

Recent Developments in Theoretical Population Genetics

Reinhard Bürger

Department of Mathematics
University of Vienna
Austria

The two main perspectives of population genetics

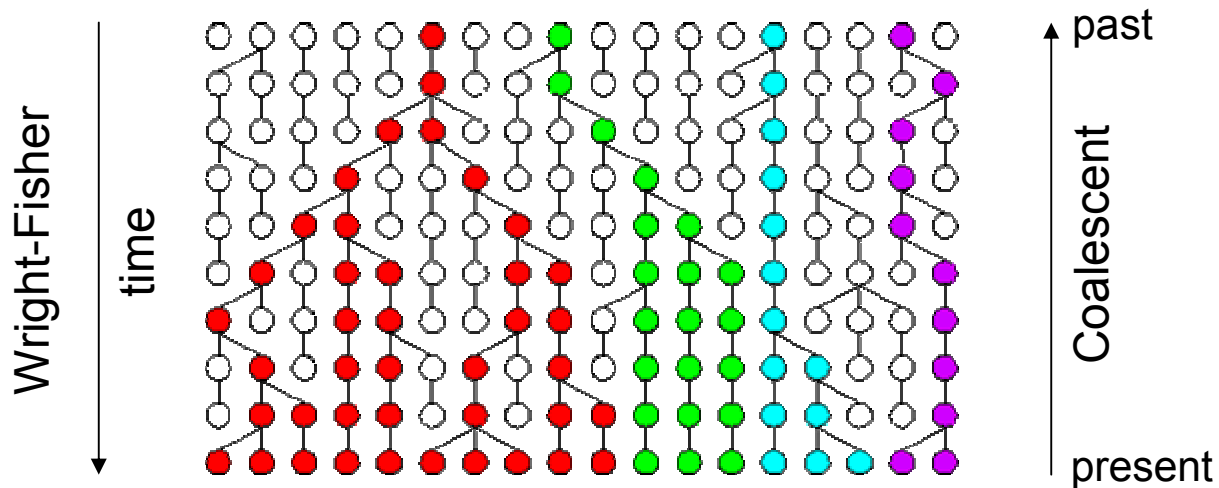
- **The forward view:**

Predicting evolution from given genetic, ecological, and evolutionary mechanisms

- **The backward view:**

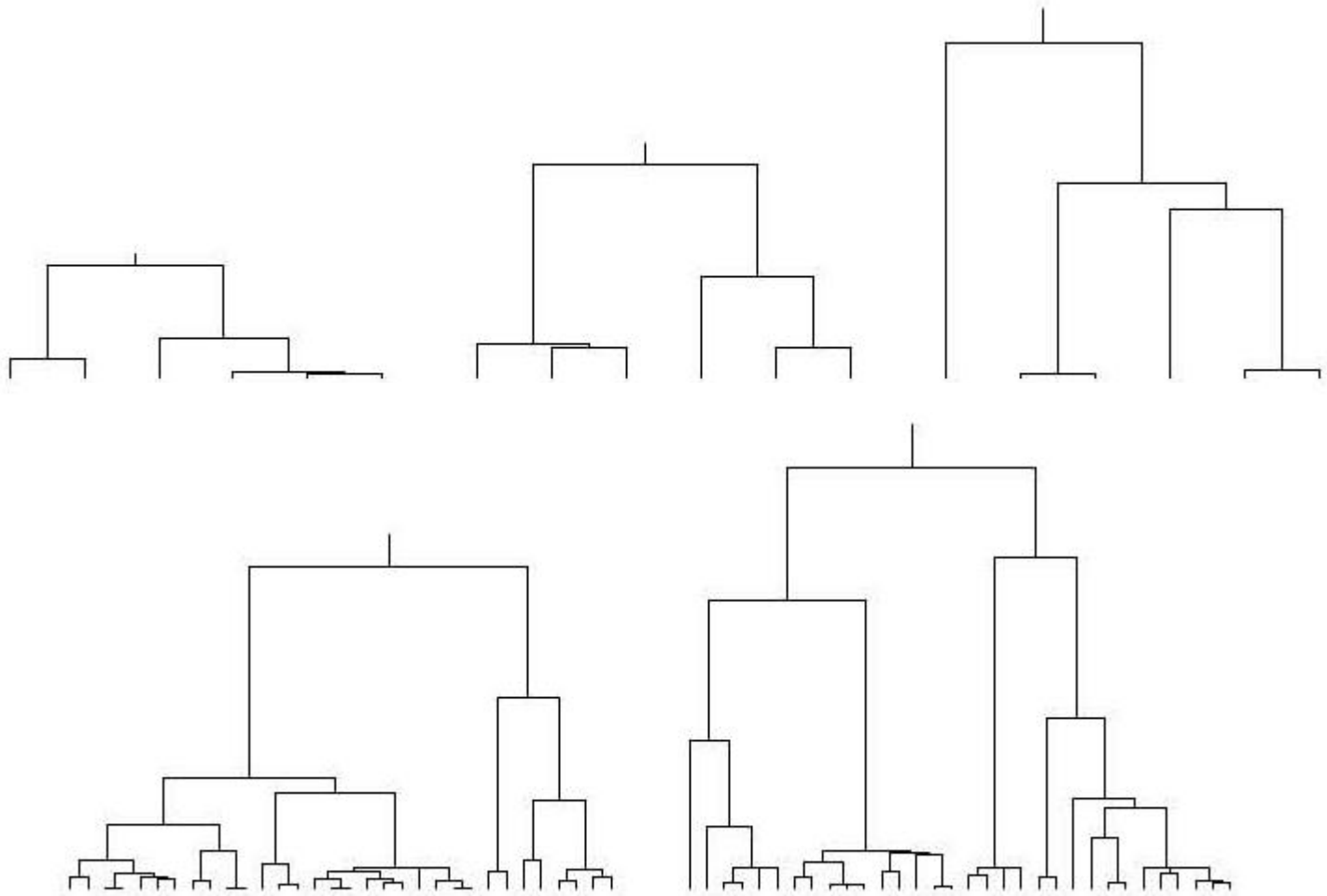
Inferring the genealogy and past evolutionary events from molecular (sequence) data of extant populations

Wright-Fisher model and the coalescent



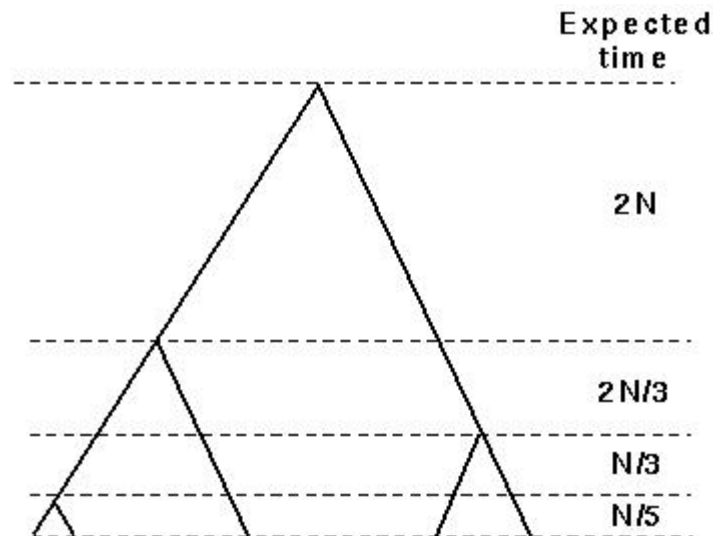
Schematic depiction of one possible realization of the coalescent process in a population with 18 haploid gametes. The eighteen alleles in our current sample are descended from only four alleles that were present in the populations ten generations ago. How far back in time do we have to go to find the **most recent common ancestor** (MRCA)?

Simulated genealogies of samples of 6 and of 32 sequences



Expected time to MRCA

$$E[T_{\text{MRCA}}] = \sum_{k=2}^n E[T_k] = \sum_{k=2}^n \frac{2}{k(k-1)} = 2\left(1 - \frac{1}{n}\right)$$



The Forward Perspective

Overview

- Review of some classical models

- Selection in single-locus systems
- Mutation and selection
- Recombination between two loci
- Quantitative genetics
- Evolution on adaptive landscapes

The Forward Perspective

Overview

- Review of some classical models
- **General results for multilocus systems**

- Recombination and linkage disequilibrium
- Weak selection
- Weak epistasis
- Exact recursions for quantitative genetic models

The Forward Perspective

Overview

- Review of some classical models
- General results for multilocus systems
- 'Hot' topics

- Maintenance of quantitative genetic variation
- Epistasis and the evolution of genetic architecture
- Evolvability and robustness
- Pleiotropy and the evolution of multivariate complex traits
- Spatially structured populations
- Speciation

Classical Models and Results

Selection in single-locus systems

- Alleles: $A_i, i = 1, \dots, k$
- Frequencies: p_i, p_i'
- Diploid genotypes: $A_i A_j$
- Fitness of $A_i A_j$: W_{ij}
- Marginal fitness of A_i : $W_i = \sum_j W_{ij} p_j$
- Fitness of $A_i A_j$: W_{ij}
- Mean (population) fitness: $\bar{W} = \sum_i W_i p_i$

Dynamics under selection

- Discrete time

$$p_i' = p_i \frac{W_i}{\bar{W}}, \quad i = 1, \dots, k$$

- Continuous time

$$\frac{dp_i}{dt} = p_i(W_i - \bar{W}), \quad i = 1, \dots, k$$

Dynamics under selection

- Equivalent formulation:

$$\frac{dp}{dt} = \frac{1}{2} \mathbf{G}_p \nabla \bar{W},$$

where $(\mathbf{G}_p)_{ij} = g_{ij} = p_i(\delta_{ij} - p_i)$.

- The continuous-time selection dynamics is a **Svirezhev-Sahshahani** gradient.

Fisher's Fundamental Theorem of Natural Selection

- Mean fitness is strictly increasing along trajectories except at equilibria.
- The rate of increase is proportional to the additive genetic variance in fitness. (Holds exactly only in continuous time.)

Further results

- An *internal* (polymorphic) equilibrium is asymptotically stable if and only if it is an isolated local maximum of \overline{W} .
- If there is an asymptotically stable internal equilibrium, then it attracts all trajectories from the interior of the simplex.
- Every trajectory converges to an equilibrium *point*.

Selection and mutation

- Mutation rate from $A_i \rightarrow A_j$: μ_{ij} ($\mu_{ii} = 0$)
- Mutation-selection dynamics:

$$p_i' = p_i \frac{W_i}{\bar{W}} + \frac{1}{\bar{W}} \sum_j (p_j W_j \mu_{ji} - p_i W_j \mu_{ij})$$

Mutation-selection dynamics

- It is, in general, not gradient like.
- Stable limit cycles may exist if there are more than two alleles.
- A unique, globally stable, internal equilibrium point exists if fitnesses are multiplicative, i.e., $W_{ij} = W_i W_j$ (additive in continuous time).

Mutation-selection dynamics

- If mutation rates satisfy the ‘house-of-cards’ condition,

$$\mu_{ij} = \mu_j \quad \text{for all } j \neq i,$$

then there exists a Lyapunov function and the dynamics is gradient like.

- If the pure selection dynamics admits an asymptotically stable, internal equilibrium, then also the house-of-cards mutation dynamics.

The generalized Haldane principle

- Mutation load:

$$L = \frac{W_{\max} - \bar{W}}{W_{\max}}$$

- At any equilibrium, the mutation load is – to first-order approximation – given by the total mutation rate to less fit alleles, thus independent of fitness.

Recombination

- Two loci, two alleles: $A_1, A_2; B_1, B_2$
- Four gametes: $A_1 B_1, A_1 B_2, A_2 B_1, A_2 B_2$
- Frequencies: x_1, x_2, x_3, x_4
- Recombination rate: r
- Linkage disequilibrium:

$$D = x_1 - (x_1 + x_2)(x_1 + x_3) = \dots = x_1 x_4 - x_2 x_3$$

Dynamics under recombination

$$x_1' = x_1 - rD$$

$$x_2' = x_2 + rD$$

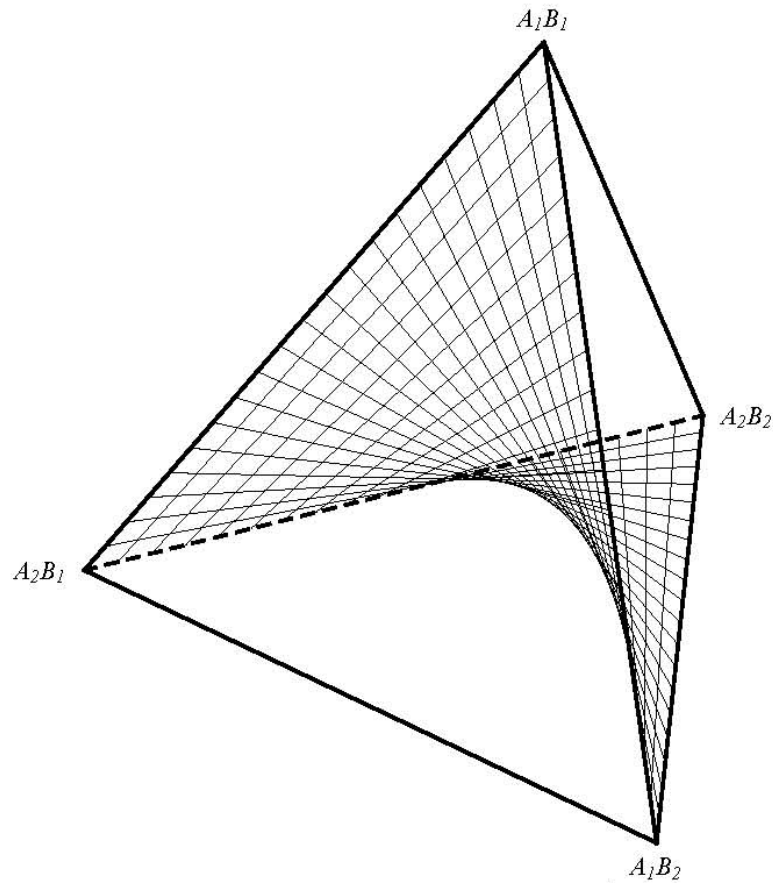
$$x_3' = x_3 + rD$$

$$x_4' = x_4 - rD$$

- Linkage disequilibrium decays geometrically

$$D' = (1 - r)D$$

The state space (= simplex) with the linkage equilibrium manifold, $D = 0$



Recombination and selection

- Dynamics may be complicated, i.e., stable limit cycles may exist

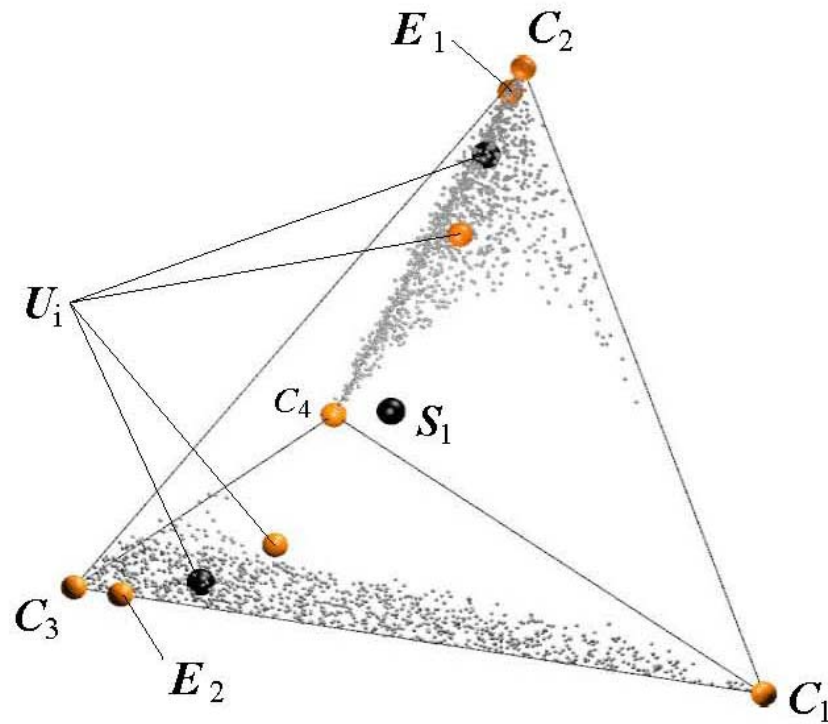
- ‘Only’ the case of additive fitnesses,

$$w(A_i B_j / A_k B_l) = w(A_i A_k) + w(B_j B_l),$$

is simple, i.e., mean fitness is increasing

- Many (internal) equilibria may coexist

Two diallelic loci under Gaussian stabilizing selection



Quantitative genetics

- **Quantitative traits** are characters that vary continuously and can be measured on a metric scale:
 - body weight
 - height
 - morphological measurements
 - (milk) yield
 - fitness

Quantitative traits

- have complex genetics, i.e., they are determined by many genes, most of them with small effects,
- have genetic and environmental (including developmental) components,
- are often normally distributed (on an appropriate scale)

The additive model

$$P = G + E = \sum_i (X_i + X_i^*) + E$$

$$\bar{P} = \bar{G} \quad \text{and} \quad V_P = V_G + V_E.$$

X_i / X_i^* : maternal/paternal effect of alleles at
locus i

\bar{P}, \bar{G} : phenotypic and genotypic mean values of
the trait,

V_P, V_G, V_E : phenotypic, genotypic and
environmental variances, resp.

The breeder's equation

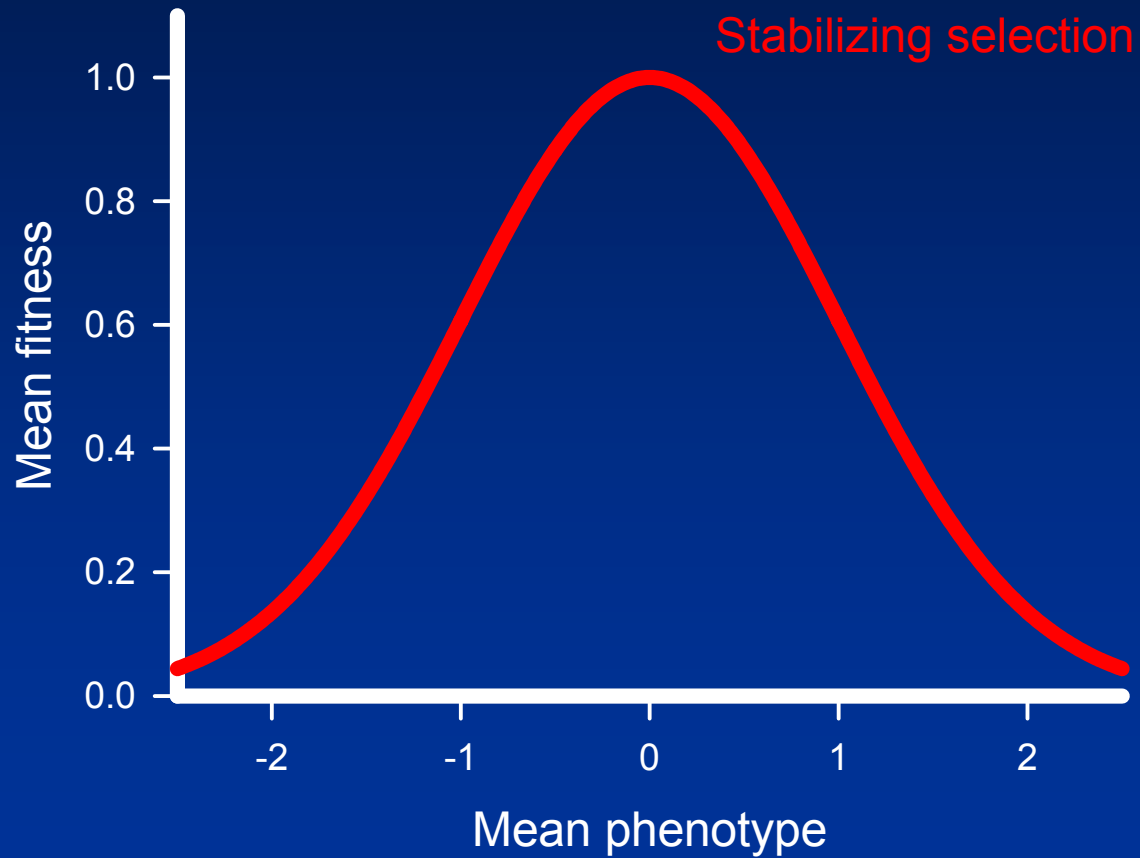
$$\Delta\bar{P} = \bar{P}' - \bar{P} = h^2 (\bar{P}_s - \bar{P}),$$

where \bar{P}_s is the mean phenotype after selection but before reproduction and

$$h^2 = V_G / V_P$$

is the **heritability**, the ratio of additive genetic to phenotypic variance.

Adaptive landscape



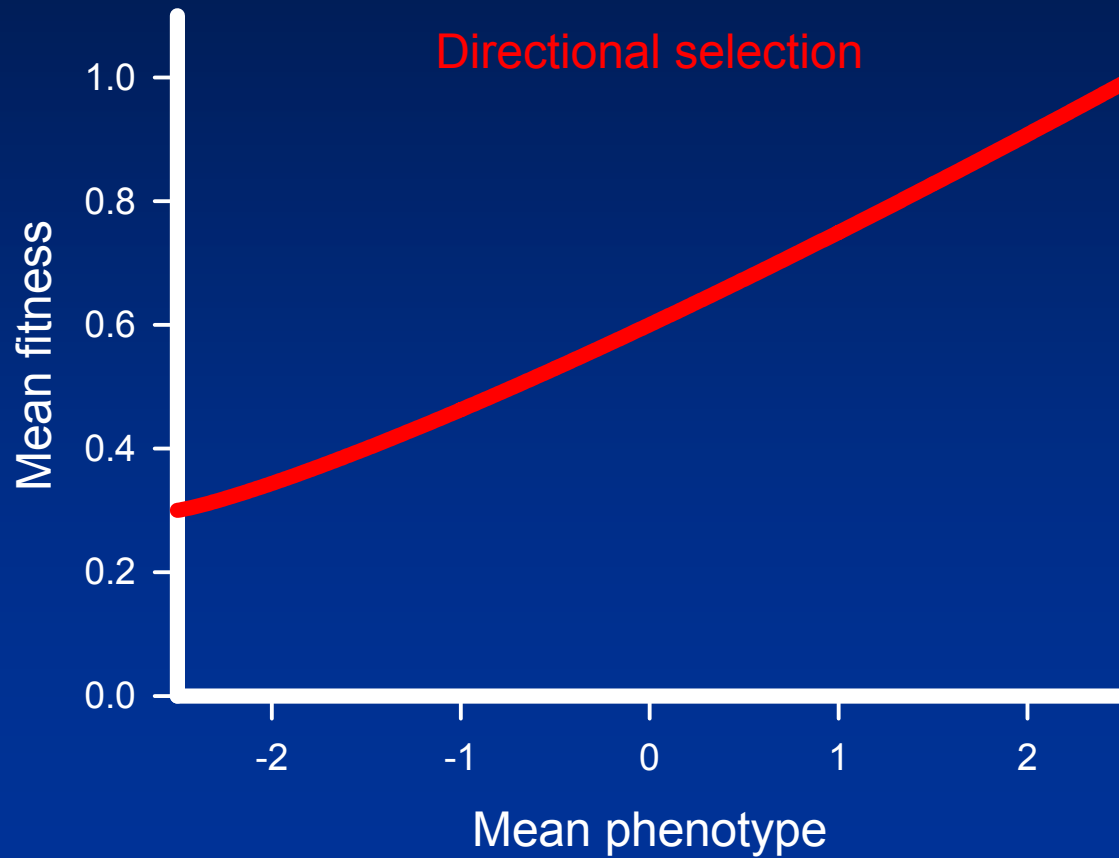
The adaptive landscape

- Summarizes the selection pressures that act on a population and direct its evolution.
- Its height represents mean fitness of the population as a function of the mean phenotype.
- Its slope and curvature determine the strength of directional and stabilizing selection.

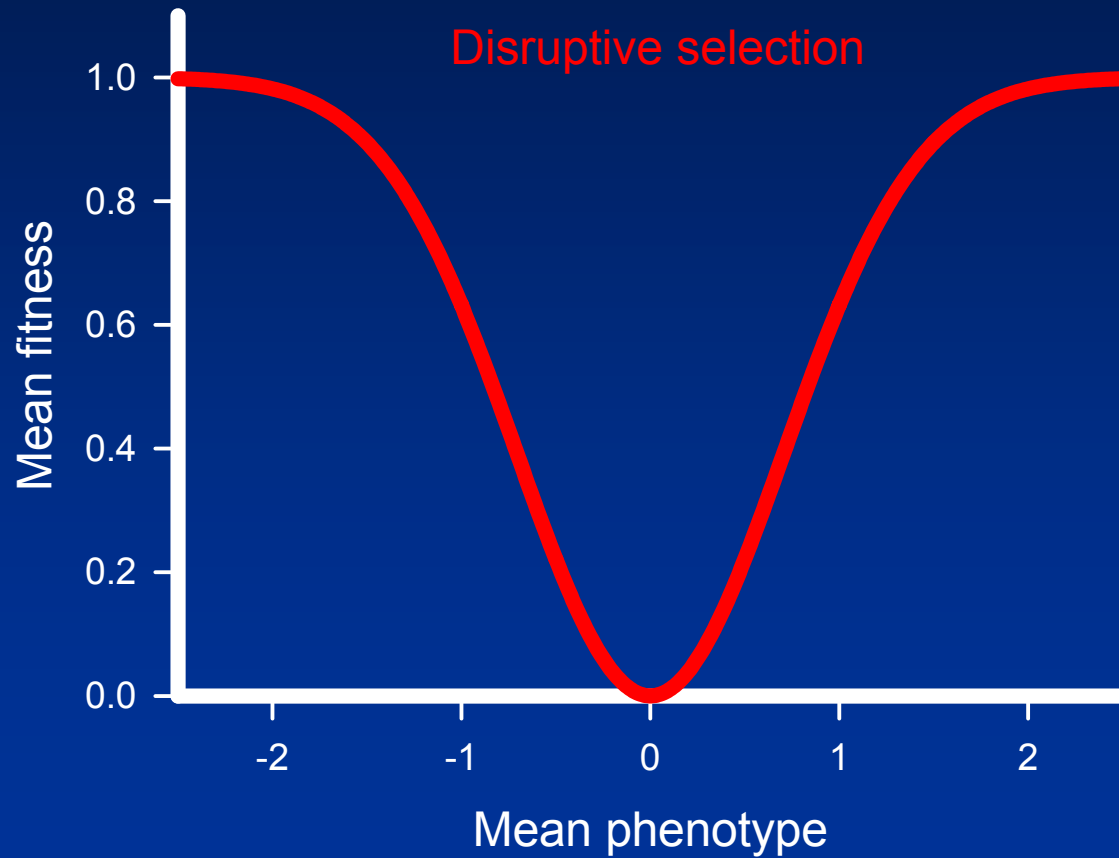
The adaptive landscape

- Has to be distinguished from the individual's fitness landscape.
- It depends strongly on the distribution of phenotypes in the population.

Adaptive landscapes



Adaptive landscapes



Evolution on adaptive landscapes

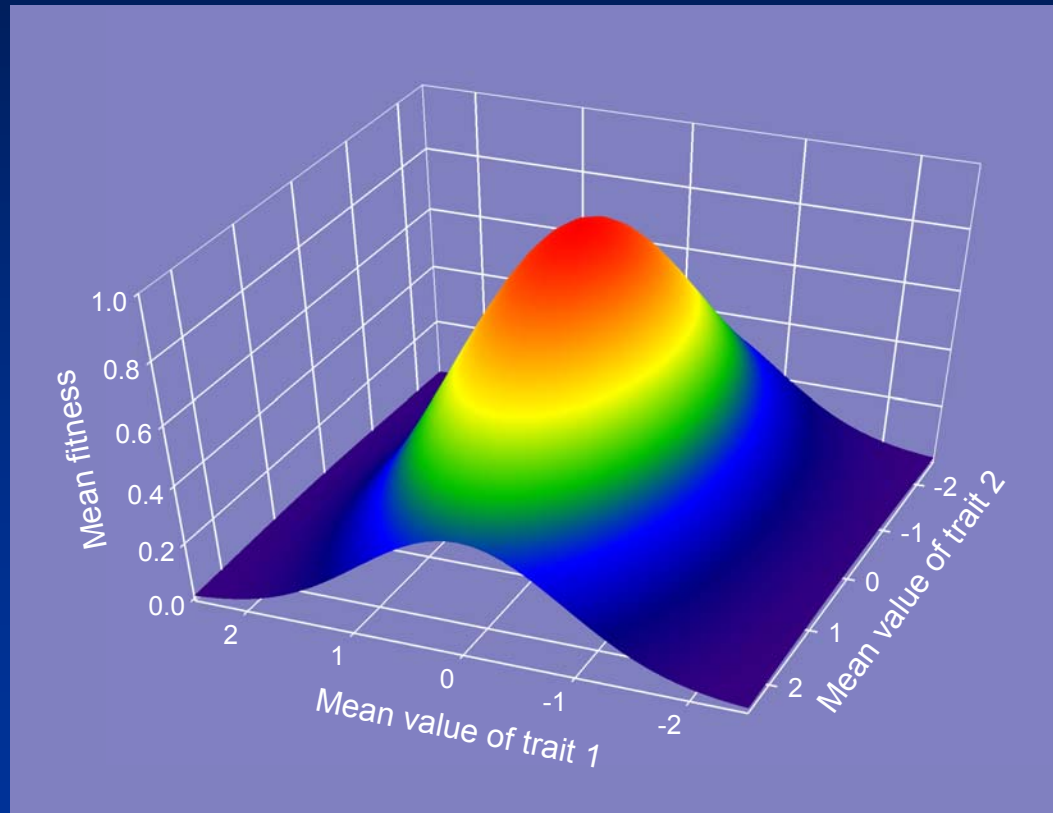
- Lande's equation:

$$\Delta \bar{P} = V_G \frac{d \ln \bar{W}}{d \bar{P}},$$

where \bar{W} is the mean fitness.

- This is derived by *assuming* a Gaussian distribution of trait values, P , and a linear parent-offspring correlation.

Multivariate adaptive landscapes



Multivariate evolutionary dynamics

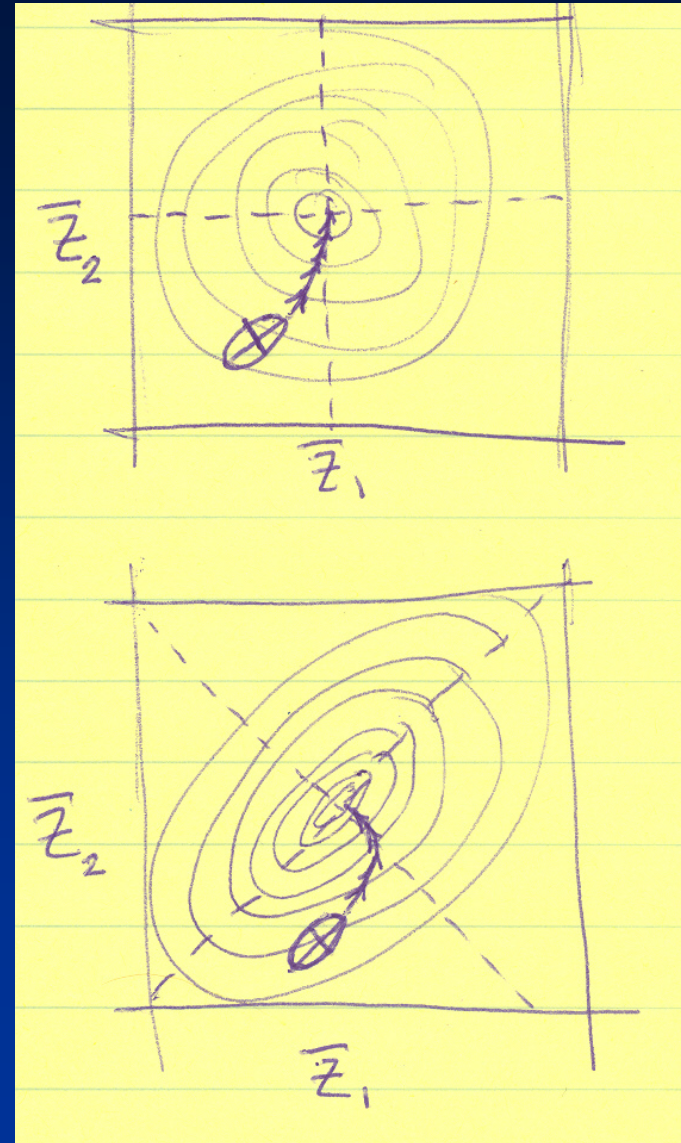
- Lande's multivariate equation:

$$\Delta \bar{\mathbf{P}} = \mathbf{G} \nabla \ln \bar{W},$$

where $\bar{\mathbf{P}} = (\bar{P}_1, \dots, \bar{P}_n)$ is the vector of mean phenotypic values, \mathbf{G} , the genetic variance-covariance matrix, and ∇ denotes the gradient vector $(\partial / \partial \bar{P}_1, \dots, \partial / \partial \bar{P}_n)$.

- Direction and magnitude of selection response depend crucially on \mathbf{G} .

Correlated evolution
of traits
governed by
pleiotropic mutation



Problems

- Can these equations be derived from ‘first principles’?
- When is G (V_G) constant over (evolutionary) time?
- How does G depend on the pattern of pleiotropic mutations?
- How is G shaped by linkage and by selection?
- Does it evolve in response to selection?