The Future of Life Sciences: Fluidics & Imaging
9 AM 9-Jun-2010 2010 Caliper Owners Group (COG) Meeting

Thanks to:

National Cancer Institute
National Human Genome Research Institute
MIT
BROAD INSTITUTE
Harvard Stem Cell Institute
ArmRev.org
LSRF
BBC
LS9, INC.
Agilent Technologies
Roche
Complete Genomics
23andMe
Helicos
Counsyl
illumina
Knome
Coriell Institute for Medical Research
Bill & Melinda Gates Foundation
Openheimer Foundation
PersonalGenomes.org
FBRC
AFYMETRIX®
Enzymatics
A2A
Complete Genomics
RBH
Joule
Affymetrix

Read = = = = = = = I/O = = = = = = = Write
The Future of Life Sciences: Fluidics & Imaging

• Personal Genomes: Environments, Traits, stem cells, microbiome/immunome, imaging, medical data
  - Diverse Apps/Sample prep bottleneck
• Synthetic Life & (bio)nanotech:
  - Integrating BI/O: Reading, Writing, & Quantitative characterization
Reading & Writing Genomes:
First semi-synthetic plasmid 1978: $10/b
CGI Human diploid genome 2009: $1500 / 6Gb
(=7 logs/30y mostly since 2005)

- Human insulin
- Human growth hormone
- Alpha-interferon
- G-CSF
- TPA
- GM-CSF
- Gamma-interferon
- IL-2
- Erythropoietin
- Hepatitis B vaccine

NAR 1978
Sutcliffe & Church
(BR: Bolivar & Rodriguez)

(Amgen, Biogen, Genentech, etc)
5 factors of 10 in in 5 years

(Moore’s law)

1.5x/yr for electronics

vs

10x/yr for DNA Sequencing

>20 years ahead of the 1970-2004 exponential
# Ultra-low-cost sequencing

1. **Illumina-GA**  
   SbP Fluorescent read-length 2*110 bp

2. **AB-SOLiD**  
   SbL Longest ligation reads

3. **CGI**  
   SbL $2000 genome, rolony grid, 100Kb haplotypes

4. **Polonator**  
   SbL/P Open-source, $170K device, 100Mb haplotypes

5. **Roche-454**  
   SbP Long reads (>0.4 kb)

6. **Helicos**  
   SbP-sm High parallelism & quantitation

7. **Ion Torrent**  
   SbP $50K, small device

8. **Pacific Bio**  
   SbP-sm Long reads (>2.0 kb)

9. **Intelligent Bio**  
   SbP hexagonal grid

10. **Halcyon**  
    EM-sm Long reads (>Mb), $100 genome

11. **Genizon BioSci**  
    SbH in situ sequencing

12. **LightSpeed**  
    SbL 16X higher density, >10X speed

13. **Bionanomatrix**  
    SbP-sm Fluorescent mapping

14. **OxfordNanopore**  
    Pore-protein-sm small device

15. **Visigen**  
    SbP-sm Pol <> dNTP FRET

16. **ZS Genetics**  
    EM-sm Iodine labels

17. **Nabsys**  
    Pore-SbH-sm small device

18. **GE Global**  
    SbP-sm

19. **IBM**  
    Pore Si-sm small device

20. **Electronic Biosci**  
    Pore-protein-sm

21. **GnuBio**  
    SbP-picoliter droplets
Next gen sequencing started with a Caliper prototype

**1994** Start of Caliper – fluidics prototype with Nikon microscope in Church lab (2003)

**1999** *NAR*: In situ localized amplification and contact replication of many individual DNA molecules.

**2003** *Analyt. Biochem.* Fluorescent in situ sequencing on polymerase colonies

**2003** *Science*: Single Molecule Profiling of Alternative Pre-mRNA Splicing
Three approaches to Mbp continuity

Zhang et al 2006: in situ seq, dilution libraries, ss-DNA-EM (Halcyon grant)
145 morphological measures → in situ sequencing

Bakal, et al. Science 316, 1753
2nd-Gen Synthesis: off chips

$500 per 15Mbp

8K Xeotron Photo-Generated Acid
12K Combimatrix Electrolytic
120K Roche, Febit Photolabile 5'protection
244K Agilent Ink-jet standard reagents

Amplify pools of 50mers using flanking universal PCR primers & 3 paths to 10X error correction

Tian et al. 2004 Nature
Carr & Jacobson 2004 NAR
Smith & Modrich 1997 PNAS
Merging Next-gen Sequencing, Synthesis & Characterization

1. Select best sequences: 3E-3 to 1E-5
2. Bio-Device characterization
3. Cell sorting (FACS)
Systems Biology Challenge

NOT going from ONLY Genome Sequence to Prediction

Genome alleles

TRAITS (Phenome)
Engineering Genomes Environments Traits

One in a life-time genome + yearly (to daily) tests
Bio-weather map: Allergens, Microbes, Viruses

PersonalGenomes.org

PERSONAL GENOME 3M alleles → Personal stem-cells epigenome (RNA, mC)

Immunome

TRAITS (Phenome)

Microbiome
1) First/only open access data
2) Avoid over-promising on de-identification
3) 100% on Exam to assure informed consent
   (*Educate pre-consent rather than post-discovery*)
4) Genome sequence + epigenome
5) Multi-traits: imaging, RNA, microbe/immune
6) Cells available for personal functional genomics
7) IRB approval for 100,000 diverse volunteers
   501(c)(3)

   16,000 so far
Medical Genomics: Individually rare collectively common 10%

1540 genes are highly predictive & medically actionable (inherited & cancer) at ~$2K per gene. **Very few of these are on DTC SNP chips.** Why? PKU, Tay Sachs, Cystic Fibrosis, BRCA1/2, etc.

Pharmacogenomic drug/allele combinations: Herceptin, Iressa, ..

Also:
Ancestry, Forensics,
Social Networking,
Education, Research
GET evidence: 32 diploid exomes & genomes: hypertrophic cardiomyopathy allele

Figure 5.4: Summary of MYL2 A13T findings from clinical testing. This variant has been seen in two families and three sporadic cases based on our enquiries to clinical testing laboratories that sequence the MYL2 gene. In one family MYL2 A13T has been seen concurrently with MYBPC3 E619K and in the other family with MYH7 N1327K.

evidence.personalgenomes.org
PersonalGenomes.org
Inherited + Environmental Genomics

One in a life-time genome + yearly (to daily) tests

Public Health Bio-weather map: Allergens, Microbes, Viruses

PERSONAL GENOME
3M alleles

Personal stem-cells
epigenome
(RNA,mC)

VDJ-ome

TRAITS
(Phenome)

Microbiome
Microbiomes: What limits diagnostics

- Standard practice: skip diagnostics; guess at pathogen & antibiotics
- Biomarkers vs causative sequences.
- Ideally target pathogenicity & resistance
- Assay 6 nanoliters or 6 liters?
  (if <1 cell / ml)
Circulating tumor/pathogen/fetal cells

Yung, Ingber et al. Lab Chip, 2009
PGP Microbiome—Resistome: 18 Antibiotics

Dantás, Sommer, Church
Science 2009
Microbiome vs Immune-ome

Microbe tests: Detect Drug resistance spectrum
Earlier warning (e.g. meningitis)

Immune tests: Focus on response to exposure
Longer times to detect exposure (e.g. HIV, TB)
Vaccination Immune-ome

Harvard/MIT: Vigneault, Laserson, Lieberman-Aiden, Church
Roche: Egholm, Simen
Time Series Vaccine Experiment

Tracking human dynamic response to vaccination to 11 strains:
Hepatitis A+B, Flu A/Brisbane/59/2007 (H1N1)-like, 10/2007 (H3N2)-like, B/Florida/4/2006-like virus
Polio, Yellow fever
Meningococcus
Typhoid, Tetanus
Diptheria, Pertussis

Collect samples at -14d, 0d,
+1d, +3d,
+7d, +14d,
+21d, +28d
PGP Immune-ome
Self Organizing Map (SOM) clustering
PersonalGenomes.org
Inherited + Environmental Genomics

One in a life-time genome + yearly (to daily) tests

Public Health Bio-weather map: Allergens, Microbes, Viruses

PERSONAL GENOME
3M alleles

Personal stem-cells
epigenome
(RNA,mC)

VDJ-ome

TRAITS
(Phenome)

Microbiome
PGP Personal Stem cells to hepatic proteins & activity

Generation of Functional Human Hepatic Endoderm from Human Induced Pluripotent Stem Cells
Gareth et al (Daley, Church, Ian Wilmut labs)
Synthetic Life … & applications

• Minimal cells vs fast/robust/useful cells
• Immediate deployment vs lowering costs
• Whole genomes vs just-in-time synthesis
• Developing new safety & security methods
• Design vs evolution
Bio-petroleum from microbes flotation -- not distillation

Fatty acid derived

Gasoline & diesel for current engines & infrastructure

LS9, INC.
the renewable petroleum company

Bar chart showing distribution of organic and aqueous fractions.
Triumph of comparative genomics: identifying alkane biosynthetic enzymes

<table>
<thead>
<tr>
<th>Yeasts</th>
<th>Fat Content %</th>
<th>Molds</th>
<th>Fat Content %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candida lipolytica</td>
<td>36</td>
<td>Eutomophthora virulenta</td>
<td>26</td>
</tr>
<tr>
<td>Trichosporum cutaneum</td>
<td>45</td>
<td>Aspergillus flavus</td>
<td>28</td>
</tr>
<tr>
<td>Candida curvata</td>
<td>58</td>
<td>Pytium ultimum</td>
<td>49</td>
</tr>
<tr>
<td>Lipomyces lipferus</td>
<td>63</td>
<td>Fusarium bulbigenum</td>
<td>50</td>
</tr>
<tr>
<td>Endomyces vernalis</td>
<td>65</td>
<td>Aspergillus fischeri</td>
<td>53</td>
</tr>
<tr>
<td>Rhodotorula glutinis</td>
<td>71</td>
<td>Penicillium lilacinum</td>
<td>56</td>
</tr>
</tbody>
</table>

Botryococcus braunii decarboxylase

Methylelecosene from *Prasiola stipitata*

http://www.biofuelsdatabase.org/map/alkane-decarboxylation_map.shtml
http://www.springerlink.com/content/p6451qx982638856/fulltext.pdf
**Multiplex Automated Genome Engineering (MAGE)**

**Allelic Replacement**
- Strain: MG1655, ΔmutS, integrated λ-Red
- Highly complex oligo pools for multiplexed multi-loci modifications
- >4 billion bp of targeted genetic variation produced per day

**Optimized Parameters**
- Oligo length: 90mer
- Oligo 2ndary structure: <12 kcal/mol
- Oligo half-life: 5’ phosphothiol bps
- Oligo conc.: up to 50 uM
- Cycle time: 2 to 2.5 hrs
- >30% efficiency per cycle

---

Wang HH, in preparation, 2009
Accelerated Evolution 23K combinations per gene

Lycopene (hydrocarbon): 20 genes up, 4 down, 2 new

Harris Wang H et al Nature 2009
New genetic code for New AA & Multi-virus resistance

<table>
<thead>
<tr>
<th>Virus</th>
<th>Length (b)</th>
<th>Type</th>
<th>TAG / total</th>
</tr>
</thead>
<tbody>
<tr>
<td>φX-174</td>
<td>5,386</td>
<td>ss-DNA</td>
<td>0 / 9</td>
</tr>
<tr>
<td>M13</td>
<td>6,407</td>
<td>ss-DNA</td>
<td>1 / 10</td>
</tr>
<tr>
<td>MS2</td>
<td>3,569</td>
<td>ss-RNA</td>
<td>2 / 4</td>
</tr>
<tr>
<td>λ</td>
<td>48,502</td>
<td>ds-DNA</td>
<td>3 / 66</td>
</tr>
<tr>
<td>T7</td>
<td>39,937</td>
<td>ds-DNA</td>
<td>6 / 60</td>
</tr>
<tr>
<td>T4</td>
<td>168,903</td>
<td>ds-DNA</td>
<td>19 / 277</td>
</tr>
<tr>
<td>E.coli</td>
<td>4,639,675</td>
<td>ds-DNA</td>
<td>314 / 1,360,152</td>
</tr>
</tbody>
</table>

Farren Isaacs

ncbi.nlm.nih.gov/nuccore/9626372
56718463 176120924 J02459 9627425 29366675 (7 tRNAs: RITSPGL)
MAGE System Upgrade Jan-May 2010

Dimension:
4’ x 3’ x 2.5’

Sample Storage

96-well Electroporator

Motion 1

DNA Queue

Motion 2

Incubation

Reagents:
H2O, EtOH, Growth Media

Expression Reagent

Microfiltration Plate

Electroporator Head Unit

Waste

Harris Wang

Copyright 2010 – Boston Engineering Corporation
Project: HAR002.P2
The Future of Life Sciences: Fluidics & Imaging

• Personal Genomes: Environments, Traits, stem cells, microbiome/immunome, imaging, medical data
  - Diverse Apps/Sample prep bottleneck

• Synthetic Life & (bio)nanotech:
  - Integrating BI/O: Reading, Writing, & Quantitative characterization