

Lecture 3: Experimental Evolution

Assignment Deadlines

Assignment 1: Monday 4th March, 5pm

Assignment 2 to follow...

Recap: Drift, selection, and mutation

Kimura's diffusion equation

$p(x, t)$ Probability that allele frequency is x at time t

$$\frac{\partial p(x, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} \frac{x(1-x)}{N} p(x, t) - \frac{\partial}{\partial x} [\sigma x(1-x) + \mu(1-2x)] p(x, t)$$

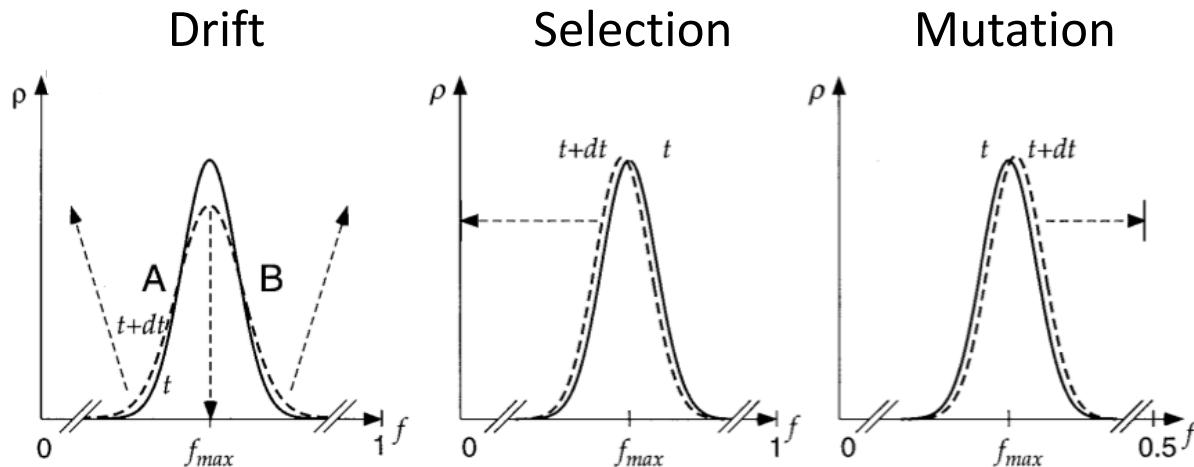
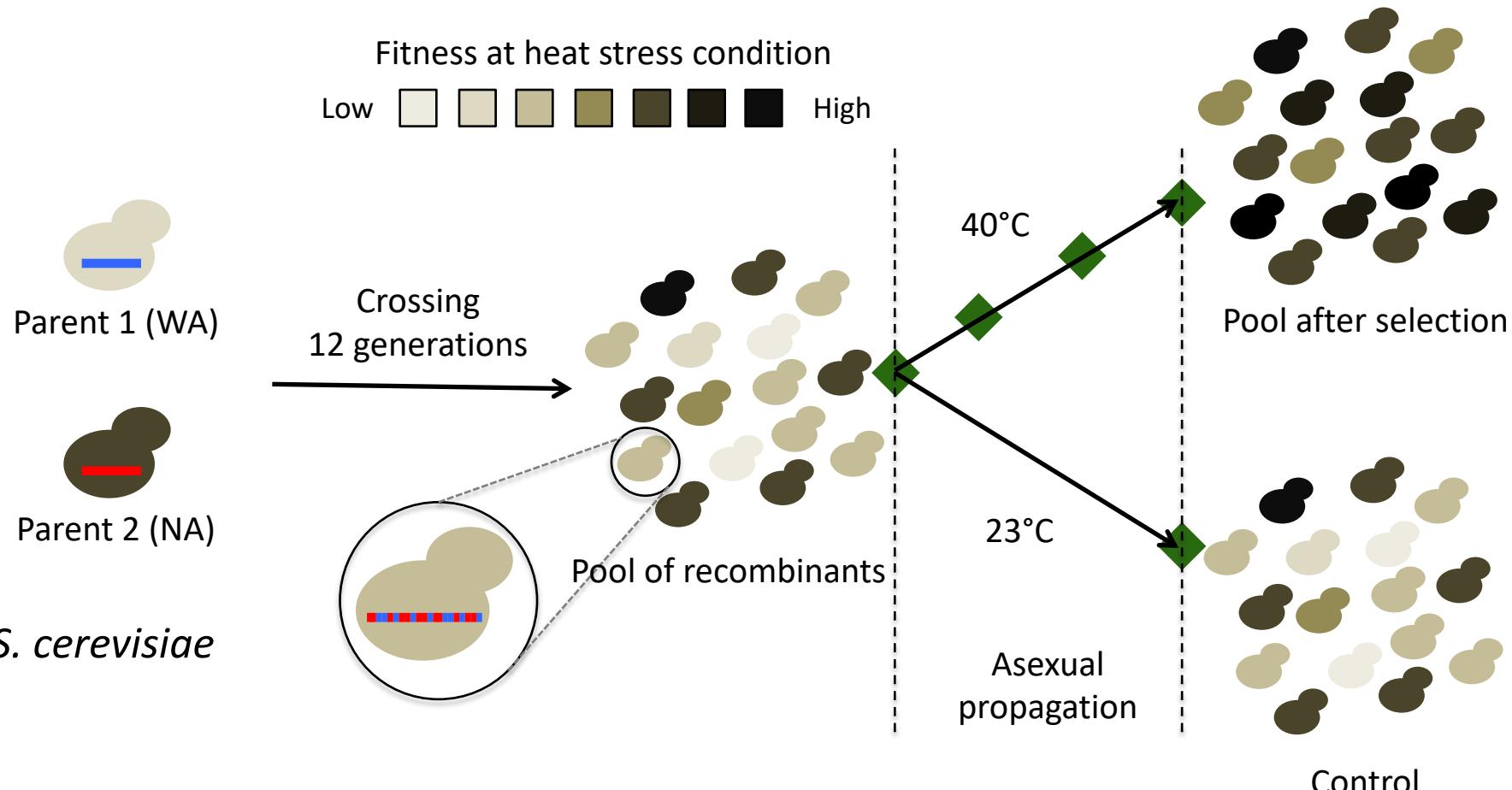


Image: Rouzine et al, 2001

Selection for heat tolerance in yeast

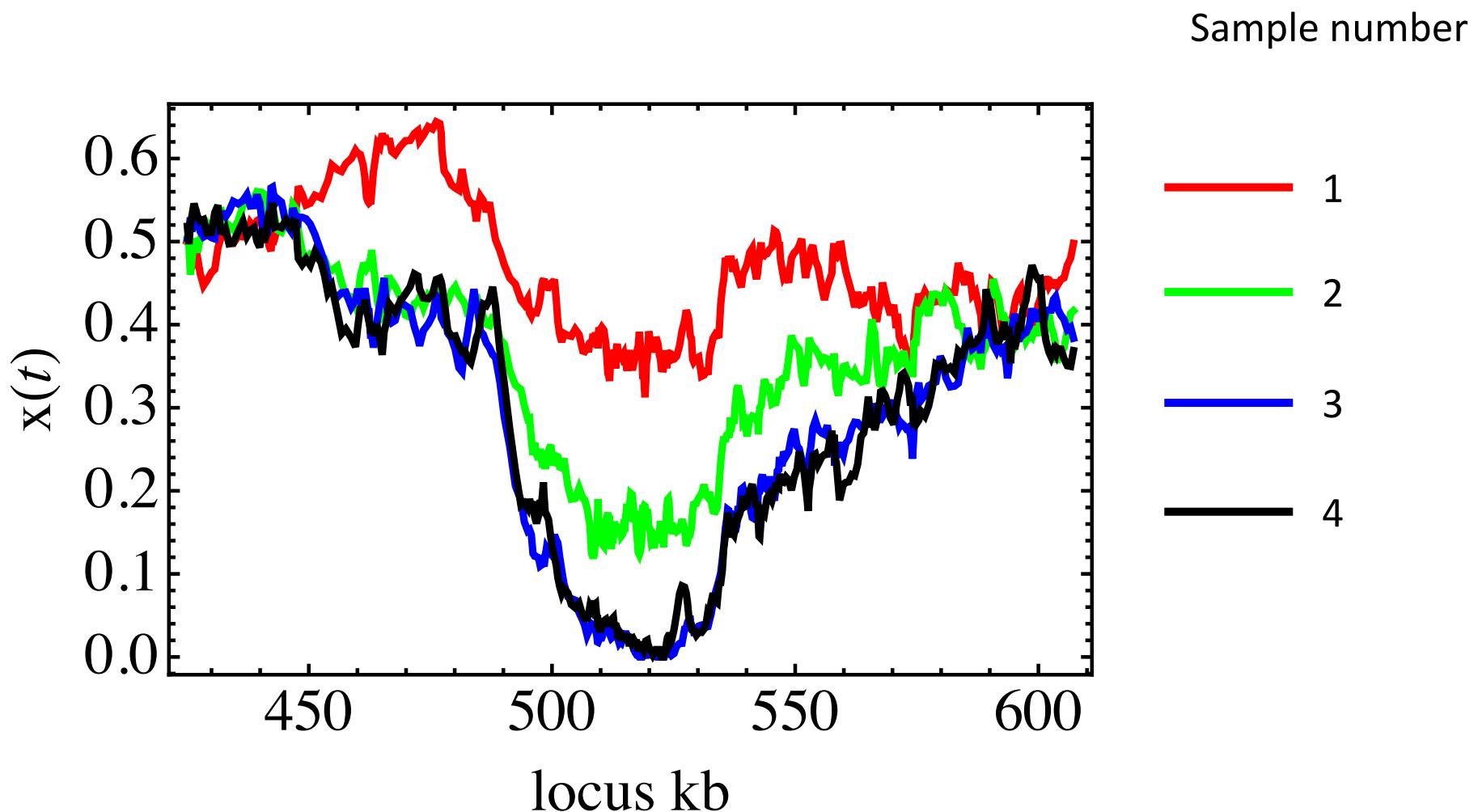


◆ Whole-genome sequencing of the pool

Figure adapted from Parts et al, Genome Research 2011

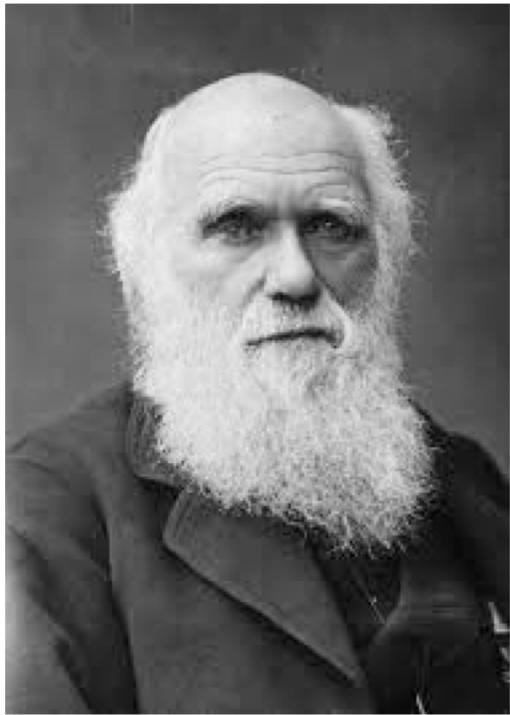
Selection for heat tolerance in yeast

Region in chromosome II



Experimental Evolution

Evolution, Fast and Slow

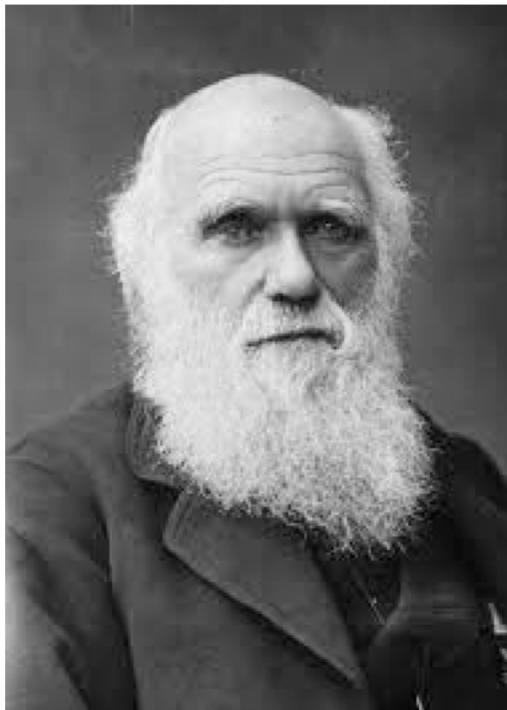


Natural selection is slow

“Natural selection will always act very slowly, often only at long intervals of time, and generally on only a few inhabitants of the same region at the same time”

The Origin of Species, 1859

Evolution, Fast and Slow



Artificial selection can produce substantial changes

“Slow though the process of selection may be, if feeble man can do much by his powers of artificial selection, I can see no limit to the amount of change ...which may be effected in the long course of time by nature's power of selection”

The Origin of Species, 1859

Evolution, Fast and Slow

Animal and plant breeding have brought about substantial phenotypic changes across the course of human history



Evolution, Fast and Slow

Animal and plant breeding have brought about substantial phenotypic changes across the course of human history



Teosinte

Modern maize

Evolution, Fast and Slow

Trait-selection experiments have shown substantial phenotypic changes on a time-scale of decades

The Illinois Long-term Evolution Experiment

Began in 1896:

Aim to improve the characteristics of maize

Protein content : Animal feed

Oil content : Corn oil

Initially 163 ears of corn. Analyse for chemical content.

Use top 24 ears to seed a high protein/oil crop

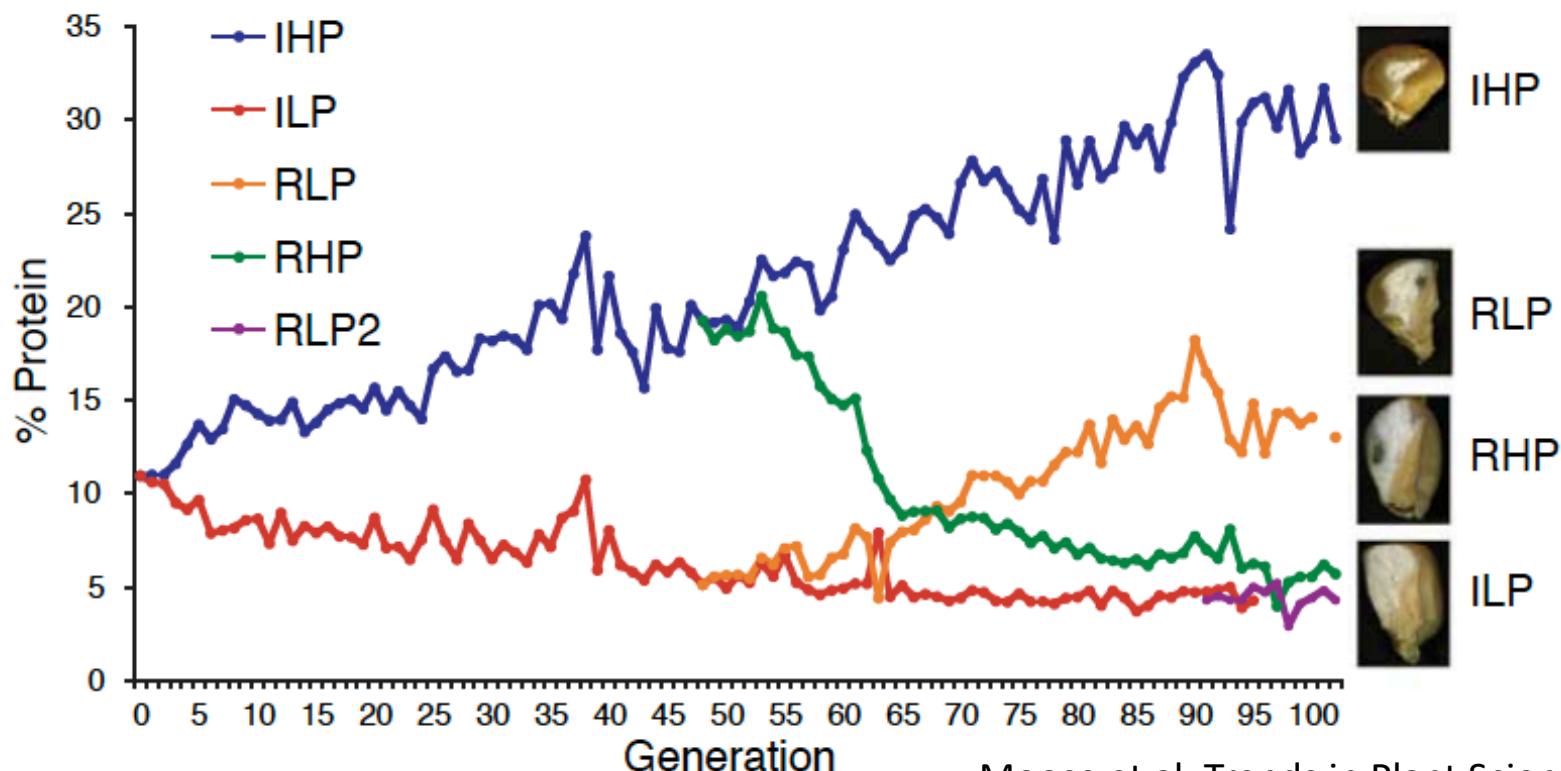
Use bottom 12 ears to seed a low protein/oil crop

Repeat each year

Evolution, Fast and Slow

The Illinois Long-Term Selection Experiment

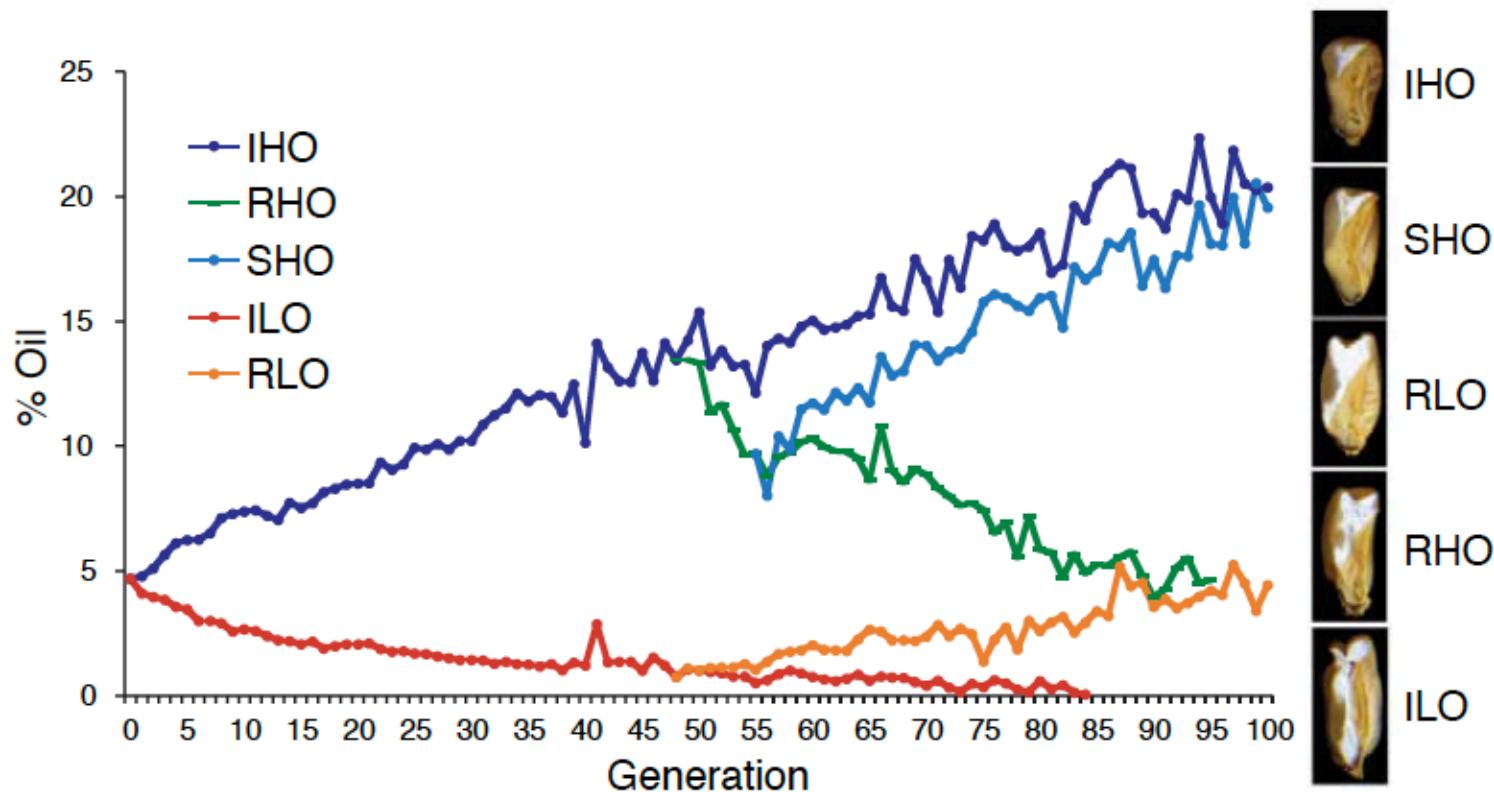
Select for high and low protein content



Evolution, Fast and Slow

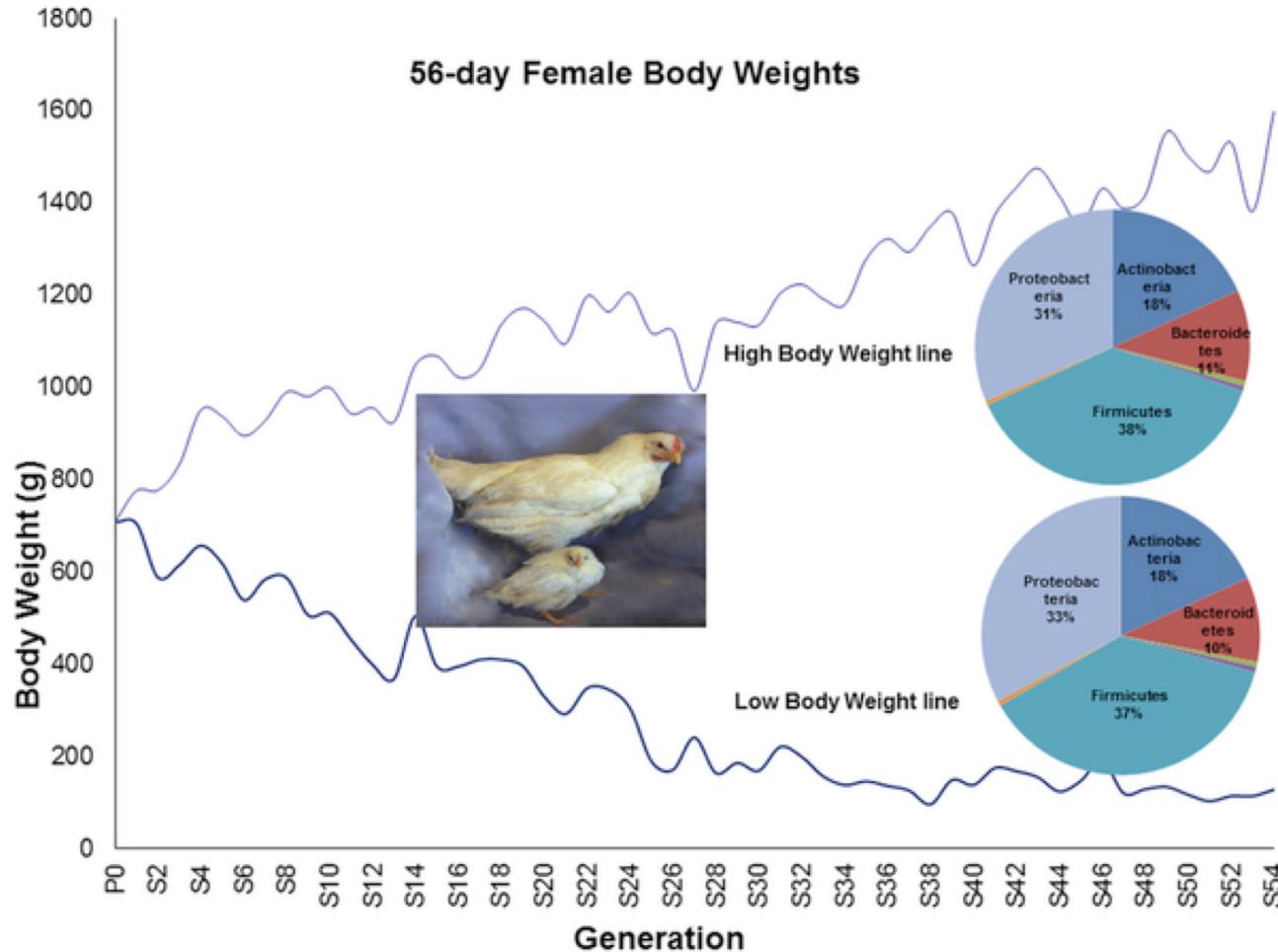
The Illinois Long-Term Selection Experiment

Select for high and low oil content



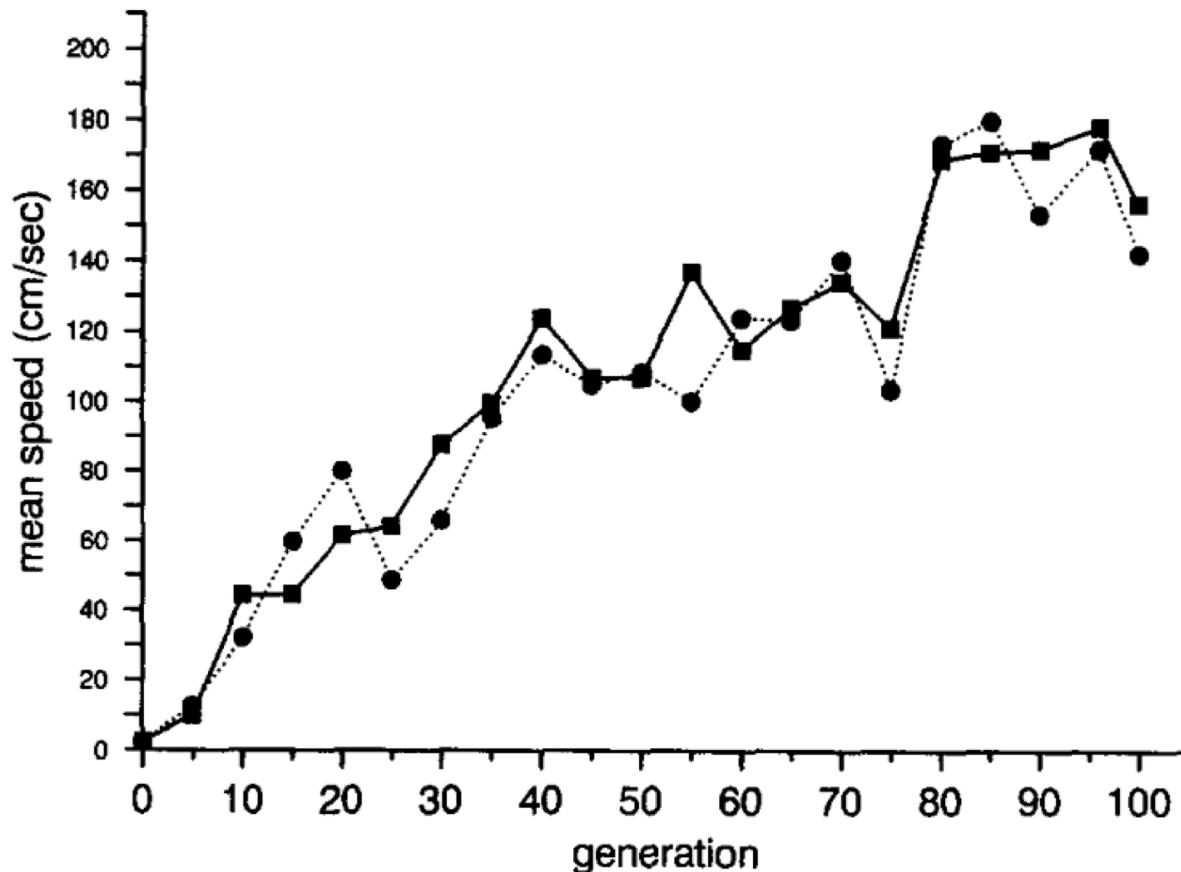
Evolution, Fast and Slow

Selection for body weight in chickens



Evolution, Fast and Slow

Selection for flight speed in *Drosophila melanogaster*



Evolutionary experiments

Definition

The study of evolutionary changes occurring in experimental populations as a consequence of conditions (environmental, demographic, genetic, social, and so forth) imposed by the experimenter

Kawecki et al, Trends in Ecology and Evolution, 2012

Evolutionary experiments

Experiments can be conducted in either wild or laboratory conditions

Pros of lab evolution:

- Environment is known and can be controlled
- Experiments can be replicated

Cons of lab evolution:

- Lab conditions are unnatural (artificial light, food, climate)
 - Adaptation to lab conditions may occur

- Phenotypes which would not be viable in the wild may occur
 - Lack of predators etc.

- What happens in the real world cannot always be replicated in the lab
 - Drosophila response to climatic variation

Evolutionary experiments

Example: Laboratory evolution of *Escherichia coli*



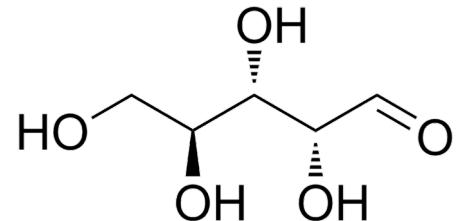
Evolutionary experiments

Example: Lab evolution of *Escherichia coli*

Begun in 1988. Twelve near-identical populations of *E. coli*.

Differ in ability to metabolise arabinose

6 Ara⁺, 6 Ara⁻

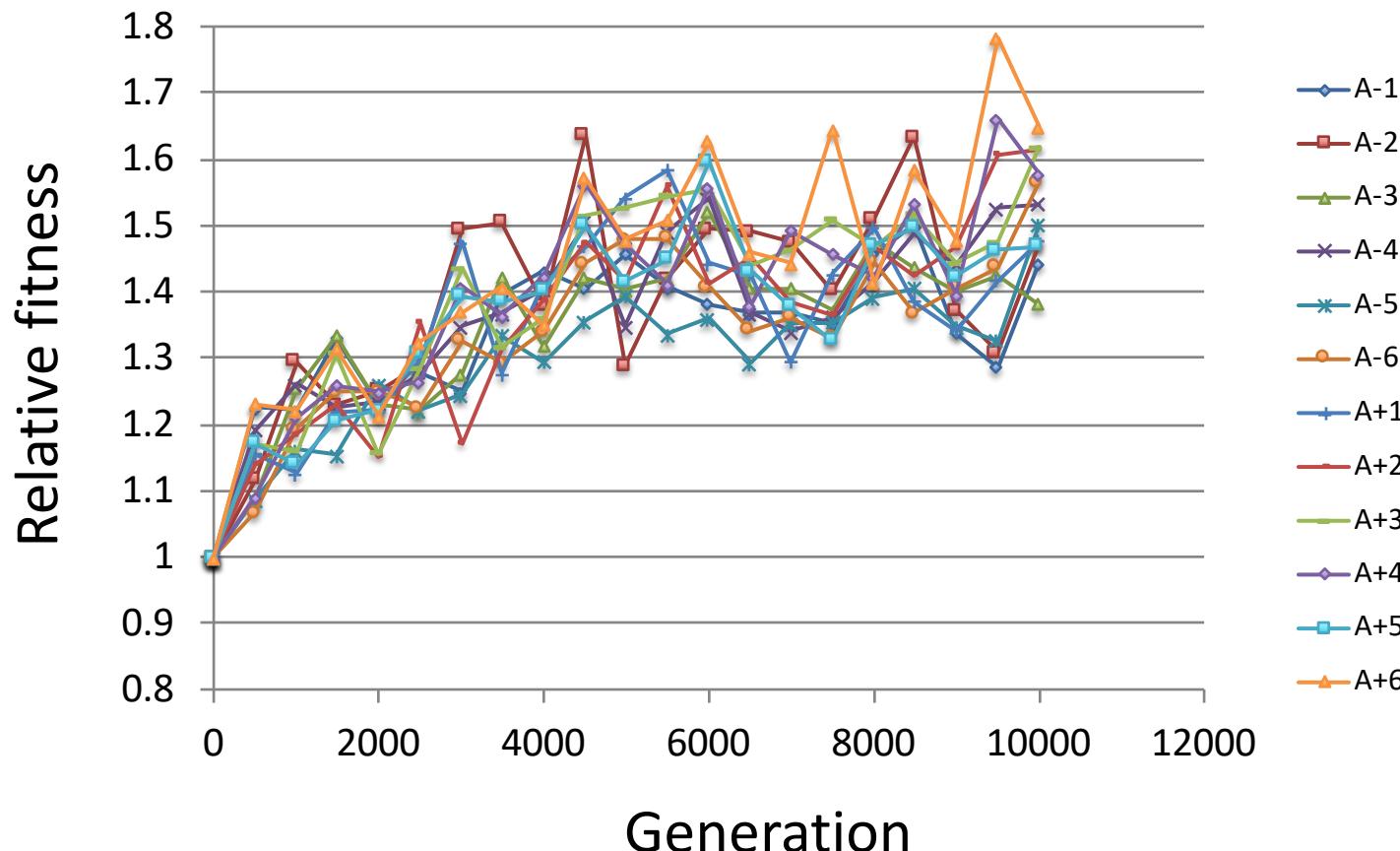


Populations grow over time: each day transfer 1% of each population to a new flask.

Evolutionary experiments

Example: Lab evolution of *Escherichia coli*

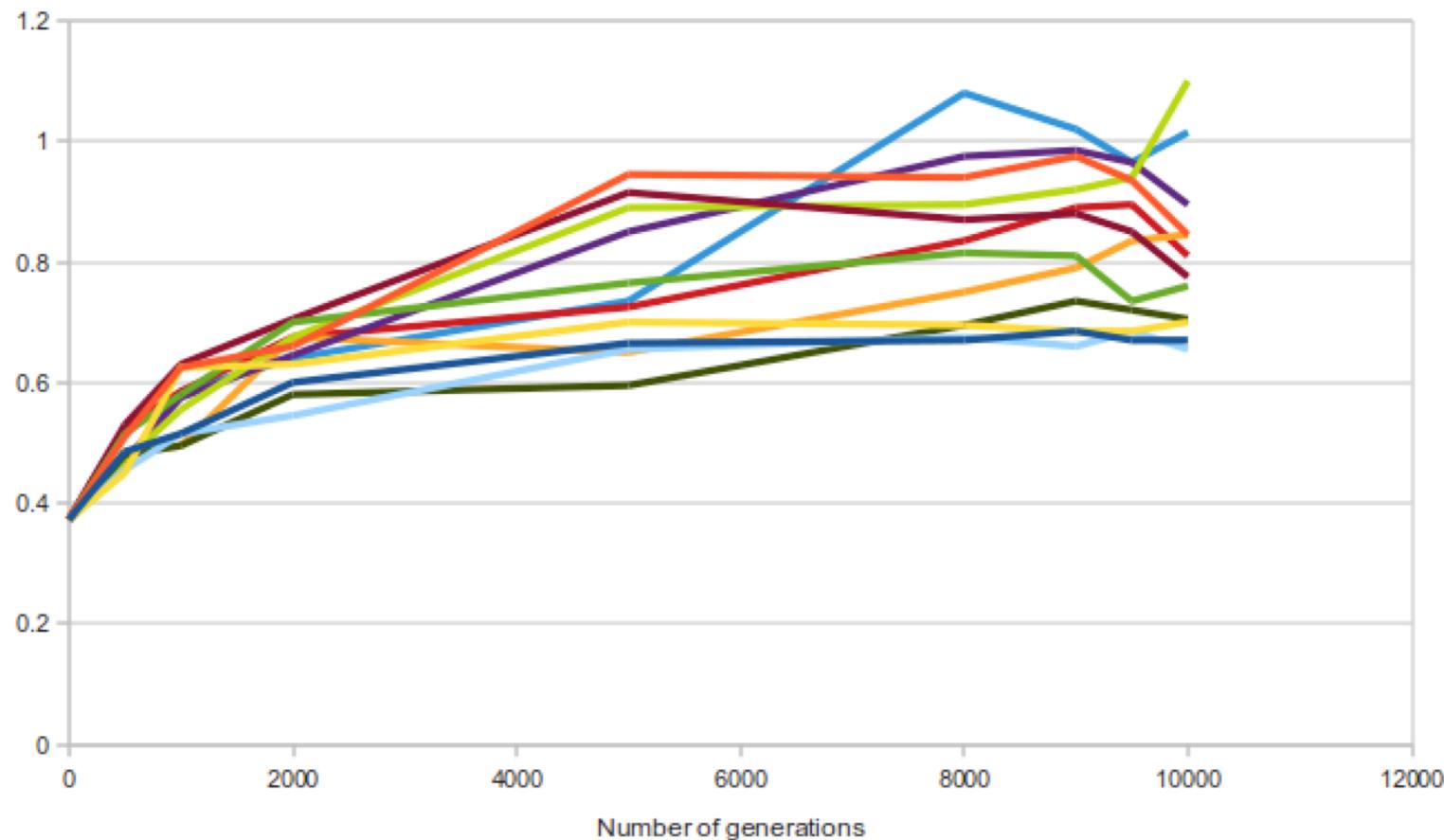
Change in fitness over first 10,000 generations



Evolutionary experiments

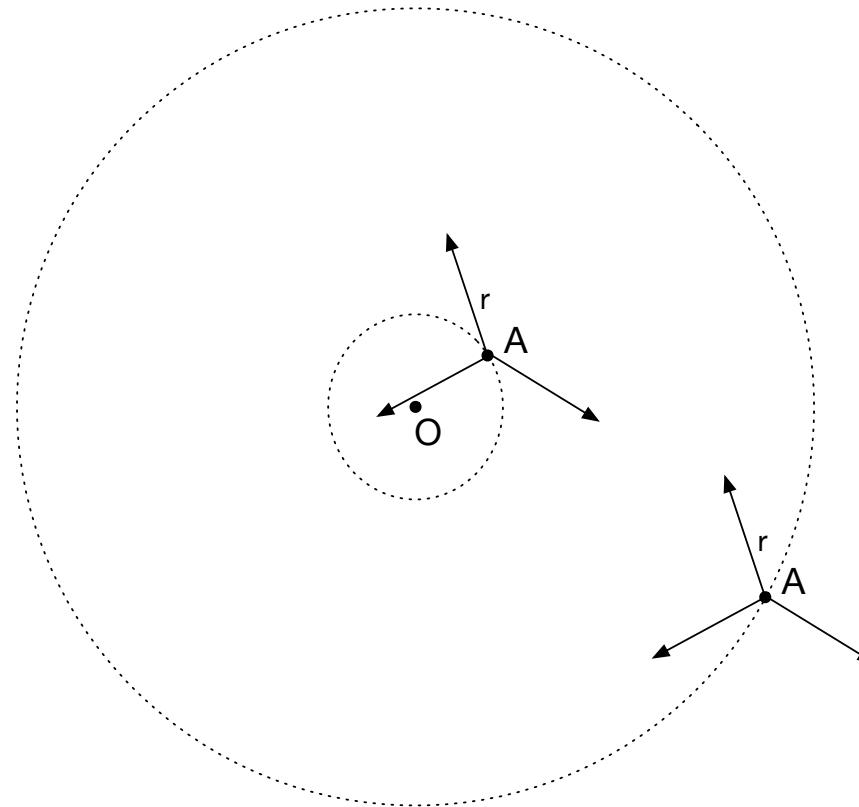
Example: Lab evolution of *Escherichia coli*

Change in cell size over first 10,000 generations



Fisher's geometric model

Generic model of fitness and adaptation



Point O represents fitness maximum

Experiments and evolutionary theory

E. coli populations with differing mutation supply rates

Two strains :

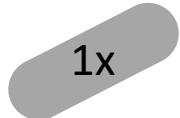


Adapted (post-10k generations)



Non-adapted

Varying mutation rates:



1x



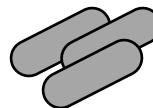
3x



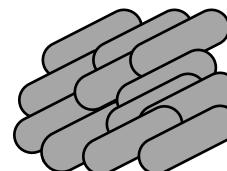
35x

Disable DNA-repair genes

Varying population sizes:



Standard size

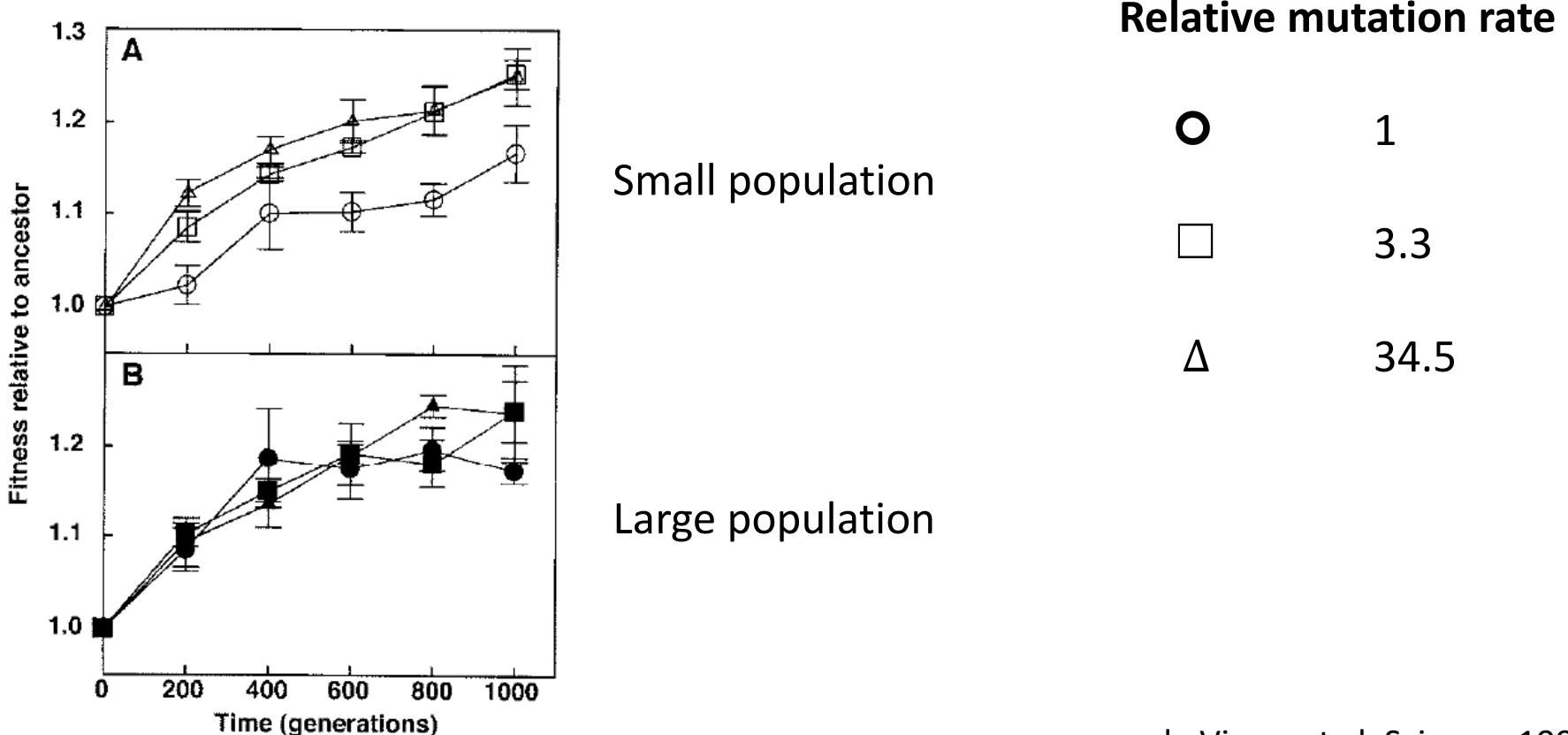


Extra-large size (50x individuals)

“Speed limit” on adaptive asexual evolution

Non-adapted population fitness

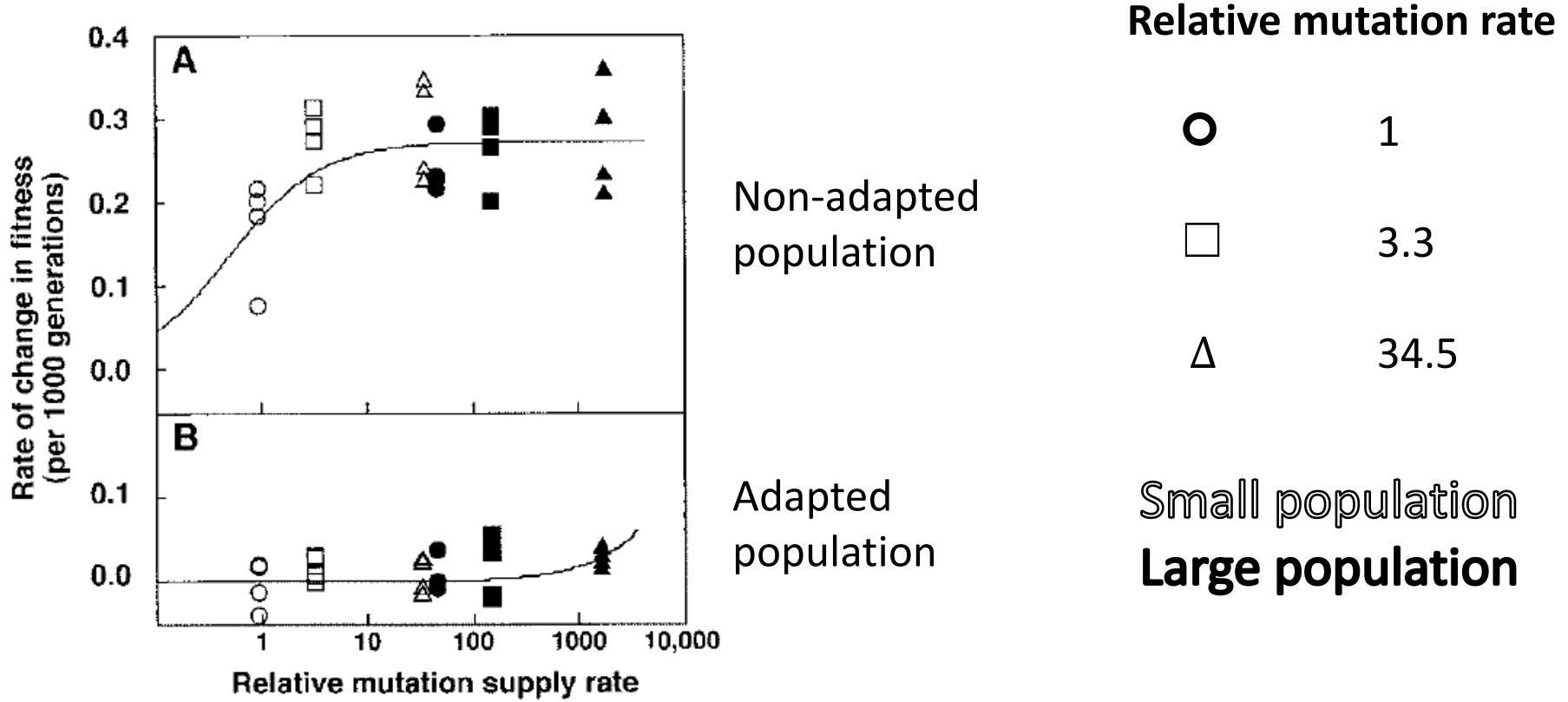
Increase seen in every case



“Speed limit” on adaptive asexual evolution

Adaptation rate vs mutation supply $N\mu$

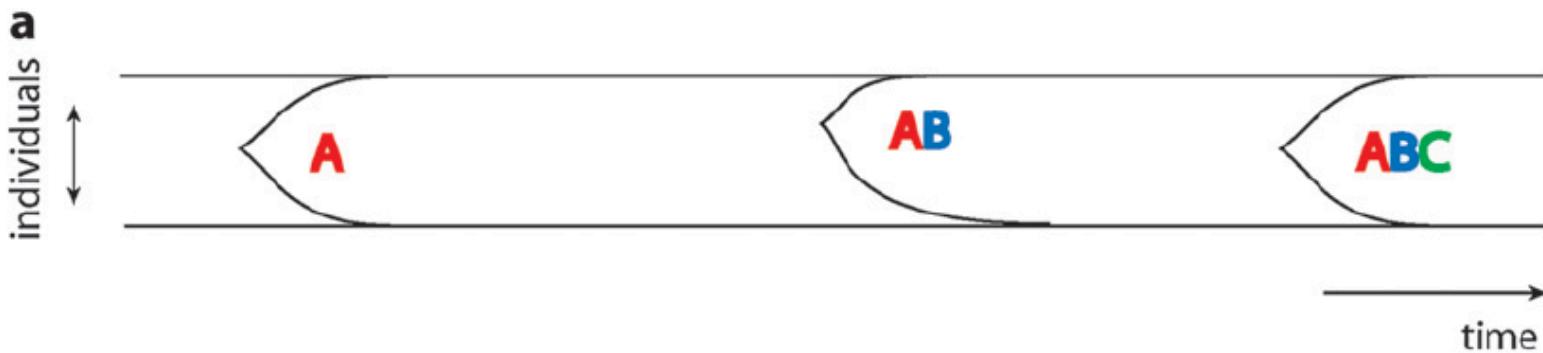
Plateau in the rate of adaptation



“Speed limit” on adaptive asexual evolution

Clonal interference slows adaptation

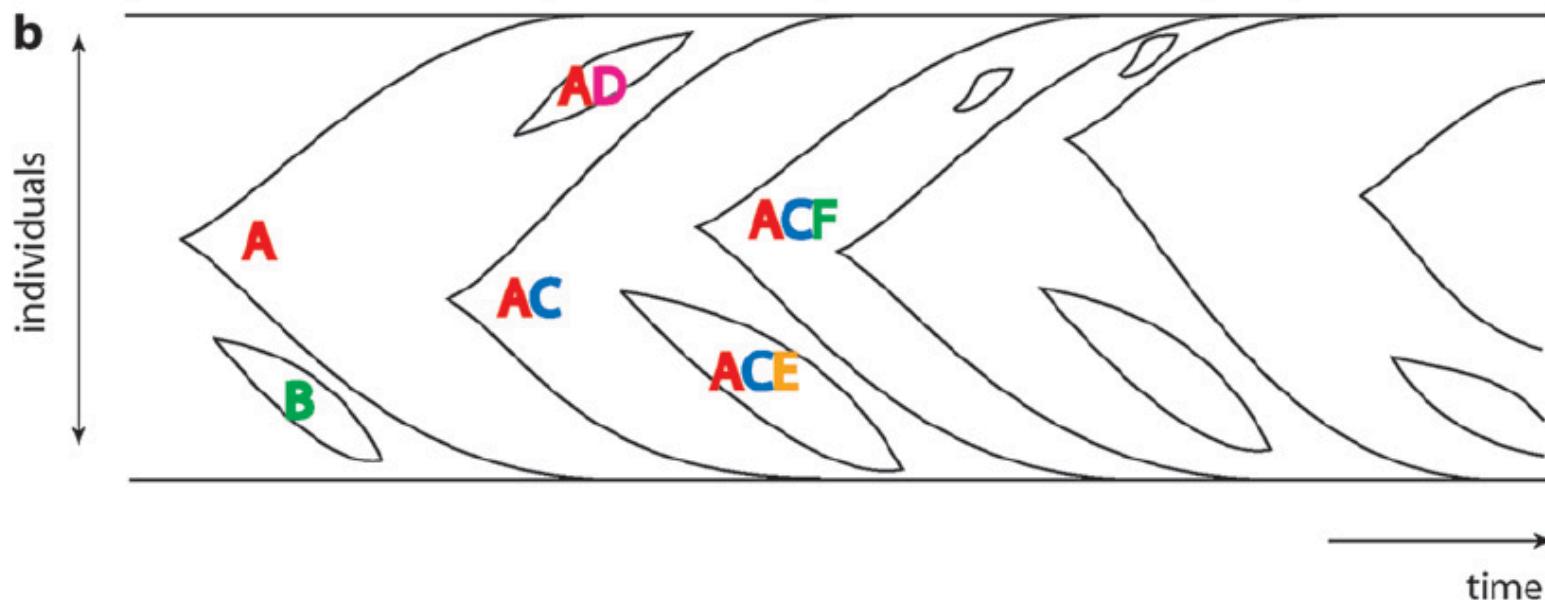
Low mutation rates: beneficial mutations appear slowly, but are likely to fix



“Speed limit” on adaptive asexual evolution

Clonal interference slows adaptation

High mutation rates: beneficial mutations appear rapidly, but outcompete one another: only the strongest survive.



Recombination

Molecular process

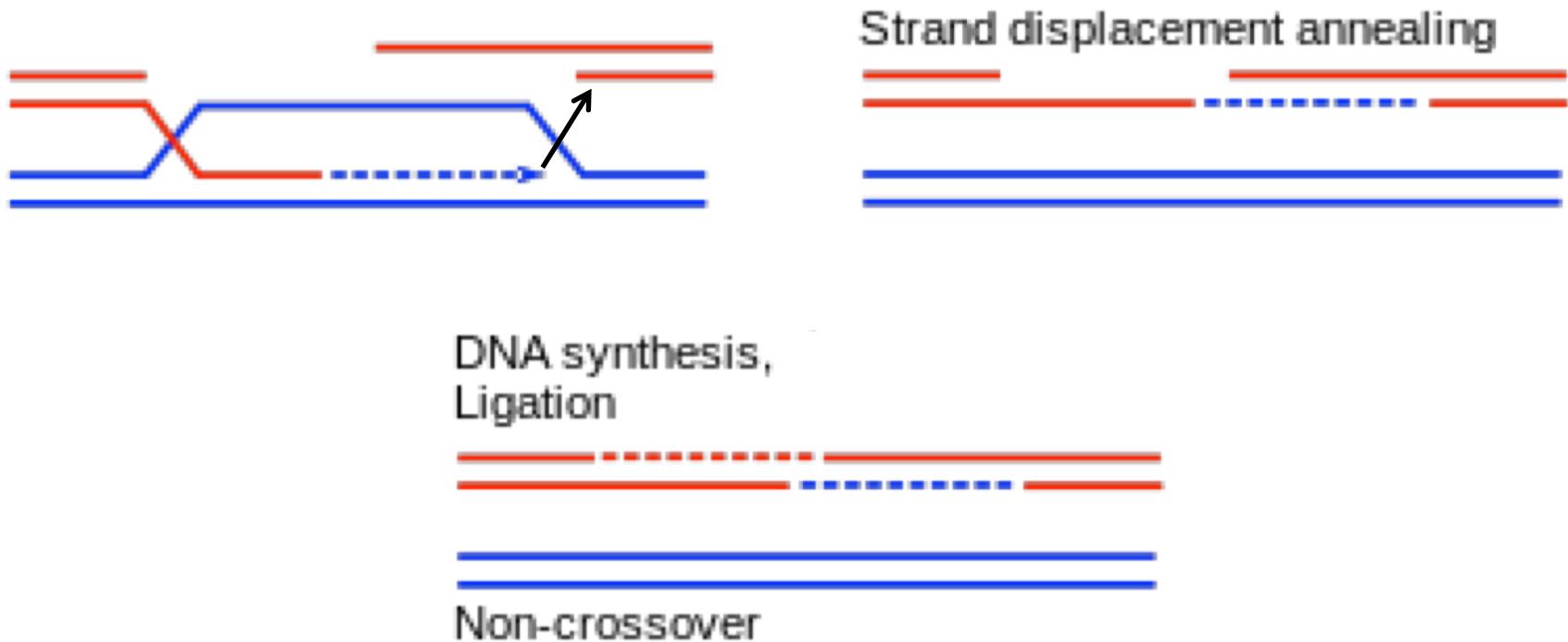
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

Molecular process

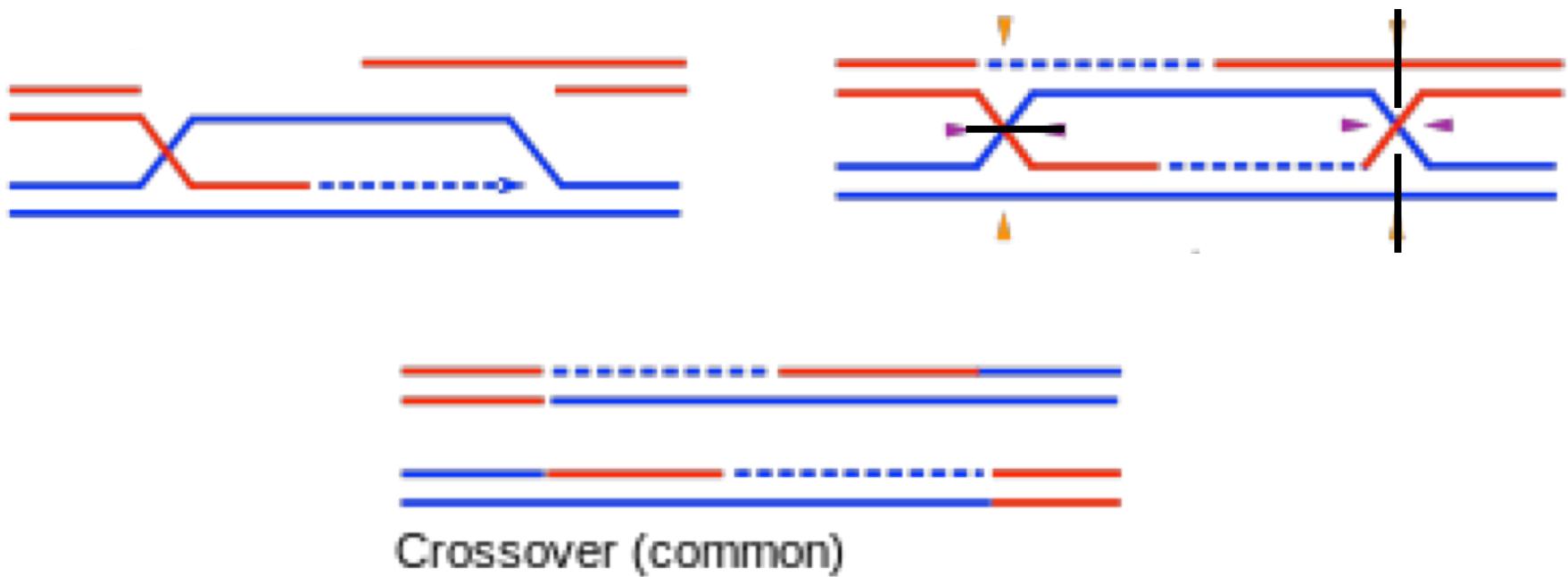
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

Molecular process

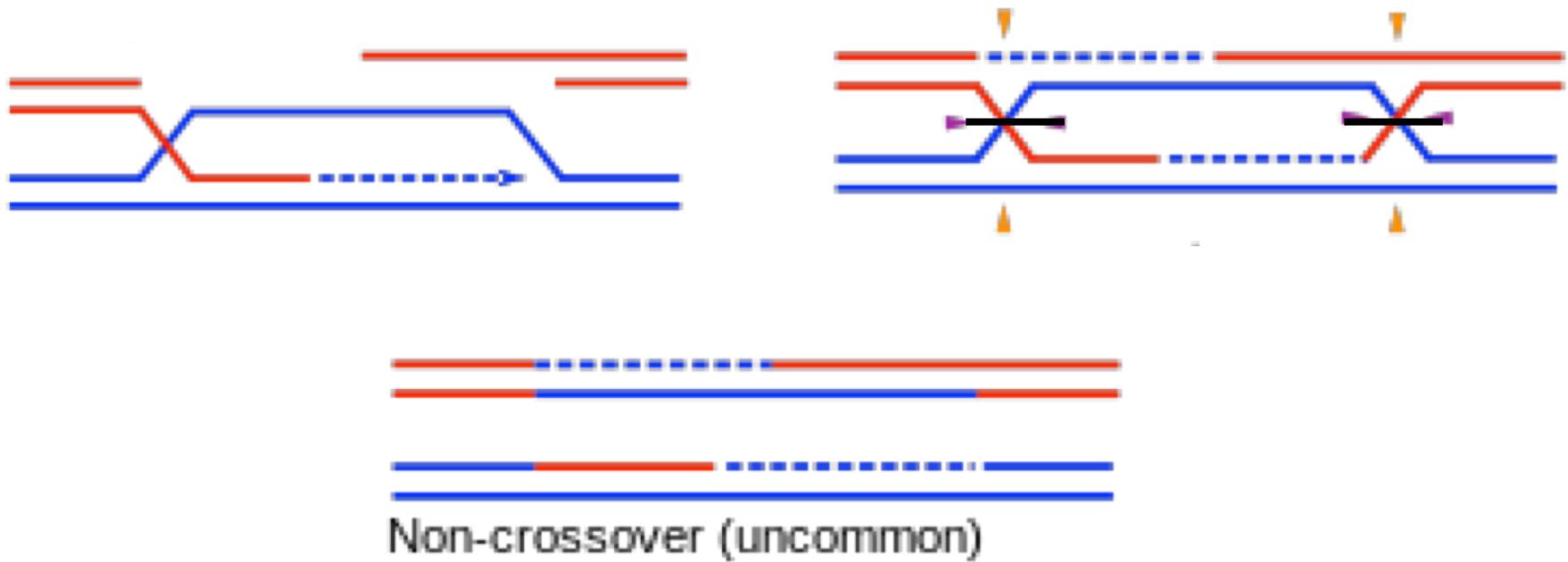
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

Molecular process

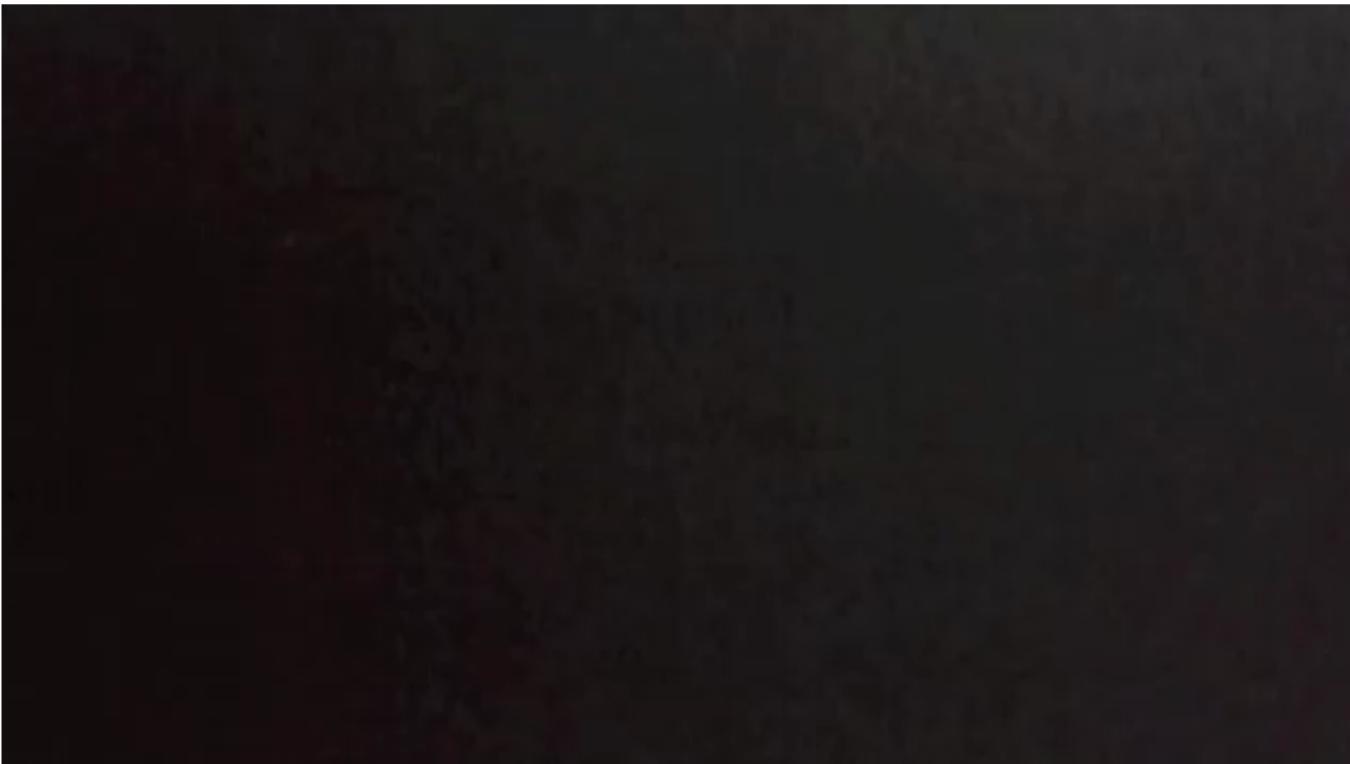
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

Molecular process

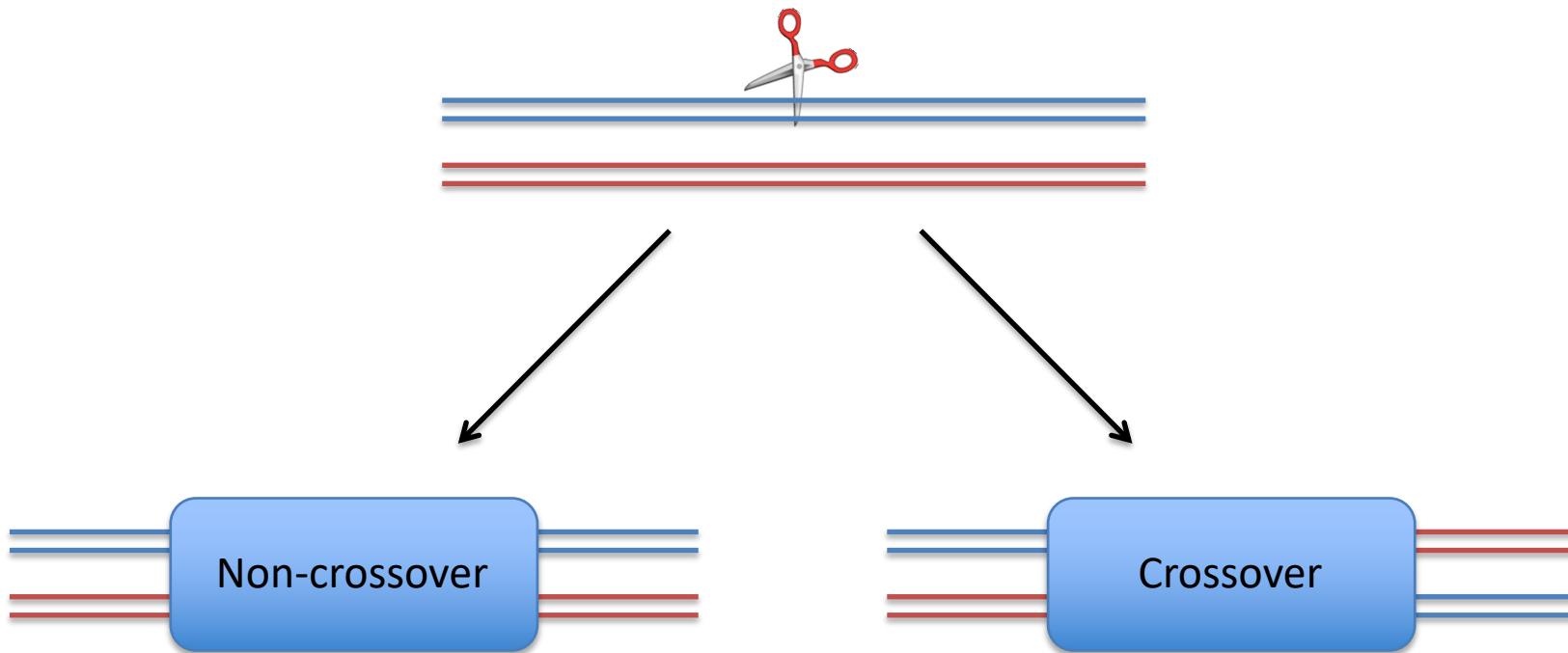
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

Molecular process

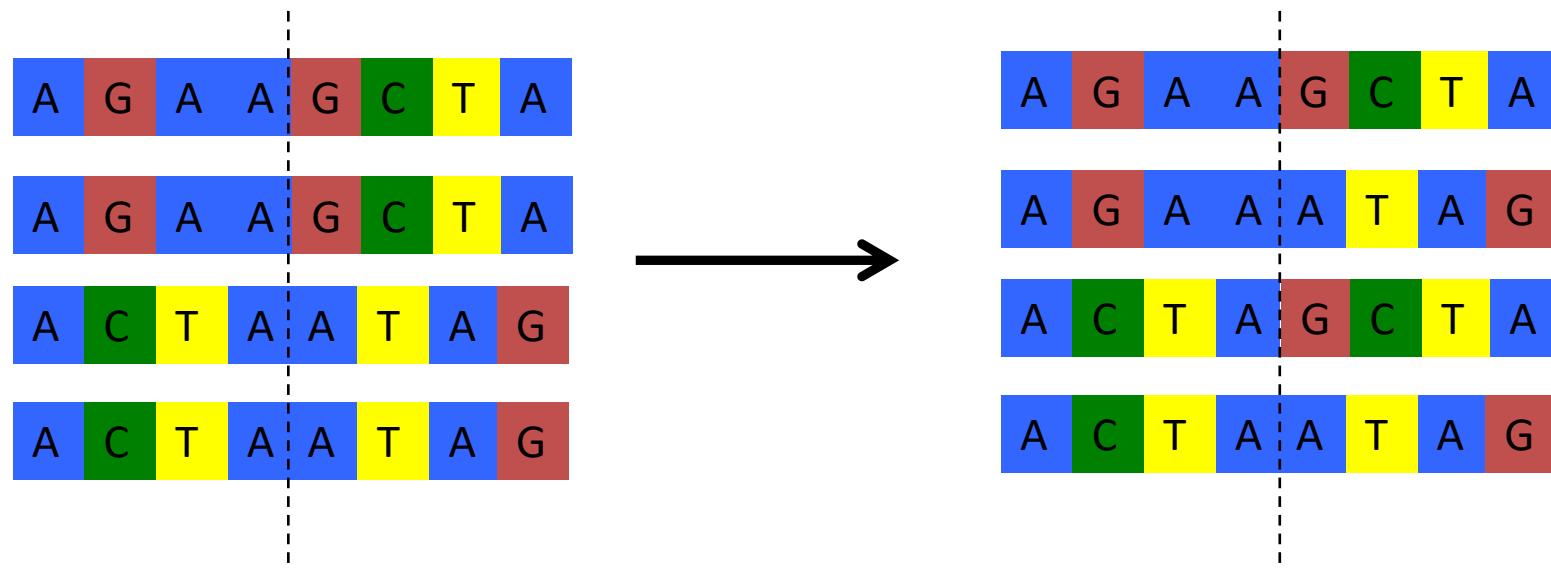
Recombination occurs through a process of DNA breakage and repair



Genetic recombination

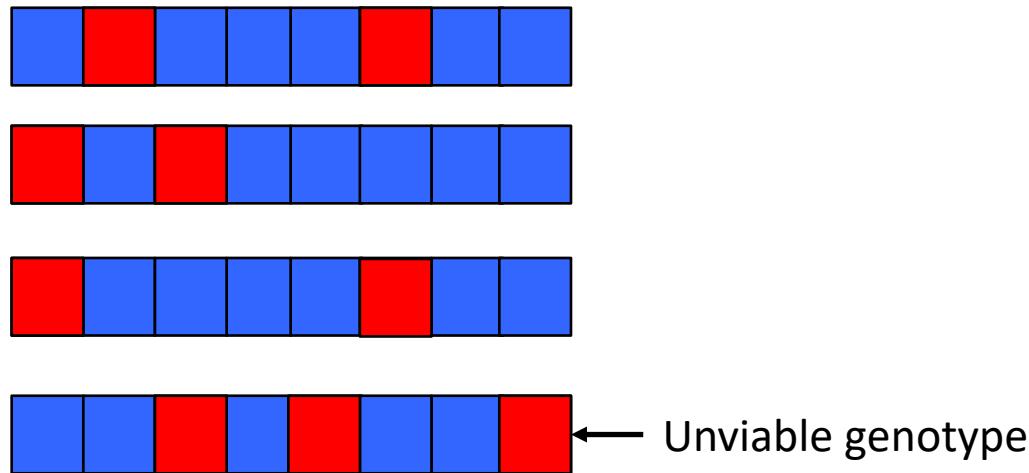
Effect of recombination

Produce new combinations of alleles



Importance of recombination: Muller's Ratchet

Accumulation of deleterious mutations

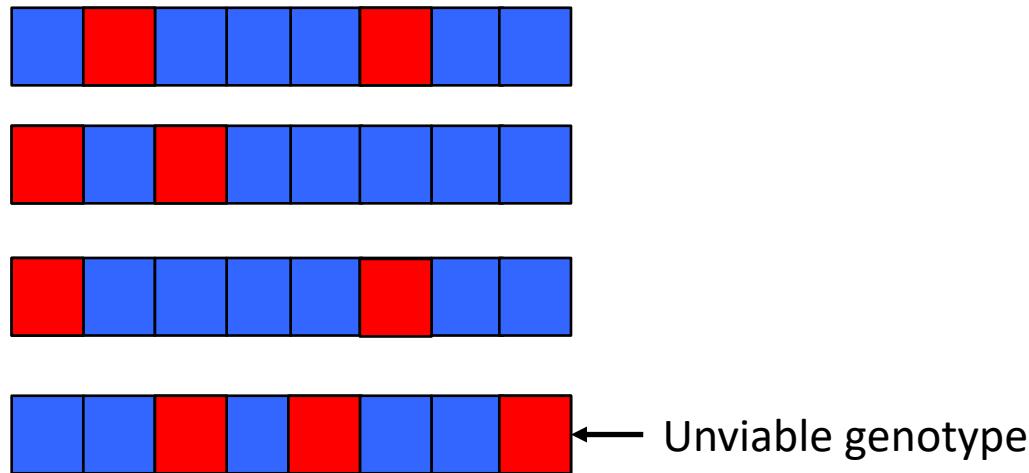


In each generation, each individual copies itself

Some rate of mutation: most mutations are deleterious

Deleterious mutations: Muller's Ratchet

Accumulation of deleterious mutations

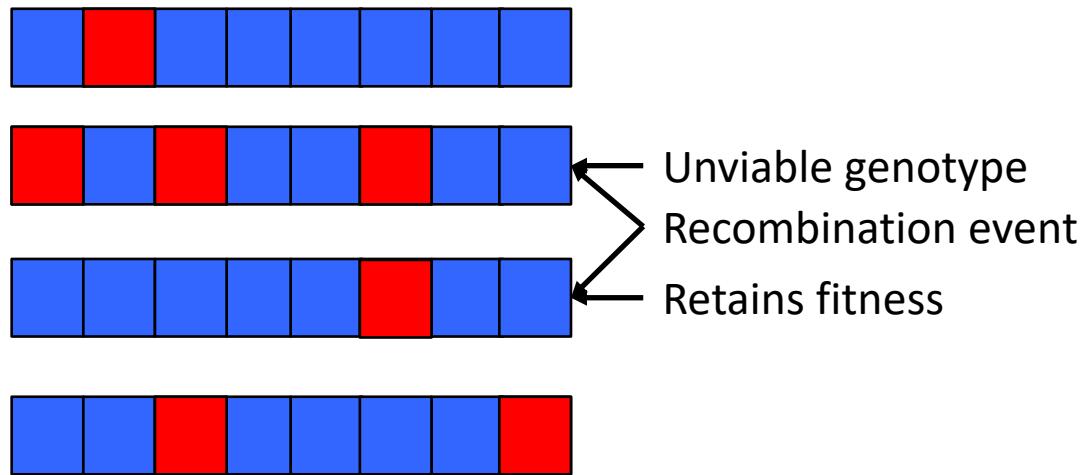


Population fitness decreases over time

Fitness can only be recovered by reverse mutation

Deleterious mutations: Muller's Ratchet

Sexual reproduction allows mutations to be recombined out

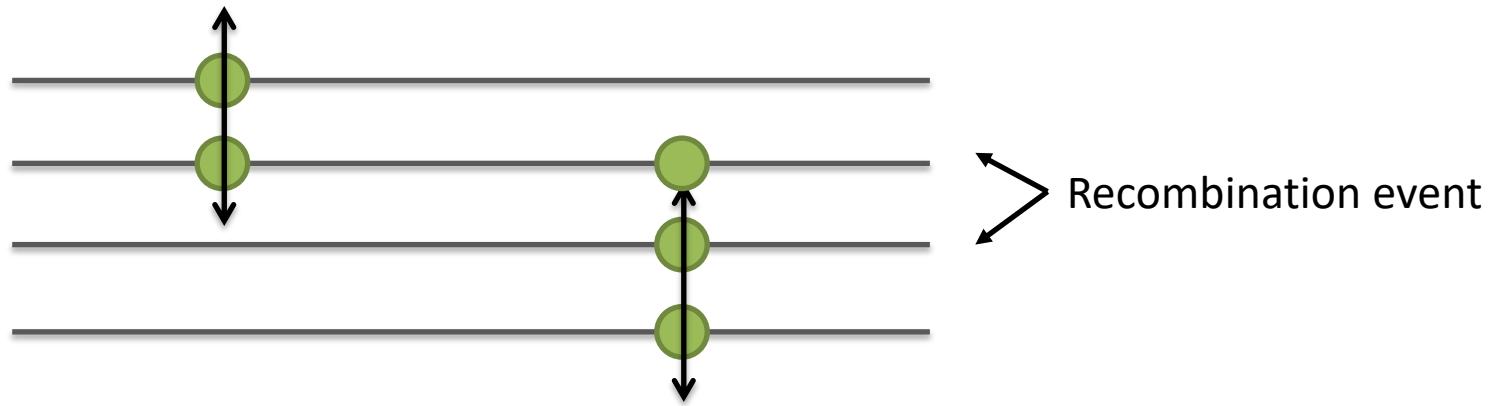


Recombination swaps combinations of alleles

Allows for fitter genotypes to be maintained; deleterious mutants recombined out of the population.

“Speed limit” on adaptive asexual evolution

Recombination breaks the speed limit



Beneficial mutations on opposing haplotypes oppose each other

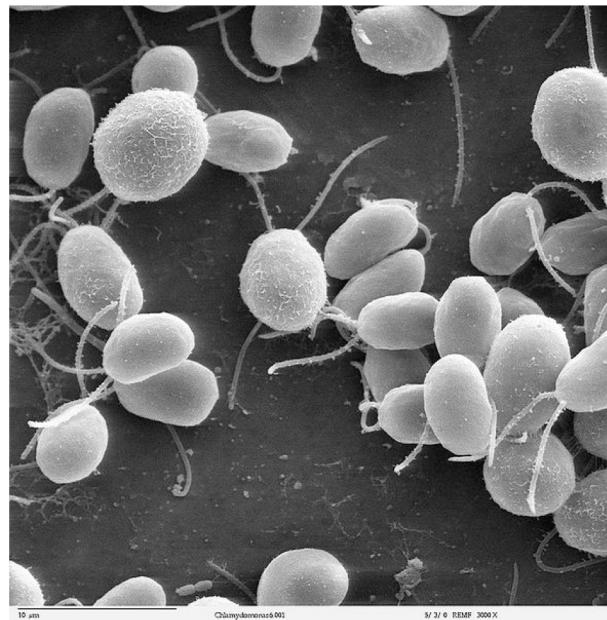
Recombination allows haplotypes to share beneficial mutations

Under asexual evolution, the double mutant could only be created by two mutations. Sex allows for more rapid adaptation

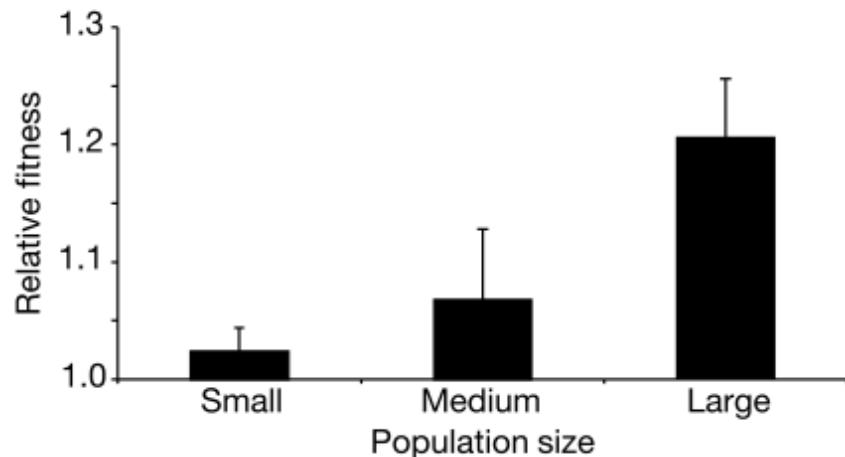
Importance of Recombination

Recombination and adaptation

Experimental evolution of algae *Chlamydomonas reinhardtii*



Difference in fitness between sexual and asexual lines



In larger populations, with greater supply of mutations, sexual reproduction is of greater benefit to the population

Pathogenic origin to sexual reproduction?

Red Queen hypothesis

Evolutionary race between pathogens and hosts:

Increased variance of offspring confers an evolutionary advantage

Sexual reproduction increases variance



Evolutionary dynamics

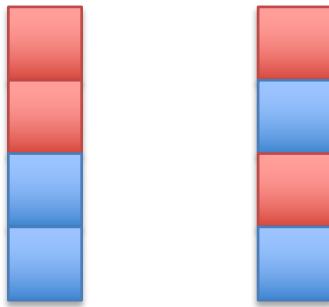
How does selection at one position in the genome affect the overall composition of the population?

Linkage disequilibrium quantifies associations

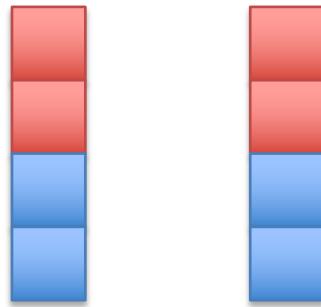
Definition

Non-random association of alleles at two or more loci

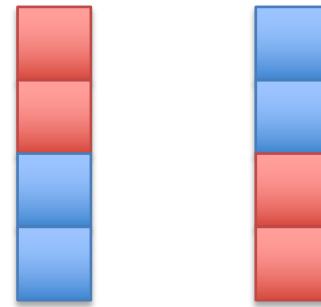
$$D_{ij} = q_{ij}^{11} - q_i^1 q_j^1$$



$$D_{ij} = 0$$



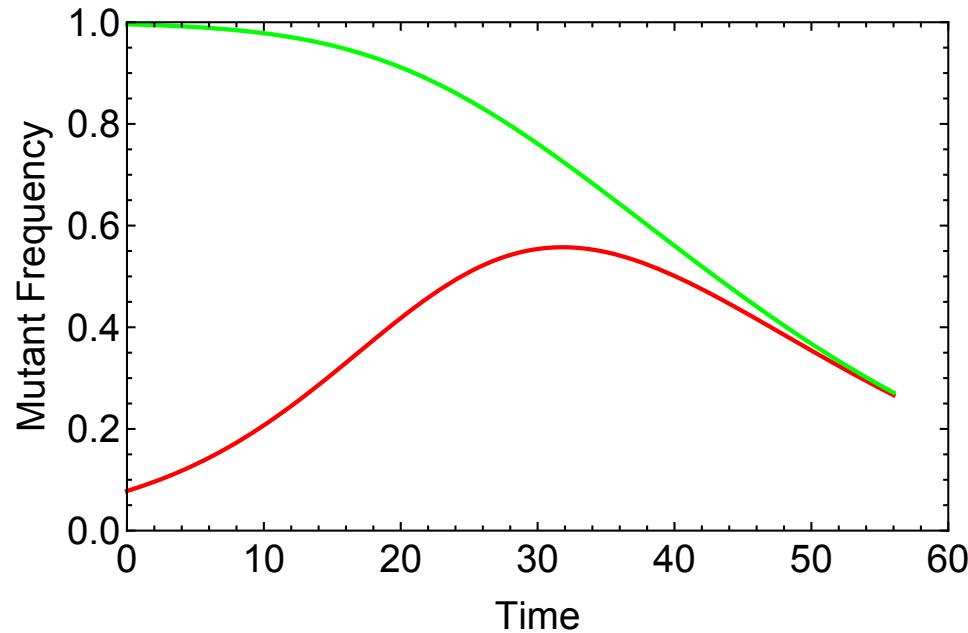
$$D_{ij} = 0.25$$



$$D_{ij} = -0.25$$

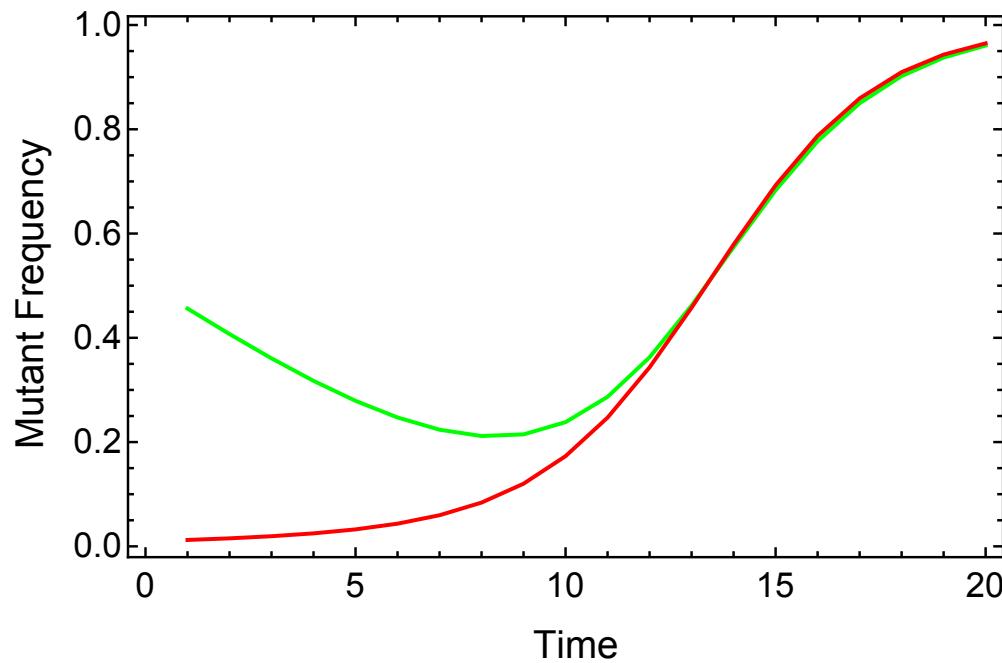
Clonal interference

A beneficial allele may be outcompeted by another



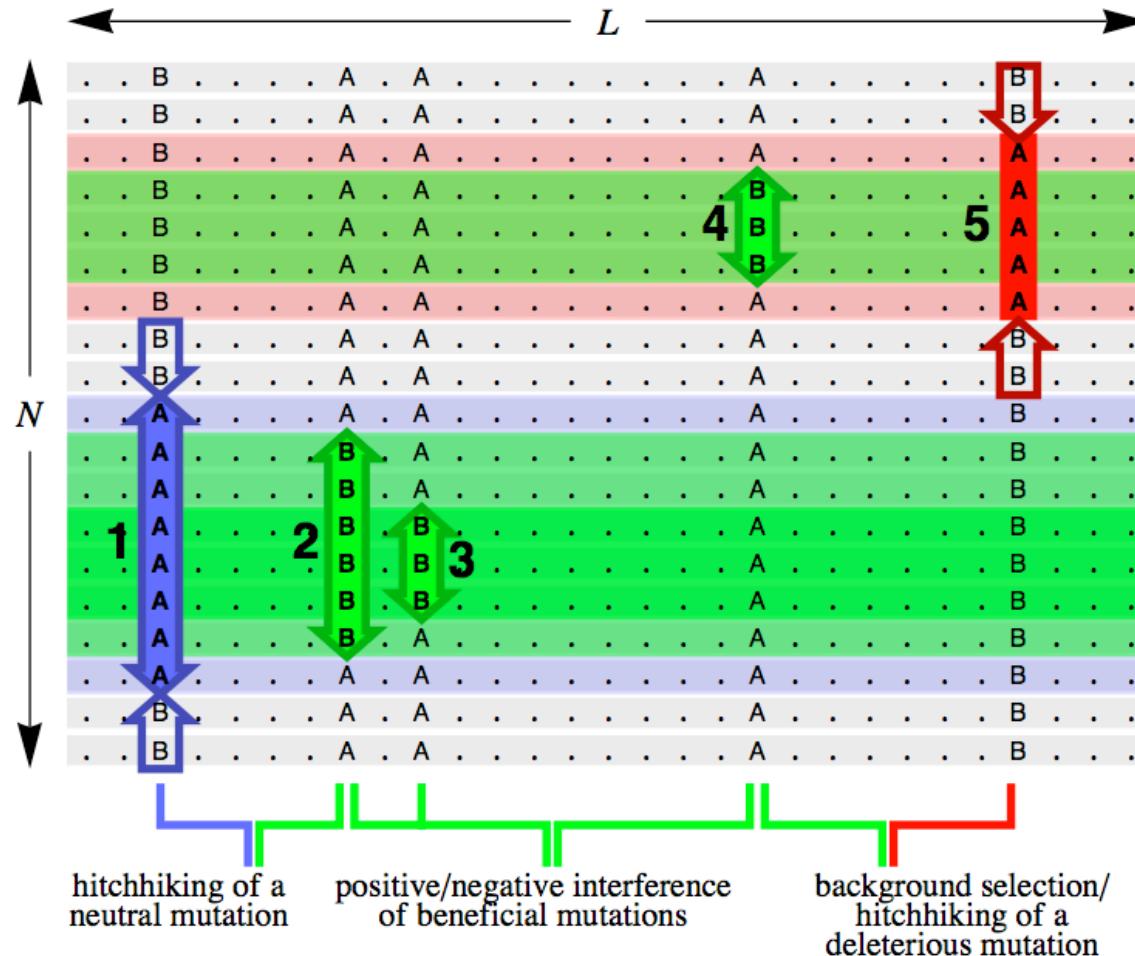
Genetic hitchhiking

A beneficial allele carries a deleterious allele to fixation



Selection across multiple loci

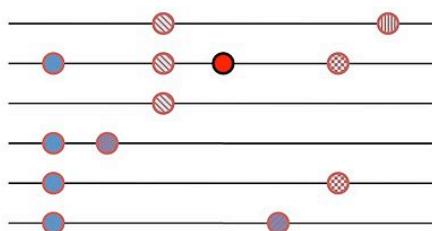
May have complex patterns of interference



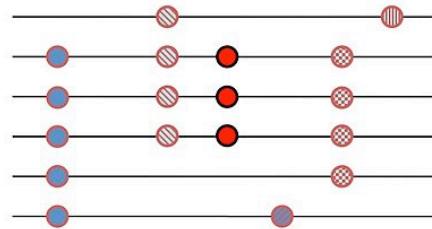
Selection across multiple loci

Selective sweep: Hard sweep

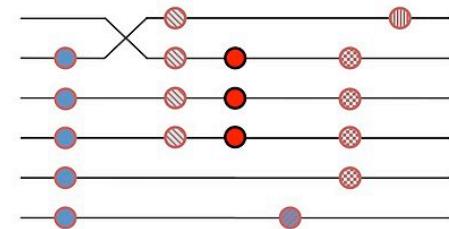
2. A beneficial mutation occurs (bright red dot on the second sequence).



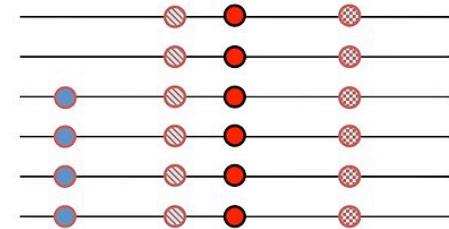
3. The beneficial mutation increases in frequency in the population, and so does the genomic background it is associated with.



4. A recombination event creates a new combination (the beneficial mutation is no longer always associated with the blue neutral mutation).



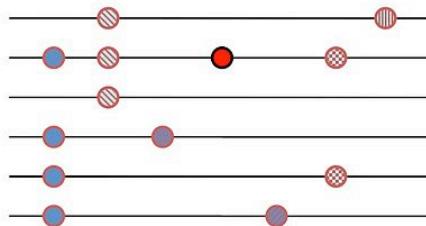
5. The beneficial mutation is fixed in the population. Close to the beneficial mutation there is now no genetic variation left, all sequences look the same. Three neutral mutations have hitchhiked along with the beneficial mutation and reached high frequency.



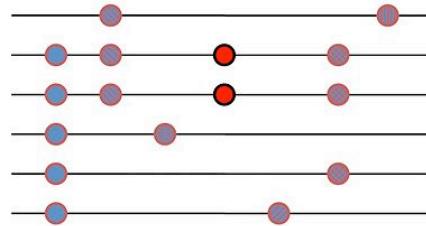
Selection across multiple loci

Selective sweep: Soft sweep

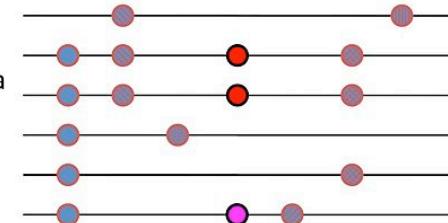
2. A beneficial mutation occurs (bright red dot on the second sequence).



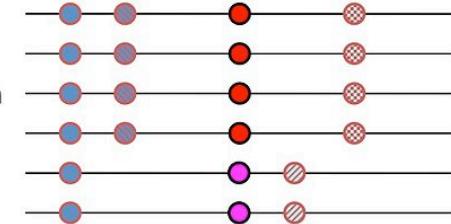
3. The beneficial mutation increases in frequency in the population, and so does the genomic background it is associated with.



4. The same (or very similar) beneficial mutation occurs on a second genomic background (pink dot on the 6th chromosome).

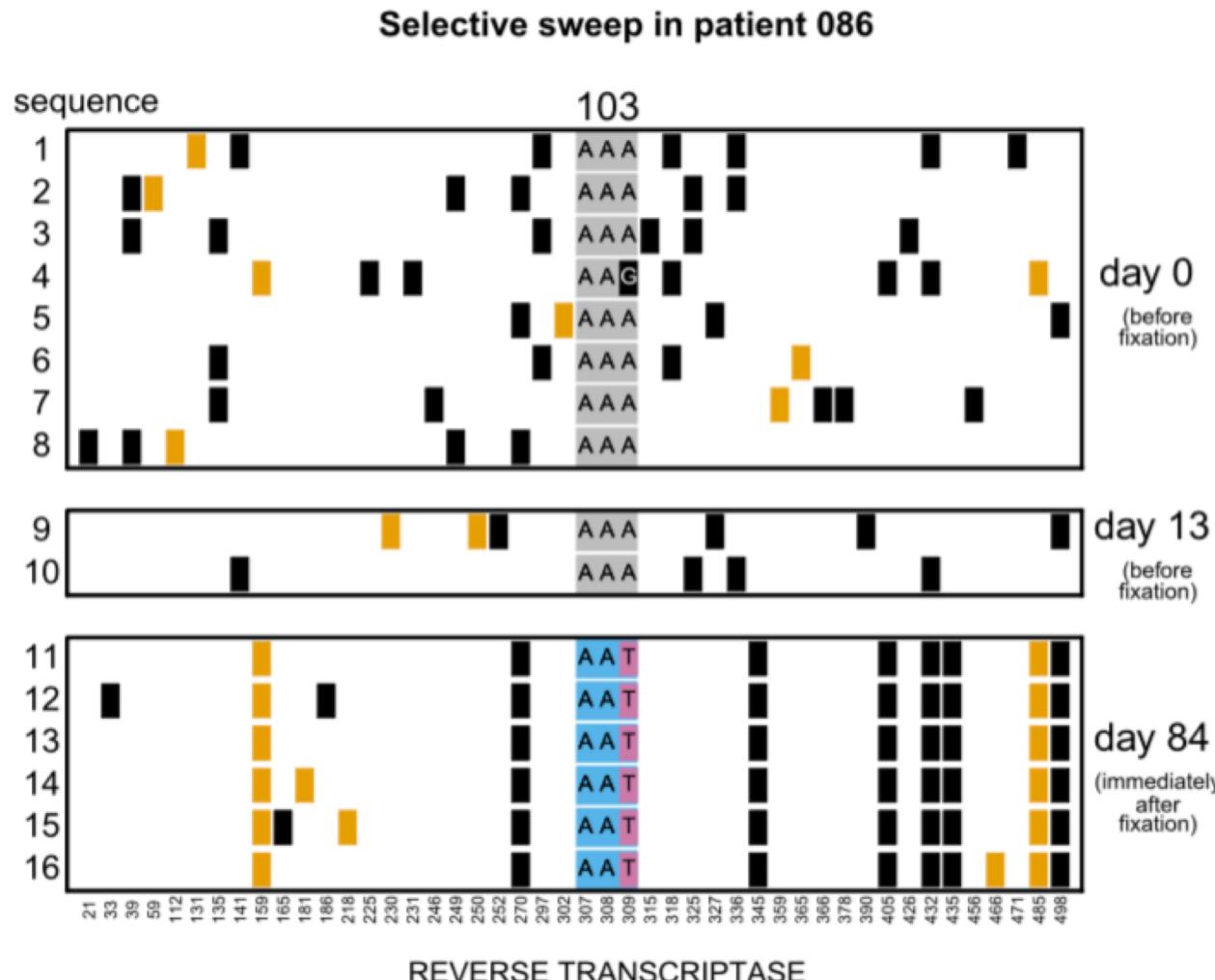


5. The beneficial mutations together fix in the population. Close to the beneficial mutation there is genetic variation left and the genomic region shows strong linkage disequilibrium.



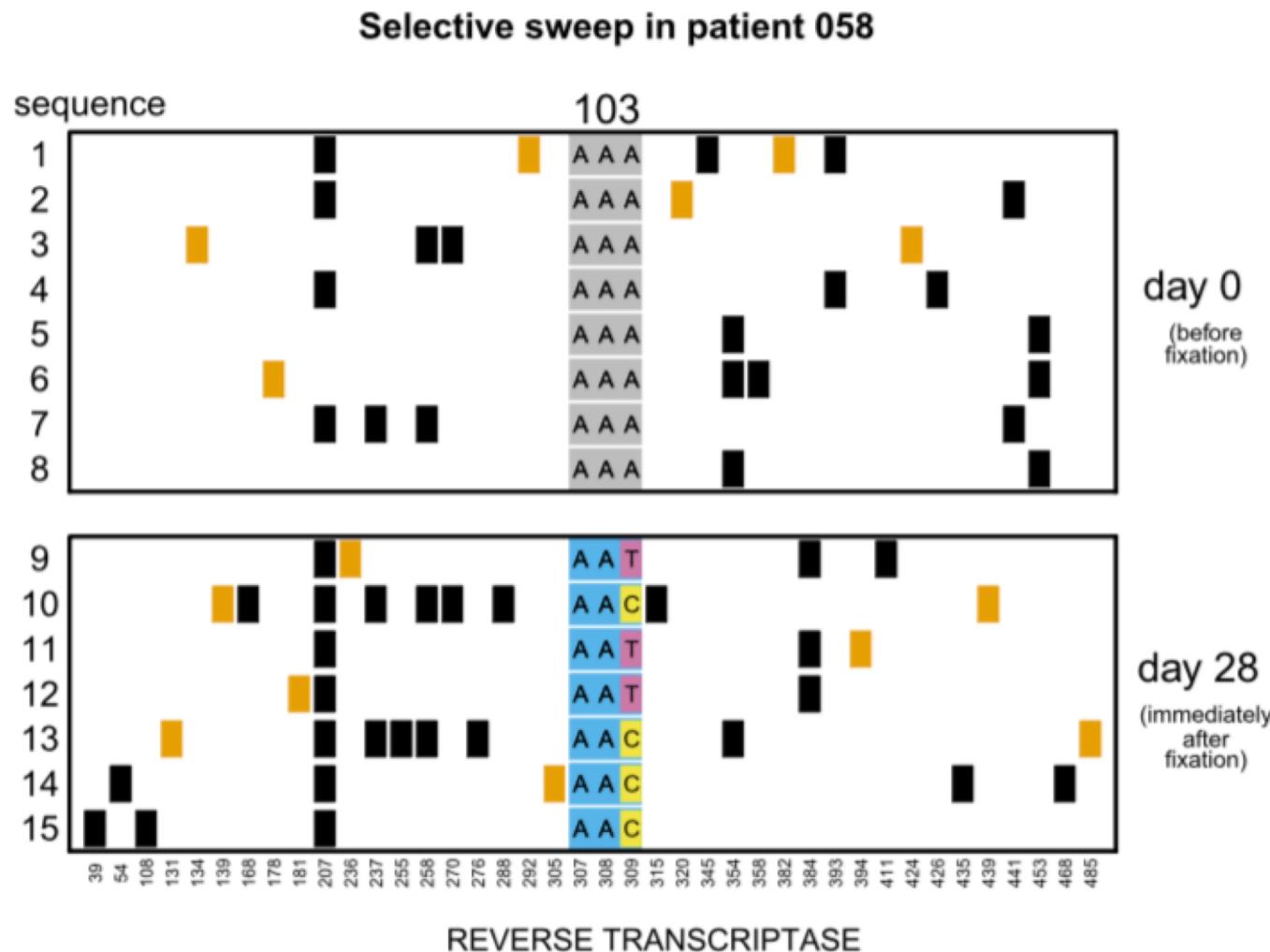
Application: Effective population size of HIV

Example of a hard sweep



Application: Effective population size of HIV

Example of a soft sweep

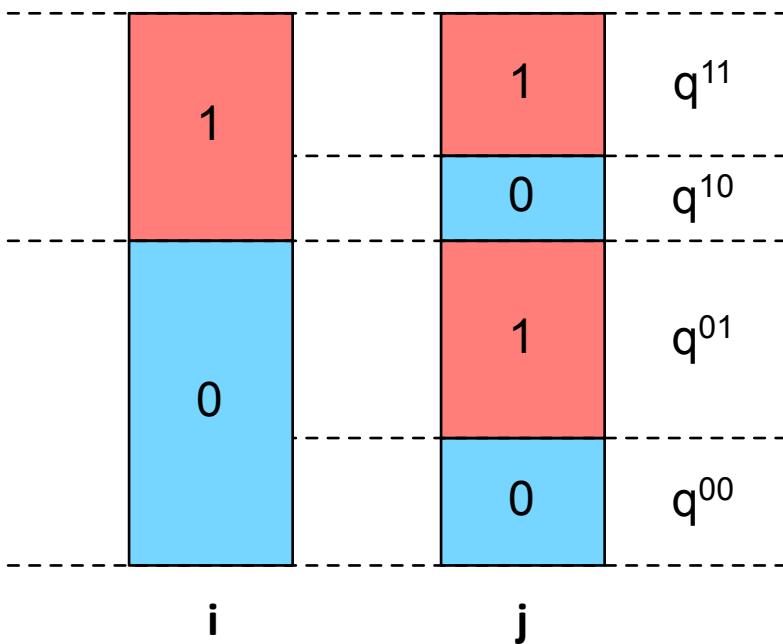


Selection across multiple loci

Special case: Driver-passenger model

Driver: $\frac{dq_i^1}{dt} = \sigma q_i^1(1 - q_i^1)$

Passengers:



$$\dot{q}_{ij}^{ab} = f_{ij}^{ab} q_{ij}^{ab} - q_{ij}^{ab} \left(\sum_{a', b' \in \{0, 1\}} f_{ij}^{a'b'} q_{ij}^{a'b'} \right)$$

$$\frac{dq_j}{dt} = \sigma D_{ij}$$

Selection across multiple loci

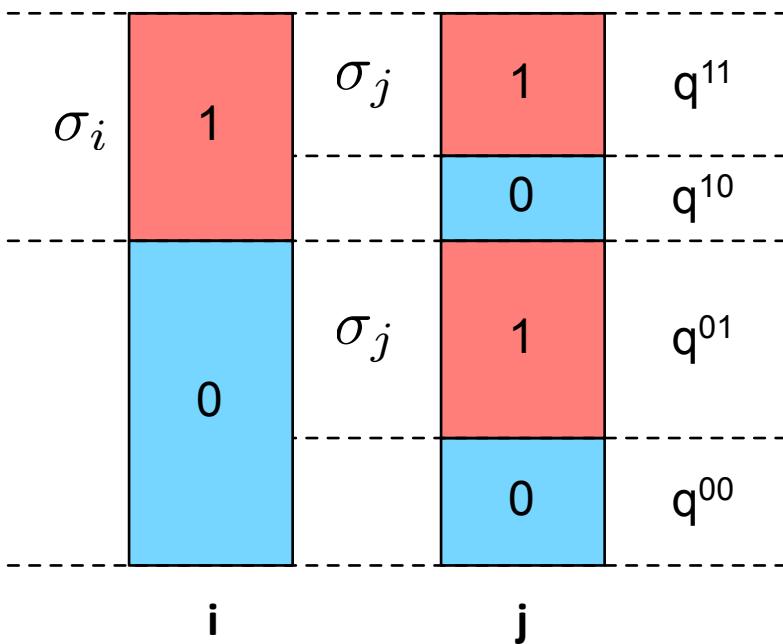
Special case: Additive selection

Interference between variants

$$\frac{dq_i^1}{dt} = \frac{d}{dt}(q_{ij}^{11} + q_{ij}^{10})$$

$$= (\sigma_i + \sigma_j)q_{ij}^{11} + \sigma_i q_{ij}^{10}$$

$$+ (q_{ij}^{11} + q_{ij}^{10})((\sigma_i + \sigma_j)q_{ij}^{11} + q_{ij}^{10} + \sigma_j q_{ij}^{01})$$



$$= \left[\sigma_i + \sigma_j \left(\frac{q_{ij}^{11}}{q_i^1} - \frac{q_{ij}^{01}}{q_i^0} \right) \right] q_i^1 (1 - q_i^1)$$

Inherent
selection

Interference

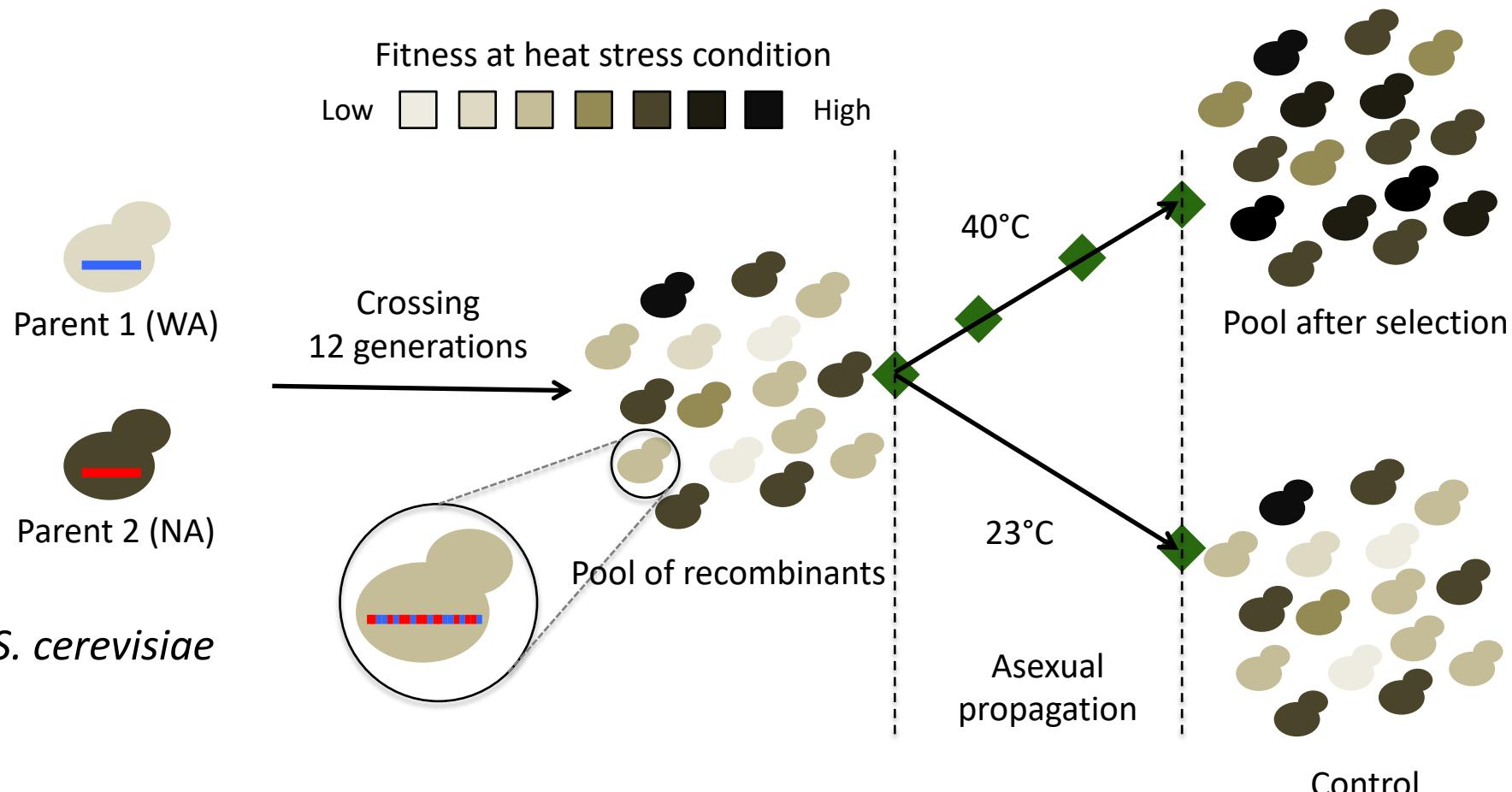
Selection across multiple loci

Special case: Additive selection

Multiple variants – additive model

$$\frac{dq_i^1}{dt} = \left[\sigma_i + \sum_k \sigma_k \left(\frac{q_{ik}^{11}}{q_i^1} - \frac{q_{ik}^{01}}{q_i^0} \right) \right] q_i^1 (1 - q_i^1)$$

Example III: Selection across multiple loci

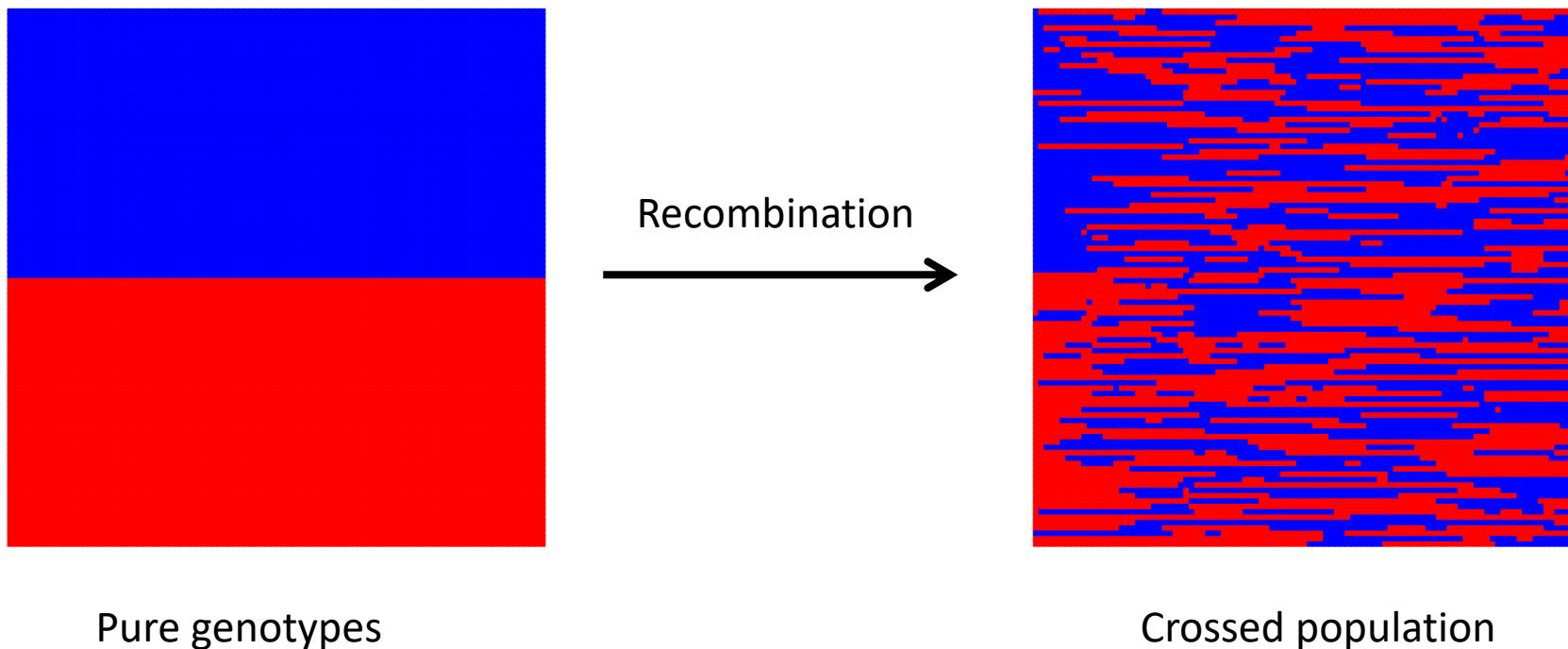


◆ Whole-genome sequencing of the pool

Figure adapted from Parts et al, Genome Research 2011

Genetic structure following recombination

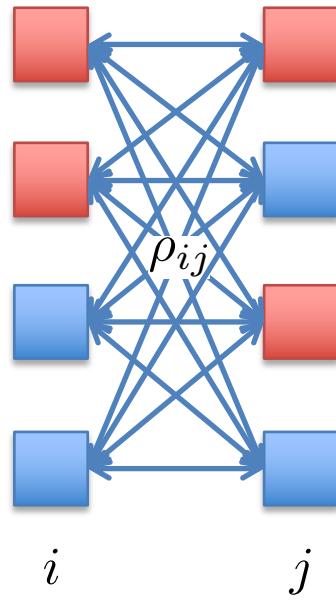
Links between alleles are statistically comprehensible



Linkage disequilibrium

LD and Recombination

Recombination breaks linkage disequilibrium



Allele frequencies unchanged by recombination

Among non-recombinant sequences:

$$q_{ij}^{11}(t+1) = q_{ij}^{11}(t)$$

Among recombinant sequences:

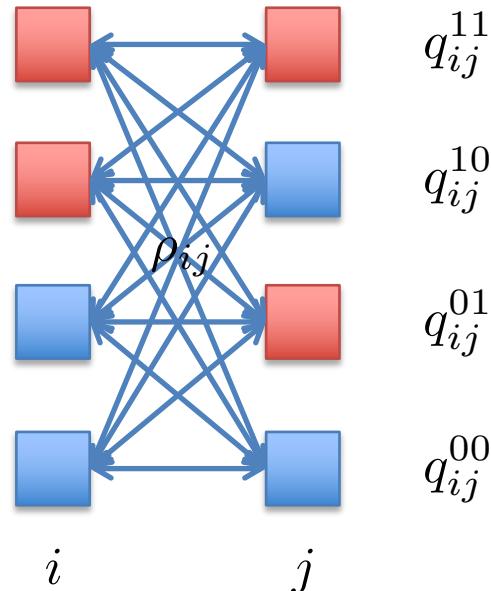
$$q_{ij}^{11}(t+1) = q_i^1(t)q_j^1(t)$$

$$D_{ij} = q_{ij}^{11} - q_i^1 q_j^1$$

Linkage disequilibrium

LD and Recombination

Recombination breaks linkage disequilibrium



$$q_{ij}^{11}$$

$$q_{ij}^{10}$$

$$q_{ij}^{01}$$

$$q_{ij}^{00}$$

$$i$$

$$j$$

$$D_{ij} = q_{ij}^{11} - q_i^1 q_j^1$$

$$q_{ij}^{11}(t+1) = (1 - \rho_{ij})q_{ij}^{11}(t) + \rho_{ij}q_i^1(t)q_j^1(t)$$

$$D_{ij}(t+1) = (1 - \rho_{ij})D_{ij}(t)$$

Linkage disequilibrium decays to zero exponentially over time

Effective rate of recombination scales with distance

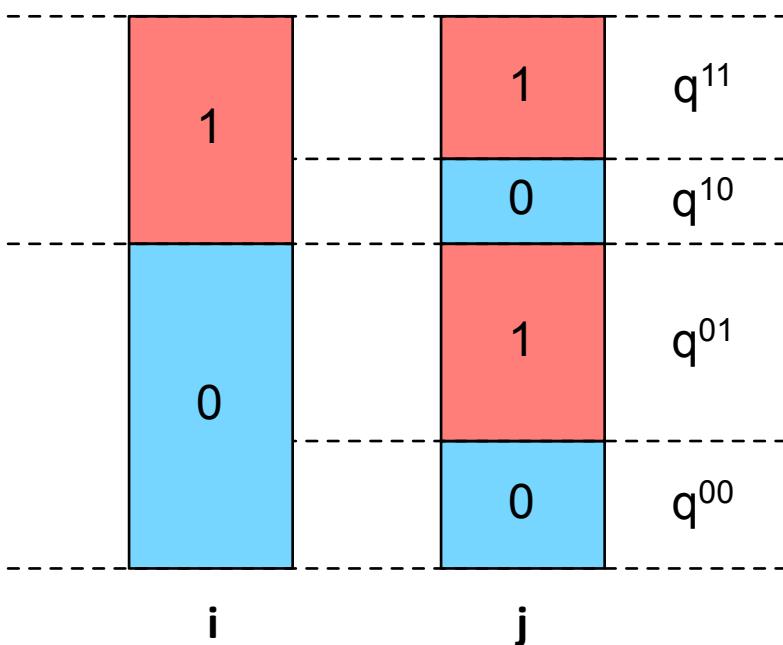
Selection across multiple loci

Driver-passenger model

Driver: $\frac{dq_i^1}{dt} = \sigma q_i^1(1 - q_i^1)$

$$\frac{dq_j^1}{dt} = \frac{d}{dt}(q_{ij}^{11} + q_{ij}^{01})$$

Passengers:



$$= \sigma D_{ij}$$

$$D_{ij}(\rho, \Delta_{ij}) = D'_{ij}(1 - \rho\Delta_{ij})^{N_c}$$

N_c : Generations of crossing

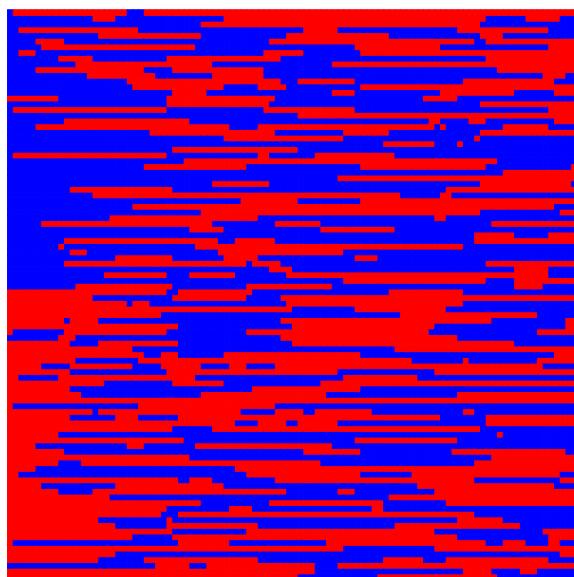
D_{ij}' : Initial LD between i and j

ρ : Recombination rate

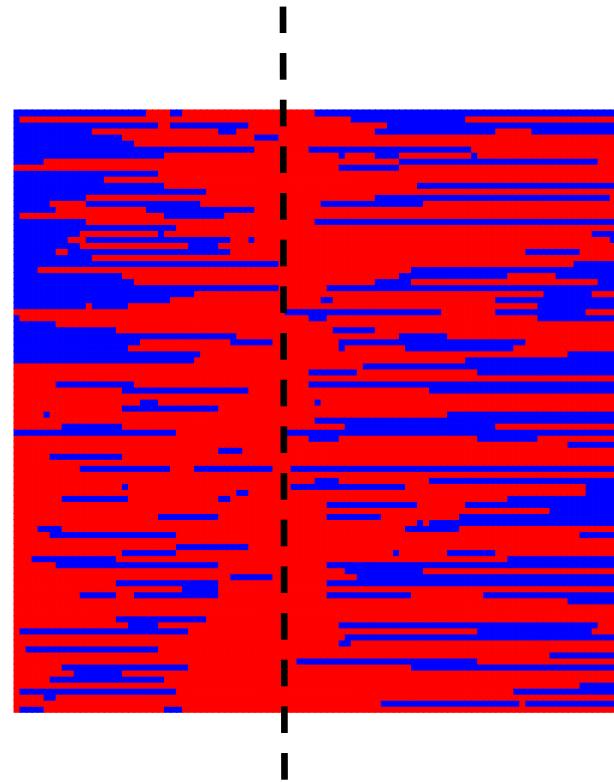
Δ_{ij} : Distance in genome between i and j

Genetic crossing experiments

Experimental design



Selection →

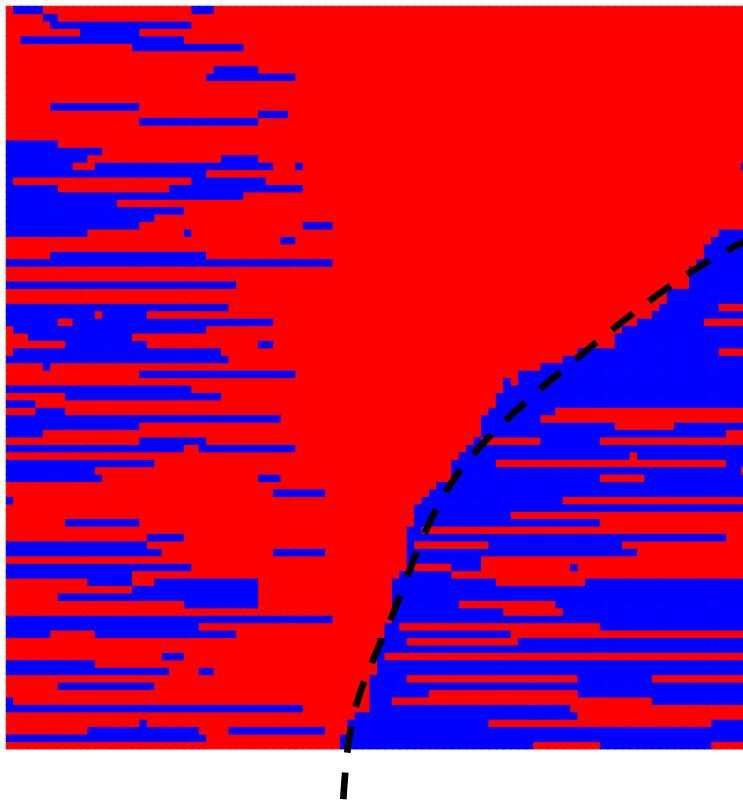


Crossed population

Selected cross

Genetic crossing experiments

Mathematical order to selected population



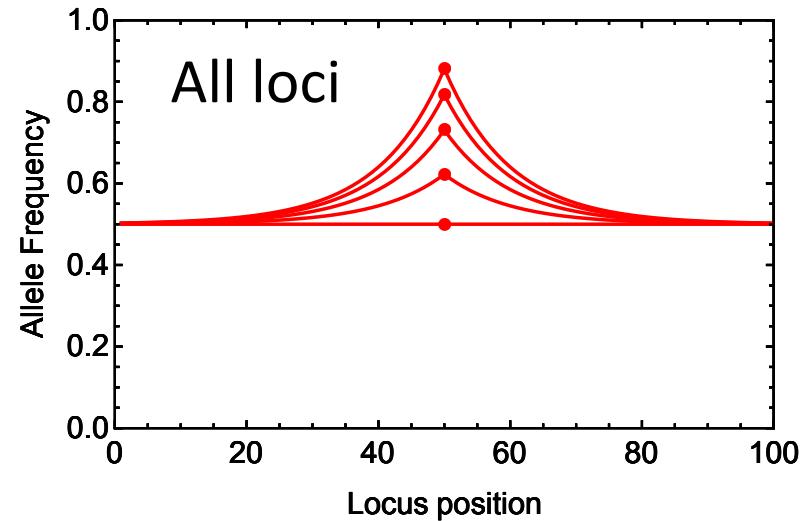
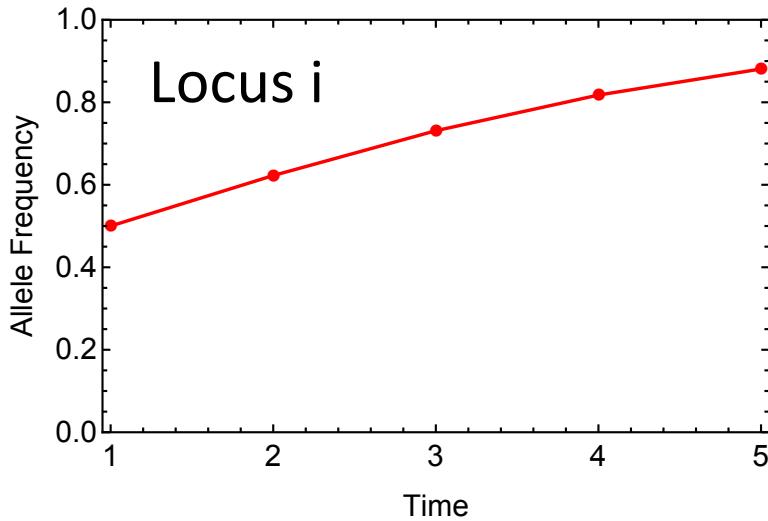
Recombination events occur at a range of distances from the selected variant

Evolutionary dynamics

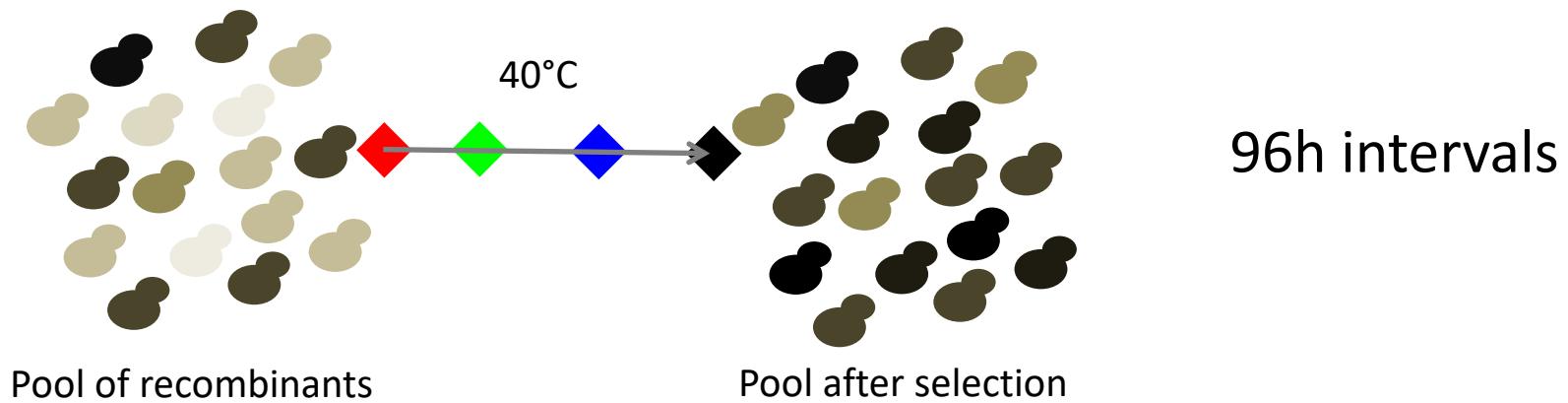
Selection for allele at locus i $\frac{dq_i^1}{dt} = \sigma q_i^1(1 - q_i^1)$

Change in frequency at locus i = Δq_i

Change in frequency at locus j = $\Delta q_i D'_{ij} (1 - \rho \Delta_{ij})^{N_c}$



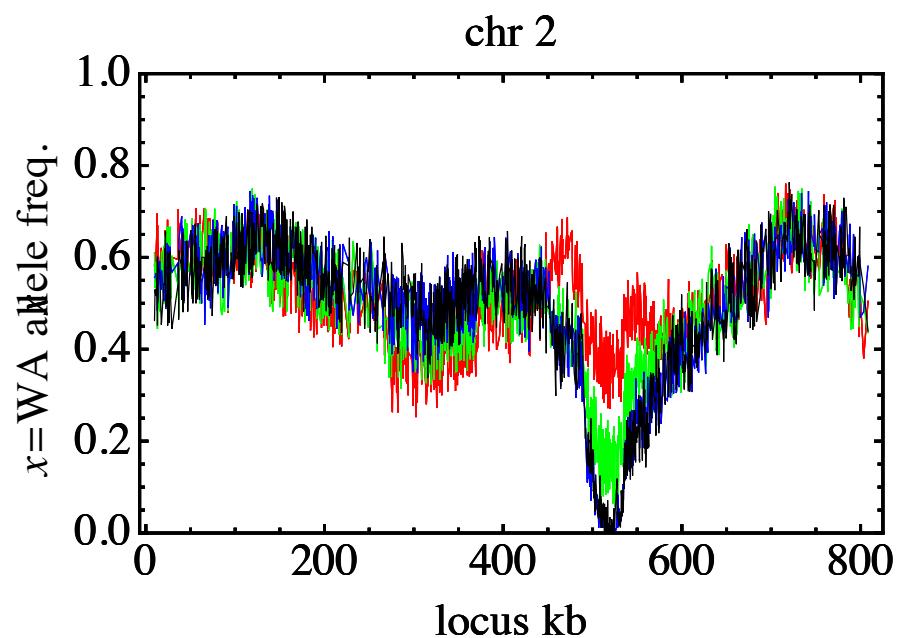
Time-resolved data



Pool of recombinants

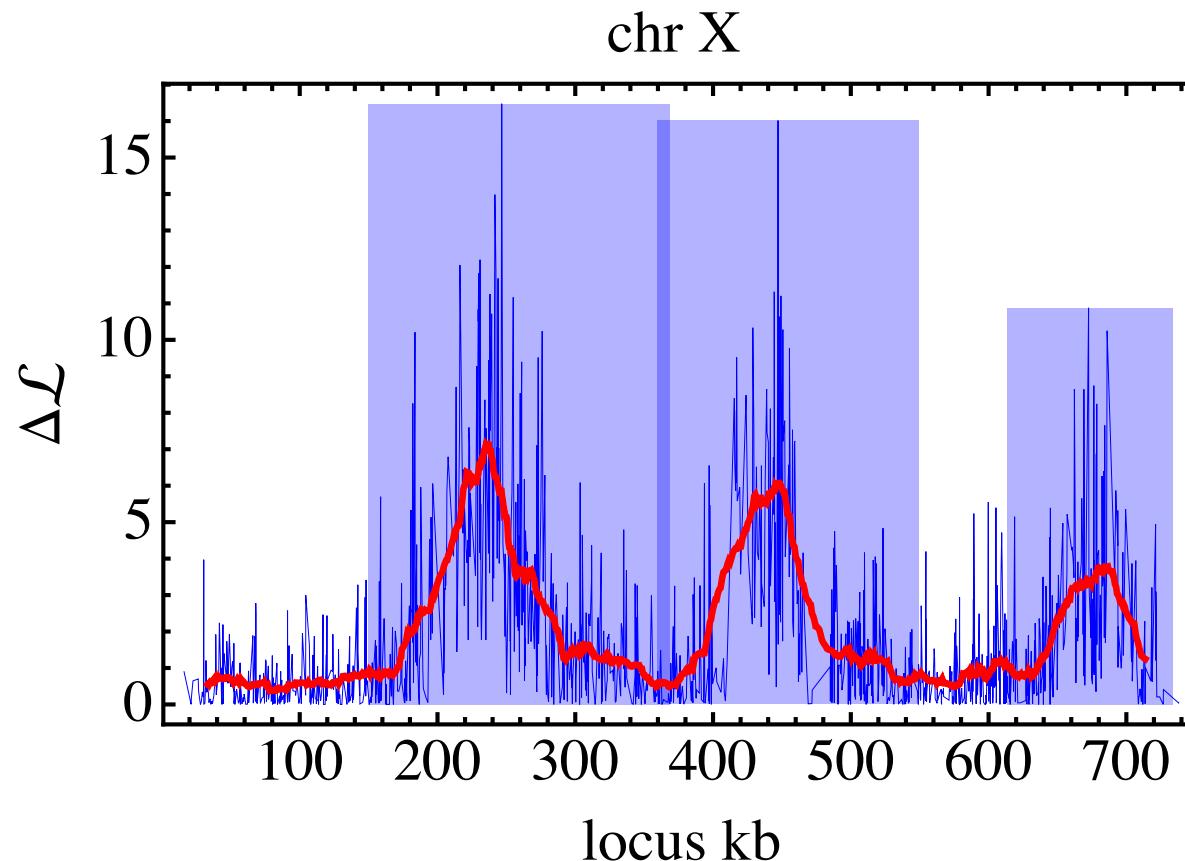
Pool after selection

Data from
chromosome 2



Example III: Selection across multiple loci

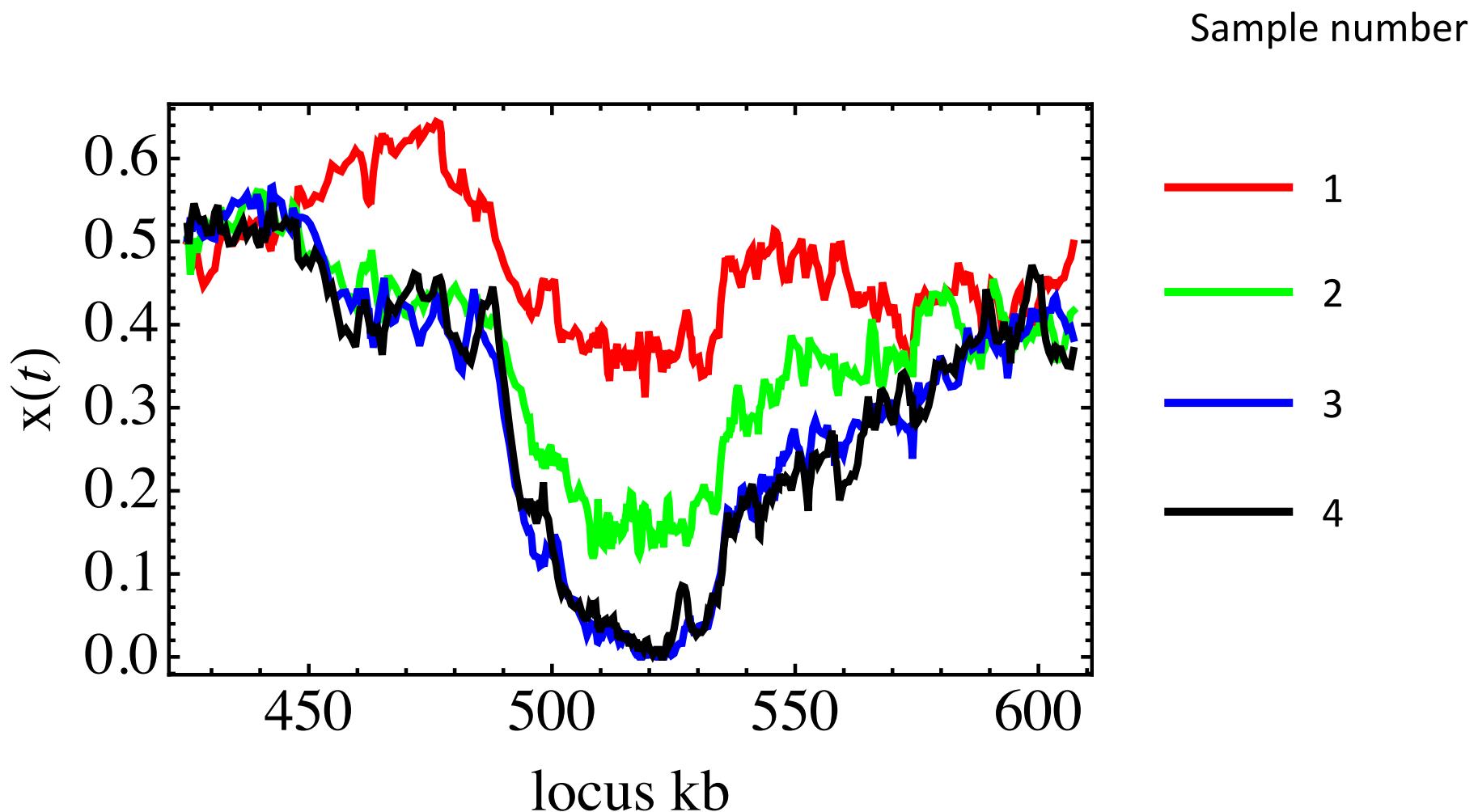
Regions of “non-neutral” behaviour



Consider each region in turn

Experimental data

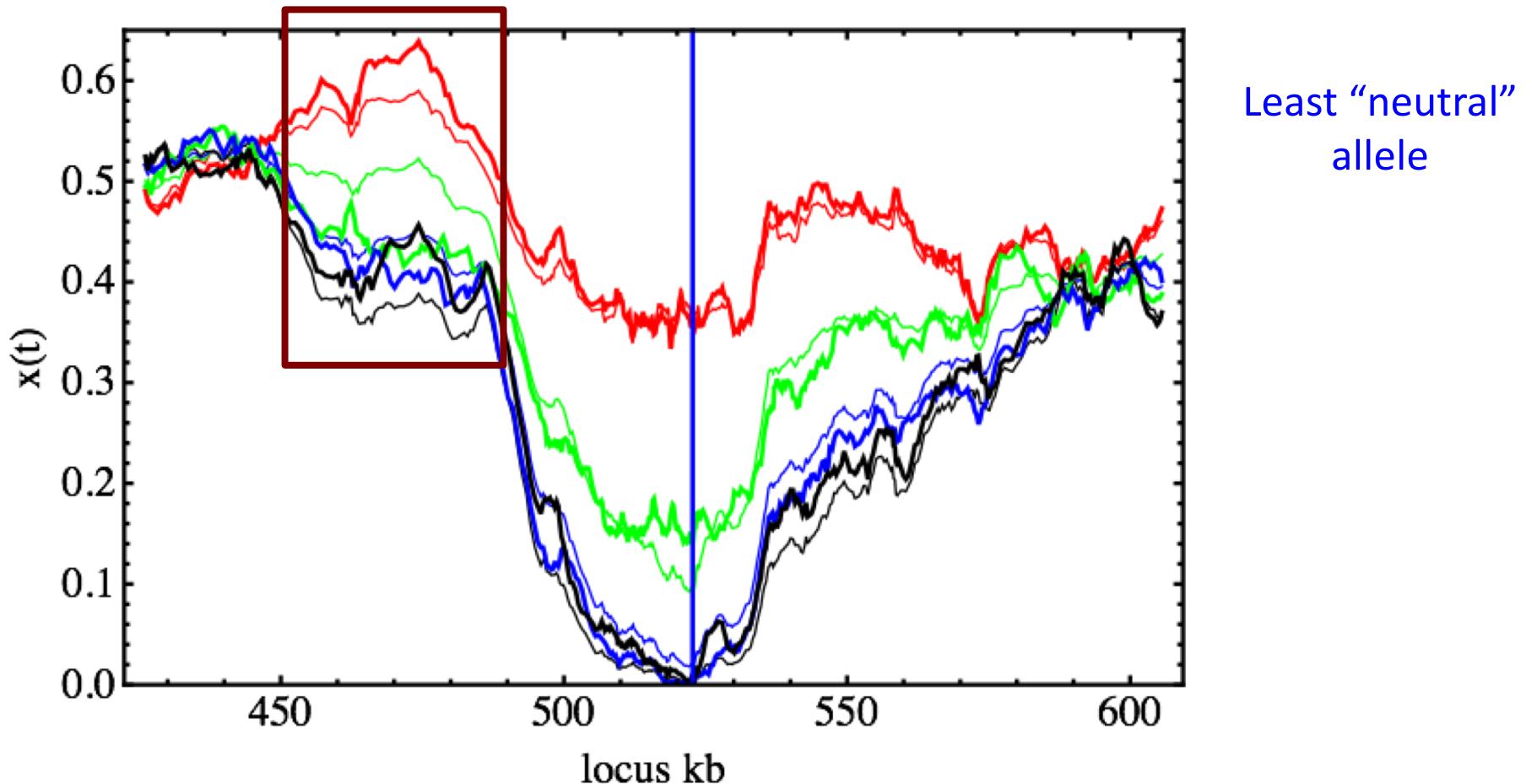
Region in chromosome II



Model comparison

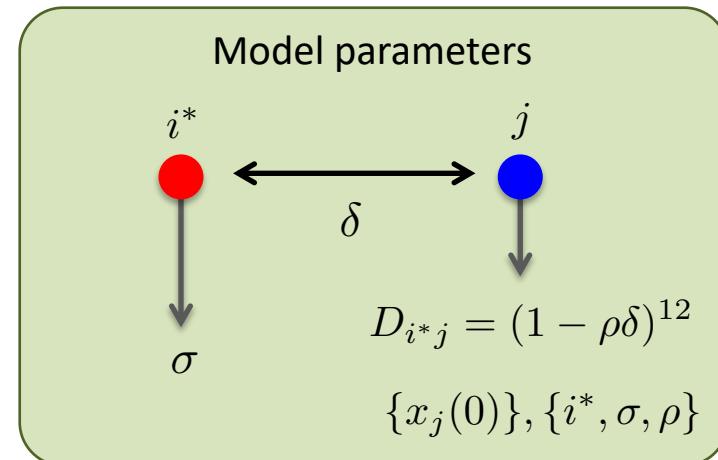
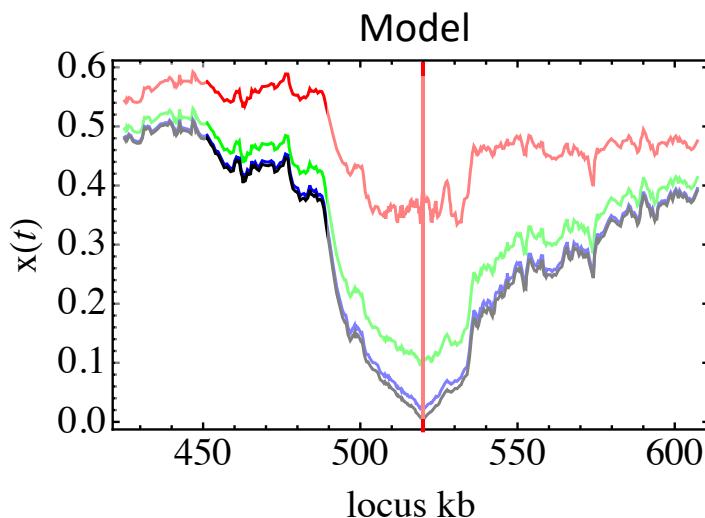
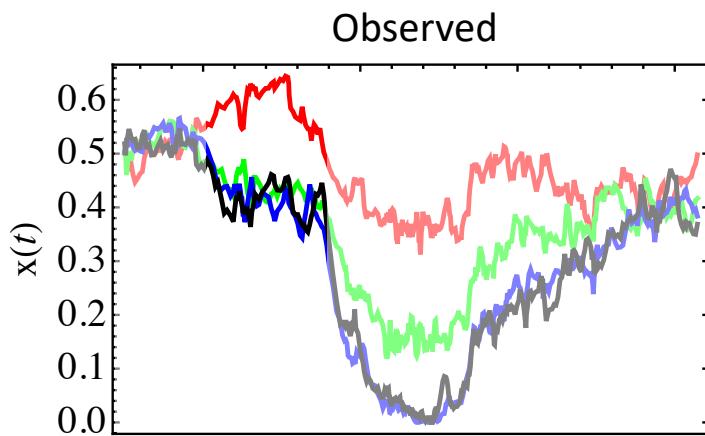
Region in chromosome II

Single-locus model: Each allele moves under its own selection



Model : Driver and passengers

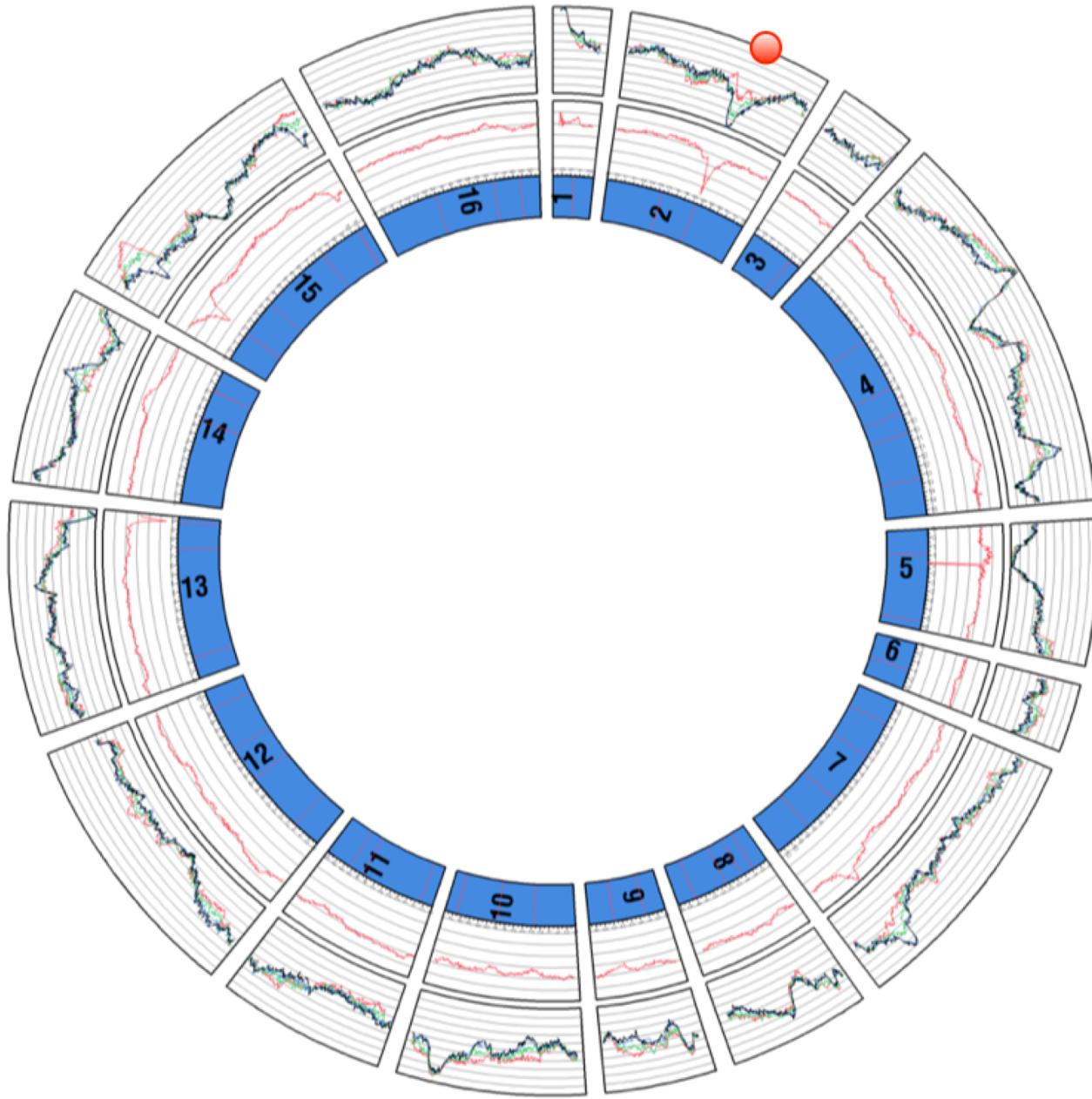
Allele frequencies change via linkage to a single 'driver' allele.



$$\Delta x_i^* = \sigma x_i^* (1 - x_i^*)$$

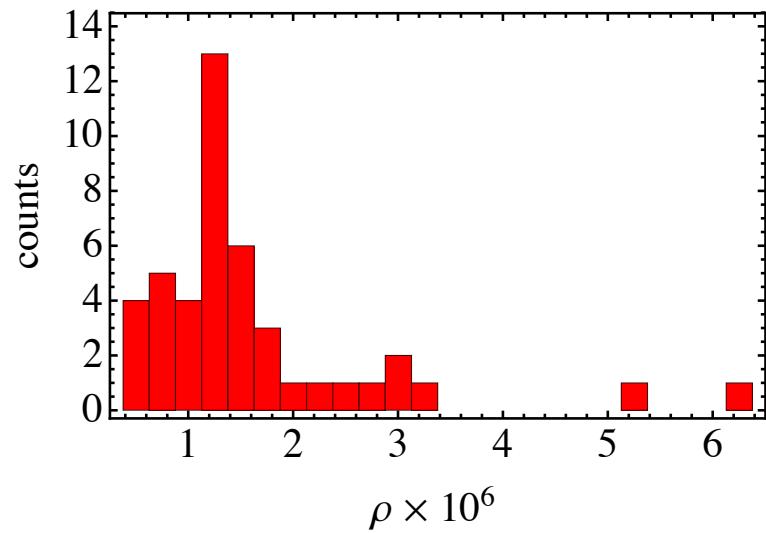
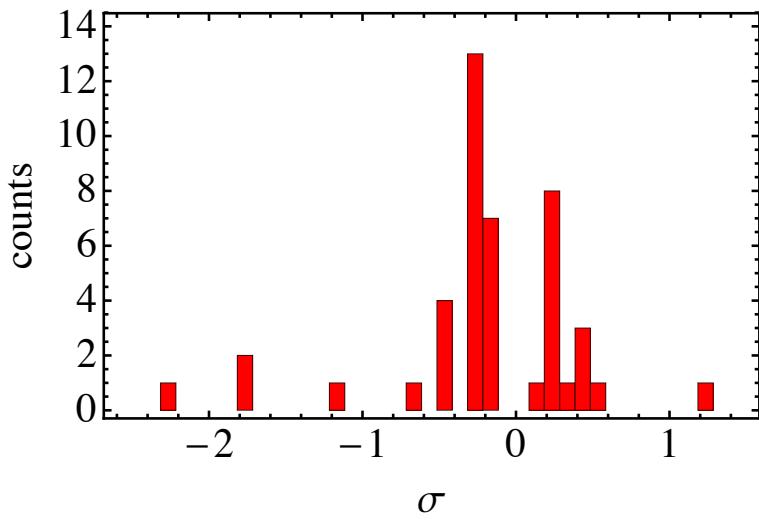
$$\Delta x_j = \sigma D_{i^*j}$$

Linkage explains fixing at intermediate frequency



Inference of selection genome-wide

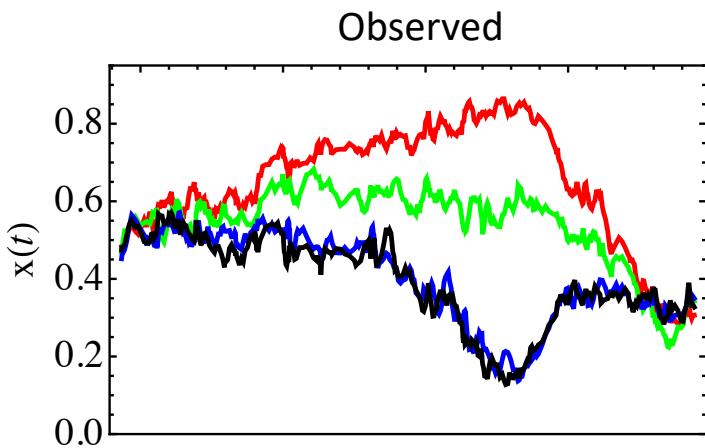
44 alleles under selection



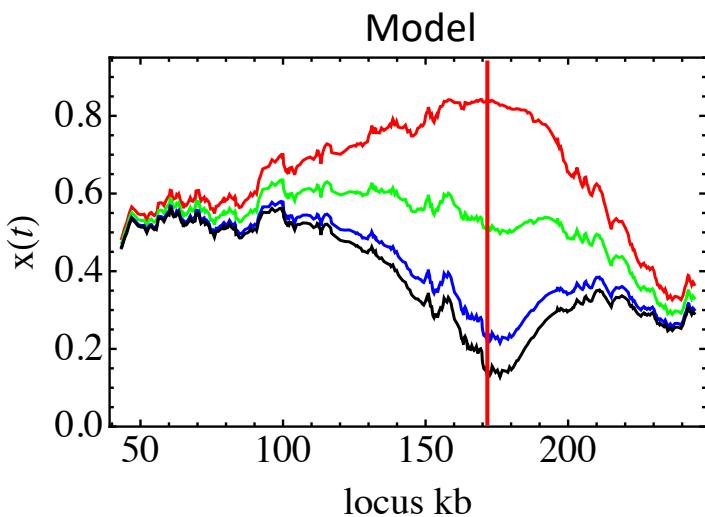
Selection for both strains: Low heat-tolerance strain has some beneficial genes for high temperatures

Limitations to the model

Some data visibly do not fit the model

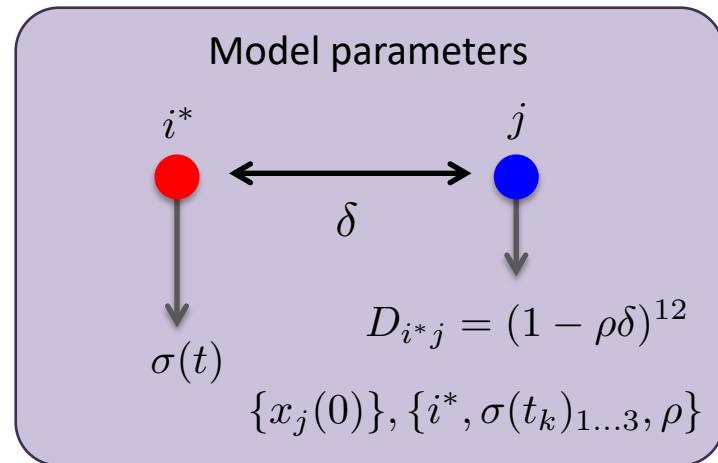
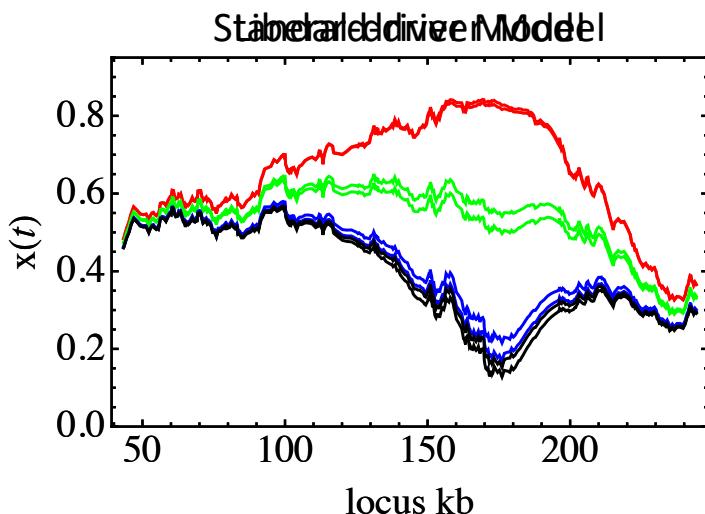
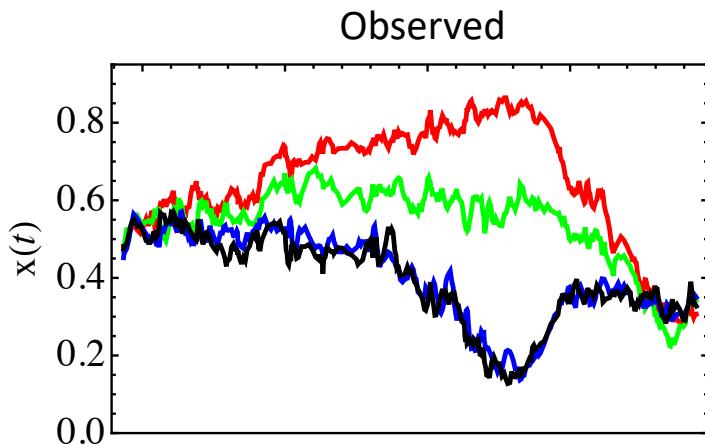


Chromosome XV: Apparent driver
stops at intermediate frequency



Liberal-driver and passengers

New Model: Allele frequencies change via linkage to a single 'liberal-driver' allele, which can move arbitrarily.



Data from chromosome XV

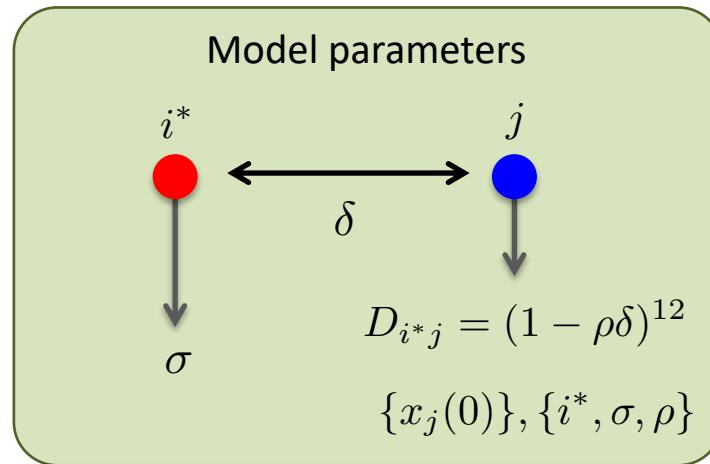
Liberal-driver model gave an improved AIC in 38/44 regions

Mean selection coefficients and recombination rates very similar

Liberal drivers

The liberal driver model detects effects going beyond our driver-passenger model

?



It does not specify their cause

Liberal drivers

The liberal driver model detects effects going beyond our driver-passenger model

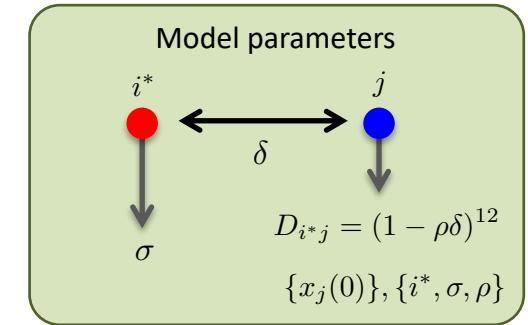
Random linkage?



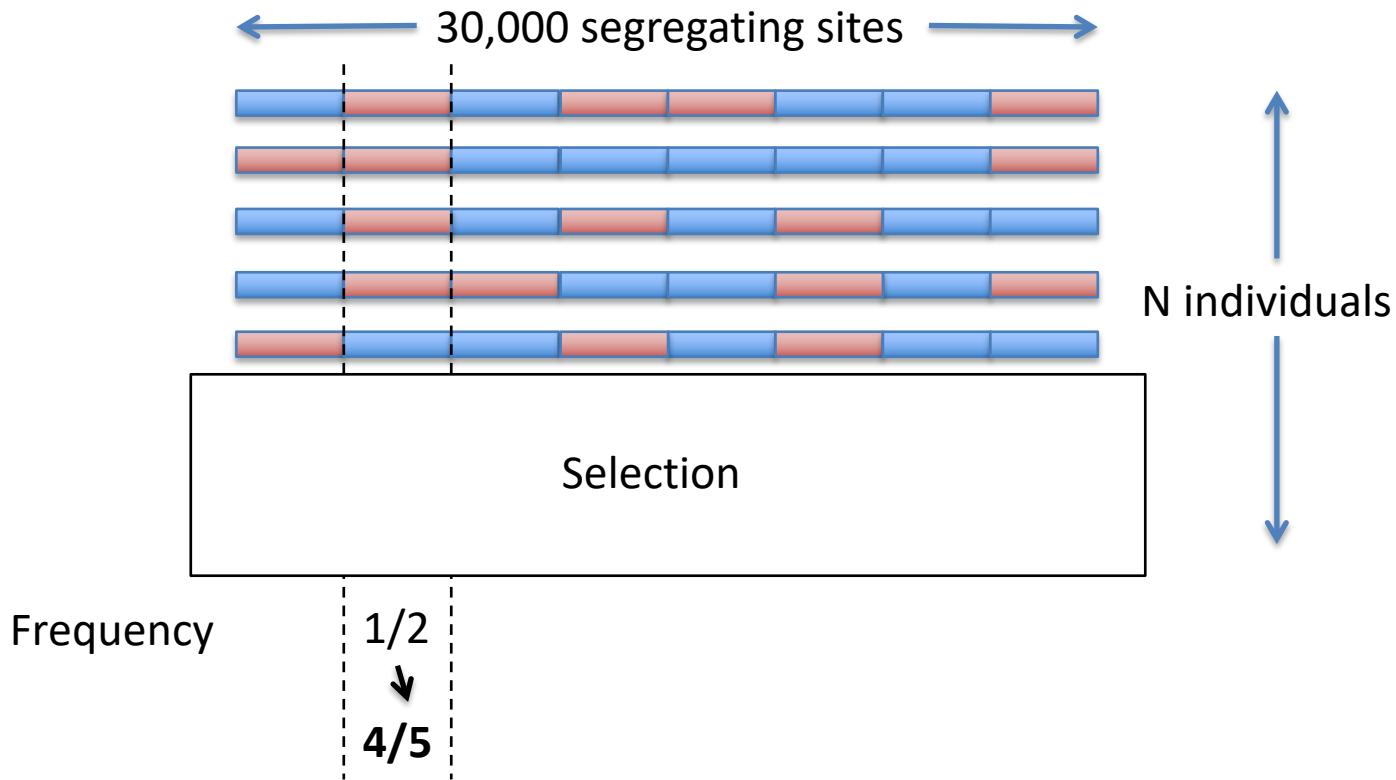
Long-range linkage between drivers?

Frequency-dependent selection?

Epistasis?



Random linkage



Initial pool : Linkage between sites is random. c. 4.5×10^8 linkages

Selection removes a number of haplotypes from the pool

Random linkage

Is the population size large enough to avoid random linkage?

Simulations run at different population sizes:
Looked for false identifications of drivers

Simulated population size	Mean # false positive identifications
10^5	21.9
10^6	8.5
5×10^6	1.4

Simulation parameters

Size $N=10^5 - 5 \times 10^6$

Genome length 12Mb

30000 equally-spaced segregating sites

Crossing protocol : Uniform recombination, 12 gen

Selection protocol: 30 generations, additive selection

Driver locations and selection coefficients as inferred

Random linkage is a substantial effect for population sizes c. 10^5

No substantial random linkage at population size 10^7

Liberal drivers

The liberal driver model detects effects going beyond our driver-passenger model

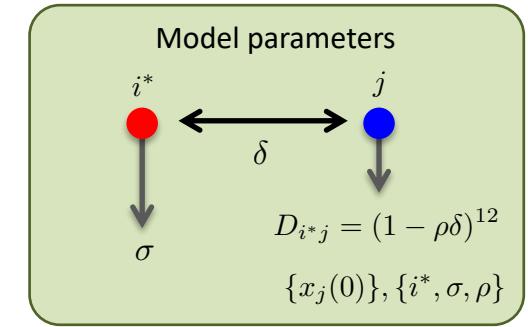
~~Random linkage?~~



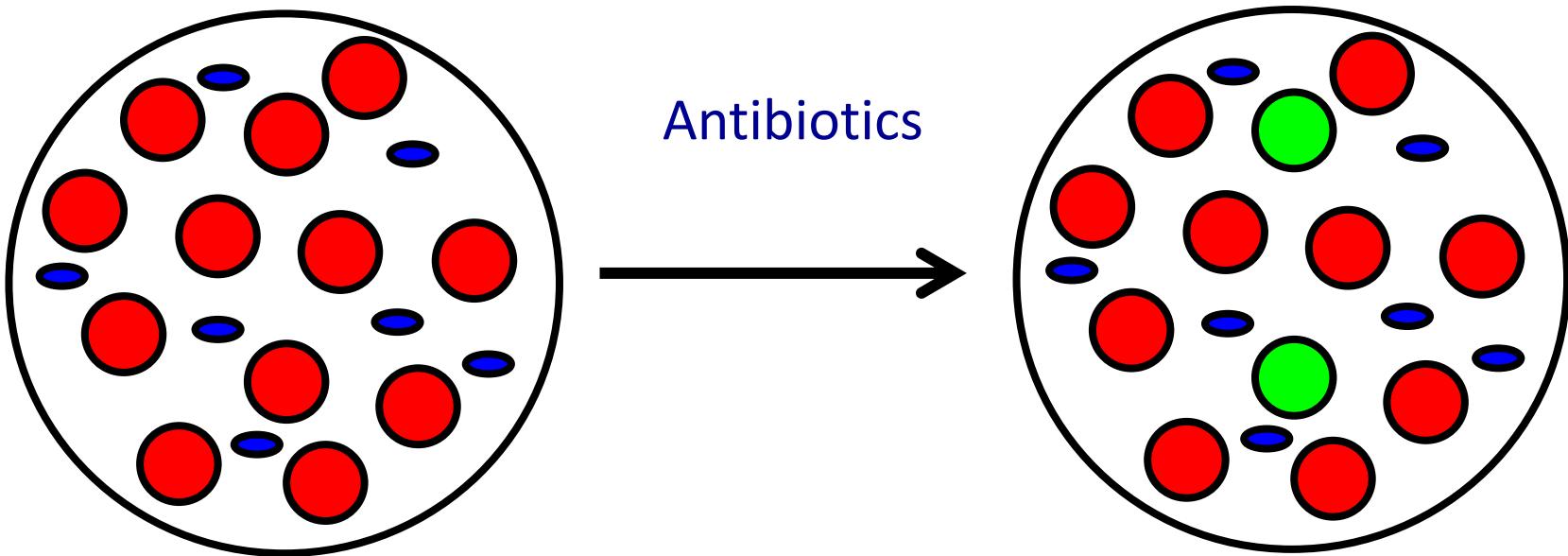
Long-range linkage between drivers?

Frequency-dependent selection?

Epistasis?



Frequency-dependent selection



A certain frequency of a given phenotype benefits the entire population

Liberal drivers

The liberal driver model detects effects going beyond our driver-passenger model

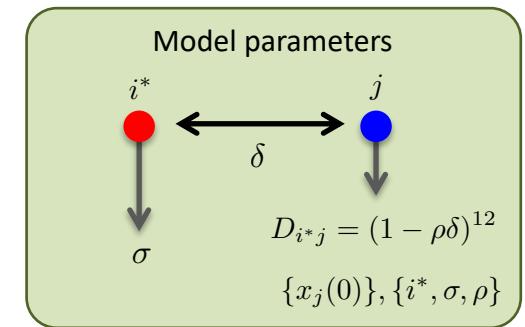
~~Random linkage?~~

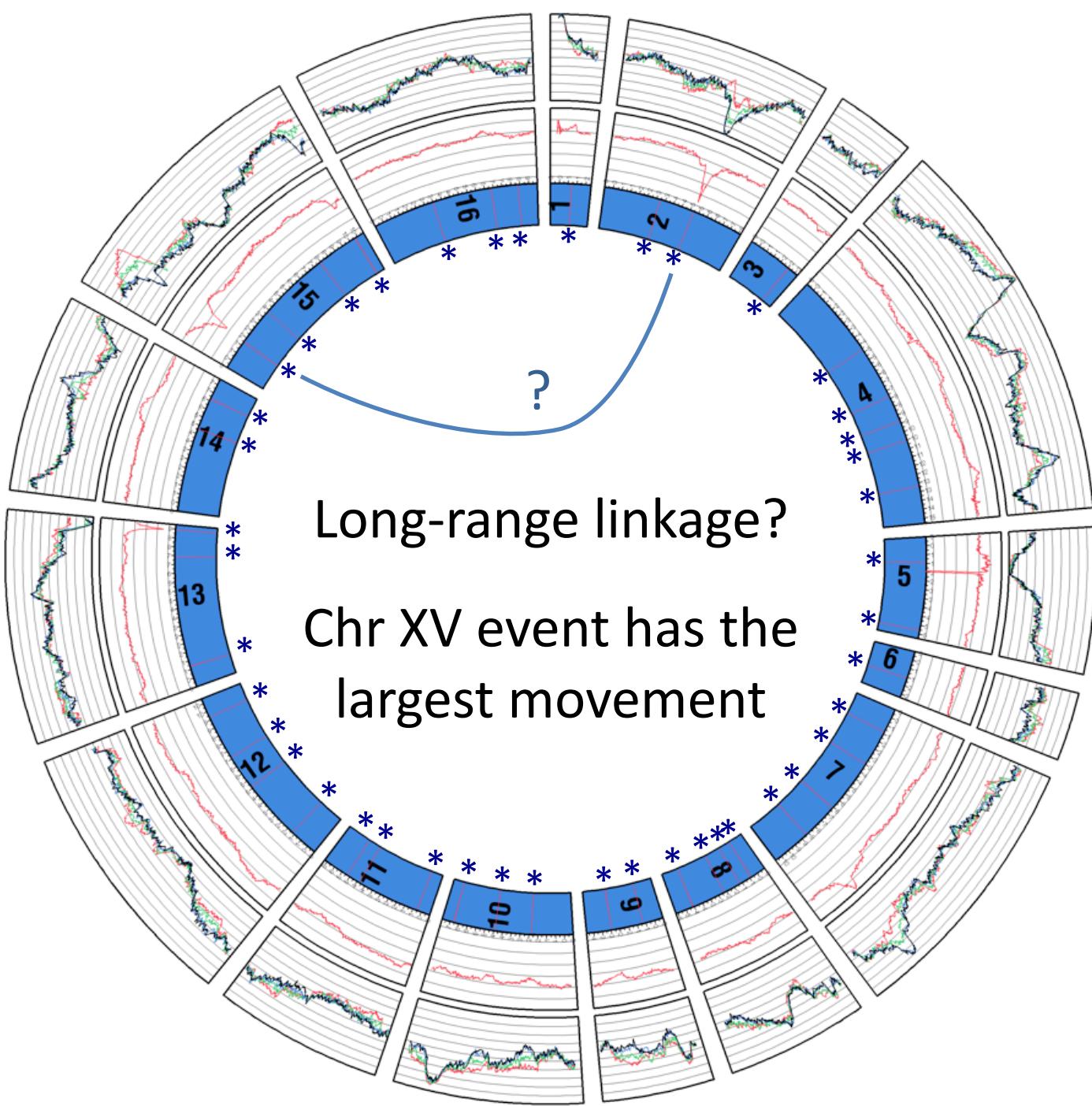


Long-range linkage between drivers?

Frequency-dependent selection?

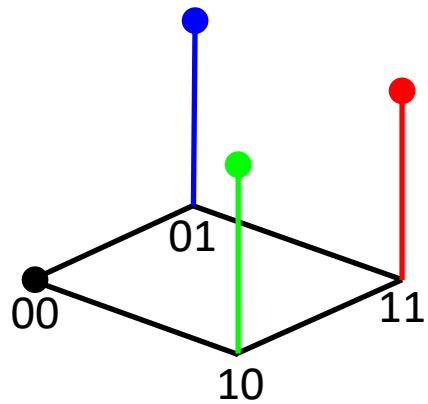
Epistasis?



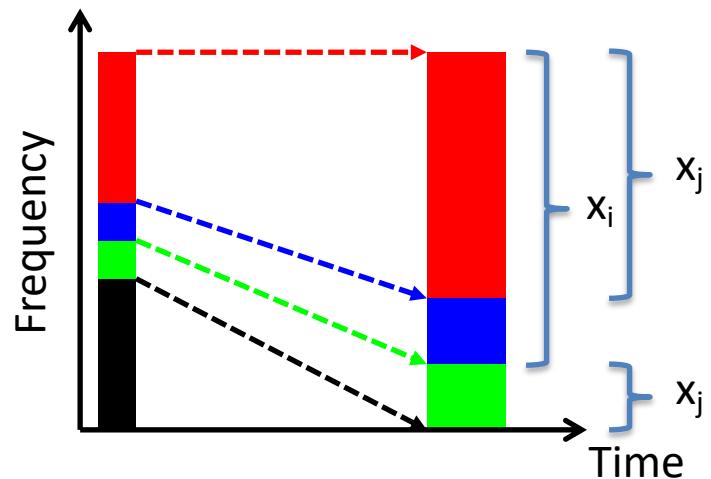


Epistasis

One explanation: epistasis can lead to changes in allele frequency which stop at intermediate values:



Example haplotype
fitnesses

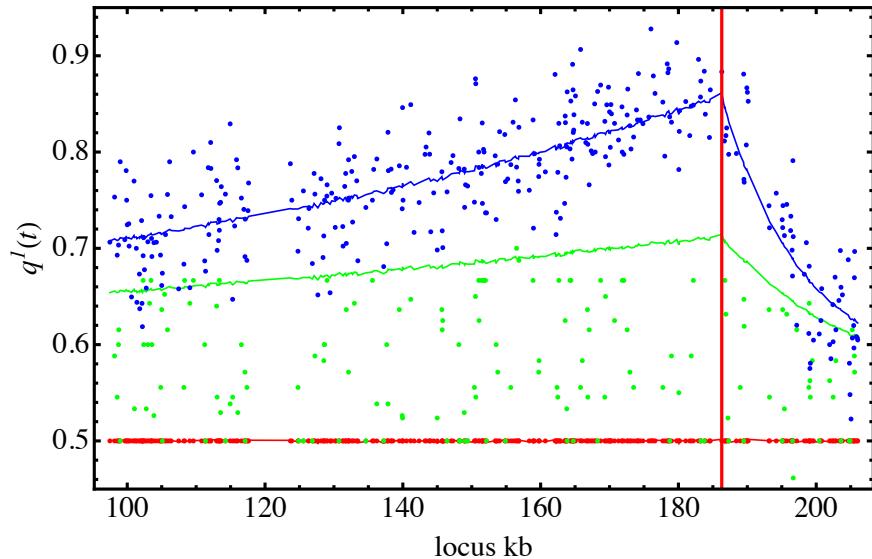


Evolution of haplotype
frequencies

Variation in recombination rate

Extension of the model to incorporate multiple recombination rates

Data from chromosome XV before, during, and after the cross

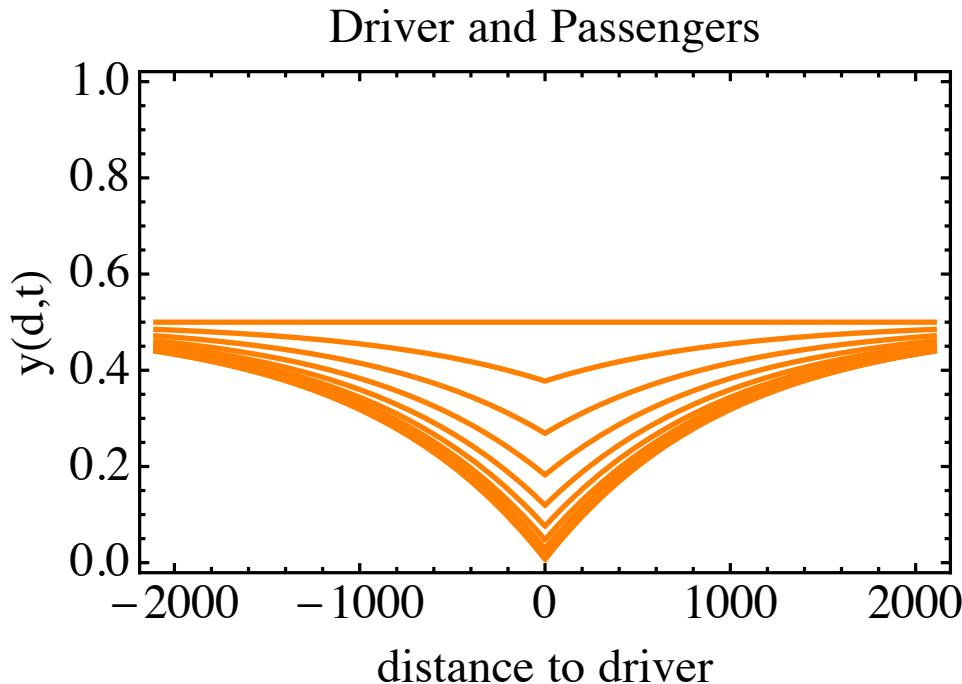


Model optimised for two recombination rates, either side of the driver

Substantial variation in recombination rate across the region of Chr XV

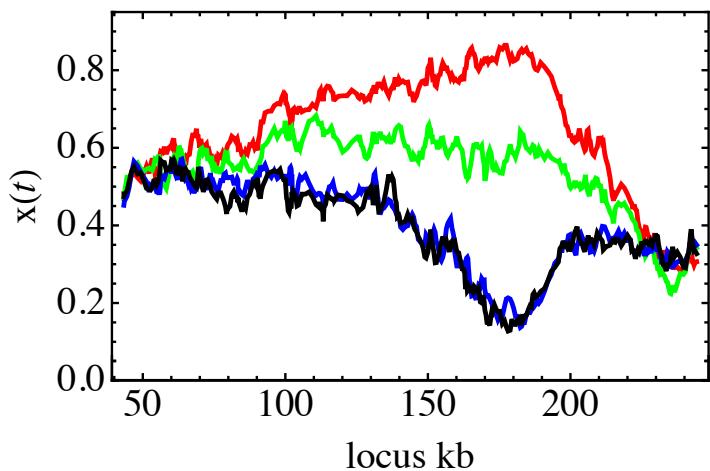
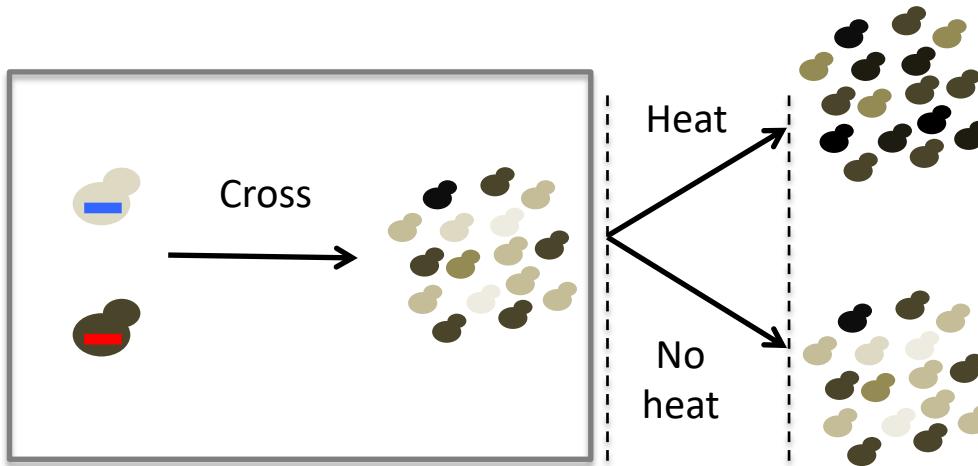
Re-examining the crossing process

Have assumed locally constant recombination in the crossing process:



Can we do better than this?

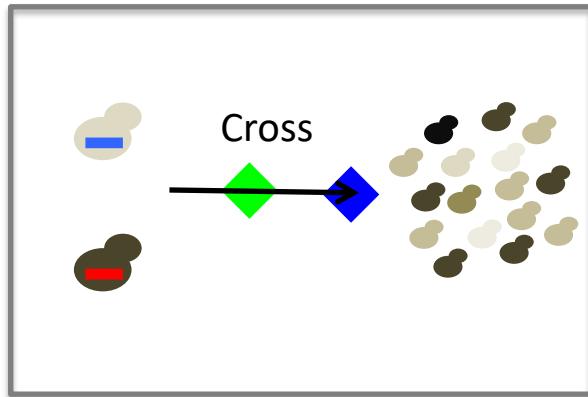
Crossing process



Initial time point (**red**)
deviates from frequency 0.5

Selection in the crossing
phase?

Driver and passengers + recombination

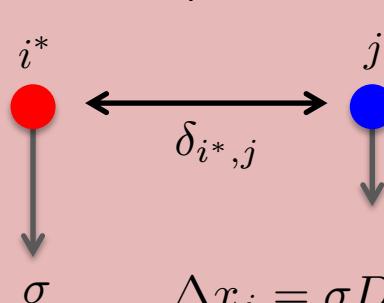


Sequencing:

- ◆ 6 generations (depth c.10)
- ◆ 12 generations (depth c.100)

Driver and passengers model: Selection and recombination

Model parameters



$$\Delta D_{i^*j} = D_{i^*j}(-\rho_{i^*j}^{tot} + \sigma - 2\sigma x_{i^*})$$

$$\rho_{i^*j}^{tot} = \sum_{k=\min\{i^*,j\}}^{k=\max\{i^*,j\}} \rho_{k,k+1} \delta_{k,k+1}$$

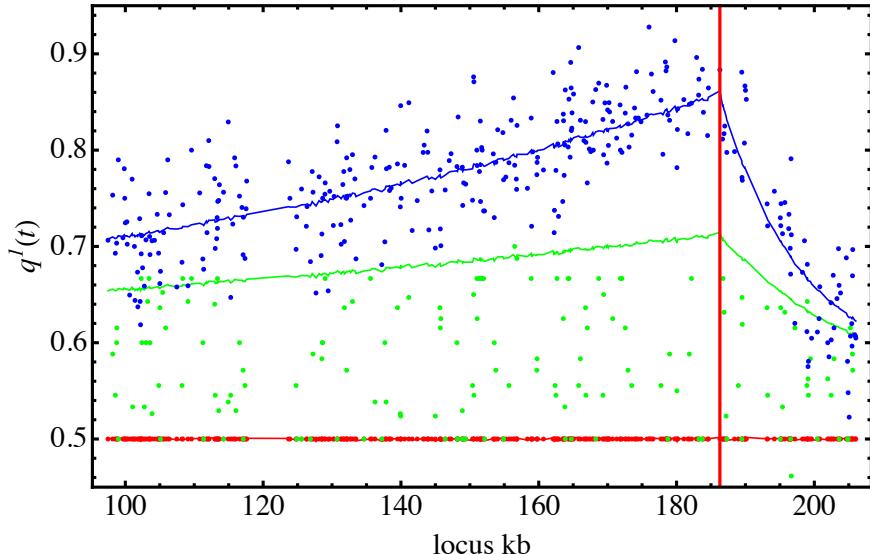
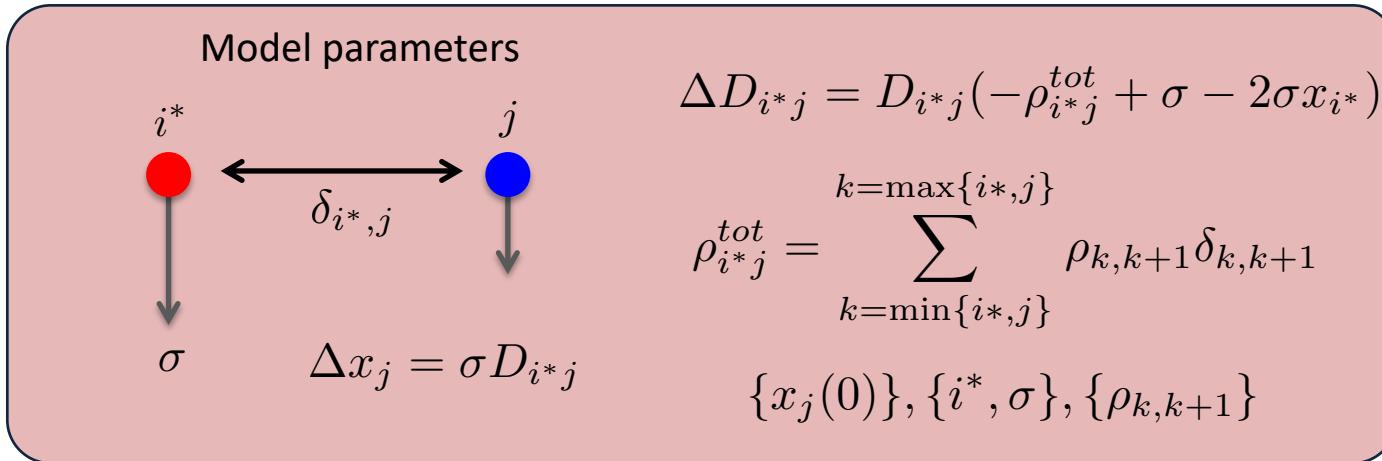
$$\{x_j(0)\}, \{i^*, \sigma\}, \{\rho_{k,k+1}\}$$

Selection affects
change in linkage

Variable rate of
recombination

Infer multiple
recombination rates

Crossing experiment : Chr XV



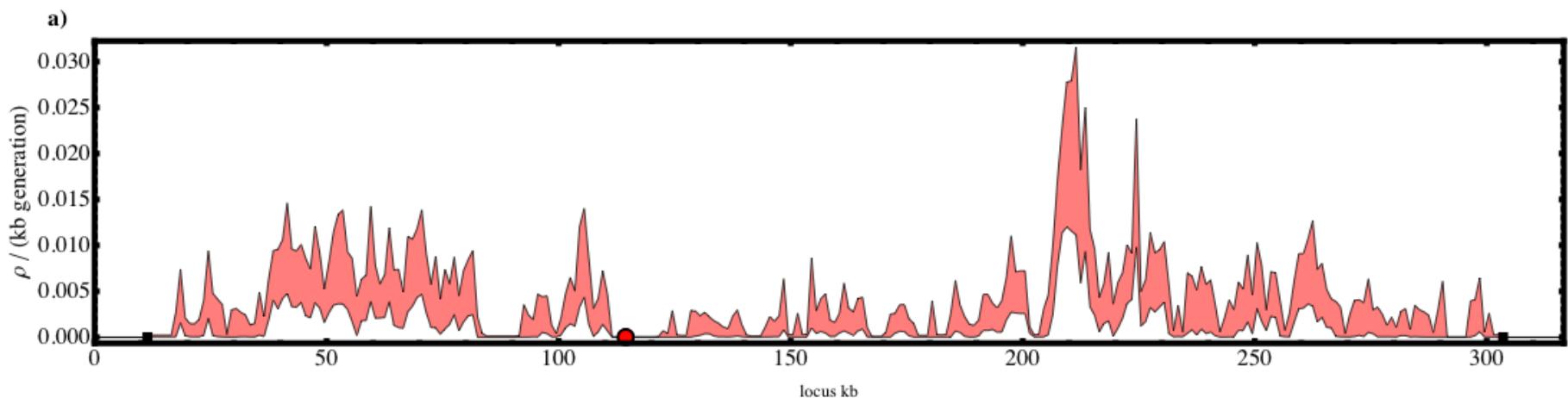
Model optimised for two recombination rates, either side of the driver

Substantial variation in recombination rate across the region of Chr XV

Variation in recombination rate

Obtain map of recombination rate genome-wide

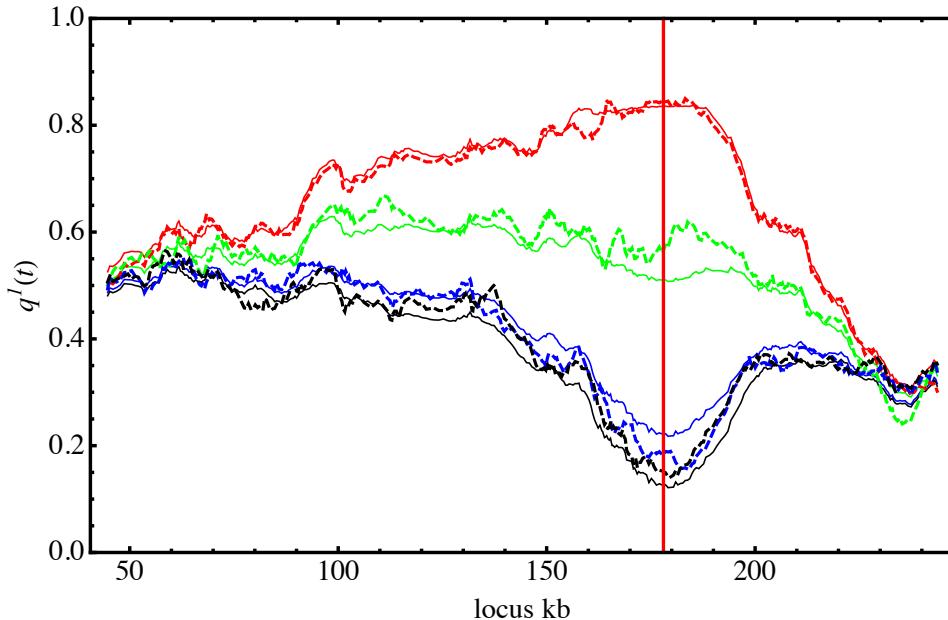
Data for chromosome III:



Account for variation in recombination rate across chromosomes

$$\rho \Delta_{ij} \rightarrow \sum_k \rho_{k,k+1} \delta_{k,k+1}$$

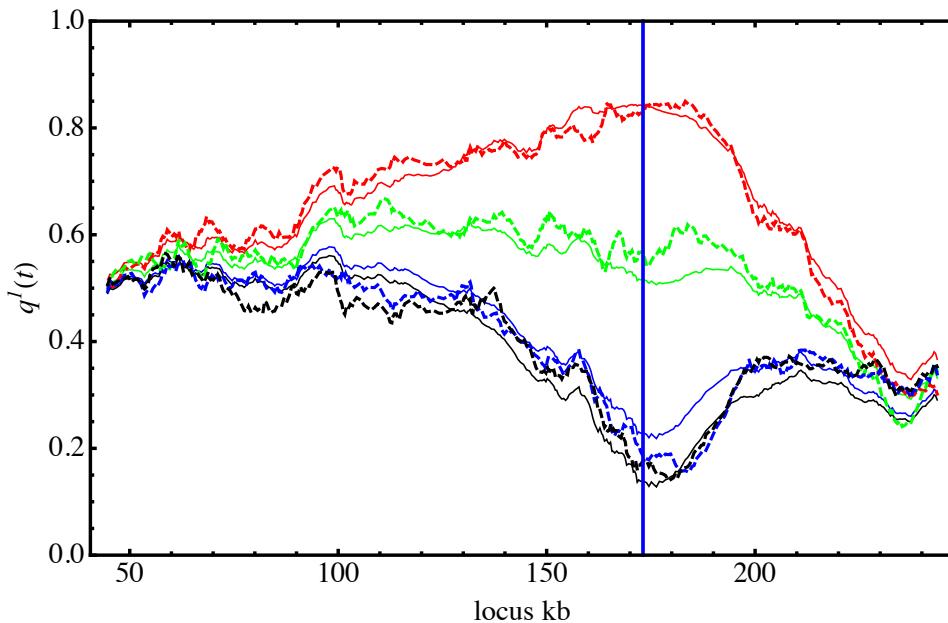
Use recombination map to find selection



Chromosome XV data

With recombination map

Driver location shifted to
the right



Improvement in overall
model fit

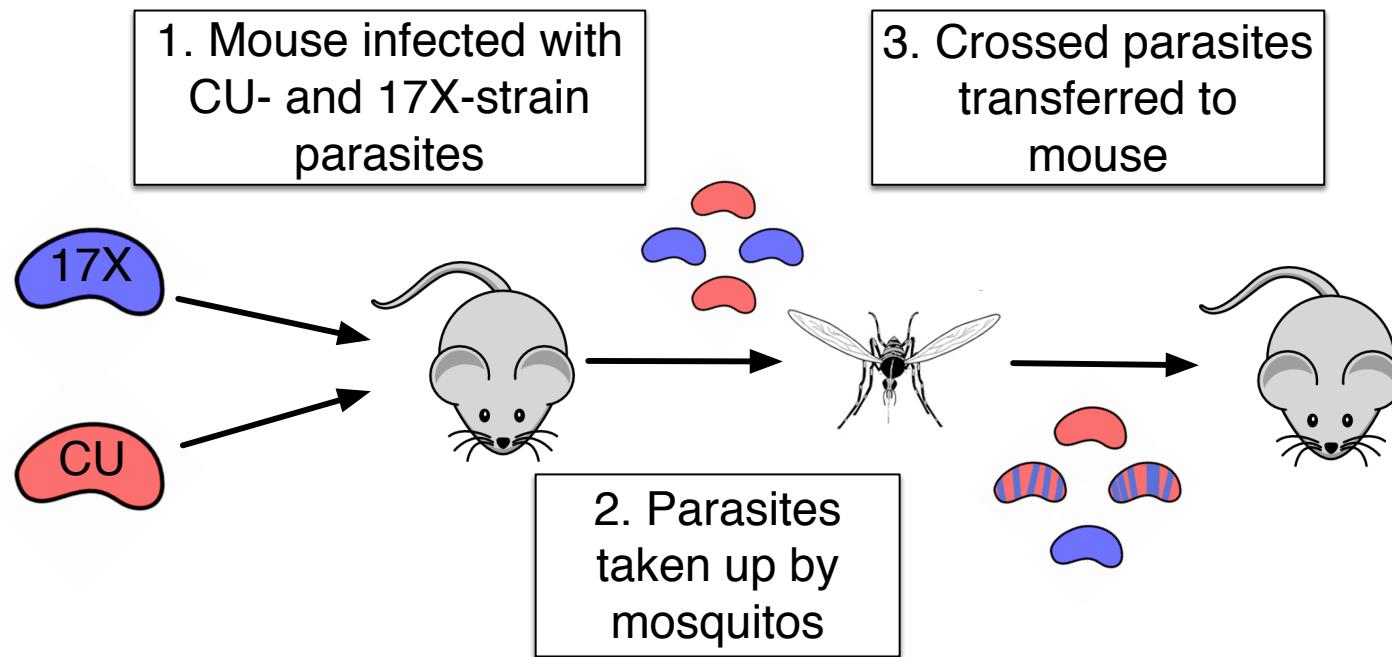
Without recombination map

Conclusions

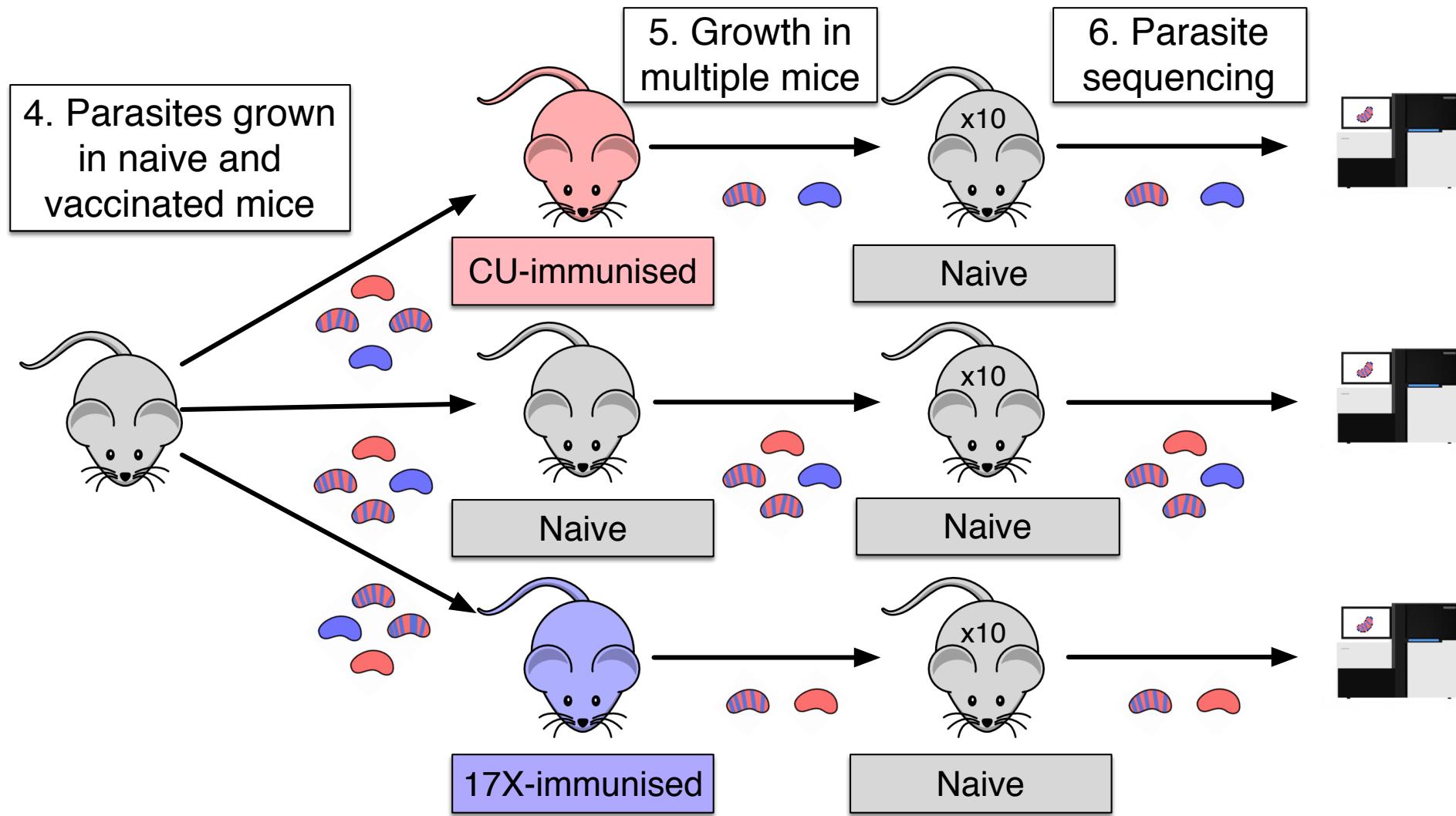
1. Using an additive model we identified 44 candidate driver loci for heat tolerance, with their selection coefficients (+ local recombination rates).
2. The liberal driver model shows the presence of significant non-additive fitness effects. Further modelling did not identify a precise cause for this.
3. Modelling the crossing phase of the experiment demonstrated the presence of substantial selection in the cross, and of local variation in the recombination rate.

Example IV: Crossed malaria parasites

Protocol:

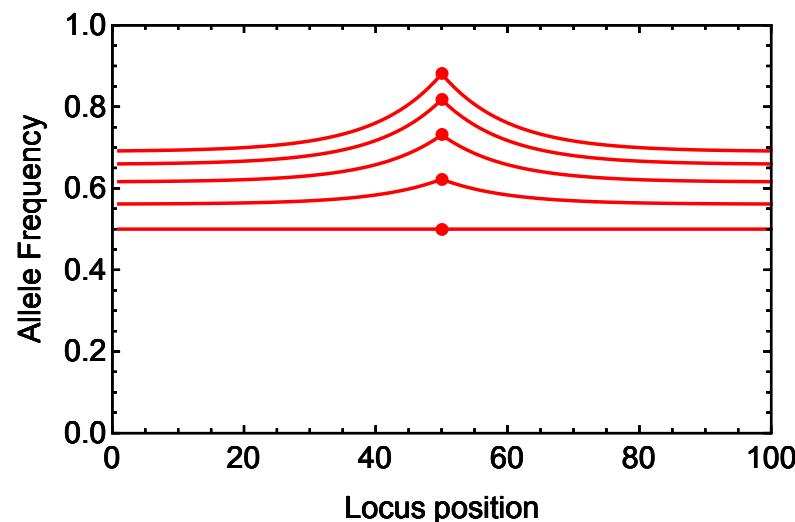
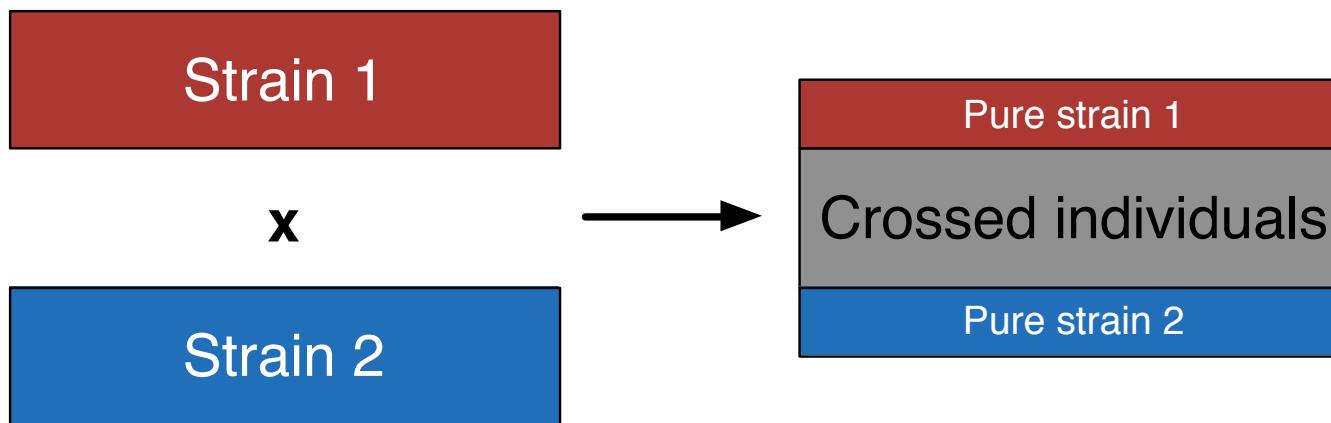


Example IV: Crossed malaria parasites



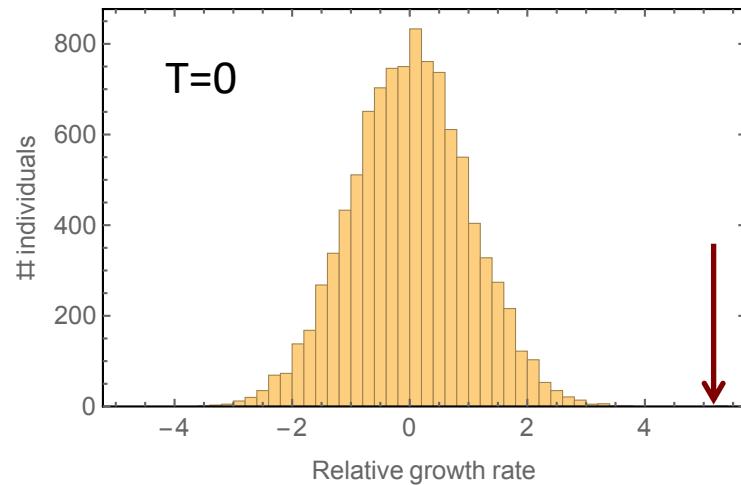
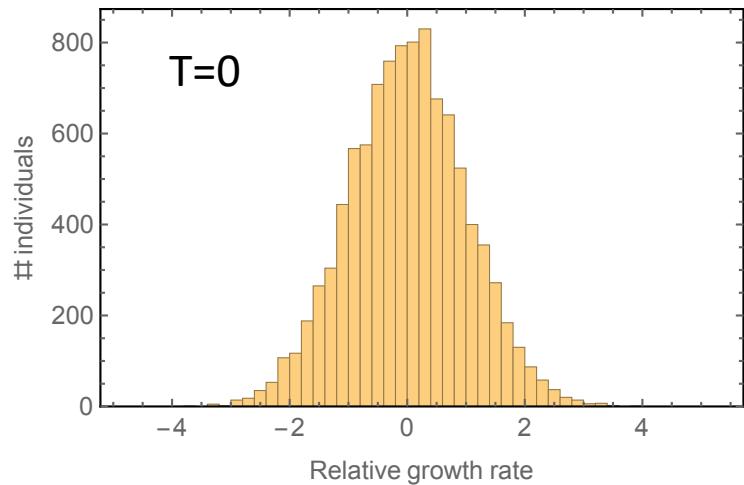
Example IV: Crossed malaria parasites

Single generation of crossing:



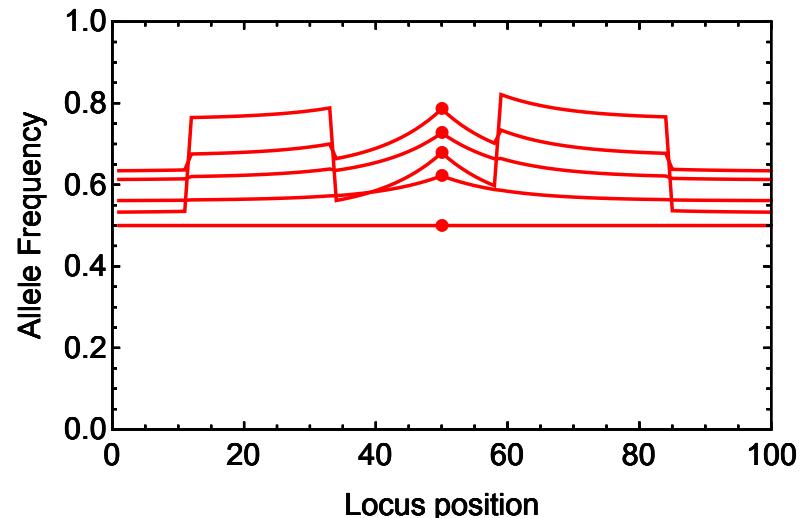
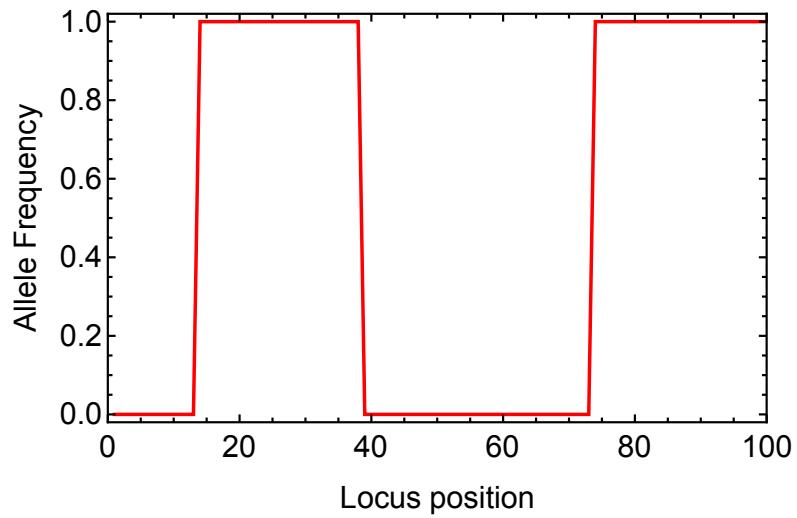
Example IV: Crossed malaria parasites

Clonal growth: high fitness individuals



Example IV: Crossed malaria parasites

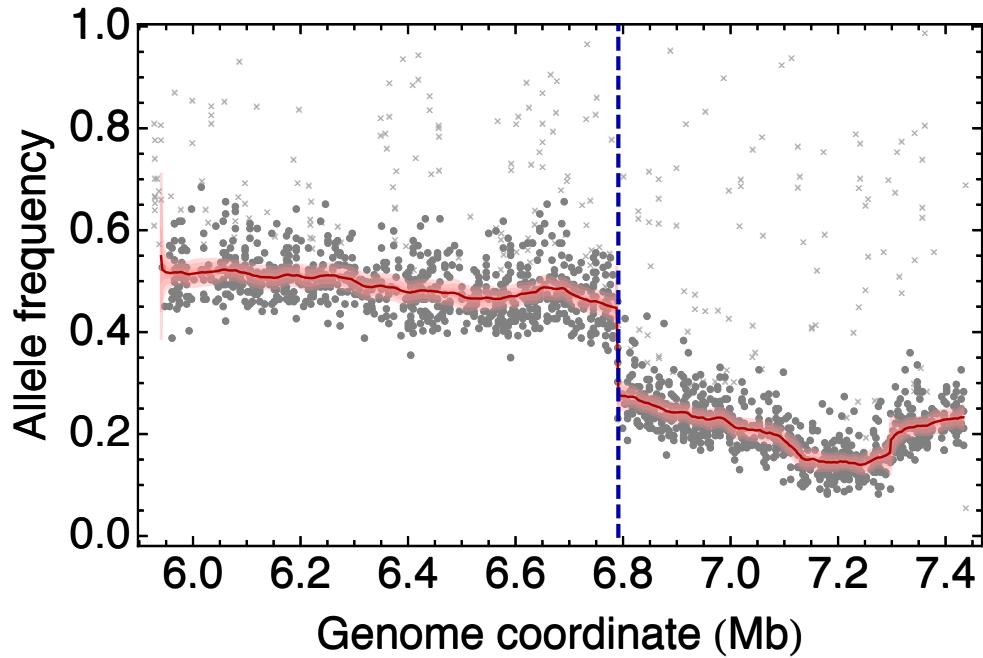
Clonal genotype switches between parental alleles



Comprises an increasing fraction of the population

Detect clonal growth

Replica 1: Apparent sudden changes in frequency



Jump-diffusion model

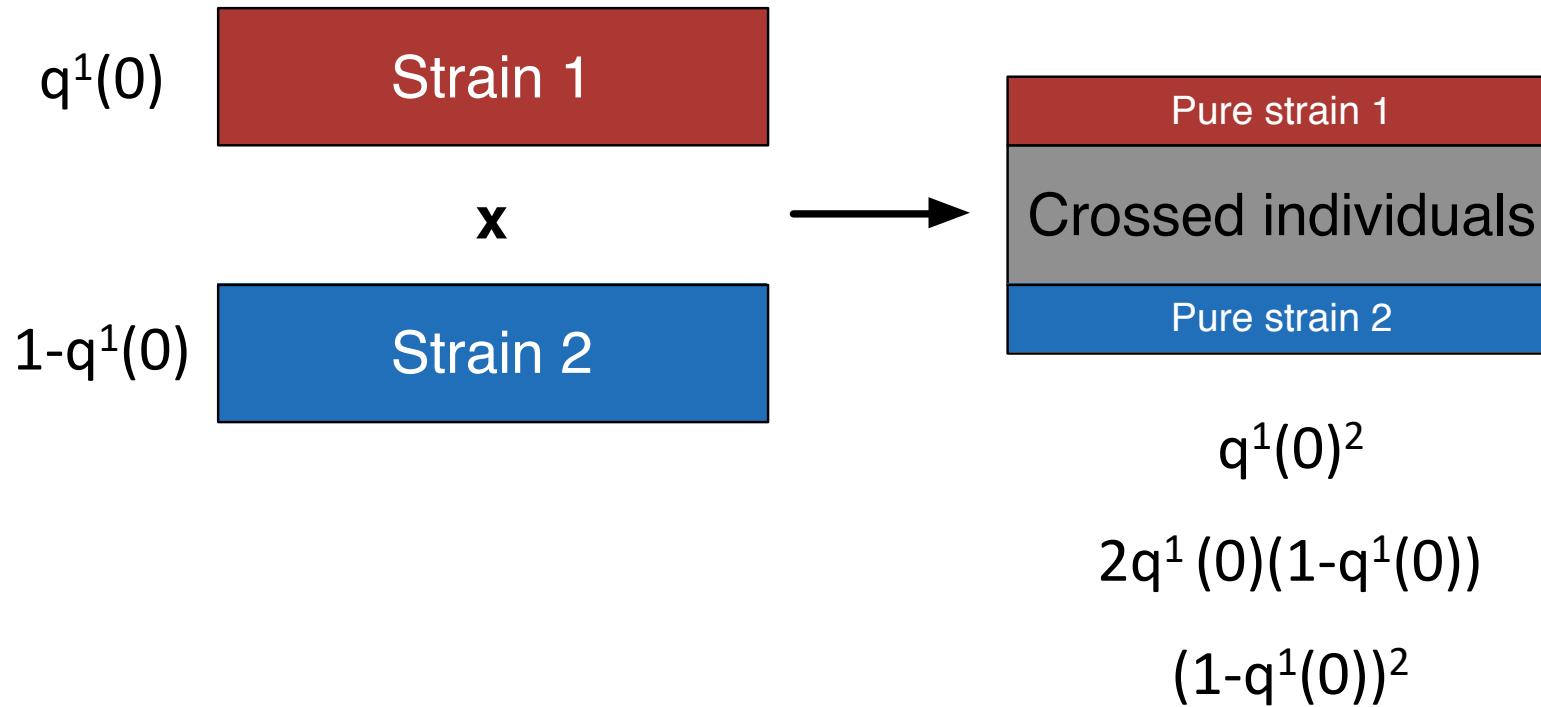
$$x_{i+1} = x_i + \mathcal{N}(0, s\sqrt{\Delta_{ij}})$$

$$x_{i+1} \sim \mathcal{U}(0, 1)$$

Consider regions separately

Location of selection

Change in selected allele (Cross at time zero):



Location of selection

Change in selected allele (1 at position i)

$$q_i^1(t) = \frac{q^1(0)e^{\sigma t}}{1 - q^1(0) + q^1(0)e^{\sigma t}}$$

Change in nearby allele (1 at position j):

$$q_j^1(t) = q_i^1(t) \frac{q_{ij}^{11}(0)}{q_i^1(0)} + q_i^0(t) \frac{q_{ij}^{01}(0)}{q_i^0(0)}$$

What are the terms $q^{11}(0)$ and $q^{01}(0)$?

Pure genotypes: contribution of $q^1(0)^2$ to q^{11}

contribution of 0 to q^{01}

Location of selection

Crossed genotypes:

$$\begin{aligned}\tilde{q}_{ij}^{11}(0) &= \tilde{q}_i^1(0)\tilde{q}_j^1(0) + D'_{ij}e^{-\rho\Delta_{ij}} \\ &= \frac{1}{4}(1 + e^{-\rho\Delta_{ij}})\end{aligned}$$

Frequency of crossed genotypes is $2q^1(0)(1-q^1(0))$

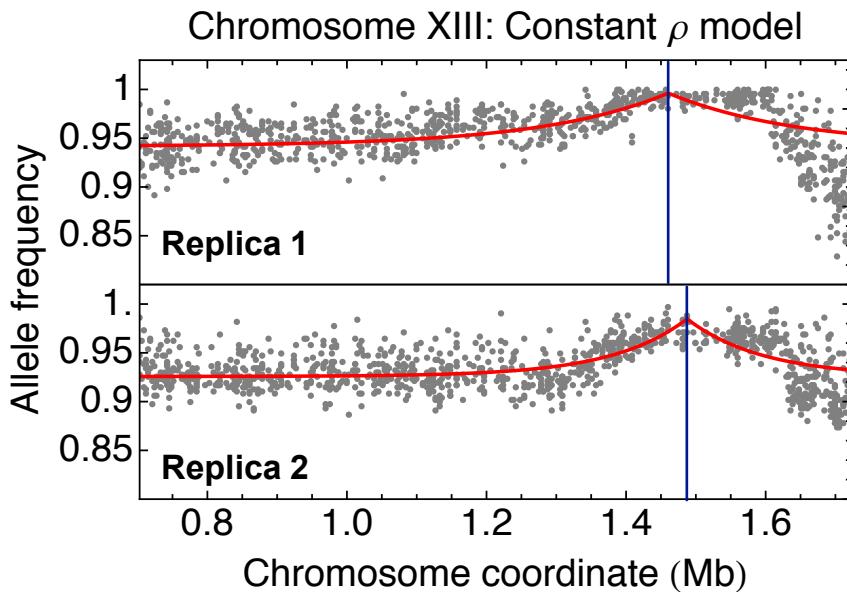
$$q_{ij}^{11}(0) = q_i^1(0)^2 + \frac{1}{2}q_i^1(0)(1 - q_i^1(0))(1 + e^{-\rho\Delta_{ij}})$$

$$q_{ij}^{01}(0) = \frac{1}{2}q_i^1(0)(1 - q_i^1(0))(1 + e^{-\rho\Delta_{ij}})$$

$$q_j(0) = \left[q_i^1(0) + \frac{1}{2}(1 - q_i^1(0))(1 + e^{-\rho\Delta_{ij}}) \right] q + \left[\frac{1}{2}q_i^1(0)(1 - e^{-\rho\Delta_{ij}}) \right] (1 - q) + e$$

Evolutionary model

Example:



Location of driver:

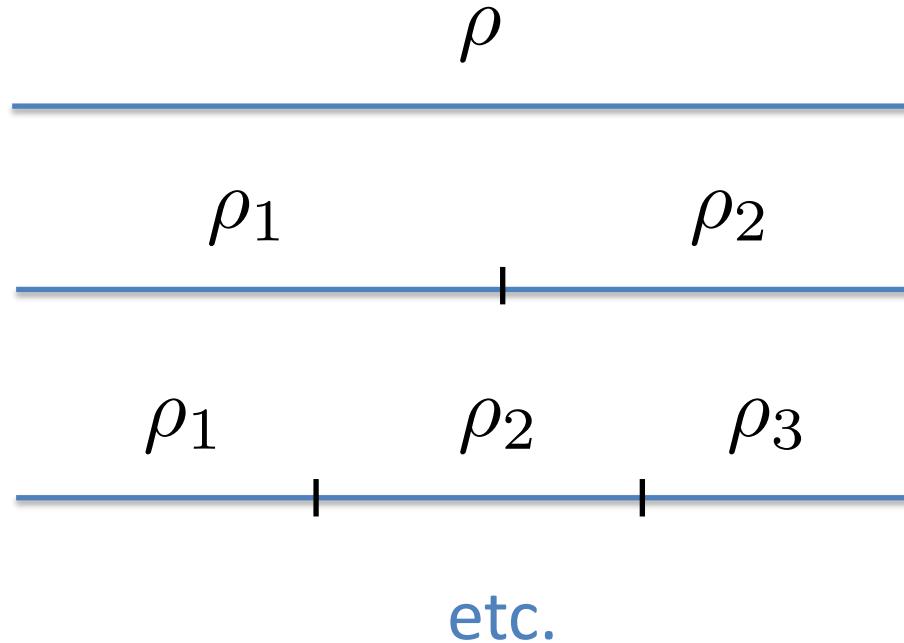
R1: 1513 kb

R2: 1529 kb

Identify candidate region for selection

More flexible models of recombination

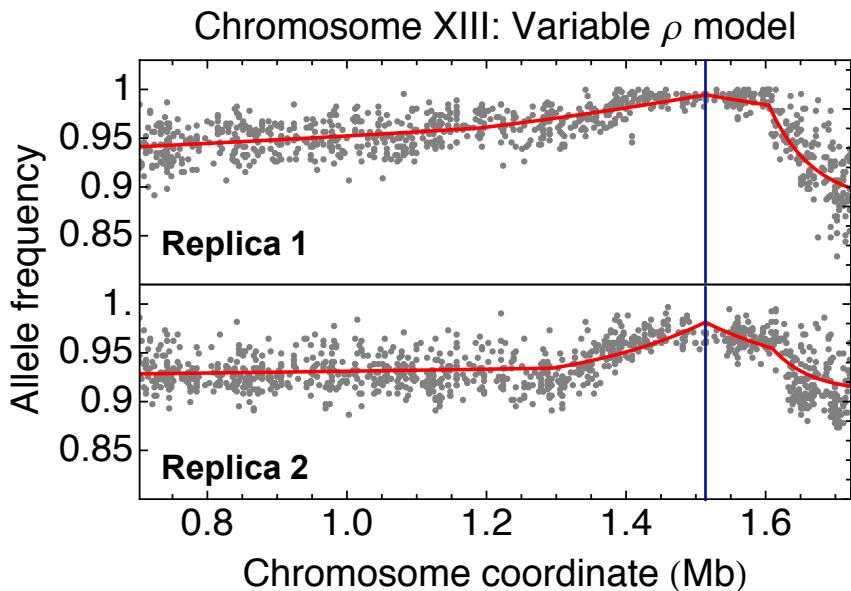
Stepwise models: Use model selection



Model selection: Bayesian Information Criterion

Evolutionary model

Refine model: Variable recombination



Location of driver:

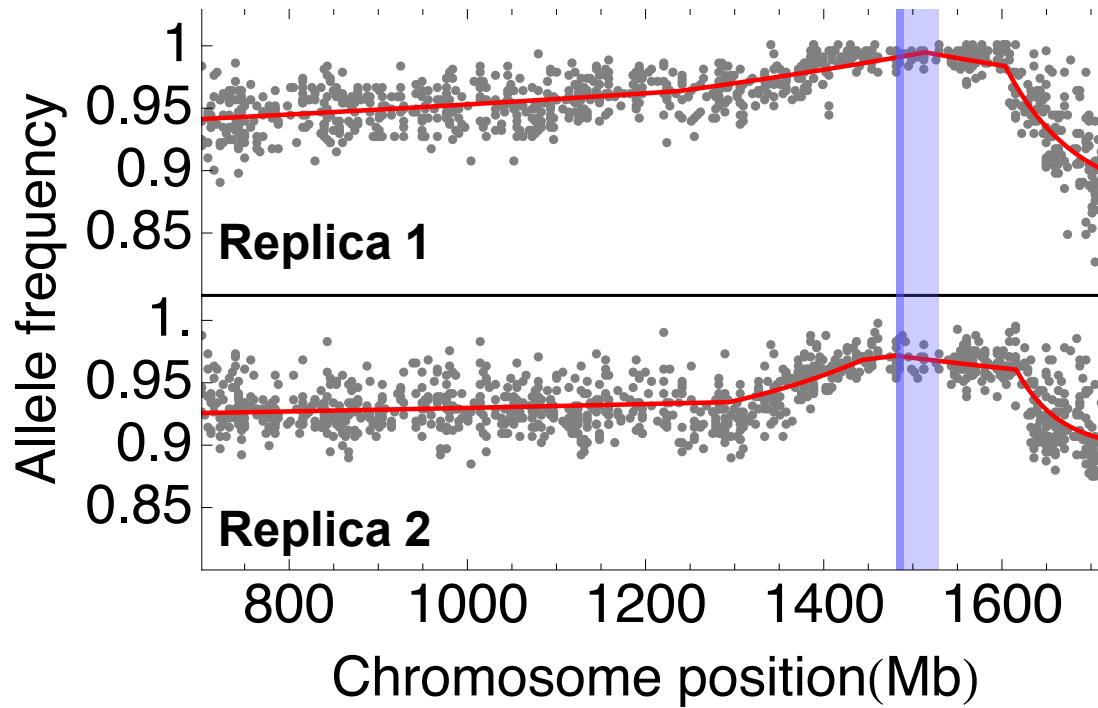
R1: 1514 kb

R2: 1514 kb

Two- and three-recombination rate models

Evolutionary model

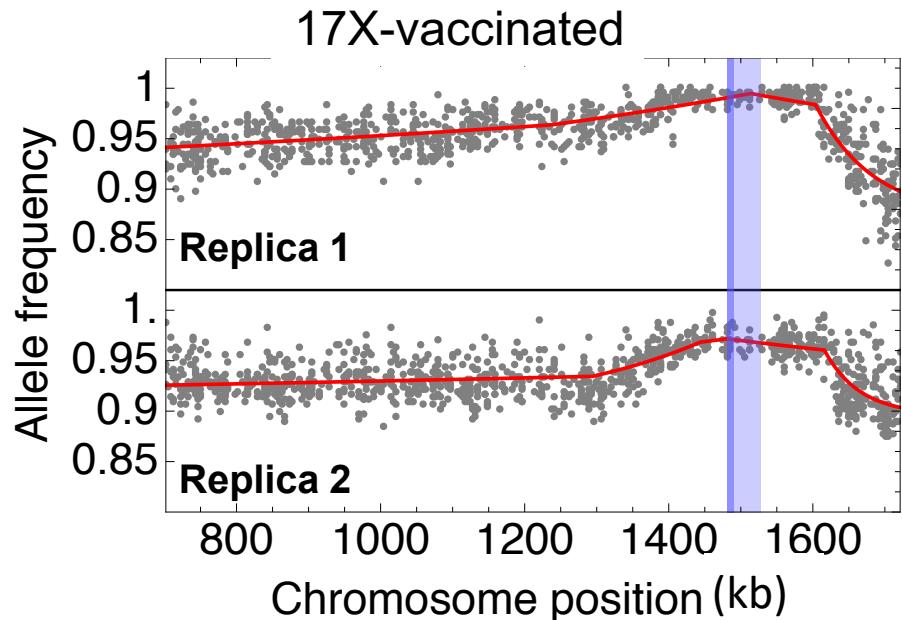
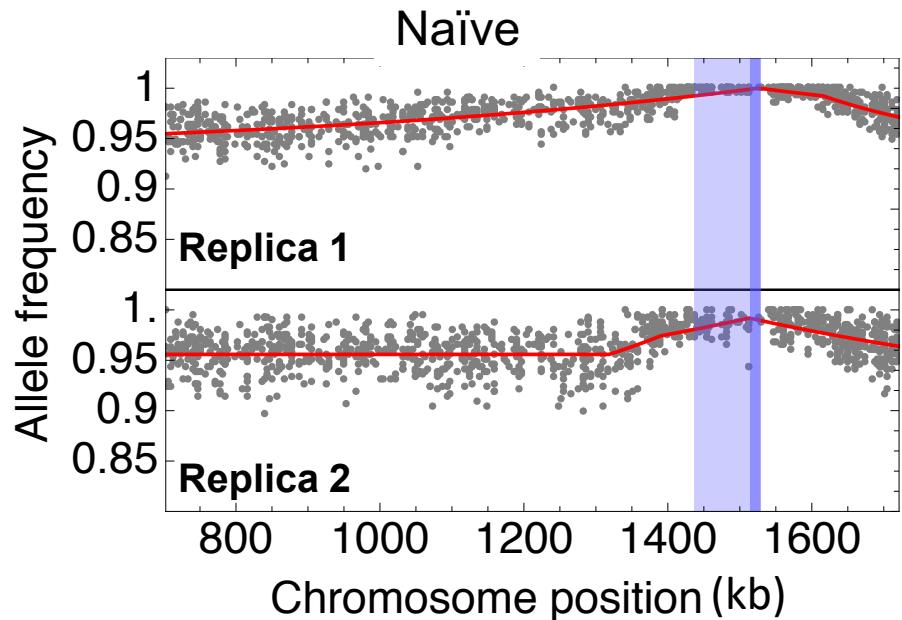
Identify confidence intervals by likelihood



More and less conservative intervals

Model outcome

Chromosome XIII



Growth allele: Position close to gene PyEBL

Erythrocyte Binding Ligand

Summary

Selection across multiple loci

Experimental evolution

Recombination

Genetic cross experiments

Evolutionary contingency

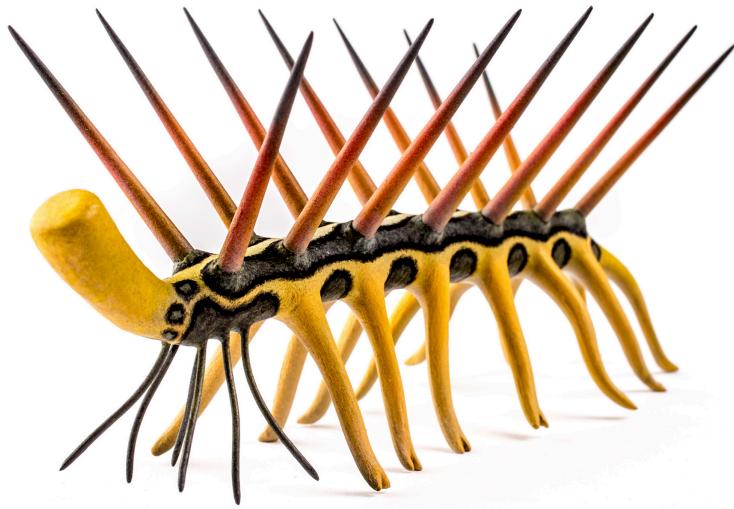
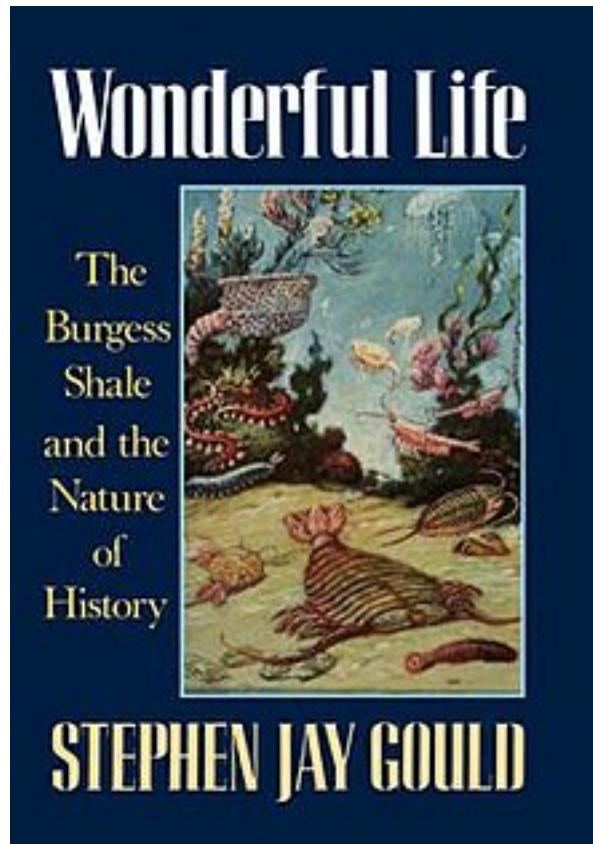
Can evolution be predicted?

$$\frac{dq}{dt} = \sigma q(1 - q)$$

1. Use data to infer selection
2. Use knowledge of selection to predict evolution?

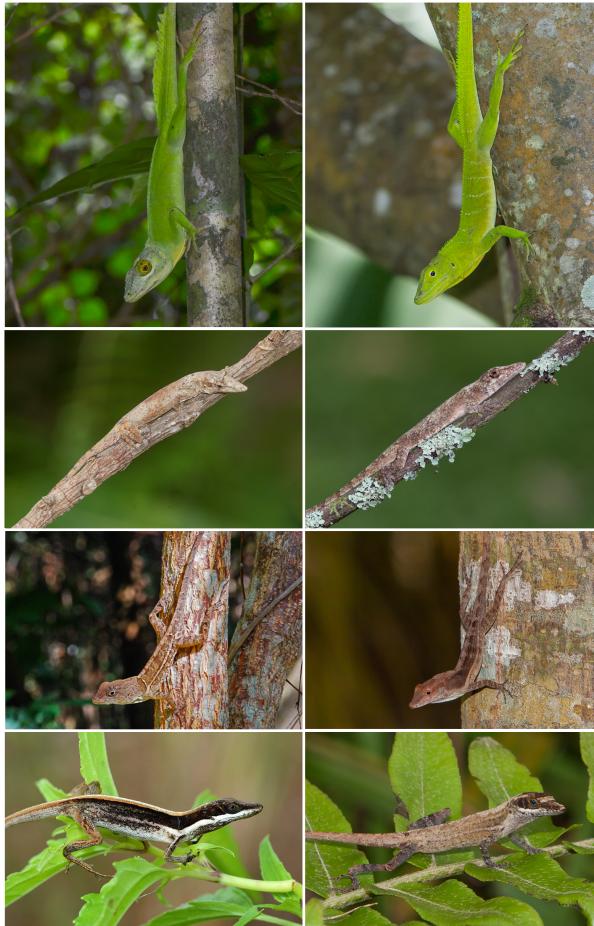
Evolutionary contingency

Replaying the tape of life: Macro-scale predictability?



Evolutionary contingency

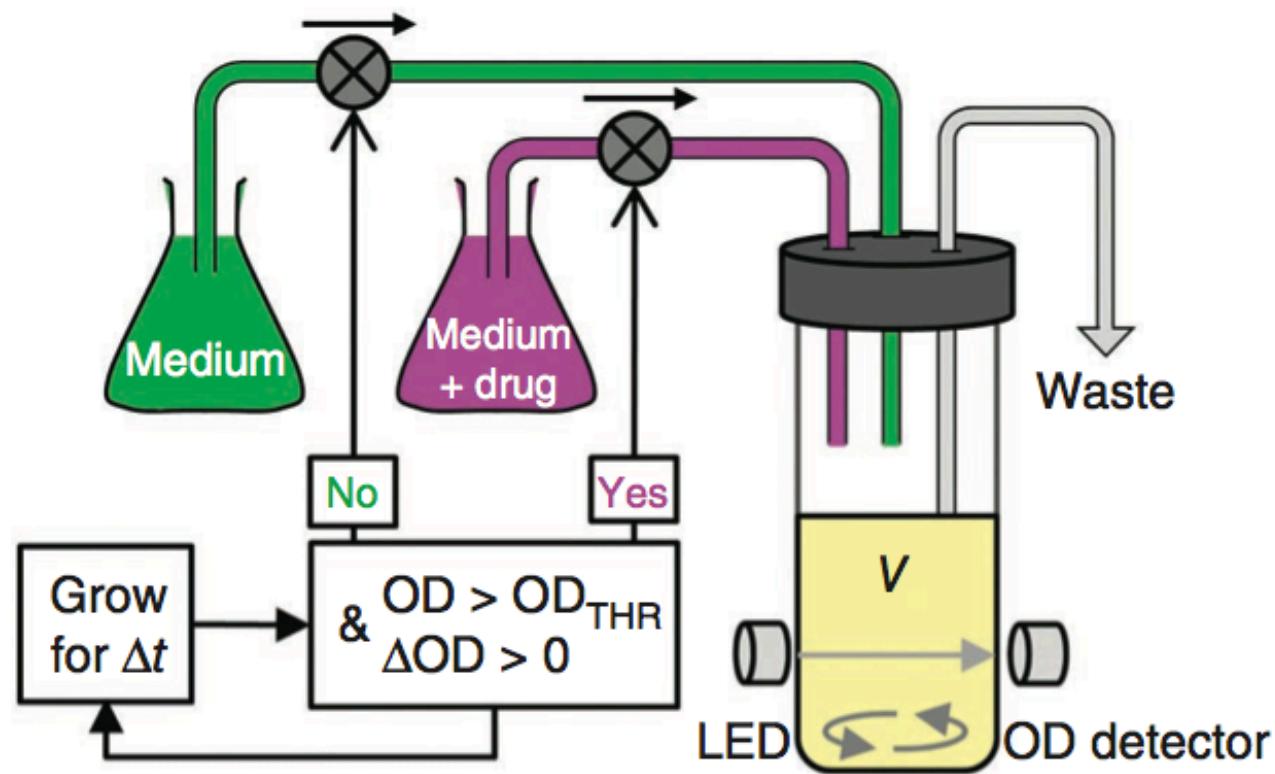
Anole lizards



Evolutionary experiments

Example: Morbidostat

Alter drug concentration with time to maintain population size



Evolutionary experiments

Observed mutations

Drug binds to DHFR: observe mutations in this protein

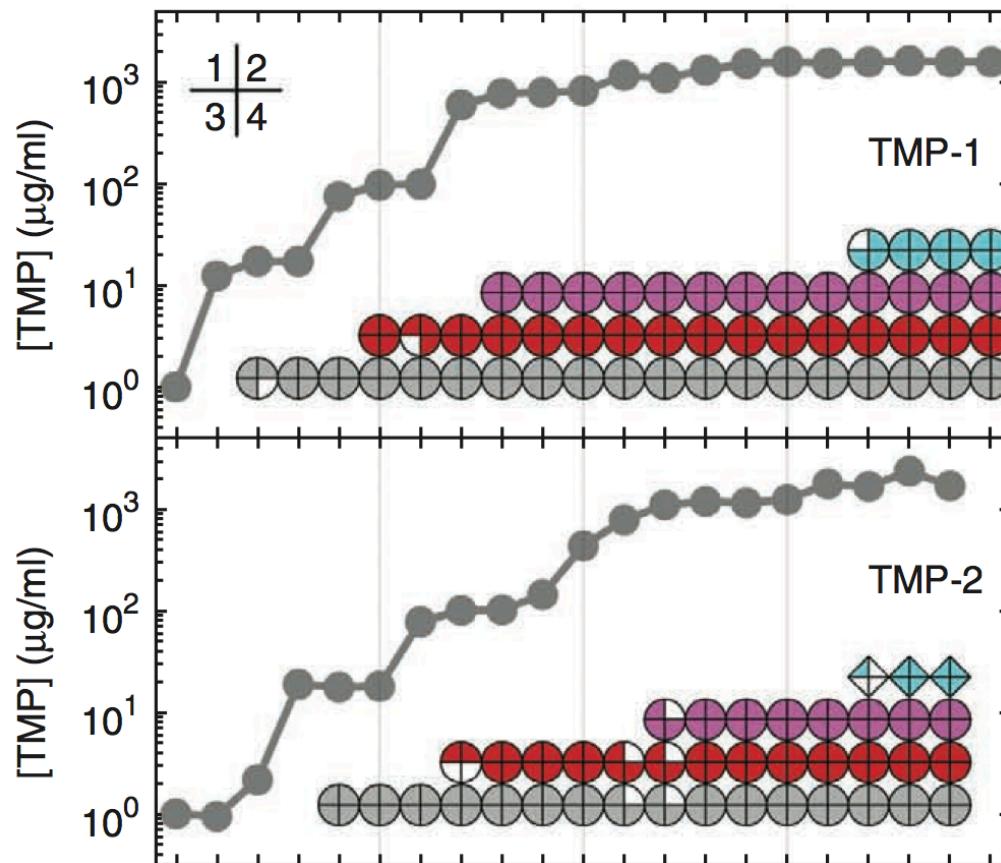
- ◆ -9G>A
- -35C>T
- P21L
- A26T
- ◆ A26V
- A26S
- L28R
- W30R
- ◆ W30G
- W30C
- I94L



Evolutionary experiments

Observed mutations

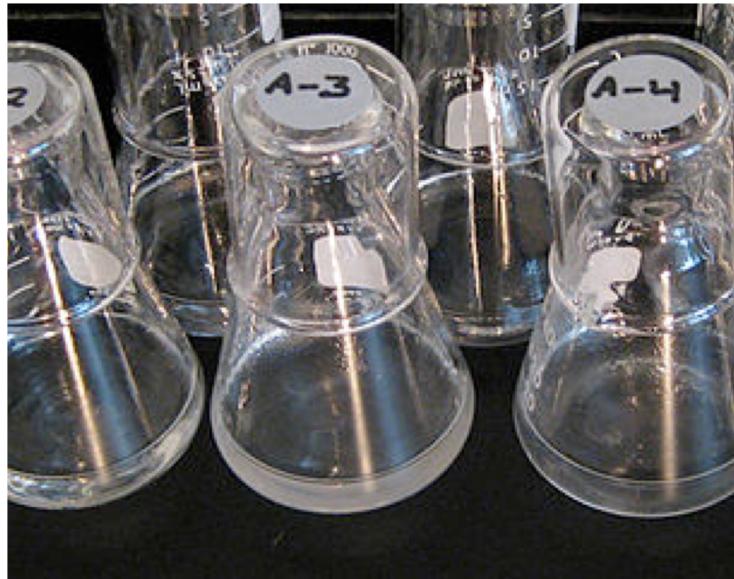
Order of mutations preserved between experiments



Evolutionary experiments

Evolution of citrate metabolism

One population has evolved the ability to digest citrate



More available nutrient: leads to an increase in the population size

Genome sequencing shows the underlying mechanisms of evolution

Evolutionary experiments

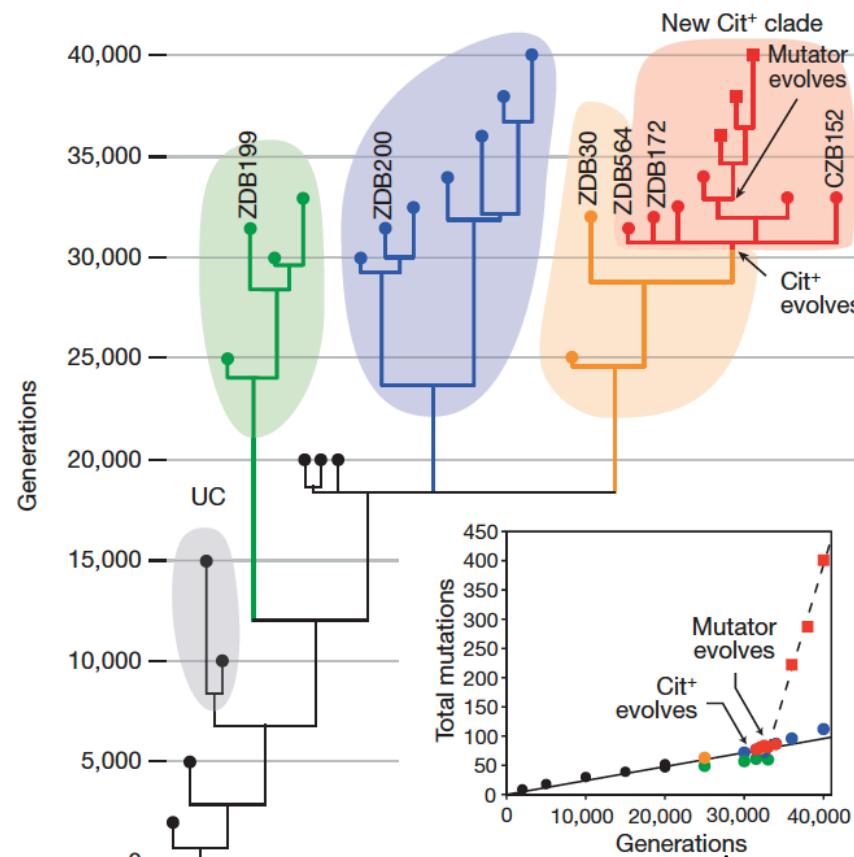
Evolution of citrate metabolism

Phylogenetics shows a structured population evolving over time

Identify three consistent clades persisting over time.

The ability to metabolise citrate occurred in one clade.

Following this, a mutator phenotype emerged.



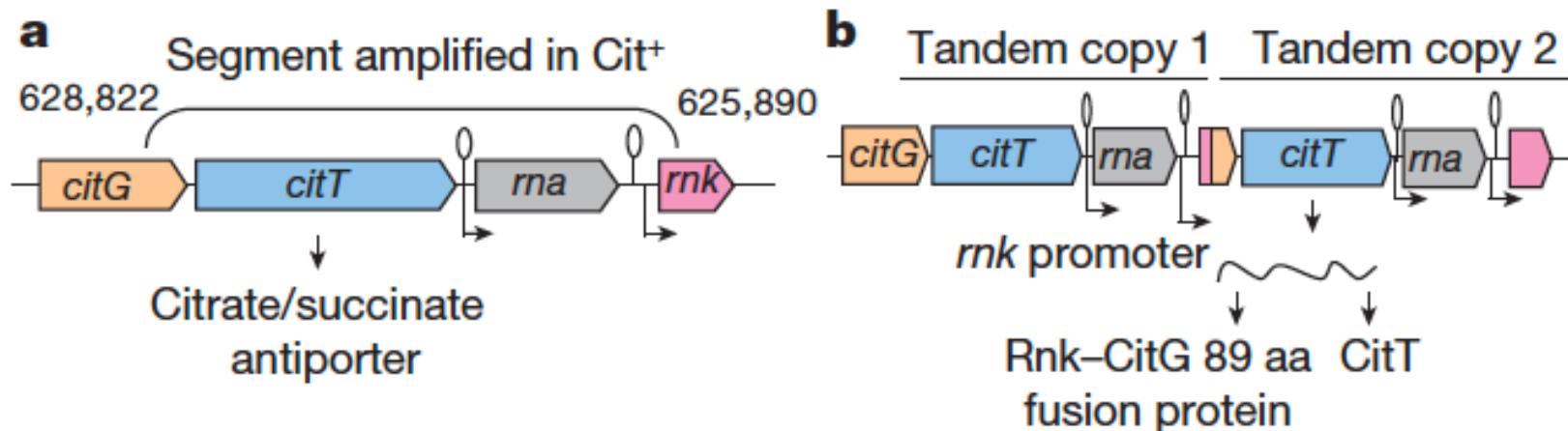
Blount et al, Nature, 2012

Evolutionary experiments

Evolution of citrate metabolism

Ability to metabolise citrate emerged from a duplication event

Duplication of part of a promoter leads to the expression of *citT* gene



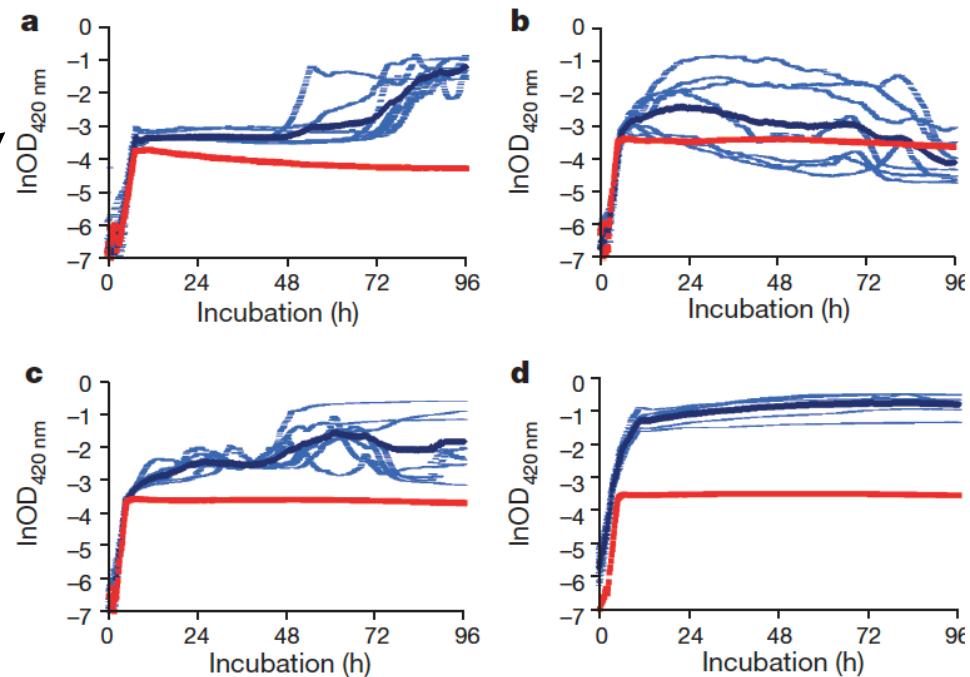
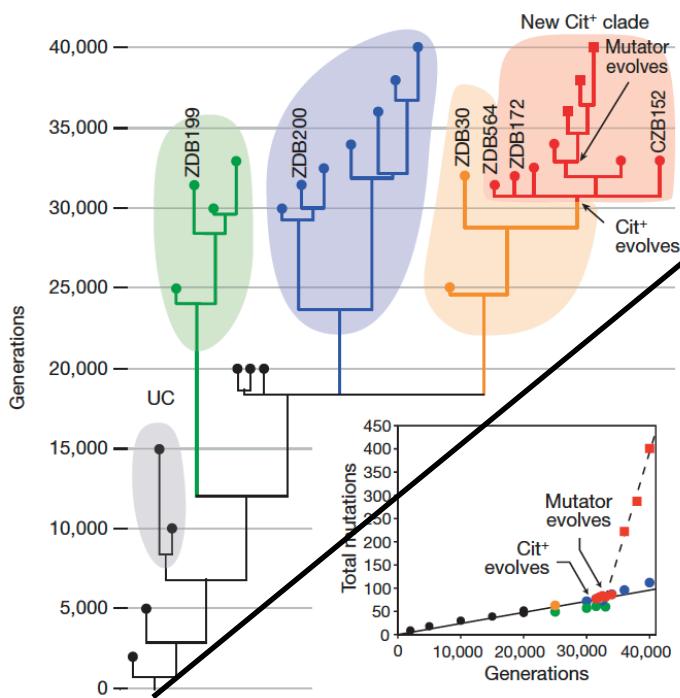
Amplification produces limited metabolism of citrate

Conveys slight fitness advantage

Evolutionary experiments

Evolution of citrate metabolism

Benefit of module is dependent on strain background : Epistasis

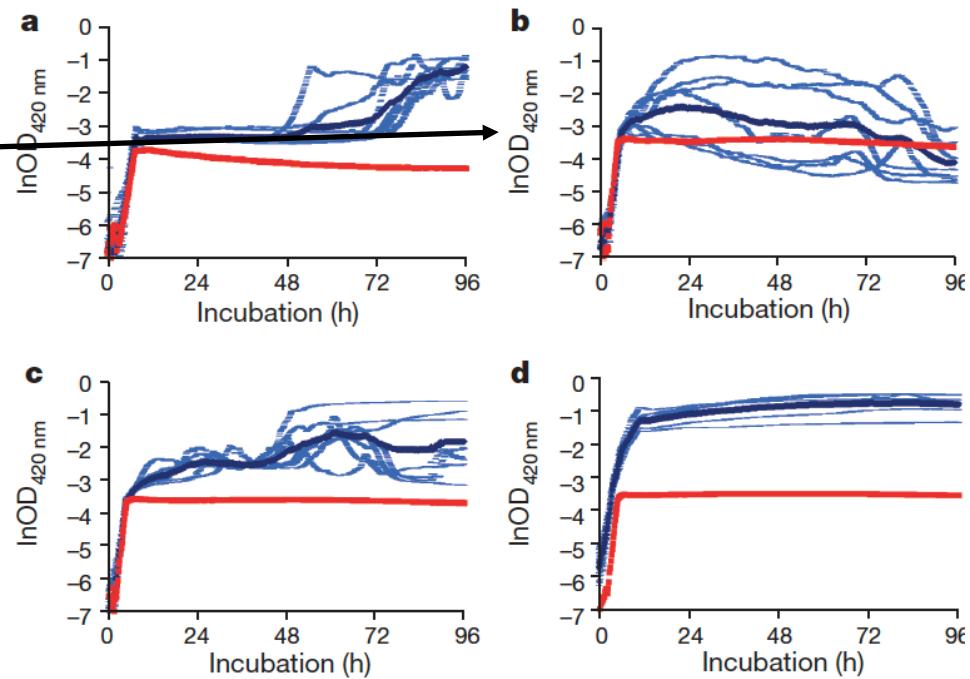
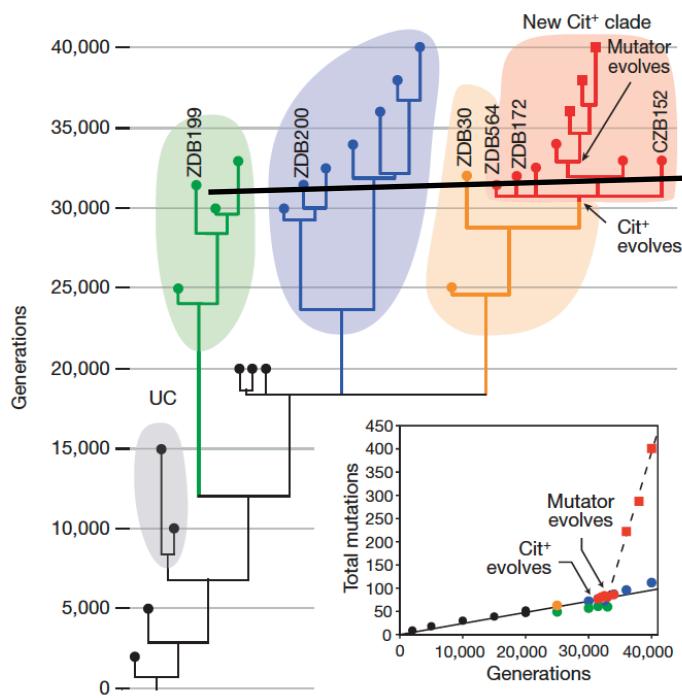


Previous mutations in the bacteria led to the module being able to convey a beneficial effect

Evolutionary experiments

Evolution of citrate metabolism

Benefit of module is dependent on strain background : Epistasis

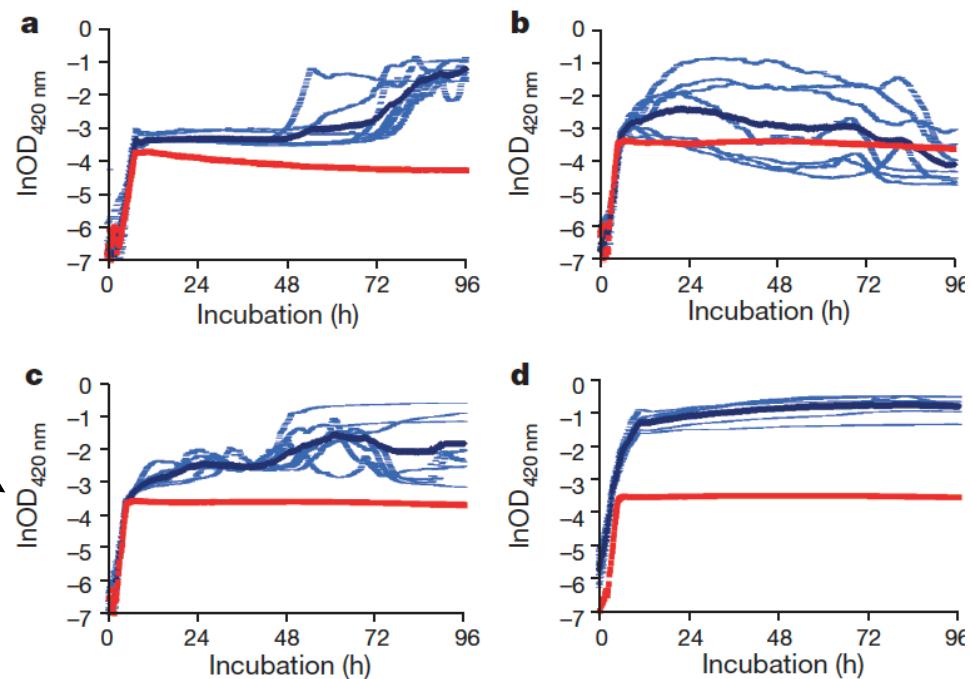
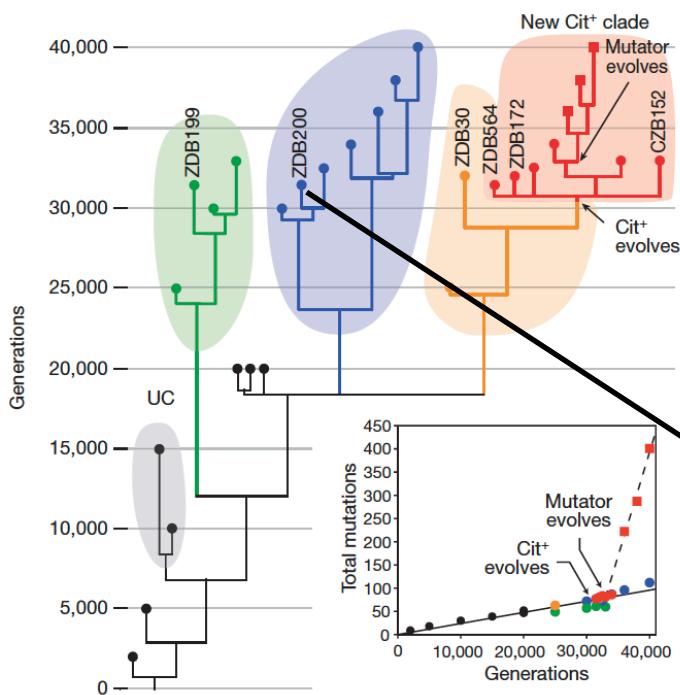


Previous mutations in the bacteria led to the module being able to convey a beneficial effect

Evolutionary experiments

Evolution of citrate metabolism

Benefit of module is dependent on strain background : Epistasis

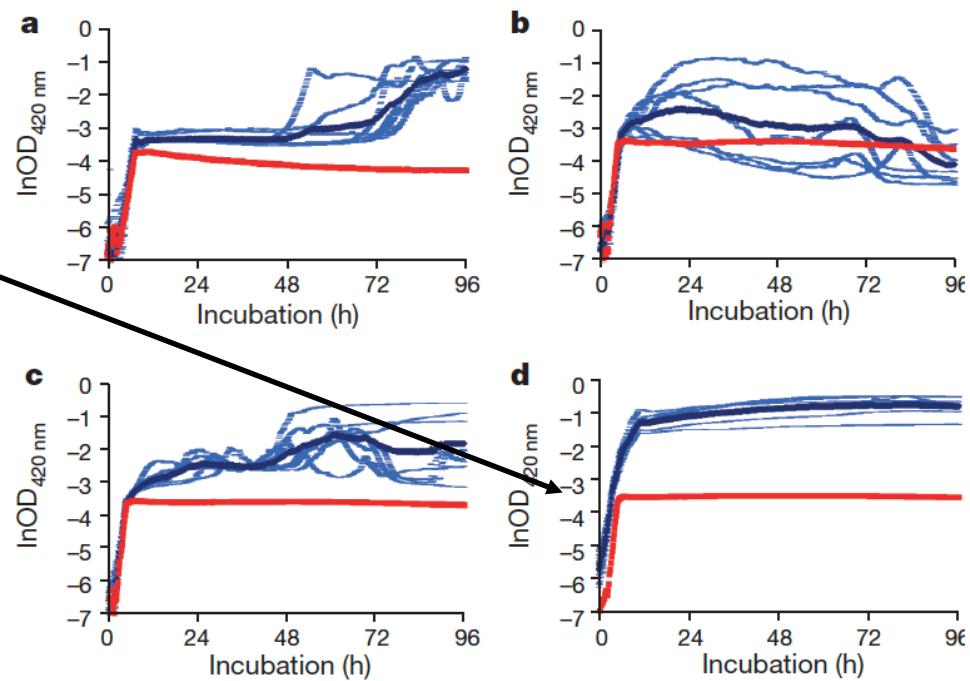
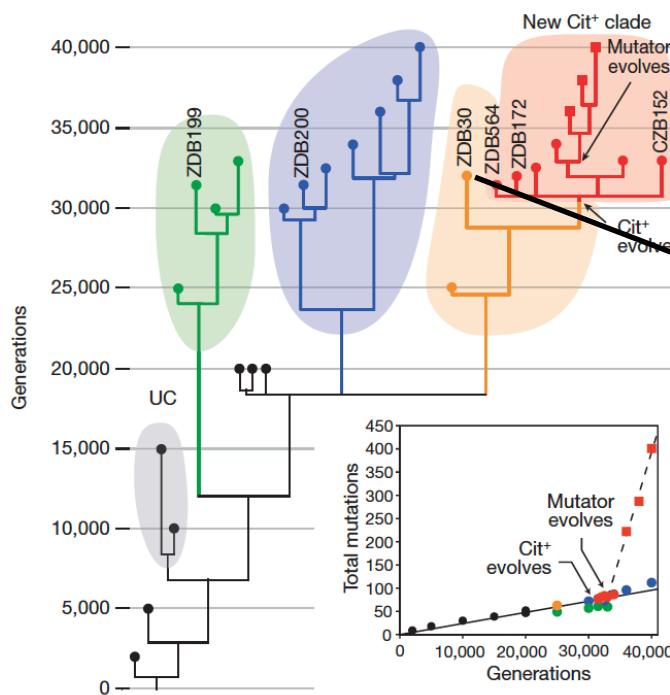


Previous mutations in the bacteria led to the module being able to convey a beneficial effect

Evolutionary experiments

Evolution of citrate metabolism

Benefit of module is dependent on strain background : Epistasis



Previous mutations in the bacteria led to the module being able to convey a beneficial effect

Evolutionary experiments

Implication: Evolution is contingent – background strain is important

