"Nothing in biology makes sense, except in the light of evolution."

Theodosius Dobzhansky 1900-1975

Overview

Biology as computation

An introduction to population genetics

Questions about evolution & integration

What will people say

What do cells do?

Harvest energy.

Self organize: isotropy -> anisotropy -> movement

Depend on history: state1 -> state 2 (-> state 1)

Call and Respond (internally and externally)

Integrate

Duplicate

Evolve

- Inherit: replicate and decode information
- Mutate: copying mistakes
- Select (on function)

Biology As Computation

Computation: rule based symbol transformation

2.15.15.27.25.1.14.11.5.5.19.28.27.25.5.25.27.18.5.4.27.19.15.24

Boo Yankees, Yay Red Sox

Most biological activities are computations

Evolution selects for :
Computations that increase fitness in current & future environments

Similarities to computer engineering

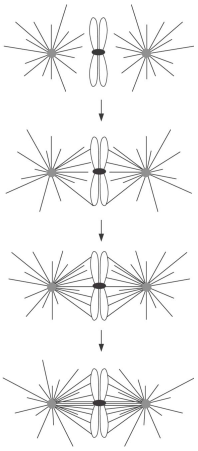
What is understanding?

Explanations that satisfy curiosity

Hypotheses that make accurate predictions

Discovering general principles (computational rules)

Exploration with selection as a design principle



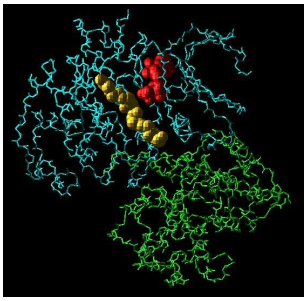
Order created by sequentially stabilizing a subset of random excursions

Scaleable

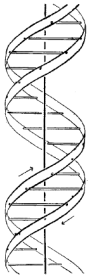
- Protein folding
- Spindle assembly
- Nervous system development
- Evolution
- Ecosystems?

What do we understand?

Biological chemistry



Nat Cell Biol **1**, 438-443



Genetic information transmission

Nature, **171,** 737-738 (1953)

We don't understand how cells...

Harvest energy

Self organize: isotropy -> anisotropy -> movement

Depend on history: state1 -> state 2 (-> state 1)

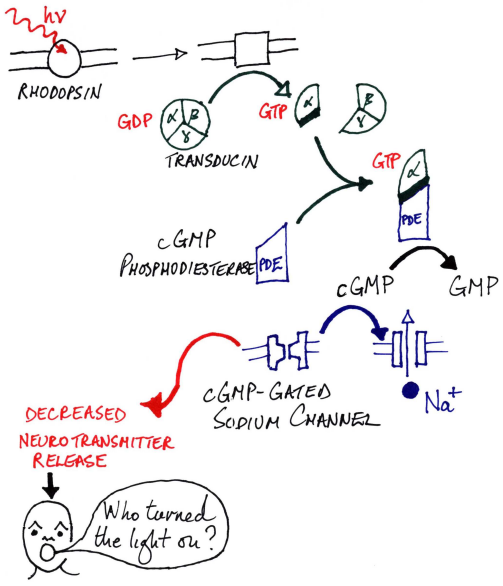
Call and Respond (internally and externally)

Integrate

Duplicate

Evolve

Signal transduction



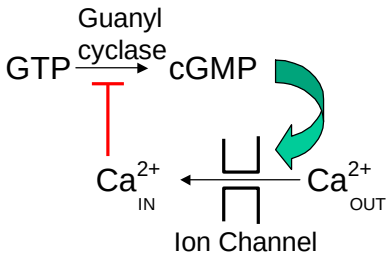
The systems view of vision

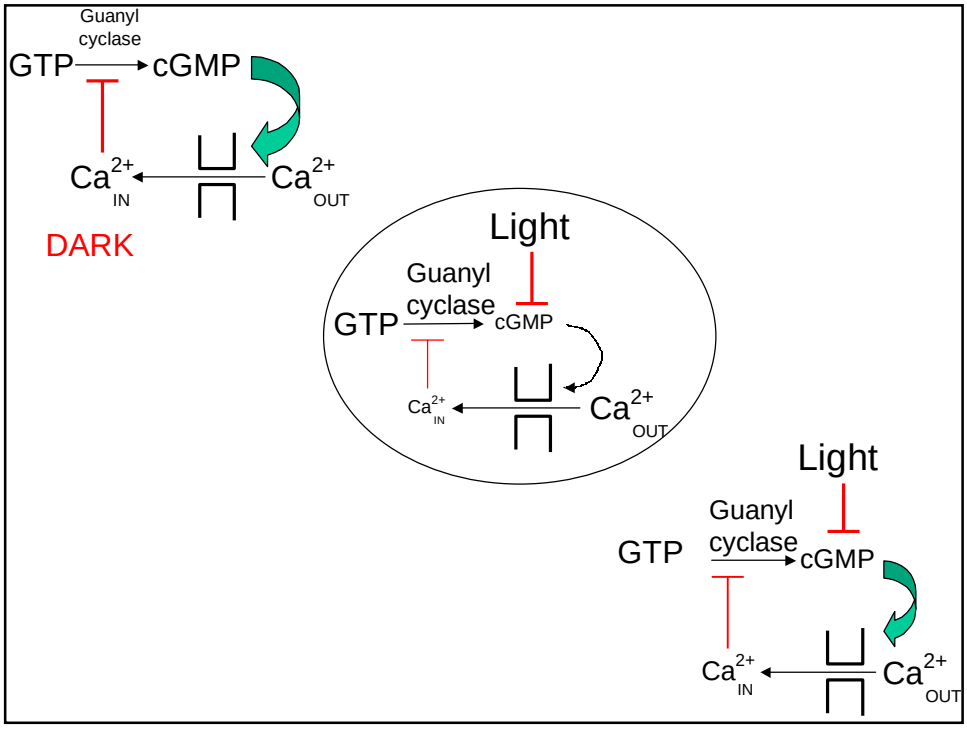
Adaptation

Total range: fastball in bright sunlight **to** candle at 5 miles

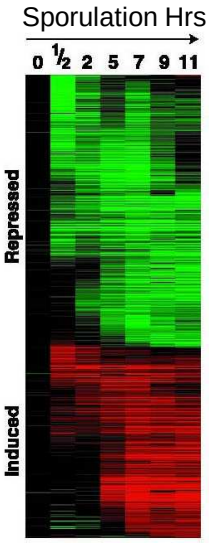
Detection range: 10^{10} **to** 1 Photons/cell/sec

Instantaneous dynamic range: 100 fold (locally adjustable)





Integration: How to make safe state changes?



10 parameters/protein (conc., location, partners,..)

6000 genes = 60000 parameters

Cell physiology = a point in 60000 dimensional space

1000 genes change = 10000 parameters change

How do cells find safe paths between states differing by so many parameters

Chu et al Science **282** 699

Population Genetics I

A statistical description of idealized creatures and genes

Virtue: mathematics forces statement of assumptions

Vice: assumed $\neq$ real world

Deals best with changes in fraction of existing mutations

Incorporates new point mutations: better, worse, neutral

Population Genetics II

Assumes:

- All mutations are point mutations
- Effects of mutations are small

Problems:

- Big mutations: gene creation, fusion, deletion, duplication
- Mutations of large effect

Population Genetics III

At each gene there can be many different forms (alleles), 1,2,3,...i

f_i = allele fraction (frequency), $\Sigma f_i = 1$

Fixation occurs when only one allele is found of a gene, $f_i \approx 1$

N = Population size

N_e = Effective population size $\quad \frac{1}{N_e} = \frac{1}{G}\sum_{1}^{G}\frac{1}{N_1} + \frac{1}{N_2} + \frac{1}{N_3} ..... + \frac{1}{N_G}$

Human: $N \approx 10^{10}$, $N_e \approx 10^4$

μ = Mutation frequency (rate)

Population Genetics IV

Fitness, w_i, is the number of viable offspring produced

Average fitness, $\overline{w}$, set to 1

s = selective coeffiecient

$w_i = 1 + s_i$

Epistasis refers to the interaction between mutations

If $w_1, w_2 < 1$

Synergistic Epistasis: $w_{12} < w_1 \times w_2$

Diminishing Returns Epistasis: $w_{12} > w_1 \times w_2$

In general, distribution of fitnesses and epistasis unknown

Population Genetics V

Chance of advantageous mutation fixing $\approx 1/s$

Time (generations) to fixation of advantageous mutation $\approx \ln N_e/s$

Rate at which advantageous mutations are fixed $\leq N_e\mu$

Chance of neutral mutation fixing $\approx 1/N_e$

Mean time to fixation of neutral mutation = N_e

Mean time to loss of neutral mutation = $\ln N_e$

Rate at which neutral mutations are fixed $\approx \mu$

Q1: Reproduction vs. Survival

All living things terminus of 4 Byr chain of being

Good times = selection on reproductive rate

$$\frac{f_1}{f_2} = \left(\frac{w_1}{w_2}\right)^G$$

If $w_1 = 1.1 \times w_2$, $N_e = 10^6$, 145 generations eliminate allele 2

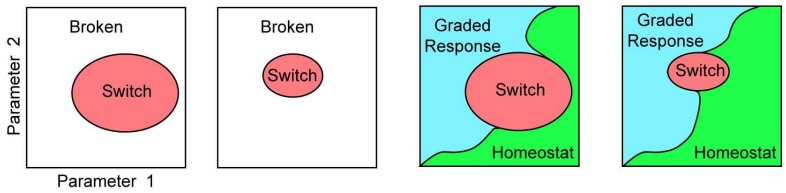
145 generations = 10^{43} progeny/founder!

Bad times = selection on scavenging and survival

Q2: Robustness vs. Evolvability

Robustness = Resistance to perturbation (genetic or environmental)

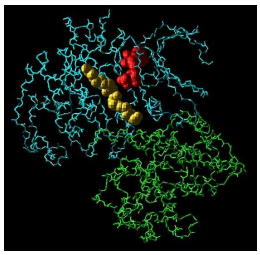
Evolvability = Ability to change function

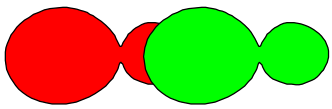


Robust Unevolvable | Frail Unevolvable | Robust Evolvable | Frail Evolvable

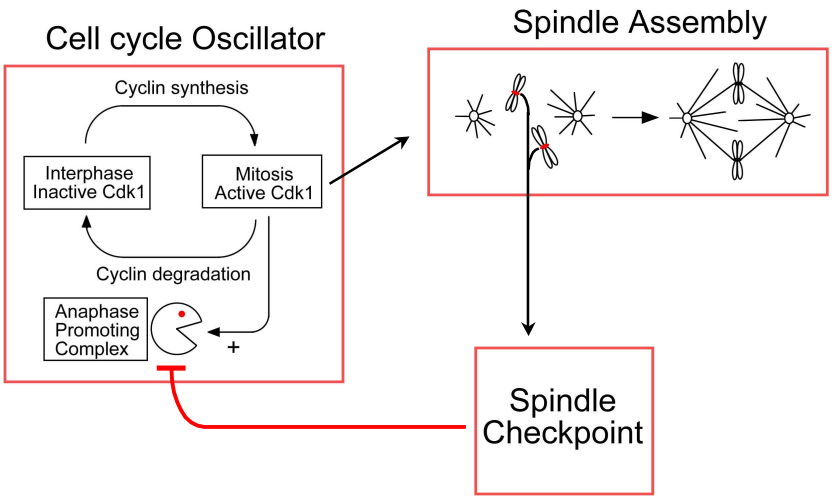
Q3: What are the units of biological function?

Protein properties → ? → Reproductive fitness





Modules as units of biological function



(Hartwell et al. *Nature* **402**, C47-52 (1999))

Modular predictions for synthetic lethality

Definition: A^+b^- and a^-B^+ alive, but a^-b^- dead

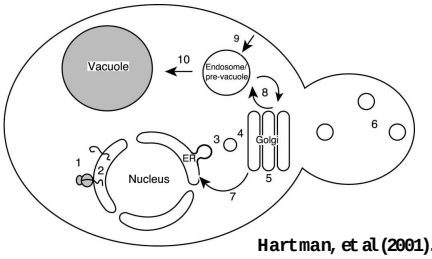
Fitness (w) is quantized to two values 1 (alive), 0 (dead)

Prediction

A, B in different modules: $w_{AB} = w_{aB} = w_{Ab} = w_{ab} = 1$

A, B in same module: $w_{AB} = w_{aB} = w_{Ab} = 1$, $w_{ab} = 0$

Synthetic Lethality in Yeast Secretion



Hartman, et al (2001). Science *291*, 1001-4.

			Number of synthetic lethal interactions		
Step	Function	Number of genes	Same step	Different step	Not secretion
1	Translocation to Golgi	23	6	0	4
2	Maturation in ER	16	10	0	2
3	Vesicle budding	12	11	3	2
4	Vesicle fusion	18	11	13	1
5	Transport to Golgi	11	2	9	3
6	Fusion w/ plasma membrane	30	43	13	12
7	Retrieval	20	15	9	1
8	Vacuolar targeting	16	6	6	4
9	Endocytosis	20	12	12	22
10	Endosome to vacuole	7	0	3	2
Total		173	116	68	53

Fraction of non-essential genes in secretion module = 173/4800 = 0.04

Fraction of synthetic lethals in secretion module = 184/227 = 0.78

What might modules be good for?

Framing & answering questions about integration

Escaping from molecular tyranny

Evolutionary issues

Independence: altering one function w/o disaster elsewhere

Novelty by making and breaking inter-module connections

Q4: Does population genetics describe evolution?

Minor errors vs. entirely wrong picture?

Are adaptive landscapes real? 

How context (environment, genotype)-depedent is fitness

Q5: How common is evolve or die?

Evidence for long-running host-parasite arms races

The thought experiment:

In an assembly of species freeze the gene pool of one

How fast does this species go extinct?

Are there stable, slowly evolving ecosystems?

Q6: Can organisms "control" evolution?

Mutation Rate: d(DNA)/dt
- Heritable Mutators
- Transient changes in mutation frequency

Phenotypic change per mutation: dϕ/d(DNA)
- Local buffering via overlapping genes
- Global buffering via chaperones (Hsp90)
- Prion based protein extensions

What is the role of pre-existing vs novel variation in evolution
- Initial response dominated by existing variation
- Long term response dominated by novel variation

Q7: How do complicated pathways evolve?

In principle, tinkering makes the watchmaker who makes God

In practice
- How much invention versus recycling?
- How well does gradualism explain metabolic pathways?
- How do developmental pathways evolve?