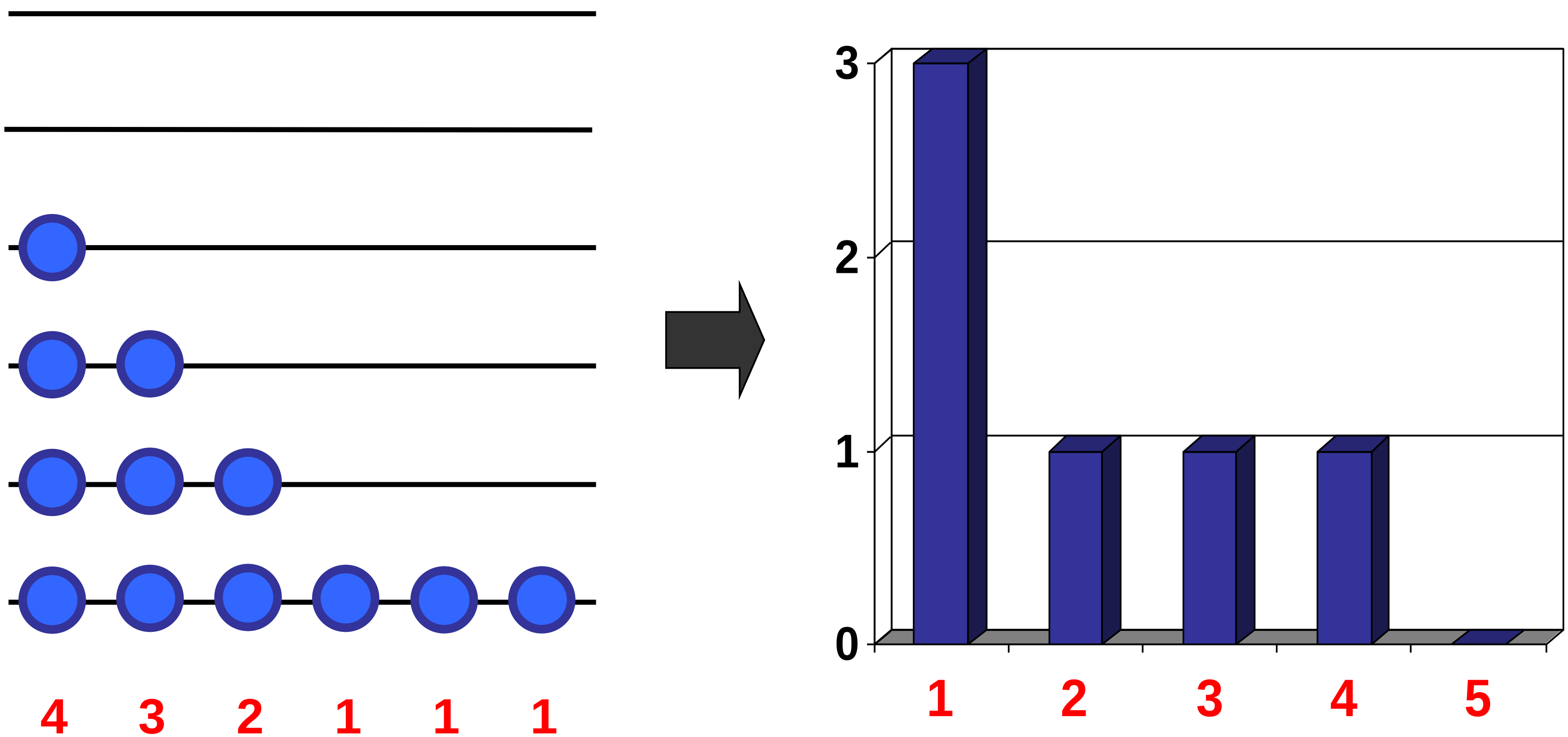
Patterns of Evolution

Summary statistics based on segregating sites

Site Frequency Spectrum



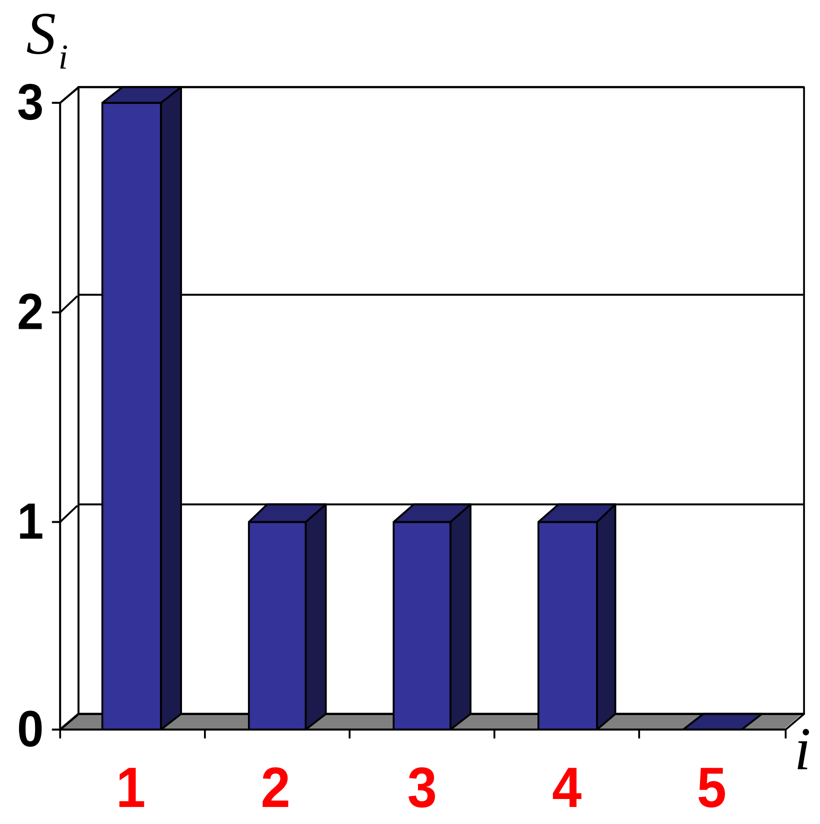
Patterns of Evolution
Summary statistics based on segregating sites

Site Frequency Spectrum

S_i : number of mutants that appear in i copies in the sample

$$S = \sum_{i=1}^{n-1} S_i :$$ total number of segregating sites in an sample of size n

$$\pi = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} i(n-i) S_i :$$ average number of pairwise differences



Patterns of Evolution

Summary statistics based on segregating sites

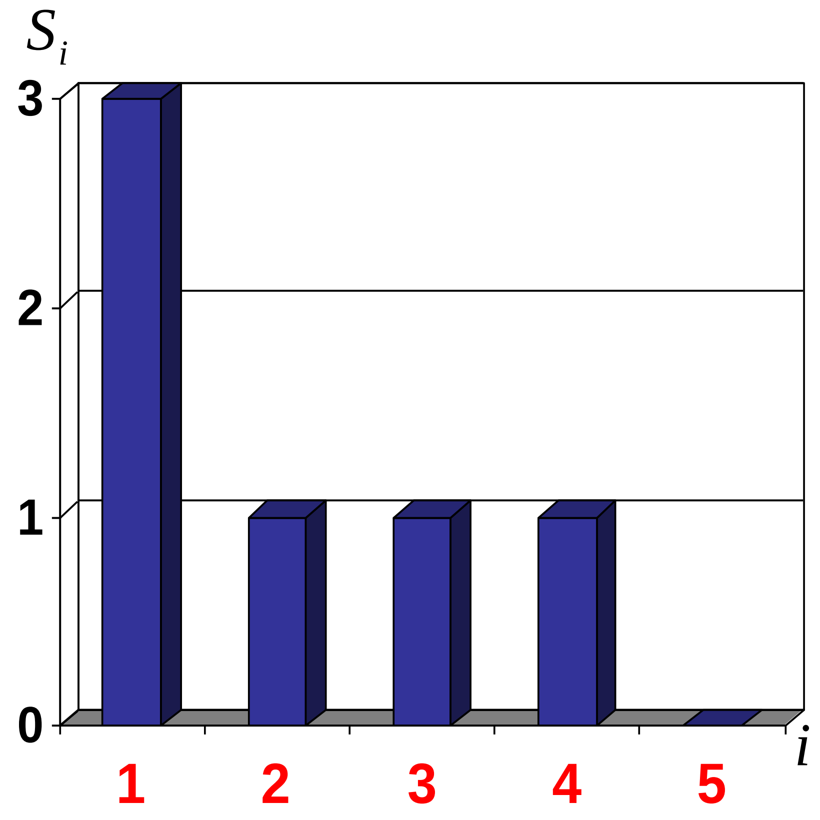
Site Frequency Spectrum

S_i : number of mutants that appear in i copies in the sample

$S = \sum_{i=1}^{n-1} S_i$: total number of segregating sites in an sample of size n

$\pi = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} i(n-i)S_i$:

Each mutation of size i contributes to divergence in $i\ (n-i)$ sequence pairs



Coalescent Theory
Estimators

Unbiased estimators of the mutation parameter $\theta = 4Nu$:

Watterson‘s estimator: $$\hat{\theta}_W = \frac{S}{a_n} = \sum_{i=1}^{n-1} S_i \Big/ \sum_{i=1}^{n-1} \frac{1}{i}$$ (equal weights)

π-based estimator: $$\hat{\theta}_\pi = \pi = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} i(n-i)\, S_i$$ (intermediate frequencies)

Fay and Wu‘s estimator: $$\hat{\theta}_H = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} i^2 S_i$$ (high frequencies)

singleton estimator: $$\hat{\theta}_s = \frac{n-1}{n} \underbrace{\left(S_1 + S_{n-1}\right)}_{\text{singletons of the folded spectrum}}$$ (extreme frequencies)

singletons of the folded spectrum

Coalescent Theory
Test statistics

Test statistics for the deviation from neutrality:

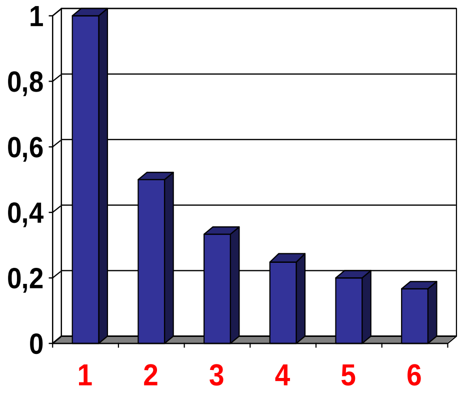
Tajima's D:

$$D_T = \frac{\hat{\theta}_\pi - \hat{\theta}_W}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_W\right]}}$$

$$D_T = 0$$

Fu and Li's D:

$$D_{FL} = \frac{\hat{\theta}_W - \hat{\theta}_S}{\sqrt{\mathrm{Var}\left[\hat{\theta}_W - \hat{\theta}_S\right]}}$$



Fay and Wu's H:

$$H_{FW} = \frac{\hat{\theta}_\pi - \hat{\theta}_H}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_H\right]}}$$

Coalescent Theory
Test statistics

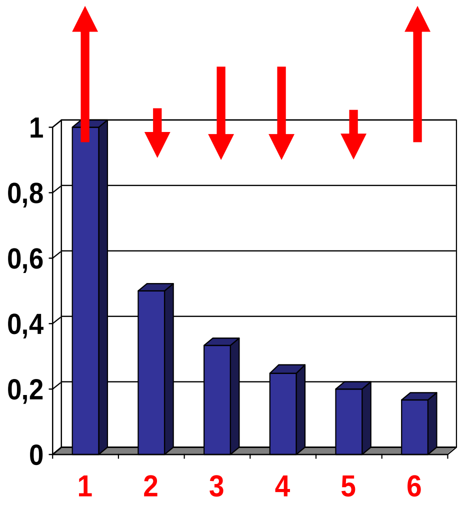
Test statistics for the deviation from neutrality:

Tajima's D: $$D_T = \frac{\hat{\theta}_\pi - \hat{\theta}_W}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_W\right]}}$$

Fu and Li's D: $$D_{FL} = \frac{\hat{\theta}_W - \hat{\theta}_S}{\sqrt{\mathrm{Var}\left[\hat{\theta}_W - \hat{\theta}_S\right]}}$$

Fay and Wu's H: $$H_{FW} = \frac{\hat{\theta}_\pi - \hat{\theta}_H}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_H\right]}}$$

$$D_T < 0$$



Coalescent Theory
Test statistics

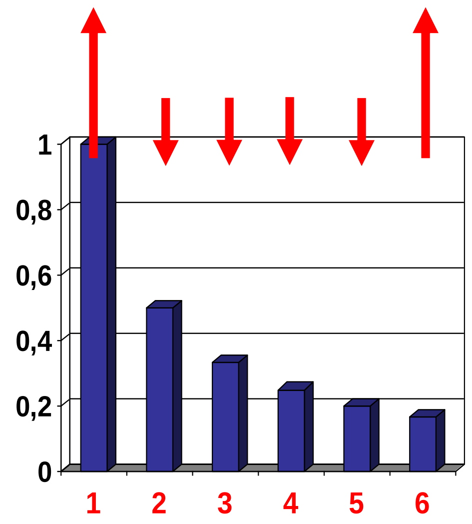
Test statistics for the deviation from neutrality:

Tajima‘s D: $$D_T = \frac{\hat{\theta}_\pi - \hat{\theta}_W}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_W\right]}}$$

Fu and Li‘s D: $$D_{FL} = \frac{\hat{\theta}_W - \hat{\theta}_S}{\sqrt{\mathrm{Var}\left[\hat{\theta}_W - \hat{\theta}_S\right]}}$$

Fay and Wu‘s H: $$H_{FW} = \frac{\hat{\theta}_\pi - \hat{\theta}_H}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_H\right]}}$$

$$D_{FW} < 0$$



Coalescent Theory
Test statistics

Test statistics for the deviation from neutrality:

Tajima's D: $$D_T = \frac{\hat{\theta}_\pi - \hat{\theta}_W}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_W\right]}}$$

Fu and Li's D: $$D_{FL} = \frac{\hat{\theta}_W - \hat{\theta}_S}{\sqrt{\mathrm{Var}\left[\hat{\theta}_W - \hat{\theta}_S\right]}}$$

Fay and Wu's H: $$H_{FW} = \frac{\hat{\theta}_\pi - \hat{\theta}_H}{\sqrt{\mathrm{Var}\left[\hat{\theta}_\pi - \hat{\theta}_H\right]}}$$

$$H_{FW} < 0$$