NSABB July 1, 2005

Synthetic Genomics
From Reading to Writing the Genetic Code

Genes are the design components of the future

J. Craig Venter
INSTITUTE

First Genome Sequenced 1995

Ten years ago this month


nature
INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima
Science
24 March 2000
Vol. 287 No. 5461
Pages 2105–2364 $8
The Drosophila Genome
SCIENCE
23 AUGUST 1996
VOL. 273 • PAGES 1017–1140
nature
International weekly journal of science
11 December 1997
The Dog Genome: Survey Sequencing and Comparative Analysis
Ewen F. Kirkness,[1] Vineet Bafna,[2*] Aaron L. Halpern,[2*] Samuel Levy,[2*] Karin Remington,[2*] Douglas B. Rusch,[2*] Arthur L. Delcher,[1] Mihai Pop,[1] Wei Wang,[1] Claire M. Fraser,[1] J. Craig Venter[2]
A survey of the dog genome sequence (6.22 million sequ
coverage) demo
RESEARCH ARTICLE
A Comparison of Whole-Genome Shotgun-Derived Mouse Chromosome 16 and the Human Genome
Science
4 October 2002
Vol. 298 No. 5591
Pages 1–310 $10
THE MOSQUITO GENOME
Anopheles gambiae
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE
nature
International weekly journal of science
www.naturejpn.com
smodium genomics
omics and proteomics pave way for controlling malaria
Antarctic ice
Flow reversals
Antigen presentation
A 'customizing' protease
naturejobs
genomics & parasitology
Science
16 February 2001
Vol. 291 No. 5507
Pages 1145–1434 $9
THE HUMAN GENOME
AMERICAN ASSOCIATION FOR THE
10 March 2000
e Sequence
ingitidis
in MC58
Vaccine
st
ngococcu

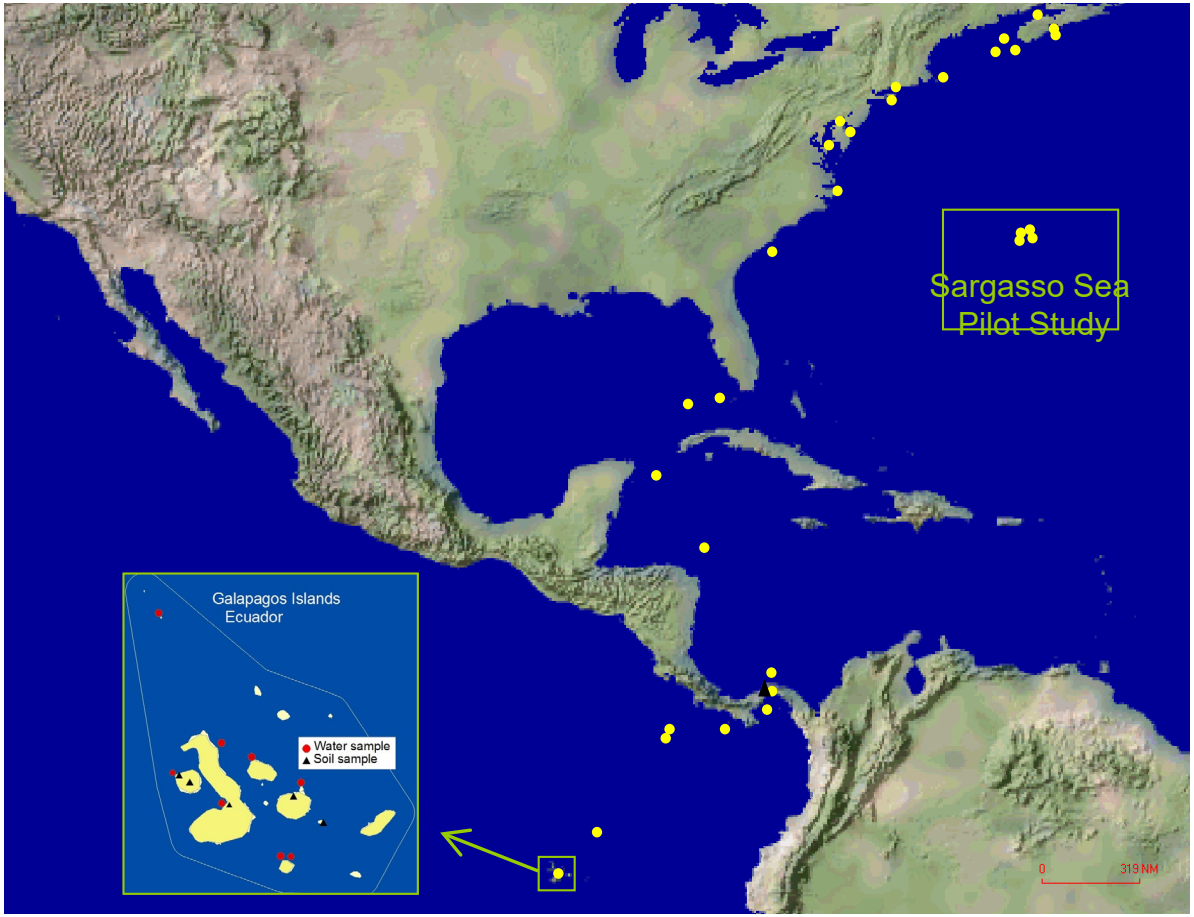
Microbial Abundance

- Microbes make up roughly ½ of the Earth's biomass
- 6 x 10^{30} microbes over entire Earth
- Animals make up 1/1000th of Earth's total biomass
- Each ml of sea water has one million bacteria and 10 million viruses

www.sorcerer2expedition.org

J. Craig Venter
INSTITUTE

Sorcerer II Global Ocean Survey:
Nova Scotia through Galapagos Islands



J. Craig Venter
INSTITUTE


100 Million letters of genetic code
every 24 hours

454 Genome Sequencing System

- Up to 100x throughput over Fluorescent Sequencing:
  - 20 megabases/4 hr instrument run (20-32 high Q)

Current vs 454

Platform	3730	454
Bases per day	945K	100-200 million
Runs per day	12	5
Bases per run	78.7K	20-40 million
Average RL	820bp	125bp
Cost	~$0.0012/base ($1.20/Kb)	$0.00025/base ($0.25/Kb)

Marine Microbe Sequencing Project

- $9M Funded by Gordon and Betty Moore Foundation
  - Sequence, Assemble and Auto-Annotate up to 130 Marine Microbes
  - Add to the current (10-20) set of sequenced Marine Microbes
    - 1000% increase
  - Synergy with Sorcerer Expedition

Sorcerer II Expedition and Moore Foundation Isolated Organisms

- Sampled stations
- Planned stations

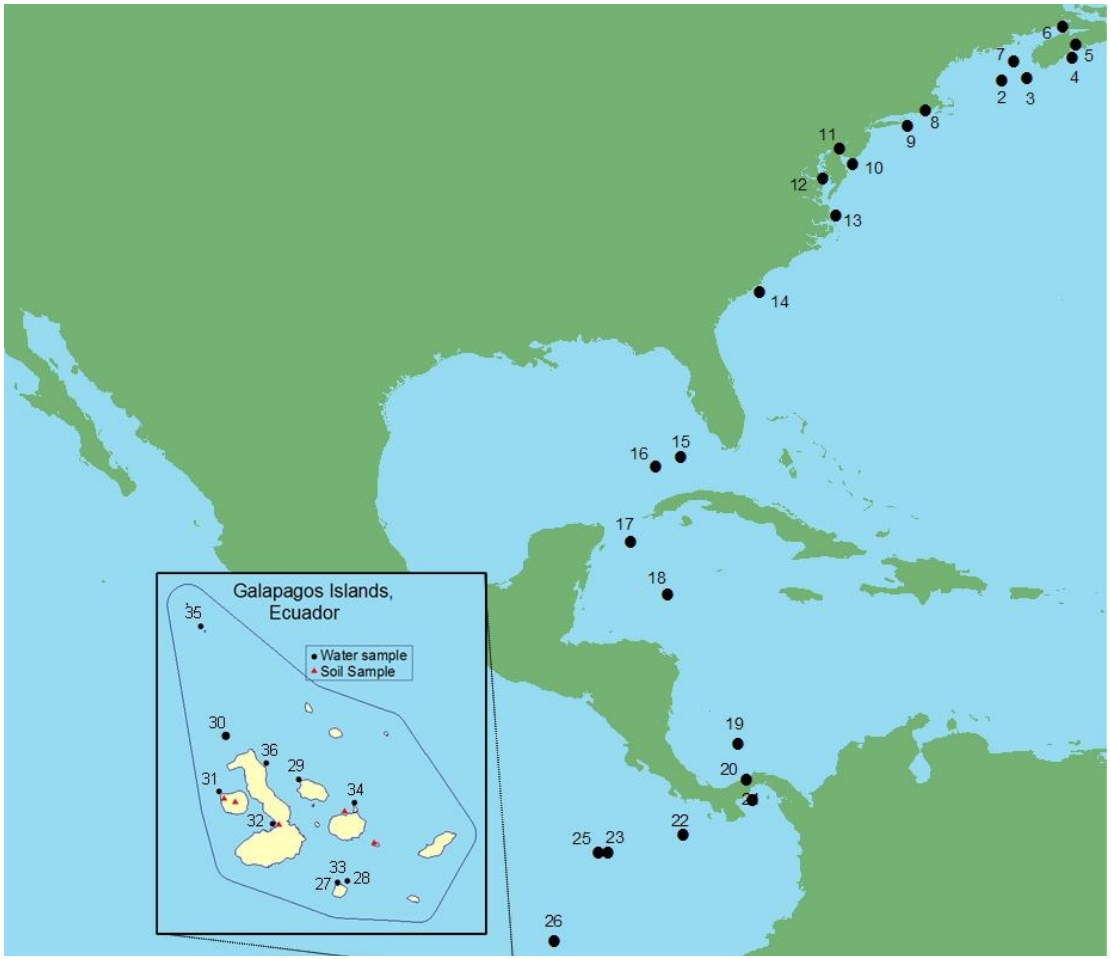
Number of organisms isolated from each region for the Moore Foundation Program

J. Craig Venter INSTITUTE

Aerosol sample collection

- Samples 1360 m^3 air in 24 hrs (58 m^3/hr)
- Average 1 X 10^9 bacteria/filter/day
- Multi-day sample averages known diurnal variation in microbial composition

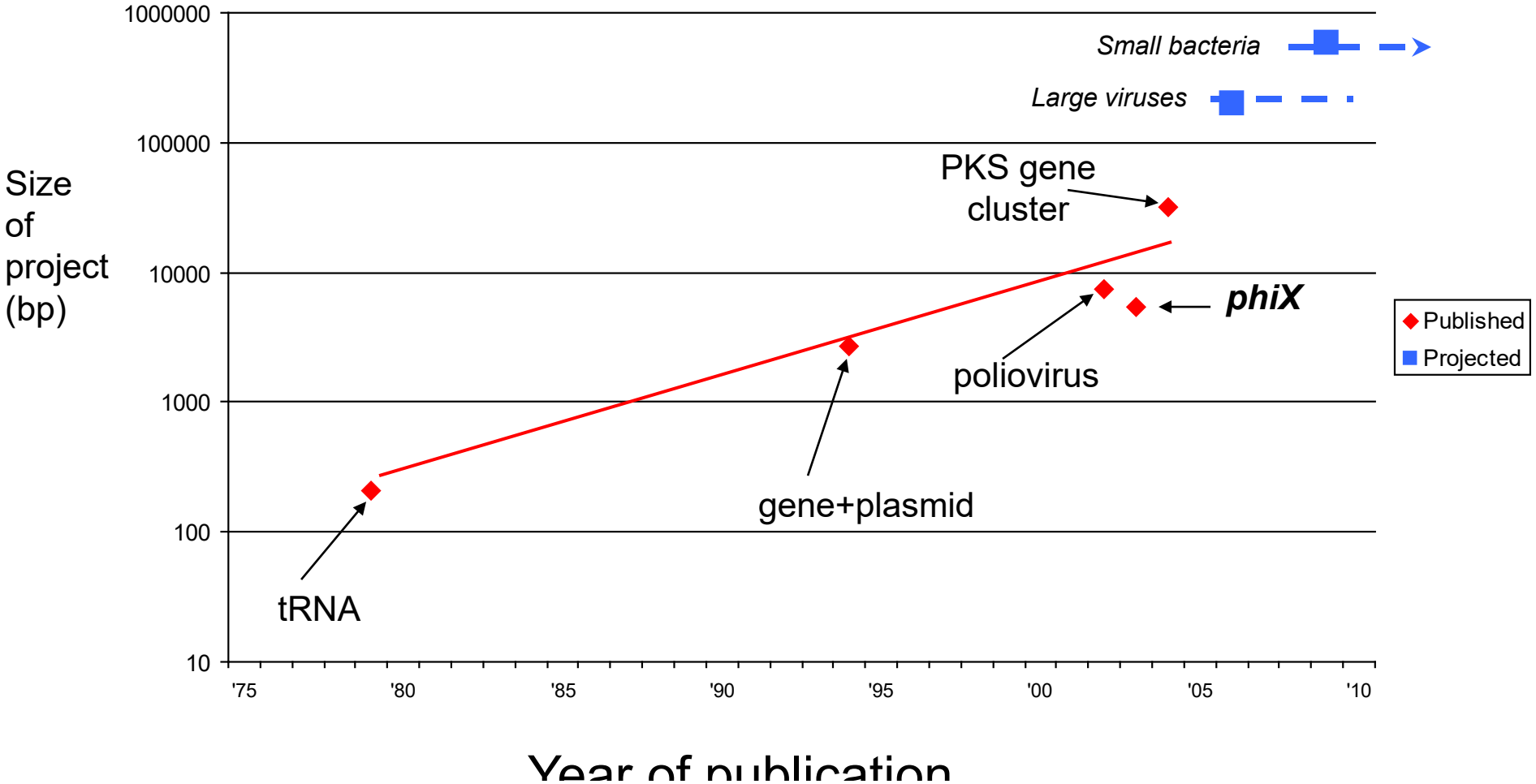
Preliminary Analysis: Nova Scotia through Galapagos



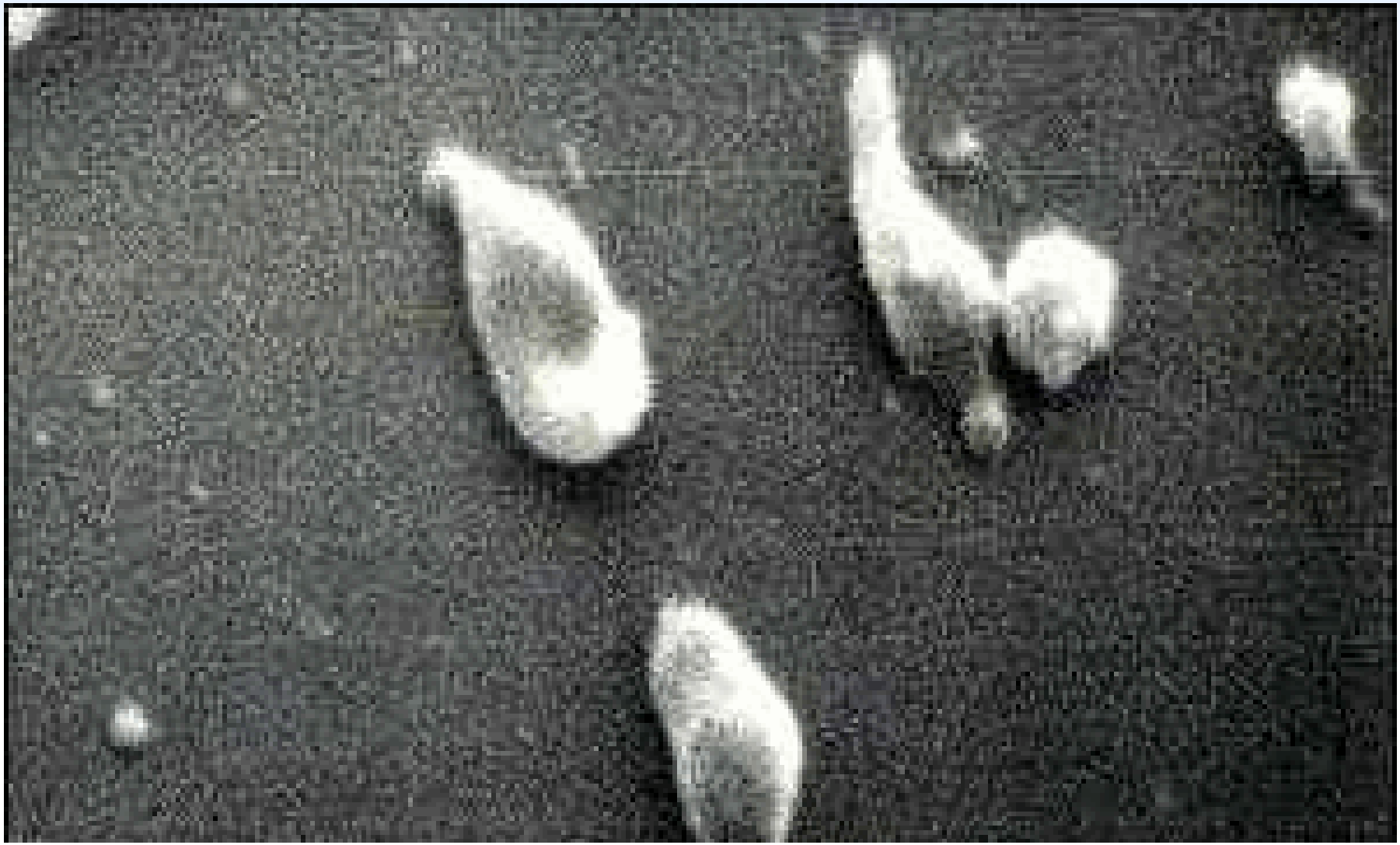
Synthetic Genomics

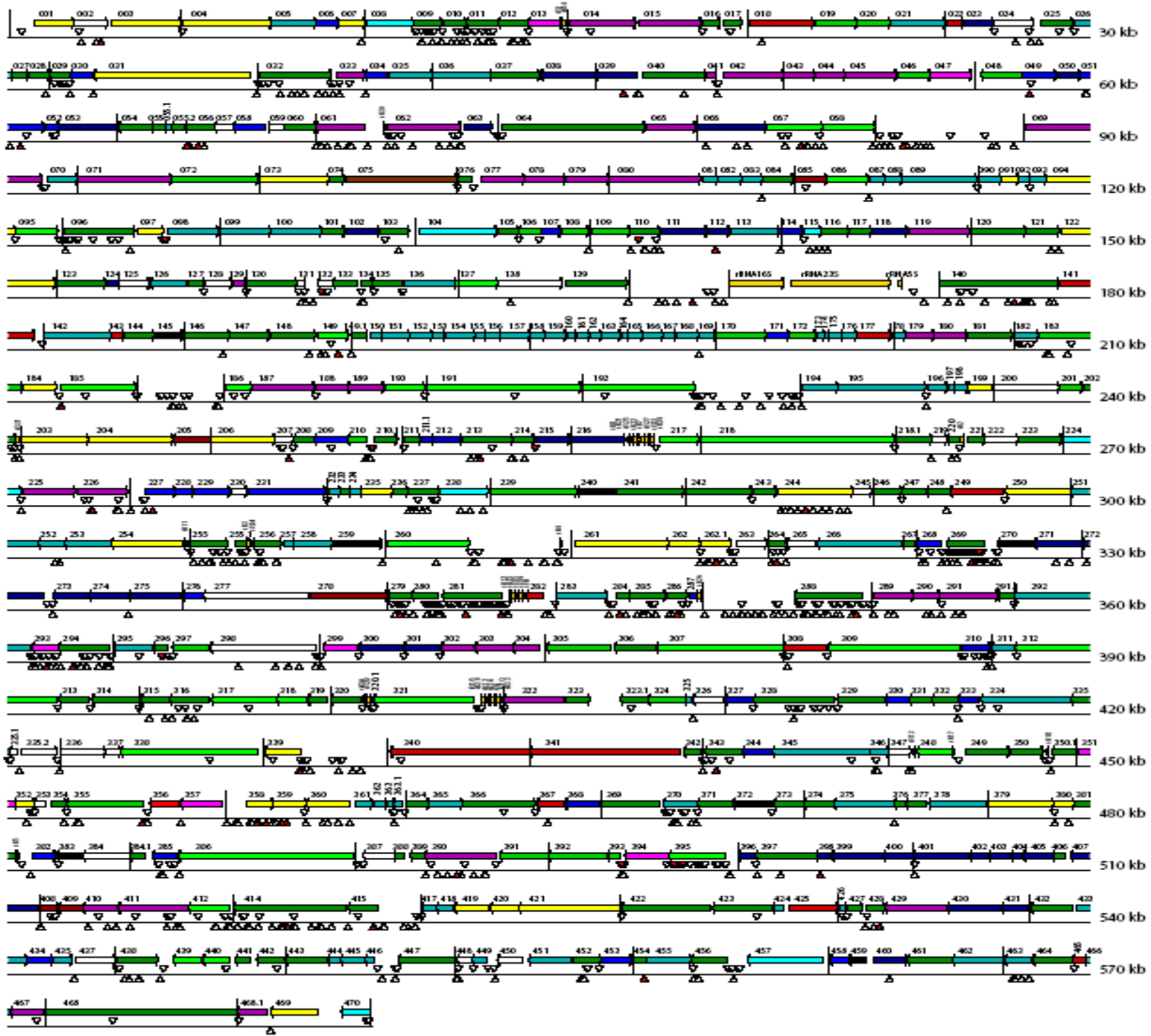
Design and construction of genomes from scratch

Chemical synthesis of DNA: from genes to genomes

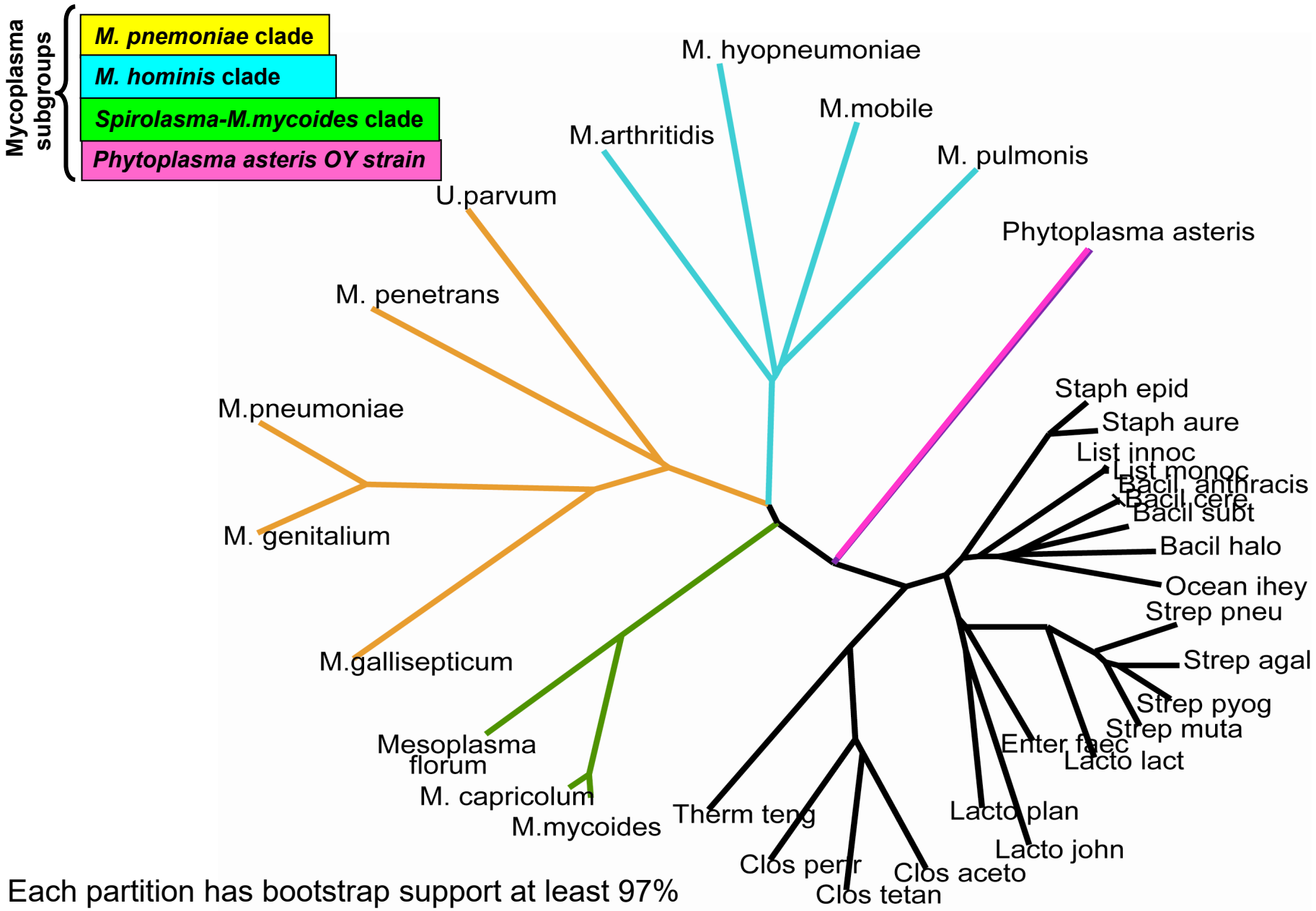


M. genitalium: Living Organism With the Smallest Genome




30 kb
60 kb
90 kb
120 kb
150 kb
180 kb
210 kb
240 kb
270 kb
300 kb
330 kb
360 kb
390 kb
420 kb
450 kb
480 kb
510 kb
540 kb
570 kb

Whole genome phylogeny of Firmicutes

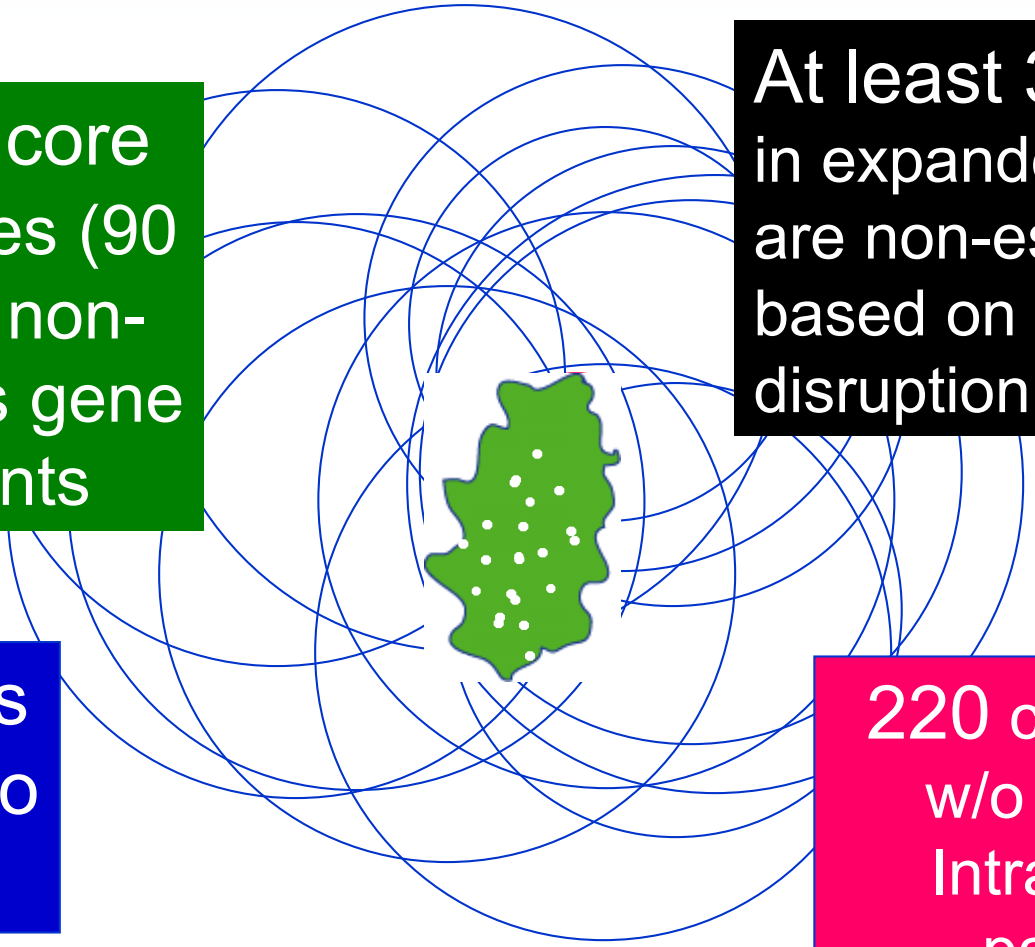


Each partition has bootstrap support at least 97%

The core mycoplasma genome is the set of genes common to all 13 complete sequences

Expanded core of 310 genes (90 because of non-orthologous gene displacements

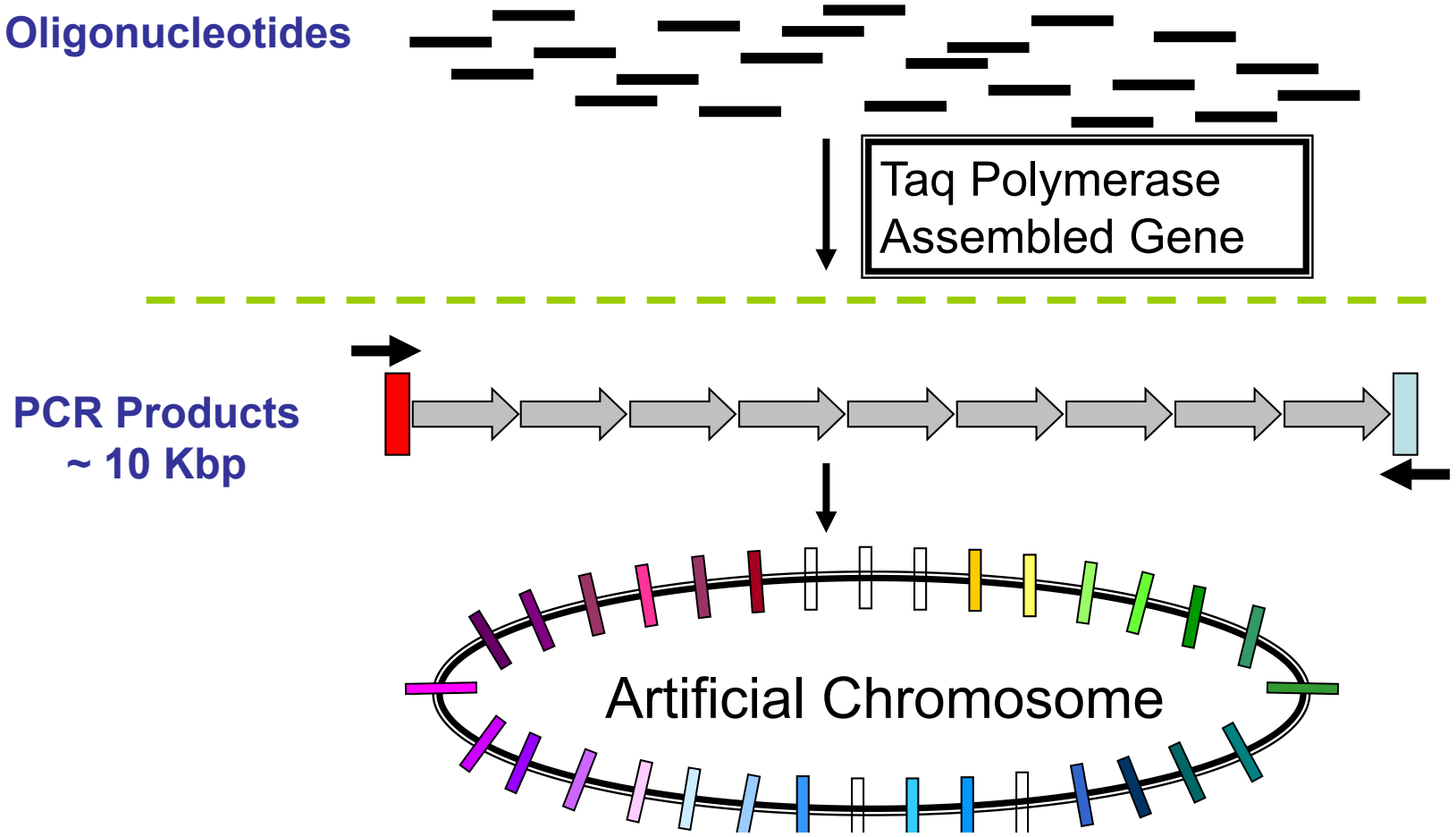
At least 36 genes in expanded core are non-essential based on gene disruption studies



173 genes common to all 13

220 core genes w/o obligate Intracellular parasite

Construction of an Artificial Chromosome



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.

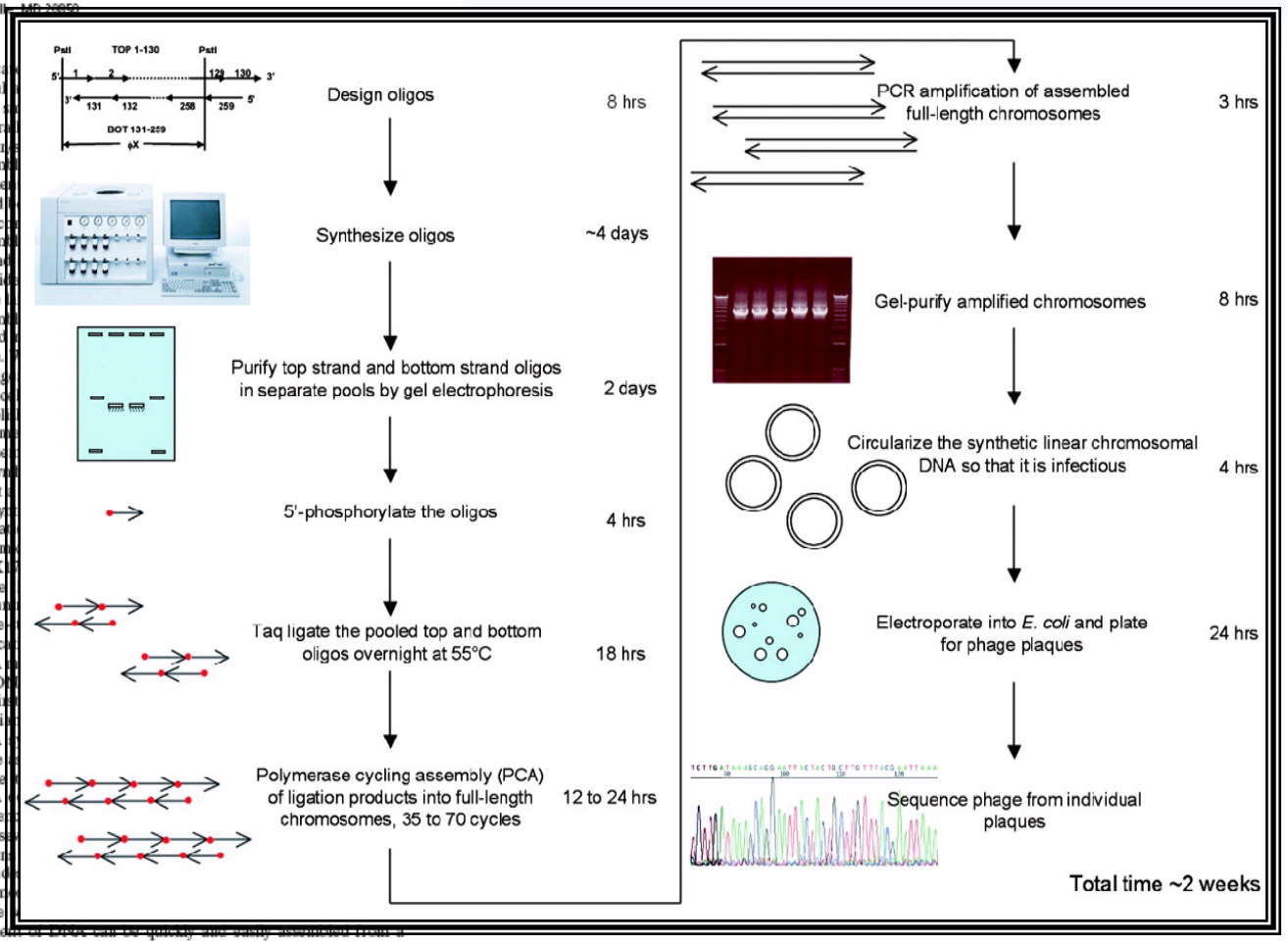
[†]Present address: Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599-7290.

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

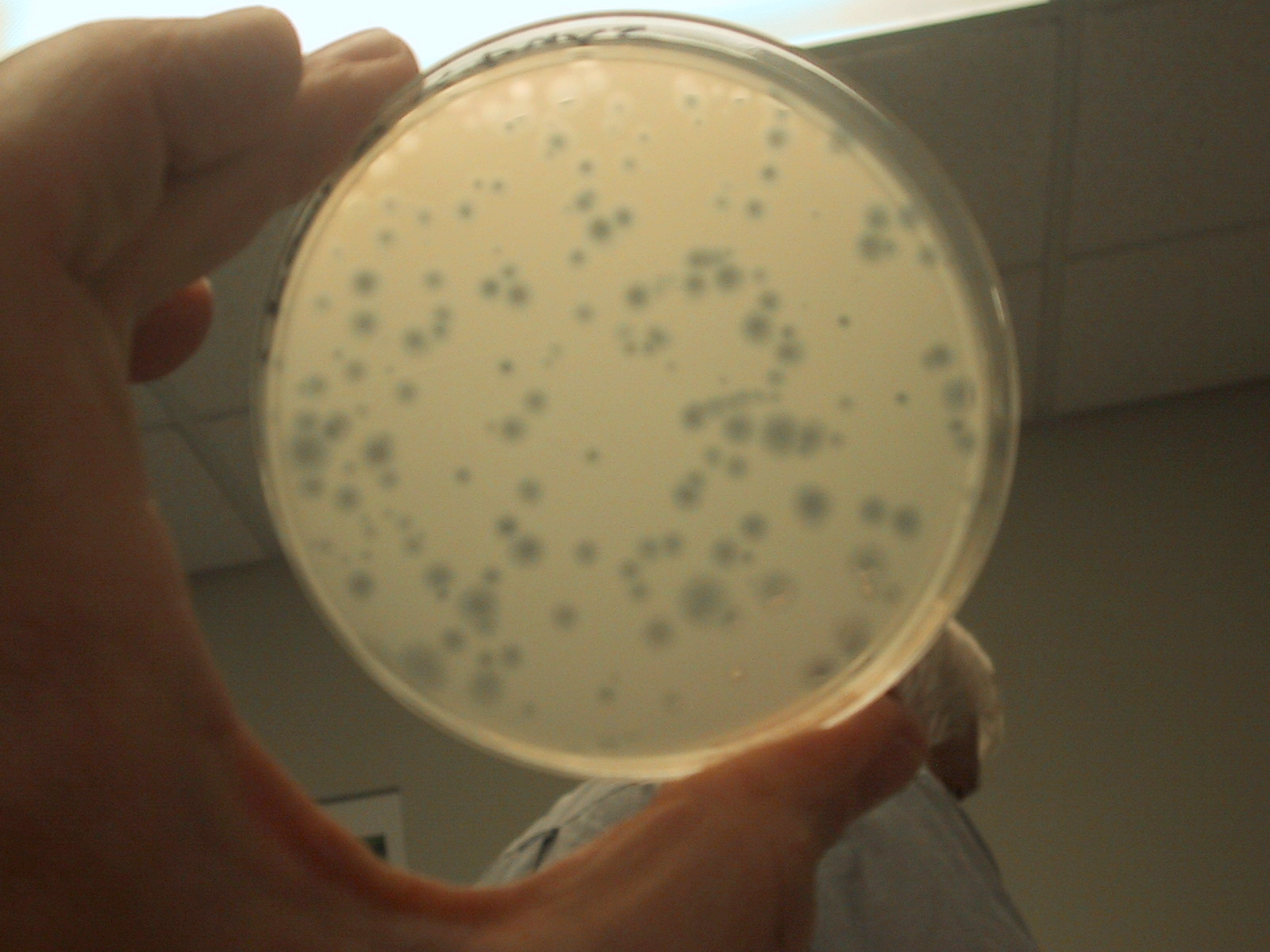
[§]Pasteur, L., Sorbonne Scientific Soiree, Apr. 7, 1864, Paris.

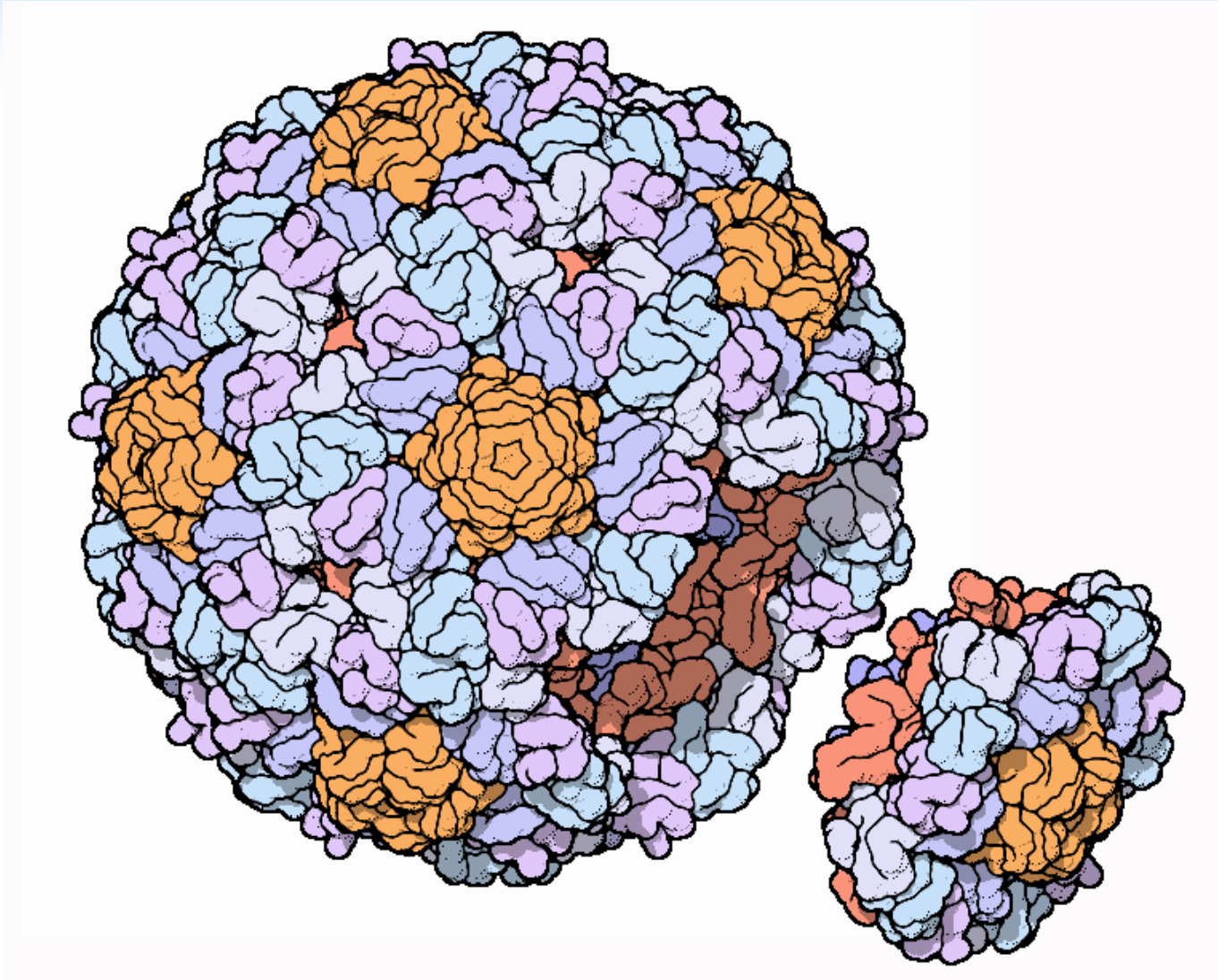
www.pnas.org/cgi/doi/10.1073/pnas.2237126100

Rapid gene synthesis from oligonucleotides



Hamilton Smith, Clyde Hutchison, Cindi Pfannkoch, Craig Venter



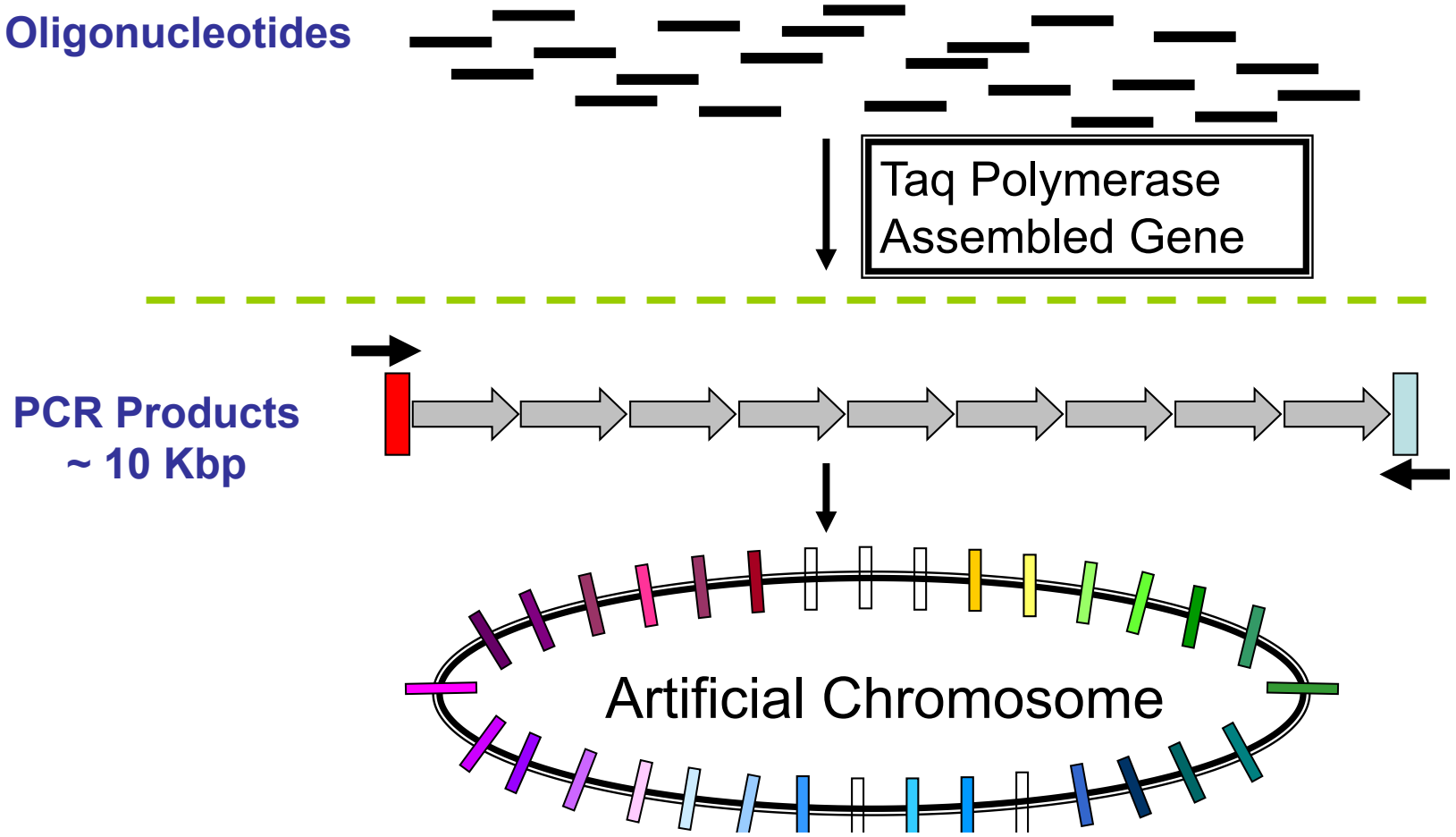


This Software Builds Its Own Hardware

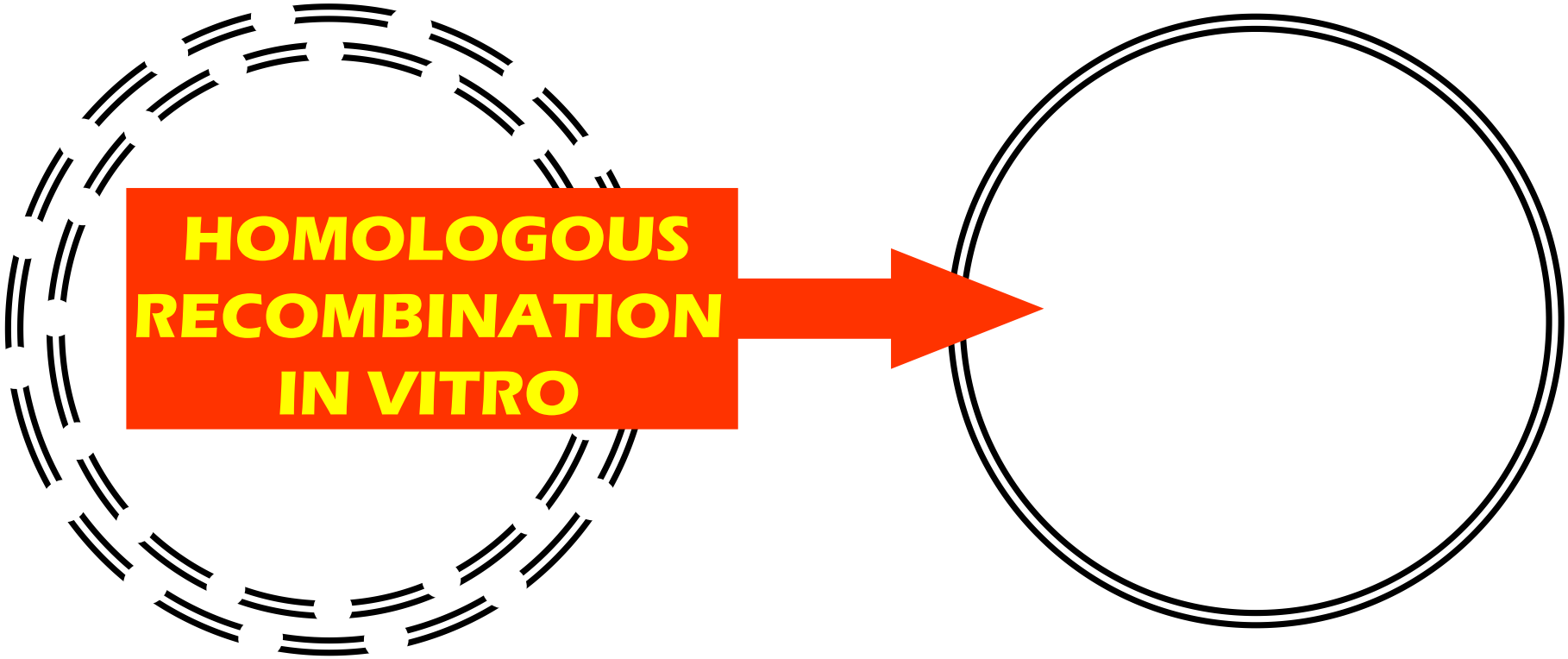
From Reading To Writing:
What can be made and when?

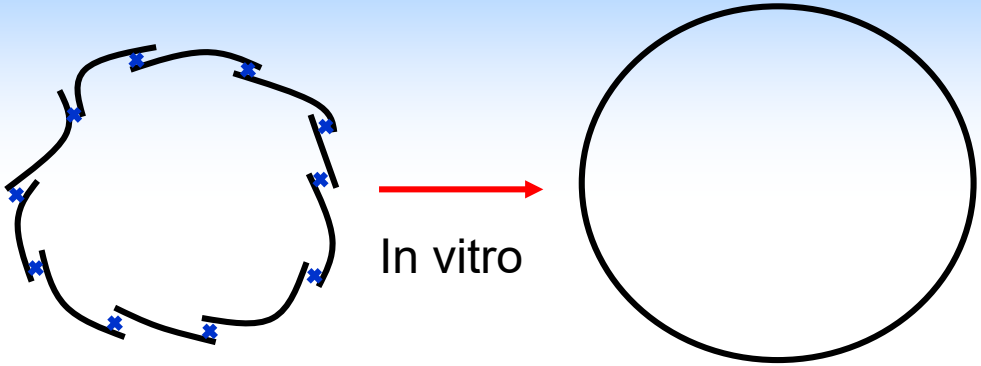
- Any sequenced viral genome including select agents can be made ***today***
- Designer viruses" are over a decade away
- Prokaryotes (bacteria): 2 years
- Single-cell eukaryotes: Within 10 years

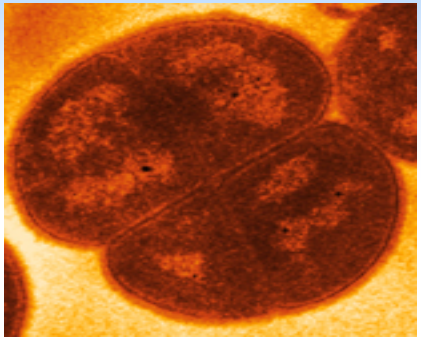
Construction of an Artificial Chromosome



How do we connect many small synthetic genome fragments into a single circular chromosome?

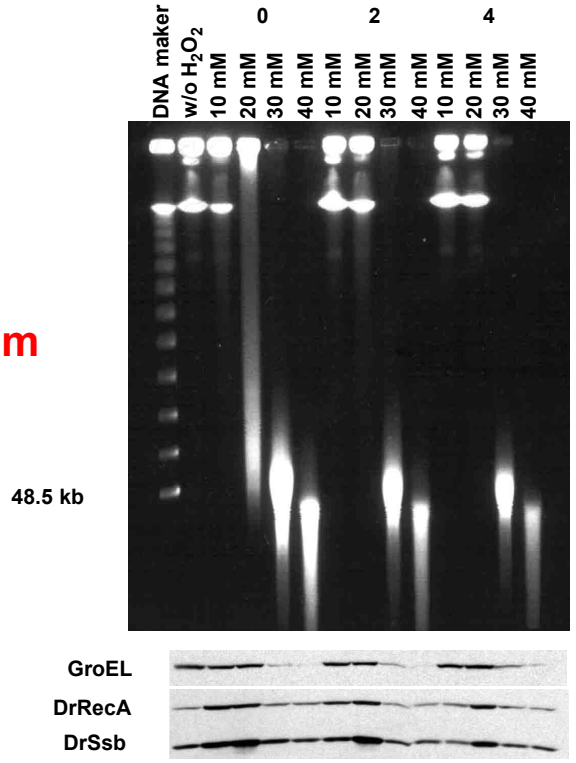




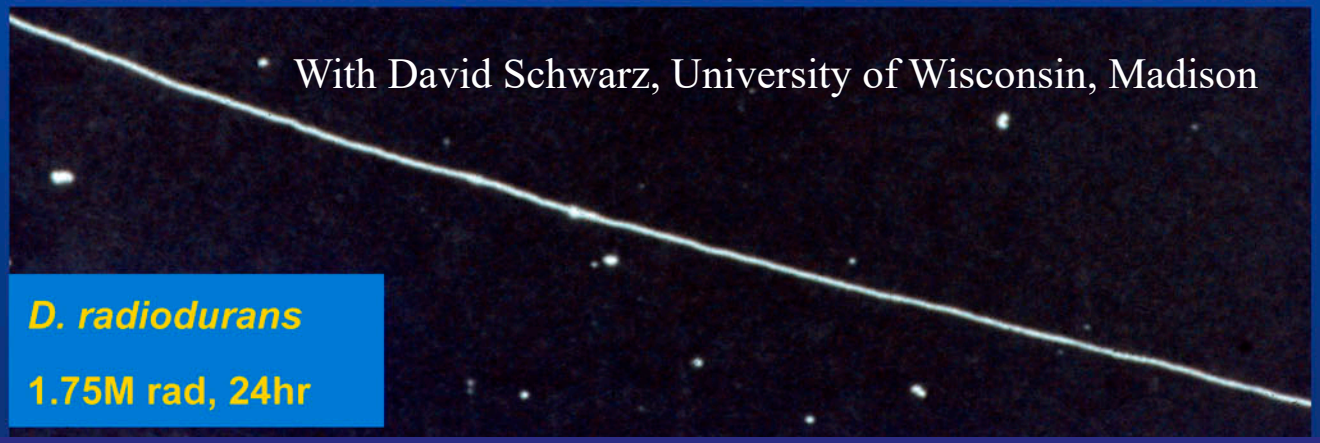


X homologous recombination

- **Develop an in vitro D. radiodurans recombination system**
- **Express and characterize recombination proteins**



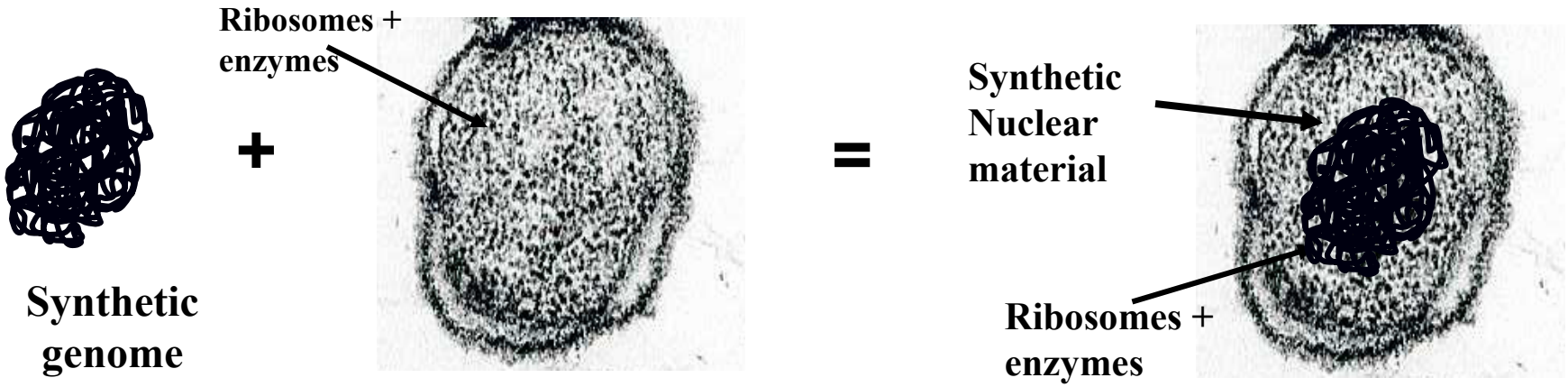
D. radiodurans: The Ultimate DNA Assembly Machine

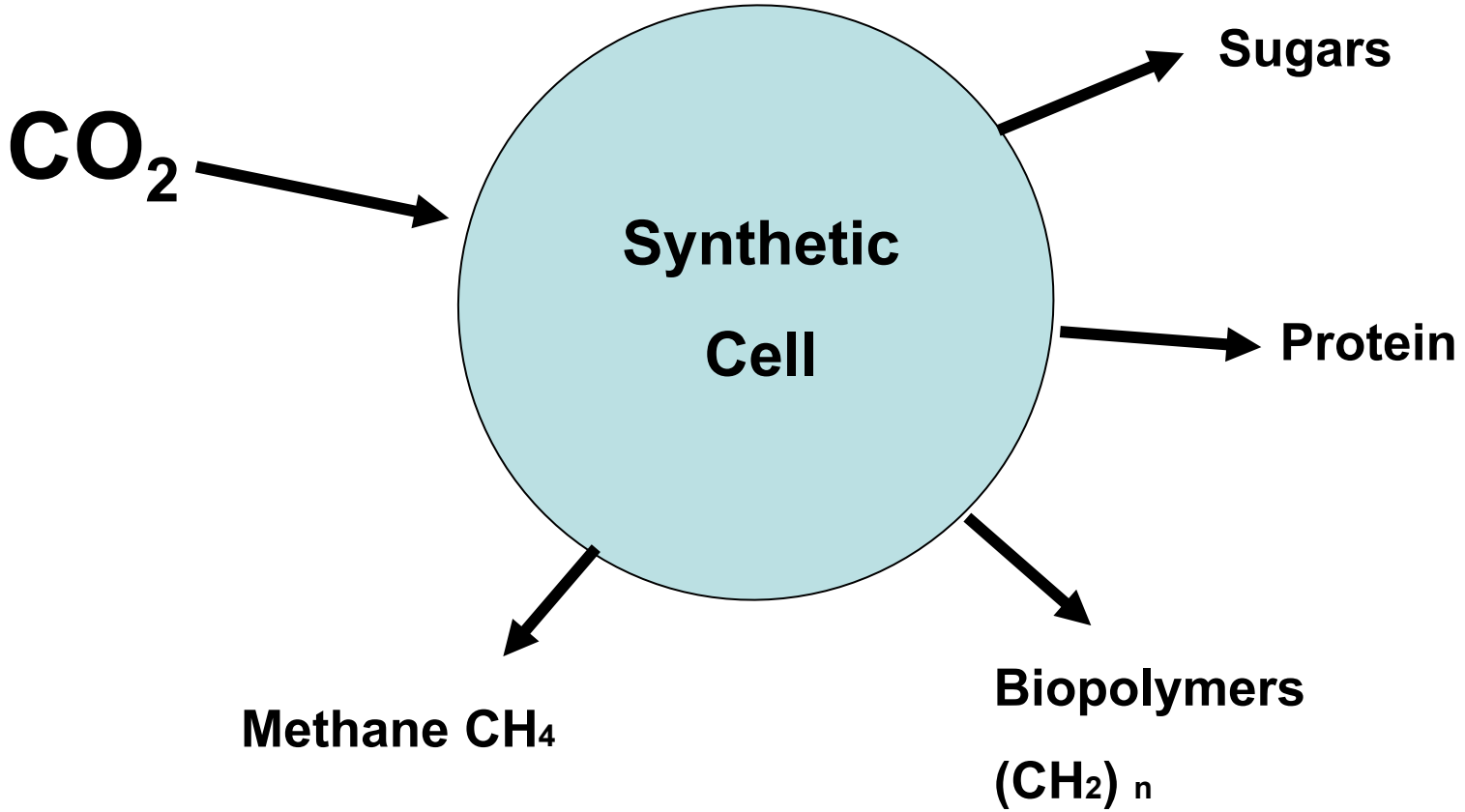


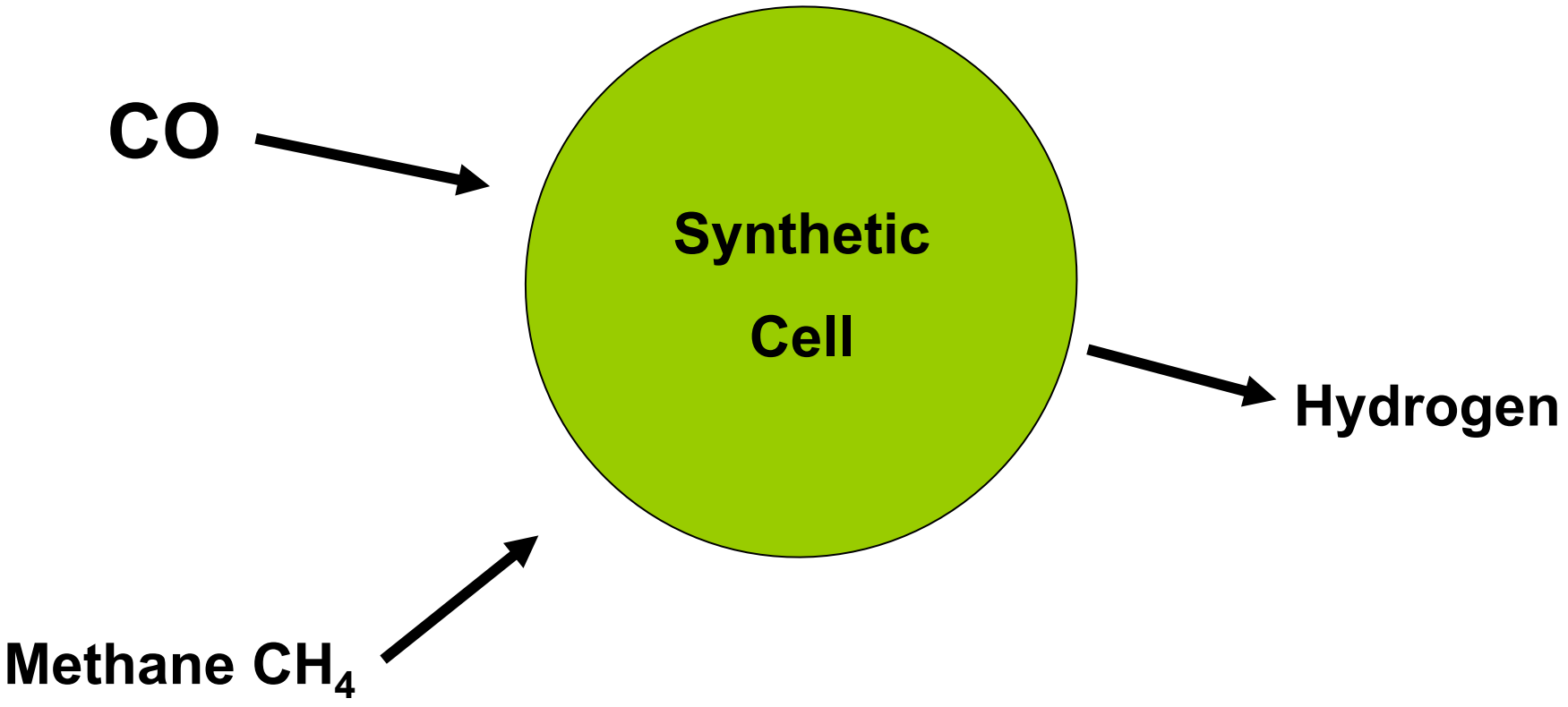
Combinatorial Genomics

- Genes can be rapidly screened via cassette based construction of thousands to millions of genomes/day
- Selection through screening for
  - Chemical production
  - Viability
  - Hydrogen production etc.

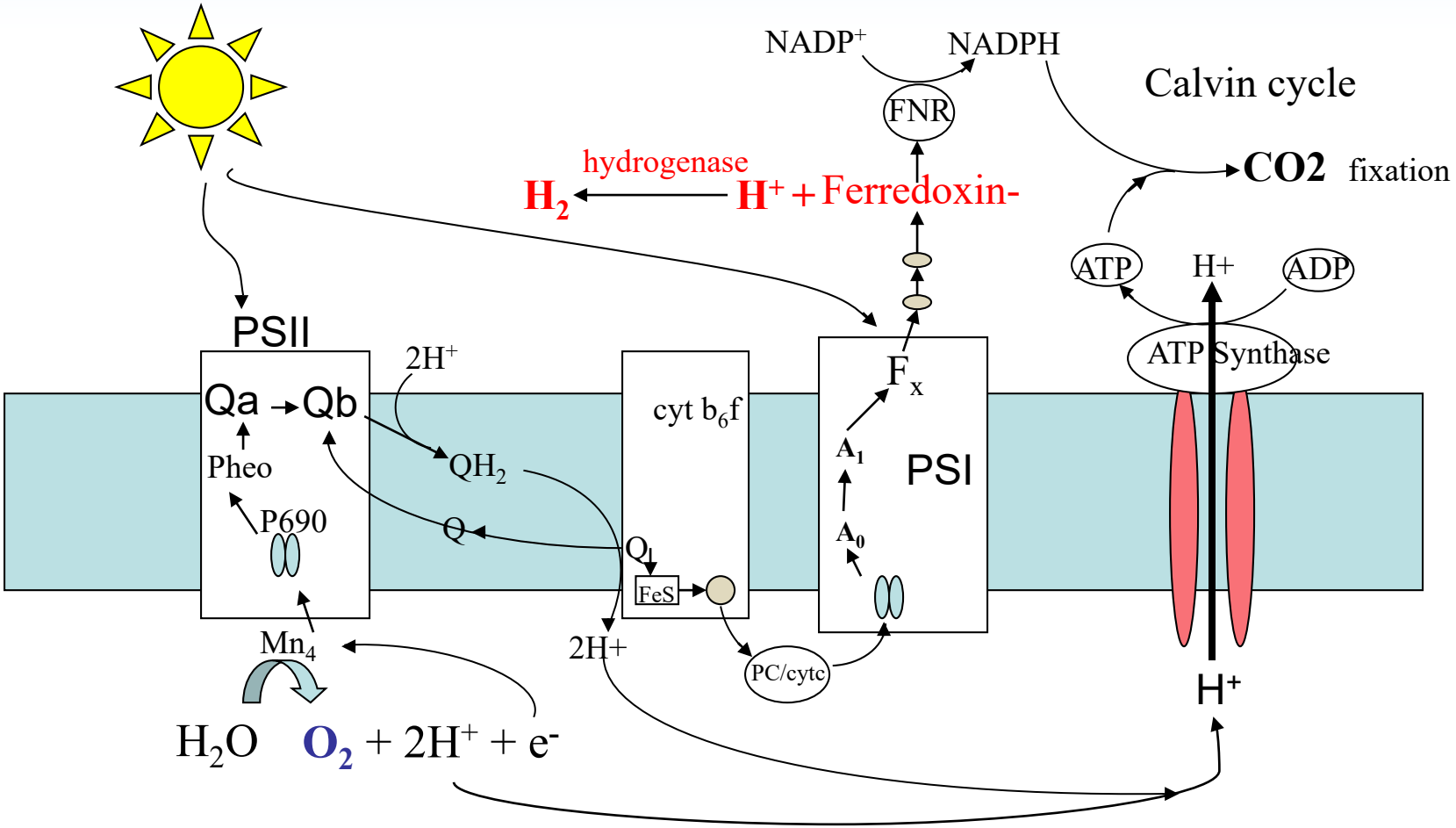
Genome transplantation




CO_2
Synthetic
Cell
Sugars
Protein
Biopolymers
$(CH_2)_n$
Methane CH_4


CO
Synthetic
Cell
Hydrogen
Methane CH_4

Photosynthetic production of hydrogen in cyanobacteria



Some Potential benefits

- Basic biology
  - Study evolution
  - Understand requirements for life
  - Confirm sequences
- Energy
  - e.g., Renewable hydrogen (a JCVI project funded by DoE Office of Science)
- **Human and animal health**
  - Vaccines (research and applications)
  - Gene therapy
  - Phage-based antibiotics
  - Drugs
- **New materials**
  - Bioplastics

Future of Engineered Species

- Designed and engineered species could replace petro-chemical industry
- Will be a source of future food
- A source of energy
- The basis of bioremediation

Regulation vs Good Scientific Standards

- **We Postponed Our research for Ethical/Policy Review in 1999**
  - **Review indicated acceptable to proceed with minimal genome efforts***
- **Good Steward Standard**
  - **Responsible lab precautions**
  - **Lab Bio Safety Level 3 for early stages**
  - **No human pathogens/No human genomes**
  - **No organism survival outside lab**
  - **Engineer out pathogenesis and self evolution**
  - **Open communication with non-science communities**
- **Tremendous Opportunity for Good**
  - **Carbon neutral energy sources**
  - **Environmental remediation**
  - **Carbon sequestration**
  - **Vaccine and other opportunities**

*** "Genetics: Ethical Considerations in Synthesizing a Minimal Genome" *Science*, 286, p. 2087**

Ethics and policy studies

- **Ethical and religious concerns surrounding work on minimal genomes (1999)**
  - ***Ethics of Genomics* group considered the implications of determining and/or synthesizing a minimal genome**
  - **Identified specific areas of concern, but determined that the research overall does not cross any bioethics boundaries**
    - **Cho et al., 1999. *Science 286*: 2087-2090.**

- **Societal implications (ongoing)**
  - ***Synthetic Genomics: Risks and Benefits for Science and Society***
  - **15-month study, funded by the Sloan Foundation**
  - **Focus on bioterrorism and environment, health, and safety**
  - **Partners**
    - **Policy Center, Venter Institute**
    - **Homeland Security Program, Center for Strategic & International Studies**
    - **Synthetic Biology Group, Massachusetts Institute of Technology**