Accelerated evolution, proliferation & mammalian development. 14-Oct-2020, 10AM George Church Cancer & Evolution Symposium

George Church
Full COI list: v.ht/PHNc

Lipper 1997
Foundation
Nucleic Acid Read / Write cost & quality improving exponentially, faster than Moore’s law

1E-3
1E-2
1E-1
1E+0
1E+1
1E+2
1E+3
1E+4
1E+5
1E+6
1E+7


60 million-fold+

Oligo bp/$

Seq bp/$

2009-12 diploid PGP
1E-7 error rate

1984 HGP haploid
1E-2 error rate

$3 billion

Multiplex NGS in 2003

$0 DigiD8
$100 BGI
$299 Nebula 2020
6Gb

Multiplex Chip-Syn 2004

$0

3Gb
### Therapies → Low cost alternatives

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Source</th>
<th>Genes</th>
<th>Disease</th>
<th>US/yr</th>
<th>$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luxturna</td>
<td>Spark</td>
<td>AAV2/RPE65</td>
<td>LCA-blindness</td>
<td>80</td>
<td>425k/eye</td>
</tr>
<tr>
<td>Imlygic</td>
<td>Amgen</td>
<td>HSV1/GM-CSF</td>
<td>Melanoma</td>
<td>7200</td>
<td>780k/6mo</td>
</tr>
<tr>
<td>Kymriah</td>
<td>Novartis</td>
<td>CAR(to CD19)</td>
<td>Leukemia</td>
<td>9600</td>
<td>475k</td>
</tr>
<tr>
<td>Yescarta</td>
<td>Kite Pharma</td>
<td>CAR(to CD19)</td>
<td>Lymphoma</td>
<td>8000</td>
<td>373k</td>
</tr>
<tr>
<td>Zolgensma</td>
<td>Novartis</td>
<td>AAV9/SMN1</td>
<td>SMA</td>
<td>500</td>
<td>2.1M</td>
</tr>
<tr>
<td>Spinraza</td>
<td>Biogen</td>
<td>ASO</td>
<td>SMA</td>
<td>500</td>
<td>4M/10yr</td>
</tr>
<tr>
<td>Multiplex</td>
<td>Rejuvenate</td>
<td>AAV/3genes</td>
<td>Multiplex</td>
<td>2,800,000</td>
<td>2k ?</td>
</tr>
</tbody>
</table>

1) Rare genetic diseases: Prevention: dating/matchmaking
2) Infectious diseases: Prevention: 24-7 monitoring
3) Common diseases: (age-affected) spread R&D costs
4) Less failure: ML+Libraries: delivery, action, testing
Preventing common diseases: Cancer

Pre-conception / pre-implantation genetics
Adult genetics: Pre-symptomatic surgeries (BRCA1/2, CDH1)
Environmental: Smoke, UV, other carcinogens
8 of 11 infectious cancer causes are viral.
(v) vaccine, (e) editing, (s) small molecules, (h)hygiene

MCPyV: Skin
EBV: Lymphoma, stomach, nasopharynx
HBV (v,e): Hepatic
HCV: Hepatic
HHV8 (h): Sarcoma
HIV (e, h, s): Sarcoma
HTLV-1 (h): Leukemia
HPV (v): Cervical, vaginal, vulvar, anal, penile, oropharyngeal
Helicobacter pylori (s): Stomach, Lymphoma
Opisthorchis viverrini (h): Cholangiocarcinoma
Schistosoma hematobium (h): Bladder
1/64 Recoding $\rightarrow$ resistance to all viruses

**Viral Relative Titers**

<table>
<thead>
<tr>
<th>Virus</th>
<th>+UAG +RF1</th>
<th>ΔUAG +RF1</th>
<th>ΔUAG ΔRF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ cl857</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1vir</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1 c1-100 Tn9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Titer (PFU/mL)*

*Cells Systems 2016 Ma, Isaacs, et al*

*Science 2016 Ostrov et al*

*Science 2019 Ostrov et al*

*Nature 2015 Mandell et al*

*Science 2013 Lajoie et al*
Combinatorial gene therapies for multiple age-related diseases

GenAge
The Ageing Gene Database

João Pedro de Magalhães

1. High fat diet obesity
2. Type 2 Diabetes
3. Osteoarthritis
4. Cardiac damage recovery
5. Kidney disease

FGF21: glucose handling, sTGFβR2: Soluble form αKlotho: intracellular calcium, serum mineral-ion homeostasis

→ 5 aging diseases via combinatorial gene therapy
AAV →

Davidsohn et al. (2019) PNAS
Martinez-Redondo et al. (2020) Protein Cell.
Accelerated Evolution via multiplex-editing, ML, libraries

Addition, subtraction, precise-editing, epigenetics

Transgenics + cisgenics

Nucleases (error-prone)

De-aminases C→T, A→G

Recombinases & Primer-Editor (precise & flexible)

Methylation, Activation, Repression/Silencing

Accelerated Basic Science → Translation
Recombinases & Integrases

λ Red recombinase: Gabe Filsinger & Tim Wannier found interaction with Single-strand binding proteins (SSB)
No custom protein or RNA (just donor DNA)

Small S/Y integrases: Ben Weinberg. ML-Syn
Custom protein (no RNA)
GP-Write: precise hyperediting (MAGE)
4 Million bp Genome Recoding: 63 & 57 codon types

1. Non-standard amino acids (NSAA)

2. Genetic & Metabolic Isolation

3. Multi-Virus resistance

- Science 2016 Ostrov, et al
- Nature 2015 Mandell et al
- Science 2013 Lajoie et al
<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Comment</th>
<th>Double hr</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Vibrio natrigens</em></td>
<td>-</td>
<td>Heterotroph</td>
<td>0.16</td>
</tr>
<tr>
<td><em>Synechococcus</em></td>
<td>~3%</td>
<td>11801 in BG11 media .04% CO₂</td>
<td>2.3</td>
</tr>
<tr>
<td><em>Aphanothece sacrum</em></td>
<td>-</td>
<td>Suizenji-nori cyanobac</td>
<td>15</td>
</tr>
<tr>
<td><em>Arthrospira platensis</em></td>
<td>60%</td>
<td>&quot;Spirulina&quot; no-N₂-fix cyanobac</td>
<td>24</td>
</tr>
<tr>
<td><em>Wolffia arrhiza</em></td>
<td>40%</td>
<td>duckweed-flower, watermeal, N₂-fix</td>
<td>30</td>
</tr>
<tr>
<td><em>Azolla pinnata</em></td>
<td>-</td>
<td>duckweed-fern, N₂-fix</td>
<td>46</td>
</tr>
<tr>
<td><em>Chlorella pyrenoidosa</em></td>
<td>-</td>
<td>no-N₂-fix</td>
<td>51</td>
</tr>
<tr>
<td><em>Nostochopsis lobatus</em></td>
<td>-</td>
<td>N₂-fix cyanobac</td>
<td>58</td>
</tr>
<tr>
<td><em>Pyropia tenera</em></td>
<td>39%</td>
<td>Red algae “Nori”</td>
<td>240</td>
</tr>
<tr>
<td><em>Palmaria palmata</em></td>
<td>-</td>
<td>Red algae</td>
<td>240</td>
</tr>
<tr>
<td><em>Nostoc commune</em></td>
<td>~3%</td>
<td>N₂-fix cyanobac</td>
<td>288</td>
</tr>
</tbody>
</table>
“Pre”-training on natural protein sequences enables learning of state-of-the-art virtual fitness landscapes

9 diverse proteins with diverse activities

Alley, Khimulya, Biswas et al. 2019
Multiplex testing thousands of biologic designs in a single animal (via barcodes)

ManifoldBio

Pierce Ogden  Gleb Kuznetsov  Shane Lofgren

Protein Quantitation Platform
Protein Design (AAV & Ab): Multiplex deep sequence testing

1.2 M capsid designs

Machine Learning

DNA Synthesis → Selection → DNA sequencing

Capsid packaging → Immune evasion → In vivo delivery → Tissue specificity

Pierce Ogden

Eric Kelsic

Science Nov 2019
Mapping the space of tissue tropism modifications

Ogden et al. Science 2019
Multiplex cell & epigenetic engineering via complete human transcription factor expression library

1,748 human TF ORFs → 290 cell type recipes so far

Oligodendrocyte
Myelin wrapping Axons
Human organoid transplant reversal & enhanced relative to original in resistance to demyelination

Ng, Khoshakhlagh, Shipman, Swiersy, Appleton, Huang, Saylor, Trono, Taipale, Hill, Vidal, Busskamp, Church (2020) in review
Universal Chimeric Antigen Receptor T-cells (UCART) (& Macrophages CAR-M)

Most B-cell malignancies express CD19, so target with CAR.

Multiplex editing (ZFN, TALEN, CRISPR): TRAC-, B2M-, PDCD1-, CD52-, DCK-

ZFEN: Torikai…Cooper Blood 2013 (U. TX)
TALEN: Valton…Poirot L. Molec Therap 2015 (Cellectis)
CRISPR: Liu…Wang Cell Research 2-Dec-2016 (Beijing)
Enhanced transplants & cell therapies.
Multiplex (42-plex) germline editing

**Sugars:** GGTA1, CMAH, β4GalNT2

**Clotting:** human TFPI, TBM, EPCR, vWF

**MHC:** HLA/SLA class I, II

**Immune functions:** human CTLA4-Ig, HLA-E/G/Cw3 inhibit NK cells

**Complement:** regulatory genes CD46, CD55, CD59

**Porcine ERVs:** 25 in normal fibroblasts

MGH:
Jim Markmann

Luhan Yang, et al 2015 Science
Dong Niu, et al. 2017 Science
Yanan Yue et al. 2020 Nat BME
# Multiplex editing to reduce somatic mutations

<table>
<thead>
<tr>
<th>Category</th>
<th>#/ hum cell</th>
<th>Length</th>
<th>Total bp</th>
<th>Previous mutation attempts</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERV</td>
<td>25</td>
<td>7500</td>
<td>2.4E+7</td>
<td>Cas9 Yang .. Science 2015, 2017</td>
</tr>
<tr>
<td>Telomeres</td>
<td>46</td>
<td>8000</td>
<td>3.7E+5</td>
<td>TERT Ramunas .. FASEB J 2015</td>
</tr>
<tr>
<td>rDNA repeats</td>
<td>300</td>
<td>43000</td>
<td>1.3E+7</td>
<td>I-CreI Paredes .. Genetics 2009</td>
</tr>
<tr>
<td>UCE</td>
<td>855</td>
<td>200</td>
<td>1.7E+5</td>
<td>HDR</td>
</tr>
<tr>
<td>LINEs</td>
<td>26,000</td>
<td>7000</td>
<td>1.8E+8</td>
<td>dABE Smith .. Biorxiv 2019</td>
</tr>
<tr>
<td>Centromeres</td>
<td>1.0E+6</td>
<td>171</td>
<td>1.7E+8</td>
<td>Cas9 Adikusuma .. Mol Ther 2017</td>
</tr>
<tr>
<td>SINEs (Alu)</td>
<td>1.5E+6</td>
<td>280</td>
<td>4.2E+8</td>
<td>Natural variants</td>
</tr>
<tr>
<td>SSR</td>
<td>3.0E+6</td>
<td>16</td>
<td>4.8E+7</td>
<td>Cas9 [14] Monteys..Mol Ther 2017</td>
</tr>
<tr>
<td>Triplex sites</td>
<td>1.7E+7</td>
<td>20</td>
<td>3.4E+8</td>
<td></td>
</tr>
</tbody>
</table>

**Repeats involved in senescence, neurogenesis, cancer, inflammation**

Wang et al. Cell Cycle 2011  
DeCecco et al. Nature 2019  
Singer et al. Trends Neurosci. 2010
Enabling large-scale genome editing by reducing DNA nicking

Deamination

C$\rightarrow$T  $5'$ –ATTCTACCAGAGGTACAAGGAGG–3'$
A$\rightarrow$G  $5'$ –ATTCTGCCAGAGGTACAAGGAGG–3'

1. No nick. dCas9
2. No MMR
3. No U-glycosylase
4. Anti-apoptotic molecules
5. Growth factors (bFGF)

Nick-less dABE targeting of LINE-1 in PGP1 hiPSCs

Single cells - Day 11

dABE: 60% = 23k of 38k sites (diploid)
Abstract: We have developed methods using human transcription factor expression libraries to accelerate some developmental processes from 300 days to 4 days and have used these semi-synthetic tissues to accurately evaluate late onset (70 year) diseases like Alzheimer's. We have easily engineerable proliferative systems with doubling times of only 15 minutes that we would like to adapt to display analogous developmental processes. We have accelerated evolution way of a combination of gigabase-scale computer-controlled DNA synthesis plus machine learning. We have used this to design and test millions of viral capsids for improved AAV gene therapies.