Computational Insights and the Theory of Evolution

Christos H. Papadimitriou
UC Berkeley

Evolution Before Darwin

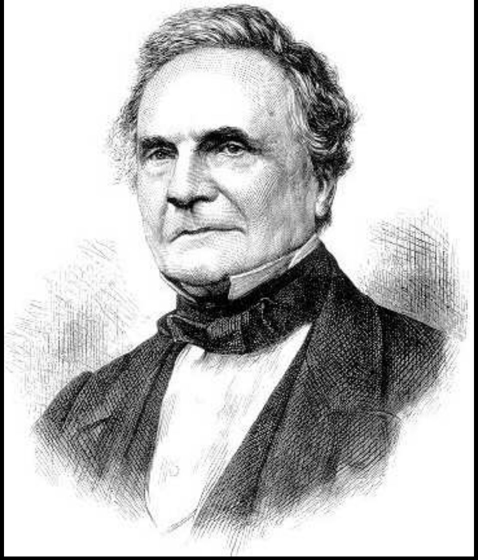
- Erasmus Darwin

Before Darwin

- J.-B. Lamarck

Before Darwin

- Charles Babbage



[Paraphrased]

"God created not species, but the Algorithm for creating species"

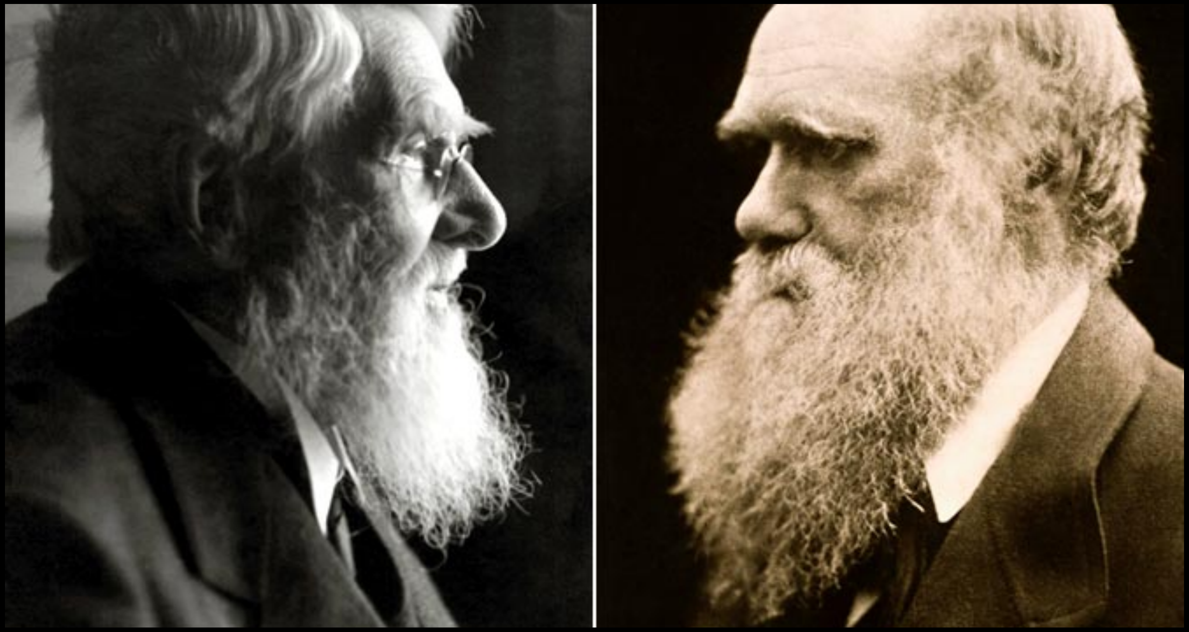
Darwin, 1858

- Common Ancestry
- Natural Selection

The Origin of Species

- Possibly the world's most masterfully compelling scientific argument
- The six editions 1859, 1860, 1861, 1866, 1869, 1872

The Wallace-Darwin papers



Brilliant argument, and yet many questions left unasked, e.g.:

- How does novelty arise?
- What is the role of sex?

After Darwin

- A. Weismann

[Paraphrased]

"The mapping from genotype to phenotype is one-way"

Genetics

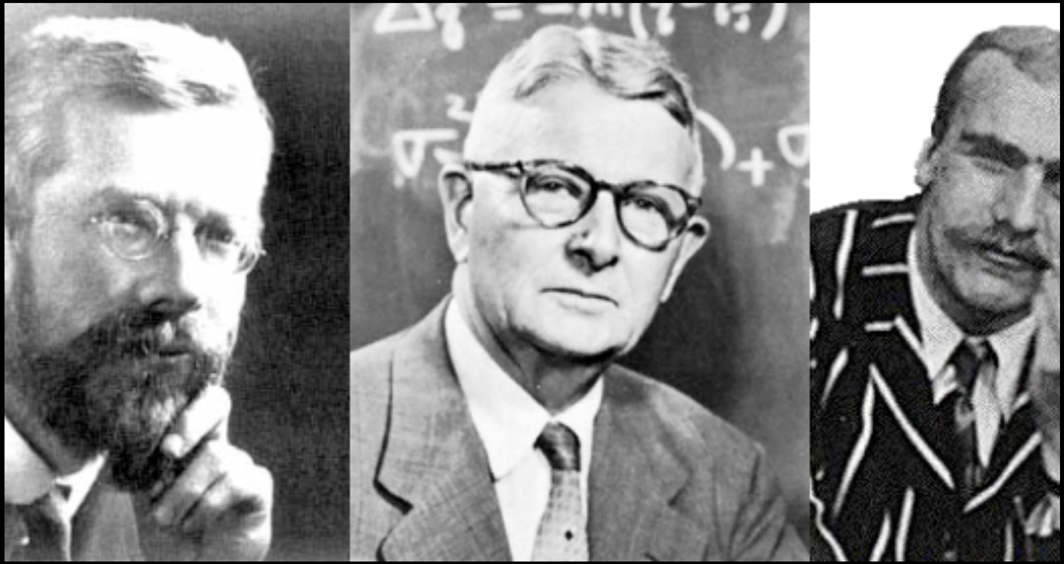
- Gregor Mendel [1866]
- Number of citations
  between 1866 and 1901:

3

The crisis in Evolution 1900 - 1920

- Mendelians vs. Darwinians
- Geneticists vs. Biometricists/ Gradualists
- Population genetics

The "Modern Synthesis"
1920 - 1950



Fisher – Wright - Haldane

Big questions remain e.g.:

- How does novelty arise?
- What is the role of sex?

Evolution and Computer Science

- *How do you find a 3-billion long string in 3 billion years?"*

L. G.Valiant

At the Wistar conference (1967), Schutzenberger asked virtually the same question

Valiant’s Theory of the Evolvable

- Which functions (traits of an organism) can evolve by natural selection?
- Properly formalized, this question leads to identifying obstacles to evolution
- For example, the function has to be learnable (actually, statistically so)
- Evolvability is a (quite) restricted form of learnability

Evolution and CS Practice: Genetic Algorithms [ca. 1980s]

- To solve an optimization problem…
- …create a population of solutions/genotypes
- …who procreate through sex/genotype recombination…
- …with success proportional to their objective function value
- Eventually, some very good solutions are bound to arise in the soup…

And in this Corner…
Simulated Annealing

- Inspired by *asexual* reproduction
- Mutations are adopted with probability increasing with fitness/objective differential
- …(and decreasing with time)

The Mystery of Sex Deepens

- Simulated annealing (asexual reproduction) works fine
- Genetic algorithms (sexual reproduction) don’t work
- In Nature, the opposite happens: Sex is successful and ubiquitous

?

A Radical Thought

- What if sex is a mediocre optimizer of fitness (= expectation of offspring)?
- What if sex optimizes something else?
- And what if this something else is its *raison d' être?*

Mixability!

- In a recent paper [LPDF, PNAS 2008] we establish through simulations that:
- Natural selection under asex optimizes fitness
- But under sex it optimizes ***mixability:***
- The ability of alleles (gene variants) to perform well with a broad spectrum of other alleles

Explaining Mixability

- Fitness landscape of a 2-gene organism

3	2	4	5	4
1	0	0	7	2
2	1	0	4	3
1	8	1	3	2

Rows: alleles of gene A

Entries: fitness of the combination

Columns: alleles of gene B

Explaining Mixability (cont)

- Asex will select the largest numbers

3	2	4	5	4
1	0	0	7	2
2	1	0	4	3
1	8	1	3	2

Explaining Mixability (cont)

- But sex will select the rows and columns with the largest average

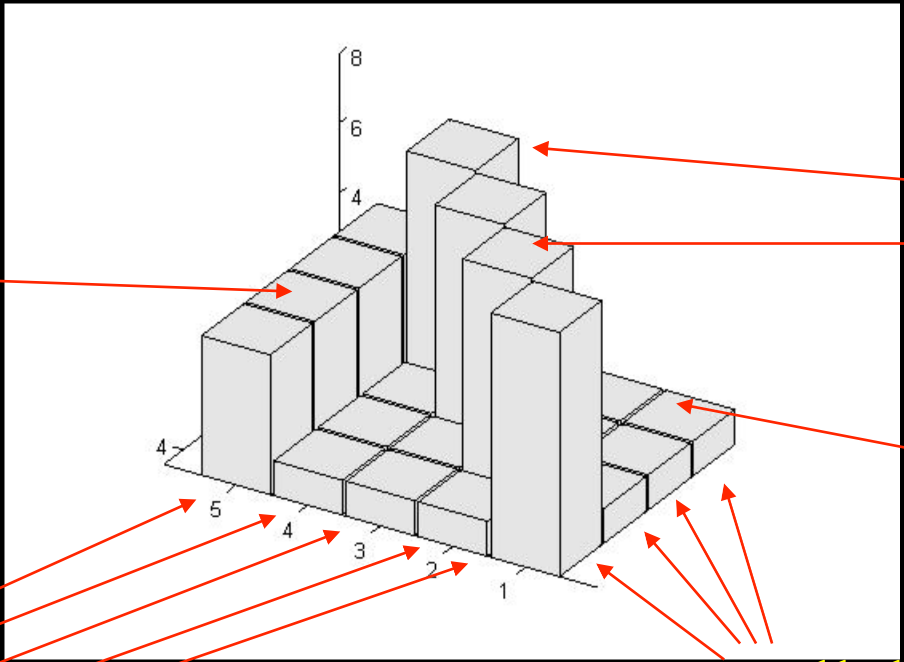
3	2	4	5	4
1	0	0	7	2
2	1	0	4	3
1	8	1	3	2

In Pictures

“plateau”

peaks

troughs



alleles (variants) of gene A

alleles of gene B

Sex favors plateaus over peaks

Theorem [Livnat, P., Feldman 11] In landscapes of this form

- Unless peak > 2 × plateau, in sexual reproduction the plateau will dominate and the peaks will become extinct
- In asexual reproduction, the peaks will always dominate and the plateau will become extinct

And plateaus accelerate evolution

- They act as springboards allowing alternatives to be explored *in parallel*…
- …and this acceleration promotes *speciation* (the creation of new species)…
- …which results in an altered landscape…
- …in which sex selects more plateaus…
- …and life goes on…

Pointer Dogs

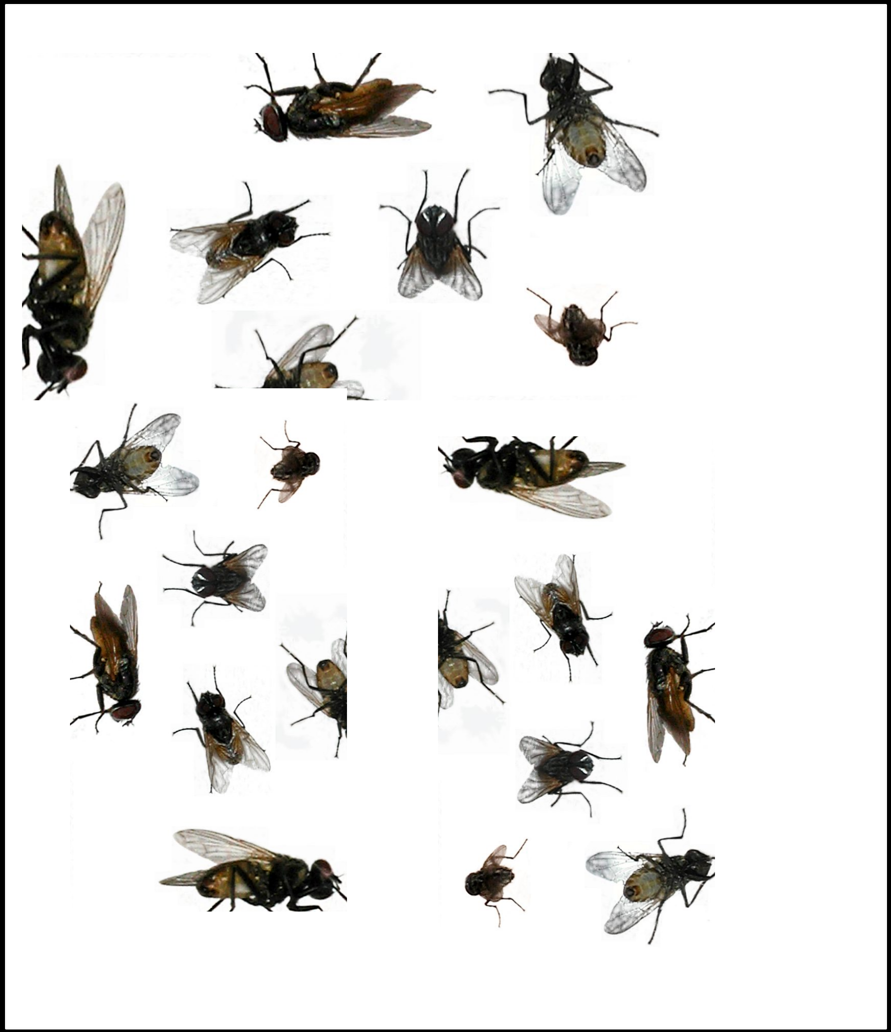
Pointer Dogs

C. H. Waddington

Waddington's Experiment (1952)

Generation 1

Temp: 20° C



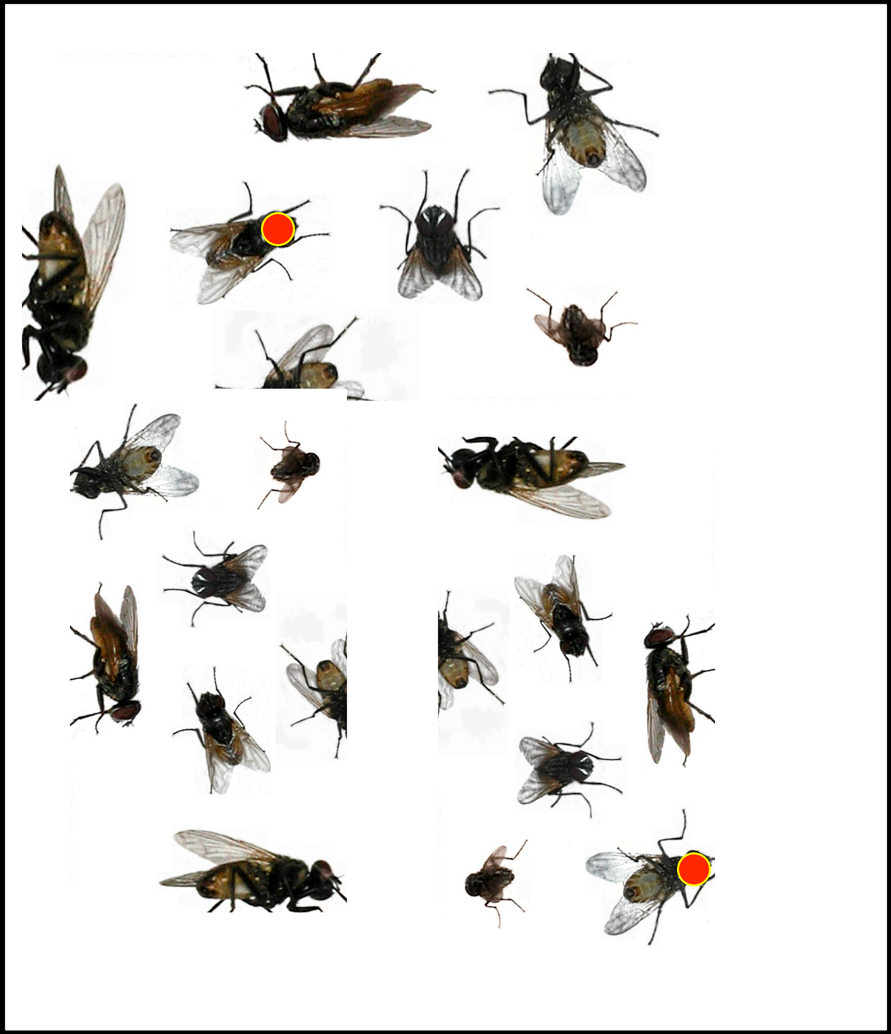
Waddington's Experiment (1952)

Generation 2-4

Temp: 40^o C

~15% changed

Select and breed those



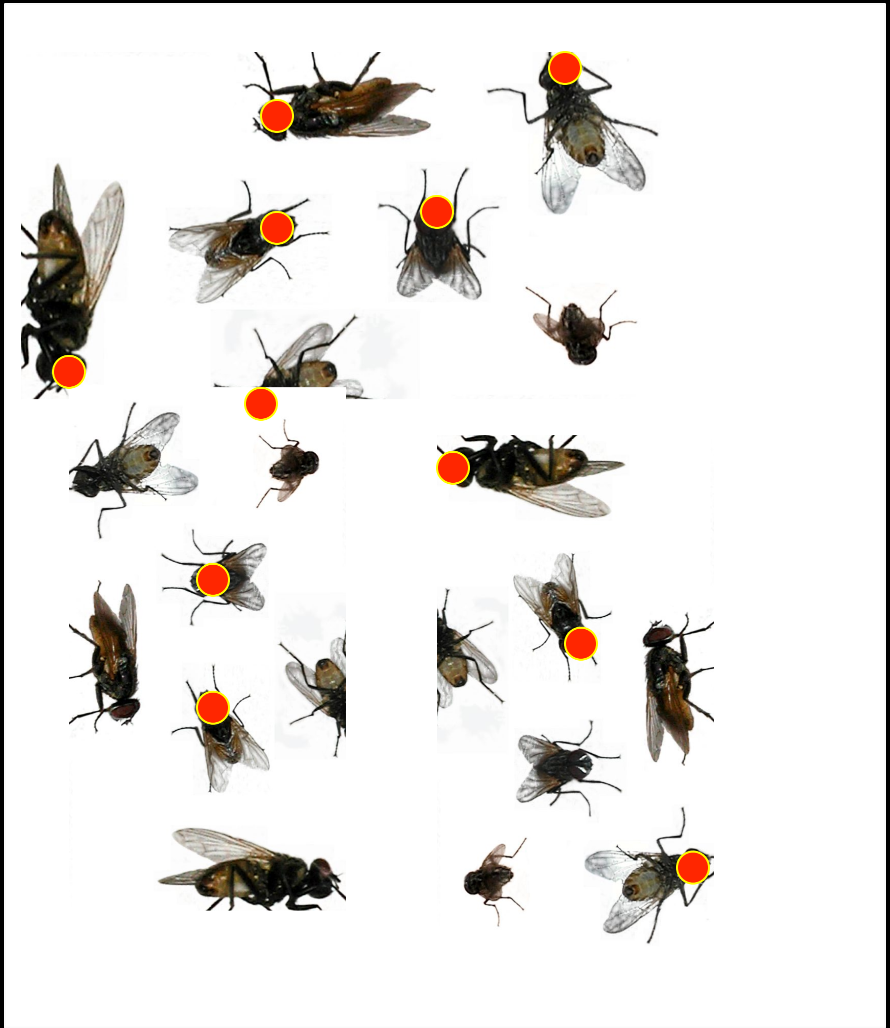
Waddington's Experiment (1952)

Generation 5

Temp: 40^{o} C

~60% changed

Select and breed those



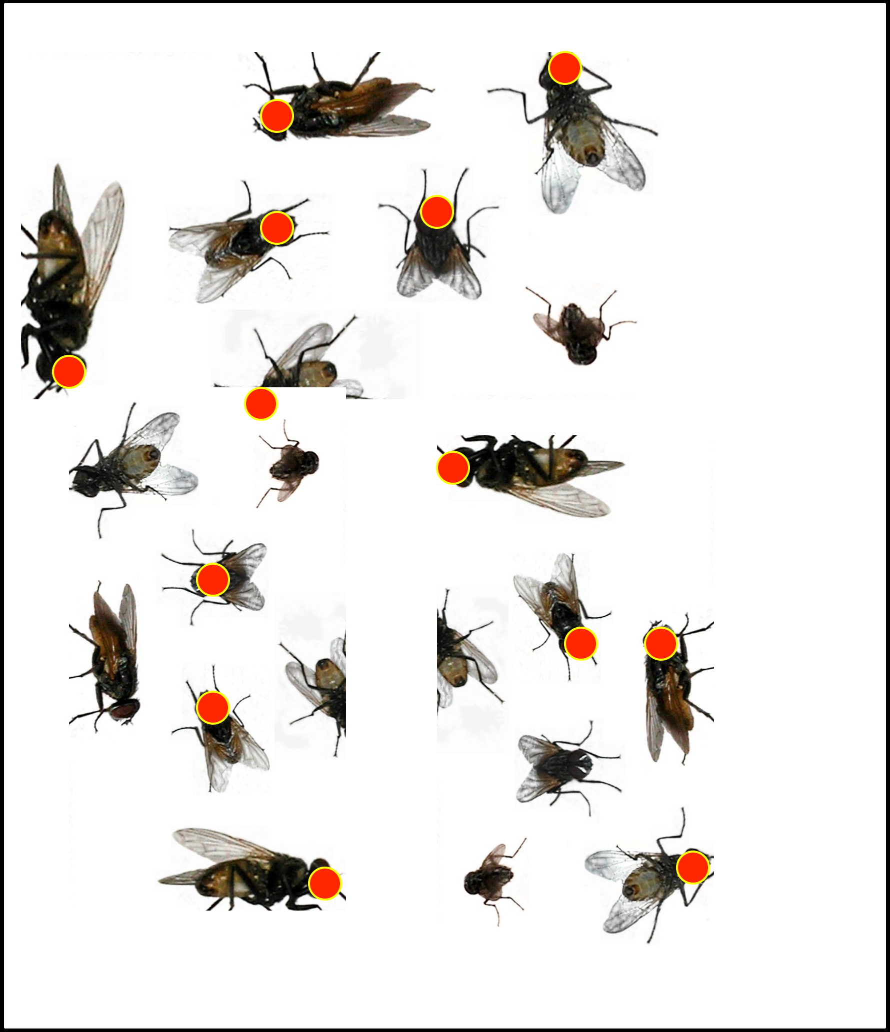
Waddington's Experiment (1952)

Generation 6

Temp: 40^{o} C

~63% changed

Select and breed those



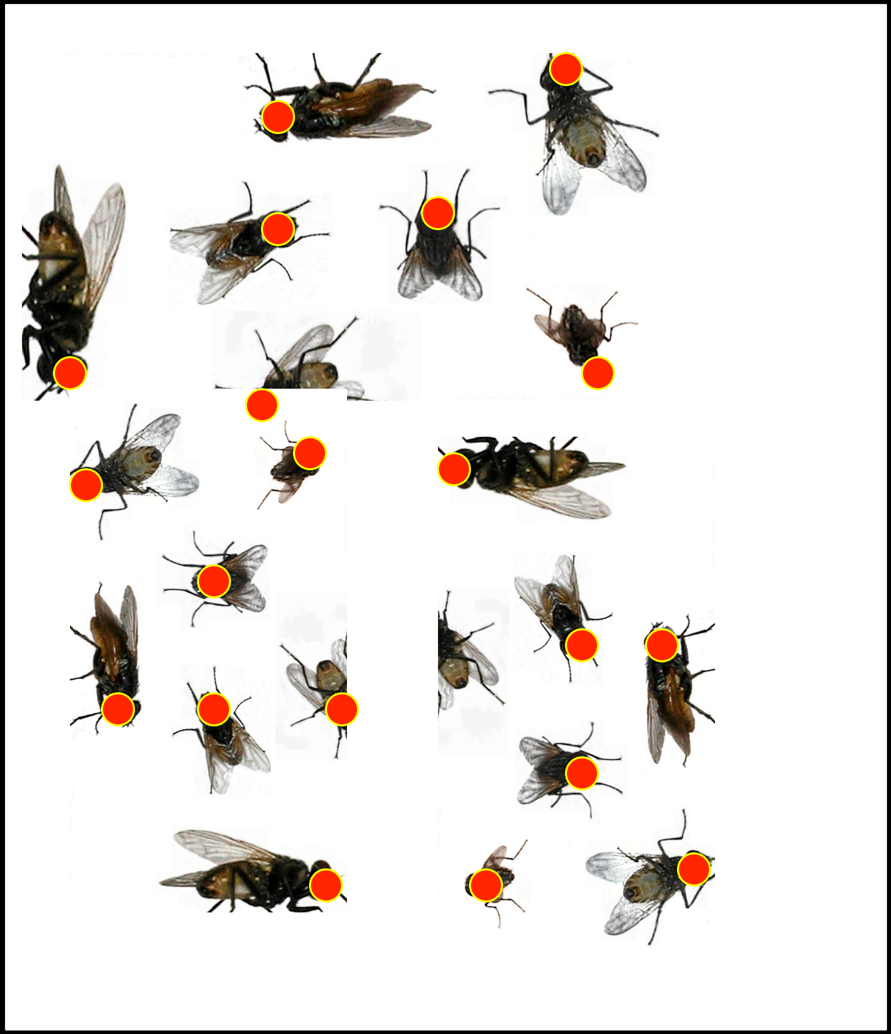
Waddington's Experiment (1952)

(…)

Generation 20

Temp: 40^{o} C

~99% changed

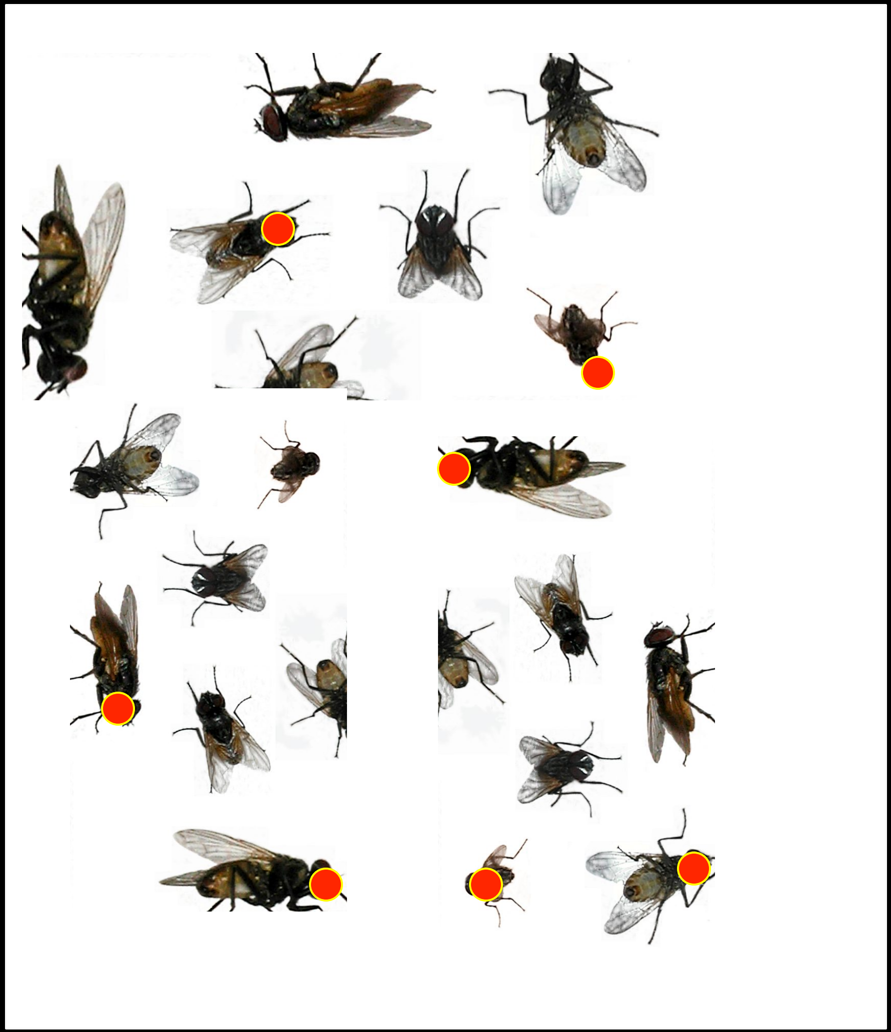


Surprise!

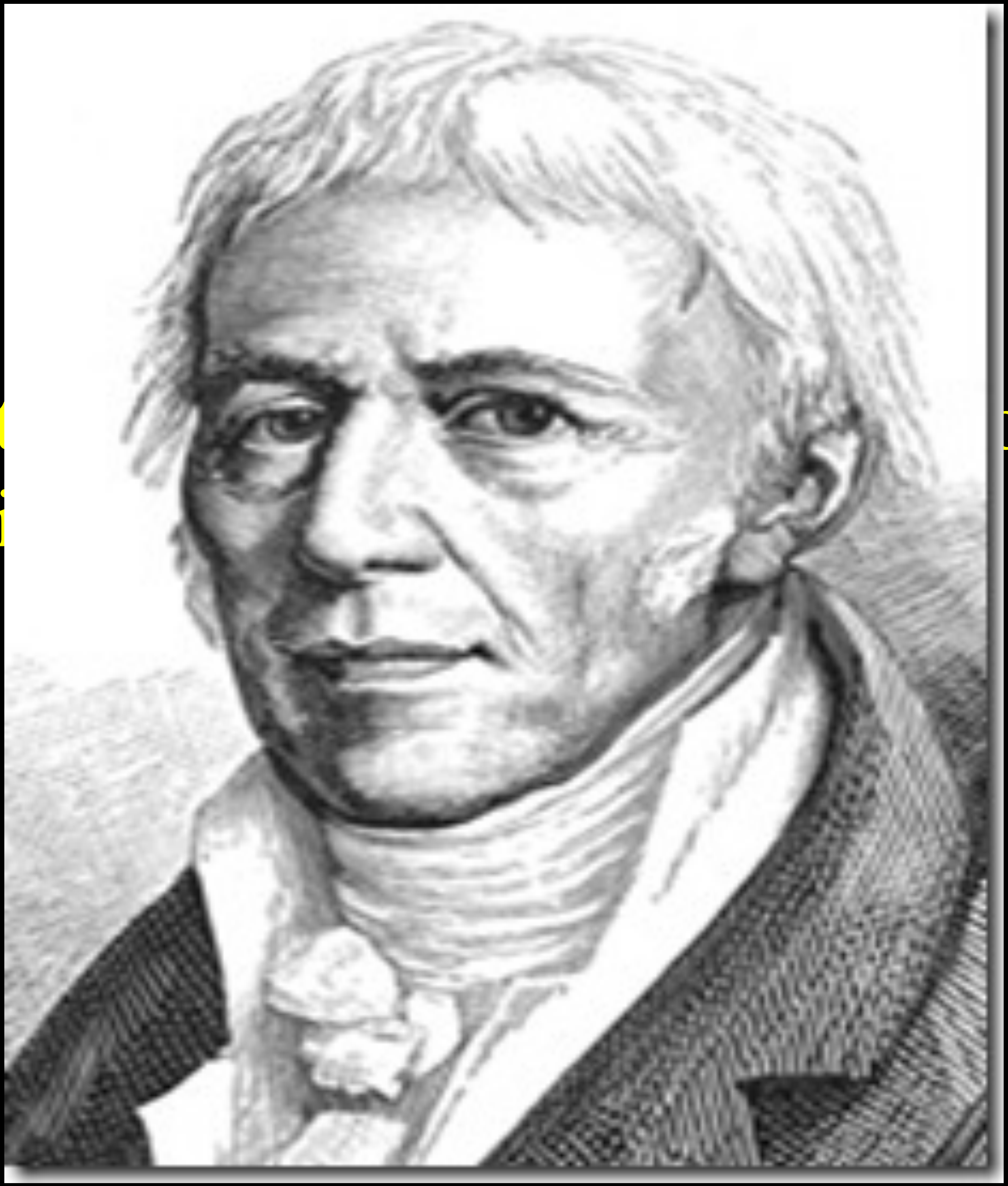
Generation 20

Temp: 20^{o} C

~25% stay changed!!



- Adapt ome
  geneti



A Genetic Explanation

- Suppose that "red head" is this Boolean function of 10 genes and "high temperature"

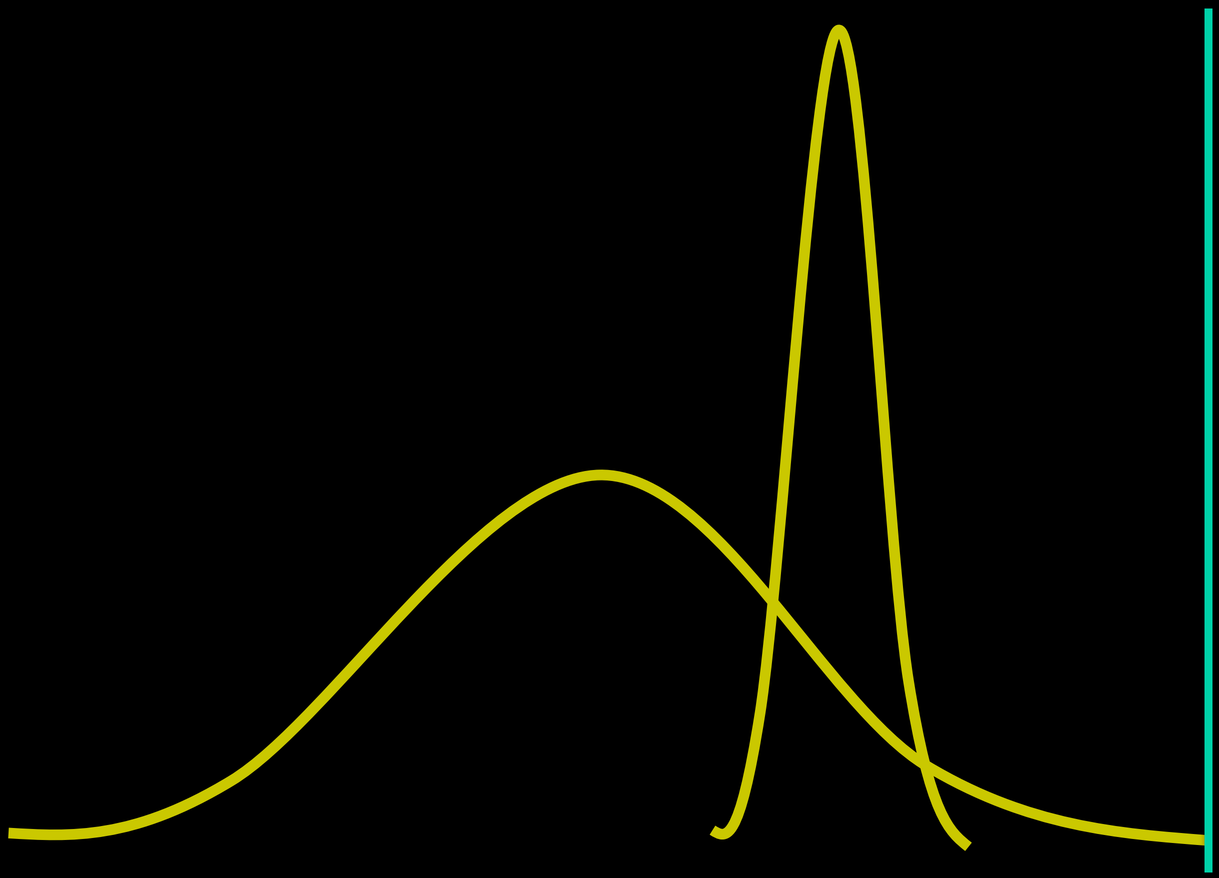
"red head" = "$x_1 + x_2 + \ldots + x_{10} + 3t \geq 10$"

- Suppose also that the genes are independent random variables, with p_i initially half, say.

A Genetic Explanation (cont.)

- In the beginning, no fly is red (the probability of being red is 2^{-n})
- With the help of t = 1, a few become red
- If you select them and breed them, ~60% will be red!

Why 60%?



A Genetic Explanation (cont.)

- Eventually, the population will be very biased towards $x_i = 1$ (the p_i's are close to 1)
- And so, a few flies will have all $x_i = 1$ for all i, and they will stay red.

Generalize!

- Let B is any Boolean function
- n variables $x_1\ x_2\ \ldots\ x_n$ (no t)
- Independent, with probabilities

  $p = (p_1\ p_2\ \ldots\ p_n)$
- Now, generate a population of bit vectors, and select the ones that make $B(x) = 1$

(cont.)

- In expectation, $p \rightarrow p'$,

  where $p_i' = \text{prob}_p (x_i = 1 \mid B(x) = 1)$

(Looking under the rug: Linkage?)

Conjecture: This solves SAT

Theorem (with Greg Valiant, 2011): If instead $p \rightarrow \varepsilon p' + (1 - \varepsilon) p$, then it converges to a satisfying assignment with prob. 1 (for large enough population and small enough ε).

Interpretation

- If there is any Boolean combination of a modestly large number of alleles that creates an unanticipated trait conferring even a small advantage, then this combination will be discovered and eventually fixed in the population.
- "With sex, all moderate-sized Boolean functions are evolvable."

Sooooo…

- The theory of life is deep and fascinating
- The point of view of a computer scientist makes it even more tantalizing
- Mixability helps understand the role of sex
- A natural stochastic process on Boolean functions may help illuminate genetic assimilation and the emergence of novel traits

Thank You!