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

My background

- Mostly software
- Some biotech, particularly through the "do-it-yourself biology" scene
- Currently working in the finance industry at LedgerX, a derivatives clearinghouse startup (Bitcoin)
- Strong interests in molecular biology, synthetic biology, gene therapy, molecular nanotechnology, neuroscience, brain uploading, cryonics, life extension research, kinematic self-replicating machines, open-source hardware, etc.
- Also a Bitcoin developer

Goal

- Share some ideas that have been cooking in my head
- Suggest some project ideas

Technique: Directed evolution & selection

- Couple genotype-phenotype in isolated compartments
- Selection pressure and measure/sort
- Eliminate some contenders, promote others
- Do many iterations
- Techniques: mRNA display, ribosome display, phage display, phage-assisted continuous evolution (PACE), cell surface display, in vitro compartmentalization, ...

Why does bruteforce directed evolution (sometimes) work?

- 10^9 to 10^13 small objects
- Billions and billions of possibilities that each respond to selection pressures
- Where has it worked?
  - green fluorescent protein (GFP), other FPs (red)
  - binding affinities, ligands/receptors, aptamers, antibodies
  - cell size
  - subtilisin thermostability (a protease for degrading proteins)

Technique: Mutagenesis and randomization

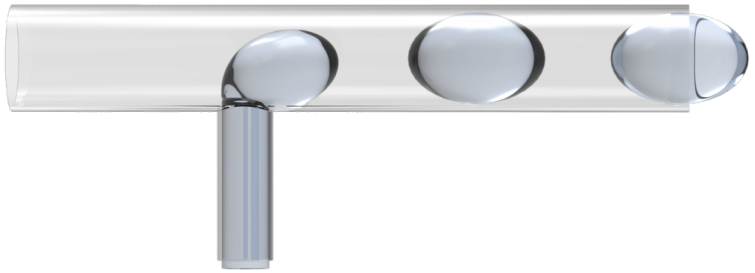
- Directed evolution needs a bit of a "push" sometimes.
- Lots of possible mutagens to use: error-prone polymerases, radiation, various toxins, expansion of the genetic alphabet, and a RNG (random number generator) on CPU before printing DNA, and many many more...
- Bottom line: speed up bruteforce search over genetic landscape by increasing the rate of mutation.
- (Also consider increasing total population size.)

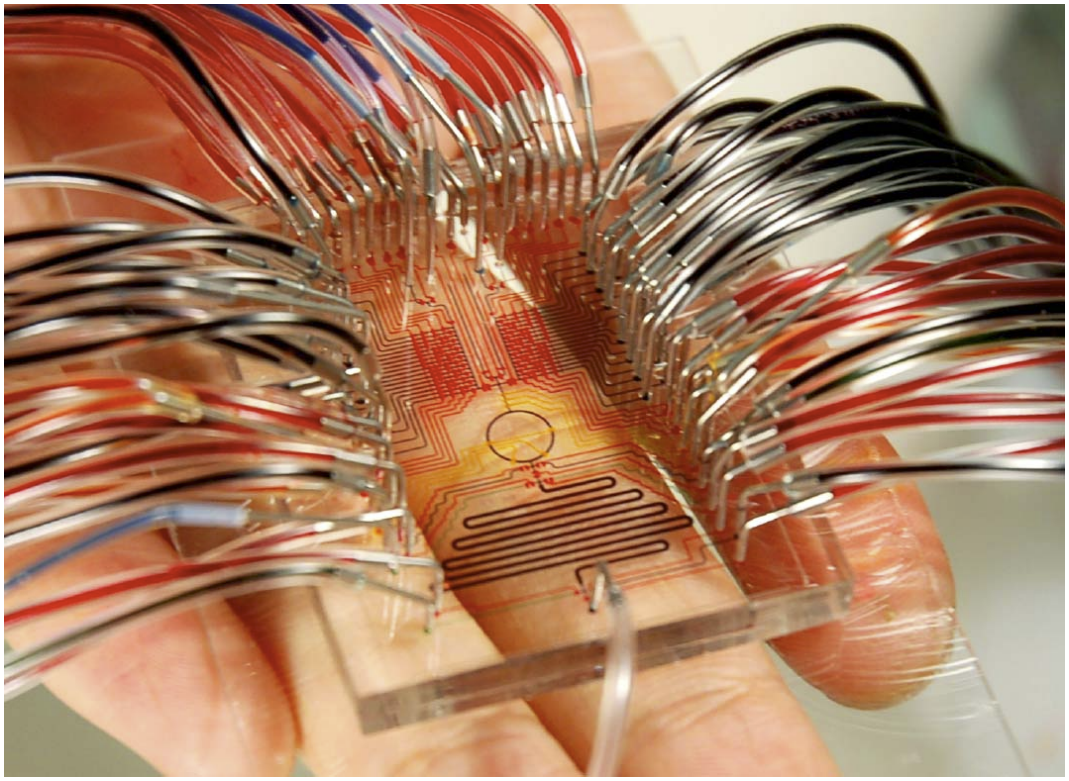
Technique: Rational design + selection

- Bruteforce mutagenesis is pretty great!
- We often already know roughly the outline that solutions might take
- Use "rational" genetic engineering to focus on areas of a genome or protein that are known to be somewhat related to the target problem.
- Sometimes "rational" approaches (and human guessing) can make interesting progress in rational protein design, even without bruteforce approaches
- If we know there needs to be a surface display mechanism, then we should insert it ourselves into the plasmid/genome.

Technique: Emulsion droplet microfluidics

- Droplet emulsion + beads + cells
- "Miniaturized lab"
- Downside: Valves tend to be difficult





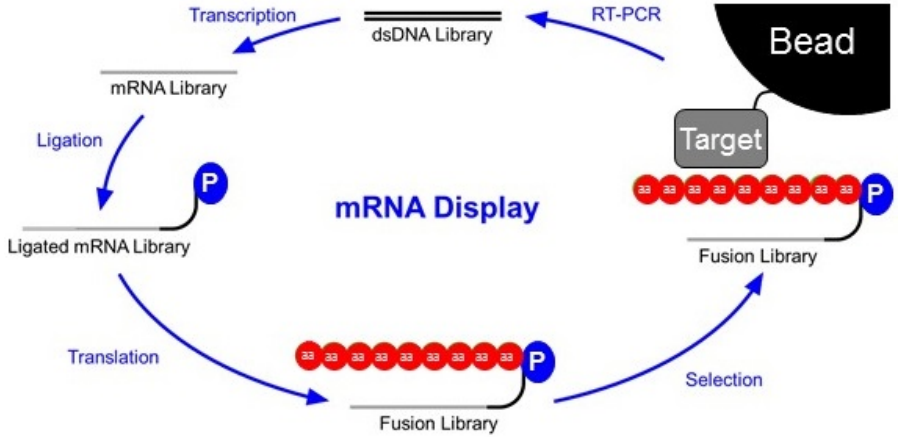
Now on to various project ideas I think about...

Selection-specific core facility?

- Either a facility at a school or maybe a company
- Offer general selection projects using common infrastructure-- mRNA display, phage display, microfluidics setups, etc.
- Besides engineering benefits, many research opportunities
- Enzyme optimization & strain optimization
- Lots of industrial opportunities-- agricultural, healthcare, plastics, oil, mining, recycling, etc. For example, insect farmers would appreciate insect size selections.

If not a core facility, then: easier mRNA display lab bench techniques

- Project: focus on reducing the difficulty of *-display techniques for directed evolution. Reach out to biologists and reduce the pain points; come up with a new protocol if necessary.
- This will save thousands of hours in the lab & better evolution will lead to better lab outcomes.



http://mrnadisplay.usc.edu/research/mrna-display/

"Young blood" rejuvenation?

- Short version: maybe "young blood" rejuvenates aged people? Some studies say yes/maybe, others say no.
- With directed evolution, perhaps "young blood" could be made to be rejuvenative?
- Possibly rejuvenative ingredients in blood: stem cells, bacteria, viruses, naked DNA, growth hormones, cell signalling molecules, …
- Selection: select for blood (or stem cells) that seem to have a repairing effect on adult members of the species.
- Increase the 'signal' of rejuvenation

Torpor? Cryonics?

- Instead of varying only the cryopreservation protocol, we should also vary the genetic material going into the cryonics procedure.
- Select for cells/tissues that survive cryoresuscitation. Already a 1% survival rate?
- Later move up to tissues, organs, multiple organs, and finally an entire animal.
- Not cryonics, but: we should sequence the genome of Rana sylvatica (Alaskan wood frog with interesting winter adaptation)

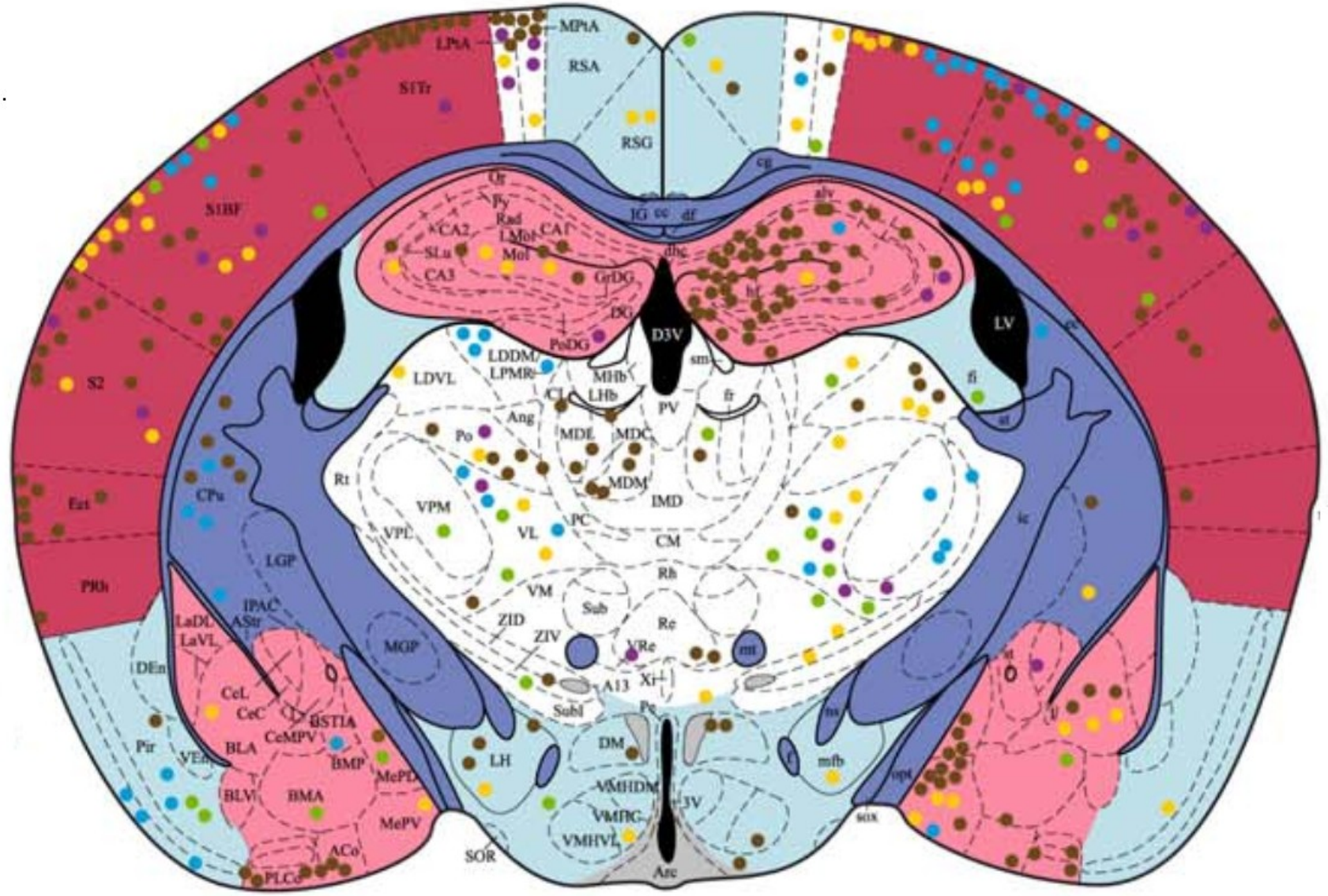
Cryonics, more examples

- Cytotoxicity – the vitrification chemicals tend to be cytotoxic, so select for resistance at increasingly higher levels
- Select (by eye?) lipid membrane structure before/after preservation, look for damage
- Post-resuscitation healing: cryonics likely to be highly damaging, need healing

Nootropics with microbes?

- My definition of a "smart drug" goal: user gets "smart" and makes a better smart drug.
- Neutralized brain eating microbe (or Toxoplasma gondii) as chasis + rationally-designed BDNF plasmids. Also intracellular parasites...
- Mouse maze solving selection
  - Recover microbes from mouse brain after evaluating maze solving performance
  - Select microbes from the better maze solvers, then mutate, then iterate rounds
  - Outcome: maybe mice that are really great at mazes? This one needs some tweaking: maze solving isn't precisely the goal of nootropics!
- Anti-sleep deprivation selection: brain microbes that keep mice awake?

Distribution of Toxoplasma gondii cysts in mouse brain coronal slice



http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0028925#s5

Technique: Aerosol atmospheric clouds?

- Use these as bioreactors for directed evolution
- Can span over multiple square/cube kilometers
- Lots of aerosol materials to choose from
- Really really large-scale directed evolution bioreactor... higher in the atmosphere you go, more volume becomes available.

DNA synthesis

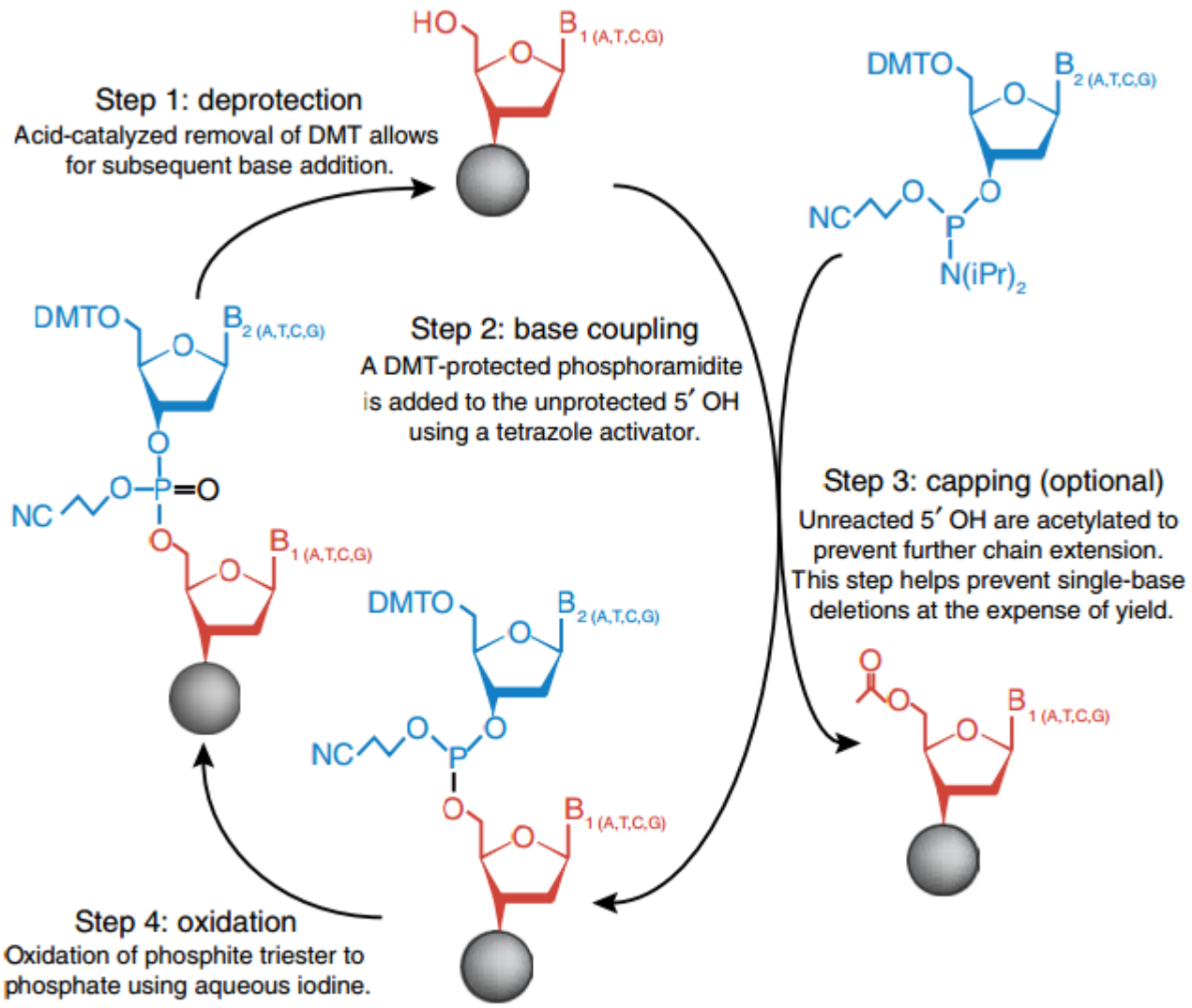


Figure 2 | Phosphoramidite chemistry. The four-step synthetic oligo synthesis is the most commonly used chemistry for the production of DNA oligos.

http://diyhpl.us/~bryan/papers2/DNA/oligonucleotide-synthesis-by-phosphoramidite-method-cycle-diagram.png

"Old school" DNA synthesizers

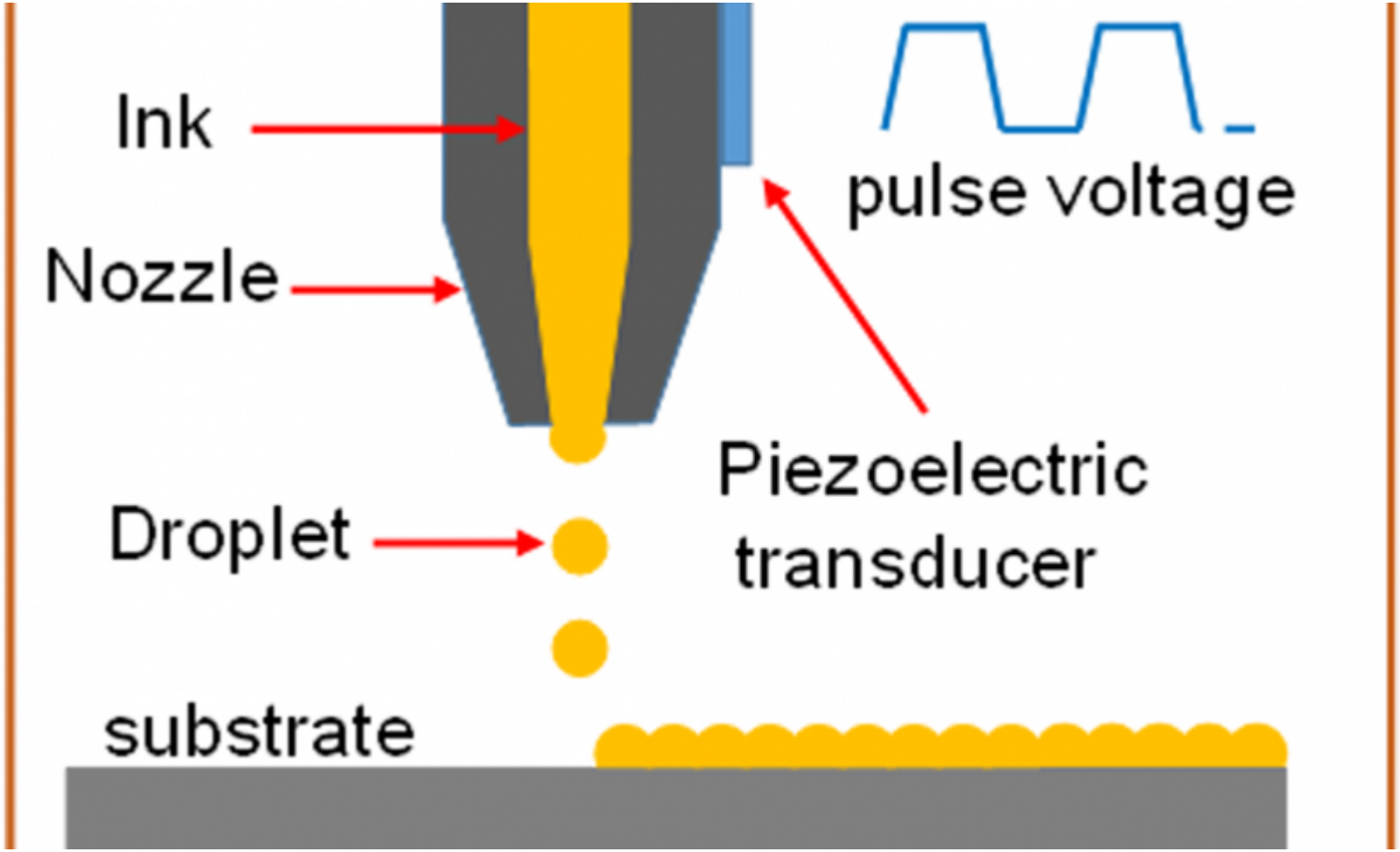
- High cost per base pair (BP), at least ~$0.10/bp

https://www.takeitapart.com/guide/94

Inkjet + phosphoramidite chemistry

- Long-term personal project: inkjet DNA synthesizer, using phosphoramidite chemistry and inkjet printer heads.
- Similar open-source inkjet oligo synthesizer already described in literature (POSAM), and others...

Inkjet oligonucleotide synthesizer (POSaM)



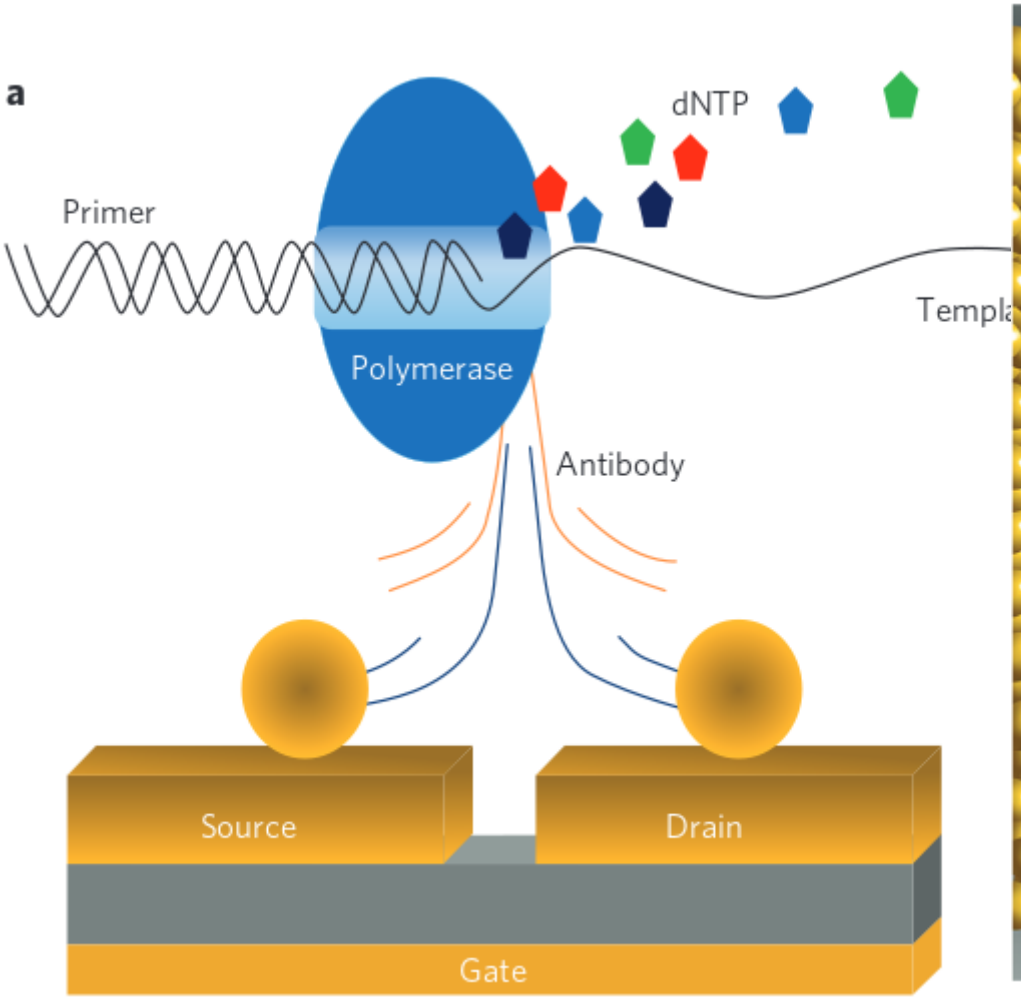
http://citizensciencequarterly.com/2011/06/open-source-inkjet-based-oligonucleotide-synthesizer-and-microarrayer/

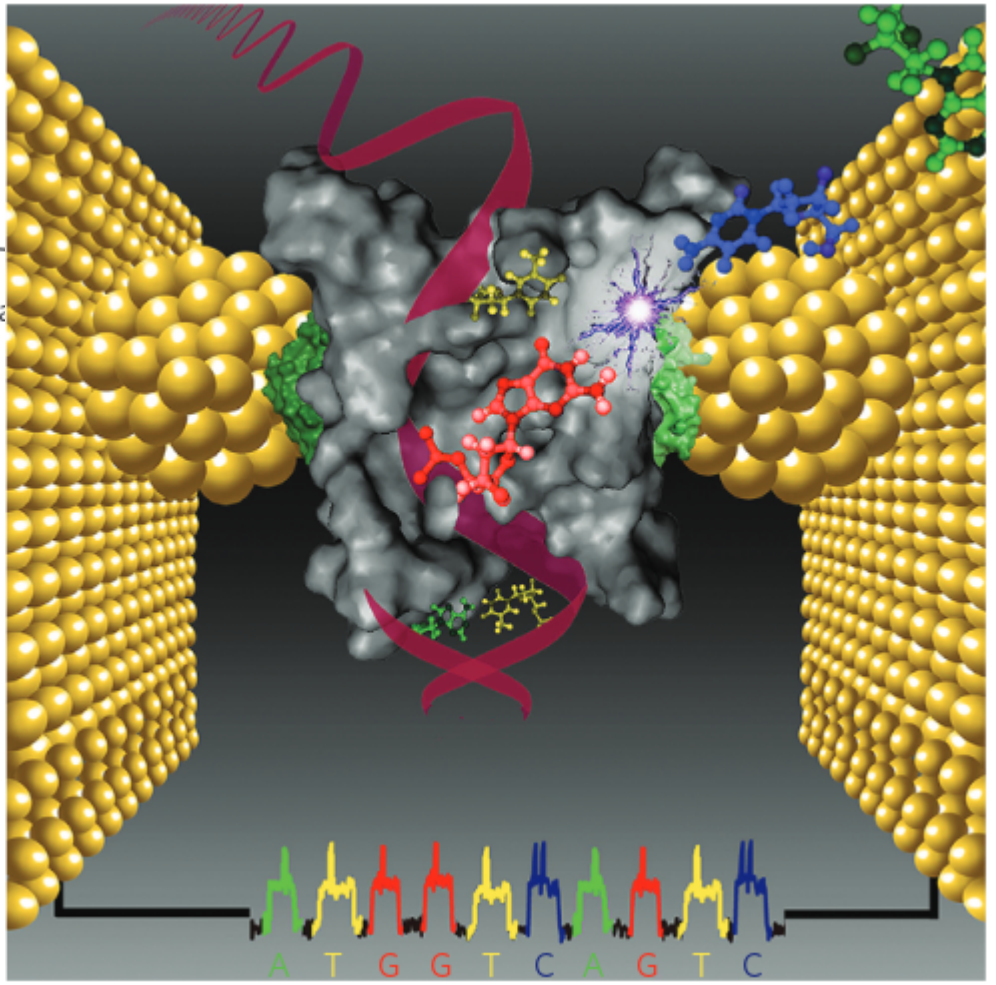
https://3c1703fe8d.site.internapcdn.net/newman/csz/news/800/2015/inkjetprinti.png

What about DNA synthesis in the future?

- I want $1 genome synthesis – billions of base pairs of DNA for $1 dollar
  - State of the art is >$10 million per custom genome
- Oligonucleotide synthesis is complex and doesn't work that well for genome-length DNA
- DNA polymerase

Electronic control of DNA polymerase





http://www.nature.com/nnano/journal/v8/n6/full/nnano.2013.71.html

Electronic control of DNA polymerase

- Idea: directly control a DNA polymerase enzyme to exactly synthesize a DNA molecule from user input.
- Small initiative at a school to explore electronic interfacing with DNA polymerase
- Something helpful: Azobenzene embedded in an amino acid in a DNA polymerase enyzme
- Applications of directed evolution to this problem
- I've been drafting a review paper on this topic. Happy to share it.

Homologous recombination for in vivo DNA assembly

- Short version: evolve a strain of yeast that can assemble genomes from millions of overlapping 100 bp fragments. Selection should start with smaller intermediate goals.
- Homologous recombination pathways not all explored yet, recA vs NHEJ vs other pathways.
- Use alternative genetic alphabet if necessary
- Select for DNA molecule length (random ligation), later select for better ligation ordering
- https://groups.google.com/d/msg/enzymaticsynthesis/uyZqtJO24RE/lApLb4JmCAAJ

Something completely different: Mechanical self-replication

- Goal: find the set of mechanical manufacturing processes that are able to reliably construct a replicate
- Graph grammars & automated design
- Vitamin parts-- what gets to be a vitamin?
- Robert Freitas – Kinematic Self-replicating machines (KSRM) and Advanced Automation for Space Missions (AASM)
- Biology already self-replicates, maybe we're stuck with biology for now.

Simulations for the directed evolution of AGI

- What were the most important selection pressures in human history?
- … Might be able to data mine selection pressures from scientific literature, then "auto" simulate these pressures in alife simulators.
- A modal sequence of selection pressures (not always simultaneous)
- Use existing results as springboard from neuroanatomy ("Consilience" paper) and connectomics (brain scanning) and computational neuroscience

Plausible sequences of selection pressures for human-level AGI?

- Foraging before coordinated hunting
- Sounds and hearing before language
- Rituals and superstition before development of scientific methods
- What were the important selection pressures in human history...? Can we replicate these?

The hplusroadmap group

- Self-funded
- We do mad science
- Recently: open-source ultrasound imaging device, open-source electroporator, etc.
- http://diyhpl.us/wiki/hplusroadmap

End

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

Quantification?

- Context: selection projects, digital or biological
- Effects of population size vs mutational load on speed towards discovering certain solutions?
- At what rate does it make sense to increase population size vs increasing mutational load?
- Is this decidable?
- If we know the answer (the phenotype landscape), then can we quantify some of these details?

Directed evolution towards novel forms?

- Q: Have we demonstrated microevolution of new forms of order? Or are we still stuck at optimization of existing proteins/forms?

The evolution of self-replicating computer organisms

A.N. Pargellis

Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA

Received 1 September 1995; revised 23 December 1995; accepted 19 March 1996
Communicated by A.M. Albano

Abstract

A computer model is described that explores some of the possible behavior of biological life during the early stages of evolution. The simulation starts with a primordial soup composed of randomly generated sequences of computer operations selected from a basis set of 16 opcodes. With a probability of about 10^{-4}, these sequences spontaneously generate large and inefficient self-replicating "organisms". Driven by mutations, these protobiotic ancestors more efficiently generate offspring by initially eliminating unnecessary code. Later they increase their complexity by adding additional subroutines as they compete for the system's two limited resources, computer memory and CPU time. The ensuing biology includes replicating hosts, parasites and colonies.