# Adaptation in polygenic traits
## Criteria for sweeps and shifts

Joachim Hermisson

Mathematics & Biology, University of Vienna

# Adaptation in polygenic traits

Ilse Hölliger
University of Vienna

Pleuni Pennings
San Francisco State University

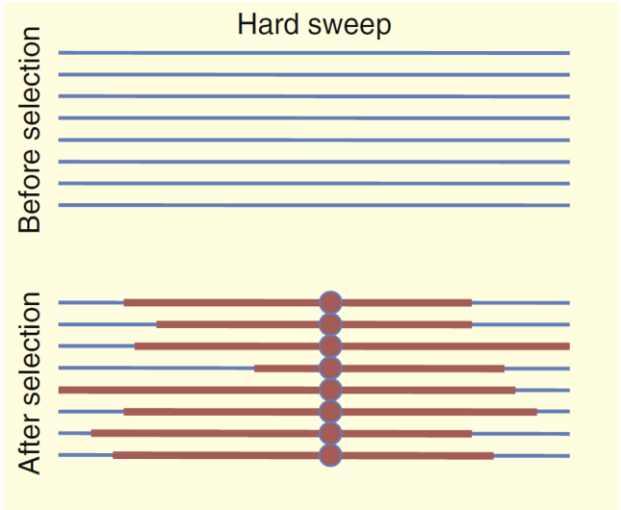
# Adaptive scenarios

*molecular popgen* ⟷ *quantitative genetics*

## "Sweeps"

- adaptation due to independent large changes at single loci

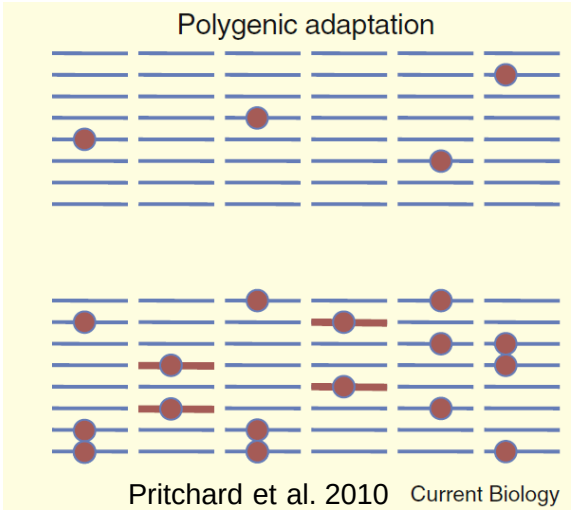
➢ clear molecular footprint




## "Shifts"

- adaptation due to small collective shifts at many loci

➢ no clear sweep pattern



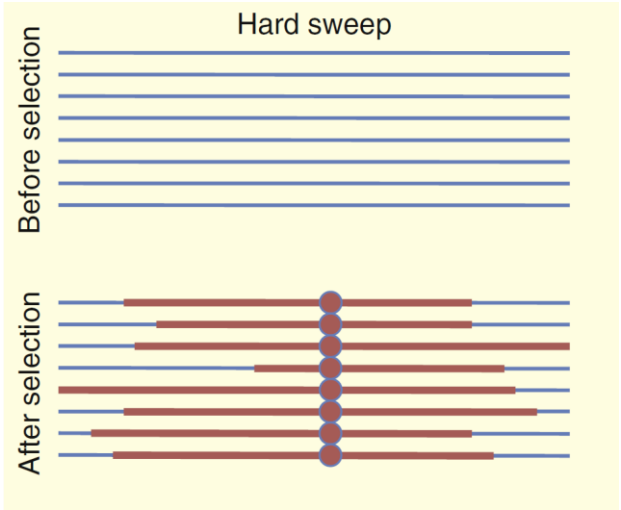
# Adaptive scenarios

*molecular popgen* ⟷ *quantitative genetics*

## "Sweeps"

- adaptation due to independent large changes at single loci

➢ clear molecular footprint



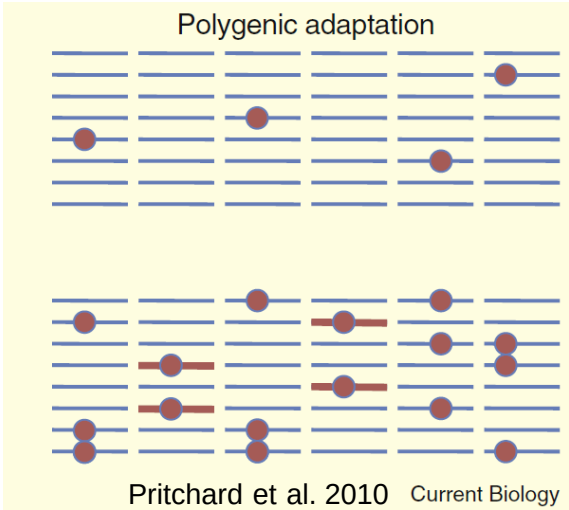

*Which scenario should we expect*

?

## "Shifts"

- adaptation due to small collective shifts at many loci

➢ no clear sweep pattern



# Adaptive scenarios

*Assume:*

- panmictic population, new selection pressure
- adaptation from mutation-selection-drift balance

*few loci* ⟷ *many loci*

*new mutation* ⟷ *standing genetic variation*

*weak selection* ⟷ *strong selection*

*weak mutation* ⟷ *strong mutation*

# Adaptive scenarios

*Assume:*

- panmictic population, new selection pressure
- adaptation from mutation-selection-drift balance

*few loci* ⟷ *many loci*

*new mutation* ⟷ *standing genetic variation*

*weak selection* ⟷ *strong selection*

*weak mutation* ⟷ *strong mutation*

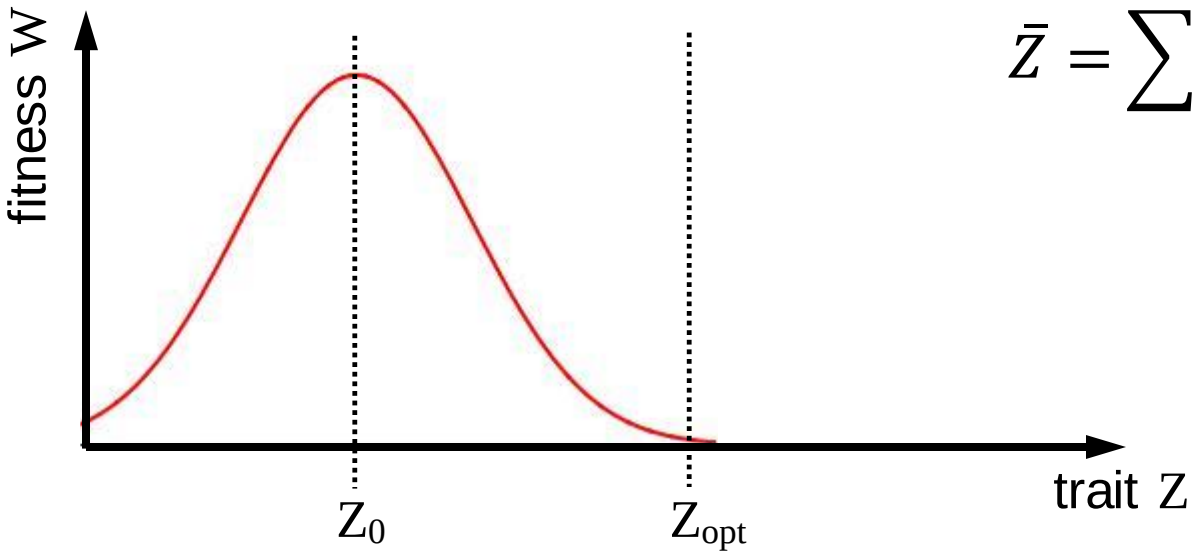
➢ *Which scenario is favored under which conditions ?*

# Additive quantitative trait
## under stabilizing selection

- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

$$W(Z) = \exp\left[-\frac{\sigma}{2}(Z - Z_{\text{opt}})^2\right]$$

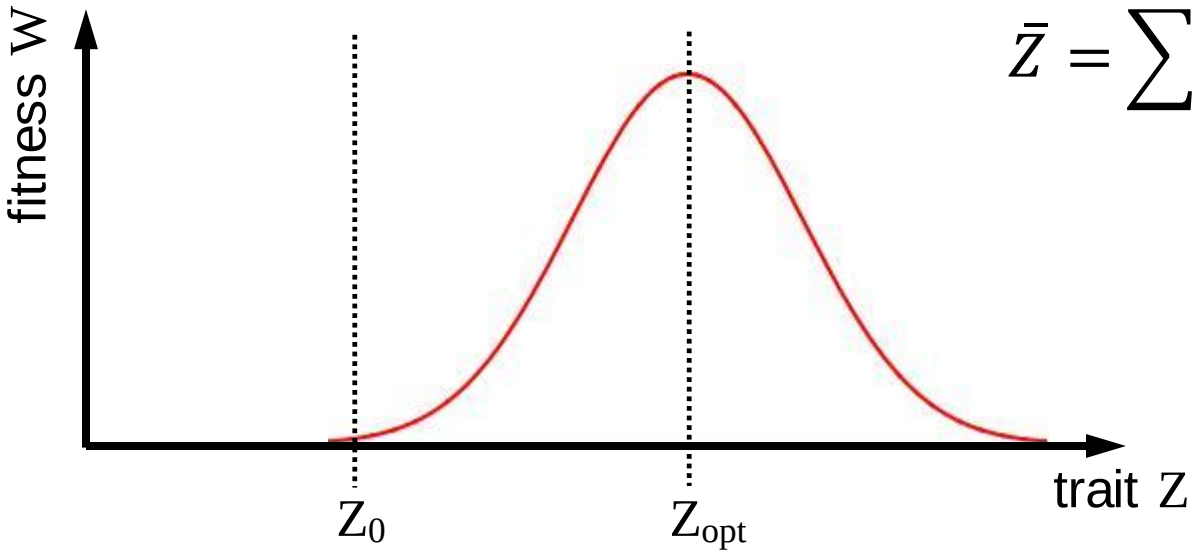
$$\bar{Z} = \sum_{i=1}^{L} \gamma p_i$$



# Additive quantitative trait
## under stabilizing selection

- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

$$W(Z) = \exp\left[-\frac{\sigma}{2}(Z - Z_{\text{opt}})^2\right]$$



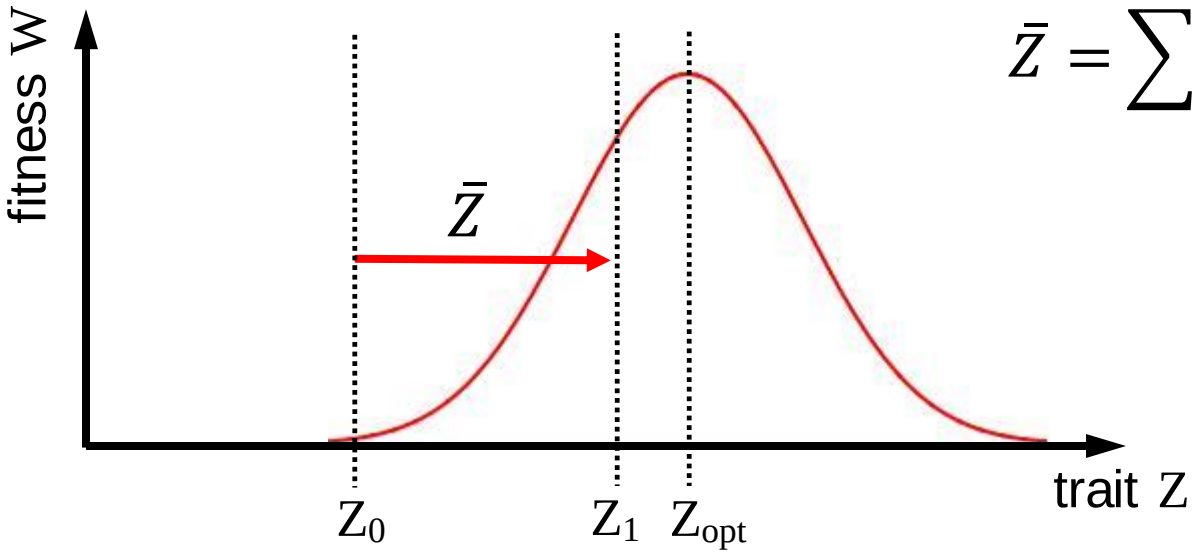

$$\bar{Z} = \sum_{i=1}^{L} \gamma p_i$$

# Additive quantitative trait
## under stabilizing selection

- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

$$W(Z) = \exp\left[-\frac{\sigma}{2}(Z - Z_{\text{opt}})^2\right]$$

$$\bar{Z} = \sum_{i=1}^{L} \gamma p_i$$

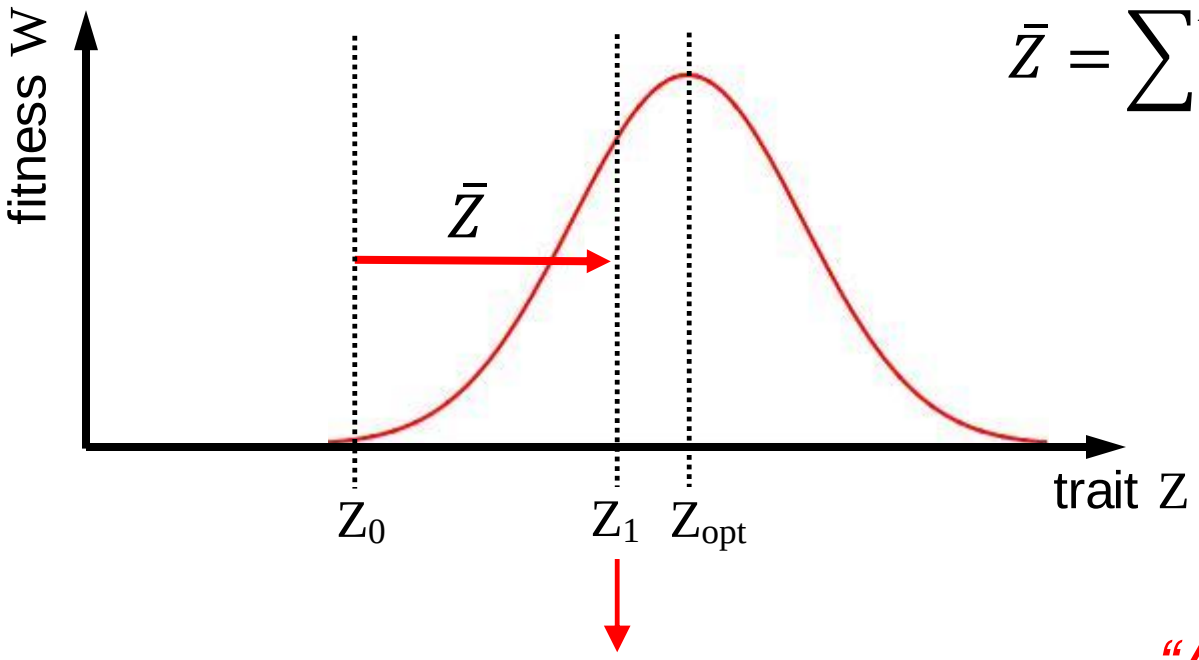


# Additive quantitative trait under stabilizing selection

- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

$$W(Z) = \exp\left[-\frac{\sigma}{2}(Z - Z_{\text{opt}})^2\right]$$

$$\bar{Z} = \sum_{i=1}^{L} \gamma p_i$$



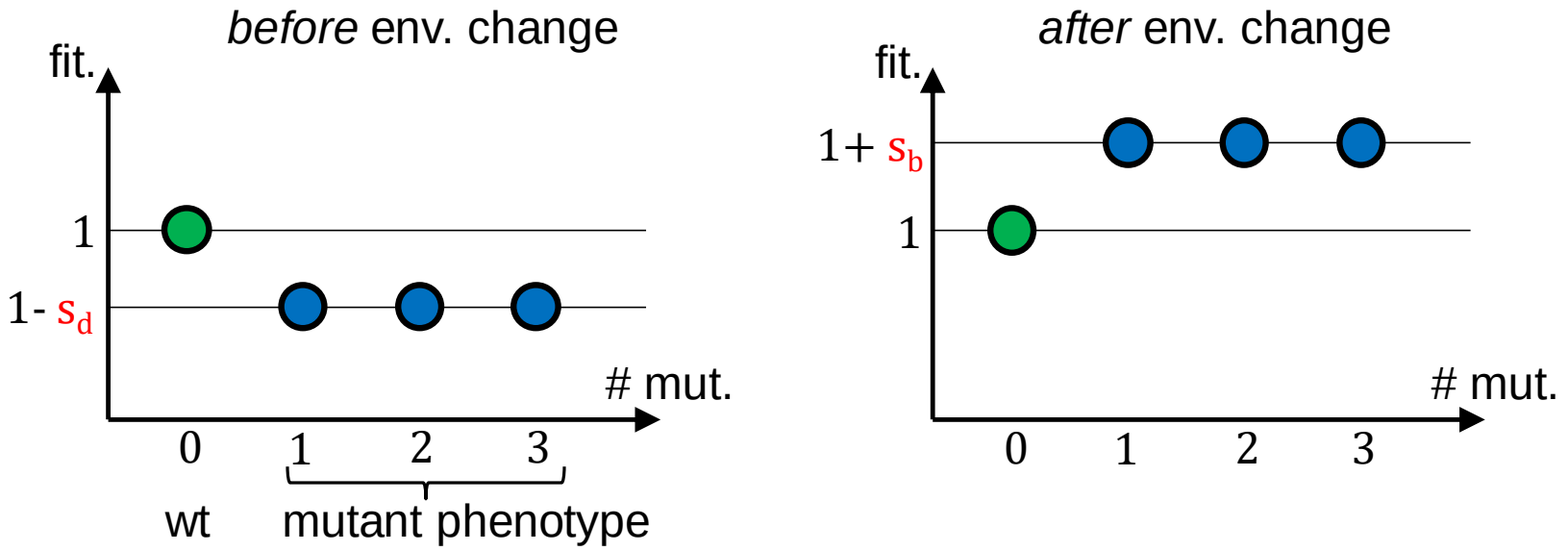

joint distribution of allele frequencies $p_i$:

*"Architecture of polygenic adaptation"*

# Binary trait
## with complete redundancy

- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

- fitness function (e.g. resistance):



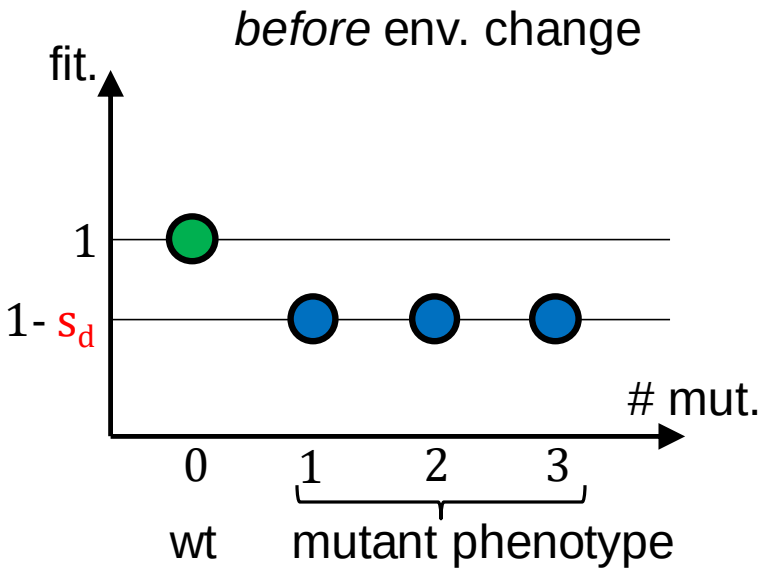
# Binary trait
## with complete redundancy

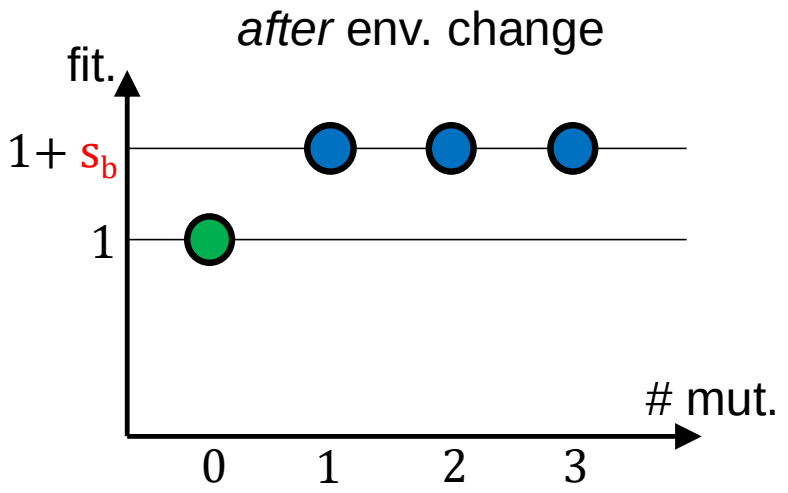
- $N$ haploids, $L$ biallelic loci
- recurrent mutation $\theta_i = 2Nu_i$

- fitness function (e.g. resistance):

$$W(Z) = 1 \pm s_{b,d}\, Z$$

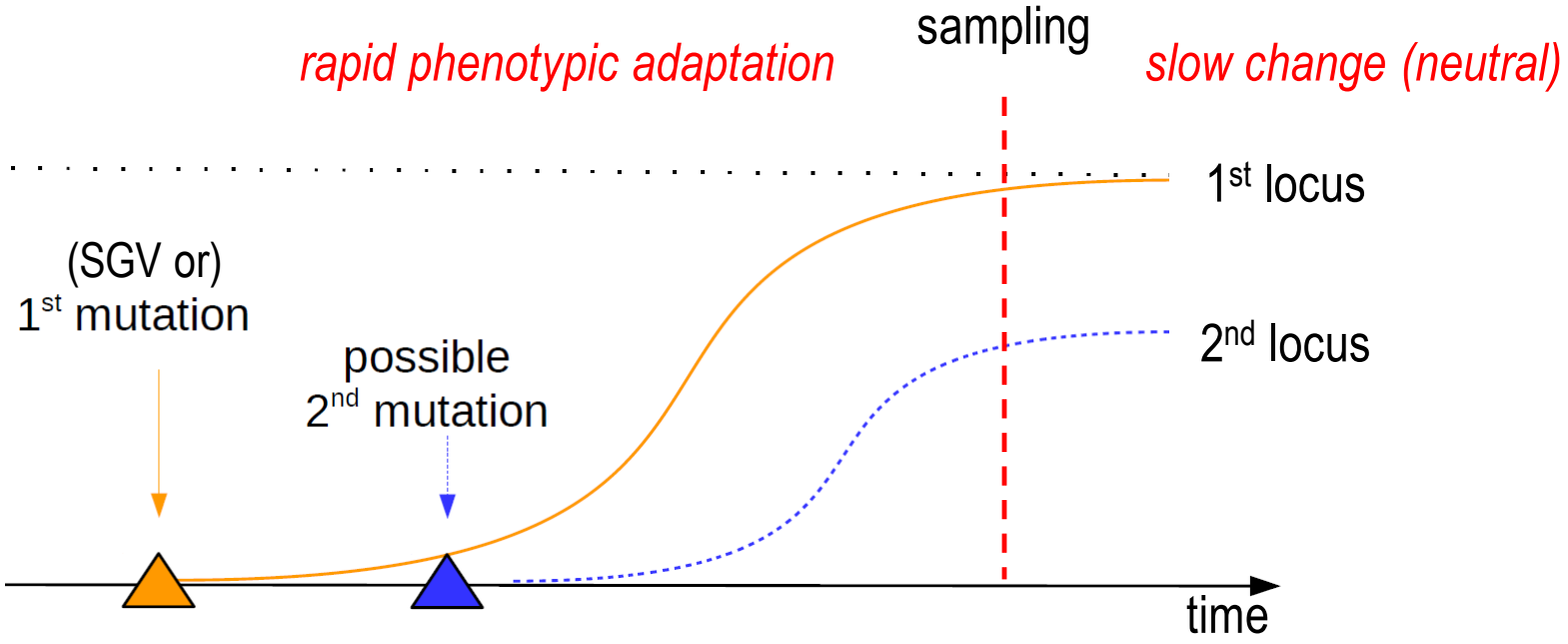
$$\bar{Z} = \underbrace{1 - \prod_{i=1}^{L}(1 - p_i)}_{\text{freq. of mutant phenotype}}$$



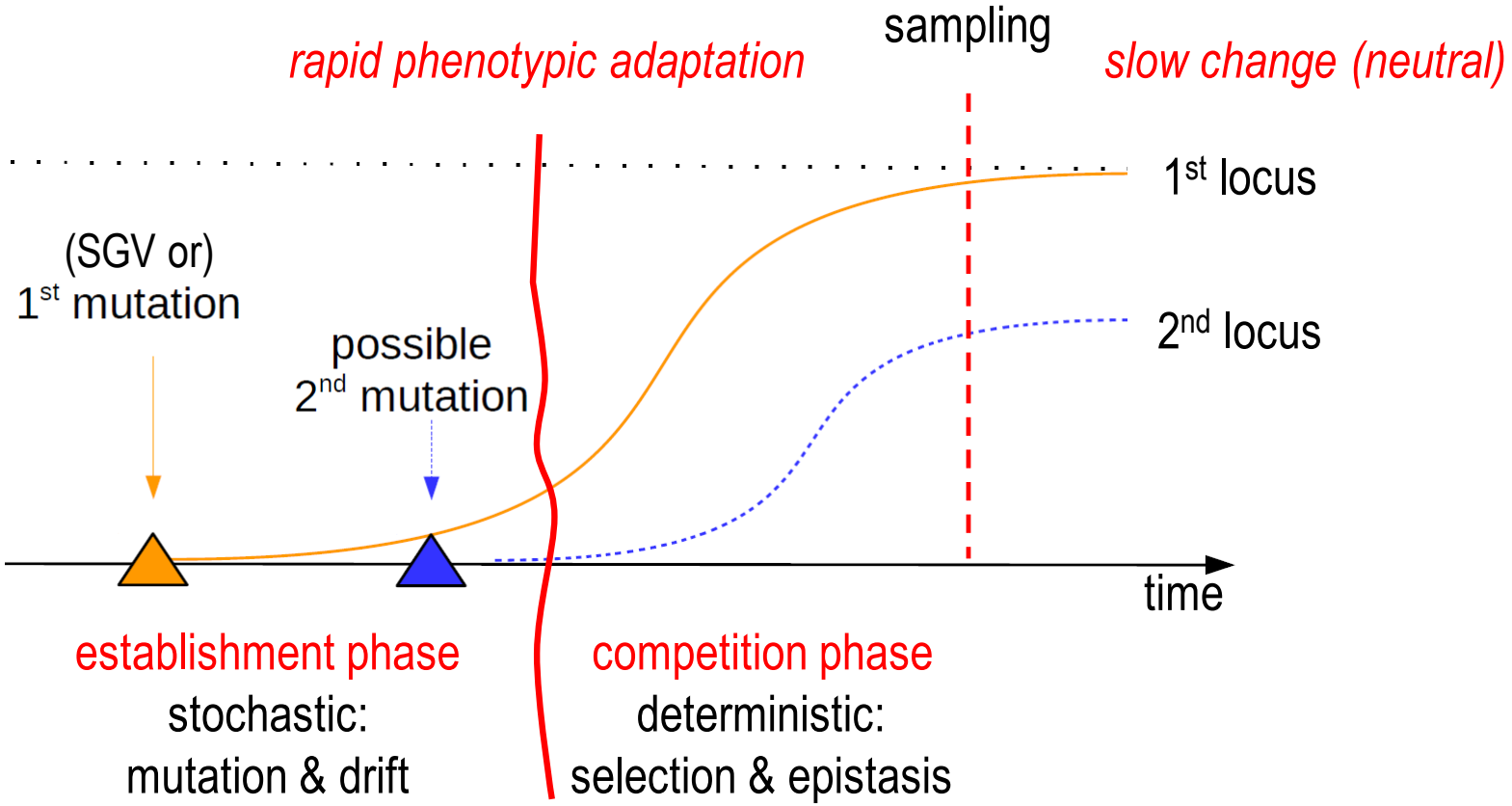
# Maths of polygenic adaptation

Evolutionary trajectories (2 loci, schematic):



# Maths of polygenic adaptation

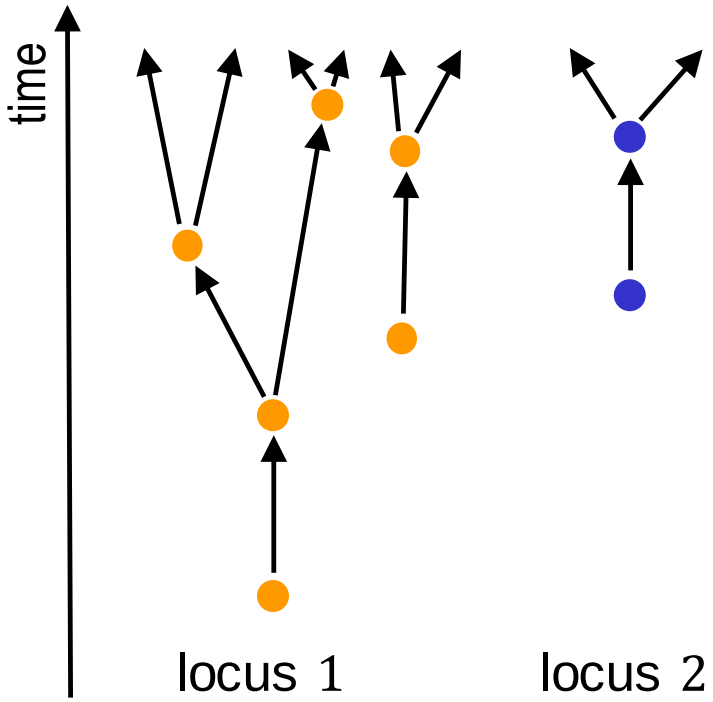
Evolutionary trajectories (2 loci, schematic):



# Maths of polygenic adaptation

Establishment phase (both models): Yule branching process

- track only mutant copies destined for establishment, prob. $p_{\mathrm{est}}(s_b, s_d; \sigma, \gamma)$




- new lines (mutation) $\sim \theta_i \cdot p_{\mathrm{est}}$ per locus
- split rate (reproduction) $\sim p_{\mathrm{est}}$ per line

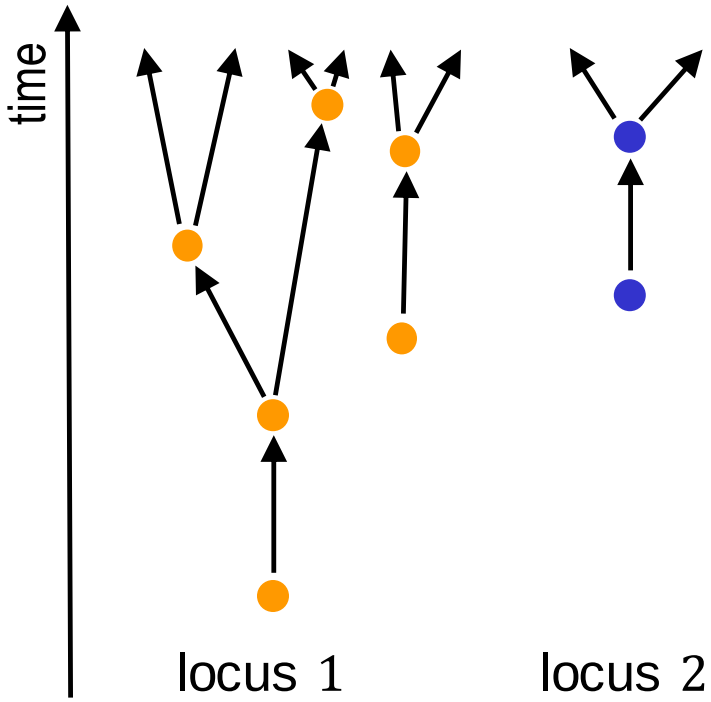
$n_1, n_2, \ldots$ copies at all loci

*mutation and drift during establishment create stochastic differences among loci*

# Maths of polygenic adaptation

Establishment phase (both models): Yule branching process

- track only mutant copies destined for establishment, prob. $p_{\mathrm{est}}(s_b, s_d; \sigma, \gamma)$




- new lines (mutation) $\sim \theta_i \cdot p_{\mathrm{est}}$ per locus
- split rate (reproduction) $\sim p_{\mathrm{est}}$ per line

$n_1, n_2, \ldots$ copies at all loci

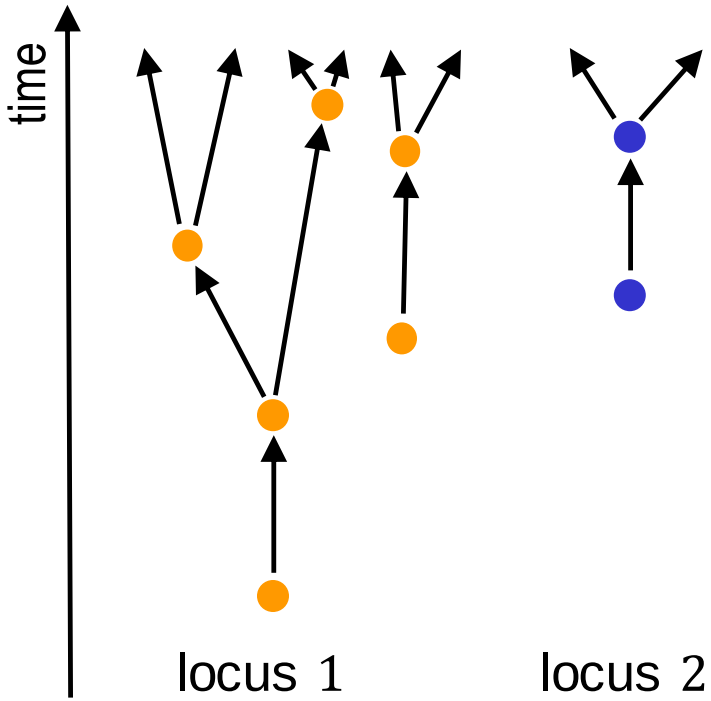
*mutation and drift during establishment create stochastic differences among loci*

- ratios $x_i = n_i/n_1$ independent of $s_{b/d}; \sigma, \gamma$

# Maths of polygenic adaptation

Establishment phase (both models): Yule branching process

- track only mutant copies destined for establishment, prob. $p_{\text{est}}(s_b, s_d; \sigma, \gamma)$




- new lines (mutation) $\sim \theta_i \cdot p_{\text{est}}$ per locus
- split rate (reproduction) $\sim p_{\text{est}}$ per line

$n_1, n_2, \ldots$ copies at all loci

*mutation and drift during establishment create stochastic differences among loci*

- ratios $x_i = n_i/n_1$ independent of $s_{b/d}; \sigma, \gamma$
- joint distribution of frequency ratios $x_i$ depends only on mutation rates $\theta_i$:

$$P_{\text{inDir}}[\mathbf{x}|\boldsymbol{\Theta}] = \frac{1}{B[\boldsymbol{\Theta}]} \prod_{j=2}^{L} x_j^{\Theta_j - 1} \left(1 + \sum_{i=2}^{L} x_i\right)^{-\sum_{i=1}^{L} \Theta_i}$$

(inverted Dirichlet distribution)

# Maths of polygenic adaptation

Competition phase: binary trait model

- deterministic dynamics

$$\dot{p}_i = p_i s_b (1 - \bar{Z}) \qquad \boxed{\Rightarrow \frac{d}{dt}\left(\frac{p_i}{p_j}\right) = 0}$$

- maintains ratios
  - ➢ zooms up differences

# Maths of polygenic adaptation

Competition phase: binary trait model

- deterministic dynamics

$$\dot{p}_i = p_i s_b (1 - \bar{Z}) \qquad \boxed{\Rightarrow \frac{d}{dt}\left(\frac{p_i}{p_j}\right) = 0}$$

- maintains ratios
  - ➢ zooms up differences

- joint distribution of mutant frequencies $p_i$ at $\bar{Z} = 1 - f_w$ :

$$\mathsf{P}_{f_w}[\mathbf{p}|\boldsymbol{\Theta}] = \frac{\delta_{\prod_{j=1}^{L}(1-p_j)-f_w}}{B[\boldsymbol{\Theta}]} \prod_{j=1}^{L} p_j^{\Theta_j - 1} \left(\sum_{i=1}^{L} p_i\right)^{-\sum_{i=1}^{L}\Theta_i} \left(\sum_{j=1}^{L} \frac{f_w p_j}{1-p_j}\right)$$

# Maths of polygenic adaptation

Competition phase: binary trait model

- deterministic dynamics

$$\dot{p}_i = p_i s_b (1 - \bar{Z}) \quad \boxed{\Rightarrow \frac{d}{dt}\left(\frac{p_i}{p_j}\right) = 0}$$

- maintains ratios
- ➢ zooms up differences

- joint distribution of mutant frequencies $p_i$ at $\bar{Z} = 1 - f_w$ :

$$\mathrm{P}_{f_w}[\mathbf{p}|\boldsymbol{\Theta}] = \frac{\delta_{\prod_{j=1}^{L}(1-p_j)-f_w}}{B[\boldsymbol{\Theta}]} \prod_{j=1}^{L} p_j^{\Theta_j - 1} \left(\sum_{i=1}^{L} p_i\right)^{-\sum_{i=1}^{L}\Theta_i} \left(\sum_{j=1}^{L} \frac{f_w p_j}{1-p_j}\right)$$

- depends only on mutation rates $\theta_i$
- independent of selection strength $s_b, s_d$

# Maths of polygenic adaptation

Competition phase: quantitative trait

- deterministic dynamics (LE and weak selection)

[DeVladar/Barton 2014
Jain/Stephan 2017]

$$\dot{p}_i = p_i(1-p_i)\Big[\underbrace{\sigma\gamma\left(Z_{\text{opt}} - \bar{Z}\right)}_{\text{directional selection}} + \underbrace{\sigma\gamma^2(2p_i - 1)}_{\text{disruptive selection}}\Big]$$

# Maths of polygenic adaptation

Competition phase: quantitative trait

- deterministic dynamics (LE and weak selection)

[DeVladar/Barton 2014
Jain/Stephan 2017]

$$\dot{p}_i = p_i(1-p_i)\,\underbrace{\sigma\gamma\left(Z_{\text{opt}} - \bar{Z}\right)}_{\text{directional selection}}$$

# Maths of polygenic adaptation

Competition phase: quantitative trait

- deterministic dynamics (LE and weak selection)

[DeVladar/Barton 2014
Jain/Stephan 2017]

$$\dot{p}_i = p_i(1-p_i)\,\sigma\gamma\left(Z_{\text{opt}} - \bar{Z}\right) \quad y_i := \frac{p_i}{1-p_i}$$

$$\Rightarrow \frac{d}{dt}\left(\frac{y_i}{y_j}\right) = 0$$

# Maths of polygenic adaptation

Competition phase: quantitative trait

- deterministic dynamics (LE and weak selection)

[DeVladar/Barton 2014
Jain/Stephan 2017]

$$\dot{p}_i = p_i(1-p_i)\,\sigma\gamma\left(Z_{\text{opt}} - \bar{Z}\right) \quad y_i := \frac{p_i}{1-p_i} \qquad \Rightarrow \frac{d}{dt}\left(\frac{y_i}{y_j}\right) = 0$$

- joint distribution of mutant frequencies $p_i$ at $\bar{Z} = Z_1 = \gamma c_Z$ :

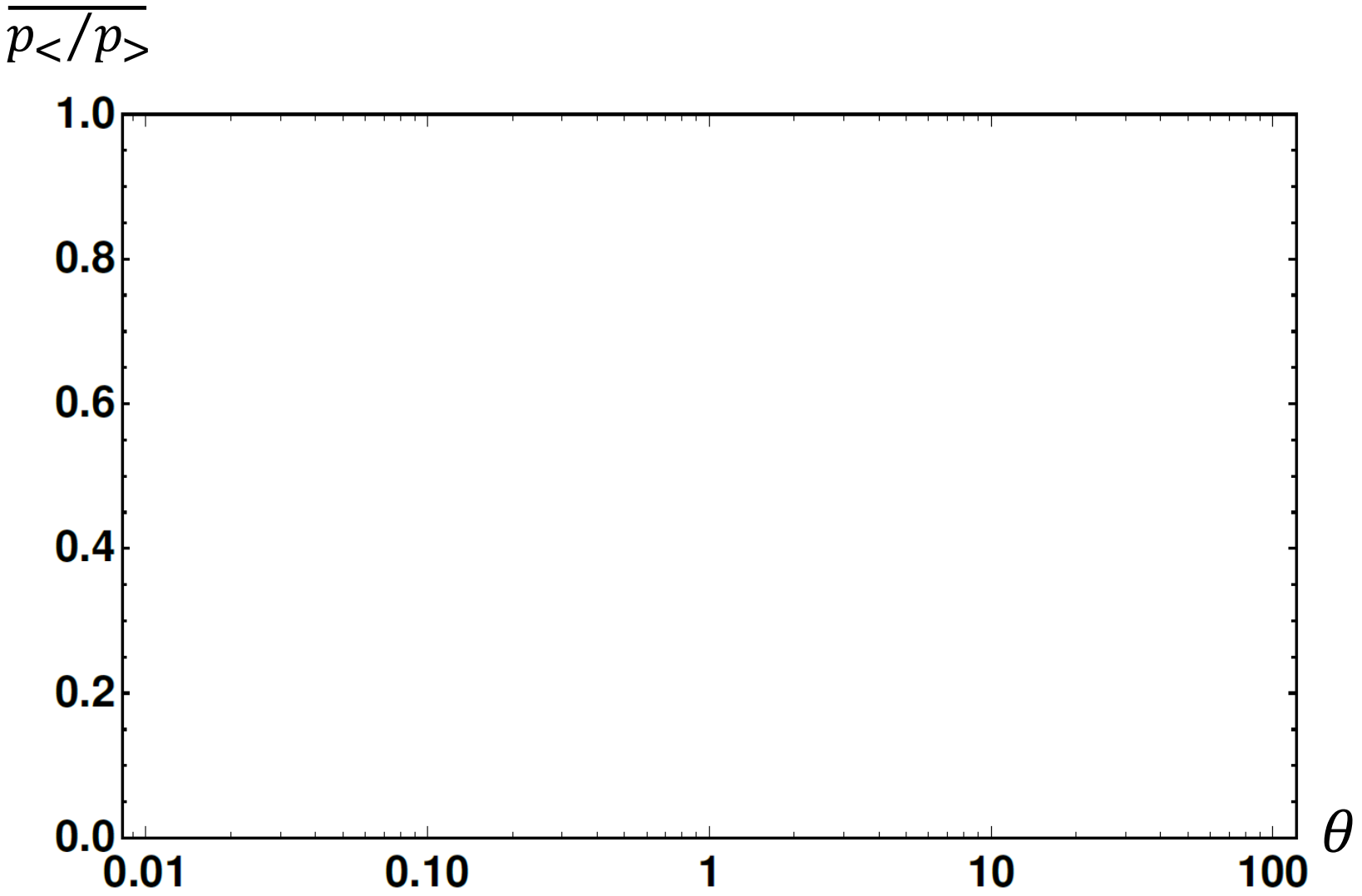
$$\mathsf{P}[\{p_i\}|c_Z] = C_\Gamma \prod_{i=2}^{L} \frac{p_i^{\Theta_i - 1}}{(1-p_i)^{\Theta_i+1}} \left(1 + \frac{\sum_{i=2}^{L} p_i(1-p_i)}{(c_Z - \sum_{k=2}^{L} p_k)(1 + \sum_{k=2}^{L} p_k - c_Z)}\right)$$
$$\cdot \left(1 + \frac{1 + \sum_{k=2}^{L} p_k - c_Z}{c_Z - \sum_{k=2}^{L} p_k} \sum_{i=2}^{L} \frac{p_i}{1-p_i}\right)^{-\Theta} \left(\frac{c_Z - \sum_{k=2}^{L} p_k}{1 + \sum_{k=2}^{L} p_k - c_Z} + \sum_{2=1}^{L} \frac{p_i}{1-p_i}\right)^{-\sum_{2=1}^{L} \Theta_i}$$

- depends only on mutation rates $\theta_i$
- independent of locus effect and selection strength $\gamma, \sigma$

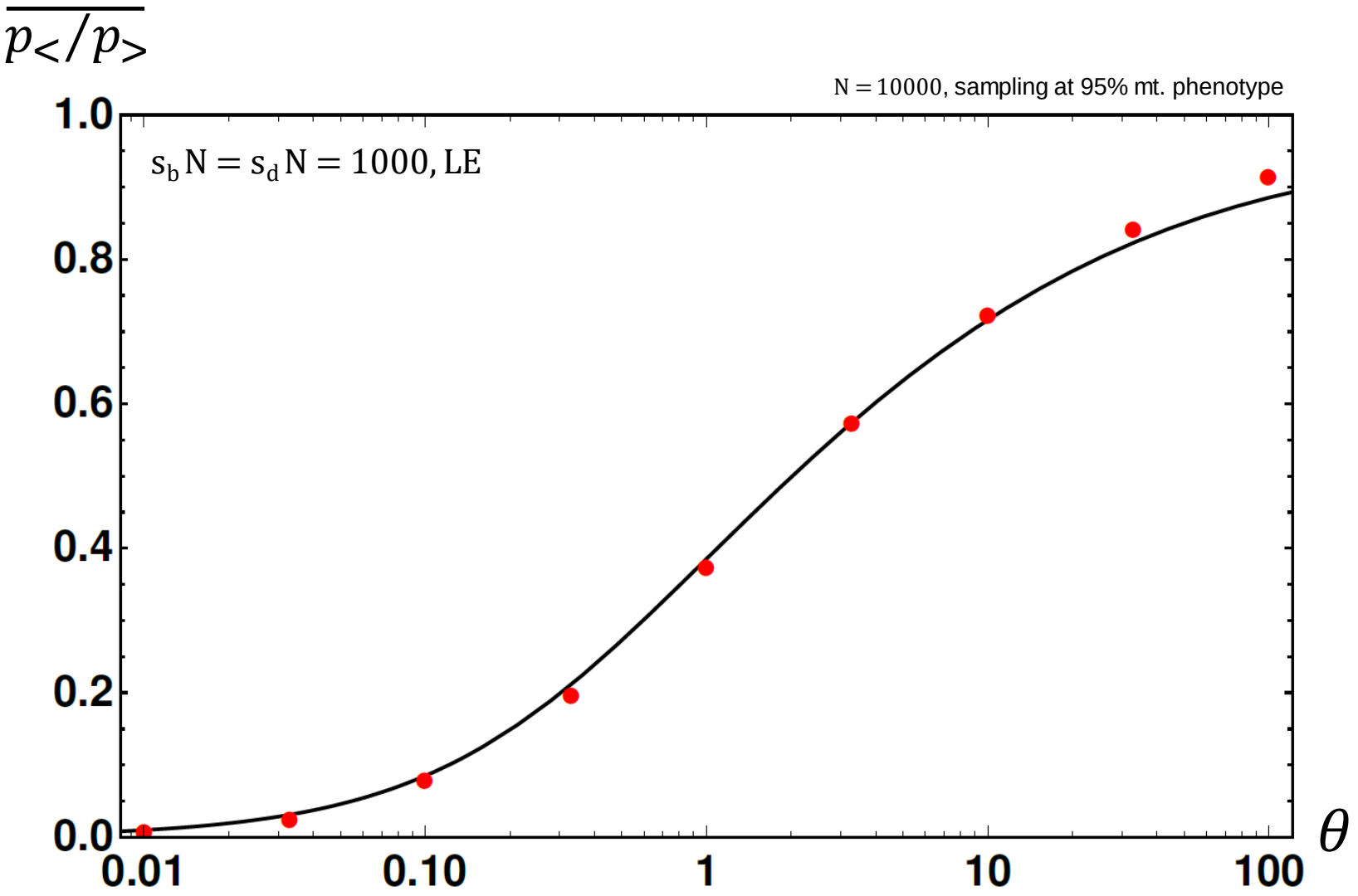
# Results: binary trait

- equal loci, $\theta_i = \theta$
- start in mutation-selection-drift balance
- adaptation until 95% mt phenotypes $(\bar{Z} = 1 - f_w = 0.95)$
- loci ordered according to their contribution to the adaptive response:
  - locus with largest frequency: *major locus*
  - all other loci: *minor loci*

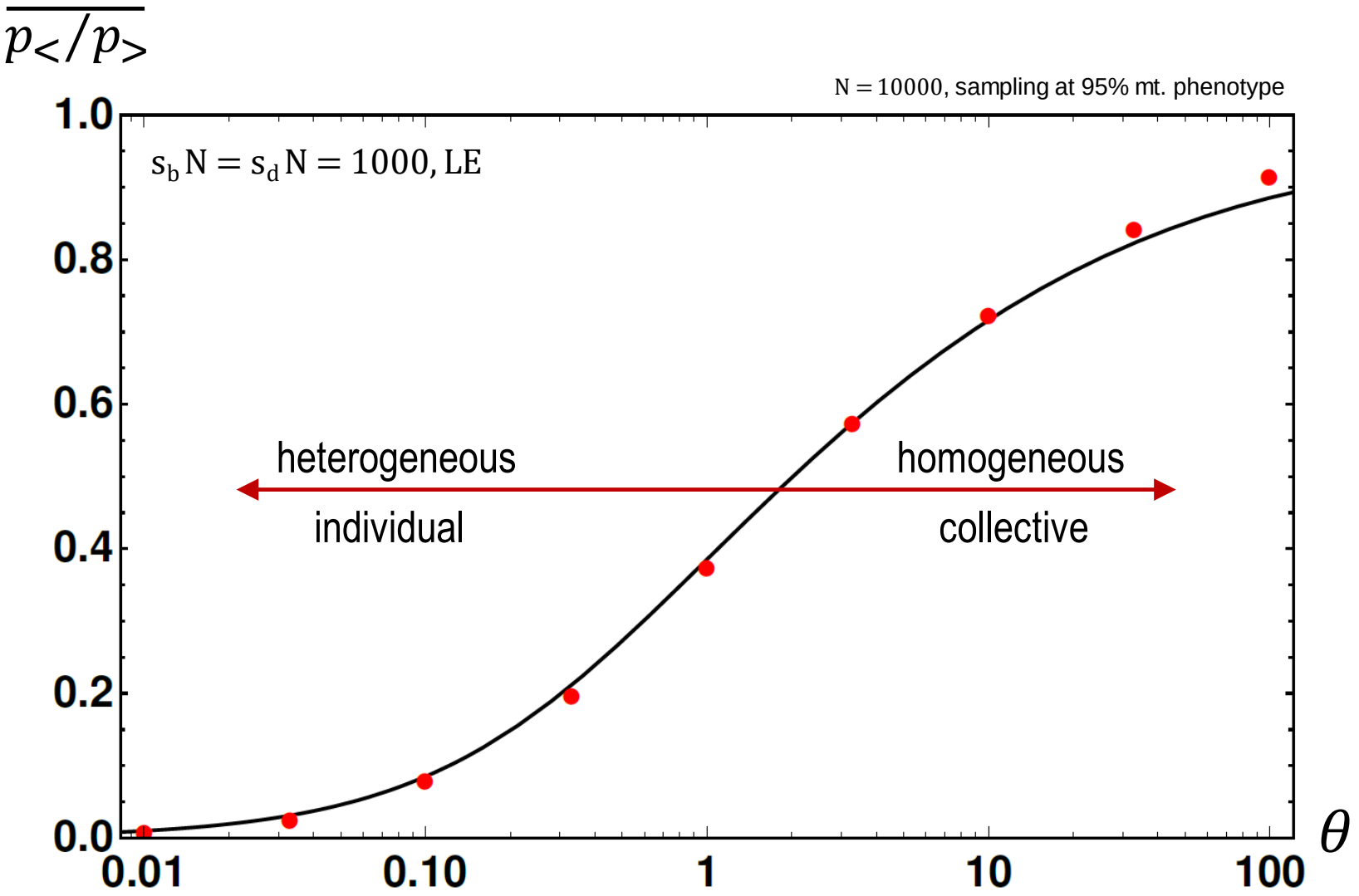
# Relative adaptive response (2 loci)



# Relative adaptive response (2 loci)



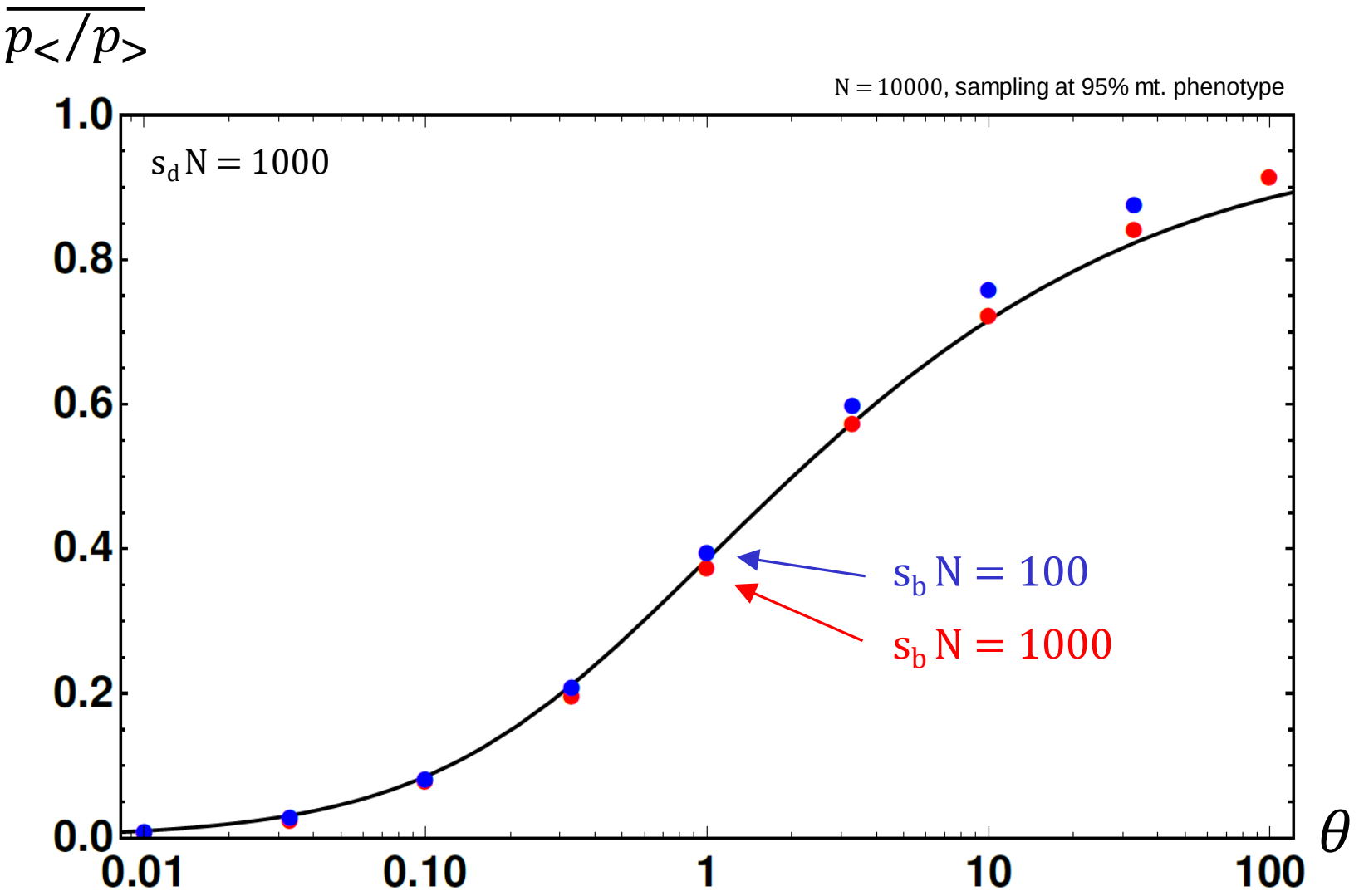
# Relative adaptive response (2 loci)



# Relative adaptive response (2 loci)

Influence of selection?

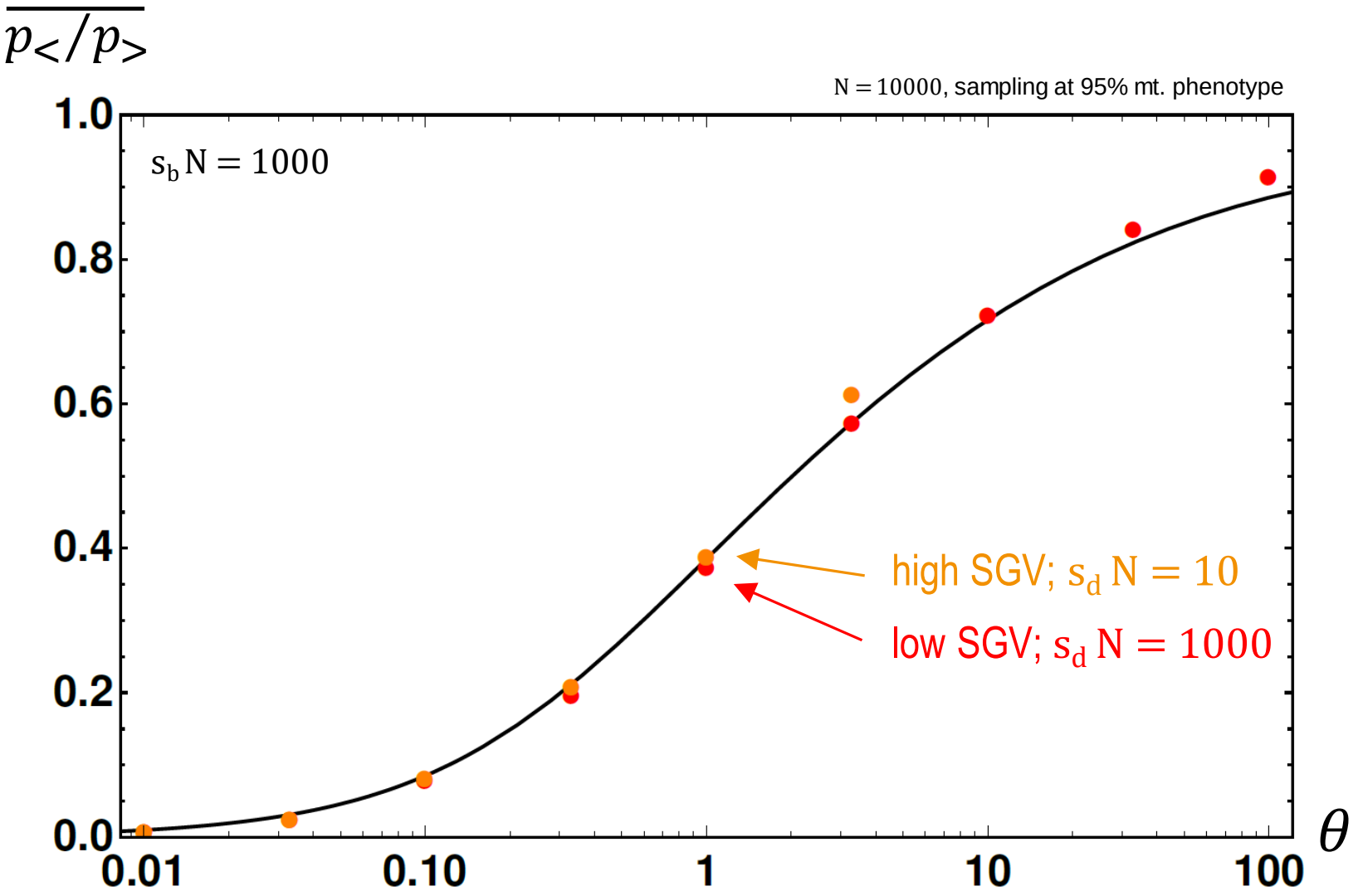
# Relative adaptive response (2 loci)



# Relative adaptive response (2 loci)

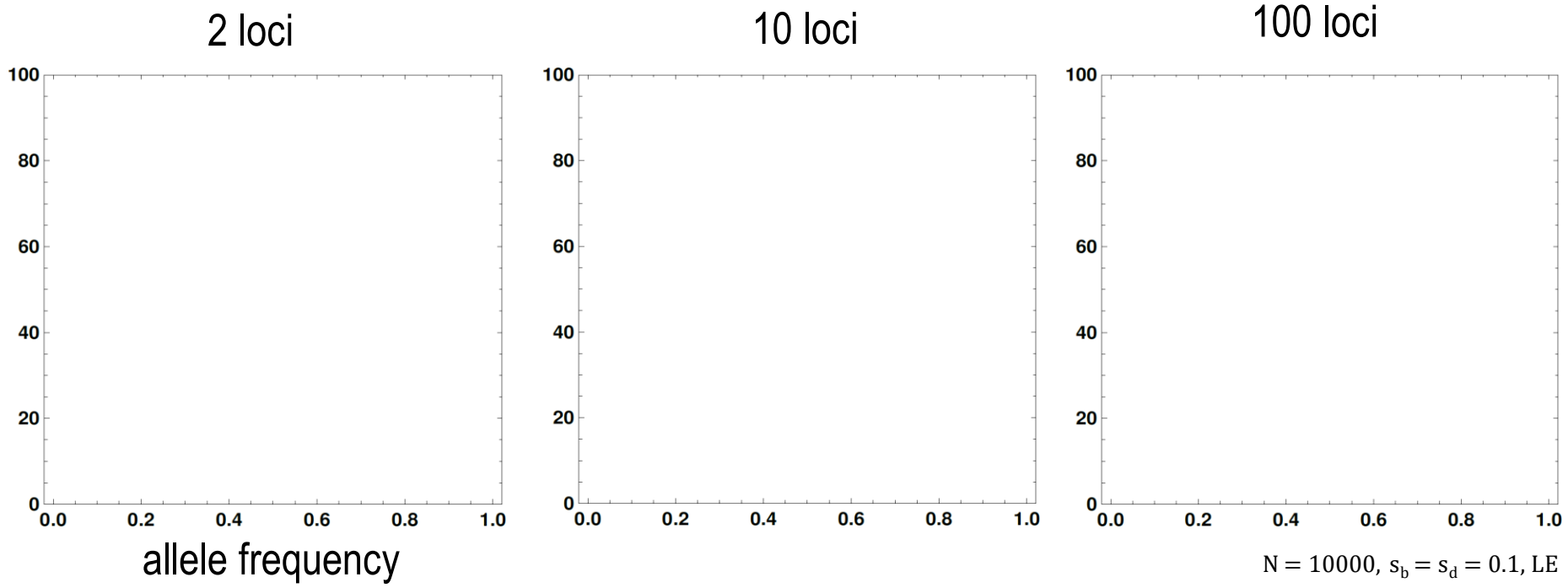
How about adaptation
from standing genetic variation?

# Relative adaptive response (2 loci)



# Architecture of polygenic adaptation

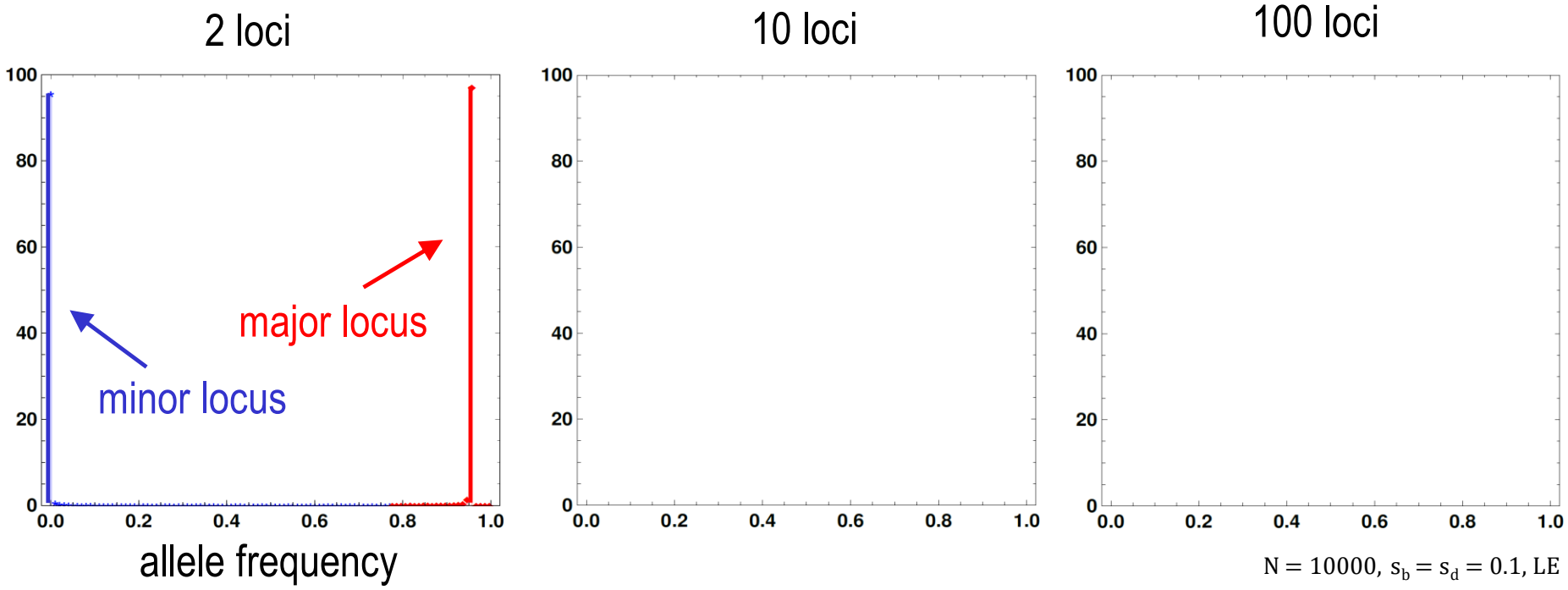
## allele frequencies with complete redundancy



# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$\theta = 0.01$

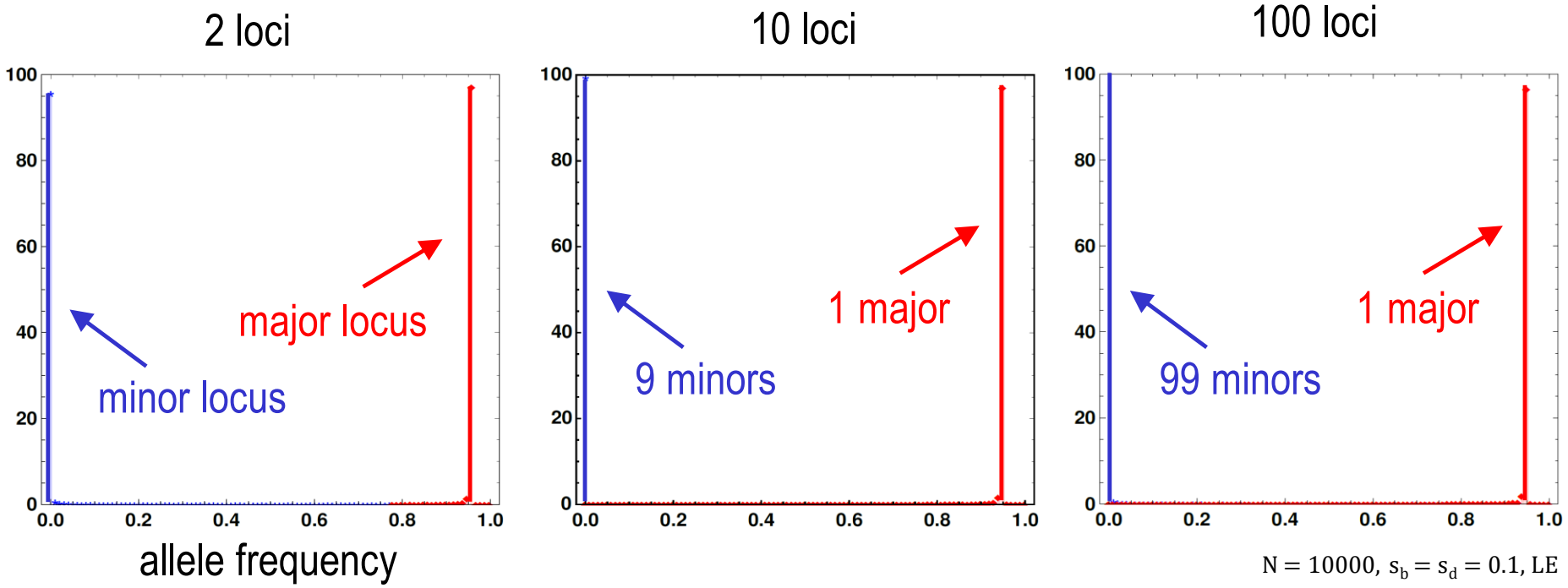



sweep @ single major locus

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$\theta = 0.01 \longrightarrow \theta_{\mathrm{bg}} = (L-1)\theta = 0.01$ (≈ "genome-wide $\theta$")

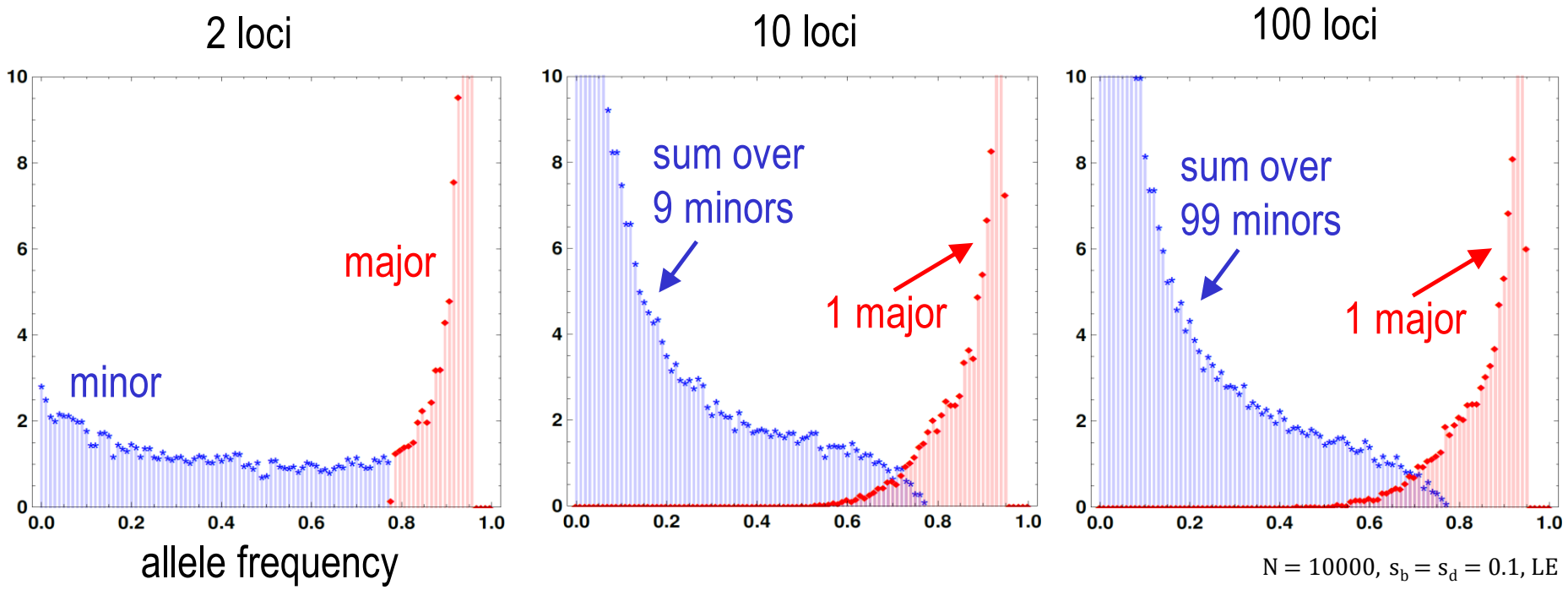



sweep @ single major locus

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$$\theta_{\mathrm{bg}} = (L-1)\theta = 1$$

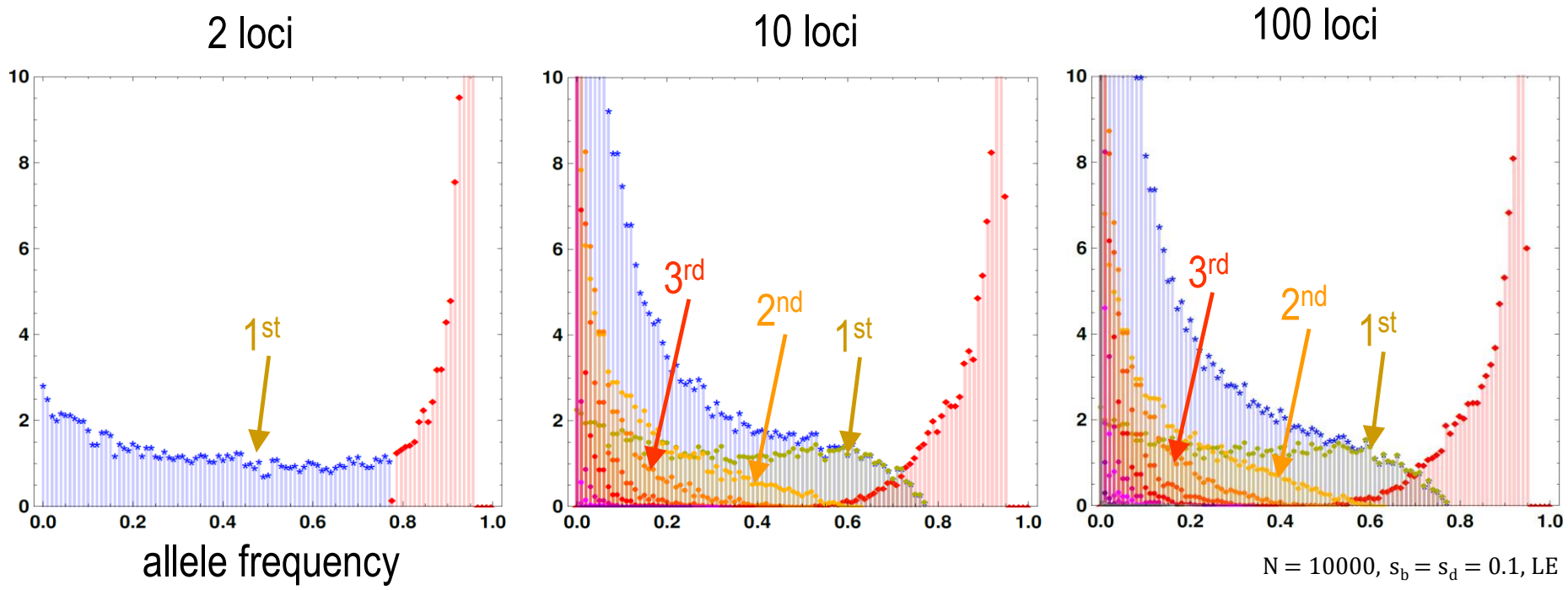



partial sweeps & strong major/minor structure

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$$\theta_{\mathrm{bg}} = (L-1)\theta = 1$$

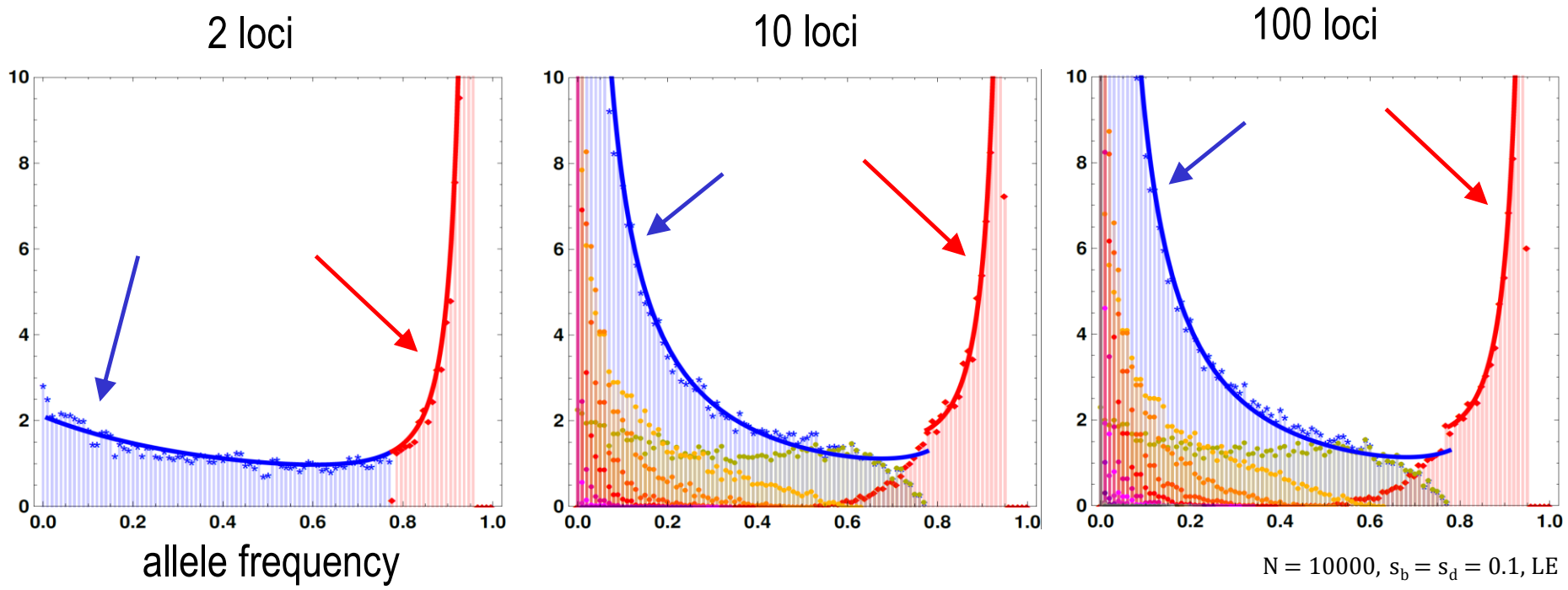



partial sweeps & strong major/minor structure

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$\theta_{\mathrm{bg}} = 1$, Yule approximation

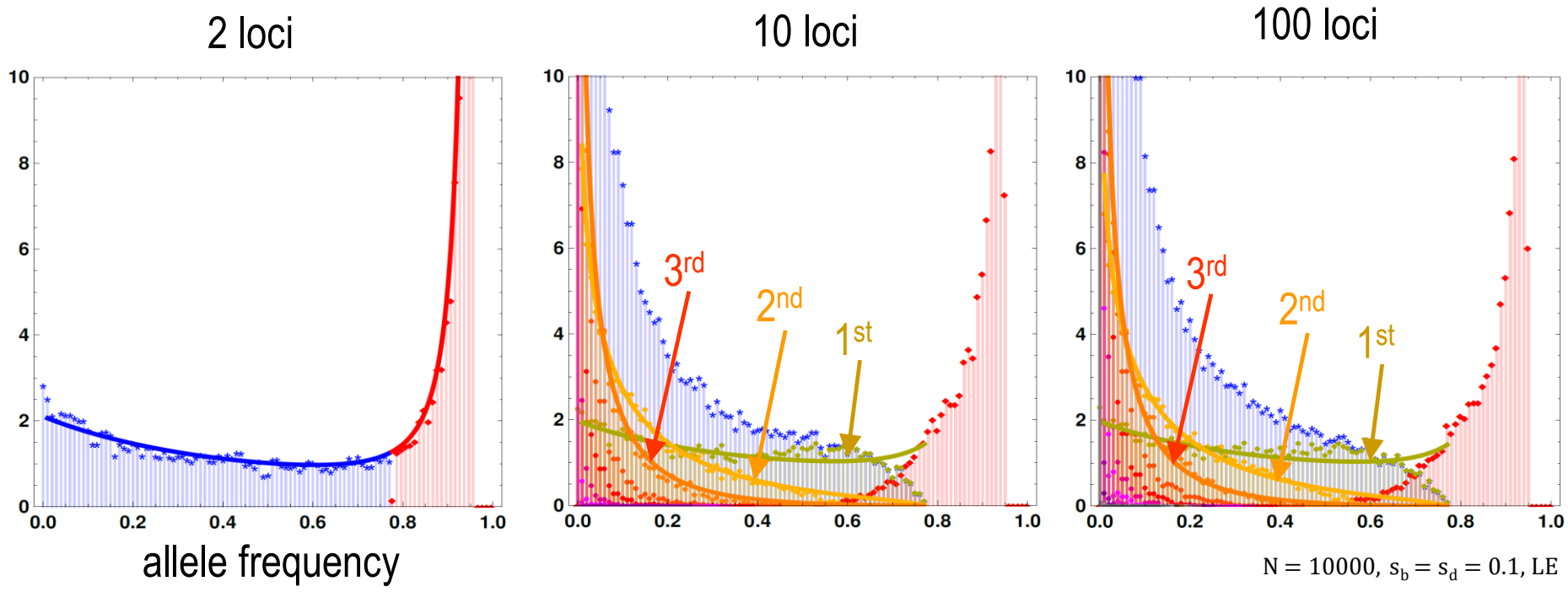



partial sweeps & strong major/minor structure

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$\theta_{\mathrm{bg}} = 1$, Yule approximation

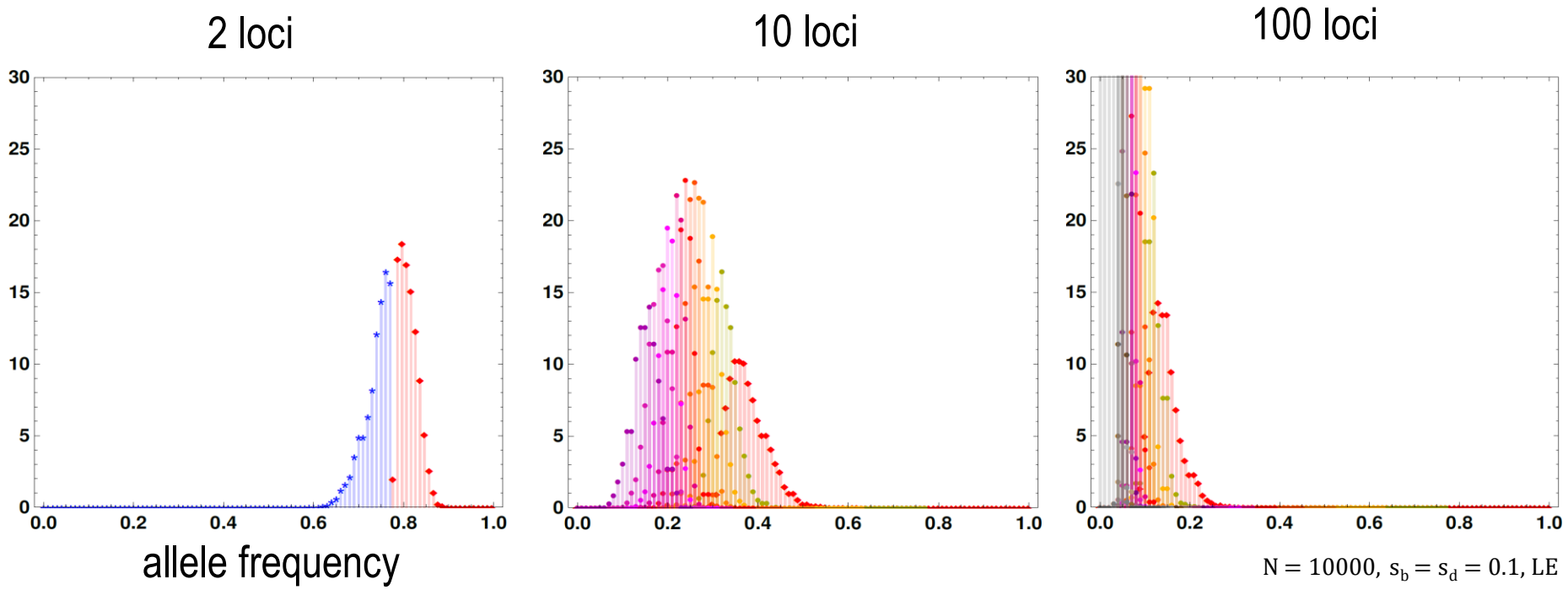



partial sweeps & strong major/minor structure

# Architecture of polygenic adaptation

## allele frequencies with complete redundancy

$$\theta_{\mathrm{bg}} = (L-1)\theta = 100$$




shifts @ many loci

# Results: quantitative trait

- equal loci, $\theta_i = \theta$
- start from mutation-selection-drift balance with some initial trait optimum $Z_0$
- adaptation from SGV and new mutations until mean phenotype reaches a threshold value, $\bar{Z} = Z_1 < Z_{\mathrm{opt}}$
- loci ordered according to their contribution to the adaptive response

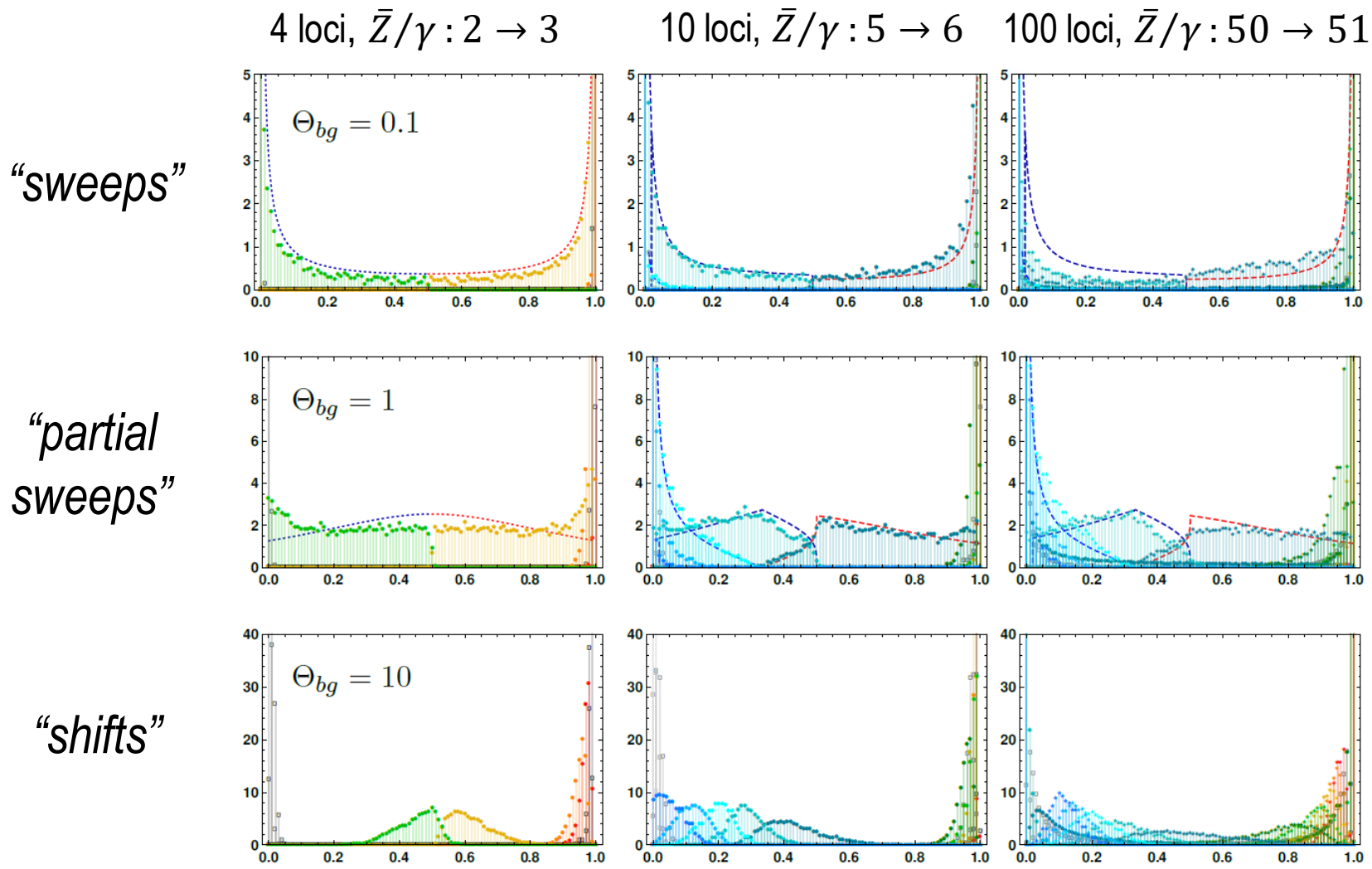
# Adaptive architecture

## additive trait: single adaptive step

4 loci, $\bar{Z}/\gamma : 2 \rightarrow 3$ 10 loci, $\bar{Z}/\gamma : 5 \rightarrow 6$ 100 loci, $\bar{Z}/\gamma : 50 \rightarrow 51$

# Adaptive architecture

## additive trait: single adaptive step



# Adaptive architecture

## additive trait: adaptation over larger distances

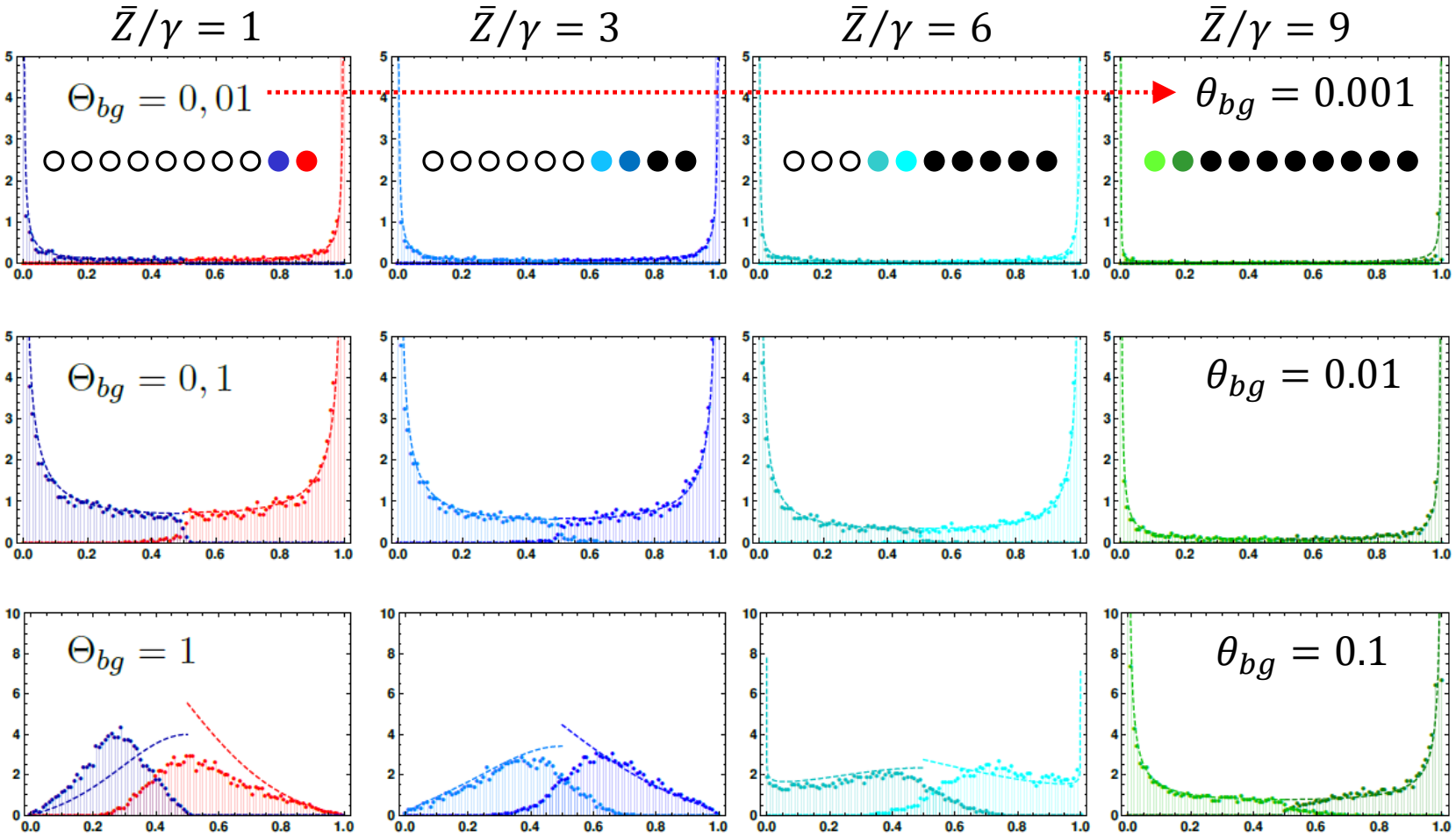
10 loci, $Z_{\mathrm{opt}}/\gamma: 0 \rightarrow 10$

$\bar{Z}/\gamma = 1$ $\bar{Z}/\gamma = 3$ $\bar{Z}/\gamma = 6$ $\bar{Z}/\gamma = 9$

# Adaptive architecture

## additive trait: adaptation over larger distances

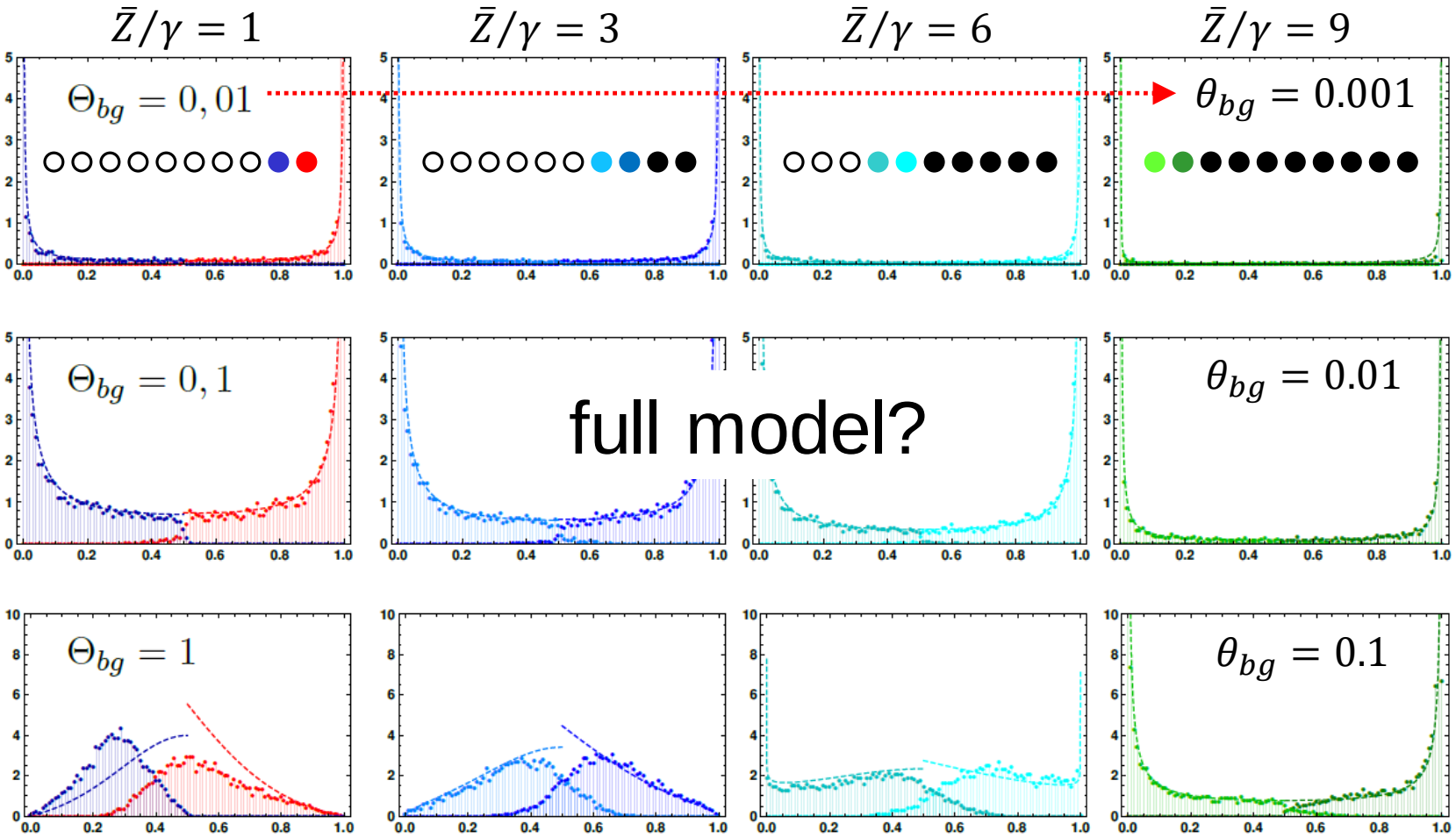
10 loci, $Z_{\mathrm{opt}}/\gamma: 0 \rightarrow 10$, effective $\theta_{bg} = (L - \bar{Z}/\gamma)\theta$ : decreasing redundancy



# Adaptive architecture

## additive trait: adaptation over larger distances

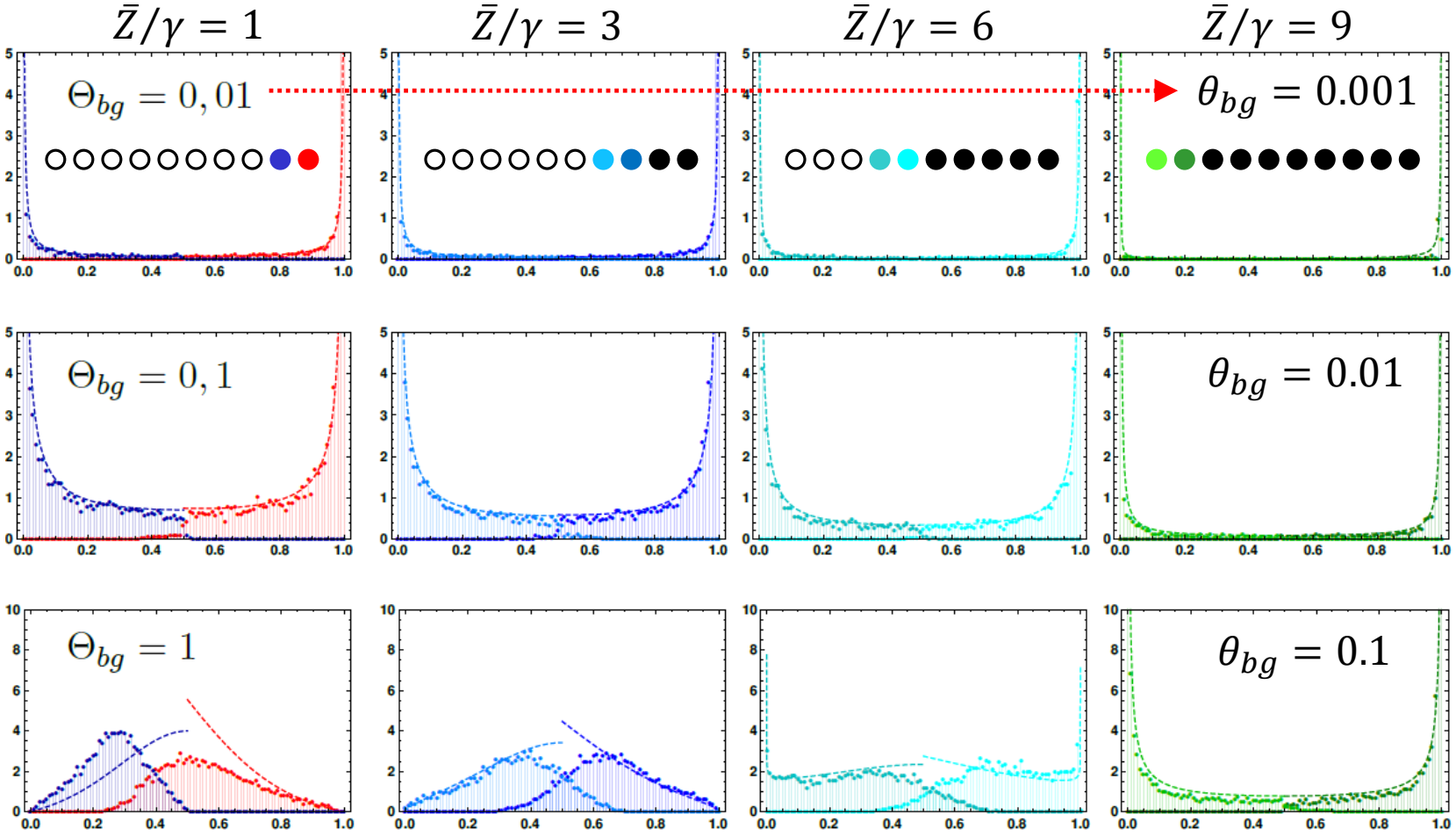
10 loci, $Z_{\mathrm{opt}}/\gamma: 0 \rightarrow 10$, effective $\theta_{bg} = (L - \bar{Z}/\gamma)\theta$ : decreasing redundancy



# Adaptive architecture

## additive trait: adaptation over larger distances

10 loci, $Z_{\mathrm{opt}}/\gamma: 0 \rightarrow 10$, effective $\theta_{bg} = (L - \bar{Z}/\gamma)\theta$ : decreasing redundancy



# Summary: limitations & generalizations

Analytical approximation for joint distribution of mutant frequencies

… at loci with equal effect on the trait assuming linkage equilibrium

# Summary: limitations & generalizations

Analytical approximation for joint distribution of mutant frequencies

… at loci with equal effect on the trait assuming linkage equilibrium

- locus mutation rates can differ
- selection strength / trait optimum can depend on time
- background variation due to loci of arbitrary effect

$$\bar{Z} = \underbrace{\sum_{i=1}^{L} \gamma p_i}_{\text{focal QTL}} + \underbrace{Z_B(t)}_{\text{background}} \qquad W(Z) = \exp\left[-\frac{\sigma(t)}{2}(Z - Z_{\text{opt}}(t))^2\right]$$

# Summary: limitations & generalizations

Analytical approximation for joint distribution of mutant frequencies

… at loci with equal effect on the trait assuming linkage equilibrium

- locus mutation rates can differ
- selection strength / trait optimum can depend on time
- background variation due to loci of arbitrary effect

$$\bar{Z} = \underbrace{\sum_{i=1}^{L} \gamma p_i}_{\text{focal QTL}} + \underbrace{Z_B(t)}_{\text{background}} \qquad W(Z) = \exp\left[-\frac{\sigma(t)}{2}(Z - Z_{\text{opt}}(t))^2\right]$$

Adaptive architecture as function of mean phenotype

- independent of selection strength $\sigma, \gamma, Z_{\text{opt}}$ (resp. $s_b, s_d$)
- only depends on the mutation rate $\theta_{\text{bg}} = (L-1)\theta$

# Three scenarios of polygenic adaptation

## for adaptation from mutation-selection-drift balance

$\theta_{\mathrm{bg}}$: "effective redundancy", measures competition due to potential adaptation at alternative loci

# Three scenarios of polygenic adaptation

## for adaptation from mutation-selection-drift balance

$\theta_{\mathrm{bg}}$: "effective redundancy", measures competition due to potential adaptation at alternative loci

- $\theta_{\mathrm{bg}} \lesssim 0.1$ : sweep @ single major locus
  - usually hard sweep from new mutation
- $0.1 \lesssim \theta_{\mathrm{bg}} \lesssim 10$ : major-minor locus pattern of adaptation
  - (almost) completed sweep from SGV at major locus
  - partial (hard or soft) sweeps at several minor loci
- $\theta_{\mathrm{bg}} \gtrsim 10$ : small frequency shifts @ many loci
  - no clear selection footprint in linked variation

# Scenarios of polygenic adaptation
## by small frequency shifts

How should we explain evidence for adaptation by small shifts (size / weight / yield traits)?

# Scenarios of polygenic adaptation
## by small frequency shifts

How should we explain evidence for adaptation by small shifts (size / weight / yield traits)?

- really “small shifts”?
  - “slow sweeps” at small-effect loci ?

# Scenarios of polygenic adaptation by small frequency shifts

How should we explain evidence for adaptation by small shifts (size / weight / yield traits)?

- really "small shifts"?
  - "slow sweeps" at small-effect loci ?
- $\theta_{\mathrm{bg}} \approx 2LN_e u$ is large:
  - large "omnigenic" basis $L > 10000$ ?
  - large "short-term $N_e$" ?

# Scenarios of polygenic adaptation by small frequency shifts

How should we explain evidence for adaptation by small shifts (size / weight / yield traits)?

- really "small shifts"?
  - "slow sweeps" at small-effect loci ?
- $\theta_{\mathrm{bg}} \approx 2LN_e u$ is large:
  - large "omnigenic" basis $L > 10000$ ?
  - large "short-term $N_e$" ?
- initial allele frequencies more homogeneous than predicted by mutation-selection-drift balance
  - balancing selection ? (but implies constraint)
  - spatial structure or admixture ?

# More than *sweeps* or *shifts*

Pattern of stalled partial sweeps:

- predicted in large and relevant parameter region
- strongly heterogeneous even among loci with identical effect
- should also be heterogeneous among replicates / for parallel adaptation (“zoomed-up stochasticity”)

# More than *sweeps* or *shifts*

Pattern of stalled partial sweeps:

- predicted in large and relevant parameter region
- strongly heterogeneous even among loci with identical effect
- should also be heterogeneous among replicates / for parallel adaptation ("zoomed-up stochasticity")

Evidence ?

- a lot of evidence for partial sweeps
  - ➢ strong completed sweeps are rare
- plateauing of allele trajectories in experimental evolution
  - ➢ strongly heterogeneous among replicates

# Thanks!

Ilse Hölliger

Pleuni Pennings

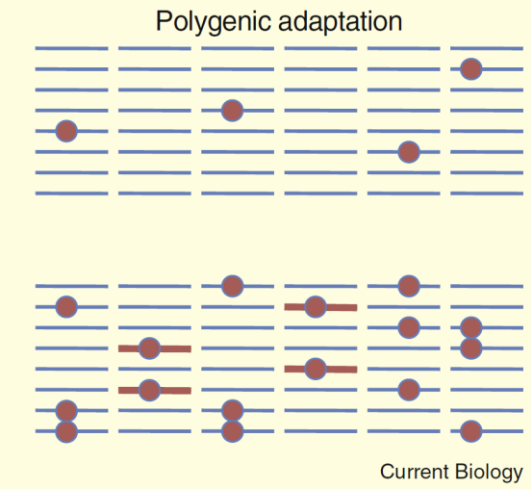
POPGEN VIENNA

VIENNA GRADUATE SCHOOL OF POPULATION GENETICS

# Relative adaptive response (2 loci)

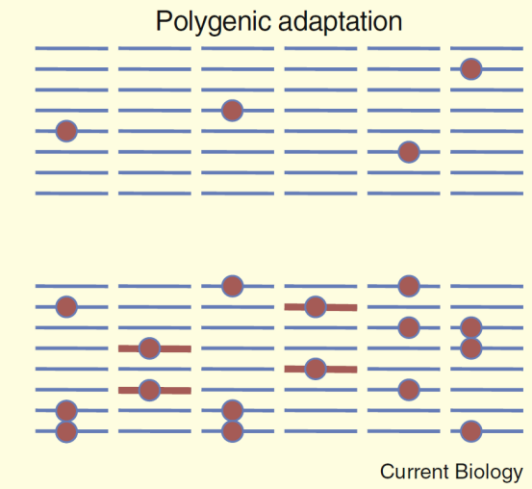
*Equal levels of SGV* $N\frac{u}{s_d} = \frac{\theta}{2s_d}$ *at many loci …*

*… should lead to equal proportions* $n_i/n_j$ *after adaptation*

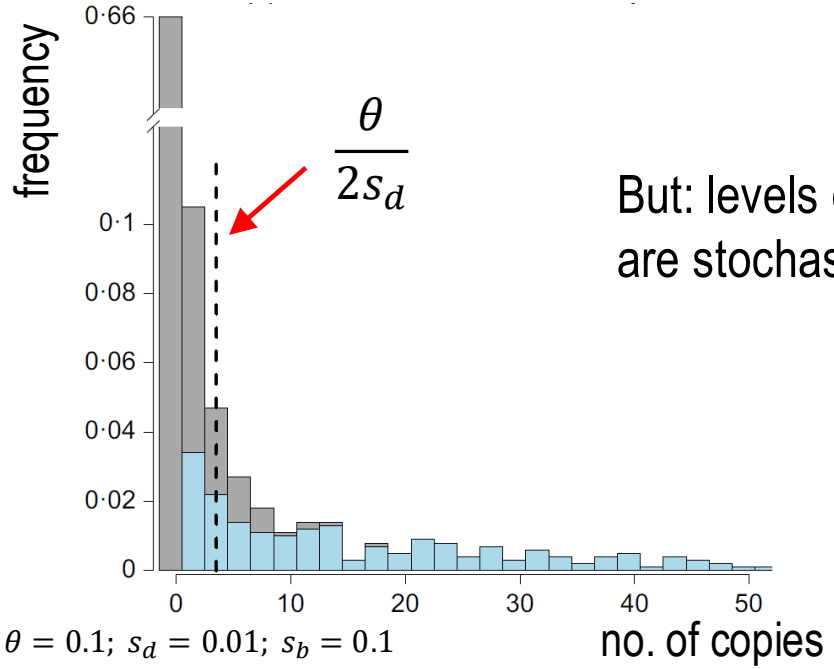


# Relative adaptive response (2 loci)

*Equal levels of SGV* $N\frac{u}{s_d} = \frac{\theta}{2s_d}$ *at many loci …*




*… should lead to equal proportions* $n_i/n_j$ *after adaptation*

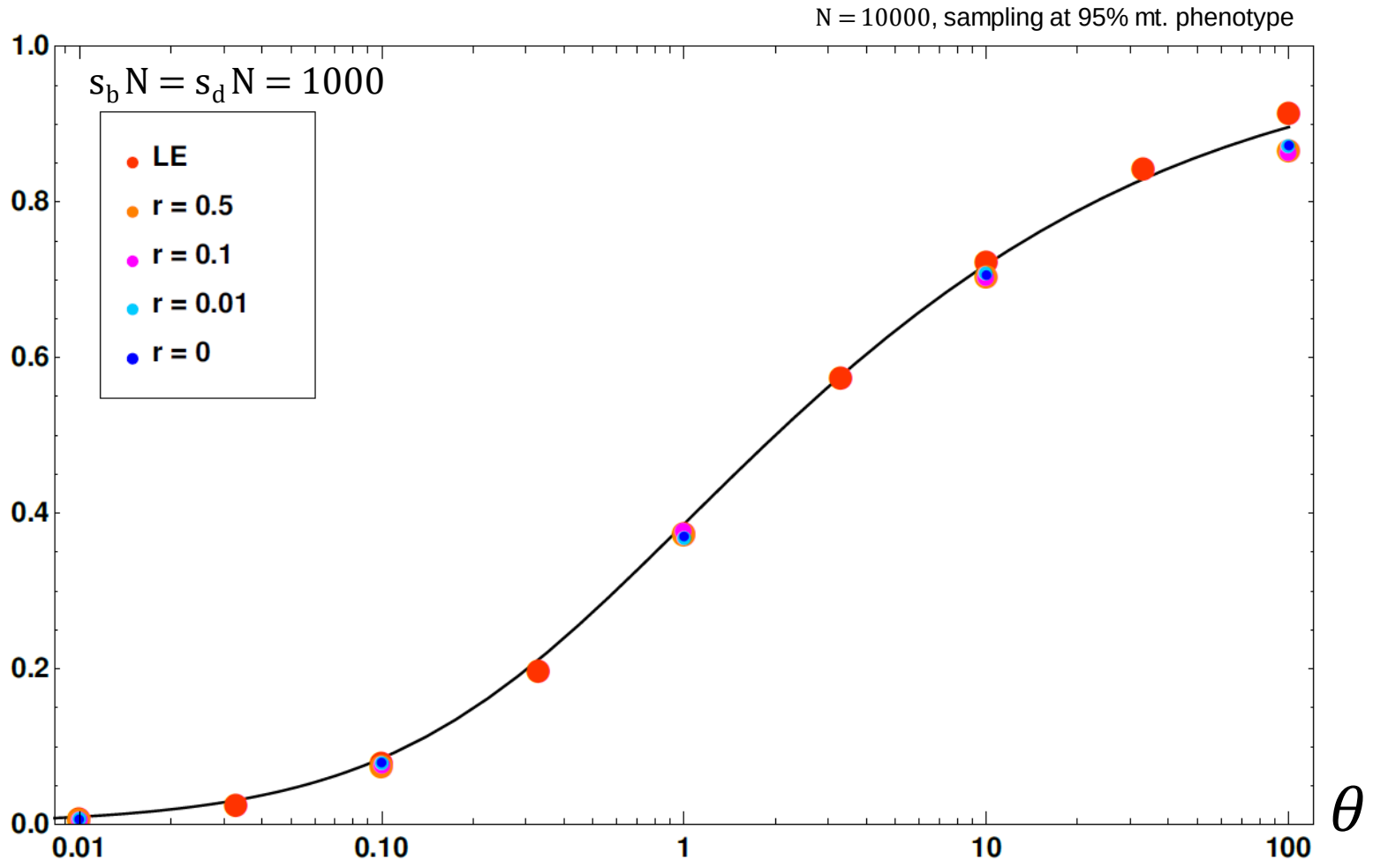



But: levels of SGV in mutation-selection-drift balance are stochastic and have a large variance

- Same major-minor locus structure as for new mutation

# Relative adaptive response (2 loci)

How about
linkage disequilibrium ?

# Relative adaptive response (2 loci)

$\overline{p_</p_>}$