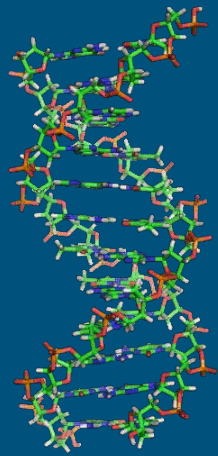
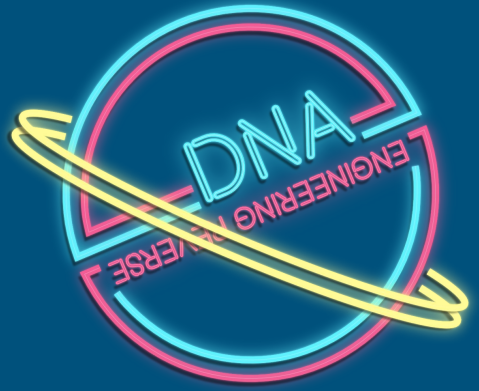
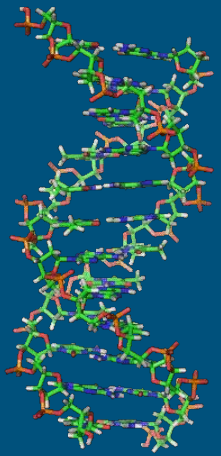


A brief 2025 recap
DNA: the code of life





Credits

- This presentation leans heavily on other people's work and graphics
- All credits are available in the **speaker notes** which you should consult to find out who made all these great movies and images
- **Thank you so much Wikipedia Commons in particular!**

https://berthub.eu/whydna

Original Article

The ISME Journal (2016) **10,** 30–38; doi:10.1038/ismej.2015.107; published online 3 July 2015

Density-dependent adaptive resistance allows swimming bacteria to colonize an antibiotic gradient

Felix J H Hol[1], Bert Hubert[1], Cees Dekker[1] and Juan E Keymer[1,2,3]

[1]Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands
[2]Department of Ecology, Faculty of Biological Sciences, P. Catholic University of Chile, Santiago, Chile
[3]Institute of Physics, Faculty of Physics, P. Catholic University of Chile, Santiago, Chile

Correspondence: FJH Hol or JE Keymer, Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, Delft 2628CJ, The Netherlands. E-mail: f.j.h.hol@tudelft.nl or jkeymer@uc.cl

Received 5 January 2015; Revised 9 April 2015; Accepted 19 May 2015
Advance online publication 3 July 2015

Abstract

During antibiotic treatment, antibiotic concentration gradients develop. Little is know regarding the effects of antibiotic gradients on populations of nonresistant bacteria. Using a microfluidic device, we show that high-

Data Descriptor | Open access | Published: 22 March 2022

SkewDB, a comprehensive database of GC and 10 other skews for over 30,000 chromosomes and plasmids

Bert Hubert

Scientific Data **9**, Article number: 92 (2022) | Cite this article

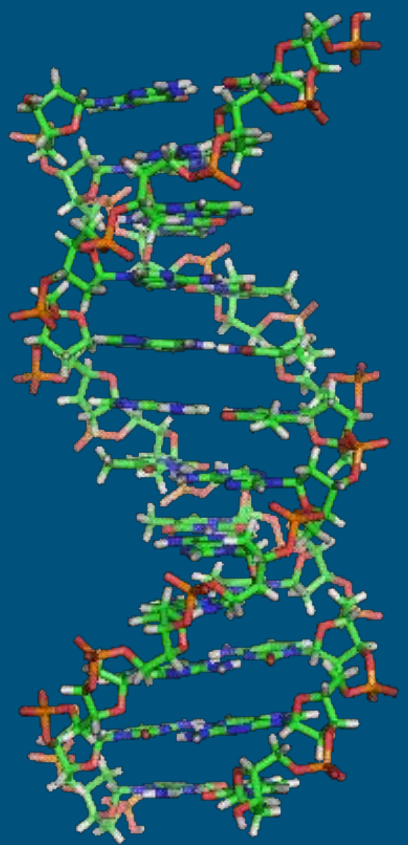
“Imagine a flashy spaceship lands in your backyard. The door opens and you are invited to investigate everything to see what you can learn. The technology is clearly millions of years beyond what we can make.

This is biology.”

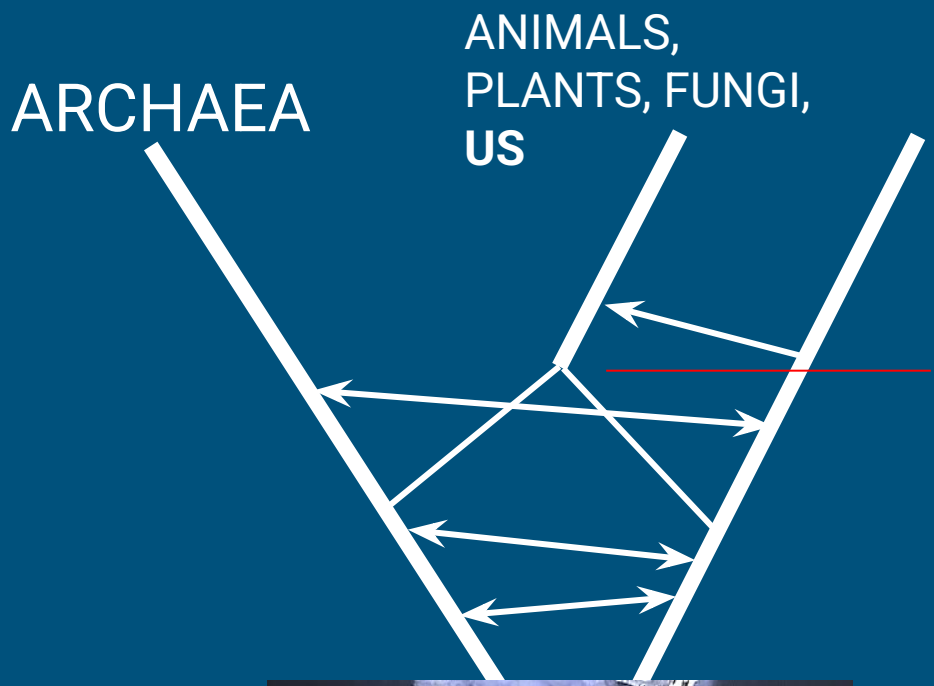
Although…

© NASA (oddly enough)

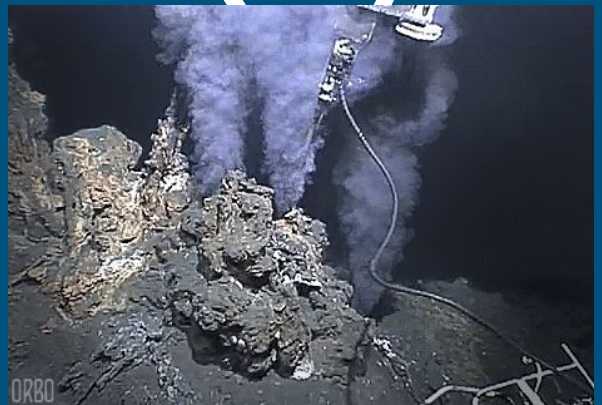
1. Information storage
2. The 3D printer of life
3. Algorithms & control
4. Hacks



- DNA: Millions, billions of nucleotides or “bases”:
  - A, C, G, T
- Organized in chromosomes & genes
- Absolutely **atom for atom** universal across all life
  - >4 billion years old
- Stable for 100s of thousands of years



2 billion years ago



4 billion years ago

First water: 4.2 billion

Age of earth: 4.5 billion

Basics: Each nucleotide is **2 (arbitrary) bits**

A	**00**
C	**01**
G	**10**
T	**11**

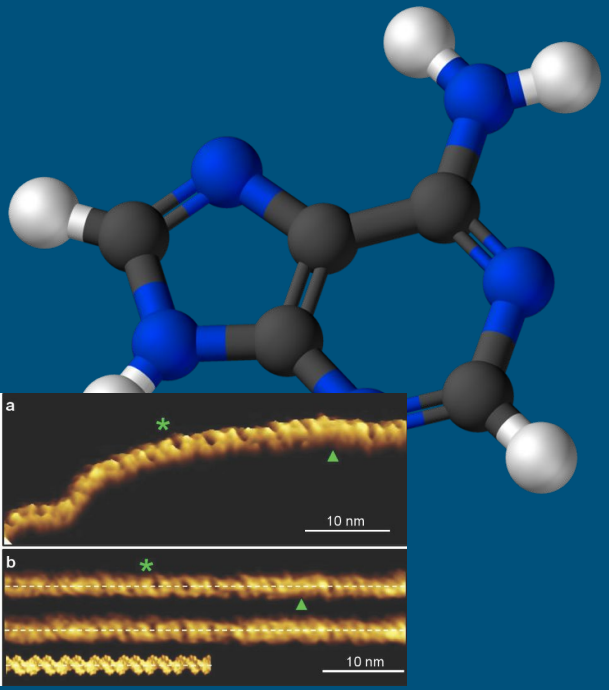
Basics: DNA

A

C

G

T

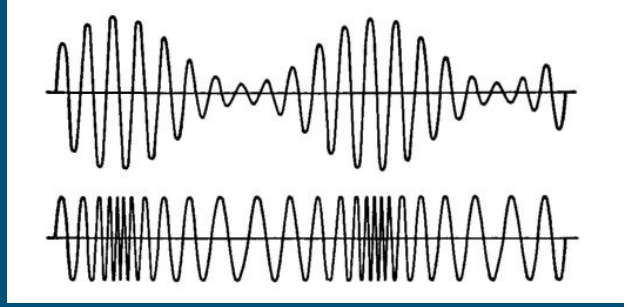


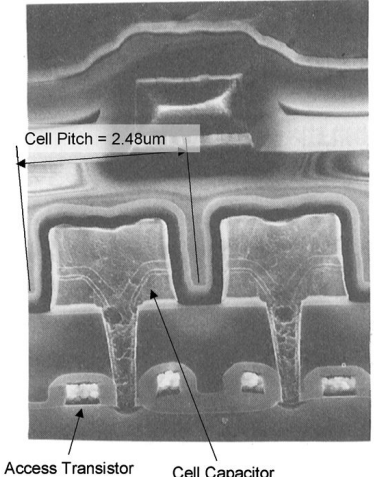
00

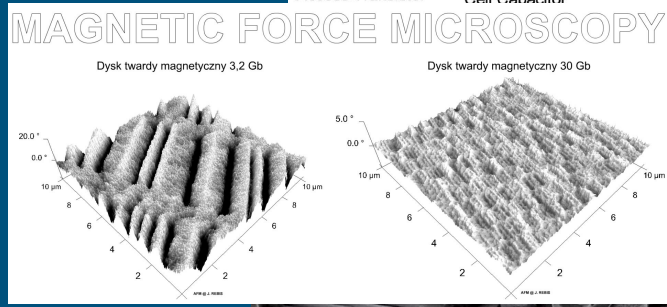
01

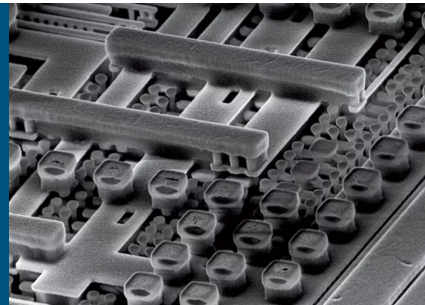
10

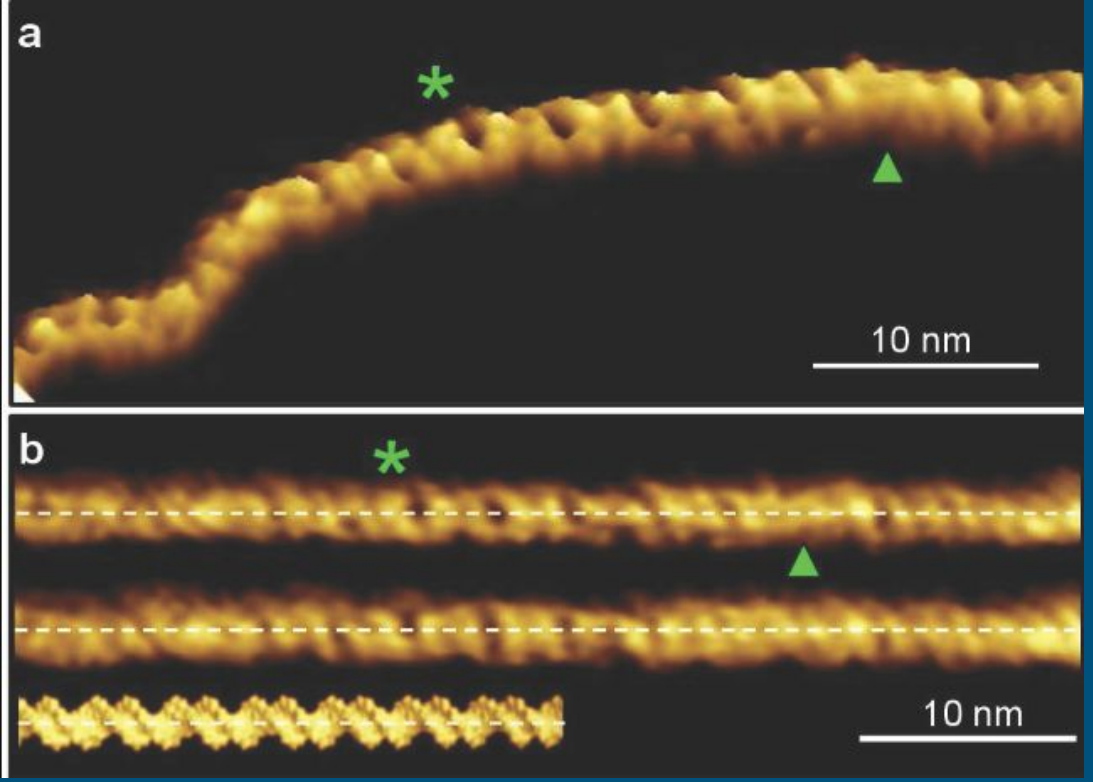
11











DNA is very much like tape

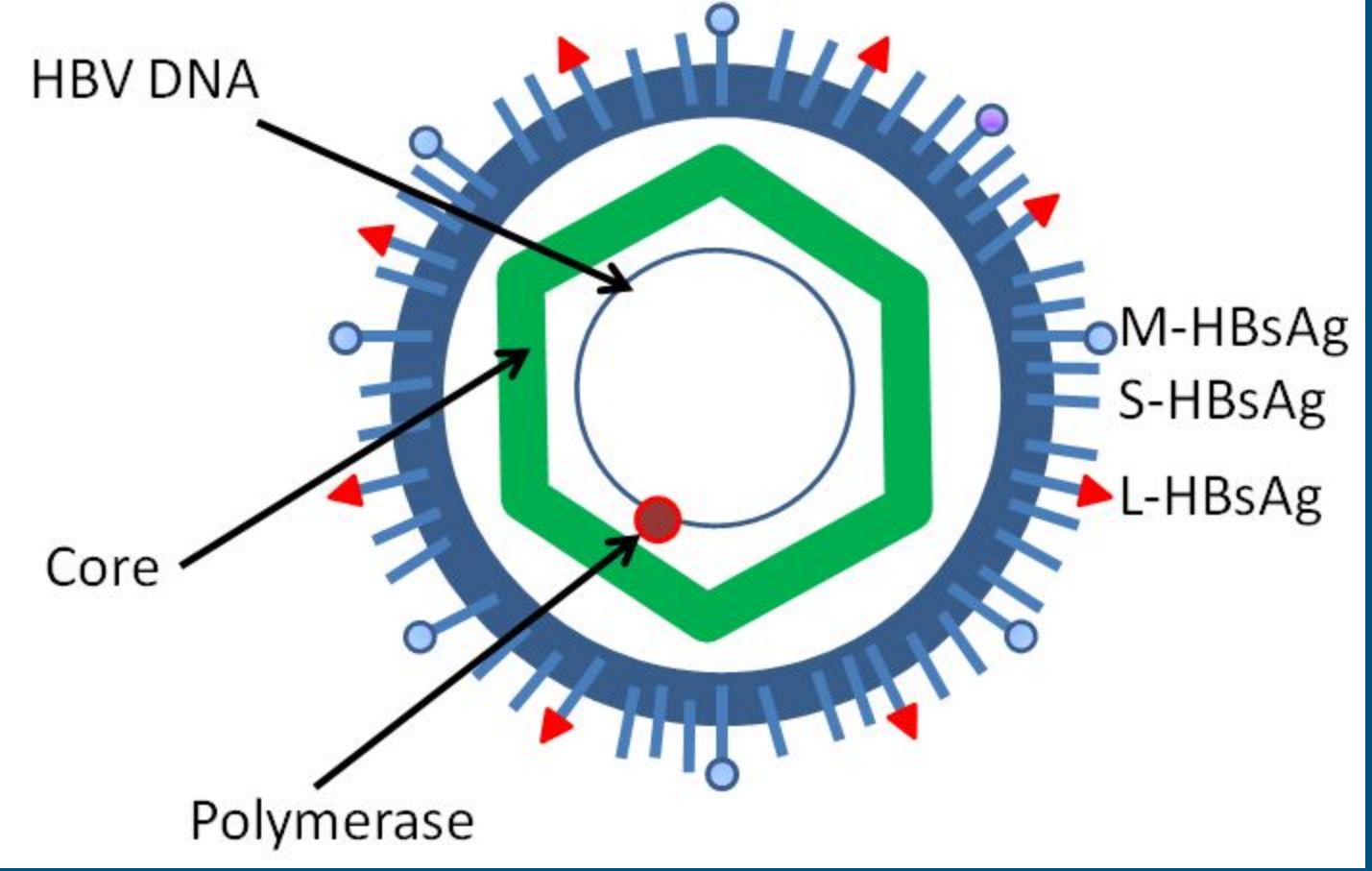
Sometimes circular tape - no beginning, no end!

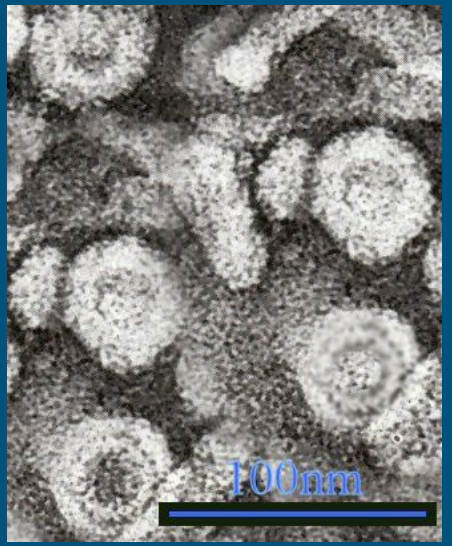
No addressing! No alignment!

It is a **nucleotide stream** which can be compared to a **bitstream**

It IS however **content addressable**!

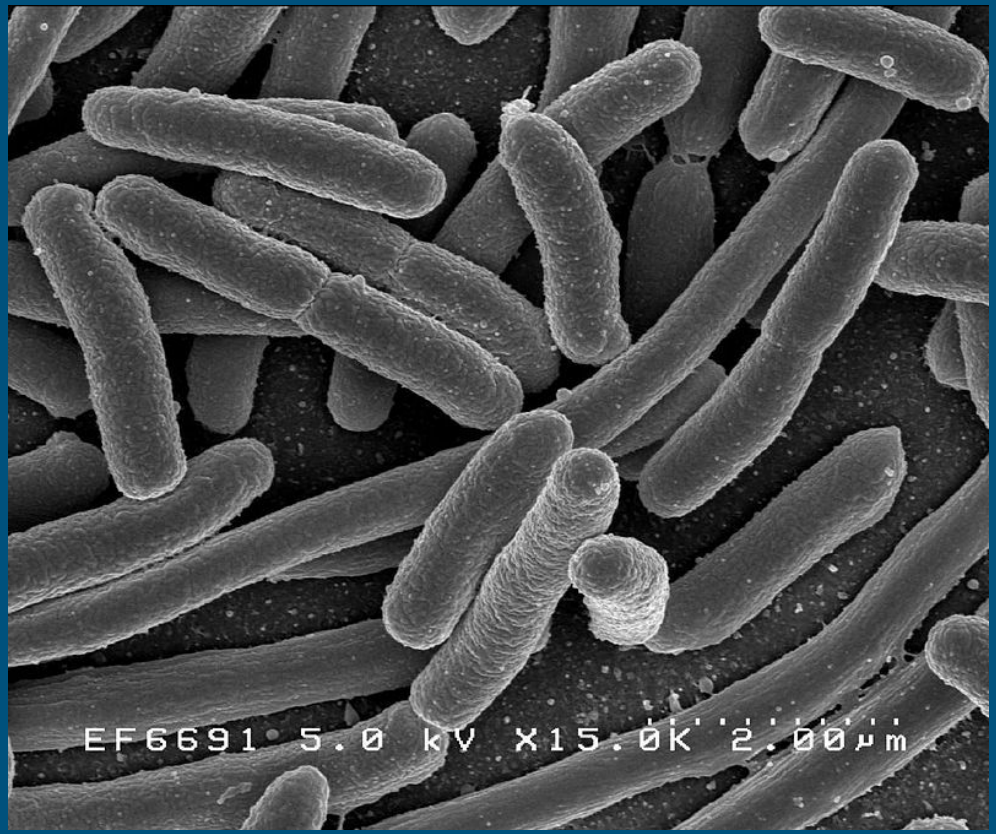
And it has FRAMING ~~challenges~~ **possibilities**

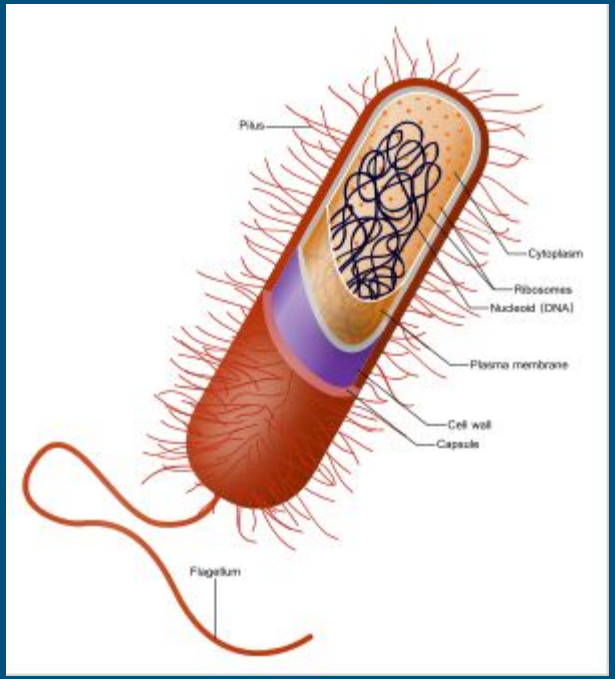




Hepatitis B - 800 bytes (unpleasant)

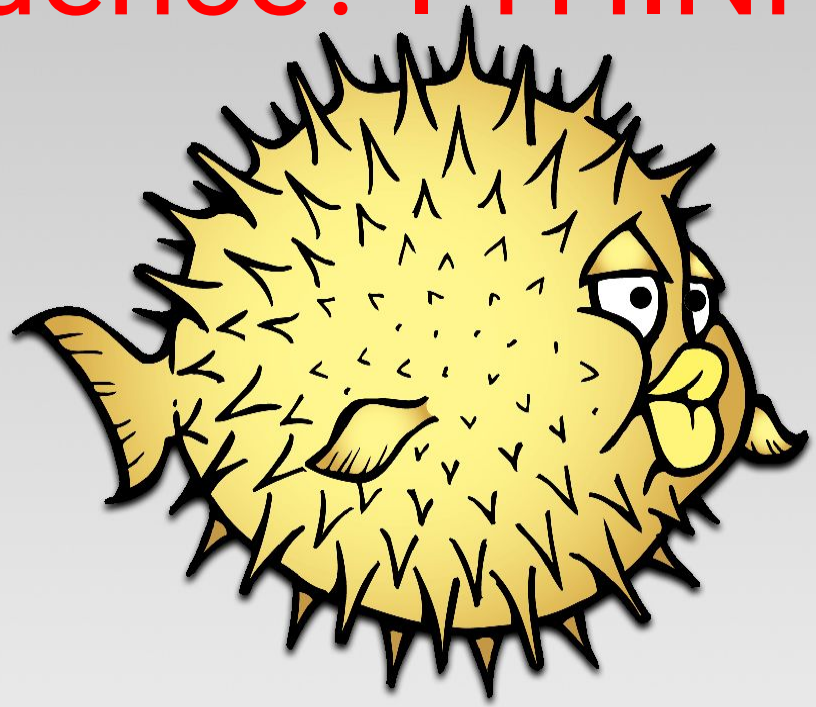
All of Hepatitis-B - 800 bytes



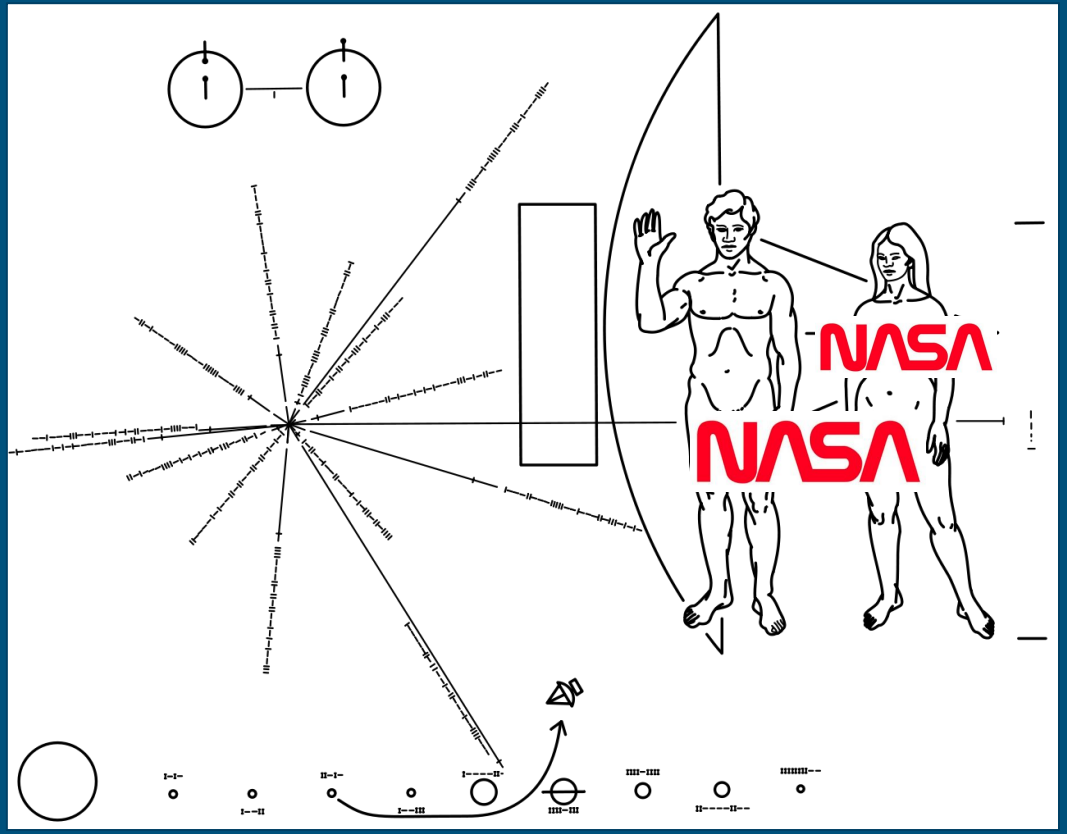


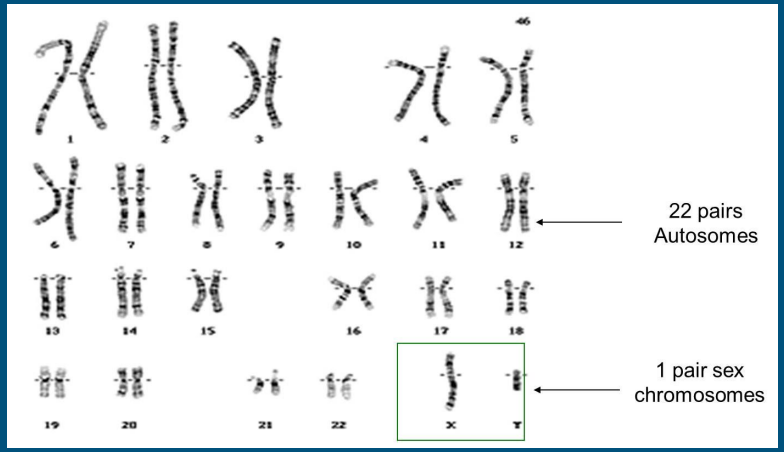
E. Coli: 750KB (can be good, can be unpleasant)

Coincidence? I THINK NOT!



Pufferfish: 100MB, smallest & tightest genome of all animals (lethal)





Your unique DNA *diff* is around 1-10 megabyte. HOWEVER - could be way more: the pangenome

Homo sapiens: **750MB** (lethal)

Paris Japonica: 37.5GB (quite pleasant)

Marbled Lungfish: 33GB

But is life REALLY digital?

RESEARCH ARTICLE

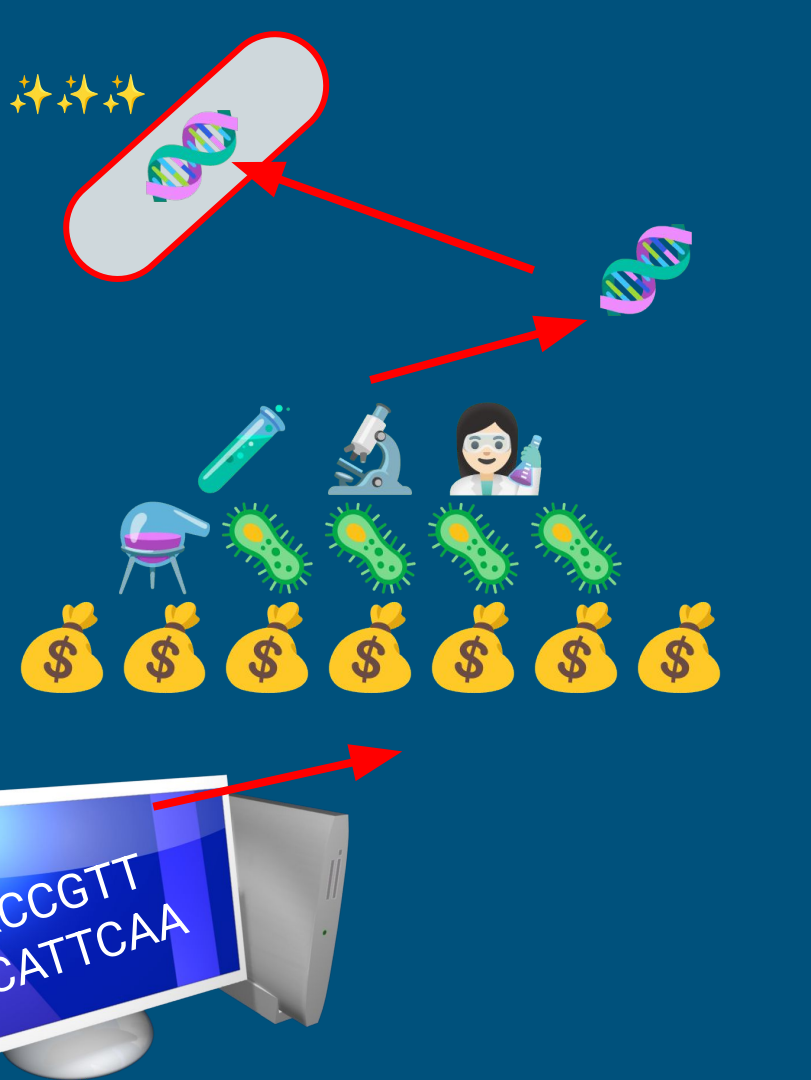
Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

DANIEL G. GIBSON, JOHN I. GLASS, CAROLE LARTIGUE, VLADIMIR N. NOSKOV, RAY-YUAN CHUANG, MIKKEL A. ALGIRE, GWYNEDD A. BENDERS, MICHAEL G. MONTAGUE, LI MA, [...], AND J. CRAIG VENTER +14 authors Authors Info & Affiliations

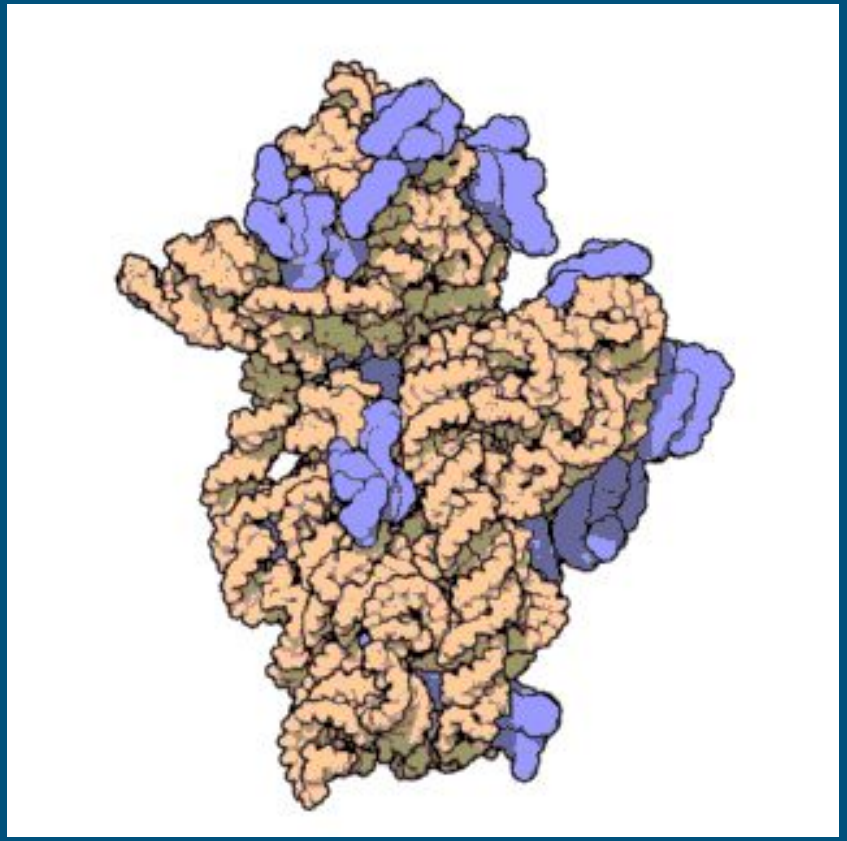
SCIENCE · 20 May 2010 · Vol 329, Issue 5987 · pp. 52-56 · DOI: 10.1126/science.1190719

Let There Be Life

The DNA sequence information from thousands of genomes is stored digitally as ones and zeros in computer memory. Now, **Gibson *et al.*** (p. 52, published online 20 May; see the cover; see the Policy Forum by **Cho and Relman**) have brought together technologies from the past 15 years to start from digital information on the genome of *Mycoplasma mycoides* to chemically synthesize the genomic DNA as segments that could then be assembled in yeast and transplanted into the cytoplasm of another organism. A number of methods were also incorporated to facilitate testing and error correction of the synthetic genome segments. The transplanted genome became established in the recipient cell, replacing the recipient genome, which was lost from the cell. The reconstituted cells were able to replicate and form colonies, providing a proof-of-principle for future developments in synthetic biology.



Behold! The source of all life & you



~10
nanometers

BEHOLD! Part of the Ribosome, the 1D/3D printer of life!
It printed you or the things that made you!

The Central Dogma

Long term storage: **DNA (*/dev/sda*)**

Converts to live form: **RNA (*RAM*)**

RNA converts to: **Proteins**

Proteins: **Sense and Do**

Twenty-One Amino Acids

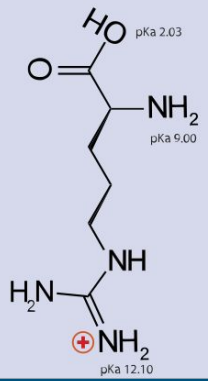
⊕ Positive ⊖ Negative

• Side chain charge at physiological **pH 7.4**

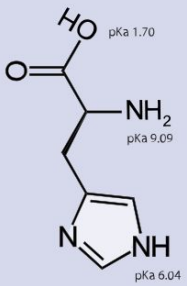
A. *Amino Acids with Electrically Charged Side Chains*

Positive

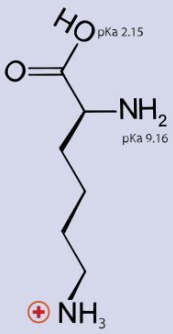
Arginine (Arg) R



Histidine (His) H

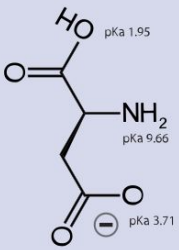


Lysine (Lys) K

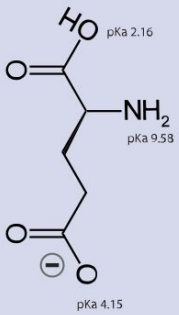


Negative

Aspartic Acid (Asp) D

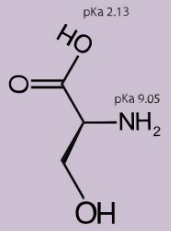


Glutamic Acid (Glu) E

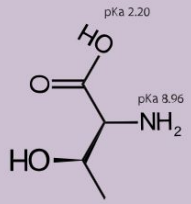


B. *Amino Acids with Polar Uncharged Side Chains*

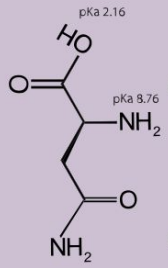
Serine (Ser) S



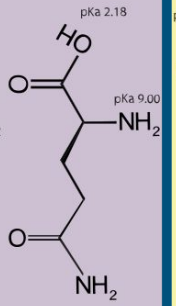
Threonine (Thr) T



Asparagine (Asn) N

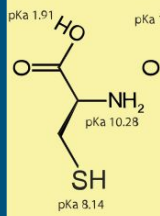


Glutamine (Gln) Q

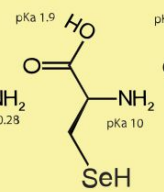


C. *Special Cases*

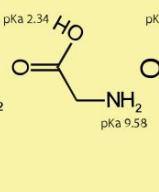
Cysteine (Cys) C



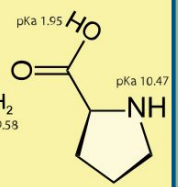
Selenocysteine (Sec) U



Glycine (Gly) G

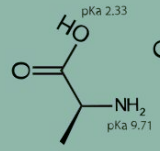


Proline (Pro) P

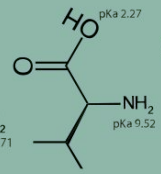


D. *Amino Acids with Hydrophobic Side Chain*

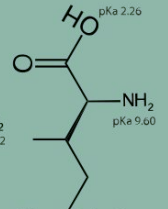
Alanine (Ala) A



Valine (Val) V



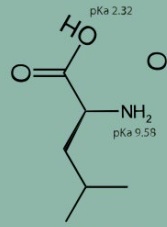
Isoleucine (Ile) I

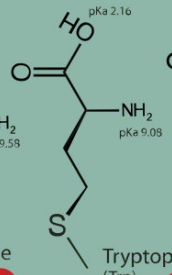


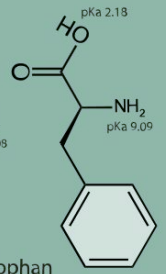
Leucine (Leu) L

Methionine (Met) M

Phenylalanine (Phe) F

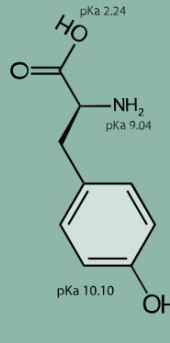


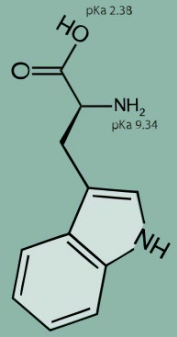




Tyrosine (Tyr) Y

Tryptophan (Trp) W





1st base	2nd base								3rd base
	T		C		A		G		
T	TTT	(Phe/F) Phenylalanine (np)	TCT	(Ser/S) Serine (p)	TAT	(Tyr/Y) Tyrosine (p)	TGT	(Cys/C) Cysteine (p)	T
	TTC		TCC		TAC		TGC		C
	TTA	(Leu/L) Leucine (np)	TCA		TAA	Stop (*Ochre*) *[note 2]	TGA	Stop (*Opal*) *[note 2]	A
	TTG ⇒		TCG		TAG	Stop (*Amber*) *[note 2]	TGG	(Trp/W) Tryptophan (np)	G
C	CTT		CCT	(Pro/P) Proline (np)	CAT	(His/H) Histidine (b)	CGT	(Arg/R) Arginine (b)	T
	CTC		CCC		CAC		CGC		C
	CTA		CCA		CAA	(Gln/Q) Glutamine (p)	CGA		A
	CTG		CCG		CAG		CGG		G
A	ATT	(Ile/I) Isoleucine (np)	ACT	(Thr/T) Threonine (p)	AAT	(Asn/N) Asparagine (p)	AGT	(Ser/S) Serine (p)	T
	ATC		ACC		AAC		AGC		C
	ATA		ACA		AAA	(Lys/K) Lysine (b)	AGA	(Arg/R) Arginine (b)	A
	ATG ⇒	(Met/M) Methionine (np)	ACG		AAG		AGG		G
G	GTT	(Val/V) Valine (np)	GCT	(Ala/A) Alanine (np)	GAT	(Asp/D) Aspartic acid (a)	GGT	(Gly/G) Glycine (np)	T
	GTC		GCC		GAC		GGC		C
	GTA		GCA		GAA	(Glu/E) Glutamic acid (a)	GGA		A
	GTG ⇒		GCG		GAG		GGG		G

Multi-billion year old table!

Multiple codons for same amino acids

This allows for **dialects** and shaping DNA

>gnl|ECOLI|G0-10439 kdpF MONOMER0-12 (complement(728732..728821)) Escherichia coli K-12 substr. MG1655

gtgAGTGCAGGCGTGATAACCGGCGTATTGCTGGTGTTTTTATTACTGGGTTATCTGGTTTATGCCCTGA
TCAATGCGGAGGCGTTCtga

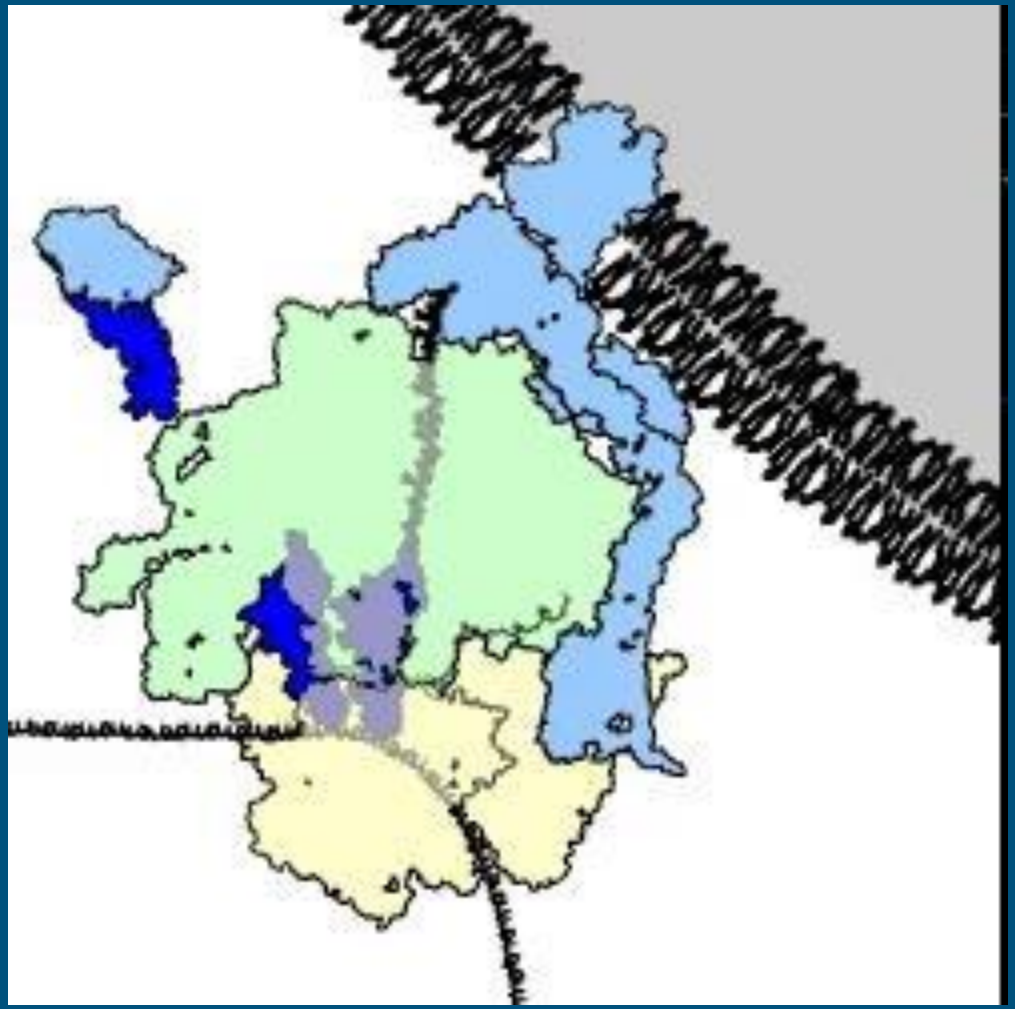
>gnl|ECOLI|G0-10439 kdpF MONOMER0-12 (complement(728732..728821)) Escherichia coli K-12 substr. MG1655

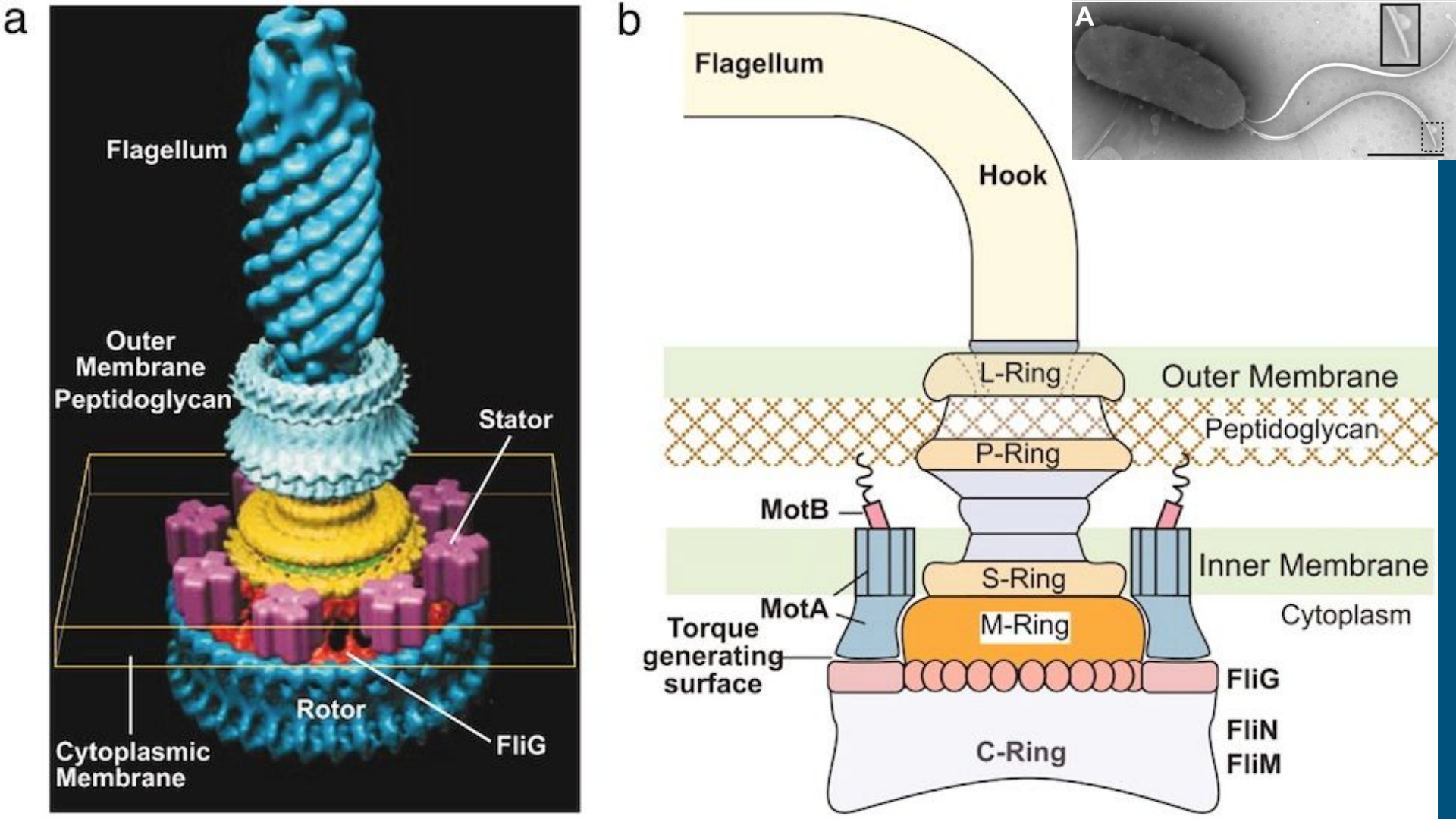
gtg AGT GCA GGC GTG ATA ACC GGC GTA TTG CTG GTG TTT TTA TTA CTG GGT TAT CTG GTT
TAT GCC CTG ATC AAT GCG GAG GCG TTC tga

>gnl|ECOLI|MONOMER0-12 gn=kdpF K+ transporting P-type ATPase subunit KdpF (complement(728732..728821)) Escherichia coli K-12 substr. MG1655

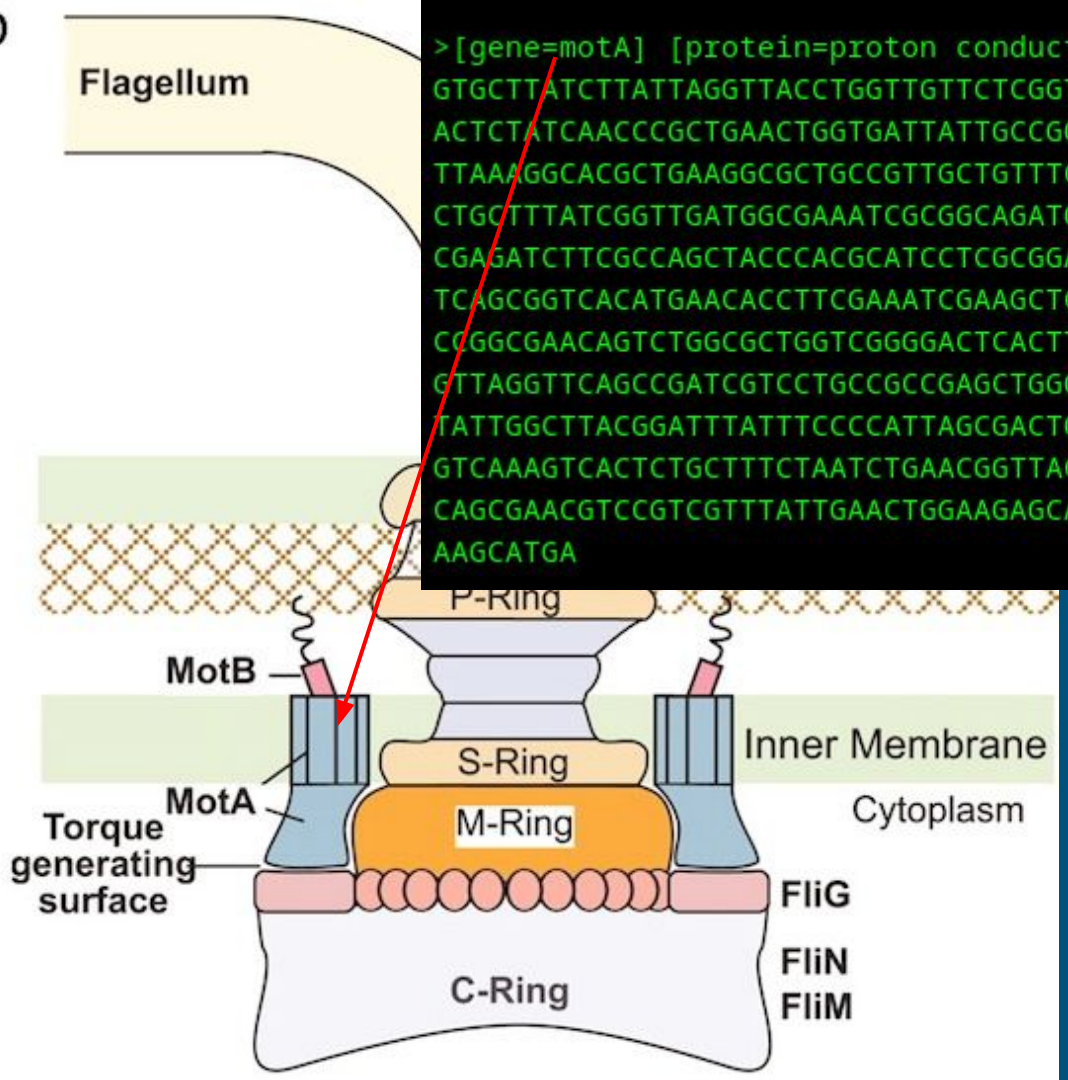
MSAGVITGVLLVFLLLGYLVYALINAEAF<EOP>

“A 1D printer that leads to 3D objects with ATOMIC resolution”




a
Flagellum
Outer Membrane
Peptidoglycan
Stator
Rotor
FliG
Cytoplasmic Membrane
b
Flagellum
Hook
A
L-Ring
Outer Membrane
Peptidoglycan
P-Ring
MotB
Inner Membrane
S-Ring
MotA
Cytoplasm
Torque generating surface
M-Ring
FliG
FliN
FliM
C-Ring

>[gene=motA] [protein=proton conductor component of flagella motor]
GTGCTTATCTTATTAGGTTACCTGGTTGTTCTCGGTACAGTTTTCGGCGGTTATTTGATGACCGGTGGAAGCCTTGGAGC
ACTCTATCAACCCGCTGAACTGGTGATTATTGCCGGTGCAGGGATTGGGTCGTTTATCGTCGGCAATAATGGCAAAGCGA
TTAAAGGCACGCTGAAGGCGCTGCCGTTGCTGTTTCGTCGCTCCAAATACACCAAAGCAATGTATATGGATCTGCTGGCT
CTGCTTTATCGGTTGATGGCGAAATCGCGGCAGATGGGGATGTTTTCGCTGGAACGTGATATTGAAAATCCCCGTGAGAG
CGAGATCTTCGCCAGCTACCCACGCATCCTCGCGGATAGCGTCATGCTTGATTTTATCGTCGATTATCTGCGCCTGATTA
TCAGCGGTCACATGAACACCTTCGAAATCGAAGCTCTGATGGATGAAGAGATTGAGACGCACGAAAGCGAGGCAGAAGTC
CGGGCGAACAGTCTGGCGCTGGTCGGGGACTCACTTCCGGCGTTTGGTATTGTTGCGGCTGTAATGGGGGTCGTTCACGC
GTTAGGTTCAGCCGATCGTCCTGCCGCCGAGCTGGGTGCGCTTATCGCACATGCGATGGTGGGGACTTTCCTCGGCATTT
TATTGGCTTACGGATTTATTTCCCCATTAGCGACTGTTTTACGTCAGAAAAGCGCCGAAACCAGCAAAATGATGCAGTGC
GTCAAAGTCACTCTGCTTTCTAATCTGAACGGTTACGCACCGCCTATCGCCGTTGAGTTTGGTCGCAAAACGCTCTATTC
CAGCGAACGTCCGTCGTTTATTGAACTGGAAGAGCATGTGCGTGCGGTGAAAAATCCGCAACAACAGACGACAACCGAGG
AAGCATGA



If you 3D print this model yourself:

- Your resolution is like 10 million times too coarse
- Your flagellar motor does not actually work
- It looks remarkably silly! (‘honestly, this is for science!’)

Gene control & how it hangs together

DNA layer
RNA layer
Amino acid layer
..
RNA layer
DNA layer

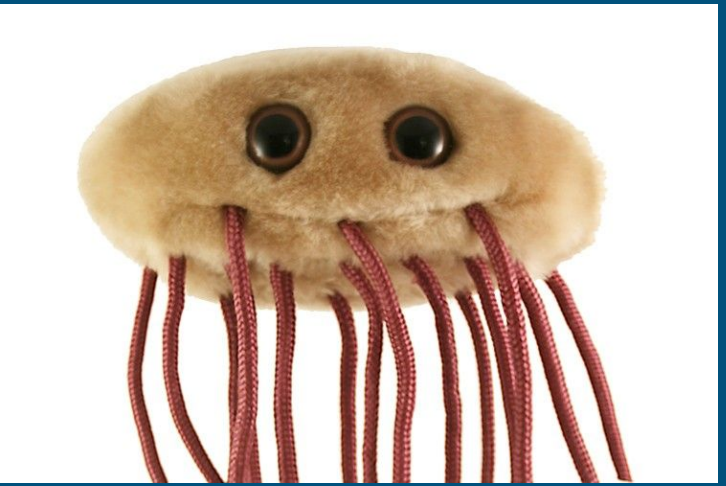
Amino Acids

RNA

Amino acids code

Promoter	-35	-10	Shine-Dalgarno	ATG…TGA
	“TTGACA”	“TATAAT”	“AGGAGG”	

E. coli

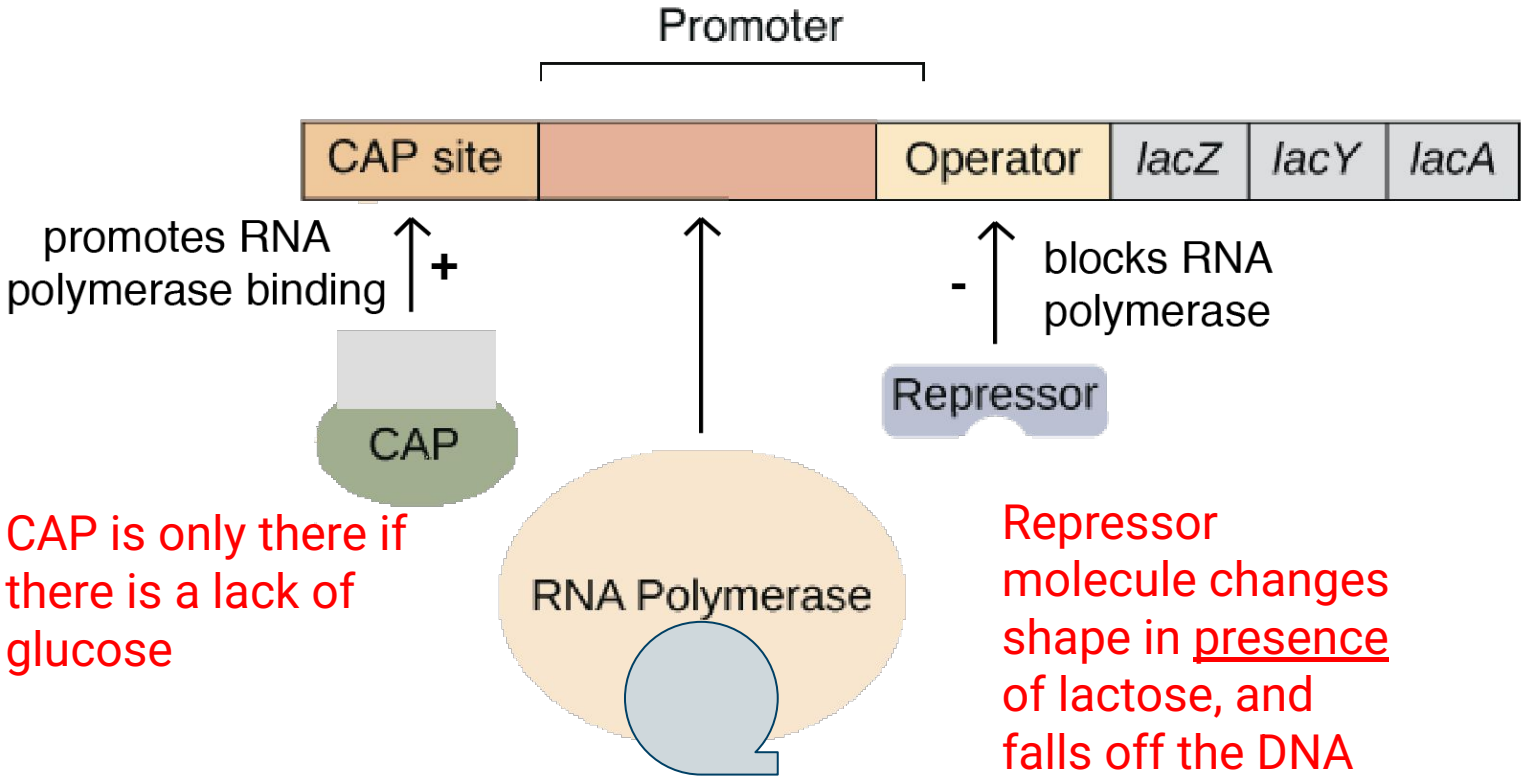


- Needs power to run
  - 0.5pW
- Loves glucose
- Can run on **lactose** if it must
  - Through conversion
  - Less efficient
- Algorithm required:

```
if(!glucosePresent && lactosePresent)

    convertLactoseToGlucose();
```

The *lac* operon:

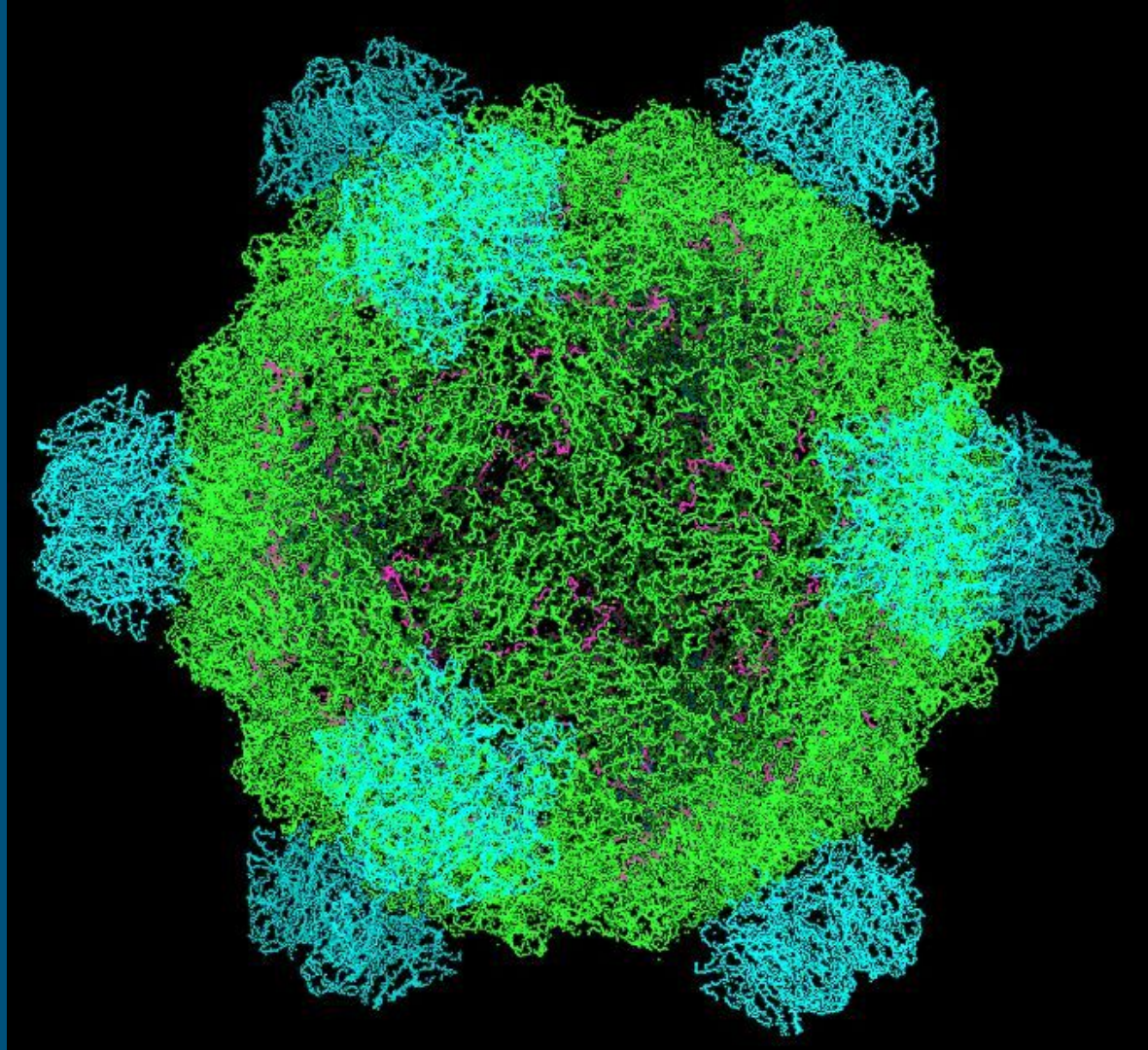


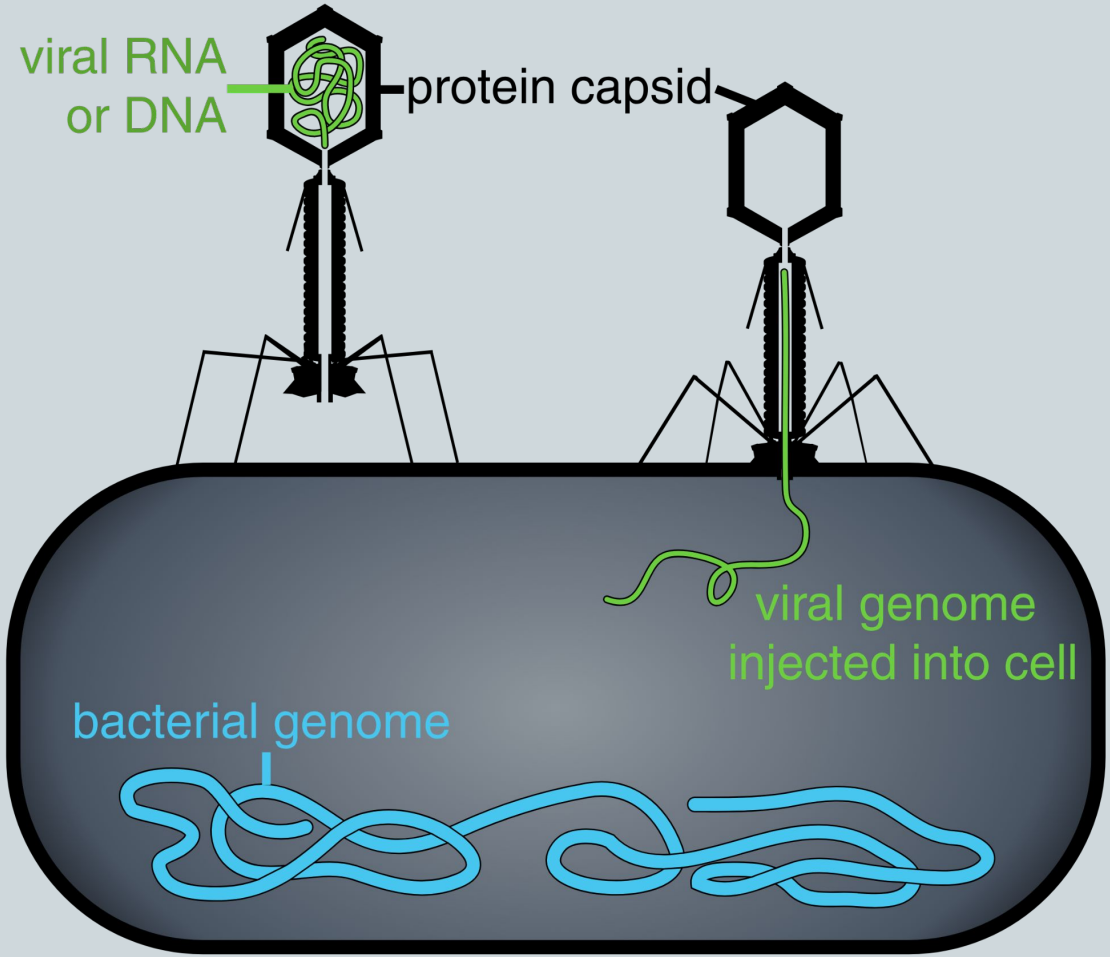
CAP is only there if there is a lack of glucose

Repressor molecule changes shape in presence of lactose, and falls off the DNA

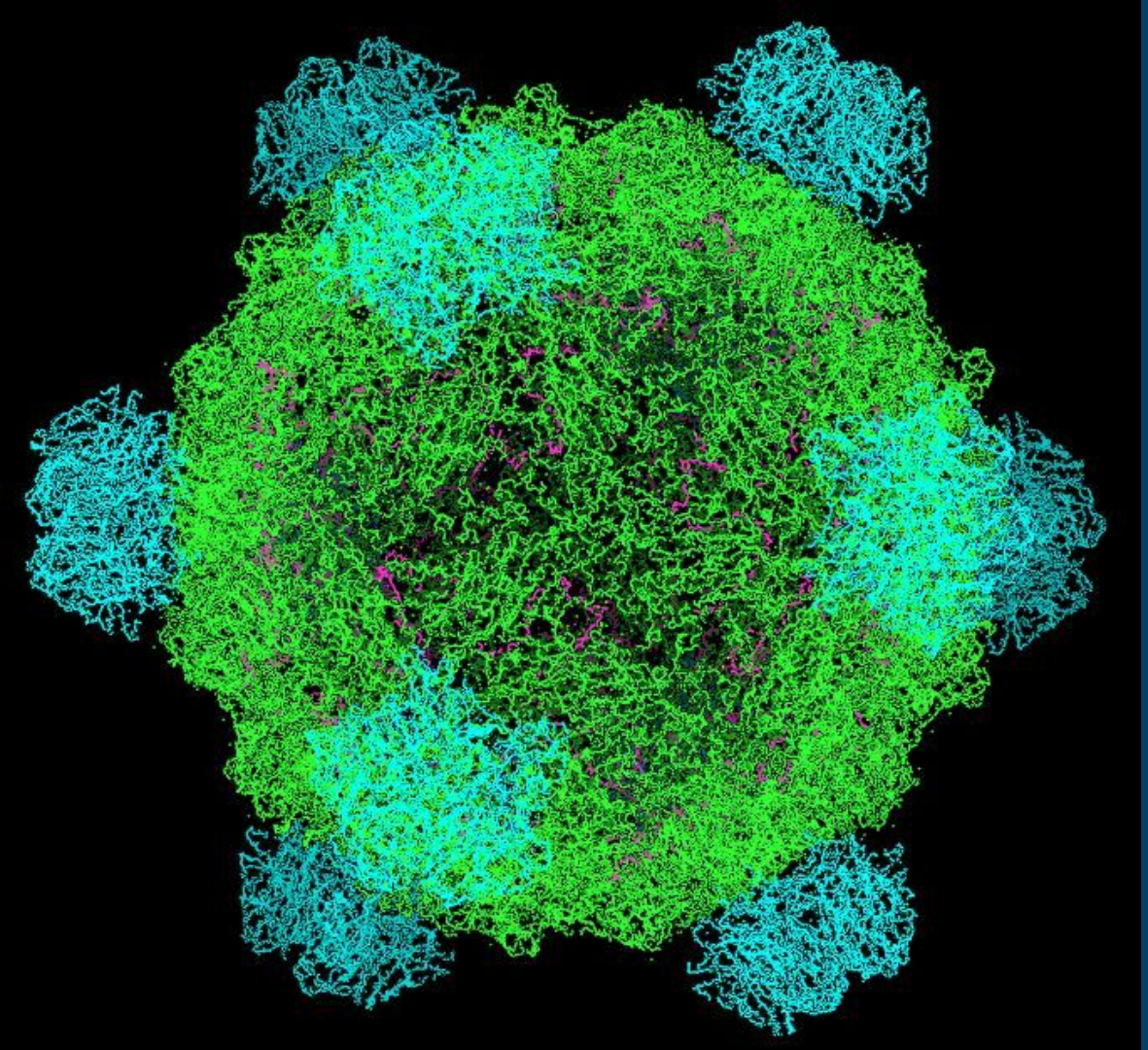
Viruses: the ultimate hackers

“Shell code”





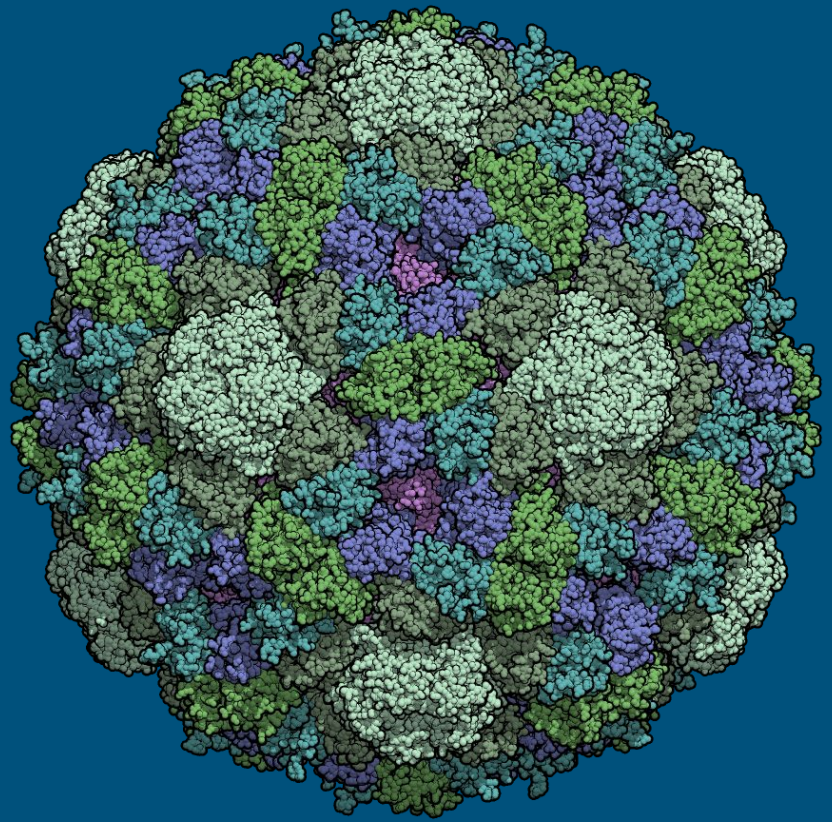
"Does not end well..."



ΦX174 infects e coli.

Having **less DNA** means faster copying, smaller virion, **evading defences**

“Shell code shaping”



https://commons.wikimedia.org/wiki/File:Phi_X_174_procapsid.png

ΦX174 has 5386 DNA letters

Enough for 5386/3 = **1795** amino acids, TOPS

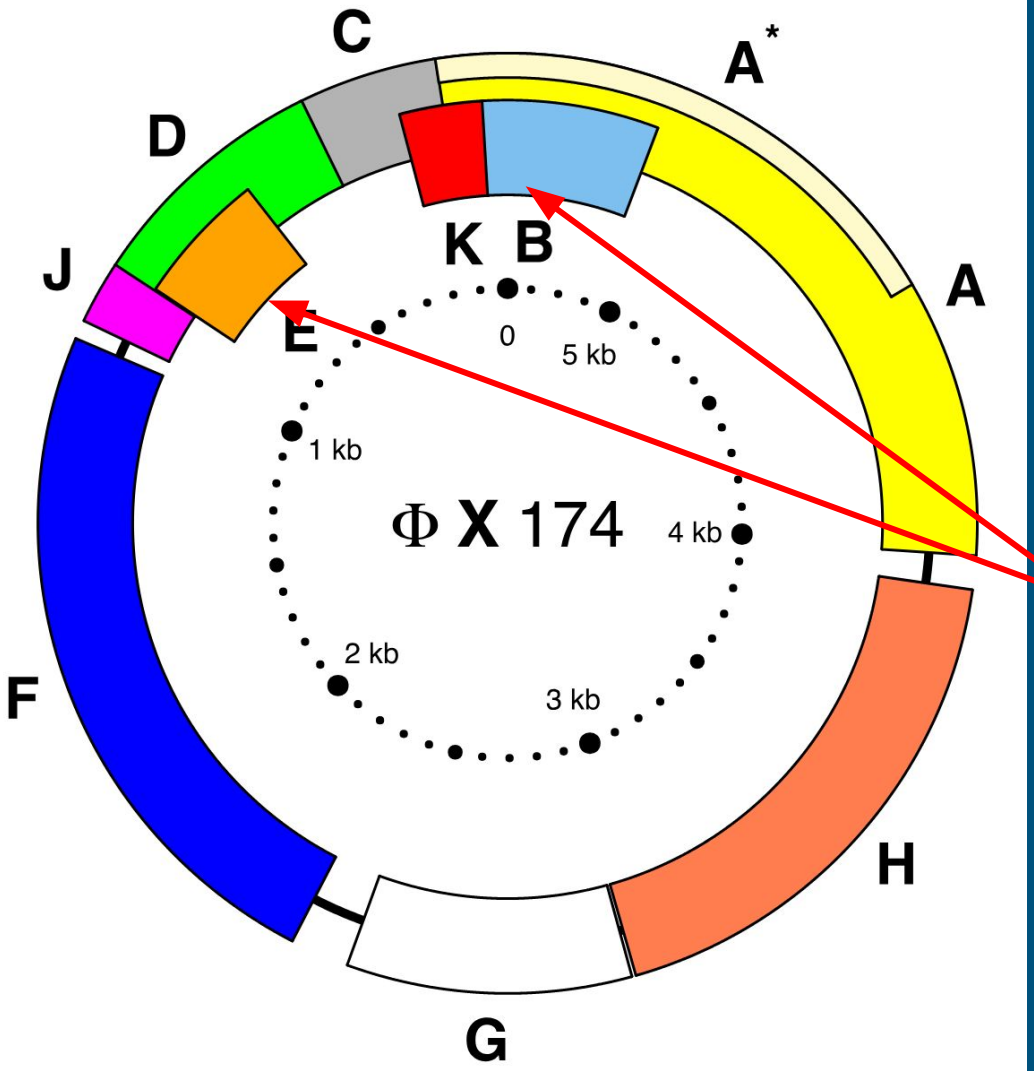
… it has 2334 amino acids!

?????

```
>lcl|NC_001422.1_cds_NP_040706.1_4 [protein=K]
ATG AGT CGA AAA ATT ATC TTG ATA AAG CAG GAA TTA CTA CTG CTT GTT TAC GAA TTA AAT CGA AGT GGA
CTG CTG GCG GAA AAT GAG AAA ATT CGA CCT ATC CTT GCG CAG CTC GAG AAG CTC TTA CTT TGC GAC CTT
TCG CCA TCA ACT AAC GAT TCT GTC AAA AAC TGA
```

0123456789

0123456789



OVERLAPPING GENES!!!

Picture © Emmanuel Douzery

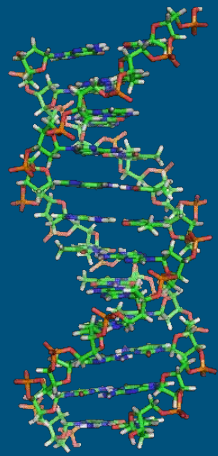
Even more DNA!!

Tomorrow, Sunday, 20:00, Andromeda (this) tent:

Reverse Engineering Life: What we can learn from the DNA

Monday, 15:00, Cassiopeia:

Afterparty for the talks, more Q&A and some bonus content.



A brief 2025 recap
DNA: the code of life