small volume of material is deposited on to a predetermined location. This control is often achieved by choosing to release a drop, or a stream of drops, from a printhead in which the diameter of the drop is usually related to the size of the nozzle. There are inkjet systems in which the volume dispensed is better described as a slug, a jet or a spray of liquid. In general, two principal attributes make inkjet systems attractive: the non-contact deposition of material and the precise control of the amount and placement of that material. Exactly what is or is not regarded as an inkjet system is sometimes a matter of context. For example a hand-held airbrush would not be considered to be an inkjet system, but a similar computer-controlled airbrush used to create a graphical image on a rotating drum might well be

CAV-2 — why a canine virus is a neurobiologist's best friend

Felix Junyent^{1,2} and Eric J Kremer^{1,2}



Canine adenovirus type 2 (CAV-2) vectors are powerful tools for fundamental and applied neurobiology due to their negligible immunogenicity, preferential transduction of neurons, widespread distribution via axonal transport, and duration of expression in the mammalian brain. CAV-2 vectors are internalized in neurons by the selective use of coxsackievirus and adenovirus receptor (CAR), which is located at the presynapse in neurons. Neuronal internalization and axonal transport is mediated by CAR, which potentiates vector biodistribution. The above characteristics, together with the ~30 kb cloning capacity of helper-dependent (HD) CAV-2 vectors, optimized CAV-2 vector creation, production and purification, is expanding the therapeutic and fundamental options for CNS gene transfer.

Addresses

¹ Institut de Génétique Moléculaire de Montpellier, Montpellier, France

² Université de Montpellier, Montpellier, France

Corresponding author: Kremer, Eric J (eric.kremer@igmm.cnrs.fr)

Current Opinion in Pharmacology 2015, 24:86–93

This review comes from a themed issue on **New Technologies**

Edited by **Andrew H Baker** and **Adrian J Thrasher**

<http://dx.doi.org/10.1016/j.coph.2015.08.004>

1471-4892/© 2015 Elsevier Ltd. All rights reserved.

Introduction

Adenoviruses (AdVs) are 150-mDa nonenveloped pathogens. They have a linear double-stranded DNA genome of 28 — 42 kilobase pairs (kbp) packaged in a 90 nm icosahedral shell. An increasing number of the >200 AdVs isolated from human and nonhuman hosts are partially characterized and are being exploited for their potential as a gene transfer tools. Unfortunately, the characteristics of vectors derived from human type 5 (HAdV-C5) in mice have subjugated the majority of global research and medical communities' perception of AdV vectors for long-term gene transfer (in contrast to the use of AdV vectors for vaccines). We believed that it was unlikely that vectors derived from other AdV types would have a similar set of characteristics, and therefore we pushed forward with the development of nonhuman AdV vectors. One of these 'other' AdVs is canine adenovirus type 2

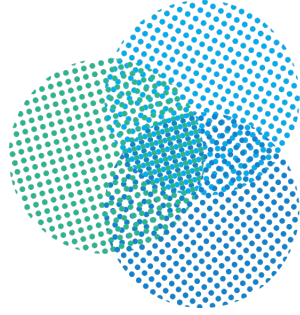
(CAV-2, or commonly referred to as CAV-2). In most hosts, AdVs cause minor infection in the epithelium. In domestic dogs, CAV-2 typically causes an upper respiratory track disease. Back in the early 1990s the impetus to treat cystic fibrosis with gene therapy seemed feasible. Why create a vector from a virus that naturally infects members of the Canidae family [1]? CAV-2 infects the respiratory track, is mass-produced as a vaccine against CAV-1, and at the time was partially sequenced. CAV-2 seemed like an ideal choice to make the first nonhuman AdV vector to avoid the pre-existing immune response and treat cystic fibrosis. Although CAV-2 vectors efficiently transduce lung epithelial cells *in vivo* [2], gene transfer to lung for long-term therapy has encountered many obstacles.

Notably though, vector tropism is not restricted to tissues that manifest disease symptoms following wild type virus infections. During the development of CAV-2 vectors we found that intranasal, intramuscular and intracerebral injections preferentially transduce neurons [3]. In rat olfactory cavity, CAV-2 vectors preferentially transduce the olfactory neurons (in contrast to the neighboring columnar epithelial cells). When CAV-2 vectors are injected in muscles, they poorly transduce myofibres — while efficiently transducing the innervating motor and sensory neurons. To transduce motor and sensory neurons, whose somas are located near the spinal cord, CAV-2 traffics via retrograde axonal transport. In brain parenchyma CAV-2 vectors preferentially transduce neurons at the site of injection as well as the neurons that projected to the injected structure via retrograde axonal transport [3,4].

Because of these characteristics CAV-2 vectors are being used to understand higher-order brain functions and anatomical organization of neural circuits [5[•],6,7,8[•],9[•],10,11[•],12]. In addition to fundamental studies, the characteristics of CAV-2 vectors permit us to treat global brain diseases as well (see below).

Vectors

In the mid-1990s, creating replication-defective, E1-deleted (Δ E1) CAV-2 vectors was not as straightforward as that for HAdV vectors [13,14]. Transfection of overlapping fragments of the CAV-2 vector genomes in canine cells and expecting homologous recombination [13] never worked (even transfection of the intact 32 kb CAV-2 genome poorly produces viruses). Now, Δ E1 CAV-2 vector genomes are created by homologous recombination in



Theory of Deep Learning III: Generalization Properties of SGD

by

Chiyuan Zhang¹ Qianli Liao¹ Alexander Rakhlin² Brando Miranda¹ Noah Golowich¹ Tomaso Poggio¹

¹Center for Brains, Minds, and Machines, McGovern Institute for Brain Research,
Massachusetts Institute of Technology, Cambridge, MA, 02139.

² University of Pennsylvania

Abstract: In Theory III we characterize with a mix of theory and experiments the consistency and generalization properties of deep convolutional networks trained with Stochastic Gradient Descent in classification tasks. A present perceived puzzle is that deep networks show good predictive performance when the classical learning theory seems to suggest overfitting. We describe an explanation of these empirical results in terms of the following new results on SGD:

1. SGD concentrates in probability - like the classical Langevin equation – on large volume, “flat” minima, selecting flat minimizers which are also global minimizers.
2. Minimization under the constraint of maximum volume (usually corresponding to flatness) yields through the jacobian wrt weights and the jacobian wrt x , large (geometrical) margin classification in the case of separable data (zero empirical error for classification loss).
3. Large geometrical margin implies classification bounds via robustness theorems. These bounds can qualitatively explain all the generalization properties empirically observed for deep networks.

Thus SGD selects minimizers corresponding to maximum geometrical margin. Within a single flat minimum the average of the asymptotic fluctuations for each of the degenerate directions (and the non-degenerate ones) is at the maximum margin (the variance however is expected to increase with time in the presence of noise). Because of its connection with robust optimization, SGD can be shown to perform a form of implicit regularization. This explains the puzzling findings about fitting randomly labeled data while performing well on natural labeled data. It also explains while overparametrization does not result in overfitting. Quantitative, non-vacuous bounds are still missing as it has almost always been the case for most practical applications of machine learning. We describe in the appendix an alternative approach that explains with tools of linear algebra the same qualitative properties and puzzles of generalization in deep polynomial networks.

Chemical screening identifies ATM as a target for alleviating senescence

Hyun Tae Kang^{1,5}, Joon Tae Park^{1,5*}, Kobong Choi^{1,5}, Yongsu Kim², Hyo Jei Claudia Choi¹, Chul Won Jung¹, Young-Sam Lee^{3,4*} & Sang Chul Park^{3,4*}

Senescence, defined as irreversible cell-cycle arrest, is the main driving force of aging and age-related diseases. Here, we performed high-throughput screening to identify compounds that alleviate senescence and identified the ataxia telangiectasia mutated (ATM) inhibitor KU-60019 as an effective agent. To elucidate the mechanism underlying ATM's role in senescence, we performed a yeast two-hybrid screen and found that ATM interacted with the vacuolar ATPase V₁ subunits ATP6V1E1 and ATP6V1G1. Specifically, ATM decreased E-G dimerization through direct phosphorylation of ATP6V1G1. Attenuation of ATM activity restored the dimerization, thus consequently facilitating assembly of the V₁ and V₀ domains with concomitant reacidification of the lysosome. In turn, this reacidification induced the functional recovery of the lysosome/autophagy system and was coupled with mitochondrial functional recovery and metabolic reprogramming. Together, our data reveal a new mechanism through which senescence is controlled by the lysosomal-mitochondrial axis, whose function is modulated by the fine-tuning of ATM activity.

Cellular senescence has been suggested to be a primary cause of aging and to account for most aging-associated chronic disorders. Hallmarks of senescence encompass changes including permanent cell-cycle arrest, telomere erosion, senescence-associated β -galactosidase (SA- β -gal) activity, and a senescence-associated secretory phenotype¹.

Lysosomes are essential for cellular homeostasis and have many crucial functions including intracellular digestion of macromolecules and organelles. The digestion of endocytosed or autophagocytosed substrates occurs in an acidic compartment whose acidity is maintained by a vacuolar ATPase (V-ATPase) proton pump. V-ATPase is a multisubunit rotary motor composed of two domains: a peripheral V₁ domain responsible for ATP hydrolysis and a membrane-bound V₀ domain responsible for proton translocation². The regulatory mechanism of V-ATPase, achieved through assembly and disassembly of the V₁ and V₀ domains, plays a crucial role in maintaining lysosomal pH³.

Mitochondria are highly dynamic organelles that undergo continuous fission and fusion cycles. However, dysfunctional mitochondria regularly do not fuse back into the mitochondrial network and instead are eliminated by autophagy, a process through which damaged organelles are degraded⁴. This mitophagy process is crucial for maintaining cellular homeostasis, because dysfunctional mitochondria are incompetent at ATP production but produce more reactive oxygen species (ROS) than do functional mitochondria⁵. Lysosomes and mitochondria are functionally interconnected during senescence⁶. Dysfunctional lysosomes hinder autophagy, thus resulting in accumulation of dysfunctional mitochondria with excessive ROS production⁵. As the major producer of ROS, mitochondria are also the primary target of ROS-induced oxidative damage⁷. This damage in turn further elevates ROS generation, thereby establishing a vicious proaging cycle⁶.

In this study, we sought to identify compounds that alleviate senescence, by using high-throughput screening (HTS). This analysis identified the ATM inhibitor KU-60019 as an effective agent.

Because the role of ATM in senescence remains elusive, we performed a yeast two-hybrid screen and identified a new interaction between ATM and the V-ATPase V₁ subunit. Here, we demonstrate that ATM modulates V₁-V₀ assembly in the V-ATPase and regulates senescence via the lysosomal-mitochondrial axis.

RESULTS

Chemical screening for agents that alleviate senescence

The present screening strategy comprised two different methods of measuring the capacity for alleviating senescence: (i) increasing cell numbers and (ii) decreasing SA- β -gal activity. For the primary screen, a DNA-content-based method was used to accurately measure the cell number⁸ (Supplementary Results, Supplementary Fig. 1a and Supplementary Tables 1 and 2). A library containing 355 kinase inhibitors was added to senescent fibroblasts, and the effects of these inhibitors on cell proliferation were determined on day 21 (Supplementary Table 3). Inhibitors leading to more than a 1.3-fold increase relative to DMSO controls were considered potential hits, and 11 compounds were identified as candidate drugs (Supplementary Fig. 1b). A second round of screening was performed to quantitatively measure SA- β -gal activity levels with a chemiluminescence method⁹. Among the 11 drugs, the ATM inhibitor KU-60019 was found to decrease SA- β -gal activity to 30% that of the DMSO control (Fig. 1a) and was selected as a candidate agent that may ameliorate the senescence phenotype.

ATM as a potential target to ameliorate senescence

To confirm the proliferation-inducing effect observed in screening, cumulative population doubling (CPD) was measured. Treatment with KU-60019, compared with the DMSO control, increased CPD but did not induce malignant growth (Fig. 1b and Supplementary Fig. 2a). To confirm the decrease in SA- β -gal activity, SA- β -gal-positive cells were counted. After KU-60019 treatment, the percentage of SA- β -gal positive cells was significantly decreased on day 8, thus suggesting that KU-60019 effectively alleviates replicative

¹Well-Aging Research Center, Samsung Advanced Institute of Technology, Samsung Electronics, Suwon, Korea. ²Department of Biomedical Sciences, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea. ³Well Aging Research Center, DGIST, Daegu, Korea. ⁴Department of New Biology, DGIST, Daegu, Korea. ⁵These authors contributed equally to this work. *e-mail: joontae.park@samsung.com, lee.yongsam@dgist.ac.kr or blueocean2016@dgist.ac.kr

INFORMATION REPORT INFORMATION REPORT

CENTRAL INTELLIGENCE AGENCY

50X1-HUM
50X1-HUM

This material contains information affecting the National Defense of the United States within the meaning of the espionage laws, Title 18, U.S.C. Secs. 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

CONFIDENTIAL
NO FOREIGN DISSEM

COUNTRY Germany (DDR)

REPORT

SUBJECT Technical Examination of German Transistors

DATE DISTR. 6 January 1964

50X1-HUM

NO. PAGES 43

REFERENCES

DATE OF INFO.
PLACE &
DATE ACQ.

50X1-HUM

THIS IS UNEVALUATED INFORMATION

For ease of discussion, we have numbered the samples individually from 7 through 16.

- a. MCN 19328 (7 and 8) - Two low-frequency, germanium, power transistors, produced during August 1960 at the VEB RFT Funkwerk Kolleda, Germany (DDR) for use in driver stages of unidentified equipment. Exceptionally low noise characteristics were claimed.
- b. MCN 19329 (9 and 10) - Two low-frequency, germanium, power transistors, produced as for MCN 19328, but for audio amplification.
- c. MCN 19335 (11 and 12) - Two Radio-frequency, germanium transistors, produced during September 1960 at the VEB RFT Funkwerk Kolleda; it is claimed that these devices are "velocity modulated".
- d. MCN 19336 (13 and 14) - Two Radio-frequency, germanium transistors, produced as for MCN 19335, but for I.F. applications.
- e. MCN 19337 (15 and 16) - Two Radio-frequency, germanium transistors, produced as for MCN 19335.

2. Examination and tests of these samples have revealed the following:

- a. MCN 19328 - These are low-frequency germanium transistors with 67 mw power capability (determined with 45°C ambient temperature in accordance with European practice). Structurally, these devices are identical with the OC 304 manufactured by Intermetall, Freiburg Brsg. West Germany; electrically they meet Intermetall specifications for the OC 304/2.
- b. MCN 19329 - These are transistors in cases like those used for the OC 318 by Intermetall (with which an aluminum cooling fin is usually supplied). The construction is typically Intermetall and the devices conform to Intermetall characteristics for the OC 318 when measured without the usual cooling fin.

CONFIDENTIAL
NO FOREIGN DISSEM

GROUP 1
Excluded from automatic
downgrading and
declassification

50X1-HUM

STATE ARMY NAVY AIR FORCE PHOENIX AEC

INFORMATION REPORT INFORMATION REPORT

DISSEM: The dissemination of this document is limited to civilian employees and active duty military personnel within the intelligence components of the USIB member agencies, and to those senior officials of the member agencies who must act upon the information. However, unless specifically controlled in accordance with paragraph 8 of DCID 1/7, it may be released to those components of the departments and agencies of the U. S. Government directly participating in the production of National Intelligence. IT SHALL NOT BE DISSEMINATED TO CONTRACTORS. It shall not be disseminated to organizations or personnel, including consultants, under a contractual relationship to the U.S. Government without the written permission of the originator.

50X1-HUM

L 46310-66 EWT(m)

ACC NR:

AP6019631

(A, N)

SOURCE CODE: UR/0048/66/030/002/0343/0348

AUTHOR: Mikhaleva, T.N.; Zazulin, V.S.; Chuprunov, D.L.; Titov, V.I.

8
B

ORG: Scientific Research Institute of Nuclear Physics, Moscow State University im. M.V. Lomonosov (Nauchno-issledovatel'skiy institut yadernoy fiziki Moskovskogo gosudarstvennogo universiteta)

19
TITLE: A scintillation spectrometer with charged particle discrimination /Report, Fifteenth Annual Conference on Nuclear Spectroscopy and Nuclear Structure, held at Minsk, 25 January to 2 February 1965/

SOURCE: AN SSSR. Izvestiya. Seriya fizicheskaya, v. 30, no. 2, 1966, 343-348

TOPIC TAGS: scintillation spectrometer, gamma spectrometer, proton, ~~spectrometer,~~
alpha particle, spectrometer, gamma ray, ~~gamma background, proton, alpha particle,~~

ABSTRACT: There is described a scintillation spectrometer employing a single CsI(Tl) crystal and a single photomultiplier with which pulses due to α rays, protons, and α particles can be distinguished, identified, and recorded in different channels of a multichannel pulse analyzer, depending on the energies of the particles producing them. The technique for identifying the particles is based on the fact that the current pulse on a dynode of the photomultiplier has the form of a decreasing exponential, of which the time constant depends on the nature of the particle producing

Card 1/2

REPORT OF, A. S.,
S. S. D. A, B. C. D. E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z (1951)



Cognitive consilience: primate non-primary neuroanatomical circuits underlying cognition

Soren Van Hout Solari^{1,2*} and Rich Stoner^{3*}

¹ Simigence Inc., Solana Beach, Cardiff, CA, USA

² Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA, USA

³ Department of Neurosciences, Autism Center of Excellence, University of California San Diego, La Jolla, CA, USA

Edited by:

Julian Budd, University of Sussex, UK

Reviewed by:

Trygve B. Leergaard, University of Oslo, Norway

Giorgio Innocenti, Karolinska Institutet, Sweden

Hugo Merchant, Universidad Nacional Autónoma de México, Mexico

*Correspondence:

Soren Van Hout Solari, Simigence Inc., 201 Lomas Santa Fe Suite 490, Solana Beach, CA 92075, USA.
e-mail: sorenсолari@gmail.com;
Rich Stoner, Department of Neurosciences, Autism Center of Excellence, University of California San Diego, 8110 La Jolla Shores Road, Suite 200, La Jolla, CA 92037, USA.
e-mail: rstoner@ucsd.edu

Interactions between the cerebral cortex, thalamus, and basal ganglia form the basis of cognitive information processing in the mammalian brain. Understanding the principles of neuroanatomical organization in these structures is critical to understanding the functions they perform and ultimately how the human brain works. We have manually distilled and synthesized hundreds of primate neuroanatomy facts into a single interactive visualization. The resulting picture represents the fundamental neuroanatomical blueprint upon which cognitive functions must be implemented. Within this framework we hypothesize and detail 7 functional circuits corresponding to psychological perspectives on the brain: consolidated long-term declarative memory, short-term declarative memory, working memory/information processing, behavioral memory selection, behavioral memory output, cognitive control, and cortical information flow regulation. Each circuit is described in terms of distinguishable neuronal groups including the cerebral isocortex (9 pyramidal neuronal groups), parahippocampal gyrus and hippocampus, thalamus (4 neuronal groups), basal ganglia (7 neuronal groups), metencephalon, basal forebrain, and other subcortical nuclei. We focus on neuroanatomy related to primate non-primary cortical systems to elucidate the basis underlying the distinct homotypical cognitive architecture. To display the breadth of this review, we introduce a novel method of integrating and presenting data in multiple independent visualizations: an interactive website (<http://www.frontiersin.org/files/cognitiveconsilience/index.html>) and standalone iPhone and iPad applications. With these tools we present a unique, annotated view of neuroanatomical consilience (integration of knowledge).

Keywords: cerebral cortex, thalamus, basal ganglia, circuitry, consilience, isocortex, cognition

1. INTRODUCTION

At the turn of the twentieth century Cajal (1899, 2002) published what is considered now as the beginning of the modern anatomical understanding of the brain. Cajal's work, entirely dependent on the Golgi staining method, analyzed the neuroanatomical circuitry of complete brains in multiple species. His work stands out from 100 years of subsequent research as a single comprehensive examination across species and brain regions. Brodmann (1909) and von Economo (1929) respectively produced what are, surprisingly still today, the most comprehensive cytoarchitectonic maps of the human cerebral cortex. By the early 1970s, axonal tracing methods were introduced to study distant neuroanatomical projections (Graham and Karnovsky, 1966; Kristensson and Olsson, 1971). Tracing studies have continued to improve and produce detailed projection and connectivity data, but in so doing, fragment knowledge across species and brain regions (Zaborszky et al., 2006).

Forming an accurate mental view of brain circuitry is difficult, yet without one we cannot understand the function of the brain. Only with a comprehensive and cohesive picture can we make accurate inferences about the function of discrete neuroanatomical circuits. Each structure imposes dependencies and

constraints on any theory that must be maintained for a working hypothesis of brain function. Several efforts are currently underway to reconcile the disparity between individual connectivity studies within a global scope. CoCoMac, a tool based on primate literature, represents the state of the art in mapping corticocortical interconnectivity between functional regions (Kotter, 2004). The Human Connectome Project (Marcus et al., 2011), along with other projects within the NIH Blueprint for Neuroscience Research, are using novel imaging methods to describe connectivity details for both primate and human brains (Stephan et al., 2000; Schmahmann et al., 2007; Hagmann et al., 2010). Unfortunately the resolution of external imaging methods is insufficient to elucidate neuroanatomical details underlying circuit organization.

This review is an attempt to form a comprehensive and cohesive understanding of the primate non-primary neuroanatomical circuitry through consilience (the integration of knowledge). Our first goal is to assemble a comprehensive neuroanatomical picture that is not *inconsistent* with known facts. We have produced an interactive visualization by synthesizing a vast number of fragmented studies into a single referenced framework that can be explored dynamically (**Figure 1**). We present this neuroanatomical picture as a detailed first-order approximation of

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/274901092>

Connectomics-Based Analysis of Information Flow in the Drosophila Brain

ARTICLE in CURRENT BIOLOGY: CB · APRIL 2015

Impact Factor: 9.57 · DOI: 10.1016/j.cub.2015.03.021 · Source: PubMed

CITATIONS

10

READS

208

10 AUTHORS, INCLUDING:



Chi-Tin Shih

Tunghai University

48 PUBLICATIONS 762 CITATIONS

SEE PROFILE



Olaf Sporns

Indiana University Bloomington

290 PUBLICATIONS 24,813 CITATIONS

SEE PROFILE



Chung-Chuan Lo

National Tsing Hua University

25 PUBLICATIONS 582 CITATIONS

SEE PROFILE



Ann-Shyn Chiang

National Tsing Hua University

153 PUBLICATIONS 3,646 CITATIONS

SEE PROFILE

Cite as: S. D. Perli *et al.*, *Science*
10.1126/science.aag0511 (2016).

Continuous genetic recording with self-targeting CRISPR-Cas in human cells

Samuel D. Perli,^{1,2,3*} Cheryl H. Cui,^{1,2,4*} Timothy K. Lu^{1,2,3,5†}

¹Synthetic Biology Group, MIT Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ²Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ³Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ⁴Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA. ⁵Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.

†Corresponding author. Email: timlu@mit.edu

The ability to record molecular events in vivo would enable monitoring of signaling dynamics within cellular niches and critical factors that orchestrate cellular behavior. We present a self-contained analog memory device for longitudinal recording of molecular stimuli into DNA mutations in human cells. This device consists of a self-targeting guide RNA (stgRNA) that repeatedly directs *Streptococcus pyogenes* Cas9 nuclease activity toward the DNA that encodes the stgRNA, thereby enabling localized, continuous DNA mutagenesis as a function of stgRNA expression. We demonstrate programmable and multiplexed memory storage in human cells triggered by exogenous inducers or inflammation, both in vitro and in vivo. This tool, Mammalian Synthetic Cellular Recorder Integrating Biological Events (mSCRIBE), provides a unique strategy for investigating cell biology in vivo and enables continuous evolution of targeted DNA sequences.

Cellular behavior is dynamic, responsive and regulated by the integration of multiple molecular signals. Biological memory devices that can record regulatory events would be useful tools for investigating cellular behavior over the course of a biological process and further our understanding of signaling dynamics within cellular niches. Earlier generations of biological memory devices relied on digital switching between two or multiple quasi-stable states based on active transcription and translation of proteins (1–3). However, such systems do not maintain their memory after the cells are disruptively harvested. Encoding transient cellular events into genomic DNA memory using DNA recombinases enables the storage of heritable biological information even after gene regulation is disrupted (4, 5). The capacity and scalability of these memory devices are limited by the number of orthogonal regulatory elements (e.g., transcription factors and recombinases) that can reliably function together. Furthermore, because they are restricted to a small number of digital states, they cannot record dynamic (analog) biological information, such as the magnitude or duration of a cellular event. We recently demonstrated a population-based technology for genomically encoded analog memory in *Escherichia coli* based on dynamic genome editing with retrons (6). Here, we present Mammalian Synthetic Cellular Recorders Integrating Biological Events (mSCRIBE), an analog memory system that enables the re-

cording of cellular events within human cell populations in the form of DNA mutations. mSCRIBE uses self-targeting guide RNAs (stgRNAs) that direct CRISPR-Cas activity to repeatedly mutagenize the DNA loci that encodes the stgRNAs. During the course of review of this work, systems with similar principles have been proposed (7, 8). Although these systems use Cas9 to record information in DNA, they pursue different applications such as lineage tracing and generating barcodes to uniquely tag multiple cells simultaneously. In contrast, we utilize our platform to build memory devices capable of recording analog biological activity into mammalian cells both in vitro and in vivo.

The *S. pyogenes* Cas9 system from the Clustered Regularly-Interspaced Short Palindromic Repeats-associated (CRISPR-Cas) family is an effective genome engineering enzyme that catalyzes double-stranded breaks and generates mutations at DNA loci targeted by a small guide RNA (sgRNA) (9–11). Normal sgRNAs are comprised of a 20 nucleotide (nt) Specificity Determining Sequence (SDS), which specifies the DNA sequence to be targeted and is immediately followed by a 80 nt scaffold sequence, which associates the sgRNA with Cas9. In addition to sequence homology with the SDS, targeted DNA sequences must possess a Protospacer Adjacent Motif (PAM) (5'-NGG-3') immediately adjacent to their 3'-end in order to be bound by the Cas9-sgRNA complex and cleaved (12). When a double-strand

CRISPR-Cas9 Reference Model
for MPEG LA's CRISPR-Cas9 Joint Licensing
Platform

Version 1.0

April 24, 2017



NEWS RELEASE

For Immediate Release

CONTACT:

Tom O'Reilly
MPEG LA, LLC
Tel: 303.200.1710
Fax: 301.986.8575
toreilly@mpegla.com

The Broad Institute of MIT and Harvard Among Those Participating in MPEG LA's CRISPR-Cas9 Joint Licensing Facilitation

Patent submissions continue to be accepted

(DENVER, CO, US – 10 July 2017) – World licensing leader MPEG LA, LLC is pleased to announce that key CRISPR patent holders have submitted patents in response to MPEG LA's [call to participate](#) in the creation of a global CRISPR-Cas9 Joint Licensing Platform and that patent submissions continue to be accepted.

“The enthusiastic response to MPEG LA's patent pool initiative is further evidence of the widely held view that CRISPR is too important to be left at risk of endless patent battles and splintered licensing regimes and that a pool providing one-stop licensing efficiency and predictability to scientists and businesses worldwide represents the best hope to unleash its life-enhancing potential,” said Larry Horn, President and CEO of MPEG LA. “As MPEG LA's objective is to provide worldwide access to as much CRISPR intellectual property as possible and the market will benefit from the viewpoints of all stakeholders in this voluntary ground floor opportunity to determine CRISPR's future, the door remains open and additional patent submissions continue to be welcomed.”

Among those who have agreed to participate, MPEG LA welcomes the Broad Institute of MIT and Harvard, which has submitted key CRISPR-Cas9 patents for consideration with joint owners Harvard University, the Massachusetts Institute of Technology, and The Rockefeller University.

“We strongly support making CRISPR technology broadly available,” said Issi Rozen, Chief Business Officer of the Broad Institute. “The Broad Institute already licenses CRISPR-Cas9 non-exclusively for all applications, with the exception of human therapeutics, where we have significantly limited the exclusivity. We look forward to working with others to ensure the widest possible access to all key CRISPR intellectual property.”

Induction of self awareness in dreams through frontal low current stimulation of gamma activity

Ursula Voss^{1,2}, Romain Holzmann³, Allan Hobson⁴,
Walter Paulus⁵, Judith Koppehele-Gossel⁶, Ansgar Klimke^{2,7} &
Michael A Nitsche⁵

Recent findings link fronto-temporal gamma electroencephalographic (EEG) activity to conscious awareness in dreams, but a causal relationship has not yet been established. We found that current stimulation in the lower gamma band during REM sleep influences ongoing brain activity and induces self-reflective awareness in dreams. Other stimulation frequencies were not effective, suggesting that higher order consciousness is indeed related to synchronous oscillations around 25 and 40 Hz.

Rapid eye movement (REM) sleep dreams are primary states of consciousness in that they are concerned with the immediate present, with only uncontrolled access to the past or the anticipated future^{1–3}. After awakening, humans—and supposedly¹ only humans—enter a secondary mode of consciousness that introduces higher order cognitive functions such as self-reflective awareness, abstract thinking, volition and metacognition^{1–4}. A state of sleep in which primary and secondary states of consciousness coexist is lucid dreaming, a phenomenon that is most likely unique to humans. In lucid dreams, elements of secondary consciousness coexist with normal REM sleep consciousness, enabling the sleeper to become aware of the fact that he is dreaming while the dream continues. Sometimes the dreamer gains control over the ongoing dream plot and, for example, is able to put a dream aggressor to flight. Scientifically, lucid dreams present the unique opportunity to watch the brain change conscious states, from primary to secondary consciousness^{2,4}, and to arrive at testable predictions about the determinants of these states. At the neurophysiological level, EEG³ and functional magnetic resonance imaging (fMRI) studies⁵ have shown that lucid dreams are accompanied by increased phase synchrony and elevated frequency-specific activity in the lower gamma frequency band centered around 40 Hz, especially in frontal^{3,5} and temporal¹⁵ parts of the brain. Fronto-temporal activity in this frequency band is related to executive ego functions and secondary consciousness, which is characteristic of the human wake state and atypical for REM sleep^{2,3,6}. The increase in gamma activity observed during lucid dreaming raises several theoretical questions.

Does lucid dreaming trigger gamma-band activity or does gamma-band activity trigger lucid dreaming? Perhaps the capacity to generate gamma oscillatory activity sets the stage for lucid dreaming, which may then further enhance gamma activity. Furthermore, is lucid dreaming dependent on the presence of gamma activity (necessary condition) or can higher order consciousness in dreams be elicited via other causal routes, such as through stimulation with other frequencies (causally enabling condition)? We tested these hypotheses via fronto-temporal transcranial alternating current stimulation (tACS) at various frequencies (2, 6, 12, 25, 40, 70 and 100 Hz) and under sham conditions (simulated stimulation, but no current flow; **Supplementary Fig. 1**). This relatively new method of brain stimulation has no such side effects as acoustic noise and tactile sensations, which are known to accompany transcranial magnetic stimulation and might result in sleep disturbance. tACS has already been shown to modify perceptual and cognitive performance in waking⁷ and, in combination with superimposed transcranial direct current stimulation (tDCS), in sleep⁸.

Brain activity was monitored by continuous EEG, electrooculography (EOG) and electromyography (EMG) (**Supplementary Fig. 2a**). tACS was applied following ~2 min of uninterrupted arousal-free REM sleep, after which subjects were awakened and asked to rate dream consciousness based on a factor analytically derived and validated scale (LuCiD scale⁹). Previous laboratory research with the LuCiD scale has shown that, in lucid dreaming, three of eight factors are substantially increased: insight into the fact that one is currently dreaming, control over the dream plot and dissociation akin to taking on a third-person perspective (**Supplementary Fig. 3**).

The EEG was quantitatively analyzed for all stimulation conditions and sham (**Fig. 1a–c**). Unchanged REMs (**Fig. 1a**) and continuous muscle atonia (**Fig. 1b**) documented the persistence of REM-like sleep preceding (phase I), throughout (phase II) and following stimulation (phase III) until awakening (phase IV) (**Supplementary Fig. 4**).

When representing EEG power as a function of frequency and time, as depicted in the wavelet transform shown in **Figure 1c**, we confirmed the findings from EOG and EMG, showing that the EEG power spectrum during stimulation remained very similar until awakening (phase IV). Wakefulness was characterized by a strong increase in the alpha frequency band, typical for waking with eyes closed.

In the analyzed EEG samples, subjects maintained typical signs of REM sleep during stimulation, as evidenced by EMG, EOG and EEG (**Fig. 1a–c** and **Supplementary Figs. 4 and 5**). Unlike in normal REM sleep, however, activity in the lower gamma frequency band increased during stimulation with 40 Hz (mean increase between 37–43 Hz = 28%, s.e. = 4.82) and, to a lesser degree, during stimulation with 25 Hz (mean increase between 22–28 Hz = 12%, s.e. = 5.82). During sham or

¹Department of Psychology, J.W. Goethe-University Frankfurt, Frankfurt, Germany. ²Department of Clinical Sleep Research, VITOS Hochtaunus Klinik, Friedrichsdorf, Germany. ³GSI Helmholtzzentrum für Schwerionenforschung, Darmstadt, Germany. ⁴Harvard Medical School, Boston, Massachusetts, USA. ⁵Department of Clinical Neurophysiology, University Medical Center, Göttingen, Germany. ⁶Department of Psychology, Bonn University, Bonn, Germany. ⁷Department of Psychiatry, Duesseldorf University, Duesseldorf, Germany. Correspondence should be addressed to U.V. (voss@psych.uni-frankfurt.de).

Received 7 January; accepted 16 April; published online 11 May 2014; doi:10.1038/nn.3719

A brush fire feedback loop for the ramp-like rise in hominin brain size that began 2.4 million years ago

WILLIAM H. CALVIN

Department of Psychiatry and Behavioral Sciences, University of Washington,
725 9th Avenue, Box 2605, Seattle WA 98104-2086, WCalvin@UW.edu, and the
UCSD/Salk Center for Academic Research and Training in Anthropogeny, La Jolla CA
92037

During the first million years of the Pleistocene climate changes, our ancestors' brain size doubled. Enlargement continued at the same rate, suggesting a self-sustaining process with a rate-limiting component. For large grazing animals and their predator *Homo erectus*, I analyzed the brush-fire cycle behind grasslands' brushy frontier, seeking a feedback loop. The burn scar's new grass is an empty niche for grass-specialized herbivores, which evolved from mixed feeders only in the early Pleistocene. The frontier subpopulation of grazers discovering the auxiliary grassland quickly multiplies. Following this boom, a bust occurs several decades later when the brush returns; it squeezes this offshoot population back into the core grasslands population. For both prey and predators, such a feedback loop can shift the core's gene frequencies toward those of the brush explorers.

Hunters of both browsers and grazers spend more time in the brush than those specializing in grazers; this versatility becomes a candidate for differential amplification. Any brush-relevant allele could benefit from amplifying feedback by such trait hitchhiking, so long as its phenotypes also concentrate near where empty niches can open up in the brush for grazers and their predators. Increased versatility likely correlates with larger brain size on the evolutionary time scale. Among the tasks likely to need the shade of brush are toolmaking and food preparation. The more versatile, larger-than-average brains need only spend more-than-average time in the catchment zone for this recursive evolutionary process to keep average brain size increasing even further, making advance room for some future functionality in the cerebral cortex.

Generating Adversarial Examples for Speech Recognition

Dan Iter

Stanford University
daniter@stanford

Jade Huang

Stanford University
jayebird@stanford

Mike Jermann

Stanford University
mjermann@stanford

Abstract

With the proliferation of natural language interfaces on mobile devices and in home personal assistants such as Siri and Alexa, many services and data are becoming available through transcription from a speech recognition system. One major risk factor in this trend is that a malicious adversary may attack this system without the primary user noticing. One way to accomplish this is to use adversarial examples that are perceived one way by a human, but transcribed differently by the Automatic Speech Recognition (ASR) system. For example, a recording that sounds like "hello" to the human ear, but is transcribed as "goodbye" by the ASR system. Recent work has shown that adversarial examples can be created for convolutional neural networks to fool vision recognition systems. We show that similar methods can be applied to neural ASR systems. We show successful results for two methods of generating adversarial examples where we fool a high quality ASR system but the difference in the audio is imperceptible to the human ear. We also present a method for converting the adversarial MFCC features back into audio.

1 Introduction

Recent work has shown that adversarial examples can be created for convolutional neural networks to fool vision recognition systems (Goodfellow et al., 2014), (Szegedy et al., 2014). For example, an adversarial image can be classified by neural networks as "gibbon", but to the human eye, the image clearly appears to be a "panda". Such adversarial examples expose a weakness in neu-

ral networks. Adversarial examples are not just the result of overfitting to a particular model or specific selection of a training set. They often remain difficult for a network to classify even with a model trained with multiple hyperparameters. With this in mind, they can be used to improve the quality of existing neural networks through adversarial training.

We show that similar methods can be applied to neural ASR systems, namely the fast gradient sign method (Goodfellow et al., 2014) and the fooling gradient sign method (Karpathy, 2015). We utilize an existing codebase for a state-of-the-art neural ASR system, WaveNet, implemented in TensorFlow (Kim and Park, 2016) for the task of speech-to-text. We integrate our two methods into the code so that we are able to gain access to the computed gradients and modify them accordingly.

As we utilized a pre-trained model of WaveNet whose input were MFCC Features (mainly from VCTKCorpus), we also operated using MFCC Features. After successfully creating adversarial samples that our neural network misclassifies, we convert the features back into audio with an approximate approximation of an inverse MFCC transform, as discussed in 3.4. While this is a lossy transform, the resulting audio is still coherent to the human ear.

The result is that a neural network classifies a set of MFCC features as "Siri call police", but the features when converted back to audio sound like "Please call Stella" to the human ear.

2 Related Work

2.1 Intriguing properties of neural networks

Szegedy et al. presents a somewhat formal definition of an "adversarial examples" in the context of two intriguing properties of neural networks. This concept is presented as part of two main properties

Genome-Wide Analyses of Working-Memory Ability: A Review

E. E. M. Knowles · S. R. Mathias · D. R. McKay ·
E. Sprooten · John Blangero · Laura Almasy ·
D. C. Glahn

© Springer International Publishing AG 2014

Abstract Working memory, a theoretical construct from the field of cognitive psychology, is crucial to everyday life. It refers to the ability to temporarily store and manipulate task-relevant information. The identification of genes for working memory might shed light on the molecular mechanisms of this important cognitive ability and—given the genetic overlap between, for example, schizophrenia risk and working-memory ability—might also reveal important candidate genes for psychiatric illness. A number of genome-wide searches for genes that influence working memory have been conducted in recent years. Interestingly, the results of those searches converge on the mediating role of neuronal excitability in working-memory performance, such that the role of each gene highlighted by genome-wide methods plays a part in ion channel formation and/or dopaminergic signaling in the brain, with either direct or indirect influence on dopamine levels in the prefrontal cortex. This result dovetails with animal models of working memory that highlight the role of dynamic network connectivity, as mediated by dopaminergic signaling, in the dorsolateral prefrontal cortex. Future work, which aims to characterize functional variants influencing working-memory ability, might choose to focus on those genes highlighted in the

present review and also those networks in which the genes fall. Confirming gene associations and highlighting functional characterization of those associations might have implications for the understanding of normal variation in working-memory ability and also for the development of drugs for mental illness.

Keywords Working memory · Genomics · Cognition · GWA · Dynamic network connectivity

Introduction

Working memory is crucial to everyday life; it plays a key role in everyday tasks—following spoken directions, reading a magazine article, calculating a tip in a restaurant—that require information to be temporarily stored and manipulated [1, 2]. It has been described as being core to reasoning and judgment in humans; in other words, working memory is crucial to other important aspects of cognitive performance, such as attention and executive functioning [3, 4], and is a determinant of an individual's level of intelligence [5, 6]. As a consequence, working memory is one of the most studied concepts in cognitive neuroscience [7]. In addition, working memory is impaired in psychiatric and neurodegenerative illnesses such as schizophrenia and Alzheimer's disease [8, 9]. Moreover, there is thought to be a substantial genetic overlap between those genes that mediate illness risk and those that influence working-memory ability [10, 11]. The importance of working memory to cognition in general, combined with the key role that working memory plays in the symptomatology of certain illnesses, behooves the research community to provide insights into the molecular underpinnings of working-memory ability. The field of behavior genetics is ideally suited to fulfill this task and, to date, several genome-wide searches for genes influencing working memory have been conducted. The present manuscript provides a qualitative review of this literature.

E. E. M. Knowles (✉) · S. R. Mathias · D. R. McKay ·
E. Sprooten · D. C. Glahn
Department of Psychiatry, Yale University School of Medicine,
New Haven, CT, USA
e-mail: emma.knowles@yale.edu

D. C. Glahn
e-mail: david.glahn@yale.edu

E. E. M. Knowles · S. R. Mathias · D. R. McKay · E. Sprooten ·
D. C. Glahn
Olin Neuropsychiatric Research Center, Institute of Living, Hartford
Hospital, Hartford, CT, USA

J. Blangero · L. Almasy
Department of Genetics, Texas Biomedical Research Institute,
San Antonio, TX, USA

Are all BSDs created equally?

A survey of BSD kernel vulnerabilities.

Ilja van Sprundel <ivansprundel@ioactive.com>

Delivery technologies for genome editing

Hao Yin¹, Kevin J. Kauffman^{1,2} and Daniel G. Anderson^{1–4}

Abstract | With the recent development of CRISPR technology, it is becoming increasingly easy to engineer the genome. Genome-editing systems based on CRISPR, as well as transcription activator-like effector nucleases (TALENs) and zinc-finger nucleases (ZFNs), are becoming valuable tools for biomedical research, drug discovery and development, and even gene therapy. However, for each of these systems to effectively enter cells of interest and perform their function, efficient and safe delivery technologies are needed. This Review discusses the principles of biomacromolecule delivery and gene editing, examines recent advances and challenges in non-viral and viral delivery methods, and highlights the status of related clinical trials.

RNA interference

(RNAi). Process by which one strand of double-stranded RNA binds to complementary mRNA and degrades or regulates the mRNA via an enzymatic process, usually resulting in a decrease in the expression of a desired protein.

More than 3,000 human genes have been identified that are associated with Mendelian diseases, and ~500 genes are associated with susceptibility to complex diseases or infections (see [Orphanet](#) and [OMIM Gene Map Statistics](#) websites). These numbers are rapidly increasing, and it is predicted that ~4,000–7,000 additional disease-associated genes will be uncovered in the next decade¹. However, despite tremendous advances in genomic sciences and substantial progress in drug development, effective therapies are still needed; for example, <5% of rare diseases have an effective treatment².

Cystic fibrosis is an example of a genetic disease with effective treatments for a large portion of patients, with the successful development of small-molecule drugs to target defective transmembrane conductance regulator (CFTR) protein³. Small molecules now can treat about 40% of patients with cystic fibrosis and eventually could hopefully benefit >90% of patients. However, small molecules cannot provide a cure for cystic fibrosis, and it is generally difficult to use small-molecule therapies, bone marrow transplantation or surgical approaches to treat most genetic diseases⁴. Protein therapeutics, including protein replacement (or augmentation) therapies and antibodies, have been investigated for treating certain genetic disorders, and some have received regulatory approval. Because administered purified proteins generally do not enter cells, protein augmentation therapies have been applied to treat disorders for which a deficient protein functions at least partially in the extracellular milieu⁵. Examples of regulatory-approved protein therapeutics include recombinant acid α -glucosidase, recombinant factors VIII and IX, and a humanized vascular endothelial growth factor A (VEGFA)-specific antibody. Recombinant acid α -glucosidase effectively treats Pompe disease, a type of lysosomal storage disorder, resulting

in improved function of multiple organs in patients⁶. Recombinant factors VIII and IX are used to treat haemophilia A and haemophilia B, respectively, to correct bleeding episodes⁷, whereas the VEGFA-specific antibody is used to treat age-related macular degeneration⁸. However, most of these protein therapeutics target a limited number of biomolecules that are involved in a small proportion of genetic diseases⁹.

Meanwhile, viral and non-viral delivery strategies of functional copies of mutated genes have been developed to treat loss-of-function genetic diseases^{10,11}, and many are currently in clinical trials (for examples, see [ClinicalTrials.gov](#)). In 2012, the adeno-associated virus (AAV)-mediated delivery of a functional gene to treat the rare disease lipoprotein lipase deficiency became the first gene therapy product based on viral gene-transfer technology to receive marketing approval in Europe.

In addition, RNA modification therapies, such as RNA interference (RNAi) and antisense oligonucleotides (ASOs), which silence the mRNA transcribed from disease genes, are moving forward in clinical trials^{12,13}. An ASO-based product has received US Food and Drug Administration approval for use in patients with homozygous familial hypercholesterolaemia, and other RNA modification therapies are in advanced stages of development; for example, alicaforsen for pouchitis (see [Ionis Pharmaceuticals' pipeline](#)). RNAi is being evaluated in one phase III clinical trial targeting transthyretin-mediated amyloidosis, as well as in multiple phase I/II trials¹⁰.

However, despite the remarkable progress and therapeutic promise of ASO and RNAi technologies, as well as AAV-facilitated gene transfer, these approaches have fundamental limitations. In principle, ASOs and RNAi can induce the knockdown of any target protein; however, delivery barriers have limited their clinical application to

¹David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

³Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology, Cambridge, Massachusetts 02139, USA.

⁴Institute of Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Correspondence to D.G.A. dgander@mit.edu

doi:10.1038/nrd.2016.280
Published online 24 Mar 2017

Discreet Log Contracts

Thaddeus Dryja

MIT Digital Currency Initiative

Abstract

Smart contracts [1] are an often touted feature of cryptographic currency systems such as Bitcoin, but they have yet to see widespread financial use. Two of the biggest hurdles to their implementation and adoption have been scalability of the smart contracts, and the difficulty in getting data external to the currency system into the smart contract. Privacy of the contract has been another issue to date. Discreet Log Contracts are a system which addresses the scalability and privacy concerns and seeks to minimize the trust required in the oracle which provides external data. The contracts are discreet in that external observers cannot detect the presence of the contract in the transaction log. They also hinge on knowledge of a *discrete* logarithm, which is a plus.

Model

There are 3 parties involved in the contract process: Alice, Bob, and Olivia. Alice and Bob are contract counterparties, while Olivia is the oracle. Alice and Bob do not trust each other and do not need to know any legal identifying information about each other, but they must be able to communicate over an authenticated channel, and they must be able to persistently recognize each other. Alice and Bob also must be able to receive signed broadcast messages from Olivia. Olivia does not need to be aware of Alice and Bob, and ideally she has no contact other than broadcasting information. The information is compact enough that broadcast could take place over the Bitcoin network itself, though this should not be necessary.

The DLC protocol can be used for a wide variety of contracts, covering most cases where payouts between parties depend on a publicly known number in the future. In this example, Alice and Bob make and execute

RESEARCH ARTICLE

DNA Assembly in 3D Printed Fluidics

William G. Patrick¹, Alec A. K. Nielsen², Steven J. Keating^{1,3}, Taylor J. Levy¹, Che-Wei Wang¹, Jaime J. Rivera², Octavio Mondragón-Palomino², Peter A. Carr⁴, Christopher A. Voigt², Neri Oxman¹, David S. Kong^{4*}

1 MIT Media Lab, School of Architecture and Planning, Massachusetts Institute of Technology, Cambridge, MA, United States of America, **2** Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States of America, **3** Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **4** Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA, United States of America

* dkong@ll.mit.edu



OPEN ACCESS

Citation: Patrick WG, Nielsen AAK, Keating SJ, Levy TJ, Wang C-W, Rivera JJ, et al. (2015) DNA Assembly in 3D Printed Fluidics. PLoS ONE 10(12): e0143636. doi:10.1371/journal.pone.0143636

Editor: Meni Wanunu, Northeastern University, UNITED STATES

Received: July 30, 2015

Accepted: November 6, 2015

Published: December 30, 2015

Copyright: © 2015 Patrick et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was sponsored by the Assistant Secretary of Defense for Research & Engineering under Air Force Contract #FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The process of connecting genetic parts—DNA assembly—is a foundational technology for synthetic biology. Microfluidics present an attractive solution for minimizing use of costly reagents, enabling multiplexed reactions, and automating protocols by integrating multiple protocol steps. However, microfluidics fabrication and operation can be expensive and requires expertise, limiting access to the technology. With advances in commodity digital fabrication tools, it is now possible to directly print fluidic devices and supporting hardware. 3D printed micro- and millifluidic devices are inexpensive, easy to make and quick to produce. We demonstrate Golden Gate DNA assembly in 3D-printed fluidics with reaction volumes as small as 490 nL, channel widths as fine as 220 microns, and per unit part costs ranging from \$0.61 to \$5.71. A 3D-printed syringe pump with an accompanying programmable software interface was designed and fabricated to operate the devices. Quick turnaround and inexpensive materials allowed for rapid exploration of device parameters, demonstrating a manufacturing paradigm for designing and fabricating hardware for synthetic biology.

Introduction

Synthetic Biology is a rapidly advancing field that is being used to create novel biotechnology applications, next-generation therapeutics, and new methods of scientific inquiry [1–3]. The commercialization and rapid decline in price of DNA sequencing and synthesis technologies have enabled much of this development. Solid-phase DNA synthesis has declined in price, enabling researchers to routinely design and order synthetic DNA up to several kilobases in length. However, assembly of these molecules into larger constructs remains an essential technique. Indeed, DNA assembly is necessary to generate complex single-gene and multi-gene constructs [4–6], to create functionally-diverse part combinations (e.g., gene clusters with libraries of RBSs and promoters [7,8]), to “shuffle” homologous proteins at specific recombination points (e.g., shuffling of three trypsinogen homologues [9]), and to explore higher-order effects of genetic architectures (e.g., the position and orientation of transcription units [7]).

DNA rendering of polyhedral meshes at the nanoscale

Erik Benson^{1,2}, Abdulmelik Mohammed³, Johan Gardell^{1,2}, Sergej Masich⁴, Eugen Czeizler³, Pekka Orponen³ & Björn Högberg^{1,2}

It was suggested¹ more than thirty years ago that Watson–Crick base pairing might be used for the rational design of nanometre-scale structures from nucleic acids. Since then, and especially since the introduction of the origami technique², DNA nanotechnology has enabled increasingly more complex structures^{3–18}. But although general approaches for creating DNA origami polygonal meshes and design software are available^{14,16,17,19–21}, there are still important constraints arising from DNA geometry and sense/antisense pairing, necessitating some manual adjustment during the design process. Here we present a general method of folding arbitrary polygonal digital meshes in DNA that readily produces structures that would be very difficult to realize using previous approaches. The design process is highly automated, using a routing algorithm based on graph theory and a relaxation simulation that traces scaffold strands through the target structures. Moreover, unlike conventional origami designs built from close-packed helices, our structures have a more open conformation with one helix per edge and are therefore stable under the ionic conditions usually used in biological assays.

The starting point of the method we present here is a 3D mesh representing the geometry one wishes to realize at the nanoscale.

Focusing only on polyhedral meshes, that is, meshes which enclose a volume inflatable to a ball, and in contrast to several previous approaches^{14,17,19} (see Extended Data Fig. 1), we aim to replace the edges of the mesh by single DNA double helices such that the scaffold strand traverses each of these edges once. This problem is closely related to the ‘Chinese postman tour’ problem²² in graph theory, and finding solutions by hand would be impossible in practice for most meshes. The main three principles underpinning our design paradigm are: first, that the technique should allow meshes to be triangulated to optimize structural rigidity; second, that each edge should be represented by one double helix to enable construction of large structures using as little DNA as possible (though some meshes require two helices to render certain edges, as discussed below); and, third, that vertices should be non-crossing (that is, the scaffold should not cross itself at the vertices, which ensures non-knotted paths with fewer topological and kinetic traps during folding, and vertex junctions should be planar, which avoids mesh protrusions caused by the stacking of crossing helices at each vertex).

The overall design scheme is split into four discrete steps, as follows. (1) Drawing of a 3D polygon mesh using 3D software; see Fig. 1a. (2) Generating an appropriate routing of the long scaffold strand through

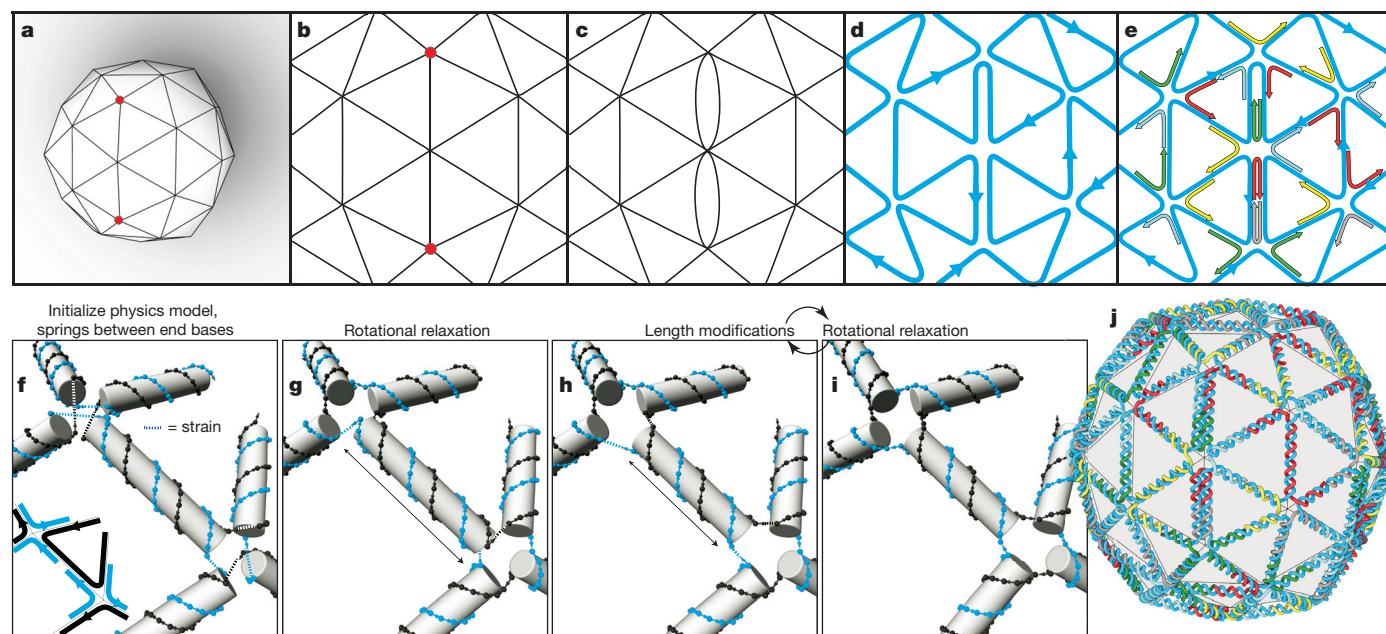


Figure 1 | Design paradigm and automated workflow for scaffold-routing sequence design of origami 3D meshes. **a**, A 3D mesh is drawn using 3D software. **b**, Using the minimum weight perfect matching algorithm, odd-degree vertices are paired. **c**, Double edges are introduced. **d**, The developed A-trails algorithm routes the scaffold according to the constraints. **e**, The staple-strand (multi-coloured) routing follows implicitly from the scaffold (blue)

routing. **f–i**, Before computation of the sequences, a physics model is used to relax and evenly distribute strain in the design. Each double helix is treated as a stiff rod with springs connecting the bases at the ends of the scaffold strand and staple strand. Iterations of rotational relaxation (**g** and **i**) and length modification of helices (**h**) leads to the final design (**j**), where sequences are calculated after importing to vHelix.

¹Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden. ²Department of Neuroscience, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

³Department of Computer Science, Aalto University, FI-00076 Aalto, Finland. ⁴Department of Cell and Molecular Biology, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

The Berkeley Out-of-Order Machine (BOOM): An Industry-Competitive, Synthesizable, Parameterized RISC-V Processor

*Christopher Celio
David A. Patterson
Krste Asanović*

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-167

<http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-167.html>

June 13, 2015



Embryo Selection for Cognitive Enhancement: Curiosity or Game-changer?

Carl Shulman and Nick Bostrom

Future of Humanity Institute, Oxford University

Abstract

Human capital is an important determinant of individual and aggregate economic outcomes, and a major input to scientific progress. It has been suggested that advances in genomics may open up new avenues to enhance human intellectual abilities genetically, complementing environmental interventions such as education and nutrition. One way to do this would be via embryo selection in the context of *in vitro* fertilization (IVF). In this article, we analyze the feasibility, timescale, and possible societal impacts of embryo selection for cognitive enhancement. We find that embryo selection, on its own, may have significant (but likely not drastic) impacts over the next 50 years, though large effects could accumulate over multiple generations. However, there is a complementary technology – stem cell-derived gametes – which has been making rapid progress and which could amplify the impact of embryo selection, enabling very large changes if successfully applied to humans.

Policy Implications

- Recent advances in embryo testing and genomics indicate that embryo selection for modest cognitive enhancement in humans may become feasible within five to ten years. This may spark wider public debate on the desirability of genetic enhancements in humans.
- The effects of conventional embryo selection on the first offspring produced would likely be small relative to past environmental improvements on cognitive ability. However, cumulative effects would be much greater over multiple generations.
- Biomedical research into human stem cell-derived gametes may enable iterated embryo selection (IES) *in vitro*, compressing multiple generations of selection into a few years or less.
- Regulators could speed or slow advance through rules on stem cell research and the private consumer genomics market. Science funding agencies can adjust their support for research into cognitive genomics and stem cell gametes.
- The extent of adoption of human genetic selection may significantly influence national competitiveness and global economic and scientific productivity in the second half of the century.

From carrier-screening to cognitive enhancement

Infantile Tay-Sachs disease is a recessive genetic condition which normally kills before the age of four. When mass genetic testing was introduced into the North American Ashkenazi Jewish community, which suffers from an elevated prevalence of the disease, rates fell by over 90 per cent (Kaback, 2000). Given genetic knowledge, most parents make use of alternative reproductive methods to have healthy children.

Advances in biotechnology are now making it possible to cheaply test for all known single-gene conditions simultaneously. US genomics firm 23andMe offers tests for most common genetic variants, as well as disease alleles, for \$99 (23andMe, 2013). As DNA sequencing costs continue to fall, it will become increasingly routine

for would-be parents to have thorough information on their own genetic makeup before having children.

The same technologies now used to avert genetic disease also seem likely to enable embryo selection for more complex heritable traits that involve many genes and environmental influences, such as height or cognitive ability. Instead of selecting an embryo based on a single variant, this would involve predicting embryo characteristics using data from many genes, and then selecting embryos using those predictions. The key technical barrier is that existing studies have failed to provide the necessary predictive power, leaving most of the heritability of these traits in study populations unexplained.

Early studies of the genetics of cognitive ability used small samples that could only detect common genetic variants with large effects. However, recent work using genome-wide complex trait analysis suggests that most

	G P O L					
				1	2	3
	Journal Code			Manuscript No.		
			WILEY			
			No. of pages: 8			PE: Sumathi
			Dispatch: 28,12,13			CE:

COMPANY BACKGROUND

OVERVIEW

Empiriko is a clinical intelligence company based in Newton, MA, that has patented “chemosynthetic liver technology” – Biomimiks™ – that mimics the *in vivo* metabolism mediated by liver enzyme (cytochrome P450), providing a chemistry-based solution for generating and predicting more comprehensive metabolic profiles of drugs. With Biomimiks™ pharmaceutical companies can identify New Chemical Entities (NCEs) more predictably, redesign more effective compounds, identify potential drug toxicity and drug-to-drug interactions, and reduce drug attrition rates. Empiriko is in a unique position to capitalize on several key market factors, including: financial and competitive pressure on pharmaceutical companies to discover novel and safer drug candidates more quickly and cost-effectively, and regulatory pressure to find alternative ways to cut down on animal testing. Our expanding patent portfolio, engaged Scientific Advisory Board (including two Nobel Laureates) and experienced management team have established the foundation for Empiriko’s short-term and long-term success.

METABOLISM PREDICTION & PROFILING

Metabolism profiling, which is conducted relatively late in the drug discovery process, is a widely-used means of identifying toxicity and potential side effects, and selecting the best drug candidates for further study. In humans and other animals, most small molecule drugs are metabolized in the liver – necessitating the current *in vivo* and *ex vivo* practice of studying liver enzymes (cytochrome P450 or CYP) mediated drug metabolism. When liver enzymes react with drugs, the chemical end-products are called “metabolites.” In many cases, toxic metabolites counteract the beneficial effects of drugs, causing harmful side effects. Since a majority of drug metabolites are implicated in adverse reactions, metabolites hold the key to understanding many critical pharmacological processes.

Today, scientists isolate and identify drug metabolites using liver slices, microsomal preparations and bodily fluids extracted from animals. Besides involving animals, the present-day methods for profiling metabolites are labor-intensive, time-consuming, costly and chemically inconclusive. The water-based or “aqueous” solutions used during testing preclude isolation and characterization of highly soluble drug metabolites. This leads to an incomplete metabolism profile which can derail drug development or present challenges with serious side effects when the drug reaches the market. Although animal sacrifice cannot be completely eliminated from drug testing, it can be reduced significantly during drug research with the use of more advanced chemical-based systems (e.g., chemosynthetic livers).

PRODUCT & TECHNOLOGY

Empiriko’s unique biomimetic platform – Biomimiks™ – accurately mimics the structure and function of liver enzymes, serving as chemically-based synthetic livers that can be used to more precisely predict metabolism patterns, pathways and profiles. We have made significant improvements in oxidative catalyst design by adding steric barriers and halogen atoms to our Biomimiks™ catalysts. These saddle-shaped catalysts are incredibly stable and fast-reacting, and they generate the large volumes of metabolites needed for further research.

Biomimiks™ solutions address the full drug development lifecycle, from lead generation to patient treatment.

- **For Drug Discovery (now available for customers):** With Biomimiks™ scientists can now conduct early metabolism studies, generating quantitative measures of toxicity and reducing the necessity of animal experimentation. This approach affords an efficient method for the systematic preparation and identification of the entire spectrum of metabolites from a series of compounds that are under non-clinical evaluation.
- **For Personalized Patient Treatment:** With Biomimiks™ clinicians will soon be able to analyze metabolic patterns to assess drug-to-drug interactions and the safety and efficacy associated with adding a new drug to a patient’s drug regimen. This enables clinicians to match the right combination of drugs to the right patient and achieve the best outcome for the patient.

PRODUCT ROADMAP

Empiriko’s short-term product strategy for Biomimiks™ is to provide chemistry-based *in vitro* and *ex vivo* solutions that can be immediately applied in drug metabolite synthesis, screening and toxicity evaluation for drug discovery, development and clinical trials.

- **Empiriko’s short-term product strategy** for Biomimiks™ is to provide chemistry-based *in vitro* and *ex vivo* solutions that can be immediately applied in drug metabolite synthesis, screening and toxicity evaluation for drug discovery, development and clinical trials.

```
main( )  
{  
    Learning by doing  
    Engineering + biology = ?  
    Now what?  
}
```

```
# endy@stanford.edu  
# Yale NUS STEM Innovation  
# 27 April 2017  
# fair use, public domain, have fun
```

Short Technical Reports

SHORT TECHNICAL REPORTS

Manuscripts published in the Short Technical Reports section are shorter and less comprehensive in scope than full Research Reports.

Error Analysis of Chemically Synthesized Polynucleotides

BioTechniques 24:256-260 (February 1998)

ABSTRACT

Two single-stranded polynucleotide constructs, 123 and 126 nucleotides in length, were chemically synthesized using standard phosphoramidite chemistry. Clonable, double-stranded DNA fragments about 100-bp long were prepared from the polynucleotides by primer extension with a DNA polymerase and end-trimming with two restriction endonucleases, then the fragments were ligated into separate plasmids. Errors in individual insert copies were determined by dideoxy sequencing after in vivo amplification of plasmids. Five of the ten inserts sequenced contained errors, including seven single-base-pair deletions, one four-base-pair deletion and one G@C transversion. The origins of the latter two errors are unclear, but single-base deletions are inconsistent with errors of polymerases; thus, the most common sequence errors of chemical synthesis are deletion mutations. Deletions are most likely to result from incomplete capping or de-tritylation. The observed error rate can become a significant limiting factor in applications that depend on the correctness of a polynucleotide sequence in individual insert clones.

INTRODUCTION

Nucleic acid synthesis is a fast, economical and reliable technology (13,14). High-quality oligonucleotides containing up to approximately 30 nucleotides (nt) are routinely used as sequencing primers, PCR primers or hybridization probes. Polynucleotides with significantly longer chains have been designed and synthesized for additional applications, such as the identification of binding motifs in nucleic acid chains (4,17) and the detection and quantitation of sequence-selective drug binding to nucleic acids (6,8,15).

DNA synthesis by the standard phosphoramidite method (2,12) is a

simple, repetitive process by which chosen nucleosides are added successively to a DNA chain immobilized on a solid support. In brief, a 3' phosphoramidite nucleoside with a protected 5'-hydroxyl group and base amino groups, supplied as a monomer in solution, is coupled to a 5'-hydroxyl group of an immobilized growing DNA chain. A subsequent capping step blocks any unreacted 5'-hydroxyl groups from the prior coupling, thereby preventing elongation of incomplete sequences in further coupling steps. Two steps are required after capping before chain elongation is begun. (i) An oxidation step converts the inter-nucleotide linkage from a phosphite to a more stable phosphotriester. (ii) Acid treatment then removes the trityl protecting group from the 5' hydroxyl of the nucleoside added last, making it available for chain elongation. Completed oligonucleotides are cleaved from the support, and blocking groups are removed with ammonium hydroxide.

While coupling-step efficiencies are routinely monitored during syntheses, the efficiencies of other steps that may influence the sequences of individual chains, such as capping, are not easily monitored. Recently, we had the opportunity to examine errors made in individual chains during the preparation of clonable (ca. 100-bp) double-stranded (ds)DNA fragments from synthetic polynucleotides. As part of our studies of sequence-selective DNA interactions with small molecules (8,15), we chemically synthesized two single-stranded polynucleotides (123 and 126 nt in length, respectively) by the standard phosphoramidite method. dsDNA fragments were prepared after isolation of full-length and near full-length synthesis products by primer extension with a DNA polymerase, then the opposing ends were trimmed with two different restriction endonucleases. Trimmed fragments were cloned directionally into plasmid vectors. After in vivo amplification of the plasmids, the sequences of inserts from several independent clones were determined by dideoxy sequencing and analyzed for errors relative to the intended sequence.

A surprisingly high fraction of inserts contained sequence errors, primarily single-base deletions. A single



Getting started with
OpenBSD device driver development

Stefan Sperling <stsp@openbsd.org>

EuroBSDcon 2017



Ex vivo DNA assembly

Adam B. Fisher¹, Zachary B. Canfield², Laura C. Hayward², Stephen S. Fong^{1,2*} and George H. McArthur IV^{2*}

¹ Integrative Life Sciences Program, Virginia Commonwealth University, Richmond, VA, USA

² Department of Chemical and Life Science Engineering, Virginia Commonwealth University, Richmond, VA, USA

Edited by:

Zhanglin Lin, Tsinghua University, China

Reviewed by:

Weiwen Zhang, Tianjin University, China

Mario Andrea Marchisio, Harbin Institute of Technology, China

*Correspondence:

Stephen S. Fong and
George H. McArthur IV, Department
of Chemical and Life Science
Engineering, Virginia Commonwealth
University, 601 West Main Street,
Richmond, VA 23284-3028, USA
e-mail: ssfong@vcu.edu;
mcarthurgh@vcu.edu

Even with decreasing DNA synthesis costs there remains a need for inexpensive, rapid, and reliable methods for assembling synthetic DNA into larger constructs or combinatorial libraries. Advances in cloning techniques have resulted in powerful *in vitro* and *in vivo* assembly of DNA. However, monetary and time costs have limited these approaches. Here, we report an *ex vivo* DNA assembly method that uses cellular lysates derived from a commonly used laboratory strain of *Escherichia coli* for joining double-stranded DNA with short end homologies embedded within inexpensive primers. This method concurrently shortens the time and decreases costs associated with current DNA assembly methods.

Keywords: DNA assembly, *ex vivo*, end joining, cellular lysates, colorimetric screen, synthetic biology, genetic engineering

INTRODUCTION

Our capacity to (re)engineer living systems is linked to our ability to physically build specific DNA molecules that encode desired functionality and behavior. Traditionally, recombinant DNA has been constructed using restriction cloning (i.e., cutting with an endonuclease and joining with a ligase). In addition to restriction-ligation approaches, site-specific recombination systems have been employed to assemble DNA with great success (Hartley et al., 2000). However, genetic engineering has recently become more flexible with the use of sequence-independent approaches that take advantage of the decreasing cost of DNA synthesis (Li and Elledge, 2007; Gibson et al., 2009; Quan and Tian, 2009, 2011; Zhang et al., 2012). For example, a completely synthetic *Mycoplasma* genome was assembled from chemically synthesized double-stranded DNA (dsDNA) fragments (5–7 kbp each) using an *in vitro* chew-back assembly method in 2008 (Gibson et al., 2008). The authors report that the assembled half-genome DNA molecules could not be propagated by *E. coli* due to their size and therefore relied on yeast transformation-associated recombination (TAR), an *in vivo* assembly in *Saccharomyces cerevisiae*, to finish constructing the genome (Ma et al., 2002). More recently, the mouse mitochondrial genome was reconstructed from overlapping synthetic single-stranded oligonucleotides (60 nucleotides each) using a three-enzyme *in vitro* isothermal DNA assembly method (ISO assembly, also known as Gibson assembly) (Smith et al., 2010).

These homology-based approaches are extremely versatile, but there are disadvantages associated with each. Although isothermal DNA assembly provides a quick and reliable way to simultaneously join together multiple pieces of DNA, the cost of the purified enzymes needed to carry out the assembly reaction is non-trivial (~US\$3/reaction formulated in-house). TAR cloning in *S. cerevisiae* is considerably cheaper but requires: (1) preparation of

yeast spheroplasts, (2) the inclusion of both a selection marker and replication origin that function in yeast, and (3) the subsequent purification of the assembled DNA product (if the DNA is intended to be used in an organism other than yeast). This is a long process that takes 8–9 days to complete (Gibson et al., 2008). To address the monetary cost of ISO assembly and the time cost of TAR cloning, we hypothesized that the DNA repair machinery endogenous to yeast and other microorganisms would remain functional within cellular lysates and should be able to catalyze the assembly of linear and circular DNA constructs in just hours.

We tested this hypothesis using cellular lysates derived from *S. cerevisiae*, *E. coli*, and *Deinococcus radiodurans*, an extremophilic bacterium known to have exceedingly efficient dsDNA repair capabilities. The overall goal of this study was to determine if this general lysate-based approach, termed *ex vivo* DNA assembly, is feasible for physically joining DNA molecules quickly, cheaply, and efficiently. The *ex vivo* assembly reactions were characterized by gel electrophoresis analysis and by a colorimetric screen of the resulting *E. coli* transformants (in which colonies housing the correctly assembled plasmid express a vibrant blue chromoprotein). Here we demonstrate that *ex vivo* dsDNA assembly is organism-dependent (*E. coli* lysate is able to perform end joining, but lysate derived from *S. cerevisiae* or *D. radiodurans* is not) and we show that *in vivo* circularization of overlapping dsDNA occurs after transformation in *E. coli*, a process that is often overlooked when characterizing *in vitro* DNA assembly methods.

RESULTS

To assess the ability of select cellular lysates to join together dsDNA, we first designed two amplicons with appropriate overlapping ends (26 and 30 bp overlaps) to be assembled into a circular plasmid (Figure 1A). Correctly assembled plasmids endow *E.*

Facilitation and restoration of cognitive function in primate prefrontal cortex by a neuroprosthesis that utilizes minicolumn-specific neural firing

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 J. Neural Eng. 9 056012

(<http://iopscience.iop.org/1741-2552/9/5/056012>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 152.11.187.85

The article was downloaded on 13/09/2012 at 17:41

Please note that [terms and conditions apply](#).

Published in final edited form as:

J Neural Eng. 2013 December ; 10(6): 066013. doi:10.1088/1741-2560/10/6/066013.

Facilitation of Memory Encoding in Primate Hippocampus by a Neuroprosthesis that Promotes Task Specific Neural Firing

Robert E. Hampson¹, Dong Song², Ioan Opris¹, Lucas M. Santos¹, Dae C. Shin², Greg A. Gerhardt³, Vasilis Z. Marmarelis², Theodore W. Berger², and Sam A. Deadwyler¹

¹Department of Physiology and Pharmacology, Wake Forest University School of Medicine, Winston-Salem, NC

²Department of Biomedical Engineering, University of Southern California, LA, CA

³Department of Anatomy and Neurobiology, University of Kentucky, Lexington, KY

Abstract

Objective—Memory accuracy is a major problem in human disease and is the primary factor that defines Alzheimer's, aging and dementia resulting from impaired hippocampal function in medial temporal lobe. Development of a hippocampal memory neuroprosthesis that facilitates normal memory encoding in nonhuman primates (NHPs) could provide the basis for improving memory in human disease states.

Approach—NHPs trained to perform a short-term delayed match to sample (DMS) memory task were examined with multi-neuron recordings from synaptically connected hippocampal cell fields, CA1 and CA3. Recordings were analyzed utilizing a previously developed nonlinear multi-input multi-output (MIMO) neuroprosthetic model, capable of extracting CA3-to-CA1 spatiotemporal firing patterns during DMS performance.

Main Results—The MIMO model verified that specific CA3-to-CA1 firing patterns were critical for successful encoding of Sample phase information on more difficult DMS trials. This was validated by delivery of successful MIMO-derived encoding patterns via electrical stimulation to the same CA1 recording locations during the Sample phase which facilitated task performance in the subsequent delayed Match phase on difficult trials that required more precise encoding of Sample information.

Significance—These findings provide the first successful application of a neuroprosthesis designed to enhance and/or repair memory encoding in primate brain.

Keywords

Hippocampal neuron; spike train; memory encoding; nonlinear model; patterned electrical stimulation; prosthesis; prosthetic; memory facilitation; memory retention; nonhuman primate

INTRODUCTION

Memory deficits in humans are constantly related to an inability to recall items previously exposed in different contexts or to utilize the same items for different purposes (Jenkins and Ranganath 2010; Tubridy and Davachi 2011). It has been known from the initial

*Corresponding Author: Dr. Sam A. Deadwyler, Wake Forest University School of Medicine, Department of Physiology and Pharmacology, Medical Center Boulevard, Winston Salem NC 27157, ph: 336-716-8540, fax: 336-716-8628, sdeadwyl@wfubmc.edu.

The authors declare no competing financial interests

Learning to Recognize Speech From Chaotically Synthesized Data

Faraz Fadavi Akhavan Bonab

Computer Science Department
Stanford University

fadavi@cs.stanford.edu

Samuel Ginn

Computer Science Department
Stanford University

samginn@cs.stanford.edu

Abstract

Academic researchers in speech recognition have for years been stymied by a lack of large data corpora of speech audio with accurate transcripts. To combat this, we present a novel method to synthetically create an arbitrarily large dataset of transcribed speech. We designed a synthetic speech generator to create speech audio from Wikipedia articles. In order to model human variation and natural noise, we injected “chaos” into the speech generation pipeline. This newly generated corpus of chaotically synthesized speech became the input data to our end-to-end deep learning speech recognizer. Through experiments, we show that our synthetic speech data can be successfully used to augment a model trained off of the Switchboard dataset. With more computation, we hope that the synthetic speech model will be able to exceed state-of-the-art results in speech recognition.

1 Introduction

In the preceding decade, the artificial intelligence community has seen breakthrough advancements with deep learning in machine learning tasks such as image recognition and natural language processing. In speech recognition as well, deep learning has spurred a revolution. Earlier this year, Saon, et. al. reached a state-of-the-art 5.5% word error rate on the canonical Switchboard dataset through an end-to-end deep learning system (Saon et al., 2017). Yet, no one has been able to match human parity yet (estimated to be 5.1% word error rate) (Saon et al., 2017). Although these deep learning techniques have enabled tremendous improvements in almost all machine learning tasks

(especially, speech recognition), they rely on a huge amount of data to adequately train. In computer vision, this challenge was met by the ImageNet team when they compiled over 14 million labeled images to train their neural networks (Deng et al., 2009).

For years, however, the state-of-the-art in academic automatic speech recognizers have relied on a small set of corpora with limited data. Furthermore, very few large datasets exist for languages other than the most popular, posing challenges for universal speech recognition across languages. Unlike image classification, speech recognition data collection is laborious and expensive. In order to be adequate for training, the data must have a speech audio file with an accurately labeled transcript. In the academic world, Table 1 shows the current datasets most often used by researchers with their respective state-of-the-art benchmarks. These paltry datasets have hindered researchers ability to truly leverage the depth of deep learning models.

We seek to solve this problem. Hannun et al. concluded their breakthrough Deep Speech paper by suggesting that “this approach will continue to improve as we capitalize on increased computing power and dataset sizes in the future” (2014). Yet, dataset sizes cannot increase without substantial monetary and human investment in order to tediously procure the data. If, however, there existed a way to artificially synthesize such a dataset, it may allow Hannun and others to utilize their deep learning innovations in such a way to truly surpass human quality in speech recognition.

We present a novel technique to acquire such large-scale data in order to properly train a deep neural network. Recognizing the vast amount of written language publicly available on the Internet and the latest improvements in speech synthesis, we developed a proof-of-concept system that

Brazing Molybdenum and Graphite with a Titanium-Based Powder Filler Metal

Examined were the different filler metal powder application amounts per brazed area as well as the texture of the graphite side of the brazed joint

BY I. V. FEDOTOV, C. E. RICHMAN, O. N. SEVRYUKOV, A. N. SUCHKOV, J. LI,
B. A. KALIN, V. T. FEDOTOV, AND A. A. IVANNIKOV

ABSTRACT

A new method of brazing Mo-C joints for mechanical performance up to 1650°C was found. A Ti-40 Zr-8.5 Nb-1.5 Be powder filler metal was created for this brazing application, and its melting range and phase composition were found. The effects of different filler metal powder application amounts per brazed area and the texture of the graphite side of the brazed joint were studied. EDS microanalysis of the brazed joint was carried out and the connections were analyzed for shear strength and porosity. EDS analysis revealed Ti, Zr, and Nb carbides were present in the brazed joint. When the graphite surface was smooth, the most high-quality joints were obtained with a powder application of 0.5 g/cm². It was found that texturing the graphite surface with concentric notches increased the shear strength of the joints by 2.5×. The strongest brazed connection, in which the graphite surface was notched, was subjected to a remelting test. The braze was maintained when heated to a temperature of 1650°C, though the composition of the brazed joint changed, exhibiting a higher concentration of carbides near the tips of the notches.

KEYWORDS

• Brazing • Joining • Molybdenum • Graphite • Microstructure

Introduction

Production of x-ray tubes for computed tomography (CT) and angiography devices requires methods of anode component joining that can ensure reliable, long-term equipment operation. The anode of an x-ray tube consists of a molybdenum disk with a tungsten layer deposited on it by vapor deposition and a heat sink made from high-density graphite MPG-6. Because the anode operates at high temperatures and tangential loads induced by rapid rotation (9000 rev/min) to achieve high x-ray fluence rates, this Mo-C joint must be very high performance.

A known method of joining the molybdenum alloy TZM and high-density graphite is by contact-reactive brazing with zirconium as an interlayer filler metal (Ref. 1). Brazing is carried out at temperatures above the eutectic temperature of Mo-Zr (1550°C), and thermal stresses can occur during postbrazing cooling, affecting the reliability of the brazed joint. Therefore, using filler metals with lower melting points than zirconium is preferred for brazing molybdenum.

Filler metal alloys based on silver activated by titanium or zirconium are used for brazing carbon (Refs. 2–4) as well as filler alloys based on titanium

or zirconium with nickel and a small amount of added copper (Refs. 5, 6). Previous research has demonstrated success in brazing molybdenum by using titanium (Ref. 7).

In this work, molybdenum and graphite MPG-6 were brazed with a Ti-40 Zr-8.5 Nb-1.5 Be filler metal powder containing a particle size less than 50 microns, ground down from a stock of rapidly quenched ribbons of the alloy. Many studies (Refs. 8, 9) have shown that the usage of rapidly quenched filler metals yields several advantages for achieving high-quality joints, such as high diffusion and capillary activity.

Experimental Setup and Procedure

Ingots of the filler metal composed with iodide titanium and zirconium, niobium NB1, and alloy Ti-6% Be were melted in an argon-arc furnace MEPHI-9 (USSR, 1980). The filler metal ribbon was produced by melt spinning in the Crystal-702 facility (USSR, 1975). An optical pyrometer Promin (USSR, 1980) was used to monitor the start and end temperatures of the alloy melting.

Then the filler metal ribbon was heat treated in a vacuum oven Xerion XRETORT600 (Germany, 2014). The resulting embrittled filler metal ribbon was ground in a planetary mill Pul-

I. V. FEDOTOV (fed_ivan@mail.ru), O. N. SEVRYUKOV, A. N. SUCHKOV, B. A. KALIN, V. T. FEDOTOV, and A. A. IVANNIKOV are with the National Research Nuclear University, Moscow Engineering Physics Institute, Moscow, Russia. C. E. RICHMAN (rcamille@mit.edu) is with MIT (Massachusetts Institute of Technology), Cambridge, Mass. J. LI is with the Department of Nuclear Science and Engineering, and Department of Materials Science and Engineering, MIT, Cambridge, Mass.



Probing Chromatin-modifying Enzymes with Chemical Tools

Item type	Article
Authors	Fischle, Wolfgang; Schwarzer, Dirk
Citation	Probing Chromatin-modifying Enzymes with Chemical Tools 2016 ACS Chemical Biology
Eprint version	Post-print
DOI	10.1021/acschembio.5b01023
Publisher	American Chemical Society (ACS)
Journal	ACS Chemical Biology
Rights	This document is the Accepted Manuscript version of a Published Work that appeared in final form in ACS Chemical Biology, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see http://pubs.acs.org/doi/abs/10.1021/acschembio.5b01023 .
Downloaded	22-Oct-2017 17:52:10
Link to item	http://hdl.handle.net/10754/600475

Securing the Datacenter with FlexVer™

Raptor Engineering, LLC
<https://www.raptorengineering.com>



Overview

In this document we explore the revolutionary FlexVer™ security technology introduced on our Talos™ mainboards, and answer the most frequently asked questions regarding usage of this technology.

What is FlexVer™?

FlexVer™ is a new, owner-controlled security technology designed to safeguard critical data and applications in the event of software or hardware tampering. FlexVer™ allows a system to be provisioned in a trusted physical environment, then deployed to an untrustworthy physical location while retaining system integrity. Provided that OS-level attack avenues are properly mitigated, for example through the use of TRESOR and similar technologies, FlexVer™ allows deployment of provisioned systems without concern of hardware and/or software tampering and subsequent extraction of sensitive material -- a provisioned system can be guaranteed to be answering only to its previously configured owner, not the owner of the physical space in which the system resides. This is a major departure from prevailing security models, which largely assume that either the possibility of physical access by a malicious actor must result in loss of trust of the affected system, or that trust must be delegated to the system vendor in all situations.

How is FlexVer™ different than Intel® Boot Guard™ and related technologies?

Unlike existing security technologies, FlexVer™ does not depend on a fully trustworthy vendor for the root of system trust. Recent events have shown that this trustworthy vendor assumption is not valid, and in fact there is strong pressure on all vendors to compromise their root of trust for financial gain, warrant-related data extraction, industrial espionage, and related purposes. Any given security technology is only as secure as the weakest link in the chain; Boot Guard™ and related technologies operate by permanently locking the hardware to a vendor-controlled signing key, not only keeping the vendor in complete control of the hardware at all times, but also creating a single point of failure by which millions of systems could potentially be compromised with a single hack or leaked key. Effectively, the vendor and their partners' software, data security processes, and key handlers have become the weakest link in the chain, offering a large attack surface and severely weakening all systems based on this security model.

In contrast, FlexVer™ abandons this centralised security model, using a distributed, locally-verified model instead. FlexVer™ becomes the local root of trust for each protected system, removing the possibility of a single data breach compromising all systems simultaneously and eliminating the capability for a vendor or its affiliates to access protected data on your system. Under the FlexVer™ security model, each system is provisioned in a secure, trusted physical environment by trusted members of an organization. The FlexVer™ hardware definition files and resultant bitstream are verified to be trustworthy, and this trusted bitstream is then loaded into the FlexVer™ hardware. Immediately upon FlexVer™ startup, a unique internal key is

How Many People Can the World Support?

John Heaver Fremlin (1913-95)

New Scientist, No. 415, 285-287, 1964

The world population is now about 3,000 million and is increasing at a rate corresponding to a doubling in 37 years. In view of the increasing importance attached to the immediate effects of the rapid growth in human numbers, it is of interest to examine ultimate technical limits to this growth. Traditionally, these limits have usually been regarded as fixed by possible food supplies. Although, in practice, at least in historical times, the actual limiting factor has more often been disease.

Diseases are now nearly, and will soon be entirely, eliminated as effective controllers of population growth but it is not at all clear that difficulties in food production will take their place. It is true that there is a limit to the improvement of agricultural output by application of existing scientific knowledge, but by the time this limit is reached other methods of food-production will have been devised. In this article I shall explore the possibility that the real limits are physical rather than biological.

I shall assume throughout an effective degree of world cooperation in the application of food technology, etc. This is quite evidently essential if the maximum world population is to be reached. There are of course many ways of not reaching the maximum, but none of these will be discussed here.

In order to give a time scale, it is supposed that the rate of increase of population remains constant at the present value -- that is to say, doubling every 37 years. In fact the rate is itself accelerating, so that, in the absence of limitations, this time scale will be too long.

Stage 1: up to 400,000 million in 260 years' time

Using existing crop plants and methods it may not be practicable to produce adequate food for more than four doublings of the world population, though the complete elimination of all land wildlife, the agricultural use of roofs over cities and roads, the elimination of meat-eating and the efficient harvesting of sea food might allow two or three further doublings -- say seven in all. That would give us, with the present doubling time of 37 years, 260 years to develop less conventional methods, and would allow the population of the world to increase to about 130 times its present size, or about 400,000 million.

Stage 2: up to 3 million million in 370 years' time

The area of ice-free sea is some three times that of land. Photosynthesis by single-celled marine organisms may be more efficient than that of the best land plants. If organisms could be found capable of the theoretical maximum efficiency (8 percent of total solar radiation, according to A. A. Niciporovic) we should gain a factor of three in yield. We could then double our numbers a further three more times if all the wildlife in the sea, too, was removed and replaced by the most useful organisms growing under controlled conditions, with the optimum concentration of carbonates, nitrates and minerals. (Of course a reserve of specimens of potentially useful species could be preserved, perhaps in a dormant state.) Again, for maximum efficiency we must harvest and consume directly the primary photosynthesis organisms, rather than allow the loss of efficiency involved in the food chains leading to such secondary organisms as zooplankton or fish.

By this stage, we should have had ten doublings, which at the present rate would take some 370 years, with a final world population of 3 million million. Since the world's surface (land and sea) is 500 million million square meters, each person would have a little over 160 square meters for his maintenance -- about a thirtieth of an acre -- which does not seem unreasonable by more than a factor of two, so long as no important human activity other than food production takes place on the surface.

No serious shortages of important elements need be envisaged so far, though extensive mining operations for phosphates might be needed, and we have not yet approached any real limit.

Stage 3: up to 15 million million in 450 years' time

At first sight, it seems that a very big leap forward could be taken if we use sources of power other than sunlight for photosynthesis. The solar power received at the earth's surface is only about 1 kilowatt per



Genome Engineering with Custom Recombinases

Thomas Gaj^{*,†,1,2}, Carlos F. Barbas III^{*,†}

^{*}Department of Chemistry, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California, USA

[†]Department of Cell and Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California, USA

¹Current address: University of California, Berkeley, California, USA.

²Corresponding author: e-mail address: gaj@berkeley.edu

Contents

1. Introduction	79
2. Target Identification	81
3. Recombinase Construction	82
4. Measurements of Recombinase Activity	85
4.1 Reporter plasmid construction	86
4.2 Luciferase assay	86
5. Site-Specific Integration	87
5.1 Donor plasmid construction	87
5.2 Cell culture methods	88
6. Conclusions	90
Acknowledgments	90
References	90

Abstract

Site-specific recombinases are valuable tools for myriad basic research and genome engineering applications. In particular, hybrid recombinases consisting of catalytic domains from the resolvase/invertase family of serine recombinases fused to Cys₂-His₂ zinc-finger or TAL effector DNA-binding domains are capable of introducing targeted modifications into mammalian cells. Due to their inherent modularity, new recombinases with distinct targeting specificities can readily be generated and utilized in a "plug-and-play" manner. In this protocol, we provide detailed, step-by-step instructions for generating new hybrid recombinases with user-defined specificity, as well as methods for achieving site-specific integration into targeted genomic loci using these systems.



1. INTRODUCTION

Hybrid recombinases composed of catalytic domains derived from the resolvase/invertase family of serine recombinases fused to engineered

Theoretical Neuroscience and Deep Learning Theory

Surya Ganguli

Dept. of Applied Physics,
Neurobiology,
and Electrical Engineering

Stanford University

Funding:

Bio-X Neuroventures
Burroughs Wellcome
Genentech Foundation
James S. McDonnell Foundation
McKnight Foundation
National Science Foundation

NIH
Office of Naval Research
Simons Foundation
Sloan Foundation
Swartz Foundation
Stanford Terman Award

<http://ganguli-gang.stanford.edu>

Twitter: @SuryaGanguli

**Genome Project-write: A Grand Challenge
Using Synthesis, Gene Editing and Other Technologies
to Understand, Engineer and Test Living Systems**

November 30, 2016

Author: Jef D Boeke, George Church, Andrew Hessel, Nancy J Kelley (GP-write Leadership Group) and The GP-Write Consortium†

†The authors and their affiliations maybe found at www.gpwrite.org

INTRODUCTION

- The Human Genome Project (HGP-read) was initiated in December 1984 at the DOE Alta meeting and launched on October 1, 1990. Aimed at a nearly complete “reading” of the human and other genomes with annotations of plausible gene functions, as well as the improvement of the technology, cost, and quality of DNA sequencing (1, 2), it was the largest life science project ever conducted and one of the great feats of exploration in history.
- The first draft of the human genome was announced in June 2000, and a more complete version in April 2003.
- Since this announcement, interpretation of that data was made possible by sequencing humans from other populations around the world (HapMap and 1KG) and by other “omics” tools yielding insights into the functionality at base pair precision (ENCODE), correlations between natural genome variants and various traits and diseases via genome-wide association studies (GWAS), and identification of genes and variants responsible for rare Mendelian traits by a combination of genomic sequencing and advanced computational methods.
- Incredibly powerful and inexpensive DNA sequencing and genetic analysis technologies, developed, in part, by the National Human Genome Research Institute’s (NHGRI) Advanced DNA Sequencing Technology Development program, reduced the cost of sequencing a human genome more than a million fold from \$3 billion to less than \$1,000 in just over a decade.
- Since then, the world has embarked on a revolution in science and healthcare that is changing the way we live and carries the promise of individualizing clinical delivery to improve health, and prevent and cure human disease. The Precision Medicine Initiative (for example), announced by the White House in 2015, would use the insights gained from “reading” each individual’s genome to personalize medical care based on an individual’s genes,

University of Groningen

The 2015 super-resolution microscopy roadmap

Hell, Stefan W.; Sahl, Steffen J.; Bates, Mark; Zhuang, Xiaowei; Heintzmann, Rainer; Booth, Martin J.; Bewersdorf, Joerg; Shtengel, Gleb; Hess, Harald; Tinnefeld, Philip

Published in:
Journal of Physics D-Applied Physics

DOI:
[10.1088/0022-3727/48/44/443001](https://doi.org/10.1088/0022-3727/48/44/443001)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hell, S. W., Sahl, S. J., Bates, M., Zhuang, X., Heintzmann, R., Booth, M. J., ... Cordes, T. (2015). The 2015 super-resolution microscopy roadmap. *Journal of Physics D-Applied Physics*, 48(44), [443001]. DOI: 10.1088/0022-3727/48/44/443001


Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



The Genome Project-write (HGP-write)

HGP-write Leadership

Jef Boeke, Ph.D.

Director, Institute for Systems Genetics
Professor, Department of Biochemistry and Molecular Pharmacology
NYU Langone Medical Center

George Church, Ph.D.

Robert Winthrop Professor of Genetics,
Harvard Medical School
Core Faculty Member,
Wyss Institute for Biologically Inspired Engineering at Harvard University
Professor of Health Sciences and Technology, Harvard and the Massachusetts Institute of Technology (MIT)
Associate Faculty Member, Broad Institute

Andrew Hessel

Distinguished Researcher,
Bio/Nano Research Group
Autodesk

Nancy J Kelley, J.D., M.P.P.

President & CEO
Nancy J Kelley & Associates
Formerly Founding Executive Director, New York Genome Center

Project Contact

Nancy J Kelley, J.D., M.P.P.
(617) 519-6144

Media Contact

Lynn Blenkhorn
lynn.blenkhorn@fkhealth.com
(617) 519-9827

A Grand Challenge: To Understand Life-write

The Genome Project-write (HGP-write) will be an open, academic, international scientific research project led by a multi-disciplinary group of scientific leaders who will oversee a reduction in the costs of engineering and testing large genomes in cell lines by over 1,000-fold within ten years. This will include whole genome engineering of human cell lines and other organisms of agricultural and public health significance. Additionally, HGP-write will develop new technologies and an ethical framework for genome-scale engineering, as well as transformative medical applications. The overarching goal of such an effort is to understand the blueprint for life provided by the Human Genome Project (HGP-read).

Biology's first large-scale project, HGP-read aimed to "read" a human genome. Successfully completed in 2003, HGP-read is now widely recognized as one of the great feats of exploration, one that sparked a global revolution in science and medicine, particularly in genomic-based diagnostics and therapeutics.

But our understanding of the human genome – and the full benefits to humanity to be obtained from this knowledge – remains far from complete. Many scientists now believe that to truly understand our genetic blueprint, it is necessary to "write" DNA and build human (and other) genomes from scratch. Such an endeavor will require research and development on a grand scale.

From Observation to Action

HGP-write is expected to generate considerable knowledge by connecting the sequence of bases in DNA with their physiological and functional behaviors. Importantly, scientists will be able to move beyond observation to action, and facilitate the use of biological engineering to address many of the global problems facing humanity.

Some applications that may arise from HGP-write that could have a significant impact on human health include, but are not limited to, growing transplantable human organs, engineering immunity to viruses in cell lines, engineering cancer resistance into therapeutic cell lines, and enabling high-productivity, cost-efficient vaccine and pharmaceutical development using human cells and organoids that makes precision medicine more affordable and universal.

Responsible Innovation

Genome synthesis is a natural extension of the genetic engineering tools that have been safely used within the biotechnology industry for the past 40 years and have provided significant benefits to society. However, recent technological advancements, such as standardized genomic parts and CRISPR/Cas9 genome editing technology are the field and creating uncertainty in how these technologies will be applied.



HGP-write: Testing Large Genomes in Cells

Meeting Summary

Jef D. Boeke, George Church, Andrew Hessel, and Nancy J Kelley

MAY 10, 2016
BOSTON, MA

PREPARED BY FEINSTEIN KEAN HEALTHCARE





CENTER *of* EXCELLENCE
for ENGINEERING BIOLOGY



Pilot Project Proposal

(Not to exceed two pages)

Name of Project:

Proposer and Contact Information:

Background:

Technical Idea:

Utility:

“Fit” For GP-write

Submit to:

info@engineeringbiologycenter.org



High molecular weight DNA assembly in vivo for synthetic biology applications

Mario Juhas & James W. Ajioka

To cite this article: Mario Juhas & James W. Ajioka (2016): High molecular weight DNA assembly in vivo for synthetic biology applications, Critical Reviews in Biotechnology, DOI: [10.3109/07388551.2016.1141394](https://doi.org/10.3109/07388551.2016.1141394)

To link to this article: <http://dx.doi.org/10.3109/07388551.2016.1141394>



© 2016 The Author(s). Published by Taylor & Francis.



Published online: 10 Feb 2016.



Submit your article to this journal [↗](#)



Article views: 542



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)

High-throughput mapping of single neuron projections by sequencing of barcoded RNA

Justus M Kebschull^{1,2}, Pedro Garcia da Silva^{2,3}, Ashlan P Reid², Ian D Peikon^{1,2}, Dinu F Albeanu², Anthony M Zador²

¹Watson School of Biological Sciences, Cold Spring Harbor, NY 11724, USA

²Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

³Champalimaud Centre for the Unknown, Lisbon, Portugal

Correspondence and requests should be addressed to Anthony Zador, zador@cshl.edu.

Summary

Neurons transmit information to distant brain regions via long-range axonal projections. In the mouse, area-to-area connections have only been systematically mapped using bulk labeling techniques, which obscure the diverse projections of intermingled single neurons. Here we describe MAPseq (Multiplexed Analysis of Projections by Sequencing), a technique that can map the projections of thousands or even millions of single neurons by labeling large sets of neurons with random RNA sequences ("barcodes"). Axons are filled with barcode mRNA, each putative projection area is dissected, and the barcode mRNA is extracted and sequenced. Applying MAPseq to the locus coeruleus (LC), we find that individual LC neurons have preferred cortical targets. By recasting neuroanatomy, which is traditionally viewed as a problem of microscopy, as a problem of sequencing, MAPseq harnesses advances in sequencing technology to permit high-throughput interrogation of brain circuits.

Introduction

Neurons transmit information to distant brain regions via long-range axonal projections. In some cases, functionally distinct populations of neurons are intermingled within a small region. For example, nearby hypothalamic nuclei regulate basic drives including appetite, aggression, and sexual attraction (Kennedy et al., 2014; Sternson, 2013), and neurons from these nuclei project to distinct targets. In visual cortical area V1, responses to visual stimuli are matched to the properties of the higher visual areas to which the neurons project (Glickfeld et al., 2013; Movshon and Newsome, 1996). Findings such as these suggest that the information transmitted by individual neurons may be tailored to their targets. Such selective routing of information requires an anatomical substrate, but there is currently no high-throughput method for determining the diverse projection patterns of individual neurons.

At present, there is a steep tradeoff between throughput and resolution in anatomical approaches to mapping long-range connections. In conventional anterograde brain mapping studies, a



Neural Decoding of Visual Imagery During Sleep

T. Horikawa *et al.*

Science **340**, 639 (2013);

DOI: 10.1126/science.1234330

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of October 16, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/340/6132/639.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2013/04/04/science.1234330.DC2.html>

<http://www.sciencemag.org/content/suppl/2013/04/03/science.1234330.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/340/6132/639.full.html#related>

This article **cites 34 articles**, 10 of which can be accessed free:

<http://www.sciencemag.org/content/340/6132/639.full.html#ref-list-1>

This article has been **cited by** 2 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/340/6132/639.full.html#related-urls>

This article appears in the following **subject collections**:

Neuroscience

<http://www.sciencemag.org/cgi/collection/neuroscience>

How Hard is Artificial Intelligence? Evolutionary Arguments and Selection Effects

(2012) REVISED VERSION

Carl Shulman

Nick Bostrom

[Forthcoming in the *Journal of Consciousness Studies*]

www.nickbostrom.com

Abstract

Several authors have made the argument that because blind evolutionary processes produced human intelligence on Earth, it should be feasible for clever human engineers to create human-level artificial intelligence in the not-too-distant future. This evolutionary argument, however, has ignored the observation selection effect that guarantees that observers will see intelligent life having arisen on their planet no matter how hard it is for intelligent life to evolve on any given Earth-like planet. We explore how the evolutionary argument might be salvaged from this objection, using a variety of considerations from observation selection theory and analysis of specific timing features and instances of convergent evolution in the terrestrial evolutionary record. We find that, depending on the resolution of disputed questions in observation selection theory, the objection can be either be wholly or moderately defused, although other challenges for the evolutionary argument remain.

1. Evolutionary arguments for easy intelligence

1.1 Introduction

What can human evolution tell us about the prospects for human-level Artificial Intelligence (AI)?¹ A number of philosophers and technologists, including David Chalmers (2010) and Hans Moravec (1976, 1988, 1998, 1999), argue that human evolution shows that such AI is not just possible but feasible within this century. On these accounts, we can estimate the relative capability of evolution and human engineering to produce intelligence, and find that human engineering is already vastly superior to evolution in some areas and is likely to become superior in the remaining areas before too long. The fact that evolution produced intelligence therefore indicates that human engineering will be able to do the same. Thus, Moravec writes:

¹ Here, we mean systems which match or exceed the cognitive performance of humans in virtually all domains of interest: uniformly “human-level” performance seems unlikely, except perhaps through close emulation of human brains (Sandberg and Bostrom, 2008), since software is already superhuman in many fields.

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

ELSEVIER INC., ELSEVIER B.V., ELSEVIER
LTD.

Plaintiffs,

v.

SCI-HUB d/b/a WWW.SCI-HUB.ORG, THE
LIBRARY GENESIS PROJECT d/b/a
LIBGEN.ORG, ALEXANDRA ELBAKYAN,
JOHN DOES 1-99,

Defendants.

Index No. 15-cv-4282 (RWS)

PROPOSED ORDER + *JB*

JUDGMENT

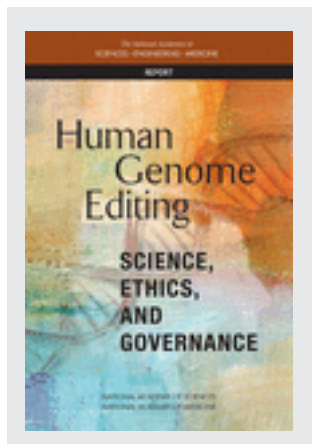
USDC SDNY
DOCUMENT
ELECTRONICALLY FILED
DOC #:
DATE FILED: 6-21-17

On June 3, 2015, Plaintiffs Elsevier Inc., Elsevier B.V., and Elsevier Ltd. (collectively "Elsevier" or "Plaintiffs") commenced this action with a Complaint seeking damages and injunctive relief against Sci-Hub and its operator Alexandra Elbakyan and (collectively the "Sci-Hub Defendants") and The Library Genesis Project and its operators (collectively the "LibGen Defendants"). On June 11, 2015, Plaintiffs moved the Court for an Order to Show Cause for a Preliminary Injunction, which this Court granted on June 18, 2015. On July 1, 2015, Plaintiffs filed Affidavits of Service attesting to service on all Defendants.

The Defendants, having been served with the Complaint, Order to Show Cause, and related papers, did not appear or oppose Plaintiffs' application for a Preliminary Injunction. On October 30, 2015, this Court granted Plaintiffs' application and entered a Preliminary Injunction against all Defendants.

This PDF is available at <http://www.nap.edu/24623>

SHARE



Human Genome Editing: Science, Ethics, and Governance

DETAILS

300 pages | 6 x 9 | PAPERBACK

ISBN 978-0-309-45288-5 | DOI: 10.17226/24623

CONTRIBUTORS

Committee on Human Gene Editing: Scientific, Medical, and Ethical Considerations; National Academy of Sciences; National Academy of Medicine; National Academies of Sciences, Engineering, and Medicine

GET THIS BOOK

FIND RELATED TITLES

Visit the National Academies Press at NAP.edu and login or register to get:

- Access to free PDF downloads of thousands of scientific reports
- 10% off the price of print titles
- Email or social media notifications of new titles related to your interests
- Special offers and discounts



Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press. ([Request Permission](#)) Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences.

Copyright © National Academy of Sciences. All rights reserved.



100-Person Mars Transfer Vehicle using Torpor-Inducing Habitats

66th International Astronautical Congress, Jerusalem, Israel
13 October 2015

Mark Schaffer

Senior Aerospace Engineer
mark.schaffer@sei.aero | +1.770.379.8013

John Bradford, Ph.D.

President and COO, SpaceWorks Enterprises, Inc.
john.bradford@sei.aero | +1.770.379.8007

Doug Talk, M.D.

Medical Advisor
doug.talk@sei.aero | +1.770.379.8000

All iGEM Teams Ever (2008-2016)

This is a list of all iGEM Teams where we could find abstracts to. 2016 and 2014 are missing unfortunately. Please use this as a reference and if you want to see if your idea has been done before, simply press CTRL+F or Apple+F and search for some key words of your idea.

You can also click the team names to get to their wiki page.

One sentence summaries:

<http://diyhp1.us/wiki/dna/projects/#igem-2013>

diybio mailing list thread:

<https://groups.google.com/d/msg/diybio/Cu8PfDh2IHY/wlkaH9iMEQAJ>

2015

Team AHUT China: Shining Sanctifier

Water, the origin of life, is the necessary and elementary component of our daily life. Various kinds of means have been developed to dispose nitrite and ammonium which are the main contaminants of this type of effluent. One of them is anaerobic ammonium oxidation bacteria (anammox) which can convert the fomite in the water into nitrogen. Our goal is to design a wastewater treatment system which can absorb the pollutant efficiently while transform it into luminous energy. We plan to use E.coli to design a bacterium that can digest the nitrite and ammonium in its interior using the disposal system from the anammox. Through the introduction of luciferase, the energy can be transformed into bioluminescence. Therefore, we named it Shining Sanctifier. This new star in synthetic biology will be applied to the sewage treatment system on a large scale while it can also be made into illuminating system.

Team AITM-Nepal: siRNA MEDIATED IMMUNE MODULATION FOR INNATE AND ADAPTIVE RESPONSE USING GENETICALLY ENGINEERED Escherichia coli

Canonical small interfering RNA (siRNA) duplexes are potent activators of the mammalian innate immune system. The induction of innate immunity by siRNA is dependent on siRNA structure and sequence, method of delivery, and cell type. The delivery of siRNA in a packaged outer membrane vesicle of gram negative bacteria is the theme of our work. The toll like receptor-7/8 activation by siRNA in order to boost the production of Interferon type -1 molecules to inhibit the viral and outer membrane LPS structure to activate Toll like receptor -4 to inhibit bacterial pathogens is the objective of this work. The delivery is made dependent on the peptide fragment which mediated the fusogenic mechanism so as to escape the endosomal compartment once endocytosed inside host(mamalian) cell. Thus freeing the siRNA to silence the myD88 transcriptin host cytoplasm making RISC complex and hence, activating TLR-7/8 in endosomal membrane formerly.

Team BIT: A New Strategy to Detect Antibiotics in Milk: Based on Sensors with Controllable Bio-enhanced Blocks

Bio-amplification, especially controllable bio-amplification is significant for biological detection. In a synthetic biological way, 2013 BIT iGEM assembled the T7 RNA polymerase gene and T7 promoter as an

3-D Depth Reconstruction from a Single Still Image

Ashutosh Saxena, Sung H. Chung, Andrew Y. Ng

Computer Science Department
Stanford University, Stanford, CA 94305
{asaxena,codedeft,ang}@cs.stanford.edu

Abstract

We consider the task of 3-d depth estimation from a single still image. We take a supervised learning approach to this problem, in which we begin by collecting a training set of monocular images (of unstructured indoor and outdoor environments which include forests, sidewalks, trees, buildings, etc.) and their corresponding ground-truth depthmaps. Then, we apply supervised learning to predict the value of the depthmap as a function of the image. Depth estimation is a challenging problem, since local features alone are insufficient to estimate depth at a point, and one needs to consider the global context of the image. Our model uses a hierarchical, multiscale Markov Random Field (MRF) that incorporates multiscale local- and global-image features, and models the depths and the relation between depths at different points in the image. We show that, even on unstructured scenes, our algorithm is frequently able to recover fairly accurate depthmaps. We further propose a model that incorporates both monocular cues and stereo (triangulation) cues, to obtain significantly more accurate depth estimates than is possible using either monocular or stereo cues alone.

1 Introduction

Recovering 3-d depth from images is a basic problem in computer vision, and has important applications in robotics, scene understanding and 3-d reconstruction. Most work on visual 3-d reconstruction has focused on binocular vision (stereopsis) [1] and on other algorithms that require multiple images, such as structure from motion [2] and depth from defocus [3]. These algorithms consider only the geometric (triangulation) differences. Beyond stereo/triangulation cues, there are also numerous *monocular* cues—such as texture variations and gradients, defocus, color/haze, etc.—that contain useful and important depth information. Even though humans perceive depth by seamlessly combining many of these

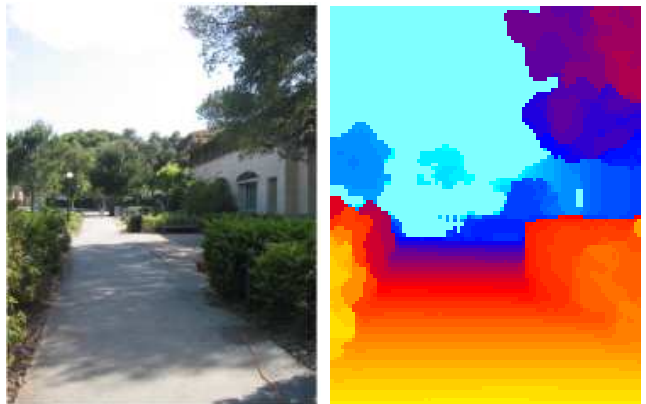


Figure 1: (a) A single still image, and (b) the corresponding (ground-truth) depthmap. Colors in the depthmap indicate estimated distances from the camera.

stereo and monocular cues, most work on depth estimation has focused on stereovision.

Depth estimation from a *single* still image is a difficult task, since depth typically remains ambiguous given only local image features. Thus, our algorithms must take into account the global structure of the image, as well as use prior knowledge about the scene. We also view depth estimation as a small but crucial step towards the larger goal of image understanding, in that it will help in tasks such as understanding the spatial layout of a scene, finding walkable areas in a scene, detecting objects, etc. In this paper, we apply supervised learning to the problem of estimating depthmaps (Fig. 1b) from a single still image (Fig. 1a) of a variety of unstructured environments, both indoor and outdoor, containing forests, sidewalks, buildings, people, bushes, etc.

Our approach is based on modeling depths and relationships between depths at multiple spatial scales using a hierarchical, multiscale Markov Random Field (MRF). Taking a supervised learning approach to the problem of depth estimation, we used a 3-d scanner to collect training data, which comprised a large set of images and their corresponding ground-truth depthmaps. (This data has been made publically available on the internet.) Using this training set, we model the conditional distribution of

THE DREAM OF ELIXIR VITAE

João Pedro de Magalhães, Ph.D.

Human aging is a universal process of loss of viability and increase in vulnerability. Although so far the underlying mechanisms of aging remain largely a mystery, it is reasonable to expect that we will eventually understand the human aging process. Possibly during this century, we will know what changes occur in a human being from ages 30 to 70 to increase the chance of dying by roughly 32-fold. Yet even if researchers detail those changes, even if researchers identify the causal molecular and cellular mechanisms responsible for human aging, this will not necessarily lead to a cure for aging. HIV was identified as the cause of AIDS over 20 years ago and we still cannot cure AIDS. [1] So to delay aging, not to mention to stop or reverse the human aging process, will be a monumental task. It is true that we still do not know in detail what changes occur as human beings age, but it is equally true that the most important question in studying aging is not why we age but how can we fix it.

Intranasal delivery of cells to the brain

Lusine Danielyan^{a,*}, Richard Schäfer^b, Andreas von Ameln-Mayerhofer^c,
Marine Buadze^a, Julia Geisler^a, Tim Klopfer^a, Ute Burkhardt^a,
Barbara Proksch^a, Stephan Verleysdonk^d, Miriam Ayturan^b,
Gayane H. Buniatian^a, Christoph H. Gleiter^a, William H. Frey II^e

^aDepartment of Clinical Pharmacology, University Hospital of Tuebingen, Otfried-Mueller Str. 45, D-72076 Tuebingen, Germany

^bInstitute of Clinical and Experimental Transfusion Medicine, University Hospital of Tuebingen, Tuebingen, Germany

^cNeuropharmacology, University of Tuebingen, Tuebingen, Germany

^dInterfaculty Institute for Biochemistry, University of Tuebingen, Tuebingen, Germany

^eAlzheimer's Research Center at Regions Hospital, HealthPartners Research Foundation, St. Paul, MN 55101, USA

Received 20 January 2009; accepted 11 February 2009

Abstract

The safety and efficacy of cell-based therapies for neurodegenerative diseases depends on the mode of cell administration. We hypothesized that intranasally administered cells could bypass the blood-brain barrier by migrating from the nasal mucosa through the cribriform plate along the olfactory neural pathway into the brain and cerebrospinal fluid (CSF). This would minimize or eliminate the distribution of cellular grafts to peripheral organs and will help to dispense with neurosurgical cell implantation. Here we demonstrate transnasal delivery of cells to the brain following intranasal application of fluorescently labeled rat mesenchymal stem cells (MSC) or human glioma cells to naive mice and rats. After cells crossed the cribriform plate, two migration routes were identified: (1) migration into the olfactory bulb and to other parts of the brain; (2) entry into the CSF with movement along the surface of the cortex followed by entrance into the brain parenchyma. The delivery of cells was enhanced by hyaluronidase treatment applied intranasally 30 min prior to the application of cells. Intranasal delivery provides a new non-invasive method for cell delivery to the CNS.

© 2009 Elsevier GmbH. All rights reserved.

Keywords: Intranasal delivery; Mesenchymal stem cells; Glioma cells; Hyaluronidase; Olfactory route; Trigeminal route; CD45; Stem cell transplantation

Introduction

The success of cell-based therapy for neurodegenerative disorders depends on the therapeutic properties of

the cell type, on the method and safety of administration, on the amount of cells delivered to the site of injury and finally on the avoidance of excessive incorporation of the therapeutic cells into other organs and systems. Transplantation of therapeutic cells into the brain is one of the most frequently used invasive approaches to promote recovery of the CNS from injury or disease. However, methodologically, transplantation may raise problems not only because of graft rejection as a result

*Corresponding author. Tel.: +49 7071 297 4926;
fax: +49 7071 295 035.

E-mail address: lusine.danielyan@med.uni-tuebingen.de
(L. Danielyan).

Hardware Architectures for Deep Neural Networks

ISCA Tutorial

June 24, 2017

Website: <http://eyeriss.mit.edu/tutorial.html>



Massachusetts
Institute of
Technology



nVIDIA®

Is mental effort exertion contagious?

Kobe Desender¹, Sarah Beurms², & Eva Van den Bussche¹

1. Department of Psychology, Vrije Universiteit Brussel, Brussels, Belgium

2. Department of Psychology, KU Leuven, Belgium

Corresponding author:

Kobe Desender

Faculty of Psychology and Educational Sciences

Vrije Universiteit Brussel

Pleinlaan 2, 1050 Brussels

Belgium

Phone: +32-(0)2 629 14 67

Fax: +32-(0)2-629 24 89

E-mail: Kobe.Desender@vub.ac.be

Keywords: effort exertion, cognitive control, contagion, social facilitation, joint-Simon

4874 words below (Abstract, References excluded)

SCIENTIFIC REPORTS

OPEN

IVA cloning: A single-tube universal cloning system exploiting bacterial *In Vivo* Assembly

Javier García-Nafría*, Jake F. Watson* & Ingo H. Greger

Received: 22 March 2016

Accepted: 13 May 2016

Published: 06 June 2016

In vivo homologous recombination holds the potential for optimal molecular cloning, however, current strategies require specialised bacterial strains or laborious protocols. Here, we exploit a *recA*-independent recombination pathway, present in widespread laboratory *E. coli* strains, to develop IVA (*In Vivo Assembly*) cloning. This system eliminates the need for enzymatic assembly and reduces all molecular cloning procedures to a single-tube, single-step PCR, performed in <2 hours from setup to transformation. Unlike other methods, IVA is a complete system, and offers significant advantages over alternative methods for all cloning procedures (insertions, deletions, site-directed mutagenesis and sub-cloning). Significantly, IVA allows unprecedented simplification of complex cloning procedures: five simultaneous modifications of any kind, multi-fragment assembly and library construction are performed in approximately half the time of current protocols, still in a single-step fashion. This system is efficient, seamless and sequence-independent, and requires no special kits, enzymes or proprietary bacteria, which will allow its immediate adoption by the academic and industrial molecular biology community.

Molecular cloning is at the heart of biomedical and biotechnological research, fundamental to protein structure-function studies, protein engineering and synthetic biology^{1–3}. Since the advent of the polymerase-chain reaction (PCR)⁴, cloning has involved selective PCR amplification and modification of DNA segments, which require directed assembly into a plasmid carrier for propagation in *E. coli*. Traditionally, restriction enzymes and ligases have been used to direct the assembly of DNA fragments¹; however, their sequence-dependency and laborious protocols led to the development of new alternatives, such as: PCR-only^{5–8}, ligation-independent cloning (LIC)⁹, recombination-based^{10,11} and multi-enzyme construction methods (such as Gibson¹², In-Fusion¹³ and USER¹⁴). LIC, multi-enzyme construction and some recombination-based approaches (such as SLiCE¹⁰) rely on *in vitro* enzymatic treatment of DNA fragments for assembly. While PCR-only methods eliminate the need for such enzymatic treatment, they involve multiple rounds of PCR and DNA purification. Both cases involve lengthy, hands-on protocols. *In vivo* assembly, where the bacterial host performs the fusion of DNA fragments, would eliminate the need for multiple steps and reagents, providing significant advantages over all current methods. The advantages of *in vivo* recombination have led molecular biologists to use the yeast gap-repair cloning system¹⁵, despite the hurdles associated with eukaryotic work. *In vivo* assembly in *E. coli* has previously been limited to the use of strains with enhanced recombinase activities, such as *recA*¹⁶ and phage recombinases Red/ET^{17,18}, yet these bacterial strains enhance plasmid instability or require specialised preparation of competent cells. As a consequence, strains with reduced recombinase activities (e.g. *recA* knockouts) are universally used for molecular cloning.

The presence of a *recA*-independent homologous recombination pathway in *E. coli* was reported more than 20 years ago^{19–21}, but has been neglected until recently, except for sporadic use in specific high throughput applications^{22,23}. The pathway is mostly uncharacterised but is most efficient at recombining linear DNA fragments, likely acting through an annealing mechanism^{20,24}, although alternative mechanisms have been suggested^{25,26}. Conveniently, the *recA*-independent pathway is responsible for the recombination of short overlapping sequences¹⁹, whereas the *recA* system requires longer homologous DNA stretches (>150–300 bp). The pathway's short homology requirements, ubiquitous presence in laboratory *E. coli* strains (in our and others' experience^{27–29}) and reduced compromises on plasmid stability, make it an optimal tool for molecular cloning.

Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.G.-N. (email: jgarcia@mrc-lmb.cam.ac.uk)

No jacket required – new fungal lineage defies dress code

Recently described zoosporic fungi lack a cell wall during trophic phase

Timothy Y. James^{1)*} and Mary L. Berbee^{2)*}

Analyses of environmental DNAs have provided tantalizing evidence for “rozellida” or “cryptomycota”, a clade of mostly undescribed and deeply diverging aquatic fungi. Here, we put cryptomycota into perspective through consideration of *Rozella*, the only clade member growing in culture. This is timely on account of the publication in *Nature* of the first images of uncultured cryptomycota from environmental filtrates, where molecular probes revealed non-motile cyst-like structures and motile spores, all lacking typical fungal chitinous cell walls. Current studies of *Rozella* can complement these fragmentary observations from environmental samples. *Rozella* has a fungal-specific chitin synthase and its resting sporangia have walls that appear to contain chitin. Cryptomycota, including *Rozella*, lack a cell wall when absorbing food but like some other fungi, they may have lost their “dinner jacket” through convergence. Rather than evolutionary intermediates, the cryptomycota may be strange, divergent fungi that evolved from an ancestor with a nearly complete suite of classical fungal-specific characters.

Keywords:

■ chitin synthase; cryptomycota; evolution; *Rozella*; rozellida

Introduction

Fungi lead hidden lives

Fungi lead cryptic lives by growing inside their food source, and if they emerge it is only to reproduce as mushrooms, cups, or other spore-producing structures. Although the reproductive structures

form the traditional basis for detection and classification, they appear only briefly in the life of a fungus. Add to this that many groups of fungi are difficult or impossible to obtain in pure culture [1, 2], and the result is that by most estimates less than 10% of all fungi have been observed and formally described (100 thousand out of an estimated 1.5 million

species). Ranging from unicellular organisms to some of the largest and most long-lived of all organisms [3, 4], the remaining fungi are hiding all around us and modern approaches to studying diversity and communities are beginning to reveal the true phylogenetic diversity of the group. After a full decade of progress in understanding fungal diversity using environmental DNA community studies, we now realize that most of the fungi in the environment do not actually match those specimens from herbarium cabinets and culture collections that were used to build the fungal tree of life [2, 5].

Most fungal sequences from environmental DNA studies can be assigned to a described class or even genus [6–8], but some represent unknown taxa on deeply diverging branches [9–12]. When lineages known exclusively from environmental DNA sequences cannot be assigned to a phylum, they challenge our understanding of the biodiversity and phylogeny (breadth and depth), and even characteristics of fungi. A recent breakthrough by Jones et al. [13] on the diversity and characteristics of one such enigmatic lineage, named “cryptomycota”, raised the possibility that the lines dividing fungus from the protozoan soup from which they evolved [14] may be fuzzier than appreciated. Specifically, Jones et al. concluded that the widespread group cryptomycota were intermediate between fungi and ancestral protists. In their words:

“Co-staining with cell wall markers demonstrates that representatives from the

DOI 10.1002/bies.201100110

¹⁾ Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

²⁾ Department of Botany, University of British Columbia, Vancouver, BC, Canada

*Corresponding authors:

Timothy Y. James
E-mail: tyjames@umich.edu

Mary L. Berbee
E-mail: berbee@interchange.ubc.ca

Symbolic Modular Deadlock Analysis

Jyotirmoy V. Deshmukh · E. Allen Emerson ·
Sriram Sankaranarayanan

Received: Jun 14 2010 / Accepted:

Abstract Methods in object-oriented concurrent libraries often encapsulate internal synchronization details. As a result of information hiding, clients calling the library methods may cause thread safety violations by invoking methods in an unsafe manner. This is frequently a cause of deadlocks. Given a concurrent library, we present a technique for inferring *interface contracts* that specify permissible concurrent method calls and patterns of aliasing among method arguments. In this work, we focus on deriving contracts that guarantee deadlock-free execution for the methods in the library. The contracts also help client developers by documenting required assumptions about the library methods. Alternatively, the contracts can be statically enforced in the client code to detect potential deadlocks in the client. Our technique combines static analysis with a symbolic encoding scheme for tracking lock dependencies, allowing us to synthesize contracts using an SMT solver. Additionally, we investigate extensions of our technique to reason about deadlocks in libraries that employ signaling primitives such as *wait-notify* for cooperative synchronization. Our prototype tool analyzes over a million lines of code for some widely-used Java libraries within an hour, thus demonstrating its scalability and efficiency. Furthermore, the contracts inferred by our approach have been able to pinpoint real deadlocks in clients, *i.e.* deadlocks that have been a part of bug-reports filed by users and developers of client code.

Keywords Deadlock Prediction · Static Analysis · Concurrent Libraries

Part of this paper was published in the proceedings of the 24th IEEE/ACM International Conference on Automated Software Engineering, 2009, [10].

J. V. Deshmukh (✉)
Department of Electrical Engineering, The University of Texas at Austin, Austin, TX 78712
E-mail: jyotirmoy@cerc.utexas.edu

E. A. Emerson
Department of Computer Science, The University of Texas at Austin, Austin, TX 78712
E-mail: emerson@cs.utexas.edu

S. Sankaranarayanan
Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309-0430
E-mail: srirams@colorado.edu

Nano-Patterning of Diffraction Gratings on Human Hair for Cosmetic Purposes

Khawar Abbas¹, Drew F. Goettler¹, Bruce C. Lamartine², Zayd C. Leleman¹

¹Mechanical Engineering Department, University of New Mexico, Albuquerque, USA

²Los Alamos National Laboratory, Los Alamos, USA

Email: zleleman@unm.edu

Received 28 March 2014; revised 26 April 2014; accepted 2 May 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A method is presented for nano-patterning a diffraction grating on human hair with a focused ion beam. Strands of brown hair are patterned with hyperbolas and Archimedean spirals whose pitches range from 540 nm to 1040 nm. Exposure of the hair strands to white light at various incident angles demonstrates that light of varying wavelengths is diffracted by the diffraction gratings. The diffraction causes the brown strands of hair to reflect light from the entire range of visible light.

Keywords

Diffraction, Hair, Nanopatterning, Focused Ion Beam Milling

1. Introduction

The practice of hair dyeing for cosmetic purposes has been around for several hundred years. Temporary hair dyes are applied on hair to change its appearance hiding graying hair or to stay in line with the latest fashion trends. Greeks and Romans used naturally occurring henna as a temporary hair dye while Arab authors described application of the paste formed by the mixture of PbO and slaked lime ($\text{Ca}(\text{OH})_2$) in water on hair [1]. Presently there are numerous commercially available products that are used in highlighting hair strands for cosmetic purposes to brighten and/or create unnatural hair colors for Halloween, football games, etc.

Adverse effects of these synthetic hair dyes often range from skin discoloration and allergic irritations [2] to toxicity and cancer [3] [4]. In this paper we present a technique wherein nano-scale diffraction gratings are patterned on hair and could be used as an alternative to common hair dyeing treatments. Diffraction gratings do occur in nature for color generation [5] [6] and similar techniques are used in optics for color separation. We mimic and adapt nature's color generation/separation technique to human hair as an alternative to chemical dyes. A Focused Ion Beam (FIB) was used to mill nano-scale diffraction gratings on individual strands of brown hu-

Join Me on a Market for Anonymity

Malte Möser¹ and Rainer Böhme²

¹ Department of Information Systems, University of Münster, Germany

² Department of Computer Science, University of Innsbruck, Austria

Abstract. We present the first measurement study of JoinMarket, a growing marketplace for more anonymous transfers in the Bitcoin ecosystem. Our study reveals that this market is funded with multiple thousand bitcoins and generated a turnover of almost 8 million USD over the course of eight months. Assessing the resilience of the market against a well-funded attacker, we discover that in a typical scenario, a selective attack with 90% success rate requires an investment of 32,000 USD (which is recoverable after the attack). We formulate stylized economic models of supply and demand to explain the existence of this novel market for anonymity and underpin some theoretical arguments with empirical data.

1 Introduction

Anonymity and economics are an odd couple. Most microeconomic models assume agents without name, and fail to predict outcomes if agents become identifiable [3, 39]. Anonymity can be defined as the state of being “not identifiable within a set of subjects, the anonymity set” [33]. With this definition, a simple measure of the quality of anonymity is the size of this set, as the probability of successful identification by random guessing is inversely proportional to the set size.

If there exist situations where the state of being anonymous improves an agent’s wealth, one would expect a market for anonymity to develop. Yet anonymity is an unconventional economic good. To produce it, other agents must behave in an indistinguishable way to an observer of the agent who seeks anonymity. This is nicely summarized in the expression “anonymity loves company” [17]. In the language of economics, the production of anonymity generates positive externalities because all agents who contribute to the supply of anonymity also receive the good in demand. Production and consumption are hard to tell apart. So, what should the market price for anonymity be?

While the economics of privacy [34] and personal data [22] have been studied for long, and empirical research has estimated price information (see [23] and [1] for reviews), remarkably little is known about the price of anonymity. Acquisti, Dingledine and Syverson [2] study the incentives to participate in anonymous communication systems based on mix networks, such as Tor. Their analysis is comprehensive, includes attacker behavior and adoption dynamics, but remains theoretical. Spiekermann [38] interprets survey data collected by self-selection among the early-adopters of an academic anonymous communication system (JAP). Köpsell [21] uses technical measurements in an experimental setup of

Paving the way for Single-molecule Sequencing of non-natural information-containing Polymers

N. F. KÖNIG¹, L. CHARLES² and J.-F. LUTZ¹

¹ Precision Macromolecular Chemistry, Institut Charles Sadron, CNRS UPR-22, 23 rue du Loess, 67034 Strasbourg Cedex 2, France

² Aix-Marseille Université – CNRS, UMR 7273, Institute of Radical Chemistry, 13397 Marseille Cedex 20, France

ABSTRACT: Storing information on the molecular level has become a flourishing discipline in recent years. Attaining an immense information density, and energy efficient storage compared to conventional storage devices, stimulates research in the field.

Artificial biopolymers like DNA set, until now, the yardstick for molecular storage systems due to already established coding and decoding methodologies.^{1,2} However, restriction to biological building blocks is clearly setting limits to future (commercial) applicability. There is a need for diversely tunable systems. Methodologies for obtaining sequence-defined non-natural polymers have recently been developed.^{3,4,5} Sequencing can be achieved by tandem mass spectrometry, yet crucial for efficient storage systems, the next objective is the development of a new single-molecule sequencing technique for non-natural polymers.

The new approach employs polyphosphodiester that comprise two different monomer units defining a 0 and a 1 bit, respectively. Polymers of that kind incorporate information as a binary code. The copolymers are accessible through the fast and convenient phosphoramidite method. To respond with information readout, every sequencing technique demands specified features within the polymer. Libraries of fine-tuned polymers are thus demanded. The approach to cope with these requirements is presented to pave the way for single-molecule sequencing of non-natural information-containing polymers.

KEY WORDS: sequence-controlled polymers, polyphosphodiester, phosphoramidite strategy, single-molecule sequencing

References

1. Zhirnov, V., Zadegan, R. M., Sandhu, G. S., Church, G. M., Hughes, W. L., *Nature Mater.* **2016**, *15*, 366.
2. Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., Birney, E., *Nature* **2013**, *494*, 77.
3. Lutz, J.-F., *Macromolecules* **2015**, *48*, 4759.
4. Al Ouahabi, A., Charles, L., Lutz, J.-F., *J. Am. Chem. Soc.* **2015**, *137*, 5629.
5. Al Ouahabi, A., Kotera, M., Charles, L., Lutz, J.-F., *ACS Macro Lett.* **2015**, *4*, 1077.

Large-scale *de novo* DNA synthesis: technologies and applications

Sriram Kosuri¹ & George M Church^{2,3}

For over 60 years, the synthetic production of new DNA sequences has helped researchers understand and engineer biology. Here we summarize methods and caveats for the *de novo* synthesis of DNA, with particular emphasis on recent technologies that allow for large-scale and low-cost production. In addition, we discuss emerging applications enabled by large-scale *de novo* DNA constructs, as well as the challenges and opportunities that lie ahead.

DNA is the predominant information carrier for life. The development of synthetic techniques to construct DNA has led to marked improvements in our ability to understand and engineer biology. For example, despite extensive efforts to unravel the genetic code using molecular genetics, modest capabilities to synthesize nucleic acids ultimately led to the code's unraveling¹. Today, reconstructions of complete viral and bacterial genomes are testaments of how far our synthetic capabilities have come.

Despite the improvements, our ability to read DNA is better than our ability to write it. Over the last decade, high-throughput sequencing technologies, here referred to as next-generation sequencing (NGS), have revolutionized the discovery and understanding of natural DNA sequence, with current installed capacity estimated at ~15 petabases per year². Large-scale data-sharing initiatives, such as GenBank, and continued improvements in bioinformatics software have made computational analyses on these data easier than ever. Such analyses help generate powerful statistical hypotheses for how genome sequence controls cellular functions across organisms and populations. In addition, NGS-based measurement tools allow for the analysis of many genetic and biochemical processes at unprecedented scale and low cost³. However, even though our ability to both generate hypotheses and measure outcomes has increased in scale owing to NGS, our ability to test such hypotheses experimentally still lags and is among the most limiting steps in the study of natural and engineered biology.

Specifically, a designed DNA construct is a physical instance of a hypothesis to be tested, whether it be a simple plasmid-based reporter or a whole-genome synthesis of an organism. Progress in large-scale, low-cost construction of desired DNA sequences could rapidly engender progress in both fundamental and applied biological research.

Although rapid modification of natural DNA sequence both *in vitro* and *in vivo* is useful for a variety of purposes, methodologies for *de novo* synthesis of DNA from nucleosides confer a number of unique advantages. First, engineering new functions often requires vastly modified or wholly new genetic sequences that are most easily accessed by *de novo* synthesis methodologies. Second, synthesized constructs are often superior to natural sequence for the study of genetic mechanisms because they can be designed to specifically test hypotheses for how sequence affects function. Finally, sequences that are targeted to be amplified or modified from natural sequences can be difficult to access (for example, from metagenomic data sets); thus, synthesis is the only practical way to experimentally study them.

Here we review technological innovations and applications for *de novo* DNA synthesis as distinct from assembly and modification of natural DNA sequence. We cover large-scale single-stranded DNA oligonucleotide (oligo) synthesis, assembly of these oligos into longer double-stranded DNA constructs, and emerging applications (Fig. 1).

¹Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California, USA. ²Wyss Institute for Biologically Inspired Engineering, Boston, Massachusetts, USA. ³Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to S.K. (sri@ucla.edu).

Optimal Brain Damage

Yann Le Cun, John S. Denker and Sara A. Solla
AT&T Bell Laboratories, Holmdel, N. J. 07733

ABSTRACT

We have used information-theoretic ideas to derive a class of practical and nearly optimal schemes for adapting the size of a neural network. By removing unimportant weights from a network, several improvements can be expected: better generalization, fewer training examples required, and improved speed of learning and/or classification. The basic idea is to use second-derivative information to make a tradeoff between network complexity and training set error. Experiments confirm the usefulness of the methods on a real-world application.

1 INTRODUCTION

Most successful applications of neural network learning to real-world problems have been achieved using highly structured networks of rather large size [for example (Waibel, 1989; LeCun et al., 1990)]. As applications become more complex, the networks will presumably become even larger and more structured. Design tools and techniques for comparing different architectures and minimizing the network size will be needed. More importantly, as the number of parameters in the systems increases, overfitting problems may arise, with devastating effects on the generalization performance. We introduce a new technique called Optimal Brain Damage (OBD) for reducing the size of a learning network by selectively deleting weights. We show that OBD can be used both as an automatic network minimization procedure and as an interactive tool to suggest better architectures.

The basic idea of OBD is that it is possible to take a perfectly reasonable network, delete half (or more) of the weights and wind up with a network that works just as well, or better. It can be applied in situations where a complicated problem must

The Saddest Moment

JAMES MICKENS



James Mickens is a researcher in the Distributed Systems group at Microsoft's Redmond lab. His current research focuses on Web applications, with an emphasis on the

design of JavaScript frameworks that allow developers to diagnose and fix bugs in widely deployed web applications. James also works on fast, scalable storage systems for datacenters. James received his PhD in computer science from the University of Michigan, and a bachelor's degree in computer science from Georgia Tech.

mickens@microsoft.com

Whenever I go to a conference and I discover that there will be a presentation about Byzantine fault tolerance, I always feel an immediate, unshakable sense of sadness, kind of like when you realize that bad things can happen to good people, or that Keanu Reeves will almost certainly make more money than you over arbitrary time scales. Watching a presentation on Byzantine fault tolerance is similar to watching a foreign film from a depressing nation that used to be controlled by the Soviets—the only difference is that computers and networks are constantly failing instead of young Kapruskin being unable to reunite with the girl he fell in love with while he was working in a coal mine beneath an orphanage that was atop a prison that was inside the abstract concept of World War II. “How can you make a reliable computer service?” the presenter will ask in an innocent voice before continuing, “It may be difficult if you can’t trust anything and the entire concept of happiness is a lie designed by unseen overlords of endless deceptive power.” The presenter never explicitly says that last part, but everybody understands what’s happening. Making distributed systems reliable is inherently impossible; we cling to Byzantine fault tolerance like Charlton Heston clings to his guns, hoping that a series of complex software protocols will somehow protect us from the oncoming storm of furious apes who have somehow learned how to wear pants and maliciously tamper with our network packets.

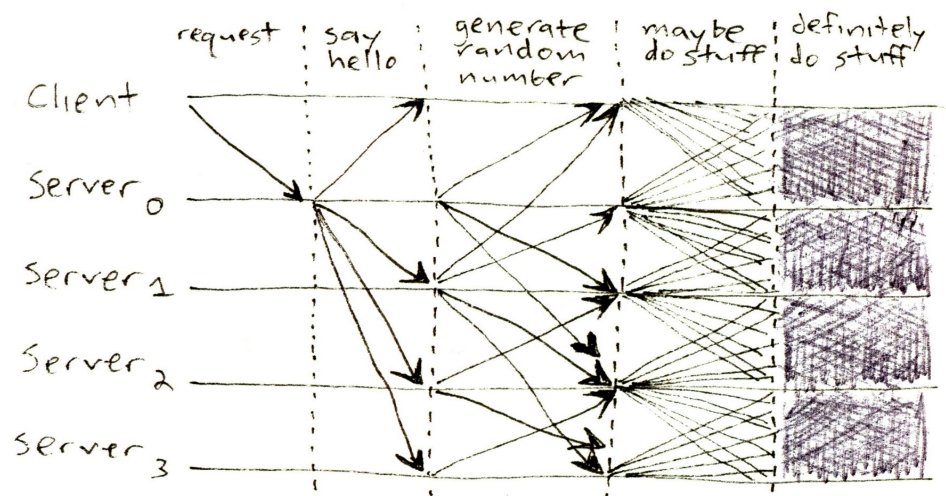


Figure 1: Typical Figure 2 from Byzantine fault paper: Our network protocol

Low Cost Design for an Orbital Ring

By David Nelson

18/7/2017

Executive Summary

This report outlines a design for an orbital ring with a cost estimate of \$9 billion that can be built using existing technology and is well within the resource capacity of over 1200 of the world's largest companies. If this design is verified it will launch a space race to construct the first orbital ring, which could be constructed within a 10 to 30 year time frame. The successful deployment of this orbital ring will reduce the cost of space travel by at least three orders of magnitude to just dollars per kilogram instead of the current \$1400/kg.

The main elements of the design include hundreds of metal bolts (40kg each) which travel between ten platforms (4500kg each) that redirect the bolts using superconducting electromagnets and a Tesla coil which is powered by solar panels and later by electricity from the tether. A Zylon tether that is 450km long and weighing 9040kg is dropped from the platforms once the ring has been stabilized. The platforms have direct line of sight between each other enabling the bolt to reach velocities of 30km/s. The entire mass of the ring will be launched into orbit by a single Falcon Heavy rocket lifting a payload of 63.8 tonnes into low Earth orbit. Everything is situated in low Earth orbit which will minimize the accumulated space junk that would result from a complete failure.

Table of Contents

Low Cost Design for an Orbital Ring	1
Executive Summary	1
Introduction	2
Deployment	4
Platform	6
Tether	10
Cost	11
Conclusion	12
Bibliography	13
Appendix	14
Calculation 1: Determining the number of platforms required to obtain line of sight (LOS)	14
Calculation 2: Calculating the angle of bolt deflection (α) for a bolt trajectory without line of sight .	14
Calculation 3: Calculating the momentum transferred to the platform	16
Calculation 4: Determining platform freefall distance between bolts	16

Considerations in planning vegan diets: Infants

ANN REED MANGELS, PhD, RD, FADA; VIRGINIA MESSINA, MPH, RD

ABSTRACT

Appropriately planned vegan diets can satisfy nutrient needs of infants. The American Dietetic Association and The American Academy of Pediatrics state that vegan diets can promote normal infant growth. It is important for parents to provide appropriate foods for vegan infants, using guidelines like those in this article. Key considerations when working with vegan families include composition of breast milk from vegan women, appropriate breast milk substitutes, supplements, type and amount of dietary fat, and solid food introduction. Growth of vegan infants appears adequate with post-weaning growth related to dietary adequacy. Breast milk composition is similar to that of non-vegetarians except for fat composition. For the first 4 to 6 months, breast milk should be the sole food with soy-based infant formula as an alternative. Commercial soymilk should not be the primary beverage until after age 1 year. Breastfed vegan infants may need supplements of vitamin B-12 if maternal diet is inadequate; older infants may need zinc supplements and reliable sources of iron and vitamins D and B-12. Timing of solid food introduction is similar to that recommended for non-vegetarians. Tofu, dried beans, and meat analogs are introduced as protein sources around 7-8 months. Vegan diets can be planned to be nutritionally adequate and support growth for infants. *J Am Diet Assoc.* 2001;101:670-677.

The period from birth to 1 year is a time of nutritional vulnerability when attention to proper nutrition is critical to support the extremely rapid growth, including brain growth, seen during this period. Throughout this first year, breast milk or infant formula provides a large portion of the energy and protein needed by both vegan and nonvegan infants. The first solid foods that are offered to most infants are infant cereals, fruits, and vegetables with meats not introduced until later. Both the American Dietetic Association and the American Academy of Pediatrics assert that well-planned vegan diets can satisfy nutrient needs of infants and promote normal growth (1, 2). It is important for parents to be aware of and provide appropriate foods for vegan infants, using guidelines such as those found in this article.

The purpose of this paper is to identify important issues in the feeding of vegan infants and to provide recommendations which will help dietetics professionals work with the families of vegan infants to plan diets which meet needs for growth and development, are age-appropriate, practical, and in keeping with the family's beliefs.

GROWTH OF VEGAN INFANTS

A limited number of studies have examined the birth weights of infants of vegan mothers. A study of close to 400 infants and children, 75% of whose mothers used vegan diets throughout pregnancy, found birth weights and incidence of low-birth-weight infants to be similar to those of well-educated US white women (3). The reported birth weights of 19 term infants born to vegan women were slightly lower than infants with non-vegetarian mothers (4). Lower birth weights of infants of Dutch women following macrobiotic diets, which exclude most animal products as well as a number of other foods, have been

A. R. Mangels is a nutrition advisor for the Vegetarian Resource Group, Baltimore, Md. V. Messina is a nutrition consultant with Nutrition Matters, Inc, Port Townsend, Wash.

Address correspondence to: Virginia Messina, Nutrition Matters, Inc, 1543 Lincoln St, Port Townsend, WA 98368.

molecular ticker-tapes

Adam Marblestone

on behalf of the Church, Kording, Boyden and Tyo labs

[original slides redacted to remove unpublished info]

3.

The microarchitecture of Intel, AMD and VIA CPUs

An optimization guide for assembly programmers and compiler makers

By Agner Fog. Technical University of Denmark.
Copyright © 1996 - 2017. Last updated 2017-05-02.

Contents

1	Introduction	6
1.1	About this manual	6
1.2	Microprocessor versions covered by this manual	7
2	Out-of-order execution (All processors except P1, PMMX)	9
2.1	Instructions are split into μ ops	9
2.2	Register renaming	10
3	Branch prediction (all processors)	12
3.1	Prediction methods for conditional jumps	12
3.2	Branch prediction in P1	18
3.3	Branch prediction in PMMX, PPro, P2, and P3	21
3.4	Branch prediction in P4 and P4E	23
3.5	Branch prediction in PM and Core2	25
3.6	Branch prediction in Intel Nehalem	27
3.7	Branch prediction in Intel Sandy Bridge and Ivy Bridge	28
3.8	Branch prediction in Intel Haswell, Broadwell and Skylake	29
3.9	Branch prediction in Intel Atom, Silvermont and Knights Landing	29
3.10	Branch prediction in VIA Nano	30
3.11	Branch prediction in AMD K8 and K10	31
3.12	Branch prediction in AMD Bulldozer, Piledriver and Steamroller	33
3.1	Branch prediction in AMD Ryzen	34
3.2	Branch prediction in AMD Bobcat and Jaguar	34
3.3	Indirect jumps on older processors	35
3.4	Returns (all processors except P1)	36
3.5	Static prediction	36
3.6	Close jumps	37
4	Pentium 1 and Pentium MMX pipeline	39
4.1	Pairing integer instructions	39
4.2	Address generation interlock	43
4.3	Splitting complex instructions into simpler ones	43
4.4	Prefixes	44
4.5	Scheduling floating point code	45
5	Pentium 4 (NetBurst) pipeline	48
5.1	Data cache	48
5.2	Trace cache	48
5.3	Instruction decoding	53
5.4	Execution units	54
5.5	Do the floating point and MMX units run at half speed?	57
5.6	Transfer of data between execution units	59
5.7	Retirement	62
5.8	Partial registers and partial flags	62
5.9	Store forwarding stalls	63
5.10	Memory intermediates in dependency chains	63
5.11	Breaking dependency chains	65
5.12	Choosing the optimal instructions	65

The Modern Myth of “Unculturable” Bacteria/ Scotoma of contemporary microbiology

Dedicated to the pioneering microbiologists who isolated pure cultures of microbes responsible for (a) infectious diseases of animals and plants, and (b) the cyclic transformations of major chemical elements on Earth. Their characterization of the biological, physiological, and genetic properties of these organisms paved the way for current research. The careers and contributions of more than 300 of the early pioneers are profiled in the classic book by William Bulloch: The History of Bacteriology (Oxford University Press, 1938).

Howard Gest

**Distinguished Professor Emeritus of Microbiology
Adjunct Professor, History and Philosophy of Science,
Indiana University, Bloomington, IN 47405
[gest@indiana.edu]**

The activities of bacteria in recycling of elements on Earth and their effects on animals and plants were unknown before techniques for isolation of pure cultures were developed. Through sustained efforts of microbiologists over many decades collections of pure cultures were established and these provided experimental systems that led to our present encyclopedic knowledge of microbiology. The isolation of pure cultures required development of appropriate growth media and this aspect of microbiological research proved to be very difficult in many instances.

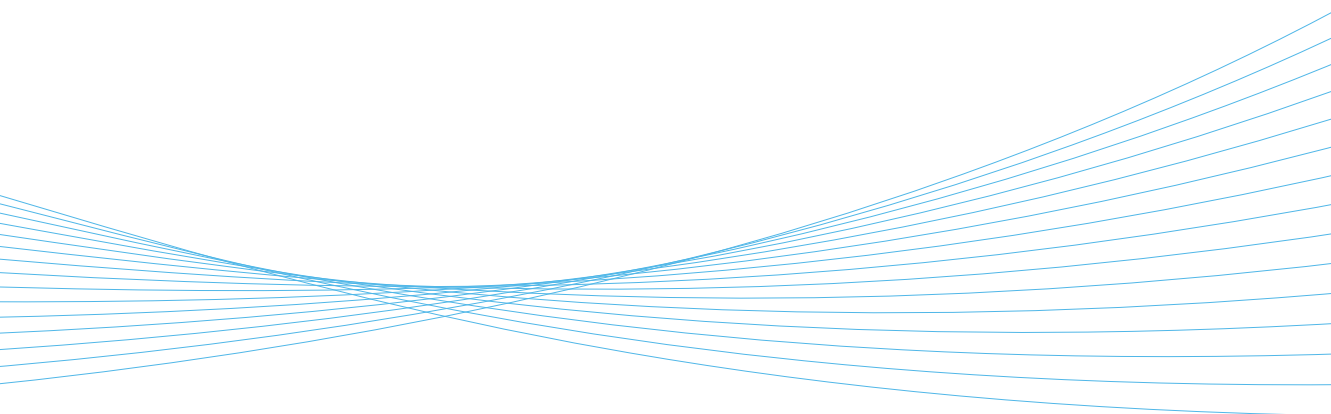


ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

119
CONTRIBUTIONS

MODELLING RADIATIVE AND CLIMATE EFFECTS OF AEROSOLS: FROM ANTHROPOGENIC EMISSIONS TO GEOENGINEERING

ANTON LAAKSO



Molecular events during translocation and proofreading extracted from 200 static structures of DNA polymerase

Zhong Ren^{1,2,*}

¹Department of Chemistry, The University of Illinois at Chicago, Chicago, IL 60607, USA and ²Renz Research, Inc., Westmont, IL 60559, USA

Received April 27, 2016; Revised June 7, 2016; Accepted June 7, 2016

ABSTRACT

DNA polymerases in family B are workhorses of DNA replication that carry out the bulk of the job at a high speed with high accuracy. A polymerase in this family relies on a built-in exonuclease for proofreading. It has not been observed at the atomic resolution how the polymerase advances one nucleotide space on the DNA template strand after a correct nucleotide is incorporated, that is, a process known as translocation. It is even more puzzling how translocation is avoided after the primer strand is excised by the exonuclease and returned back to the polymerase active site once an error occurs. The structural events along the bifurcate pathways of translocation and proofreading have been unwittingly captured by hundreds of structures in Protein Data Bank. This study analyzes all available structures of a representative member in family B and reveals the orchestrated event sequence during translocation and proofreading.

INTRODUCTION

Eukaryotic replicative DNA polymerases δ and ϵ belong to the polymerase family B. They are workhorses of genome replication during cell division (1,2). In bacteriophage T4, and a phylogenic variant RB69 in the T4 family, the gene product 43 (gp43) is a DNA polymerase in family B. RB69 DNA polymerase carries built-in 3'-exonuclease activity in a single polypeptide chain of 903 amino acid residues (3). This DNA polymerase incorporates hundreds of nucleotides to a primer strand in each second under instructions of a template strand (4–6). At the most fundamental level of the polymerase mechanism, this enzyme transfers a phosphoryl bond on P α of an incoming deoxyribonucleoside triphosphate (dNTP) to the 3' hydroxyl on the terminal ribose of the primer so that the primer strand becomes extended by one more nucleotide (7). It is well

known that the phosphoryl transfer reaction utilizes two or three divalent metal ions, often Mg²⁺ (8,9). This chemical reaction is highly regulated. First, a correct dNTP that satisfies Watson–Crick base pairing preferentially binds in the polymerase active site prior to the reaction. Second, the enzyme translocates by one nucleotide space along the single-stranded DNA template after a correct incorporation. Third, if an incorrect nucleotide is accidentally incorporated into the primer, the 3' terminus of the primer is switched to the exonuclease active site more than 30 Å away for an excision. When the corrected primer returns back to the polymerase active site, the translocation step is either bypassed or undone. This is equivalent to the function of the backspace key and termed proofreading (10). It is remarkable that RB69 DNA polymerase of 903 residues is fully capable of accomplishing such complex tasks. Good processivity in DNA synthesis requires additional accessory proteins to work jointly with the polymerase (3). The conformational changes associated with the catalytic and regulatory events are not yet clear due to the lack of suitable biophysical methods to directly observe the process at the atomic resolution. Therefore, the structural events associated with translocation and active site switching for processive proofreading (11) are largely unknown except for the chemical reaction of phosphoryl transfer.

This work employs a strategy of joint structural analysis to examine all known structures of RB69 DNA polymerase, and to reverse engineer the working mechanics of this polymerase. Nearly 170 independently observed structures of this polymerase are available in the Protein Data Bank (PDB). More than 40 other polymerase structures in family B are also available. This joint analysis is able to extract a conformational pathway that clearly shows detailed structural changes during translocation along a DNA template. A further extension of the translocation pathway leads to processive switching between the polymerase and exonuclease active sites. This joint analysis of more than 200 polymerase structures suggests a unified driving mechanism for translocation and proofreading. Extensive structural and functional data from mutants of this polymerase

*To whom correspondence should be addressed. Tel: +1 630 430 6190; Email: zren@uic.edu

Molecular Threading: Mechanical Extraction, Stretching and Placement of DNA Molecules from a Liquid-Air Interface

Andrew C. Payne^{1,2,3}, Michael Andregg^{2,3}, Kent Kemmish², Mark Hamalainen², Charlotte Howell², Andrew Bleloch², Nathan Klejwa², Wolfgang Lehrach², Ken Schatz², Heather Stark², Adam Marblestone³, George Church^{3,4*}, Christopher S. Own^{2*}, William Andregg²

1 Wyss Institute, Harvard University, Boston, Massachusetts, United States of America, **2** Halcyon Molecular, Inc., Redwood City, California, United States of America, **3** Biophysics Program and Wyss Institute, Harvard University, Boston, Massachusetts, United States of America, **4** Department of Genetics, Harvard Medical School, Boston, Massachusetts, United States of America

Abstract

We present “molecular threading”, a surface independent tip-based method for stretching and depositing single and double-stranded DNA molecules. DNA is stretched into air at a liquid-air interface, and can be subsequently deposited onto a dry substrate isolated from solution. The design of an apparatus used for molecular threading is presented, and fluorescence and electron microscopies are used to characterize the angular distribution, straightness, and reproducibility of stretched DNA deposited in arrays onto elastomeric surfaces and thin membranes. Molecular threading demonstrates high straightness and uniformity over length scales from nanometers to micrometers, and represents an alternative to existing DNA deposition and linearization methods. These results point towards scalable and high-throughput precision manipulation of single-molecule polymers.

Citation: Payne AC, Andregg M, Kemmish K, Hamalainen M, Howell C, et al. (2013) Molecular Threading: Mechanical Extraction, Stretching and Placement of DNA Molecules from a Liquid-Air Interface. PLoS ONE 8(7): e69058. doi:10.1371/journal.pone.0069058

Editor: Jerome Mathe, Université d'Evry val d'Essonne, France

Received: March 4, 2013; **Accepted:** June 3, 2013; **Published:** July 31, 2013

Copyright: © 2013 Payne et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors thank the U.S. Department of Energy Grant No. DE-FG02-02ER63445 and NIH Grant No. RC2 HG005592 for financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Two US patents has been filed on related work by authors of this manuscript: Sequencing Nucleic Acid Polymers with Electron Microscopy, inventors: William Andregg, Michael Andregg, US patent no. 8,153,438; Scanning Transmission Electron Microscopy, inventors: Christopher Su-Yan Own, William Andregg, Michael Andregg, Application number: 13/303,121 (pending publication), Filing date: 11/22/2011. Several of the authors are employed by Halcyon Molecular Inc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: gchurch@genetics.med.harvard.edu (GC); cown@voxa.co (CO)

These authors contributed equally to this work.

Introduction

In solution, a double stranded DNA (dsDNA) double-helix adopts an entropically favourable compact random-coil conformation, and a single stranded DNA (ssDNA) molecule, having a smaller persistence length, is more compact still [1]. By applying tension to the strand, a DNA molecule can be stretched into an entropically unfavorable, elongated state.

The first method developed for mechanical elongation of DNA involved stretching fibres of its precipitated sodium salt [2] in air, and a number of techniques for stretching DNA from or in solution have been developed since. Existing bulk methods for stretching DNA include molecular combing [3,4,5,6], transfer printing [7], and shear-induced stretching in micro- and nano-channels [8]. These techniques have been used to prepare DNA for sequence motif mapping and to characterize chromosomal abnormalities [9,10]. Molecular combing - where DNA molecules are elongated on compatible surfaces by the action of a receding meniscus - combined with transfer printing and electron microscopy, shows promise as a tool to enable direct reading of genetic and epigenetic information from stretched DNA molecules [11]. Existing single-molecule manipulation techniques, such as optical

and magnetic tweezers [12,13], atomic force microscopy [14], and micro-needle manipulation [15], have had success characterizing the behaviour of DNA and other bio-polymers under tension and torsion, as well as their interactions with other biomolecules [16,17].

Existing methods for stretching DNA have inherent limitations. Bulk methods such as molecular combing or transfer printing require either liquid (buffer) or solid (stamp) contact with the substrate, both of which are incompatible with sensitive or fragile surfaces. Additionally, stretching in molecular combing is sensitive to both the surface and the buffer [18,19,20]. A method to elongate DNA molecules while avoiding bulk contact of the carrier with the substrate is desirable because it circumvents these constraints. While single-molecule manipulation techniques can stretch DNA molecules independent of a surface, existing single-molecule methods do not enable the transfer of stretched molecules to a target substrate and are generally slow.

To overcome these limitations, we have developed an efficient tip-based method to mechanically stretch DNA molecules and deposit them onto a surface [21]. In this novel method, stretching is accomplished by mechanically pulling DNA segments from a droplet using a pulled-glass micro-needle and suspending them in

Is mental effort exertion contagious?

Kobe Desender¹, Sarah Beurms², & Eva Van den Bussche¹

1. Department of Psychology, Vrije Universiteit Brussel, Brussels, Belgium

2. Department of Psychology, KU Leuven, Belgium

Corresponding author:

Kobe Desender

Faculty of Psychology and Educational Sciences

Vrije Universiteit Brussel

Pleinlaan 2, 1050 Brussels

Belgium

Phone: +32-(0)2 629 14 67

Fax: +32-(0)2-629 24 89

E-mail: Kobe.Desender@vub.ac.be

Keywords: effort exertion, cognitive control, contagion, social facilitation, joint-Simon

4874 words below (Abstract, References excluded)

THE MICROSOFT 2017 CONVERSATIONAL SPEECH RECOGNITION SYSTEM

W. Xiong, L. Wu, F. Alleva, J. Droppo, X. Huang, A. Stolcke

Microsoft AI and Research
Technical Report MSR-TR-2017-39
August 2017

ABSTRACT

We describe the 2017 version of Microsoft’s conversational speech recognition system, in which we update our 2016 system with recent developments in neural-network-based acoustic and language modeling to further advance the state of the art on the Switchboard speech recognition task. The system adds a CNN-BLSTM acoustic model to the set of model architectures we combined previously, and includes character-based and dialog session aware LSTM language models in rescoring. For system combination we adopt a two-stage approach, whereby subsets of acoustic models are first combined at the senone/frame level, followed by a word-level voting via confusion networks. We also added a confusion network rescoring step after system combination. The resulting system yields a 5.1% word error rate on the 2000 Switchboard evaluation set.

1. INTRODUCTION

We have witnessed steady progress in the improvement of automatic speech recognition (ASR) systems for conversational speech, a genre that was once considered among the hardest in the speech recognition community due to its unconstrained nature and intrinsic variability [1]. The combination of deep networks and efficient training methods with older neural modeling concepts [2, 3, 4, 5, 6, 7, 8] have produced steady advances in both acoustic modeling [9, 10, 11, 12, 13, 14, 15] and language modeling [16, 17, 18, 19]. These systems typically use deep convolutional neural network (CNN) architectures in acoustic modeling, and multi-layered recurrent networks with gated memory (long-short-term memory, LSTM [8]) models for both acoustic and language modeling, driving the word error rate on the benchmark Switchboard corpus [20] down from its mid-2000s plateau of around 15% to well below 10%. We can attribute this progress to the neural models’ ability to learn regularities over a wide acoustic context in both time and frequency dimensions, and, in the case of language models, to condition on unlimited histories and learn representations of functional word similarity [21, 22].

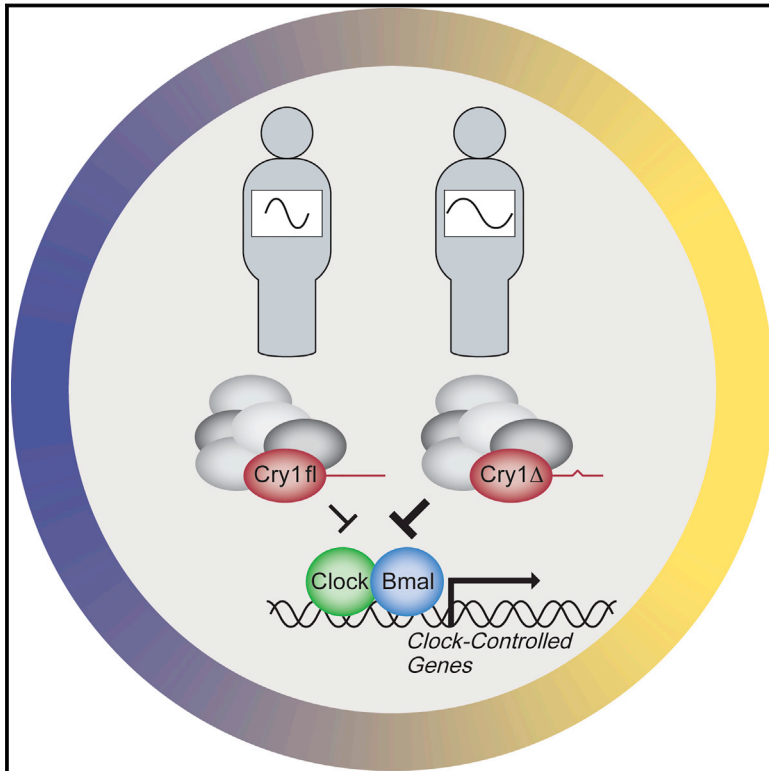
Given these developments, we carried out an experiment last year, to measure the accuracy of a state-of-the-art con-

versational speech recognition system against that of professional transcribers. We were trying to answer the question whether machines had effectively caught up with humans in this, originally very challenging, speech recognition task. To measure human error on this task, we submitted the Switchboard evaluation data to our standard conversational speech transcription vendor pipeline (who was left blind to the experiment), postprocessed the output to remove text normalization discrepancies, and then applied the NIST scoring protocol. The resulting human word error was 5.9%, not statistically different from the 5.8% error rate achieved by our ASR system [23]. In a follow-up study [24], we found that qualitatively, too, the human and machine transcriptions were remarkably similar: the same short function words account for most of the errors, the same speakers tend to be easy or hard to transcribe, and it is difficult for human subjects to tell whether an errorful transcript was produced by a human or ASR. Meanwhile, another research group carried out their own measurement of human transcription error [25], while multiple groups reported further improvements in ASR performance [25, 26]. The IBM/Appen human transcription study employed a more involved transcription process with more listening passes, a pool of transcribers, and access to the conversational context of each utterance, yielding a human error rate of 5.1%. Together with a prior study by LDC [27], we can conclude that human performance, unsurprisingly, falls within a range depending on the level of effort expended.

In this paper we describe a new iteration in the development of our system, pushing well past the 5.9% benchmark we measured previously. The overall gain comes from a combination of smaller improvements in all components of the recognition system. We added an additional acoustic model architecture, a CNN-BLSTM, to our system. Language modeling was improved with an additional utterance-level LSTM based on characters instead of words, as well as a dialog session-based LSTM that uses the entire preceding conversation as history. Our system combination approach was refined by combining predictions from multiple acoustic models at both the senone/frame and word levels. Finally, we added an LM rescoring step after confusion network creation, bringing us to an overall error rate of 5.1%, thus surpassing the human

Mutation of the Human Circadian Clock Gene *CRY1* in Familial Delayed Sleep Phase Disorder

Graphical Abstract



Authors

Alina Patke, Patricia J. Murphy, Onur Emre Onat, Ana C. Krieger, Tayfun Özçelik, Scott S. Campbell, Michael W. Young

Correspondence

patkea@rockefeller.edu (A.P.), young@mail.rockefeller.edu (M.W.Y.)

In Brief

A variation in the human circadian clock gene *CRY1* is associated with a familial form of delayed sleep phase disorder, providing genetic underpinnings for “night owls.”

Highlights

- A human subject with DSPD with a variation in *CRY1* has altered circadian rhythms
- Proband kindred and unrelated carrier families display aberrant sleep patterns
- The allele alters circadian molecular rhythms
- The genetic variation enhances *CRY1* function as a transcriptional inhibitor



Spontaneous Internalization of Cell Penetrating Peptide-Modified Nanowires into Primary Neurons

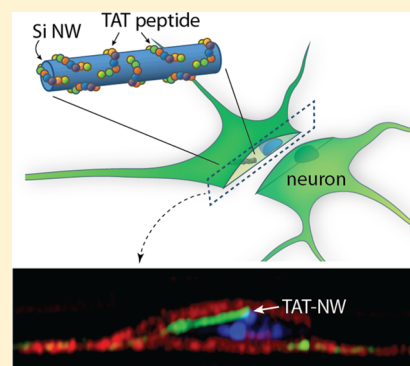
Jae-Hyun Lee,[†] Anqi Zhang,[†] Siheng Sean You,[†] and Charles M. Lieber^{*,†,‡}

[†]Department of Chemistry and Chemical Biology and [‡]John A. Paulson School of Engineering and Applied Science, Harvard University, Cambridge, Massachusetts 02138, United States

S Supporting Information

ABSTRACT: Semiconductor nanowire (NW) devices that can address intracellular electrophysiological events with high sensitivity and spatial resolution are emerging as key tools in nanobioelectronics. Intracellular delivery of NWs without compromising cellular integrity and metabolic activity has, however, proven difficult without external mechanical forces or electrical pulses. Here, we introduce a biomimetic approach in which a cell penetrating peptide, the trans-activating transcriptional activator (TAT) from human immunodeficiency virus 1, is linked to the surface of Si NWs to facilitate spontaneous internalization of NWs into primary neuronal cells. Confocal microscopy imaging studies at fixed time points demonstrate that TAT-conjugated NWs (TAT-NWs) are fully internalized into mouse hippocampal neurons, and quantitative image analyses reveal an ca. 15% internalization efficiency. In addition, live cell dynamic imaging of NW internalization shows that NW penetration begins within 10–20 min after binding to the membrane and that NWs become fully internalized within 30–40 min. The generality of cell penetrating peptide modification method is further demonstrated by internalization of TAT-NWs into primary dorsal root ganglion (DRG) neurons.

KEYWORDS: Silicon nanowires, surface modification, TAT peptide transfection, hippocampal neurons, dorsal root ganglion cells, live cell confocal imaging, membrane penetration



The application of nanomaterials in bioelectronics,¹ including ultrasensitive sensing,^{2–6} single-cell electrophysiological probes,^{7–12} flexible, stretchable, and/or degradable electronics,^{13–15} and macroporous three-dimensional electronics^{16,17} have enabled research ranging from disease marker detection⁵ to cell electrophysiology¹¹ and brain mapping.¹⁸ Surface modification of nanodevices is well-established in the area of sensing where linkage of receptors is crucial for selective detection,⁵ but also represents a critical challenge for creating bio–nano interfaces that are stable at the cellular level. For example, internalization of Si NW devices across the cell membrane in cardiac cells has demonstrated the capability for nanoscale sensing elements to record intracellular action potentials without compromising cell viability.^{7–9} Similar results have been difficult to achieve in neuronal systems, where the capability to detect stable full-amplitude intracellular action potentials without interference from the cell membrane, could open up unique research opportunities.

To date, the use of NWs for interrogating the intracellular environment of neurons has been limited by challenges in achieving stable internalization of the highly anisotropic NW devices. A limited number of approaches have been reported for the internalization of highly anisotropic NWs in neurons, with an emphasis on physical methods.^{10,19–23} Specifically, neurons were cultured on substrates with vertical NW arrays, which delivered biomolecules to the cytoplasm as a result of gravity and/or adhesive force driven internalization,¹⁹ and

electroporation was used to achieve transient access to intracellular action potentials.¹⁰ Concurrently, these same physical strategies have been used in a variety of other cell types to demonstrate internalization of vertical NW^{8,24,25}/nanotube arrays^{26–29} or single NWs³⁰ into the cells.

In another approach, chemically based alteration of the cell membrane and subsequent NW internalization has been achieved by treating Chinese hamster ovary cells with the polar organic solvent dimethyl sulfoxide (DMSO).³¹ Given that the typical size of NWs used in electronic devices is comparable to large macromolecules, viruses, and nanoparticles, it is reasonable to ask whether a biomimetic strategy used for targeted delivery of these structures might be developed for driving internalization of NWs. Indeed, surface modification of Si NW devices with cell membrane-like phospholipids has been shown to be effective in inducing internalization of the NWs into cardiac cells and enabling intracellular recording of stable action potentials.^{7,9} However, successful application of this biomimetic approach has not yet been reported for the internalization of NWs into neurons.

Here, we demonstrate a new strategy for complete internalization of NWs into neurons using cell penetrating peptides (CPPs). CPPs are short peptides capable of delivering cargo into cells.³² TAT peptide is a well-known CPP derived

Received: January 3, 2016

Published: January 8, 2016



Next generation 1536-well oligonucleotide synthesizer with on-the-fly dispense



Michael Jensen^{a,*}, Lester Roberts^a, Andrew Johnson^b, Marilyn Fukushima^a, Ronald Davis^a

^a Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94304, USA

^b Jova Solutions, Berkeley, CA 94710, USA

ARTICLE INFO

Article history:

Received 27 October 2013

Received in revised form

21 November 2013

Accepted 25 November 2013

Available online 16 December 2013

Keywords:

Oligonucleotide

Synthesis automation

Chemical DNA synthesis

Synthetic biology

ABSTRACT

Here we report the development of our Next Generation Automated Multiplexed Oligonucleotide Synthesizer (NG-1536-AMOS), capable of producing 1536 samples in a single run using a multi-well filtered titer plate. With the potential to synthesize up to 3456 samples per plate, we converted the BioRAPTR Flying Reagent Dispenser into an open-well system where spent reagents are drained to waste under vacuum. During synthesis, reagents are delivered on-the-fly to each micro-titer well at volumes $\leq 5 \mu\text{l}$ with plate speeds up to 150 mm/s. Using gas-phase cleavage and deprotection, a full plate of 1536 60mers may be processed with same-day turnaround with an average yield per well at 3.5 nmol. Final product at only \$0.00277/base is eluted into a low-volume collection plate for immediate use in downstream application (e.g. Biomek FX for versatile sample handling). Also, crude oligonucleotide quality is comparable to that of commercial synthesis instrumentation, with an error rate on the NG-1536-AMOS platform of 1.53/717 bases. Furthermore, mass spectral analysis on strands synthesized up to 80 bases showed high purity with an average coupling efficiency of 99.5%.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The Stanford Genome Technology Center (SGTC) has been at the forefront of oligonucleotide DNA synthesis automation since the early 1990s with its development of the first 96-well Automated Multiplexed Oligonucleotide Synthesizer (96-AMOS) (Lashkari et al., 1995). Capable of generating 6000 25-base primers per month on a single machine, it made large-scale projects such as sequencing the *Saccharomyces cerevisiae* and human genomes timely and economically feasible (Dietrich et al., 1997; Lander et al., 2001).

While industry costs at the time for low-throughput column synthesis were \$2.00–3.00 per base (Hager et al., 1999; Goforth, 2002), 96-AMOS was synthesizing DNA 10–15-fold cheaper within the first year of production. Currently at \$0.06 per base, this platform remains competitive with other column and titer plate-based synthesizers that have since been introduced (Cheng et al., 2002; Livesay et al., 2002; Rayner et al., 1998; Lebl et al., 2007; Gan, 1990; Jensen et al., 2012).

However, with today's multiplexed, small volume assays (Hardenbol et al., 2003; Dahl et al., 2007), there is a greater demand

for even higher throughput and lower cost oligonucleotide synthesis. As such high density synthesis arrays and micro-fluidic devices (Gao et al., 2004; Livesay et al., 2002; Lee et al., 2010; Kong et al., 2007) offer mass production at a fraction of the cost. And though array densities can be as high as 4 million features, low yield and heterogeneous oligonucleotide pooling may present a challenge for specific downstream applications, in particular synthetic biology (Kim et al., 2012; Borovkov et al., 2010; Ma et al., 2012a,b; Quan et al., 2011; Mueller et al., 2009).

In contrast to the microarray which may be limited to 10^6 molecules per feature, multi-well/column platforms generally give very high yields ($>25 \text{ nmol}$); but this is often a wasteful over-production of sample. As a consequence, excess product is either disposed of or kept in deep-well collection plates (volumes up to 1.2 ml/well) for long term storage thus taking up valuable freezer space.

Focusing on the strengths of both the array and multi-well/column platforms, we describe the development of our Next Generation AMOS capable of producing 1536 samples on a single titer plate (NG-1536-AMOS). Based on the original 96-AMOS, we combined its technology with that of the commercial BioRAPTR Flying Reagent (non-contact) Dispenser (BioRAPTR FRD, Beckman Coulter), intended for ultra high-throughput delivery (up to 3456 wells) of aqueous buffers, diluents and dye solutions (Yasgar et al., 2012; Auld et al., 2008).

* Corresponding author. Tel.: +1 650 812 2744; fax: +1 650 812 1975.
E-mail address: m.a.jensen@stanford.edu (M. Jensen).

	Name	username	email
2	Josef Davies-Coates	josefdavies-coates	josef@uniteddiversity.com
3	tav	tav	tav@espians.com
4	Bas Reus	basreus	bas.reus@gmail.com
5	mose	mose	p2p@mose.fr
6	Chris Cook	chriscook	cojock@hotmail.com
7	AGNUcius	agnucius	AGNUcius@Gmail.com
8	Michel Bauwens	michelbauwens	michelsub2003@yahoo.com
9	James Burke	jamesburke	lifesized@gmail.com
10	C4Chaos	c4chaos	coolmel@gmail.com
11	Vinay Gupta	vinaygupta	hexayurt@gmail.com
12	Valentin Spirik	valentinspirik	realdigirev@wolke7.net
13	Tia Carr Williams	tiacarrwilliams	Tia@novuminstitute.org
14	James	james	james.froggatt@gmail.com
15	Christian Siefkes	christiansiefkes	christian@siefkes.net
16	Robin Good	robingood	Robin.Good@masternewmedia.org
17	Karen Eliot	kareneliot	indy@burntout.org
18	Stan Rhodes	stanrhodes	stanleyrhodes@gmail.com
19	Colby Stuart	colbystuart	colbymedia@gmail.com
20	Paul B. Hartzog	paulbhartzog	paulbhartzog@gmail.com
21	Marcin Jakubowski	marcinjakubowski	joseph.dolittle@gmail.com
22	david cozens	davidcozens	davidcozens@yahoo.co.uk
23	Tom Murray	tommurray	tmurray@cs.umass.edu
24	Matt Cooperrider	mattcooperrider	mattcooperrider@gmail.com
25	Edward Berge	edwardberge	theurj@netzero.net
26	Sureshbabu Basavayya	sureshbabubasavayya	sureshbabu.basavayya@gmail.com
27	Georg Pleger	georgpleger	georg@pleger.at
28	Sepp Hasslberger	sepphasslberger	sepp@lastrega.com
29	Eric Prenen	ericprenen	eprenen@gmail.com
30	Teemu Arina	teemuarina	teemu@dicole.com
31	Jadon	jadon	ning@hangerhead.com
32	Tia Carr Williams	tiacarrwilliams1	tiacarrwilliams@gmail.com
33	Kare Anderson	kareanderson	kare@sayitbetter.com
34	Ian Bruk	ianbruk	chiefouthousecorrespondent@gmail.com
35	Vasilis Kostakis	vasiliskostakis	kostakis.b@gmail.com
36	Peter	peter	p.a.scott@btinternet.com
37	Stephan Dohrn	stephandohrn	stephandohrn@gmail.com
38	Tom Loeber	tomloeber	TomLoeber@gmail.com
39	Dante-Gabryell Monson	dante-gabryellmonson	dante.monson@gmail.com
40	eggy1943	eggy1943	bruce.eggum@gmail.com
41	Stefan Meretz	stefanmeretz	stefan@meretz.de
42	Julian Dibbell	juliandibbell	julian@juliandibbell.com
43	SamRose	samrose	samuel.rose@gmail.com
44	OlivierAuber	olivierauber	olivier.auber@km2.net
45	Michael Riversong	michaelriversong	rivedu@earthlink.net
46	Natalie	natalie	nataliep@gmx.at

Node-By-Node Greedy Deep Learning for Interpretable Features

Ke Wu

Department of Computer Science, 110 8th Street, Troy, NY 12180 USA

Malik Magdon-Ismail

Department of Computer Science, 110 8th Street, Troy, NY 12180 USA

WUK3@RPI.EDU

MAGDON@GMAIL.COM

Abstract

Multilayer networks have seen a resurgence under the umbrella of deep learning. Current deep learning algorithms train the layers of the network sequentially, improving algorithmic performance as well as providing some regularization. We present a new training algorithm for deep networks which trains *each node in the network* sequentially. Our algorithm is orders of magnitude faster, creates more interpretable internal representations at the node level, while not sacrificing on the ultimate out-of-sample performance.

1. Introduction

Multilayer neural networks have gone through ups and downs since their arrival in (Rosenblatt, 1958; Widrow, 1960; Hoff Jr, 1962). The resurgence in “deep” networks is largely due to the efficient greedy layer by layer algorithms for training, that create meaningful hierarchical representations of the data. Particularly in the era of “big data” from diverse applications, efficient training to create data representations that provide insight into the complex features captured by the neurons are important. We explore these two dimensions of training a deep network. Assume a standard machine learning from data setup (Abu-Mostafa et al., 2012), with N datapoints $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$; $\mathbf{x}_n \in \mathbb{R}^d$ and $y_n \in \{0, 1, \dots, c - 1\}$ (multi-class setting).

We refer to Abu-Mostafa et al. (2012, Chapter 7) for the basics of multilayer networks, including notation which we very quickly summarize here. On the right, we show a feedforward network architecture. Such a network is “deep” because it has many ($\gg 2$) layers. We assume that a network architecture has been fixed. The network implements a function whereby in each layer (ℓ), the output of the previous layer ($\ell - 1$) is transformed into the output of the layer ℓ

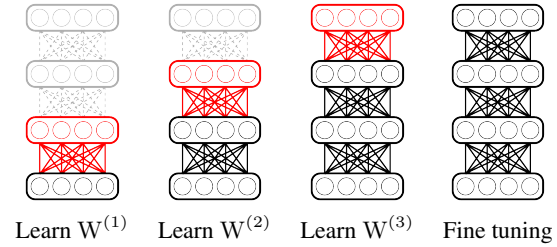
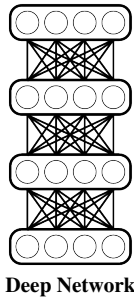


Figure 1. Layer-by-layer greedy deep learning algorithm.

until one reaches the final layer on top, which is the output of the network. The function implemented by layer ℓ is

$$\mathbf{x}^{(\ell)} = \tanh(W^{(\ell)}\mathbf{x}^{(\ell-1)}),$$

where $\mathbf{x}^{(\ell)}$ is the output of layer ℓ , and the weight-matrix $W^{(\ell)}$ (of appropriate dimensions to map a vector from layer $\ell - 1$ to a vector in ℓ) are parameters to be learned from data. The training phase uses the data to identify all the parameters $\{W^{(1)}, W^{(2)}, \dots, W^{(L)}\}$ of the deep network.

Backpropagation which trains all the weights simultaneously, allowing for maximum flexibility, was the popular approach to training a deep network (Rumelhart et al., 1986). The current approach is layer-by-layer: train the first layer weights $W^{(1)}$; then train the second layer weights $W^{(2)}$, *keeping the first layer weights fixed*; and so on until all the weights are learned. In practice, once all the weights have been learned in the greedy-layer-by-layer manner (often referred to as pre-training), the best results are obtained by fine tuning all the weights using a few iterations of backpropagation (Figure 1).

How one should train the internal layers? The two popular approaches are: (1) Each layer is an unsupervised nonlinear auto-encoder (Cottrell & Munro, 1988; Bengio et al., 2007); this approach is appealing to build meaningful hierarchical representations of the data in the internal layers. (2) Each layer is a supervised encoder; this approach primarily targets performance. Deep learning enjoys success in several applications and hence considerable effort



Non-Viral Nucleic Acid Delivery Strategies to the Central Nervous System

James-Kevin Y. Tan¹, Drew L. Sellers¹, Binhlan Pham¹, Suzie H. Pun¹ and Philip J. Horner^{2*}

¹ Department of Bioengineering and Molecular Engineering & Sciences Institute, University of Washington, Seattle, WA, USA,

² Center for Neuroregenerative Medicine, Houston Methodist Research Institute, Houston, TX, USA

With an increased prevalence and understanding of central nervous system (CNS) injuries and neurological disorders, nucleic acid therapies are gaining promise as a way to regenerate lost neurons or halt disease progression. While more viral vectors have been used clinically as tools for gene delivery, non-viral vectors are gaining interest due to lower safety concerns and the ability to deliver all types of nucleic acids. Nevertheless, there are still a number of barriers to nucleic acid delivery. In this focused review, we explore the *in vivo* challenges hindering non-viral nucleic acid delivery to the CNS and the strategies and vehicles used to overcome them. Advantages and disadvantages of different routes of administration including: systemic injection, cerebrospinal fluid injection, intraparenchymal injection and peripheral administration are discussed. Non-viral vehicles and treatment strategies that have overcome delivery barriers and demonstrated *in vivo* gene transfer to the CNS are presented. These approaches can be used as guidelines in developing synthetic gene delivery vectors for CNS applications and will ultimately bring non-viral vectors closer to clinical application.

OPEN ACCESS

Edited by:

George Smith,
Temple University School of
Medicine, USA

Reviewed by:

Eduardo Fernandez,
Universidad Miguel Hernández de
Elche, Spain
Marianna Foldvari,
University of Waterloo, Canada

*Correspondence:

Philip J. Horner
pjhorner@houstonmethodist.org

Received: 13 July 2016

Accepted: 11 October 2016

Published: 01 November 2016

Citation:

Tan J-KY, Sellers DL, Pham B,
Pun SH and Horner PJ
(2016) Non-Viral Nucleic Acid
Delivery Strategies to the Central
Nervous System.
Front. Mol. Neurosci. 9:108.
doi: 10.3389/fnmol.2016.00108

Keywords: central nervous system, delivery, *in vivo*, non-viral, nucleic acid

INTRODUCTION

The incidence of neurological diseases and injuries is increasing with the rising life expectancy (Mattson and Magnus, 2006). Nucleic acid therapeutics, such as genes and small interfering RNA (siRNA) oligonucleotides have emerged as a promising treatment strategy to preserve neuron function, enhance neurogenesis and prevent the progression of neurological diseases. The delivery of nucleic acids encoding brain-derived neurotrophic factor (Huang et al., 2012a), epidermal growth factor (Sugiura et al., 2005), fibroblast growth factor-2 (Matsuoka et al., 2003), Huntingtin (Burgess et al., 2012), neurogenin-2 (Zhang et al., 2013; Masserdotti et al., 2015), insulin growth factor-1 (Kaspar et al., 2003), and vascular endothelial growth factor (Dodge et al., 2010) have been shown to increase neuron regeneration or delay the progression of neurological diseases in mice, rats and gerbils. Targeting gene delivery vehicles to the appropriate cells and proper protein regulation remain the primary challenges to making these pathways feasible. While viral vectors such as the adeno-associated virus have typically been used clinically, interest in non-viral nucleic acid delivery remains high due to lower safety concerns, greater customizability and an ease in manufacturing (Pack et al., 2005; Burke et al., 2013). In fact, the number of synthetic vectors used in gene therapy clinical trials has been steadily increasing over the last 10 years (Gene Therapy Clinical Trials Worldwide, Wiley).

A Survey of Space Settlement Designs

Thomas Marotta

December 19, 2016

I. Summary

The author compiled every known orbital space settlement design into a database. Grouped into chronological 'eras,' the database describes basic information for each design: population capacity, dimensions, gravity level, energy source, etc. Using this information one can conclude that interest in space settlement is increasing, 1g is the preferred gravity level, solar power is the preferred energy source, and a torus is the preferred geometry. As for location, Earth-Moon Lagrange points dominate but there is a budding movement to place settlements in low Earth orbits. The database is accessible at <http://www.nss.org/settlement/journal/Space-Settlement-Designs-Database-12.19.16.pdf>.

II. Introduction

Interest in space settlement is on the rise. The National Aeronautics and Space Administration (NASA) is making slow and steady progress on its journey to Mars¹. Elon Musk, billionaire founder of commercial space company SpaceX, recently announced a plan to colonize Mars². When a marketing company announced plans to offer a one-way trip to the Red Planet, over 200,000 people volunteered³.

While efforts to settle Mars garner most media coverage, there is a significant, albeit less publicized, level of interest in orbital space settlement—that is, permanent human settlement in structures orbiting around a celestial body rather than on its surface. Well-heeled startups Bigelow Aerospace, Virgin Galactic and Blue Origin have all made commitments to expanding the human presence in orbit. It's not just the startups, though: established aerospace giant United Launch Alliance is planning for hundreds of people living and working in orbital space in the coming decades⁴. Even President Obama is strongly in favor of the concept of space settlement⁵.

In light of the burgeoning interest in orbital space settlement, this survey was assembled to facilitate education and research on the topic. The accompanying database collects all known, detailed plans for orbital space settlements. The database is a work in progress

¹ NASA's Journey to Mars, <http://www.nasa.gov/content/nasas-journey-to-mars>, accessed 14 Dec 2016

² "IAC Mars Talk, Revised," SpaceX, http://www.spacex.com/sites/spacex/files/mars_presentation.pdf, accessed 14 Dec 2016

³ Landau, Elizabeth, "200,000 people apply to live on Mars," CNN, 10 Dec 2013, <http://www.cnn.com/2013/12/10/tech/innovation/mars-one-plan/>, accessed 14 Dec 2016

⁴ "Transportation Enabling a Robust Cislunar Economy," United Launch Alliance, 9 April 2016, http://www.ulalaunch.com/uploads/docs/Published_Papers/Commercial_Space/2016_Cislunar.pdf, accessed 14 Dec 2016

⁵ "We have set a clear goal vital to the next chapter of America's story in space: sending humans to Mars by the 2030s and returning them safely to Earth, with the ultimate ambition to one day remain there for an extended time." President Barack Obama, 11 Oct 2016, CNN, <http://www.cnn.com/2016/10/11/opinions/america-will-take-giant-leap-to-mars-barack-obama/index.html?adkey=bn> accessed 14 Dec 2016

Nucleic acid memory

Victor Zhirnov, Reza M. Zadegan, Gurtej S. Sandhu, George M. Church and William L. Hughes

Nucleic acid memory has a retention time far exceeding electronic memory. As an alternative storage media, DNA surpasses the information density and energy of operation offered by flash memory.

Information and communication technologies generate vast amounts of data that will far eclipse today's data flows (Fig. 1). Memory materials must therefore be suitable for high-volume manufacturing. At the same time, they must have elevated information stability and limit the energy consumption and trailing environmental impacts that such flows will demand. Analysts estimate that global memory demand — at 3×10^{24} bits — will exceed projected silicon supply in 2040 (Fig. 1b and Supplementary Information sections 1 and 2). To meet such requirements, flash-memory manufacturers would need $\sim 10^9$ kg of silicon wafers even though the total projected wafer supply is $\sim 10^7$ – 10^8 kg (Supplementary Figs 1 and 2). Such forecasts motivate an exploration of unconventional materials with cost-competitive performance attributes. With information retention times that range from thousands to millions of years, volumetric density 10^3 times greater than flash memory and energy of operation 10^8 times less, we believe that DNA used as a memory-storage material in nucleic acid memory (NAM) products promises a viable and compelling alternative to electronic memory.

In this Commentary, we discuss the information retention, density and energetics of NAM — specifically related to DNA — for non-biological and non-volatile memory applications, ranging from letters to libraries. The potential of NAM has often been dismissed, as nucleic acids are believed by some to be fragile and therefore unreliable. This is not the case. For example, the room-temperature half-life of ancient DNA exceeds 100 years^{1,2}. Indeed, the complete genomes of an $\sim 50,000$ -year-old Neanderthal³ recovered from Siberia and an $\sim 700,000$ -year-old horse⁴ recovered from the Arctic permafrost (approximate average temperature -4°C) have been sequenced. Still, the long-term stability of DNA and its decay kinetics are poorly understood at a per-bit (that is, base) level. As an energy-barrier model shows (Methods), DNA has a retention time far exceeding electronic memory, and it can store information reliably over time. Through first-principle calculations, DNA has been validated as a model material for future NAM products (Supplementary Information section 8). Therefore, we call for increased cooperation between the biotechnology and semiconductor sectors to pair previously

unfathomable technological advances — such as those from the Human Genome Project — with the scaling expertise of the semiconductor industry.

Nucleic acid memory as a material

As a material, nucleic acids are negatively charged polyelectrolytes with four monomers (the nucleotides A, T or U, C and G). Monomers are covalently bonded to form polymer chains. Once polymerized, an individual chain can hydrogen-bond with itself or with other chains that satisfy base complementarity. These attributes endow nucleic acids with the power of molecular self-assembly, which is made possible by the thermal fluctuations between complementary hydrogen bonds during Watson–Crick hybridization. During DNA hybridization, adenine (A) forms a base-pair with thymine (T), and guanine (G) pairs with cytosine (C). In RNA, thymine is substituted by uracil (U). By encoding sequence complementarity, molecular self-assembly can be exploited to pull nucleic acids like a rope⁵, weave them like a fabric^{6,7}, decorate them like a scaffold^{8,9} and recycle¹⁰ them like a thermoplastic. Beyond their recyclability, nucleic acids and potential

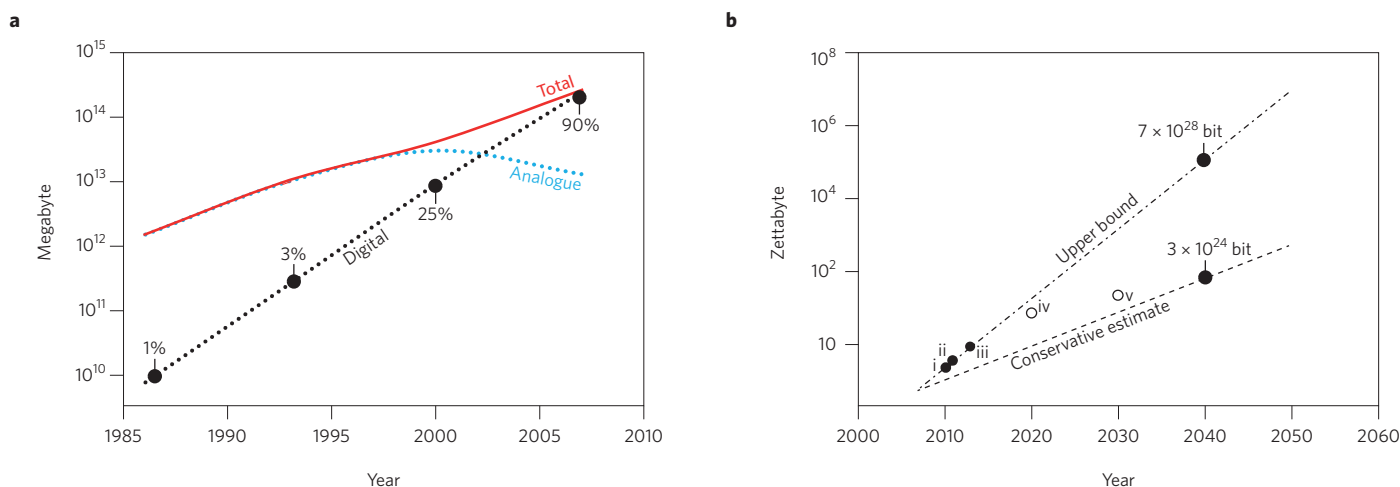
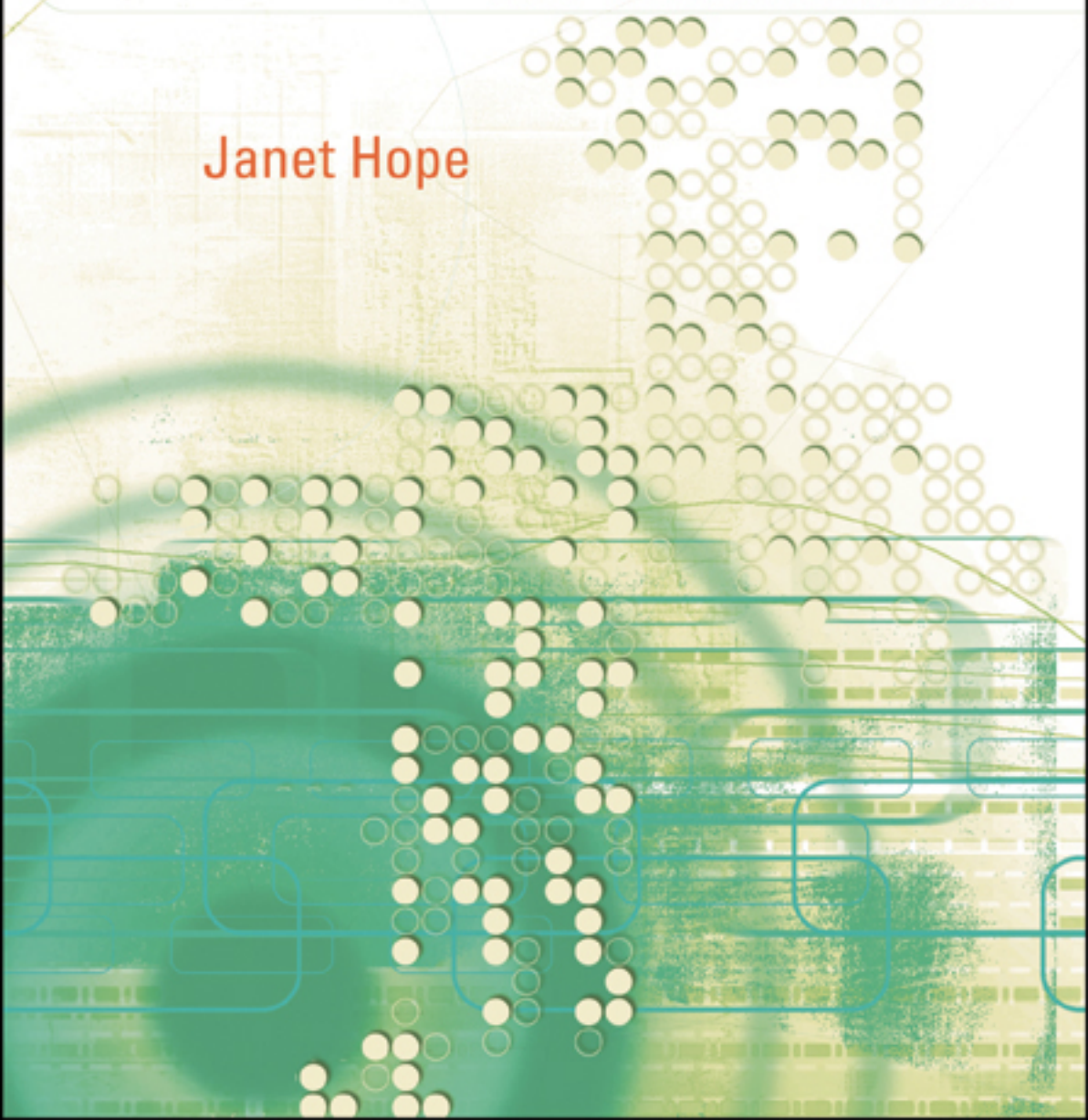


Figure 1 | Change of storage needs over time. **a**, Timeline of stored analogue, digital and total data (ref. 48) where the percentage values refer to the fraction of stored digital data. **b**, Projected global memory demand. Actual (filled circles: i, ref. 49; ii, ref. 50; iii, ref. 51) and projected (open circles: iv, ref. 51; v, ref. 52) data points fall between the conservative estimate and the upper bound. See also Supplementary Information section 1.

BIOBAZAR

the Open Source Revolution
and BIOTECHNOLOGY

Janet Hope



Parallel *In Vivo* DNA Assembly by Recombination: Experimental Demonstration and Theoretical Approaches

Zhenyu Shi^{1,3*}, Anthony G. Wedd^{1,3}, Sally L. Gras^{2,3}

1 School of Chemistry, University of Melbourne, Parkville, Victoria, Australia, **2** Department of Chemical & Biomolecular Engineering, University of Melbourne, Parkville, Victoria, Australia, **3** Bio21 Molecular Science and Biotechnology Institute, Parkville, Victoria, Australia

Abstract

The development of synthetic biology requires rapid batch construction of large gene networks from combinations of smaller units. Despite the availability of computational predictions for well-characterized enzymes, the optimization of most synthetic biology projects requires combinational constructions and tests. A new building-brick-style parallel DNA assembly framework for simple and flexible batch construction is presented here. It is based on robust recombination steps and allows a variety of DNA assembly techniques to be organized for complex constructions (with or without scars). The assembly of five DNA fragments into a host genome was performed as an experimental demonstration.

Citation: Shi Z, Wedd AG, Gras SL (2013) Parallel *In Vivo* DNA Assembly by Recombination: Experimental Demonstration and Theoretical Approaches. PLoS ONE 8(2): e56854. doi:10.1371/journal.pone.0056854

Editor: Mark Isalan, Center for Genomic Regulation, Spain

Received: October 24, 2012; **Accepted:** January 17, 2013; **Published:** February 28, 2013

Copyright: © 2013 Shi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported from funds provided by the University of Melbourne and the Bio21 Molecular Science and Biotechnology Institute, School of Chemistry, University of Melbourne (<http://www.chemistry.unimelb.edu.au>) and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne (<http://www.bio21.unimelb.edu.au>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: shiz@student.unimelb.edu.au

Introduction

Recombinant DNA technology was the first technique proposed for targeted manipulation of DNA and has had a wide-ranging and lasting impact on many fields [1]. However, its limitations include the appearance of operational restriction sites inside the DNA sequence of interest and the challenges of *in vitro* manipulation of long DNA fragments. As a result, a defined set of rules has evolved for assembly from standardized units (bricks). The BioBrick system was one of the first attempts to standardize assembly using recombinant DNA technology [2] and has been applied in a number of synthetic biology projects [3,4]. Improvements allowed the issues of junction scars in translation to be avoided (BglBricks; [5]) and permitted ligation of multiple DNA fragments in a single step [6].

Recently, *in vitro* recombination techniques such as sequence and ligation – independent cloning (SLIC) have been developed [7] and applied to prepare small fragments as part of the synthesis of bacterial genomes [8,9]. It is possible to apply SLIC to the building brick strategy, but the relatively long boundaries favour pre-designed seamless assembly as SLIC is sequence independent. The efficiency of the method is similar to that of recombinant DNA technology. Hence the longer boundary is a disadvantage for small standard unit systems. An improved SLIC approach (the Gibson isothermal method) has been commercialized [10,11]. The In-Fusion [12] and uracil-specific excision reagent (USER) [13] techniques are based on *in vitro* single-strand annealing. They employ relatively short overhangs of 15 to 30 bp, and can be modified into *in vitro* building brick systems [14].

Site-specific recombination methods have been applied widely to *in vitro* cloning. The Invitrogen Gateway system [15–18] uses

Lambda phage attachment sequence recombination and is one of the most popular approaches as it allows multiple sites to be accessed [19]. The Cre-loxP [20–23] and FLP-*frt* [24–26] systems are also very efficient. However, they are less favoured for *in vitro* cloning as they do not have the binary recombination features of the attachment sequence system (such as the attachment sequence recombination: attB + attP = attL + attR) but remain the same after recombination. The Integron system has a similar binary mechanism to that of the Lambda phage attachment system and has been applied to the assembly of multiple genes [27].

All the above strategies are *in vitro* technologies. They rely on both high reaction and high transformation efficiencies. Standardized building brick systems aim to simplify the construction process, both for design and experiments. Since complex designs can be managed via a software platform, the reduction of experimental complexity is an important issue. One possibility is to carry out the assembly *in vivo* to avoid all *in vitro* operations. *In vivo* recombination, such as homologous recombination in *Bacillus subtilis* [28,29] and in yeast [8,9,30–32], are more popular for the parallel assembly of large fragments. Although *in vivo* site-specific recombination is more efficient than *in vitro* homologous recombination, it has rarely been applied to DNA assembly except for some iterative systems [33–35] and individual integrations [36]. Lambda red-based recombinations have been applied widely in *E. coli* for gene knockout [37–40], for the generation of mutations [41] and for the integration of fragments into plasmids or genomes [35,42–45]. The full length prophage protein recE was also reported to be able to assemble multiple fragments *in vivo* [46]. It is apparent that these extremely efficient *in vivo* recombination methods have potential in parallel DNA assembly.

Concurrency and Privacy with Payment-Channel Networks*

Giulio Malavolta[†]

Friedrich-Alexander University Erlangen-Nürnberg
malavolta@cs.fau.de

Pedro Moreno-Sanchez[†]

Purdue University
pmorenos@purdue.edu

Aniket Kate

Purdue University
aniket@purdue.edu

Matteo Maffei

TU Wien
matteo.maffei@tuwien.ac.at

Srivatsan Ravi

University of Southern California
srivatsanravi@purdue.edu

Abstract

Permissionless blockchains protocols such as Bitcoin are inherently limited in transaction throughput and latency. Current efforts to address this key issue focus on off-chain payment channels that can be combined in a Payment-Channel Network (PCN) to enable an unlimited number of payments without requiring to access the blockchain other than to register the initial and final capacity of each channel. While this approach paves the way for low latency and high throughput of payments, its deployment in practice raises several privacy concerns as well as technical challenges related to the inherently concurrent nature of payments, such as race conditions and deadlocks, that have been understudied so far.

In this work, we lay the foundations for privacy and concurrency in PCNs, presenting a formal definition in the Universal Composability framework as well as practical and provably secure enforcement solutions. In particular, we present Fulgor and Rayo. Fulgor is the first payment protocol for PCNs that provides provable privacy guarantees for PCN and is fully compatible with the Bitcoin scripting system. Nevertheless, Fulgor is a blocking protocol and therefore prone to deadlocks of concurrent payments as in currently available PCNs. Instead, Rayo is the first protocol for PCNs that enforces *non-blocking progress* (i.e., at least one of the concurrent payments terminates). We show through a new impossibility result that the latter property necessarily comes at the cost of weaker privacy. At the core of Fulgor and Rayo is Multi-Hop HTLC, a new smart contract, compatible with Bitcoin, that provides conditional payments while reducing running time and communication overhead with respect to previous approaches. Our performance evaluation of Fulgor and Rayo shows that a payment with 10 intermediate users takes as few as 5 seconds, thereby demonstrating their feasibility to be deployed in practice.

1 Introduction

Bitcoin [58] is a fully decentralized digital cryptocurrency network that is widely adopted today as an alternative monetary payment system. Instead of accounting payments in a ledger locally maintained by a trusted financial institute, these are logged in the Bitcoin blockchain, a database replicated among mutually distrusted users around the world who update it by means of a global consensus algorithm based on proof-of-work. Nevertheless, the permissionless nature of this consensus algorithm

limits the transaction rate to tens of transactions per second whereas other payment networks such as Visa support peaks of up to 47,000 transactions per second [16].

In the forethought of a growing number of Bitcoin users and most importantly payments about them, scalability is considered today an important concern among the Bitcoin community [67, 3]. Several research and industry efforts are dedicated today to overcome this important burden [3, 63, 61, 32, 5, 2].

The use of Bitcoin *payment channels* [11, 32] to realize off-chain payments has flourished as a promising approach to overcome the Bitcoin scalability issue. In a nutshell, a pair of users open a payment channel by adding a single transaction to the blockchain where they lock their bitcoins in a

*This is the revision 16 June 2017. The most recent version is available at <https://www.cs.purdue.edu/homes/pmorenos/public/paychannels.pdf>

[†]Both authors contributed equally and are considered to be co-first authors.



Samba G3L Printhead

Features:

- Advanced Silicon Micro-Electro-Mechanical Systems
- 2048 nozzles per module
- 1200 native dpi/600 dpi in redundant mode
- 2.4 picoliter native drop size
- RediJet® continuous ink recirculation
- VersaDrop™ multi-level grayscale plus multi-drop and fixed drop sizes in binary mode
- Scalable, narrow to wide printbars possible
- Robust design and high speed operation
- Multi-ink compatibility

The Samba G3L printhead technology delivers the breakthrough quality, speed and scalability required for variable width, single pass production inkjet printing and materials deposition applications. Based on state-of-the-art silicon micro-electro-mechanical systems manufacturing (Si-MEMS), the Samba G3L printhead is the most advanced technology available today.

Fujifilm Dimatix continues building upon its long history of innovation with the Samba G3L printhead delivering excellent dot placement accuracy, channel-to-channel uniformity, low cross talk and high frequency/high productivity. With 2048 individually addressable nozzles, the Samba G3L printhead has 1,200 dpi native resolution (600 dpi in redundant mode) and a native ink drop size of 2.4 picoliters (pl) delivering outstanding print quality that can jet a wide range of fluids including UV curable and aqueous inks.

Scalability is the key to the Samba G3L technology, having a single head print width of 43 mm (1.7 inches) it can be easily configured to make a print bar of any width, making it an ideal choice for print application such as high quality commercial print, labels, packaging and textiles. The unique parallelogram design of the Samba nozzle plate enables simplified printhead stitching, resulting in a very narrow, 1200 dpi printbar.

Designed in collaboration with Fujifilm Graphic Systems, the highly reliable design has been field proven having been used, until now, exclusively in the Fujifilm Jet Press digital color press. The new Samba G3L printhead design is now available for use by original equipment manufacturers and integrators to incorporate into new designs with the confidence of using a highly reliable and technologically advanced printhead.

RediJet® Recirculation

Through continuous ink recirculation directly behind the nozzle and several innovations including specialized nozzle plate geometry and coatings, and waveforms tailored to specific fluids, RediJet unlocks the full productive potential of the Samba G3L printhead. RediJet allows the printheads to be quickly and easily primed, resulting in faster time to print, minimal ink waste and greater reliability.

VersaDrop™

Fujifilm Dimatix' breakthrough VersaDrop allows multiple fixed drop sizes in binary mode and grayscale capability from one printhead, with no loss to productivity. In binary operating mode, the Samba G3L printhead is designed to eject adjustable drop sizes from 2.4 up to 13.2 picoliters and can support grayscale levels as defined by the user's control electronics.



CENTER *of* EXCELLENCE
for ENGINEERING BIOLOGY



Pilot Project Proposal

(Not to exceed two pages)

Name of Project:

Proposer and Contact Information:

Background:

Technical Idea:

Utility:

“Fit” For GP-write

Submit to:

info@engineeringbiologycenter.org

$P \stackrel{?}{=} NP$

Scott Aaronson*

Abstract

In 1955, John Nash sent a remarkable letter to the National Security Agency, in which—seeking to build theoretical foundations for cryptography—he all but formulated what today we call the $P \stackrel{?}{=} NP$ problem, considered one of the great open problems of science. Here I survey the status of this problem in 2017, for a broad audience of mathematicians, scientists, and engineers. I offer a personal perspective on what it’s about, why it’s important, why it’s reasonable to conjecture that $P \neq NP$ is both true and provable, why proving it is so hard, the landscape of related problems, and crucially, what progress has been made in the last half-century toward solving those problems. The discussion of progress includes diagonalization and circuit lower bounds; the relativization, algebrization, and natural proofs barriers; and the recent works of Ryan Williams and Ketan Mulmuley, which (in different ways) hint at a duality between impossibility proofs and algorithms.

Contents

1	Introduction	3
1.1	The Importance of $P \stackrel{?}{=} NP$	4
1.2	Objections to $P \stackrel{?}{=} NP$	6
1.2.1	The Asymptotic Objection	6
1.2.2	The Polynomial-Time Objection	7
1.2.3	The Kitchen-Sink Objection	7
1.2.4	The Mathematical Snobbery Objection	8
1.2.5	The Sour Grapes Objection	8
1.2.6	The Obviousness Objection	9
1.2.7	The Constructivity Objection	9
1.3	Further Reading	10
2	Formalizing $P \stackrel{?}{=} NP$ and Central Related Concepts	10
2.1	NP-Completeness	12
2.2	Other Core Concepts	16
2.2.1	Search, Decision, and Optimization	16
2.2.2	The Twilight Zone: Between P and NP-complete	17
2.2.3	coNP and the Polynomial Hierarchy	17

*University of Texas at Austin. Email: aaronson@cs.utexas.edu. Supported by a Vannevar Bush / NSSEFF Fellowship from the US Department of Defense. Much of this work was done while the author was supported by an NSF Alan T. Waterman award.

A Hippocampal Marker of Recollection Memory Ability among Healthy Young Adults: Contributions of Posterior and Anterior Segments

Jordan Poppenk^{1,2,3,*} and Morris Moscovitch^{1,2}

¹Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada

²Rotman Research Institute, Baycrest, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada

³Princeton Neuroscience Institute, Princeton University, Green Hall, Princeton, NJ 08540, USA

*Correspondence: jpoppenk@princeton.edu

DOI 10.1016/j.neuron.2011.10.014

SUMMARY

The hippocampus is known to support recollection memory, but the relation between its structure and recollection in healthy adults has not been established. Here we show that the hippocampus (including subiculum, DG, and CA1–CA4), when separated into posterior and anterior segments, can reliably predict recollection in healthy young adults. Better memory was associated with larger posterior and smaller anterior segments, as evaluated relative to the uncus apex. Overall hippocampal volume, however, did not predict memory. This pattern was confirmed in four separate data sets from different studies and laboratories. The relationship between the posterior hippocampus and memory was mediated by the structure's functional connectivity with a neocortical network identified during a postencoding resting-state scan. The relationship was also weakest in an experiment involving no appreciable study-test interval. These findings suggest that enhanced posterior-hippocampal post-encoding processes may account for the memory benefit associated with larger posterior hippocampi.

INTRODUCTION

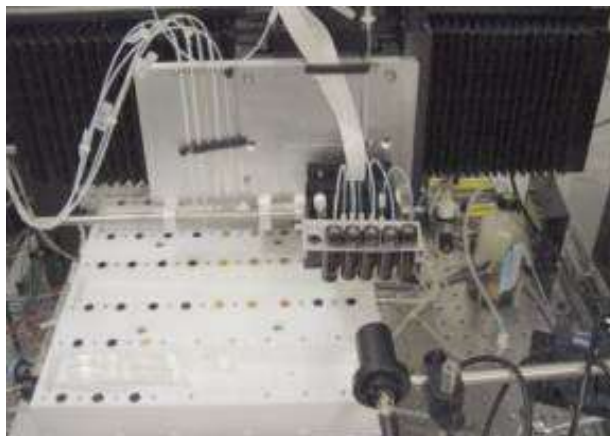
Efforts to explain individual differences in human memory using brain anatomy have centered on the hippocampus (defined here as the subiculum, dentate gyrus, and cornu ammonis regions, including fields CA1–CA4). This structure has known functional importance for the encoding, storage, and, many argue, retrieval of recollection memory (RM), a form of memory involving a detailed reexperiencing of individual episodes that is characterized by retrieval of an item and its context (Moscovitch et al., 2005; Eichenbaum et al., 2007). Indeed, among dementia and amnesic patients, smaller hippocampi predict worse memory (Van Petten, 2004), just as hippocampal volume and memory decline together with age in older adults (Raz, 2000). However, among healthy young adults, published correlations between hippocampal volume and memory approach a normal distribution of about zero (Van Petten, 2004). Nonetheless,

Maguire et al. (2000) have shown that extensive spatial memory acquisition leads to enlargement of the posterior hippocampus at the expense of anterior hippocampal volume (pHPC and aHPC; dorsal ventral in nonprimate mammals). This suggests that the crucial predictor of individual differences in recollection may not be overall hippocampal volume (HPC) but the separate contributions of pHPC and aHPC segments, a hypothesis we test in this paper.

This hypothesis is supported by neuroanatomical and functional evidence that pHPC and aHPC have dissociable properties. For instance, in primates, the segments connect with different bands of the entorhinal cortex, a key link between the hippocampus and cerebral cortex (Fanselow and Dong, 2010). Also, hippocampal connections with the retrosplenial cortex and the mammillary bodies arise primarily from pHPC in primates (Aggleton et al., 2005; Kobayashi and Amaral, 2003), and the link between these structures and the hippocampus has been shown to be important specifically for RM (Vann et al., 2009). This anatomical link is suggestive of favorable conditions for recollection in pHPC, and consistent with this notion, damage to the dorsal, but not ventral, portion of the rodent hippocampus impairs Morris water maze performance (Moser and Moser, 1998). In humans, Scoville and Milner (1957) and Penfield and Milner (1958) noted that global amnesia in patients with medial temporal lobe resection was evident only when pHPC was affected bilaterally, and Smith and Milner (1981) observed a similar drop in performance on tests of object-location memory following right pHPC lesions in patients with unilateral temporal lobectomy, although all of these patient observations were confounded with the amount of resected tissue. More recently, high-resolution neuroimaging and single-unit recordings have hinted at greater pHPC involvement in spatial and verbal memory (Maguire et al., 2000; Ludowig et al., 2008). Finally, pHPC in particular has been found to be sensitive to spatial information, which is thought to play a role in RM (Ryan et al., 2010).

To the extent that pHPC is more closely associated with RM and that pHPC and aHPC volumes trade off against one another, relatively large pHPC volumes—and conversely, small aHPC volumes—might be expected to predict enhanced RM, even in the absence of any effect of HPC. To test this hypothesis, we collected anatomical magnetic resonance imaging (MRI) and functional MRI (fMRI) scans from healthy young adults, derived various measures of hippocampal volume and connectivity (see Table S1 available online), and examined their correlations

Assembly manual for the POSaM: THE ISB Piezoelectric Oligonucleotide Synthesizer and Microarrayer



Version 1.2
28 May 2004
The Institute for Systems Biology © 2004
1441 North 34th Street
Seattle, WA 98103-8904

The Hood Laboratory “Beta Group”

Table of Contents

User manual part II: running the POSaM Platform.....	3
Overview.....	3
Slide Surface Preparation.....	3
In-House Epoxysilane Slide Preparation.....	3
Materials:.....	4
Methods:.....	5
Commercial Slide Alternatives.....	7
Reagent Preparation.....	7
Overview:.....	7
Solvent Preparation.....	7
Materials:.....	7
Methods:.....	8
Phosphoramidites and Tetrazole.....	9
Materials:.....	10
Methods:.....	10
Base Deprotection Solution.....	11
Materials:.....	11
Methods:.....	11
Acetonitrile Wash Solvent.....	12
Materials:.....	12
Methods:.....	12
Oxidizer and Deprotection Acid.....	13
Materials:.....	13
Methods:.....	13
Arrayer Operation.....	13
Overview.....	13
Loading the Phosphoramidites and Tetrazole.....	14
Materials:.....	14
Method:.....	15
Preparing and Loading the Substrates.....	16
Materials:.....	16
Method:.....	16
Priming the Bulk Reagents.....	17
Method:.....	17
Testing and Clearing the Nozzles.....	18
Method:.....	18
Drying Down the Enclosure.....	18
Materials:.....	19
Method:.....	19
Pre-Washing the Slides.....	19
Method:.....	19
Initiating Synthesis.....	19

On Stake and Consensus

Andrew Poelstra
2015-03-22
fe81626

This work is released into the public domain.

Note. This document was significantly revised in March 2015. The old version can be found at <https://download.wpsoftware.net/bitcoin/old-pos.pdf>.

1 Introduction

In 2009, Satoshi Nakamoto introduced the Bitcoin cryptocurrency[Nak09], an online currency system which allowed peer-to-peer transfer of digital tokens. To ensure a consistent view of token ownership, Nakamoto used a public ledger which can be replicated and validated by all network participants. To avoid a single point of failure, this ledger is authenticated using a *dynamic membership multiparty signature (DMMS)*[BCD⁺14] consisting of an expensive (but cheaply verifiable) computation done on the entire ledger history every “heartbeat”.

Unlike a traditional digital signature, there is no notion of “forgeability” for a DMMS. Instead, every DMMS is costly to produce (in Bitcoin, by requiring a large energy expenditure) and rewarded by introduction of new coins on the ledger. Since these coins are only useful if others recognize them, participants are incentivized to extend one “true ledger” rather than attempting to create their own version of history¹.

Because Bitcoin’s DMMS is computationally, and therefore thermodynamically[Poe14a], very expensive, alternatives have been proposed which seek to be economically and environmentally more efficient. One popular alternative, *proof-of-stake*, is frequently proposed as a mechanism for a cheap distributed consensus. As argued by the author[Poe14b] in 2014, this is simply not workable, but nonetheless the idea continues to arise in various forms. Meanwhile, the author’s argument is commonly asserted on various forums to be “debunked” or “wrong”, despite the author having never been made aware of any workable counterexamples or mistakes. This, combined with (correct) accusations that the paper is obtuse and unreadable, demonstrate that its exposition leaves much to be desired. Although this author is not aware of any inaccuracies in his former work, he has taken the opportunity to continue and elaborate his argument more formally.

Further, there has been significant progress in scientific understanding of Bitcoin’s consensus[MLJ14, BMC⁺15] which was not available when the original paper was written.

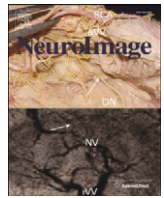
This paper aims to be an updated version of the author’s original paper, which gives more explication on the problem Bitcoin solves, why it makes the design decisions that it does, and why

¹To ensure that the “true ledger” is visible to all participants, we require a synchronous network such that all (valid) data reaches all participants in some characteristic time λ , and that the network heartbeat time is much larger than λ . Without a synchronous network, distributed consensus is much harder. (There is an oft-cited impossibility result for distributed consensus using deterministic algorithms [FLP85]; this is easy to evade simply by using probabilistic algorithms, and therefore doesn’t say much about the difficulty of designing consensus systems. Thanks to Dominic Williams for pointing this out; an earlier version of this document erroneously quoted this result as blocking any distributed consensus in an asynchronous network.)



Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Review

A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading

Cathy J. Price

Wellcome Trust Centre for Neuroimaging, UCL, London, WC1N 3BG, UK

ARTICLE INFO

Article history:

Accepted 30 April 2012

Available online xxxx

Keywords:

PET

fMRI

Language

Auditory speech

Reading

Comprehension

Speech production

ABSTRACT

The anatomy of language has been investigated with PET or fMRI for more than 20 years. Here I attempt to provide an overview of the brain areas associated with heard speech, speech production and reading. The conclusions of many hundreds of studies were considered, grouped according to the type of processing, and reported in the order that they were published. Many findings have been replicated time and time again leading to some consistent and undisputable conclusions. These are summarised in an anatomical model that indicates the location of the language areas and the most consistent functions that have been assigned to them. The implications for cognitive models of language processing are also considered. In particular, a distinction can be made between processes that are localized to specific structures (e.g. sensory and motor processing) and processes where specialisation arises in the distributed pattern of activation over many different areas that each participate in multiple functions. For example, phonological processing of heard speech is supported by the functional integration of auditory processing and articulation; and orthographic processing is supported by the functional integration of visual processing, articulation and semantics. Future studies will undoubtedly be able to improve the spatial precision with which functional regions can be dissociated but the greatest challenge will be to understand how different brain regions interact with one another in their attempts to comprehend and produce language.

© 2012 Elsevier Inc. All rights reserved.

Contents

Introduction	0
Early PET studies	0
Early fMRI studies	0
Methods	0
Inclusion criteria for the review	0
Organization of the conclusions	0
Results	0
Discussion	0
Auditory processing that is common to speech and nonspeech sounds	0
Auditory processing of speech and nonspeech. Time era: 1992–1996	0
Auditory processing of speech and nonspeech. Time era: 1997–2001	0
Auditory processing of speech and nonspeech. Time era: 2002–2006	0
Auditory processing of speech and nonspeech. Time era: 2007–2011	0
Auditory processing of speech and nonspeech: 20 year summary	0
Speech selective auditory responses (= phonological processing)	0
Speech selective auditory responses. Time era: 1992–1996	0
Speech selective auditory responses. Time era: 1997–2001	0
Speech selective auditory responses. Time era: 2002–2006	0
Speech selective auditory responses. Time era: 2007–2011	0
Speech selective auditory responses: 20 year summary	0
Speech comprehension	0
Speech comprehension. Time era: 1992–1996	0
Speech comprehension. Time era: 1997–2001	0

E-mail address: c.j.price@ucl.ac.uk.

1053-8119/\$ – see front matter © 2012 Elsevier Inc. All rights reserved.

doi:[10.1016/j.neuroimage.2012.04.062](https://doi.org/10.1016/j.neuroimage.2012.04.062)

Please cite this article as: Price, C.J., A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading, NeuroImage (2012), doi:[10.1016/j.neuroimage.2012.04.062](https://doi.org/10.1016/j.neuroimage.2012.04.062)

Print your atomic force microscope

Ferdinand Kühner,^{a)} Robert A. Lugmaier, Steffen Mihatsch, and Hermann E. Gaub

Applied Physics and Center for Nano Science, Ludwig-Maximilians Universität München, München 80799, Germany

(Received 30 April 2007; accepted 29 May 2007; published online 9 July 2007)

Progress in scanning probe microscopy profited from a flourishing multitude of new instrument designs, which lead to novel imaging modes and as a consequence to innovative microscopes. Often these designs were hampered by the restrictions, which conventional milling techniques impose. Modern rapid prototyping techniques, where layer by layer is added to the growing piece either by light driven polymerization or by three-dimensional printing techniques, overcome this constraint, allowing highly concave or even embedded and entangled structures. We have employed such a technique to manufacture an atomic force microscopy (AFM) head, and we compared its performance with a copy milled from aluminum. We tested both AFM heads for single molecule force spectroscopy applications and found little to no difference in the signal-to-noise ratio as well as in the thermal drift. The lower E modulus seems to be compensated by higher damping making this material well suited for low noise and low drift applications. Printing an AFM thus offers unparalleled freedom in the design and the rapid production of application-tailored custom instruments. © 2007 American Institute of Physics. [DOI: 10.1063/1.2751099]

INTRODUCTION

A surprisingly wide spectrum of techniques evolved from the basic concept introduced by Binnig and Rohrer in 1986:^{1,2} scanning a probe with Angstrom precision on a sample while recording a local interaction. Boosted was this boom by the relative ease at which new ideas could be implemented in new microscopes.^{3–10} Nevertheless only a very limited range of design principles could be implemented by conventional manufacturing techniques. Here we explore and evaluate new manufacturing techniques, which not only allow entirely new topologies but also support fast and inexpensive prototyping strategies.

The most versatile implementation of the above mentioned scanned probe principle is the atomic force microscope (AFM). Due to its high resolution in force, position, and time as well as its broad range of possible ambient conditions, it opened the widest spectrum of applications. The essential part of an AFM is a sharp tip on a cantilever spring. This cantilever serves as a sensor and reacts to local interactions between the tip and the surface. Its deflection is detected by a laser beam, which is focused on the cantilever and reflected onto a segmented photodiode.¹¹

By scanning the AFM only in the z direction, vertical to the surface, it can be operated as a single molecule force spectroscopy spectrometer. While approaching and retracting the cantilever from the surface, the interaction force between the tip and the surface is recorded as a function of the separation distance. Single molecule force spectroscopy has become a very important method to analyze intra- and intermolecular interactions of biological molecules.^{7,12} The main intent is to identify the fundamental interactions in single

molecules. In recent years many intrinsic properties of single molecules were determined with this approach. For example, the entropic, enthalpic, and conformational¹³ properties of single polymers such as dsDNA,^{14–16} polysaccharides,^{17–19} and other polymers^{20–22} have been evaluated. Furthermore, the unfolding and folding pathways of proteins such as titin,^{23,24} green fluorescent protein (GFP),^{25–27} and others^{28,29} have been analyzed. Beyond intramolecular forces also intermolecular forces of different receptor-ligand systems^{30–32} have been measured, e.g., the interaction between antibody-antigen,³³ protein-dsDNA,³⁴ protein-aptamer,³⁵ and ssDNA-ssDNA,^{36–38} and also of covalent bonds.³⁹

RAPID PROTOTYPING

New techniques allow the production of the parts of an AFM head in a very fast but nevertheless very precise way. Three-dimensional (3D) printing techniques can build up complex three-dimensional structures by a layer-by-layer method. The four most favorable rapid prototyping methods are the fused deposition modeling (FDM), stereolithography (SL), laminated object manufacturing (LOM) 3D printing, and selective laser sintering (SLS).

Selective laser sintering⁴⁰ is an additive rapid manufacturing technique, which uses a high power laser to selectively fuse polymers, metal, or ceramic powders. The laser selectively smelts powdered material by tracing cross sections of the object on the surface of a powder bed with an accuracy of 100 μm . After finishing the cross section, the powder bed is lowered. Now a new layer of material is applied on top, and the process is repeated until the part is completed (see Fig. 1).

Compared with other rapid manufacturing methods, SLS can produce parts from a relatively wide range of commercially available powder materials, including polymers (ny-

^{a)} Author to whom correspondence should be addressed; electronic mail: ferdinand.kuehner@physik.uni-muenchen.de



Mini Review

Programmable Genome Editing Tools and their Regulation for Efficient Genome Engineering

Tuhin Kumar Guha, Alvan Wai, Georg Hausner *

Department of Microbiology, University of Manitoba, Winnipeg, Manitoba R3T2N2, Canada

ARTICLE INFO

Article history:

Received 12 October 2016

Received in revised form 24 December 2016

Accepted 27 December 2016

Available online 12 January 2017

Keywords:

Meganuclease

TALEN

Zinc finger nuclease

CRISPR/Cas9

Regulatory switch

Hammerhead ribozyme

ABSTRACT

Targeted genome editing has become a powerful genetic tool for studying gene function or for modifying genomes by correcting defective genes or introducing genes. A variety of reagents have been developed in recent years that can generate targeted double-stranded DNA cuts which can be repaired by the error-prone, non-homologous end joining repair system or via the homologous recombination-based double-strand break repair pathway provided a suitable template is available. These genome editing reagents require components for recognizing a specific DNA target site and for DNA-cleavage that generates the double-stranded break. In order to reduce potential toxic effects of genome editing reagents, it might be desirable to control the in vitro or in vivo activity of these reagents by incorporating regulatory switches that can reduce off-target activities and/or allow for these reagents to be turned on or off. This review will outline the various genome editing tools that are currently available and describe the strategies that have so far been employed for regulating these editing reagents. In addition, this review will examine potential regulatory switches/strategies that can be employed in the future in order to provide temporal control for these reagents.

© 2017 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction	146
2. Genome Editing Reagents	147
3. Current Regulatable DNA-cutting Enzymes	152
4. Alternative Strategies for Developing Regulatable Genome Editing Reagents	155
4.1. The Utility of Hammerhead Ribozymes and Engineered Variants	155
4.2. Utility of Riboswitches and Allosteric Ribozymes	156
5. Conclusion	156
Competing interests	156
Acknowledgments	156
References	156

1. Introduction

One of the challenges in biotechnology has been developing efficient and reliable ways to make targeted changes within the genome of cells. Traditional approaches of mutagenesis utilizing chemical agents or transposons can require extensive screening in order to recover desired mutations [1–6]. Genome editing strategies using double-stranded (ds) DNA viral vectors in differentiated human cells and RNA interference (RNAi) mediated targeted gene knockdown approaches also have

some pitfalls [7–10]. For example, the protein composition of the viral capsid can be potentially immunogenic. Moreover, abnormal gene expression along with insertional mutagenesis may be triggered if there are random mutations in the viral sequences. On the other hand, the use of exogenously introduced dsRNA in RNAi technology can disrupt the “homeostasis” of the cellular machinery involved in gene silencing. Currently, the most popular genome engineering techniques apply DNA-cutting enzymes/complexes that generate targeted double-strand cuts [11–13], which are repaired by the host cells by either the error-prone, non-homologous end joining repair system (NHEJ), or the homologous recombination-based double-strand break repair pathway (HDR) [14–18]. The most frequent application of these

* Corresponding author.

E-mail address: hausnerg@cc.umanitoba.ca (G. Hausner).

Progress in sequencing deoxyribonucleic acid with an atomic force microscope

H. G. Hansma,^{a)} A. L. Weisenhorn, S. A. C. Gould, R. L. Sinsheimer,^{b)}
H. E. Gaub,^{c)} G. D. Stucky,^{d)} C. M. Zaremba,^{d)} and P. K. Hansma
Department of Physics, University of California, Santa Barbara, California 93106

(Received 24 July 1990; accepted 18 September 1990)

Atomic force microscope (AFM) images of single-stranded deoxyribonucleic acid (DNA) showing nucleotide resolution have been obtained using two sample preparation methods: DNA covalently attached to a polymerized lipid monolayer and then imaged under water and DNA on mica rinsed with a $\text{Ba}(\text{NO}_3)_2$ solution and then imaged under ethanol. Some of the bases were identified in the AFM image of DNA on polymerized lipid. Current problems in sequencing DNA with the AFM include movement of the DNA during imaging and the difficulty of reproducing experiments.

Imaging deoxyribonucleic acid (DNA) with a scanning probe microscope has captured the imagination of researchers in the field starting with Binnig, an inventor of the atomic force microscope¹ (AFM) and scanning tunneling microscope² (STM), who first imaged DNA with the STM over 6 years ago.³ In the last 2 years, many groups have published STM pictures of DNA, e.g., Refs. 4–9.

Our group has been working for over a year on imaging DNA with the AFM.^{10–12} Since the AFM can image non-conducting surfaces, it shows great promise for imaging non-conducting biological macromolecules and cells.

Why are so many groups interested in imaging DNA with scanning probe microscopes? One major goal of this research is to assist in sequencing the DNA of the human genome¹³ in order to bring people relief from some of the physical and mental disease that is of genetic origin. Rough estimates suggest that sequencing DNA with a scanning probe microscope would be at least 100 times faster than sequencing DNA with the current technology based on gel electrophoresis. Scanning probe microscopes also offer the possibility for doing much other exciting DNA research in areas such as sequence-dependent binding of regulatory proteins, sequence-dependent kinking, and B- to Z-DNA transitions.

What resolution has been obtained with DNA in scanning probe microscopes? At least three groups have resolved nucleotides of DNA with the STM^{8,9} or AFM.¹⁴ Dunlap and Bustamante⁷ have resolved the two rings of the adenine base in poly A. Nucleotides are subunits of the DNA polymer and consist of a base covalently bonded to a sugar, in turn covalently bonded to a phosphate group. The four bases, A (adenine), G (guanine), C (cytosine), and T (thymine), are arranged in different sequences in the human genome, which is 3 billion bases long.

The next step after resolving the nucleotides of DNA is to identify them in the AFM image. We have recently made some progress in this. A synthetic single-stranded DNA 25 bases long (25-mer) was covalently bonded with EDC [1-ethyl-3-(3-Dimethylaminopropyl)-carbodiimide]¹⁵ to a UV-polymerized monolayer of the fatty-acid 10,12 pentacosadiynoic acid.¹⁶ The polymerized monolayer was supported on a monolayer of cadmium arachidate on mica.¹² This sample was imaged under water in the AFM, as described previously.¹² Figure 1 shows an AFM image of this 25-mer

DNA on the polymerized fatty-acid monolayer. There appear to be two different DNA ends in this figure, perhaps belonging to a single molecule of 25-mer; they are labeled No. 1 and No. 25 in Fig. 1.

The 25-mer shown in Fig. 1 was a custom-made DNA designed to facilitate the sequencing of DNA with the AFM. It was synthesized with five aliphatic amino groups inserted along the chain using Amino-modifier II from Clontech, Palo Alto, CA. These aliphatic amino groups were reacted with fluorescein isothiocyanate to produce a 25-mer with five large fluorescein groups in five of the 25 base positions. The AFM image of this DNA in Fig. 1 shows some bands that are longer than others. One might expect these long bands to be fluorescein, since fluorescein is larger than the DNA bases. Studying the image in Fig. 1 closely, one of us predicted that there was a fluorescein 3 or 4 bases from one end of the molecule, in the location identified as No. 22. The answer was that there is indeed a fluorescein in position No.

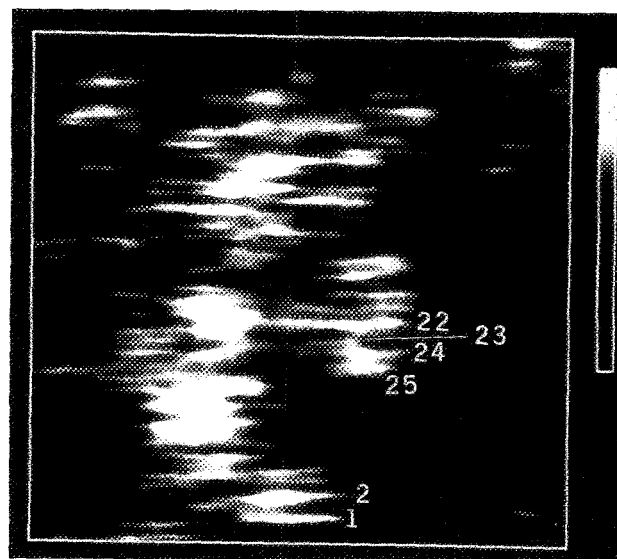


FIG. 1. AFM image of a single-stranded 25-mer DNA covalently attached to a polymerized monolayer of 10,12-pentacosadiynoic acid, imaged under water. Numbers indicate bands identified in the AFM image as being bases at the beginning and end of the 25-mer. Image size is 9.5 nm \times 9.5 nm.

THE WORLD ANTI-DOPING CODE
**INTERNATIONAL
STANDARD**



PROHIBITED LIST

JANUARY 2018



**WORLD
ANTI-DOPING
AGENCY**
play true

The official text of the *Prohibited List* shall be maintained by WADA and shall be published in English and French.
In the event of any conflict between the English and French versions, the English version shall prevail.

This List shall come into effect on 1 January 2018

A Simple and Convenient Synthesis of Pseudoephedrine From *N*-Methylamphetamine

O. Hai,^{*a} and I. B. Hakkenshit^{a,b}

Received 12th December 2011, Accepted 21st January 2012

First published on the web 23rd February 2012

DOI: 10.1039/c1ay00000f

A novel and straightforward synthesis of pseudoephedrine from readily available *N*-methylamphetamine is presented. This practical synthesis is expected to be a disruptive technology replacing the need to find an open pharmacy.

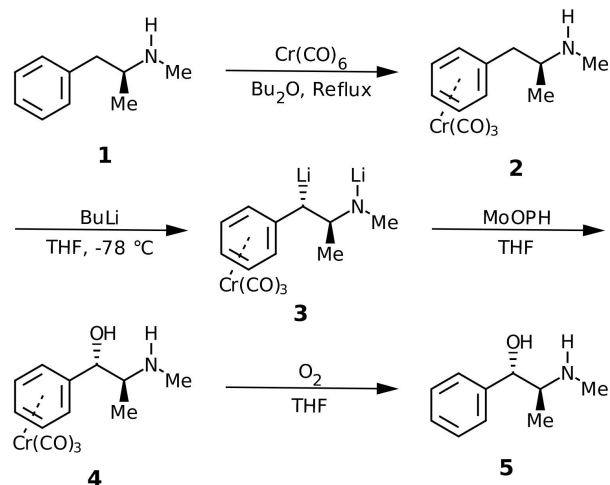
Pseudoephedrine, active ingredient of Sudafed®, has long been the most popular nasal decongestant in the United States due to its effectiveness and relatively mild side effects [1]. In recent years it has become increasingly difficult to obtain pseudoephedrine in many states because of its use as a precursor for the illegal drug *N*-methylamphetamine (also known under various names including crystal meth, meth, ice, etc.) [1,2]. While in the past many stores were able to sell pseudoephedrine, new laws in the United States have restricted sales to pharmacies, with the medicine kept behind the counter. The pharmacies require signatures and examination of government issued ID in order to purchase pseudoephedrine. Because the hours of availability of such pharmacies are often limited, it would be of great interest to have a simple synthesis of pseudoephedrine from reagents which can be more readily procured.

A quick search of several neighborhoods of the United States revealed that while pseudoephedrine is difficult to obtain, *N*-methylamphetamine can be procured at almost any time on short notice and in quantities sufficient for synthesis of useful amounts of the desired material. Moreover, according to government maintained statistics, *N*-methylamphetamine is becoming an increasingly attractive starting material for pseudoephedrine, as the availability of *N*-methylamphetamine has remained high while prices have dropped and purity has increased [2]. We present here a convenient series of transformations using reagents which can be found in most well stocked organic chemistry laboratories to produce pseudoephedrine from *N*-methylamphetamine.

While *N*-methylamphetamine itself is a powerful decongestant, it is less desirable in a medical setting because of its severe side effects and addictive properties [3]. Such side effects may include insomnia, agitation, irritability, dry mouth, sweating, and heart palpitations. Other side effects may include violent urges or, similarly, the urge to be successful in business or finance.

In our search for sources of *N*-methylamphetamine we have found that, similar to research grade chemicals purchased from the major chemical supply houses, the purity of the reagent varies greatly between suppliers and even between batches despite the above cited overall increase in purity. Unfortunately, and again similar to suppliers of fine

chemicals, relative cost is not strongly correlated to sample quality. We therefore found it necessary to purify the starting material before use. This may be accomplished by precipitating the amphetamine from isopropanol with HCl followed by deprotonation with sodium hydroxide and extraction into chloroform, which after removal of the solvent in vacuo yields *N*-methylamphetamine, **1**. In the majority of the samples obtained for this study, **1** was greater than 95% enantiomerically pure, with the *S* enantiomer being the major isomer present. This is consistent with reduction of commercially available ephedrine or pseudoephedrine as the origin of the casually procured material used in this study.



Scheme 1 Synthesis of pseudoephedrine from *N*-methylamphetamine

The synthetic procedure is shown in scheme 1. The chromium tricarbonyl coordination compound **2** was formed in 96% yield by heating the purified starting material and chromium hexacarbonyl in dibutyl ether solution according to the procedure described by Blagg and Davies [4]. While in the cited work a single equivalent of tBuLi was used to deprotonate a similar chromium complex, in the case of **2** two equivalents of tBuLi were required due to the presence of the relatively acidic amine proton. After successful formation of the dianion **3**, which was not isolated but taken on immediately in the same reaction vessel, a single hydroxy group was introduced at the alpha position by addition of one equivalent of oxodiperoxymolybdenum(pyridine)(HMPA), commonly known as MoOPH. The series of transformations producing **4** from **2** went in 87% yield overall. Chromium

Quadratic Span Programs and Succinct NIZKs without PCPs

Rosario Gennaro* Craig Gentry† Bryan Parno‡ Mariana Raykova§

Abstract

We introduce a new characterization of the NP complexity class, called *Quadratic Span Programs* (QSPs), which is a natural extension of *span programs* defined by Karchmer and Wigderson. Our main motivation is the construction of succinct arguments of NP-statements that are quick to construct and verify. QSPs seem well-suited for this task, perhaps even better than Probabilistically Checkable Proofs (PCPs).

In 2010, Groth constructed a NIZK argument in the common reference string (CRS) model for Circuit-SAT consisting of only 42 elements in a bilinear group. Interestingly, his argument does not (explicitly) use PCPs. But his scheme has some disadvantages – namely, the CRS size and prover computation are both *quadratic* in the circuit size. In 2011, Lipmaa reduced the CRS size to quasi-linear, but with prover computation still quadratic.

Using QSPs we construct a NIZK argument in the CRS model for Circuit-SAT consisting of just *7 group elements*. The CRS size is *linear* in the circuit size, and prover computation is *quasi-linear*, making our scheme seemingly quite practical. (The prover only needs to do a linear number of group operations; the quasi-linear computation is a multipoint evaluation and interpolation.)

Our results are complementary to those of Valiant (TCC 2008) and Bitansky et al. (2012), who use “bootstrapping” (recursive composition) of arguments to reduce CRS size and prover and verifier computation. QSPs also provide a crisp mathematical abstraction of some of the techniques underlying Groth’s and Lipmaa’s constructions.

*IBM T.J.Watson Research Center. rosario@us.ibm.com

†IBM T.J.Watson Research Center. cbgentry@us.ibm.com

‡Microsoft Research. parno@microsoft.com

§Columbia University. mariana@cs.columbia.edu

Quasilinear-Size Zero Knowledge from Linear-Algebraic PCPs

Eli Ben-Sasson

eli@cs.technion.ac.il

Technion

Alessandro Chiesa

alexch@berkeley.edu

UC Berkeley

Ariel Gabizon

ariel.gabizon@gmail.com

Technion

Madars Virza

madars@mit.edu

MIT

April 28, 2016

Abstract

The seminal result that every language having an interactive proof also has a zero-knowledge interactive proof assumes the existence of one-way functions. Ostrovsky and Wigderson (ISTCS 1993) proved that this assumption is necessary: if one-way functions do not exist, then only languages in BPP have zero-knowledge interactive proofs.

Ben-Or et al. (STOC 1988) proved that, nevertheless, every language having a multi-prover interactive proof also has a zero-knowledge multi-prover interactive proof, unconditionally. Their work led to, among many other things, a line of work studying zero knowledge without intractability assumptions. In this line of work, Kilian, Petrank, and Tardos (STOC 1997) defined and constructed zero-knowledge probabilistically checkable proofs (PCPs).

While PCPs with quasilinear-size proof length, but without zero knowledge, are known, no such result is known for zero knowledge PCPs. In this work, we show how to construct “2-round” PCPs that are zero knowledge and of length $\tilde{O}(K)$ where K is the number of queries made by a malicious polynomial time verifier. Previous solutions required PCPs of length at least K^6 to maintain zero knowledge. In this model, which we call *duplex PCP* (DPCP), the verifier first receives an oracle string from the prover, then replies with a message, and then receives another oracle string from the prover; a malicious verifier can make up to K queries in total to both oracles.

Deviating from previous works, our constructions do not invoke the PCP Theorem as a blackbox but instead rely on certain algebraic properties of a specific family of PCPs. We show that if the PCP has a certain linear algebraic structure — which many central constructions can be shown to possess, including [BFLS91, ALMSS98, BS08] — we can add the zero knowledge property at virtually no cost (up to additive lower order terms) while introducing only minor modifications in the algorithms of the prover and verifier. We believe that our linear-algebraic characterization of PCPs may be of independent interest, as it gives a simplified way to view previous well-studied PCP constructions.

Keywords: zero knowledge, probabilistically-checkable proofs

Reading the primary structure of a protein with 0.07 nm³ resolution using a subnanometre-diameter pore

Eamonn Kennedy^{1†}, Zhuxin Dong^{1†}, Clare Tennant² and Gregory Timp^{3*}

The primary structure of a protein consists of a sequence of amino acids and is a key factor in determining how a protein folds and functions. However, conventional methods for sequencing proteins, such as mass spectrometry and Edman degradation, suffer from short reads and lack sensitivity, so alternative approaches are sought. Here, we show that a subnanometre-diameter pore, sputtered through a thin silicon nitride membrane, can be used to detect the primary structure of a denatured protein molecule. When a denatured protein immersed in electrolyte is driven through the pore by an electric field, measurements of a blockade in the current reveal nearly regular fluctuations, the number of which coincides with the number of residues in the protein. Furthermore, the amplitudes of the fluctuations are highly correlated with the volumes that are occluded by quadromers (four residues) in the primary structure. Each fluctuation, therefore, represents a read of a quadromer. Scrutiny of the fluctuations reveals that the subnanometre pore is sensitive enough to read the occluded volume that is related to post-translational modifications of a single residue, measuring volume differences of ~0.07 nm³, but it is not sensitive enough to discriminate between the volumes of all twenty amino acids.

Proteins are the machinery that makes biology work. The 3D structure of a protein—how a protein binds to itself—determines its function. One of the key factors that dictates the protein's 3D structure is the primary structure, which consists of a linear sequence of amino acids (AAs) linked by peptide bonds. Thus, sequencing a protein is essential to proteomics, the next step beyond genomics, in the analysis of biology^{1,2}. However, the two methods commonly used for sequencing proteins, mass spectrometry and Edman degradation, are subject to limitations. Edman degradation does not work if the N-terminal AA is chemically modified or buried in the folded protein and it only sequences peptides that are about 30–50 residues long. On the other hand, mass spectrometry can sequence a protein of any size, but it does not provide information on the complete sequence. As mass spectrometry relies on enzymatic digestion, it becomes computationally demanding to reassemble the sequence from its constituents as the size increases. Moreover, mass spectrometry lacks sensitivity³.

Sequencing a single protein molecule with a nanopore is among the alternatives that have been proffered^{4–17}. Single molecule sensitivity develops when a protein that is immersed in an electrolyte is impelled through the nanopore by an electric field, producing a blockade in the current that depends on the occluded volume and informs on the chemical constituency in the pore. So far nanopores have been used to detect and analyse proteins, but not to sequence them. This is because the higher-order (secondary, tertiary, quaternary) structure of the protein confounds the interpretation of the blockade current. Moreover, the charge distribution along the native protein is not uniform, which frustrates the systematic control of the translocation kinetics by the electric field in the pore¹⁷. Recent developments have focused mainly on the use of biological pores for sequencing DNA^{18–21}. Nanopore sequencing of DNA is distinguished from other methodologies by kilobase-long

reads of single molecules²⁰. Although the read fidelity is low^{20,21}—for example, the Oxford MinION v7 chips show only about a 68% correct per-read average—with multiple reads per site, that is, high coverage (30×), it is practicable to sequence this way. This methodology for sequencing DNA cannot be extended to proteins, however, because the pores are too large—lacking chemical specificity—and the chemical agents needed for denaturation would adversely affect a biological nanopore.

To sequence a protein with a pore, several technical hurdles have to be overcome. First the protein has to be denatured to eliminate the higher-order structure, leaving only the primary structure, to facilitate the interpretation of the blockade current associated with the AAs²². Second, the deficient chemical sensitivity of a pore (which is related to the fractional volume occluded by the molecule²³) has to be improved. Third, if the electric force field in the pore is to be used to systematically impel the molecule through the pore, the charge distribution along the protein has to be uniform. To overcome these hurdles, we sputtered subnanometre-pores (subnanopores) through thin inorganic silicon nitride membranes and used them to analyse single protein molecules that were denatured by heat, sodium dodecyl sulfate (SDS) and β-mercaptoethanol (BME).

Subnanopore fabrication and characterization

Electron beam-induced sputtering²⁴ in a scanning transmission electron microscope (STEM) was used to produce pores smaller than the size of an α-helix (which has a diameter of <0.5 nm and a rise per residue of 0.15 nm)—a common secondary structure found in proteins that is comparable in size to a hydrated ion²⁵. The small size was doubtless the key to improved chemical specificity. The topographies of the subnanopores were inferred from transmission electron microscopy (Fig. 1a–c, (i) and Supplementary Fig. 1). To accurately assess the topography, each micrograph was

¹Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA. ²Chemical Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA. ³Electrical Engineering and Biological Science, University of Notre Dame, Notre Dame, Indiana 46556, USA.

[†]These authors contributed equally to this work. *e-mail: gtimp@nd.edu

MINIREVIEW

Recent advances in DNA assembly technologies

Ran Chao¹, Yongbo Yuan¹ and Huimin Zhao^{1,2,*}

¹Department of Chemical and Biomolecular Engineering, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA and ²Departments of Chemistry, Biochemistry, and Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

*Corresponding author. Department of Chemical and Biomolecular Engineering, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. Tel.: (217) 333 2631; Fax: (217) 333 5052; E-mail: zhao5@illinois.edu
Editor. Huimin Zhao

ABSTRACT

DNA assembly is one of the most important foundational technologies for synthetic biology and metabolic engineering. Since the development of the restriction digestion and ligation method in the early 1970s, a significant amount of effort has been devoted to developing better DNA assembly methods with higher efficiency, fidelity, and modularity, as well as simpler and faster protocols. This review will not only summarize the key DNA assembly methods and their recent applications, but also highlight the innovations in assembly schemes and the challenges in automating the DNA assembly methods.

Key words: metabolic engineering; synthetic biology; pathway construction; pathway optimization; genome synthesis; automation

INTRODUCTION

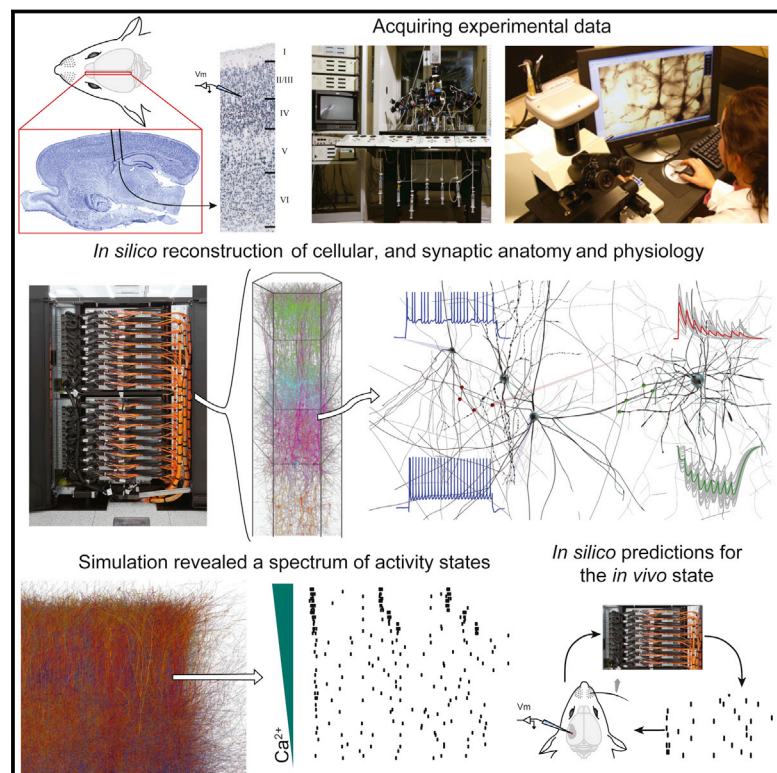
In the endeavors of metabolic engineering and synthetic biology, methods for putting genetic parts together into replicable and expressible DNA molecules are critical to rapid prototyping of the metabolic pathways or genetic circuits of interest. Although DNA can be chemically synthesized to a certain length, construction of larger fragments still relies on enzymatic assembly methods (Kosuri and Church, 2014). The first developed DNA assembly method is the restriction digestion and ligation method (Cohen et al., 1973), which led to the biotechnology revolution. Although popular, the limitations of this 40-year-old method restrict our ability of synthesizing complex DNA molecules. Increasingly, complicated DNA construct design involving multiple genes and intergenic components requires higher efficiency and fidelity in DNA assembly that is beyond the capability of the traditional cloning methods based on restriction digestion and ligation (Cobb et al., 2013). Moreover, the large size of these DNA constructs makes the selection of unique restriction sites extremely difficult. Even if restriction sites could be selected for a specific construct, they would most likely not be applicable for a different construct, which cripples

the modularity of DNA assembly, a hallmark of synthetic biology. Actually, modularity is highly desired in various research projects. For example, combinatorial pathway library construction and screening have been demonstrated as an effective approach for pathway optimization, in which various genetic parts of similar functions are mixed and matched to search for particular combinations that improve the metabolic flux or other traits (Kim et al., 2013; Xu et al., 2013; Yuan et al., 2013). Similarly, in genetic circuit design, expedited prototyping requires assembly of many characterized prefabricated parts (Canton et al., 2008; Purnick and Weiss, 2009). Therefore, broadly applicable and highly efficient DNA assembly methods are desirable. Furthermore, to survey a greater range of combinations or designs, DNA assembly will most likely be performed in large scales via automation. High-throughput DNA assembly requires robust and standardized protocols, which necessitates improvements in assembly methods for even higher efficiency, fidelity, and modularity.

A number of new DNA assembly methods have been developed in the past decade. Some of them were built upon the traditional restriction digestion and ligation method, such

Reconstruction and Simulation of Neocortical Microcircuitry

Graphical Abstract



Authors

Henry Markram, Eilif Muller,
Srikanth Ramaswamy,
Michael W. Reimann, ..., Javier DeFelipe,
Sean L. Hill, Idan Segev, Felix Schürmann

Correspondence

henry.markram@epfl.ch

In Brief

A digital reconstruction and simulation of the anatomy and physiology of neocortical microcircuitry reproduces an array of in vitro and in vivo experiments without parameter tuning and suggests that cellular and synaptic mechanisms can dynamically reconfigure the state of the network to support diverse information processing strategies.

Highlights

- The Blue Brain Project digitally reconstructs and simulates a part of neocortex
- Interdependencies allow dense in silico reconstruction from sparse experimental data
- Simulations reproduce in vitro and in vivo experiments without parameter tuning
- The neocortex reconfigures to support diverse information processing strategies



Markram et al., 2015, Cell 163, 456–492
October 8, 2015 ©2015 Elsevier Inc.
<http://dx.doi.org/10.1016/j.cell.2015.09.029>



Stem Cell Treatments – Frequently Asked Questions

DECEMBER 2013 | NHMRC REF# RM01a



Are you considering stem cell therapies to treat your condition?

If so, this document may assist you to:

- tell the difference between proven and unproven treatments
- understand the risks involved in undergoing treatments that have not been established as safe and effective.

This resource has been adapted from existing patient handbooks^{1,2}, and provides some frequently asked questions and answers regarding stem cell treatments.

1. What are stem cells?

Our bodies are made up of about 200 different types of cells, such as muscle cells, skin cells and nerve cells. Each cell type specialises in performing a certain function. Stem cells are cells that are not specialised. Their role is to replace specialised cells that have been lost through injury, disease, or the normal course of events (such as regeneration of the lining of our bowel every few days).

Stem cells have two major features, as illustrated in Figure 1. First, they have the potential to become a range of different cell types. The process of stem cells becoming particular cell types is called 'differentiation'. Second, stem cells can also replace themselves, or 'self-replicate'.

The ability of stem cells to differentiate and self-replicate is what has stimulated such enormous interest in the use of stem cells for the treatment of disease.

Stem cells can be recovered from embryonic, fetal and some adult tissues and have different properties depending on their source.

For more information on stem cells visit <http://www.closerlookatstemcells.org>

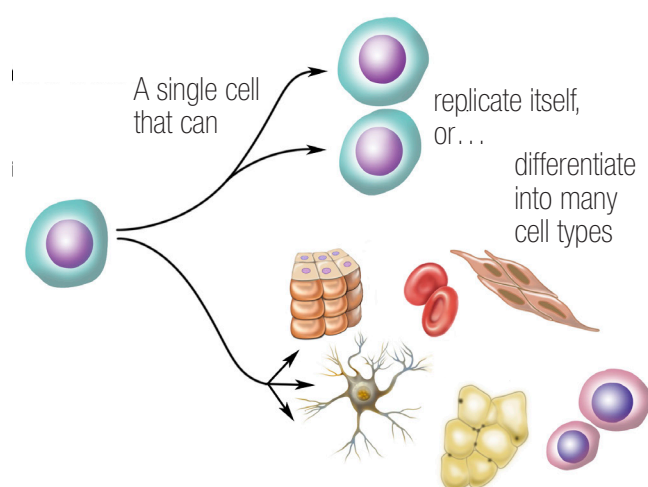


Figure 1: Simplified diagram demonstrating the features of a stem cell


Used with permission from Understanding Stem Cells: An Overview of the Science and Issues from the National Academies by the National Academy of Sciences, Courtesy of the National Academies Press, Washington, D.C.

ARTICLE

DOI: 10.1038/s41467-017-01104-3

OPEN

Mass spectrometry sequencing of long digital polymers facilitated by programmed inter-byte fragmentation

Abdelaziz Al Ouahabi¹, Jean-Arthur Amalian², Laurence Charles² & Jean-François Lutz ¹

In the context of data storage miniaturization, it was recently shown that digital information can be stored in the monomer sequences of non-natural macromolecules. However, the sequencing of such digital polymers is currently limited to short chains. Here, we report that intact multi-byte digital polymers can be sequenced in a moderate resolution mass spectrometer and that full sequence coverage can be attained without requiring pre-analysis digestion or the help of sequence databases. In order to do so, the polymers are designed to undergo controlled fragmentations in collision-induced dissociation conditions. Each byte of the sequence is labeled by an identification tag and a weak alkoxyamine group is placed between 2 bytes. As a consequence of this design, the NO-C bonds break first upon collisional activation, thus leading to a pattern of mass tag-shifted intact bytes. Afterwards, each byte is individually sequenced in pseudo-MS³ conditions and the whole sequence is found.

¹ Université de Strasbourg, CNRS, Institut Charles Sadron UPR22, 23 rue du Loess, 67034 Strasbourg, France. ² Aix-Marseille Université, CNRS, UMR 7273, Institute of Radical Chemistry, 13397 Marseille, France. Correspondence and requests for materials should be addressed to L.C. (email: laurence.charles@univ-amu.fr) or to J.-F.L. (email: jflutz@unistra.fr)

Self-Deployed Extremely Large Low Mass Space Structures

Devon G. Crowe, Prakash Joshi,
Ed Rietman, and Kophu Chang*

Physical Sciences, Inc.

March 7, 2007

*With a little help from our friends W. Cash and A. Labeyrie

Sequence control in polymer synthesis

Nezha Badi and Jean-François Lutz*

Received 25th March 2009

First published as an Advance Article on the web 23rd July 2009

DOI: 10.1039/b806413j

The control over comonomer sequences is barely studied in macromolecular science nowadays. This is an astonishing situation, taking into account that sequence-defined polymers such as nucleic acids and proteins are key components of the living world. In fact, fascinating biological machines such as enzymes, transport proteins, cytochromes or sensory receptors would certainly not exist if evolution had not favored chemical pathways for controlling chirality and sequences. Thus, it seems obvious that synthetic polymers with controlled monomer sequences have an enormous role to play in the materials science of the next centuries. The goal of this *tutorial review* is to shed light on this highly important but embryonic field of research. Both biological and synthetic mechanisms for controlling sequences in polymerization processes are critically discussed herein. This state-of-the-art overview may serve as a source of inspiration for the development of new generations of synthetic macromolecules.

Introduction

The decoding of gene and protein sequences is certainly one of the biggest scientific achievements of the 20th century.^{1–3} Indeed, the ordered monomer sequences contained in biopolymers such as nucleic acids and polypeptides are in large part responsible for the diversity, complexity and adaptability of living organisms. Thus, macromolecules such as DNA or RNA are not only crucial scientific subjects but also modern cultural icons. For instance, nucleotide sequences have escaped biology laboratories and invaded our daily lives through popular magazines, art galleries, paternity lawsuits, and forensic TV shows.

In such a context, one could logically expect that the control over monomer sequences in polymerization processes is an established research field. Yet, even if molecular biologists and biochemists have understood for decades the importance of ordered monomer sequences, this topic is far from being central in other chemical disciplines. In particular, and rather surprisingly, this crucial subject is largely ignored in contemporary polymer chemistry. In fact, the most relevant discovery to date in the field of sequence-controlled macromolecules is due to a biochemist. In early 1963, Robert Bruce Merrifield introduced the solid-phase synthesis of peptides, which relies on the step-by-step attachment of protected monomers.⁴ This method remains so far the most reliable pathway for synthesizing oligomers with tailored monomer sequences.

Nevertheless, it would be rather unfair to pretend that polymer chemists never showed interest in understanding macromolecular sequences. In fact, in the late 1960's and early

Research Group Nanotechnology for Life Science, Fraunhofer Institute for Applied Polymer Research, Geiselbergstrasse 69, 14476 Potsdam-Golm, Germany. E-mail: lutz@iap.fhg.de; Fax: +49 331 568 3000; Tel: +49 331 568 1127



Nezha Badi

Nezha Badi obtained her MSc degree from the Université Paris VI in 2003. In 2006, she received her PhD in chemistry and physico-chemistry of polymers from the University of Evry-Val-D'Essonne, under the guidance of Prof. Philippe Guégan. After working as a postdoctoral researcher in the group of Dr. Jean-Paul Behr at the Université Strasbourg I, she joined in September 2008 the research group of Dr. Jean-François Lutz at the Fraunhofer Institute for Applied Polymer Research.

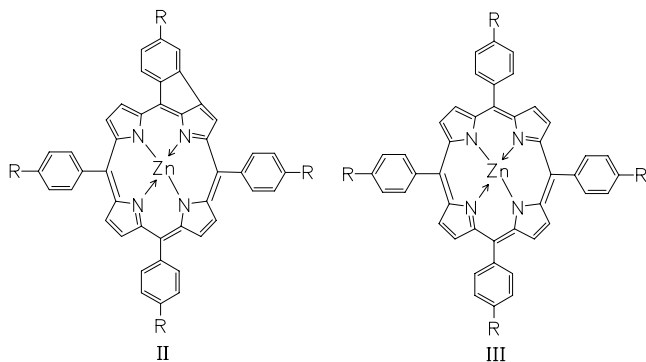
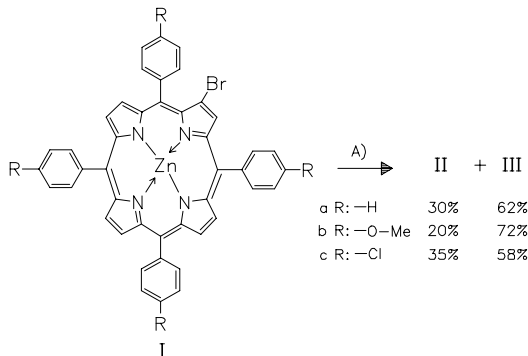


Jean-François Lutz

Jean-François Lutz received his BSc and MSc degrees from the Université Paris VI and his doctoral degree from the Université Montpellier II. Afterwards, he joined the research group of Prof. Krzysztof Matyjaszewski at Carnegie-Mellon University, where he worked for two years as a postdoctoral fellow. In late 2003, he moved to Potsdam in Germany and became team leader at the Fraunhofer Institute for Applied Polymer research.

His current research interests include the synthesis of sequence-ordered polymers, the design of smart biocompatible polymers and the self-organization of macromolecules in aqueous media.

A Novel and Facile Zn-Mediated Intramolecular Five-Membered Cyclization of β -Tetraarylporphyrin Radicals from β -Bromotetraarylporphyrins. — The title reaction affords novel fused five-membered porphyrin systems, such as (II), (III), (V) and (VI). — (SHEN, D.-M.; LIU, C.; CHEN*, Q.-Y.; Chem. Commun. (Cambridge) 2005, 39, 4982-4984; State Key Lab. Organofluorine Chem., Shanghai Inst. Org. Chem., Chin. Acad. Sci., Shanghai 200032, Peop. Rep. China; Eng.) — M. Paetzel



Understanding Understanding Source Code with Functional Magnetic Resonance Imaging

Janet Siegmund^{π,*}, Christian Kästner^ω, Sven Apel^π, Chris Parnin^β, Anja Bethmann^θ,
Thomas Leich^δ, Gunter Saake^σ, and André Brechmann^θ
^πUniversity of Passau, Germany ^ωCarnegie Mellon University, USA
^βGeorgia Institute of Technology, USA ^θLeibniz Inst. for Neurobiology Magdeburg, Germany
^δMetop Research Institute, Magdeburg, Germany ^σUniversity of Magdeburg, Germany

ABSTRACT

Program comprehension is an important cognitive process that inherently eludes direct measurement. Thus, researchers are struggling with providing suitable programming languages, tools, or coding conventions to support developers in their everyday work. In this paper, we explore whether *functional magnetic resonance imaging (fMRI)*, which is well established in cognitive neuroscience, is feasible to soundly measure program comprehension. In a controlled experiment, we observed 17 participants inside an fMRI scanner while they were comprehending short source-code snippets, which we contrasted with locating syntax errors. We found a clear, distinct activation pattern of five brain regions, which are related to working memory, attention, and language processing—all processes that fit well to our understanding of program comprehension. Our results encourage us and, hopefully, other researchers to use fMRI in future studies to measure program comprehension and, in the long run, answer questions, such as: Can we predict whether someone will be an excellent programmer? How effective are new languages and tools for program understanding? How should we train programmers?

Categories and Subject Descriptors

H.1.2 [Information Systems]: User/Machine Systems

General Terms

Experimentation, human factors

Keywords

Functional magnetic resonance imaging, program comprehension

1. INTRODUCTION

As the world becomes increasingly dependent on the billions lines of code written by software developers, little comfort can be taken in the fact that we still have no fundamental understanding of how programmers understand source code.

Understanding program comprehension is not limited to theory building, but can have real downstream effects in improving education, training, and the design and evaluation of tools and languages

*This author published previous work as Janet Feigenspan.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICSE '14, May 31 – June 7, 2014, Hyderabad, India

Copyright 14 ACM 978-1-4503-2756-5/14/05 ...\$10.00.

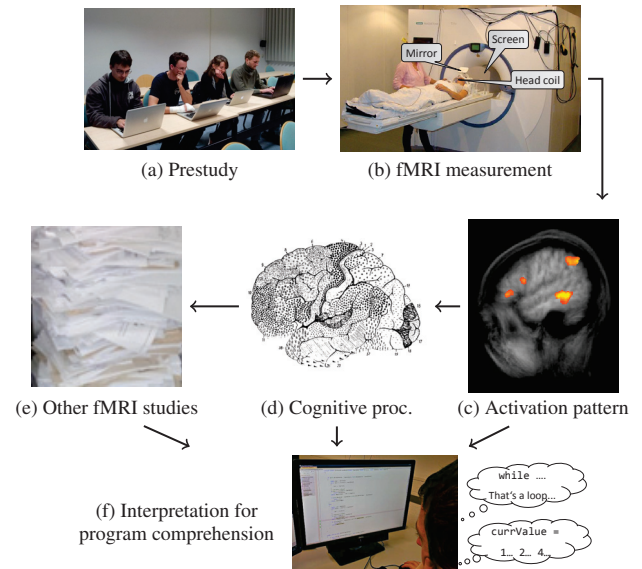


Figure 1: Workflow of our fMRI study.

for programmers. If direct measures of cognitive effort and difficulty could be obtained and correlated with programming activity, then researchers could identify and quantify which types of activities, segments of code, or kinds of problem solving are troublesome or improved with the introduction of a new language or tool.

In studying programmers, decades of psychological and observational experiments have relied on indirect techniques, such as comparing task performance or having programmers articulate their thoughts in think-aloud protocols. Each method, when skillfully applied, can yield important insights. However, these common techniques are not without problems. In human studies of programming, individual [13] and task variance [19] in performance often mask any significant effects hoping to be found when evaluating, say, a new tool. Think-aloud protocols and surveys rely on self-reporting and require considerable manual transcription and analysis that garner valuable but indefinite and inconsistent insight.

In the past few decades, psychologists and cognitive neuroscientists have collectively embraced methods that measure physiological correlates of cognition as a standard practice. One such method is *functional magnetic resonance imaging (fMRI)*, a non-invasive means of measuring blood-oxygenation levels that change as a result of localized brain activity.

Parsing Natural Scenes and Natural Language with Recursive Neural Networks

Richard Socher
Cliff Chiung-Yu Lin
Andrew Y. Ng
Christopher D. Manning

RICHARD@SOCHER.ORG
CHIUNGYU@STANFORD.EDU
ANG@CS.STANFORD.EDU
MANNING@STANFORD.EDU

Computer Science Department, Stanford University, Stanford, CA 94305, USA

Abstract

Recursive structure is commonly found in the inputs of different modalities such as natural scene images or natural language sentences. Discovering this recursive structure helps us to not only identify the units that an image or sentence contains but also how they interact to form a whole. We introduce a max-margin structure prediction architecture based on recursive neural networks that can successfully recover such structure both in complex scene images as well as sentences. The same algorithm can be used both to provide a competitive syntactic parser for natural language sentences from the Penn Treebank and to outperform alternative approaches for semantic scene segmentation, annotation and classification. For segmentation and annotation our algorithm obtains a new level of state-of-the-art performance on the Stanford background dataset (78.1%). The features from the image parse tree outperform Gist descriptors for scene classification by 4%.

1. Introduction

Recursive structure is commonly found in different modalities, as shown in Fig. 1. The syntactic rules of natural language are known to be recursive, with noun phrases containing relative clauses that themselves contain noun phrases, e.g., ... *the church which has nice windows* ... Similarly, one finds nested hierarchical structuring in scene images that capture both part-of and proximity relationships. For instance, cars are often on top of street regions. A large car region

Appearing in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011. Copyright 2011 by the author(s)/owner(s).

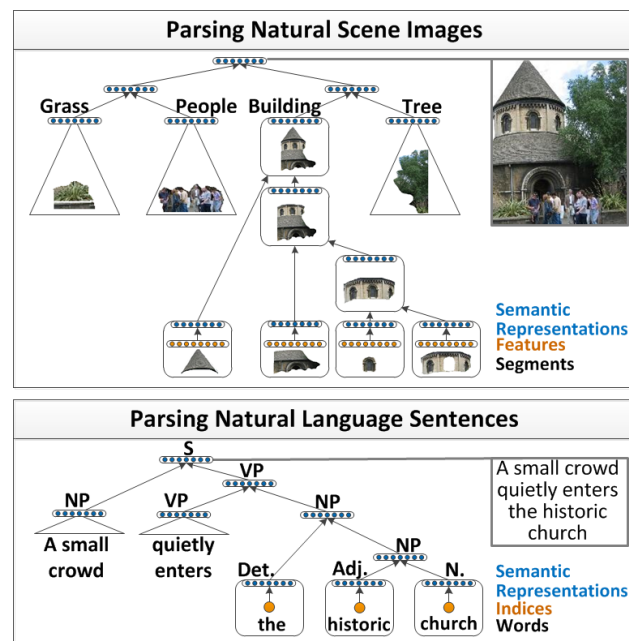


Figure 1. Illustration of our recursive neural network architecture which parses images and natural language sentences. Segment features and word indices (orange) are first mapped into semantic feature space (blue) and then recursively merged by the same neural network until they represent the entire image or sentence. Both mappings and mergings are learned.

can be recursively split into smaller car regions depicting parts such as tires and windows and these parts can occur in other contexts such as beneath airplanes or in houses. We show that recovering this structure helps in understanding and classifying scene images. In this paper, we introduce recursive neural networks (RNNs) for predicting recursive structure in multiple modalities. We primarily focus on scene understanding, a central task in computer vision often subdivided into segmentation, annotation and classification of scene images. We show that our algorithm is a general tool

9.04 Magma Oceans and Primordial Mantle Differentiation

V. Solomatov, Washington University, St. Louis, MO, USA

© 2007 Elsevier B.V. All rights reserved.

9.04.1	Earth Accretion and the Giant Impact Hypothesis	92
9.04.2	Geochemical Evidence for Magma Ocean	94
9.04.3	Thermal Structure of a Convecting Magma Ocean	95
9.04.3.1	Adiabats	95
9.04.3.2	Upper Mantle	96
9.04.3.3	Lower Mantle	96
9.04.4	Viscosity of the Magma Ocean	98
9.04.5	Convection in the Magma Ocean	100
9.04.5.1	Convective Heat Flux	100
9.04.5.2	Convective Velocities	101
9.04.6	Fractional versus Equilibrium Crystallization	102
9.04.6.1	How Are Crystals Suspended by Convection?	102
9.04.6.2	Energetics of Convective Suspension	103
9.04.6.3	Conditions for Fractional Crystallization	104
9.04.6.4	Conditions for Equilibrium Crystallization	104
9.04.7	Crystal Size in the Magma Ocean	105
9.04.7.1	What Processes Control the Crystal Size in the Magma Ocean	105
9.04.7.2	Nucleation	105
9.04.7.3	Ostwald Ripening	107
9.04.8	Crystallization beyond the Rheological Transition	108
9.04.9	The Last Stages of Crystallization	109
9.04.9.1	Cessation of Suspension	109
9.04.9.2	Formation of Thin Crust within the Thermal Boundary Layer	110
9.04.9.3	Cessation of Liquid-State Convection	110
9.04.9.4	Percolation	111
9.04.9.5	Remelting due to Melt Extraction	111
9.04.9.6	Solid-State Convection	111
9.04.10	Summary	112
References		113

Nomenclature

a	prefactor in nucleation function	d_{nucl}	crystal diameter after nucleation
a_u	prefactor in scaling law for velocity	d_{ost}	crystal diameter controlled by Ostwald ripening
c_p	isobaric specific heat	d'_{ost}	crystal diameter in shallow magma ocean
c'_p	apparent isobaric specific heat in two-phase regions	f_ϕ	hindering settling function
d	crystal diameter	g	acceleration due to gravity
d_{crit}	crystal diameter separating equilibrium and fractional crystallization	k_B	Boltzmann's constant
d_e	critical crystal diameter for equilibrium crystallization	n_i	mole fraction of component i
d_f	critical crystal diameter for fractional crystallization	n_i^l	mole fraction of liquid component i
		n_i^s	mole fraction of solid component i
		n_{mw}	mole fraction of magnesiowüstite
		t	time

Designs to be included on this list: Design must intend for permanent human habitation. Design must be published or widely publicized or be status easily found by Google search only. Illustrating scale only, may be inexact, use notes for each entry												For the human habitation portion of the settlement					
Name of Design	Year of Publication/Release	Author(s)	Era	Cost (metric USD millions)	Mass (metric tons)	Population Target	Location	Diameter (meters)	Length (meters)	Basic Shape	Energy Notes	Artificial Gravity Notes	RPM	Website link(s)	Notes	Associated with NASA Ames Contest?	Associated with GLOBUS?
1 Brick Moon	1869	Edward Everett Hale	early (before 1975)	1,754	132,000	200	MEO	61	n/a	Sphere	solar			http://www.searchoptegine.org.uk/docs/73157.pdf	Work of fiction but it is the first detailed thought experiment on orbital space settlement.	no	no
2 Tsiolkovsky "Greenhouse"	1903	Konstantin Tsiolkovsky	early							Cylinder	solar			http://www.aerospacespace.com/Forum/index.php?topic=5993.0.html	First scientific description of artificial gravity and solar radiation to aid agriculture and permanent space habitation	no	no
3 Oberth Wheel	1923	Hermann Oberth	early							Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Unclear if intended to be a home or a base, published in his seminal work, "The Rocket into Interplanetary Space," his station design rotates to provide artificial gravity and is a fuel depot.	no	no
4 Wehrard	1928	Hermann Noordung Pototskiy	early				GEO	30		Disk	solar			http://www.astronautix.com/noordung.htm	unclear if intended to be a home, published in his seminal work, "The Problem of Space Flight"	no	no
5 Bernal Sphere	1929	John Desmond Bernal	early			30,000	L4	16,000		Sphere	solar			http://en.wikipedia.org/wiki/Bernal_sphere	unclear if intended to be a home, this station has a detailed description of space solar thermal power and separates gravitic and zero n rotating portions	no	no
6 Orbital Base	1949	Ross and Smith	early			14	GEO	60		Torus	solar			https://www.bbc.com/news/science-technology-1949-01-01-orbital-base-station	unclear if intended to be a home,	no	no
7 Wheel Station	1952	Wehrner Von Braun et al	early	36,036,036										http://www.nasa.gov/pdf/1952-01-01-wehrner-von-braun-et-al-wheel-station		no	no
8 Bubbleworld	1964	Dandridge Cole	modern (1975 - 1994)							Hollow Asteroid	solar			http://discoveryengine.wordpress.com/2007/08/08/films-in-space-the-bubbleworld/		yes	no
9 Sunflower	1974	O'Neill et al	modern			10,000	L4/L5	457		Cylinder	solar	1g		http://www.nasa.org/settlement/CofC/moon/Sunflower/chaos08.html	brief reference in link provided, "stubby cylinder", envisioned as a retirement home	no	no
10 Stanford Torus	1975	O'Neill et al	modern			10,000	L4/L5	1,600		Torus	solar	1g		http://www.settlement.nasa.gov/75SummerStudy/Chapter1.html#history		no	no
11 Bernal Sphere	1976	O'Neill et al	modern			10,000	L4/L5	10,000		Sphere	solar	1.9		http://www.settlement.nasa.gov/75SummerStudy/Chapter1.html#history		no	no
12 O'Neill Cylinder	1976	O'Neill et al	modern			2,000,000	L4/L5	8,000	32,000	Cylinder	solar	1g		http://en.wikipedia.org/wiki/O'Neill_Cylinder	population target "1 million" i.e. at least 2 million people	no	no
13 Space Flower	1985	Nagatoka	modern			10,000				Cylinder	solar			http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/1985-01-01/1985-01-01.pdf	was only able to find a graphic made for the Mitsubishi Pavilion at the Tsukuba World Expo in 1985, not sure if this is an actual full design	no	no
14 Lewis One Dome	1991	GlobeOne	modern			10,000		534	1,921	Cylinder	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history		yes	no
15 DSH, Inc. McKendree Cylinder	1994	Diaz, Horn, Leong	contemporary (1994 - present)	2,000,000,000		1,000,000				Cylinder	solar	1g		http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest grand prize winner	yes	no
16 Hollow Asteroid	1995	Herder, Shepherd, Hirst	contemporary				L4/L5	820	950	Hollow Asteroid	fission			http://www.nasa.gov/content/pdf/1995-01-01/1995-01-01.pdf	reimagines much larger O'Neill cylinders using CNTs instead of steel	yes	no
17 Tango III	1996	Podesta et al	contemporary				L4/L5			Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest first place for technical merit	yes	no
18 Bishop Ring	1997	Bishop	contemporary				L4/L5	2,000,000	500	Torus	solar			http://www.lase.csl.llnwd.net/bishop/	Ames Contest grand prize winner	yes	no
20 SCORE	1997	Woods	contemporary			300,000		7,120	12,500	Hollow Asteroid	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest grand prize winner	yes	no
21 Babylon Project	1998	Beatty and Peters	contemporary			100,000	L4	3,500,000	LS	25,000	31,000			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest grand prize winner	yes	no
22 Clementina I	2000	Clementina I	contemporary			100,000	L4	2,112		Torus	solar			http://www.danishland.com/clementina/	See central sphere idea, Diameter = major radius x 2, Ames Contest Grand Prize Winner	yes	no
23 Avalon	2001	O'Connell and Thaker	contemporary			163,200		5	644	O'Neill	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 2, Ames Contest Grand Prize co-winner	yes	no
24 Centurion	2001	Armi	contemporary			8,000	LS	1,712		Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 2, Ames Contest Grand Prize co-winner	yes	no
25 Antler	2002	Hsiang, Hsiang, Te, Nemov	contemporary			100,000	L4/L5	4000	500	Torus	solar	1g	0.66	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 3, length = minor radius x 2, Ames Contest Grand Prize Winner	yes	no
26 Teba I	2003	Teodorescu, Bahin	contemporary			25,000	L4/L5	1760	30	Torus	solar			http://www.nss.gov/settlement/nasa/Contests/Results/2003/teba/TEBA-I.html	Interesting report on vibrations in the structure, Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
27 Leda	2004	Valentin, Andres, Maria	contemporary			250,000	L4	8022	1000	Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
28 Intercom	2006	Irish	contemporary	352,700,000		11,500	L4/L5	1600	100	Torus	solar			http://www.nss.gov/settlement/nasa/Contests/Results/2006/intercom/Intercom.pdf	Interesting "Twelve foot long, huge 600m radius ball, Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
29 Apis	2007	Ovidius HS	contemporary			10,000	L4/L5	8022	1000	Cylinder	solar			http://www.nss.gov/settlement/nasa/Contests/Results/2007/apis/Apis.pdf	Honeycomb Construction, Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
30 Pinta	2007	Becker, Craydon	contemporary			10,000	L4/L5	50	1000	Torus	solar/fission			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
31 Kalena One	2008	Arora, Bajora, Glicks, Strout	contemporary		x	3,000		50	325	Torus	solar			http://www.settlement.nasa.gov/75SummerStudy/Chapter1.html#history	There are actually Four Kalena designs: One through Four	yes	no
32 DACIA's	2008	Alexandrea et al	contemporary	44,849,360,000		60,000	L4/L5	9000	1500	Torus	solar			http://www.nss.gov/settlement/nasa/Contests/Results/2008/DACIA's	3 Tori, Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize Winner	yes	no
33 Asten	2009	Yam	contemporary	627,990,000		22,500	GEO	1000	1700	Cylinder	solar			http://www.nss.gov/settlement/nasa/Contests/Results/2009/ASTEN.pdf	interesting cost data, Ames Contest grand prize winner	yes	no
34 Distant Hope	2010	Distant Hope	contemporary	354,412,450		22,000	Moon Orbit	2400		Torus/Fission	solar	966	1.022g	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	reads like a link novel, huge 600m radius ball, Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize winner	yes	no
35 Hyperion	2011	Guazary et al	contemporary	243,900,060	>555,000	20,000	L4	1864	320	Torus	solar	1g	0.97	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	4 tori, 2 small, 2 large, Diameter = major radius/2 of large tori, length = minor radius/2 of large tori, Ames Contest Grand Prize Winner	yes	no
36 Kon Tiki IVS	2012	Gitten et al	contemporary	21,200,000		1,200	Interstellar	669	250	Torus	fusion/mhd		1.53	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Interstellar Cyclone, aka quasi generation ship, diameter refers to habitat torus, length refers to central acceleration truss, Ames Contest grand prize winner	yes	no
37 Aurora	2012	Tadler	contemporary	21,109,490		1,4				Torus	fission	1.35	1.53	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest grand prize winner	yes	no
38 Maui	2013	Gitten et al	contemporary			25,900,000	1000	Saturn Orbit	1400	1750	Torus	fission	1.35	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Very speculative, phased H3 mining colony in Saturn orbit in 22nd century, Ames Contest grand prize winner	yes	no
39 Greenspacer	2014	Kovachev et al	contemporary							Torus/fusion	solar/fusion			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Ames Contest grand prize winner	yes	no
40 Fire	2014	Singh et al	contemporary	485,365,039		16,640	Moon Orbit	1544	270	Torus	solar	1g	0.900	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	W8F found by sifting interstellar resources, diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize Winner	yes	no
41 Fire	2015	Reeves	contemporary	140,000,000	>1,620,000	20,000	Moon Orbit	830	200	Fusion	fission	1g	1.409	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = major radius x 2, length = minor radius x 2, Ames Contest Grand Prize Winner	yes	no
42 Divinity	2016	Do et al	contemporary			10,000	ELEO	400	65	Torus	fission	1g	2.114	http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	Diameter = diameter of torus, length is width of torus.	yes	no
43 Kalena Two	2016	Globus and Versteeg	contemporary				GLOBUS			Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	There are actually Four Kalena designs: One through Four	no	yes
44 Kalena Three	2016	Globus	contemporary		<x		ELEO			Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	There are actually Four Kalena designs: One through Four	no	yes
45 Kalena Four	2016	Globus	contemporary				ELEO		50	Torus	solar			http://settlement.nasa.gov/75SummerStudy/Chapter1.html#history	There are actually Four Kalena designs: Four is similar to Two but has a 50 m length rather than 100.	no	yes
46 StarPort 1	2016	International Space, British Interplanetary Society	contemporary			10,000	200	L4O		Torus	fission	0.8g		http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/2016-01-01/2016-01-01.pdf	commissioned/sponsored by Axion Space, a startup commercial space station module development company	no	yes

**Genome Project-write: A Grand Challenge
Using Synthesis, Gene Editing and Other Technologies
to Understand, Engineer and Test Living Systems**

October 31, 2016

Author: Jef D Boeke, George Church, Andrew Hessel, Nancy J Kelley (GP-write Leadership Group) and The GP-Write Consortium†

†The authors and their affiliations maybe found at www.gpwrite.org

INTRODUCTION

- The Human Genome Project (HGP-read) was initiated in December 1984 at the DOE Alta meeting and launched on October 1, 1990. Aimed at a nearly complete “reading” of the human and other genomes with annotations of plausible gene functions, as well as the improvement of the technology, cost, and quality of DNA sequencing (1, 2), it was the largest life science project ever conducted and one of the great feats of exploration in history.
- The first draft of the human genome was announced in June 2000, and a more complete version in April 2003.
- Since this announcement, interpretation of that data was made possible by sequencing humans from other populations around the world (HapMap and 1KG) and by other “omics” tools yielding insights into the functionality at base pair precision (ENCODE), correlations between natural genome variants and various traits and diseases via genome-wide association studies (GWAS), and identification of genes and variants responsible for rare Mendelian traits by a combination of genomic sequencing and advanced computational methods.
- Incredibly powerful and inexpensive DNA sequencing and genetic analysis technologies, developed, in part, by the National Human Genome Research Institute’s (NHGRI) Advanced DNA Sequencing Technology Development program, reduced the cost of sequencing a human genome more than a million fold from \$3 billion to less than \$1,000 in just over a decade.
- Since then, the world has embarked on a revolution in science and healthcare that is changing the way we live and carries the promise of individualizing clinical delivery to improve health, and prevent and cure human disease. The Precision Medicine Initiative (for example), announced by the White House in 2015, would use the insights gained from “reading” each individual’s genome to personalize medical care based on an individual’s genes,

LETTER

doi:10.1038/nature20777

Synthetic recording and *in situ* readout of lineage information in single cells

Kirsten L. Frieda, James M. Linton, Sahand Hormoz, Joonhyuk Choi, Ke-Huan K. Chow, Zakary S. Singer, Mark W. Budde, Michael B. Elowitz & Long Cai

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. *Nature* is providing this early version of the typeset paper as a service to our customers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Cite this article as: Frieda, K. L. *et al.* Synthetic recording and *in situ* readout of lineage information in single cells. *Nature* <http://dx.doi.org/10.1038/nature20777> (2016).

Received 30 June; accepted 11 November 2016.

Accelerated Article Preview Published online 21 November 2016.



Published in final edited form as:

Nat Nanotechnol. 2015 July ; 10(7): 629–636. doi:10.1038/nnano.2015.115.

Syringe injectable electronics

Jia Liu^{#1}, Tian-Ming Fu^{#1}, Zengguang Cheng^{#1,2}, Guosong Hong¹, Tao Zhou¹, Lihua Jin³, Madhavi Duvvuri¹, Zhe Jiang¹, Peter Kruskal¹, Chong Xie¹, Zhigang Suo³, Ying Fang², and Charles M. Lieber^{1,3,*}

¹Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

²National Center for Nanoscience and Technology, 11 Beiyitiao Street, Zhongguancun, Beijing 100190, P.R. China.

³School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.

[#] These authors contributed equally to this work.

Abstract

Seamless and minimally-invasive three-dimensional (3D) interpenetration of electronics within artificial or natural structures could allow for continuous monitoring and manipulation of their properties. Flexible electronics provide a means for conforming electronics to non-planar surfaces, yet targeted delivery of flexible electronics to internal regions remains difficult. Here, we overcome this challenge by demonstrating syringe injection and subsequent unfolding of submicrometer-thick, centimeter-scale macroporous mesh electronics through needles with a diameter as small as 100 micrometers. Our results show that electronic components can be injected into man-made and biological cavities, as well as dense gels and tissue, with > 90% device yield. We demonstrate several applications of syringe injectable electronics as a general approach for interpenetrating flexible electronics with 3D structures, including (i) monitoring of internal mechanical strains in polymer cavities, (ii) tight integration and low chronic immunoreactivity with several distinct regions of the brain, and (iii) *in vivo* multiplexed neural recording. Moreover, syringe injection enables delivery of flexible electronics through a rigid shell, delivery of large volume flexible electronics that can fill internal cavities and co-injection of electronics with other materials into host structures, opening up unique applications for flexible electronics.

Reprints and permissions information is available at www.nature.com/reprints.

* Corresponding author. cml@cmliris.harvard.edu.

Author Contributions

J.L., Z.C. and C.M.L. designed the experiments. J.L., Z.C., T.F., G.H. and M.D. performed the experiments. L.J. and Z.S. performed FEM analysis. J.L., Z.C., T.F. and C.M.L. analyzed the data and wrote the paper. All authors discussed the results and commented on the manuscript.

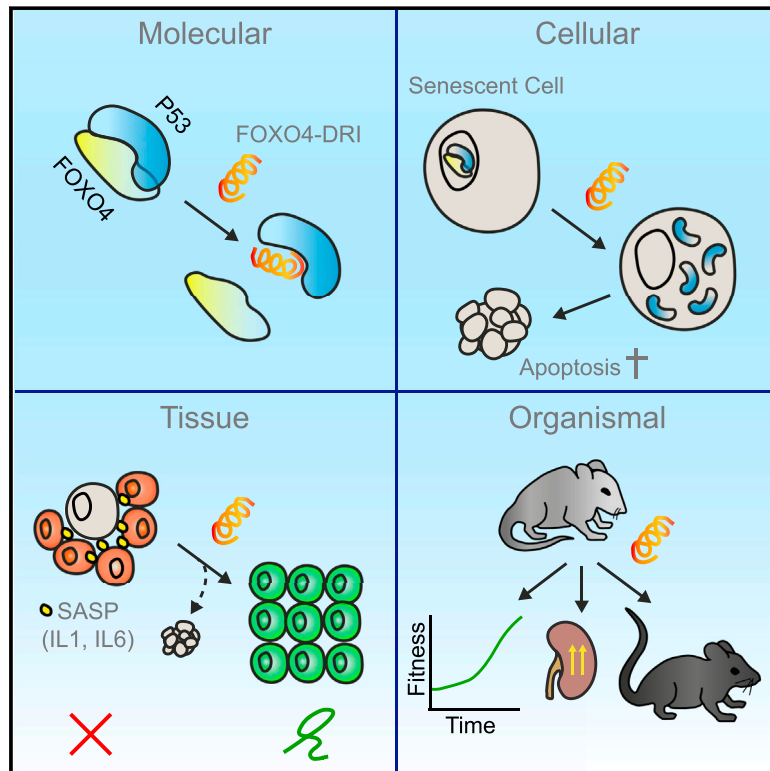
Supplementary Information is available in the online version of the paper.

Competing financial interests

The authors declare no competing financial interests.

Targeted Apoptosis of Senescent Cells Restores Tissue Homeostasis in Response to Chemotoxicity and Aging

Graphical Abstract



Authors

Marjolein P. Baar, Renata M.C. Brandt, Diana A. Putavet, ..., Jan H.J. Hoeijmakers, Judith Campisi, Peter L.J. de Keizer

Correspondence

p.dekeizer@erasmusmc.nl

In Brief

A FOXO4 peptide that selectively induces apoptosis of senescent cells reverses effects of chemotoxicity and aging in mice.

Highlights

- A modified FOXO4-p53 interfering peptide causes p53 nuclear exclusion in senescent cells
- This FOXO4 peptide induces targeted apoptosis of senescent cells (TASC)
- TASC neutralizes murine liver chemotoxicity from doxorubicin treatment
- TASC restores fitness, hair density, and renal function in fast and naturally aged mice



The A-Design approach to managing automated design synthesis

Matthew I. Campbell, Jonathan Cagan, Kenneth Kotovsky

12

Abstract This paper presents an approach to managing the complexities of design synthesis techniques. Often computational approaches model design as a search process or as an optimization problem. The nature of such techniques, however, does not allow the user to interact with the search process once it has begun. Furthermore, traditional computational search lacks the ability to learn from experience. While computational search techniques have the ability to search many design alternatives quickly, the human engineer can often arrive at a more elegant and robust solution by applying heuristics learned from past experiences. The method introduced here improves the capabilities of design synthesis methods by allowing for user input and by making decisions based on previous experience. In studying past candidate designs the process learns to be more effective in searching for solutions. Results show how such a technique improves the quality of designs and efficiency of an existing search process.

Keywords Learning, Agents, Multi-objective optimization, Design synthesis, Stochastic optimization

Introduction

As design synthesis techniques find their way into engineering practice, the designer has at his/her disposal computational processes to alleviate the need to perform tedious or time-intensive design tasks by hand. While the computer has become an indispensable companion to the engineering designer because of its speed and accuracy, the most significant decisions of the design process remain a product of human decision-making. Often several challenges within a larger engineering design problem can be restated as a computational search problem. The computer exhaustively, or by following heuristics, searches for a solution that best meets the designer's goal. From this model, numerous techniques have developed to enhance engineering design quality and efficiency. While the model of design as a search process has proven to be useful, it appears that challenges still exist in integrating the human and computer as a symbiotic design team. In order for the computer to tackle more ambitious and conceptual design tasks several theoretical and computational hurdles must be overcome.

In Fig. 1, a general model of computational synthesis shows four basic aspects of modeling design as a search process. The process initiates with some starting point description of the design problem and some qualities for making better designs. Prior to the search process, a language or representation needs to be developed for the domain of the problem. Based on the representation, the search process is performed as an iterative loop of generating alternatives, evaluating their worth, and determining how to perform redesign in subsequent iterations. While each of these tasks could be the impetus for future research, the method described here focuses on improving the guidance task of the iterative process. Guidance in human design processes results from learning from past experiences and understanding how to balance the tradeoffs in a design problem.

This paper establishes a method to overcome two shortcomings of current computational synthesis techniques. These are (1) the lack of information exchanged during the search process between man and machine, and (2) the lack of learning that occurs throughout computational search. The strategies developed here to overcome these challenges are together enveloped in a software agent that interacts with other software agents responsible for separate aspects of the search process. Altogether, the

Received: 12 May 2001 / Revised: 15 October 2002 /
Accepted: 15 October 2002 / Published online: 28 November 2002
© Springer-Verlag 2002

M.I. Campbell (✉)
Department of Mechanical Engineering,
The University of Texas at Austin, 1 University Station,
C2200 Austin, TX 78712-0292, USA
E-mail: mc1@mail.utexas.edu
Tel.: +1-512-2329122
Fax: +1-512-4717682

J. Cagan
Department of Mechanical Engineering,
Carnegie Mellon University, Pittsburgh, PA 15213, USA

K. Kotovsky
Department of Psychology, Carnegie Mellon University,
Pittsburgh, PA 15213, USA

The research effort was partially sponsored by the Defense Advanced Research Projects Agency (DARPA) and Rome Laboratory, Air Force Materiel Command, USAF, under agreement number F30602-96-2-0304. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, Rome Laboratory, or the U.S. Government. The authors would also like to acknowledge the efforts of Jarrod Moss in the testing and development of the learning algorithm and the financial support of the National Science Foundation under grant EID-9256665.

Accepted Manuscript

The Cingulate Cortex of Older Adults with Excellent Memory Capacity

Feng Lin, PhD, Ping Ren, Mark Mapstone, Steven P. Meyers, Anton Porsteinsson, Timothy M. Baran



PII: S0010-9452(16)30323-9

DOI: [10.1016/j.cortex.2016.11.009](https://doi.org/10.1016/j.cortex.2016.11.009)

Reference: CORTEX 1880

To appear in: *Cortex*

Received Date: 18 July 2016

Revised Date: 3 October 2016

Accepted Date: 4 November 2016

Please cite this article as: Lin F, Ren P, Mapstone M, Meyers SP, Porsteinsson A, Baran TM, for the Alzheimer's Disease Neuroimaging Initiative, The Cingulate Cortex of Older Adults with Excellent Memory Capacity, *CORTEX* (2016), doi: 10.1016/j.cortex.2016.11.009.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The New England Journal of Medicine

© Copyright, 1996, by the Massachusetts Medical Society

VOLUME 335

JULY 4, 1996

NUMBER 1



THE EFFECTS OF SUPRAPHYSIOLOGIC DOSES OF TESTOSTERONE ON MUSCLE SIZE AND STRENGTH IN NORMAL MEN

SHALENDER BHASIN, M.D., THOMAS W. STORER, Ph.D., NANCY BERMAN, Ph.D., CARLOS CALLEGARI, M.D.,
BRENDA CLEVINGER, B.A., JEFFREY PHILLIPS, M.D., THOMAS J. BUNNELL, B.A., RAY TRICKER, Ph.D., AIDA SHIRAZI, R.Ph.,
AND RICHARD CASABURI, Ph.D., M.D.

ABSTRACT

Background Athletes often take androgenic steroids in an attempt to increase their strength. The efficacy of these substances for this purpose is unsubstantiated, however.

Methods We randomly assigned 43 normal men to one of four groups: placebo with no exercise, testosterone with no exercise, placebo plus exercise, and testosterone plus exercise. The men received injections of 600 mg of testosterone enanthate or placebo weekly for 10 weeks. The men in the exercise groups performed standardized weight-lifting exercises three times weekly. Before and after the treatment period, fat-free mass was determined by underwater weighing, muscle size was measured by magnetic resonance imaging, and the strength of the arms and legs was assessed by bench-press and squatting exercises, respectively.

Results Among the men in the no-exercise groups, those given testosterone had greater increases than those given placebo in muscle size in their arms (mean [\pm SE] change in triceps area, 424 ± 104 vs. -81 ± 109 mm²; $P < 0.05$) and legs (change in quadriceps area, 607 ± 123 vs. -131 ± 111 mm²; $P < 0.05$) and greater increases in strength in the bench-press (9 ± 4 vs. -1 ± 1 kg, $P < 0.05$) and squatting exercises (16 ± 4 vs. 3 ± 1 kg, $P < 0.05$). The men assigned to testosterone and exercise had greater increases in fat-free mass (6.1 ± 0.6 kg) and muscle size (triceps area, 501 ± 104 mm²; quadriceps area, 1174 ± 91 mm²) than those assigned to either no-exercise group, and greater increases in muscle strength (bench-press strength, 22 ± 2 kg; squatting-exercise capacity, 38 ± 4 kg) than either no-exercise group. Neither mood nor behavior was altered in any group.

Conclusions Supraphysiologic doses of testosterone, especially when combined with strength training, increase fat-free mass and muscle size and strength in normal men. (N Engl J Med 1996;335:1-7.)

©1996, Massachusetts Medical Society.

ANABOLIC-ANDROGENIC steroids are widely abused by athletes and recreational bodybuilders because of the perception that these substances increase muscle mass and strength,¹⁻⁹ but this premise is unsubstantiated. Testosterone replacement increases nitrogen retention and fat-free mass in castrated animals and hypogonadal men,¹⁰⁻¹⁵ but whether supraphysiologic doses of testosterone or other anabolic-androgenic steroids augment muscle mass and strength in normal men is unknown.¹⁻⁹ Studies of the effects of such steroids on muscle strength have been inconclusive,¹⁶⁻³³ and several reviews have emphasized the shortcomings of the studies.^{1-5,8-10} Some of the studies were not randomized; most did not control for intake of energy and protein; the exercise stimulus was often not standardized; and some studies included competitive athletes whose motivation to win may have kept them from complying with a standardized regimen of diet and exercise.

We sought to determine whether supraphysiologic doses of testosterone, administered alone or in conjunction with a standardized program of strength-training exercise, increase fat-free mass and muscle size and strength in normal men. To overcome the pitfalls of previous studies, the intake of energy and protein and the exercise stimulus were standardized. Because some previous studies had demonstrated significant increases in muscle strength and hyper-

From the Department of Medicine, Charles R. Drew University of Medicine and Science, Los Angeles (S.B., C.C., B.C.); the Exercise Science Laboratory, El Camino College, Torrance, Calif. (T.W.S., T.J.B.); the Department of Medicine, Harbor-UCLA Medical Center, Torrance, Calif. (N.B., J.P., R.C.); and the Department of Public Health, Oregon State University, Corvallis (R.T., A.S.). Address reprint requests to Dr. Bhasin at the Division of Endocrinology, Metabolism and Molecular Medicine, Charles R. Drew University of Medicine and Science, 1621 E. 120th St., MP #2, Los Angeles, CA 90059.

The evolution of superstitious and superstition-like behaviour

Kevin R. Foster^{1,*} and Hanna Kokko²

¹Center for Systems Biology, Harvard University, 7 Divinity Avenue, Harvard, MA 02138, USA

²Laboratory of Ecological and Evolutionary Dynamics, Department of Biological and Environmental Sciences, University of Helsinki, 00014 Helsinki, Finland

Superstitious behaviours, which arise through the incorrect assignment of cause and effect, receive considerable attention in psychology and popular culture. Perhaps owing to their seeming irrationality, however, they receive little attention in evolutionary biology. Here we develop a simple model to define the condition under which natural selection will favour assigning causality between two events. This leads to an intuitive inequality—akin to an amalgam of Hamilton's rule and Pascal's wager—that shows that natural selection can favour strategies that lead to frequent errors in assessment as long as the occasional correct response carries a large fitness benefit. It follows that incorrect responses are the most common when the probability that two events are really associated is low to moderate: very strong associations are rarely incorrect, while natural selection will rarely favour making very weak associations. Extending the model to include multiple events identifies conditions under which natural selection can favour associating events that are never causally related. Specifically, limitations on assigning causal probabilities to pairs of events can favour strategies that lump non-causal associations with causal ones. We conclude that behaviours which are, or appear, superstitious are an inevitable feature of adaptive behaviour in all organisms, including ourselves.

Keywords: optimality theory; behavioural ecology; animal behaviour

1. INTRODUCTION

Although the concept of superstition encompasses a wide range of beliefs and behaviours, most can be united by a single underlying property—the incorrect establishment of cause and effect: ‘a belief or practice resulting from ignorance, fear of the unknown, trust in magic or chance, or a false conception of causation’ (Merriam-Webster online dictionary). In a world increasingly dominated by science, superstitious and indeed religious thinking typically take a back seat in academic affairs. However, superstitions play a central role in many small-scale societies, and indeed remain prevalent in the popular culture of all societies. Why is this? Can science rationalize this seemingly most irrational aspect of human behaviour?

Superstitions receive considerable attention in several fields including popular psychology (Shermer 1998; Vyse 2000; Wheen 2004), philosophy (Scheibe & Sarbin 1965), abnormal psychology (Devenport 1979; Brugger *et al.* 1994; Shaner 1999; Nayha 2002) and medicine (Hira *et al.* 1998; Diamond 2001), which typically frame superstitions as irrational mistakes in cognition. A notable exception, however, is found in the introduction to the popular book of Shermer (1998). This argues that superstitions are the adaptive outcome of a general ‘belief engine’, which evolved to both reduce anxiety (proximate cause) and enable humans to make causal associations (ultimate cause) (Tinbergen 1963; West *et al.* 2007). Specifically, Shermer argued that in making causal associations,

humans are faced with the option to minimize one of two types of statistical error: type I errors whereby they believe a falsehood or type II errors whereby they reject a truth. And as long as the cost of type II errors is high enough, natural selection can favour strategies that frequently make type I errors and generate superstitions (see Beck & Forstmeier (2007) for a similar argument). Our goal here is to explore Shermer's idea that superstitions are adaptive.

Previous biological accounts of superstition have focused upon the classic work of the behavioural psychologist Skinner who reported superstitious behaviour in pigeons (Skinner 1948; Morse & Skinner 1957). In one of his experiments on operant conditioning, Skinner presented the pigeons with food at random intervals and noted that they still displayed ritualized behaviours that he interpreted as superstitious, i.e. the pigeon was behaving as though its actions were causing the food to arrive. However, these behaviours were later reinterpreted as behaviours that improve foraging efficacy (analogous to salivation in Pavlov's dogs), which suggests that the pigeons' behaviour does not correspond to Skinner's intended meaning of superstition (Staddon & Simmelhag 1971; Timberlake & Lucas 1985; Moore 2004). Nevertheless, Skinner's early account is notable in two respects. First, it recognized the possibility of superstition occurring outside the human realm. Second, and linked to this, Skinner emphasized the behavioural aspect of superstition: ‘The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking.’ (Skinner 1948). That is, he focused on there being an incorrect response

* Author for correspondence (kfoster@cgr.harvard.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2008.0981> or via <http://journals.royalsociety.org>.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/312055204>

The Fountain of Youth by Targeting Senescent Cells?

Article in *Trends in Molecular Medicine* · December 2016

DOI: 10.1016/j.molmed.2016.11.006

CITATION

1

READS

669

1 author:



Peter de Keizer

Erasmus MC

21 PUBLICATIONS 474 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Forkhead Box O as a Sensor, Mediator, and Regulator of Redox Signaling [View project](#)

All content following this page was uploaded by [Peter de Keizer](#) on 10 January 2017.

The user has requested enhancement of the downloaded file. All in-text references [underlined in blue](#) are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Cite as: J. D. Boeke *et al.*, *Science*
10.1126/science.aaf6850 (2016).

The Genome Project-Write

Jef D. Boeke,*† George Church,* Andrew Hessel,* Nancy J. Kelley,* Adam Arkin, Yizhi Cai, Rob Carlson, Aravinda Chakravarti, Virginia W. Cornish, Liam Holt, Farren J. Isaacs, Todd Kuiken, Marc Lajoie, Tracy Lessor, Jeantine Lunshof, Matthew T. Maurano, Leslie A. Mitchell, Jasper Rine, Susan Rosser, Neville E. Sanjana, Pamela A. Silver, David Valle, Harris Wang, Jeffrey C. Way, Luhan Yang

*These authors contributed equally to this work.

†Corresponding author. Email: jef.boeke@nyumc.org

The list of author affiliations is available in the supplementary materials.

We need technology and an ethical framework for genome-scale engineering

The Human Genome Project (“HGP-read”) nominally completed in 2004 aimed to sequence the human genome and improve technology, cost, and quality of DNA sequencing (1, 2). It was biology’s first genome-scale project, and at the time was considered controversial by some. Now it is recognized as one of the great feats of exploration, one that has revolutionized science and medicine.

Although sequencing, analyzing, and editing DNA continue to advance at breakneck pace, the capability to construct DNA sequences in cells is mostly limited to a small number of short segments, restricting the ability to manipulate and understand biological systems. Further understanding of genetic blueprints could come from construction of large, gigabase (Gb)-sized animal and plant genomes, including the human genome, which would in turn drive development of tools and methods to facilitate large-scale synthesis and editing of genomes. To this end, we propose the Human Genome Project-Write (HGP-write).

Responsible innovation

Genome synthesis is a logical extension of the genetic engineering tools that have been used safely within the biotech industry for ~40 years and have provided important societal benefits. However, recent technological advancements—e.g., standardized gene parts, whole-genome synthesis, and clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 genome editing technology (3, 4)—are revolutionizing the field (5). Some applications are controversial; human germline editing in particular has raised intense moral debate (6). As human genome-scale synthesis appears increasingly feasible, a coordinated scientific effort to understand, discuss, and apply large-genome engineering technologies is timely. HGP-write will require public involvement and consideration of ethical, legal, and social implications (ELSI) from the start. Responsible innovation requires more than ELSI, though, and involves identifying common goals important to scientists and the wider public through timely and detailed consultation among diverse

stakeholders.

We will enable broad public discourse on HGP-write; having such conversations well in advance of project implementation will guide emerging capabilities in science and contribute to societal decision-making. Through open and ongoing dialogue, common goals can be identified. Informed consent must take local and regional values into account and enable true decision-making on particularly sensitive use of cells and DNA from certain sources. Finally, the highest biosafety standards should guide project work, and safety for lab workers, research participants, and ecosystems should pervade the design process. A priority will be cost reduction of both genome engineering and testing tools to aid in equitable distribution of benefits—e.g., enabling research on crop plants and infectious agents and vectors in developing nations.

To ensure responsible innovation and ongoing consideration of ELSI, a percentage of all research funds could be dedicated to these issues, enabling inclusive decision-making on the topics mentioned above (7). In addition, there should be equitable distribution of any early and future benefits in view of diverse and pressing needs in different global regions. The broad scope and novelty of the project calls for consideration of appropriate regulation alongside development of the science and societal debates. National and international laws and regulations differ and, as in stem cell research, a major burden of responsibility of setting standards rests with the scientists and their community. Existing stem cell research guidelines (8) may serve as a useful template.

From observation to action

The primary goal of HGP-write is to reduce the costs of engineering and testing large (0.1 to 100 billion base pairs) genomes in cell lines by over 1000-fold within 10 years. This will include whole-genome engineering of human cell lines and other organisms of agricultural and public health significance, or those needed to interpret human biological func-

The Selective Value of Bacterial Shape

Kevin D. Young*

*Department of Microbiology and Immunology, University of North Dakota School of Medicine,
 Grand Forks, North Dakota 58202-9037*

INTRODUCTION	662
The Best Arguments	662
The Perfect Example	662
The Perfect Experiment	662
The Imperfect Science	664
To Protect and To Serve	664
NUTRIENT ACCESS	664
Why Are Prokaryotic Cells Small?	665
Theoretical limits	665
Surface-to-volume ratio	665
The diffusion sphere	665
Intracell mixing	666
How Diffusion Affects Cell Shape	666
Contrasting examples	666
Conclusions	667
Morphological Variation	667
Variation with growth rate	667
Filamentation with nutritional status	667
Nutritionally deficient streptococci	668
True to form?	668
Prosthecae as Nutrient Whiskers?	668
Filaments and Blimps	669
Miscellaneous Shape Effects	669
Summary	669
CELL DIVISION AND SEGREGATION	670
Shape Uniformity	670
The Cell Cycle Resists Shape Changes	670
Summary	671
ATTACHMENT	671
Physicochemical Considerations	671
Shear Forces	672
Poles Apart: Polar Localization	673
Cell-Cell Interactions	674
Safety in numbers	674
Antisocial shapes	674
Biofilms: where no one stands alone	674
Summary	675
DISPERSAL	675
Float Like a Butterfly	675
Hovercraft	675
Subterranean Explorers: Geological Transport	676
Rock Solid	677
Summary	677
MOTILITY	678
Energetics of Motility	678
Brownian motion	678
Chemotaxis: efficiencies of stalking	678
Size vise	679
Sleds and saucers	679
Side Effects: Motility Near Surfaces	680
Motility versus Viscosity	680

* Mailing address: Department of Microbiology and Immunology,
 University of North Dakota School of Medicine, Grand Forks, ND
 58202-9037. Phone: (701) 777-2624. Fax: (701) 777-2054. E-mail:
 kyoung@medicine.nodak.edu.

From Aging Research Center, Department of Neurobiology, Care
Sciences, and Society
Karolinska Institutet, Stockholm, Sweden

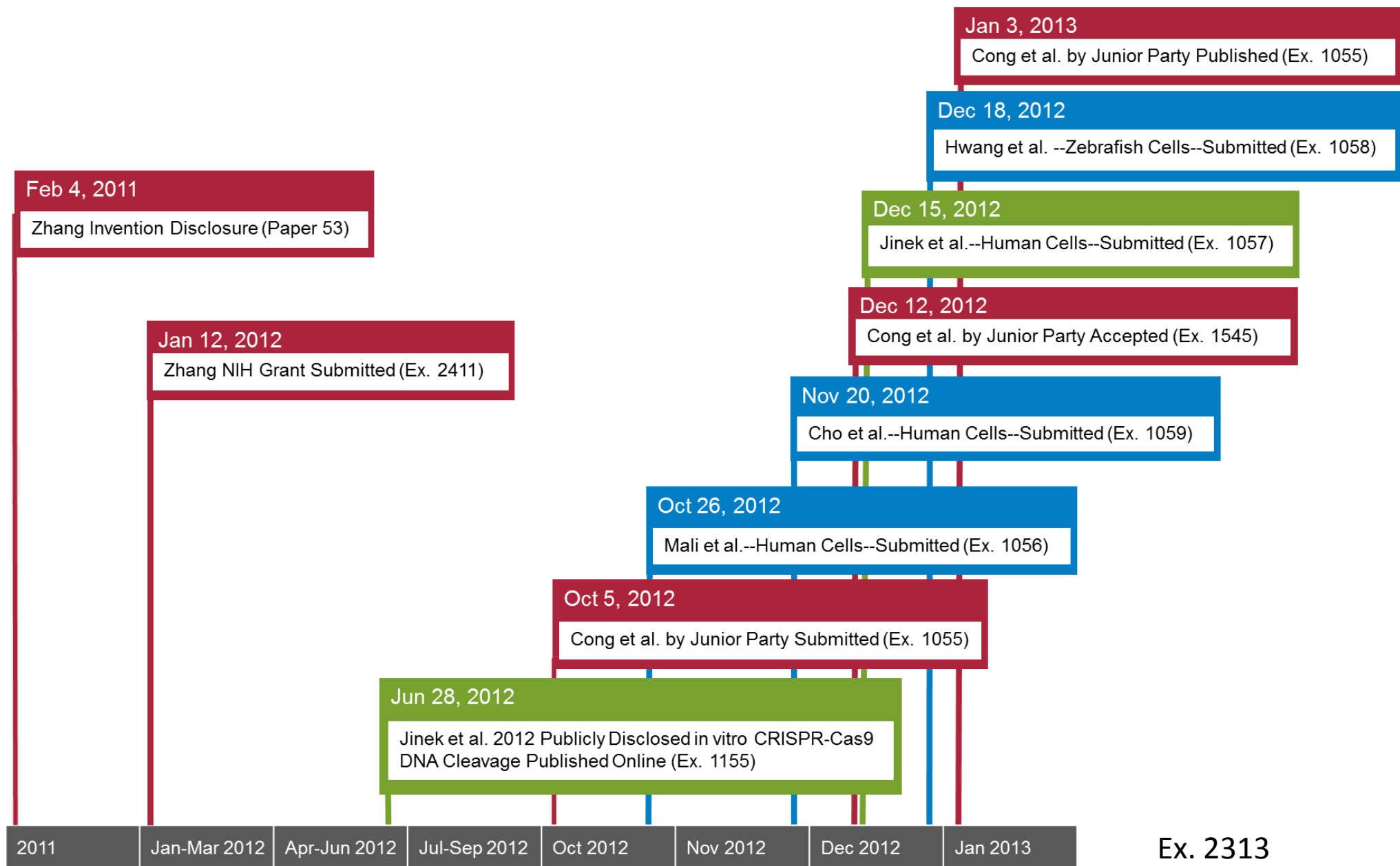
GENETIC AND LIFESTYLE INFLUENCES ON MEMORY, BRAIN STRUCTURE, AND DEMENTIA

Beata Ferencz



**Karolinska
Institutet**

Stockholm 2017



Ex. 2313



JENNER & BLOCK

To Shape a Cell: an Inquiry into the Causes of Morphogenesis of Microorganisms†

FRANKLIN M. HAROLD

Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523

INTRODUCTION	381
STRATEGY OF THE GENES	383
Mutants Defective in Morphogenesis	384
Morphogenetic mutants of filamentous fungi	384
Genetic control of the yeast cell cycle	384
Is Morphogenesis the Expression of a Genetic Program?	386
PERSISTENCE OF FORM	387
Structural Inheritance in Ciliates	387
Gene Mutations Can Alter Large-Scale Order	389
Varieties of Spatial Memory	390
Self-Organization	392
D'ARCY THOMPSON'S LEGACY	393
Self-Assembly: Phage Model of Morphogenesis	393
Surface Tension	394
Surface Stress Theory of Bacterial Morphogenesis	394
Gram-positive cocci	395
Gram-positive rods	396
Gram-negative rods	397
In search of simplicity	398
VECTORIAL PHYSIOLOGY	399
Apical Growth of Fungal Hyphae	399
Extension by polarized secretion	399
Ultrastructure of hyphal tips	402
Shape of hyphal tips	402
So how do hyphal tips grow?	403
Branching and chemotropism	404
Budding and Cell Division in Yeasts	404
Localized and delocalized wall growth	405
Septum formation	407
Germination of Furoid Embryos	407
Mechanical Design and Morphogenesis in Algal Cells	408
How <i>Nitella</i> shapes a tube	408
Ordered cellulose deposition generates diverse forms	409
Questions of causality	412
Searching for a Paradigm	413
SELF-ORGANIZATION	413
Morphogens and Fields	414
Kinetic basis of pattern formation	414
All manner of fields	415
Polarization of the Furoid Embryo	417
Localizing Tips, Branches, and Buds	419
Ionic signals	420
Spatial memory and the cytoskeleton	420
Patterns and Fields in the Cortex of Ciliates	421
DEVELOPMENTAL DYNAMICS OF MICROBIAL CELLS	423
ACKNOWLEDGMENTS	424
LITERATURE CITED	424

The Word is an edged tool, equivalent to the sculptor's chisel, for chipping form out of the obstinate material of consciousness.

Philip Glazebrook
Journey to Kars

† Dedicated to Arthur L. Koch, as a token of admiration and friendship, on the occasion of his 65th birthday.

INTRODUCTION

Of cellular morphogenesis it can justly be said that we know much but understand little. Thanks to the labors of biologists over many generations, a huge body of literature now records the form, anatomy, and life cycle of innumerable single-celled creatures, procaryotes as well as eucaryotes. However, the student of morphology will be hard put

Bacterial contraception „The liquid Condom“

Rüdiger Trojok, Copenhagen 02.04.2013

For questions and more details please send an email to trojok@openbioprojects.net
I can also recommend a list of related publications

If you want to use my work further, please invite me to contribute.
This project is ment to become an international open source drug discovery cooperation.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.



TumbleBit: An Untrusted Bitcoin-Compatible Anonymous Payment Hub

Ethan Heilman, Leen Alshenibr, Foteini Baldimtsi, Alessandra Scafuro, Sharon Goldberg
Boston University

Abstract—This paper presents *TumbleBit*, a new unidirectional unlinkable payment hub that is fully compatible with today’s Bitcoin protocol. TumbleBit allows parties to make fast, anonymous, off-blockchain payments through an untrusted intermediary called the Tumbler. TumbleBit’s anonymity properties are similar to classic Chaumian eCash: no one, not even the Tumbler, can link a payment from its payer to its payee. Every payment made via TumbleBit is backed by bitcoins, and comes with a guarantee that Tumbler can neither violate anonymity, nor steal bitcoins, nor “print money” by issuing payments to itself. We prove the security of TumbleBit using the real/ideal world paradigm and the random oracle model. Security follows from the standard RSA assumption and ECDSA unforgeability. We implement TumbleBit, mix payments from 800 users and show that TumbleBit’s off-blockchain payments can complete in seconds.

I. INTRODUCTION

One reason for Bitcoin’s initial popularity was the perception of anonymity. Today, however, the sheen of anonymity has all but worn off, dulled by a stream of academic papers [37], [49], and a blockchain surveillance industry [31], [26], that have demonstrated weaknesses in Bitcoin’s anonymity properties. As a result, a new market of anonymity-enhancing services has emerged [43], [22], [39], [1]; for instance, 1 million USD in bitcoins are funneled through JoinMarket each month [43]. These services promise to mix bitcoins from a set of *payers* (aka, input Bitcoin addresses \mathcal{A}) to a set of *payees* (aka, output bitcoin addresses \mathcal{B}) in a manner that makes it difficult to determine which payer transferred bitcoins to which payee.

To deliver on this promise, anonymity must also be provided in the face of the anonymity-enhancing service itself—if the service knows exactly which payer is paying which payee, then a compromise of the service leads to a total loss of anonymity. Compromise of anonymity-enhancing technologies is not unknown. In 2016, for example, researchers found more than 100 Tor nodes snooping on their users [45]. Moreover, users of mix services must also contend with the potential risk of “exit scams”, where an established business takes in new payments but stops providing services. Exit scams have been known to occur in the Bitcoin world. In 2015, a Darknet Marketplace stole 11.7M dollars worth of escrowed customer bitcoins [51], while btcmixers.com mentions eight different scam mix services. Thus, it is

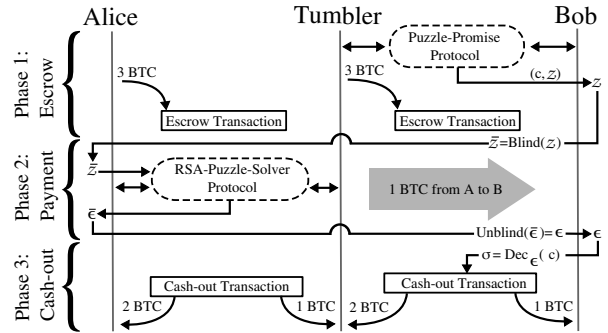


Fig. 1. Overview of the TumbleBit protocol.

crucial that anonymity-enhancing services be designed in a manner that prevents bitcoin theft.

TumbleBit: An unlinkable payment hub. We present TumbleBit¹, a *unidirectional unlinkable payment hub* that uses an *untrusted* intermediary, the *Tumbler* \mathcal{T} , to enhance anonymity. Every payment made via TumbleBit is backed by bitcoins. We use cryptographic techniques to guarantee Tumbler \mathcal{T} can neither violate anonymity, nor steal bitcoins, nor “print money” by issuing payments to itself. TumbleBit allows a payer Alice \mathcal{A} to send fast off-blockchain payments (of denomination one bitcoin) to a set of payees ($\mathcal{B}_1, \dots, \mathcal{B}_Q$) of her choice. Because payments are performed off the blockchain, TumbleBit also serves to scale the volume and velocity of bitcoin-backed payments. Today, on-blockchain bitcoin transactions suffer a latency of ≈ 10 minutes. Meanwhile, TumbleBit payments are sent off-blockchain, via the Tumbler \mathcal{T} , and complete in seconds. (Our implementation completed a payment in 1.2 seconds, on average, when \mathcal{T} was in New York and \mathcal{A} and \mathcal{B} were in Boston.)

TumbleBit Overview. TumbleBit replaces on-blockchain payments with off-blockchain puzzle solving, where Alice \mathcal{A} pays Bob \mathcal{B} by providing \mathcal{B} with the solution to a puzzle. The puzzle z is generated through interaction between \mathcal{B} and \mathcal{T} , and solved through an interaction between \mathcal{A} and \mathcal{T} . Each time a puzzle is solved, 1 bitcoin is transferred from Alice \mathcal{A} to the Tumbler \mathcal{T} and finally on to Bob \mathcal{B} .

¹ Proof-of-concept implementation of TumbleBit as a classic tumbler: <https://github.com/BUSEC/TumbleBit/>

Capital Reporting Company
Noninvasive Neurostimulation Devices & Cognitive Function (11/20/2015)

1

FOOD AND DRUG ADMINISTRATION (FDA)
CENTER FOR DEVICES AND RADIOLOGICAL HEALTH (CDRH)

PUBLIC MEETING ON
NONINVASIVE NEUROSTIMULATION DEVICES & COGNITIVE FUNCTION

Friday, November 20, 2015

FDA White Oak Campus
10903 New Hampshire Avenue
Bldg. 31, Great Room A (W031)
Silver Spring, MD 20993

Reported by: Michael Farkas
Capital Reporting Company

Group B: 2015-2016

Executive summary of the unit

(In case of units structured by team or theme, fill this 'executive summary' by team or theme)

Research unit

Team 01 : Structural Microbiology - Pedro ALZARI

Unit name: Structural Microbiology

Name of the unit for the current contract: Structural Microbiology

Name of the unit for the future contract: Structural Microbiology

Unit workforce (at the start of the current contract; please specify if the unit was set up during the evaluation period).

The team was created in 1998, and re-created after evaluation by an external committee and the scientific council of the Institut Pasteur as a full scientific Pasteur Unit in February 2010. At the time, it was composed of 5 researchers; 5 technicians, engineers and other staff; 3 post-docs and 3 doctoral students.

Staff who have left the unit during the current contract (and number of total months spent in the unit during this period).

2 staff scientists (36 months), 3 technicians IP (90 months), 2 post-docs (24 months), 4 PhD students (144 months).

Number of recruitments carried out during the period in question and where the staff come from

4 staff scientists (2 DR2 CNRS, 1 PR1 U.Paris7, 1 MCF U.Paris7), 1 engineer IP, 2 technicians IP, 3 post-docs (EMBO, FRM, ANR & CE fellowships), 1 PhD student (co-tutelle)

Research products and achievements over the previous period (1 January 2010 - 30 June 2015):

- Major progress has been achieved in the structural and functional studies of mycobacterial protein kinases (PknG) and two-component systems (CpxA), the cell wall glycosyltransferase PimA, mycobacterial DNA gyrase and *Plasmodium* SUB1 protease. See below for a representative list of major publications describing these results.

Quantitative overview of the unit's publications.

Team members have produced 79 publications during the evaluation period (10 in journals with IF>9, and 30 in journals with 4<IF<9)

Please state the unit's 5 major publications (giving their title and underlining the name of any unit members in the event of joint publications).

- Wagner T, Bellinzoni M, Wehenkel A, O'Hare HM, Alzari PM (2011) Functional plasticity and allosteric regulation of alpha-ketoglutarate decarboxylase in central mycobacterial metabolism. *Chem. Biol.* **18**, 1011-1020.
- Mechaly AE, Sassoon N, Betton JM, Alzari PM (2014) Segmental helical motions and dynamical asymmetry modulates histidine kinases autophosphorylation. *PLoS Biol.* **12**, e1001776.
- Giganti D, Bouillon A, Tawk L, Robert F, Martinez M, Crublet E, Weber P, Girard-Blanc C, Petres S, Haouz A, Hernandez JF, Mercereau-Puijalon O, Alzari PM, Barale JC (2014) A novel *Plasmodium*-specific prodomain fold regulates the malaria drug target SUB1 subtilase. *Nat. Commun.* **5**, 4833.
- Lisa MN, Gil M, André-Leroux G, Barilone N, Durán R, Biondi RM, Alzari PM (2015) Molecular basis of the activity and the regulation of the eukaryotic-like S/T protein kinase PknG from *Mycobacterium tuberculosis*. *Structure* **23**, 1039-1048. Previewed by Reckel S, Hantschel O. (2015) Kinase regulation in *Mycobacterium tuberculosis*:

Using high-throughput barcode sequencing to efficiently map connectomes

Ian D Peikon^{*1,2}, Justus M Kebschull^{*1,2}, Vasily V Vagin^{*2}, Diana I Ravens², Yu-Chi Sun², Eric Brouzes^{3,4}, Ivan R. Corrêa Jr.⁵, Dario Bressan^{1,2,6}, Anthony M Zador²

¹Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

²Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

³Department of Biomedical Engineering, Stony Brook University, Stony Brook, New York 11794, USA

⁴Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA

⁵New England Biolabs, Inc., Ipswich, Massachusetts 01938, USA

⁶Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, United Kingdom

*these authors contributed equally to the work

Correspondence and requests should be addressed to Anthony Zador, zador@cshl.edu.

Abstract

The function of a neural circuit is determined by the details of its synaptic connections. At present, the only available method for determining a neural wiring diagram with single synapse precision—a “connectome”—is based on imaging methods that are slow, labor-intensive and expensive. Here we present SYNseq, a method for converting the connectome into a form that can exploit the speed and low cost of modern high-throughput DNA sequencing. In SYNseq, each neuron is labeled with a unique random nucleotide sequence—an RNA “barcode”—which

Using high-throughput barcode sequencing to efficiently map connectomes

Ian D. Peikon^{1,2,†}, Justus M. Kebschull^{1,2,†}, Vasily V. Vagin^{2,†}, Diana I. Ravens², Yu-Chi Sun², Eric Brouzes^{3,4}, Ivan R. Corrêa, Jr⁵, Dario Bressan^{1,2,6} and Anthony M. Zador^{2,*}

¹Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA, ²Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA, ³Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794, USA, ⁴Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY 11794, USA, ⁵New England Biolabs, Inc., Ipswich, MA 01938, USA and ⁶Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, UK

Received February 11, 2017; Revised March 20, 2017; Editorial Decision April 10, 2017; Accepted April 13, 2017

ABSTRACT

The function of a neural circuit is determined by the details of its synaptic connections. At present, the only available method for determining a neural wiring diagram with single synapse precision—a ‘connectome’—is based on imaging methods that are slow, labor-intensive and expensive. Here, we present SYNseq, a method for converting the connectome into a form that can exploit the speed and low cost of modern high-throughput DNA sequencing. In SYNseq, each neuron is labeled with a unique random nucleotide sequence—an RNA ‘barcode’—which is targeted to the synapse using engineered proteins. Barcodes in pre- and postsynaptic neurons are then associated through protein-protein crosslinking across the synapse, extracted from the tissue, and joined into a form suitable for sequencing. Although our failure to develop an efficient barcode joining scheme precludes the widespread application of this approach, we expect that with further development SYNseq will enable tracing of complex circuits at high speed and low cost.

INTRODUCTION

The brain is extraordinarily complex, consisting of myriad neurons connected by even larger numbers of synapses. Disruption of these connections contributes to many neuropsychiatric disorders including autism, schizophrenia and depression. Understanding how the brain processes information and produces actions requires knowledge of both the structure of neural circuits, and of the patterns of neural activity. Sophisticated technology for recording ever-larger numbers of neurons is now widely available and is providing

unprecedented insight into the physiological responses of brain circuits (1,2). In contrast, circuit-mapping technologies with synaptic resolution remain very slow, expensive and labor intensive.

Mapping neural connectivity is traditionally viewed as a problem of microscopy. Electron microscopy (EM) allows direct imaging of synaptic contacts between neurons, so in principle circuit mapping with EM is trivial. In practice, however, it is complicated by a mismatch of scales. Imaging synapses requires nanometer resolution. In contrast, brain circuits span macroscopic distances, from millimeters in small organisms to tens of centimeters in humans. Circuit reconstruction using EM thus needs to bridge these scales, resulting in the requirement that thin axonal processes be traced across thousands of sections at an exceedingly low error rate. For example, for a 5 mm axon, and EM sections 50 nm thick, the required accuracy per single axon section would need, under simple assumptions, to exceed 99.999% in order to achieve a 36% chance of assigning a correct connection. Several major efforts are underway to increase the throughput and autonomy of EM and have resulted in impressive improvements of speed and scale (3–10). Unfortunately, most of these advancements require very expensive instruments, and the challenge of automatically tracing axonal processes through EM stacks remains unsolved.

Electrophysiological approaches allow probing the connectivity of pairs or small groups of nearby neurons (11–13). These efforts have uncovered elements of high-order structure within neural circuits, as well as spatially intertwined but non-interconnected networks (12,14). However, such physiological methods are labor-intensive, and cannot readily be scaled for the analysis of larger neural circuits or a full nervous system (see however ref (15)).

We have been developing high-throughput sequencing as a fast and efficient alternative to microscopy or physiology for probing neuroanatomical connectivity (16,17). To

*To whom correspondence should be addressed. Tel: +1 516 367 6950; Email: zador@cshl.edu

†These authors contributed equally to the work as first authors.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/317096535>

Visualization and characterization of individual type III protein secretion machines in live bacteria

Article in *Proceedings of the National Academy of Sciences* · May 2017

DOI: 10.1073/pnas.1705823114

CITATIONS

0

READS

79

4 authors, including:



Yongdeng Zhang

Yale University

37 PUBLICATIONS **308 CITATIONS**

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Vesicle traffic [View project](#)



Super-resolution microscopy [View project](#)

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/261765493>

Whole-Brain Imaging with Single-Cell Resolution Using Chemical Cocktails and Computational Analysis

Article in *Cell* · April 2014

Impact Factor: 32.24 · DOI: 10.1016/j.cell.2014.03.042 · Source: PubMed

CITATIONS

103

READS

281

16 authors, including:



[Kazuki Tainaka](#)

RIKEN

51 PUBLICATIONS 812 CITATIONS

SEE PROFILE



[Dimitri Perrin](#)

Queensland University of Technology

36 PUBLICATIONS 219 CITATIONS

SEE PROFILE



[Hirotaka Onoe](#)

RIKEN

186 PUBLICATIONS 2,414 CITATIONS

SEE PROFILE



[Hideo Yokota](#)

RIKEN

134 PUBLICATIONS 582 CITATIONS

SEE PROFILE

WHO STEERS WHO STEERS?

A NOTE ON IDENTIFYING VULNERABLE MORAL PROPENSITIES

Steven Kaas & Steve Rayhawk
stevenkaas@gmail.com

There are a variety of processes that steer the future; that is, they move it toward certain states and away from others dynamically, with changing behaviors in response to changing conditions. Our decisions now don't just steer the future directly, but influence what the major steering processes in the future will be. Certain dangers attend such a project. Often the replacement of part of an ecosystem or a society with an engineered substitute, designed on the basis of a partial understanding, meets with severe drawbacks from unrecognized missing functionality that was demolished or displaced. In the same way, a project to take into hand the steering of the future, to fulfill its potential according to some moral vision, risks demolishing or displacing some unrecognized steering processes that generated and preserved the correctness of the moral sense behind the vision. While this risk cannot be avoided entirely, it can be mitigated by developing better tools for identifying or avoiding interference with unrecognized steering processes. In light of modern physical ontology, and in light of the abstractness of some of the plausible processes (such as the relative market success of firms, or historical evolutionary selection), we suggest that such tools should be rooted in a very conservatively general theoretical framework, based on finding factorizations of the world's state space into potential "steering" and "steered" state subspaces, with partially coupled dynamics.

Quis custodiet ipsos custodes? (Who will guard the guardians themselves?)

Juvenal, Roman poet, on the problem of engineering stably moral governance

Yo dawg, I heard you like cars, so we put a car in your car, so you can drive while you drive.

Xzibit, television presenter, on upgrading a car to contain a racing simulation

Just as our evolutionary history has selected for brains that select courses of action expected to achieve goals previously associated with reproductive success, and just as the authors of constitutions have written laws designed to cause good laws to be written, when we attempt to think about "steering the future", we engage in a project to "drive while we drive": to strategically influence the main factors that will strategically influence the future.



(43) International Publication Date
13 April 2017 (13.04.2017)

(10) International Publication Number
WO 2017/062311 A1

(51) International Patent Classification:

A61P 39/00 (2006.01) A23L 33/175 (2016.01)
A23L 33/13 (2016.01) A61K 36/48 (2006.01)

(21) International Application Number:

PCT/US2016/055173

(22) International Filing Date:

3 October 2016 (03.10.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/238,338 7 October 2015 (07.10.2015) US

(72) Inventor; and

(71) Applicant : HUIZENGA, Joel [US/US]; 354 Gravilla Street, La Jolla, California 92037 (US).

(74) Agents: CURFMAN, Christopher L. et al.; Meunier Carlin & Curfman LLC, 999 Peachtree Street, NE, Suite 1300, Atlanta, Georgia 30309 (US).

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, — with international search report (Art. 21(3))

(54) Title: RESETTING BIOLOGICAL PATHWAYS FOR DEFENDING AGAINST AND REPAIRING DETERIORATION FROM HUMAN AGING

(57) Abstract: Compositions for addressing one or more of the effects of aging are described. The compositions comprise a first component comprising repair system activator(s) such as nicotinamide adenine dinucleotide (NAD⁺), nicotinamide mononucleotide (NMN), nicotinamide riboside (NR), nicotinic acid adenine mononucleotide (NaMN), nicotinic acid adenine dinucleotide (NaAD), nicotinic acid riboside (NAR), 1-methylnicotinamide (MNM), cyclic adenosine monophosphate (cAMP) and combinations thereof; a second component comprising methyl donor(s) such as S-5'-adenosyl-L-methionine (SAM), methionine, betaine, choline, folate, vitamin B12, or combinations thereof; and a third component comprising antioxidant defense activators such as H₂O₂, N₂S, NaSH, N₂S, and several others, including combinations thereof. Methods of administering the disclosed compositions or separate formulations of repair system activator, methyl donors, and antioxidant defense activators are also disclosed.



WO 2017/062311 A1

Enhanced biomass production of duckweeds by inoculating a plant growth-promoting bacterium, *Acinetobacter calcoaceticus* P23, in sterile medium and non-sterile environmental waters

T. Toyama, M. Kuroda, Y. Ogata, Y. Hachiya, A. Quach, K. Tokura, Y. Tanaka, K. Mori, M. Morikawa and M. Ike

ABSTRACT

Duckweed offers the promise of co-benefit culture combining water purification with biomass production. *Acinetobacter calcoaceticus* P23 is a plant growth-promoting bacterium isolated from a duckweed, *Lemna aequinoctialis*. This study quantified its growth-promoting effect on three duckweeds (*L. aoukikusa*, *Lemna minor*, and *Spirodela polyrhiza*) in sterile Hoagland solution and evaluated its usefulness in duckweed culture under non-sterile conditions. P23 promoted growth of three duckweeds in sterile Hoagland solution at low to high nutrient concentrations (1.25–10 mg NO₃-N/L and 0.25–2.0 mg PO₄-P/L). It increased the biomass production of *L. aequinoctialis* 3.8–4.3-fold, of *L. minor* 2.3–3.3-fold, and of *S. polyrhiza* 1.4–1.5-fold after 7 days compared with uninoculated controls. P23 also increased the biomass production of *L. minor* 2.4-fold in pond water and 1.7-fold in secondary effluent of a sewage treatment plant under non-sterile conditions at lab-scale experiments. P23 rescued *L. minor* from growth inhibition caused by microorganisms indigenous to the pond water. The results demonstrate that the use of P23 in duckweed culture can improve the efficiency of duckweed biomass production, and a positive effect of P23 on duckweed-based wastewater treatment be assumed.

Key words | *Acinetobacter calcoaceticus*, biomass production, duckweed, plant growth-promoting bacteria

T. Toyama (corresponding author)

Y. Tanaka

K. Mori

Graduate Faculty of Interdisciplinary Research,
University of Yamanashi,
4-3-11 Takeda, Kofu,
Yamanashi 400-8511,
Japan
E-mail: ttohyama@yamanashi.ac.jp

M. Kuroda

Y. Hachiya

A. Quach

K. Tokura

M. Ike

Division of Sustainable Energy and Environmental
Engineering,
Osaka University,
2-1 Yamadaoka,
Suita, Osaka 565-0871,
Japan

Y. Ogata

National Institute for Environmental Studies,
16-2 Onogawa,
Tsukuba, Ibaraki 305-8506,
Japan

M. Morikawa

Division of Biosphere Science, Graduate School of
Environmental Science,
Hokkaido University,
Kita-10 Nishi-5,
Kita-ku, Sapporo 060-0810,
Japan

INTRODUCTION

Duckweeds are the smallest and fastest-growing aquatic plants, classified in the Araceae subfamily Lemnoideae, which includes five genera: *Lemna*, *Landoltia*, *Spirodela*, *Wolffia*, and *Wolffiella* (Landolt 1986). They are useful agents for removing nitrogen (N) and phosphorus (P) from municipal (Dalu & Ndamba 2003; Ran *et al.* 2004), livestock

(Cheng *et al.* 2002; Xu & Shen 2011), and industrial (Ozengin & Elmaci 2007) wastewaters because of their high growth rate and high nutrient uptake capabilities. They are also used to clean up eutrophied water bodies (Ansari & Khan 2008, 2009). Wastewater treatment and purification of polluted waters using duckweeds offers successful, cost-effective, low-energy, and environmentally friendly options around the world.

In addition, duckweeds have recently attracted attention as a good alternative feedstock for biofuel production owing to their high growth rate and high starch accumulation

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).



Voltage imaging with genetically encoded indicators

Yongxian Xu^{1,2}, Peng Zou^{1,2} and Adam E Cohen^{3,4}

Membrane voltages are ubiquitous throughout cell biology. Voltage is most commonly associated with excitable cells such as neurons and cardiomyocytes, although many other cell types and organelles also support electrical signaling. Voltage imaging *in vivo* would offer unique capabilities in reporting the spatial pattern and temporal dynamics of electrical signaling at the cellular and circuit levels. Voltage is not directly visible, and so a longstanding challenge has been to develop genetically encoded fluorescent voltage indicator proteins. Recent advances have led to a profusion of new voltage indicators, based on different scaffolds and with different tradeoffs between voltage sensitivity, speed, brightness, and spectrum. In this review, we describe recent advances in design and applications of genetically-encoded voltage indicators (GEVIs). We also highlight the protein engineering strategies employed to improve the dynamic range and kinetics of GEVIs and opportunities for future advances.

Addresses

¹ Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

² Peking-Tsinghua Center for Life Sciences, PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

³ Departments of Chemistry and Chemical Biology and of Physics, Harvard University, Cambridge, MA 02138, USA

⁴ Howard Hughes Medical Institute

Corresponding authors: Zou, Peng (zoupeng@pku.edu.cn), Cohen, Adam E (cohen@chemistry.harvard.edu)

Current Opinion in Chemical Biology 2017, 39:1–10

This review comes from a themed issue on **Molecular imaging**

Edited by **Xing Chen** and **Yanyi Huang**

<http://dx.doi.org/10.1016/j.cbpa.2017.04.005>

1367-5931/© 2017 Elsevier Ltd. All rights reserved.

Introduction

Membrane voltage is an important signal that affects many fundamental aspects of cellular physiology. The transmembrane electric field perturbs the energy landscape of biomolecules embedded in the lipid membrane, including ion channels, G protein-coupled receptors and membrane-associated enzymes [1,2], and is affected in turn by the dynamics of voltage- and ligand-gated ion channels, as well as electrogenic transporters and pumps.

While bioelectrical signaling is most commonly associated with neurons and cardiomyocytes, membrane voltage also forms the basis for signaling in other cell types including bacteria [3[•],4], fungi [5] and plants [6]. Compared with electrode-based techniques such as patch clamp electrophysiology, optical measurements offer spatial resolution, non-invasiveness, ease of operation, and high measurement throughput. Much effort has gone toward development of fluorescent genetically encoded voltage indicators (GEVIs), though current-generation GEVIs are still far from optimal. Here we review recent advances in GEVI designs and their applications in voltage imaging of bioelectric phenomena. We hope that this article will help the reader select the right GEVI for his or her application, and will inspire novel ideas for better GEVIs.

The family of GEVIs

Fluorescent protein-based GEVIs

The history of GEVI development is marked by occasional introduction of qualitatively new scaffold designs, followed by periods of iterative optimization and refinement (Figure 1). The earliest GEVI designs comprised naturally occurring ion channel voltage sensor domains fused to fluorescent proteins (FPs) (FlaSh [7], SPARC [8]) or Förster resonance energy transfer (FRET) FP pairs (VSFP1 [9]). GEVIs based on ion channel scaffolds typically exhibited modest voltage sensitivity (<5% $\Delta F/F$ per 100 mV) and slow kinetics (typically 10–200 ms), and many suffered from poor membrane trafficking [10].

The transmembrane voltage-sensing domain (VSD) of the ascidian *Ciona intestinalis* voltage-sensing phosphatase (Ci-VSP) [11], turned out to be a good transducer for many GEVIs (VSFP2.x [12,13], Mermaid [14,15], Butterfly [16], VSFP-CR [17], ArcLight [18], ASAP [19^{••},20[•]], Bongwoori [21], FlicR1 [22]). GEVIs based on Ci-VSP had much improved membrane trafficking, presumably due to the monomeric nature of the VSD. Ci-VSP VSDs fused with FRET pairs showed good voltage sensitivity ($\Delta R/R$ approaching 48% per 100 mV for Mermaid2 [14,15]) and have been applied to report large-area membrane voltage fluctuations *in vivo* (optical EEG with Butterfly 1.2 [16]). While the ratiometric nature of FRET signals is useful for canceling motion and blood flow artifacts, the broad spectrum of FRET GEVIs often precludes multiplex imaging with other sensors. Surprisingly, a VSD fused with a single FP also reported membrane voltage (VSFP3.1), suggesting that a mechanism other than FRET was responsible for some of the

The glEnd() of Zelda

Dr. Tom Murphy VII Ph.D.*

1 April 2016

Abstract

3D ZELDA

Keywords: small keys, boss keys, dungeon keys

1 Introduction

1986. Hyrule. The Legend of *frickin'* Zelda for the Nintendo *freakin'* Entertainment System. Need I say more? A *god damn* gold cartridge. Fortunes made just from melting the gold cartridge down to make gold teeth grilles, after carefully extracting and preserving the even more precious ROM inside. A die-cut hole in the box so that you could get a peek of the cartridge and presage that you were getting some solid gold. A die cut little window that you could palpate through the wrapping paper on Christmas Eve, presaging some epic thumb blisters in store for the coming weeks. Koji *freckin'* Kondo. Koji Kondo whipping up a nice 8-bit arrangement of Boléro as the theme music until realizing at the last minute that this music was copyrighted¹ and so instead composing its epic theme in *one day??*

A gold cartridge that contained ROMs and a little swallowable battery to keep the onboard SRAM powered up so that it could retain your epic save game. A battery designed to last 70 years. Nothing could cause you to lose your save game, even once you were half way through the Second Quest. Unless your little brother starts a completely new game and saves over your slot, earning him one of the most righteously deserved clobberings this side of Inigo Montoya. Saves right on top of your slot with a player called just A . Right over your



Figure 1: Technical diagram of mathematical equations.

*Copyright © 2016 the Regents of the Wikipia Foundation. Appears in SIGBOVIK 2016 with the danger of going alone of the Association for Computational Heresy; *IEEEEEE!* press, Verlag-Vergag volume no. 0x2016. 255 Rupees

¹Perhaps ironic, since Ravel's Boléro itself was composed as a result of Ravel getting cop-blocked.[1] And surely some Zelda knock-off since then has included a Muzak ersatz of the Zelda theme! What is the longest documented sequence of compositions due to Copyright restrictions?

slot, erasing it, and hasn't even picked up the *sword* yet. All you have to do is step on the black square in the first room and the guy gives you the sword and then dies. Jeez pooleaze. Saves right over yours with no sword even though there are TWO OTHER UNUSED SLOTS and you even wisely put your game in the third slot *exactly to avoid this kind of calamity*.

Bobtail: A Proof-of-Work Target that Minimizes Blockchain Mining Variance

George Bissias Brian N. Levine

College of Information and Computer Sciences, UMass Amherst

ABSTRACT

Blockchain systems are designed to produce blocks at a constant average rate. The most popular systems currently employ a Proof of Work (PoW) algorithm as a means of creating these blocks. Bitcoin produces, on average, one block every 10 minutes. An unfortunate limitation of all deployed PoW blockchain systems is that the time between blocks has high variance. For example, 5% of the time, Bitcoin’s inter-block time is at least 40 minutes. This variance impedes the consistent flow of validated transactions through the system. We propose an alternative process for PoW-based block discovery that results in an inter-block time with significantly lower variance. Our algorithm, called Bobtail, generalizes the current algorithm by comparing the mean of the k lowest order statistics to a target. We show that the variance of inter-block times decreases as k increases. If our approach were applied to Bitcoin, about 80% of blocks would be found within 7 to 12 minutes, and nearly every block would be found within 5 to 18 minutes; the average inter-block time would remain at 10 minutes. The cost of our approach is a larger block header.

1 INTRODUCTION

Blockchain systems are designed to produce blocks at a constant average rate. The most popular systems currently employ a *Proof of Work (PoW)* algorithm as a means of creating these blocks. Bitcoin [25] (and Bitcoin Cash [1]) produce, on average, one block every 10 minutes, and will self-adjust the difficulty of producing a block every two weeks if too many or too few have been produced. Unfortunately, a limitation of all deployed PoW blockchain systems is that the time between blocks has high variance and the distribution of inter-block times has a very long tail. For example, 5% of the time, Bitcoin’s inter-block time is at least 40 minutes. This variance impedes the consistent flow of validated transactions through the system.

The high inter-block time variance is a direct consequence of the PoW algorithms that are at the core of blockchains, including Bitcoin [1, 25], Litecoin [2], and Ethereum [17]. In all these systems, generally, the miners repeatedly craft block headers by changing a nonce until the hash of the header is less than a target value t . In other words, the hash of each header is a sample taken randomly from a discrete uniform distribution that ranges between $[0, S]$, where $S = 2^b - 1$

and typically $b = 256$. A block is discovered when the *first order statistic* (i.e., the minimum value) of all sampled values is less than target $0 < t < S$.

In this paper, we propose an alternative process for PoW-based block discovery that results in an inter-block time with significantly lower variance. Our algorithm generalizes the current algorithm by comparing the mean of the k lowest order statistics to a target. We show that the variance of inter-block times decreases as k increases. For example, if our approach were applied to Bitcoin, about 80% of blocks would be found within 7 to 12 minutes, and nearly every block would be found within 5 to 18 minutes; the average inter-block time would remain at 10 minutes. The cost of our approach is a larger block header. We call our approach *Bobtail*¹ mining.

Problem Statement. Consider a fixed interval of time during which the entire network produces θ hashes generating a sequence of hash values $\mathbf{Z} = Z_1, \dots, Z_\theta$. Let Z be an arbitrary random variable from the sequence \mathbf{Z} ; note that $Z \sim \text{Uniform}(0, S)$. Define V_i to be the i th lowest order statistic of \mathbf{Z} , i.e. $V_i = Z_{(i)}$ in standard notation. And let random variable W_k be the mean of the k lowest order statistics:

$$W_k = \frac{1}{k} \sum_{i=1}^k V_i. \quad (1)$$

W_k constitutes the collective mining proof (*proof*, for short) for the entire network. Our Bobtail mining criterion says that a new block is discovered when a realized value of W_k meets the target t :

$$w_k \leq t. \quad (2)$$

Notably, this approach is a generalization of current systems, which are the special case of $k = 1$.

Our primary goals are therefore to show, given values of $k > 1$, that: (i) there is a significantly reduced inter-block time variance; and (ii) the costs are relatively small, which include an increase in block header size and a slight increase in network traffic.

Contributions.

- We derive the statistical characteristics of our approach and validate each empirically. For example,

¹A *bobtail* refers to an animal’s tail that is unusually short or is missing completely [9].

Layer: Layer 2 ideas

Title: Multi-hop payment packetization on Lightning Network channel

Author: Takaya Imai <takaya.imai@unitedbitcoiners.com, takaya.imai@frontier-ptnrs.com>

[Abstract]

Lighting Network(LN) was invented by Joseph Poon and Thaddeus Dryja in 2015 \cite{ln}. LN makes Bitcoin very scalable by lower transaction fee and rapid bitcoin transfer. There are four major LN products \cite{products}, lit, lnd, c-lightning and eclair, and RFC \cite{bolt}. These has been developped steadily and LN can be extended and connected (atomic swap) to other blockchains simply.

LN Routing algorisms like flare \cite{flare} and gossip is proposed and developed but LN has two problems.

One is that it is easy to be centralized. Some whose channel has large money as a deposit can dominate a whole LN. This depends on routing algorism though.

Flare is one of proposals for a routing algorism but nothing about the centralization. The other is that nodes in the middle have a possibility to take a long time to get money if a preimage does not propagate successfully.

I propose one solution. That is a packet-like payment.

[Anti-centralization]

This is a way to reduce centralization of LN and provide higher processing ability of payment by distributed nodes on whole LN.

Consider to send 10,000 satoshis from Alice to Bob on the following channel network.

```
Alice - Coulomb - Bob
      \          /
      Dirac
      \          /
      Einstein-
      \          /
      Faraday
```

Alice can send at once through Coulomb but this case promote a centralization because only nodes which have enough deposit on a channel can become one of nodes on a route. So Alice divides a big payment into small payments such as 10,000 satoshis into 2,500 satoshis * 4 etc.

Even if a channel between Dirac and Bob has 4,000 satoshis as a direction of Dirac to Bob, this channel can join this payment.

I think smaller amount is better but it might need excess payments so it is important to keep a balance between process overhead loss and decentralization.

[Payment damage reduction]

This is a way to reduce damage in case that a preimage propagation problem occurs.

Consider to send 10,000 satoshis from Alice to Bob on the following channel network.

```
Alice - Coulomb - Dirac - Bob
      \          /
```


Catch-up Mining

CYRIL GRUNSPAN

Léonard de Vinci Pôle Universitaire
Research Center

Paris-La Défense Cedex, France

Email: cyril.grunspan@devinci.fr

RICARDO PÉREZ-MARCO

CNRS, IMJ-PRG
Labex Refi , Labex MME-DDII

Paris, France

Email: ricardo.perez.marco@gmail.com

September 24, 2017

ABSTRACT. The rules of the Bitcoin protocol are enforced by economic self-interest. We prove that the protocol is unstable when a miner has over 43% of the total hashrate. In this situation, the economic self-interest of the miner is not always to accept the longest mined chain. Under adequate circumstances, it can be profitable to mine a shorter chain hoping to catch-up the main chain (catch-up mining strategy).

Keywords: Bitcoin, blockchain, mining, proof-of-work, gambling problem, Dyck paths

1. INTRODUCTION

Decentralization of the Bitcoin protocol [N] requires that the rules of the protocol cannot be enforced by a regulatory body. Instead, they are enforced by the economic self-interest of the participants in the network: Any deviation from the rules of the protocol must be economically penalized.

One of the principal rules is that miners must adopt newly mined blocks and mine only on top of the longest chain¹. At first look, and this seems to be a widespread belief, a miner with less than 50% of hashrate², seems to not have any interest in mining on top of a shorter chain hoping to surpass the chain accepted by the majority of the network. The reason for this belief is that he will be wasting his hashrate chasing the network longest chain. But this is false.

Theorem 1. *For a miner with more than 43% of the total hashrate it can be profitable to mine a chain that is one or more blocks shorter.*

We assume that the rogue miner has a hashrate less than 50%, otherwise he can mine any shorter chain and take over any longer chain in the long run. The reasons behind the theorem is that although the probability of mining 2 or more blocks before the rest of the network is small, the reward for taking over the main chain is also larger since the miner reaps up the reward (and fees) of the invalidated blocks. Thus this strategy has a positive Expectation Value (EV) when there are enough blocks to invalidate, or the sum of blocks fees and rewards surpass a certain threshold.

Note that this rogue behaviour, that we call “catch up mining”, is different from selfish mining when the miner keeps a mined block to himself and starts mining on top of it.

We study the general problem of when it makes economic sense to try to catch up m blocks from the main chain. We study $E_n^m(v)$ which is the EV of the optimal strategy in order to catch up a delay of m blocks in n validation rounds and v is the pay-off.

The function $v \mapsto E_n^m(v)$ is a convex increasing affine by pieces

$$E_n^m(v) = \sum_l \pi(l) (v - v(l))_+ \quad (1)$$

where the sum runs over all winning paths shorter than n , $\pi(l)$ is the probability of the path l to occur, and $v(l)$ is the « minimal reward » over the path l .

Let $v_n^m = \inf_l v_n^m$. The sequence v_n^m is non-decreasing on m and decreasing on n . The strategy is profitable if and only if $v > v_n^m$.

1. “longest chain” meaning the chain with most work.

2. With more than 50% of hashrate the protocol is well known to be unstable



Bitcoin Logic: The mathematical path to an universal financial logic

YIANNIS KIOUVREKIS

Athens, Galatsi, Kiknon 67, 11146

☎ (+30) 6944347515 | ✉ yiannis.kiouvrekis@gmail.com | 📱 [john.kioubrekis](#)

Scaling Bitcoin

September 25, 2017

STANFORD

Abstract

Many experts in the field of Digital Currency and Computer Science in general have highlighted the need for convergence with the field of Formal Methods.

The article of Herihy and Moir, *Blockchains and the Logic of Accountability: Keynote Address*, highlights the necessity of creating a logical system that can express propositions such as “Because A (an agent) endorsed false statement, A can no longer be trust with nuclear codes” and properties like authorization, fairness, incentives as well as behaviors of miners.

In his presentation *How Formal Analysis and Verification Add Security to Blockchain-based Systems* during the last Blockchain Protocol Analysis and Security Engineering Conference 2017 Shinichiro Matsuo (MIT), reports the need for a logical system that is sound and complete, and within which we can describe the notion of security, privacy and thus prove the desirable security specifications.

A first step was made by Brunnler et al with blockchain epistemic logic, but as described by the writers, it is at a rather early stage,. In addition Joseph Y. Halpern and Rafael Pass in *A Knowledge-Based Analysis of the Blockchain Protocol* provide a complete characterization of agent’s knowledge. With this work our goal is:

First, to examine whether it is possible to create a logic that can describe properties such as privacy, security or Common Prefix Property, Chain Quality Property e.t.c. Can there be only one logical system, a universal logic? What knowledge do we need from model theory and mathematical logic in order to be sure that we can express properties in the logical system.? Furthermore, we should question the implications of other properties of a logical system like Craig’s interpolation property and compactness property.

Moreover, what knowledge can we get from the community of formal methods and the work done on the expressiveness of properties such as privacy in logical systems? The second thing we intend to explore is the relationship between Bitcoin, Game Theory and Mathematical Logic. Can we translate the fundamentals of Bitcoin Game in Formal Languages? It is known that the tools of modal logic have enriched the game-theoretic language by making it possible to express concepts that were previously either informally claimed to be captured by a solution concept. A famous theorem of this theory is the notion of com-

Proposal: Weak-Signal Radio Communications for Bitcoin Network Resilience

Nick Szabo, Elaine Ou
{nick, elaine}@globalfinancialaccess.com

Censorship-resistant bandwidth is crucial to maintaining a functional Bitcoin network. Prior work by Apostolaki, et al. has demonstrated Bitcoin’s vulnerability to routing attacks, where a BGP hijack or traffic interception can isolate parts of the network [1]. More recently, the Chinese Network Bureau has threatened to block Bitcoin network access by employing deep packet inspection through internet service providers [2].

The Blockstream Satellite was recently deployed to broadcast the Bitcoin blockchain [3], enabling users to receive blocks without incurring the bandwidth cost of running a full node. However, the Blockstream Satellite represents a single fallible source of information, and currently serves as only a one-way relay.

In this work, we explore the feasibility of using high-frequency (HF) radio transmissions to complement satellite receiver nodes for full-node network access. HF radio broadcasts have a range of hundreds of kilometers. For SPV clients, HF radio transmissions can provide sufficient bandwidth to broadcast merkle blocks and signed transactions, giving users the ability to participate in the network without any internet connection at all. Furthermore, the radio transmission of block headers by full nodes can provide early detection of network partitions, preventing the execution of potential exploits.

1 Bitcoin Routing Attacks

BGP routing vulnerabilities have been explored by Apostolaki, et al. [1]. Internet traffic associated with different IP prefixes is exchanged between neighboring networks, or Autonomous Systems (AS). Examples of ASes include broadband service providers such as Comcast or China Telecom, or cloud providers like AWS. In China, there are basically only two state-owned providers, China Telecom and China Unicom. Parts of the Bitcoin network are heavily centralized from a routing perspective.

Routing centralization can be exploited through the isolation of some subset of the network, leading to the creation of two different versions of the blockchain. Individual nodes are also vulnerable to an eclipse attack, which could delay the overall propagation of blocks towards the victim node [5]. For example, a

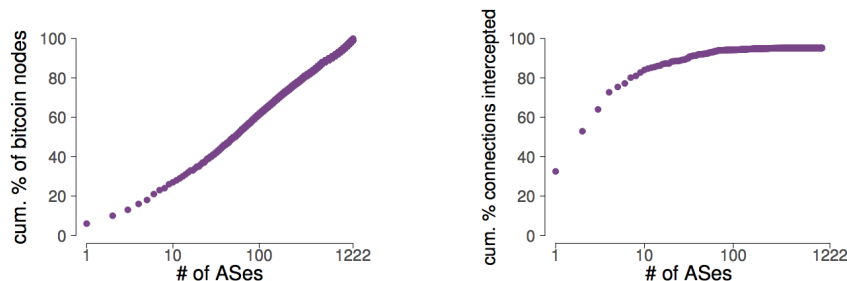


Figure 1: Left: Only 50 ASes host 50% of the Bitcoin network. Right: 3 ASes intercept 60% of all possible Bitcoin connections [1].

Bolt: Anonymous Payment Channels for Decentralized Currencies

Matthew Green
Johns Hopkins University
Baltimore, Maryland
mgreen@cs.jhu.edu

Ian Miers
Johns Hopkins University
Baltimore, Maryland
imiers@cs.jhu.edu

ABSTRACT

Bitcoin owes its success to the fact that transactions are transparently recorded in the blockchain, a global public ledger that removes the need for trusted parties. Unfortunately, recording every transaction in the blockchain causes privacy, latency, and scalability issues. Building on recent proposals for “micropayment channels” — two party associations that use the ledger only for dispute resolution — we introduce techniques for constructing *anonymous* payment channels. Our proposals allow for secure, instantaneous and private payments that substantially reduce the storage burden on the payment network. Specifically, we introduce three channel proposals, including a technique that allows payments via untrusted intermediaries. We build a concrete implementation of our scheme and show that it can be deployed via a soft fork to existing anonymous currencies such as ZCash.

ACM Reference format:

Matthew Green and Ian Miers. 2017. Bolt: Anonymous Payment Channels for Decentralized Currencies. In *Proceedings of CCS’17, Oct. 30–Nov. 3, 2017, Dallas, TX, USA.*, 17 pages.
<https://doi.org/10.1145/3133956.3134093>

1 INTRODUCTION

Bitcoin has become increasingly popular as a decentralized electronic currency. In Bitcoin, each transaction is recorded in the *blockchain*, a public transaction ledger maintained by a set of decentralized peers. While this design has proven successful at low transaction volumes, the reliance on a globally-shared ledger has caused serious scaling issues. Since in Bitcoin 1MB blocks are added to the blockchain every ten minutes on average, the Bitcoin transaction rate is limited to fewer than ten new transactions per second across the entire Bitcoin user base [1].¹ Several proposals to increase blockchain bandwidth are being debated in the Bitcoin community today, but none are likely to produce a transaction rate that competes with centralized services such as payment card networks.

A promising approach to addressing the scaling problem is to move the bulk of Bitcoin transactions *off chain*, while preserving the system’s decentralized structure and strong integrity guarantees. The leading proposal for off-chain payments is to use *payment channels*, exemplified by the Lightning Network [45] and Duplex Micropayment Channels [30]. Rather than posting individual payment transactions to the blockchain, channels employ the blockchain to first establish a shared deposit between two parties. The parties interact directly to make payments — adjusting the respective ownership shares of the deposit — and communicate with the blockchain only to agree on the final split of escrowed funds. In cases where no

direct payment channel exists between two parties, these proposals also allow participants to route transactions via intermediate peers [45]. The main benefit of the payment channel paradigm is that it dramatically reduces the transaction volume arriving at the blockchain, without adding new trusted and centralized parties.

While payment channels offer a solution to the scaling problem, they inherit many of the well-known privacy weaknesses of Bitcoin [40, 46]. Although payments are conducted off chain, any party may learn the pseudonymous identities and initial (resp. final) channel balances of the participants. More critically, payment channels provide few privacy protections against transaction counterparties. By establishing a channel to pay for *e.g.*, Tor bandwidth or web content, a user implicitly links each payment on a given channel to all of her other payments on this channel. This is particularly problematic in the likely event that payments are routed via a common intermediate peer — such as a currency exchange — since the intermediary must now be trusted to keep private your full payment history. Some proposals, such as the Lightning Network, have proposed to work around this problem by routing the payment via *multiple* intermediary nodes; however (as we discuss in §6) this approach substantially increases the complexity of establishing payment channels, and reveals payment information in the event that even a subset of the intermediaries collude.

Several academic works have recently proposed solutions that address the privacy problems of Bitcoin-type currencies [29, 41, 42, 47]. Some of the resulting systems been publicly deployed, notably ZCash [3] (an implementation of the Zerocash protocol [47]) and Monero [2]. Unfortunately, the privacy mechanisms contained in these systems apply to the privacy of transactions *on the blockchain*, and do not address the setting of payment channels. Indeed privacy for payment channels seems fundamentally challenging due to channels’ pairwise structure. Even when a channel is funded with anonymous currency, repeated payments within the same channel are inherently linkable. This is concerning, given that one of the main proposed applications of channels is for *web micropayments* — which are often described as a more private alternative to tracking and online behavioral advertising.

We stress that concerns about privacy are not theoretical. Several commercial ventures [11, 23, 32] have been founded around the task of analyzing and tracing blockchain transactions. It is reasonable to expect that surveillance will be applied to payment channel systems if they become widely deployed.

Our Contribution. In this paper we propose Blind Off-chain Lightweight Transactions, or *Bolt*. Bolt consists of a set of techniques for constructing *privacy-preserving* payment channels for a decentralized currency. These techniques ensure that multiple payments on

¹As of early May 2017, this has resulted in a backlog of nearly 165,000 transactions [15].

Hybrid architecture model to increase bootstrapping capability on Bitcoin

Richard Dennis

School of Computing
University of Portsmouth
Portsmouth, United Kingdom
Richard.dennis@port.ac.uk

Gareth Owenson

School of Computing
University of Portsmouth
Portsmouth, United Kingdom
Gareth.owenson@port.ac.uk

Abstract— How to scale Bitcoin is still an open research question, while most of the research currently focuses on increasing the number of transaction Bitcoin can process, this paper takes a different view, and looks at the bootstrapping method. We demonstrate an effective and proven attack on the current DNS protocol which enables a low resourced attacker to partition new nodes joining the network. We then conduct analysis on how well the current DNS model can scale, before suggesting a hybrid P2P architecture model comparing this model with the current protocols, in terms of resistance to the DNS attack and scalability.

Blockchain, scalability, cryptographic protocols, distributed networks, peer-to-peer, Bittorent

I. INTRODUCTION

Bitcoin is the first worldwide, mass adopted cryptocurrency and digital payment system to be implemented and deployed without a requirement of a centralized repository system or administrator. It was invented by an unknown programmer, or a group of programmers, under the name Satoshi Nakamoto, published in a white paper in 2008, before being released as open-source software in 2009.

The key invention made by Nakamoto was the blockchain is a novel peer-to-peer approach which links a sequence of transactions or events together in a way that makes them immutable.

The blockchain is a public ledger of all transactions that have ever been completed since the first “genesis” block. Each transaction from the Bitcoin protocol is broadcast to all nodes in the network which are maintaining the blockchain.

A blockchain-node and a miner are two types of nodes on the network, which while can be conducted on the same node, is usually separated. A blockchain-node can be classed as node which maintains and updates the blockchain, with valid blocks received from miners on the network.

Each blockchain-node confirms if each transaction is valid and can be added to a block. Each blockchain-node confirms every transaction made on the network, to do so, each blockchain-node will search through the blockchain they store and maintain to see if the user requesting the transaction has got enough funds to process the transaction, and this transaction has not previously been conducted. Only once this process has happened will each node compile a block (a group of

transactions) and send this to the miners. There are not incentives to run a blockchain-node.

A miner participates in the process by which transactions are verified and added to the public ledger, known as the block chain, and also the means through which new bitcoin are released. The mining process involves compiling recent transactions into blocks and trying to solve a computationally difficult puzzle. The participant who first solves the puzzle gets to place the next block on the block chain and claim the rewards. The rewards, which incentivize mining, are both the transaction fees associated with the transactions compiled in the block as well as newly released bitcoin

Bitcoin is the most successful blockchain-based network; it has a market cap of over USD 8.5 billion and sees an average of 214,000 transactions being conducted on its network every day.

Blockchain-based networks have not properly addressed the issue of scalability; this causes the original decentralized nature of the blockchain to become increasingly centralized, as only the highest-resourced users are able participate in the network. This is because each node on the network is required to store the entire blockchain, which stores every transaction since its deployment and consequently low-resourced users; such as mobile users – are excluded from the network.

There has been several other peer-to-peer (P2P) and decentralized networks such as BitTorrent which have face similar scalability issues, overcame some of these issues with the use of a Hybrid architecture model, combining both the client-server model and P2P architecture.

This paper proposes a new approach in the way lower resourced nodes can be included in the network. It will examine how trusted “super nodes” can be utilized enabling a global view of the network, while providing monitoring facilities of all nodes on the network. We will conduct a thorough analysis on how the introduction of super nodes can aid in the scalability of the bootstrap process, by first demonstrating a unique attack against the currently implementation of the bootstrap protocol, and then analysing how the introduction of super nodes not only prevents such an attack from occurring, but also allows for a greater number of simultaneous nodes to bootstrap at the same time.

How to Charge Lightning*

The Economics of Bitcoin Transaction Channels

Simina Brânzei[†]

Erel Segal-Halevi[‡]

Aviv Zohar[§]

September 25, 2017

Abstract

Off-chain transaction channels represent one of the leading techniques to scale the transaction throughput in cryptocurrencies. However, the economic effect of transaction channels on the system has not been fully explored up until now. We present a framework for economic analysis of the lightning network and its effect on transaction fees on the blockchain. Our framework allows us to reason about different patterns of demand for transactions, different topologies of the lightning network and to derive the resulting fees for transacting both on and off the blockchain. Our initial results (that should be considered carefully) indicate that while the lightning network does allow for a substantially greater number of transactions to pass through the system, it does not necessarily provide higher fees to miners, and as a result may in fact lead to lower participation in mining within the system.

1 Introduction

A main approach to solve the scalability problem in Bitcoin is to use off-chain transaction channels that allow parties to communicate and transfer funds while communicating directly, and only occasionally to settle on the blockchain. The recent deployment of SegWit, a solution to transaction malleability (among other benefits) opens the path for better constructions of off-chain transaction channels. While transaction channels themselves are limited to exchanges between pairs of individuals, further developments like the lightning network allow to route payments over longer paths and thus can allow the construction of a well connected network of payment channels that can be used to transfer money quickly and with relatively little interaction with the blockchain.

One of the key unknowns regarding fast payment networks is the economic effect that they will have on the Bitcoin fee market. If the blockchain is used less often, fees to miners are paid less frequently and competition for space in blocks declines. Bitcoin's security depends heavily on having a large amount of computational power invested in solving proof-of-work puzzles, making it hard for attackers to double spend or censor transactions in the currency. As the

*The authors are in alphabetical order. This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 740282), from the Israel Science Foundation (grant 616/13 and grant 1083/13) and from the HUJI Cyber Security Research Center in conjunction with the Israel National Cyber Bureau (grant 039-9230). Simina was also supported by the ISF grant 1435/14 administered by the Israeli Academy of Sciences and Israel-USA Bi-national Science Foundation (BSF) grant 2014389 and the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation. Erel was supported by the ISF grant 1083/13.

[†]The Hebrew University of Jerusalem, Israel. E-mail: simina.branzei@gmail.com

[‡]Ariel University, Israel. E-mail: erelsgl@gmail.com

[§]The Hebrew University of Jerusalem, Israel. E-mail: avivz@cs.huji.ac.il

Redesigning Bitcoin’s fee market

Ron Lavi¹, Or Sattath^{2,3}, and Aviv Zohar²

¹Technion

²The Hebrew University

³MIT

September 25, 2017

Abstract

The security of the Bitcoin system is based on having a large amount of computational power in the hands of honest miners. Such miners are incentivized to join the system and validate transactions by the payments issued by the protocol to anyone who creates blocks. As new bitcoins creation rate decreases (halving approximately every 4 years), the revenue derived from transaction fees start to have an increasingly important role. We argue that Bitcoin’s current fee market does not extract revenue well when blocks are not congested. This effect has implications for the scalability debate: revenue from transaction fees may decrease if block size is increased.

The current mechanism is a “pay your bid” auction in which included transactions pay the amount they suggested. We propose two alternative auction mechanisms: The Monopolistic Price Mechanism, and the Random Sampling Optimal Price (RSOP) Mechanism (due to Goldberger *et al.*). In the monopolistic price mechanism, the miner chooses the number of accepted transactions in the block, and all transactions pay exactly the smallest bid included in the block. The mechanism thus sets the block size dynamically (up to a bound required for fast block propagation and other security concerns). We show, using analysis and simulations, that this mechanism extracts revenue better from users, and that it is nearly incentive compatible: the profit due to strategic bidding relative to honest bidding decreases as the number of bidders grows. Users can then simply set their bids truthfully to exactly the amount they are willing to pay to transact, and do not need to utilize fee estimate mechanisms, do not resort to bid shading and do not need to adjust transaction fees (via replace-by-fee mechanisms) if the mempool grows.

We discuss these and other properties of our mechanisms, and explore various desired properties of fee market mechanisms for crypto-currencies.

Using the Chain for what Chains are Good For

This proposal is for a high-level introductory talk. The list below can be compressed into a paragraph if that fits the abstract format better.

The blockchain is simultaneously Bitcoin's core innovation, letting it succeed where no other system had before, and its greatest weakness, requiring miner discretion in choosing transactions, while no other part of the system has any third-party dependence. In exchange for this dependence, miners produce an increasingly-immutable proof-of-publication medium, allowing anyone at any time to see what transactions occurred and in what order, and to be assured that their view matches the view of all other validators.

Bitcoin's essential use of the blockchain is to prevent *double-spending*: to publish transaction outputs as they are created and later consumed as inputs. This provides an unambiguous beginning and end of each output, which is an otherwise unattainable goal in a relativistic world. However Bitcoin uses the blockchain for much more than this: it has a script system which allows users to set arbitrary spend conditions on their coins; it allows transactions to be time-locked and invalid until some time has passed; it ensures transactions are executed atomically and not peeled apart on the network. All of these conditions are published on the chain and verified by all validators. It is the thesis of this talk that these "non-essential" uses of the blockchain can often be done with significantly reduced (or eliminated) use of the blockchain, and that this has tremendous benefits for the transactors themselves as well as the network as a whole.

First, we describe the costs of blockchain usage.

- Blocks appear only every ten minutes on average, meaning long and unpredictable latency for users of the chain.
- During this time transactions are published to the network, leaking private timing and source information, plus the transaction data itself, to anyone who cares to analyze it. This exposure undermines users' privacy, businesses' confidentiality, and the fungibility of the currency itself.
- This public data can be seen by miners before they include transactions, which poses a censorship risk for users as well as an incentive for adversaries to pressure miners into censorship.
- Blockchain space is limited, forcing users to pay for this data even as its existence harms them.
- This data must be validated by all participants in the system who want to verify that their view of its state is untampered with. Since supporting these users is a core component of the Bitcoin ethos, the result is a limitation on the scalability of the entire system.
- Finally, the rules for validating data on the blockchain are system-wide rules that cannot be changed without agreement from all users. This is not even possible for conflicting rules, but when it is possible it is extremely hard to measure and hard to achieve.

1. Introduction

Since its introduction Bitcoin received much attention as a peer-to-peer cryptocurrency based on the blockchain technology (Bonneau et al 2015, Narayanan et al 2016). Adoption of Bitcoin may exhibit advantages as well as critical aspects (Böhme et al, 2015; Athey et al, 2016). From an economic perspective its use may facilitate exchange and possibly save on transaction costs. Because of its exchangeability with fiat currencies such as the dollar, advantages could also come from a speculative activity based on oscillations of the exchange rate.

However, one of its most distinguishing features is that registration of transactions is done through the so called *mining* activity undertaken by some subjects. Such activity consists in solving a puzzle requiring high computational power, since registration of a block of transactions can only take place once the puzzle has been solved. Providing the right economic incentives to solve the puzzle is very important for the transactions to be registered on the underlying ledger. This is why for this activity *miners* are compensated with two types of rewards: first, for any solved puzzle the miner will receive a fixed sum of bitcoins by the protocol and, moreover, individuals behind a transaction may offer a fee to the miner for its registration. The larger such fee the higher the incentive for the miner to enclose a transaction in the next registered block. The fixed sum received by the protocol for each block of registrations will tend to decline over the years until its disappearance, after which only fees paid for transactions registration will reward the miners.

In this paper we focus on the mining activity as a source of economic profitability (Narayan et al, 2015), where the main strategic decision taken by miners is how much to invest in computational power to solve the puzzle. Since mining costs are increasing, the chosen level of power is becoming a critical issue for the Bitcoin community, which is what motivates our analysis. Within a very simple static game theoretic framework, our model provides some interesting insights. Due to the assumption of exponential waiting time for the solution to a puzzle, the mining activity can be characterised as an all-pay *contest* (Konrad, 2009, Vojnovic, 2015) a conceptual framework widely adopted in social sciences. All-pay contests are competitions where winners are awarded a prize specified in advance by the organizer. They require investments to participate, and this is what makes them all-pay, and victory by a participant typically occurs probabilistically. Therefore, those who obtain no prize lose their investments, unless these could be re-used in other contests. Winning probabilities are often called contest functions.

Indeed, mining activity can be seen as a contest where participants are trying to come first in the competition for the solution of the puzzle, receiving as prize a given amount of Bitcoins as well as some fees from the other participants.

At the Nash Equilibrium of the mining game with perfect information, while the level of computational power chosen by an active miner depends also on how many bitcoins could be obtained solving the puzzle, the decision to become an active miner depends only on his own marginal costs as compared to the opponents' cost structure. That is, the decision to be an active miner would only depend upon how efficient are the competitors but not on how many bitcoins will be obtained as reward.

Topic (tentative): Bitcoin script 2.0 and strengthened payment channels

Johnson Lau, Olaoluwa Osuntokun

Bitcoin uses a scripting system to imbue users with the power to designate the conditions under which UTXO's they create can be spent. However, the capabilities of Script to date are limited in many ways. Amongst these limitations include:

- Several useful opcodes were disabled in a series of emergency softforks in 2010
- Users are unable to commit to additional scripts or conditions after a UTXO is created. Scripts in the scriptSig (prior to segwit) were malleable by third parties and cannot be utilized in any meaningful way
- Script operations have very restricted access to the rest of the transaction. Only the 6 SIGHASH types allow limited introspection into the execution environment (inputs, scripts, transaction spending, etc)
- Numerical operations accept up to 32-bit signed integers, while 51-bit is needed to cover the full range of bitcoin supply ($2.1E15$)

The activation of segwit has paved the way for the evolution of contracts on top of Bitcoin due to its long overdue malleability fix and added Script versioning capabilities. The malleability fix allows contract designers to safely nest pre-signed contract execution pathways typically heavily utilized in the Bitcoin model of contract design. Additionally, the Script versioning features allows for rapid evolution of Script, as it's possible to entirely re-design portions of Script with a single version bump. Within the development community several new features have been proposed for the next generation of Bitcoin Script including:

- OP_PUSHTXDATA (Covenant)
- Signature-time commitment to additional scripts
- OP_CHECKSIGFROMSTACK
- Re-enabling all the disabled operations
- More SIGHASH types
- Addition of op-codes for primitive EC operations (group op, scalar mult)

The set of proposed additions to Script outlined above have the potential to significantly broaden the expressivity, and power of Bitcoin's Script in the domain of smart contract design. In this talk, we'll begin by briefly overview the past history of the evolution of Script. With the necessary historical context explored, will then detail the new anticipated proposed additions to Script which include the augmented power of Script introspection. Finally as a case study, we'll utilize the new Bitcoin contract functionality to design a new version of payment channels for Lightning which improve upon the scalability, safety, and privacy of prior iteration of channel design.

References:

BIP draft: Merklized Script

<https://github.com/jl2012/bips/blob/vault/bip-0114.mediawiki>

BIP draft: Pay-to-witness-public-key

<https://github.com/jl2012/bips/blob/vault/bip-0VVV.mediawiki>



David Vorick

[Follow](#)

Draft · 14 min read

Microchains—massive blockchain scalability, among other significant advantages

Proposal Note: I would boil this down to 30 minutes by selecting a few of the biggest ideas, most importantly the scorched-earth defense idea, and then presenting on those.

I am about to propose a system which I claim solves many of blockchain's problems at the same time. We are able to achieve this by turning some of our core assumptions about blockchain upside down, and by leveraging new advances in game theory and crypto-economics, without which we would not be able to achieve any remote semblance of security.

There are four major problems in blockchain today which I believe a microchains ecosystem can solve:

- **Scalability.** This is the biggest one, and the real reason that everyone should take a close look at microchains. I claim that we can get somewhere between 1,000x and 10,000,000x scalability over traditional blockchains, primarily because we end up with a secure system where people no longer need to validate every transaction to achieve full security and full trustlessness—and can do so without finality, proof of stake, or any of the other flaky mechanisms used by other scalable system
- **Upgradeability.** Today, upgrading a blockchain is a headache. Either you have to do a fancy soft-fork which gains widespread adoption, or you have to perform a hardfork, which either needs absolute adoption or creates a network split. In the microchains system, you can easily perform incremental hardfork upgrades, where only a small fraction of the network needs to upgrade at a time.
- **Blockchain Factions.** Bitcoin today requires everyone to run the exact same consensus code. Multiple implementations are highly frowned upon due to the risk of tiny hardfork bugs that cause

The Future of Proof of Work

Min Chen, CTO, Canaan Creative China (Avalon)

30 Minute Presentation

In this presentation, Min Chen, CTO for Canaan Creative China - the inventors of the AvalonMiner, the first Bitcoin mining ASIC microprocessor - describes where Proof of Work is heading, along with problems approaching this solution, and how to best solve them. Particular emphasis will be placed on the evolution of ASIC microprocessor development from the perspective of the Canaan Avalon series of chips towards where Bitcoin mining and development is going, and how to do this in silicon.

In her first presentation on the subject publicly, Min Chen will cover some unwritten history of the evolution she led from FPGA based mining Icarus and Lancelot, to the first Avalon Bitcoin mining ASIC microprocessors up to the current scaling issues facing Bitcoin, power requirements, and paths for future currency development using hardware.

Measuring maximum sustained transaction throughput on a global network of Bitcoin nodes

Andrea Suisani,¹ Andrew Clifford,¹ Andrew Stone,¹ Erik Beijnoff,¹ Peter Rizun,¹ Peter Tschipper,¹ Alexandra Fedorova,² Chen Feng,² Victoria Lemieux,² Stefan Matthews³

¹ Bitcoin Unlimited, ² University of British Columbia, ³ nChain

Overview

Although it is well understood that increasing Bitcoin’s block size limit (currently 1 MB) would immediately reduce transaction fees and improve confirmation reliability, concern exists regarding the network’s ability to safely and reliably handle the associated increase in transaction throughput.

To investigate this concern, we set up a global network of Bitcoin mining nodes¹ configured to accept blocks up to one thousand times larger (1 GB) than the current limit. To those nodes we connected transaction generators, each capable of generating and broadcasting 200 transactions per second (tx/sec) sustained.² We performed (and are continuing to perform) a series of “ramps,” where the transaction generators were programmed to increase their generation rate following an exponential curve starting at 1 tx/sec and concluding at 1000 tx/sec—as illustrated in Fig. 1—to identify bottlenecks and measure performance statistics.

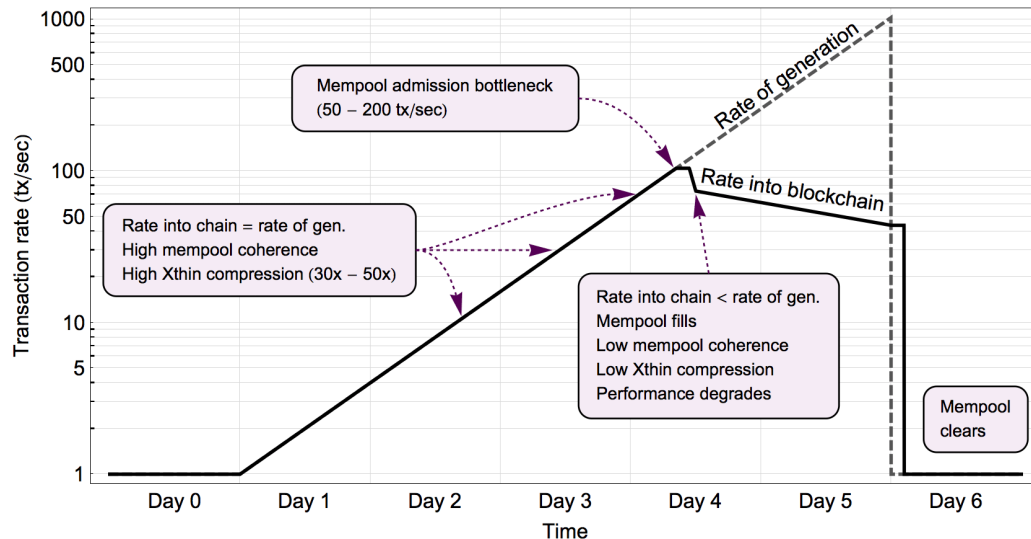


Fig. 1. Ramp input and typical node response.

¹ At the time of writing, there were mining nodes in Toronto (64 GB, 20 core VPS), Frankfurt (16 GB, 8 core VPS), Munich (64 GB, 10-core rack-mounted server with 1 TB SSD), Stockholm (64 GB, 4 core desktop with 500 GB SSD), and central Washington State (16 GB, 4 core desktop). With the passing of BUIP065 and the associated \$300,000 per year funding for the Gigablock Testnet Initiative, additional mining nodes will be deployed in Beijing, Bangalore, Sao Paulo, Sydney and Vancouver. The results we present at Stanford will include data from this larger test network as well.

² At the time of writing, there were generators in San Francisco, New York, London, Amsterdam, Singapore and Bangalore (all 8 GB, 4 core VPS). Generators are Python applications interacting with a local instance of `bitcoind`.

Graphene: A New Protocol for Block Propagation Using Set Reconciliation

A. Pinar Ozisik[†] Gavin Andresen George Bissias[†] Amir Houmansadr[†] Brian N. Levine[†]

[†]College of Information and Computer Sciences, UMass Amherst

1 INTRODUCTION

We propose an efficient method of announcing new blocks called *Graphene*. This document summarizes a more detailed description of Graphene that has been published previously [7].

Graphene blocks are a fraction of the size of related methods, such as Compact Blocks [3] and Xtreme Thinblocks [9]. For example, while a 17.5 KB Xtreme Thinblock can be encoded in 10 KB with Compact Blocks, the same information can be encoded in 2.6 KB with Graphene. We use a novel interactive combination of Bloom filters [2] and Invertible bloom lookup tables (IBLTs) [5], providing an efficient solution to the problem of set reconciliation in Bitcoin’s p2p network.

Block announcements are validated using the transaction content comprising the block. However, it is likely that the majority of peers have already received these transactions, and they only need to discern them from those in their mempool. In principle, a block announcement needs to include only the IDs of those transactions. For example, Corallo’s *Compact Block* design [3] significantly reduces block size by including a transaction ID list, though the cost is increased coordination between peers to 5 messages. *Xtreme Thinblocks* [9] works similarly to Compact Blocks but has greater data overhead. Specifically, if an *inv* is sent for a block that is not in the receiver’s mempool, the receiver sends a Bloom filter of her IDpool along with the request for the missing block. As a result, Xtreme Thinblocks are larger than Compact Blocks but require just 3 messages. The community has discussed in forums the use of IBLTs (without Bloom Filters) for reducing block announcements [1, 8], but these schemes have not been formally evaluated and are less efficient than our approach. Our method is novel; we have proved and demonstrated that it is smaller than all of these recent works, and still requires 3 messages between sender and receiver for coordination.

2 THE GRAPHENE PROTOCOL

Unlike other approaches, Graphene never sends an explicit list of transaction IDs. Instead it sends a small Bloom filter

PROTOCOL 1: Graphene	
1: Sender:	Sends <i>inv</i> for a block.
2: Receiver:	Requests unknown block; includes count of txns in her IDpool, m .
3: Sender:	Sends Bloom filter \mathcal{S} and IBLT \mathcal{I} (each created from the set of n txn IDs in the block) and essential Bitcoin header fields. The FPR of the filter is $f = \frac{a}{m-n}$, where $a = n/(c\tau)$.
4: Receiver:	Creates IBLT \mathcal{I}' from the txn IDs that pass through \mathcal{S} . She decodes the <i>subtraction</i> [4] of the two blocks, $\mathcal{I} \Delta \mathcal{I}'$.

Figure 1: A summary of the Graphene protocol.

and a very small IBLT. The intuition behind Graphene is as follows; a summary appears in Figure 1.

The sender creates an IBLT \mathcal{I} from the set of transaction (txn) IDs in the block. To help the receiver create the same IBLT (or similar), he also creates a Bloom filter \mathcal{S} of the transaction IDs in the block. The receiver uses \mathcal{S} to filter out transaction IDs from her pool of received transaction IDs (which we call the IDpool) and creates her own IBLT \mathcal{I}' . She then attempts to use \mathcal{I}' to *decode* \mathcal{I} , which, if successful, will yield the transaction IDs comprising the block. The number of transactions that falsely appear to be in \mathcal{S} , and therefore are wrongly added to \mathcal{I}' , is determined by a parameter controlled by the sender. Using this parameter, he can create \mathcal{I} such that it will decode with very high probability.

In sum, the Bloom filter from the sender allows the receiver to determine which transactions from its mempool are in the block. Other approaches require a much larger Bloom filter to keep the false positive rate small; in Graphene, the Bloom Filter FPR is high because the IBLT recovers any mistakes made. Similarly, if only the IBLT was used, it would be much larger than our use of the two mechanisms.

A Bloom filter is an array of x bits representing y items. Initially, the x bits are cleared. Whenever an item is added to the filter, k bits, selected using k hash functions, in the bit-array are set. The number of bits required by the filter is $x = y \frac{-\ln(f)}{\ln^2(2)}$, where f is the intended false positive rate (FPR). For Graphene, we set $f = \frac{a}{m-n}$, where a is the expected difference between \mathcal{I} and \mathcal{I}' . Since the Bloom filter

Bobtail: A Proof-of-Work Target that Reduces Blockchain Mining Variance

George Bissias Brian N. Levine
College of Information and Computer Sciences
University of Massachusetts Amherst

Introduction. Blockchain systems are designed to produce blocks at a constant average rate. Bitcoin [1] produces, on average, one block every 10 minutes. Unfortunately, the time between blocks has high variance and the distribution of inter-block times has a very long tail. For example, 5% of the time, Bitcoin’s inter-block time is at least 40 minutes. This variance impedes the consistent flow of validated transactions through the system.

The high inter-block time variance is a direct consequence of Bitcoin’s Proof-of-Work (PoW) algorithm. Generally, the miners repeatedly craft block headers by changing a nonce until the hash of the header is less than a target value t . In other words, the hash of each header is a sample taken randomly from a discrete uniform distribution. A block is discovered when the *first order statistic* (i.e., the minimum value) of all sampled values is less than target t .

Proposal. We propose an alternative process for PoW-based block discovery that results in an inter-block time with significantly lower variance. Our algorithm generalizes the current algorithm by comparing the mean of the k lowest order statistics to a target. As a result, the variance of inter-block times decreases as k increases. For example, if our approach were applied to Bitcoin, about 80% of blocks would be found within 7 to 12 minutes, and nearly every block would be found within 5 to 18 minutes; the average inter-block time would remain at 10 minutes. The cost of our approach is a larger block header. We call our approach *Bobtail*¹ mining. Figure 1 shows results from experiments with this method.

This proposal summarizes a full paper available from <https://forensics.umass.edu/bobtail>.

Bobtail Mining. Consider a fixed interval of time during which the entire network produces θ hashes generating a sequence of hash values $\mathbf{Z} = Z_1, \dots, Z_\theta$. Let Z be an arbitrary random variable from the sequence \mathbf{Z} ; note that $Z \sim \text{Uniform}(0, S)$. Define V_i to be the i th lowest order statistic of \mathbf{Z} , i.e. $V_i = Z_{(i)}$ in standard notation. And let random variable W_k be the mean of the k lowest order statistics:

$$W_k = \frac{1}{k} \sum_{i=1}^k V_i. \quad (1)$$

W_k constitutes the collective mining proof (*proof*, for short) for the entire network. Our Bobtail mining criterion says that a new block is discovered when a realized value of W_k meets the target t :

$$w_k \leq t. \quad (2)$$

Notably, this approach is a generalization of current systems, which are the special case of $k = 1$.

¹A *bobtail* refers to an animal’s tail that is unusually short or is missing completely (https://en.wikipedia.org/wiki/Natural_bobtail).

BlockSci: a Platform for Blockchain Science and Exploration

Harry Kalodner, Malte Möser, Steven Goldfeder, Alishah Chator, and Arvind Narayanan

Abstract

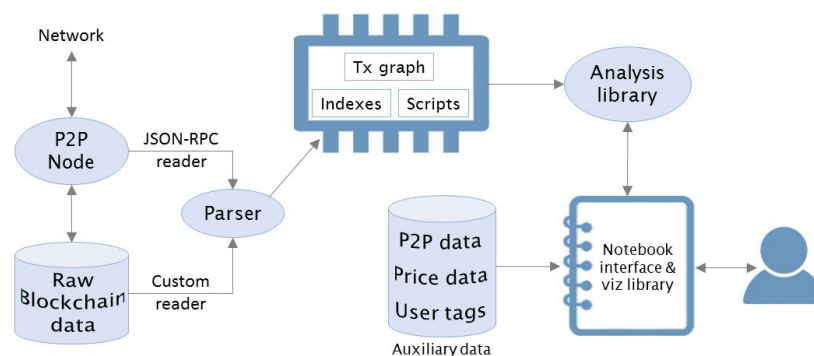
In order to understand the scaling demands of Bitcoin, we need to understand the usage demands on Bitcoin. Using BlockSci, a high-performance tool for blockchain science and exploration, we can answer questions about the nature of transactions that are being included in the blockchain. This knowledge, in turn, can help build better scaling solutions for Bitcoin.

Talk Proposal

The Bitcoin blockchain — currently 140GB and growing — contains a massive amount of data that can give us insights into the Bitcoin ecosystem, including how users, businesses, and miners operate. In this talk we present BlockSci, an open-source software tool that enables fast and expressive analyses of Bitcoin’s and many other blockchains. BlockSci has already been used in multiple academic papers, and we have released it as open-source software to encourage open and reproducible blockchain science.

An open source cryptocurrency analytics framework can provide valuable data to evaluate changes to Bitcoin. Often ideas and heuristics are discussed in the abstract, without concrete numbers attached. If numbers are cited, the details of the analysis are often private and unverifiable. Ideally, decisions impacting Bitcoin users should be made based on reliable data that was interpreted in a reproducible way.

BlockSci enables the science of blockchains. It addresses three pain points of existing tools: poor performance, limited capabilities, and a cumbersome programming interface. BlockSci is 15x–600x faster than existing tools, comes bundled with analytic modules such as address clustering, exposes different blockchains through a common interface, imports exchange rate data and “mempool” data, and gives the programmer a choice of interfaces: a Jupyter notebook for intuitive exploration and C++ for performance-critical tasks. We show an overview of BlockSci’s architecture below.



Overview of BlockSci’s architecture

BlockSci’s design starts with the observation that blockchains are append-only databases; further, the snapshots used for research are static. This makes an in-memory analytical database the natural choice. On top of the obvious speed gains of memory, we apply a number of tricks such as converting hash

1. Changes without unanimous consent

"By expecting a few developers to make controversial decisions you are breaking the expectations, as well as making life dangerous for those developers. I'll jump ship before being forced to merge an even remotely controversial hardfork."

-- <https://lists.linuxfoundation.org/pipermail/bitcoin-dev/2015-June/009137.html>

If proposed consensus changes have unanimous consent, then life is easy: developers are happy to accept the patches, miners are happy to deploy the updated software and can quickly indicate that they are enforcing the changes making the risk of chain splits arbitrarily low, validating nodes are happy to enforce the new rule set, the bitcoin economy is happy to continue respecting the resulting ledger, and regulators are happy to refrain from making anyone unhappy.

However there are many reasons why any of those people may not be happy for any given proposed change:

- Different people in the bitcoin ecosystem may have different goals; as a trivial example, regulators might propose a change due to a strong desire to prevent bitcoin from being usable for criminal purposes, while others may oppose that change due to its impact on anonymity, fungibility, or efficiency. as bitcoin grows, it is likely to attract more people with conflicting goals, increasing the frequency of this form of conflict arising.
- There may be a lack of understanding of the impact of a change; perhaps if everyone understood a change's actual effects, everyone would support (or oppose) it, but it may be that many people are unable to reach this conclusion with confidence, whether due to ignorance, or lack of time to fully study the issue are unable. If bitcoin adoption increases amongst people with less technical background or with different attitudes to monetary policy, this class of controversy is more likely to occur.
- Bitcoin upgrades are generally designed as Pareto-improvements, that is either everyone in the ecosystem is made better off, or at least are not made any worse off. It may be that some changes will not be able to be made in this manner, and some changes will cause a net benefit to the ecosystem, despite a net loss to some individuals.

Incentives and Trade-offs in Transaction Selection in DAG-based Protocols

Yoad Lewenberg and Yonatan Sompolsky

Introduction

In Bitcoin, and in the many cryptocurrencies that followed its path, the public ledger consists of a single chain of blocks. Every block embeds a pointer to a single parent block, and only blocks on a single chain -- the longest, usually -- are considered; blocks off this chain (aka orphans) are omitted from the ledger and discarded. In contrast, Lewenberg *et al.* [1] proposed a new setup where blocks may reference several parent blocks as their predecessors. The resulting data structure is a Directed Acyclic Graph of blocks (a block DAG). DAG-based protocols are increasingly gaining traction (see, e.g., the work on the SPECTRE protocol [2], which utilizes all blocks avoiding the notion of chain altogether; or Ethereum, which uses a DAG in order to reward off chain blocks and mitigate centralization thus), which justifies reexamining the ways in which these protocols can be deployed in reality and deliver on their promise.

One primary advantage of a DAG ledger is that it improves the scalability of the system in terms of transaction throughput, by allowing contribution of content from more blocks (not only those contained in a single chain). The high level idea is to incentivize miners to avoid embedding the same transaction in multiple blocks in the DAG, since that would waste potential throughput and at the same time harm the miners themselves; indeed, the fee from a transaction may only be distributed once, either to one block or divided somehow.

Concretely, assume that two blocks were created by honest miners at about the same time, and therefore they do not reference one another (directly or even indirectly). It is possible that some conflicting transactions appear in these blocks, in which case one of the conflicting transactions must be omitted, according to some rule. Either way, the remaining sets of transactions, which are compatible, can be co-accepted and considered part of the UTxO. Of course, if these blocks merely duplicate the same transaction-set then nothing will be gained by accepting both and we might as well discard one of them. However, if these blocks contain mutually unique transactions, integrating both blocks into the ledger and accepting their non-conflicting content improves the overall throughput.

The Inclusive paper [1] analyzes the dynamics that unfold from a game theoretic point of view. Overall, we observe that miners will indeed be incentivized to avoid collisions and include unique contents in their blocks. They can do so by randomizing over their local mempools and selecting transactions for their next blocks according to some probability distribution (the equilibrium distribution). The distribution takes into account the fee that each transaction entails, as well as the probability that the miner will indeed earn it in case of a collision.

A New Blockchain-as-a-Service Paradigm

Yonatan Sompolsky and Yoad Lewenberg

I. Introduction

Since the inception of Bitcoin, many blockchain based systems have been proposed and deployed. Typically, the underlying protocol specifies which transactions should be considered valid, and how to create and add new blocks of transactions to the ledger. Importantly, most if not all of current blockchain systems follow Satoshi's design in that the validity of blocks and that of transactions is coupled: a valid block must not contain invalid transactions.

We propose revisiting this design choice. Our objective is to increase the scalability of the system as well as to introduce a more soft and decentralized governance model for innovation on the application-layer. These two objectives are achieved by treating separately the different layers of blockchain, and in particular decoupling the validity of blocks from that of transactions. The result is a system where miners can contribute their mining resources to secure several application-layer protocols simultaneously. Our solution is closely connected to the concept of merged mining, yet is different in some core aspects.

II. Blockchain layers

In our design, blockchain consists of two primary layers, the mining layer and the application layer. The mining layer is where the formation of the ledger takes place: users publish data (or: transactions) with an associated fee, and miners collect these data and embed them in blocks. The rules of block creation are dictated and enforced by the mining protocol. However, this mining protocol is agnostic to the data itself---the validity of a block is determined solely by its header and structure, and does not depend on the transaction data embedded in it. The result of the mining layer is a public ledger of transactions, or simply *the ledger*.

Next, the application layer is where the ledger is interpreted by users, or *clients*. An interpretation protocol P specifies how to read the ledger, and specifically how to treat those transactions that were labelled as belonging to P ; every transaction in the ledger specifies in its own header the name/ID of the application-layer protocol according to which it ought be interpreted. For instance, transactions can be labeled by their creators as <bitcoin txn>, <zcash txn>, <ethereum smart contract>, <ethereum classic smart contract>, etc. The user can then decide which protocol(s) to run on his client, according to his own economic value and preference. For instance, the client can ignore all but transactions labeled as <bitcoin txn>, and interpret the rest of the ledger data as blank data. In this way, we allow multiple application-layer protocols to be co-hosted on the same ledger and mining platform.

We stress the following crucial design-rule: A block that was mined correctly (according to the mining protocol's instructions) remains valid regardless of its data. In particular, it may contain invalid <bitcoin txns> transactions, and while users running the Bitcoin client will ignore these

DOUBLE SPEND RACES

CYRIL GRUNSPAN AND RICARDO PÉREZ-MARCO

ABSTRACT. We correct the double spend race analysis given in Nakamoto’s foundational Bitcoin article and give a closed-form formula for the probability of success of a double spend attack using the Regularized Incomplete Beta Function. We give a proof of the exponential decay on the number of confirmations, often cited in the literature, and find an asymptotic formula. Larger number of confirmations are necessary compared to those given by Nakamoto. We also compute the probability conditional to the known validation time of the blocks. This provides a finer risk analysis than the classical one.

To the memory of our beloved teacher André Warusfel who taught us how to have fun with the applications of mathematics.

1. INTRODUCTION.

The main breakthrough in [7] is the solution to the *double spend problem*. Before this discovery no one knew how to avoid the double spending of an electronic currency unit without the supervision of a central authority. This made Bitcoin the first form of *peer-to-peer* (P2P) electronic currency.

A double spend attack can only be attempted with a substantial fraction of the hashrate used in the *Proof-of-Work* of the Bitcoin network. The attackers will start a *double spend race* against the rest of the network to replace the last blocks of the blockchain by secretly mining an alternate blockchain. The last section of [7] computes the probability that the attackers catch up. However Nakamoto’s analysis is not accurate since he makes the simplifying assumption that honest miners validate blocks at the expected rate. We present a correct analysis and give a closed-form formula for the exact probability.

Date: February 9th 2017.

2010 Mathematics Subject Classification. 68M01, 60G40, 91A60, 33B20.

Key words and phrases. Bitcoin, blockchain, double spend, mining, proof-of-work, Regularized Incomplete Beta Function.

Acknowledgements: We are grateful to N. Emerson for his comments and reading over the article.

Discreet Log Contracts

Thaddeus Dryja

MIT Digital Currency Initiative

Abstract

Smart contracts [1] are an often touted feature of cryptographic currency systems such as Bitcoin, but they have yet to see widespread financial use. Two of the biggest hurdles to their implementation and adoption have been scalability of the smart contracts, and the difficulty in getting data external to the currency system into the smart contract. Privacy of the contract has been another issue to date. Discreet Log Contracts are a system which addresses the scalability and privacy concerns and seeks to minimize the trust required in the oracle which provides external data. The contracts are discreet in that external observers cannot detect the presence of the contract in the transaction log. They also hinge on knowledge of a *discrete* logarithm, which is a plus.

Model

There are 3 parties involved in the contract process: Alice, Bob, and Olivia. Alice and Bob are contract counterparties, while Olivia is the oracle. Alice and Bob do not trust each other and do not need to know any legal identifying information about each other, but they must be able to communicate over an authenticated channel, and they must be able to persistently recognize each other. Alice and Bob also must be able to receive signed broadcast messages from Olivia. Olivia does not need to be aware of Alice and Bob, and ideally she has no contact other than broadcasting information. The information is compact enough that broadcast could take place over the Bitcoin network itself, though this should not be necessary.

The DLC protocol can be used for a wide variety of contracts, covering most cases where payouts between parties depend on a publicly known number in the future. In this example, Alice and Bob make and execute

Higher-level Scaling in Layer 2: Blockstack Subdomains

Aaron Blankstein
aaron.blankstein.com
Blockstack PBC

September 22, 2017

Blockstack implements a naming system on top of Bitcoin. This is achieved through translating name operations, e.g., registrations, transfers, or name data updates, into Bitcoin transactions, using the `OP_RETURN` field of Bitcoin transactions to hold the operation's data. Rather than storing data associated with a name (domain name in the Blockstack Naming System) directly in transactions, data (zonefiles) are stored in a gossip-replicated network called Atlas. The security of the mapping from domain to zonefile is ensured by storing the hash of that data on the Bitcoin chain.

While this separation of data storage from the blockchain itself enables a large amount of scalability, each domain name registration requires *three* bitcoin transactions: a `PREORDER`, which reserves a hashed name, a `REGISTER`, which reveals that name, and an `UPDATE` which sets the associated zonefile hash for a name. This three-step process poses both usability and scalability concerns. With respect to usability, this means that a new user of Blockstack would not only need to pay three transaction fees, which can be prohibitive, but would also need to wait for all of these transactions to confirm, and other indexers in the naming network to read and integrate that information. Blockstack's indexers are, by default, configured to only process transactions which have at least 6 confirmations, which means a new user would have to wait up to 9 blocks before possessing a usable name. For scalability, this process limits the rate of new user registrations. This system can handle roughly 4000 new users per hour, and would need to account for nearly all Bitcoin transactions in that hour.

To address these concerns, Blockstack developed a system of subdomains which allows for a weaker form of name ownership, but one that dramatically reduces the costs of a registration, both to the network and the individual. A subdomain of the form `foo.bar.id` is a name that is resolved by a network node by first resolving `bar.id` and then finding a valid subdomain entry for `foo`. Subdomains may be owned by bitcoin addresses unrelated to the domain operator's bitcoin address. Because of this, the operations that flow out of name ownership can be performed by the *subdomain owner*. For example, the owner

Atomically Trading with Roger: Gambling on the success of a hardfork

Patrick McCorry¹, Ethan Heilman² and Andrew Miller^{3,4}

¹ University College London p.mccorry@ucl.ac.uk

² Boston University heilman@bu.edu

³ University of Illinois at Urbana-Champaign soc1024@illinois.edu

⁴ Initiative for Cryptocurrencies and Contracts, initc3.org

Abstract. We present atomic trade protocols for Bitcoin and Ethereum that can bind two parties to swap coins in the event that two blockchains emerge from a single “pre-fork” blockchain. This work is motivated by a bet between two members of the Bitcoin community, Loaded and Roger Ver, to trade 60,000 bitcoins in the event that Bitcoin Unlimited’s planned hardfork occurs and the blockchain splits into two distinct forks. Additionally we study several ways to provide replay protection in the event of hardfork alongside a novel mechanism called migration inputs. We provide a detailed survey and history of previous softforks and hardforks in Ethereum and Bitcoin.

1 Introduction

Bitcoin [30] is the world’s first successful and most valuable cryptocurrency. In June 2017, it reached a market cap of \$43bn USD [11] and processed $\approx 250,000$ transactions per day [5]. However, Bitcoin’s future is uncertain; it is reaching its capacity limits, and so far the community has failed to reach consensus on how best to increase its capacity.

One proposed approach for increasing capacity, called Bitcoin Unlimited (BU), involves removing the 1-megabyte-per-block parameter that most directly effects the capacity limit [35]. A competing approach, the Core Roadmap [27], calls for a technical upgrade called SegWit [25], followed by deployment of the overlay payment network, Lightning [31]. Both approaches require changing the network’s consensus rules; however there is a critical difference between them, BU is implemented as a hardfork upgrade, whereas Core relies on softforks. These two approaches are mutually incompatible: unlike a hardfork, a softfork is “forward-compatible” in the sense that blocks mined using the new rules can still be processed by non-upgraded clients (for additional details see Section 2.3).

If the community remains divided on which approach to support, then the result may be a schism, where each faction maintains a distinct fork of Bitcoin with mutually incompatible consensus rules.⁵ Both blockchains will diverge post-fork,

⁵ A schism has previously occurred in the case of Ethereum, whose *TheDAO* hardfork precipitated a split into Ethereum and Ethereum Classic.

ValueShuffle: Mixing Confidential Transactions*

Tim Ruffing
Saarland University

Pedro Moreno-Sanchez
Purdue University

In Bitcoin’s initial design, privacy plays only a minor role. The initial perception of Bitcoin providing some anonymity and fungibility has been refuted by a vast set of academic works [1]–[4]. This state of affairs has led to a plethora of privacy-enhancing technologies [5]–[10] aiming at overcoming these shortcomings either by defining entirely new cryptocurrencies [11], [12] or without breaking with the fundamental design of Bitcoin.

A prominent example for the latter approach is Confidential Transactions (CT) [5], a novel transaction format to enforce payment value privacy in Bitcoin by hiding transacted values in homomorphic commitments. Another example is Stealth Addresses (SA) [6], a mechanism for payers to generate unique one-time addresses for improved payee anonymity.

To achieve payer anonymity, the most prevalent approach that retains compatibility with Bitcoin is coin mixing: A group of users exchange their coins with each other, effectively hiding the relations between funds and owners. Users jointly generate a CoinJoin [13] transaction that enables an atomic transfer of funds to fresh output addresses and prevents theft by design. If users exchange their output addresses by means of an anonymous P2P broadcast protocol [10], [14], mixing is possible in a decentralized fashion (*P2P coin mixing*) while making sure inputs cannot be linked to outputs even by malicious users in the mixing, and such malicious users cannot prevent the honest users from successfully completing the protocol.

However, the aforementioned approaches focus typically on just one aspect of privacy (payer anonymity, payee anonymity or payment value privacy), leaving the privacy landscape in Bitcoin orphan of a comprehensive privacy solution.

a) *Challenge:* To achieve comprehensive privacy, it is necessary to combine CT, SA, and coin mixing into a single solution. SA or other means to generate one-time addresses can be easily combined with coin mixing, but while CT has in fact been designed with CoinJoin mixing in mind, it is not clear that the trust models of CT and anonymous P2P broadcast protocols, which are required for decentralized coin mixing, can be made compatible. The design of CT assumes that a transaction is created by just one user, whereas in coin mixing it is a group of users who jointly need to create a CoinJoin transaction in a decentralized P2P fashion. A naive non-solution is that the users reveal their balances and the corresponding secrets (the blinding factors for the homomorphic commitments used in CT) to each other, and then create a CoinJoin transaction. However, this is

not an option for coin mixing because the users typically do not trust each other.

I. MIXING CONFIDENTIAL TRANSACTIONS

In this talk, we present ValueShuffle, the first coin mixing protocol compatible with CT. It enables a group of mutually distrusting users to create a CoinJoin confidential transaction, without revealing the relation between inputs and outputs or their payment values to each other. Since ValueShuffle successfully combines coin mixing, SA and the CT proposal, the resulting currency provides comprehensive privacy, i.e., payer anonymity, payee anonymity and value privacy. Since it builds upon CoinJoin, ValueShuffle inherits a variety of features crucial to its practical deployment in the Bitcoin ecosystem, e.g., compatibility with blockchain pruning.

By combining coin mixing with SA and CT, we exploit synergies which make P2P coin mixing both more efficient and more practical, thereby releasing the full potential of coin mixing. We achieve that goal by overcoming the two main limitations of current coin mixing approaches.

First, all forms of coin mixing have been heavily restricted to mixing funds of the same value, because otherwise it is trivial for an observer to link inputs and outputs together just based on their monetary value. Adding value privacy to coin mixing removes this restriction entirely but comes with the challenge of proving to the network that no money is created in the mixing, since payment values are no longer in clear.

Second, current P2P coin mixing protocols suffer from the problem that users are required to mix their funds (in a CoinJoin transaction) by sending them to a fresh address of their own first, which removes the trace to the owner. Only afterwards can users spend the mixed funds to a payee in a second transaction.

This two-step process renders mixing expensive for users, who pay additional fees and need to wait longer, and for the entire Bitcoin network, which has to process essentially twice the amount of transaction data. This is highly undesirable and creates a conflict between privacy and efficiency.

In ValueShuffle, instead, we rely on SA and CT to enable users to send their funds directly to the expected receivers in the CoinJoin transaction, which is arguably the most desirable mode of use of CoinJoin.

A. Overview of ValueShuffle

ValueShuffle is an extension of the P2P coin mixing protocol CoinShuffle++ [10], which is the result of instantiating the efficient message mixing protocol DiceMix [10] in the setting of CoinJoin-based coin mixing. To connect CoinShuffle++ with

*This is an extended abstract of the work “Mixing Confidential Transactions: Comprehensive Transaction Privacy for Bitcoin” which appeared at the Bitcoin Workshop 2017 and is available at <https://eprint.iacr.org/2017/238.pdf>.

ZeroLink: The Bitcoin Fungibility Framework



Authors

nopara73,
[Hidden Wallet](#),
adam.ficsor73@gmail.com

TDevD,
[Samourai Wallet](#),
[PGP](#)

Acknowledgements

Special thanks for Adam Gibson and Chris Belcher from [JoinMarket](#), Ethan Heilman from [TumbleBit](#), Dan Gershony from [Breeze Wallet](#) and Kristov Atlas from [Open Bitcoin Privacy Project](#) for tolerating my constant bugging and bothering to acquire their invaluable reviews, suggestions and feedbacks.

Support

186n7me3QKajQZJnUsVsezVhVrSwyFCCZ

Concurrency and Privacy with Payment-Channel Networks*

Giulio Malavolta[†]
Friedrich-Alexander University
Erlangen-Nürnberg

Pedro Moreno-Sanchez[†]
Purdue University

Aniket Kate
Purdue University

Matteo Maffei
TU Wien

Srivatsan Ravi
University of Southern
California

1. INTRODUCTION

Permissionless blockchains protocols such as Bitcoin are inherently limited in transaction throughput and latency. Therefore, in the forethought of a growing number of Bitcoin users and most importantly payments about them, scalability is considered today an important concern in the Bitcoin community. Among alternative proposals, the use of Bitcoin *payment channels* [1, 8] to realize off-chain payments has flourished as a promising approach to overcome the Bitcoin scalability issue. As a generalization, current efforts leverage a path of opened payment channels from the payer to the payee with enough capacity to settle their payments, effectively creating a *payment-channel network (PCN)* [8].

Many challenges, such as liquidity, routing scalability, privacy or concurrency, must be overcome before a PCN is widely deployed. Here, we study privacy and concurrency in PCNs and show an inherent trade-off between them.

The Privacy Challenge. Although it might seem that payment channels inherently improve the privacy of Bitcoin payments as they are no longer logged in the blockchain, the actual privacy guarantees are not clear among the community. Recent research works [3–5] propose privacy preserving protocols for payment hub networks, where all users perform off-chain payments through an unique intermediary. However, it is not clear how to extend these solutions to multi-hop PCNs. The Lightning Network [8], a multi-hop PCN, does not provide all the privacy guarantees of interest in a PCN. For instance, a payment routed through a path of payment channels includes an identifier that can be used by intermediate users to derive who is paying to whom [3]. In summary, the lack of rigorous definitions for the protocols, threat model and the privacy notions, hinders an in-depth security and privacy analysis of PCNs.

The Concurrency Challenge. The consensus algorithm eases the serialization of concurrent on-chain payments. However, the bulk of off-chain payments in a PCN are no longer required to be added to the blockchain and therefore no user has a view of all concurrent off-chain payments at any time for their serialization. On the other hand, individual users cannot easily avoid concurrency issues either since a payment might involve several intermediate users. As PCNs scale to a large number of payments, concurrent payments are likely to happen and concurrency issues must be thoroughly investigated.

2. OUR CONTRIBUTIONS

In this talk, we plan to present the following results from our research on the concurrency and privacy issues of PCNs:

- We formalize for the first time the security and privacy notions of interest for a PCN, namely *correct balance*, *value privacy* and *relationship anonymity*. Intuitively, correct balance ensures that no honest user loses coins as intermediate hop in the payment path. Value privacy ensures that the payment value is not leaked to off-path users. Finally, relationship anonymity guarantees that payer and payee of a payment remain anonymous among a set of possible payers and payees, even in the presence of an on-path adversary.

- We study the concurrency issues in PCNs and present two protocols Fulgor and Rayo that tackle this issue differently. Fulgor is a blocking protocol similar to other payment networks such as credit networks [6] that can lead to deadlocks where none of the concurrent payments go through. Overcoming this challenge, Rayo is the first protocol for PCNs guaranteeing non-blocking progress: At least one of the concurrent payments terminates.

- We characterize an arguably surprising tradeoff between privacy and concurrency in PCNs. In particular, we demonstrate that any PCN that enforces non-blocking progress inevitably reduces the anonymity set for sender and receiver of a payment, thereby weakening the privacy guarantees. Therefore, in contrast to Fulgor, Rayo inevitably provides relationship anonymity only against an off-path adversary.

- We formally describe Multi-Hop HTLC, a smart contract that lies at the core of Fulgor and Rayo and which ensures privacy properties even against users in the payment path from payer to payee. We provide an efficient instantiation based on the recently proposed system ZK-Boo [2], improving on previous proposals [9] by reducing the data required from 650 MB to 17 MB, the running time for the prover from 600 ms to 309 ms and the running time for verifying from 500 ms to 130 ms. Moreover, Multi-Hop HTLC is compatible with the current Bitcoin scripting system.

- We have implemented and evaluated a prototype of Fulgor and Rayo. Our results show that a privacy-preserving payment in a path with 10 intermediate users can be carried out in as few as 5 seconds and incurs on 17 MB of communication overhead. These results are in line with other privacy preserving payment systems [6, 7]. Additionally, our evaluation shows that Fulgor and Rayo can scale to cater a growing number of users with a small overhead that can be further reduced with an optimized implementation.

*A full version of this work has been accepted at CCS'17. A preprint version is available at <https://eprint.iacr.org/2017/820>

[†]Both authors are considered co-first authors.

Optimizing fee estimation via the mempool state

Karl-Johan Alm <karl@dglab.com>
DG Lab

September 12, 2017

Gigapixel microscopy using a flatbed scanner

Guoan Zheng,* Xiaoze Ou, and Changhui Yang

Department of Electrical Engineering, California Institute of Technology, Pasadena, CA
91125, USA

*gazheng@caltech.edu

Abstract: Microscopy imaging systems with a very wide field-of-view (FOV) are highly sought in biomedical applications. In this paper, we report a wide FOV microscopy imaging system that uses a low-cost scanner and a closed-circuit-television (CCTV) lens. We show that such an imaging system is capable to capture a 10 mm * 7.5 mm FOV image with 0.77 micron resolution, resulting in 0.54 gigapixels (10^9 pixels) across the entire image (26400 pixels * 20400 pixels). The resolution and field curve of the proposed system were characterized by imaging a USAF resolution target and a hole-array target. A 1.6 gigapixel microscopy image (0.54 gigapixel with 3 colors) of a pathology slide was acquired by using such a system for application demonstration.

OCIS codes: (170.0110) Imaging systems; (170.4730) Optical pathology; (170.0180) Microscopy

References and links

1. M. Oheim, "Advances and challenges in high-throughput microscopy for live-cell subcellular imaging," *Expert Opinion on Drug Discovery*, 1-17 (2011).
2. A. VanderLugt, *Optical signal processing* (Wiley New York, 1992).
3. J. Gilbertson, J. Ho, L. Anthony, D. Jukic, Y. Yagi, and A. Parwani, "Primary histologic diagnosis using automated whole slide imaging: a validation study," *BMC clinical pathology* 6, 4 (2006).
4. <http://www.dmetrix.net/techtutorial1.shtml>.
5. G. Zheng, S. A. Lee, Y. Antebi, M. B. Elowitz, and C. Yang, "The ePetri dish, an on-chip cell imaging platform based on subpixel perspective sweeping microscopy (SPSM)," *Proceedings of the National Academy of Sciences* 108, 16889-16894 (2011).
6. W. Bishara, T. Su, A. Coskun, and A. Ozcan, "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution," *Optics Express* 18, 11181-11191 (2010).
7. J. Wu, X. Cui, G. Zheng, Y. M. Wang, L. M. Lee, and C. Yang, "Wide field-of-view microscope based on holographic focus grid illumination," *Optics letters* 35, 2188-2190 (2010).
8. J. Wu, G. Zheng, Z. Li, and C. Yang, "Focal plane tuning in wide-field-of-view microscope with Talbot pattern illumination," *Optics letters* 36, 2179-2181 (2011).
9. J. Di, J. Zhao, H. Jiang, P. Zhang, Q. Fan, and W. Sun, "High resolution digital holographic microscopy with a wide field of view based on a synthetic aperture technique and use of linear CCD scanning," *Applied optics* 47, 5654-5659 (2008).
10. M. Lee, O. Yaglidere, and A. Ozcan, "Field-portable reflection and transmission microscopy based on lensless holography," *Biomedical Optics Express* 2, 2721-2730 (2011).
11. K. Fife, A. Gamal, and H. Wong, "A 3mpixel multi-aperture image sensor with 0.7 um pixels in 0.11 um cmos," in *IEEE ISSCC Digest of Technical Papers*, 2008), 48-49.
12. M. Ben-Ezra, "Large-Format Tile-Scan Camera," *IEEE Computer Graphics and Applications*, 49-61 (2011).
13. S. Wang and W. Heidrich, "The Design of an Inexpensive Very High Resolution Scan Camera System," *Computer Graphics Forum* 23, 441-450 (2004).
14. S. A. Lee, R. Leitao, G. Zheng, S. Yang, A. Rodriguez, and C. Yang, "Color Capable Sub-Pixel Resolving Optofluidic Microscope and Its Application to Blood Cell Imaging for Malaria Diagnosis," *PloS one* 6, e26127 (2011).
15. J. Mallon and P. F. Whelan, "Calibration and removal of lateral chromatic aberration in images," *Pattern recognition letters* 28, 125-135 (2007).
16. http://www.linhofstudio.com/products/cameras/anagramm_digital_reproduction/anagramm_digital_reproduction.html.
17. G. Zheng, C. Kolner, and C. Yang, "Microscopy refocusing and dark-field imaging by using a simple LED array," *Optics letters* 36, 3987-3989 (2011).

On the Technology Prospects and Investment Opportunities for Scalable Neuroscience

Thomas Dean^{**1,2,3**} Biafra Ahanonu^{**3**} Mainak Chowdhury^{**3**}
Anjali Datta^{**3**} Andre Esteva^{**3**} Daniel Eth^{**3**} Nobie Redmon^{**3**}
Oleg Rumyantsev^{**3**} Ysis Tarter^{**3**}

1 Google Research, **2** Brown University, **3** Stanford University

MASS EXTINCTION AND THE STRUCTURE OF THE MILKY WAY

M. D. Filipović¹, J. Horner^{2,3}, E. J. Crawford¹, N. F. H. Tothill¹¹*University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW 1797, Australia
E-mail: m.filipovic@uws.edu.au n.tothill@uws.edu.au e.crawford@uws.edu.au*²*School of Physics, University of New South Wales, Sydney 2052, Australia*³*Australian Centre for Astrobiology, University of New South Wales, Sydney 2052, Australia
E-mail: j.a.horner@unsw.edu.au*

(Received: August 2013; Accepted: September 2013)

SUMMARY: We use the most up to date Milky Way model and solar orbit data in order to test the hypothesis that the Sun's galactic spiral arm crossings cause mass extinction events on Earth. To do this, we created a new model of the Milky Way's spiral arms by combining a large quantity of data from several surveys. We then combined this model with a recently derived solution for the solar orbit to determine the timing of the Sun's historical passages through the Galaxy's spiral arms. Our new model was designed with a symmetrical appearance, with the major alteration being the addition of a spur at the far side of the Galaxy. A correlation was found between the times at which the Sun crosses the spiral arms and six known mass extinction events. Furthermore, we identify five additional historical mass extinction events that might be explained by the motion of the Sun around our Galaxy. These five additional significant drops in marine genera that we find include significant reductions in diversity at 415, 322, 300, 145 and 33 Myr ago. Our simulations indicate that the Sun has spent ~60% of its time passing through our Galaxy's various spiral arms. Also, we briefly discuss and combine previous work on the Galactic Habitable Zone with the new Milky Way model.

Key words. Galaxy: structure – Physical data and processes: Astrobiology – Solar system: general – Galaxy: solar neighbourhood

1. INTRODUCTION

Mass extinctions have the effect of wiping the biological slate clean, freeing up ecological niches and thus producing explosions in biodiversity (e.g. McElwain and Punyasena 2007; Alroy 2008). In the past, several explanations have been proposed to resolve ancient mass extinctions, including vast outpourings of flood basalt (such as the Deccan and Siberian Traps; e.g. Wignall 2001), periods of global glaciation (Mayhew et al. 2008) and the impact of large

asteroids and comets upon the Earth (e.g. Alvarez et al. 1980; Bottke et al. 2007). Of these, extreme geological and climate phenomena such as flood basalt outpouring and “snowball Earth” glaciations appear to be very rare and randomly-occurring events in the Earth's history. Overholt et al. (2009) investigate Earth's climate as a function of location in the Galaxy, however, no obvious correlation could be drawn. On the other hand, it is well established that the Earth has been continually pummelled by asteroidal and cometary impactors throughout its history, a process that will continue well into the fu-

A Roadmap to Interstellar Flight

Philip Lubin

Physics Dept, UC Santa Barbara

lubin@deepspace.ucsb.edu

submitted to JBIS April 2015

JBIS Vol. 69, pp. 40-72 Feb 2016

Current version 15-w7-4 (10/18/16)

Abstract – In the nearly 60 years of spaceflight we have accomplished wonderful feats of exploration that have shown the incredible spirit of the human drive to explore and understand our universe. Yet in those 60 years we have barely left our solar system with the Voyager 1 spacecraft launched in 1977 finally leaving the solar system after 37 years of flight at a speed of 17 km/s or less than 0.006% the speed of light. As remarkable as this, to reach even the nearest stars with our current propulsion technology will take 100 millennium. We have to radically rethink our strategy or give up our dreams of reaching the stars, or wait for technology that does not currently exist. While we all dream of human spaceflight to the stars in a way romanticized in books and movies, it is not within our power to do so, nor it is clear that this is the path we should choose. We posit a path forward, that while not simple, it is within our technological reach. We propose a roadmap to a program that will lead to sending relativistic probes to the nearest stars and will open up a vast array of possibilities of flight both within our solar system and far beyond. Spacecraft from gram level complete spacecraft on a wafer (“wafersats”) that reach more than $\frac{1}{4} c$ and reach the nearest star in 20 years to spacecraft with masses more than 10^5 kg (100 tons) that can reach speeds of greater than 1000 km/s. These systems can be propelled to speeds currently unimaginable with existing propulsion technologies. To do so requires a fundamental change in our thinking of both propulsion and in many cases what a spacecraft is. In addition to larger spacecraft, some capable of transporting humans, we consider functional spacecraft on a wafer, including integrated optical communications, imaging systems, photon thrusters, power and sensors combined with directed energy propulsion. The costs can be amortized over a very large number of missions beyond relativistic spacecraft as such planetary defense, beamed energy for distant spacecraft, sending power back to Earth, stand-off composition analysis of solar system targets, long range laser communications, SETI searches and even terra forming. Exploring the nearest stars and exo-planets would be a profound voyage for humanity, one whose non-scientific implications would be enormous. It is time to begin this inevitable journey far beyond our home.

Generating Natural Language Inference Chains

Vladyslav Kolesnyk and Tim Rocktäschel and Sebastian Riedel

University College London

London, UK

vladyslav.kolesnyk.12@ucl.ac.uk, {t.rocktaschel,s.riedel}@cs.ucl.ac.uk

Abstract

The ability to reason with natural language is a fundamental prerequisite for many NLP tasks such as information extraction, machine translation and question answering. To quantify this ability, systems are commonly tested whether they can recognize textual entailment, *i.e.*, whether one sentence can be inferred from another one. However, in most NLP applications only single source sentences instead of sentence pairs are available. Hence, we propose a new task that measures how well a model can *generate* an entailed sentence from a source sentence. We take entailment-pairs of the Stanford Natural Language Inference corpus and train an LSTM with attention. On a manually annotated test set we found that 82% of generated sentences are correct, an improvement of 10.3% over an LSTM baseline. A qualitative analysis shows that this model is not only capable of shortening input sentences, but also inferring new statements via paraphrasing and phrase entailment. We then apply this model recursively to input-output pairs, thereby generating natural language inference chains that can be used to automatically construct an entailment graph from source sentences. Finally, by swapping source and target sentences we can also train a model that given an input sentence invents additional information to generate a new sentence.

1 Introduction

The ability to determine entailment or contradiction between natural language text is essential

for improving the performance in a wide range of natural language processing tasks. Recognizing Textual Entailment (RTE) is a task primarily designed to determine whether two natural language sentences are independent, contradictory or in an entailment relationship where the second sentence (the hypothesis) can be inferred from the first (the premise). Although systems that perform well in RTE could potentially be used to improve question answering, information extraction, text summarization and machine translation [6], only in few of such downstream NLP tasks sentence-pairs are actually available. Usually, only a single source sentence (*e.g.* a question that needs to be answered or a source sentence that we want to translate) is present and models need to come up with their own hypotheses and commonsense knowledge inferences.

The release of the large Stanford Natural Language Inference (SNLI) corpus [2] allowed end-to-end differentiable neural networks to outperform feature-based classifiers on the RTE task [3, 4, 12, 15, 17].

In this work, we go a step further and investigate how well recurrent neural networks can produce true hypotheses given a source sentence. Furthermore, we qualitatively demonstrate that by only training on input-output pairs and recursively generating entailed sentence we can generate natural language inference chains (see Figure 1 for an example). Note that every inference step is interpretable as it is mapping one natural language sentence to another one.

Our contributions are fourfold: (i) we propose

The Renewed Case for the Reduced Instruction Set Computer: Avoiding ISA Bloat with Macro-Op Fusion for RISC-V

Christopher Celio, Palmer Dabbelt, David Patterson, Krste Asanović

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
celio@eecs.berkeley.edu

Abstract—This report makes the case that a well-designed Reduced Instruction Set Computer (RISC) can match, and even exceed, the performance and code density of existing commercial Complex Instruction Set Computers (CISC) while maintaining the simplicity and cost-effectiveness that underpins the original RISC goals [12].

We begin by comparing the dynamic instruction counts and dynamic instruction bytes fetched for the popular proprietary ARMv7, ARMv8, IA-32, and x86-64 Instruction Set Architectures (ISAs) against the free and open RISC-V RV64G and RV64GC ISAs when running the SPEC CINT2006 benchmark suite. RISC-V was designed as a very small ISA to support a wide range of implementations, and has a less mature compiler toolchain. However, we observe that on SPEC CINT2006 RV64G executes on average 16% more instructions than x86-64, 3% more instructions than IA-32, 9% more instructions than ARMv8, but 4% fewer instructions than ARMv7.

CISC x86 implementations break up complex instructions into smaller internal RISC-like *micro-ops*, and the RV64G instruction count is within 2% of the x86-64 retired micro-op count. RV64GC, the compressed variant of RV64G, is the densest ISA studied, fetching 8% fewer dynamic instruction bytes than x86-64. We observed that much of the increased RISC-V instruction count is due to a small set of common multi-instruction idioms.

Exploiting this fact, the RV64G and RV64GC *effective instruction* count can be reduced by 5.4% on average by leveraging *macro-op fusion*. Combining the compressed RISC-V ISA extension with macro-op fusion provides both the densest ISA and the fewest dynamic operations retired per program, reducing the motivation to add more instructions to the ISA. This approach retains a single simple ISA suitable for both low-end and high-end implementations, where high-end implementations can boost performance through microarchitectural techniques.

Compiler tool chains are a continual work-in-progress, and the results shown are a snapshot of the state as of July 2016 and are subject to change.

I. INTRODUCTION

The Instruction Set Architecture (ISA) specifies the set of instructions that a processor must understand and the expected effects of each instruction. One of the goals of the RISC-V project was to produce an ISA suitable for a wide range of implementations from tiny microcontrollers to the largest supercomputers [14]. Hence, RISC-V was designed with a much smaller number of simple standard instructions compared to other popular ISAs, including other RISC-inspired ISAs. A simple ISA is clearly a benefit for a small resource-constrained microcontroller, but how much performance is lost for high-performance implementations by not supporting the numerous instruction variants provided by popular proprietary ISAs?

A casual observer might argue that a processor's performance increases when it executes fewer instructions for a given

program, but in reality, the performance is more accurately described by the Iron Law of Performance [8]:

$$\frac{\text{seconds}}{\text{program}} = \frac{\text{cycles}}{\text{instruction}} * \frac{\text{seconds}}{\text{cycle}} * \frac{\text{instructions}}{\text{program}}$$

The ISA is just an abstract boundary; behind the scenes the processor may choose to implement instructions in any number of ways that trade off $\frac{\text{cycles}}{\text{instruction}}$, or *CPI*, and $\frac{\text{seconds}}{\text{cycle}}$, or *frequency*.

For example, a fairly powerful x86 instruction is the *repeat move* instruction (`rep movs`), which copies *C* bytes of data from one memory location to another:

```
// psuedo-code for a 'repeat move' instruction
for (i=0; i < C; i++)
    d[i] = s[i];
```

Implementations of the x86 ISA break up the *repeat move* instruction into smaller operations, or *micro-ops*, that individually perform the required operations of loading the data from the old location, storing the data to the new location, incrementing the address pointers, and checking to see if the end condition has been met. Therefore, a raw comparison of instruction counts may hide a significant amount of work and complexity to execute a particular benchmark.

In contrast to the process of generating many micro-ops from a single ISA instruction, several commercial micro-processors perform *macro-op fusion*, where several ISA instructions are fused in the decode stage and handled as one internal operation. As an example, compare-and-branch is a very commonly executed idiom, and the RISC-V ISA includes a full register-register magnitude comparison in its branch instructions. However, both ARM and x86 typically require two ISA instructions to specify a compare-and-branch. The first instruction performs the *comparison* and sets a condition code, and the second instruction performs the *jump-on-condition-code*. While it would seem that ARM and x86 would have a penalty of one additional instruction on nearly every loop compared to RISC-V, the reality is more complicated. Both ARM and Intel employ the technique of macro-op fusion, in which the processor front-end detects these two-instruction compare-and-branch sequences in the instruction stream and “fuses” them together into a single *macro-op*, which can then be handled as a single compare-and-branch instruction by the processor back-end to reduce the effective dynamic instruction count.¹

¹The reality can be even more complicated. Depending on the micro-architecture, the front-end may fuse the two instructions together to save decode, allocation, and commit bandwidth, but break them apart in the execution pipeline for critical path or complexity reasons [6].

THE MICROSOFT 2016 CONVERSATIONAL SPEECH RECOGNITION SYSTEM

W. Xiong, J. Droppo, X. Huang, F. Seide, M. Seltzer, A. Stolcke, D. Yu and G. Zweig

Microsoft Research

ABSTRACT

We describe Microsoft’s conversational speech recognition system, in which we combine recent developments in neural-network-based acoustic and language modeling to advance the state of the art on the Switchboard recognition task. Inspired by machine learning ensemble techniques, the system uses a range of convolutional and recurrent neural networks. I-vector modeling and lattice-free MMI training provide significant gains for all acoustic model architectures. Language model rescoring with multiple forward and backward running RNNLMs, and word posterior-based system combination provide a 20% boost. The best single system uses a ResNet architecture acoustic model with RNNLM rescoring, and achieves a word error rate of 6.9% on the NIST 2000 Switchboard task. The combined system has an error rate of 6.2%, representing an improvement over previously reported results on this benchmark task.

Index Terms— Conversational speech recognition, convolutional neural networks, recurrent neural networks, VGG, ResNet, LACE, BLSTM.

1. INTRODUCTION

Recent years have seen a rapid reduction in speech recognition error rates as a result of careful engineering and optimization of convolutional and recurrent neural networks. While the basic structures have been well known for a long period [1, 2, 3, 4, 5, 6, 7], it is only recently that they have dominated the field as the best models for speech recognition. Surprisingly, this is the case for both acoustic modeling [8, 9, 10, 11, 12, 13] and language modeling [14, 15]. In comparison to standard feed-forward MLPs or DNNs, these acoustic models have the ability to model a large amount of acoustic context with temporal invariance, and in the case of convolutional models, with frequency invariance as well. In language modeling, recurrent models appear to improve over classical N-gram models through the generalization ability of continuous word representations [16]. In the meantime, ensemble learning has become commonly used in several neural models [17, 18, 15], to improve robustness by reducing bias and variance.

In this paper, we use ensembles of models extensively, as well as improvements to individual component models, to advance the state-of-the-art in conversational telephone speech recognition (CTS), which has been a benchmark speech recognition task since the 1990s. The main features of this system are:

1. An ensemble of two fundamental acoustic model architectures, convolutional neural nets (CNNs) and long-short-term memory nets (LSTMs), with multiple variants of each
2. An attention mechanism in the LACE CNN which differentially weights distant context [19]
3. Lattice-free MMI training [20, 21]
4. The use of i-vector based adaptation [22] in all models

5. Language model (LM) rescoring with multiple, recurrent neural net LMs [14] running in both forward and reverse direction
6. Confusion network system combination [23] coupled with search for best system subset, as necessitated by the large number of candidate systems.

The remainder of this paper describes our system in detail. Section 2 describes the CNN and LSTM models. Section 3 describes our implementation of i-vector adaptation. Section 4 presents our lattice-free MMI training process. Language model rescoring is a significant part of our system, and described in Section 5. Experimental results are presented in Section 6, followed by a discussion of related work and conclusions.

2. CONVOLUTIONAL AND LSTM NEURAL NETWORKS

We use three CNN variants. The first is the VGG architecture of [24]. Compared to the networks used previously in image recognition, this network uses small (3x3) filters, is deeper, and applies up to five convolutional layers before pooling. The second network is modeled on the ResNet architecture [25], which adds highway connections [26], i.e. a linear transform of each layer’s input to the layer’s output [26, 27]. The only difference is that we move the Batch Normalization node to the place right before each ReLU activation.

The last CNN variant is the LACE (layer-wise context expansion with attention) model [19]. LACE is a TDNN [3] variant in which each higher layer is a weighted sum of nonlinear transformations of a window of lower layer frames. In other words, each higher layer exploits broader context than lower layers. Lower layers focus on extracting simple local patterns while higher layers extract complex patterns that cover broader contexts. Since not all frames in a window carry the same importance, an attention mask is applied. The LACE model differs from the earlier TDNN models e.g. [3, 28] in the use of a learned attention mask and ResNet like linear pass-through. As illustrated in detail in Figure 1, the model is composed of 4 blocks, each with the same architecture. Each block starts with a convolution layer with stride 2 which sub-samples the input and increases the number of channels. This layer is followed by 4 RELU-convolution layers with jump links similar to those used in ResNet. Table 1 compares the layer structure and parameters of the three CNN architectures.

While our best performing models are convolutional, the use of long short-term memory networks is a close second. We use a bidirectional architecture [29] without frame-skipping [9]. The core model structure is the LSTM defined in [8]. We found that using networks with more than six layers did not improve the word error rate on the development set, and chose 512 hidden units, per direction, per layer, as that provided a reasonable trade-off between training time and final model accuracy. Network parameters for different configurations of the LSTM architecture are summarized in Table 2.

Quantum-Chemical Insights from Deep Tensor Neural Networks

Kristof T. Schütt¹, Farhad Arbabzadah¹, Stefan Chmiela¹, Klaus R. Müller^{1,2,*} and Alexandre Tkatchenko^{3,4,†}

¹*Machine Learning Group, Technische Universität Berlin, Marchstr. 23, 10587 Berlin, Germany*

²*Department of Brain and Cognitive Engineering, Korea University,
Anam-dong, Seongbuk-gu, Seoul 136-713, Republic of Korea*

³*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195, Berlin, Germany*

⁴*Physics and Materials Science Research Unit, University of Luxembourg, L-1511 Luxembourg*

Learning from data has led to paradigm shifts in a multitude of disciplines, including web, text, and image search, speech recognition, as well as bioinformatics. Can machine learning spur similar breakthroughs in understanding quantum many-body systems? Here we develop an efficient deep learning approach that enables spatially and chemically resolved insights into quantum-mechanical observables of molecular systems. We unify concepts from many-body Hamiltonians with purpose-designed deep tensor neural networks (DTNN), which leads to size-extensive and uniformly accurate (1 kcal/mol) predictions in compositional and configurational chemical space for molecules of intermediate size. As an example of chemical relevance, the DTNN model reveals a classification of aromatic rings with respect to their stability – a useful property that is not contained as such in the training dataset. Further applications of DTNN for predicting atomic energies and local chemical potentials in molecules, reliable isomer energies, and molecules with peculiar electronic structure demonstrate the high potential of machine learning for revealing novel insights into complex quantum-chemical systems.

Chemistry permeates all aspects of our life, from the development of new drugs to the food that we consume and materials we use on a daily basis. Chemists rely on empirical observations based on creative and painstaking experimentation that leads to eventual discoveries of molecules and materials with desired properties and mechanisms to synthesize them. Many discoveries in chemistry can be guided by searching large databases of experimental or computational molecular structures and properties by using concepts based on chemical similarity. Because the structure and properties of molecules are determined by the laws of quantum mechanics, ultimately chemical discovery must be based on fundamental quantum principles. Indeed, electronic structure calculations and intelligent data analysis (machine learning, ML) have recently been combined aiming towards the goal of accelerated discovery of chemicals with desired properties [1–8]. However, so far the majority of these pioneering efforts have focused on the construction of reduced models trained on large datasets of density-functional theory calculations. In this work, we develop an efficient deep learning approach that enables spatially and chemically resolved insights into quantum-mechanical properties of molecular systems beyond those trivially contained in the training dataset. Obviously, computational models are not predictive if they lack accuracy. In addition to being interpretable, size extensive and efficient, our deep tensor neural network (DTNN) approach is uniformly accurate (1 kcal/mol) throughout compositional and configurational chemical space. On the more fundamental side, the mathematical construction of the DTNN model provides statistically rigorous partitioning of extensive molecular properties into atomic contributions – a long-standing challenge for quantum-mechanical calculations of molecules.

MOLECULAR DEEP TENSOR NEURAL NETWORKS

It is common to use a carefully chosen representation of the problem at hand as a basis for machine learning [9–11]. For example, molecules can be represented as Coulomb matrices [7, 12, 13], scattering transforms [14], bags of bonds (BoB) [15], smooth overlap of atomic positions (SOAP) [16, 17], or generalized symmetry functions [18, 19]. Kernel-based learning of molecular properties transforms these representations non-linearly by virtue of kernel functions. In contrast, deep neural networks [20] are able to infer the underlying regularities and learn an efficient representation in a layer-wise fashion [21].

Molecular properties are governed by the laws of quantum mechanics, which yield the remarkable flexibility of chemical systems, but also impose constraints on the behavior of bonding in molecules. The approach presented here utilizes the many-body Hamiltonian concept for the construction of the DTNN architecture (see Fig. 1), embracing the principles of quantum chemistry, while maintaining the full flexibility of a complex data-driven learning machine.

DTNN receives molecular structures through a vector of nuclear charges Z and a matrix of atomic distances D ensuring rotational and translational invariance by construction (Fig. 1A). The distances are expanded in a Gaussian basis, yielding a feature vector \mathbf{d}_{ij} , which accounts for the different nature of interactions at various distance regimes.

The total energy E_M for the molecule M composed of N atoms is written as a sum over N atomic energy contributions E_i , thus satisfying permutational invariance with respect to atom indexing. Each atom i is represented by a coefficient vector $\mathbf{c} \in \mathbb{R}^B$, where B is the number of basis functions, or features. Motivated by quantum-chemical atomic basis set expansions, we assign an atom type-specific descriptor vector \mathbf{c}_i to these coefficients $\mathbf{c}_i^{(0)}$. Subsequently, this atomic expansion is repeatedly refined by pairwise interactions with the

A computer program for simulating time travel and a possible 'solution' for the grandfather paradox

Doron Friedman, The Interdisciplinary Center, Herzliya, Israel

doronf@idc.ac.il

Abstract

While the possibility of time travel in physics is still debated, the explosive growth of virtual-reality simulations opens up new possibilities to rigorously explore such time travel and its consequences in the digital domain. Here we provide a computational model of time travel and a computer program that allows exploring digital time travel. In order to explain our method we formalize a simplified version of the famous grandfather paradox, show how the system can allow the participant to go back in time, try to kill their ancestors before they were born, and experience the consequences. The system has even come up with scenarios that can be considered consistent "solutions" of the grandfather paradox. We discuss the conditions for digital time travel, which indicate that it has a large number of practical applications.

1. Introduction

In the principal paradox of time travel a person travels back in time and kills his grandfather before the grandfather meets the time traveler's grandmother. As a consequence, one of the traveler's parents, and therefore the traveler himself, would never have been born. This would imply that the traveller could not have travelled back in time, which means that the grandfather would still be alive, which now makes it possible for the traveler to be born, travel back in time and kill his grandfather, hence a paradox.

In this paper we describe a computer program that allows us to interactively explore the consequences of "changing the past" in a narrative. Our system uses well-known techniques from automated reasoning to compute the specific consequences of modifying history. The contribution of this paper is in the application of automated reasoning to the experience of time travel and in the implementation of this method in a computer

HYPERNETWORKS

David Ha*, Andrew Dai, Quoc V. Le

Google Brain

{hadavid, adai, qvl}@google.com

ABSTRACT

This work explores hypernetworks: an approach of using a one network, also known as a hypernetwork, to generate the weights for another network. Hypernetworks provide an abstraction that is similar to what is found in nature: the relationship between a genotype – the hypernetwork – and a phenotype – the main network. Though they are also reminiscent of HyperNEAT in evolution, our hypernetworks are trained end-to-end with backpropagation and thus are usually faster. The focus of this work is to make hypernetworks useful for deep convolutional networks and long recurrent networks, where hypernetworks can be viewed as relaxed form of weight-sharing across layers. Our main result is that hypernetworks can generate non-shared weights for LSTM and achieve near state-of-the-art results on a variety of sequence modelling tasks including character-level language modelling, handwriting generation and neural machine translation, challenging the weight-sharing paradigm for recurrent networks. Our results also show that hypernetworks applied to convolutional networks still achieve respectable results for image recognition tasks compared to state-of-the-art baseline models while requiring fewer learnable parameters.

1 INTRODUCTION

In this work, we consider an approach of using a small network (called a “hypernetwork”) to generate the weights for a larger network (called a main network). The behavior of the main network is the same with any usual neural network: it learns to map some raw inputs to their desired targets; whereas the hypernetwork takes a set of inputs that contain information about the structure of the weights and generates the weight for that layer (see Figure 1).

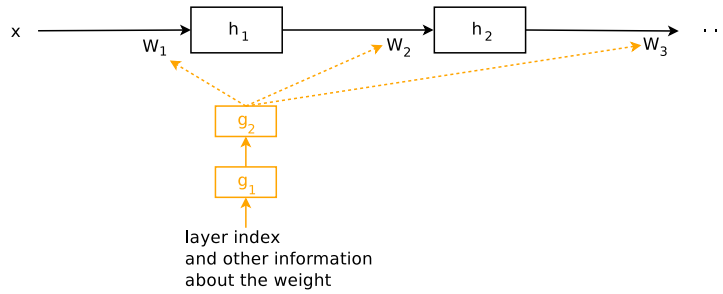


Figure 1: A hypernetwork generates the weights for a feedforward network. Black connections and parameters are associated the main network whereas orange connections and parameters are associated with the hypernetwork.

HyperNEAT (Stanley et al., 2009) is an example of hypernetworks where the inputs are a set of virtual coordinates for each weight in the main network. In this work, we will focus on a more powerful approach where the input is an embedding vector that describes the entire weights of a given layer. Our embedding vectors can be fixed parameters that are also learned during end-to-end training, allowing approximate weight-sharing within a layer and across layers of the main network. In

*Work done as a member of the Google Brain Residency program (g.co/brainresidency).

A low-cost, Arduino-like development kit for single-element ultrasound imaging*

Luc Jonveaux¹

Abstract—An open-source software ecosystem for ultrasound imaging is widely available to developers, however, limited resources can be found on the open-hardware side. The focus of this work was to develop an easy-to-use platform kit (hardware and software) for providing the community a complete experimental setup for ultrasound imaging at a low cost, without the need of specific equipment. The goal of this work resembles the needs of medical systems in the 80's where analog techniques using single-sensor devices were prominent.

To this end, two open-source, arduino-like modules have been developed for building a simple, yet complete, single-channel analog front-end system, where all the intermediary signals are readily accessible by the user. A single-channel architecture avoids the beamforming overhead, though it limits the quality of the captured image, and brings simplicity to the system.

The modules were tested using re-purposed ultrasound mechanical probes, as well as non-medical transducers. Furthermore, different digital acquisition systems were utilized for providing the images of interest. The developed modules can also be used in Radio Frequency (RF) projects, non-destructive testing and control projects, as well as in low-cost medical imaging projects.

I. INTRODUCTION

A. Approach

Ultrasound imaging has evolved since the first ultrasound machine appeared. The first devices were using single-sensor (transducers) techniques, coupled with mechanical scanning [22]. The architecture of such systems, as shown in Fig. 1, is well-known and formed the basis of ultrasound imaging.

Mechanical scanning has its limitation, but also its strengths: a single signal channel, linked to a single sensor, means that the corresponding electronics are simplified, and the cost is reduced. Moreover, with progress made in different technical fields, mechanical probes are seen on the market again. Search in academic literature, and open-electronics resources, yielded little to no documentation of previous research to rebuild these mechanical ultrasound imaging devices.

To the best of the author's knowledge, there are no open-source hardware designs nor electronics accessible online for the analog-processing component. To bridge this gap, this work provides modules to the community to understand and recreate the electronic core of an ultrasound device.

B. This work

This kit consists of several modules mainly built from easily available components. **Two electronic modules were**

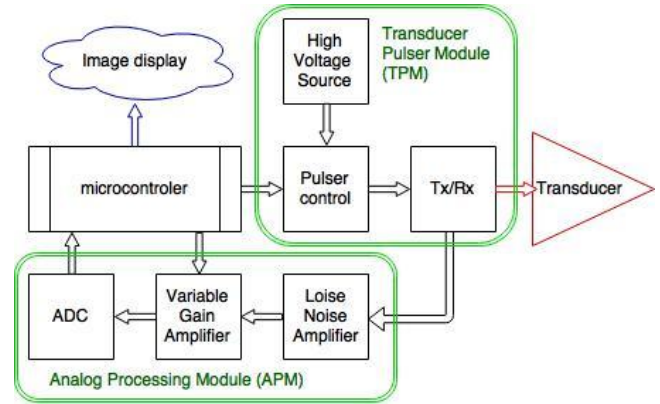


Fig. 1. Architecture of the ultrasound imaging system. The two custom boards, TPM and APM, are shown in green.

specifically designed to provide the basic development kit. These two modules, called the Ultrasound Imaging Analog Core (UIAC), as shown in Fig. 1, are:

- the **Transducer Pulser Module (TPM)**: designed to provide a precise high-voltage pulse, necessary to excite the sensor, while remaining robust enough to be controlled by an Arduino;
- the **Analog Processing Module (APM)**: designed to correctly process the raw ultrasound electric signal, while easily exposing all intermediary signals, and exposing a digital output to the user.

C. A module approach

We have chosen a modular approach to ensure that each key component inside the ultrasound image processing can easily be replaced and compared with another module. Each electronic module takes the place of a function in the signal processing chain or allows tapping into the different signals circulating between the blocks.

We have considered readily available open-source modules and recycled components (probes) to provide the user with building bricks for the basic non-medical ultrasound imaging tool.

We remind that the aim of this work is not to design an ultrasound probe: *the goal of the present article is rather to provide a basic open-source tool to understand ultrasound imaging technique and provide the analog core, unavailable today, as well as selecting proper off-the-shelf components for the other elements.*

This paper also discusses several possible options, keeping in mind that none are preferred and that the modular

*This work was not supported by any organization

¹Luc is just a independant maker, reachable at kelu124@gmail.com

Strong Federations: An Interoperable Blockchain Solution to Centralized Third-Party Risks

Johnny Dilley*, Andrew Poelstra*, Jonathan Wilkins*, Marta Piekarska*, Ben Gorlick*, and Mark Friedenbach*
Email: johnny, andrew, jonathan, marta, ben, mark @blockstream.com

Abstract—Bitcoin, the first peer-to-peer electronic cash system, opened the door to permissionless, private, and trustless transactions. Attempts to repurpose Bitcoin’s underlying blockchain technology have run up against fundamental limitations to privacy, faithful execution, and transaction finality. We introduce *Strong Federations*: publicly verifiable, Byzantine-robust transaction networks that facilitate movement of any asset between disparate markets, without requiring third-party trust. *Strong Federations* enable commercial privacy, with support for transactions where asset types and amounts are opaque, while remaining publicly verifiable. As in Bitcoin, execution fidelity is cryptographically enforced; however, *Strong Federations* significantly lower capital requirements for market participants by reducing transaction latency and improving interoperability. To show how this innovative solution can be applied today, we describe *Liquid*: the first implementation of *Strong Federations* deployed in a Financial Market.

I. INTRODUCTION

Bitcoin, proposed by Satoshi Nakamoto in 2008, is based on the idea of a *blockchain* [1]. A blockchain consists of a series of blocks, each of which is composed of time-stamped sets of transactions and a hash of the previous block, which connects the two together, as presented in Figure 1.

The underlying principle of Bitcoin’s design is that all participants in its network are on equal footing. They jointly trust proof-of-work [2] to validate and enforce the network’s rules, which obviates the need for central authorities such as clearinghouses. As a result, Bitcoin empowers a wide range of participants to be their own banks – storing, transacting, and clearing for themselves without the need for a third-party intermediary. Bitcoin’s network automatically enforces settlements between participants using publicly verifiable algorithms that avoid security compromises, expensive (or unavailable) legal infrastructure, third-party trust requirements, or the physical transportation of money. For the first time, users of a system have the ability to cryptographically verify other participants’ behaviors, enforcing rules based on mathematics that anyone can check and no one can subvert.

Due to its design, Bitcoin has characteristics that make it a vehicle of value unlike anything that previously existed. First, it eliminates most counterparty risk from transactions [3]. Second, it offers cryptographic proof of ownership of assets, as the knowledge of a cryptographic key defines ownership [4]. Third, it is a programmable asset, offering the ability to pay to a program, or a “smart contract”, rather than a passive account or a singular public key [5]. Fourth, and finally, it is a disruptive market mechanism for use cases such as point-

to-point real-time transfers, accelerated cross-border payment, B2B remittance, asset transfers, and micropayments [6].

A. Problem Statement

Because it is a global consensus system, Bitcoin’s decentralized network and public verifiability come with costs. Speed of execution and insufficient guarantees of privacy are two of Bitcoin’s limitations.

Bitcoin’s proof-of-work methodology was designed to process transactions on average only once every ten minutes, with large variance. As a result, Bitcoin is slow from a real-time transaction processing perspective. This creates spontaneous illiquidity for parties using bitcoin¹ as an intermediary, volatility exposure for those holding bitcoin for any length of time, and obstacles for the use of Bitcoin’s contracting features for fast settlements. Even after a transaction is processed, counterparties must generally wait until several additional blocks have been created before considering their transaction settled. This is because Bitcoin’s global ledger is at constant risk of *reorganization*, wherein very recent history can be modified or rewritten. This latency undermines many commercial applications, which require real-time, or nearly instant, execution². Today, solving this requires a centralized counterparty, which introduces a third-party risk.

Despite issues of short-term validation, Bitcoin excels on settlement finality, providing strong assurance against transaction reversals after adequate block confirmations. In contrast, legacy payment networks leave absolute final settlement in limbo for up to 120 days typically, though chargebacks have been allowed up to 8 years late [7], depending on policies imposed by the centralized network owner [8] [9].

While a popular prevailing belief is that Bitcoin is anonymous [10], its privacy properties are insufficient for many commercial use cases. Every transaction is published in a global ledger, which allows small amounts of information about users’ financial activity (e.g., the identities of the participants in a single transaction [11]) to be amplified by statistical analysis [12]. This limits the commercial usefulness of the network and also harms individual privacy [13], as user behavior frequently reflects the pervasive assumption that Bitcoin is an anonymous system. Further, it can damage the

¹The capitalized “Bitcoin” is used to talk about the technology and the engine, while the lowercase “bitcoin” is used to refer to the currency.

²In most traditional systems, the speed of transaction is achieved by instant execution and delayed settlement.

FAST RADIO BURSTS FROM EXTRAGALACTIC LIGHT SAILS

MANASVI LINGAM^{1,2} AND ABRAHAM LOEB²

¹*Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

²*Harvard-Smithsonian Center for Astrophysics, The Institute for Theory and Computation, 60 Garden Street, Cambridge, MA 02138, USA*

ABSTRACT

We examine the possibility that Fast Radio Bursts (FRBs) originate from the activity of extragalactic civilizations. Our analysis shows that beams used for powering large light sails could yield parameters that are consistent with FRBs. The characteristic diameter of the beam emitter is estimated through a combination of energetic and engineering constraints, and both approaches intriguingly yield a similar result which is on the scale of a large rocky planet. Moreover, the optimal frequency for powering the light sail is shown to be similar to the detected FRB frequencies. These ‘coincidences’ lend some credence to the possibility that FRBs might be artificial in origin. Other relevant quantities, such as the characteristic mass of the light sail, and the angular velocity of the beam, are also derived. By using the FRB occurrence rate, we infer upper bounds on the rate of FRBs from extragalactic civilizations in a typical galaxy. The possibility of detecting fainter signals is briefly discussed, and the wait time for an exceptionally bright FRB event in the Milky Way is estimated.

arXiv:1701.01109v2 [astro-ph.HE] 27 Feb 2017

INCREMENTAL NETWORK QUANTIZATION: TOWARDS LOSSLESS CNNs WITH LOW-PRECISION WEIGHTS

Aojun Zhou*, Anbang Yao, Yiwen Guo, Lin Xu, and Yurong Chen

Intel Labs China

{aojun.zhou, anbang.yao, yiwen.guo, lin.x.xu, yurong.chen}@intel.com

ABSTRACT

This paper presents incremental network quantization (INQ), a novel method, targeting to efficiently convert any pre-trained full-precision convolutional neural network (CNN) model into a low-precision version whose weights are constrained to be either powers of two or zero. Unlike existing methods which are struggled in noticeable accuracy loss, our INQ has the potential to resolve this issue, as benefiting from two innovations. On one hand, we introduce three interdependent operations, namely weight partition, group-wise quantization and re-training. A well-proven measure is employed to divide the weights in each layer of a pre-trained CNN model into two disjoint groups. The weights in the first group are responsible to form a low-precision base, thus they are quantized by a variable-length encoding method. The weights in the other group are responsible to compensate for the accuracy loss from the quantization, thus they are the ones to be re-trained. On the other hand, these three operations are repeated on the latest re-trained group in an iterative manner until all the weights are converted into low-precision ones, acting as an incremental network quantization and accuracy enhancement procedure. Extensive experiments on the ImageNet classification task using almost all known deep CNN architectures including AlexNet, VGG-16, GoogleNet and ResNets well testify the efficacy of the proposed method. Specifically, at 5-bit quantization (a variable-length encoding: 1 bit for representing zero value, and the remaining 4 bits represent at most 16 different values for the powers of two)¹, our models have improved accuracy than the 32-bit floating-point references. Taking ResNet-18 as an example, we further show that our quantized models with 4-bit, 3-bit and 2-bit ternary weights have improved or very similar accuracy against its 32-bit floating-point baseline. Besides, impressive results with the combination of network pruning and INQ are also reported. We believe that our method sheds new insights on how to make deep CNNs to be applicable on mobile or embedded devices. The code is available at <https://github.com/Zhouaojun/Incremental-Network-Quantization>.

1 INTRODUCTION

Deep convolutional neural networks (CNNs) have demonstrated record breaking results on a variety of computer vision tasks such as image classification (Krizhevsky et al., 2012; Simonyan & Zisserman, 2015), face recognition (Taigman et al., 2014; Sun et al., 2014), semantic segmentation (Long et al., 2015; Chen et al., 2015a) and object detection (Girshick, 2015; Ren et al., 2015). Regardless of the availability of significantly improved training resources such as abundant annotated data, powerful computational platforms and diverse training frameworks, the promising results of deep CNNs are mainly attributed to the large number of learnable parameters, ranging from tens of millions to even hundreds of millions. Recent progress further shows clear evidence that CNNs could easily enjoy the accuracy gain from the increased network depth and width (He et al., 2016; Szegedy et al., 2015; 2016). However, this in turn lays heavy burdens on the memory and

*This work was done when Aojun Zhou was an intern at Intel Labs China, supervised by Anbang Yao who proposed the original idea and is responsible for correspondence. The first three authors contributed equally to the writing of the paper.

¹This notation applies to our method throughout the paper.

Protein bioelectronics: a review of what we do and do not know

Christopher D. Bostick,^{#1,2} Sabyasachi Mukhopadhyay,^{#3‡} Israel
Pecht,^{3*} Mordechai Sheves,^{3*} David Cahen,^{3*} David Lederman^{*4}

¹ *Dept. of Pharmaceutical Sciences, West Virginia University, Morgantown, WV 26506, USA*

² *Inst. for Genomic Medicine, Columbia University Medical Center, New York, NY 10032, USA*

³ *Departments of Materials & Interfaces, of Organic Chemistry and of Immunology,
Weizmann Institute of Science, Rehovot, Israel 76100*

⁴ *Department of Physics, University of California, Santa Cruz, CA 95060, USA*

[#] These authors contributed equally to the work

^{*} Corresponding authors, email addresses: david.cahen@weizmann.ac.il,
dlederma@ucsc.edu, mudi.sheves@weizmann.ac.il, israel.pecht@weizmann.ac.il

Abstract

We review the status of protein-based molecular electronics. First, we define and discuss fundamental concepts of electron transfer and transport in and across proteins and proposed mechanisms for these processes. We then describe the immobilization of proteins to solid-state surfaces in both nanoscale and macroscopic approaches, and highlight how different methodologies can alter protein electronic properties. Because immobilizing proteins while retaining biological activity is crucial to the successful development of bioelectronic devices, we discuss this process at length. We briefly discuss computational predictions and their connection to experimental results. We then summarize how the biological activity of immobilized proteins is beneficial for bioelectronic devices, and how conductance measurements can shed light on protein properties. Finally, we consider how the research to date could influence the development of future bioelectronic devices.

[‡] Current affiliation - Department of Physics, School of Engineering & Applied Sciences, SRM University-AP, Amaravati, Andhra Pradesh, India - 522502

Sprites: Payment Channels that Go Faster than Lightning

Andrew Miller
UIUC

Iddo Bentov
Cornell University

Ranjit Kumaresan
Microsoft Research

Patrick McCorry
Newcastle University

Abstract

It is well known that Bitcoin, Ethereum, and other blockchain-based cryptocurrencies are facing hurdles in scaling to meet user demand. One of the most promising approaches is to form a network of “off-chain payment channels,” which are backed by on-chain currency but support rapid, optimistic transactions and use the blockchain only in case of disputes.

We develop a novel construction for payment channels that reduces the worst-case “collateral cost” for off-chain payments. In existing proposals, particularly the Lightning Network, a payment across a path of ℓ channels requires locking up collateral for $O(\ell\Delta)$ time, where Δ is the time to commit a on-chain transaction. Our construction reduces this cost to $O(\ell + \Delta)$. We formalize our construction in the simulation-based security model, and provide an implementation as an Ethereum smart contract. Our construction relies on a general purpose primitive called a “state channel,” which is of independent interest.

1 Introduction

Cryptocurrencies such as Bitcoin, Ethereum, and others, derive their security from wide replication, which unfortunately comes at the expense of limited scalability. A leading proposal for improved scaling of cryptocurrencies is to form a network of “off-chain” rapid payment channels, which act like credit lines secured by “on-chain” currency. In this vision of the future, cryptocurrencies will largely be used as collateral, so interaction with the blockchain directly will rarely be needed.

A chief concern for the feasibility of payment networks is whether the “collateral costs” will be prohibitive. In general, collateral cost is the lost opportunity (in dimensions of money \times time) occurring when funds are held in escrow instead of being invested profitably. Currency deposited in a payment channel can earn fees when it is used to facilitate linked payments. However, each in-flight payment along a channel must reserve a portion of that channel’s available collateral, preventing its use elsewhere until the payment is settled. In the optimistic case, payments complete quickly, requiring only off-chain point-to-point messages; but if some party fails, the collateral can be tied up for a significant duration, until the balance can be settled on-chain. The more hops in a payment path, the more collateral must

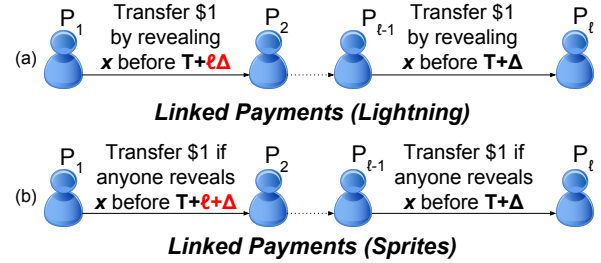


Figure 1: Cryptocurrencies like Bitcoin and Ethereum can serve as collateral for a scalable payment network (i.e. a credit network without counterparty risk) [26, 8]. Payment channels (a) let one party rapidly pay another using available collateral, requiring a blockchain transaction only in case of dispute. Payments across multiple channels (b) can be linked using a common condition (such as revealing the preimage of a hash, h). We contribute a novel payment channel (c) improving the worst case delay for ℓ -hop payments from $O(\ell\Delta)$ to $O(\ell + \Delta)$.

be reserved: $O(\ell\$X)$ in total for a $O(X)$ sized payment spanning ℓ channels. Furthermore, due to limitations of the current state-of-the-art payment channels, each link in the path adds an additional worst-case delay. This additional delay is determined by the worst-case confirmation time for an on-chain transaction, which we denote by Δ (i.e., an on-chain transaction may take Δ times longer than an ordinary off-chain message). Thus the worst-case delay is $O(\ell\Delta)$, and so the total collateral cost of a $\$X$ payment over a path of length ℓ is $O(\ell^2\$X\Delta)$.

In this paper we present an improved construction of payment channels, called “Sprites,” that reduce the timeout delay from $O(\ell\Delta)$ to $O(\ell + \Delta)$, resulting in a total collateral cost of $O(\ell\$X\Delta)$ for a payment over ℓ hops. Our solution makes use of a feature available today in Ethereum smart contracts, but that cannot (we conjecture) be implemented in Bitcoin — in particular, the ability for a transaction to depend on a “global” event recorded in the blockchain.

Our construction is highly modular; a key contribution of our work is the development of a useful general primitive called a “state channel,” which allows two or more parties to maintain an arbitrary off-chain shared process, which can be synchronized on demand (or in case of a dispute) with the blockchain. This abstraction neatly encapsulates the underlying cryptography; by making use of it, our payment channel constructions do not mention

Deep Forest: Towards an Alternative to Deep Neural Networks*

Zhi-Hua Zhou and Ji Feng

National Key Lab for Novel Software Technology, Nanjing University, Nanjing 210023, China
{zhouzh, fengj}@lamda.nju.edu.cn

Abstract

In this paper, we propose gcForest, a decision tree ensemble approach with performance highly competitive to deep neural networks in a broad range of tasks. In contrast to deep neural networks which require great effort in hyper-parameter tuning, gcForest is much easier to train; even when it is applied to different data across different domains in our experiments, excellent performance can be achieved by almost same settings of hyper-parameters. The training process of gcForest is efficient, and users can control training cost according to computational resource available. The efficiency may be further enhanced because gcForest is naturally apt to parallel implementation. Furthermore, in contrast to deep neural networks which require large-scale training data, gcForest can work well even when there are only small-scale training data.

1 Introduction

In recent years, deep neural networks have achieved great success in various applications, particularly in tasks involving visual and speech information [Krizhevsky *et al.*, 2012; Hinton *et al.*, 2012], leading to the hot wave of deep learning [Goodfellow *et al.*, 2016].

Though deep neural networks are powerful, they have apparent deficiencies. First, it is well known that a huge amount of training data are usually required for training, disabling deep neural networks to be directly applied to tasks with small-scale data. Note that even in the big data era, many real tasks still lack sufficient amount of *labeled* data due to high cost of labeling, leading to inferior performance of deep neural networks in those tasks. Second, deep neural networks are very complicated models and powerful computational facilities are usually required for the training process, encumbering individuals outside big companies to fully exploit the learning ability. More importantly, deep neural networks are with too many hyper-parameters, and the learning performance depends seriously on careful tuning of them. For ex-

ample, even when several authors all use convolutional neural networks [LeCun *et al.*, 1998; Krizhevsky *et al.*, 2012; Simonyan and Zisserman, 2014], they are actually using different learning models due to the many different options such as the convolutional layer structures. This fact makes not only the training of deep neural networks very tricky, like an art rather than science/engineering, but also theoretical analysis of deep neural networks extremely difficult because of too many interfering factors with almost infinite configurational combinations.

It is widely recognized that the *representation learning* ability is crucial for deep neural networks. It is also noteworthy that, to exploit large training data, the capacity of learning models should be large; this partially explains why the deep neural networks are very complicated, much more complex than ordinary learning models such as support vector machines. We conjecture that if we can endow these properties to some other suitable forms of learning models, we may be able to achieve performance competitive to deep neural networks but with less aforementioned deficiencies.

In this paper, we propose gcForest (multi-Grained Cascade Forest), a novel decision tree ensemble method. This method generates a deep forest ensemble, with a cascade structure which enables gcForest to do representation learning. Its representational learning ability can be further enhanced by multi-grained scanning when the inputs are with high dimensionality, potentially enabling gcForest to be contextual or structural aware. The number of cascade levels can be adaptively determined such that the model complexity can be automatically set, enabling gcForest to perform excellently even on small-scale data. Moreover, users can control training costs according to computational resources available. The gcForest has much fewer hyper-parameters than deep neural networks; even better news is that its performance is quite robust to hyper-parameter settings, such that in most cases, even across different data from different domains, it is able to get excellent performance by using the default setting. This makes not only the training of gcForest convenient, but also theoretical analysis, although beyond the scope of this paper, potentially easier than deep neural networks (needless to say that tree learners are typically easier to analyze than neural networks). In our experiments, gcForest achieves highly competitive performance to deep neural networks, whereas the training time cost of gcForest is smaller than that of deep

*This research was supported by NSFC (61333014), 973 Program (2014CB340501) and the Collaborative Innovation Center of Novel Software Technology and Industrialization.

Draft version 0.2, 2017-08-11.
This paper has not been peer
reviewed. Please do not copy or
cite without author's permission.

A Computational Model of Systems Memory Reconsolidation and Extinction

Peter Helfer and Thomas R. Shultz

McGill University

Enhancement of human color vision by breaking the binocular redundancy

Bradley S. Gundlach¹, Alireza Shahsafi¹, Gregory Vershbow², Chenghao Wan^{1,3}, Jad Salman¹, Bas Rokers^{4,5}, Laurent Lessard¹, Mikhail A. Kats^{1,3,5*}

¹Department of Electrical and Computer Engineering, University of Wisconsin-Madison

²Department of Art, University of Wisconsin-Madison

³Department of Materials Science and Engineering, University of Wisconsin-Madison

⁴Department of Psychology, University of Wisconsin-Madison

⁵McPherson Eye Research Institute, University of Wisconsin-Madison

*Corresponding Author: Mikhail A. Kats

Address: 1415 Engineering Dr, Madison, WI 53706

Email: mkats@wisc.edu

Phone: (608)890-3984

ABSTRACT

To see color, the human visual system combines the responses of three types of cone cells in the retina – a process that discards a significant amount of spectral information. We present an approach that can enhance human color vision by breaking the inherent redundancy in binocular vision, providing different spectral content to each eye. Using a psychophysical color model and thin-film optimization, we designed a wearable passive multispectral device that uses two distinct transmission filters, one for each eye, to enhance the user’s ability to perceive spectral information. We fabricated and tested a design that “splits” the response of the short-wavelength cone of individuals with typical trichromatic vision, effectively simulating the presence of four distinct cone types between the two eyes (“tetrachromacy”). Users of this device were able to differentiate metamers (distinct spectra that resolve to the same perceived color in typical observers) without apparent adverse effects to vision. The increase in the number of effective cones from the typical three reduces the number of possible metamers that can be encountered, enhancing the ability to discriminate objects based on their emission, reflection, or transmission spectra. This technique represents a significant enhancement of the spectral perception of typical humans, and may have applications ranging from camouflage detection and anti-counterfeiting to art and data visualization.

Experimental certification of millions of genuinely entangled atoms in a solid

Florian Fröwis,^{*} Peter C. Strassmann,[†] Alexey Tiranov,[†] Corentin Gut,[‡] Jonathan Lavoie,[§] Nicolas Brunner, Félix Bussi eres, Mikael Afzelius, and Nicolas Gisin
Groupe de Physique Appliqu ee, Universit  de Gen ve, CH-1211 Gen ve, Switzerland
 (Dated: March 16, 2017)

Quantum theory predicts that entanglement can also persist in macroscopic physical systems, albeit difficulties to demonstrate it experimentally remain. Recently, significant progress has been achieved and genuine entanglement between up to 2900 atoms was reported. Here we demonstrate 16 million genuinely entangled atoms in a solid-state quantum memory prepared by the heralded absorption of a single photon. We develop an entanglement witness for quantifying the number of genuinely entangled particles based on the collective effect of directed emission combined with the nonclassical nature of the emitted light. The method is applicable to a wide range of physical systems and is effective even in situations with significant losses. Our results clarify the role of multipartite entanglement in ensemble-based quantum memories as a necessary prerequisite to achieve a high single-photon process fidelity crucial for future quantum networks. On a more fundamental level, our results reveal the robustness of certain classes of multipartite entangled states, contrary to, e.g., Schr dinger-cat states, and that the depth of entanglement can be experimentally certified at unprecedented scales.

A clear picture of large-scale entanglement with its complex structure is so far not developed. It is however important to understand the role of different facets of multipartite entanglement in nature and in technical applications [1, 2]. For example, the so-called Schr dinger cat states [3] are fundamentally different from a single photon coherently absorbed by a large atomic ensemble; even though both are instances of multipartite entanglement [4, chapter 16.5]. The theoretical study of large-scale entanglement has to be followed by an experimental demonstration, which consists of two basic steps: the preparation of an entangled system and a subsequent appropriate measurement verifying the presence of entanglement. In the context of entanglement in large systems, the preparation of entanglement is generally much simpler than its verification. For example, single-particle measurements are often not possible and collective measurements are typically restricted to certain types and are of finite resolution. These limitations call for new witnesses that allow one to certify entanglement based on accessible measurement data.

The concept of entanglement depth [5] was shown to be meaningful for and applicable to large quantum systems. It is defined as the smallest number of genuinely entangled particles that is compatible with the measured data. This allows one to witness at least one subgroup of genuinely entangled particles in a state-independent and scalable way. Large entanglement depth was successfully demonstrated with so-called spin-squeezed and oversqueezed states by measuring first and second moments of collective spin operators [6–9]; lately up of 680 atoms [10]. Recently, a witness was proposed that is designed for the W state, which is a coherent superposition of a single excitation shared by many atoms [11]. Based on this witness, an entanglement depth of around 2900 was measured [12]. However, these witnesses do not de-

tect entanglement when the vacuum component of the state is dominant [11], even though the W state is known to be quite robust against various sources of noise, in particular, against loss of particles and excitation [13]. Hence, much larger values for the entanglement depth could be expected.

In this paper, we present theoretical methods and experimental data that verify a large entanglement depth in a solid-state quantum memory. A rare-earth-ion-doped crystal spectrally shaped to an atomic frequency comb (AFC) is used to absorb and re-emit light at the single-photon level [14–17], where at least 40 billion atoms collectively interact with the optical field. Using the measured photon number statistics of the re-emitted light we collect partial information about the quantum state of the atomic ensemble before emission. Then, we show that certain combinations of emission probabilities for one and two photons imply entanglement between a large number of atoms. With the measured data from our solid-state quantum memory we demonstrate inseparable groups of entangled particles containing at least 16 million atoms.

RESULTS

Before discussing the experiment, we give an intuitive explanation for the appearance of large entanglement depth when a large atomic ensemble coherently interacts with a single photon (see Fig. 1(a)). Suppose that N two-level atoms ($|g\rangle$ and $|e\rangle$ denote ground and excited state, respectively), couple to a light field. The quantised interaction in the dipole approximation is described by [18]

$$H_{\text{int}} = \sum_{j,\vec{k}} e^{-i\vec{k}\cdot\vec{r}_j} a_{\vec{k}} \sigma_+^{(j)} + e^{i\vec{k}\cdot\vec{r}_j} a_{\vec{k}}^\dagger \sigma_-^{(j)}, \quad (1)$$

DEEP LSTM FOR LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION

Xu Tian, Jun Zhang, Zejun Ma, Yi He, Juan Wei, Peihao Wu, Wenchang Situ, Shuai Li, Yang Zhang

Alibaba Shenma Search, Beijing, China

{xu.tian, zj102217, zejun.mamzj, heyi.hy, wj80290, peihao.wph, wenchang.situwc, voolc.li, zy80232}@alibaba-inc.com

ABSTRACT

Recurrent neural networks (RNNs), especially long short-term memory (LSTM) RNNs, are effective network for sequential task like speech recognition. Deeper LSTM models perform well on large vocabulary continuous speech recognition, because of their impressive learning ability. However, it is more difficult to train a deeper network. We introduce a training framework with layer-wise training and exponential moving average methods for deeper LSTM models. It is a competitive framework that LSTM models of more than 7 layers are successfully trained on Shenma voice search data in Mandarin and they outperform the deep LSTM models trained by conventional approach. Moreover, in order for on-line streaming speech recognition applications, the shallow model with low real time factor is distilled from the very deep model. The recognition accuracy have little loss in the distillation process. Therefore, the model trained with the proposed training framework reduces relative 14% character error rate, compared to original model which has the similar real-time capability. Furthermore, the novel transfer learning strategy with segmental Minimum Bayes-Risk is also introduced in the framework. The strategy makes it possible that training with only a small part of dataset could outperform full dataset training from the beginning.

1. INTRODUCTION

Recently, deep neural network has been widely employed in various recognition tasks. Increasing the depth of neural network is a effective way to improve the performance, and convolutional neural network (CNN) has benefited from it in visual recognition task[1]. Deeper long short-term memory (LSTM) recurrent neural networks (RNNs) are also applied in large vocabulary continuous speech recognition (LVCSR) task, because LSTM networks have shown better performance than Fully-connected feed-forward deep neural network[2, 3, 4, 5].

Training neural network becomes more challenge when it goes deep. A conceptual tool called linear classifier probe is introduced to better understand the dynamics inside a neural network [6]. The discriminating features of linear classifier is the hidden units of a intermediate layer. For deep neural networks, it is observed that deeper layer's accuracy is lower

than that of shallower layers. Therefore, the tool shows the difficulty of deep neural model training visually.

Layer-wise pre-training is a successful method to train very deep neural networks [7]. The convergence becomes harder with increasing the number of layers, even though the model is initialized with Xavier or its variants [8, 9]. But the deeper network which is initialized with a shallower trained network could converge well.

The size of LVCSR training dataset goes larger and training with only one GPU becomes high time consumption inevitably. Therefore, parallel training with multi-GPUs is more suitable for LVCSR system. Mini-batch based stochastic gradient descent (SGD) is the most popular method in neural network training procedure. Asynchronous SGD is a successful effort for parallel training based on it [10, 11]. It can many times speed up the training time without decreasing the accuracy. Besides, synchronous SGD is another effective effort, where the parameter server waits for every works to finish their computation and sent their local models to it, and then it sends updated model back to all workers [12]. Synchronous SGD converges well in parallel training with data parallelism, and is also easy to implement.

In order to further improve the performance of deep neural network with parallel training, several methods are proposed. Model averaging method achieves linear speedup, as the final model is averaged from all parameters of local models in different workers [13, 14], but the accuracy decreases compared with single GPU training. Moreover, blockwise model-updating filter (BMUF) provides another almost linear speedup approach with multi-GPUs on the basis of model averaging. It can achieve improvement or no-degradation of recognition performance compared with mini-batch SGD on single GPU [15].

Moving averaged (MA) approaches are also proposed for parallel training. It is demonstrated that the moving average of the parameters obtained by SGD performs as well as the parameters that minimize the empirical cost, and moving average parameters can be used as the estimator of them, if the size of training data is large enough [16]. One pass learning is then proposed, which is the combination of learning rate schedule and averaged SGD using moving average [17]. Exponential moving average (EMA) is proposed as a non-interference method[18]. EMA model is not broadcasted to

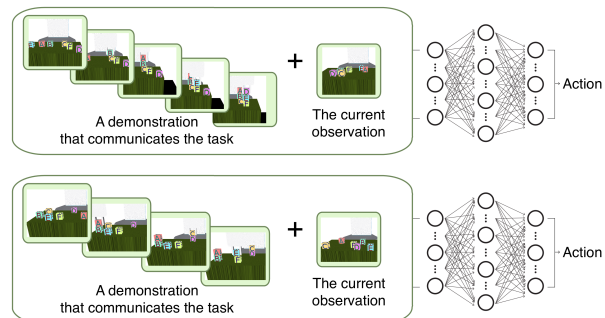
One-Shot Imitation Learning

Yan Duan^{1,2} Marcin Andrychowicz¹ Bradley C. Stadie^{1,2} Jonathan Ho^{1,2} Jonas Schneider¹ Ilya Sutskever¹
Pieter Abbeel^{1,2} Wojciech Zaremba¹

Abstract

Imitation learning has been commonly applied to solve different tasks in isolation. This usually requires either careful feature engineering, or a significant number of samples. This is far from what we desire: ideally, robots should be able to learn from very few demonstrations of any given task, and instantly generalize to new situations of the same task, without requiring task-specific engineering. In this paper, we propose a meta-learning framework for achieving such capability, which we call *one-shot imitation learning*.

Specifically, we consider the setting where there is a very large (maybe infinite) set of tasks, and each task has many instantiations. For example, a task could be to stack all blocks on a table into a single tower, another task could be to place all blocks on a table into two-block towers, etc. In each case, different instances of the task would consist of different sets of blocks with different initial states. At training time, our algorithm is presented with pairs of demonstrations for a subset of all tasks. A neural net is trained that takes as input one demonstration and the current state (which initially is the initial state of the other demonstration of the pair), and outputs an action with the goal that the resulting sequence of states and actions matches as closely as possible with the second demonstration. At test time, a demonstration of a single instance of a new task is presented, and the neural net is expected to perform well on new instances of this new task. Our experiments show that the use of soft attention allows the model to generalize to conditions and tasks unseen in the training data. We anticipate that by training this model on a much greater variety of tasks and settings, we will obtain a general system that can turn any demonstrations into robust policies that can accomplish an overwhelming variety of tasks.



One-shot policy. A single policy trained to solve many tasks.



(left) Task-specific policy. This policy is trained to stack blocks into two towers, each of height 3. (right) A separate task-specific policy. This policy is trained to stack blocks into three towers, each of height 2.

Figure 1. Traditionally, policies are task-specific. For example, a policy might have been trained (through imitation or reinforcement learning) to stack blocks into towers of height 3, and then another policy would be trained to stack blocks into towers of height 2, etc. In this paper, we are interested in policies that are *not* specific to one task, but rather can be told (through a single demonstration) what the current new task is, and be successful at this new task. As illustrative examples, we would want to be able to provide a single demonstration of each task, and from that the one-shot policy would know what to do when faced with a new situation of the task, where the blocks are randomly rearranged. Videos of the illustrated tasks are available at <http://bit.ly/one-shot-imitation>.

1. Introduction

We are interested in robotic systems that are able to perform a variety of complex useful tasks, e.g. tidying up a home or preparing a meal. The robot should be able to learn new tasks without long system interaction time. To accomplish this, we must solve two broad problems:

- The first problem is that of dexterity: robots should learn how to approach, grasp and pick up complex un-

¹OpenAI ²University of California, Berkeley. Correspondence to: Yan Duan <rocky@openai.com>. Copyright 2017 by the author(s).

Deep brain fluorescence imaging with minimally invasive ultra-thin optical fibers

Shay Ohayon^(1,2), Antonio M. Caravaca-Aguirre⁽³⁾, Rafael Piestun⁽³⁾, James J. DiCarlo^(1,2)

- (1) McGovern Institute for Brain Research, Massachusetts Institute of Technology, 43 Vassar Street Cambridge, MA 02139
- (2) Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139
- (3) Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, Colorado, 80309, USA

Abstract

A major open challenge in neuroscience is the ability to measure and perturb neural activity in vivo from well-defined neural sub-populations at cellular resolution anywhere in the brain. However, limitations posed by scattering and absorption prohibit non-invasive (surface) multiphoton approaches^{1,2} for deep (>2mm) structures, while Gradient Refractive Index (GRIN) endoscopes²⁻⁴ are thick and cause significant damage upon insertion. Here, we demonstrate a novel microendoscope to image neural activity at arbitrary depths via an ultrathin multimode optical fiber (MMF) probe that is 5-10X thinner than commercially available microendoscopes. We demonstrate micron-scale resolution, multispectral and volumetric imaging. In contrast to previous approaches^{1,5-8} we show that this method has an improved acquisition speed that is sufficient to capture rapid neuronal dynamics in-vivo in rodents expressing a genetically encoded calcium indicator. Our results emphasize the potential of this technology in neuroscience applications and open up possibilities for cellular resolution imaging in previously unreachable brain regions.

1. Introduction

The main potential advantage of MMF for biological endoscopy is their thin diameter (50-150 μm) which induces less damage compared to thicker probes, such as GRIN lens^{4,9-11}. Fiber bending remains a major challenge to MMF imaging techniques that are based on Wavefront shaping (WFS) since fiber deformation changes how modes are coupled and invalidates precomputed transformations critical for image formation¹²⁻¹⁴. However, in some domains, such as neuroscience, major fiber bending may not always pose a direct problem since in many experimental designs fibers can be inserted into the brain along a straight trajectory to target specific brain regions. Furthermore, minor perturbations at the distal end of the fiber have only minor effects on the imaging capability if a proper fiber is selected¹⁵. Other challenges, on the other hand, need to be addressed to make this technology useful for neuroscience. First, acquisition speed needs to be sufficiently high to capture rapid neural firing events (ms scale), while maintaining sufficient spatial resolution to resolve cellular level details. Previous approaches^{1,5-8} that use liquid crystal spatial light modulators (LC-SLM) have not been shown to be capable of this temporal resolution which would prohibit rapid sampling of neural signals. Second, the system needs

to have high collection efficiency to capture small fluorescence changes evoked by calcium transients without causing photobleaching. To date, only fixed tissue has been successfully imaged^{5,15}. Third, the system should image some distance away from the fiber tip since tissue in close proximity may not be functioning properly due to possible damage inflicted by fiber insertion.

Here, we focus on addressing these challenges and propose a novel microendoscope system that is based on a digital mirror device (DMD).

2. Methods

2.1 Generating phase modulation with a digital mirror device

The basic idea behind the Lee hologram¹⁶ technique is to create a binary mask that creates a diffraction pattern with the desired phase modulation. Following the work of^{16,17}, to create a 2D spatial phase mask (x,y) , we first define a spatial carrier wave with frequency f and rotation θ over the mirror domain $[x,y]$. The mask is then defined as:

$$(1) \quad X = \cos(\theta) x + \sin(\theta) y$$

$$(2) \quad \text{Mask} = \frac{1}{2} (1 + \cos(2\pi X - \Phi(x, y))) > 0.5$$

This mask generates three peaks on the Fourier plane since

$$(3) \quad \frac{1}{2} (1 + \cos(2\pi X - \Phi(x, y))) = \frac{1}{2} + \frac{1}{4} e^{2\pi j(x-y)f_0} e^{-j\Phi(x,y)} + \frac{1}{4} e^{-2\pi j(x-y)f_0} e^{j\Phi(x,y)}$$

By blocking the first two terms (DC and one diffraction order) with an iris and allowing only one diffraction order to pass through, we create a 2D phase array that is up to a constant shift from the desired $\Phi(x,y)$. In the limit, each phase $\Phi(x,y)$ cannot be represented by a single mirror on the DMD, hence multiple mirrors are grouped to represent a single desired phase. We have experimented with varying sampling (number of mirrors per phase) and found that 10 is a good compromise between the size of each block and sufficient remaining mirrors to deliver the constant reference phase (see below).

2.2 Transmission matrix estimation and spot generation

On the Impossibility of Supersized Machines

Ben Garfinkel , Miles Brundage¹, Daniel Filan^{2,3}, Carrick Flynn³,
Jelena Luketina , Michael Page , Anders Sandberg³, Andrew
Snyder-Beattie³, and Max Tegmark⁴

¹*School for the Future of Innovation in Society, Arizona State University*

²*Department of Computer Science, University of California, Berkeley*

³*Future of Humanity Institute, University of Oxford*

⁴*Department of Physics, Massachusetts Institute of Technology*

April 1, 2017

Abstract

In recent years, a number of prominent computer scientists, along with academics in fields such as philosophy and physics, have lent credence to the notion that machines may one day become as large as humans. Many have further argued that machines could even come to exceed human size by a significant margin. However, there are at least seven distinct arguments that preclude this outcome. We show that it is not only implausible that machines will ever exceed human size, but in fact impossible.

Introduction

The history of life is often understood as a story of growth. If one takes the long view, then one can trace an exponential curve from our minuscule earliest ancestors, which were little more than self-replicating molecules, to the substantial creatures that we are today (Payne, 2009).

Although humanity became aware of this story only in the 19th century, through the work of Charles Darwin, we have long had the privilege of witnessing a partial recapitulation every time someone new comes into the world (Darwin, 1859). Before each person is a full-sized adult, they are first an invisibly small cell.

It is perhaps no surprise, then, that human largeness has for thousands of years fascinated many of our greatest thinkers. While some have sought to understand the nature and origins of largeness, others have anxiously inquired: *Could there ever be something larger than a human?*

Evidence of this anxiety can be found as far back as humanity's oldest recorded myth, *The Epic of Gilgamesh*, in which the monstrous giant Humbaba is appointed by Enlil, the king of the gods, to terrorize mankind (Sandars,

Associative content-addressable networks with exponentially many robust stable states

Rishidev Chaudhuri¹ and Ila Fiete¹

¹Center for Learning and Memory and Department of Neuroscience,
The University of Texas at Austin, Austin, Texas, USA.

Abstract

The brain must robustly store a large number of memories, corresponding to the many events and scenes a person encounters over a lifetime. However, the number of memory states in existing neural network models either grows weakly with network size or recall performance fails catastrophically with vanishingly little noise. Here we show that it is possible to construct an associative content-addressable memory (ACAM) with *exponentially* many stable states and robust error-correction. The network possesses expander graph connectivity on a restricted Boltzmann machine architecture. The expansion property allows simple neural network dynamics to perform at par with modern error-correcting codes. Appropriate networks can be constructed with sparse random connections combined with glomerular nodes and a local associative learning rule, using low dynamic-range weights. Thus, sparse quasi-random constraint structures – characteristic of an important class of modern error-correcting codes – may provide for high-performance computation in artificial neural networks and the brain.

Introduction

Neural long-term memory systems have high capacity, by which we mean that the number of memory states is large. Such systems are also able to recover the correct memory state from partial or noisy cues, the definition of an associative memory. If the memory state can be addressed by its content, it is furthermore called content-addressable.

Classic studies of the dynamics and capacity of associative content-addressable neural memory (ACAM) have focused on connectionist neural network models commonly called Hopfield networks^{1–3}, which provide a powerful conceptual framework for thinking about pattern completion and associative memory in the brain. Here we continue in this tradition and examine constructions of ACAM networks in the form of Hopfield networks and their stochastic equivalents, Boltzmann machines. We show that it is possible to construct associative content-addressable memory networks with an unprecedented combination of robustness and capacity.

Transit Detection of a “Starshade” at the Inner Lagrange Point of an Exoplanet

E. Gaidos,^{1*}

¹*Department of Geology & Geophysics, University of Hawaii at Mānoa, Honolulu, Hawaii 96822 USA*

Submitted to MNRAS

ABSTRACT

All water-covered rocky planets in the inner habitable zones of solar-type stars will inevitably experience a catastrophic runaway climate due to increasing stellar luminosity and limits to outgoing infrared radiation from wet greenhouse atmospheres. Reflectors or scatterers placed near Earth’s inner Lagrange point (\mathcal{L}_1) have been proposed as a “geo-engineering” solution to anthropogenic climate change and an advanced version of this could modulate incident irradiation over many Gyr or “rescue” a planet from the interior of the habitable zone. The distance of the starshade from the planet that minimizes its mass is 1.6 times the Earth- \mathcal{L}_1 distance. Such a starshade would have to be similar in size to the planet and the mutual occultations during planetary transits could produce a characteristic maximum at mid-transit in the light-curve. Because of a fortuitous ratio of densities, Earth-size planets around G dwarf stars present the best opportunity to detect such an artifact. The signal would be persistent and is potentially detectable by a future space photometry mission to characterize transiting planets. The signal could be distinguished from natural phenomenon, i.e. starspots or cometary dust clouds, by its shape, persistence, and transmission spectrum.

Key words: techniques: photometric – planets and satellites: terrestrial planets – astrobiology – extraterrestrial intelligence –

1 INTRODUCTION

Like every liquid water-covered planet around a solar-mass star, Earth has a serious greenhouse problem. As the Sun converts hydrogen to helium and becomes denser, hotter, and more luminous (Gough 1981) the inexorable increase in incident irradiation on Earth will, all else being equal, cause surface temperatures to rise and the atmosphere to contain more water vapor. Water is an efficient greenhouse gas, creating a positive feedback in Earth’s climate system as more water vapor leads to elevated temperatures, and vice versa. Walker, Hays & Kasting (1981) proposed that temperature-dependent aqueous weathering and precipitation of carbonate minerals acts as a negative climate feedback that adjusts atmospheric CO_2 to maintain weathering at a rate that balances volcanic degassing, i.e. surface temperatures permissive of abundant liquid water.

However this planetary “thermostat” has its limits. As irradiance continues to increase, CO_2 will eventually disappear from the atmosphere causing a crisis for land plant life and any indeed any autotrophic life relying on atmospheric

CO_2 as a source of carbon (Caldeira & Kasting 1992). Beyond this point, increasing irradiance cannot be compensated by diminished CO_2 and temperatures rise. Climate models predict an asymptotic (maximum) value for the outgoing infrared radiation of an Earth-like atmosphere as a function of temperature; if the absorbed incident radiation exceeds this value, temperatures will increase until the oceans evaporate (Ingersoll 1969). At that point multiple “runaway” climate feed-backs, not mutually exclusive, can occur: water vapor in the upper atmosphere will be photolyzed and the escape of hydrogen will lead to loss of water. Continued de-gassing of CO_2 from the mantle will not be compensated by aqueous weathering of silicates and formation of carbonates, enhancing the greenhouse effect and causing higher surface temperatures. Eventually, surface temperatures will exceed the stability of carbonate minerals, leading to the breakdown of carbonate rocks in the crust, the release of ≈ 90 bars of CO_2 into the atmosphere (Tuck 1980), a Venus-like climate, and the complete extinction of the biosphere. Because of stellar luminosity evolution, all Earth-like planets within the “conservative” habitable zone of solar-mass stars (0.95–1.67 AU Kopparapu et al. 2013) will eventually orbit interior to the habitable zone and meet this fate.

* Visiting Scientist, Center for Space and Habitability, University of Bern, Bern, Switzerland. E-mail: gaidos@hawaii.edu (EG)

Improving Nanopore Reads Raw Signal Alignment

Vladimír Boža, Broňa Brejová, and Tomáš Vinař

Faculty of Mathematics, Physics, and Informatics, Comenius University,
Mlynská dolina, 842 48 Bratislava, Slovakia
{boza,brejova,vinar}@fmph.uniba.sk

Abstract. We investigate usage of dynamic time warping (DTW) algorithm for aligning raw signal data from MinION sequencer. DTW is mostly using for fast alignment for selective sequencing to quickly determine whether a read comes from sequence of interest.

We show that standard usage of DTW has low discriminative power mainly due to problem with accurate estimation of scaling parameters. We propose a simple variation of DTW algorithm, which does not suffer from scaling problems and has much higher discriminative power.

1 Introduction

In this paper, we propose improvements to algorithms for aligning raw signals from MinION nanopore sequencer. The MinION device by Oxford Nanopore (Mikheyev and Tin, 2014), weighing only 90 grams, is currently the smallest high-throughput DNA sequencer. Thanks to its low capital costs, small size and the possibility of analyzing the data in real time as they are produced, MinION is very promising for clinical applications, such as monitoring infectious disease outbreaks (Quick et al., 2015, 2016), and characterizing structural variants in cancer (Norris et al., 2016).

One of the MinION advantages is selective sequencing, where we only sequence "interesting" reads by rejecting all other reads after reading the first few hundred bases. The idea was mostly explored in Readuntil tool (Loose et al., 2016).

Hard part of selective sequencing is deciding which reads to reject. Standard algorithm for selective sequencing would be to base call the first few hundreds of events from the read and then align them to the reference sequence. After aligning, we would accept the read if we find a match with the acceptable similarity.

Unfortunately, the speed of the base calling algorithm on a reasonable computer is much lower than the rate at which the sequencer produces the data. This means that we need to use other approaches than base calling followed by alignment. We will explore an approach which does not translate electric signal into DNA, but instead it works directly with the electric signal data.

Loose et al. (2016) solved this problem using dynamic time warping algorithm (DTW), but he only align reads to small target sequences with size up to hundred kilobases.

Morphological Error Detection in 3D Segmentations

David Rolnick^{1*}, Yaron Meirovitch¹, Toufiq Parag², Hanspeter Pfister²

Viren Jain³, Jeff W. Lichtman², Edward S. Boyden¹, Nir Shavit¹

¹Massachusetts Institute of Technology, ²Harvard University, ³Google, Inc.

Abstract

Deep learning algorithms for connectomics rely upon localized classification, rather than overall morphology. This leads to a high incidence of erroneously merged objects. Humans, by contrast, can easily detect such errors by acquiring intuition for the correct morphology of objects. Biological neurons have complicated and variable shapes, which are challenging to learn, and merge errors take a multitude of different forms. We present an algorithm, MergeNet, that shows 3D ConvNets can, in fact, detect merge errors from high-level neuronal morphology. MergeNet follows unsupervised training and operates across datasets. We demonstrate the performance of MergeNet both on a variety of connectomics data and on a dataset created from merged MNIST images.

1 Introduction

The neural network of the brain remains a mystery, even as engineers have succeeded in building artificial neural networks that can solve a wide variety of problems. Understanding the brain at a deeper level could significantly impact both biology and artificial intelligence [3, 8, 19, 23, 35]. Perhaps appropriately, artificial neural networks are now being used to map biological neural networks. However, humans still outperform computer vision algorithms in segmenting brain tissue. Deep learning has not yet attained the intuition that allows humans to recognize and trace the fine, intermingled branches of neurons.

The field of *connectomics* aims to reconstruct three-dimensional networks of biological neurons from high-resolution microscope images. Automated segmentation is a necessity due to the quantities of data involved. In one recent study [9], the brain of a larval zebrafish was annotated by hand, requiring more than a year of human labor. It is estimated that mapping a single human brain would require a zettabyte (one billion terabytes) of image data [17], clearly more than can be manually segmented.

State-of-the-art algorithms apply a convolutional neural network (ConvNet) to predict, for each voxel of an image, whether it is on the boundary (*cell membrane*) of a neuron. The predicted membranes are then filled in by subsequent algorithms [10]. Such methods are prone both to *split errors*, in which true objects are subdivided, and to *merge errors*, in which objects are fused together. The latter pose a particular challenge. Neurons are highly variable, unpredictably sprouting thousands of branches, so their correct shapes cannot be catalogued. Erroneously merged neurons are obvious to trained humans because they simply don't look right, but it has hitherto been impossible to make such determinations automatically.

*Correspondence should be addressed to: drolnick@mit.edu.

Genuine Multipartite Entanglement in the 3-Photon Decay of Positronium

Beatrix C. Hiesmayr¹ and Pawel Moskal²

¹*Faculty of Physics, University of Vienna, Boltzmannngasse 5, 1090 Vienna, Austria*

²*Institute of Physics, Jagiellonian University, Cracow, Poland*

The electron-positron annihilation into two photons is a standard technology in medicine to observe e.g. metabolic processes in human bodies. A new tomograph will provide the possibility to observe not only direct e^+e^- annihilations but also the 3 photons from the decay of ortho-positronium atoms formed in the body. We show in this contribution that the three-photon state with respect to polarisation degrees of freedom depends on the angles between the photons and exhibits various specific entanglement features. In particular genuine multipartite entanglement, a type of entanglement involving all degrees of freedoms, is subsistent if the positronium was in a definite spin eigenstate. Remarkably, when all spin eigenstates are mixed equally, entanglement – and even stronger genuine multipartite entanglement – survives. Due to a “*symmetrization*” process, however, *Dicke*-type of entanglement remains whereas *GHZ*-type of entanglement vanishes. The survival of particular entanglement properties in the mixing scenario may make it possible to extract quantum information in form of distinct entanglement features, e.g., from metabolic processes in human bodies.

I. INTRODUCTION

The detection of the two high energetic photons coming from the annihilation of an electron and a positron is a well-established successful technology to image metabolic processes in living bodies (PET: Positron Emission Tomography). PET application are used in many different fields of medicine, e.g. in oncology, in cardiology, in radiation therapy or in neurology. In recent years, PET instrumentation has undergone a steady multifaceted evolution and the improvements include new hardware, new reconstruction methods and implementation of time-of-flight techniques [1–7].

With no doubt PET serves as an important tool in imaging metabolic processes based on the sensitivity to tracers (positron-emitting radionuclides) injected into the body or tissue.

Electron-positron annihilations may occur either directly or via the creation of positronium atoms (a bound state of electron and positron). Positronium [8–11] can be in an anti-symmetric spin state (para-positronium) or a symmetric spin state (ortho-positronium). Charge conjugation implies that in the first case it decays into an even number of photons (2γ , 4γ , ...) and in the other case into an odd number of photons (3γ , 5γ , ...). Due to kinematics and smallness of the fine-structure constant the 2γ and 3γ cases are the two most likely options. Since positronium atoms are formed copiously inside the human body during routine PET imaging 3γ -decays occur also frequently. Even in water the production of ortho-positronium has a probability equal to about 25% [12] and is expected to be more than 38% in a tissue [13]. Three-photon events, however, have never been used in tomography because of technical limitations of standard PET devices. A new prototype, called J-PET (Jagiellonian-PET) [14–19], has shown to meet all technical requirements in performing such a measurement for the first time.

This paper investigates the entanglement in the polar-

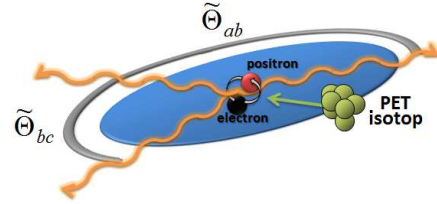


FIG. 1: (Color online) This graphic shows schematically how from an isotope typically used in standard PET-therapy, e.g. FDG-18 (fludeoxyglucose), positronium is generated that decays into three photons which wave vectors have to lay in one plane due to energy and momentum conservation.

isation degrees of freedom of the three photons resulting from the decay of the ortho-positronium. Both for a fixed spin quantization direction of the positronium as well as the case of equal mixing. Photons are fascinating quantum systems, having spin one, but due to their massless property there is a nontrivial coupling between the spin and momentum properties. The most appropriate single-photon description remains controversial. A recent framework describing all single-photon states and single-photon observables by POVMs (positive-operator valued measurements) can be found in Ref. [20]. In this contribution we restrict ourselves to the polarisation degrees of freedom and are interested in the correlation of three photons with energies that ranges from 0 to 511keV. Entanglement and in particular multipartite entanglement is a highly investigated field that has the potential to become a new technology. This paper makes a step towards investigating what type of entanglement is present in the three-photon state generated by the decay of ortho-positronium. This may one day result in obtaining not only the local information where in a tissue the positronium decays, but as well revealing the quantum information which may serve as a new quantum marker for specific biological processes.

Note that entanglement seems to play an important

Quantum Resource Estimates for Computing Elliptic Curve Discrete Logarithms

Martin Roetteler, Michael Naehrig, Krysta M. Svore, and Kristin Lauter

Microsoft Research, USA

Abstract. We give precise quantum resource estimates for Shor’s algorithm to compute discrete logarithms on elliptic curves over prime fields. The estimates are derived from a simulation of a Toffoli gate network for controlled elliptic curve point addition, implemented within the framework of the quantum computing software tool suite LIQUi|. We determine circuit implementations for reversible modular arithmetic, including modular addition, multiplication and inversion, as well as reversible elliptic curve point addition. We conclude that elliptic curve discrete logarithms on an elliptic curve defined over an n -bit prime field can be computed on a quantum computer with at most $9n + 2\lceil\log_2(n)\rceil + 10$ qubits using a quantum circuit of at most $448n^3 \log_2(n) + 4090n^3$ Toffoli gates. We are able to classically simulate the Toffoli networks corresponding to the controlled elliptic curve point addition as the core piece of Shor’s algorithm for the NIST standard curves P-192, P-224, P-256, P-384 and P-521. Our approach allows gate-level comparisons to recent resource estimates for Shor’s factoring algorithm. The results also support estimates given earlier by Proos and Zalka and indicate that, for current parameters at comparable classical security levels, the number of qubits required to tackle elliptic curves is less than for attacking RSA, suggesting that indeed ECC is an easier target than RSA.

Keywords: Quantum cryptanalysis, elliptic curve cryptography, elliptic curve discrete logarithm problem.

1 Introduction

Elliptic curve cryptography (ECC). Elliptic curves are a fundamental building block of today’s cryptographic landscape. Thirty years after their introduction to cryptography [31,27], they are used to instantiate public key mechanisms such as key exchange [11] and digital signatures [17,23] that are widely deployed in various cryptographic systems. Elliptic curves are used in applications such as transport layer security [10,5], secure shell [46], the Bitcoin digital currency system [33], in national ID cards [22], the Tor anonymity network [12], and the WhatsApp messaging app [52], just to name a few. Hence, they play a significant role in securing our data and communications.

Different standards (e.g., [8,49]) and standardization efforts (e.g., [13,35]) have identified elliptic curves of different sizes targeting different levels of security. Notable curves with widespread use are the NIST curves P-256, P-384, P-521, which are curves in Weierstrass form over special primes of size 256, 384, and 521 bits respectively, the Bitcoin curve `secp256k1` from the SEC2 [8] standard and the Brainpool curves [13]. More recently, Bernstein’s Curve25519 [53], a Montgomery curve over a 255-bit prime field, has seen more and more deployment, and it has been recommended to be used in the next version of the TLS protocol [29] along with another even more recent curve proposed by Hamburg called Goldilocks [20].

The security of elliptic curve cryptography relies on the hardness of computing discrete logarithms in elliptic curve groups, i.e. the difficulty of the Elliptic Curve Discrete Logarithm Problem (ECDLP). Elliptic curves have the advantage of relatively small parameter and key sizes in comparison to other cryptographic schemes, such as those based on RSA [40] or finite field discrete logarithms [11], when compared at the same security level. For example, according to NIST recommendations from 2016, a 256-bit elliptic curve provides a similar resistance against classical

Learning Macromanagement in StarCraft from Replays using Deep Learning

Niels Justesen
IT University of Copenhagen
Copenhagen, Denmark
noju@itu.dk

Sebastian Risi
IT University of Copenhagen
Copenhagen, Denmark
sebr@itu.dk

Abstract—The real-time strategy game StarCraft has proven to be a challenging environment for artificial intelligence techniques, and as a result, current state-of-the-art solutions consist of numerous hand-crafted modules. In this paper, we show how macromanagement decisions in StarCraft can be learned directly from game replays using deep learning. Neural networks are trained on 789,571 state-action pairs extracted from 2,005 replays of highly skilled players, achieving top-1 and top-3 error rates of 54.6% and 22.9% in predicting the next build action. By integrating the trained network into UAlbertaBot, an open source StarCraft bot, the system can significantly outperform the game’s built-in Terran bot, and play competitively against UAlbertaBot with a fixed rush strategy. To our knowledge, this is the first time macromanagement tasks are learned directly from replays in StarCraft. While the best hand-crafted strategies are still the state-of-the-art, the deep network approach is able to express a wide range of different strategies and thus improving the network’s performance further with deep reinforcement learning is an immediately promising avenue for future research. Ultimately this approach could lead to strong StarCraft bots that are less reliant on hard-coded strategies.

I. INTRODUCTION

Artificial neural networks have been a promising tool in machine learning for many tasks. In the last decade, the increase in computational resources as well as several algorithmic improvements, have allowed deep neural networks with many layers to be trained on large datasets. This approach, also re-branded as *deep learning*, has remarkably pushed the limits within object recognition [13], speech recognition [8], and many other domains. Combined with reinforcement learning, these techniques have surpassed the previous state-of-the-art in playing Atari games [16], the classic board game Go [23] and the 3D first-person shooter Doom [15].

An open challenge for these methods are real-time strategy (RTS) games such as StarCraft, which are highly complex on many levels because of their enormous state and actions space with a large number of units that must be controlled in real-time. Furthermore, in contrast to games like Go, AI algorithms in StarCraft must deal with hidden information; the opponent’s base is initially hidden and must be explored continuously throughout the game to know (or guess) what strategy the opponent is following. The game has been a popular environment for game AI researchers with several StarCraft AI competitions such as the *AIIDE StarCraft AI*

*Competition*¹, *CIG StarCraft RTS AI Competition*² and the *Student StarCraft AI Competition*³.

However, bots participating in these competitions rely mainly on hard-coded strategies [6, 20] and are rarely able to adapt to the opponent during the game. They usually have a modular control architecture that divides the game into smaller task areas, relying heavily on hand-crafted modules and developer domain knowledge. Learning to play the entire game with end-to-end deep learning, as it was done for Atari games [16], is currently an unsolved challenge and perhaps an infeasible approach. A simpler approach, which we follow in this paper, is to apply deep learning to replace a specific function in a larger AI architecture.

More specifically, we focus on applying deep learning to macromanagement tasks in *StarCraft: Brood War* in the context of deciding what to produce next. A neural network is trained to predict these decisions based on a training set extracted from replay files (i.e. game logs) of highly skilled human players. The trained neural network is combined with the existing StarCraft bot UAlbertaBot, and is responsible for deciding what unit, building, technology, or upgrade to produce next, given the current state of the game. While our approach does not achieve state-of-the-art results on its own, it is a promising first step towards self-learning methods for macromanagement in RTS games. Additionally, the approach presented here is not restricted to StarCraft and can be directly applied to other RTS games as well.

II. STARCRAFT

StarCraft is a real-time strategy (RTS) game released by Blizzard in 1998. The same year an expansion set called *StarCraft: Brood War* was released, which became so popular that a professional StarCraft gamer scene emerged. The game is a strategic military combat simulation in a science fiction setting. Each player controls one of three races; Terran, Protoss and Zerg. During the game, they must gather resources to expand their base and produce an army. The winner of a game is the player that manages to destroy the opponent’s base. Figure 1 shows a screenshot from a player’s perspective controlling the Protoss. The screenshot shows numerous workers

¹<http://www.cs.mun.ca/~dchurhill/starcraftaicom/>

²http://cilab.sejong.ac.kr/sc_competition/

³<http://sscaitournament.com/>

Centimeter-Scale Suspended Photonic Crystal Mirrors

João P. Moura,^{1,*} Richard A. Norte,^{1,†,*} Jingkun Guo,¹ Clemens Schäfermeier,¹ and Simon Gröblacher^{1,‡}

¹*Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628CJ Delft, The Netherlands*

Demand for lightweight, highly reflective and mechanically compliant mirrors for optics experiments has recently seen a significant surge. Due to their bulky geometry, standard mirror solutions have a high mass, which severely limits their use in applications such as light sails [1], evanescent field sensors [2, 3], or deformable mirrors [4]. Advances in nanofabrication have shown that photonic crystal (PhC) membranes [5] are an ideal alternative to conventional mirrors, as they provide high reflectivity with only a single layer of dielectric material. In particular, devices made of silicon nitride constitute the state-of-the-art in PhC mirrors with low optical absorption and mechanical loss [6–8]. However, fabrication technology has constrained their effective area to a few square-micrometers. Here we experimentally demonstrate the first example of suspended PhC mirrors spanning areas up to $10 \times 10 \text{ mm}^2$. We overcome the limitations imposed by the finite size of the PhC, which allows us to measure reflectivities greater than 99 % on 210 nm thin devices at 1550 nm wavelength and beyond 90 % on 56 nm thick mirrors – an unrivaled performance compared to PhC mirrors with micro scale diameters. We also consider their use as mirrors in gravitational wave detectors where they could potentially reduce mirror coating noise at cryogenic temperatures. These structures bridge the gap between nano scale technologies and macroscopic optical elements.

Photonic crystal (PhC) membranes are suspended dielectric sheets patterned with sub-wavelength, low-index two-dimensional periodic structures [5]. These patterns give rise to resonances that can couple out-of-plane radiation to in-plane guided modes, and can be engineered to transform a flat membrane into a mirror [9], a lens [10], or even a curved mirror [10–12]. Here we study a PhC consisting of a periodic lattice of holes in a membrane, whose hole radius and lattice constant can be tuned to reflect light at a wavelength of choice. When fabricated from materials with low optical absorption such as low-pressure chemically vapor-deposited silicon nitride (LPCVD SiN), one can realize mirrors with sub-wavelength thicknesses and reflectivities $> 99 \%$, only limited by scattering losses [8]. LPCVD SiN thin films also enable the combination of PhC mirrors with low thermal noise mechanical oscillators, due to their high intrinsic stress, thin geometry, and weak coupling to undesired thermal modes [6, 13].

Limitations in microfabrication processes have so far restricted suspended PhC mirrors to areas not much bigger than $100 \times 100 \mu\text{m}^2$. This size sets an upper bound to the waist of incident Gaussian beams, since wider waists do not completely interact with the PhC, resulting in decreased reflectivity. But the beams also have a lower bound, since very small waists have a high divergence and couple to undesired PhC modes, which leads to shifting, broadening and shallowing of the reflectivity. These adverse finite-size effects have been consistently measured in very thin mirrors with thicknesses below 0.1λ , where λ is the wavelength of the reflected light [6, 7, 14, 15].

The ability to fabricate larger PhC mirrors with increasingly thinner membranes could greatly facilitate the combination of high reflectivity and low mechanical clamping losses [6]. These properties are crucial for reducing thermal mirror coating noise which stands as a limit in precision measurements such as atomic clocks [16], frequency-stabilized lasers [17], and gravitational wave detectors [18]. At the centimeter scale, PhC mirrors could have more immediate applications as deformable mirrors with adjustable wavefront [4], or evanescent field sensors with a large interaction area [2, 3].

Scaling up suspended PhC mirrors even further to meter scales would make this technology compatible with a number of next-generation experiments. Some of the most promising avenues for interstellar exploration, for example, rely on the development of low-mass light sails (i.e. lightweight reflectors), which could be accelerated to $1/5$ the speed of light using radiation pressure [1]. Initiatives like the Starshot Breakthrough [19] require meter-sized light sails with a thickness of only tens of nanometers – an aspect-ratio that is far beyond current nanotechnology and that stands out as one of the most daunting challenges of this ambitious project.

In this letter we experimentally demonstrate free-standing SiN photonic crystal mirrors with thicknesses of 56 and 210 nm and diameters of up to 10 mm. Not only do we increase the area of suspended PhC mirrors by nearly 4 orders of magnitude compared to previous works, we also show that these large aspect-ratios allow us to achieve high reflectivity from membranes 3 times thinner than previously measured. We observe greater than 90 % reflectivity of 1550 nm light from mirrors with a thickness of 0.038λ (56 nm) – an experimental first as PhC mirrors have consistently been limited to thicknesses above 0.13λ (210 nm at 1550 nm) to attain high reflectivity [6]. Such large structures allow studying the spec-

* These authors contributed equally to this work.

† r.a.norte@tudelft.nl

‡ s.groeblicher@tudelft.nl

Photographic Image Synthesis with Cascaded Refinement Networks

Qifeng Chen^{†‡}

Vladlen Koltun[†]

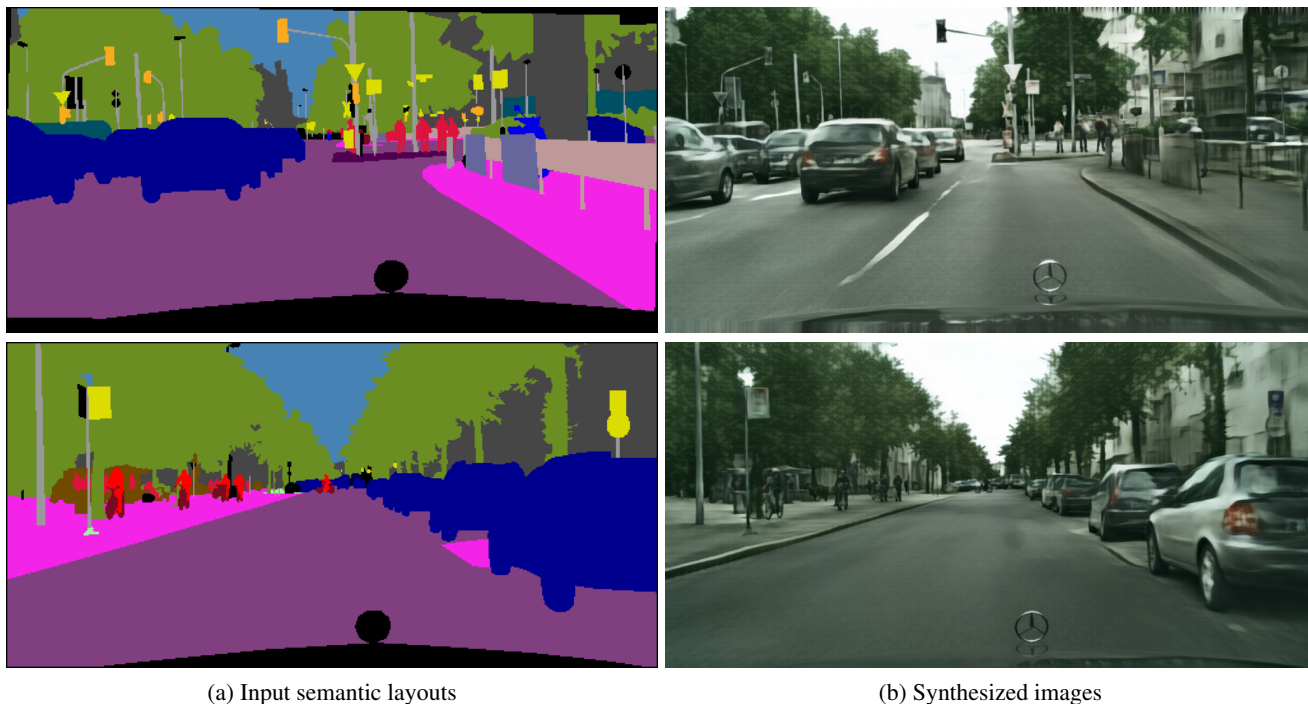


Figure 1. Given a pixelwise semantic layout, the presented model synthesizes an image that conforms to this layout. (a) Semantic layouts from the Cityscapes dataset of urban scenes; semantic classes are coded by color. (b) Images synthesized by our model for these layouts. The layouts shown here and throughout the paper are from the validation set and depict scenes from new cities that were never seen during training. Best viewed on the screen.

Abstract

We present an approach to synthesizing photographic images conditioned on semantic layouts. Given a semantic label map, our approach produces an image with photographic appearance that conforms to the input layout. The approach thus functions as a rendering engine that takes a two-dimensional semantic specification of the scene and produces a corresponding photographic image. Unlike recent and contemporaneous work, our approach does not rely on adversarial training. We show that photographic images can be synthesized from semantic layouts by a single feedforward network with appropriate structure, trained end-to-end with a direct regression objective. The presented approach scales seamlessly to high resolutions; we

demonstrate this by synthesizing photographic images at 2-megapixel resolution, the full resolution of our training data. Extensive perceptual experiments on datasets of outdoor and indoor scenes demonstrate that images synthesized by the presented approach are considerably more realistic than alternative approaches.

1. Introduction

Consider the semantic layouts in Figure 1. A skilled painter could draw images that depict urban scenes that conform to these layouts. Highly trained craftsmen can even create paintings that approach photorealism [20]. Can we train computational models that have this ability? Given a semantic layout of a novel scene, can an artificial system synthesize an image that depicts this scene and looks like a photograph?

[†]Intel Labs

[‡]Stanford University

Natural Language Processing with Small Feed-Forward Networks

Jan A. Botha Emily Pitler Ji Ma Anton Bakalov
 Alex Salcianu David Weiss Ryan McDonald Slav Petrov

Google Inc.
 Mountain View, CA

{jabot,epitler,maji,abakalov,salcianu,djweiss,ryanmcd,slav}@google.com

Abstract

We show that small and shallow feed-forward neural networks can achieve near state-of-the-art results on a range of unstructured and structured language processing tasks while being considerably cheaper in memory and computational requirements than deep recurrent models. Motivated by resource-constrained environments like mobile phones, we showcase simple techniques for obtaining such small neural network models, and investigate different tradeoffs when deciding how to allocate a small memory budget.

1 Introduction

Deep and recurrent neural networks with large network capacity have become increasingly accurate for challenging language processing tasks. For example, machine translation models have been able to attain impressive accuracies, with models that use hundreds of millions (Bahdanau et al., 2014; Wu et al., 2016) or billions (Shazeer et al., 2017) of parameters. These models, however, may not be feasible in all computational settings. In particular, models running on mobile devices are often constrained in terms of memory and computation.

Long Short-Term Memory (LSTM) models (Hochreiter and Schmidhuber, 1997) have achieved good results with small memory footprints by using character-based input representations: e.g., the part-of-speech tagging models of Gillick et al. (2016) have only roughly 900,000 parameters. Latency, however, can still be an issue with LSTMs, due to the large number of matrix multiplications they require (eight per LSTM cell): Kim and Rush (2016) report speeds of only 8.8 words/second when running a two-layer LSTM translation system on an Android phone.

Feed-forward neural networks have the potential to be much faster. In this paper, we show that small feed-forward networks can achieve results at or near the state-of-the-art on a variety of natural language processing tasks, with an order of magnitude speedup over an LSTM-based approach.

We begin by introducing the network model structure and the character-based representations we use throughout all tasks (§2). The four tasks that we address are: language identification (Lang-ID), part-of-speech (POS) tagging, word segmentation, and preordering for translation. In order to use feed-forward networks for structured prediction tasks, we use transition systems (Titov and Henderson, 2007, 2010) with feature embeddings as proposed by Chen and Manning (2014), and introduce two novel transition systems for the last two tasks. We focus on *budgeted* models and ablate four techniques (one on each task) for improving accuracy for a given memory budget:

1. Quantization: Using more dimensions and less precision (Lang-ID: §3.1).
2. Word clusters: Reducing the network size to allow for word clusters and derived features (POS tagging: §3.2).
3. Selected features: Adding explicit feature conjunctions (segmentation: §3.3).
4. Pipelines: Introducing another task in a pipeline and allocating parameters to the auxiliary task instead (preordering: §3.4).

We achieve results at or near state-of-the-art with small (< 3 MB) models on all four tasks.

2 Small Feed-Forward Network Models

The network architectures are designed to limit the memory and runtime of the model. Figure 1 illustrates the model architecture:

STARDATA: A StarCraft AI Research Dataset

Zeming Lin

Facebook
770 Broadway
New York, NY, 10003

Jonas Gehring

Facebook
6, rue Ménars
75002 Paris, France

Vasil Khalidov

Facebook
6, rue Ménars
75002 Paris, France

Gabriel Synnaeve

Facebook
770 Broadway
New York, NY, 10003

Abstract

We release a dataset of 65646 StarCraft replays that contains 1535 million frames and 496 million player actions. We provide full game state data along with the original replays that can be viewed in StarCraft. The game state data was recorded every 3 frames which ensures suitability for a wide variety of machine learning tasks such as strategy classification, inverse reinforcement learning, imitation learning, forward modeling, partial information extraction, and others. We use TorchCraft to extract and store the data, which standardizes the data format for both reading from replays and reading directly from the game. Furthermore, the data can be used on different operating systems and platforms. The dataset contains valid, non-corrupted replays only and its quality and diversity was ensured by a number of heuristics. We illustrate the diversity of the data with various statistics and provide examples of tasks that benefit from the dataset. We make the dataset available at <https://github.com/TorchCraft/StarData>. En Taro Adun!

Introduction

Real time strategy games as an AI research problem is attracting substantial attention (Ontañón et al. 2013; Usunier et al. 2016; Peng et al. 2017) due to their complex game dynamics, partial observability, and existing expert games in the form of human replays. These games are a good test bed for various reinforcement learning algorithms on a domain with higher complexity than toy robotics tasks and turn-based board games. Due to recent advances in deep learning, we see a trend of improved model performance with larger datasets. As learning capacity of these models increases, there is a growing need for data, especially in order to apply deep learning methods to control in RTS games.

Although learning in StarCraft can be performed through playing, the dynamics of the game are extremely complex, and it is beneficial to speed up learning by using existing games. The availability of datasets of recorded games between experienced players is therefore desirable.

StarCraft allows one to record replays of games which contain all commands issued by players. A number of online resources contain collections of replays from various tournaments (see Table 1). Some information can be directly

inferred from the replay file; however, reconstructing the full game state requires playback in StarCraft.

There are several aspects that make it difficult to use the replays directly for machine learning purposes. *Firstly*, the reconstruction speed of StarCraft is limited and would impose an upper threshold on training speed. *Secondly*, incompatibility between replays produced by different StarCraft versions makes it impossible to use the same game engine for all the replays or might result in corrupted data. *Finally*, the reconstruction process can only be reliably run on Windows, which adds additional unnecessary restrictions. Hence, the utility of a replay dataset can be increased by extracting game states, validating them and storing them as a separate dataset.

For a dataset to serve as a good base for learning models, it should fulfill a number of requirements:

Universality: the data stored in the dataset can be used to learn different aspects of game strategy and at different levels. Thus the dataset should provide data which is not specific to any particular context and should be as close to the full game state as possible.

Diversity: the dataset should cover a variety of game scenarios in terms of match-ups, maps, player strategies, etc.

Validity: the dataset should be representative of the distribution of StarCraft matches where both sides are trying to win.

Interfacing: one should be able to easily substitute game states received from the game engine with game states recorded in the dataset.

Portability: dataset access should be supported on a variety of platforms and operating systems.

With these requirements in mind, we constructed a new dataset of StarCraft replays from games among humans that can be used for StarCraft AI research. The following are our major contributions.

We provide a large set of StarCraft human replays, which is about 10x bigger than any of the comparable datasets currently available. The dataset includes a variety of scenarios and thus ensures the *diversity* requirement. Detailed statistics on matchups, maps etc. can be found in further sections.

All replays are checked for playability in StarCraft and BWAPI. We used additional scripted rule-based checks for

DeepRebirth: Accelerating Deep Neural Network Execution on Mobile Devices

Dawei Li
Samsung Research America
dawei.l@samsung.com

Xiaolong Wang
Samsung Research America
xiaolong.w@samsung.com

Deguang Kong
doogkong@gmail.com

Abstract

Deploying deep neural networks on mobile devices is a challenging task due to computation complexity and memory intensity. Current model reduction methods (e.g., matrix approximation using SVD) cannot satisfy real-time processing requirement. This paper first discovers that the major obstacle is the excessive execution time of non-tensor layers (with tensor-like parameters) such as pooling and normalization. This motivates us to design a novel acceleration framework: DeepRebirth through “slimming” existing consecutive and parallel non-tensor and tensor layers. The layer slimming is executed at different substructures: (a) streamline slimming by merging the consecutive non-tensor and tensor layer vertically; (b) branch slimming by merging non-tensor and tensor branches horizontally. These different optimization operations accelerate the model execution and reduce the run-time memory cost significantly. To maximally avoid accuracy loss, the parameters in new generated layers are learned with layer-wise fine-tuning based on both theoretical analysis and empirical verification. As observed in the experiment, DeepRebirth achieves 3x-5x speed-up and energy saving on GoogLeNet with only 0.4% accuracy drop on top-5 categorization in ImageNet. Further, by combining with other model compression techniques, DeepRebirth offers an average of 65ms model forwarding time on a single image using Samsung Galaxy S6 with 86.54% top-5 accuracy with 2.5x run-time memory saving.

1. Introduction

Recent years have witnessed the breakthrough of deep learning techniques for many computer vision tasks such as image classification and object detection. More and more mobile applications adopt deep learning techniques to provide accurate, intelligent and effective services. However, the execution speed of deep learning models on mobile devices becomes a bottleneck for deployment of many applications due to limited computing resources.

In this paper, we focus on improving the execution efficiency of deep learning models on mobile CPUs, which

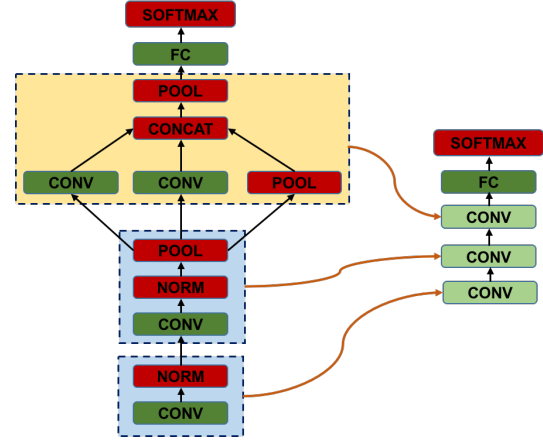


Figure 1: An illustration of proposed DeepRebirth model acceleration pipeline. DeepRebirth optimizes a trained deep learning model (left) to an accelerated “slim” model (right). Such optimization is achieved with two operations: Streamline Slimming which absorbs non-tensor layers (i.e., pooling and normalization) to their bottom convolutional layer (in light blue background) and Branch Slimming which absorbs non-tensor branches and convolutional branches with small convolution filters (e.g., 1x1) to a convolutional branch with large convolution filter (e.g., 5x5) (in light yellow background). We name new generated layers as slim layers.

is a highly intriguing feature. On one hand, a large majority of mobile devices are equipped with mobile GPUs, however, the speed-up achieved is quite limited when compared to CPU [18], not to mention the complexity caused by different mobile GPU architectures. On the other hand, major deep learning frameworks such as Caffe [14] and Tensorflow [1] only support CPU implementation on mobile devices currently, and therefore an efficient CPU-friendly model is highly desirable. In reality, it takes more than 651ms to recognize an image using GoogleNet on Samsung S5 (Table 4) with 984mJ energy costs (Table 5). The effective solution is expected to provide minimum accuracy loss by leveraging widely used deep learning framework (such as GoogLeNet and ResNet) with support of deep model acceleration on different types of layers.

Visions of Human Futures in Space and SETI

Jason T. Wright

Department of Astronomy & Astrophysics and
Center for Exoplanets and Habitable Worlds
525 Davey Laboratory, The Pennsylvania State University,
University Park, PA, 16802, USA, astrowright@gmail.com

Visiting Associate Professor, Department of Astronomy
Breakthrough Listen Laboratory
501 Campbell Hall #3411, University of California, Berkeley, CA,
94720, USA

PI, NASA Nexus for Exoplanet System Science

Michael P. Oman-Reagan

Department of Anthropology, Memorial University, St. John's, NL
A1C 5S7, Canada

August 22, 2017

Abstract

We discuss how visions for the futures of humanity in space and SETI are intertwined, and are shaped by prior work in the fields and by science fiction. This appears in the language used in the fields, and in the sometimes implicit assumptions made in discussions of them. We give examples from articulations of the so-called Fermi Paradox, discussions of the settlement of the Solar System (in the near future) and the Galaxy (in the far future), and METI. We argue that science fiction, especially the campy variety, is a significant contributor to the “giggle factor” that hinders serious discussion and funding for SETI and Solar System settlement projects. We argue that humanity’s long-term future in space will be shaped by our short-term visions for who goes there and how. Because of the way they entered the fields, we recommend avoiding the term “colony” and its cognates when discussing the settlement of space, as well as other terms with similar pedigrees. We offer examples of science fiction and other writing that broaden and challenge our visions of human futures in space and SETI. In an appendix, we use an analogy with

SMASH: One-Shot Model Architecture Search through HyperNetworks

Andrew Brock, Theodore Lim, & J.M. Ritchie
 School of Engineering and Physical Sciences
 Heriot-Watt University
 Edinburgh, UK
 {ajb5, t.lim, j.m.ritchie}@hw.ac.uk

Nick Weston
 Renishaw plc
 Research Ave, North
 Edinburgh, UK
 Nick.Weston@renishaw.com

Abstract

Designing architectures for deep neural networks requires expert knowledge and substantial computation time. We propose a technique to accelerate architecture selection by learning an auxiliary HyperNet that generates the weights of a main model conditioned on that model’s architecture. By comparing the relative validation performance of networks with HyperNet-generated weights, we can effectively search over a wide range of architectures at the cost of a single training run. To facilitate this search, we develop a flexible mechanism based on memory read-writes that allows us to define a wide range of network connectivity patterns, with ResNet, DenseNet, and FractalNet blocks as special cases. We validate our method (SMASH) on CIFAR-10 and CIFAR-100, STL-10, ModelNet10, and Imagenet32x32, achieving competitive performance with similarly-sized hand-designed networks.

1 Introduction

The high performance of deep neural nets is tempered by the cost of extensive engineering and validation to find the best architecture for a given problem. High-level design decisions such as depth, units per layer, and layer connectivity are not always obvious, and the success of models such as Inception [39], ResNets [13], FractalNets [20] and DenseNets [15] demonstrates the benefits of intricate design patterns. Even with expert knowledge, determining which design elements to weave together requires ample experimental time.

In this work, we propose to bypass the expensive procedure of fully training candidate models by instead training an auxiliary model, a HyperNet [12], to dynamically generate the weights of a main model with variable architecture. Though these generated weights are worse than freely learned weights for a fixed architecture, we leverage the observation [21] that the relative performance of different networks early in training (i.e. some distance from the eventual optimum) often provides a meaningful indication of performance at optimality. By comparing validation performance for a set of architectures using generated weights, we can approximately rank numerous architectures at the cost of a single training run.

To facilitate this search, we develop a flexible scheme based on memory read-writes that allows us to define a diverse range of architectures, with ResNets, DenseNets, and FractalNets as special cases. We validate our one-Shot Model Architecture Search through Hypernetworks (SMASH) for Convolutional Neural Networks (CNN) on CIFAR-10 and CIFAR-100 [19], Imagenet32x32 [7], ModelNet10 [41], and STL-10 [8], achieving competitive performance with similarly-sized hand-designed networks.

Opto-magnetic imaging of neural network activity in brain slices at high resolution using color centers in diamond

Mürsel Karadas¹, Adam M. Wojciechowski², Alexander Huck², Nils Ole Dalby^{2,3}, Ulrik Lund Andersen², Axel Thielscher^{1,4}

Institutions:

¹Department of Electrical Engineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

²Department of Physics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

³Department of Drug Design and Pharmacology, Copenhagen University, 2100 Copenhagen, Denmark

⁴Danish Research Center for Magnetic Resonance, Copenhagen University Hospital, 2650 Hvidovre, Denmark

Correspondence by:

Dr. Axel Thielscher
Danish Research Center for Magnetic Resonance
Copenhagen University Hospital Hvidovre
2650 Hvidovre, Denmark
Email: axelt@drcmr.dk
Phone #: 0045-38623326

HUMAN EXPERTS VS. MACHINES IN TAXA RECOGNITION

J. Ärje^a, V. Tirronen^b, S. Kärkkäinen^a, K. Meissner^c, J. Raitoharju^d, M. Gabbouj^d, S. Kiranyaz^e

^a University of Jyväskylä, Department of Mathematics and Statistics, johanna.arje@jyu.fi

^b University of Jyväskylä, Department of Information Technology

^c Finnish Environment Institute

^d Tampere University of Technology

^e Qatar University

ABSTRACT

Biomonitoring of waterbodies is vital as the number of anthropogenic stressors on aquatic ecosystems keeps growing. However, the continuous decrease in funding makes it impossible to meet monitoring goals or sustain traditional manual sample processing. In this paper, we review what kind of statistical tools can be used to enhance the cost efficiency of biomonitoring: We explore automated identification of freshwater macroinvertebrates which are used as one indicator group in biomonitoring of aquatic ecosystems. We present the first classification results of a new imaging system producing multiple images per specimen. Moreover, these results are compared with the results of human experts. On a data set of 29 taxonomical groups, automated classification produces a higher average accuracy than human experts.

1. INTRODUCTION

Benthic macroinvertebrates are a diverse group of species that quickly react to changes in their environment (Rosenberg and Resh, 1993). Their community composition can reflect even subtle human-induced changes in their environment, making them an ideal indicator group for aquatic biomonitoring (Wright et al., 1984; Karr and Chu, 2000). In many countries, biomonitoring of benthic macroinvertebrates is a key part of ecological status assessment of surface waters required by the European Union's Water Framework Directive (WFD, 2000).

The traditional process of macroinvertebrate biomonitoring is presented in Fig. 1. First, macroinvertebrates are sampled, usually by using a kick-net method (e.g. Brua et al., 2011). Second, the specimen are sorted out from the detritus and identified manually by an expert. Third, the observed taxa abundancies are used to calculate several biological indices indicating changes compared to previous sampling or a reference community. Finally, the index values are combined to evaluate the ecological status of the sampled waterbody.

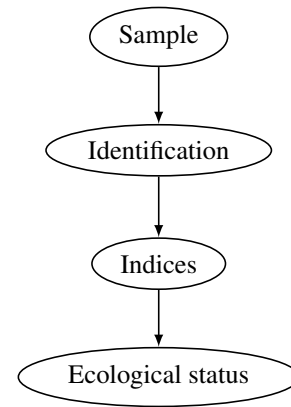


Fig. 1. The process of biomonitoring

In macroinvertebrate biomonitoring a large proportion of the total cost and time is spent on manual identification by highly trained experts. It takes several years to train an expert and manually identifying a sample of few thousand individuals can take hours. The monitoring process could be expedited substantially by shifting from manual to automated identification and in recent years there have been many studies on the automated identification of benthic macroinvertebrates (Tirronen et al., 2009; Lytle et al., 2010; Kiranyaz et al., 2010a,b; Ärje et al., 2010; Kiranyaz et al., 2011; Joutsijoki and Juhola, 2012; Joutsijoki et al., 2014; Ärje et al., 2013, 2017). Many biologists tend to oppose the shift to automated identification of macroinvertebrates due to fear of it not being accurate enough. However, manual identification has been found to be surprisingly error prone as well (Haase et al., 2010). While there exist studies on the automated classification of macroinvertebrates, to our knowledge, none of them include a comparison between manual and automated identification accuracy.

In this article we gather classification results achieved in previous studies on automated identification of macroinvertebrates on single image data. We introduce a new imaging

A Deep Structured Learning Approach Towards Automating Connectome Reconstruction from 3D Electron Micrographs

Jan Funke*, Fabian David Tschopp*, William Grisaitis, Arlo Sheridan,
Chandan Singh, Stephan Saalfeld, Srinivas C. Turaga



Abstract—We present a deep structured learning method for neuron segmentation from 3D electron microscopy (EM) which improves significantly upon the state of the art in terms of accuracy and scalability. Our method consists of a 3D U-NET architecture, trained to predict affinity graphs on voxels, followed by a simple and efficient iterative region agglomeration. We train the U-NET using a new structured loss function based on MALIS that encourages topological correctness. Our MALIS extension consists of two parts: First, we present an $O(n \log(n))$ method to compute the loss gradient, which improves over the originally proposed $O(n^2)$ algorithm. Second, we compute the gradient in two separate passes to avoid spurious gradient contributions in early training stages. Our affinity predictions are accurate enough that simple learning-free percentile-based agglomeration outperforms more involved methods used earlier on inferior predictions. We present results on three EM datasets (CREMI, FIB-25, and SEGEM) of different imaging techniques and animals where we achieve relative improvements over previous results of 27%, 15%, and 250%, respectively. Our findings suggest that a single 3D segmentation strategy can be applied to both nearly isotropic block-face EM data and anisotropic serial sectioned EM data. The runtime of our method scales with $O(n)$ in the size of the volume and achieves a throughput of about 2.6 seconds per megavoxel, qualifying our method for the processing of very large datasets.

1 INTRODUCTION

Precise reconstruction of neural connectivity is of great importance to understand the function of biological nervous systems. 3D electron microscopy (EM) is the only available imaging method with the resolution necessary to visualize and reconstruct dense neural morphology without ambiguity. At this resolution, however, even moderately small neural circuits yield image volumes that are too large for manual reconstruction. Therefore, automated methods for neuron tracing are needed to aid human analysis.

We present a structured deep learning based image segmentation method for reconstructing neurons from 3D electron microscopy which improves significantly upon state of the art in terms of accuracy and scalability. For an overview, see Fig. 1, top row. The main components of our method are: (1) Prediction of 3D affinity graphs using a 3D U-NET architecture [1], (2) a structured loss based on MALIS [2] to train the U-NET to minimize topological errors, and (3) an

efficient $O(n)$ agglomeration scheme based on quantiles of predicted affinities.

The choice of using a 3D U-NET architecture to predict voxel affinities is motivated by two considerations: First, U-NETs have already shown superior performance on the segmentation of 2D [3] and 3D [1] biomedical image data. One of their favourable properties is the multi-scale architecture which enables computational and statistical efficiency. Second, U-NETs efficiently predict large regions. This is of particular interest in combination with training on the MALIS structured loss, for which we need affinity predictions in a region.

We train our 3D U-NET to predict affinities using an extension of the MALIS loss function [2]. Like the original MALIS loss, we minimize a topological error on hypothetical thresholding and connected component analysis on the predicted affinities. We extended the original formulation to derive the gradient with respect to all predicted affinities (as opposed to sparsely sampling them), leading to denser and faster gradient computation. Furthermore, we compute the MALIS loss in two passes: In the *positive pass*, we constrain all predicted affinities between and outside of ground-truth regions to be 0, and in the *negative pass*, we constrain affinities inside regions to be 1 which avoids spurious gradients in early training stages.

Although the training is performed assuming subsequent thresholding, we found iterative agglomeration of *fragments* (or “supervoxels”) to be more robust to small errors in the affinity predictions. To this end, we extract fragments running a watershed algorithm on the predicted affinities. The fragments are then represented in a region adjacency graph (RAG), where edges are scored to reflect the predicted affinities between adjacent fragments: edges with small scores will be merged before edges with high scores. We discretize edge scores into k evenly distributed bins, which allows us to use a bucket priority queue for sorting. This way, the agglomeration can be carried out with a worst-case linear runtime.

The resulting method (prediction of affinities, watershed, and agglomeration) scales favourably with $O(n)$ in the size n of the volume, a crucial property for neuron segmentation from EM volumes, where volumes easily reach several

* these authors contributed equally

**Complete absence of thebaine biosynthesis under
home-brew fermentation conditions**

Drew Endy^{†^}, Stephanie Galanie^{†#}, and Christina D. Smolke^{^*}

[^]Department of Bioengineering; 443 Via Ortega, MC 4245
Stanford University; Stanford, CA 94305

[#]Department of Chemistry; 443 Via Ortega, MC 4245
Stanford University; Stanford, CA 94305

[†]These authors contributed equally to this work.

**Correspondence should be addressed to Christina D. Smolke
Phone: 650.721.6371
FAX: 650.721.6602
E-mail: csmolke@stanford.edu*

The aptabot: an inducibly affinity-switching, minimally invasive *in vivo* contrast agent

Elleard Felix Webster Heffern^{1,2,3}, Jason Fuller³, Russell W. Hanson^{3,4,5,†}

¹Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

²Department of Biochemistry, University of California in Los Angeles, Los Angeles, California, USA

³*Department of Genetics & Genomic Sciences*, Icahn Institute for Genomics & Multiscale Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, Box 1498, New York, NY 10029-6574

⁴Children's Hospital Informatics Program at the Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology, Harvard Medical School, Boston, Massachusetts, USA

⁵Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Lyngby, Denmark

†Corresponding author, russell.hanson@mssm.edu

Abstract

The ambitions of current neuroscience—understanding neurological disease progression and mapping the connectome—demonstrate a need for safe *in vivo* tools for creating intricate maps of brain circuitry. Present *in vivo* contrast agents are often limited by their specificity, uptake, resolvability, and/or clearance.

We describe an aptamer-functionalized sensor for high-resolution imaging that can switch imaging targets by an induced multi-stage aptamer reaction. Included are synthetic methods as well as calculations of sensor efficacy based on known kinetics. Calculations show that 10 distinct targets may be imaged in a living brain at the submicron scale within 42 hours.

Introduction

In order to map fine details of brain circuitry, agents must be developed that both bind specific targets of interest and provide highly resolvable contrast. MRI contrast agents have long been used to image brain activity and structure, and in recent years, specifically targeted PET, MRI, and radio imaging agents have been developed for expressed cell surface markers (Xue 2009). Current whole brain contrast agents are often constrained either by radiolabels that provide inadequate resolution to identify individual receptors and emit radiation in their environment, or metal nanoparticles with inefficient clearance or degradation that impedes imaging of subsequent targets. A multi-functionalized contrast agent that attaches to successive targets can provide the key to more intricate brain maps.

Rapidly evolving homing CRISPR barcodes

Reza Kalhor¹, Prashant Mali², George M. Church¹

¹*Department of Genetics, Harvard Medical School, Boston, MA 02115.*

²*Department of Bioengineering, University of California San Diego, La Jolla, CA 92093.*

Correspondence should be addressed to G. M. C.

(gchurch@genetics.med.harvard.edu) or to P. M. (pmali@ucsd.edu).

Keywords: DNA barcodes, lineage tracing, CRISPR, Cas9, homing CRISPR, homing guide RNA, brain mapping, fluorescent in situ sequencing, FISSEQ, genome engineering.

Abstract: We present here an approach for engineering evolving DNA barcodes in living cells. The methodology entails use of a homing guide RNA (hgRNA) scaffold that directs the Cas9-hgRNA complex to target the DNA locus of the hgRNA itself. We show this homing CRISPR-Cas9 system acts as an expressed evolving genetic barcode, and corresponding small RNAs can be assayed as single molecules *in situ*. This integrated approach will have wide ranging applications, such as in deep lineage tracing, cellular barcoding, molecular recording, dissecting cancer biology, and connectome mapping.

Main Text

A single totipotent zygote has the remarkable ability to generate an entire multicellular organism. Methodologies to comprehensively map and modulate the parameters that govern this transformation will have far ranging impact on our understanding of human development and our ability to restore normal function in damaged or diseased tissues. Precise lineage history of cells during development is one of the parameters that can shed important insights into developmental processes (Sulston et al. 1983; Kretzschmar and Watt 2012). Contemporary lineage-tracing approaches, however, do not scale readily to the model organisms, such as mice, that are most relevant to human development (Weisblat, Sawyer, and Stent 1978; Dymecki and Tomaszewicz 1998; Walsh and Cepko 1992; Porter et al. 2014; Lu et al. 2011). Precise mapping of lineage history in these organisms may be facilitated by combining modern genome engineering and DNA sequencing technologies (Mali, Esvelt, and Church 2013; Lee et al. 2014; Church, Marblestone, and Kalhor, 2015.): if every cell in an organism contained a

Biologically plausible learning in recurrent neural networks reproduces neural dynamics observed during cognitive tasks

Thomas Miconi

The Neurosciences Institute

La Jolla, CA 92037, USA

miconi@nsi.edu

Abstract

Neural activity during cognitive tasks exhibits complex dynamics that flexibly encode task-relevant variables. Chaotic recurrent networks, which spontaneously generate rich dynamics, have been proposed as a model of cortical computation during cognitive tasks. However, existing methods for training these networks are either biologically implausible, and/or require a continuous, real-time error signal to guide learning. Here we show that a biologically plausible learning rule can train such recurrent networks, guided solely by delayed, phasic rewards at the end of each trial. Networks endowed with this learning rule can successfully learn nontrivial tasks requiring flexible (context-dependent) associations, memory maintenance, nonlinear mixed selectivities, and coordination among multiple outputs. The resulting networks replicate complex dynamics previously observed in animal cortex, such as dynamic encoding of task features and selective integration of sensory inputs. We conclude that recurrent neural networks offer a plausible model of cortical dynamics during both learning and performance of flexible behavior.

Suite2p: beyond 10,000 neurons with standard two-photon microscopy

Marius Pachitariu^{1,2*}, Carsen Stringer³, Mario Dipoppa^{1,2}, Sylvia Schröder⁴, L. Federico Rossi⁴, Henry Dalglish⁵, Matteo Carandini⁴ and Kenneth D. Harris^{1,2}

¹UCL Institute of Neurology, London WC1N 3BG, UK.

²UCL Department of Neuroscience, Physiology, and Pharmacology, London WC1E 6BT, UK.

³Gatsby Computational Neuroscience Unit, London W1T 4JG, UK.

⁴UCL Institute of Ophthalmology, London EC1V 9EL, UK.

⁵UCL Wolfson Institute for Biomedical Research, London WC1E 6BT, UK.

*Correspondence to marius10p@gmail.com

Abstract

Two-photon microscopy of calcium-dependent sensors has enabled unprecedented recordings from vast populations of neurons. While the sensors and microscopes have matured over several generations of development, computational methods to process the resulting movies remain inefficient and can give results that are hard to interpret. Here we introduce Suite2p: a fast, accurate and complete pipeline that registers raw movies, detects active cells, extracts their calcium traces and infers their spike times. Suite2p runs on standard workstations, operates faster than real time, and recovers ~2 times more cells than the previous state-of-the-art method. Its low computational load allows routine detection of ~10,000 cells simultaneously with standard two-photon resonant-scanning microscopes. Recordings at this scale promise to reveal the fine structure of activity in large populations of neurons or large populations of subcellular structures such as synaptic boutons.

Introduction

Standard resonance-scanning two-photon microscopes readily image the activity of large numbers of neurons, but pipelines for processing the resulting data still suffer from significant limitations. Ideally, such a pipeline should satisfy several criteria. First, it should be fast, to keep up with ever-larger data sets produced by next-generation microscopes^{1,2}. Second, the pipeline should be transparent and its results interpretable, so that the original data undergo minimal processing, and a human curator can recognize mistakes or biases in the pipeline's output. Third, the pipeline should be accurate, so that its results require only brief curation by a human operator. Fourth, the pipeline should generalize to recordings of multiple cell types, subcellular structures, and brain regions, which can exhibit widely different activity patterns. Fifth, the pipeline should appropriately model and handle experimental confounds such as neuropil contamination³. Finally, it would be ideal if the pipeline could run on inexpensive workstations rather than requiring a cluster of servers, as some current software packages do⁴.

To fulfil these criteria, we developed Suite2p, an end-to-end pipeline made of fast and accurate algorithms (Figure 1a). The pipeline involves four independent stages: 1) image registration; 2) region-of-interest (ROI) detection; 3) ROI labelling and quality control; 4) activity extraction with neuropil correction and spike deconvolution (Figure 1b-e). Each stage can be separately adapted

A Rationally Designed Aminoacyl-tRNA Synthetase for Genetically Encoded Fluorescent Amino Acids

Ximena Steinberg¹, Jason Galpin², Gibran Nasir², Jose Sepulveda-Ugarte¹, Romina V. Sepúlveda³, Fernando Gonzalez-Nilo^{3,4}, Leon D. Islas⁵, Christopher A. Ahern² and Sebastian Brauchi^{1*}.

¹ Physiology Department, Faculty of Medicine, Universidad Austral de Chile, Campus Isla Teja, Valdivia, 5110566, Chile. ² Department of Molecular Physiology and Biophysics, University of Iowa, Iowa City, IA 52242, US. ³ Universidad Andres Bello, Center for Bioinformatics and Integrative Biology, Facultad de Ciencias Biológicas, Santiago 8370146, Chile. ⁴ Centro Interdisciplinario de Neurociencia de Valparaíso, Universidad de Valparaíso, Valparaíso, Chile ⁵ Departamento de Fisiología, Facultad de Medicina, Universidad Nacional Autónoma de México, Distrito Federal 04510, México

*Correspondence should be addressed to Dr. Sebastian Brauchi: sbrauchi@uach.cl

Abstract: The incorporation of non-canonical amino acids into proteins has emerged as a promising strategy to manipulate and study protein structure-function relationships with superior precision *in vitro* and *in vivo*. To date, fluorescent non-canonical amino acids (f-ncAA) have been successfully incorporated in proteins expressed in bacterial systems, *Xenopus* oocytes, and HEK-293T cells. Here, we describe the rational generation of an orthogonal aminoacyl-tRNA synthetase based on the *E. coli* tyrosine synthetase that is capable of encoding the f-ncAA tyr-coumarin in HEK-293T cells.

Keywords: Unnatural amino acids, aminoacyl-tRNA synthetase, coumarin

INTRODUCTION

The emergence of chemical conjugation and genetic encoding techniques to label proteins with fluorescent probes has enabled significant advances in the mechanistic understanding of proteins in biochemical and cellular environments¹. Encoding large fluorescent proteins (e.g. GFP) as fusion protein products is experimentally straightforward, however, the relative size of the probes can alter the function and biology of the protein being studied. Alternatively, chemical conjugation of an expressed protein requires the labeling sites are solvent accessible and labeling of cytoplasmic sites often comes with significant background reactivity². A possible solution to these issues is the use of genetic code expansion to introduce a relatively compact fluorescent side chain as a non-canonical amino acid directly into the target protein in a site-specific fashion³. Indeed, the rapid development of genetically encoded fluorophores as a non-canonical amino acids (ncAA) is emerging as a promising strategy to describe protein function under minimal perturbations in eukaryotic cells^{3,4}. The experimental strategy employs an orthogonal suppressor tRNA and an evolved aminoacyl tRNA synthetase

(RS), often based upon the tyrosine pair, which can be used to encode the ncAA at virtually any site in the reading frame of the target gene. This pair is orthogonal to the translation system employed, meaning that an evolved TyrRs cannot acylate endogenous Tyr-tRNA molecules and the suppressor tRNA^{tyr} is minimally acylated by host cell synthetases⁵. The orthogonal tRNA has the appropriate anticodon to suppress the nonsense codon, thus allowing for an introduced amber codon of target genes in both prokaryote and eukaryotic systems^{3,5,6}. This approach has been successfully used for site-specific incorporation of f-ncAA into a soluble proteins in prokaryote cells (dansyl and hydroxycoumarin)^{4,7}, membrane proteins expressed in *Xenopus laevis* oocytes (bodipy and anap)^{8,9}, and proteins expressed in mammalian cells (dansyl and anap)^{10,11}. Hydroxycoumarin is notable because of its small size and high environmental sensitivity^{12,13}, however, no system yet exists for its incorporation via expanding genetic code in mammalian cells.

Genome Editing With Targeted Deaminases

Luhan Yang^{1,2±}, Adrian W. Briggs^{1*}, Wei Leong Chew^{1,2*}, Prashant Mali¹, Marc Guell¹,
John Aach¹, Daniel Bryan Goodman^{1,3}, David Cox³, Yinan Kan¹, Emal Lesha¹
Venkataramanan Soundararajan¹, Feng Zhang⁵, George Church^{1,3,4,±}

¹ Department of Genetics, Harvard Medical School, Boston, MA

² Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA

³ Harvard-MIT Division of Health Science and Technology, Cambridge, MA

⁴ Wyss Institute for Biologically Inspired Engineering, Boston, MA

⁵ Broad Institute of MIT and Harvard, Cambridge, MA

McGovern Institute for Brain Research, MIT, Cambridge, MA

Department of Brain and Cognitive Sciences, MIT Cambridge, MA

* Contributed equally to this work as joint second authors

± To whom correspondence should be addressed.

E-mail: yang_luhan@genetics.med.harvard.edu, gmc@harvard.edu

Luhan Yang

Department of Genetics, Harvard Medical School

New Research Building, Rm 232

77 Avenue Louis Pasteur, Boston, MA 02115

E-mail: yang_luhan@genetics.med.harvard.edu

George Church, Ph.D.

Department of Genetics, Harvard Medical School

New Research Building, Rm 238

77 Avenue Louis Pasteur, Boston, MA 02115

E-mail: gmc@harvard.edu

***Easi*-CRISPR: Efficient germline modification with long ssDNA donors**

Rolen M. Quadros^{1,12}, Masato Ohtsuka^{2,3,12}, Donald W Harms^{1,12}, Tomomi Aida^{4,5}, Ronald Redder⁶, Hiromi Miura⁷, Guy P. Richardson⁸, Mark A. Behlke⁹, Sarah A. Zeiner⁹, Ashley M. Jacobi⁹, Lisa D. Urness¹⁰, Suzanne L. Mansour^{10,13}, Channabasavaiah B. Gurumurthy^{1,11,13}

- 1) Mouse Genome Engineering Core Facility, Vice Chancellor for Research Office, University of Nebraska Medical Center, Omaha, NE, USA
- 2) Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, School of Medicine, Tokai University, Kanagawa 259-1193, Japan
- 3) The Institute of Medical Sciences, Tokai University, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan
- 4) Laboratory of Molecular Neuroscience, Medical Research Institute (MRI), Tokyo Medical and Dental University (TMDU), 1-5-45, Yushima, Bunkyo, Tokyo, 113-8510, Japan
- 5) Laboratory of Recombinant Animals, MRI, TMDU, 2-3-10, Surugadai, Kanda, Chiyoda, Tokyo, 101-0062, Japan
- 6) High-Throughput DNA Sequencing and Genotyping Core Facility, Vice Chancellor for Research Office, University of Nebraska Medical Center, Omaha, NE, USA
- 7) Department of Regenerative Medicine, Basic Medical Science, Tokai University School of Medicine
- 8) Sussex Neuroscience, University of Sussex, Falmer, Brighton, BN1 9QG, UK
- 9) Integrated DNA Technologies, Inc. Coralville, IA, 52241, USA
- 10) Department of Human Genetics, University of Utah, Salt Lake City, UT, 84112, USA
- 11) Developmental Neuroscience, Munroe Meyer Institute for Genetics and Rehabilitation, University of Nebraska Medical Center, Omaha, NE, USA
- 12) Contributed equally
- 13) Corresponding authors: cgurumurthy@unmc.edu and suzi.mansour@genetics.utah.edu

Abstract: CRISPR/Cas9 technology efficiently produces short insertions or deletions (*indels*) and can insert short exogenous sequences at Cas9 cut sites. However, targeting long inserts is still a major technical challenge. To overcome this challenge, we developed *Easi*-CRISPR (Efficient additions with ssDNA inserts-CRISPR), a method that uses long, *in vitro*-synthesized, single-stranded DNAs with 50-100 base homology arms as repair templates. We demonstrate that *Easi*-CRISPR can generate knock-in and floxed alleles in mice with an efficiency at many loci as high as 100%. The simple design requirements for donor DNAs and the reproducibly high-efficiency of *Easi*-CRISPR enables rapid development of many types of commonly used animal and cell models.

Tractography-based connectomes are dominated by false-positive connections

Klaus H. Maier-Hein^{*,a}, Peter Neher^a, Jean-Christophe Houde^b, Marc-Alexandre Côté^b, Eleftherios Garyfallidis^b, Jidan Zhong^g, Maxime Chamberland^b, Fang-Cheng Yeh^h, Ying-Chia Linⁱ, Qing Ji^j, Wilburn E. Reddick^j, John O. Glass^j, David Qixiang Chen^k, Yuanjing Feng^l, Chengfeng Gao^l, Ye Wu^l, Jieyan Ma^m, H Renjie^m, Qiang Li^{m,n}, Carl-Fredrik Westin^o, Samuel Deslauriers-Gauthier^b, J. Omar Ocegueda González^p, Michael Paquette^b, Samuel St-Jean^b, Gabriel Girard^b, François Rheault^b, Jasmeen Sidhu^b, Chantal M.W. Tax^r, Fenghua Guo^r, Hamed Y. Mesri^r, Szabolcs Dávid^r, Martijn Froeling^s, Anneriet M. Heemskerk^r, Alexander Leemans^r, Arnaud Boré^q, Basile Pinsard^{q,zg}, Christophe Bedetti^{q,zh}, Matthieu Desrosiers^q, Simona Brambati^q, Julien Doyon^q, Alessia Sarica^t, Roberta Vasta^t, Antonio Cerasa^t, Aldo Quattrone^u, Jason Yeatman^v, Ali R. Khan^w, Wes Hodges^x, Simon Alexander^x, David Romascano^d, Muhamed Barakovic^d, Anna Auría^d, Oscar Esteban^{zd}, Alia Lemkaddem^d, Jean-Philippe Thiran^{d,ze}, H. Ertan Cetinçul^y, Benjamin L. Odry^y, Boris Mailhe^y, Mariappan S. Nadar^y, Fabrizio Pizzagalli^z, Gautam Prasad^z, Julio E. Villalon-Reina^z, Justin Galvis^z, Paul M. Thompson^z, Francisco De Santiago Requejo^{za}, Pedro Luque Laguna^{za}, Luis Miguel Lacerda^{za}, Rachel Barrett^{za}, Flavio Dell'Acqua^{za}, Marco Catani^{za}, Laurent Petit^{zb}, Emmanuel Caruyer^e, Alessandro Daducci^d, Tim B. Dyrby^{f,zf}, Tim Holland-Letz^{zc}, Claus C. Hilgetag^{zi}, Bram Stieltjes^c, Maxime Descoteaux^{*,b}

*indicates corresponding authors.

- a. Medical Image Computing Group (MIC), German Cancer Research Center (DKFZ), Heidelberg, Germany
- b. Sherbrooke Connectivity Imaging Lab (SCIL), Université de Sherbrooke, Sherbrooke, Quebec, Canada
- c. University Hospital Basel, Radiology & Nuclear Medicine Clinic, Basel, Switzerland.
- d. Signal Processing Lab (LTS5), Ecole Polytechnique Federale de Lausanne, Switzerland
- e. Centre national de la recherche scientifique (CNRS), Institute for Research in IT and Random Systems (IRISA), UMR 6074 VISAGES project-team, Rennes, France
- f. Danish Research Centre for Magnetic Resonance, Center for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark
- g. Krembil Research Institute, University Health Network, Toronto, Canada
- h. Department of Psychology, Carnegie Mellon University, USA
- i. IMT - Institute for Advanced Studies, Lucca, Italy
- j. Department of Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, USA
- k. University of Toronto Institute of Medical Science, Toronto, Canada
- l. Institute of Information Processing and Automation, Zhejiang University of Technology, Hangzhou, Zhejiang, China
- m. United Imaging Healthcare Co., Shanghai, China
- n. Shanghai Advanced Research Institute, Shanghai, China
- o. Laboratory of Mathematics in Imaging, Harvard Medical School, Boston, MA, United States
- p. Center for Research in Mathematics, Guanajuato, Mexico
- q. Centre de recherche institut universitaire de geriatrie de Montreal (CRIUGM), Université de Montréal, Montreal, Quebec, Canada
- r. PROVIDI Lab, Image Sciences Institute, University Medical Center Utrecht, Utrecht, Netherlands
- s. Department of Radiology, University Medical Center Utrecht, Utrecht, Netherlands
- t. Neuroimaging Unit, Institute of Bioimaging and Molecular Physiology (IBFM), National Research Council (CNR), Policlinico Magna Graecia, Germaneto (CZ), Italy
- u. Institute of Neurology, University Magna Graecia, Germaneto (CZ), Italy

1 **Structure of the *Legionella* Dot/Icm type IV secretion system *in situ* by**
2 **electron cryotomography**

3

4 Debnath Ghosal¹, Yi-Wei Chang¹, Kwangcheol C. Jeong^{2,a}, Joseph P. Vogel², Grant J.
5 Jensen^{1,3*}

6

7 Affiliations:

8 ¹California Institute of Technology, Pasadena, CA 91125, USA.

9 ²Washington University School of Medicine, St. Louis, MO 63110, USA.

10 ³Howard Hughes Medical Institute, Pasadena, CA 91125, USA.

11 Present address:

12 ^aUniversity of Florida, Gainesville, FL 32611, USA.

13

14 *Correspondence to: jensen@caltech.edu.

In vivo magnetic recording of neuronal activity

Laure Caruso¹, Thomas Wunderle², Christopher Murphy Lewis², Joao Valadeiro^{3,4}, Vincent Trauchessec¹, Josué Trejo Rosillo¹, José Pedro Amaral^{3,4}, Jianguang Ni², Patrick Jendritza², Claude Fermon¹, Susana Cardoso^{3,4}, Paulo Peixeiro Freitas^{3,4}, Pascal Fries^{2,5,6,*}, Myriam Pannetier-Lecoeur^{1,6,7,*}

¹SPEC, CEA, CNRS, Université Paris-Saclay, CEA Saclay 91191 Gif-sur-Yvette Cedex, France.

²Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Deutschordenstraße 46, 60528 Frankfurt, Germany.

³Instituto de Engenharia de Sistemas de Computadores-Microsystems and Nanotechnology (INESC-MN), Rua Alves Redol, No. 9, Lisboa 1000-029, Portugal.

⁴Instituto Superior Técnico IST, Physics Department, Universidade de Lisboa, Lisbon 1049-001, Portugal.

⁵Donders Institute for Brain, Cognition and Behaviour, Kapittelweg 29, 6525 EN Nijmegen, Netherlands.

⁶These authors contributed equally.

⁷Lead contact.

*Correspondence: myriam.lecoeur@cea.fr, pascal.fries@esi-frankfurt.de

KEYWORDS

Magnetic fields, magnetoencephalography, MEG, spin electronics, magnetic sensors.

Using high-throughput barcode sequencing to efficiently map connectomes

Ian D Peikon^{*1,2}, Justus M Kebschull^{*1,2}, Vasily V Vagin^{*2}, Diana I Ravens², Eric Brouzes³, Ivan R. Corrêa Jr.⁴, Dario Bressan^{1,2,5}, Anthony M Zador²

¹Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

²Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

³Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA

⁴New England Biolabs, Inc., Ipswich, Massachusetts 01938, USA

⁵Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, United Kingdom

*these authors contributed equally to the work

Correspondence and requests should be addressed to Anthony Zador, zador@cshl.edu.

Abstract

The function of a neural circuit is determined by the details of its synaptic connections. At present, the only available method for determining a neural wiring diagram with single synapse precision—a “connectome”—is based on imaging methods that are slow, labor-intensive and expensive. Here we present SYNseq, a method for converting the connectome into a form that can exploit the speed and low cost of modern high-throughput DNA sequencing. In SYNseq, each neuron is labeled with a unique random nucleotide sequence—an RNA “barcode”—which is targeted to the synapse using engineered proteins. Barcodes in pre- and postsynaptic neurons are then associated through protein-protein crosslinking across the synapse, extracted from the

A systematic comparison of error correction enzymes by next-generation sequencing

Nathan B. Lubock^{1,2,3†}, Di Zhang^{4†}, George M. Church^{5,6}, Sriram Kosuri^{1,2,3*}

¹ Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California, USA

² UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, California, USA

³ Molecular Biology Institute, University of California, Los Angeles, California, USA

⁴ Genomics and Computational Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁵ Wyss Institute for Biologically Inspired Engineering, Boston, Massachusetts, USA

⁶ Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

January 16, 2017

Abstract

Gene synthesis, the process of assembling gene-length fragments from shorter groups of oligonucleotides (oligos), is becoming an increasingly important tool in molecular and synthetic biology. The length, quality, and cost of gene synthesis is limited by errors produced during oligo synthesis and subsequent assembly. Enzymatic error correction methods are cost-effective means to ameliorate errors in gene synthesis. Previous analyses of these methods relied on cloning and Sanger sequencing to evaluate their efficiencies, limiting quantitative assessment and throughput. Here we develop a method to quantify errors in synthetic DNA by next-generation sequencing. We analyzed errors in a model gene assembly and systematically compared six different error correction enzymes across 11 conditions. We find that ErrASE and T7 Endonuclease I are the most effective at decreasing average error rates (up to 5.8-fold relative to the input), whereas MutS is the best for increasing the number of perfect assemblies (up to 25.2-fold). We are able to quantify differential specificities such as ErrASE preferentially corrects C/G → G/C transversions whereas T7 Endonuclease I preferentially corrects A/T → T/A transversions. More generally, this experimental and computational pipeline is a fast, scalable, and extensible way to analyze errors in gene assemblies, to profile error correction methods, and to benchmark DNA synthesis methods.

Introduction

Synthetic DNA is a central tool for biological research [1]. Notably, the initial development of nucleic acid synthesis led directly to the cracking of the genetic code [2]. Today, progress in biology is often limited by the difficulty in producing long, high-quality synthetic DNA [3, 4]. This bottleneck is particularly apparent in the assembly of gene-sized fragments of DNA known as gene synthesis [5].

*To whom correspondence should be addressed. Tel: +1 310 825-8931; Email: sri@ucla.edu

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

Novel Abundant Oceanic Viruses of Uncultured Marine Group II Euryarchaeota Identified by Genome-Centric Metagenomics

Alon Philoso^{1*}, Natalya Yutin², José Flores-Urbe¹, Itai Sharon³, Eugene V. Koonin²,
 5 and Oded Béja^{1*}

¹Faculty of Biology, Technion - Israel Institute of Technology, Haifa, Israel.

²National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, 20894, USA.

10 ³Migal Galilee Research Institute, Kiryat Shmona, 11016, Israel. Tel Hai College, Upper Galilee 12210, Israel.

*To whom correspondence should be addressed. [E-mail: aphilosof@gmail.com](mailto:aphilosof@gmail.com) and

E-mail: beja@tx.technion.ac.il

An inducible CRISPR-ON system for controllable gene activation in human pluripotent stem cells

Jiaying Guo¹, Dacheng Ma², Rujin Huang¹, Jia Ming¹, Min Ye¹, Kehkooi Kee¹, Zhen Xie², and Jie Na^{1*}

¹Center for Stem Cell Biology, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China.

²MOE Key Laboratory of Bioinformatics and Bioinformatics Division, Center for Synthetic and System Biology, TNLIST/Department of Automation, Tsinghua University, Beijing 100084, China.

Corresponding author

Jie Na

jie.na@tsinghua.edu.cn

Running title: CRISPR-ON gene activation system in human pluripotent stem cells

Keywords: CRISPR, transcription activation, human pluripotent stem cells, NANOG, pluripotency

Engineered shell proteins confer improved encapsulated pathway behavior in a bacterial microcompartment

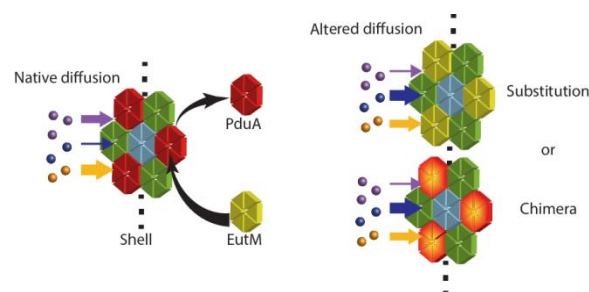
Marilyn F. Slininger Lee[†], Christopher M. Jakobson[†], Danielle Tullman-Ercek^{‡*}

[†] Department of Chemical and Biomolecular Engineering, University of California Berkeley, United States

[‡] Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan Rd, Technological Institute E136, Evanston, IL 60208-3120, USA; Chemistry of Life Processes Institute, Northwestern University, 2170 Campus Drive, Evanston, IL 60208-3120

*Corresponding author

TOC Abstract Graphic



Abstract: Bacterial microcompartments are a class of proteinaceous organelles comprising a characteristic protein shell enclosing a set of enzymes. Compartmentalization can prevent escape of volatile or toxic intermediates, prevent off-pathway reactions, and create private cofactor pools. Encapsulation in synthetic microcompartment organelles will enhance the function of heterologous pathways, but to do so, it is critical to understand how to control diffusion in and out of the microcompartment organelle. To this end, we explored how small differences in the shell protein structure result in changes in the diffusion of metabolites through the shell. We found that the ethanolamine utilization (Eut) protein EutM properly incorporates into the 1,2-propanediol utilization (Pdu) microcompartment, altering native metabolite accumulation and the resulting growth on 1,2-propanediol as the sole carbon source. Further, we identified a single pore-lining residue mutation that confers the same phenotype as substitution of the full EutM protein, indicating that small molecule diffusion through the shell is the cause of growth enhancement. Finally, we show that the hydropathy index and charge of pore amino acids are important indicators to predict how pore mutations will affect growth on 1,2-

Multidimensional encoding of brain connectomes

Cesar F. Caiafa^{1,2,3} and Franco Pestilli¹

¹ Department of Psychological and Brain Sciences, Programs in Neuroscience and Cognitive Science, Indiana University Bloomington, 1101 E 10th Street, Bloomington, Indiana USA 47401

² Instituto Argentino de Radioastronomía (CCT-La Plata, CONICET; CICPBA), CC5 V. Elisa, ARGENTINA, 1894.

³ Facultad de Ingeniería - Universidad de Buenos Aires, Buenos Aires, ARGENTINA, C1063ACV.

The ability to map brain networks at the macroscale in living individuals is fundamental in efforts to chart the relation between human behavior, health and disease. We present a framework to encode structural brain connectomes and diffusion-weighted magnetic resonance data into multidimensional arrays (tensors). The framework overcomes current limitations in building connectomes; it prevents information loss by integrating the relation between connectome nodes, edges, fascicles and diffusion data. We demonstrate the utility of the framework for *in vivo* white matter mapping and anatomical computing. The framework reduces dramatically storage requirements for connectome evaluation methods, with up to 40x compression factors. We apply the framework to evaluate 1,980 connectomes, thirteen tractography methods, and three data sets. We describe a general equation to predicts connectome resolution (number of fascicles) given data quality and tractography model parameters. Finally, we provide open-source software implementing the method and data to reproduce the results.

INTRODUCTION

A fundamental goal of neuroscience is to develop methods to understand how brain networks support function and behavior in individuals across human populations [1–3]. The recent increase in availability of neuroimaging data and large scale projects has the potential to empower new ways of discovery by studying large populations of human brains [4–22]. Exploiting these large-scale data sets will require advances in measurement, computational approaches and theories [23].

Innovation in measurement and computational methods for human brain mapping is shifting the *in vivo* study of the white matter and large-scale brain networks beyond qualitative characterization (such as *camera lucida* drawings), toward structural and functional quantification [24–30]. Tractography and diffusion-weighted magnetic resonance imaging (dMRI) are the primary methods for mapping structural brain connectivity and white matter tissue properties in living human brains. These *in vivo* investigations have shown that there is much to learn about the macrostructural organization of the human brain such that network neuroscience has become one of the fastest-growing fields [3,25,28,29,31–39].

Tractography algorithms use dMRI data to estimate the three-dimensional trajectory of neuronal axons bundles wrapped by myelin sheaths – the white matter fascicles. Fascicles are normally represented as sets of brain coordinates, with coordinates segments spanning anything between 0.01 to 1 mm in length (**Fig. 1a top**). Fascicles have historically been clustered into anatomically cohesive groups called white matter tracts. The largest tracts in the human brain are relatively well characterized and associated with names – such as the corticospinal tract (CST) and the arcuate fasciculus (**Fig. 1b top** [40,41]). White matter tracts communicate between cytoarchitectonically and functionally distinct areas – such as Broca’s or Wernicke’s areas involved in human language processing (**Fig. 1c top** [42–44]). White matter tracts and brain areas together compose a large-scale network called the connectome [45]. Within this network, white-matter tracts represent communication pathways (the edges; **Fig. 1b top**) and brain areas units of information processing (the nodes; **Fig. 1c-top**).

Title: Individual nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP) to determine protein-RNA interactions.

Author:

Christopher R. Sibley

Correspondence:

Address: Division of Brain Sciences
Department of Medicine
Imperial College London
London
UK

E-mail: c.sibley@imperial.ac.uk

Tel: +44 (0)207 5947019

Uncharacterized bacterial structures revealed by electron cryotomography

Megan J. Dobro¹, Catherine M. Oikonomou², Aidan Piper¹, John Cohen¹, Kylie Guo², Taylor Jensen², Jahan Tadayon², Joseph Donermeyer², Yeram Park², Benjamin A. Solis³, Andreas Kjær⁴, Andrew I. Jewett², Alasdair W. McDowall², Songye Chen², Yi-Wei Chang², Jian Shi⁵, Poorna Subramanian², Cristina V. Iancu⁶, Zhuo Li⁷, Ariane Briegel⁸, Elitza I. Tocheva⁹, Martin Pilhofer¹⁰, Grant J. Jensen^{2,11,*}

¹ Hampshire College, 893 West St., Amherst, MA 01002

² California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125

³ University at Albany, SUNY, 135 Western Avenue, Albany, NY. 12203

⁴ University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

⁵ National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077

⁶ Rosalind Franklin University of Medicine and Science, 3333 Green Bay Rd., North Chicago, IL 60064

⁷ City of Hope, 1500 E. Duarte Road, Duarte, CA 91010

⁸ Leiden University, Sylvius Laboratories, Sylviusweg 72, 2333 BE, Leiden, Netherlands

⁹ University of Montreal, C.P. 6128, succursale Centre-ville, Montreal, Quebec, Canada

¹⁰ ETH Zurich, Otto-Stern-Weg 5, 8093 Zurich, Switzerland

¹¹ Howard Hughes Medical Institute, 1200 E. California Blvd., Pasadena, CA 91125

* To whom correspondence should be addressed. Tel: (626) 395-8827. Email:

jensen@caltech.edu

Running title: ECT of novel bacterial structures

Keywords: bacteria, electron cryotomography, cryo-EM, bacterial ultrastructure, uncharacterized structures

Yi-Wei Chang¹, Lee A. Rettberg², Grant J. Jensen^{1,2,*}

¹California Institute of Technology, Pasadena, CA 91125, USA.

*Correspondence to: jensen@caltech.edu.

Broadband spectral responses in visual cortex revealed by a new MEG denoising algorithm

Eline R. Kupers¹, Helena X. Wang¹, Kaoru Amano², Kendrick N. Kay³, David J. Heeger¹, Jonathan Winawer¹

Author institutions

1. Department of Psychology and Center for Neural Science, New York University
2. Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology
3. Center for Magnetic Resonance Research, University of Minnesota

Short title: Broadband spectral responses in visual cortex

Contact information:

Eline Kupers
Department of Psychology, New York University
New York University
6 Washington Place, Room 959
New York, NY 10003
eline.kupers@nyu.edu

Acknowledgments: This study was supported NEI grant R00-EY022116 (J.W.), and NIMH grant R01-MH111417 (J.W.).

Whole exome sequencing reveals a mutation in *ARMC9* as a cause of mental retardation, ptosis and polydactyly

Anjana Kar^{1,2}, Shubha R Phadke³, Aneek Das Bhowmik², Ashwin Dalal^{1,4*}

1) Diagnostics Division, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India;

2) Graduate Studies, Manipal University, Manipal, India;

3) Department of Medical Genetics, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, India;

4) Department of Medical Genetics, Kasturba Medical College, Manipal University, Manipal, India.

Conflict of Interest: The authors declare no conflict of interests.

Correspondence to:

Dr Ashwin Dalal

Head, Diagnostics Division

Centre for DNA Fingerprinting and Diagnostics

4-1-714, Tuljaguda Complex Mozamzahi Road, Nampally

Hyderabad Telangana 500001 INDIA

Ph. 040-24749335

Fax: 040 24749448

email: ashwindalal@gmail.com, adalal@cdfd.org.in

Running title: Exome sequencing reveals cause of mental retardation, ptosis and polydactyly syndrome.

Key Words: Intellectual disability, *ARMC9*, Exome sequencing, Homozygosity mapping, Splice site assay.

ABSTRACT

Intellectual disability (ID) refers to deficits in mental abilities, social behaviour and motor skills to perform activities of daily living as compared to peers. Numerous genetic and environmental factors may be responsible for ID. We report on identification of a novel gene for syndromic ID, using homozygosity mapping followed by exome sequencing in a family with mental retardation, ptosis and polydactyly. The analysis revealed a synonymous mutation c.879G>A which leads to a splicing defect in *ARMC9* gene. The variant is present in conserved region of ARM domain of *ARMC9* protein which is predicted to form a platform for protein interaction. This domain is likely to be altered in patients due to splicing defect caused by this synonymous mutation. Our study was helpful in elucidation of molecular basis of mental retardation, ptosis and polydactyly phenotype and addition of *ARMC9* to group of genes leading to syndromic ID.

INTRODUCTION

Intellectual disability has a prevalence of 1-3% in population and can result from heterogeneous causes like environmental/nutritional effect, chromosomal or monogenic causes¹. More than 230 genes have been reported to be involved in causation of syndromes of intellectual disability. Mental retardation, ptosis and polydactyly² are a distinctive combination of clinical features reported by us. This is a rare type of intellectual disability syndrome, reported only in a single consanguineous Muslim family from India, where three individuals were affected².

Intellectual disability, short stature and polydactyly was suggestive of possibility of Bardet Biedl syndrome (BBS, MIM: 209900), but absence of obesity, renal abnormality, retinopathy and normal sexual development were some of the phenotypes that were different from BBS. Another possibility

was 3MC syndrome, which was formally known as Carnevale syndrome (MIM: 265050), however patients did not had hip dysplasia, cryptorchidism and abdominal muscle defect. For identification of candidate gene in this family we have employed homozygosity mapping followed by exome sequencing in all the three affected siblings with mental retardation, ptosis and polydactyly phenotype.

MATERIALS AND METHODS

Patients Details

We described the clinical features in three affected siblings born out of consanguineous union, characterized as mental retardation, ptosis and polydactyly phenotype². The proband showed severe intellectual disability (ID), bilateral ptosis, downslanting palpebral fissures, hypertelorism, round face, high arched palate, clinodactyly, tapering of fingers and

Patterns of shared signatures of recent positive selection across human populations

Kelsey Elizabeth Johnson¹, Benjamin F. Voight^{2,3,4}

¹ Cell and Molecular Biology Graduate Group, Genetics and Gene Regulation Program,
Perelman School of Medicine, University of Pennsylvania, Philadelphia PA 19104

² Department of Systems Pharmacology and Translational Therapeutics, Perelman School of
Medicine, University of Pennsylvania, Philadelphia, PA 19104

³ Department of Genetics, Perelman School of Medicine, University of Pennsylvania, PA 19104

⁴ Institute for Translational Medicine and Therapeutics, Perelman School of Medicine, University
of Pennsylvania, Philadelphia, PA 19104

Correspondence to:

Benjamin F. Voight, PhD

Assistant Professor of Systems Pharmacology and Translational Therapeutics

Assistant Professor of Genetics

University of Pennsylvania - Perelman School of Medicine

3400 Civic Center Boulevard

10-126 Smilow Center for Translational Research

Philadelphia, PA 19104

bvoight@upenn.edu

Noname manuscript No.
(will be inserted by the editor)

Toxicity and applications of internalised magnetite nanoparticles within live *Paramecium caudatum* cells

Richard Mayne · James Whiting · Chris Melhuish · Andrew Adamatzky

Received: date / Accepted: date

Abstract The nanotechnology revolution has allowed us to speculate on the possibilities of hybridising nanoscale materials with live substrates, yet significant doubt still remains pertaining to the effects of nanomaterials on biological matter. In this investigation we cultivate the ciliated protistic pond-dwelling microorganism *Paramecium caudatum* in the presence of excessive quantities of magnetite nanoparticles in order to assess both potential beneficial applications for this technique as well as any deleterious effects on the organisms' health. Our findings indicate that these nanoparticles are well-tolerated by paramecia, who were observed to consume in quantities exceeding 10% of their body volume: cultivation in the presence of magnetite nanoparticles does not alter *P. caudatum* cell volume, swim speed, growth rate or peak colony density and cultures may persist in nanoparticle-contaminated medium for many weeks. We demonstrate that *P. caudatum* cells ingest starch coated magnetite nanoparticles which facilitates their being magnetically immobilised whilst maintaining apparently normal ciliary dynamics, thus demonstrating that nanoparticle biohybridisation is a viable alternative to conventional forms of ciliate quieting. Ingested magnetite nanoparticle deposits appear to aggregate, suggesting that (a) the process of being internalised concentrates and therefore detoxifies nanomaterial suspensions in aquatic environments and (b) *P. caudatum* is a candidate organism for programmable nanomaterial manipulation and delivery.

R. Mayne, J. Whiting & A. Adamatzky
Unconventional Computing Laboratory
University of the West of England, Bristol, UK
Tel.: +44 (0)117 32 87861
E-mail: {Richard.Mayne}{James.Whiting}{Andrew.Adamatzky}@uwe.ac.uk

C. Melhuish
Bristol Robotics Laboratory
University of the West of England, Bristol, UK
E-mail: Chris.Melhuish@brl.ac.uk

Harnessing optogenetics to probe sub-cellular mechanics

Patrick W. Oakes^{1,4*}, Elizabeth Wagner^{2*}, Christoph A. Brand³⁺, Dimitri Probst³⁺, Marco Linke³⁺, Ulrich S. Schwarz^{3*}, Michael Glotzer^{2*}, Margaret L. Gardel^{1*}

¹Institute for Biophysical Dynamics, James Franck Institute and the Department of Physics, University of Chicago, Chicago, IL 606037

²Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637

³Institute for Theoretical Physics and BioQuant, Heidelberg University, Heidelberg, Germany

⁴Department of Physics & Astronomy, and Department of Biology, University of Rochester, Rochester, NY 14627

^{*,+}Contributed equally

*Corresponding Authors

Modeling allele-specific gene expression by single-cell RNA sequencing

Yuchao Jiang¹, Nancy R Zhang^{2,*}, Mingyao Li^{3,*}

¹ Genomics and Computational Biology Graduate Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

² Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA

³ Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

* To whom correspondence should be addressed. Tel: (+1) 215-746-3916; Fax: (+1) 215-573-1050; Email: nzh@wharton.upenn.edu, mingyao@mail.med.upenn.edu

Holocene selection for variants associated with cognitive ability:

Comparing ancient and modern genomes.

Michael A. Woodley of Menie^{1,2*}, Shameem Younuskunja³, Bipin Balan⁴, Davide Piffer⁵

¹Scientist in Residence, Technische Universität Chemnitz, Germany

²Center Leo Apostel for Interdisciplinary Studies, Vrije Universiteit Brussel, Belgium

³Weill Cornell Medicine, Cornell University, Qatar

⁴Department of Agriculture and Forestry Science, University of Palermo, Italy.

⁵Department of Psychology, Ben Gurion University of the Negev, Israel

*Email for correspondence: M.A.Woodley@vub.ac.be

Repressing Integrase attachment site operation with CRISPR-Cas9 in *E. coli*

Andrey Shur¹, Richard M. Murray¹

1. Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125.

Abstract

Serine integrases are bacteriophage proteins responsible for integrating the phage genome into that of the host. Synthetic biologists have co-opted these proteins into useful tools for permanent DNA logic, utilizing their specific DNA recombination abilities to build synthetic cell differentiation and genetic memory systems. Each integrase has a specific pair of DNA sequences (attP/attB sites) that it recombines, but multiple identical sites can result in unpredictable recombination. We have developed a way to control integrase activity on identical attP/attB sites by using catalytically dead Cas9 (dCas9) as a programmable binding protein that can compete with integrase for binding to specific attachment sites. Utilizing a plasmid that contains two identical Bxb1 attP sites, integration can be repressed up to 8 fold at either one of the two attP sites when guide RNA and dCas9 are present. Guide RNA sequences that bind specifically to attB, or either of two attP sites, have been developed. Future goals are to utilize this technology to construct larger and more complex integrase logic circuits.



Communication

How to name and classify your phage: an informal guide

Evelien M. Adriaenssens^{1,2,*} and J. Rodney Brister³

¹ Microbiology Research Group, Institute of Integrative Biology, University of Liverpool, UK; evelien.adriaenssens@liv.ac.uk

² Vice Chair of the Bacterial and Archaeal Virus Subcommittee of the International Committee on the Taxonomy of Viruses; evelien.adriaenssens@gmail.com

³ National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20892, USA; jamesbr@ncbi.nlm.nih.gov

* Correspondence: evelien.adriaenssens@gmail.com; Tel.: +44-151-795-4576

Academic Editor: name

Received: date; Accepted: date; Published: date

Abstract: With this informal guide, we try to assist both new and experienced phage researchers through two important stages that follow phage discovery, i.e. naming and classification. Providing an appropriate name for a bacteriophage is not as trivial as it sounds and the effects might be long-lasting in databases and in official taxon names. Phage classification is the responsibility of the Bacterial and Archaeal Viruses Subcommittee (BAVS) of the International Committee on the Taxonomy of Viruses (ICTV). While the BAVS aims at providing a holistic approach to phage taxonomy, for individual researchers who have isolated and sequenced a new phage, this can be a little overwhelming. We are now providing these researchers with an informal guide to phage naming and classification, taking a “bottom-up” approach from the phage isolate level.

Keywords: Bacteriophages; phage taxonomy; phage classification; naming guide; classification guide

1. Introduction

Virus taxonomy is currently the responsibility of the International Committee on the Taxonomy of Viruses (ICTV, [1]), which published its first report in 1971. The Bacterial and Archaeal Viruses Subcommittee (BAVS) within ICTV holds the responsibility of classifying new prokaryotic viruses. New proposals can be submitted year round by the public. These Taxonomy Proposals (TaxoProps) are evaluated by relevant Study Groups (SGs) and the BAVS [2], and are then discussed and voted on by the Executive Committee (EC) during the yearly meeting. All ICTV-accepted proposals are finally ratified by the members of the IUMS (International Union of Microbiological Societies) Virology Division through an email vote.

Bacterial virus taxonomy has undergone a number of changes since the discovery of bacteriophages in the early 20th century. Electron microscopy, which led to the recognition of different phage morphologies, and nucleic acid content provided the basis for the first classification scheme [3,4]. Ever since, genome composition and morphology have been the major criterion for classification at the family rank, with the current taxonomy comprising 22 families grouping bacterial or archaeal viruses.

For many years, the grouping of prokaryotic viruses in lower rank taxa such as genus and subfamily, happened at a minimal pace. Taking the tailed phage families as an example, the 5th Report of ICTV (1991) described one genus in each of the families *Myoviridae*, *Podoviridae* and *Siphoviridae* [5]. This increased to 16 genera spread over the three families by the 7th Report [6] and 18 genera by the 8th Report [7]. As nucleotide sequencing techniques improved the number of publically available

Multiplexed confocal and super-resolution fluorescence imaging of cytoskeletal and neuronal synapse proteins

Syuan-Ming Guo¹, Remi Veneziano¹, Simon Gordonov¹, Li Li², Demian Park³, Anthony B. Kulesa^{1,4}, Paul C. Blainey^{1,4}, Jeffrey R. Cottrell², Edward S. Boyden^{1,3,5}, Mark Bathe^{1,4*}

¹Department of Biological Engineering, MIT, Cambridge, MA, USA

²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA

³Media Lab, MIT, Cambridge, MA, USA

⁴Broad Institute of MIT and Harvard, Cambridge, MA, USA

⁵McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA

*To whom correspondence should be addressed. E-mail: mark.bathe@mit.edu

ABSTRACT

Neuronal synapses contain dozens of protein species whose expression levels and localizations are key determinants of synaptic transmission and plasticity. The spectral properties of fluorophores used in conventional microscopy limit the number of measured proteins to four species within a given sample. The ability to perform high-throughput confocal or super-resolution imaging of many proteins simultaneously without limitation in target number imposed by this spectral limit would enable large-scale characterization of synaptic protein networks in situ. Here, we introduce PRISM: Probe-based Imaging for Sequential Multiplexing, a method that sequentially utilizes either high affinity Locked Nucleic Acid (LNA) or low affinity DNA probes to enable diffraction-limited confocal and PAINT-based super-resolution imaging. High-affinity LNA probes offer high-throughput, confocal-based imaging compared with PAINT, which uses low affinity probes to realize localization-based super-resolution imaging. Simultaneous immunostaining of all targets is performed prior to imaging, followed by sequential LNA/DNA probe exchange that requires only minutes under mild wash conditions. We apply PRISM to quantify the co-expression levels and nanometer-scale organization of one dozen cytoskeletal and synaptic proteins within individual neuronal synapses. Our approach is scalable to dozens of target proteins and is compatible with high-content screening platforms commonly used to interrogate phenotypic changes associated with genetic and drug perturbations in a variety of cell types.

Rapid sequential *in situ* multiplexing with DNA-Exchange-Imaging

Yu Wang,^{1,2,3,9} Johannes B. Woehrstein^{1,2,15}, Noah Donoghue,^{1,3,13} Mingjie Dai,^{1,2,10} Maier S. Avendaño,^{1,2} Ron C.J. Schackmann,⁴ Jason J. Zoeller,⁴ Shan Shan H. Wang,^{5,11} Paul W. Tillberg,^{7,12} Demian Park⁷, Sylvain W. Lapan³, Edward S. Boyden,^{6,7} Joan S. Brugge,⁴ Pascal S. Kaeser,⁵ George M. Church,^{1,3} Sarit S. Agasti,^{1,2,14,*} Ralf Jungmann,^{1,2,15*} Peng Yin.^{1,2,*}

1. Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts, 02115, USA.
 2. Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 3. Department of Genetics, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 4. Department of Cell Biology, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 5. Department of Neurobiology, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 6. Department of Biological Engineering, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, USA.
 7. Media Lab, MIT, Cambridge, Massachusetts, 02139, USA.
 8. Department of Brain and Cognitive Sciences, MIT, Cambridge, Massachusetts, 02139, USA.
 9. Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 10. Program in Biophysics, Harvard University, Boston, Massachusetts, 02138, USA.
 11. Program in Neuroscience, Harvard Medical School, Boston, Massachusetts, 02115, USA.
 12. Department of Electrical Engineering and Computer Science, MIT, Cambridge, Massachusetts 02139, USA.
 13. Warren Alpert Medical School, Brown University, Providence, Rhode Island, 02903, USA.
 14. Present address: New Chemistry Unit and Chemistry & Physics of Materials Unit, Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR), Bangalore, India.
 15. Present address: Department of Physics and Center for Nanoscience, Ludwig Maximilian University, 80539 Munich, Germany, Max Planck Institute of Biochemistry, 82152 Martinsried near Munich, Germany.
- *Emails: P.Y. (py@hms.harvard.edu), R.J. (jungmann@biochem.mpg.de), and S.S.A. (sagasti@jncasr.ac.in)

Abstract

To decipher the molecular mechanism of biological function, it is critical to map the molecular composition of individual cells in the context of their biological environment *in situ*. Immunofluorescence (IF) provides specific labeling for molecular profiling. However, conventional IF methods have finite multiplexing capabilities due to spectral overlap of the fluorophores. Various sequential imaging methods have been developed to circumvent this spectral limit, but are not widely adopted due to the common limitation of requiring multi-rounds of slow (typically over 2 hours at room temperature to overnight at 4 °C in practice) immunostaining. DNA-Exchange-Imaging is a practical platform for rapid *in situ* spectrally-unlimited multiplexing. This technique overcomes speed restrictions by allowing for single-step immunostaining with DNA-barcoded antibodies, followed by rapid (less than 10 minutes) buffer exchange of fluorophore-bearing DNA imager strands. By eliminating the need for multiple rounds of immunostaining, DEI enables rapid spectrally unlimited sequential imaging. The programmability of DNA-Exchange-Imaging allows us to further adapt it to diverse microscopy platforms (with Exchange-Confocal, Exchange-SIM, Exchange-STED, and Exchange-PAINT demonstrated here), achieving highly multiplexed *in situ* protein visualization in diverse samples (including neuronal and tumor cells as well as fresh-frozen or paraffin-embedded tissue sections) and at multiple desired resolution scales (from ~300 nm down to sub-20-nm). Validation highlights include 8-target imaging using single-channel Exchange-Confocal in tens of micron thick retina tissue sections in 2-3 hours (as compared to days required in principle by previous methods using comparable equipment), and 8-target super-resolution imaging with ~20 nm resolution using Exchange-PAINT in primary neurons. These results collectively suggest DNA-Exchange as a versatile, practical platform for rapid, highly

***orco* mutagenesis causes loss of antennal lobe glomeruli and impaired social behavior in ants**

Authors: Waring Tribble,^{1†} Ni-Chen Chang,^{1*} Benjamin J Matthews,^{2,3*} Sean K McKenzie,^{1*} Leonora Olivos-Cisneros,^{1*} Peter R Oxley,^{1*} Jonathan Saragosti,^{1*} Daniel JC Kronauer¹

Affiliations:

¹Laboratory of Social Evolution and Behavior, The Rockefeller University, New York, NY 10065, USA.

²Laboratory of Neurogenetics and Behavior, The Rockefeller University, New York, NY 10065, USA.

³Howard Hughes Medical Institute.

*Co-authors are listed alphabetically and individual contributions are listed in the Acknowledgments section.

†Correspondence to: wtribble@rockefeller.edu

CRISPR/Cas9 screening using unique molecular identifiers

Bernhard Schmierer^{1,#}, Sandeep K. Botla^{1,#}, Jilin Zhang¹, Mikko Turunen², Teemu Kivioja² and Jussi Taipale^{1,2,*}

¹Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden.

²Genome-Scale Biology Research Program, Faculty of Medicine, University of Helsinki, PO Box 63 FI-00014 Helsinki, Finland.

* corresponding author

equal contribution

Loss of function screening by CRISPR/Cas9 gene knockout with pooled, lentiviral guide libraries is a widely applicable method for systematic identification of genes contributing to diverse cellular phenotypes. Here, random sequence labels (RSLs) were incorporated into the guide-library. RSLs function as internal replicates for robust and reproducible hit calling, and act as unique molecular identifiers (UMIs) to allow massively parallel lineage tracing (MPLT) and true dropout screening.

Pooled CRISPR/Cas9 loss of function screening is a powerful approach to identify genes contributing to a wide range of phenotypes. Most commonly, a library of guide sequences is transduced lentivirally into a population of Cas9-expressing cells, which are then subjected to some form of selection pressure. Relative guide frequencies in the genome of the population before and after selection are quantified by next generation sequencing (NGS).

The approach has been applied successfully, but suffers from two major shortcomings: First, the presence of a guide in a cell does not necessarily lead to loss of function of the corresponding gene, as the total read count for a guide reflects cells with distinct genotypes, which are the result of mono- or bi-allelic frameshifts or in-frame deletions¹ (**Supplementary Fig. 1a**), as well as off target effects. These different genotypes also result in a range of phenotypes for the cells, and optimal identification of hit genes ideally requires a method that would individually track each clonal cell lineage derived from a single editing event. Secondly, identification of fitness genes whose guides are under negative selection can be statistically challenging, because of confounders such as random drift or undersampling.

Finally, creating a sufficient number of replicates in this type of experiment is labor intensive and costly. As a consequence, the read count variance in CRISPR/Cas9 pooled screening is commonly estimated globally from a single data set^{2,3}, similar to statistical methods developed for RNA-Seq and ChIP-Seq^{4,5}. This is only valid if the vast majority of guides lack a detectable effect, which might not always be the case, for instance in smaller, targeted libraries. With an insufficient number of replicates, outliers tend to be called as hits and technical artefacts such as PCR bias or other random effects cannot be distinguished from real biological effects.

To address these issues, we have developed a method that allows tracing of individual virus-transduced cell lineages during a CRISPR/Cas9 screen. Depending on the kinetics of genome editing, we can either follow up single clones of identically edited cells, or small populations of sublineages with

A marker-free co-selection strategy for high efficiency human genome engineering

Daniel Agudelo^{1*}, Lusiné Bozoyan^{1,2*}, Alexis Durringer^{1*}, Caroline C. Huard¹, Sophie Carter¹, Jeremy Loehr¹, Dafni Synodinou¹, Mathieu Drouin², Jayme Salsman³, Graham Dellaire³, Josée Laganière², and Yannick Doyon^{1,4}.

¹ Centre Hospitalier Universitaire de Québec Research Center and Faculty of Medicine, Laval University, Quebec City, QC G1V 4G2, Canada.

² Research and Development, Héma-Québec, Quebec City, QC G1V 5C3, Canada.

³ Department of Pathology, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada

*These authors contributed equally to this work

⁴Address correspondence to:

Yannick Doyon, Ph.D.

Centre de recherche du CHU de Québec – Université Laval

2705, boulevard Laurier, T-3-67

Québec, QC G1V 4G2

CANADA

Tel: 418-525-4444 ext. 46264

e-mail: Yannick.Doyon@crchudequebec.ulaval.ca

Intracellular production of hydrogels and synthetic RNA granules by multivalent enhancers

Hideki Nakamura^{1,2*}, Albert A. Lee^{1,2,8*}, Ali Sobhi Afshar^{3‡}, Shigeki Watanabe¹, Elmer Rho², Shiva Razavi^{1,4}, Allison Suarez^{1,2}, Yu-Chun Lin^{1,2}, Makoto Tanigawa^{1,4}, Brian Huang², Robert DeRose^{1,2}, Diana Bobb^{1,2}, William Hong⁵, Sandra B. Gabelli^{5,6,7}, John Goutsias³, Takanari Inoue^{1,2,4†}

¹ Department of Cell Biology, School of Medicine, The Johns Hopkins University, Baltimore, MD, 21205

² Center for Cell Dynamics, Institute for Basic Biomedical Sciences, The Johns Hopkins University, Baltimore, MD, 21205

³ Center for Imaging Science, Whitaker Biomedical Engineering Institute, The Johns Hopkins University, Baltimore, MD, 21218

⁴ Department of Biomedical Engineering, Whitaker Biomedical Engineering Institute, The Johns Hopkins University, Baltimore, MD 21218

⁵ Department of Biophysics and Biophysical Chemistry, School of Medicine, The Johns Hopkins University, Baltimore, MD, 21205

⁶ Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, 21205

⁷ Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, MD, 21205

⁸ Current address: Department of Chemistry, National Taiwan University, Taiwan

*These authors contributed equally.

†To whom general correspondence should be addressed: ictinoue@jhmi.edu (T.I.)

‡ To whom correspondence regarding the computational analysis should be addressed: aas.afshar@gmail.com (A.S.A)

Light Sheet Theta Microscopy for High-resolution Quantitative Imaging of Large Biological Systems

Bianca Migliori^{1,4,#}, Malika S. Datta^{1,#}, Mehmet C. Apak¹ & Raju Tomer^{1,2,3,*}

¹Department of Biological Sciences

²Neurotechnology Center

³Data Science Institute

Columbia University, New York, NY 10027, USA.

⁴Department of Neuroscience, Karolinska Institut, Stockholm, Sweden

[#]Equal contributions

*Correspondence: raju.tomer@columbia.edu

Two-photon calcium imaging of medial prefrontal cortex and hippocampus without cortical invasion

Masashi Kondo^{1,2}, Kenta Kobayashi³, Masamichi Ohkura⁴, Junichi Nakai⁴, and Masanori Matsuzaki^{1,2}

¹Department of Physiology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

²Division of Brain Circuits, National Institute for Basic Biology, Okazaki, Japan

³Section of Viral Vector Development, National Institute for Physiological Sciences, Okazaki, Japan

⁴Brain Science Institute, Saitama University, Saitama, Japan

Corresponding author: Professor Masanori Matsuzaki, Ph.D.

Address: Department of Physiology, Graduate School of Medicine, The University of Tokyo,

7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

Tel: +81-3-5841-3471

Fax: +81-3-5841-3471

E-mail: mzakim@m.u-tokyo.ac.jp

Decoding directional genetic dependencies through orthogonal CRISPR/Cas screens

Michael Boettcher¹, Ruilin Tian², James Blau¹, Evan Markegard³, David Wu¹, Anne Biton⁴, Noah Zaitlen⁴,
Frank McCormick³, Martin Kampmann², Michael T. McManus^{1*}

1. Department of Microbiology and Immunology; Diabetes Center at UCSF, WM Keck Center for Noncoding RNAs, University of California, San Francisco, San Francisco, California, USA.

2. Institute for Neurodegenerative Diseases, Department of Biochemistry and Biophysics, University of California, San Francisco, and Chan-Zuckerberg Biohub, San Francisco, California, USA.

3. Helen Diller Family Comprehensive Cancer Center, Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, California, USA.

4. Department of Medicine, University of California, San Francisco, San Francisco, California, USA.

* Correspondence to: michael.mcmanus@ucsf.edu

Neuronal brain region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric disease heritability

Lindsay F. Rizzardi^{1,2,+}, Peter F. Hickey^{3,+}, Varenka Rodriguez DiBlasi^{1,2}, Rakel Tryggvadóttir¹, Colin M. Callahan¹, Adrian Idrizi¹, Kasper D. Hansen^{1,3,4,*}, Andrew P. Feinberg^{1,2,5,*}

¹Center for Epigenetics, Johns Hopkins University School of Medicine, 855 N. Wolfe St., Baltimore, MD, 21205

²Department of Medicine, Johns Hopkins University School of Medicine

³Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 N. Wolfe St, Baltimore, MD 21205

⁴McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine

⁵Departments of Biomedical Engineering and Mental Health, Johns Hopkins University Schools of Medicine, Engineering, and Public Health

⁺both authors contributed equally to this work

^{*}co-corresponding authors

Anatomical and functional organization of the human substantia nigra and its connections

Yu Zhang¹, Kevin Larcher¹, Bratislav Misic¹, Alain Dagher¹

Affiliations:

¹Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada.

***Corresponding author:**

Alain Dagher MD,

Montreal Neurological Institute, 3801 University St., Montreal, Quebec H3A 2B4, Canada.

alain.dagher@mcgill.ca

Conflict of interest:

The authors declare no competing financial interests.

Acknowledgments

This work was supported by funding from the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council of Canada.

REGULATION OF LIFE SPAN BY THE GUT MICROBIOTA IN THE SHORT-LIVED AFRICAN TURQUOISE KILLIFISH

Patrick Smith^{1,5}, David Willemsen^{1,5}, Miriam Lea Popkes^{1,5}, Franziska Metge¹, Edson
Gandiwa², Martin Reichard³, and Dario Riccardo Valenzano^{1,4,*}

¹ Max Planck Institute for Biology of Ageing, Cologne, Germany

² Chinhoyi University of Technology, Chinhoyi, Zimbabwe

³ Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic

⁴ CECAD, University of Cologne, Cologne, Germany

⁵ These authors equally contributed to the paper

* Correspondence: dvalenzano@age.mpg.de

Motor-Like Properties of Non-Motor Enzymes

David R. Slochower¹ and Michael K. Gilson^{1*}

Affiliations:

¹ Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093-0736, United States.

*Correspondence to: mgilson@ucsd.edu.

Abstract: Biological molecular motors use chemical free energy to drive mechanical motion in a specific direction. This function appears to require high molecular complexity, and it is interesting to consider how the evolutionary leap from non-motor enzymes to molecular motors occurred. Here, atomistic simulations coupled with kinetic modeling show that conformational switching of non-motor enzymes, induced by substrate binding and catalysis, induces motor-like, directional torsional motions, as well as oar-like, reciprocating motions, which should be detectable experimentally. Such directional motions in the earliest enzymes would have been starting point for the evolution of motor proteins. Additionally, driven molecular motions in catalytically active enzymes may help explain why the apparent diffusion constants of some enzymes increase with enzyme velocity (1-3).

One Sentence Summary: Analysis of protein simulations shows that catalytically active non-motor enzymes can execute motor-like motions.

Main Text:

A biological molecular motor is an enzyme that uses the free energy of an out-of-equilibrium chemical reaction to drive mechanical motion. This motion must have a specific direction to fulfill the motor's functional role; for example, a helical flagellum must rotate in the appropriate sense to propel the organism. The ability to generate directional motion may appear to be a complex protein property, so it is interesting to consider how non-motor enzymes could have evolved to molecular motors. Here, we use computational and theoretical methods to test a hypothesis that essentially any enzyme catalyzing an out-of-equilibrium reaction executes directional motions. The results bear on motor evolution, the importance of chirality, and recent experimental observations of enzyme diffusion and motility.

Gamma and beta bursts during working memory read-out suggest roles in its volitional control

Mikael Lundqvist¹, Pawel Herman², Melissa R. Warden^{1,3}, Scott L. Brincat¹, and Earl K. Miller^{1*}

1 The Picower Institute for Learning & Memory and Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology, 43 Vassar Street, Cambridge, MA 02139, USA.

2 Computational Brain Science Lab, Dept. Comp. Sci. & Tech, KTH Royal Institute of Technology, Stockholm, Sweden.

3 Dept. of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853

* Corresponding author

Abstract

Working memory (WM) activity is not as stationary or sustained as previously thought. There are brief bursts of gamma (~55–120 Hz) and beta (~20–35 Hz) oscillations, the former linked to stimulus information in spiking. We examine these dynamics in relation to read-out from WM, which is still not well understood. Monkeys held a sequence of two objects and had to decide if they matched a subsequent sequence. Changes in the balance of beta/gamma suggested their role in WM control. In anticipation of having to use an object for the match decision, there was an increase in spiking information about that object along with an increase in gamma and a decrease in beta. When an object was no longer needed, beta increased and gamma as well as spiking information about that object decreased. Deviations from these dynamics predicted behavioral errors. Thus, turning up or down beta could regulate gamma and the information in working memory.

Introduction

Sustained spiking activity has been the dominant neural model of working memory (WM)¹⁻⁵. The idea is that, once activated by a stimulus, neurons keep spiking, sustaining the representation of that stimulus. Recent closer examinations have revealed that complex dynamics underlie sustained activity, with brief, discrete narrow-band oscillatory bursts in the gamma and beta bands⁶. Gamma bursts (~55–120 Hz) were tied to spiking carrying information about the remembered items. Beta bursts (~20–35 Hz) were associated with suppression of both informative spiking and gamma. These data are consistent with a model in which gamma-associated spiking stores memories by short-term changes in synaptic weights⁷. Multiple items can be held in WM without mutual interference because different gamma bursts store different items.

Multiplex genome editing for synthetic biology in *Vibrio natriegens*

Triana N. Dalia¹, Chelsea A. Hayes¹, Sergey Stolyar², Christopher J. Marx², James B. McKinlay¹, and Ankur B. Dalia^{1,*}

¹Department of Biology, Indiana University, Bloomington, IN 47401. ²Department of Biological Sciences, University of Idaho, Moscow, ID 83844.

*Author for correspondence: Ankur B. Dalia, ankdalia@indiana.edu

Vibrio natriegens has recently emerged as an alternative to *Escherichia coli* for molecular biology and biotechnology, but low-efficiency genetic tools hamper its development. Here, we uncover how to induce natural competence in *V. natriegens* and describe methods for multiplex genome editing by natural transformation (MuGENT). MuGENT promotes integration of large genome edits at high-efficiency on unprecedented timescales, which will extend the utility of this species for diverse applications.

V. natriegens is the fastest growing organism known, with a doubling time of <10 min^{1,2}. With broad metabolic capabilities, lack of pathogenicity, and its rapid growth rate, it is an attractive alternative to *E. coli* for diverse molecular biology and biotechnology applications³. Methods for classical genetic techniques have been developed for *V. natriegens*, but these are relatively laborious, require multiple steps, and must be used sequentially to generate multiple genome edits³. The challenges of these techniques contrast with the ease of genetics in *Vibrio* species that are naturally transformable. Competent *Vibrios* can take up DNA from the environment and integrate it into their genome by homologous recombination; processes known as natural competence and natural transformation, respectively⁴⁻⁷. The inducing cue for natural transformation in competent *Vibrios* is growth on the chitinous shells of crustacean zooplankton, which are commonly found in the aquatic environment where these microbes reside⁴. Chitin induces

The 100 € lab: A 3-D printable open source platform for fluorescence microscopy, optogenetics and accurate temperature control during behaviour of zebrafish, *Drosophila* and *C. elegans*.

Andre Maia Chagas^{1-3,5,7\$}, Lucia Prieto Godino^{3,4}, Aristides B. Arrenberg^{1,6,7}, Tom Baden^{1,3,5,7,8\$}

1: Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 2: Graduate school for Neural and Behavioural Neuroscience, 3: TReND in Africa gUG, 4: CIG, University of Lausanne, 5: Institute of Ophthalmic Research, 6: Institute of Neurobiology, 7: University of Tübingen, Germany, 8: School of Life Sciences, University of Sussex, Brighton, UK.

\$: Correspondence at andremaia.chagas@gmail.com and t.baden@sussex.ac.uk

36 pages

6787 Words (Summary, Main text and Main Figure Legends)

7 Colour Figures

1 Supplementary Figure

1 Supplementary Table

1 Supplementary Assembly Manual

10 Supplementary videos

Hypocretin underlies the evolution of sleep loss in the Mexican cavefish

James B. Jaggard¹, Bethany A. Stahl¹, Evan Lloyd¹, Erik R. Duboue² and Alex C. Keene^{1*}

1. Department of Biological Sciences, Florida Atlantic University, Jupiter, FL 33458

2. Department of Embryology, Carnegie Institution for Science, Baltimore, MD, 21218

***Address Correspondence to:**

Alex C. Keene
Department of Biological Sciences
Florida Atlantic University
5353 Parkside Drive
Jupiter, FL 33458, USA
Email: KeeneA@FAU.edu

Local Connectome Phenotypes Predict Social, Health, and Cognitive Factors

Michael A. Powell

Department of Mathematical Sciences, United States Military Academy, West Point, NY

Javier O. Garcia

U.S. Army Research Laboratory, Aberdeen Proving Ground, MD

Department of Bioengineering, University of Pennsylvania, Philadelphia, PA

Fang-Cheng Yeh

Department of Neurological Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA

Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA

Jean M. Vettel

U.S. Army Research Laboratory, Aberdeen Proving Ground, MD

Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA

Department of Bioengineering, University of Pennsylvania, Philadelphia, PA

Timothy Verstynen

Department of Psychology & Center for the Neural Basis of Cognition, Carnegie Mellon

University, Pittsburgh, PA

Corresponding Author:

Name: Timothy Verstynen

Address: Department of Psychology, 342C Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213

Telephone: 412-533-2961

Fax:

Short Title: Local Connectome Phenotypes

Keywords: local connectome, white matter, individual differences, behavior prediction, structural connectivity

Acknowledgements: The research was sponsored by the Army Research Laboratory and accomplished under Cooperative Agreement Number W911NF-10-2-0022. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

A Systematic Nomenclature for the *Drosophila* Ventral Nervous System

Robert Court^{1*†¶§}, Douglas Armstrong^{1‡}, Jana Börner^{4‡}, Gwyneth Card^{2‡§}, Marta Costa^{5‡}, Michael Dickinson^{6‡}, Carsten Duch^{7‡}, Wyatt Korff^{2‡§}, Richard Mann^{8‡}, David Merritt^{9‡}, Rod Murphey^{4‡}, Shigehiro Namiki^{2‡§}, Andrew Seeds^{2‡}, David Shepherd^{3*†**§}, Troy Shirangi^{2‡}, Julie Simpson^{2‡}, James Truman^{2‡}, John Tuthill^{10‡}, Darren Williams^{11‡}

*For correspondence:

rcourt@ed.ac.uk (FMS);
d.shepherd@bangor.ac.uk (FS)

[†]RC worked under supervision of DS to produce this work

[‡]These authors helped organise the working group

[§]These authors contributed as part of the working group

Present address: [¶]Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, UK;
^{**}School of Biological Sciences, Bangor University, UK

¹School of Informatics, U. of Edinburgh; ²HHMI- Janelia Research Campus; ³School of Biol. Sci, Bangor University; ⁴Biological Sciences, Florida Atlantic University; ⁵Dept. of Genetics, U. of Cambridge; ⁶Dept. of Biology, U. of Washington; ⁷Univ. of Mainz; ⁸Biochemistry and Molecular Biophysics, Columbia University; ⁹School of Biological Sciences, The University of Queensland; ¹⁰Department of Neurobiology, Harvard Medical School; ¹¹MRC Centre for Developmental Neurobiology, King's College London

Abstract

Insect nervous systems are proven and powerful model systems for neuroscience research with wide relevance in biology and medicine. However, descriptions of insect brains have suffered from a lack of a complete and uniform nomenclature. Recognising this problem the Insect Brain Name Working Group produced the first agreed hierarchical nomenclature system for the adult insect brain, using *Drosophila melanogaster* as the reference framework, with other insect taxa considered to ensure greater consistency and expandability (Ito et al., 2014). Ito et al. (2014) purposely focused on the gnathal regions that account for approximately 50% of the adult CNS. We extend this nomenclature system to the sub-gnathal regions of the adult *Drosophila* nervous system to provide a nomenclature of the so-called ventral nervous system (VNS), which includes the thoracic and abdominal neuromeres that was not included in the original work and contains the neurons that play critical roles underpinning most fly behaviours.

Background

Insect nervous systems are proven and powerful model systems for neuroscience research with wide relevance for biology and medicine. Although vast anatomical, physiological and molecular data are already available, integrating this information into a common analytical framework would generate an even more powerful resource. Computational analysis combined with digital microscopy now make it possible to consolidate data from multiple techniques and transform how we analyse nervous system function (Jenett et al., 2006; Dance, 2015; Boettiger et al., 2016). It is no longer sufficient to use 2D labelled diagrams and photomicrographs to identify and define anatomical structures, as it is now possible to use multilayer microscopy with computational reconstruction to precisely define and allocate boundaries and structures in 3D. This requires a systematic and consistent nomenclature to precisely define anatomical structures and boundaries. Furthermore the definitions and nomenclature need to be rationalised as multiple names can be used for the same structure. The precise inter relationships between structures also need to be specified. Once this is complete, new findings can be more easily added to this framework allowing

Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization

Brock Roberts^{*†}, Amanda Haupt^{*}, Andrew Tucker, Tanya Grancharova, Joy Arakaki, Margaret A. Fuqua, Angelique Nelson, Caroline Hookway, Susan A. Ludmann, Irina A. Mueller, Ruian Yang, Alan R. Horwitz, Susanne M. Rafelski, and Ruwanthi N. Gunawardane[†]

Allen Institute for Cell Science, 615 Westlake Ave North, Seattle, WA 98109

[†] Address correspondence to: Brock Roberts (brockr@alleninstitute.org), Ruwanthi Gunawardane (rug@alleninstitute.org)

^{*} These authors contributed equally

Running title: Methods for endogenous FP tagging in stem cells

Keywords: CRISPR/Cas9, genome editing, GFP, live imaging

Number of characters: 61,613

1 Rapid and Programmable Protein Mutagenesis Using Plasmid

2 Recombineering

3 Sean A. Higgins¹, Sorel Ouonkap¹, David F. Savage^{1,2,*}

4

5 ¹ Department of Molecular and Cell Biology, UC Berkeley, Berkeley, CA, 94720, USA

6 ² Department of Chemistry, UC Berkeley, Berkeley, CA, 94720, USA

7

8 * To whom correspondence should be addressed. Email: savage@berkeley.edu.

9 Address: 2151 Berkeley Way, Berkeley, CA 94720; (510) 643-7847

10

11

12

13

Covalent protein labeling by SpyTag-SpyCatcher in fixed cells for super-resolution microscopy

Veronica Pessino¹, Rose Citron¹, Siyu Feng², Bo Huang^{3,4,*}

¹Graduate Program of Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA

²The UC Berkeley-UCSF Graduate Program in Bioengineering, University of California, San Francisco, San Francisco, CA 94143, USA

³Department Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA

⁴Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA

* Correspondence should be addressed to bo.huang@ucsf.edu.

Integrated computational guide design, execution, and analysis of arrayed and pooled CRISPR genome editing experiments

Matthew C. Canver^{1,*}, Maximilian Haeussler^{2,*}, Daniel E. Bauer¹, Stuart H. Orkin^{1,3}, Neville E. Sanjana⁴, Ophir Shalem⁵, Guo-Cheng Yuan⁶, Feng Zhang^{7,8}, Jean-Paul Concordet⁹, Luca Pinello^{10,#}

¹Division of Hematology/Oncology, Boston Children's Hospital, Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Stem Cell Institute, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA

²Santa Cruz Genomics Institute, MS CBSE, University of California, Santa Cruz, California, USA

³Howard Hughes Medical Institute, Boston, Massachusetts, USA

⁴New York Genome Center and Department of Biology, New York University, New York City, New York, USA

⁵Raymond G. Perelman Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Department of Genetics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

⁶Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

⁷The Broad Institute, Cambridge, Massachusetts, USA.

⁸McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

⁹INSERM U1154, CNRS UMR 7196, Muséum National d'Histoire Naturelle, Paris, France.

¹⁰Department of Molecular Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

*Co-first authors

#Correspondence to Luca Pinello, lpinello@mgh.harvard.edu

Keywords: CRISPR genome editing; Cas9/Cpf1; CRISPR pooled screening; Saturating mutagenesis; sgRNA design; Lentivirus; CRISPOR; CRISPResso; Docker; Locus-specific deep sequencing; Off-target effect; On-target effect; Amplicon sequencing; Targeted sequencing; Sequence analysis; Computational tools; Software pipeline

The Developing Human Connectome Project: a Minimal Processing Pipeline for Neonatal Cortical Surface Reconstruction

Antonios Makropoulos^{a,1,*}, Emma C. Robinson^{a,1}, Andreas Schuh^a, Robert Wright^b, Sean Fitzgibbon^c, Jelena Bozek^d, Serena J. Counsell^b, Johannes Steinweg^b, Jonathan Passerat-Palmbach^a, Gregor Lenz^a, Filippo Mortari^a, Tencho Tenev^a, Eugene P. Duff^c, Matteo Bastiani^c, Lucilio Cordero-Grande^b, Emer Hughes^b, Nora Tusor^b, Jacques-Donald Tournier^b, Jana Hutter^b, Anthony N. Price^b, Maria Murgasova^b, Christopher Kelly^b, Mary A. Rutherford^b, Stephen M. Smith^c, A. David Edwards^b, Joseph V. Hajnal^b, Mark Jenkinson^c, Daniel Rueckert^a

^a*Biomedical Image Analysis Group, Department of Computing, Imperial College London, London, United Kingdom.*

^b*Centre for the Developing Brain, Division of Imaging Sciences and Biomedical Engineering, King's College London, London, United Kingdom.*

^c*FMRIB Centre, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom.*

^d*Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia.*

Abstract

The Developing Human Connectome Project (dHCP) seeks to create the first 4-dimensional connectome of early life. Understanding this connectome in detail may provide insights into normal as well as abnormal patterns of brain development. Following established best practices adopted by the WU-MINN Human Connectome Project and pioneered by FreeSurfer, the project utilises cortical surface-based processing pipelines. In this paper, we propose a fully automated processing pipeline for the structural Magnetic Resonance Imaging (MRI) of the developing neonatal brain. This proposed pipeline consists of a refined framework for cortical and sub-cortical volume segmentation, cortical surface extraction and cortical surface inflation of neonatal subjects, which has been specifically designed to address considerable differences between adult and neonatal brains, as imaged using MRI. Using the proposed pipeline our results demonstrate that images collected from

Recording action potential propagation in single axons using multi-electrode arrays

Kenneth R. Tovar¹, Daniel C. Bridges^{1, 3}, Bian Wu¹, Connor Randall³, Morgane Audouard^{1, 2}, Jiwon Jang^{1, 2}, Paul K. Hansma^{1, 3}, Kenneth S. Kosik^{1, 2}

¹Neuroscience Research Institute, the ²Department of Molecular, Cellular and Developmental Biology and the ³Department of Physics
University of California, Santa Barbara
Santa Barbara, CA 93106

Running Title: Non-invasive, multi-site recording from axons

Correspondence:

Kenneth R. Tovar
Neuroscience Research Institute
UC Santa Barbara
Santa Barbara, CA 93016
ken.tovar@lifesci.ucsb.edu

7 Figures
no tables

Acknowledgements:

This research was sponsored by the U.S. Army Research Laboratory and Defense Advanced Research Projects Agency under Cooperative Agreement Number W911NF-15-2-0056. The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. Additional support was also provided by the California NanoSystems Institute (CNSI). We thank Bridget N. Queenan and Carol A. Vandenburg for their thoughtful reading and insightful comments on this manuscript.

High Accuracy Base Calls in Nanopore Sequencing

Philippe Faucon, Robert Trevino
School of Computing, Informatics,
and Decision Systems Engineering
Arizona State University
Tempe, AZ 85282, USA
{pfaucon, rptrevin}@asu.edu

Parithi Balachandran, Kylie Standage-Beier, and Xiao Wang
School of Biological
and Health Systems Engineering
Arizona State University
Tempe, AZ 85287-9709, USA
{pbalach1, ksstanda, xiaowang}@asu.edu

Abstract—Nanopore sequencing has introduced the ability to sequence long stretches of DNA, enabling the resolution of repeating segments, or paired SNPs across long stretches of DNA. Unfortunately significant error rates $>15\%$, introduced through systematic and random noise inhibit downstream analysis. We propose a novel method, using unsupervised learning, to correct biologically amplified reads before downstream analysis proceeds. We also demonstrate that our method has performance comparable to existing techniques without limiting the detection of repeats, or the length of the input sequence.

I. INTRODUCTION

DNA sequencing has become a critical part of most biological research, in tasks ranging from gene network identification, to biological engineering, to organism identification and the generation of phylogenies. Most of these applications have 2 primary foci: the length of DNA sequence reads, and the per-base error in those reads. Third generation sequencing technologies presented by Oxford Nanopore Technologies(ONT) and PacBio offer read lengths 10-100x what was possible with previous sequencing technologies, but at a per-base read accuracy near 85%, down from 99.99% using second generation technologies. While the increased read length enables many new biological applications [1], [2] the low read accuracy hinders others. Extensive research has gone into developing techniques that can robustly improve Nanopore read accuracy and analyzing the trade-offs [3]–[5].

We propose an approach to correcting errors that does not make assumptions about the presence of a reference genome and is robust to mixed biological populations where reads with significant overlap must be identified as different. To accomplish this task multiple copies of the same original DNA sequence must be read, resulting in a large number reads, of which a small number are high-error copies of one another. We then use the K-means algorithm to cluster reads based on an engineered feature space. This technique does not require replicated reads to be biologically attached [6], meaning that the individual sequence lengths are not constrained. Additionally our approach groups reads with their own replicates, removing the need for a reference genome. Our contributions are summarized below:

- We demonstrate that while some base calling error is random there is a significant component that is systematic and can be modeled. in section mbox IV-A:mbox we

present a simple model for predicting easily identifiable k-mers and show that our model correlates well with data found empirically.

- We propose a feature space based on the easily identifiable k-mers that accurately groups copies of unique DNA strands. In section mbox V-A:mbox , we demonstrate comparable clustering performance to MinHash [7] with a number of simulated reads similar to biological experiments.
- We demonstrate that a simple K-means approach can replace a more complex biological process to group unique DNA strands together.
- The proposed method allows for the analysis of larger DNA strands compared to INC-Seq [6].

II. RELATED WORKS

Given the importance of sequence read accuracy, a variety of methods have been proposed for increasing read accuracy. These approaches centered around either detecting overlapping regions in genomic alignments to provide read corrections [4], [8], [9], or correcting reads before alignment [3], [6], [10], [11].

Read overlapping [9], and correction from genomic alignments [4], [8] has shown significant potential for read correction. These techniques have shown the ability to increase sequence accuracy into the high 90%'s, with relatively minimal requirements [4]. One major shortcoming is that they are limited by error rate and genomic similarity, with large genomes or high error rates resulting in excessive correction times or erroneous results [8]. Some of these shortcomings can be softened by the presence of a reasonable reference genome, as the error rate for read-genome alignments is approximately half of that seen in read-read alignments.

Read pre-correction has gained significant steam recently, and particular advantages have been seen with small genomes which contain more than 10x even within a single sequencing run. Early attempts at pre-correction focused on aligning high accuracy second generation reads to long nanopore reads [3], [11]. This provides the possibility to resolve many features but introduces systematic error within repeating regions of the reads. More recent attempts have focused on nanopore reads exclusively [6], trading a significantly reduced read length for increased read accuracy. To a large extent both of these

Modulation of Genome Editing Outcomes by Cell Cycle Control of Cas9 Expression

Yuping Huang, Caitlin McCann, Andrey Samsonov, Dmitry Malkov, Gregory D. Davis*,
Qingzhou Ji*

MilliporeSigma, St. Louis, USA

* Corresponding authors at: MilliporeSigma, 2909 Laclede Ave., St. Louis, MO 63103, USA. Tel.: +1 314 289 8496.

E-mail addresses: qingzhou.ji@sial.com (Q.Ji), greg.davis@sial.com (G. Davis)

Enhancing Multiplex Genome Editing by Natural Transformation (MuGENT) via inactivation of ssDNA exonucleases

Triana N. Dalia¹, Soo H. Yoon², Elisa Galli³, Francois-Xavier Barre³, Christopher M. Waters², and Ankur B. Dalia^{1,*}

¹Department of Biology, Indiana University, Bloomington, IN USA. ²Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, USA.

³Institute for Integrative Biology of the Cell (I2BC), Université Paris-Saclay, CEA, CNRS, Université Paris Sud, Gif sur Yvette, France.

*Author for correspondence – Ankur B. Dalia, ankdalia@indiana.edu

ABSTRACT

Recently, we described a method for multiplex genome editing by natural transformation (MuGENT). Mutant constructs for MuGENT require large arms of homology (>2000 bp) surrounding each genome edit, which necessitates laborious *in vitro* DNA splicing. In *Vibrio cholerae*, we uncover that this requirement is due to cytoplasmic ssDNA exonucleases, which inhibit natural transformation. In ssDNA exonuclease mutants, one arm of homology can be reduced to as little as 40 bp while still promoting integration of genome edits at rates of ~50% without selection *in cis*. Consequently, editing constructs are generated in a single PCR reaction where one homology arm is oligonucleotide encoded. To further enhance editing efficiencies, we also developed a strain for transient inactivation of the mismatch repair system. As a proof-of-concept, we used these advances to rapidly mutate 10 high-affinity binding sites for the nucleoid occlusion protein SlmA and generated a duodecuple mutant of 12 diguanylate cyclases in *V. cholerae*. Whole genome sequencing revealed little to no off-target mutations in these strains. Finally, we show that ssDNA exonucleases inhibit natural transformation in *Acinetobacter baylyi*. Thus, rational removal of ssDNA exonucleases may be broadly applicable for enhancing the efficacy and ease of MuGENT in diverse naturally transformable species.

A general method to fine-tune fluorophores for live-cell and *in vivo* imaging

Jonathan B. Grimm, Anand K. Muthusamy, Yajie Liang, Timothy A. Brown, William C. Lemon, Ronak Patel, Rongwen Lu, John J. Macklin, Phillip J. Keller, Na Ji, and Luke D. Lavis*

Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

*Corresponding author: lavisl@janelia.hhmi.org

Nanopore sequencing and assembly of a human genome with ultra-long reads

M Jain^{1,§}, S Koren^{2,§}, J Quick^{3,§}, AC Rand^{1,§}, TA Sasani^{4,5,§}, JR Tyson^{7,§}, AD Beggs⁸, AT Dilthey², IT Fiddes¹, S Malla⁹, H Marriott⁹, KH Miga¹, T Nieto⁸, J O'Grady¹⁰, HE Olsen¹, BS Pedersen^{4,5}, A Rhie², H Richardson¹⁰, AR Quinlan^{4,5,6}, TP Snutch⁷, L Tee⁸, B Paten¹, AM Phillippy², JT Simpson^{11,12}, NJ Loman^{3,*}, M Loose^{9,*}

§ These authors contributed equally to this work.

* Authors for correspondence n.j.loman@bham.ac.uk, matt.loose@nottingham.ac.uk

Affiliations:

1. UC Santa Cruz Genomics Institute, University of California, Santa Cruz, CA, USA
2. Genome Informatics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, Bethesda, Maryland, USA
3. Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, UK
4. Department of Human Genetics, University of Utah, Salt Lake City, UT, USA
5. USTAR Center for Genetic Discovery, University of Utah, Salt Lake City, UT, USA
6. Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, USA
7. Michael Smith Laboratories and Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, Canada
8. Surgical Research Laboratory, Institute of Cancer & Genomic Science, University of Birmingham, UK
9. DeepSeq, School of Life Sciences, University of Nottingham, UK
10. Norwich Medical School, University of East Anglia, Norwich, UK
11. Ontario Institute for Cancer Research, Toronto M5G 0A3, Canada
12. Department of Computer Science, University of Toronto, Toronto M5S 3G4, Canada

Contributions:

N.J.L, M.L, J.T.S, and J.R.T conceived the study, J.Q developed the long read protocol, A.D.B, M.J, M.L, H.M, S.M, T.N, J.O'G, J.Q, H.R, J.R.T and L.T prepared materials and/or performed sequencing, A.T.D, I.T.F, M.J, S.K, N.J.L, M.L, K.H.M, H.E.O, B.P, B.S.P, A.M.P, A.R.Q, A.C.R, A.R, T.A.S, J.T.S and J.R.T performed bioinformatics analysis and wrote or modified software, I.T.F, M.J, S.K, N.J.L, M.L, K.H.M, J.O'G, H.E.O, B.P, A.M.P, J.Q, A.R.Q, A.C.R, T.A.S, J.T.S, T.P.S and J.R.T wrote and edited the manuscript. All authors approved the manuscript and provided strategic oversight for the work.

Recombineering in *Vibrio natriegens*

Henry H. Lee^{1*}, Nili Ostrov^{1*}, Michaela A. Gold², George M. Church^{1,3}

¹Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

²Department of Biology, Chemical and Biological Engineering, Tufts University, Medford, Massachusetts, 02155.

³Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115, USA.

*These authors contributed equally to this work.

Abstract

Here, we show that λ -Red homologs found in the *Vibrio*-associated SXT mobile element potentiate allelic exchange in *V. natriegens* by ~10,000-fold. Specifically, we show SXT-Beta (s065), SXT-Exo (s066), and λ -Gam proteins are sufficient to enable recombination of single- and double-stranded DNA with episomal and genomic loci. We characterize and optimize episomal oligonucleotide-mediated recombineering and demonstrate recombineering at genomic loci. We further show targeted genomic deletion of the extracellular nuclease gene *dns* using a double-stranded DNA cassette. Continued development of this recombination technology will advance high-throughput and large-scale genetic engineering efforts to domesticate *V. natriegens* and to investigate its rapid growth rate.

Main text

We have previously proposed the fast growing marine bacterium *Vibrio natriegens* as a powerful bacterial host and reported foundational genomics resources for its utilization (1). While methods for genome modification of *V. natriegens* by homologous recombination have recently been reported, they are laborious and protract experimental time. Plasmid-based integration methods require extensive cloning, conjugation, and a strong negative selection for elimination of the plasmid backbone (2). Similarly, recombination of double-stranded DNA cassettes by natural competence, though attractive due to its efficiency, requires cloning of unwieldy homology arms up to 3 kb and extended incubation times (3). Development of one-step recombineering method which tolerates short homology arms, particularly with oligonucleotides, would be an attractive advancement for genomic manipulation of *V. natriegens*.

Recombineering is a powerful method for precise DNA editing, enabling *in vivo* construction of mutant alleles and structural changes such as insertions and deletion of genes (4–9). These mutations can be introduced by allelic exchange between the target sequence and recombinant single- or double-stranded DNA, potentiated by expression of powerful homologous recombination (HR) proteins found in bacteriophages (5). λ -Beta, the most well-studied phage recombinase, has been shown to enhance HR in *E. coli* by ~10,000-fold over the basal mutation rate (10). Unfortunately, λ -phage Red Beta does not sustain this efficiency in diverse bacteria

A practical guide to CRISPR/Cas9 genome editing in Lepidoptera

Linlin Zhang and Robert D. Reed

Department of Ecology and Evolutionary Biology, Cornell University, 215 Tower Rd., Ithaca, NY 14853-7202, USA

Corresponding Authors: lz355@cornell.edu, robertreed@cornell.edu

Abstract

CRISPR/Cas9 genome editing has revolutionized functional genetic work in many organisms and is having an especially strong impact in emerging model systems. Here we summarize recent advances in applying CRISPR/Cas9 methods in Lepidoptera, with a focus on providing practical advice on the entire process of genome editing from experimental design through to genotyping. We also describe successful targeted GFP knock-ins that we have achieved in butterflies. Finally, we provide a complete, detailed protocol for producing targeted long deletions in butterflies.

Introduction

The order Lepidoptera represents a tenth of the world's described species and contains many taxa of scientific and agricultural importance. Despite major interest in this group, however, there has been a frustrating lack of progress in developing routine approaches for manipulative genetic work. While the last decade has seen examples of transgenesis and targeted knockouts using methods like transposon insertion¹, zinc-finger nucleases^{2,3}, and TALENs^{4,5}, especially in the silk moth *Bombyx mori*, these approaches have resisted widespread application due to their laborious nature. We see two other main reasons manipulative genetics has failed to become routine in Lepidoptera. The first is that many lepidopterans are sensitive to inbreeding, and in some species it can be difficult to maintain experimental lines without special effort. The second is that lepidopterans appear to have an unusual resistance to RNAi^{6,7}, a method that has dramatically accelerated work in many other groups of insects.

Classification: Biological Sciences, Neuroscience

Arc restores juvenile plasticity in adult mouse visual cortex

Short title: Arc restores juvenile plasticity in adult mice

Kyle R. Jenks^{1,a}, Taekeun Kim^{1,b}, Elissa D. Pastuzyn^{1,a}, Hiroyuki Okuno^{c,d},
Andrew V. Taibi^a, Haruhiko Bito^d, Mark F. Bear^{2,b}, and Jason D. Shepherd^{2,a}

a. Department of Neurobiology and Anatomy, University of Utah, Salt Lake City
Utah, 84112. b. The Picower Institute for Learning and Memory, Massachusetts
Institute of Technology, Cambridge Massachusetts, 02139. c. Medical Innovation
Center, Kyoto University Graduate School of Medicine,
Sakyo-ku, Kyoto 606-8507, Japan. d. Department of Neurochemistry, Graduate
School of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo
113-0033, Japan.

¹Contributed equally

²Correspondence, mbear@mit.edu and Jason.Shepherd@neuro.utah.edu

Keywords: Arc, plasticity, visual cortex, ocular dominance plasticity,
amblyopia, critical period

Title:

Fused dorsal-ventral cerebral organoids model human cortical interneuron migration.

Authors:

Joshua A Bagley¹, Daniel Reumann¹, Shan Bian¹, Juergen A Knoblich^{1*}

Affiliations:

¹Institute of Molecular Biotechnology of the Austrian Academy of Science (IMBA), Vienna
1030, Austria

*corresponding author

Rapid whole brain imaging of neural activities in freely behaving larval zebrafish

Lin Cong^{1,4}, Zeguan Wang^{2,4}, Yuming Chai^{2,4}, Wei Hang^{1,4}, Chunfeng Shang¹, Wenbin Yang², Lu Bai¹, Jiulin Du¹, Kai Wang^{1,3}, Quan Wen²

1. Institute of Neuroscience, State Key Laboratory of Neuroscience, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China.

2. Chinese Academy of Sciences Key Laboratory of Brain Function and Disease, School of Life Sciences, Center for Excellence in Brain Science and Intelligence Technology, University of Science and Technology of China, Hefei, 230027, China.

3. University of Chinese Academy of Sciences, Beijing 100049, China.

4. These authors contributed equally to this work.

Correspondence should be addressed to K.W. (wangkai@ion.ac.cn) or Q.W. (qwen@ustc.edu.cn).

A high resolution whole brain imaging using Oblique Light Sheet Tomography

Arun Narasimhan, Judith Mizrachi, Kannan Umadevi Venkatraju, Dinu F. Albeanu,

Pavel Osten

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724.

albeanu@cshl.edu, osten@cshl.edu

Deep Neural Networks in Computational Neuroscience

Tim C Kietzmann¹, Patrick McClure¹, & Nikolaus Kriegeskorte^{1,2}

¹ MRC Cognition and Brain Science Unit, Cambridge, UK

² Department of Psychology, Columbia University

Keywords: deep neural networks, deep learning, convolutional neural networks, objective functions, recurrence, black box, levels of abstraction, modelling the brain, input statistics, biological detail

Summary

The goal of computational neuroscience is to find mechanistic explanations of how the nervous system processes information to support cognitive function and behaviour. At the heart of the field are its models, i.e. mathematical and computational descriptions of the system being studied. These models typically map sensory stimuli to neural responses and/or neural to behavioural responses and range for simple to complex. Recently, deep neural networks (DNNs), using either feedforward and recurrent architectures, have come to dominate several domains of artificial intelligence (AI). As the term “neural network” suggests, these models are inspired by biological brains. However, current DNN models abstract from many details of biological neural networks. Their abstractions contribute to their computational efficiency, enabling to perform complex feats of intelligence, ranging from perceptual tasks (e.g. visual object and auditory speech recognition) to cognitive tasks (e.g. machine translation), and on to motor control tasks (e.g. playing computer games or controlling a robot arm). In addition to their ability to model complex intelligent behaviours, DNNs have been shown to predict neural responses to novel sensory stimuli that cannot be predicted with any other currently available type of model. DNNs can have millions of parameters (connection strengths), which are required to capture the domain knowledge needed for task performance. These parameters are often set by task training using stochastic gradient descent. The computational properties of the units are the result of four directly manipulable elements: input statistics, network structure, functional objective, and learning algorithm. The advances with neural nets in engineering provide the technological basis for building task-performing models of varying degrees of biological realism that promise substantial insights for computational neuroscience.

BioRxiv Preprint

A Survey of Genome Editing Activity for 16 Cpf1 orthologs

Bernd Zetsche^{1-5,8}, Jonathan Strecker^{1-4,8}, Omar O. Abudayyeh^{1-4,6}, Jonathan S. Gootenberg^{1-4,7}, David A. Scott^{1,2}, Feng Zhang^{1-4*}

¹Broad Institute of MIT and Harvard
Cambridge, MA 02142, USA

²McGovern Institute for Brain Research at MIT,
Cambridge, MA 02139, USA

³Department of Brain and Cognitive Science

⁴Department of Biological Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

⁵Department of Developmental Pathology
Institute of Pathology
Bonn Medical School, Bonn, Germany

⁶Harvard-MIT Division of Health Science and Technology
Cambridge, Massachusetts, USA

⁷Department of Systems Biology Harvard Medical School
Boston, Massachusetts, USA

⁸These authors contributed equally to this work.

*Correspondence should be addressed to zhang@broadinstitute.org (F.Z.).

Whole-brain serial-section electron microscopy in larval zebrafish

David Grant Colburn Hildebrand^{1,2,3,4,5,†,*}, Marcelo Cicconet⁵, Russel Miguel Torres^{2,4,†}, Woohyuk Choi⁶, Tran Minh Quan⁶, Jungmin Moon⁶, Arthur Willis Wetzel⁷, Andrew Scott Champion⁸, Brett Jesse Graham⁴, Owen Randlett^{2,†}, George Scott Plummer^{2,†}, Ruben Portugues^{2,†}, Isaac Henry Bianco^{2,†}, Stephan Saalfeld⁸, Alex Baden⁹, Kunal Lillaney⁹, Randal Burns⁹, Joshua Tzvi Vogelstein¹⁰, Alexander Franz Schier^{2,3,11,12,13}, Wei-Chung Allen Lee⁴, Won-Ki Jeong⁶, Jeff William Lichtman^{2,3,14}, Florian Engert^{2,3,14,*}

¹ Graduate Program in Neuroscience, Division of Medical Sciences, Graduate School of Arts and Sciences, Harvard University, Cambridge, Massachusetts, United States of America

² Department of Molecular and Cell Biology, Harvard University, Cambridge, Massachusetts, United States of America

³ Center for Brain Science, Harvard University, Cambridge, Massachusetts, United States of America

⁴ Department of Neurobiology, Harvard Medical School, Boston, Massachusetts, United States of America

⁵ Image and Data Analysis Core, Harvard Medical School, Boston, Massachusetts, United States of America

⁶ School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology (UNIST), Republic of Korea

⁷ Pittsburgh Supercomputing Center, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America

⁸ Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, United States of America

⁹ Department of Computer Science, Johns Hopkins University, Baltimore, Maryland, United States of America

¹⁰ Department of Biomedical Engineering and Institute for Computational Medicine, Johns Hopkins University, Baltimore, Maryland, United States of America

¹¹ Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America

¹² Harvard Stem Cell Institute, Harvard University, Cambridge, Massachusetts, United States of America

¹³ FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts, United States of America

¹⁴ These authors contributed equally to this work.

*Corresponding authors:

D.G.C.H.: e-mail: david@hildebrand.name; telephone: +1 (857) 600-1770

F.A.E.: e-mail: florian@mcb.harvard.edu; telephone: +1 (617) 384-9773

†Present addresses:

D.G.C.H.: Laboratory of Neural Systems, Rockefeller University, New York, New York, United States of America

R.M.T.: Allen Institute for Brain Science, Seattle, Washington, United States of America

G.S.P.: Tufts University School of Medicine, Boston, Massachusetts, United States of America

O.R.: Department of Cell and Developmental Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, United States of America

R.P.: Max Planck Institute of Neurobiology, Martinsried, Germany

I.H.B.: Department of Neuroscience, Physiology, and Pharmacology, University College London, London, United Kingdom

Title: Intercellular mRNA trafficking via membrane nanotubes in mammalian cells

Authors: Gal Haimovich^{a,b}, Christopher M. Ecker^{c,d}, Margaret C. Dunagin^c, Elliott Eggan^c, Arjun Raj^c, Jeffrey E. Gerst^{a*} and Robert H. Singer^{b*}

Affiliations

^aDepartment of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.

^bDepartment of Anatomy & Structural Biology, Albert Einstein College of Medicine, Bronx, NY, USA.

^cDepartment of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA.

^dCurrent address: Department of Microbiology, University of Pennsylvania, Philadelphia, PA, USA.

*Correspondence to: robert.singer@einstein.yu.edu; jeffrey.gerst@weizmann.ac.il

Key words: β -actin mRNA, membrane nanotubes, MS2 aptamer, single-molecule FISH, mammalian cells

A thermostable Cas9 with increased lifetime in human plasma

Lucas B. Harrington¹, David Paez-Espino², Janice S. Chen¹, Enbo Ma¹, Brett T. Staahl¹,
Nikos C. Kyrpides² & Jennifer Doudna^{1, 3-6}

¹Department of Molecular and Cell Biology, University of California, Berkeley, California, 94720, USA.

²Department of Energy, Joint Genome Institute, Walnut Creek, California 94598, USA

³Department of Chemistry, University of California, Berkeley, California, 94720, USA.

⁴Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA.

⁵Innovative Genomics Institute, University of California, Berkeley, California 94720, USA.

⁶MBIB Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

TITLE PAGE

Title: Removal of inhibition uncovers latent movement preparation dynamics

Abbreviated title: Motor potential during saccade preparation

Authors: Uday K. Jagadisan^{1,3} and Neeraj J. Gandhi^{1,2,3}

Author affiliations:

¹Department of Bioengineering

²Department of Neuroscience

³Center for the Neural Basis of Cognition

University of Pittsburgh, Pittsburgh, PA 15213

Corresponding author: Uday K. Jagadisan, Ph.D.

153 Eye and Ear Institute

203 Lothrop St

Pittsburgh, PA 15213

USA

Tel: +1-315-4099934

E-mail: kj.udayakiran@gmail.com

Author Contributions: U.K.J and N.J.G designed the study. U.K.J. performed the experiments and analyzed the data. U.K.J and N.J.G wrote the manuscript.

Number of pages: 40

Number of figures: 7 + 2 supplementary figures

Number of words: Abstract - 209, Main text - 5859, Methods - 2144

Conflict of interest: The authors declare no competing financial interests.

Acknowledgements: The study was funded by NIH grants EY022854 and EY024831 awarded to N.J.G.

Keywords: sensorimotor, superior colliculus, eye movements, threshold, parallel processing, inhibitory gating

Dynamic blue light-inducible T7 RNA polymerases (Opto-T7RNAPs) for precise spatiotemporal gene expression control

Armin Baumschlager, Stephanie K. Aoki and Mustafa Khammash*

Department of Biosystems Science and Engineering (D-BSSE), ETH–Zürich, Mattenstrasse 26, 4058 Basel, Switzerland

KEYWORDS: *optogenetics, light control, photoreceptor, LOV domain, transcription, dynamic gene expression, T7 RNA polymerase, split protein*

ABSTRACT: Light has emerged as control input for biological systems due to its precise spatiotemporal resolution. The limited toolset for light control in bacteria motivated us to develop a light-inducible transcription system that is independent from cellular regulation through the use of an orthogonal RNA polymerase. Here, we present our engineered blue light-responsive T7 RNA polymerases (Opto-T7RNAPs) that show properties such as low leakiness of gene expression in the dark-state, high expression strength when induced with blue light, or an inducible range of more than 300-fold. Following optimization of the system to reduce expression variability, we have created a variant, which returns to the inactive dark-state within minutes, once blue light is turned off. This allows for precise dynamic control of gene expression, which is a key aspect for most applications using optogenetic regulation. The regulators were developed and tested in the bacterium *Escherichia coli*, which is a crucial cell factory for biotechnology due to its fast and inexpensive cultivation and well understood physiology and genetics. However, minor alterations should be sufficient to allow their use in other species in which the T7 RNAP polymerase and the light-inducible Vivid regulator were shown to be functional, which comprises other bacterial species and eukaryotes such as mammalian cells or yeast. We anticipate that our approach will expand the applicability of using light as an inducer for gene expression independent from cellular regulation, and allow for a more reliable dynamic control of synthetic and natural gene networks.

INTRODUCTION

Small molecule induced gene expression systems are a key component in synthetic biology¹ and biotechnological applications². However, chemical inducers are limited in their application in space and time. Spatiotemporal control is of increasing interest, as biological systems are regulated dynamically and respond to intracellular stimuli and changes in internal states³. Although static perturbations, such as growth media variation and gene knockouts, have been extensively and successfully used to elucidate gene network structure and function, approaches using dynamic perturbations are providing new insight into the organizing principles of biology and the study of gene networks⁴. Dynamic regulation is also starting to be explored by metabolic engineers^{4–6}. However, very recent work addresses the problem that few broadly applicable tools are available for dynamic pathway regulation, and further show that dynamic regulation can significantly increase product titers through dynamic pathway regulation⁷.

Light-based regulation is superior to conventional small-molecule inducers in this regard, displaying better temporal properties, as removal of small molecules might be challenging in scenarios such as batch or fed-batch processes. Further, it allows for spatial control of individual cells (whereas small molecules are diffusion-limited), and is minimally invasive⁸, a desired feature for basic research. These distinguishing properties of light over small molecules led to the development of numerous light-controlled devices^{9–13}. Light-inducible dimerization domains were successfully exploited in two-hybrid-like systems¹⁴ to create optogenetic gene expression systems in

eukaryotes^{8–10,15}, and for reconstitution of functional proteins from their split parts^{16–19}.

Dynamically light-inducible systems allow new regulation schemes, by moving the controller out of the cell and using light as input signal for control. Both biology and engineering make use of feedback control to achieve robust regulation, which in turn allows natural and engineered systems to function reliably in the face of disturbances or changing environmental conditions. However, the design of synthetic biological feedback controllers remains challenging due to the fact that biological parts do not behave as reproducibly as electronic ones. To overcome this obstacle, *in silico* feedback control was introduced by our group, which allows for electronic control of biological responses^{20,21}.

Another challenge for precise control is that depending on the growth phases, nutrient conditions, and other extrinsic factors, the concentration of RNA polymerase varies, ranging from 1,800 to 10,200 molecules per cell²². Along with fluctuations in the ribosome concentration, this can result in changes of expression levels^{23–25} and reduce the performance of constitutive promoters. This poses a challenge to systems that require precise balances in expression levels²⁶, especially when media and growth conditions change such as during industrial scale-up²⁷. To decouple expression of a gene of interest from cellular RNAP concentrations, the heterologous T7 DNA-dependent RNA polymerase (T7RNAP), originating from the T7 bacteriophage^{28,29}, is commonly used for protein overexpression. The polymerase shows high processivity, a high selectivity for the T7 promoter, and does not transcribe se-

1 **A complete electron microscopy volume of the brain of adult *Drosophila*** 2 ***melanogaster***

3 Zhihao Zheng^{1,‡}, J. Scott Lauritzen^{1,‡}, Eric Perlman^{1,6}, Camenzind G. Robinson¹, Matthew
 4 Nichols¹, Daniel Milkie^{2,7}, Omar Torrens², John Price³, Corey B. Fisher¹, Nadiya Sharifi¹, Steven
 5 A. Calle-Schuler¹, Lucia Kmecova¹, Iqbal J. Ali¹, Bill Karsh¹, Eric T. Trautman¹, John Bogovic¹,
 6 Philipp Hanslovsky¹, Gregory S. X. E. Jefferis⁴, Michael Kazhdan⁵, Khaled Khairy¹, Stephan
 7 Saalfeld¹, Richard D. Fetter^{1,8}, Davi D. Bock^{1*}

8 1. Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147

9 2. Coleman Technologies, Inc., Newtown Square, PA 19073

10 3. Hudson Price Designs, LLC, Hingham, MA 02043

11 4. Division of Neurobiology, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK

12 5. Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218

13 6. Present address: Center for Imaging Science, Johns Hopkins University, Baltimore, MD
 14 21218

15 7. Present address: Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA
 16 20147

17 8. Present address: Department of Biochemistry & Biophysics, University of California, San
 18 Francisco 94158

19 * bockd@janelia.hhmi.org

20 [‡]These authors contributed equally to this work.

An Efficient CRISPR protocol for generating Conditional and Knock-in mice using long single-stranded DNA donors.

Hiromi Miura ^{1,2,6}, Rolan M. Quadros ^{3,6}, Channabasavaiah B. Gurumurthy ^{3,4,7}, Masato Ohtsuka ^{1,2,5,7}

- 1) Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, School of Medicine, Tokai University, Kanagawa 259-1193, Japan
- 2) Center for Matrix Biology and Medicine, Graduate School of Medicine, Tokai University, Kanagawa 259-1193, Japan
- 3) Mouse Genome Engineering Core Facility, Vice Chancellor for Research Office, University of Nebraska Medical Center, Omaha, NE, USA
- 4) Developmental Neuroscience, Munroe Meyer Institute for Genetics and Rehabilitation, University of Nebraska Medical Center, Omaha, NE, USA
- 5) The Institute of Medical Sciences, Tokai University, Kanagawa 259-1193, Japan
- 6) Contributed equally
- 7) Corresponding authors: masato@is.icc.u-tokai.ac.jp and cgurumurthy@unmc.edu,

The CRISPR/Cas9 tool can easily generate knockout mouse models by disrupting the gene sequence, but its efficiency for creating models that require either insertion of exogenous DNA (knock-in) or replacement of genomic segments is very poor. The majority of mouse models used in research are knock-in (reporters or recombinases) or gene-replacement (for example, conditional knockout alleles containing *LoxP* sites flanked exons). A few methods for creating such models are reported using double-stranded DNA as donors, but their efficiency is typically 1-10% and therefore not suitable for routine use. We recently demonstrated that long single-stranded DNAs serve as very efficient donors, both for insertion and for gene replacement. We call this method *Easi*-CRISPR (efficient additions with ssDNA inserts-CRISPR), a highly efficient technology (typically 25%-50%, and up to 100% in some cases), one that has worked at over a dozen loci thus far. Here, we provide detailed protocols for *Easi*-CRISPR.

Genetically engineered orange petunias on the market

Hany Bashandy^{1,2}, Teemu H. Teeri^{1,*}

¹Department of Agricultural Sciences, Viikki Plant Science Centre, P.O. Box 27, 00014 University of Helsinki, Finland.

²Department of Genetics, Cairo University, 13 Gamaa St., Giza 12619, Egypt.

*Correspondence to: teemu.teeri@helsinki.fi

Abstract: Genetic engineering of petunia was shown to lead to novel flower color some twenty years ago. Here we show that petunia lines with orange flowers, generated for scientific purposes, apparently found their way to petunia breeding programmes, intentionally or unintentionally. Today they are widely available, but have not been registered for commerce.

The pathway to the colored anthocyanins in the ornamental plant petunia (*Petunia hybrida*) is a well-known example of substrate specificity of one enzyme limiting the spectrum of possible products of the pathway¹. Anthocyanins are water soluble pigments giving flowers, fruits and sometimes vegetative parts of plants colours ranging from orange and red to blue and purple². Anthocyanins are extensively glycosylated and acylated, the molecular decoration affecting their spectral properties. At the aglycone level the three most common variants of the molecule are the anthocyanidins pelargonidin, cyanidin and delphinidin, differing by the number of hydroxyl groups (one, two or three, respectively) in the B-ring of the molecule. Hydroxylation takes place at the level of dihydroflavonols in the pathway (possibly earlier in some cases) by two enzymes, flavonoid 3'-hydroxylase (F3'H) and flavonoid 3',5'-hydroxylase (F3'5'H). The enzyme dihydroflavonol reductase (DFR) converts dihydroflavonols to corresponding leucoanthocyanidins, which then are oxidized to anthocyanidins by anthocyanidin synthase (syn. leucoanthocyanidin oxidase). In petunia, the DFR enzyme does not react with the simplest precursor (dihydrokaempferol), therefore the natural range of petunia flower colours lack orange hues typical to pelargonidin derivatives. Flowers of petunia cultivars that have mutations in the two hydroxylases are therefore white.

It was shown few decades ago that by introducing a gene encoding DFR from a species where the enzyme does not show substrate specificity into a petunia line that lacks F3'H and F3'5'H activity, one can open up the pathway to pelargonidin. Using the maize gene *A1* Meyer and colleagues generated brick red colored flowers in petunia³ and using the gene from the ornamental plant *Gerbera hybrida*, our own laboratory generated petunia lines with bright orange flowers⁴.

These petunia flowers were investigated concerning factors relating to stability of the transgene (and therefore the novel colour)^{5,6}, but they were never commercialized. The list of registered genetically modified petunia plants is very short and includes a single line transgenic for a chalcone synthase encoding gene approved for cultivation in China (<http://www.isaaa.org/gmapprovaldatabase/>). Therefore, it was a great surprise and a delight from the point of view of maybe gaining insight in the ways petunia germplasm changes under breeding, when we encountered bright orange coloured petunias in flower boxes decorating the Helsinki railway station during summers of 2015 and 2016 (Figure 1). Indeed, orange petunias are widely on

Tandem histone-binding domains enhance the activity of a synthetic chromatin effector

Stefan J. Tekel¹, Daniel Vargas¹, Lusheng Song², Joshua LaBaer², and Karmella A. Haynes¹

1. School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ
2. Biodesign Institute, Arizona State University, Tempe, AZ

Keywords: chromatin, histones, transcription factor, protein engineering, epigenetics

Corresponding author:

Karmella A. Haynes

karmella.haynes@asu.edu

ABSTRACT

Fusion proteins that specifically interact with biochemical marks on chromosomes represent a new class of synthetic transcriptional regulators that decode cell state information rather than DNA sequences. In multicellular organisms, information relevant to cell state, tissue identity, and oncogenesis is often encoded as biochemical modifications of histones, which are bound to DNA in eukaryotic nuclei and regulate gene expression states. We have previously reported the development and validation of the “Polycomb-based transcription factor” (PcTF), a fusion protein that recognizes histone modifications through a protein-protein interaction between its polycomb chromodomain (PCD) motif and trimethylated lysine 27 of histone H3 (H3K27me3) at genomic sites. We demonstrated that PcTF activates genes at methyl-histone-enriched loci in cancer-derived cell lines. However, PcTF induces modest activation of a methyl-histone associated reporter compared to a DNA-binding activator. Therefore, we modified PcTF to enhance its target affinity. Here, we demonstrate the activity of a modified regulator called Pc₂TF, which has two tandem copies of the H3K27me3-binding PCD at the N-terminus. Pc₂TF shows higher affinity for H3K27me3 *in vitro* and shows enhanced gene activation in HEK293 cells compared to PcTF. These results provide compelling evidence that the intrinsic histone-binding activity of the PCD motif can be used to tune the activity of synthetic histone-binding transcriptional regulators.

Highly parallel genome variant engineering with CRISPR/Cas9 in eukaryotic cells

Meru J. Sadhu^{1,*†}, Joshua S. Bloom^{1,*†}, Laura Day¹, Jake J. Siegel^{1‡}, Sriram Kosuri²,
Leonid Kruglyak^{1,†}

¹ Department of Human Genetics, Department of Biological Chemistry, Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA.

² Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90095, USA.

* These authors contributed equally to this work.

†Corresponding author. Email: msadhu@mednet.ucla.edu (M.J.S.); jbloom@mednet.ucla.edu (J.S.B.); lkruglyak@mednet.ucla.edu (L.K.)

‡Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Abstract:

Direct measurement of functional effects of DNA sequence variants throughout a genome is a major challenge. We developed a method that uses CRISPR/Cas9 to engineer many specific variants of interest in parallel in the budding yeast *Saccharomyces cerevisiae*, and to screen them for functional effects. We used the method to examine the functional consequences of premature termination codons (PTCs) at different locations within all annotated essential genes in yeast. We found

Chimeric CRISPR guides enhance Cas9 target specificity

Noah Jakimo^{1,2}, Pranam Chatterjee^{1,2} & Joseph M Jacobson^{1,2}

¹MIT Media Lab, Cambridge, United States and

²MIT Center for Bits and Atoms, Cambridge, United States.

Oligonucleotide-guided nucleases (OGNs) have enabled rapid advances in precision genome engineering. Though much effort has gone into characterizing and mitigating mismatch tolerance for the most widely adopted OGN, *Streptococcus pyogenes* Cas9 (SpCas9), potential off-target interactions may still limit applications where on-target specificity is critical. Here we present a new axis to control mismatch sensitivity along the recognition-conferring spacer sequence of SpCas9's guide RNA (gRNA). We introduce mismatch-evading lowered-thermostability guides (melt-guides) and exhibit how nucleotide-type substitutions in the spacer can reduce cleavage of sequences mismatched by as few as a single base pair. Co-transfecting melt-guides into human cell culture with an exonuclease involved in DNA repair, we observe indel formation on a standard genomic target at approximately 70% the rate of canonical gRNA and undetectable on off-target data.

The recent discoveries, characterizations, and modifications of natural oligonucleotide-guided nucleases associated with CRISPR and RNAi have empowered a genome-editing revolution¹⁻⁴. Low barriers for OGNs' cost and design drive their widespread adoption over alternatives, including modular base-recognition domains (i.e., transcription activator like effector, zinc finger, and pumilio assemblies), which can be hard to synthesize, or meganucleases, which are difficult

A single-cell anatomical blueprint for intracortical information transfer from primary visual cortex

Yunyun Han^{1,2,3,*}, Justus M Kebschull^{4,5,*}, Robert AA Campbell^{1,*}, Devon Cowan¹, Fabia Imhof¹, Anthony M Zador^{5,†}, Thomas D Mrsic-Flogel^{1,6,†}

¹Biozentrum, University of Basel, 4056 Basel, Switzerland

²Department of Neurobiology School of Basic Medicine and Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

³The Institute for Brain Research, Collaborative Innovation Center for Brain Science, Huazhong University of Science and Technology, Wuhan, China

⁴Watson School of Biological Sciences, Cold Spring Harbor, NY 11724, USA

⁵Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

⁶Sainsbury Wellcome Centre, University College London, London, UK

*These authors contributed equally to the work as first authors.

†These authors contributed equally to the work as senior authors.

Correspondence and requests should be addressed to Anthony Zador, zador@cshl.edu, or Thomas Mrsic-Flogel, tom@mouse.vision.

The wiring diagram of the neocortex determines how information is processed across dozens of cortical areas. Each area communicates with multiple others via extensive long-range axonal projections¹⁻⁵, but the logic of inter-area information transfer remains unresolved. In sensory neocortex, previous work suggests that neurons typically innervate single cortical areas^{3,5,6}, implying that information is distributed via ensembles of dedicated pathways. Alternatively, single neurons could broadcast information to multiple cortical targets⁶⁻⁹. Distinguishing between these models has been challenging because the projection patterns of only a few individual neurons have been reconstructed. Here we map the projection patterns of axonal arbors from 595 individual neurons in mouse primary visual cortex (V1) using two complementary methods: whole-brain fluorescence-based axonal tracing^{10,11} and high-throughput DNA sequencing of genetically barcoded neurons (MAPseq)¹². Although our results confirm the existence of dedicated projections to certain cortical areas, we find these are the exception, and that the majority of V1 neurons broadcast information to multiple cortical targets. Furthermore, broadcasting cells do not project to all targets randomly, but rather comprise subpopulations that either avoid or preferentially innervate specific subsets of cortical areas. Our data argue against a model of dedicated lines of inter-areal information transfer via “one neuron – one target area” mapping. Instead, long-range communication between a sensory cortical area and its targets is based on a principle whereby individual neurons copy information to, and potentially coordinate activity across, specific subsets of cortical areas.

**Correlated Light-Serial Scanning Electron Microscopy (CoLSSEM)
for ultrastructural visualization of single neurons *in vivo***

Yusuke Hirabayashi^{1,2,3,4}, Juan Carlos Tapia^{1,5 *} and Franck Polleux^{1,2,3 *}

¹ Columbia University Medical Center - Department of Neuroscience

² Mortimer B. Zuckerman Mind Brain Behavior Institute

³ Kavli Institute for Brain Science

⁴ JST, PRESTO

550 W. 120th Street, Northwest Corner Building #1104, New York, NY 10027

⁵ University of Talca, Talca, Chile – Department of Health Sciences

* Co-corresponding authors:

Franck Polleux, Ph.D.
Columbia University
Department of Neuroscience
Mortimer B. Zuckerman Mind Brain Behavior Institute
Kavli Institute for Brain Science

1108 NWC Building - MC 4862
550 West 120th Street
New York, N.Y. 10027
fp2304@cumc.columbia.edu

Tel: 212-853-0407

Juan Carlos Tapia, Ph.D.
University of Talca
Department of Health Sciences
Avda. Lircay S/N
Talca, Chile 3460000
juantapia@utalca.cl

Tel: [56-712-418855](tel:56-712-418855) extension 2855

Inducible, tunable and multiplex human gene regulation using CRISPR-Cpf1-based transcription factors

Yu Gyoung Tak^{1,2}, Benjamin P. Kleinstiver^{1,2}, James K. Nuñez³, Jonathan Y. Hsu¹, Jingyi Gong¹, Jonathan S. Weissman³, J. Keith Joung^{1,2}


¹Molecular Pathology Unit, Center for Cancer Research, and Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, MA 02129 USA

²Department of Pathology, Harvard Medical School, Boston, MA 02115 USA

³Department of Cellular & Molecular Pharmacology University of California, San Francisco, San Francisco, CA 94158 USA and Howard Hughes Medical Institute


*Correspondence to: JJOUNG@MGH.HARVARD.EDU

Impact of fluorescent protein fusions on the bacterial flagellar motor

M Heo¹, AL Nord¹, D Chamousset¹, E van Rijn², HJE Beaumont², F Pedaci^{1,*}

1 Single-molecule Biophysics dept, Centre de Biochimie Structurale, CNRS UMR5048 UM INSERM U1054, 29 Rue de Navacelles, 34090 Montpellier, France

2 Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, 2628 CJ, Delft, The Netherlands

 These authors contributed equally to this work.

* francesco.pedaci@cbs.cnrs.fr

The experimental design and data interpretation in “Unexpected mutations after CRISPR–Cas9 editing *in vivo*” by Schaefer et al. are insufficient to support the conclusions drawn by the authors

The recent correspondence to the Editor of Nature Methods by Schaefer et. al.¹ has garnered significant attention since its publication as a result of its strong conclusions contradicting numerous publications in the field using similar analytical approaches and methods²⁻⁴. The authors suggest that the CRISPR-Cas9 system is highly mutagenic in genomic regions not expected to be targeted by the gRNA. We believe that the conclusions drawn from this study are unsubstantiated by the disclosed experiments as they were designed and carried out. Further, it is impossible to ascribe the observed differences in the subject mice to the effects of CRISPR *per se*. The genetic differences seen in this comparative analysis were likely present prior to editing with CRISPR.

In our view, the experiments, observations, and subsequent assertions in Schaefer et al.¹ can be summarized as follows. Two mice created using CRISPR-based genome editing in the zygote stage, when compared to a single “co-housed FVB/NJ mouse without CRISPR-mediated correction”, showed a significant number of single nucleotide variants (SNVs) and insertions and deletions (indels) across the genome. The number of mutations common to the two independently generated CRISPR edited mice was 1,397 SNVs and 117 indels. Surprisingly, these apparent mutations all arose from sequences in the genome that contain poor homology to the gRNA (between 5% – 65%). Furthermore, none of the 50 closest, predicted off-target sites (based on gRNA sequence homology) had any observed activity (SNVs or indels). The authors speculate that there is an unreported activity where “certain sgRNAs may target loci independently of their target *in vivo*.”

Our opinion is that the conclusions drawn from this study are unsubstantiated by the disclosed experiments and that it is impossible to ascribe the observed differences in the subject mice to the effects of CRISPR *per se* is based upon the following observations:

Firstly, the overall number of the study subjects is low (n = 2 treated mice and n = 1 untreated mouse) and the depth of sequencing applied to the treated and untreated mice is not equivalent. An underpowered study may prove limiting when attempting to understand statistical reproducibility and reliability of scientific observations.

Secondly, the selection of a co-housed mouse (as opposed to the parents or *bona fide* littermates) as the control is insufficient to attribute the observed differences between the treated mice and control mouse to CRISPR. The design of the experiment makes it impossible for the authors to rule out the possibility that the reported genomic differences between the experimental animals and the single control existed prior to experimental manipulation with CRISPR. In fact, published literature has shown that differences in the genomes of *littermates* analyzed by whole genome sequencing (WGS) can be significant (985 SNVs were identified by

Response to Editas: Unexpected mutations after CRISPR-Cas9 editing *in vivo*

Kellie A. Schaefer^{1,2*}, Wen-Hsuan Wu^{3,4*}, Diana F. Colgan¹, Stephen H. Tsang M.D., Ph.D.^{3,4}, Alexander G. Bassuk M.D., Ph.D.^{5,6§}, Vinit B. Mahajan M.D. Ph.D.^{1,7§}

¹Omics Laboratory, Stanford University, Palo Alto, CA.

²Interdisciplinary Graduate Program in Genetics, University of Iowa, Iowa City, IA.

³Jonas Children's Vision Care, and Bernard & Shirlee Brown Glaucoma Laboratory, Departments of Ophthalmology, Pathology & Cell Biology, Institute of Human Nutrition, College of Physicians and Surgeons, Columbia University, New York, NY 10032.

⁴Edward S. Harkness Eye Institute, New York-Presbyterian Hospital, New York, NY 10032

⁵Department of Pediatrics, University of Iowa, Iowa City, IA.

⁶Neurology, University of Iowa, Iowa City, IA.

⁷Byers Eye Institute, Stanford University, Palo Alto, CA.

*Contributed equally

§Co-corresponding Authors:

Vinit B. Mahajan, M.D., Ph.D., Byers Eye Institute, Stanford University, Palo Alto, CA, 94303.

Phone: 650-721-6888. E-mail: mahajanlab@gmail.com

Alexander G. Bassuk, M.D., Ph.D., Department of Pediatrics, The University of Iowa, Iowa City, IA, 52242. Phone: 319-356-7726. E-mail: alexander-bassuk@uiowa.edu

Fluidic microactuation of flexible electrodes for neural recording

Flavia Vitale^{1†§}, Daniel G. Vercosa^{2,3†}, Alexander V. Rodriguez³, Sushma Sri Pamulapati¹, Frederik Seibt⁴, Eric Lewis³, J. Stephen Yan⁵, Krishna Badhiwala⁵, Mohammed Adnan^{1¶}, Gianni Royer-Carfagni⁶, Michael Beierlein⁴, Caleb Kemere^{3,5,7*}, Matteo Pasquali^{1, 8*} and Jacob T. Robinson^{2,3,5,7*}

1. Department of Chemical and Biomolecular Engineering, Rice University, Houston TX 77005
2. Applied Physics Program, Rice University, Houston TX 77005
3. Department of Electrical and Computer Engineering, Rice University, Houston TX 77005
4. McGovern Medical School at UTHealth, Department of Neurobiology and Anatomy, Houston TX 77030
5. Department of Bioengineering, Rice University, Houston TX 77005
6. Department of Engineering and Architecture, University of Parma, Parma, I-43100, Italy
7. Department of Neuroscience, Baylor College of Medicine, Houston TX 77030
8. Department of Chemistry, The Smalley-Curl Institute, Rice University, Houston, TX 77005

†These authors contributed equally to this work

*Correspondence:

Jacob T. Robinson, Department of Electrical and Computer Engineering, Rice University, Duncan Hall 2041, 6100 Main Street, MS 380 Houston, TX 77005, USA e-mail: jtrobinson@rice.edu

Matteo Pasquali, Department of Chemical and Biomolecular Engineering, Rice University, Keck Hall 229, 6100 Main Street, MS 369 Houston, TX 77005, USA e-mail: mp@rice.edu

Caleb Kemere, Department of Electrical and Computer Engineering, Rice University, BRC 727, 6100 Main Street, MS 384 Houston, TX 77005, USA e-mail: caleb.kemere@rice.edu

Current address:

[§]Center for Neuroengineering and Therapeutics, Department of Neurology, University of Pennsylvania, Philadelphia PA 19104

[¶]Abu Dhabi Financial Group, Abu Dhabi, UAE

Mechanistic Implications of Enhanced Editing by a HyperTRIBE RNA-binding protein

Weijin Xu, Reazur Rahman, Michael Rosbash

Department of Biology, Howard Hughes Medical Institute and National
Center for Behavioral Genomics, Brandeis University, Waltham,
Massachusetts, USA.

Corresponding author: Michael Rosbash, rosbash@brandeis.edu

Variation in neuronal activity state, axonal projection target, and position principally define the transcriptional identity of individual neocortical projection neurons

*Maxime Chevé^{1,2}, Johanna D. Robertson³, Gabrielle H. Cannon⁴
Solange P. Brown^{2,*} and Loyal A. Goff^{2,4,*}*

Affiliations:

¹Biochemistry, Cellular and Molecular Biology Graduate Program, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

²Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

³Human Genetics Training Program, McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

⁴McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, 21205, USA.

*Corresponding Authors: spbrown@jhmi.edu (SPB) and loyaloff@jhmi.edu (LAG)

Questioning unexpected CRISPR off-target mutations in vivo

To the Editor: Recently, Schaefer et al.¹ reported that whole genome sequencing (WGS) of two Cas9-treated, gene-corrected mice and a wild-type control mouse unveiled 1,397 single-nucleotide variations (SNVs) and 117 small insertions and deletions (indels) present commonly in the two Cas9-treated mice “but absent in the uncorrected control” and from a database of mouse SNVs and indels. There was essentially no sequence homology between the on-target site and these SNVs and indel sites, most of which lacked a protospacer-adjacent motif (PAM) sequence, suggesting that these variations were both small guide RNA (sgRNA)-independent and Cas9-independent, respectively. Nevertheless, the authors made a bold claim that these variations were caused by CRISPR-Cas9 without validating these unexpected off-target effects even at a single SNV or indel site by performing an independent experiment in vitro or in vivo. Another major concern in Schaefer et al. is the absence of analysis of variants that are present in the control mice but absent in the two gene-edited mice.

Target specificities of CRISPR systems have been extensively studied in animals and cell lines. For example, we showed that certain CRISPR-Cas9 nucleases did not induce detectable off-target mutations at sites with just two or three-nucleotide mismatches in human cells, first using T7 endonuclease I assays² and then using targeted deep sequencing³. We and others also performed WGS to show that Cas9 rarely induced off-target indels in a clonal population of cells⁴⁻⁷ or a gene-edited animal⁸. Note that Schaefer et al. did not find any off-target mutations in the two Cas9-treated mice at the top 50 most likely off-target sites with 3- to 4-nucleotide mismatches, in line with these previous reports. Given the remarkable specificity of CRISPR-Cas9, it is difficult to believe that Cas9 can cleave sites that differ by > 10 nucleotides from the on-target sequence, as suggested in Schaefer et al.

The authors did not articulate whether the unexpected off-target effects were limited to the particular target site or their method or FVB/NJ zygotes used in their experiments. In silico off-target predicting algorithms and cell-based, genome-wide off-target profiling methods such as GUIDE-seq⁹ and HTGTS¹⁰ cannot identify off-target sites with no sequence homology. However, Digenome-seq, an in vitro method of capturing in vitro cleavage sites using Cas9-digested, cell-free genomic DNA via WGS, does not rely on sequence homology⁴, because DNA double-strand break ends remain intact in vitro and are protected from digestion by endogenous exonucleases in vivo. Note that we did not find any unusual, non-homologous off-target sites in human genomic DNA using Digenome-seq⁴, showing that such sites cannot

CRISPR/Cas9-APEX-mediated proximity labeling enables discovery of proteins associated with a predefined genomic locus in living cells

Samuel A Myers^{1*}, Jason Wright¹, Feng Zhang¹, and Steven A Carr^{1*}

Affiliations

1. Broad Institute of MIT and Harvard, Cambridge, MA 02142

Correspondence: sammyers@broadinstitute.org and scarr@broadinstitute.org

Abstract

The activation or repression of a gene's expression is primarily controlled by changes in the proteins that occupy its regulatory elements. The most common method to identify proteins associated with genomic loci is chromatin immunoprecipitation (ChIP). While having greatly advanced our understanding of gene expression regulation, ChIP requires specific, high quality, IP-competent antibodies against nominated proteins, which can limit its utility and scope for discovery. Thus, a method able to discover and identify proteins associated with a particular genomic locus within the native cellular context would be extremely valuable. Here, we present a novel technology combining recent advances in chemical biology, genome targeting, and quantitative mass spectrometry to develop genomic locus proteomics, a method able to identify proteins which occupy a specific genomic locus.

Body

Transcriptional regulation is a highly-coordinated process largely controlled by changes in protein occupancy at regulatory elements of the modulated genes. Chromatin immunoprecipitation (ChIP) followed by quantitative polymerase chain reaction (-qPCR), microarrays (-chip), or massively parallel next-generation sequencing (-seq) has been invaluable for our understanding of transcriptional regulation and chromatin structure, both at the individual locus and genome-wide levels (1-6). However, because ChIP requires the use of antibodies, its utility can often be limited by the presupposition of a suspected protein's occupancy, and lack of highly specific and high affinity reagents. Previously developed "reverse ChIP" type methods suffer from several drawbacks including loss of cellular and/or chromatin context, extensive engineering and locus disruption, reliance on repetitive DNA sequences, and the need for chemical crosslinking, which reduces sensitivity for mass spectrometric-based approaches (7-11). Therefore, we sought to develop a method to identify proteins associated with a specific, non-repetitive genomic locus in the native cellular context without the need for crosslinking or genomic alterations. Here, we utilized recent advances in sequence-specific DNA targeting and affinity labeling in cells to develop genomic locus proteomics (GLOPro) to characterize proteins associated with a particular genomic locus.

“Unexpected mutations after CRISPR-Cas9 editing *in vivo*” are most likely pre-existing sequence variants and not nuclease-induced mutations

Caleb A. Lareau,^{1,2,3,*} Kendell Clement,^{1,3,4,5,*} Jonathan Y. Hsu,^{1,6,*} Vikram Pattanayak,^{1,4} J. Keith Joung,^{1,4,+} Martin J. Aryee,^{1,2,3,4,+} Luca Pinello^{1,4,+}

¹Molecular Pathology Unit, Massachusetts General Hospital, Charlestown, MA 02129 USA

²Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115 USA

³Cell Circuits and Epigenomics Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142 USA

⁴Department of Pathology, Harvard Medical School, Boston, MA 02115

⁵Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138 USA

⁶Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

*These authors contributed equally to this work

+Correspondence should be addressed to jjoung@mgh.harvard.edu, aryee.martin@mgh.harvard.edu, or lpinello@mgh.harvard.edu

Competition for a limited supply of synaptic building blocks predicts multiplicative synaptic normalization and heterosynaptic plasticity

Jochen Triesch^{1,*}, Anne-Sophie Hafner²

1 Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany

2 Max-Planck Institute for Brain Research, Frankfurt am Main, Germany

*** E-mail: triesch@fias.uni-frankfurt.de**

Abstract

We present a mathematical model of synaptic normalization and heterosynaptic plasticity based on competition for limited synaptic resources. In the model, afferent synapses on a part of the dendritic tree of a neuron compete for a limited supply of synaptic building blocks such as AMPA receptors or other postsynaptic components, which are distributed across the dendritic tree. These building blocks form a pool of parts that are ready for incorporation into synapses. Using minimal assumptions, the model produces fast multiplicative normalization behavior and leads to a homeostatic form of heterosynaptic plasticity. It therefore supports the use of such rules in neural network models. Furthermore, the model predicts that the amount of heterosynaptic plasticity is small when many building blocks are available in the pool. The model also suggests that local production and/or assembly of postsynaptic building blocks across the dendritic tree may be necessary to maintain a neuron's proper function, because it facilitates their homogeneous distribution across the dendritic tree. Because of its simplicity and analytical tractability, the model provides a convenient starting point for the development of more detailed models of the molecular mechanisms underlying different forms of synaptic plasticity.

Author Summary

Changes in the efficacies of synapses are thought to be the neurobiological basis of learning and memory. When a synapse is strengthened, new neurotransmitter receptors are added to the postsynaptic membrane. Recent experiments have shown that the behavior of these receptors is highly dynamic, with receptors moving back and forth between synapses on time scales of seconds and minutes, reflecting a competition of synapses for available receptors. Here we propose a mathematical model of this competition of synapses for neurotransmitter receptors or other synaptic building blocks. Using minimal assumptions the model produces a multiplicative normalization behavior of synapses and it predicts a well-known form of so-called heterosynaptic plasticity, where changes in stimulated synaptic pathways induce changes of opposite sign in neighboring, unstimulated pathways. Thus, the model offers a parsimonious mechanistic explanation of these forms of synaptic plasticity.

Introduction

Simple mathematical models of Hebbian learning exhibit an unconstrained growth of synaptic efficacies. To avoid runaway dynamics, some mechanism for limiting weight growth needs to

Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription-translation system

Ryan Marshall^{1*}, Colin S. Maxwell^{2*}, Scott P. Collins², Michelle L. Luo², Thomas Jacobsen²,
Chase L. Beisel^{2†}, Vincent Noireaux^{1†}

¹School of Physics and Astronomy

University of Minnesota, Minneapolis, MN, 55455

²Department of Chemical and Biomolecular Engineering

North Carolina State University, Raleigh, NC 27695

*These authors contributed equally to this work

†Correspondence directed to cbeisel@ncsu.edu (C.L.B.) and noireaux@umn.edu (V.N.)

Key words: Cas9, Cascade, Cpf1, PAM, prototyping, synthetic biology, TXTL

Running title: Characterizing CRISPR with TXTL

Frequent lack of repressive capacity of promoter DNA methylation identified through genome-wide epigenomic manipulation

**Ethan Ford¹, Matthew R. Grimmer^{2,3,4}, Sabine Stolzenburg⁵, Ozren Bogdanovic^{1,6,7}, Alex de
Mendoza¹, Peggy J. Farnham², Pilar Blancafort^{5,8}, Ryan Lister^{1,8 *}**

Affiliations:

¹Australian Research Council Centre of Excellence in Plant Energy Biology, School of Molecular Sciences, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia.

²Department of Biochemistry and Molecular Medicine, University of Southern California, 1450 Biggy St, Los Angeles, CA 90089, USA.

³Integrated Genetics and Genomics, University of California, Davis, 451 Health Sciences Dr, Davis, CA 95616, USA.

⁴Department of Neurological Surgery, University of California, San Francisco, 1450 3rd St, San Francisco, CA 94158, USA.

⁵School of Anatomy, Physiology and Human Biology, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia.

⁶Genomics and Epigenetics Division, Garvan Institute of Medical Research, Sydney, New South Wales, Australia.

⁷St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia.

⁸Harry Perkins Institute of Medical Research, 6 Verdun St, Nedlands, WA 6009, Australia.

*Corresponding author. Email: ryan.lister@uwa.edu.au (R.L.)

1 **Title: Functional consequences of genetic loci associated with intelligence in a** 2 **meta-analysis of 87,740 individuals**

3

4 Jonathan R. I. Coleman^{1,2}, Julien Bryois³, H  l  na A. Gaspar¹, Philip R. Jansen^{4,5},
5 Jeanne Savage⁴, Nathan Skene⁶, Robert Plomin¹, Ana B. Mu  oz-Manchado⁶, Sten
6 Linnarsson⁶, Greg Crawford^{7,8}, Jens Hjerling-Leffler⁶, Patrick F. Sullivan^{3,9}, Danielle
7 Posthuma^{4,10,  ,*}, Gerome Breen^{1,2,  ,*}

8

9 Author Affiliations:

10 ¹ MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry,
11 Psychology & Neuroscience, King's College London, London SE5 8AF, UK.

12 ² NIHR Biomedical Research Centre for Mental Health, South London and Maudsley
13 NHS Trust, London SE5 8AF, UK.

14 ³ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-
15 17177 Stockholm, Sweden.

16 ⁴ Department of Complex Trait Genetics, VU University, Center for Neurogenomics
17 and Cognitive Research, Amsterdam, 1081 HV, The Netherlands.

18 ⁵ Department of Child and Adolescent Psychiatry, Erasmus University Medical
19 Center, Rotterdam, the Netherlands

20 ⁶ Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and
21 Biophysics, Karolinska Institutet, SE-17177 Stockholm, Sweden.

22 ⁷ Center for Genomic and Computational Biology, Duke University, Durham, NC
23 27708, USA

24 ⁸ Department of Pediatrics, Medical Genetics Division, Duke University, Durham, NC
25 27708, USA

26 ⁹ Departments of Genetics, University of North Carolina, Chapel Hill, NC, 27599-
27 7264, USA.

28 ¹⁰ Department of Clinical Genetics, VU University Medical Center (VUMC),
29 Neuroscience Campus Amsterdam, Amsterdam, 1081 HV, The Netherlands.

30 ^{  } These authors contributed equally to the work

31 * Correspondence to: gerome.breen@kcl.ac.uk or danielle.posthuma@vu.nl

32

33 Running title: Secondary analysis of IQ GWAS

34

35 Conflict of Interest Statement

36 PF Sullivan reports the following potentially competing financial interests: Lundbeck
37 (advisory committee), Pfizer (Scientific Advisory Board member), and Roche (grant
38 recipient, speaker reimbursement). G Breen reports consultancy and speaker fees

Exploration and generalization in vast spaces

Charley M. Wu^{1,2,*,+}, Eric Schulz^{3,+}, Maarten Speekenbrink³, Jonathan D. Nelson^{1,4}, and Björn Meder^{1,5}

¹Center for Adaptive Behavior & Cognition, Max Planck Institute for Human Development; Lentzeallee 94, 14195 Berlin, Germany

²Center for Adaptive Rationality, Max Planck Institute for Human Development; Lentzeallee 94, 14195 Berlin, Germany

³Department of Experimental Psychology, University College London; 26 Bedford Way, London WC1H 0AP, United Kingdom

⁴School of Psychology, University of Surrey, 388 Stag Hill, Guildford GU2 7XH, UK

⁵MPRG iSearch, Max Planck Institute for Human Development; Lentzeallee 94, 14195 Berlin, Germany

*cwu@mpib-berlin.mpg.de

+these authors contributed equally to this work

ABSTRACT

Foraging for food, developing new medicines, and learning complex games are search problems with vast numbers of possible actions. Under time or resource constraints, optimal solutions are generally unobtainable. How do humans generalize and learn which actions to take when not all outcomes can be explored? We present two behavioural experiments and competitively test 27 models for predicting individual search decisions. We find that a Bayesian function learning model, combined with an optimistic sampling strategy, robustly captures how humans use generalization to guide search behaviour. Taken together, these two form a model of exploration and generalization that leads to reproducible and psychologically meaningful parameter estimates, providing novel insights into the nature of human search in vast spaces. Importantly, our modelling results and parameter estimates are recoverable, and can be used to simulate human-like performance, bridging a critical gap between human and machine learning.

Introduction

From engineering proteins for medical treatment¹ to mastering a game like Go², many complex tasks can be described as search problems³. Frequently, these tasks come with a vast space of possible actions, each corresponding to some reward that can only be observed through experience. In such problems, one must learn to balance the dual goals of exploring unknown options, while also exploiting existing knowledge for immediate returns. This frames the *exploration-exploitation dilemma*, typically studied using the multi-armed bandit framework^{4,5}, with the assumption that each option has its own reward distribution to be learned independently. Yet under real-world constraints of limited time or resources, it is not enough to know *when* to explore, but also *where*. How could an intelligent agent, biological or machine, learn which actions to take when not all outcomes can be explored?

There is an intriguing gap between human and machine learning, since humans are able to quickly learn and adapt to unfamiliar environments, where the same situation is rarely encountered twice^{6,7}. This contrasts with traditional approaches to reinforcement learning, which learn about the distribution of rewards for each state independently⁸. Such an approach falls short in more realistic scenarios where it is

*The multi-armed bandit is a metaphor for a row of slot machines in a casino, where each slot machine has an independent payoff distribution. Solutions to the problem propose different policies for how to learn about which arms are better to play (exploration), while also playing known high-value arms to maximize reward (exploitation).

DeepATAC: A deep-learning method to predict regulatory factor binding activity from ATAC-seq signals

Naozumi Hiranuma¹, Scott Lundberg¹, Su-In Lee^{1,2}

¹ Paul G. Allen School of Computer Science and Engineering,

² Department of Genome Sciences, School of Medicine, University of Washington

Abstract

Determining the binding locations of *regulatory factors*, such as transcription factors and histone modifications, is essential to both basic biology research and many clinical applications. Obtaining such genome-wide location maps directly is often invasive and resource-intensive, so it is common to impute binding locations from DNA sequence or measures of chromatin accessibility. We introduce DeepATAC, a deep-learning approach for imputing binding locations that uses both DNA sequence and chromatin accessibility as measured by ATAC-seq. DeepATAC significantly outperforms current approaches such as FIMO motif predictions overlapped with ATAC-seq peaks, and models based only on DNA sequence, such as DeepSEA. Visualizing the input importances for the DeepATAC model reveals DNA sequence motifs and ATAC-seq signal patterns that are important for predicting binding events. The Keras implementation and analysis pipelines of DeepATAC are available at <https://github.com/hiranumn/deepatac>.

1 Introduction

Knowing when and where proteins such as transcription factors interact with DNA is important for both clinical and research purposes. Biological assays such as ChIP-seq [5] are designed to directly measure these interactions, but they require significant resources and a large biological sample. To address these limitations several methods have been proposed to impute these binding locations from raw DNA sequence. The FIMO software [4] from the MEME suite, and deep-learning approaches such as DeepSea [11] and Basset [6] have been successful in this regard. However, DNA sequence alone does not contain any cell-type specific information, which is important for making more accurate predictions. This motivates combining DNA based predictions with predictions from cell type specific data sources.

There are several biological assays that measure cell type specific organization of genome, namely, DNase-seq, MNase-seq, HiC, and ATAC-seq [2, 7, 1, 3]. ATAC-seq is the most recent method and is rapidly gaining popularity due to its cost-efficiency and simplicity. In particular, ATAC-seq only requires 500 to 50,000 cells to measure in-vivo open chromatin signal, while other assays require millions of cells. This is a particularly important feature in clinical settings where you cannot sample a large number of cells from patients when performing personal level analysis. Traditionally, researchers have used putative binding locations predicted by motif finding algorithms (e.g FIMO) overlapped with ATAC-seq peaks to determine where transcription factors are bound. DeepATAC, a deep-learning model that is jointly trained on both ATAC-seq and DNA sequence, significantly outperforms this traditional approach.

Title: Neural entrainment determines the words we hear

Anne Kösem^{1,2*}, Hans Rutger Bosker^{1,2}, Atsuko Takashima^{1,2}, Antje Meyer^{1,2}, Ole Jensen^{2,3},
Peter Hagoort^{1,2}

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Radboud University, Donders Institute for Brain, Cognition, and Behaviour, Nijmegen, The Netherlands

³University of Birmingham, Centre for Human Brain Health, Birmingham, United Kingdom

*Corresponding author: a.kosem@donders.ru.nl

Ninety-nine independent genetic loci influencing general cognitive function include genes associated with brain health and structure (N = 280,360)

Gail Davies^{1,182}, Max Lam^{2,182}, Sarah E Harris^{1,3}, Joey W Trampush⁴, Michelle Luciano¹, W David Hill¹, Saskia PHagenaars^{1,5}, Stuart J Ritchie¹, Riccardo E Marioni^{1,3}, Chloe Fawns-Ritchie¹, David CM Liewald¹, Judith A Okely¹, Ari V Ahola-Olli^{6,7}, Catriona LK Barnes⁸, Lars Bertram⁹, Joshua C Bis¹⁰, Katherine E Burdick^{11,12,13}, Andrea Christoforou^{14,15}, Pamela DeRosse^{2,16}, Srdjan Djurovic^{14,17}, Thomas Espeseth^{18,19}, Stella Giakoumaki²⁰, Sudheer Giddaluru^{14,15}, Daniel E Gustavson^{21,22}, Caroline Hayward^{23,24}, Edith Hofer^{25,26}, M Arfan Ikram^{27,28,29}, Robert Karlsson³⁰, Emma Knowles³¹, Jari Lahti^{32,33}, Markus Leber³⁴, Shuo Li³⁵, Karen A Mather³⁶, Ingrid Melle^{14,18}, Derek Morris³⁷, Christopher Oldmeadow³⁸, Teemu Palviainen³⁹, Antony Payton⁴⁰, Raha Pazoki⁴¹, Katja Petrovic²⁵, Chandra A Reynolds⁴², Muralidharan Sargurupremraj⁴³, Markus Scholz^{44,45}, Jennifer A Smith^{46,47}, Albert V Smith^{48,49}, Natalie Terzikhan^{27,50}, Anbu Thalamuthu³⁶, Stella Trompet⁵¹, Sven J van der Lee²⁷, Erin B Ware⁴⁷, Beverly G Windham⁵², Margaret J Wright^{53,54}, Jingyun Yang^{55,56}, Jin Yu¹⁶, David Ames^{57,58}, Najaf Amin²⁷, Philippe Amouyel^{59,60,61}, Ole A Andreassen^{18,62}, Nicola Armstrong⁶³, John R Attia⁶⁴, Deborah Attix^{65,66}, Dimitrios Avramopoulos^{67,68}, David A Bennett^{55,56}, Anne C Böhmer^{69,70}, Patricia A Boyle^{55,71}, Henry Brodaty^{36,72}, Harry Campbell⁸, Tyrone D Cannon⁷³, Elizabeth T Cirulli⁷⁴, Eliza Congdon⁷⁵, Emily Drabant Conley⁷⁶, Janie Corley¹, Simon R Cox¹, Anders M Dale^{21,77,78,79}, Abbas Dehghan^{41,80}, Danielle Dick⁸¹, Dwight Dickinson⁸², Johan G Eriksson^{83,84,85,86}, Evangelos Evangelou^{41,83}, Jessica D Faul⁴⁷, Ian Ford⁸⁸, Nelson A Freimer⁷⁵, He Gao⁴¹, Ina Giegling⁸⁹, Nathan A Gillespie⁹⁰, Scott D Gordon⁹¹, Rebecca F Gottesman^{92,93}, Michael E Griswold⁹⁴, Vilmundur Gudnason^{48,49}, Tamara B Harris⁹⁵, Alex Hatzimanolis^{96,97,98}, Gerardo Heiss⁹⁹, Elizabeth G Holliday⁶⁴, Peter K Joshi⁸, Mika Kähönen^{100,101}, Sharon LR Kardina⁴⁶, Ida Karlsson³⁰, Luca Kleindam¹⁰², David S Knopman¹⁰³, Nicole Kochan^{36,104}, Bettina Konte⁸⁹, John B Kwok^{105,106}, Stephanie Le Hellard^{14,15}, Teresa Lee^{36,104}, Terho Lehtimäki^{107,108}, Shu-Chen Li^{109,110}, Tian Liu^{9,109}, Marisa Koini²⁵, Edythe London⁷⁵, Will T Longstreth,

Chemical control of a CRISPR-Cas9 acetyltransferase

Jonathan H. Shrimp¹, Carissa Grose², Stephanie R. T. Widmeyer², Ajit Jadhav³, Jordan L. Meier¹

¹Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA.

²Protein Expression Laboratory, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, MD 21702, USA.

³Division of Preclinical Innovation, National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland 20850, USA

Abstract:

Lysine acetyltransferases (KATs) play a critical role in the regulation of transcription and other genomic functions. However, a persistent challenge is the development of assays capable of defining KAT activity directly in living cells. Towards this goal, here we report the application of a previously reported dCas9-p300 fusion as a transcriptional reporter of KAT activity. First we benchmark the activity of dCas9-p300 relative to other dCas9-based transcriptional activators, and demonstrate its compatibility with second generation short guide RNA architectures. Next, we repurpose this technology to rapidly identify small molecule inhibitors of acetylation-dependent gene expression. These studies validate a recently reported p300 inhibitor chemotype, and reveal a role for p300's bromodomain in dCas9-p300-mediated transcriptional activation. Comparison with other CRISPR-Cas9 transcriptional activators highlights the inherent ligand tuneable nature of dCas9-p300 fusions, suggesting new opportunities for orthogonal gene expression control. Overall, our studies highlight dCas9-p300 as a powerful tool for studying gene expression mechanisms in which acetylation plays a causal role, and provide a foundation for future applications requiring spatiotemporal control over acetylation at specific genomic loci.

Introduction

Lysine acetyltransferases (KATs) catalyze protein acetylation, a reversible posttranslational modification (PTM) that plays a critical role in many processes, including gene expression.¹ Two of the most well-studied KATs are EP300 and its homolog CREBBP (commonly referred to jointly as p300/CBP). These two KATs possess a versatile substrate scope which includes histones, transcription factors, and members of the transcriptional regulatory apparatus itself.² Accordingly, disruption of p300/CBP is associated with substantial changes in gene expression, and has been linked to several diseases.³⁻⁴ Besides its KAT domain, p300 and CBP additionally contain several non-catalytic modules including zinc fingers, acetylysine readers (bromodomain, BRD), methyllysine readers (PHD domain), and protein-protein interaction domains.² Thus, a significant challenge in the study of p300/CBP lies in defining the specific role of the KAT domain in gene expression, as well as its targetable role in disease.

Considering methods to study cellular KAT activity, we were inspired by a recent report by Gersbach et al. which found that p300 could be delivered to specific genomic loci using the genomic-targeting methodology CRISPR-Cas9.⁵ Specifically, this study engineered a catalytically inactive variant of *S. pyogenes* Cas9 (dCas9) fused to truncated p300 module containing the BRD and KAT domains (dCas9-300) (Figure 1). Expression of this fusion in combination with chimeric short guide RNAs (sgRNAs) targeted to promoter regions led to

Simulating extracted connectomes

Jonathan Gornet and Louis K. Scheffer

Howard Hughes Medical Institute

(Dated: August 16, 2017)

Connectomes derived from volume EM imaging of the brain can generate detailed physical models of every neuron, and simulators such as NEURON or GENESIS are designed to work with such models. In principal, combining these technologies, plus transmitter and channel models, should allow detailed and accurate simulation of real neural circuits. Here we experiment with this combination, using a well-studied system (motion detection in *Drosophila*). Since simulation requires both the physical geometry (which we have) and the models of the synapses (which are not currently available), we built approximate synapses corresponding to their known and estimated function. Once we did so, we reproduced direction selectivity in T4 cells, one of the main functions of this neural circuit. This verified the basic functionality of both extraction and simulations, and provided a biologically relevant computation we could use in further experiments. We then compared models with different degrees of physical realism, from full detailed models down to models consisting of a single node, to examine the tradeoff of simulation resources required versus accuracy achieved.

Our results show that much simpler models may be adequate, at least in the case of medulla neurons in *Drosophila*. Such models can be easily derived from fully detailed models, and result in simulations that are much smaller, much faster, and accurate enough for many purposes. Biologically, we show that a lumped neuron model reproduces the main motion detector operation, confirming the result of Gruntman[1], that dendritic computation is not required for this function.

INTRODUCTION

Connectomes, maps of biological neural networks in a computer, are derived from volume imaging of the brain and include very detailed physical models of each neuron, such as the portion of an extracted neuron shown in Fig. 1(a). Simulators such as NEURON[2] or GENESIS[3] are explicitly designed to work with physical models, and compute results that depend on physical parameters. Driving a simulator such as NEURON with the output of EM reconstruction should therefore be able to reproduce the operation of biological circuits. However, straightforward attempts to do this run into several obstacles.

First, in addition to geometry, concrete numerical models of synapse operation are required. EM reconstruction can give synapse locations, but does not tell how they operate (or in *Drosophila*, even the sign - inhibitory and excitatory synapses look the same). Second, the use of detailed physical models brings additional concerns. The sheer number of nodes, plus the wide range of time constants between short and long segments, creates systems of equations that are hard to solve efficiently with numerical techniques. This not a problem unique to biological systems - circuit networks extracted from integrated circuits share the same concerns, and explicit techniques to avoid this problem have been used[4].

We investigated these problems by using the results from EM reconstruction to drive the simulator NEURON, to try to reproduce a known circuit operation - motion selectivity of the T4 cells in the medulla of *Drosophila*. We ran directly into the problems described above. First, we could not find in the literature detailed models for the graded synapses found in this circuit. Therefore we created analytic synapse models, tuned to

get approximately the responses shown in the literature. Next, we found that if we used the fully detailed geometrical models, then the run times of the simulator were prohibitive. To proceed, we had to reduce the geometrical complexity of the extracted neurons. On a positive note, once we added plausible synapse models to our simplified geometrical models, we were able to reproduce major portions of the known network function.

To examine the tradeoff of geometrical complexity versus accuracy, we compared fully detailed simulations with several simpler models. These included both a simpler branched model and a model with a single lumped node. These simplified models are much smaller, much faster to simulate, and give nearly the same results for the neurons we consider here.

EXPERIMENTAL DESIGN

For this experiment, we chose a portion of the visual pathway of the *Drosophila* fly brain, since it has both a detailed connectome[5], and a wide variety of detailed experimental and theoretical data. In particular we decided to try to reproduce the motion selectivity of the T4 cells. These cells react strongly to motion in the sensitive direction, and less strongly to other stimuli, including motion in the opposing direction, motion at right angles to the sensitive direction, or a uniform flash across the visual field. The T4 circuit is complex, with at least 8 differing cell types providing input, and the operation is still not fully understood[6][7][8][9].

The physical structure of each neuron in the network was imported as an SWC file, generated by the reconstruction of Takemura, et al.[5]. This network contains

Efficient *in situ* barcode sequencing using padlock probe-based BaristaSeq

Xiaoyin Chen¹, Yu-Chi Sun¹, George M Church^{2,3}, Je Hyuk Lee¹, and Anthony M Zador¹

¹ Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

² Wyss Institute, Harvard Medical School, Boston, Massachusetts, USA.

³ Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

Correspondence and requests should be addressed to Xiaoyin Chen, xichen@cshl.edu.

Inter-homologue repair in fertilized human eggs?

Dieter Egli^{1,*}, Michael V. Zuccaro², Michal Kosicki³, George M. Church⁴, Allan Bradley³, and Maria Jasin^{5,*}

¹Department of Obstetrics and Gynecology and Department of Pediatrics, Columbia University, New York NY 10032, USA; ²Graduate Program, Department of Physiology and Cellular Biophysics, Columbia University, New York NY 10032, USA; ³Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, United Kingdom; ⁴Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA, ⁵Developmental Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA

* correspondence: de2220@cumc.columbia.edu, m-jasin@ski.mskcc.org

Many human diseases have an underlying genetic component. The development and application of methods to prevent the inheritance of damaging mutations through the human germline could have significant health benefits, and currently include preimplantation genetic diagnosis and carrier screening. Ma et al. take this a step further by attempting to remove a disease mutation from the human germline through gene editing¹. They assert the following advances: (i) the correction of a pathogenic gene mutation responsible for hypertrophic cardiomyopathy in human embryos using CRISPR-Cas9 and (ii) the avoidance of mosaicism in edited embryos. In the case of correction, the authors conclude that repair using the homologous chromosome was as or more frequent than mutagenic nonhomologous end-joining (NHEJ). Their conclusion is significant, if validated, because such a “self-repair” mechanism would allow gene correction without the introduction of a repair template. While the authors’ analyses relied on the failure to detect mutant alleles, here we suggest approaches to provide direct evidence for inter-homologue recombination and discuss other events consistent with the data. We also review the biological constraints on inter-homologue recombination in the early embryo.

In their first approach, Ma et al. used donor sperm from a patient heterozygous for the *MYBPC3*^{ΔGAGT} mutation to fertilize wild-type oocytes, such that half of the embryos started out as wild type at the *MYBPC3* locus and half heterozygous. Fertilized zygotes were injected with Cas9 and an sgRNA directed to create a double-strand break (DSB) in the mutant paternal allele. The authors report that 24% of the embryos at day 3 of development were mosaic, with some cells of the embryo containing the mutant paternal locus, either intact or modified by NHEJ, together with a wild-type locus. Remaining cells of the embryo contained only a detectable wild-type allele. While some zygotes were also co-injected with a wild-type, exogenous, single-stranded oligodeoxynucleotide template (ssODN) with two synonymous mutations, no mutations consistent with ssODN-templated repair were detected. Furthermore, ‘wild-type only’ cells were present at a similar frequency both in the presence and absence of the ssODN. The authors infer that these cells arose by homology-directed repair (HDR) of the mutant paternal allele using the wild-type maternal allele as a template, i.e., inter-homologue recombination, leading to gene correction.

In a second approach, earlier, MII-phase oocytes were coinjected with Cas9 complexes and donor sperm. In this case, mosaicism was not detected, except in a single embryo, which contained both ‘wild-type only’ cells and ones heterozygous for wild-type and ssODN-templated alleles. Although wild-type embryos were expected at 50% frequency, they appeared to comprise 72% of embryos. The authors

Reconstruction of developmental landscapes by optimal-transport analysis of single-cell gene expression sheds light on cellular reprogramming.

Geoffrey Schiebinger,^{1,11,*} Jian Shu,^{1,2,*†} Marcin Tabaka,^{1,*} Brian Cleary,^{1,3,*} Vidya Subramanian,¹ Aryeh Solomon,^{1,@} Siyan Liu,^{1,15} Stacie Lin,^{1,6} Peter Berube,¹ Lia Lee,¹ Jenny Chen,^{1,4} Justin Brumbaugh,^{5,7,8,9,10} Philippe Rigollet,^{11,12} Konrad Hochedlinger,^{7,8,9,13} Rudolf Jaenisch,^{2,3} Aviv Regev,^{1,6,13,†} Eric S. Lander^{1,6,14,†}

***These authors contributed equally to this work.**

†Corresponding author.

Email: lander@broadinstitute.org (E.S.L.), aregev@broadinstitute.org (A.R.); jianshu@broadinstitute.org (J.S.)

Author Affiliations

¹Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

²Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

³Computational and Systems Biology Program, MIT, Cambridge, MA 02142, USA

⁴Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139 USA

⁵Cancer Center, Massachusetts General Hospital, Boston, MA 02114 USA

⁶Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁷Department of Molecular Biology, Center for Regenerative Medicine and Cancer Center, Massachusetts General Hospital, Boston, MA 02114, USA

⁸Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

⁹Harvard Stem Cell Institute, Cambridge, MA 02138, USA

¹⁰Harvard Medical School, Boston, MA 02115, USA

¹¹MIT Center for Statistics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

¹²Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

¹³Howard Hughes Medical Institute, Chevy Chase, MD, USA

¹⁴Department of Systems Biology Harvard Medical School, Boston, MA 02125, USA

¹⁵Biochemistry Program, Wellesley College, Wellesley, MA, 02481, USA

Present Address:

@Weizmann Institute of Science, Rehovot, Israel

WHOLE-BRAIN VASCULATURE RECONSTRUCTION AT THE SINGLE CAPILLARY LEVEL

Antonino Paolo Di Giovanna¹, Alessandro Tibo², Ludovico Silvestri^{3,1}, Marie Caroline Müllenbroich^{3,1}, Irene Costantini¹, Anna Letizia Allegra Mascaro^{4,1}, Leonardo Sacconi^{3,1}, Paolo Frasconi², and Francesco Saverio Pavone^{1,5,3*}

*email: pavone@lens.unifi.it

¹European Laboratory for Non-linear Spectroscopy, University of Florence, Italy;

²Department of Information Engineering (DINFO), University of Florence, Italy;

³National Institute of Optics, National Research Council, Italy;

⁴Neuroscience Institute, National Research Council, Pisa, Italy

⁵Department of Physics and Astronomy, University of Florence, Italy

ABSTRACT

The distinct organization of the brain's vascular network ensures that it is adequately supplied with oxygen and nutrients. However, despite this fundamental role, a detailed reconstruction of the brain-wide vasculature at the capillary level remains elusive, due to insufficient image quality using the best available techniques. Here, we demonstrate a novel approach that improves vascular demarcation by combining CLARITY with a vascular staining approach that can fill the entire blood vessel lumen and imaging with light-sheet fluorescence microscopy. This method significantly improves image contrast, particularly in depth, thereby allowing reliable application of automatic segmentation algorithms, which play an increasingly important role in high-throughput imaging of the terabyte-sized datasets now routinely produced. Furthermore, our novel method is compatible with endogenous fluorescence, thus allowing simultaneous investigations of vasculature and genetically

Title: Large-scale screening of rare genetic variants in humans reveals frequent splicing disruptions

Authors:

Rocky Cheung^{1†}, Kimberly D. Insigne^{2†}, David Yao³, Christina P. Burghard², Eric M. Jones¹, Daniel B. Goodman⁴, Sriram Kosuri^{1,5*}

Affiliations:

¹ Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA

² Bioinformatics Interdepartmental Graduate Program, University of California, Los Angeles, CA 90095, USA

³ Genetics Graduate Program, Stanford University, Stanford, CA 94035, USA

⁴ Department of Microbiology and Immunology, University of California, San Francisco, CA 94143, USA

⁵ UCLA-DOE Institute for Genomics and Proteomics, Molecular Biology Institute, Quantitative and Computational Biology Institute, Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA 90095, USA

*To whom correspondence should be addressed. Tel: +1 310 825 8931; Email: sri@ucla.edu

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Droplet-based microfluidic analysis and screening of single plant cells

Ziyi Yu,^{a,*} Christian R. Boehm,^{b,*‡} Julian M. Hibberd,^b Chris Abell,^a Jim Haseloff,^b Steven J. Burgess,^{b,#} and Ivan Reyna-Llorens^{b,#}

a Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK. Email: ZY- zy251@cam.ac.uk, CA- ca26@cam.ac.uk

b Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK. Email: CB- cboehm@mpimp-golm.mpg.de, JMH- jmh65@cam.ac.uk, JH- jh295@cam.ac.uk, SJB- sburgess011@gmail.com, IRL- iar28@cam.ac.uk

* These authors contributed equally to this work.

Corresponding authors.

‡ Present address: *Max Planck Institute for Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam, Germany.*

Abstract

Droplet-based microfluidics has been used to facilitate high throughput analysis of individual prokaryote and mammalian cells. However, there is a scarcity of similar workflows applicable to rapid phenotyping of plant systems. We report on-chip encapsulation and analysis of protoplasts isolated from the emergent plant model *Marchantia polymorpha* at processing rates of >100,000 protoplasts per hour. We use our microfluidic system to quantify the stochastic properties of a heat-inducible promoter across a population

Enzymatic synthesis of gene-length single-stranded DNA

Rémi Veneziano^{1†*}, Tyson R. Shepherd^{1†}, Sakul Ratanalet^{1,2}, Leila Bellou¹, Chaoqun Tao¹, Mark Bathe^{1*}

¹Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*To whom correspondence should be addressed.

e-mail: mark.bathe@mit.edu and rvenezia@mit.edu

†These authors contributed equally to this work.

Single-stranded DNA (ssDNA) increases the likelihood of homology directed repair with reduced cellular toxicity, yet ssDNA synthesis strategies are limited by the maximum length attainable, as well control over nucleotide composition. Here, we apply purely enzymatic synthesis to generate ssDNA greater than 15 kb using asymmetric PCR, and illustrate the incorporation of diverse modified nucleotides for therapeutic and imaging applications.

Efficient ssDNA synthesis on the 10+ kb-scale is a major need for numerous biotechnology applications including templated homology directed repair for genome editing (1-4), systems-scale gene synthesis and cloning (5-9), and scaffolded DNA origami (10,11). Conventional ssDNA synthesis is performed using either chemical or enzymatic approaches. Chemical synthesis is currently limited to approximately 98% incorporation efficiency for each base addition and therefore limited to the production of ssDNA oligos up to only 200 bases(5). Standard enzymatic synthesis through ligation or polymerization yields double-stranded DNA (dsDNA) that requires additional steps to generate ssDNA. Commercially available ssDNA synthesis is limited to 2 kb from Integrated DNA Technologies, Inc. (IDT, Coralville, IA) or recommended up to 5 kb using a strandase enzyme-based approach from Takara Biosciences, Inc. (Mountain View, CA). Enzymatic or chemical approaches to denaturation of dsDNA to form ssDNA is an alternative approach to ssDNA production, but limited by low yield and high cost.

In contrast, asymmetric polymerase chain reaction (aPCR) offers the direct synthesis of ssDNA from an underlying dsDNA template and has been applied to generate ssDNA ranging from several hundred to several thousand nucleotides in length (12-15). aPCR differs from traditional PCR by having one primer (the forward primer) in molar excess over the second primer (the reverse primer). This approach has previously been applied to short ssDNA synthesis for aptamers and gene detection (12,15), and more recently to kb-scale ssDNA for scaffolded DNA origami (11). However, previous work was limited to 3.3 kb due to low enzyme processivity. Here, we overcome this limitation by using a highly-processive LongAmp Taq polymerase to achieve 15+ kb length ssDNA. Additionally, using a standardized protocol and rules-based primer design, we achieve pure product yields up to 690 ng per 50 µl reaction volume and demonstrate direct incorporation of chemically modified nucleotides for ssDNA applications in therapeutics and imaging that require base or backbone modifications.

High-fidelity polymerases such as Phusion® allow for long dsDNA synthesis in standard PCR, however, Phusion polymerase was unable to synthesize fragment large than 1kb ssDNA (**Fig. 1A**

Biological Insights Into Muscular Strength: Genetic Findings in the UK Biobank

Emmi Tikkanen, PhD¹, Stefan Gustafsson, PhD², David Amar, PhD¹, Anna Shcherbina, M.Eng¹, Daryl Waggott, PhD¹, Euan A. Ashley, MD, PhD¹, Erik Ingelsson MD, PhD¹

¹ Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, CA

² Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Sweden

Word count: 299 (abstract), 3690 (main text)

Address for Correspondence:

Erik Ingelsson, MD, PhD, FAHA

300 Pasteur Dr, mail code: 5773; Stanford, CA 94305; USA

Phone: +1-650-656-0089; E-mail: eriking@stanford.edu

Short Title: Biological Insights Into Muscular Strength

Key words: Genome-wide association, genetics, grip strength, fitness

A cell-free synthetic biochemistry platform for raspberry ketone production

Simon J Moore^{1,3}

Tomasso Tosi³

David Bell^{1,3}

Yonek B Hleba^{1,2}

Karen M Polizzi^{1,2}

Paul S Freemont^{1,3}

¹Centre for Synthetic Biology and Innovation, Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK

²Department of Life Sciences, Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK

³Department of Medicine; Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK

Keywords

Synthetic biology, fine chemical, raspberry ketone, polyketide synthase, synthetic biochemistry

Efficient generation of targeted large insertions in mouse embryos using 2C-HR-CRISPR

Authors

Bin Gu^{1*}, Eszter Posfai^{1*} and Janet Rossant^{1,2**}

Affiliations

¹ Program in Developmental and Stem Cell Biology, Hospital for Sick Children, Toronto, ON M5G 0A4, Canada.

² Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada.

*Equal contribution

**Correspondence to Janet Rossant <janet.rossant@sickkids.ca>

Rapid and efficient generation of large fragment targeted knock-in mouse models is still a major hurdle in mouse genetics. Here we developed 2C-HR-CRISPR, a highly efficient gene editing method based on introducing CRISPR reagents into mouse embryos at the 2-cell stage, taking advantage of the likely increase in HR efficiency during the long G2 phase and open chromatin structure of the 2-cell embryo. With 2C-HR-CRISPR and a modified biotin-streptavidin approach to localize repair templates to target sites, we rapidly targeted 20 endogenous genes that are expressed in mouse blastocysts with fluorescent reporters and generated reporter mouse lines. We showcase the