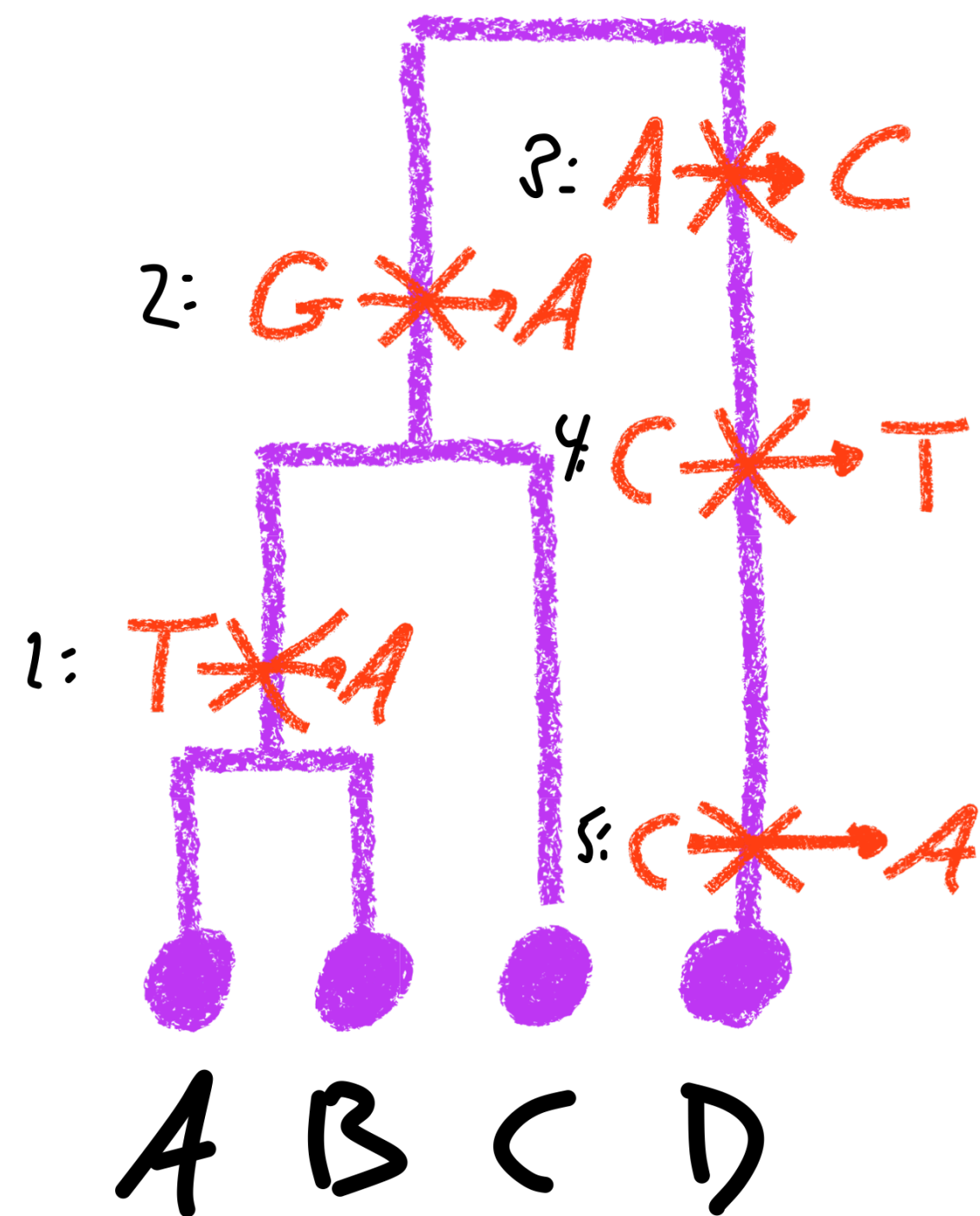


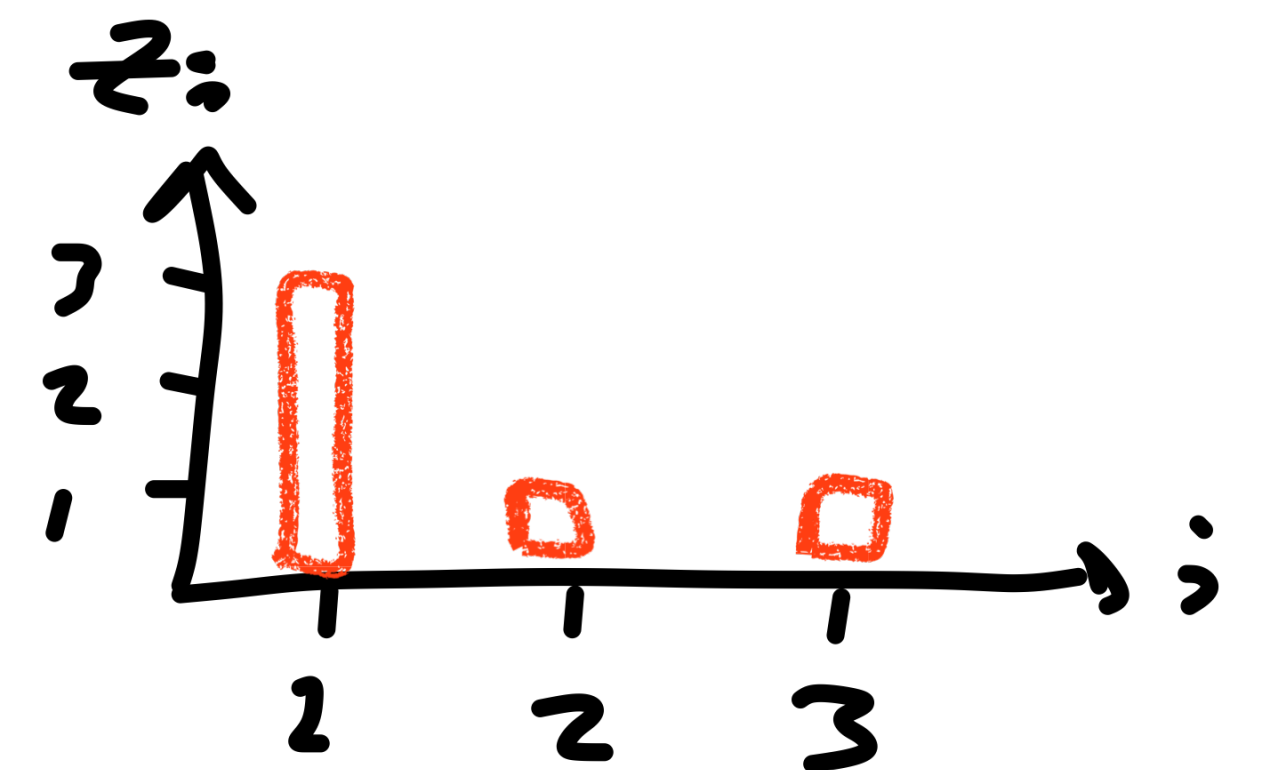
Example: sample frequency spectrum (SFS)



| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| A | A | A | A | C | C |
| B | A | A | A | C | C |
| C | T | A | A | C | C |
| D | T | G | C | T | A |

SFS: histogram of mutant allele frequencies

$$\bar{z} = [z_1, z_2, \dots, z_{n-1}], \quad z_i = \text{\# mutations with frequency } i \text{ in sample}$$



Theory: sample frequency spectrum (SFS)

SFS: histogram of mutant allele frequencies

$\vec{z} = [z_1, z_2, \dots, z_{n-2}]$, $z_i = \# \text{mutations with frequency } i \text{ in sample}$

With a bit of work...

$$E[z_i] = \frac{4Nm}{i}$$

(for constant N) $i=1, 2, \dots, n-1$

$$\log z_i = \log \left(\frac{4Nm}{i} \right)$$

$$= \log 4Nm - \log i$$