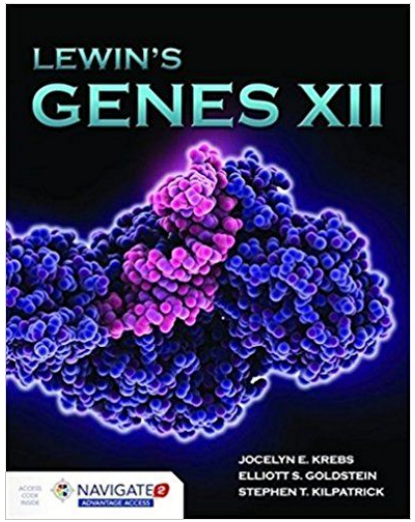
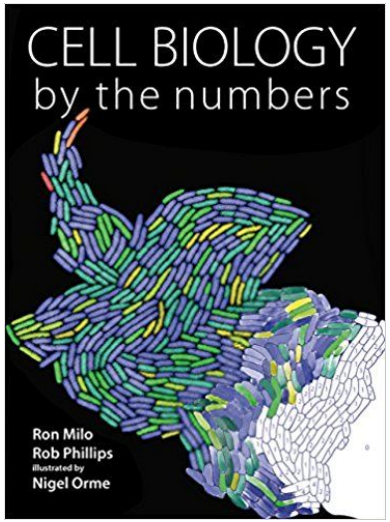
The cell as a computer

Turing complete and massively parallel.

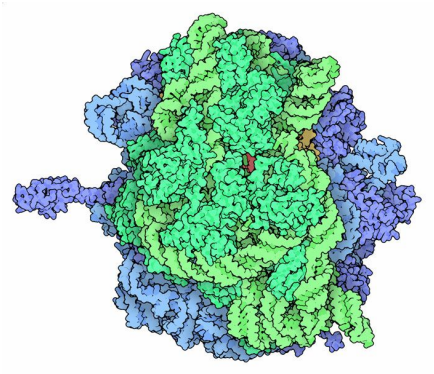
Reading list



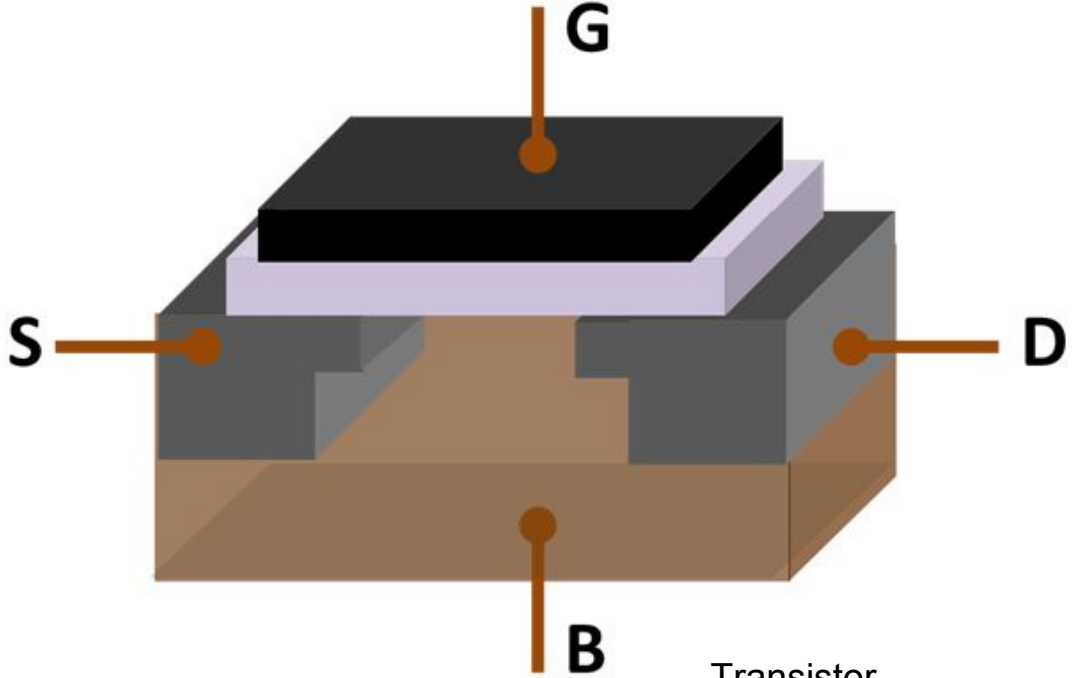


https://www.meetup.com/Oxford-Biohackers/
https://www.meetup.com/ACCU-Oxford/

Biology vs lithography



Ribosome (CPU)



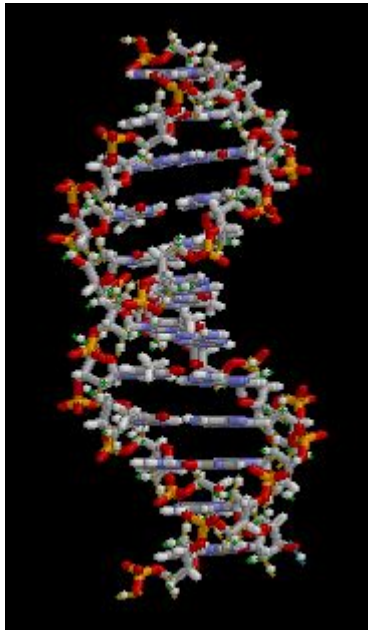
Transistor

http://pdb101.rcsb.org/motm/121

https://upload.wikimedia.org/wikipedia/commons/a/a5/MOSFET_Structure.png

Data storage

CD-ROM: 720MB



Human DNA: 700MB per cell

By Marus - Marus, Public Domain, https://commons.wikimedia.org/w/index.php?curid=1604269

DNA - read only (ish) memory.

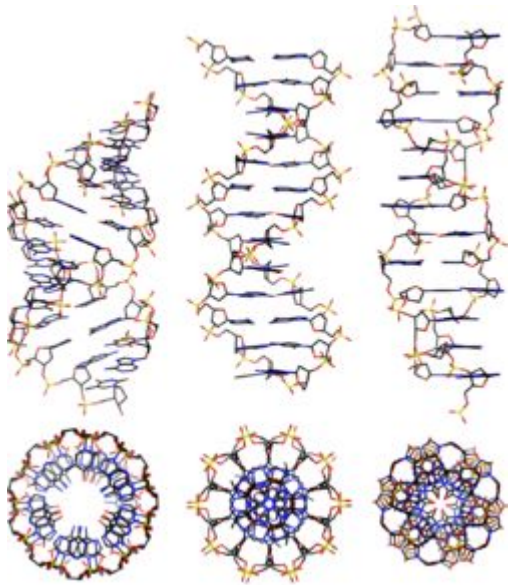
A C G T - 2 bits per base, 4 bases per byte.

Diameter of helix: 2nm

5 bases per nm.

Is modified by recombination, transposons and retroviruses (like HIV).

Encodes protein recipes and implements a complex program of transcription factor logic.



https://en.wikipedia.org/wiki/Nucleic_acid_double_helix

Data storage

How much data is in a pint of beer?

https://www.beeradvocate.com/community/threads/how-many-cells.77320/ http://www.broadberry.co.uk/petarack.php

Data storage

How much data is in a pint of beer?

Assume 40 billion yeast cells (Saccharomyces cerevisiae)

https://www.beeradvocate.com/community/threads/how-many-cells.77320/ http://www.broadberry.co.uk/petarack.php

Data storage

How much data is in a pint of beer?

Assume 40 billion yeast cells (Saccharomyces cerevisiae)

Each cell has 12 million bases = 3MB.

https://www.beeradvocate.com/community/threads/how-many-cells.77320/ http://www.broadberry.co.uk/petarack.php

Data storage

How much data is in a pint of beer?

Assume 40 billion yeast cells (Saccharomyces cerevisiae)

Each cell has 12 million bases = 3MB.

40 billion x 3MB = 120,000,000,000,000,000 bytes = 120 PB.

https://www.beeradvocate.com/community/threads/how-many-cells.77320/ http://www.broadberry.co.uk/petarack.php

Data storage

How much data is in a pint of beer?

Assume 40 billion yeast cells (Saccharomyces cerevisiae)

Each cell has 12 million bases = 3MB.

40 billion x 3MB = 120,000,000,000,000,000 bytes = 120 PB.

X 120

https://www.beeradvocate.com/community/threads/how-many-cells.77320/ http://www.broadberry.co.uk/petarack.php

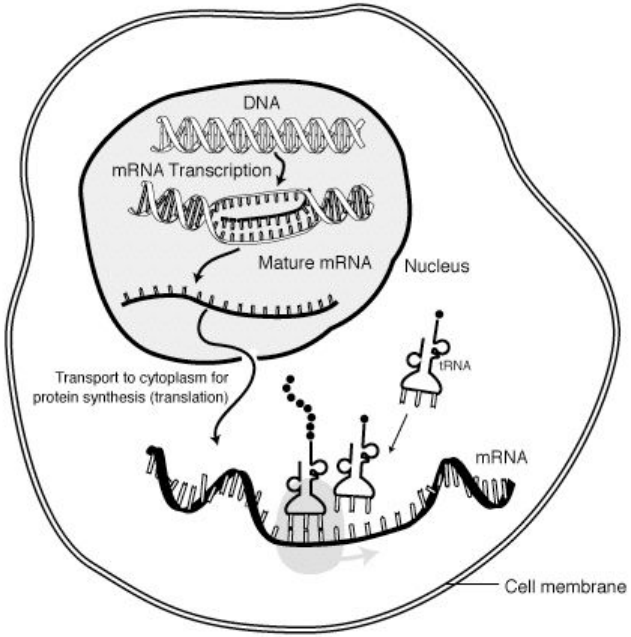
RAM = mRNA

In cells messenger RNA is used as RAM

It is initially copied from the DNA (ROM) but gets modified substantially before it is read by the ribosomes.

RNA is used to make proteins, to form the backbones of some structures (rRNA) and to implement logic in the nucleus (miRNAs).

Probably about half the dry mass of a cell is RNA.



Ribosomes - the high latency CPU

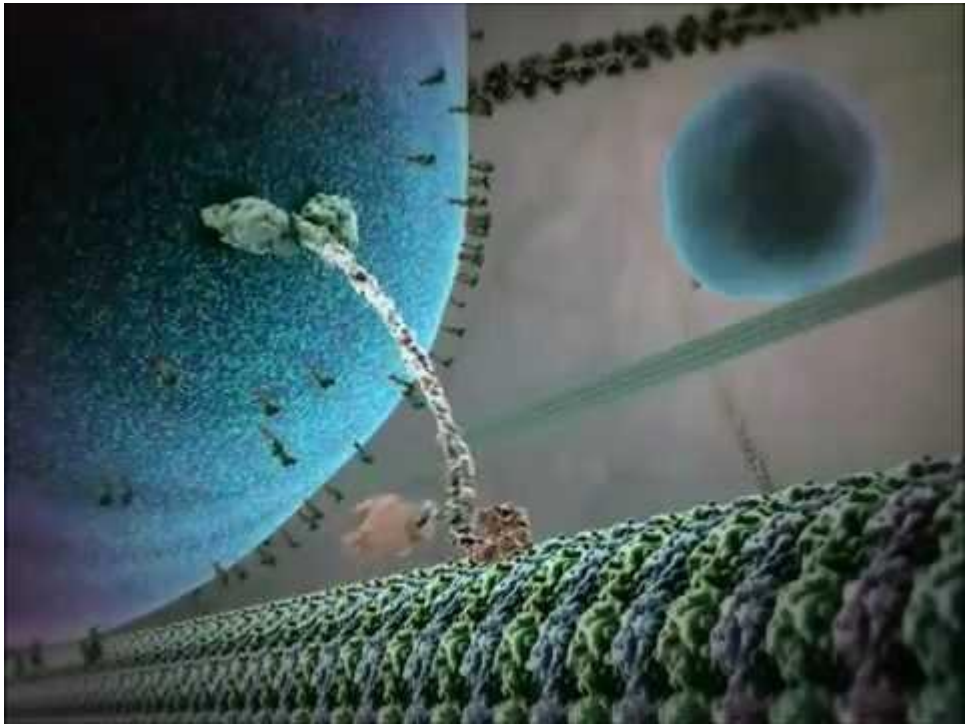
Ribosomes execute the DNA protein code - the instruction set of the cell.

Standard genetic code

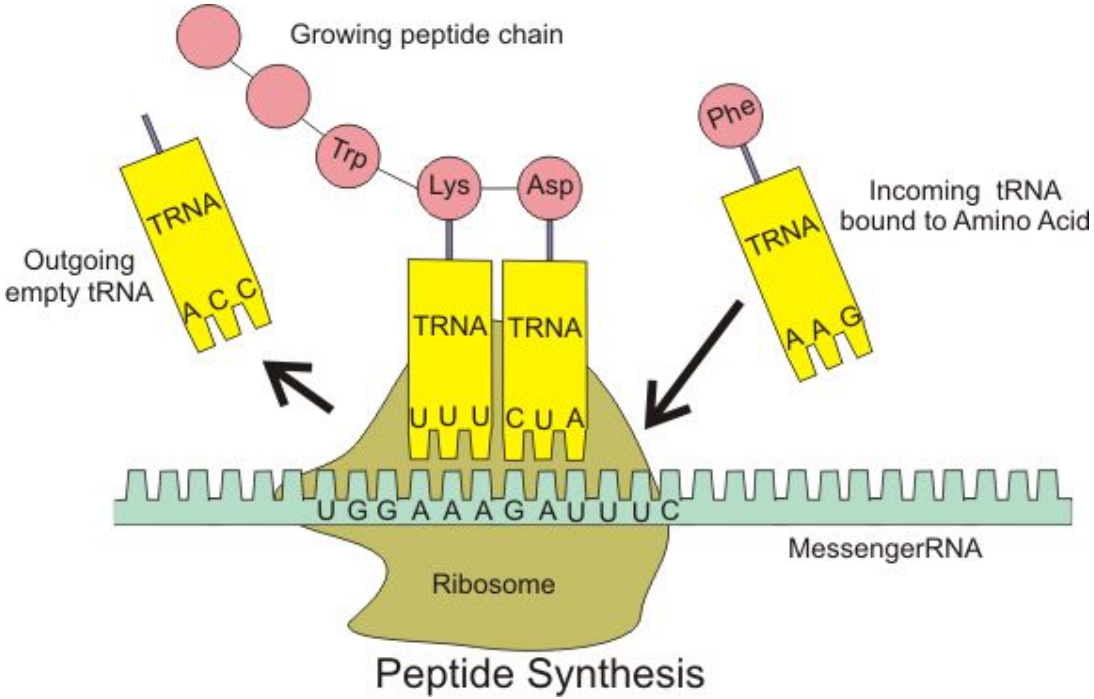
1st base	2nd base								3rd base
	U		C		A		G		
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	(Leu/L) Leucine	UCA		UAA[B]	Stop (*Ochre*)	UGA[B]	Stop (*Opal*)	A
	UUG		UCG		UAG[B]	Stop (*Amber*)	UGG	(Trp/W) Tryptophan	G
C	CUU		CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	(Gln/Q) Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine	A
	AUG[A]	(Met/M) Methionine	ACG		AAG		AGG		G
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	(Glu/E) Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G

https://en.wikipedia.org/wiki/Genetic_code

Example proteins



Ribosomes - the high latency CPU



About 10 bases per second per ribosome.

https://en.wikipedia.org/wiki/Ribosome#/media/File:Peptide_syn.png

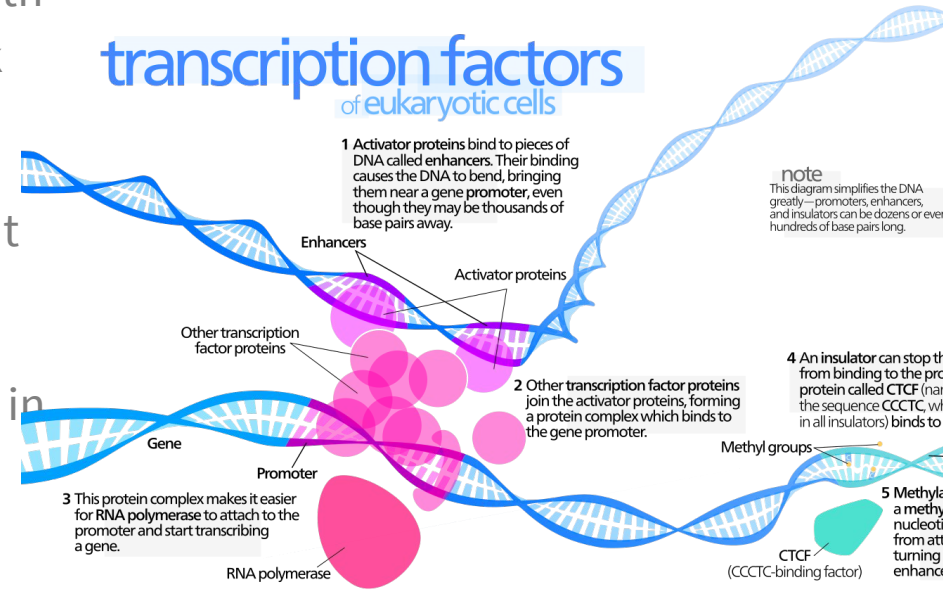
Ribosomes - the high latency CPU

Proteins generated by the ribosomes bind with the DNA to create low performance feedback loops.

Transcription factors are special proteins that bind the DNA and form feedback loops.

Transcription factors determine the cell type in humans - ie. eye vs. skin.

Transcription factors form complex groups implementing logic - and, or, not.

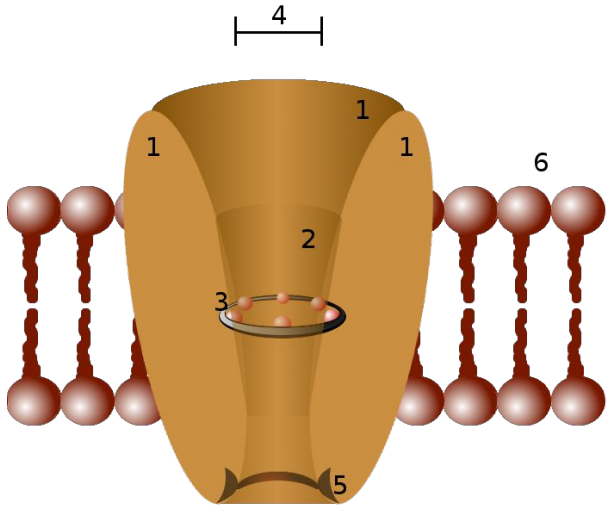


Low latency logic - Ion channels

High performance logic is implemented by modification of proteins and by ion channels

Ion channels transport ions (Calcium, Sodium etc.) very quickly and can switch extremely fast.

Nerve cells use ion channels to transmit electrical signals.



Low latency logic - proteins

High performance logic is implemented by modification of proteins and by ion channels

Phosphorylation, Methylation and acetylation and other processes modify proteins.

Modified proteins form networks in cells to carry signals very quickly through the cell.

Protein A -> protein B -> protein C

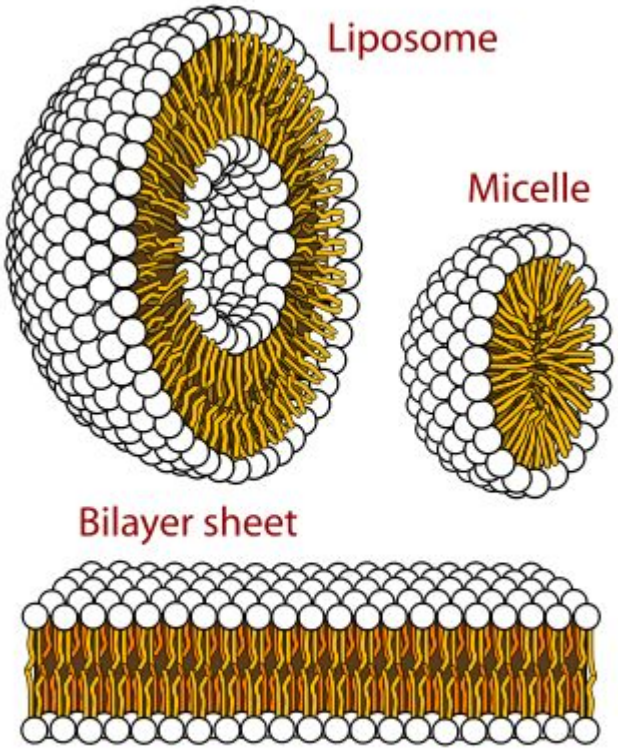
Cells can respond very quickly to external signals.

Lipids - The circuit board

Lipid layers (soap bubbles) form an insulating layer for the placement of channels and receptors.

Like a circuit board they both insulate and form a mechanical placement for components.

Lipid layers hold about 50% proteins.

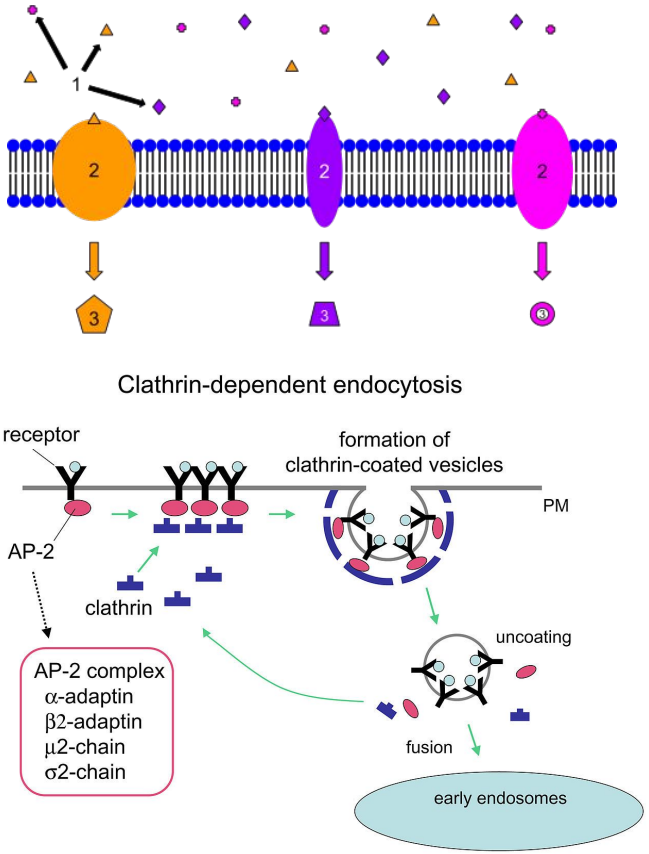


Receptors - the network

Cells communicate by releasing signals (such as neurotransmitters) that bind with receptors on the surface of other cells.

Bacteria exchange DNA information via plasmids.

Clathrin pinches off parts of the membrane to import signals and chemicals.



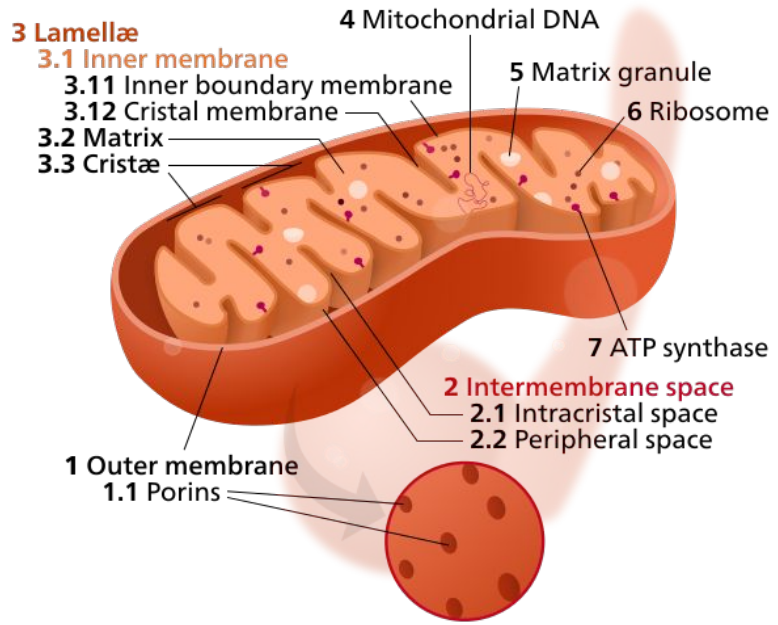
Mitochondria - the power supply

Electrical energy in cells comes from phosphate ions.

Phosphates are transported by ATP.

The Mitochondria use ATP synthase to convert ADP to ATP, adding a phosphate.

RNA and DNA also have phosphate backbones which are used to drive the transcription process - ie. contain their own batteries.



Mitochondria - the power supply

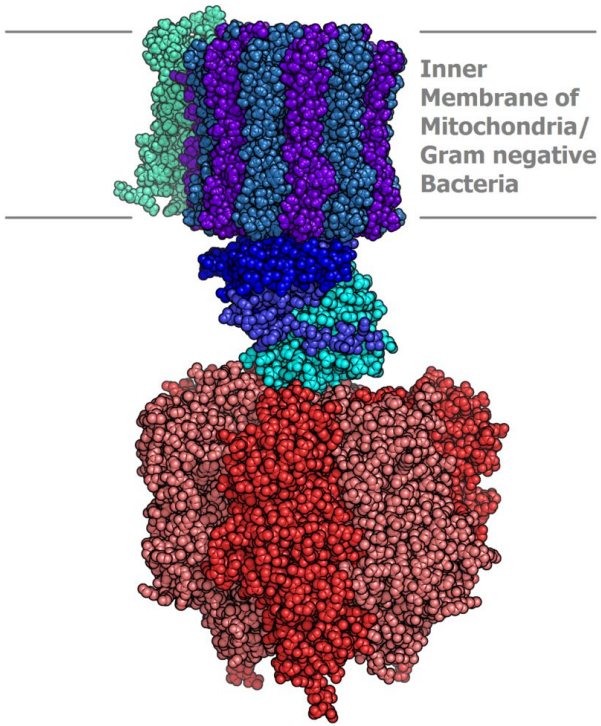
ATP synthase in mitochondria turn electrical energy (protons) into phosphates.

The protons come from glucose breakdown - each stage makes one proton.

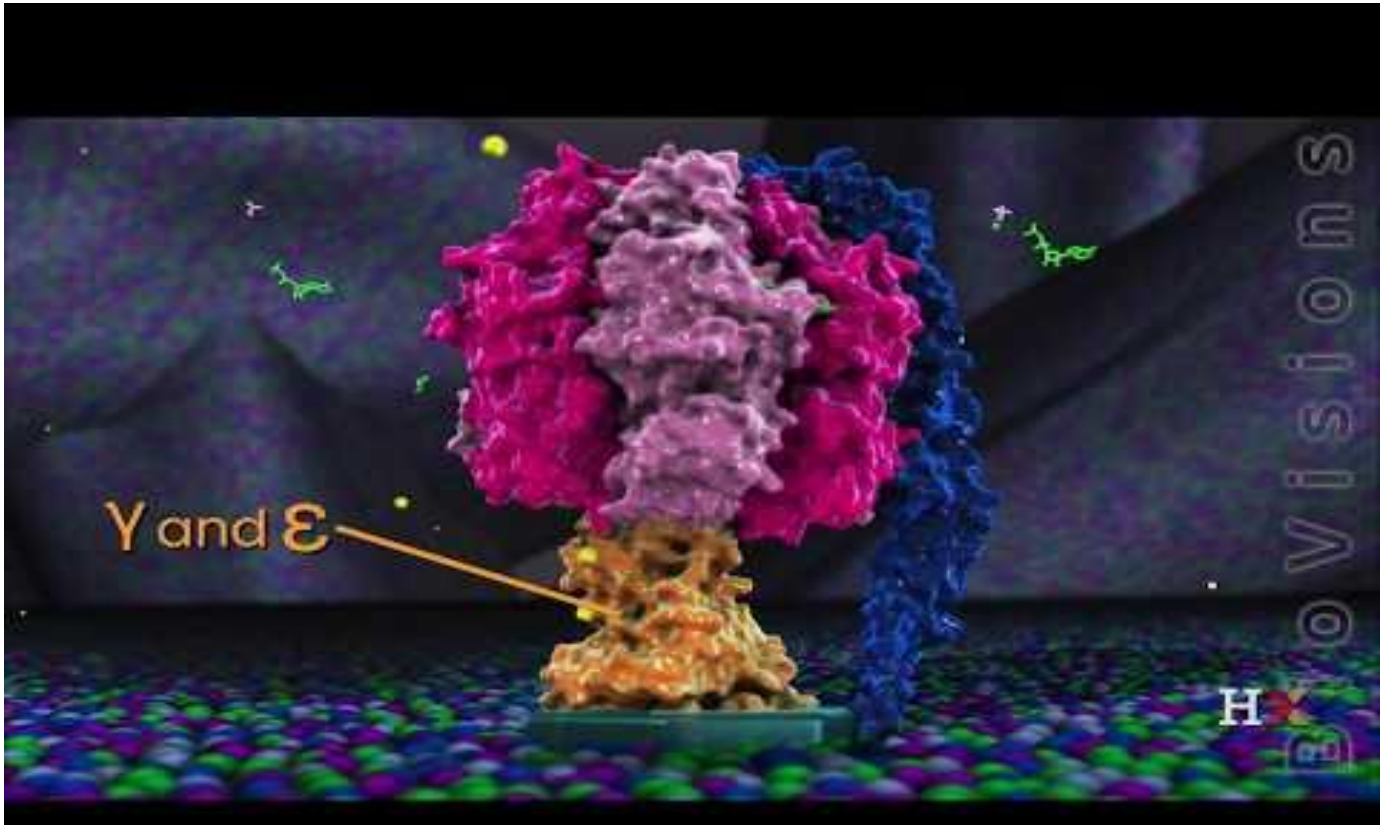
ATP synthase is an electric motor plus a chemical engine (ie. like a generator).

It can work in reverse also, generating electricity.

https://gfycat.com/gifs/detail/MildImaginaryCougar



ATP Synthase



The future of computing

https://en.wikipedia.org/wiki/Biological_computing

Using biological components

Creating structures from DNA, transcription factors and ion channels.

DNA very cheap to “print”.

Synthetic biology can create components outside cells.

Millions of tons of enzymes are produced this way every year (washing powder).

Oxford nanopore - use proteins in membranes to read DNA.

Which is smarter? Your beer or your smartphone?

Beer	Smartphone
120PB DNA 100PB mRNA 500 Quadrillion ribosomes 	32GB Flash 1GB RAM 8 cores 

Computers in numbers

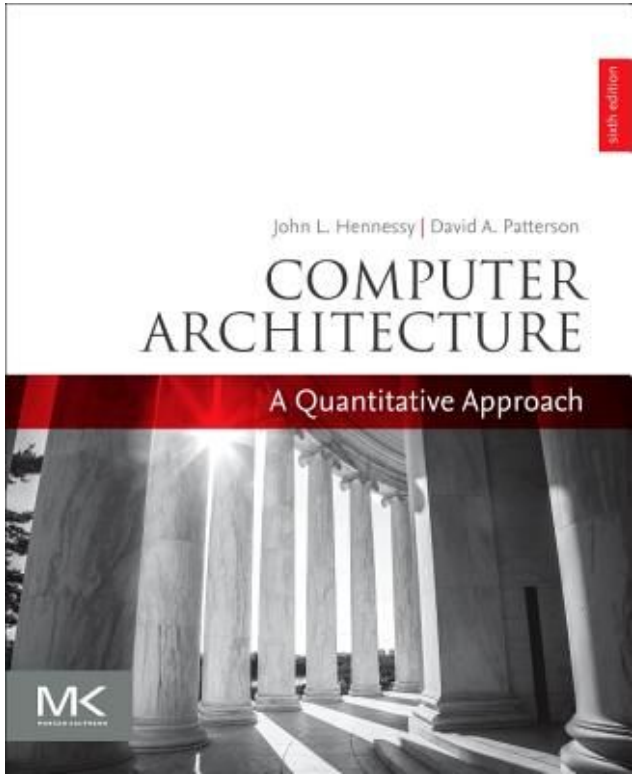
Writing compilers requires intricately detailed knowledge of computer systems.

Run endless stats.

Generate random programmes and time them hot and cold.

Map pipelines and latencies for different CPU architectures.

Ignore the manuals!



Biology in numbers

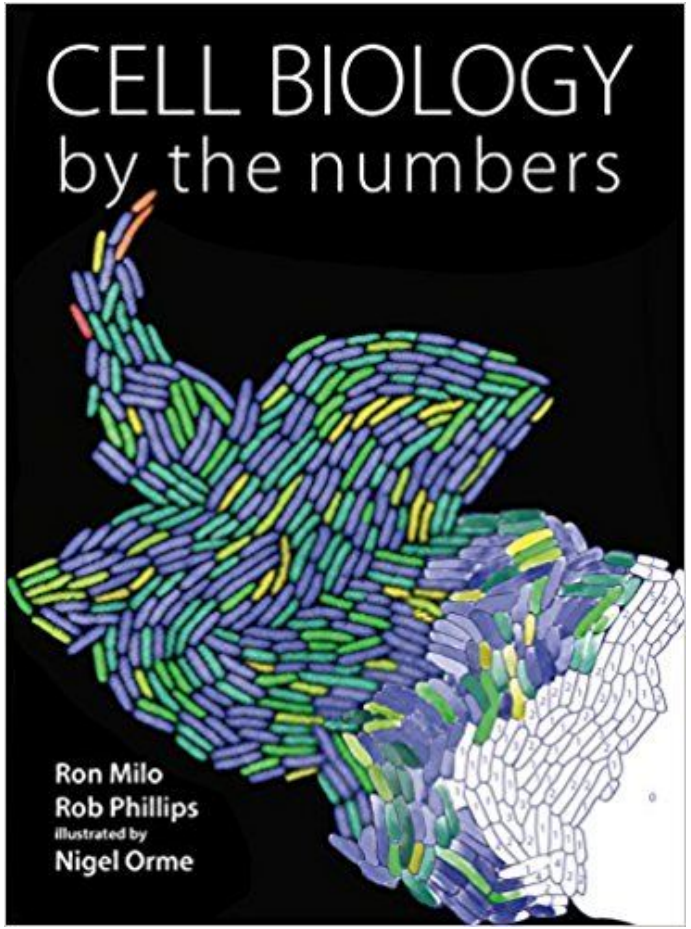
Bioinfomatics requires intricately detailed knowledge of computer systems.

Run endless stats.

Use random humans and measure them hot and cold.

Build pathway maps and models for many different cell types.

Ignore the manuals - there are none!



Quantum computing vs Biocomputing

Quantum computing has same timeline as fusion power.

Biocomputing can be done with existing technology.

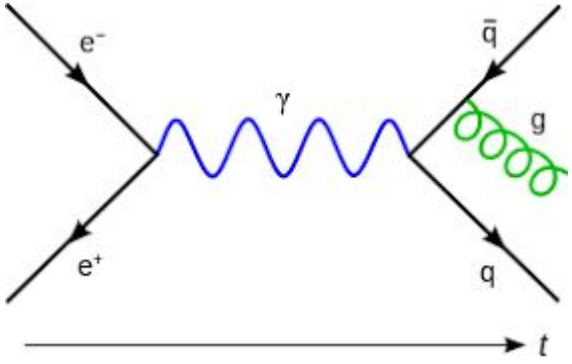
We already have DNA printed on chips (array sequencing).

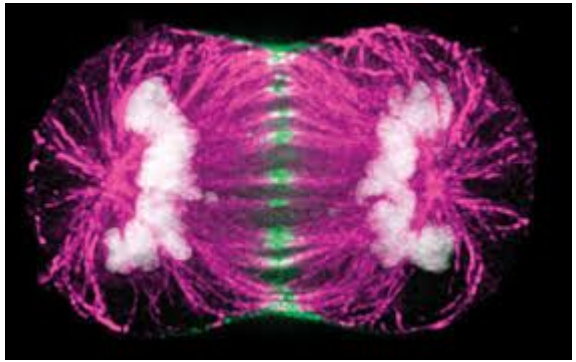
We will be able to design proteins for any problem.

We already use biological components everywhere.

Biology is *very very* 3D with 10^{26} components in a PC case.

Biology is low power.





The first hybrid silicon-DNA technology

https://en.wikipedia.org/wiki/DNA_microarray

How to get your DNA sequenced

Many options

- Microarrays - cheap as chips at ~€20 a go ~ 6M variants.

  Ancestry.com, 23 and me etc.

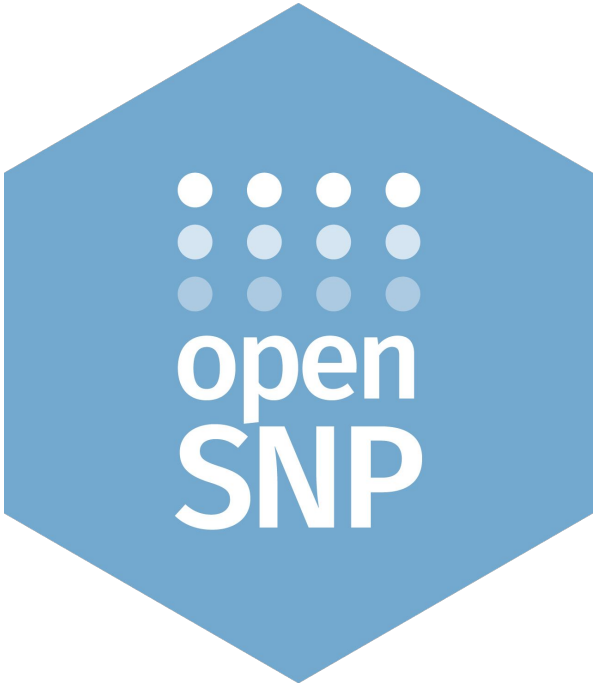
- Full next generation sequencing at ~€500 a go ~3G base pairs.
- Oxford nanopore - squeeze DNA through a pore, measure the current.
- PCR - quick and inexpensive - single sites, viruses, ancient DNA etc.

Share your DNA

Download your VCF file

Share on OpenSNP

https://opensnp.org/



Biology is not just about health

Because of the large rewards, most biotech firms make drugs and diagnoses.

Nanotechnology with DNA-based components on the rise.

Molecule-sized motors, pumps and sensors are possible.

Still looking for the perfect circuit board.

Invest now while it is still young.

And - oh, you need a bit of software to do this!