Computing the Future of Biology & Biotechnology

Drew Endy
Stanford Department of Bioengineering
The BioBricks Foundation

9 March 2009

http://i.pbase.com/u47/hodero/large/30691408.jungle.jpg

Questions

What is this stuff?

How does it work?

Where did it come from?

What is it doing?

Where is it going?

How much DNA on Earth?

Image c/o http://www.rps.psu.edu/probing/graphics/earth2.jpg

How much DNA on Earth?

~1E10 tons bacteria (5% DNA)
~5E9 tons plants (1% DNA, made up number)
~1E9 tons animals (1% DNA, made up number)
~1E31 viruses (50% DNA)

The pangenome concept: a unifying view of genetic information. 1: Med Sci Monit. 2005 Jul;11(7):HY24-9. Epub 2005 Jun 29.

How much DNA on Earth?

4E35 bp bacterial DNA
4E34 bp plant DNA
1E34 bp animal DNA
5E30 bp virus DNA

5E35 bp DNA on the planet

How quickly can we sequence?

1E10 bp per sequencing center per month

1E12 bp per month per Earth

5E23 months to sequence all DNA on earth

Image c/o http://picasaweb.google.com/dipanjanbhattacharya/AroundMITD80#5213062475793120674

Carlson, Pace & Proliferation of Biological Technologies, Biosec. & Bioterror. 1(3):1 (2003)

Typical NIH R01 c.2090

Specific Aim #1:

Sequence all DNA on earth in first month of project

Reviewer Comments:

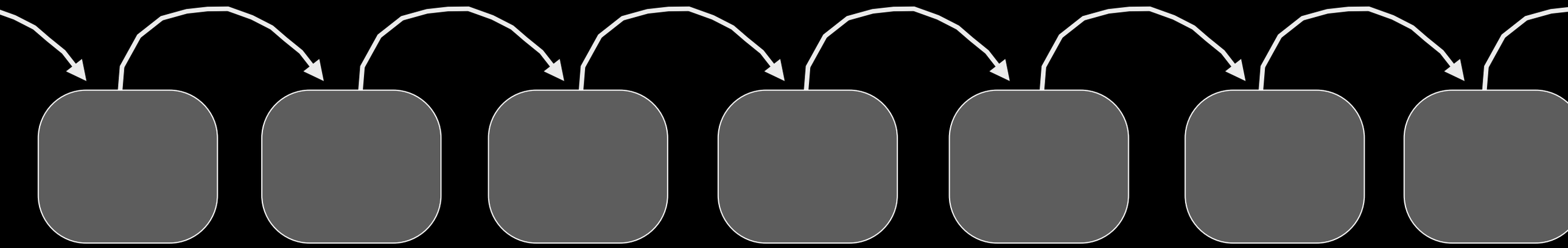
Data storage & sharing plans not adequate

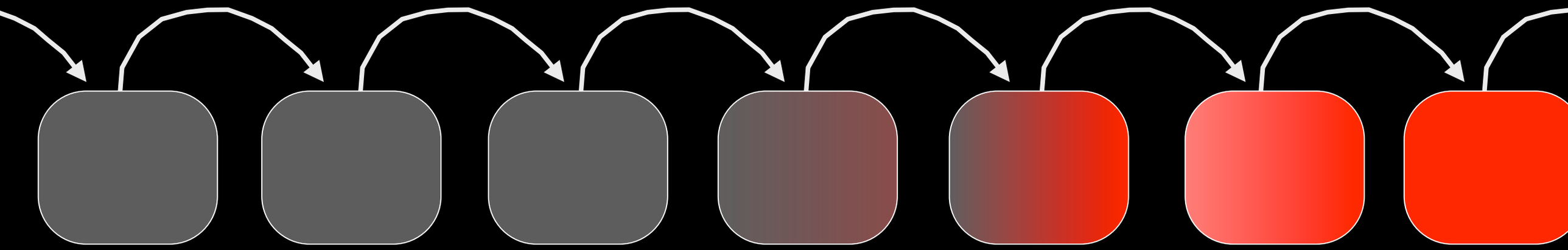
Functional annotation tools insufficient

Questions.

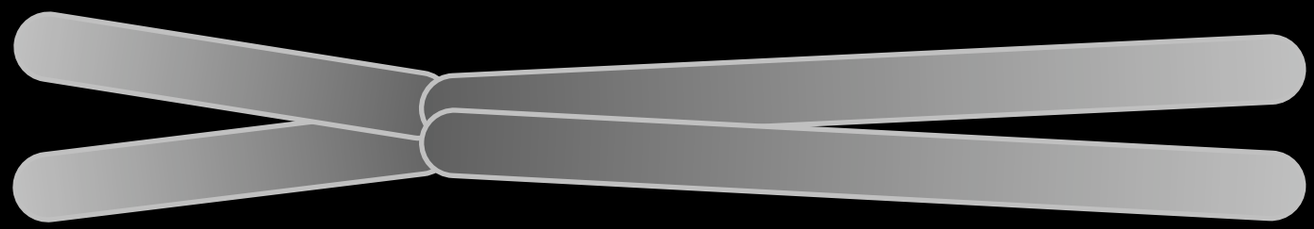
Can we finish genetics?

Could we finish early?



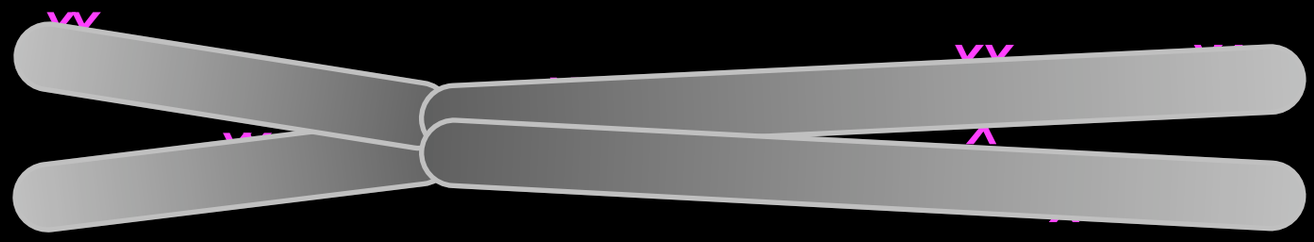


Genetics is bounded?!



1. Genomes have finite lengths.
2. Fixation rates of mutants in populations are finite.
3. Atrophy rates of functional genetic elements are greater than zero.

Pre-Sequencing Genetics

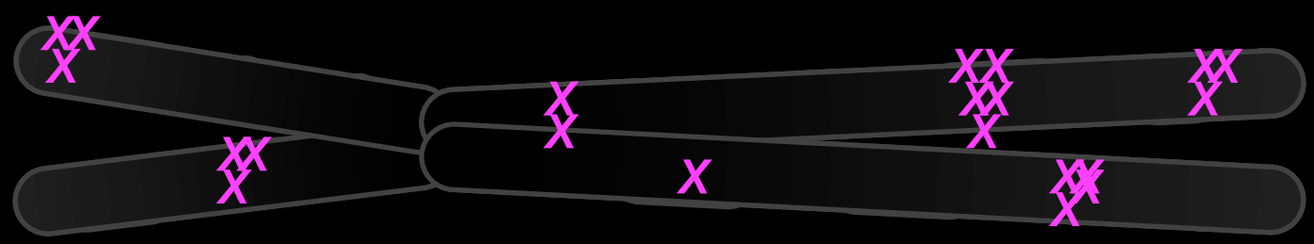


1. Collect or make libraries of mutants; characterize or screen.

2. Map mutants to loci.

Underlying Math = Logic

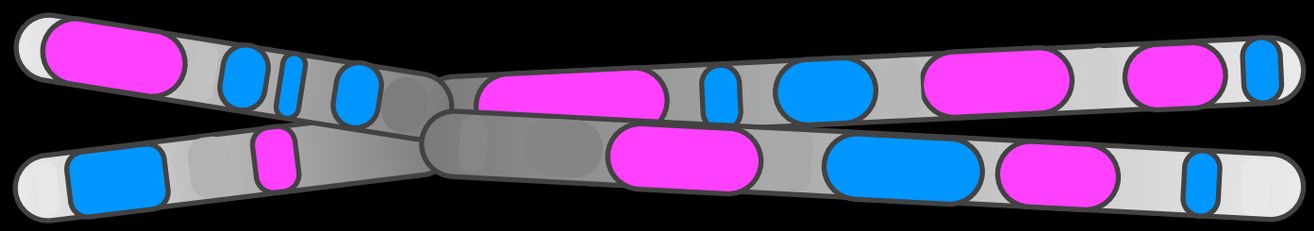
Post-Sequencing Genetics



1. Read DNA.
2. Identify putative functional genetic elements via models.

Underlying Math = Pattern Recognition

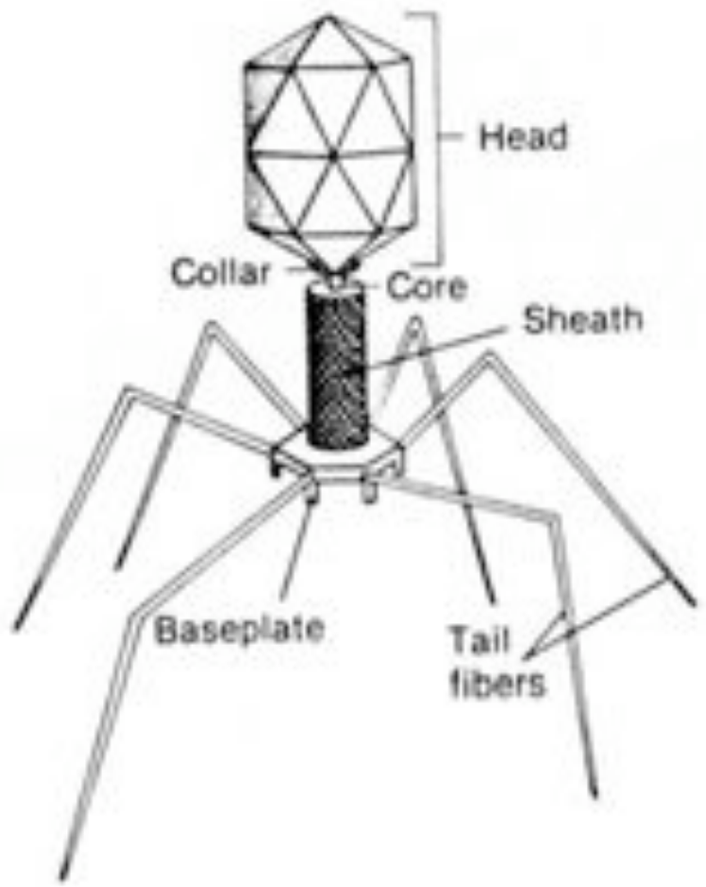
Post-Synthesis Genetics?

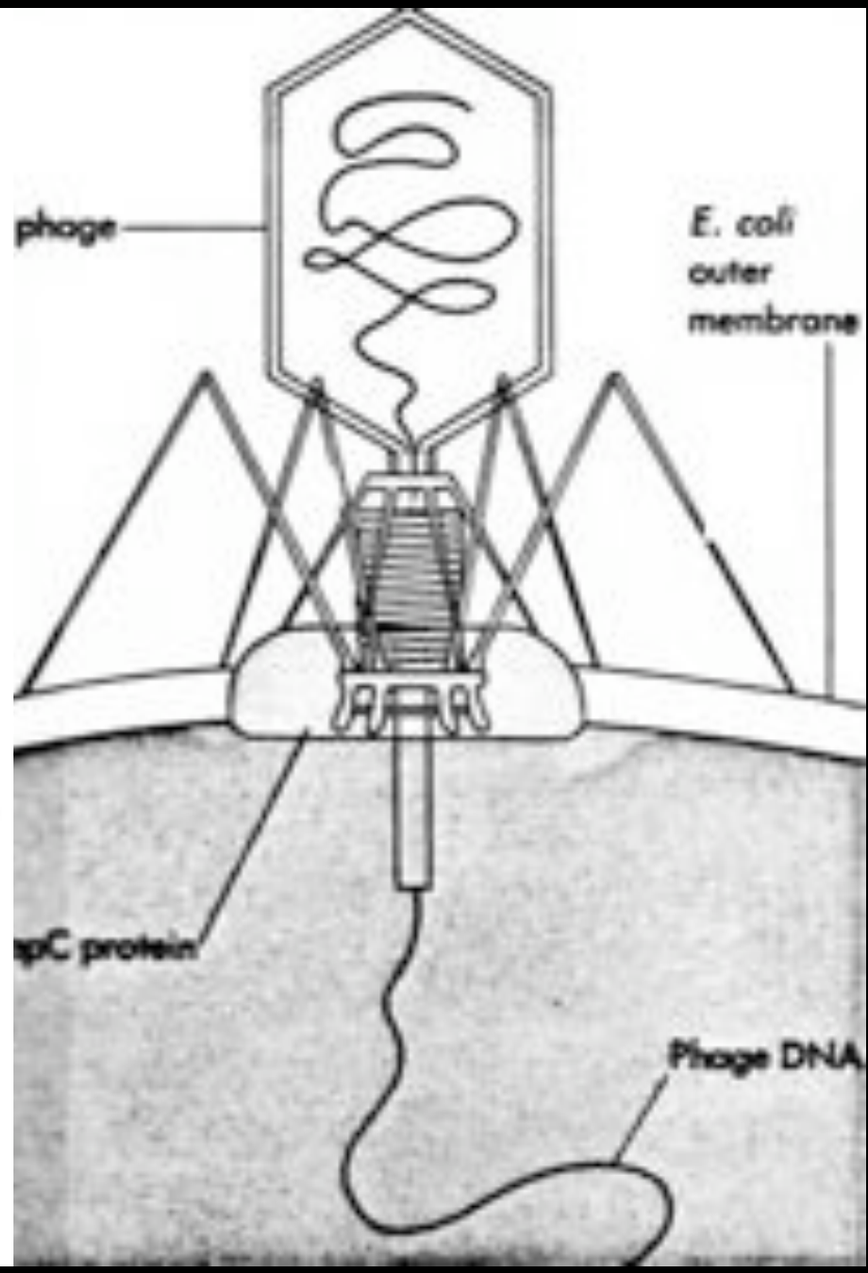


1. Design and construct any mutant or combinations thereof; characterize.
2. Reveal / confirm functions or demonstrate absence thereof.

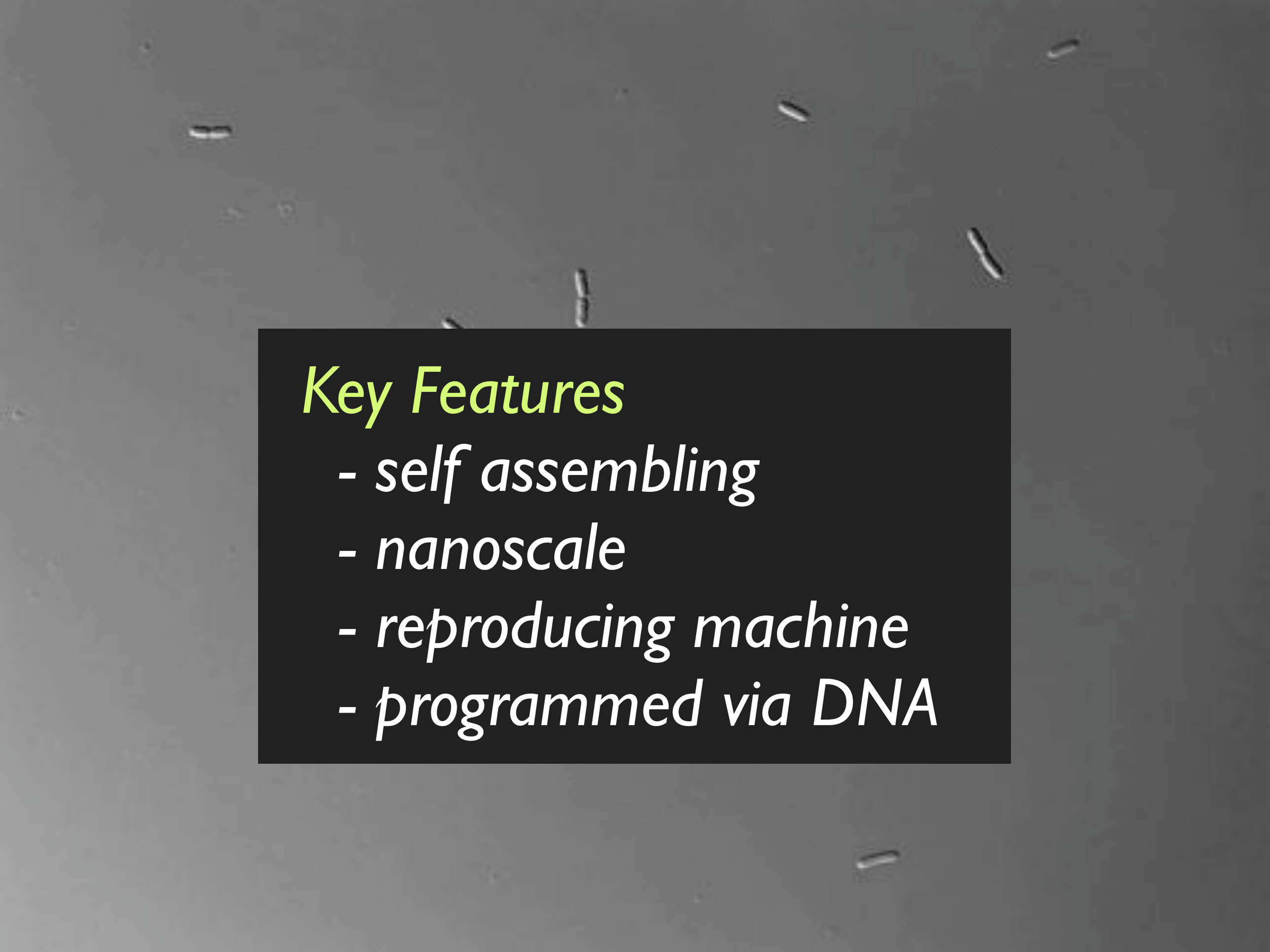
Underlying Math =? Perturbation Design

Bacteriophage





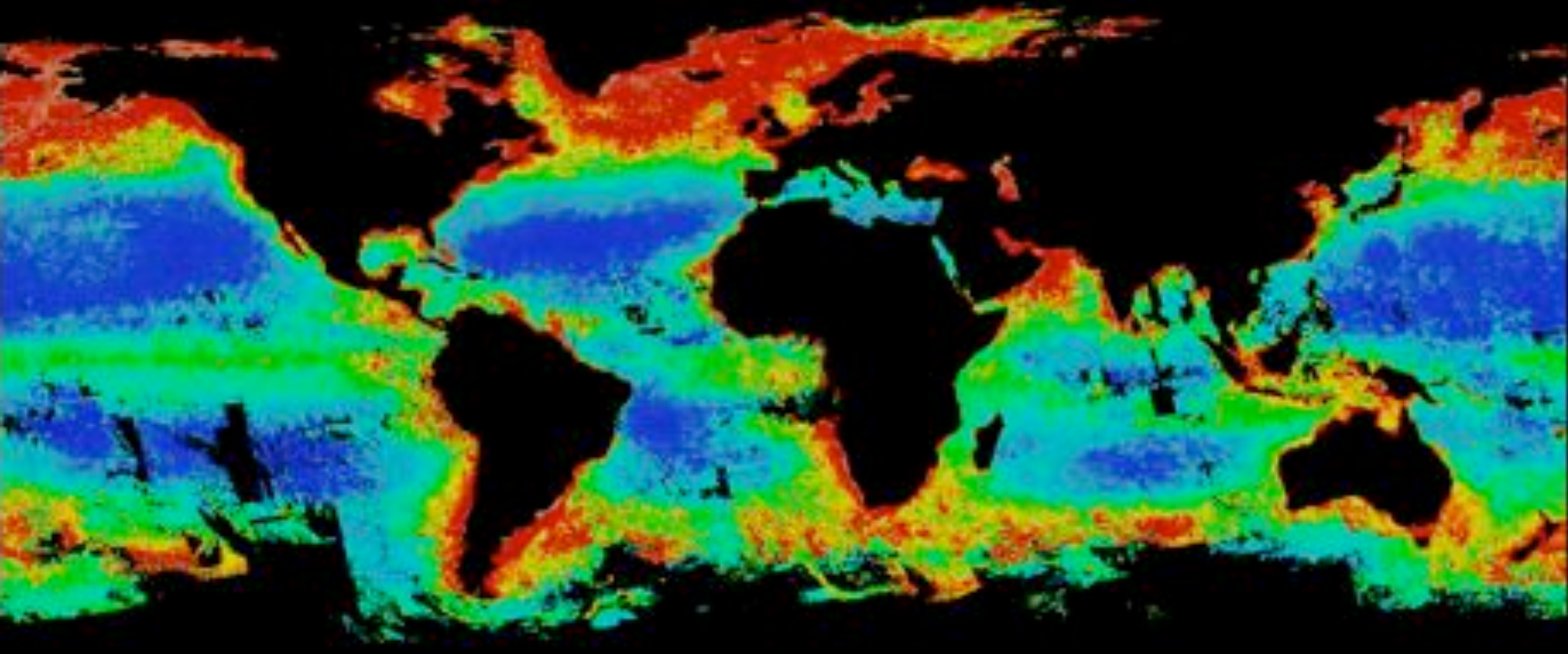
http://www.bact.wisc.edu/themicrobialworld/Phage.html



Key Features

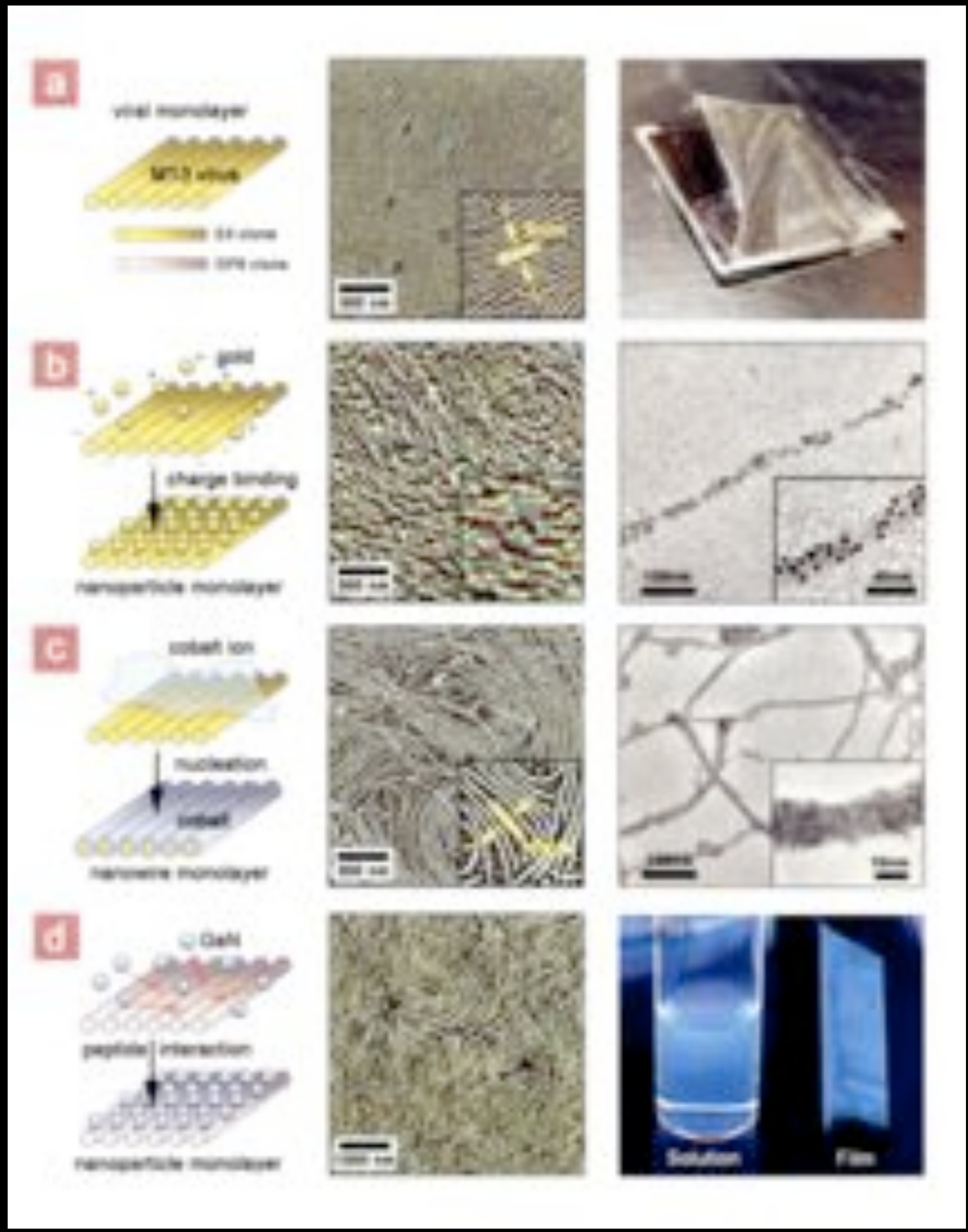
- *self assembling*
- *nanoscale*
- *reproducing machine*
- *programmed via DNA*

Ecological importance



http://svs.gsfc.nasa.gov/vis/a000000/a002400/a002497/index.html

Technical importance



http://dspace.mit.edu/handle/1721.1/39677

Phage T7: *Pre-Sequencing*

Virology
Volume 39, Issue 3, November 1969, Pages 562-574

▸ Abstract

doi:10.1016/0042-6822(69)90104-4 Cite or Link Using DOI

The genetics and physiology of bacteriophage T7[*1]

F. William Studier

Biology Department, Brookhaven

Accepted 9 July 1969. ; Available

Abstract

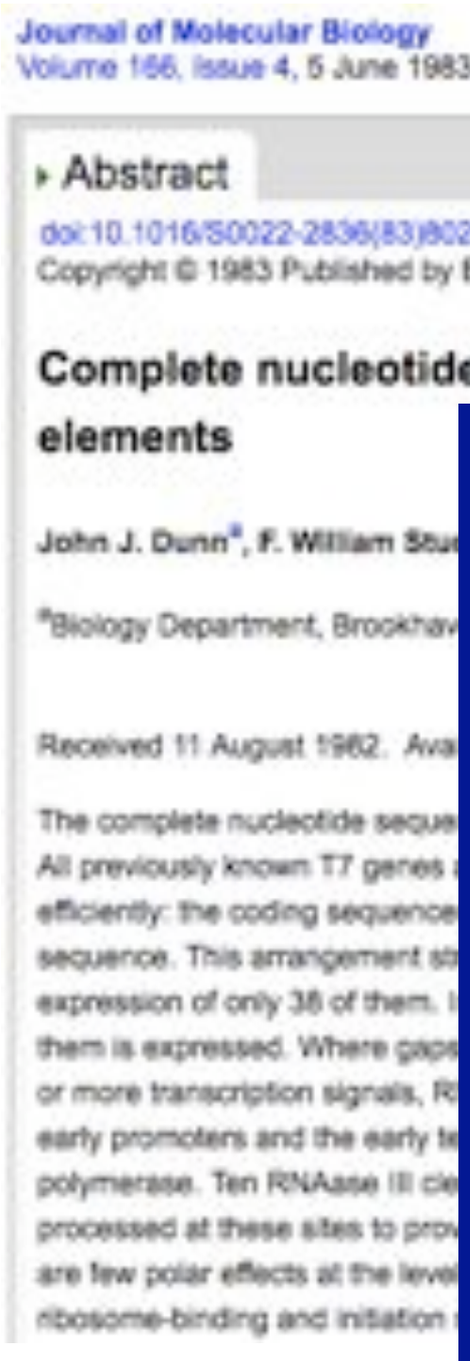
Nineteen genes have been identif
been ordered on a linear genetic
map is over 200 map units long.
another 10 or so genes remain to

The ability of mutants from each g
has been determined. Mutants fro
appear to be essentially normal in
DNA-containing particles when gr

The development of detailed gene
one in which to study a number of

"19 genes have been identified in T7 by isolation & characterization of amber mutants... ... it is estimated that not more than another 10 or so genes remain to be discovered."

Phage T7: *Post-Sequencing*



Journal of Molecular Biology
Volume 166, Issue 4, 5 June 1983, Pages 477-535

▸ Abstract

doi:10.1016/S0022-2836(83)80282-4 Cite or Link Using DOI
Copyright © 1983 Published by Elsevier Ltd.

Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements

John J. Dunn[a], F. William Stu

[a]Biology Department, Brookhav

Received 11 August 1982. Ava

The complete nucleotide seque
All previously known T7 genes
efficiently: the coding sequence
sequence. This arrangement st
expression of only 38 of them.
them is expressed. Where gaps
or more transcription signals, R
early promoters and the early te
polymerase. Ten RNAase III cle
processed at these sites to prov
are few polar effects at the leve
ribosome-binding and initiation

"...coding sequences for 50 genes..."

"...evidence for expression of only 38..."

~100 genetic elements total

PubMed Home

http://www.ncbi.nlm.nih.gov/pubmed/

MIT OWW

NCBI

PubMed
www.pubmed.gov

A service of the U.S. National Library of Medicine and the National Institutes of Health

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journa

Search Genome for bacteriophage Go Clear

Limits | Preview/Index | History | Clipboard | Details

About Entrez
Text Version

Entrez PubMed
Overview
Help | FAQ
Tutorials
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI

Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

NOTE: Performing your original search, ***pubmed***, in PubMed will retrieve 7802 citations.

To get started with PubMed, enter one or more search terms.

Search terms may be topics, authors or journals.

MyNCBI

Set up an automated PubMed update in fewer than five minutes.

1. Create a My NCBI account.
2. Save your search.
3. Your PubMed updates can be e-mailed directly to you.

Read the My NCBI Help material to explore other options, such as automated updates of other databases, setting search filters, and highlighting search terms.

PubMed is a service of the U.S. National Library of Medicine that includes over 17 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources.

NCBI Sequence Viewer v2.0

http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=NC_001604

MIT OWW

NCBI Entrez Nucleotide

PubMed Nucleotide Protein Genome Structure PMC

Search Nucleotide for Go Clear

Limits Preview/Index History Clipboard

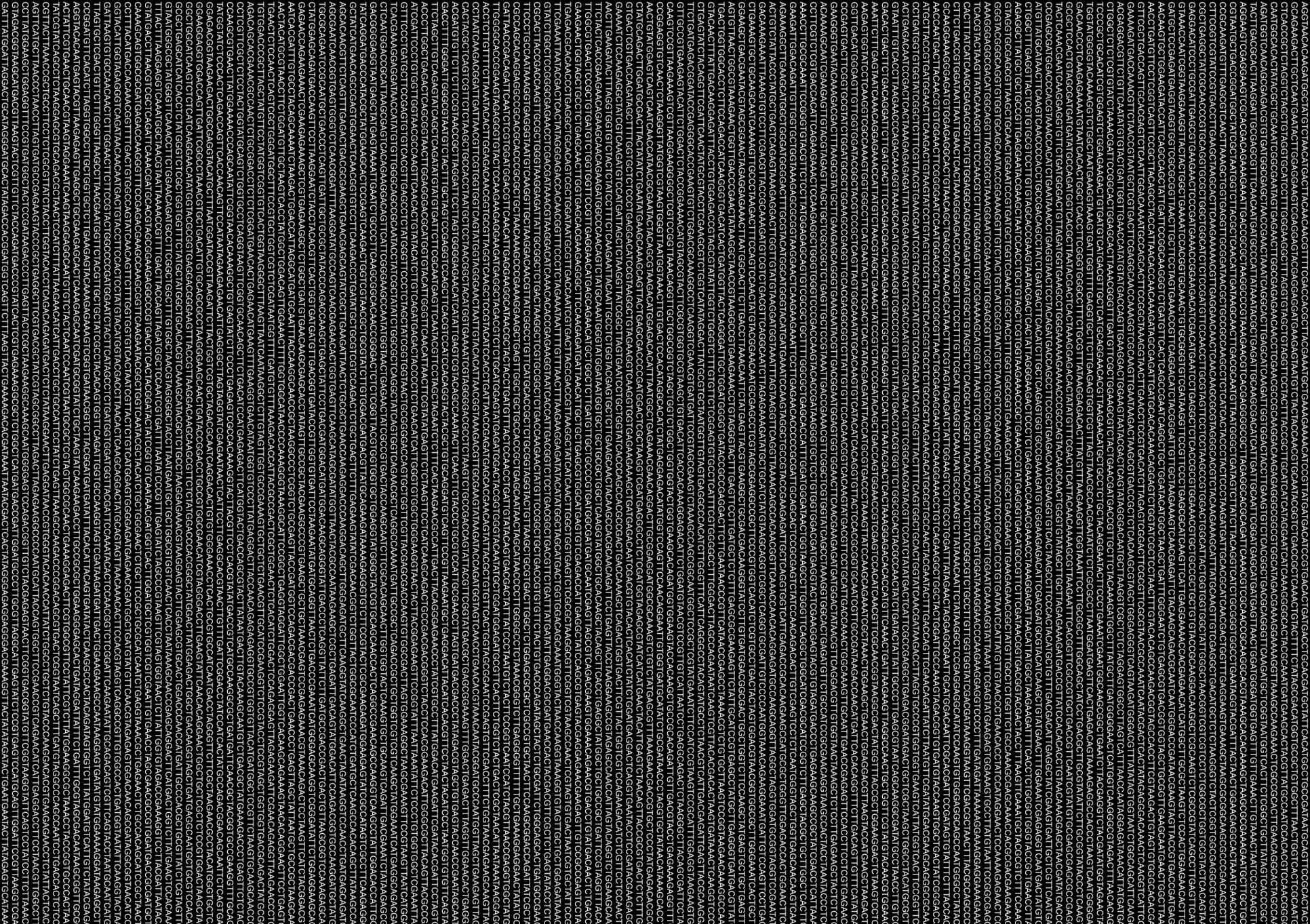
Display GenBank Show 5 Send to Hide: sequence all but gene, CDS and mRNA features

Range: from begin to end Reverse complemented strand Features: + Refresh

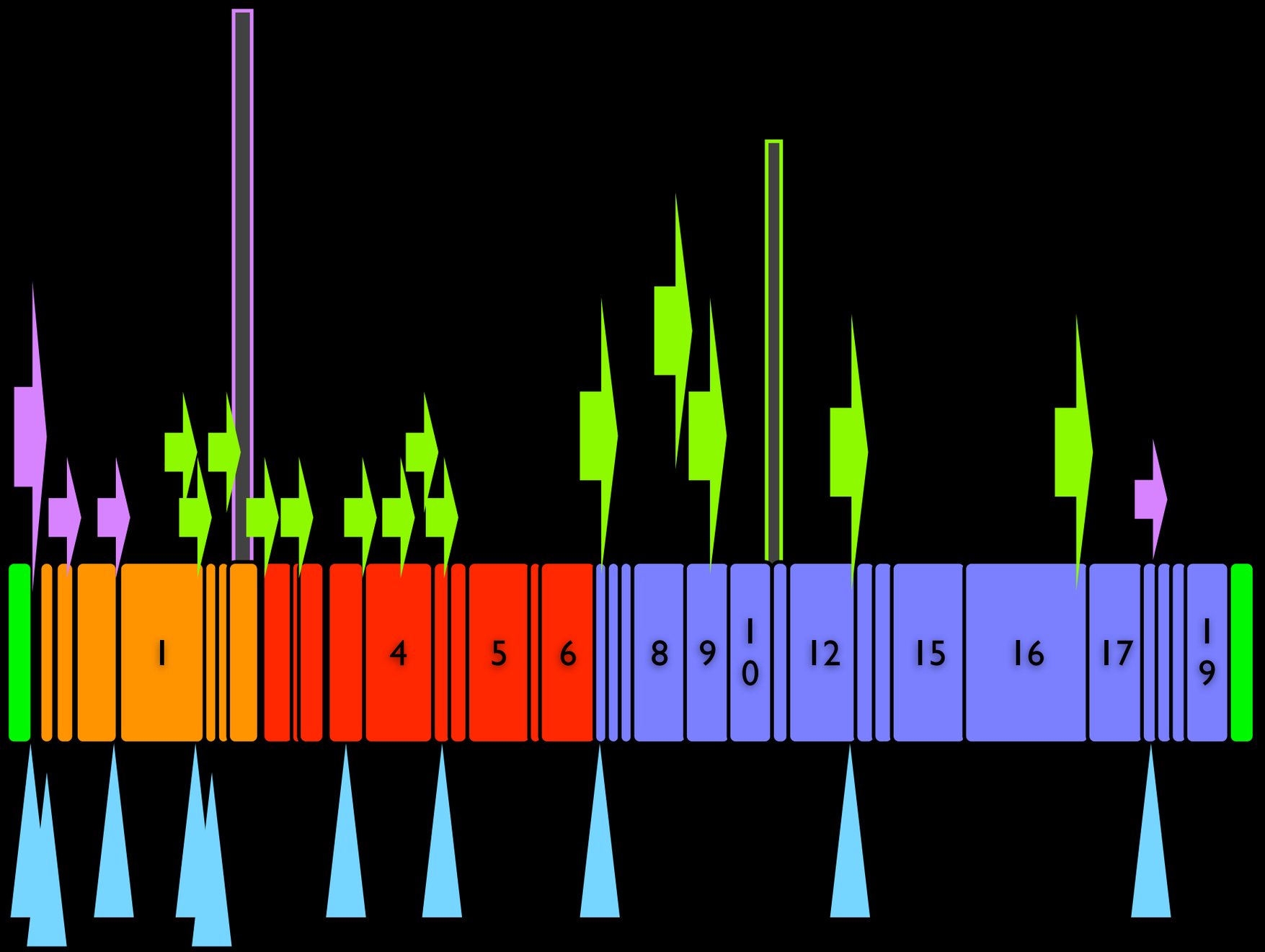
1: NC_001604. Reports Enterobacteria ph...[gi:9627425]

Comment Features Sequence

```
LOCUS       NC_001604               39937 bp    DNA     linear   PHG 04-MAY-2006
DEFINITION  Enterobacteria phage T7, complete genome.
ACCESSION   NC_001604
VERSION     NC_001604.1  GI:9627425
PROJECT     GenomeProject:14460
KEYWORDS    capsid; dCTPase; DNA ligase; DNA polymerase; E. coli; E. coli RNA
            polymerase; EcoB; EcoK; endonuclease; exonuclease; F plasmid; H-NS
            protein; head-tail connector; helicase; internal virion; lysozyme;
            primase; promoter; protein; protein kinase; recBCD nuclease; RNase
            III site; scaffolding protein; single-stranded DNA binding protein;
            T7 RNA polymerase; tail; tail fiber; terminase; transcription
            terminator.
SOURCE      Enterobacteria phage T7
  ORGANISM  Enterobacteria phage T7
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
            T7-like viruses.
REFERENCE   1  (bases 11061 to 11062)
  AUTHORS   Cheng,X., Zhang,X., Pflugrath,J.W. and Studier,F.W.
  TITLE     The structure of bacteriophage T7 lysozyme, a zinc amidase and an
            inhibitor of T7 RNA polymerase
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 91 (9), 4034-4038 (1994)
   PUBMED   8171031
REFERENCE   2  (bases 1 to 39937)
  AUTHORS   Liu,Q. and Richardson,C.C.
```



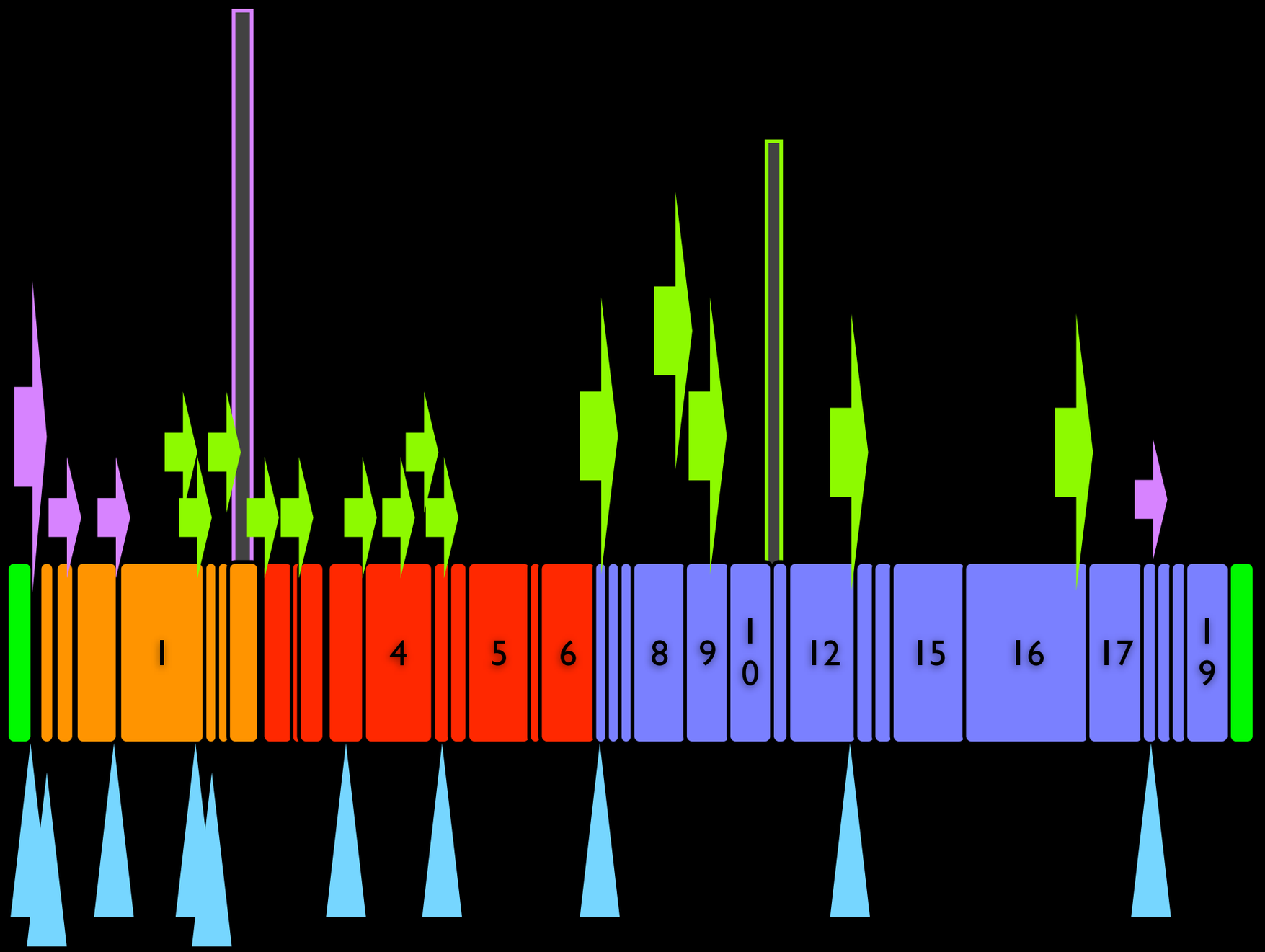
Functional depiction of T7 genome

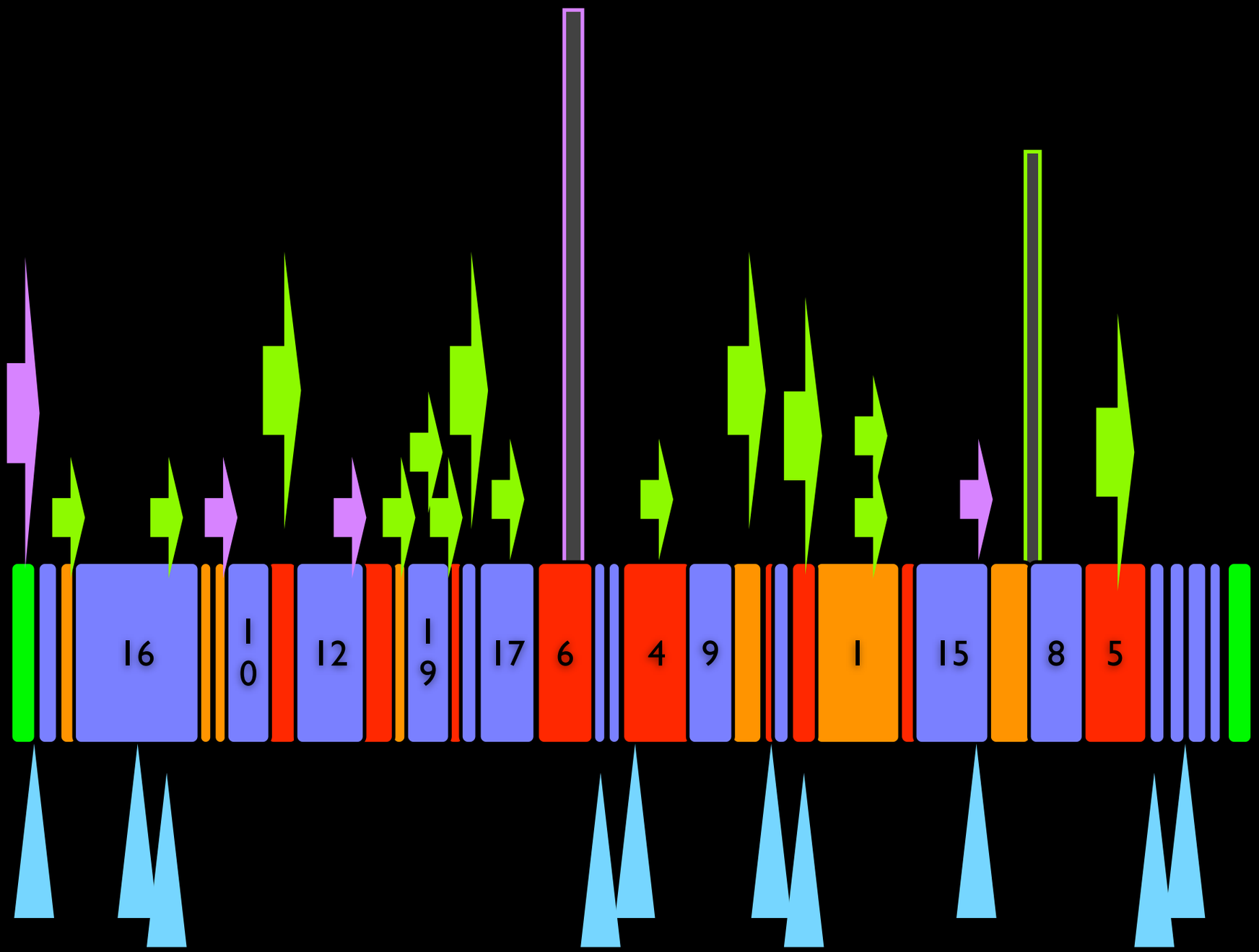


	Gene[a]	Amino[a] acids	M_r[a]	Function[a]
Class I	*0.3*	116	13,676	Inactivates host restriction
	0.4	59	6621	
	0.5	47	4744	
	0.6.1	53	6293	
	0.6.0	111	(13,204)	
	0.7	309	41,124	Protein kinase
	1	883	99,062	T7 RNA polymerase
	2.1	42	5190	
	2.0	84	10,009	Replication
	2.5	359	41,133	DNA ligase
Class II	*3.4*	31	3446	
	3.5	29	3174	
	3.6	90	9946	
	3.7	186	22,063	
	3.8	68	8783	
	3'	63	7043	Inactivates host RNA polymerase
	3.5	233	25,502	Single-stranded DNA-binding protein
	3.9	139	15,817	
	3	148	17,049	Endonuclease
	3.5	139	16,406	Amidase (lysozyme)
	3.4	125	14,329	
	4.1	566	62,606	Primase
	4.5	393	50,743	Primase
	(4.7)	39	4265	
	(4.0)	112	12,653	
	4.3	70	7927	
	4.5	86	9690	
	4.7	135	15,206	
	5	794	79,695	DNA polymerase
	5.2	118	13,062	
	5.4	98	11,074	
	5.7	68	7280	Permits growth on λ lysogens
	6	347	39,995	Exonuclease
	6.2	37	4068	
	6.5	84	9474	
	6.7	87	9597	
	7	132	15,300	Host range
	7.3	86	9697	Host range
	7.7	130	14,737	
	8	523	58,999	Head-tail protein
	9	306	33,796	Head assembly protein
	10.4	344	36,414	Major head protein
	10.8	397	(41,890)	Minor head protein
	11	196	22,289	Tail protein
	12	783	89,265	Tail protein
	13	139	15,852	Internal virion protein
	14	186	20,836	Internal virion protein
	15	746	84,259	Internal virion protein
	16	1216	143,943	Internal virion protein
	17	532	61,441	Tail fiber protein
	17.5	67	7594	
	18	99	10,145	DNA maturation
	(18.7)	82	9195	
	19	585	66,130	DNA maturation
	(19.2)	84	9264	
	(19.3)	56	6429	
	19.5	49	5434	

Dunn & Studier (1983)
Journal of Molecular Biology
v166 p477

Functional depiction of T7 genome





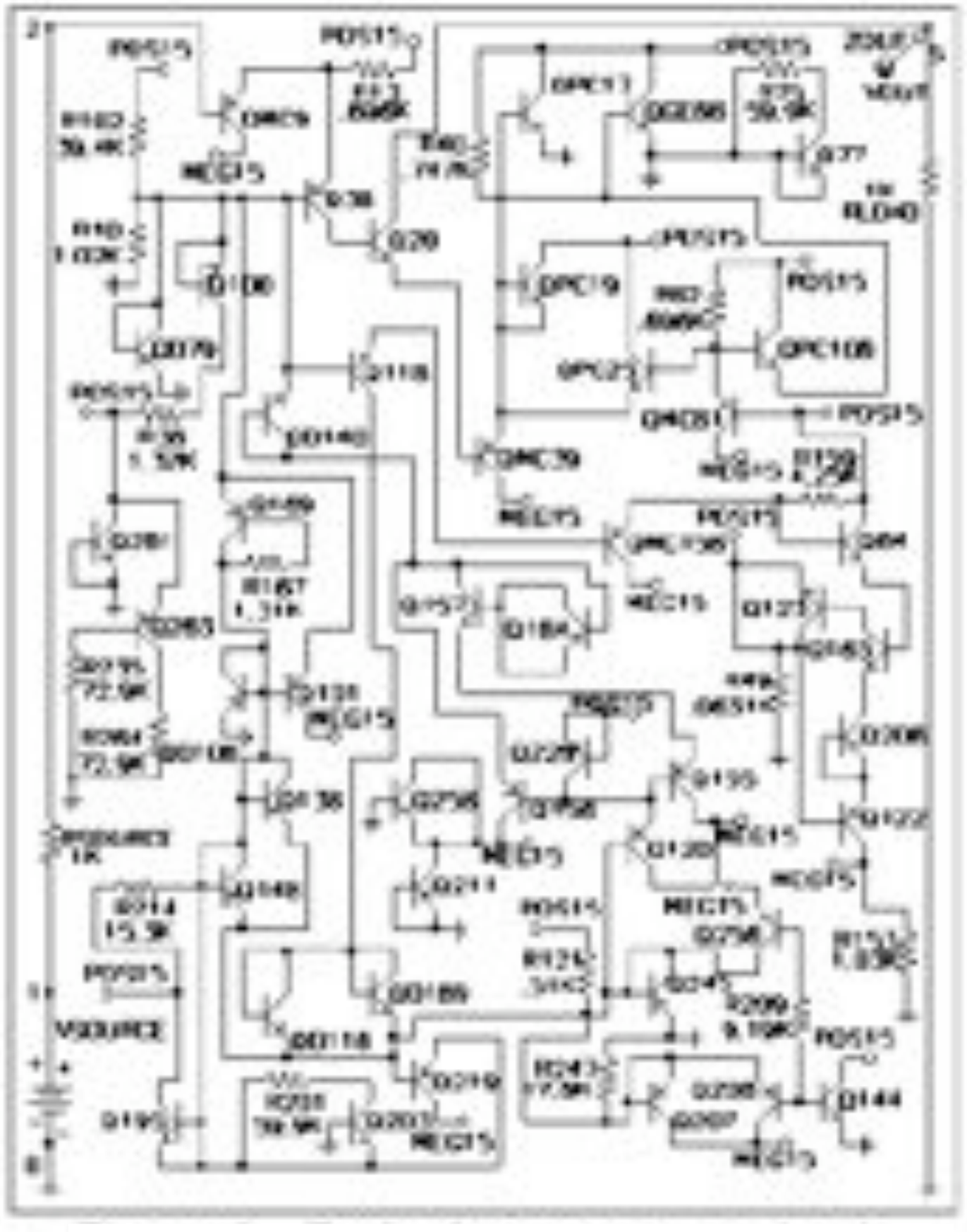
I have no idea what this genetic program will do

Natural DNA sequence

```
----------------2.8---------------->

...acgcaaagggaggcgacatggcaggttacggcgctaaaggaatccgaaa...

<--3-RBS---><----------------3----------------
```



Evolved Circuit:

1. What does it do?
2. How does it work?
3. Why this design?

J.R. Koza et al.
Automated Synthesis of Computational Circuits using GeneticProgramming,
1997 IEEE International Conference on Evolutionary Computation

Design & build new genome?

```
----------------2.8----------------->

acgcaaagggaggcgacatggcaggttacggcgctaaaggaatccgaaa

      <--3-RBS---><----------------3--------------
```

```
acgcaaGgggagAcgacaCggcaggttacggcgctaaggatccggccgcaaagggaggcgacatggcaggttacggcgctaaa

--------------2.8------------------><D28R|D29L><--3RBS------><--------------3----
```

Phage T7: *Post-Synthesis*

Molecular Systems Biology (2005) doi:10.1038/msb4100035

www.molecularsystemsbiology.com

molecular systems biology

Refactoring bacteriophage T7

Leon Y Chan[1,3], Sriram Kosuri[2,3] and Drew Endy[2,*]

[1] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA and [2] Division of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[3] These authors contributed equally to this work
* Corresponding author. Division of Biological Engineering, Massachusetts Institute of Technology, 68-580, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Tel.: +1 617 258 5152; Fax: +1 617 253 5865; E-mail: endy@mit.edu

Received 15.7.05; accepted 22.7.05

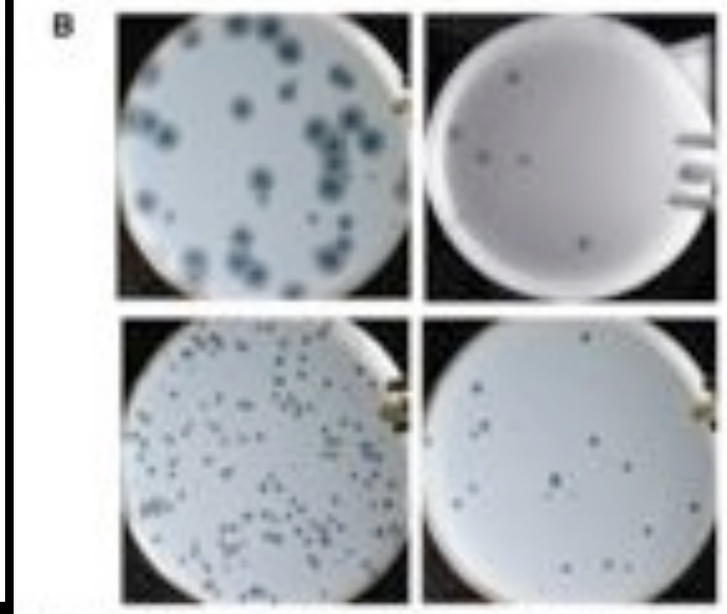


Figure 4 Characterization of T7.1 (**A**) Lysis of log-phase liquid cultures of *E. coli* BL21 (30°C) by wild-type T7 (black), alpha-WT chimera (red), WT-beta-WT chimera (blue), alpha-beta-WT chimera (orange); absorbance of 0.275 is ~ 20.8 cells/mL. Vertical bars show standard deviation at each time point (based on four replicates) (Supplementary information). (**B**) T7 plaques on *E. coli* BL21 (DE3, 37°C, 10 cm Petri dish). Clockwise from top left: wild-type (WT) T7, alpha-WT chimera, WT-beta-WT chimera, alpha-beta-WT chimera (Supplementary information).

```
----------------2.8----------------->

acgcaaagggaggcgacatggcaggttacggcgctaaaggaatccgaaa

      <--3-RBS---><----------------3-------------

acgcaaGgggagAcgacaCggcaggttacggcgctaaggatccggccgcaaagggaggcgacatggcaggttacggcgctaaa

---------------2.8------------------><D28R|D29L><--3RBS------><---------------3----
```

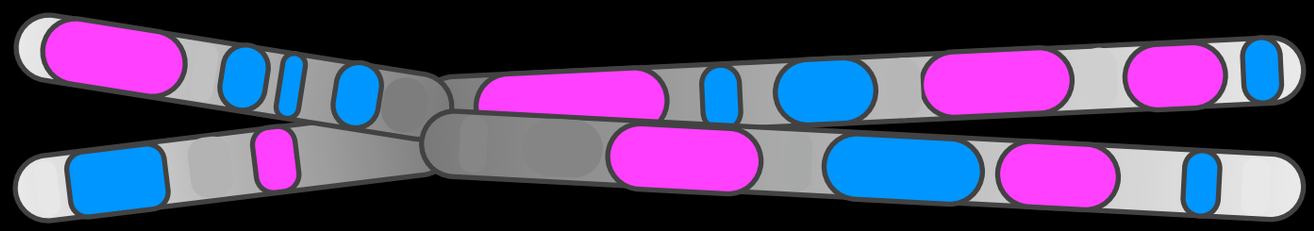
Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome

Note added in proof: While this paper was in press, we realized that the TARBAC vector in our sMgTARBAC37 clone interrupts the gene for the RNA subunit of RNase P (rnpB). This confirms our speculation that the vector might not be at a suitable site for subsequent transplantation experiments.

References and Notes

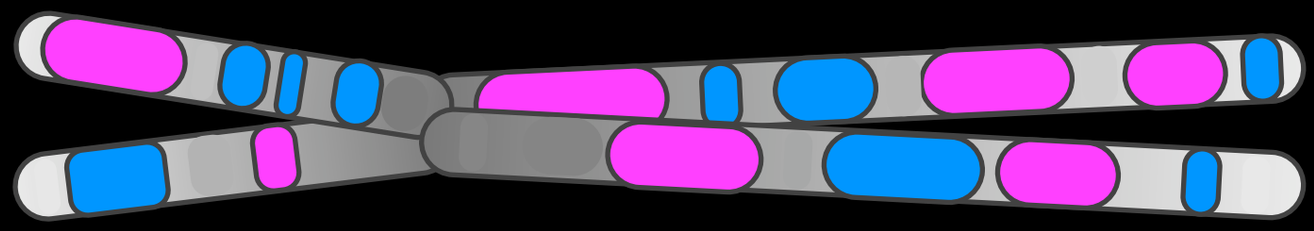
1. S. D. Colman, P. C. Hu, W. Litaker, K. F. Bott, *Mol. Microbiol.* **4**, 683 (1990).

Post-Synthesis Genetics?



1. Design and construct any mutant or combinations thereof; characterize.
2. Reveal / confirm functions or demonstrate absence thereof.

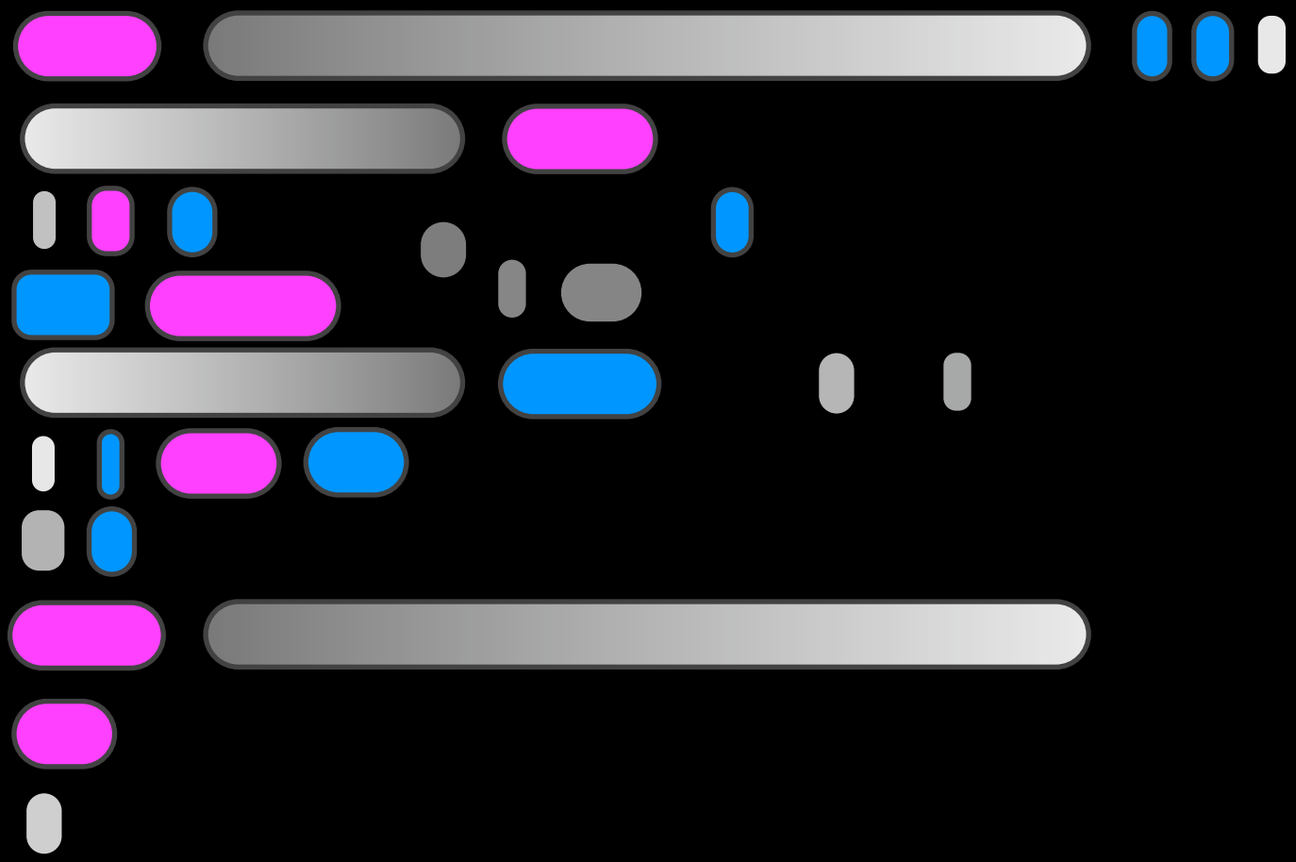
0. Given a newly sequenced & annotated genome...



1. Synthesize and assemble DNA fragments of length, L (say 1.5x max CDS length), offset so as to provide 3x coverage, only including known elements, deleting, scrambling, or otherwise destroying all unannotated sequence (i.e., simplons).

2. Recombine simplons back onto native genome, or into generic testbed (perhaps a generic deletion collection), screen.

4. As a biproduct of preceding, produce full collection of native genetic elements: CDSs, regulatory elements, and so on.



Letters to Nature

Nature **429**, 314-318 (20 May 2004) | doi:10.1038/nature02535; Received 22 January 2004; Accepted 30 March 2004

A highly active synthetic mammalian retrotransposon

Jeffrey S. Han[1] and Jef D. Boeke[1]

1. Department of Molecular Biology and Genetics and High Throughput Biology Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

Correspondence to: Jef D. Boeke[1] Email: jboeke@jhmi.edu

LINE-1 (L1) elements are retrotransposons that comprise large fractions of mammalian genomes[1]. Transcription through L1 open reading frames is inefficient owing to an elongation defect[2], inhibiting the robust expression of L1 RNA and proteins, the substrate and enzyme(s) for retrotransposition[3,4,5]. This elongation defect probably controls L1 transposition frequency in mammalian cells. Here we report bypassing this transcriptional defect by synthesizing the open reading frames of L1 from synthetic oligonucleotides, altering 24% of the nucleic acid sequence without changing the amino acid sequence. Such resynthesis led to greatly enhanced steady-state L1 RNA and protein levels. Remarkably, when the synthetic open reading frames were substituted for the wild-type open reading frames in an established retrotransposition assay[4], transposition levels increased more than 200-fold. This indicates that there are probably no large, rigidly conserved *cis*-acting nucleic acid sequences required for retrotransposition within L1 coding regions. These synthetic retrotransposons are also the most highly active L1 elements known so far and have potential as practical tools for manipulating mammalian genomes.

Published online before print October 31, 2006
Genome Research, DOI: 10.1101/gr.5565706

Letter

Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements

Marie Dewannieux[1,3], Francis Harper[2,4], Aurélien Richaud[1,4], Claire Letzelter[1], David Ribet[1], Gérard Pierron[2], and Thierry Heidmann[1,5]

[1] Unité des Rétrovirus Endogènes et Eléments Rétroïdes des Eucaryotes Supérieurs, UMR 8122 CNRS, Institut Gustave Roussy, 94805 Villejuif Cedex, France; [2] Laboratoire de Réplication de l'ADN et Ultrastructure du Noyau, UPR1983 Institut André Lwoff, 94801 Villejuif Cedex, France

Human Endogenous Retroviruses are expected to be the remnants of ancestral infections of primates by active retroviruses that have thereafter been transmitted in a Mendelian fashion. Here, we derived in silico the sequence of the putative ancestral "progenitor" element of one of the most recently amplified family—the HERV-K family—and constructed it. This element, *Phoenix*, produces viral particles that disclose all of the structural and functional properties of a bona-fide retrovirus, can infect mammalian, including human, cells, and integrate with the exact signature of the presently found endogenous HERV-K progeny. We also show that this element amplifies via an extracellular pathway involving reinfection, at variance with the non-LTR-retrotransposons (LINEs SINEs) or LTR-retrotransposons, thus recapitulating ex vivo the molecular events responsible for its dissemination in the host genomes. We also show that in vitro recombinations among present-day human HERV-K loci can similarly generate functional HERV-K elements, indicating that human cells still have the potential to produce infectious retroviruses.

[3] Present address:

Virus Attenuation by Genome-Scale Changes in Codon Pair Bias

J. Robert Coleman,[1] Dimitris Papamichail,[2]* Steven Skiena,[2] Bruce Futcher,[1] Eckard Wimmer,[3]† Steffen Mueller[1]

As a result of the redundancy of the genetic code, adjacent pairs of amino acids can be encoded by as many as 36 different pairs of synonymous codons. A species-specific "codon pair bias" provides that some synonymous codon pairs are used more or less frequently than statistically predicted. We synthesized de novo large DNA molecules using hundreds of over- or underrepresented synonymous codon pairs to encode the poliovirus capsid protein. Underrepresented codon pairs caused decreased rates of protein translation, and polioviruses containing such amino acid–independent changes were attenuated in mice. Polioviruses thus customized were used to immunize mice and provided protective immunity after challenge. This "death by a thousand cuts" strategy could be generally applicable to attenuating many kinds of viruses.

"I think you're just ignoring the facts.

I'm a geneticist, I write code, okay?

A, G, T, P in different combinations; hacking the human genome, okay?

I choose what I choose. Either your kidneys work or you grow six fingers.

I do that."

Tom Sizemore, *Red Planet* (2000).

Proposed Computing Priorities

- Primary sequence assembly, annotation, analysis
- Functional primary networks (graphs)
- Secondary structure (2D physics)
- Tertiary structures, networks, dynamics (3D physics)
- Reverse engineering tools
  (for reproducing machines;
  time-scale decoupling of control systems;
  pattern forming systems)
- Post-synthesis biology tools (perturbation design)

http://i.pbase.com/u47/hodero/large/30691408.jungle.jpg

Challenge

Make biology easy to engineer.

Opportunity

Enable all of biotechnology.

People

Platforms

Tools

Translation

People

Platforms

Tools

Translation

Main Page - partsregistry.org

http://www.partsregistry.org/

Most Visited · MIT OWW 20.20 iGEM MIT 2008 team Estonia PubMed Cal PiP iKitesurf UAL Reload BBF Swimming

Registry of Standard Biological Parts

Go Search

discussion view source history Log in / create account

The Registry of Standard Biological Parts has moved from parts.mit.edu to partsregistry.org.
References to the Registry at parts.mit.edu will be automatically redirected to the new site.

Registry Tools

Add a part

Over 3500 BioBrick DNA parts

Freely available*

Today

- We are starting an editorial board for promoting well-defined and useful parts to BioBrick™ part status. To join this effort check the BioBrick™ Part Program
- There is a problem with using primers VR and VF2 to PCR parts containing B0015 or B0010.
- News archive...

Report any bugs here | Request new features here | See new features here | See old bugs,requests, and features here

Recent changes What links here Related changes Upload file Special pages My preferences
Recent part changes Printable version Permanent link Privacy policy Disclaimers

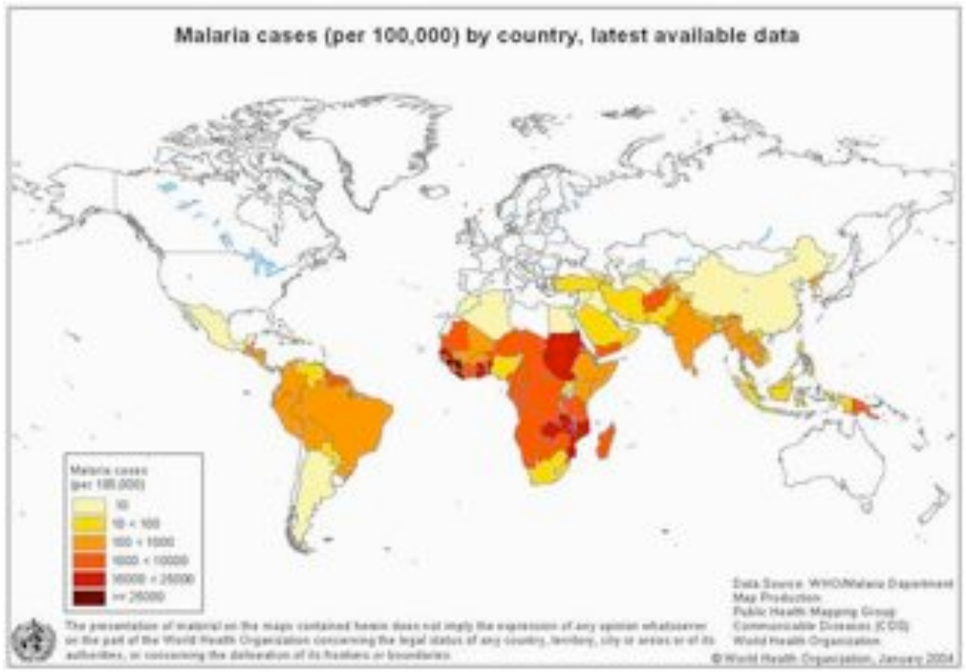
People

Platforms

Tools

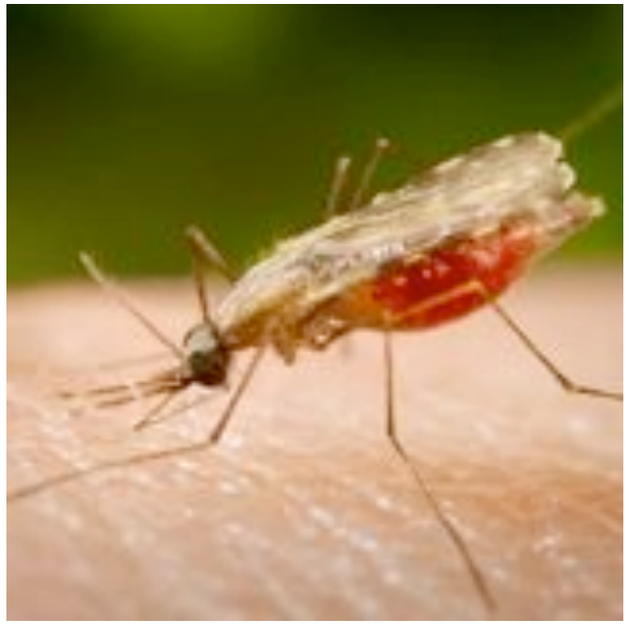
Translation

Today, each project is Herculean



1) Malaria is a global problem, artemisinin offers a cure.

2) Jay Keasling's team spent $25M to make artemisinin via biotechnology.

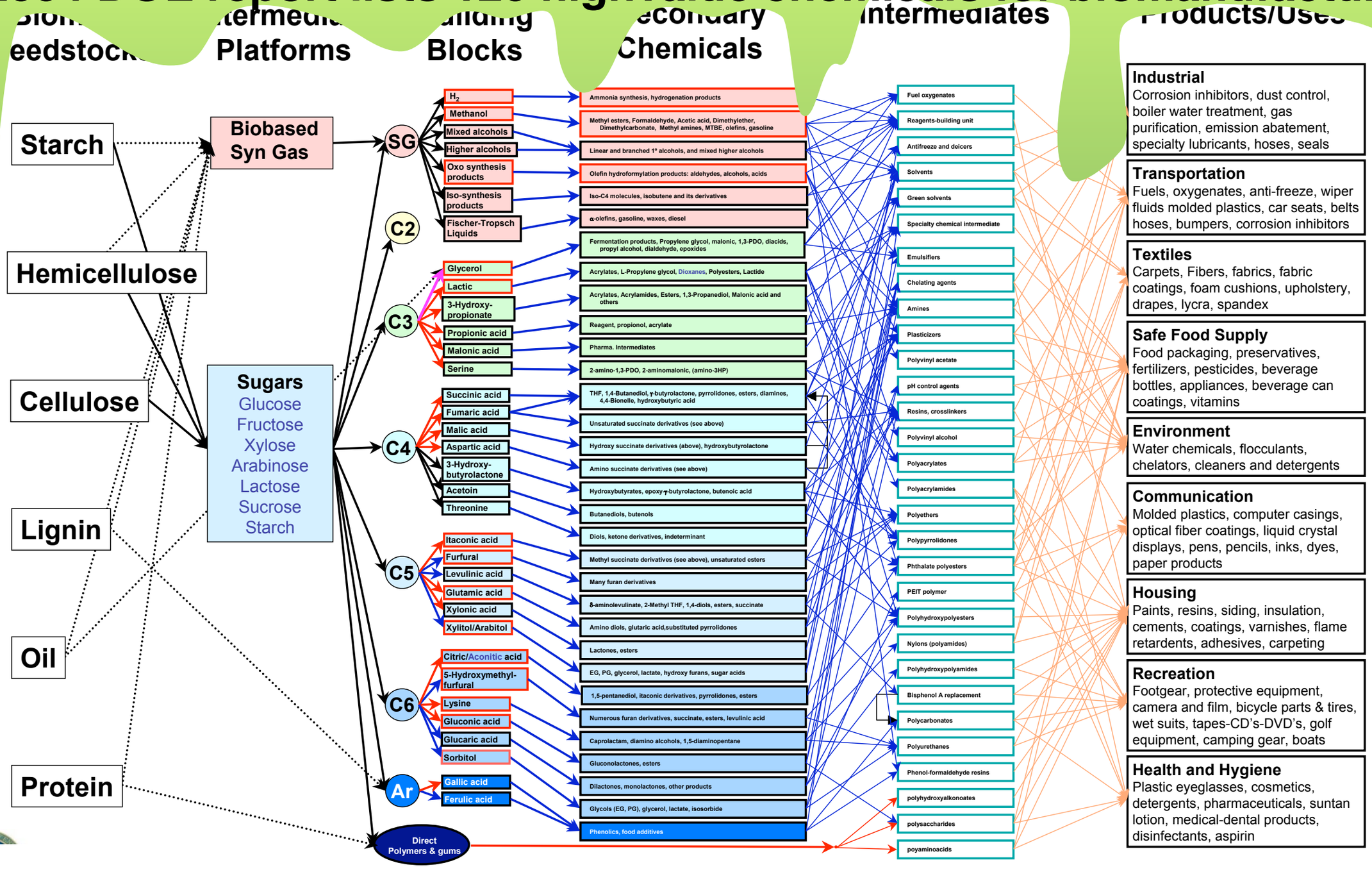


3) But artemisinin resistance is already occurring.

Must we always spend many years and $25M for each pressing biotech project?

Food | Energy | Environment | Agriculture | Health | Chemicals | Security

2004 DOE report lists 120 highvalue chemicals for biomanufacturing



Biotechnology is everywhere

Food	Energy	Environment	Agriculture	Health	Chemicals	Security
$2.5B	$2B	$2.1B	$3B	$91B	$20B	$14B

Total U.S. R&D spending by sector

NIH spends $1.5B just on manipulating DNA

<u>Investment in open technology platforms non-existent</u>

AAAS 2008 Federal R&D Budget Analysis; R&D Magazine 2008; John Mulligan, Blue Heron

We need new tools

1973

Construction of biologically functional bacterial plasmids in vitro

Cohen et al., PNAS, 1973

1985

Cloning and expression of the human erthropoietin gene

Lin et al., PNAS,1985

2006

Production of the antimalarial drug precursor artemisinic acid in engineered yeast

Ro et al., Nature, 2006

MATERIALS AND METHODS

E. coli strain W1485 containing the RSF1000 plasmid, which carries resistance to streptomycin and sulfonamide, was obtained from S. Falkow. Other bacterial strains and R factors and procedures for DNA isolation, electron microscopy, and transformation of *E. coli* by plasmid DNA have been described (1, 7, 8). Purification and use of the *Eco*RI restriction endonuclease have been described (8). Plasmid heteroduplex studies were performed as previously described (9, 10). *E. coli* DNA ligase was a gift from P. Modrich and R. L. Lehman and was used as described (11). The detailed procedures for gel electrophoresis of DNA will be described elsewhere (Helling, Goodman, and Boyer, in preparation); in brief, duplex DNA was subjected to electrophoresis in a tube-type apparatus (Hoefer Scientific Instrument) (0.6 × 15-cm gel) at about 20° in 0.7% agarose at 22.5 V with 40 mM Tris-acetate buffer (pH 8.05) containing 20 mM sodium acetate, 2 mM EDTA, and 18 mM sodium chloride. The gels were then soaked in ethidium bromide (5 μg/ml) and the DNA was visualized by fluorescence under long wavelength ultraviolet light ("black light"). The molecular weight of each fragment in the range of 1 to 200 × 10⁶ was determined from its

Assembly of Expression Vector for the Epo Gene. For direct expression of the genomic Epo gene, the 4.8-kilobase (kb) *Bst*EII–*Bam*HI fragment of λHE1 (see *Results*), which contains the entire Epo gene, was used. After converting the *Bst*EII site into a *Bam*HI site with a synthetic linker, the fragment was inserted into the unique *Bam*HI site of the expression vector pDSVL (unpublished data), which contains a dihydrofolate reductase (DHFR) minigene from pMg1 (24). The resulting plasmid pDSVL-gHuEPO (Fig. 1A) was then used to transfect Chinese hamster ovary (CHO) DHFR⁻ cells (25) by the calcium phosphate microprecipitate method (26). The transformants were selected by growth in medium lacking hypoxanthine and thymidine. The culture medium used was Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, penicillin, streptomycin, and glutamine (25).

Plasmid construction. To create plasmid pRS425ADS for expression of *ADS* with the *GAL1* promoter, *ADS* was PCR amplified from pADS[7] using primer pair 9 and 10. (Supplementary Table 1). Using these primers the nucleotide sequence 5'-AAAACA-3' was cloned immediately upstream of the start codon of *ADS*. This consensus sequence was used for efficient translation[8,9] of *ADS* and the other galactose-inducible genes used in this study. The amplified product was cleaved with SpeI and HindIII and cloned into SpeI and HindIII digested pRS425GAL1[10].

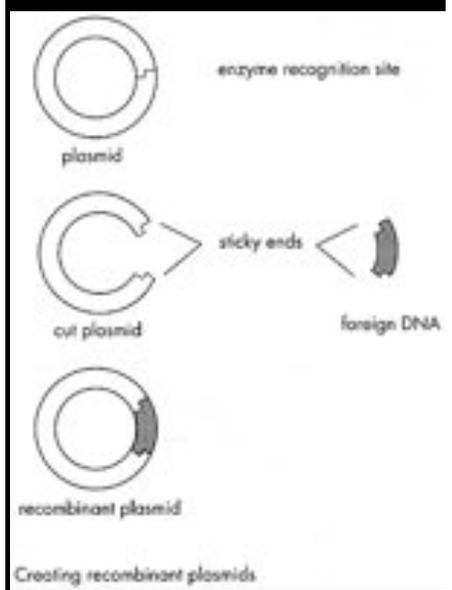
For integration of an expression cassette for *tHMGR*, plasmid p8-HMGR was constructed. First SacII restriction sites were introduced into pRS426GAL1[10] at the 5' end of the *GAL1* promoter and 3' end of the *CYC1* terminator. To achieve this, the promoter-multiple cloning site-terminator cassette of pRS426GAL1 was PCR amplified using primer pair 11 and 12. The amplified product was cloned directly into PvuII-

Genetic engineering basics unchanged past 30+ years

Synthetic biology = tools revolution

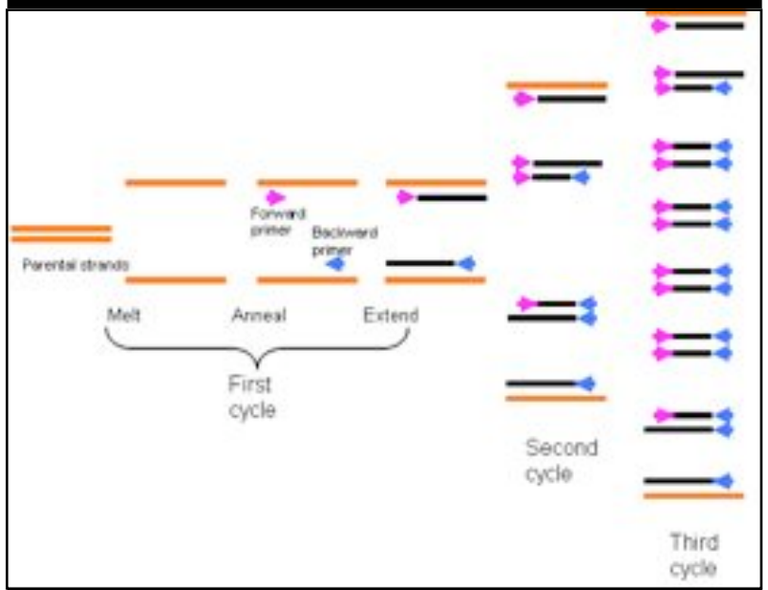
First Gen. Biotech =

Recombinant DNA



Basic "Cut" & "Paste"

Polymerase Chain Reaction



Amplify & Make Simple Changes

DNA Sequencing

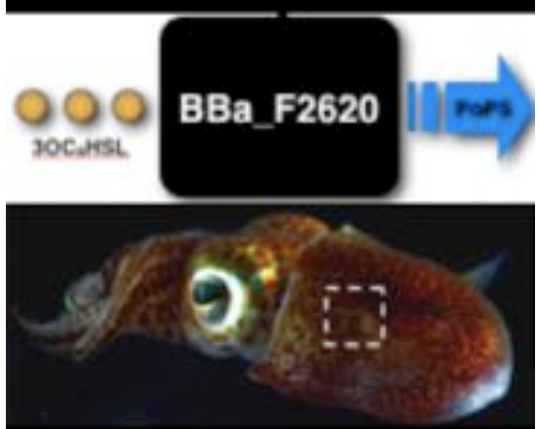
Read Out the Genetic Code

Next Gen. Biotech Adds New Tools =

Synthesis

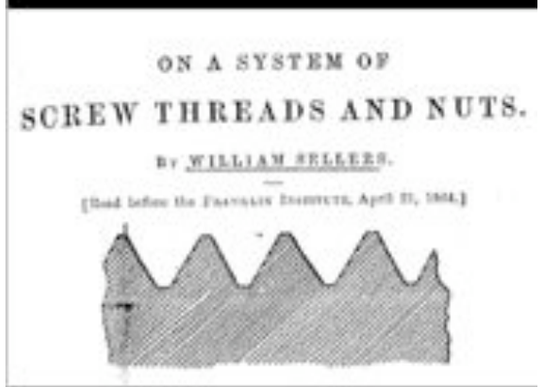
Decoupling of design & fabrication, leading to CAD and EDA.

Abstraction



Engineered simplicity enabling many component systems.

Standardization



Refined genetic components supporting "off the shelf" reuse.

•••

What is the BioFAB?

A public-benefit facility producing the parts, tools & standards powering biotechnology.

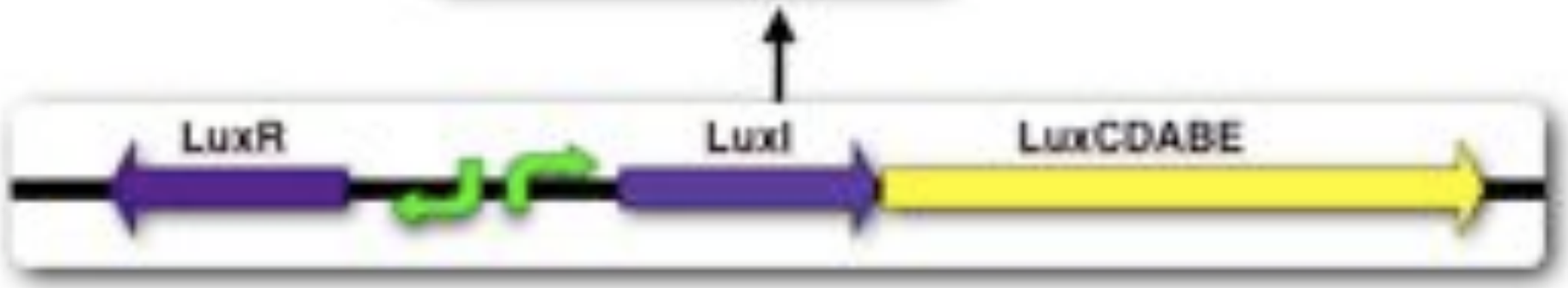
"Off-the-shelf genetic parts cut our project development times and overall costs by half."

- Jack Newman, Founder & SVP Research, Amyris Biotechnologies

Device Engineers

Biologists

3



2

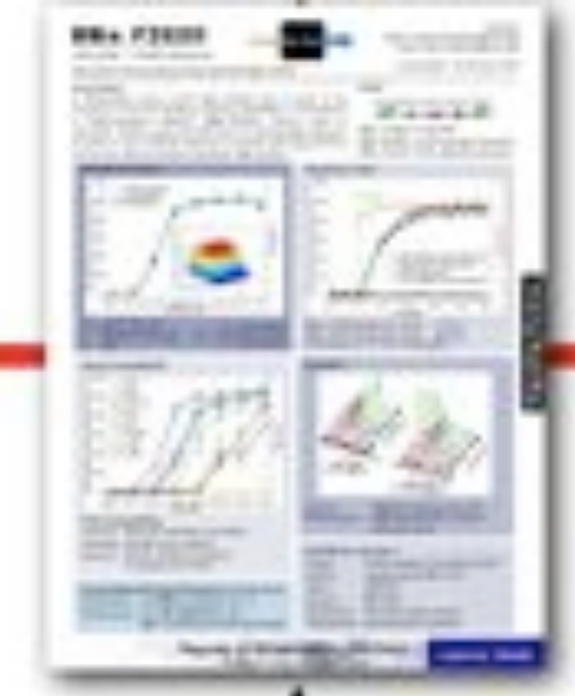
1

Canton, Labno & Endy
Nature Biotechnology v26:7, 787-93 July 2008

System Engineers

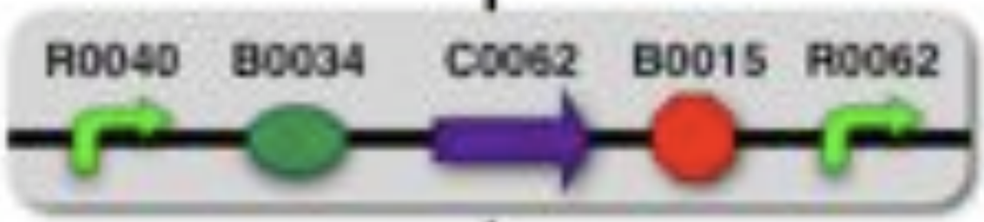
Device Engineers

7

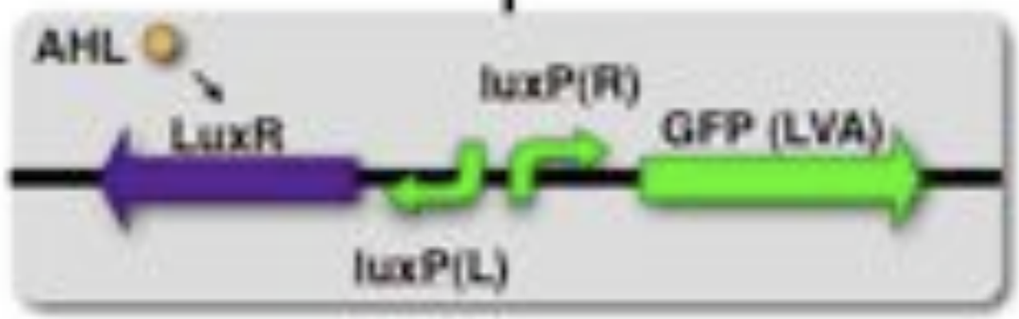


6

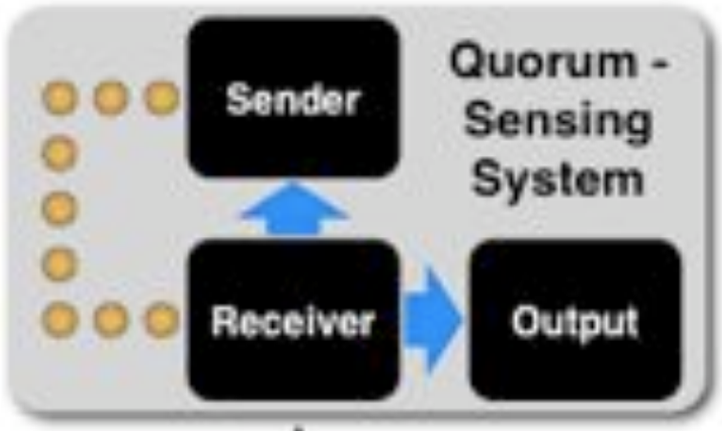
5



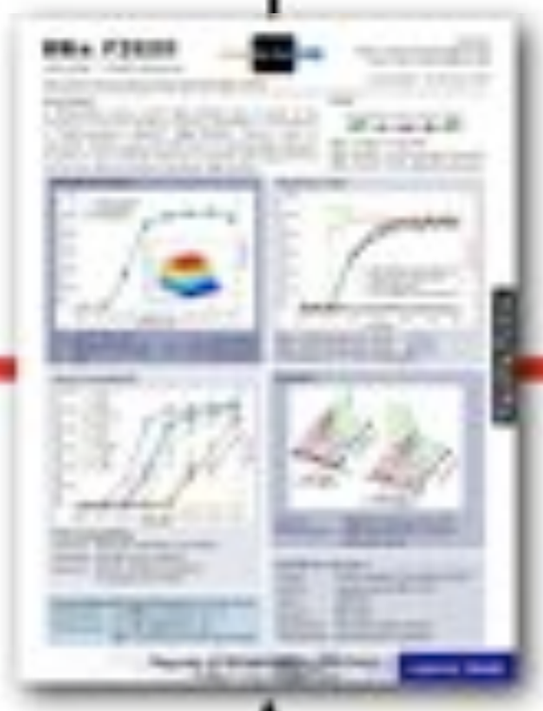
4



Canton, Labno & Endy
Nature Biotechnology v26:7, 787-93 July 2008



8



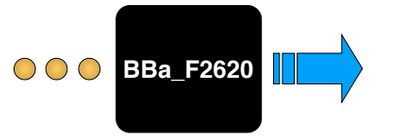
System Engineers

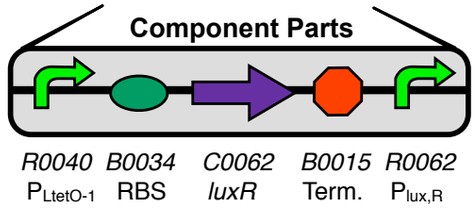
7

Device Engineers

BBa_F2620

3OC$_6$HSL → PoPS Receiver

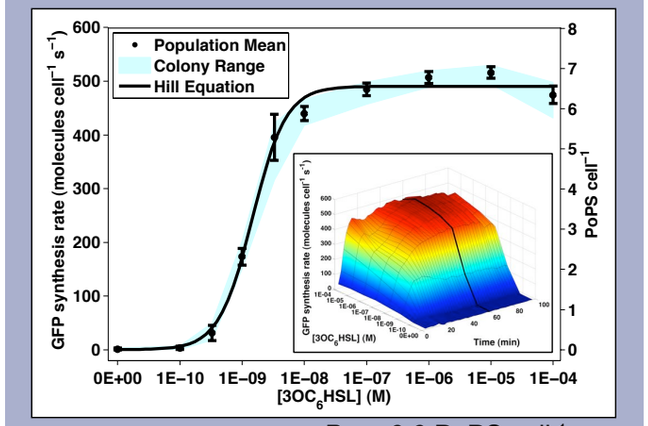




Mechanism & Function

A transcription factor (LuxR) that is active in the presence of a cell-cell signaling molecule (3OC$_6$HSL) is controlled by a regulated operator (P$_{LtetO-1}$). Device input is 3OC$_6$HSL. Device output is PoPS from a LuxR-regulated operator. If used in a cell containing TetR then a second input such as aTc can be used to produce a Boolean AND function.

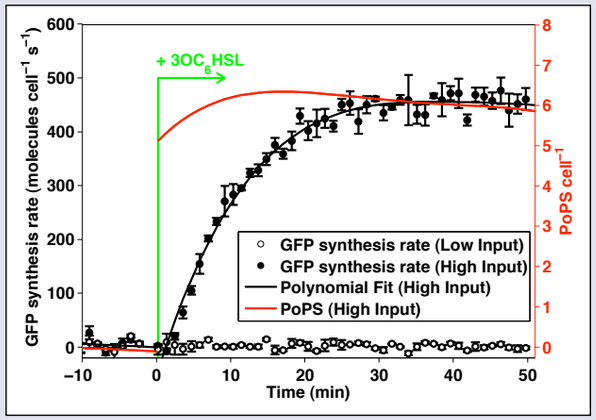
Static Performance*



$$P_{out} = \frac{P_{max}[3OC_6HSL]^n}{K^n + [3OC_6HSL]^n}$$

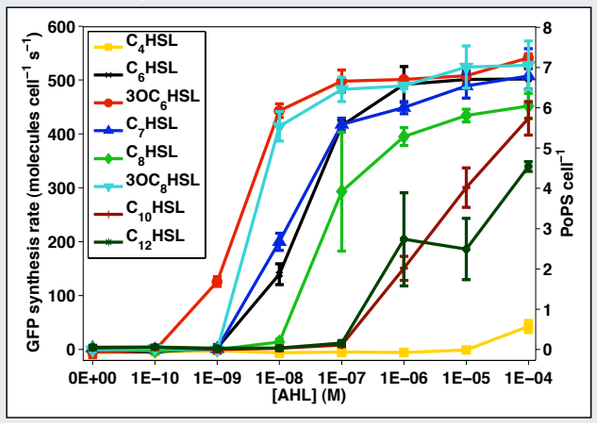
P_{max}: **6.6** PoPS cell^{-1}
K: **1.5E-09** M 3OC$_6$HSL
n: **1.6**

Dynamic Performance*

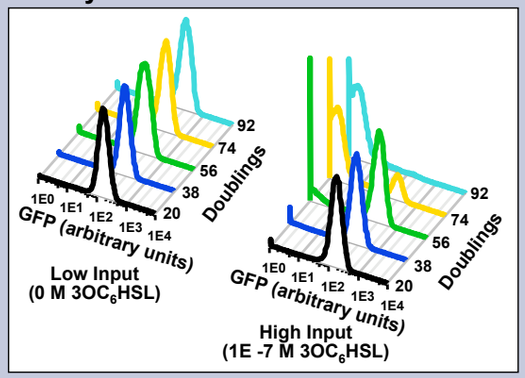


BBa_F2620 Response Time: ***<1** min*
BBa_T9002 Response Time: ***6±1** min*
Inputs: 0 M (Low), 1E-07 M (High) 3OC$_6$HSL

Input Compatibility*



Reliability**



Genetic: **>92/>56** culture doublings
Performance: **>92/>56** culture doublings
(low/high input during propagation)

Part Compatibility (qualitative)

Chassis: *MC4100*, *MG1655*, and *DH5α*
Plasmids: *pSB3K3* and *pSB1A2*
Devices: *E0240*, *E0430* and *E0434*

Transcriptional Output Demand (low/high input)

Nucleotides: **0 / 6xNt** nucleotides cell^{-1} s^{-1}
Polymerases: **0 / 1.5E-1xNt** RNAP cell^{-1}
(**Nt** = downstream transcript length)

Conditions (abridged)

Output:	PoPS measured via *BBa_E0240*
Culture:	Supplemented M9, 37ºC
Plasmid:	pSB3K3
Chassis:	MG1655
**Equipment:*	PE Victor3 multi-well fluorimeter
***Equipment:*	BD FACScan cytometer

http://parts.mit.edu/registry/index.php/Part:BBa_F2620

Signaling Devices

Canton, Labno & Endy
Nature Biotechnology v26:7, 787-93 July 2008

Authors: Barry Canton
Ania Labno
Updated: March 2008

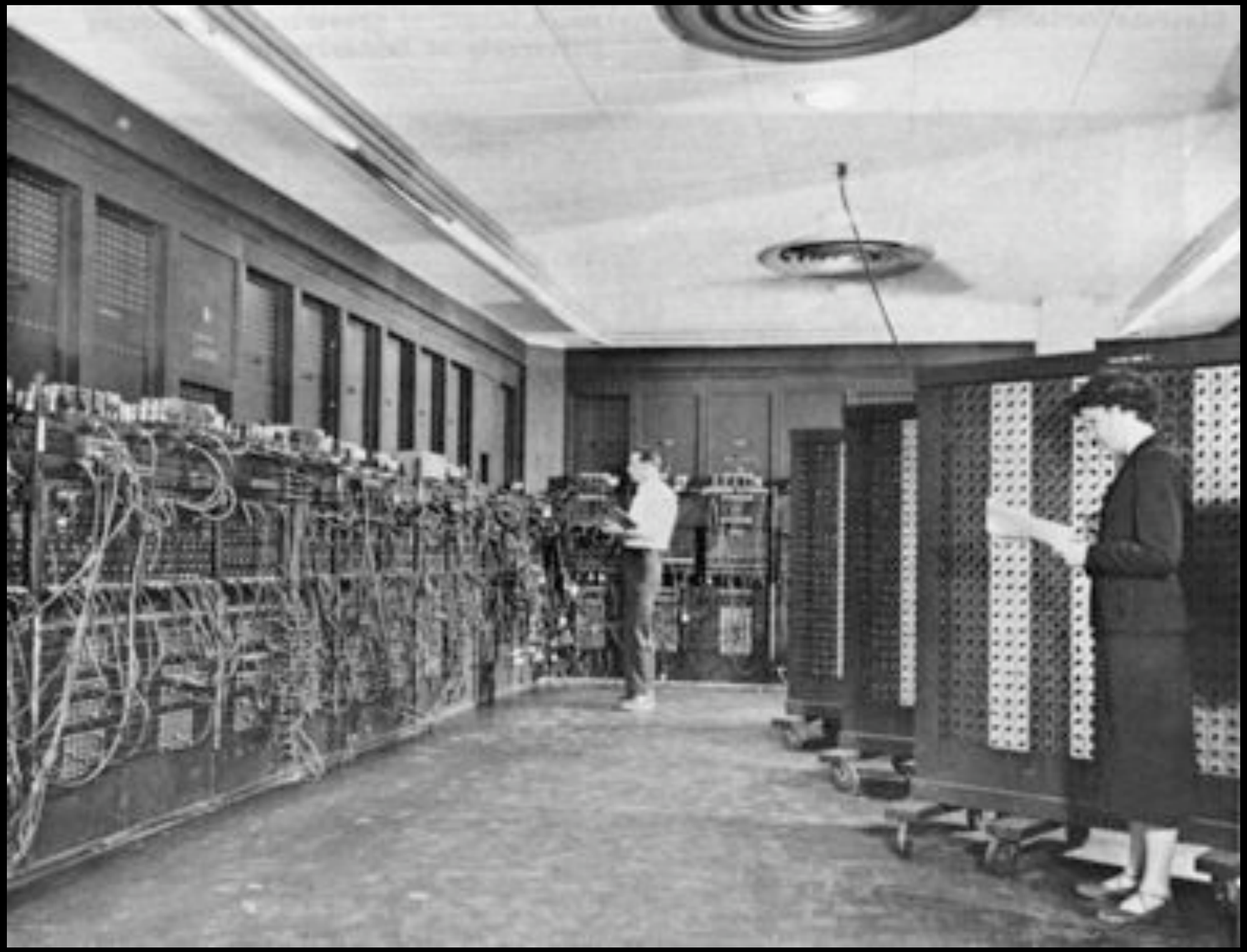
Unknown
Our Lord's Candle (Yucca whipplei), Lakeview Mountains, 22 Apr 2006
Smugmug.com

Unknown
http://www.rootsweb.com/~usgenweb/mn/stearns/postcards/quarry.jpg

Washington National Cathedral

http://www.cathedral.org/cathedral/discover/1930photos/5.shtml

http://www.nicmyers.com/media/hongkong_large_pano_nicmyers.jpg



http://en.wikipedia.org/wiki/Image:Eniac.jpg

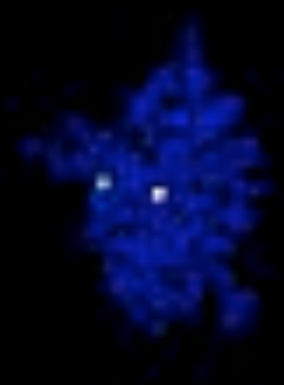
http://en.wikipedia.org/wiki/Image:Apple_I.jpg

http://en.wikipedia.org/wiki/Internet

Apps

Tools

7. Control & dyn. systems
6. Reverse engineering
5. Fab, CAD & EDA
4. Standards & abstraction
3. Languages & grammars
2. Device design
1. Info. theory & signal proc.



Nagpal, Kondacs, Chang, Programming Methodology for Biologically-Inspired Self-Assembling Systems, in the *AAAI Spring Symposium on Computational Synthesis: From Basic Building Blocks to High Level Functionality*, March 2003.

Movie online via http://www.swiss.csail.mit.edu/projects/amorphous/Robust/cross.mpeg

6. Standardization
5. Abstraction
4. DNA construction
3. DNA sequencing
2. Polymerase chain reaction
1. Recombinant DNA

Proposed Computing Priorities

- Advocacy for tools!!!
- Advocacy for open technology platforms.
- More tools, including:
    - Languages (e.g., for programming dynamic shapes)
    - CAD and EDA (what's going into the synthesizers?)
    - Coding schemes (signatures, fail fast, et cetera)
- Theory supporting of design of understandable reproducing machines
- Tools that enable humanity (e.g., Games, Community)