J. CRAIG VENTER

LIFE AT THE SPEED OF LIGHT

FROM THE

DOUBLE HELIX

TO THE

DAWN OF DIGITAL LIFE

BY THE AUTHOR OF **A LIFE DECODED**

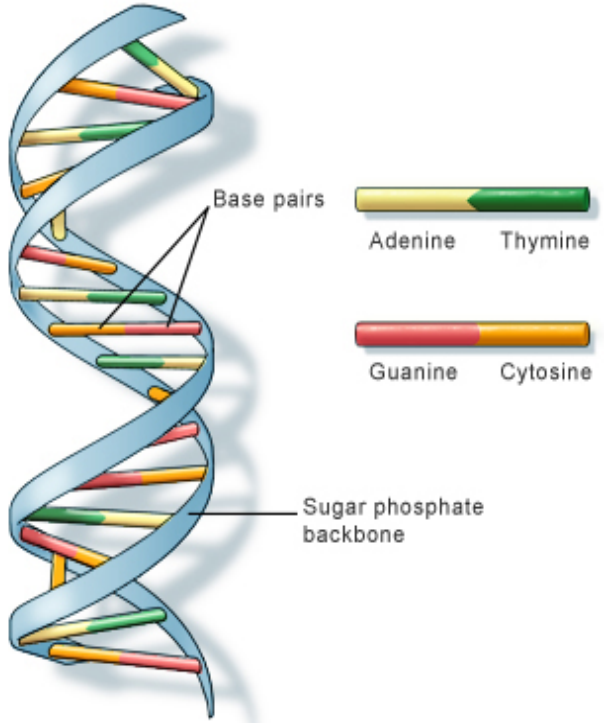
From Genotype to Phenotype

Can the Genetic Code Predict Present and Future Phenotypes?

CONVERTING THE ANALOG GENETIC CODE INTO DIGITAL CODE

GENETIC CODE

ATGAATTCGACTTACCGTAGG

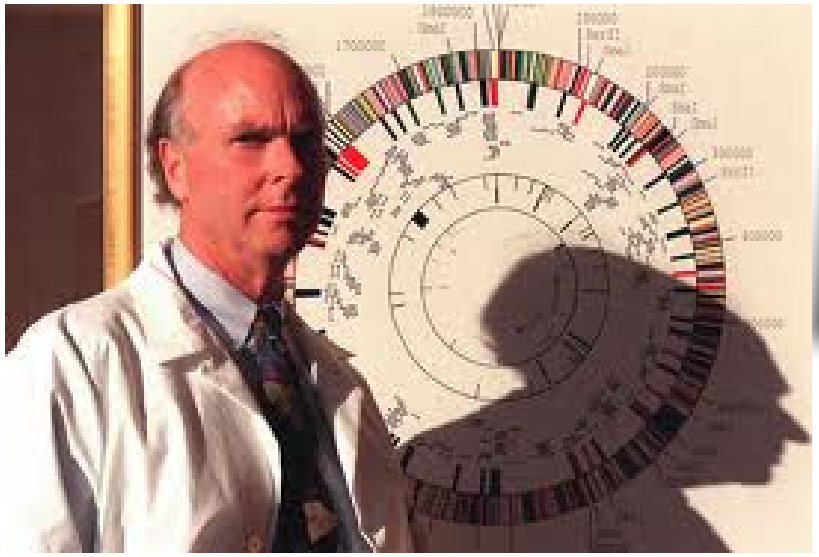


U.S. National Library of Medicine

COMPUTER BINARY CODE

01000001 01010100 01000111
01000001 01000001 01010100
01010100 01000011 01000111
01000001 01000011 01010100
01010100 01000001 01000011
01000011 01000111 01010100
01000001 01000111 01000111

First Genome Sequenced 1995



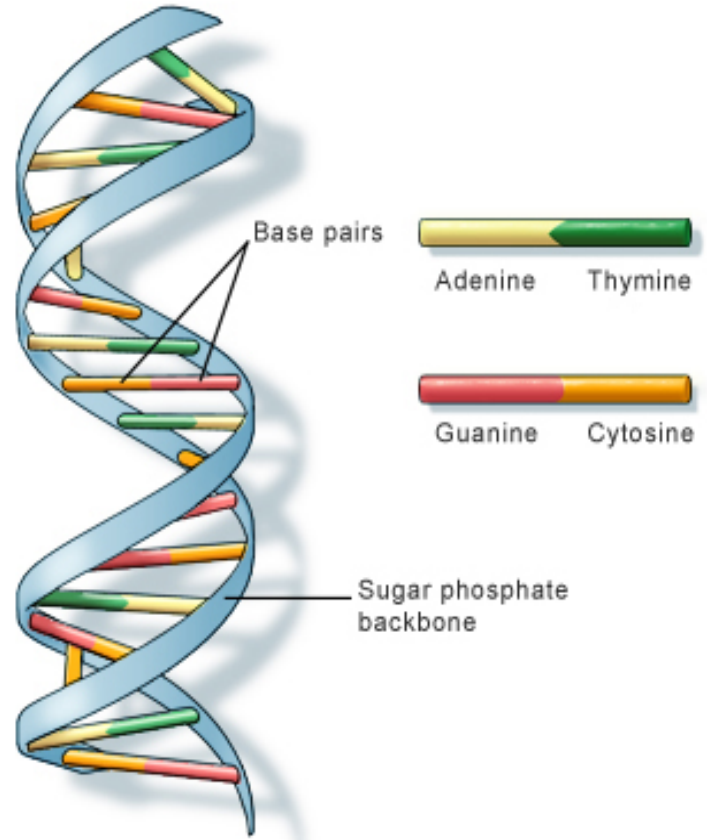
CONVERTING THE DIGITAL CODE BACK INTO THE ANALOG GENETIC CODE

GENETIC CODE

ATGAATTCGACTTACCGTAGG

COMPUTER BINARY CODE

01000001 01010100 01000111
01000001 01000001 01010100
01010100 01000011 01000111
01000001 01000011 01010100
01010100 01000001 01000011
01000011 01000111 01010100
01000001 01000111 01000111



U.S. National Library of Medicine

Does DNA contain all the necessary information for cellular life?

What is the smallest number of genes needed to run the machinery of a cell?

Can we design and construct a minimal genome/cell?

Technical Questions

- Can the chemistry permit synthesizing entire cellular chromosomes?
- If we make a synthetic chromosome can we boot it up and create a new living cell?

Rapid gene synthesis from oligonucleotides

Generating a synthetic genome by whole genome assembly: ϕX174 bacteriophage from synthetic oligonucleotides

Hamilton O. Smith, Clyde A. Hutchison III[†], Cynthia Pfannkoch, and J. Craig Venter[‡]

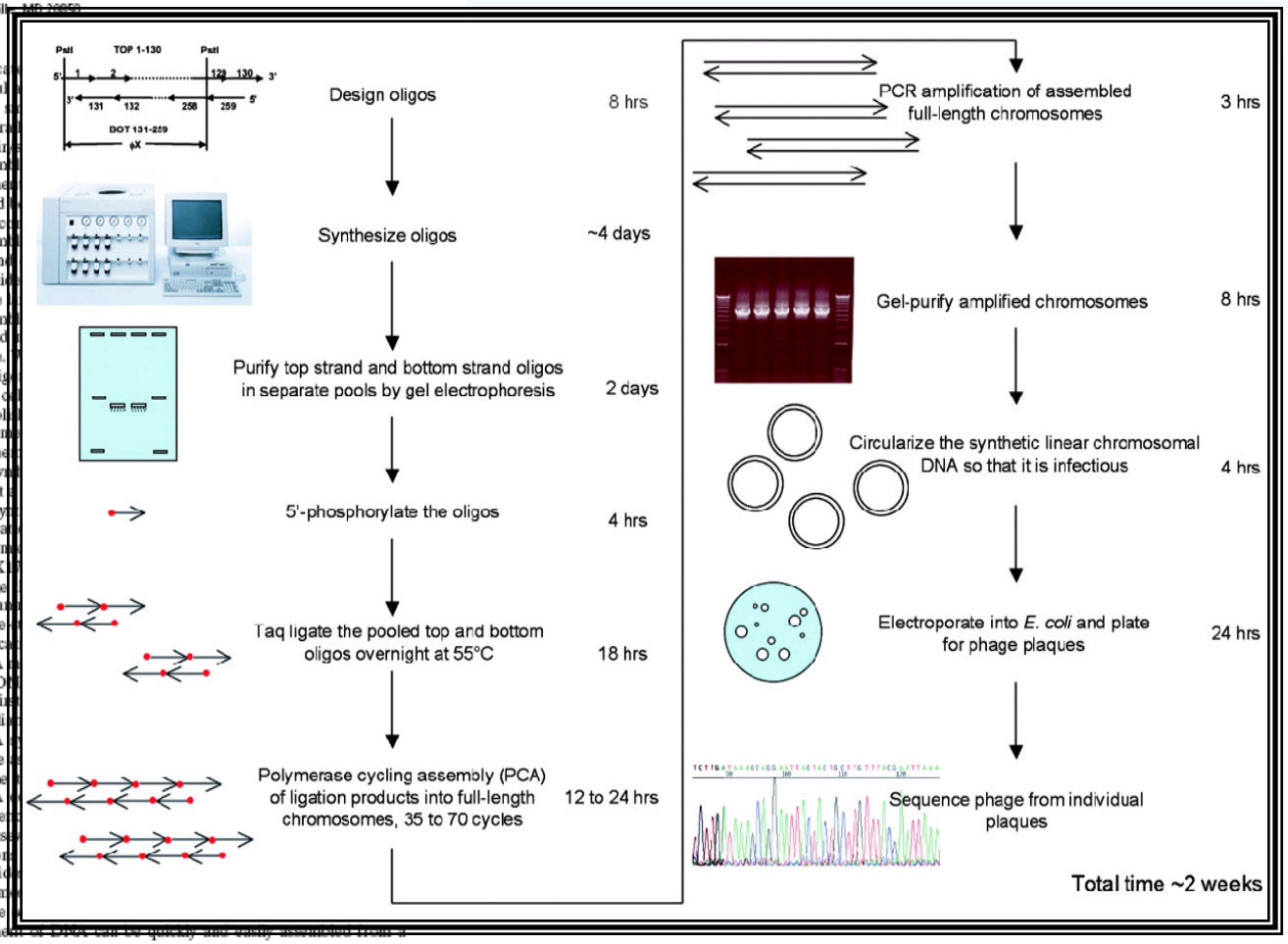
Institute for Biological Energy Alternatives, 1901 Research Boulevard, Suite 600, Rockville, MD 20850

Contributed by J. Craig Venter, November 3, 2003

We have improved upon the methodology and dramatically shortened the time required for accurate assembly of 5- to 6-kb segments of DNA from synthetic oligonucleotides. As a test of this methodology, we have established conditions for the rapid (14-day) assembly of the complete infectious genome of bacteriophage ϕX174 (5,386 bp) from a single pool of chemically synthesized oligonucleotides. The procedure involves three key steps: (*I*) gel purification of pooled oligonucleotides to reduce contamination with molecules of incorrect chain length, (*II*) ligation of the oligonucleotides under stringent annealing conditions (55°C) to select against annealing of molecules with incorrect sequences, and (*III*) assembly of ligation products into full-length genomes by polymerase cycling assembly, a nonexponential reaction in which each terminal oligonucleotide can be extended only once to produce a full-length molecule. We observed a discrete band of full-length assemblies upon gel analysis of the polymerase cycling assembly product, without any PCR amplification. PCR amplification was then used to obtain larger amounts of pure full-length genomes for circularization and infectivity measurements. The synthetic DNA had a lower infectivity than natural DNA, indicating approximately one lethal error per 500 bp. However, fully infectious ϕX174 virions were recovered after electroporation into *Escherichia coli*. Sequence analysis of several infectious isolates verified the accuracy of these synthetic genomes. One such isolate had exactly the intended sequence. We propose to assemble larger genomes by joining separately assembled 5- to 6-kb segments; ~60 such segments would be required for a minimal cellular genome.

Chemical synthesis of life in the laboratory has been a standing challenge to synthetic organic chemistry since Wöhler's synthesis of urea in 1828 (1), and the doctrine of spontaneous generation was put to rest by an address by Louis Pasteur in 1864.[§] With an understanding of the genetic role of DNA, much work has focused on the synthesis of oligonucleotides and genes. The synthesis of the 207-bp gene for tyrosine suppressor tRNA in 1979 by Khorana and 17 coworkers (2) was a monumental undertaking. Since then, the automated DNA synthesizer has been developed based on fundamental advances in synthetic methods from the laboratories of Letsinger (3, 4) and Caruthers (5, 6).

In 1999 we described a minimal prokaryotic genome based on results from random whole genome transposon mutagenesis that inactivated one gene per cell (7). By using this approach, ≈300 essential genes for self-replicating cellular life were described, and we proposed to make a synthetic chromosome to test the viability of this hypothesis (7). Before attempting synthesis of a microbial chromosome, we commissioned an independent bioethical review of our proposed scientific plan (8). After >1 year of deliberation, the reviewers concluded that we were taking a reasonable scientific approach to an important biological question. The broader implications of the creation of life in the laboratory can now be considered a realistic possibility. However, there are several technical barriers to the synthesis of microbial chromosome-sized stretches of DNA that are hundreds of thousands to millions of nucleotides long, the most notable being the contamination of the oligonucleotides by



Abbreviations: RF, replicative form of DNA; PCA, polymerase cycling assembly; synφX, synthetic ϕX174.

[†]Present address: Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599-7290.

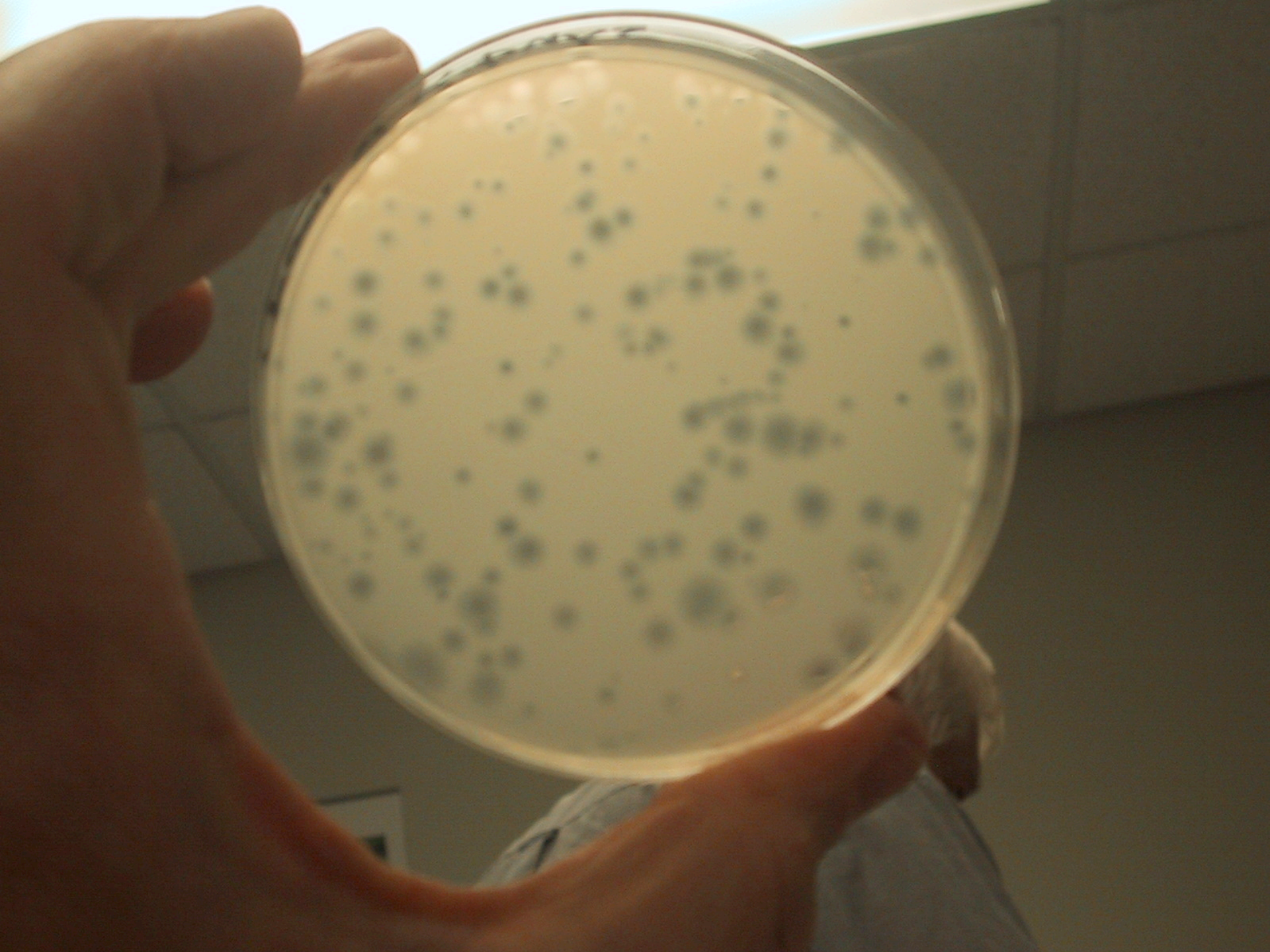
[‡]To whom correspondence should be addressed. E-mail: jcventer@tcag.org.

[§]Pasteur, L., Sorbonne Scientific Soirée, Apr. 7, 1864, Paris.

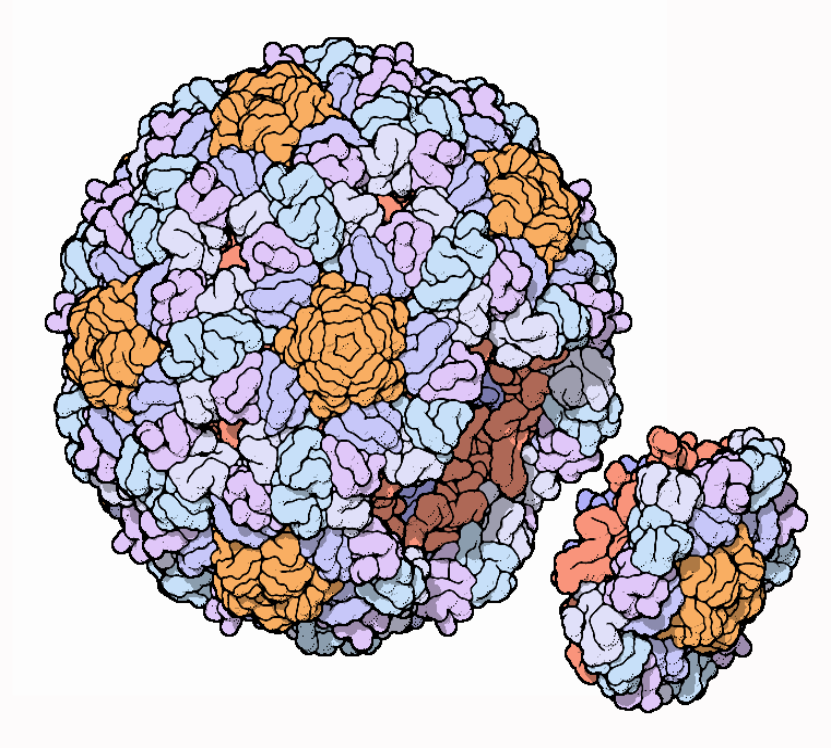
© 2003 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.2237126100

Hamilton Smith, Clyde Hutchison, Cindi Pfannkoch, Craig Venter

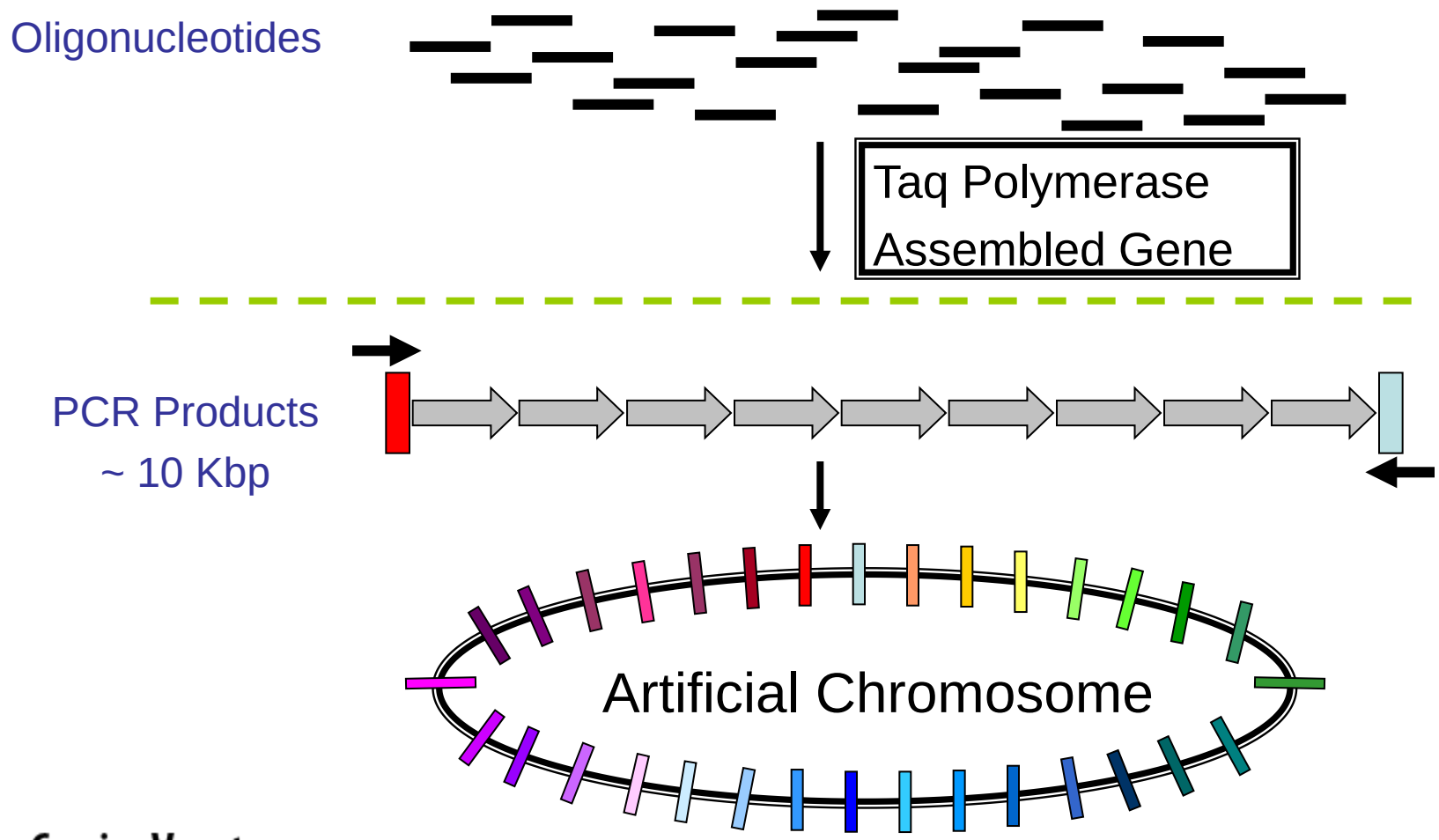


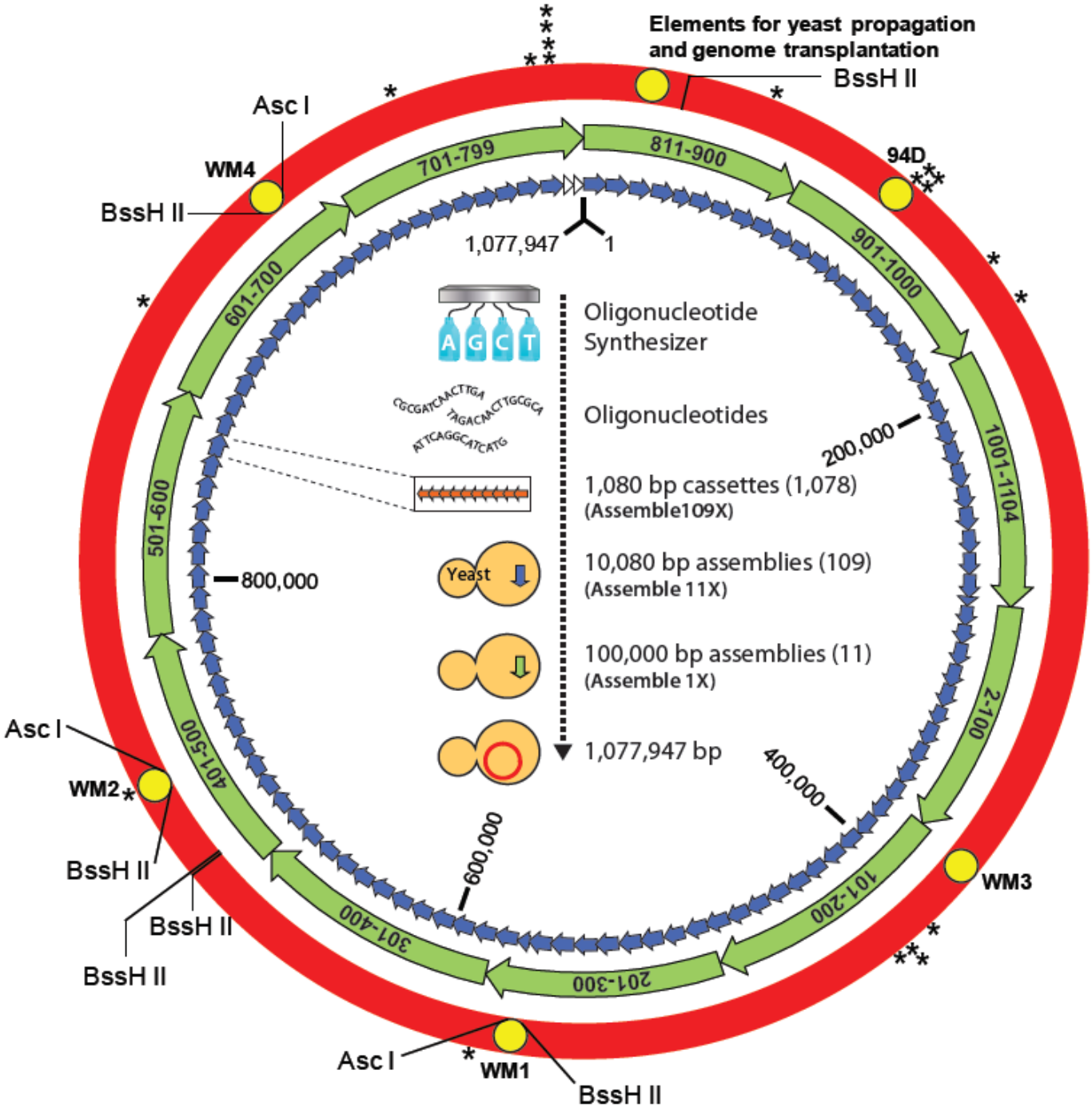
This Software Builds Its Own Hardware



J. Craig Venter
INSTITUTE

Construction of an Artificial Chromosome




Elements for yeast propagation
and genome transplantation
BssH II
Asc I
WM4
BssH II
94D
701-799
811-900
601-700
901-1000
1,077,947
1
Oligonucleotide
Synthesizer
A G C T
CGCGATCAACTTGA
TAGACAACTTGCGCA
ATTCAGGCATCATG
Oligonucleotides
200,000
1001-1104
1,080 bp cassettes (1,078)
(Assemble109X)
501-600
800,000
Yeast
10,080 bp assemblies (109)
(Assemble 11X)
100,000 bp assemblies (11)
(Assemble 1X)
2-100
Asc I
401-500
1,077,947 bp
400,000
WM2
600,000
BssH II
WM3
101-200
BssH II
301-400
BssH II
201-300
Asc I
WM1
BssH II

Science, August 2007

Genome Transplantation in Bacteria: Changing One Species to Another

Carole Lartigue, John I. Glass,* Nina Alperovich, Rembert Pieper, Prashanth P. Parmar, Clyde A. Hutchison III, Hamilton O. Smith, J. Craig Venter

As a step toward propagation of synthetic genomes, we completely replaced the genome of a bacterial cell with one from another species by transplanting a whole genome as naked DNA. Intact genomic DNA from *Mycoplasma mycoides* large colony (LC), virtually free of protein, was transplanted into *Mycoplasma capricolum* cells by polyethylene glycol–mediated transformation. Cells selected for tetracycline resistance, carried by the *M. mycoides* LC chromosome, contain the complete donor genome and are free of detectable recipient genomic sequences. These cells that result from genome transplantation are phenotypically identical to the *M. mycoides* LC donor strain as judged by several criteria.

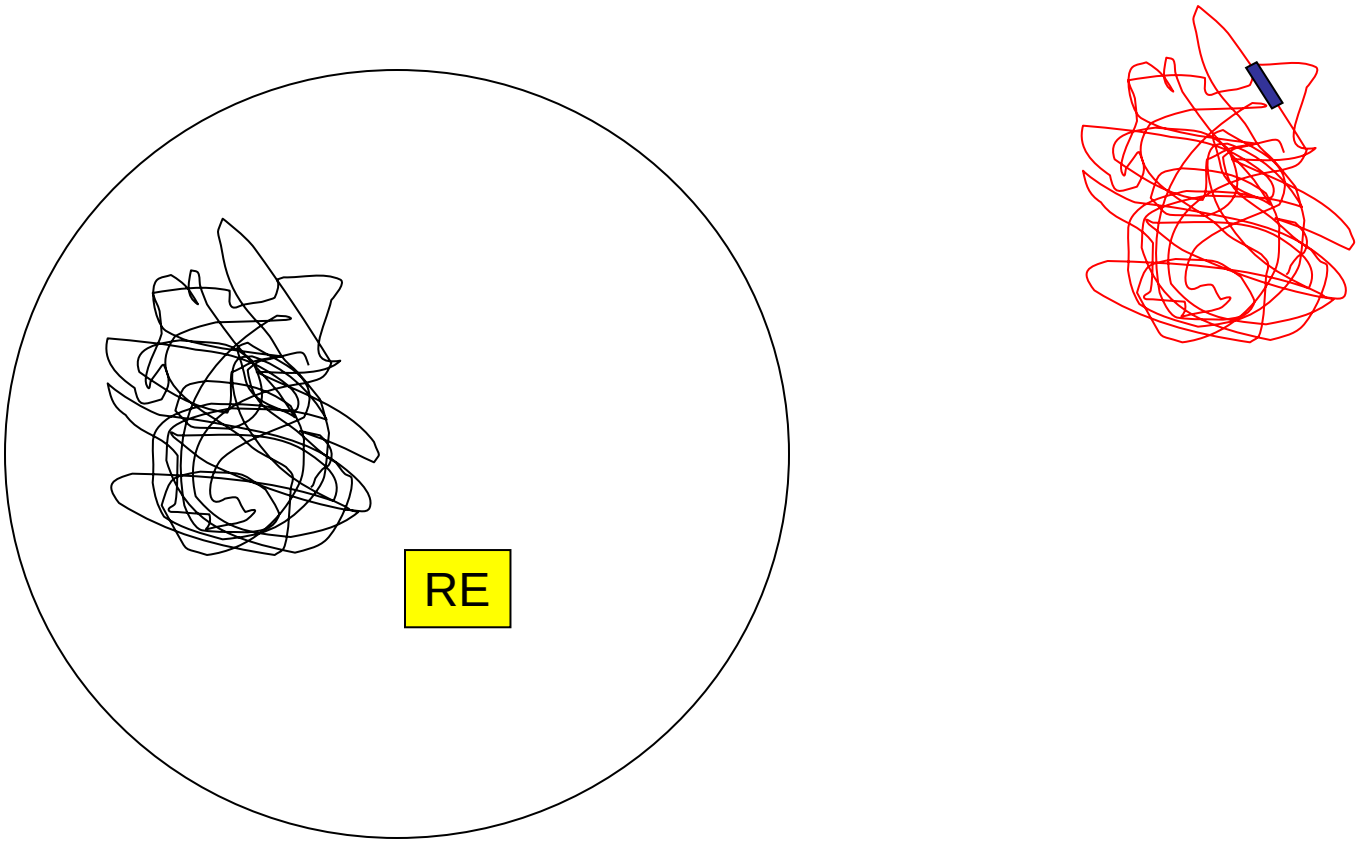
from incompatibility between the two genomes (*6*). Transplantation of nuclei as intact organelles into enucleated eggs is a well-established procedure in vertebrates (*7–9*). Our choice of the term "genome transplantation" comes from the similarity to eukaryotic nuclear transplantation in which one genome is cleanly replaced by another.

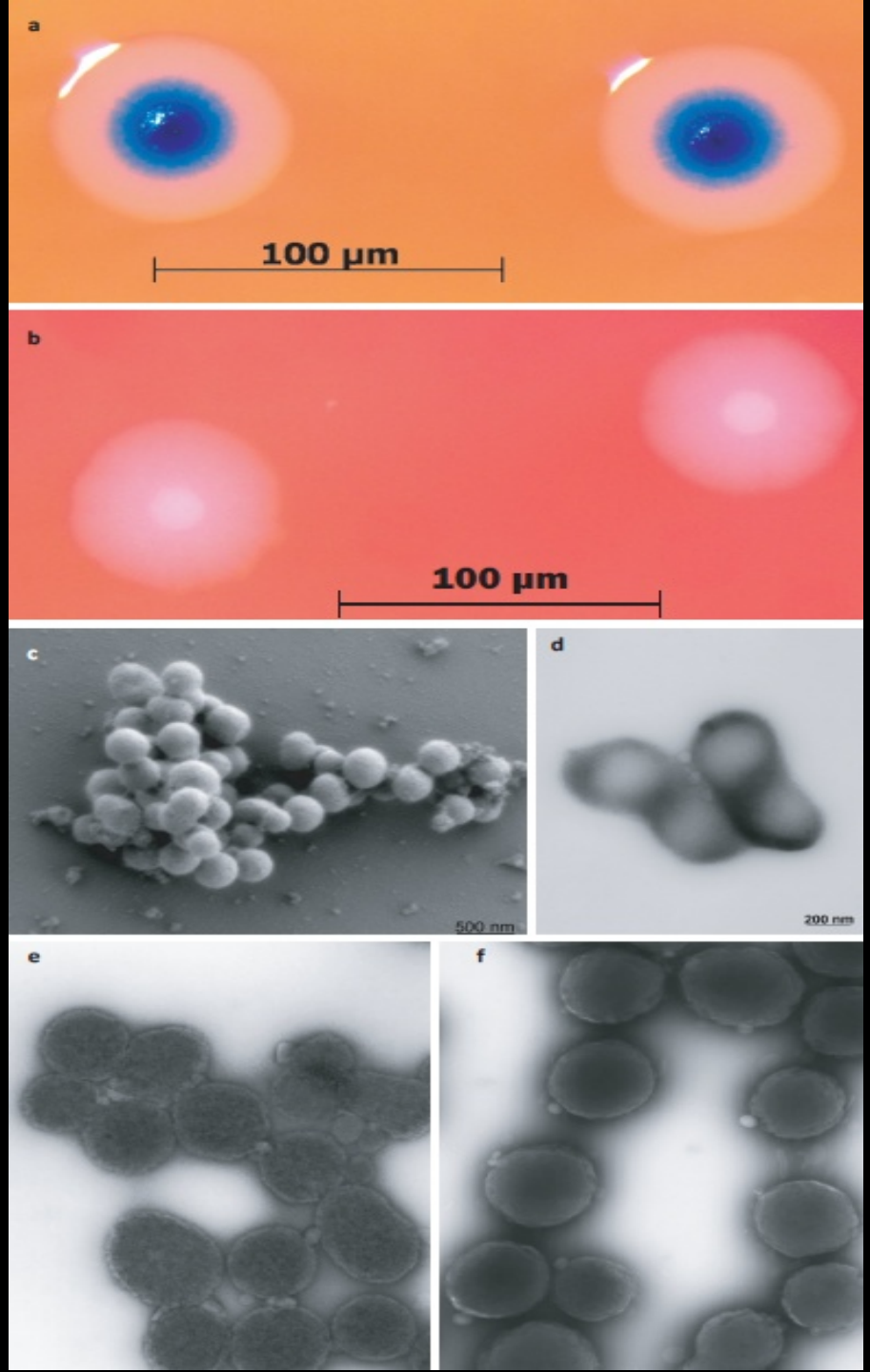
Genome transplantation is a requirement for the establishment of the new field of synthetic genomics. It may facilitate construction of useful microorganisms with the potential to solve pressing societal problems in energy production, environmental stewardship, and medicine. Chemically synthesized chromosomes must eventually be transplanted into a cellular milieu where the encoded instructions can be expressed. We have long been interested in defining a minimal genome that is just sufficient for cellular

J. Craig Venter

INSTITUTE

The donor genome encoded restriction enzyme degrades the recipient cell genome




a
100 µm
b
100 µm
c
500 nm
d
200 nm
e
f

Synthetic Life

Watermarks

-1, 321 unencoded characters, 1246 base pairs

J. CRAIG VENTER INSTITUTE 2009

ABCDEFGHIJKLMNOPQRSTUVWXYZ

0123456789#@()-+\=/:<;>$&}{*]"[%!'.,

SYNTHETIC GENOMICS, INC.

<!DOCTYPE HTML><HTML><HEAD><TITLE>GENOME TEAM</TITLE></HEAD><BODY><A HREF="HTTP://WWW.JCVI.ORG/">THE JCVI</A><P>PROVE YOU'VE DECODED THIS WATERMARK BY EMAILING US <A HREF="MAILTO:MROQSTIZ@JCVI.ORG">HERE!</A></P></BODY></HTML>

TTAACTAGCTAAGTTCGAATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAATACTGTGCGCTTGACTGTGATCCTGATAAATAACTTCTTCTGTAGGGTAGAGTTTTATTTAAGGCTACTCACTGGTTGCAAACCAATGCCGTACATTACTAGCTTGATCCTTGGTCGGTCATTGGGGGATATCTCTTACTAATAGAGCGGCCTATCGCGTATTCTCGCCGGACCCCCCTCTCCCACACCAGCGGTGTAGCATCACCAAGAAAATGAGGGGAACGGATGAGGAACGAGTGGGGGCTCATTGCTGATCATAATGACTGTTTATATACTAATGCCGTCAACTGTTTGCTGTGATACTGTGCTTTCGAGGGCGGGAGATTCGTTTTTGACATACATAAATATCATGACAAAACAGCCGGTCATGACAAAACAGCCGGTCATAATAGATTAGCCGGTGACTGTGAAACTAAAGCTACTAATGCCGTCAATAAATATGATAATAGCAACGGCACTGACTGTGAAACTAAAGCCGGCACTCATAATAGATTAGCCGGAGTCGTATTCATAGCCGGTAGATATCACTATAAGGCCCAGGATCATGATGAACACAGCACCACGTCGTCGTCCGAGTTTTTTTGCTGCGACGTCTATACCACGGAAGCTGATCATAAATAGTTTTTTTGCTGCGGCACTAGAGCCGGACAAGCACACTACGTTTGTAAATACATCGTTCCGAATTGTAAATAATTTAATTTCGTATTTAAATTATATGATCACTGGCTATAGTCTAGTGATAACTACAATAGCTAGCAATAAGTCATATATAACAATAGCTGAACCTGTGCTACATATCCGCTATACGGTAGATATCACTATAAGGCCCAGGACAATAGCTGAACTGACGTCAGCAACTACGTTTAGCTTGACTGTGGTCGGTTTTTTTGCTGCGACGTCTATACGGAAGCTCATAACTATAAGAGCGGCACTAGAGCCGGCACACAAGCCGGCACAGTCGTATTCATAGCCGGCACTCATGACAAAACAGCGGCGCGCCTTAACTAGCTAA

Watermark-2 326 unencoded characters, 1081 base pairs

MIKKEL ALGIRE, MICHAEL MONTAGUE, SANJAY VASHEE, CAROLE LARTIGUE, CHUCK MERRYMAN, NINA ALPEROVICH, NACYRA ASSAD-GARCIA, GWYN BENDERS, RAY-YUAN CHUANG, EVGENIA DENISOVA, DANIEL GIBSON, JOHN GLASS, ZHI-QING QI.

"TO LIVE, TO ERR, TO FALL, TO TRIUMPH, TO RECREATE LIFE OUT OF LIFE." - JAMES JOYCE

TTAACTAGCTAACAACTGGCAGCATAAAACATATAGAACTACCTGCTATAAGTGATACAACTGTTTTCATAGTAAAACATACAACGTTGCTGATAGTACTCCTAAGTGATAGCTTAGTGCGTTTAGCATATATTGTAGGCTTCATAATAAGTGATATTTTAGCTACGTAACTAAATAAACTAGCTATGACTGTACTCCTAAGTGATATTTTCATCCTTTGCAATACAATAACTACTACATCAATAGTGCGTGATATGCCTGTGCTAGATATAGAACACATAACTACGTTTGCTGTTTTCAGTGATATGCTAGTTTCATCTATAGATATAGGCTGCTTAGATTCCCTACTAGCTATTTCTGTAGGTGATATACGTCCATTGCATAAGTTAATGCATTTAACTAGCTGTGATACTATAGCATCCCCATTCCTAGTGCATATTTTCATCCTAGTGCTACGTGATATAATTGTACTAATGCCTGTAGATAATTTAATGCCTGGCTCGTTTGTAGGTGATAATTTAGTGCCTGTAAAACATATACCTGAGTGCTCGTTGCGTGATAGTTCGTTCATGCATATACAACTAGGCTGCTGTGATATGGTCACTGCCCTTACTGTGCTACATATTACTGCGAGGGGGATGACGTATAAACCTGTTGTAAGTGATATGACGTATATAACTACTAGTGATATGACGTATAGGCTAGAACAACGTGATATGACGTATATGACTACTGTCCCAAACATCAGTGATATGACGTATACTATAATTTCTATAATAGTGATAAATAAACCTGGGCTAAATACGTTCCTGAATACGTGGCATAAACCTGGGCTAACGAGGAATACCCATAGTTTAGCAATAAGCTATAGTTCGTCATTTTTAAGGCGCGCCTTAACTAGCTAA

Watermark-3 335 unencoded characters, 1109 base pairs

CLYDE HUTCHISON, ADRIANA JIGA, RADHA KRISHNAKUMAR, JAN MOY, MONZIA MOODIE, MARVIN FRAZIER, HOLLY BADEN-TILSON, JASON MITCHELL, DANA BUSAM, JUSTIN JOHNSON, LAKSHMI DEVI VISWANATHAN, JESSICA HOSTETLER, ROBERT FRIEDMAN, VLADIMIR NOSKOV, JAYSHREE ZAVERI.

"SEE THINGS NOT AS THEY ARE, BUT AS THEY MIGHT BE."

TTAACTAGCTAATTTAACCATATTTAAATATCATCCTGATTTTCACTGGCTCGTTGCGTGATATAGATTCTACTGTAGTGCTAGATAGTTCTGTACTAGGTGATACTATAGATTTCATAGATAGCACTACTGGCTTCATGCTAGGCATCCCAATAGCTAGTGATAGTTTAGTGCATACAACGTCATGTGATACAACGTTGCTGGCTGTAGATACAACGTCGTATTCTGTAAGTGATACAATAGCTATTGCTGTGCATAGGCCTATAGTGGCTGTAACTAGTGATATCACGTAACAACCATATAAGTTAGATTTAATGCCCCTGACTGAACGCTCGTTGCGTGATAGTTTAGGCTCGTTGCATACAACTGTGATTTTCATAAAACAACGTGATAATTTAGTGCTAGATAAGTTCCGCTTAGCAAGTGATAGTTTCCGCTTGACTGTGCATAGTTCGTTCATGCGCTCGTTGCGTGATAACTAGGCAGCTTCACAACTGATAATTTAATTGCTGATATTGCTGGCTGTCTAGTGCTAGTGATCATAGTGCGTGATAGTTTAAGCTGCTCTGTTTTAGATATCACGTGCTTGATAATGAAACTAACTAGTGATACTACGTAGTTAACTATGAATAGGCCTACTGTAAATTCAATAGTGCGTGATATTGAACTAGATTCTGCAACTGCTAATATGCCGTGCTGCACGTTTGGTGATAGTTTAGCATGCTTCACTATAATAAATATGGTAGTTGTAACTACTGCGAATAGGGGGAGCTTAATAAATATGATCACTGTGCTACGCTATATGCCGTTGAATATAGGCTATATGATCATAACATATATAGCTATAAGTGATAAGTTCCTGAATATAGGCTATATGATCATAACATATACAACTGTACTCATGAATAAGTTAACGAGGATTAACTAGCTAA

Watermark-4 338 unencoded characters, 1222 base pairs

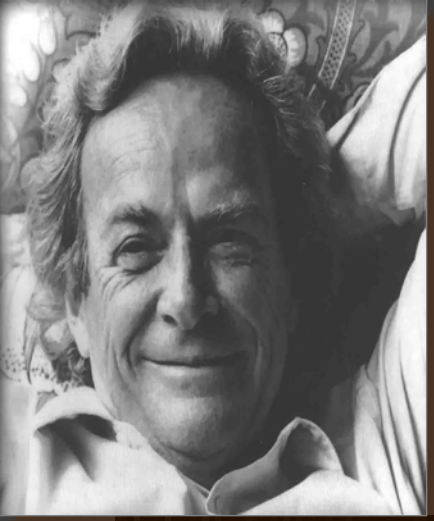
CYNTHIA ANDREWS-PFANNKOCH, QUANG PHAN, LI MA, HAMILTON SMITH, ADI RAMON, CHRISTIAN TAGWERKER, J CRAIG VENTER, EULA WILTURNER, LEI YOUNG, SHIBU YOOSEPH, PRABHA IYER, TIM STOCKWELL, DIANA RADUNE, BRIDGET SZCZYPINSKI, SCOTT DURKIN, NADIA FEDOROVA, JAVIER QUINONES, HANNA TEKLEAB.

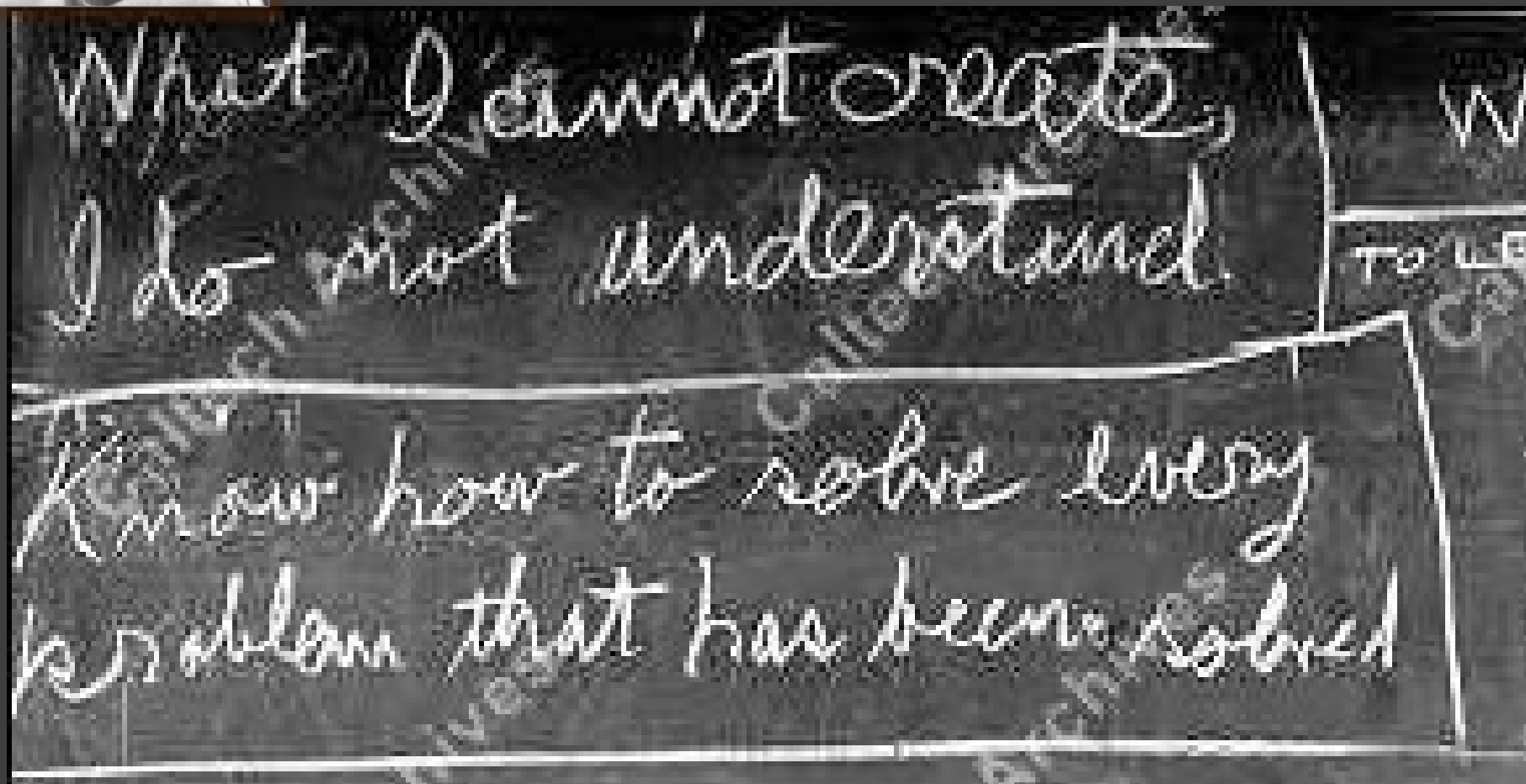
"WHAT I CANNOT BUILD, I CANNOT UNDERSTAND." - RICHARD FEYNMAN

TTAACTAGCTAATTTCATTGCTGATCACTGTAGATATAGTGCATTCTATAAGTCGCTCCCACAGGCTAGTGCTGCGCACGTTTTTCAGTGATATTATCCTAGTGCTACATAACATCATAGTGCGTGATAAACCTGATACAATAGGTGATATCATAGCAACTGAACTGACGTTGCATAGCTCAACTGTGATCAGTGATAGATTCTGATACTATAGCAACGTTGCGTGATATTTTCACTACTGGCTTGACTGTAGTGCATATGATAGTACGTCTAACTAGCATAACTAGTGATAGTTATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAGTGATATAATCCAACTAGATAGTCCTGAACTGATCCCTATGCTAACTAGTGATAAACTAACTGATACATCGTTCCTGCTACGTGATAGCTTCACTGAGTTCCATACATCGTCGTGCTTAAACATCAGTGATAACACTATAGAGTTCATAGATACTGCATTAACTAGTGATATGACTGCAAATAGCTTGACGTTTTGCAGTCTAAAACAACGTGATAATTCTGTAGTGCTAGATACTATAGATTTCCTGCTAAGTGATAAGTCTACTGATTTACTAATGAATAGCTTGGTTTTGGCATACACTGTGCGCTGCACTGGTGATAGCTTTTCGTTGATGAATAATTTCCCTAGCACTGTGCGTGATATGCTAGATTCTGTAGATAGGCTAAATTCGTCTACGTTTGTAGGTGATAGTTTAGTTGCTGTAACTAATATTATCCCTGTGCCGTTGCTAAGCTGTGATATCATAGTGCTGCTAGATATGATAAGCAAACTAATAGAGTCGAGGGGGAGTCTCATAGTGAATACTGATATTTTAGTGCTGCCGTTGAATAAGTTCCCTGAACATTGTGATACTGATATTTTAGTGCTGCCGTTGAATATCCTGCATTTAACTAGCTTGATAGTGCATTCGAGGAATACCCATACTACTGTTTTCATAGCTAATTATAGGCTAACATTGCCAATAGTGCGGCGCGCCTTAACTAGCTAA

J. Craig Venter

INSTITUTE

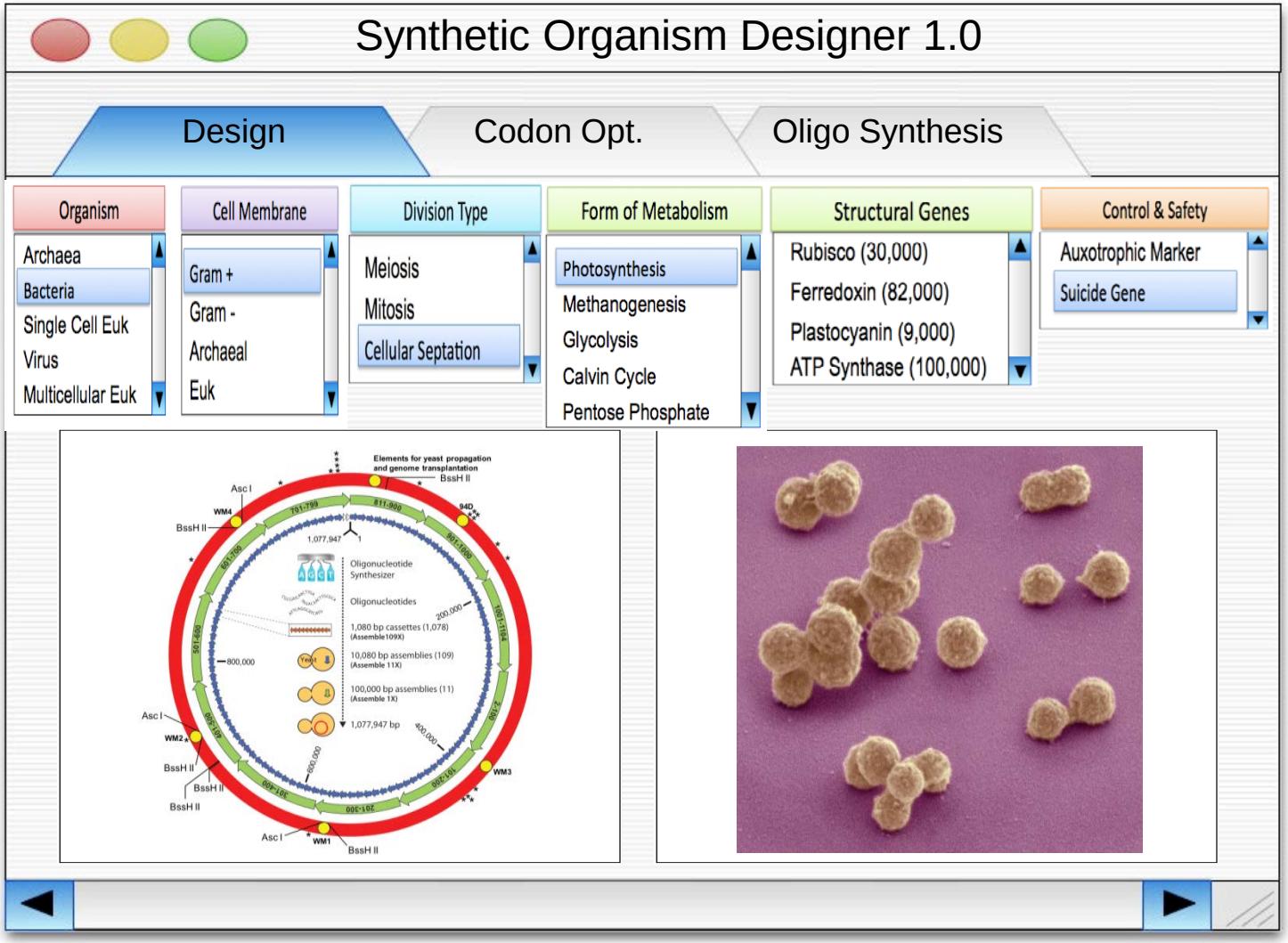



What I cannot create,
I do not understand.
Know how to solve every
problem that has been solved

Life is a DNA software system

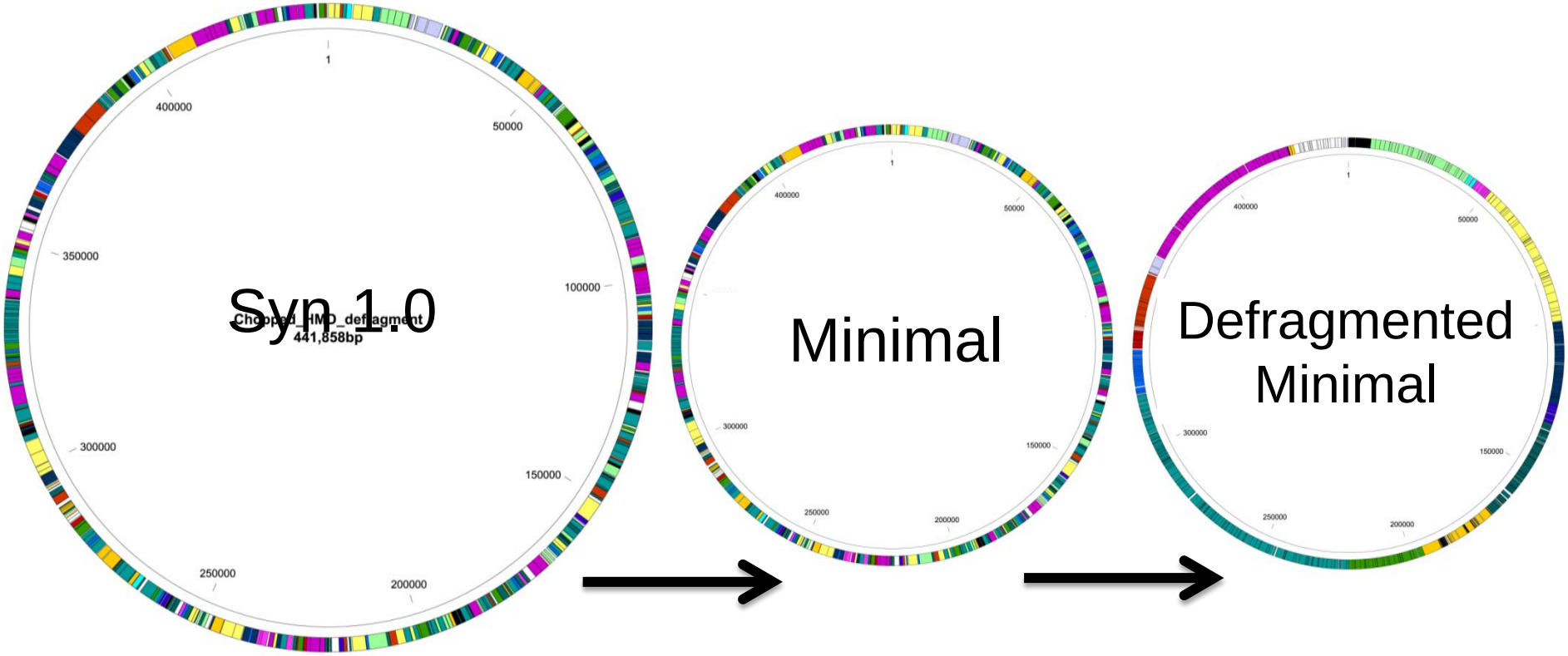
Change the DNA software change the species

Learning to build cells from scratch



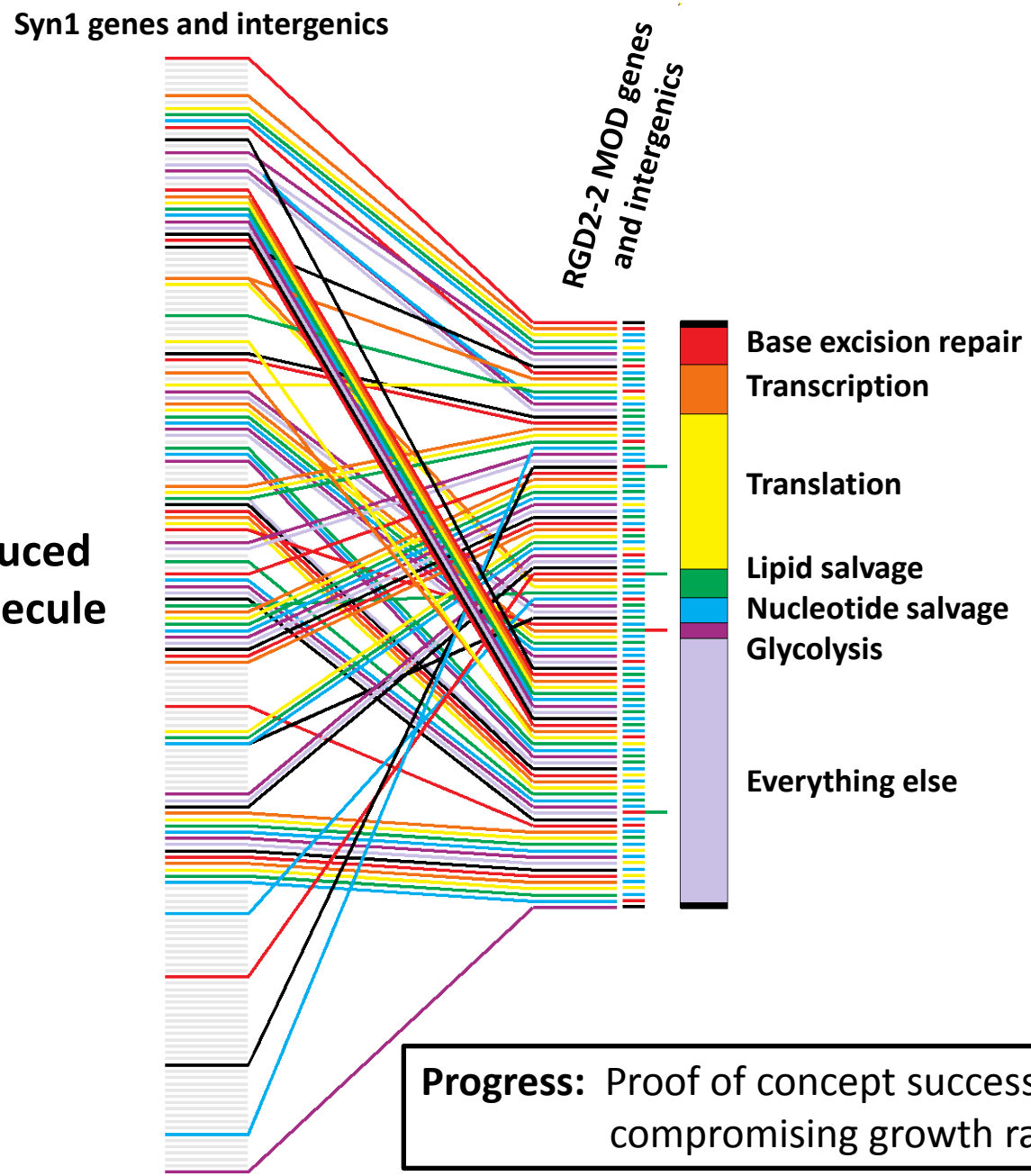
Genome minimization and modularization

Progress on Genome Design



- Synthesized eight functional 1/8th molecules
- Combined 6 ½ of these to produce a viable cell
  - Identified 25 genes that we need to add back
- 1.1 Mbp → 0.65 Mbp reduction (59%)
- Initiated defragmentation experiments
- Investigating genome design principles

Progress on Genome Design (modularization)



Non-reduced 1/8th molecule

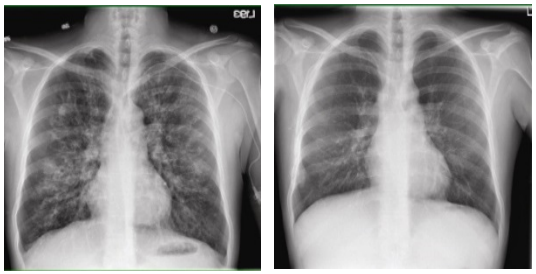
Reduced and organized 1/8th molecule

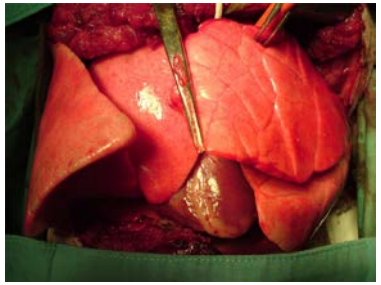
Progress: Proof of concept successfully completed without compromising growth rate

Synthetic Genomics to rewrite the Pig genome

Lung Transplantation Overview

- US Market size: $25B
  - 250,000 deaths* from end-stage lung disease
  - Current supply of allograft lungs satisfy only
  - <1% of demand (2000/yr)
  - 2014 cost: $100K/lung
- Xenolung production: Transformative
  - Increase transplants 100-fold
- Xenotransplantion Opportunity
  - Other organs same demand/supply mismatch
  - Xenografts enable scheduled procedures
  - Scheduled procedures enable tolerance-induction
    - Tolerance induction can eliminate chronic immunosuppression
    - Proven by UT SAB Member Thomas Starzl
- SGI-UT Collaboration announced May 6, 2014

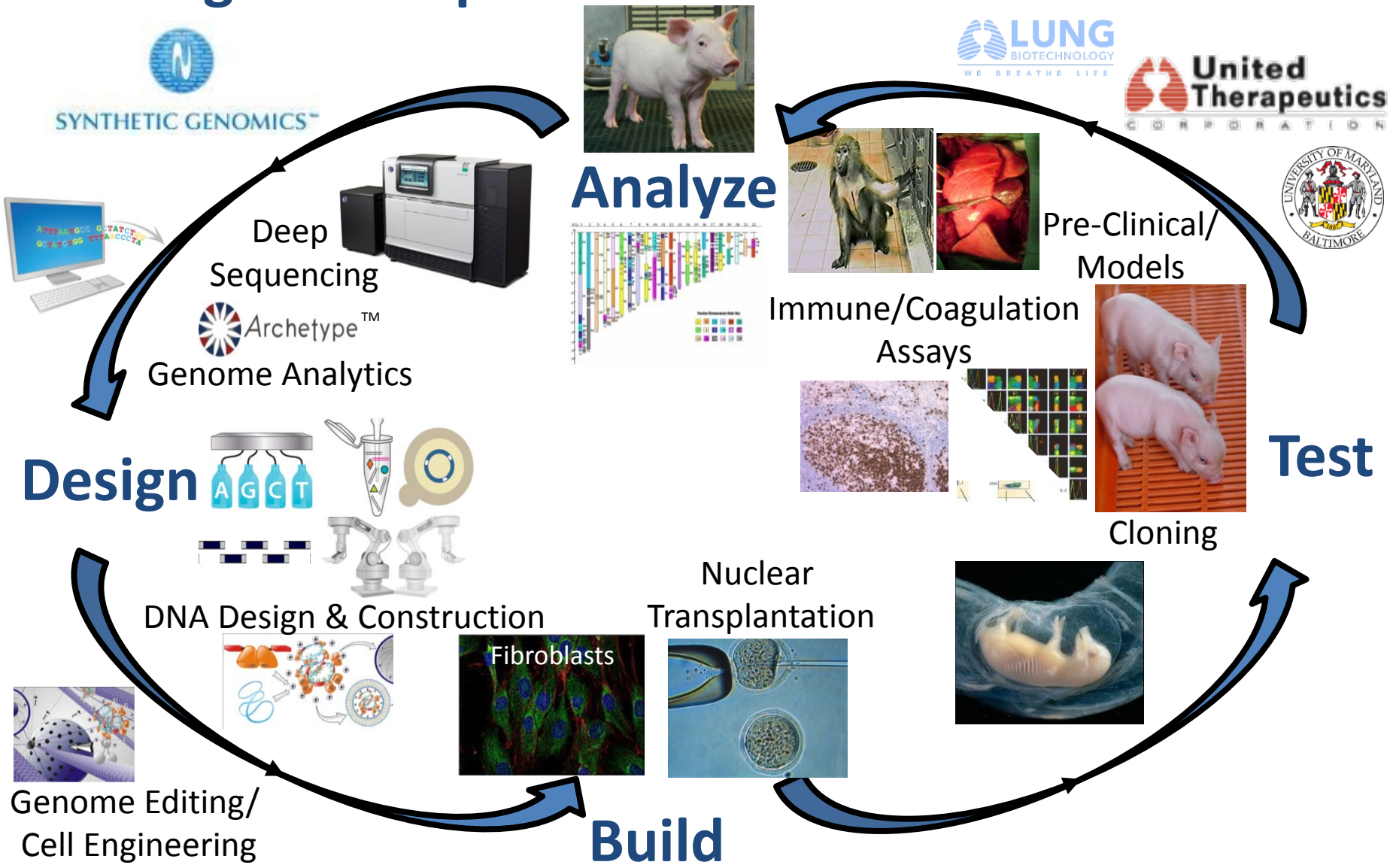


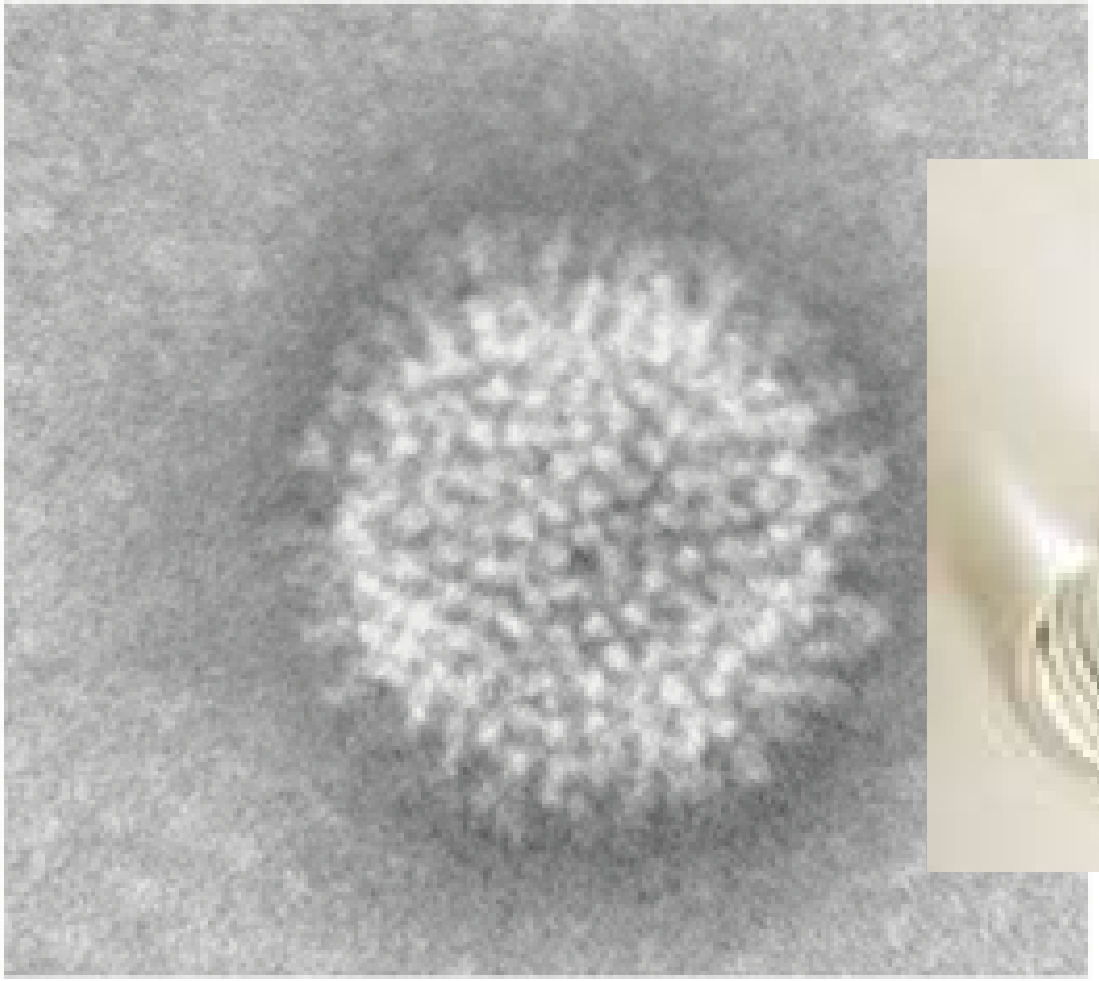


*does not include potential for lung cancer at 150K/year

United Therapeutics CORPORATION

SGI-UT R&D Collaboration to Humanize Pigs for Organ Transplants





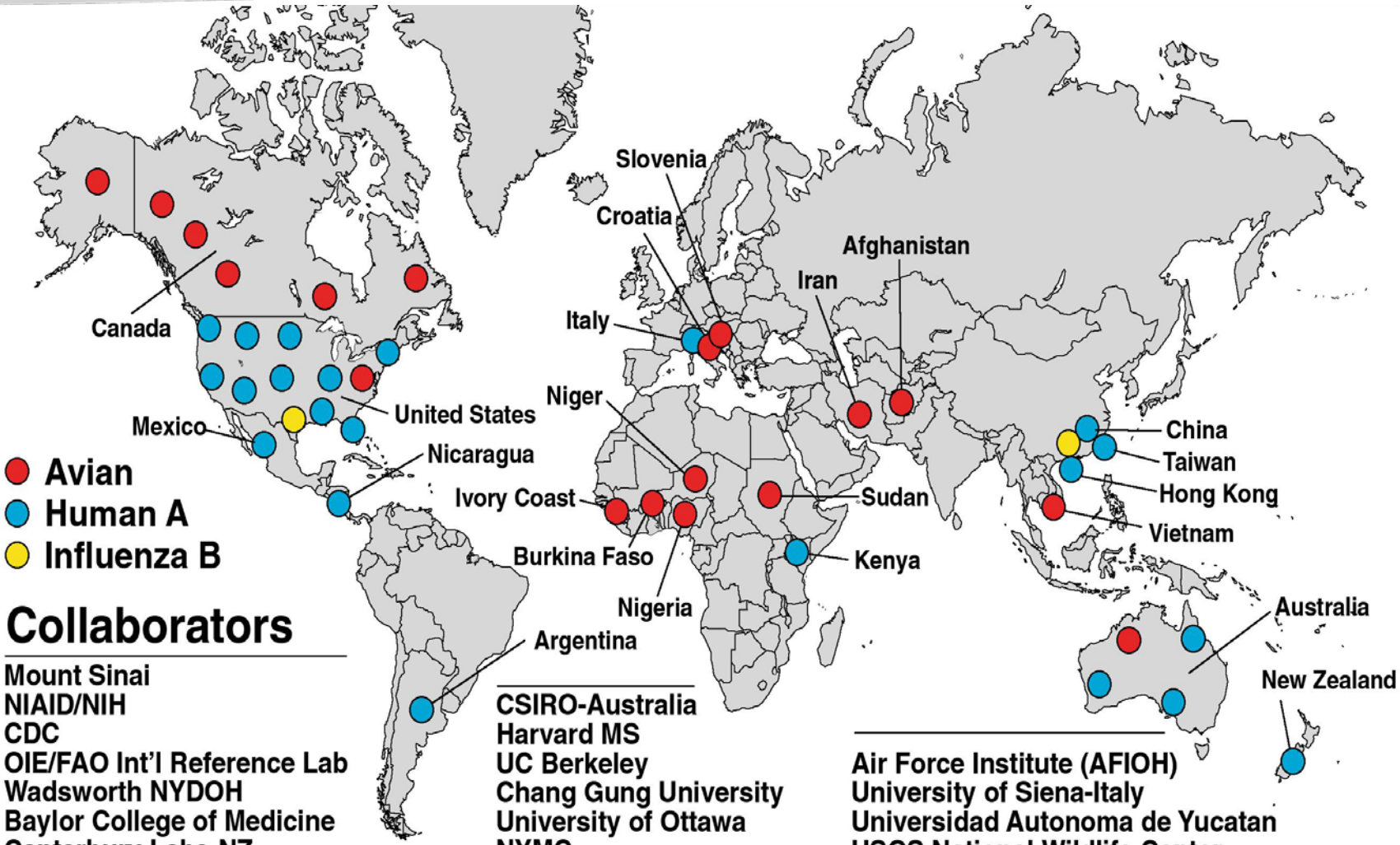
50 nm

Electron micrograph
of an influenza virus

Expert Reviews in Molecular Medicine
© 2001 Cambridge University Press

J. Craig Venter™
INSTITUTE

Influenza Genome Sequencing Project Collections



Collaborators

Mount Sinai
NIAID/NIH
CDC
OIE/FAO Int'l Reference Lab
Wadsworth NYDOH
Baylor College of Medicine
Canterbury Labs-NZ
Ohio State
JCVI

CSIRO-Australia
Harvard MS
UC Berkeley
Chang Gung University
University of Ottawa
NYMC
UPitt MC
University of Hong Kong

Air Force Institute (AFIOH)
University of Siena-Italy
Universidad Autonoma de Yucatan
USGS National Wildlife Center
Canadian Wild Bird Survey
National Influenza Lab., Argentina

Human Cases of Avian influenza A Virus H7N9

- First human cases reported March 31
- As of April 10 2013, 33 confirmed cases and 9 deaths
- Unlike H5N1, H7N9 not causing serious illness in birds – complicating detection and control
- Closely monitoring for evidence of human-to-human transmission

Two die in China from bird flu strain not previously seen in humans-Xinhua

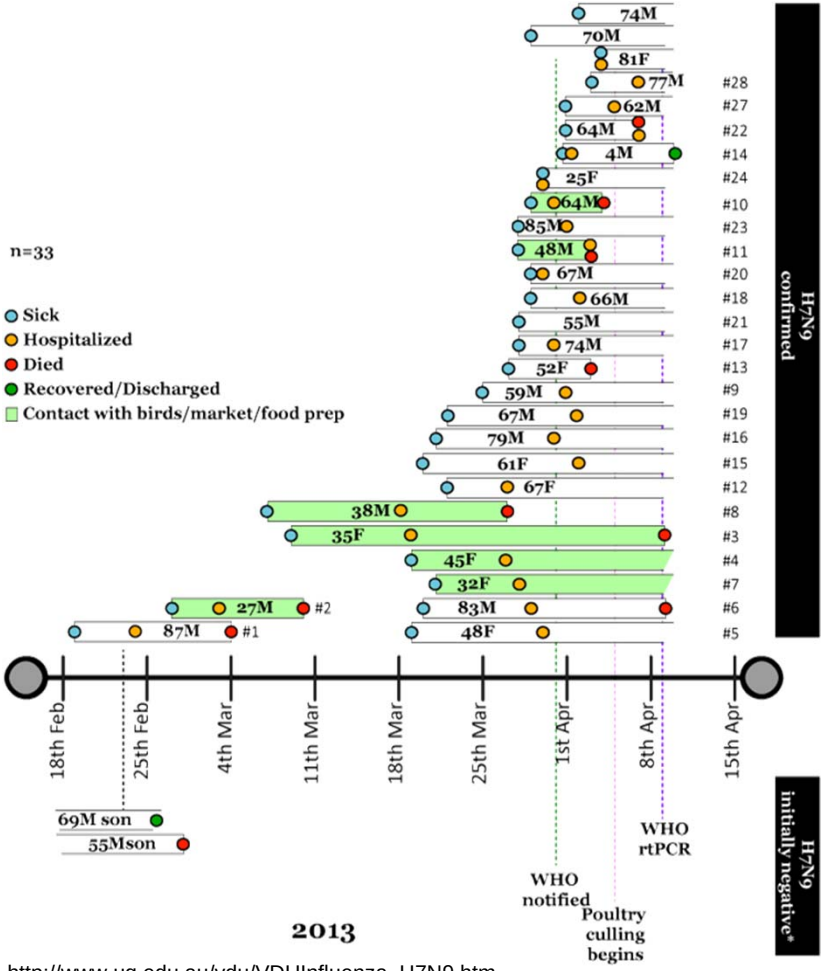
Sun, 31 Mar 2013 07:56 GMT
Source: Reuters // Reuters

SHANGHAI, March 31 (Reuters) - Two people in Shanghai, one of China's largest cities, died this month after contracting a strain of avian influenza that had never been passed to humans before, the official Xinhua News Agency reported on Sunday.

The two men, aged 87 and 27, became sick late February and died in early March. Another woman in nearby Anhui province also contracted the virus in early March and is in a critical condition, Xinhua said, quoting the National Health and Family Planning Commission (NHFPC).

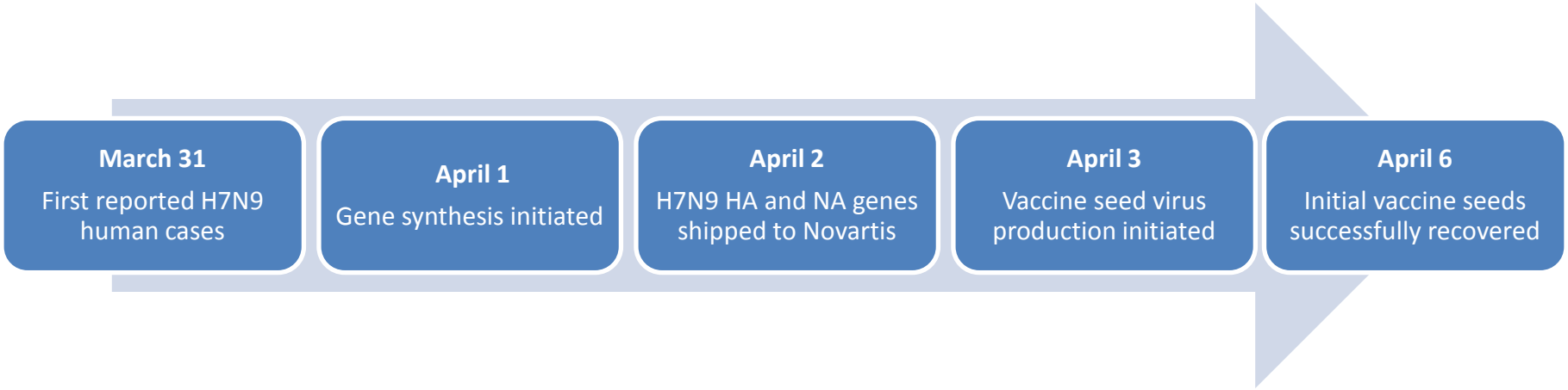
The strain of the bird flu virus found in all three people was identified as H7N9, which had not been transmitted to humans before, the commission said.

http://www.nature.com/news/urgent-search-for-flu-source-1.12762
SOURCES: WHO/ECDC/XINHUA STATE MEDIA



http://www.uq.edu.au/vdu/VDUInfluenza_H7N9.htm

Demonstrating Pandemic Preparedness in Real Time



- Real time demonstration of pandemic preparedness with synthetic gene assembly process
- Initial vaccine seeds generated within 6 days of first reported cases
- Genes provided to CDC on April 2 (2 days after first human cases reported)
- Vaccine seeds undergoing characterization, testing and preparation for vaccine production

➥ Synthetic process provides significant advantages for making diagnostic and testing reagents, speed of vaccine production and overall vaccine yield

Novartis MDCK Cell Vaccine Facility

Can produce 300 million Flu vaccine doses

Novartis announces positive clinical trial results for novel H7N9 vaccine

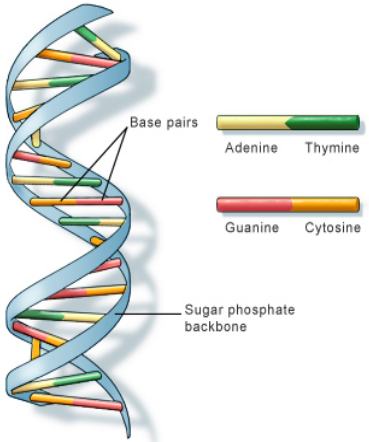
November 14, 2013

- ***85% of subjects immunologically protected after receiving second dose of investigational cell culture vaccine when combined with proven MF59® adjuvant***
- ***Vaccine now in large scale production highlighting rapid response capability of novel FDA licensed cell culture technology***
- ***135 confirmed cases and 45 deaths from H7N9 virus since emergence in March according to the World Health Organization[1]***

SYNTHETIC GENOMICS™

Biological Teleportation

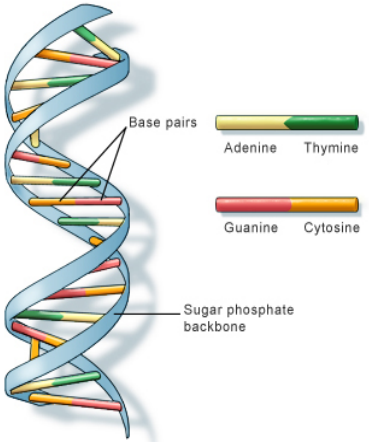
INTER-CONVERTING THE GENETIC CODE and the DIGITAL CODE



U.S. National Library of Medicine

COMPUTER BINARY CODE

01000001 01010100 01000111 01000001
01000001 01010100 01010100 01000011
01000111 01000001 01000011 01010100
01010100 01000001 01000011 01000011
01000111 01010100 01000001 01000111
01000111



U.S. National Library of Medicine

J. Craig Venter
INSTITUTE

Developing a Fax Machine to Copy Life on Mars

Sample Collection Sites

Environmental Samples

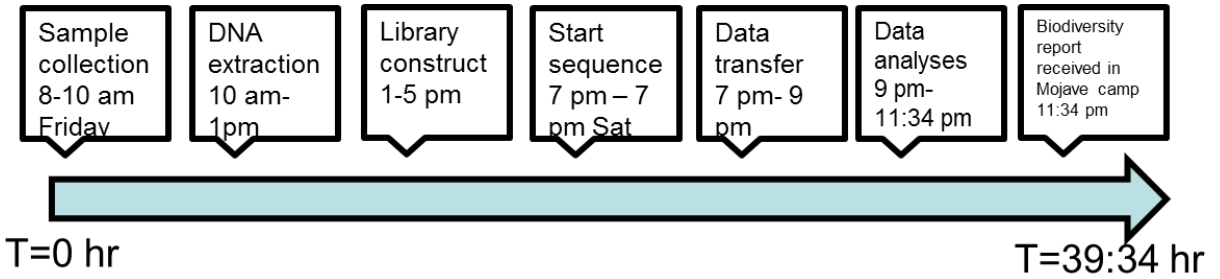
Soil

Quartz

Crust

Granite

Bio-teleportation of Microbial Diversity in the Mojave Desert



- **Timeline significantly short for biodiversity analyses**
- **Data consistent with other molecular-based studies**
- **Technology needs to be miniaturized**
- **Visible differences at sequence level between samples**

percent sequences in dataset

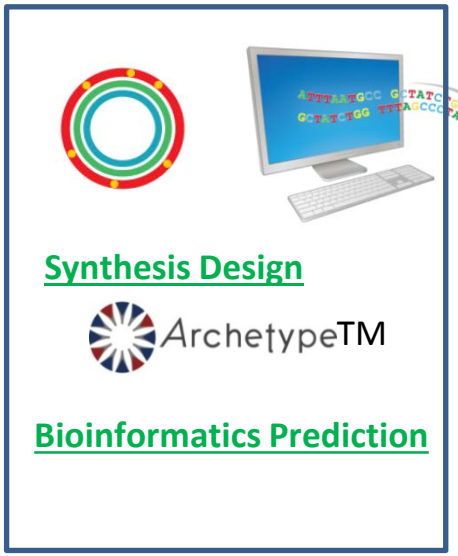
	fungi	cyanobac	actinobac	acidobac	proteobac	sum
soil	2	5	41	4	18	70
quartz	9	25	22	7	9	72
crust	32	20	15	0.8	13	80.8
granite	2	25	23	8	14	72

illumina®

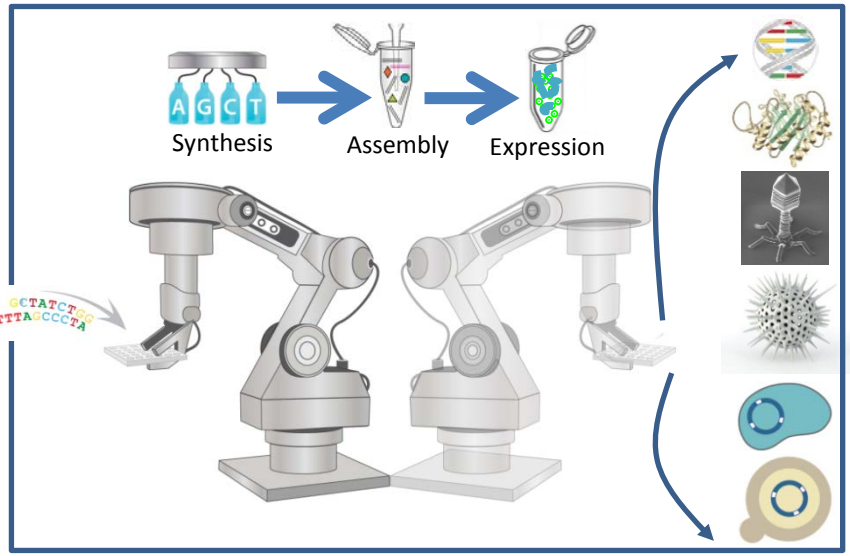
SYNTHETIC GENOMICS®

Digital-Biological Converter (Receiving Unit)

A device that can receive digital genetic code as electromagnetic waves and convert it into a biological entity such as DNA, proteins, viruses, and microbial cells



0-100,000,000 kilometers



Potential Applications:

Vaccines (including DNA/RNA)
Peptide, protein or antibody therapies
Bacteriophage therapy

Digital-Biological Converter

- Genome engineering/transplantation
- Emergency response
- Diagnostics
- Product and Lead Discovery

DBC: Progress to Date

Made possible by our advances in rapid, robust, and accurate DNA synthesis and assembly methods

Program Initiation → Jun

DBC Design Finalized → July

Installation → Aug

DNA Assembly → Sep

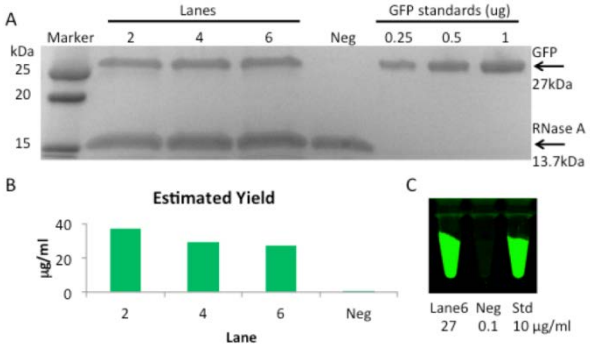
Protein Production → Oct

HA/NA Production and Viral Rescue → Nov, Dec

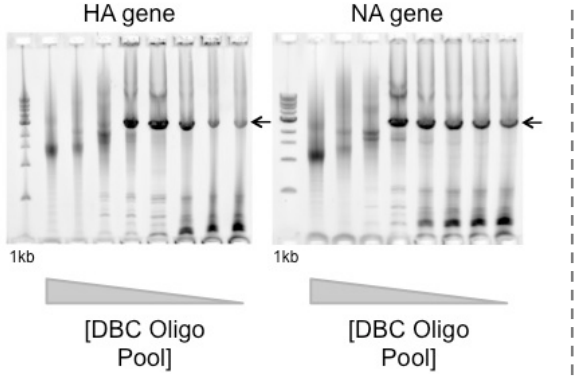
Phi X 174 Phage Production → Mar

Jun > July > Aug > Sep > Oct > Nov > Dec > Jan > Feb > Mar > Apr

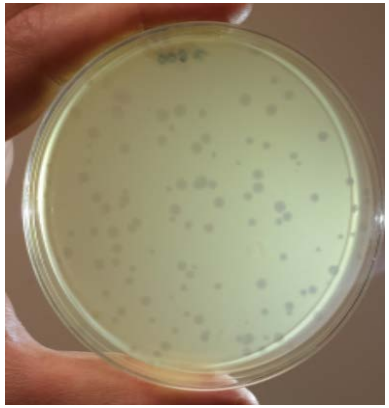
Automated Protein Expression: GFP



DNA Synthesis and Susequent Viral Resuce



Infectious Phi X 174 Phage Particles

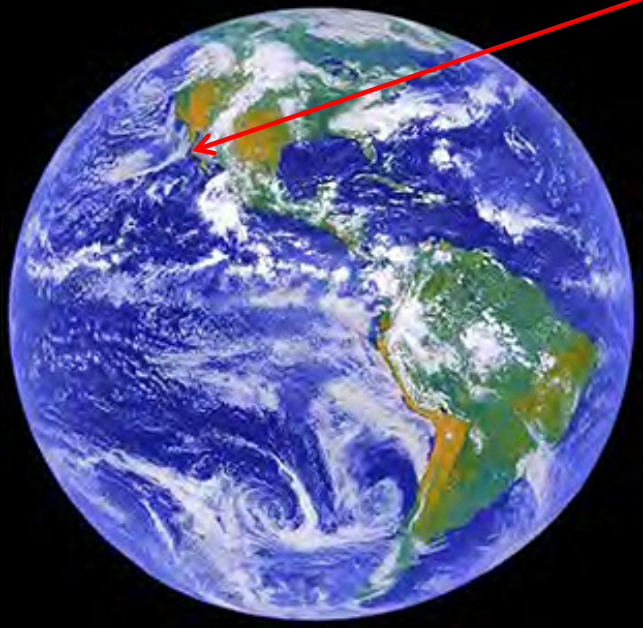


SGI-Gen/Genome- bot

Will automatically assemble DNA up to 100kb from input oligos or DNA segments

First versions will be available in 2014

Biological Teleportation



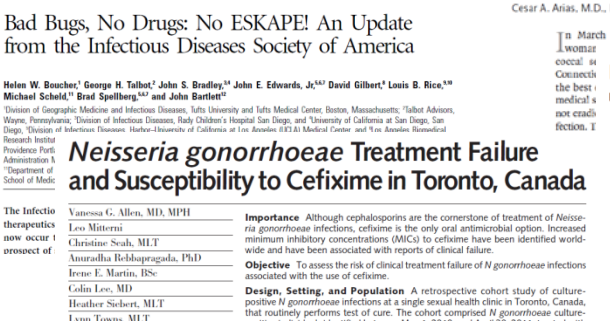
DNA sequence sent from Mars as an electromagnetic wave could arrive on Earth in as little as 4.3 min

Reconstruction of Alien Life in Secure Labs

Antibiotic Resistance

- **Antibiotic resistant pathogens cause >2M infections and >100K deaths per year**
- **Resistance to almost all available antibiotics has been developed by multiple pathogens; some have mortality rates of 30-40%**
  - *i.e.* Recent spread of cefixime-resistant gonorrhea to North America – only one drug left
- Outbreaks are challenging to control once established
  - *i.e. Klebsiella* outbreak in NIH hospital in 2011 – 19 seriously ill and 7 deaths
- Clear evidence that resistance genes are being shared between pathogens
- New antibiotic pipeline is very limited

National Institutes of Health via WikiMedia Commons



Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America

Helen W. Boucher,[1] George H. Talbot,[2] John S. Bradley,[3,4] John E. Edwards, Jr,[5,6,7] David Gilbert,[8] Louis B. Rice,[9,10] Michael Scheld,[11] Brad Spellberg,[6,8,7] and John Bartlett[12]

[1]Division of Geographic Medicine and Infectious Diseases, Tufts University and Tufts Medical Center, Boston, Massachusetts; [2]Talbot Advisors, Wayne, Pennsylvania; [3]Division of Infectious Diseases, Rady Children's Hospital San Diego, and [4]University of California at San Diego, San Diego, [5]Division of Infectious Diseases, Harbor–University of California at Los Angeles (UCLA) Medical Center, and [6]Los Angeles Biomedical Research Institut
Providence Porti
Administration M
[7]Department of
School of Medic

The Infectio
therapeutics
now occur
prosecect of

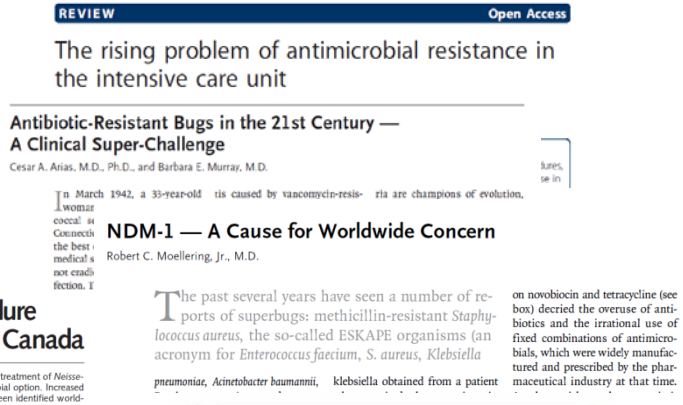
Neisseria gonorrhoeae Treatment Failure and Susceptibility to Cefixime in Toronto, Canada

Vanessa G. Allen, MD, MPH
Leo Mitterni
Christine Seah, MLT
Anuradha Rebbapragada, PhD
Irene E. Martin, BSc
Colin Lee, MD
Heather Siebert, MLT
Lynn Towns, MLT

Importance Although cephalosporins are the cornerstone of treatment of *Neisseria gonorrhoeae* infections, cefixime is the only oral antimicrobial option. Increased minimum inhibitory concentrations (MICs) to cefixime have been identified worldwide and have been associated with reports of clinical failure.

Objective To assess the risk of clinical treatment failure of *N gonorrhoeae* infections associated with the use of cefixime.

Design, Setting, and Population A retrospective cohort study of culture-positive *N gonorrhoeae* infections at a single sexual health clinic in Toronto, Canada, that routinely performs test of cure. The cohort comprised *N gonorrhoeae* culture-



REVIEW Open Access

The rising problem of antimicrobial resistance in the intensive care unit

Antibiotic-Resistant Bugs in the 21st Century — A Clinical Super-Challenge

Cesar A. Arias, M.D., Ph.D., and Barbara E. Murray, M.D.

In March 1942, a 33-year-old tis caused by vancomycin-resis- ria are champions of evolution,
woman
coccal
Connecti
the best
medical
not eradi
fection.

NDM-1 — A Cause for Worldwide Concern

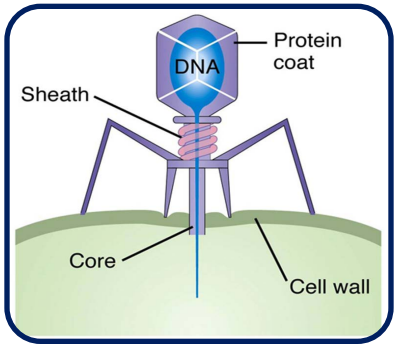
Robert C. Moellering, Jr., M.D.

The past several years have seen a number of reports of superbugs: methicillin-resistant *Staphylococcus aureus*, the so-called ESKAPE organisms (an acronym for *Enterococcus faecium*, *S. aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, klebsiella obtained from a patient

on novobiocin and tetracycline (see box) decried the overuse of antibiotics and the irrational use of fixed combinations of antimicrobials, which were widely manufactured and prescribed by the pharmaceutical industry at that time.

"There are few public health issues of potentially greater importance for society than antibiotic resistance.."- Dame Sally Davies, England's Chief Medical Officer

Bacteriophage Therapy



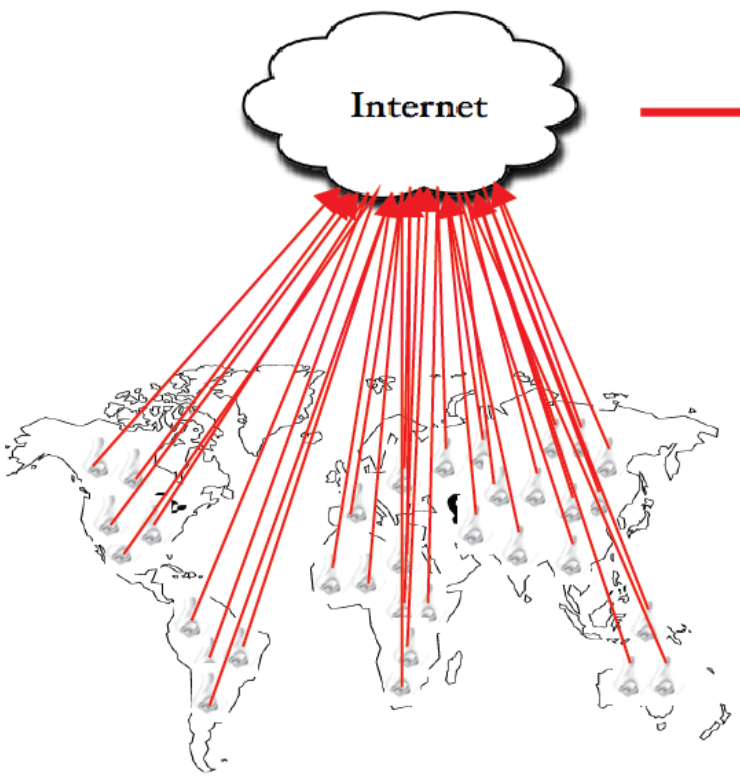
- Phages are most abundant biological entities on earth; they are viruses that target and kill bacterial cell hosts with high specificity
- Resurgence of interest in phage therapy given increasing antibiotic resistance and dearth of new antibiotics in development
- Potential advantages over antibiotics include targeting specificity, self-producing and self-limiting, and few side effects or adverse events

Current Applications

- Staphylococcal bacteriophage lysate approved in US for human use in 1959 (later limited to veterinary use in 1990)
- More recent approval of ListShield and EcoShield in US to reduce or eliminate risk of foodborne listeria and E. coli O157:H7
- Significant development activity today in areas such as agriculture, industrial biofilms, veterinary, diagnostics and food safety
- Multiple human therapeutic clinical programs in progress with natural occurring phages for organisms including S. aureus, P. aeruginosa and E. coli

Next Steps to Realizing the Complete Vision

Speed and accessibility in a pandemic response will be fully realized when we automate seq

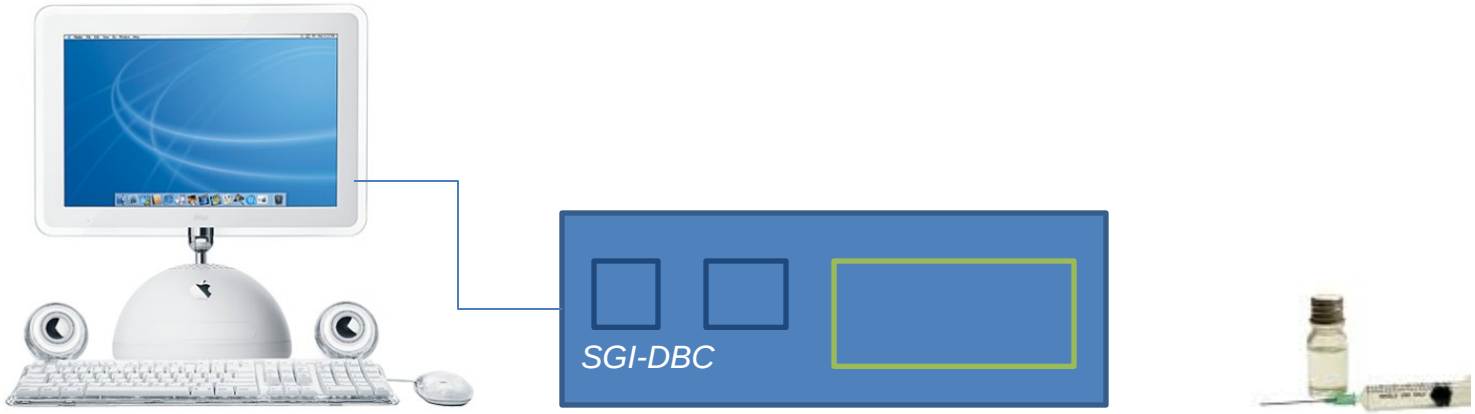


Mobile 'sending' units that directly sequence strain samples from around the world and automatically upload the sequence information to the internet

Scaled-down 'receiving' units that receive the sequence information and automatically make corresponding vaccine seeds for local distribution to manufacturers around the world

SGI Digital Biological Converter

Home Vaccine module (future)



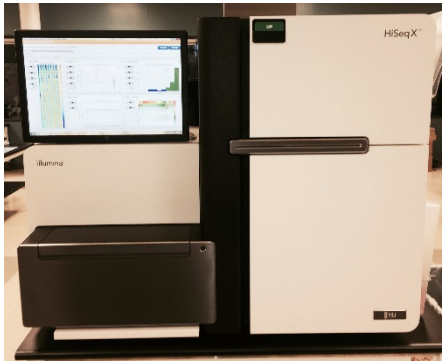
The key to gaining control over Human biology and disease depends on understanding our DNA software

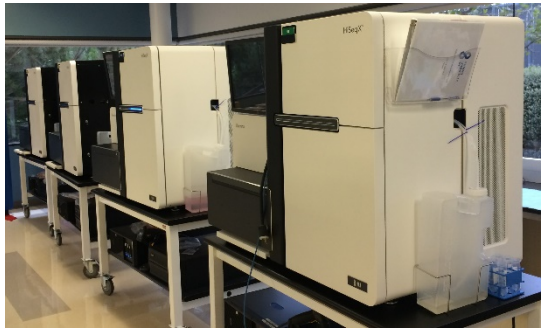
HUMAN LONGEVITY, INC.

The Research Park at Torrey Pines

- 10835 Road to the Cure, San Diego, CA
- 23,000 SQF
- 0.5 miles from UCSD Medical Center
- Planned for up to 70 FTE

Human Genome, Microbiome and Proteome Sequencing Facility and Operations

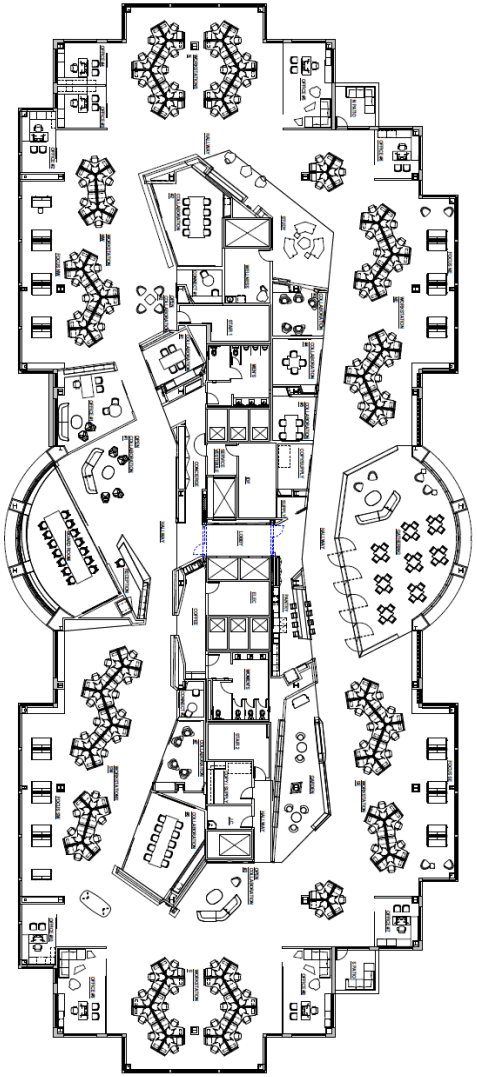




Future Corporate Headquarters

- Leadership
- Bioinformatics
- Technology
- Direct Medical Services (Future Site)
- Shared Services

- Move-in on Feb 1, 2015
- 4570 Executive Drive, San Diego, CA
- 55,000 SQF
- 0.25 miles from UCSD Medical Center
- Planned for up to 120 FTE



Global Footprint

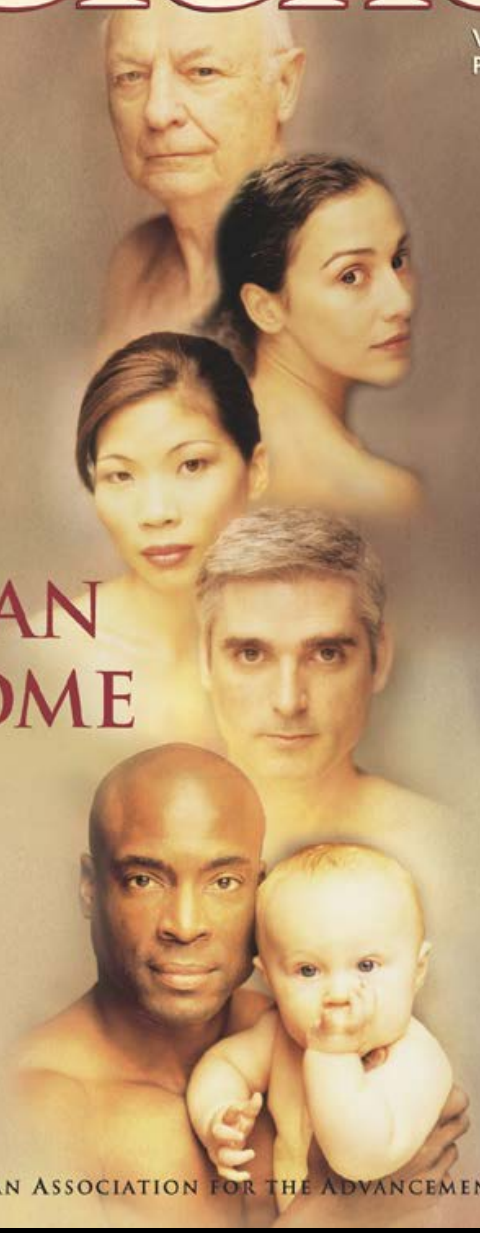
300+ teraflops; 300 terabytes flash memory; top 50 fastest computers in 2012

Amazon/Azure/Google Cloud Service Regions

16 February 2001

Science

Vol. 291 No. 5507
Pages 1145–1434 $9



THE HUMAN GENOME

American Association for the Advancement of Science

JTC Sequencer Lab

Capacity: 240,000 sequences/day or 80 million lanes/year at 24 runs per day

J. Craig Venter
INSTITUTE

Celera Computing Facility

- Celera's system was one of the most powerful civilian computing facilities in the world in 1999
- Over 1.5 teraflop of computing power in a virtual compute farm of Compaq processors with 100 terabytes storage: Cost over $50 million

OPEN ACCESS Freely available online PLoS BIOLOGY

The Diploid Genome Sequence of an Individual Human

Samuel Levy[1*], Granger Sutton[1], Pauline C. Ng[1], Lars Feuk[2], Aaron L. Halpern[1], Brian P. Walenz[1], Nelson Axelrod[1], Jiaqi Huang[1], Ewen F. Kirkness[1], Gennady Denisov[1], Yuan Lin[1], Jeffrey R. MacDonald[2], Andy Wing Chun Pang[2], Mary Shago[2], Timothy B. Stockwell[1], Alexia Tsiamouri[1], Vineet Bafna[3], Vikas Bansal[3], Saul A. Kravitz[1], Dana A. Busam[1], Karen Y. Beeson[1], Tina C. McIntosh[1], Karin A. Remington[1], Josep F. Abril[4], John Gill[1], Jon Borman[1], Yu-Hui Rogers[1], Marvin E. Frazier[1], Stephen W. Scherer[2], Robert L. Strausberg[1], J. Craig Venter[1]

1 J. Craig Venter Institute, Rockville, Maryland, United States of America, 2 Program in Genetics and Genomic Biology, The Hospital for Sick Children, and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada, 3 Department of Computer Science and Engineering, University of California San Diego, La Jolla, California, United States of America, 4 Genetics Department, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

Presented here is a genome sequence of an individual human. It was produced from ~32 million random DNA fragments, sequenced by Sanger dideoxy technology and assembled into 4,528 scaffolds, comprising 2,810 million bases (Mb) of contiguous sequence with approximately 7.5-fold coverage for any given region. We developed a modified version of the Celera assembler to facilitate the identification and comparison of alternate alleles within this individual diploid genome. Comparison of this genome and the National Center for Biotechnology Information human reference assembly revealed more than 4.1 million DNA variants, encompassing 12.3 Mb. These variants (of which 1,288,319 were novel) included 3,213,401 single nucleotide polymorphisms (SNPs), 53,823 block substitutions (2–206 bp), 292,102 heterozygous insertion/deletion events (indels)(1–571 bp), 559,473 homozygous indels (1–82,711 bp), 90 inversions, as well as numerous segmental duplications and copy number variation regions. Non-SNP DNA variation accounts for 22% of all events identified in the donor, however they involve 74% of all variant bases. This suggests an important role for non-SNP genetic alterations in defining the diploid genome structure. Moreover, 44% of genes were heterozygous for one or more variants. Using a novel haplotype assembly strategy, we were able to span 1.5 Gb of genome sequence in segments >200 kb, providing further precision to the diploid nature of the genome. These data depict a definitive molecular portrait of a diploid human genome that provides a starting point for future genome comparisons and enables an era of individualized genomic information.

HUMAN LONGEVITY, INC.

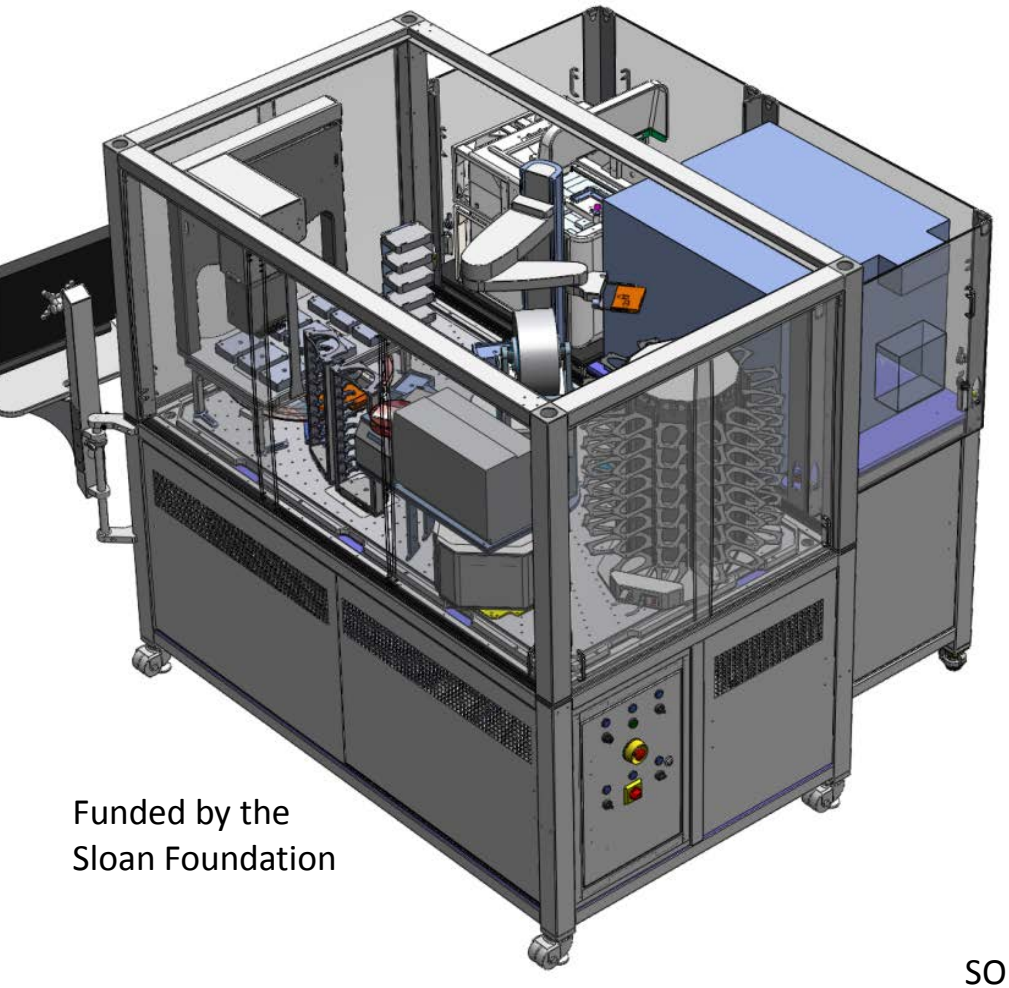
Illumina HiSeq X 10 x4

HLI plans to sequence up to 80,000 human genomes per year, with plans to rapidly scale beyond 100,000

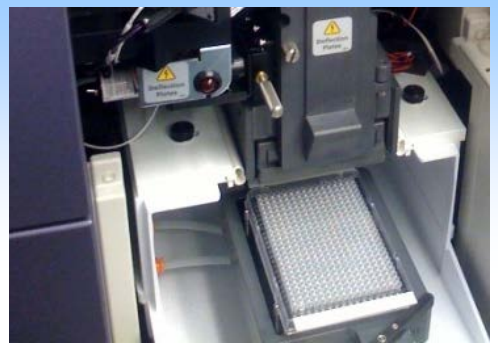
OBI-WAN
CORUSCANT

High Throughput Pipeline for Single Cell Genomics
Process for gene expression analysis

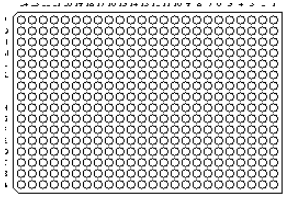
Roger Lasken, JCVI



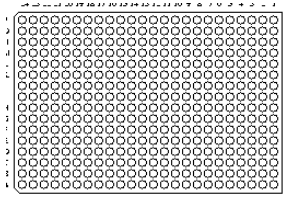
Flow sorted cells in 384-well format

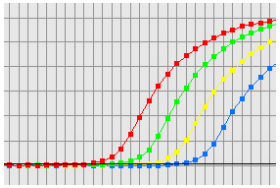


Cell lysis
cDNA synthesis
PCR amplification

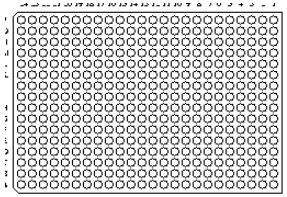


Real time qPCR for high throughput screening of Individual transcripts

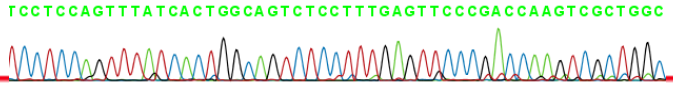




Single cell cDNA of interest moved to new plate

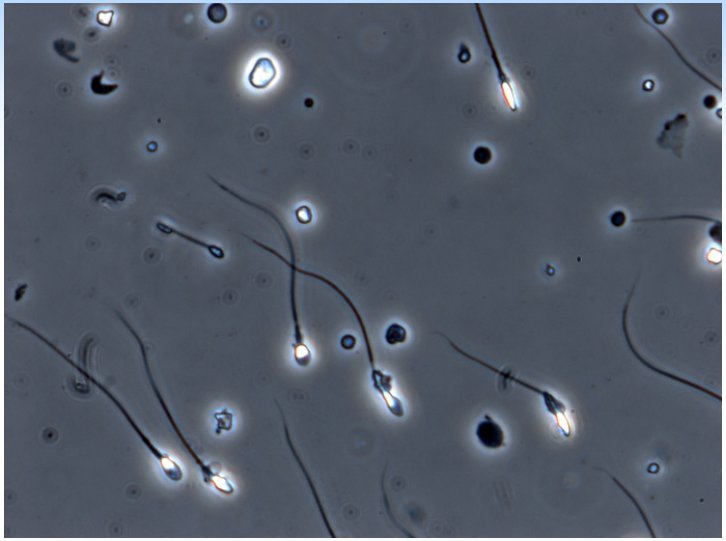


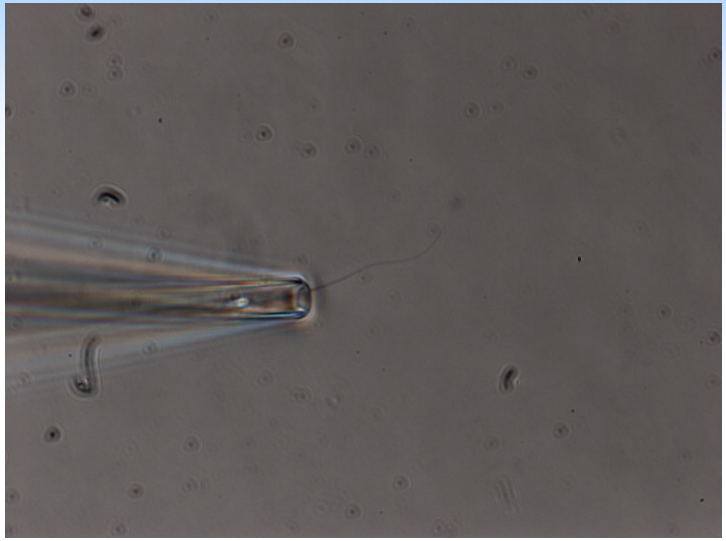
SOLiD sequencing of cDNA for global transcription analysis of selected cells

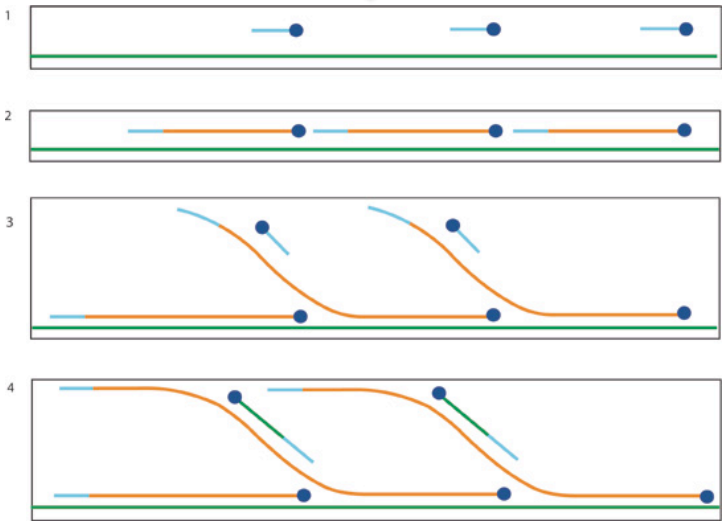


Funded by the Sloan Foundation

J. Craig Venter INSTITUTE







J. Craig Venter
INSTITUTE

Genome Research 2013

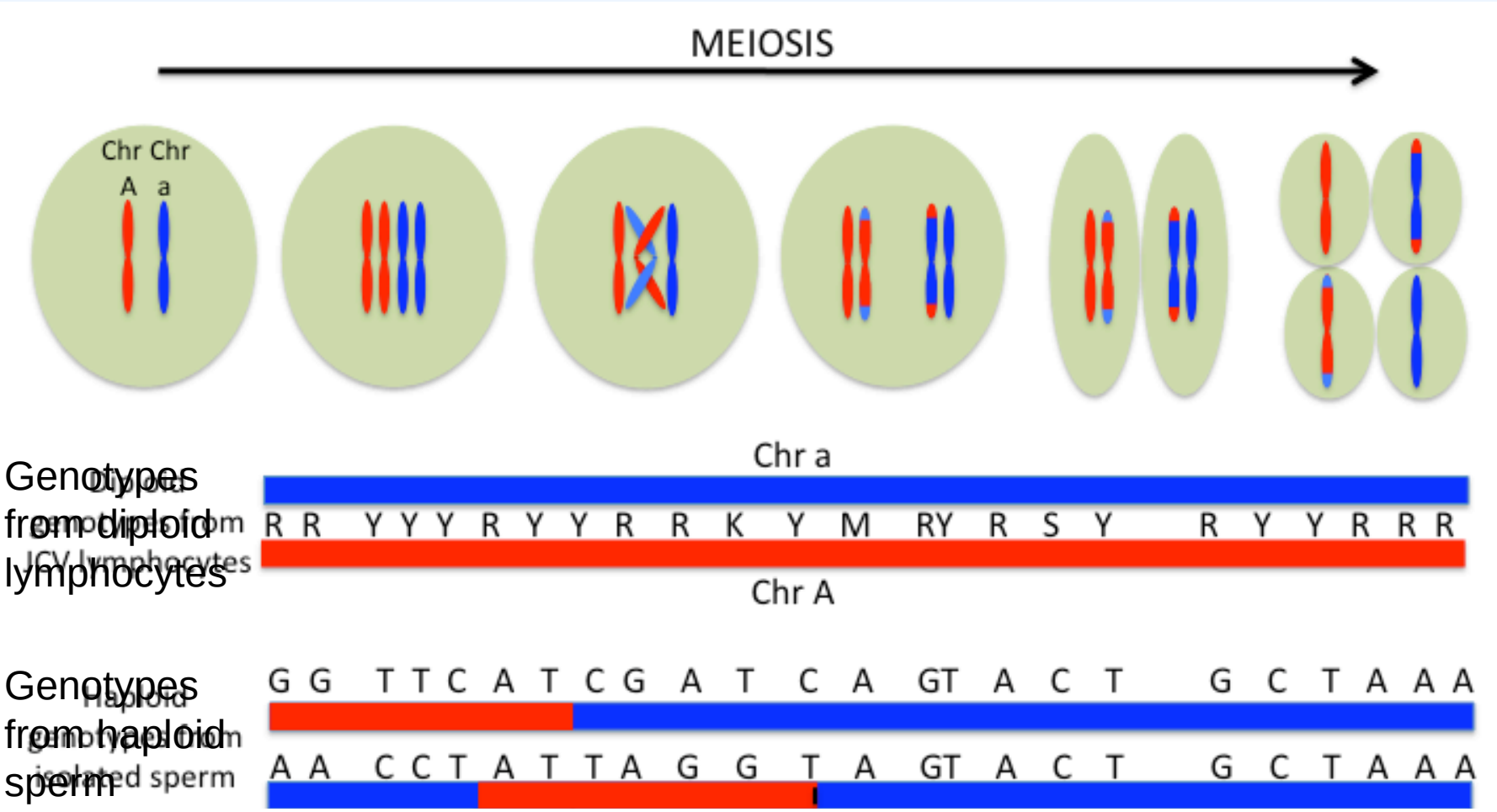
Sequencing of isolated sperm cells for direct haplotyping of a human genome

Ewen F. Kirkness, Rashel V. Grindberg, Joyclyn Yee-Greenbaum,Christian R. Marshall, Stephen W. Scherer, Roger S. Lasken, and J. Craig Venter

J. Craig Venter Institute, Rockville, Maryland 20850, USA; J. Craig Venter Institute, San Diego, California 92121, USA; University of Toronto McLaughlin Centre and The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario M5G 1L7, Canada

There is increasing evidence that the phenotypic effects of genomic sequence variants are best understood in terms of AU1 variant haplotypes rather than as isolated polymorphisms. Haplotype analysis is also critically important for uncovering 10 population histories and for the study of evolutionary genetics. Although the sequencing of individual human genomes to reveal personal collections of sequence variants is now well established, there has been slower progress in the phasing of these variants into pairs of haplotypes along each pair of chromosomes. Here, we have developed a distinct approach to haplotyping that can yield chromosome-length haplotypes, including the vast majority of heterozygous single-nucleotide AU2 polymorphisms (SNPs) in an individual human genome. This approach exploits the haploid nature of sperm cells and 15 employs a combination of genotyping and low-coverage sequencing on a short-read platform. In addition to generating chromosome-length haplotypes, the approach can directly identify recombination events (averaging 1.1 per chromosome)with a median resolution of <100 kb.

Haploid cells for genome haplotyping



Genotypes from diploid lymphocytes

Genotypes from haploid sperm

J. Craig Venter
INSTITUTE

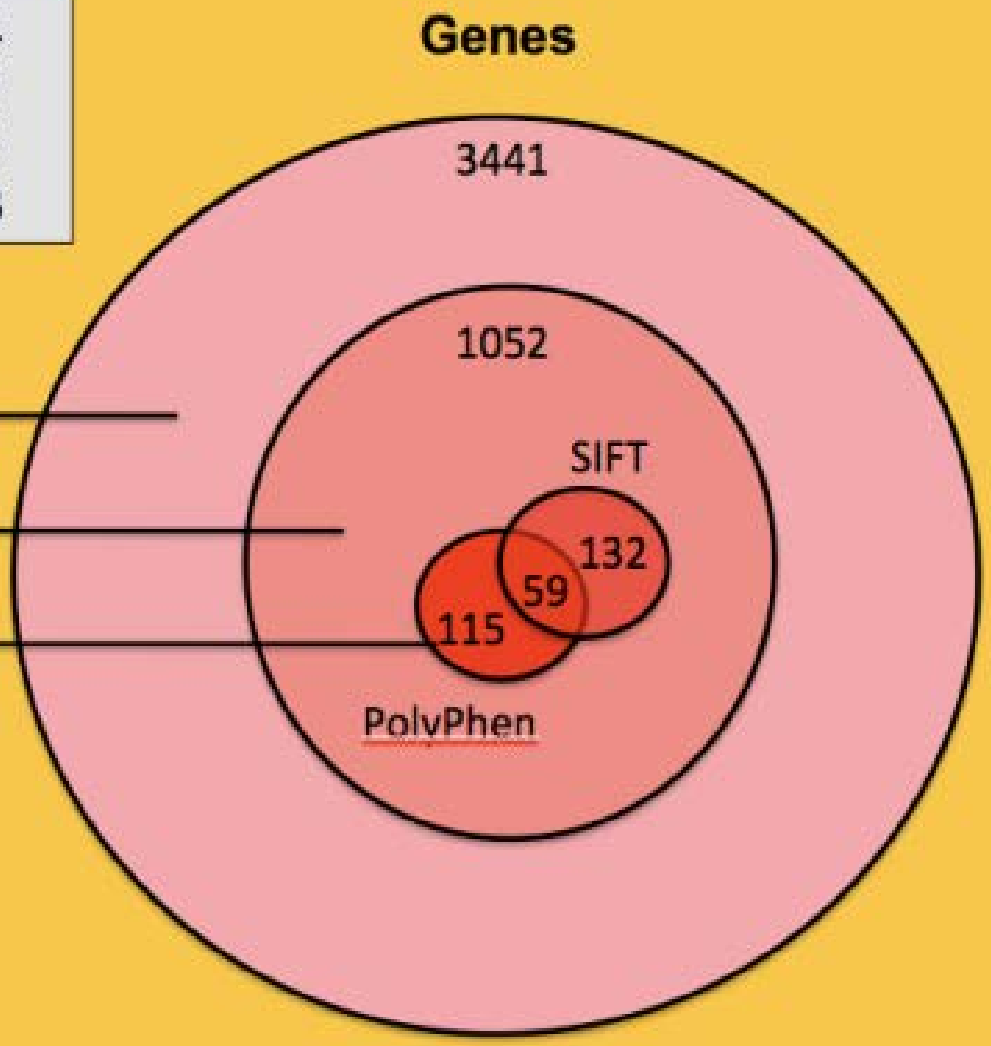
Heterozygous SNPs and multiple mutations that are predicted to be damaging in the HuRef genome

# Het SNPs	1,948,924
# In cds	12,353
# Syn	6,280
# Non-syn	6,073

Genes with non-syn SNPs

With multiple non-syn SNPs

Multiple predicted damaging



Compound Mutations

- With a majority of HuRef heterozygous SNPs phased into haplotypes, it was possible to identify genes with potentially harmful mutations on both chromosomal copies (compound mutations).
- Among the 1.9 million heterozygous SNPs, 12,353 fall within protein-coding sequences, and 6,073 involve nonsynonymous substitutions in 3,441 genes.
- For genes that contain multiple nonsynonymous mutations, predict multiple damaging mutations in 59 genes.

HUMAN LONGEVITY, INC.

UC San Diego
MOORES CANCER CENTER

Collaboration

GOAL: Champion the application of whole genome sequencing (WGS) in cancer care.

Initial Activities

- Provide germline WGS for all consenting MCC patients, via an 'opt-out' consent, for research purposes
- Leverage HLI collaborations for the provision of clinically-actionable tumor genomic profiling
- Enable UCSD researchers to explore connections between germline variants and phenotypic endpoints

Niche Blockbusters

- **Pfizer's Crizotinib**
- Treats Lung Cancer with a mutation in ALK gene
- Rare genetic change in 2-3% of 200,000 Lung Cancer patents
- Shrunk tumors in 60% of patients
- $9,600/month or $1 billion in world wide sales

- **Roche's Herceptin**
- Breast and Stomach Cancer patients with unique gene mutation
- $50,400/treatment or $6.8 billion in 2010 sales

- **Roche and Daiichi Sankyo's Zelboraf**
- Skin cancer with genetic mutation
- $56,400 for 6 month treatment: Estimated 5000 patients

Enrollment Savings if Genetic Screening Had Been Pursued for Preventive Trials: Lipid Lower Drug Trials and Cardiovascular Events, Leveraging 9p21 Genotype Data

Identifying individuals at genetic risk for CV events can increase event rates and thereby increase power to detect an effect of an intervention relative to placebo or standard

Table 4 Projected sample sizes for genotype-based trials that test cholesterol lowering treatments for preventing myocardial infarction (MI)

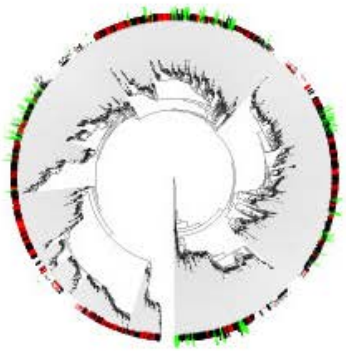
Trial	Drug	Ref	I Rate	C Rate	N Int	N Ctrl	Pow N	Gen N	Screen	R Pow
HOPE	Ramipril	1	0.099	0.123	4645	4652	2115	1554	14134 ± 223	0.735
APTC	Aspirin	2	0.024	0.035	48,736	48,882	2922	2012	18298 ± 254	0.689
FTT*	Thrombolytics	3	0.096	0.115	29,315	29,285	3228	2362	21475 ± 275	0.732
4S*	Simvastatin	4	0.189	0.291	2275	2134	216	176	1604 ± 75	0.815
EPIC	Abciximab	5	0.057	0.086	1403	696	974	693	6306 ± 149	0.711
CURE	Clopidogrel	6	0.052	0.067	6258	6303	3071	2166	19694 ± 264	0.705
14 Total trials	Many statins	7	0.044	0.062	45,054	45,002	1913	1342	12201 ± 207	0.701

Actual Number of People Enrolled

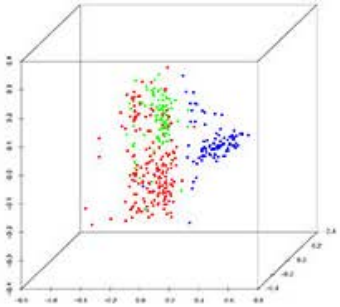
Number of Enrollees if Genotype Information Had Been Used

Schork and Topol, J BioPharm Stat, 2010

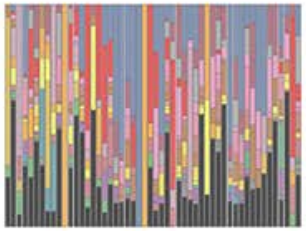
Human Longevity Microbiome Efforts



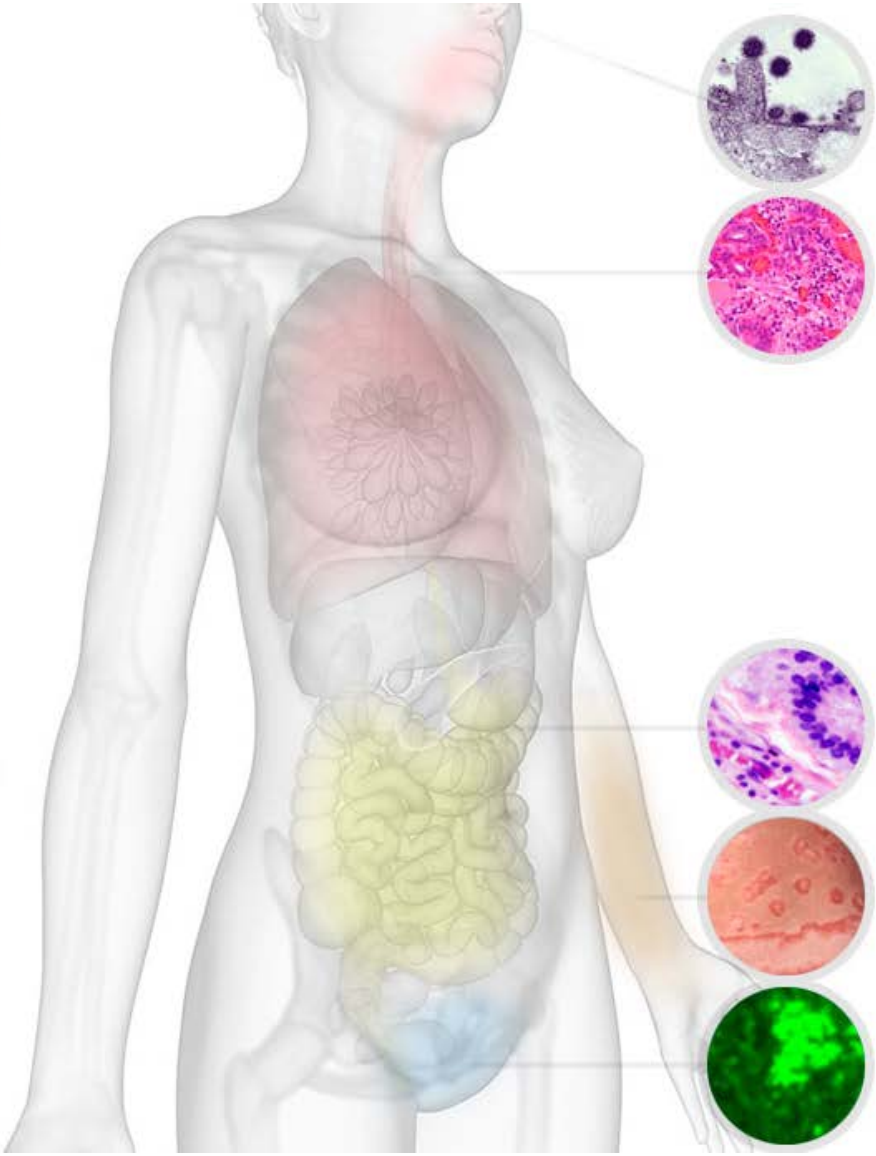
Microbial Taxonomy



Microbiome Sample Comparison



Taxonomic Composition of Microbiome



Mouth, Pharynx, Respiratory System

Stomach and Intestines

Skin

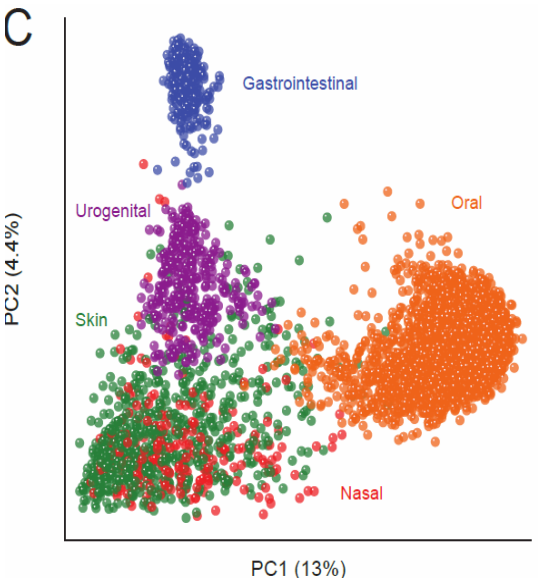
Urogenital Tract

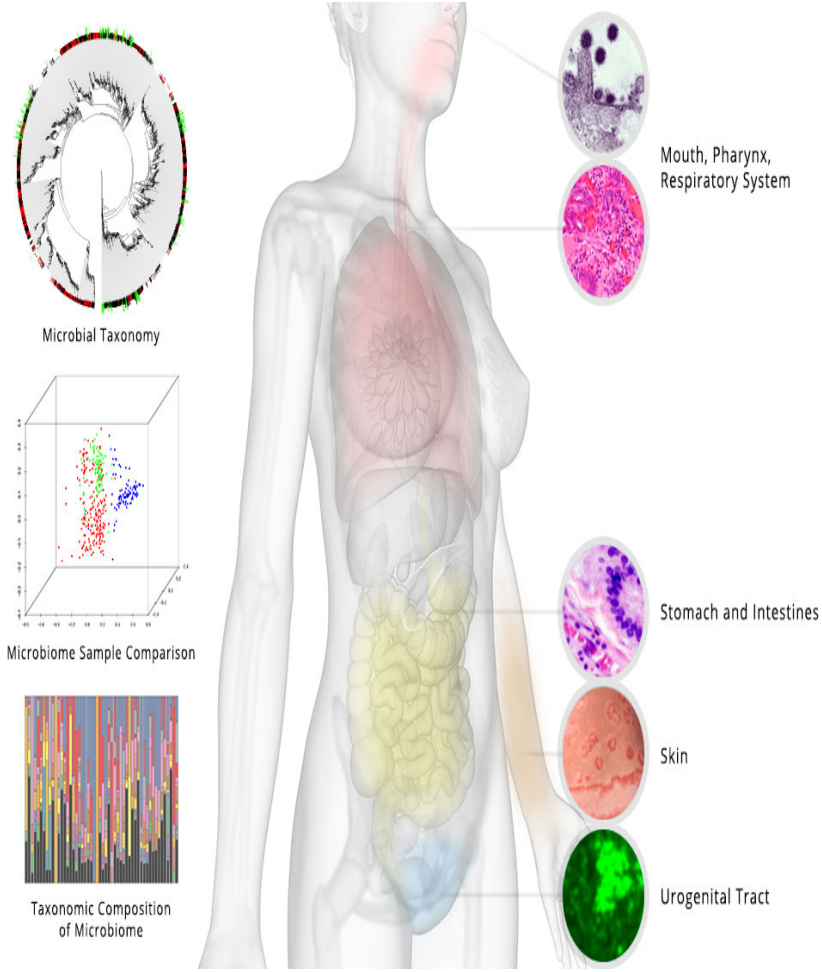
HUMAN LONGEVITY, INC.

Human microbiome – untapped component of the supraorganism

- Healthy people have different communities
- up to 3% of body mass
- Vast majority uncultured
- Essential to immune development
- Need new approaches to modulate during disease

Skin – 100-200 species
Gut – 500- 1,000 species
Oral cavity - ~700 species





Changes in the Microbiome

- Perturbations of the typical phylogenetic structure of a microbiome can alter its ecological function leading to numerous negative influences on health.
- Causative links have been made between perturbed microbiomes and numerous diseases and disorders that include, but are not limited to:
- Peptic ulcers
- Cancer
- Cardiovascular disease
- Pre-term birth
- Kidney stones
- Osteoporosis
- Obesity
- Diabetes
- Neurological disorders (including memory dysfunction, and obsessive compulsive disorder).

A Synthetic Microbiome?

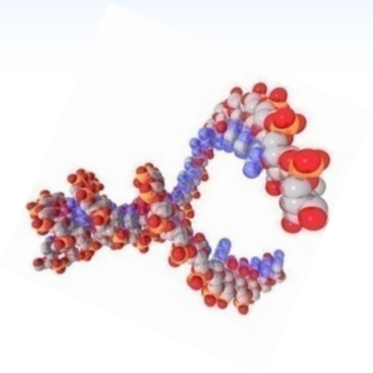
Creating a synthetic microbiome could have multiple healthcare implications

- Eliminate disease organisms which should eliminate future infections and dental decay
- Eliminate methanogens
- Eliminate sulfur producers
- Eliminate organisms associated with body odor
- Add cells that produce specific nutrients and vitamins
- Add unique metabolic systems eg. Cellulose degradation, Calcium retention, etc.

Synthetic cells will create unique IP positions for companies in the healthcare space

Metabolomics

DNA



Genomics
23,000 Genes

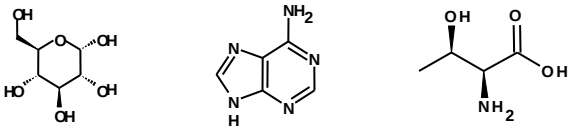
RNA

Transcriptomics
100,000 Transcripts

Protein

Proteomics
300,000 Proteins

Biochemicals
(Metabolites)



glucose adenine L-threonine

Metabolomics
2,400 Compounds

450-550 Chemicals in blood plasma

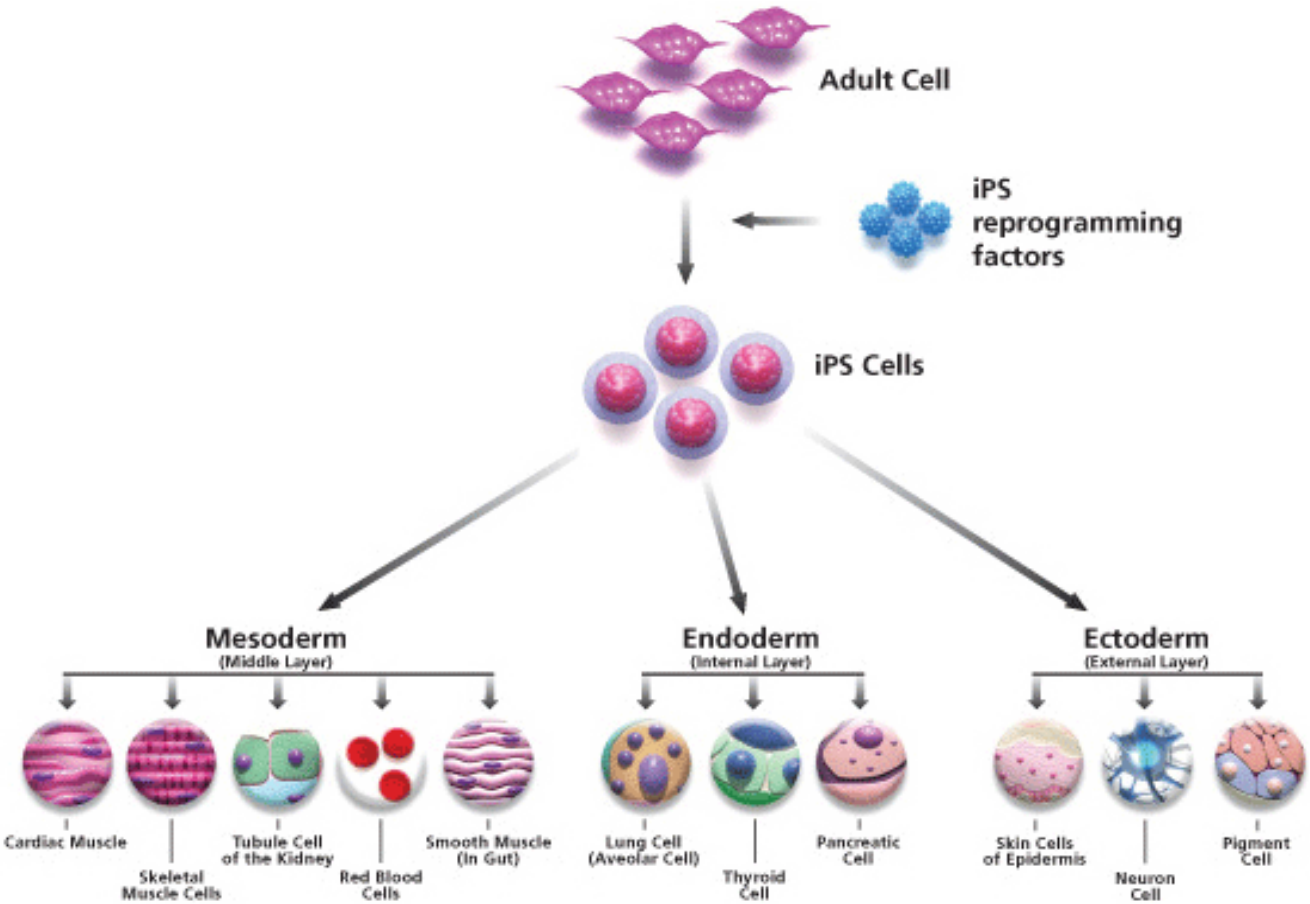
Approximately-

60% human

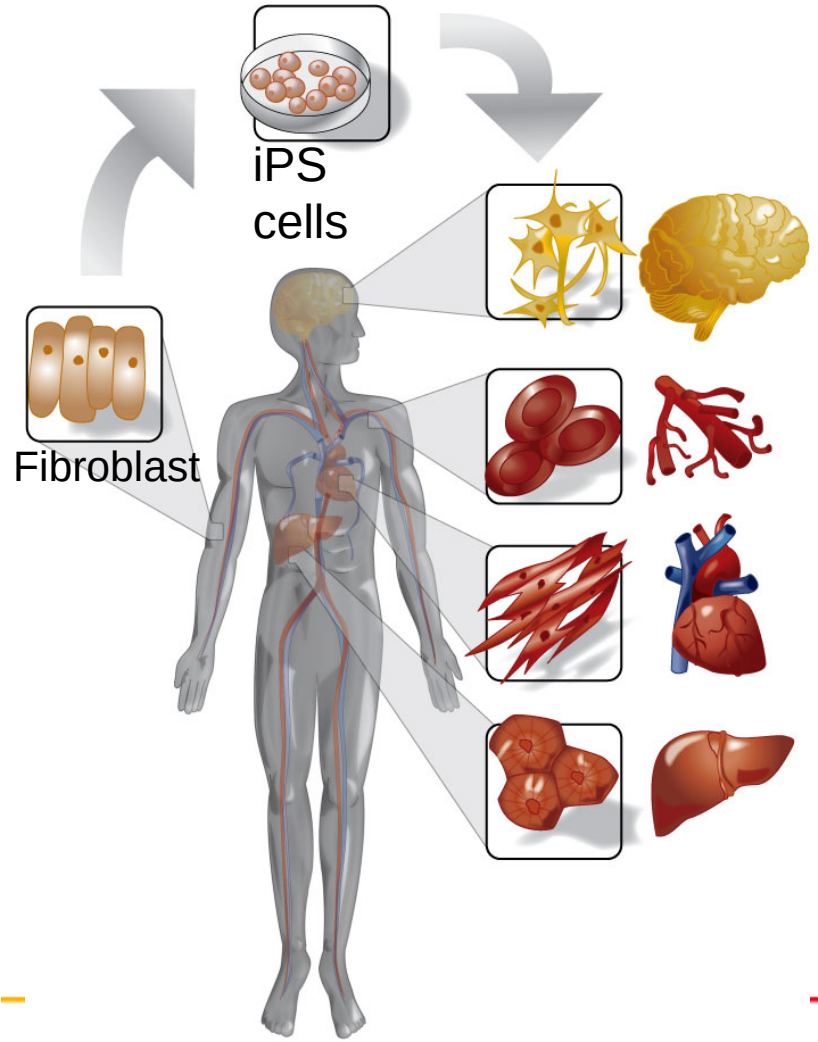
30% diet/(other species)

10% bacterial

iPS Cells


Adult Cell
iPS reprogramming factors
iPS Cells
Mesoderm
(Middle Layer)
Endoderm
(Internal Layer)
Ectoderm
(External Layer)
Cardiac Muscle
Skeletal Muscle Cells
Tubule Cell of the Kidney
Red Blood Cells
Smooth Muscle (In Gut)
Lung Cell (Aveolar Cell)
Thyroid Cell
Pancreatic Cell
Skin Cells of Epidermis
Neuron Cell
Pigment Cell

Genetic and/or Epigenetic remodeling during reprogramming



LETTER

doi:10.1038/nature10135

Immunogenicity of induced pluripotent stem cells

Tongbiao Zhao[1], Zhen-Ning Zhang[1], Zhili Rong[1] & Yang Xu[1]

Rejection of reprogrammed tissue back into the host (UCSD, Nature 2011)

ARTICLE

doi:10.1038/nature09805

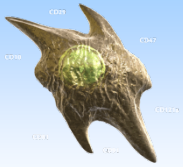
Somatic coding mutations in human induced pluripotent stem cells

Athurva Gore[1]*, Zhe Li[1]*, Ho-Lim Fung[1], Jessica E. Young[2], Suneet Agarwal[3], Jessica Antosiewicz-Bourget[4], Isabel Canto[2], Alessandra Giorgetti[5], Mason A. Israel[2], Evangelos Kiskinis[6], Je-Hyuk Lee[7], Yuin-Han Loh[3], Philip D. Manos[3], Nuria Montserrat[5], Athanasia D. Panopoulos[8], Sergio Ruiz[8], Melissa L. Wilbert[2], Junying Yu[4], Ewen F. Kirkness[9], Juan Carlos Izpisua Belmonte[5,8], Derrick J. Rossi[10], James A. Thomson[4], Kevin Eggan[6], George Q. Daley[1], Lawrence S. B. Goldstein[2] & Kun Zhang[1]

Estimated six protein-coding point mutations per exome

Mutations were enriched in cancer-causing genes (UCSD and JCVI, Nature 2011)

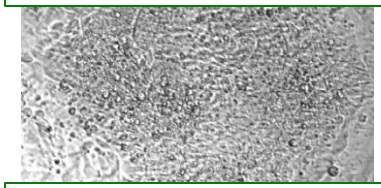
Fibroblasts



ROUTE A

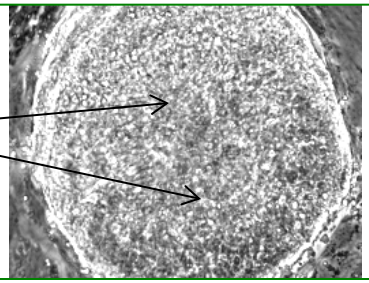
ROUTE B

Reprogramming



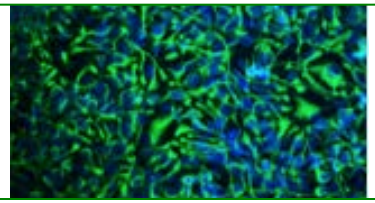
iPSC colony

Differentiation



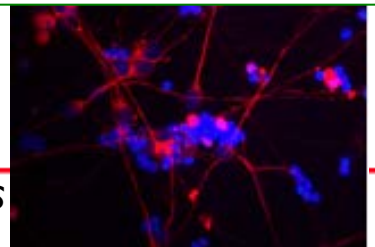
neural rosettes

Expansion



(Nestin+) NSCs

Differentiation



(Map2+) neurons

iPSC
897,680

NSC
906,218

Neurons
899,864

iPSC
879,935

NSC
917,524

Neurons
1,101,231

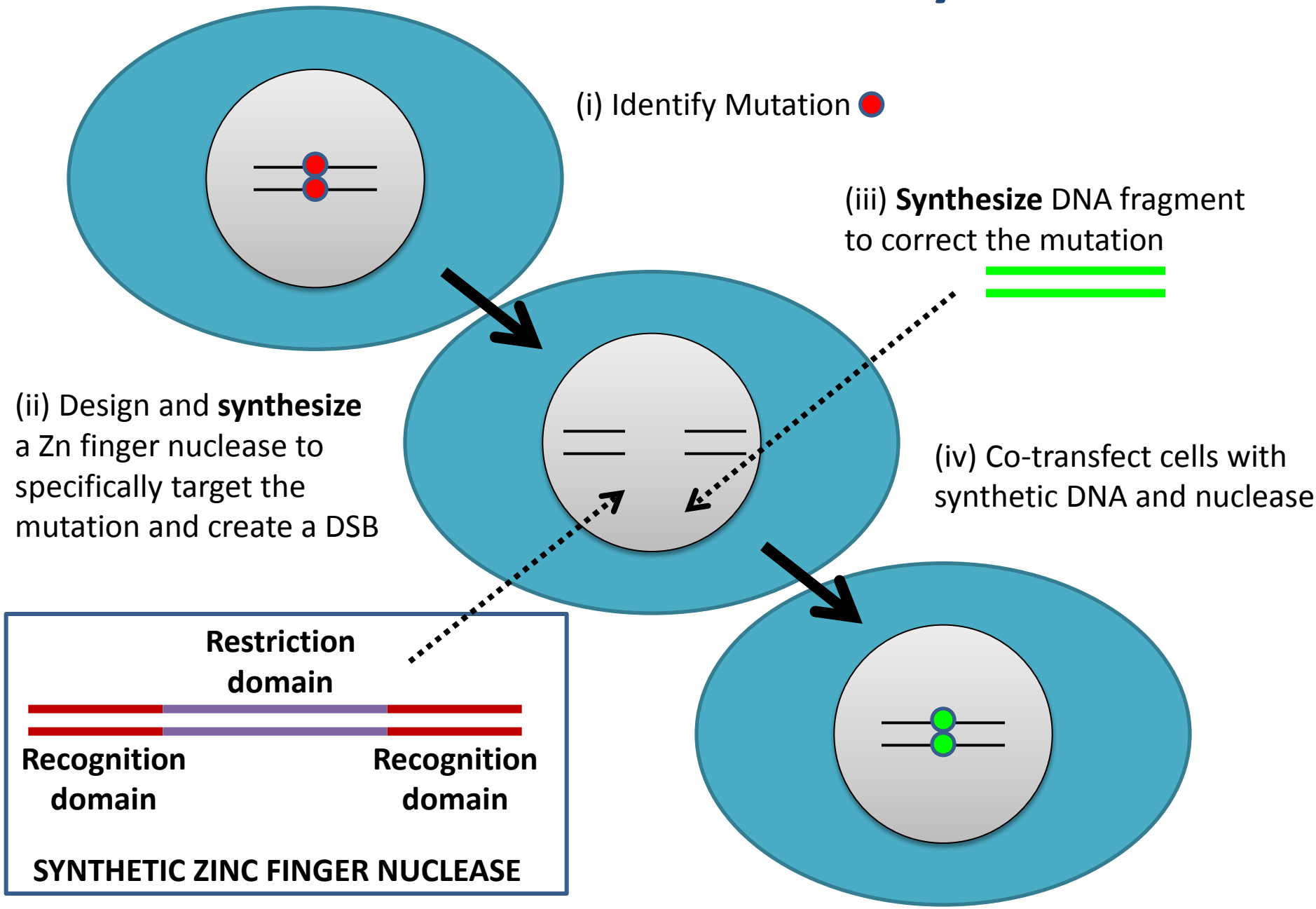
NSC

Neurons
1,125,530

J. Craig Venter INSTITUTE

Goldstein and Venter Lab

Stem Cell Error Correction With Synthetic DNA



Human Phenotype

We are developing, in a digitized format, a complete description of the human phenotype on at least 1,000,000 people by 2020

HUMAN LONGEVITY, INC.

Human Phenotype

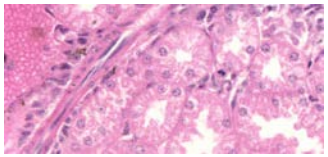
- Complete history and physical examination
- Whole genome sequencing
- 2400 blood metabolites
- Oral, urinary and fecal microbiome sequencing
- Structural brain MRI for volumetric assessments
- Whole body MRI for visceral adipose tissue, subcutaneous adipose tissue, skeletal muscle mass and intrahepatic triglycerides.
- Bone density
- Ocular, visual and hearing acuity
- Over 40 blood biomarkers
- Urinalysis
- Resting Energy Expenditure
- Maximal oxygen consumption (VO2Max)
- Echocardiography
- Electrocardiography
- Carotid Intimal Media Thickness (CIMT)
- Full pulmonary spirometry
- Sleep apnea testing with the ApneaLink™ Air
- Balance, gait and strength assessment
- Neurocognitive, mood and behavior assessment
- Sleep and activity for one-week
- 7-day environmental toxin exposure with the MyExposome® passive sampler wristband

HUMAN LONGEVITY, INC.

HUMAN LONGEVITY, INC.

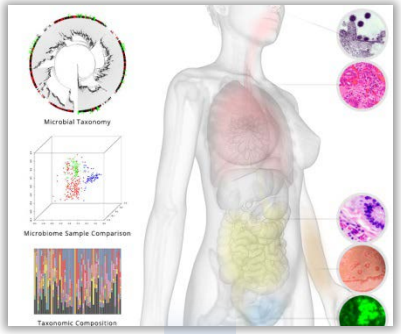
An Integrative Approach to Medicine

Clinical Information/Samples



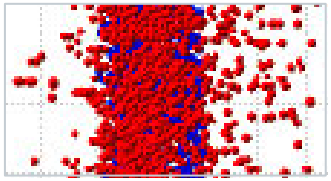
UC San Diego
MOORES CANCER CENTER

Microbiome



Human Genomics

Metabolomics

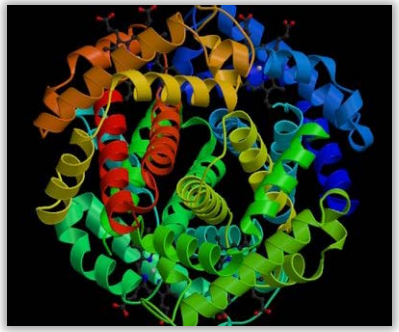


METABOLON®

Phenotyping

Subject Database

Proteomics



Novel, Actionable Discoveries

HUMAN LONGEVITY, INC.

The Premier Human Health Database

Multiple Sources

- Universities
- Hospitals / Clinics
- Governments
- Pharmaceutical companies

Multiple Targets

- Cancer and beyond

Bipolar Disorder

Neck & Head Cancers

Alzheimer's

Autism

Schizophrenia

ALS

Other CNS Disorders

Respiratory Pathogens

Twins

Bone Related Diseases

Age & Older People

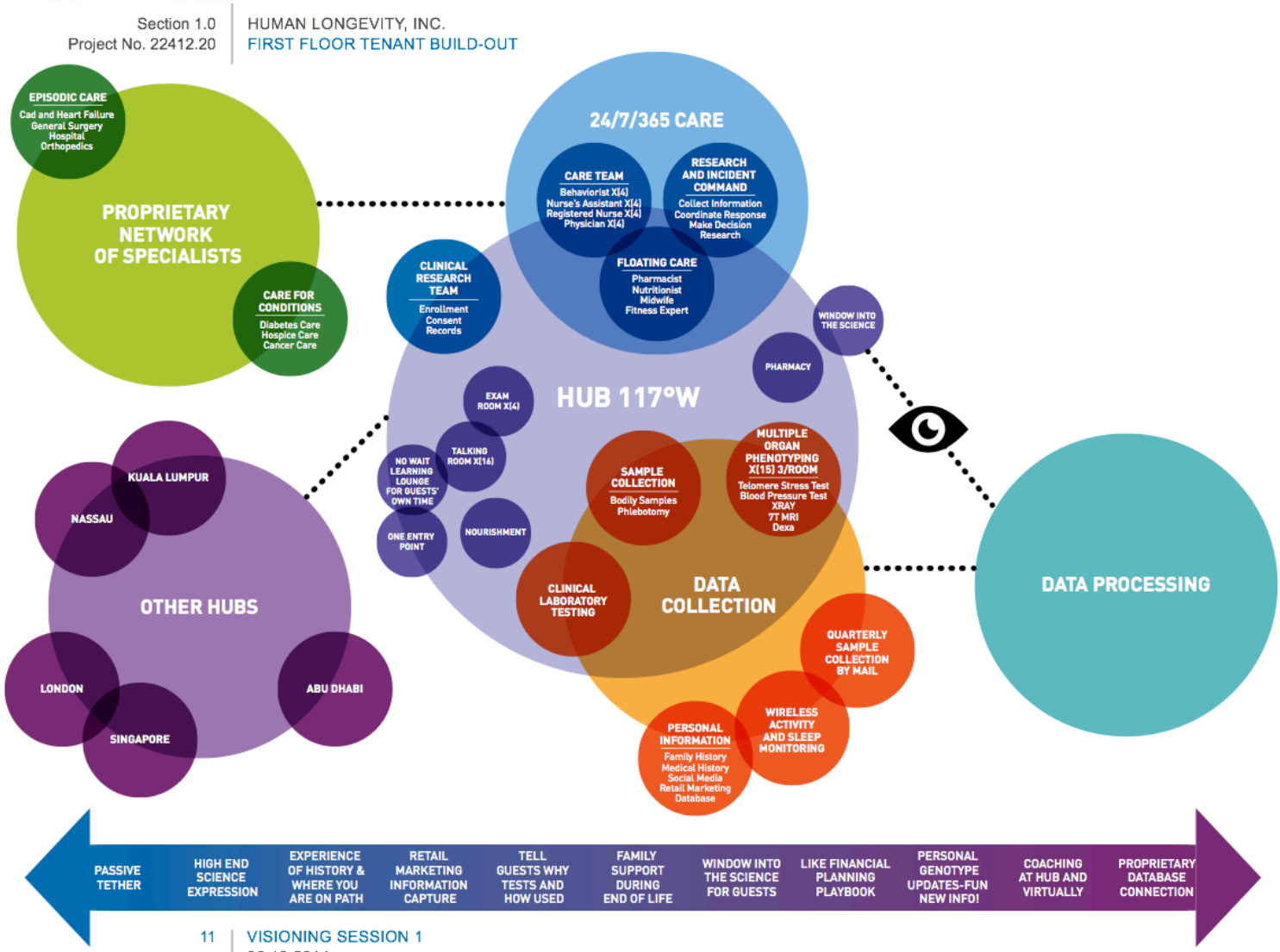
Coronary Artery Disease

Fatty Liver Disease

1M Integrated Human Records by 2020

Medical Services: "Health Hub"

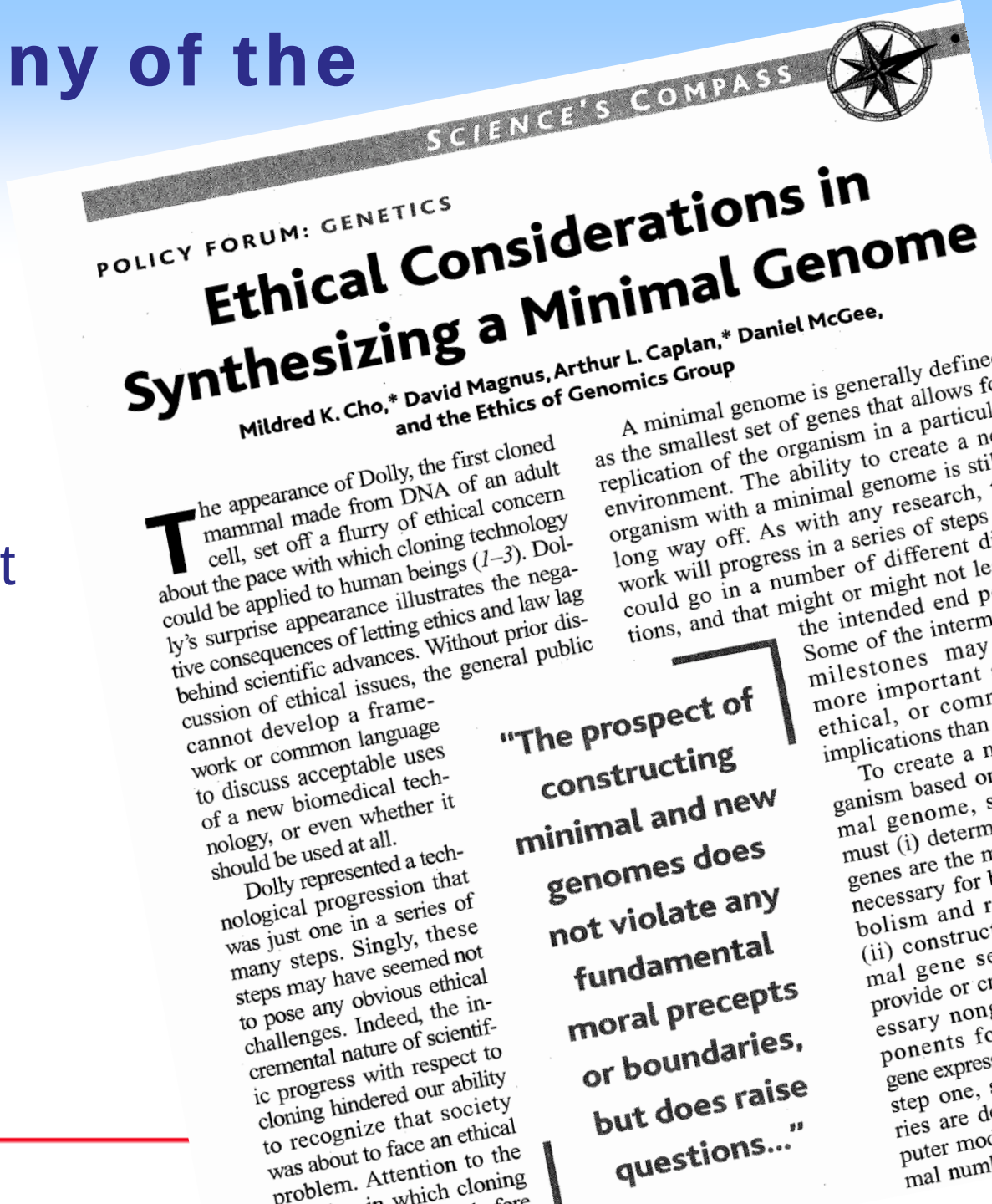
Care Model
Phenotyping
Client Experience
Informatics
Telemedicine
Remote Monitoring
Marketing
Business Planning
Architecture



Earliest Scrutiny of the Technology:

December 1999 review:

- Ethical and religious concerns
- Environment impact
- Laboratory safety
- Distribution of benefits and risks
- Biological warfare



SCIENCE'S COMPASS

POLICY FORUM: GENETICS

Ethical Considerations in Synthesizing a Minimal Genome

Mildred K. Cho,* David Magnus, Arthur L. Caplan,* Daniel McGee, and the Ethics of Genomics Group

The appearance of Dolly, the first cloned mammal made from DNA of an adult cell, set off a flurry of ethical concern about the pace with which cloning technology could be applied to human beings (1–3). Dolly's surprise appearance illustrates the negative consequences of letting ethics and law lag behind scientific advances. Without prior discussion of ethical issues, the general public cannot develop a framework or common language to discuss acceptable uses of a new biomedical technology, or even whether it should be used at all.

Dolly represented a technological progression that was just one in a series of many steps. Singly, these steps may have seemed not to pose any obvious ethical challenges. Indeed, the incremental nature of scientific progress with respect to cloning hindered our ability to recognize that society was about to face an ethical problem. Attention to the

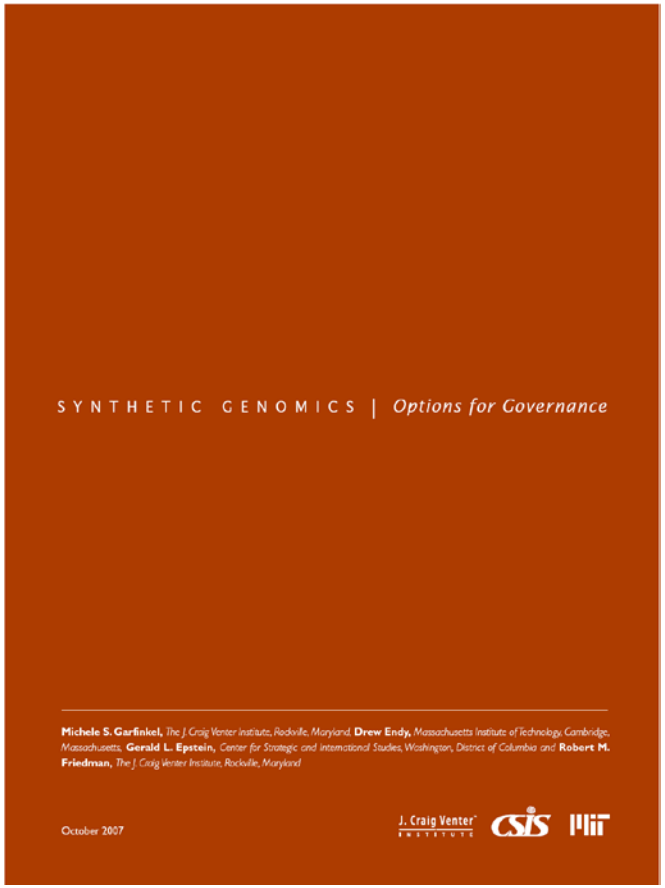
"The prospect of constructing minimal and new genomes does not violate any fundamental moral precepts or boundaries, but does raise questions..."

A minimal genome is generally define as the smallest set of genes that allows fo replication of the organism in a particul environment. The ability to create a n organism with a minimal genome is sti long way off. As with any research, work will progress in a series of steps could go in a number of different di tions, and that might or might not le the intended end p Some of the interm milestones may more important ethical, or comr implications than To create a n ganism based o mal genome, s must (i) determ genes are the n necessary for bolism and r (ii) construc mal gene se provide or c essary non ponents fo gene expres step one, ries are d puter mod mal num

J. Craig Venter INSTITUTE

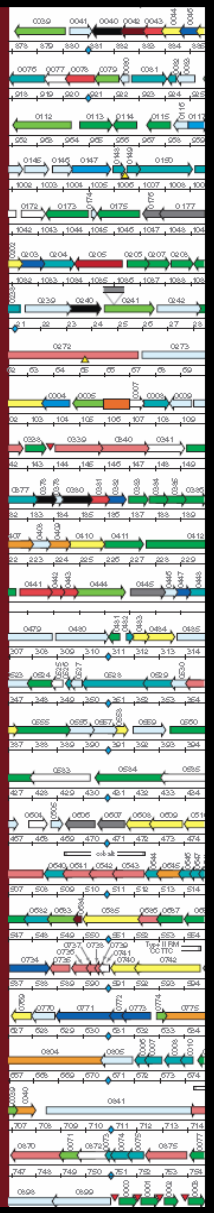
Governance efforts: biosecurity and biosafety

- ***Synthetic Genomics: Options for Governance***
  - 20 month study funded by the Alfred P. Sloan Foundation
- Organizations
  - Venter Institute
  - Center for Strategic & International Studies
  - MIT Synthetic Biology Group
- Core Group
  - Additional 14 scientists, policy analysts, etc.



NEW DIRECTIONS

The Ethics of Synthetic Biology and Emerging Technologies



Presidential Commission for the Study of Bioethical Issues

December 2010

J. CRAIG VENTER

LIFE AT THE SPEED OF LIGHT

FROM THE

DOUBLE HELIX

TO THE

DAWN OF DIGITAL LIFE

BY THE AUTHOR OF A LIFE DECODED