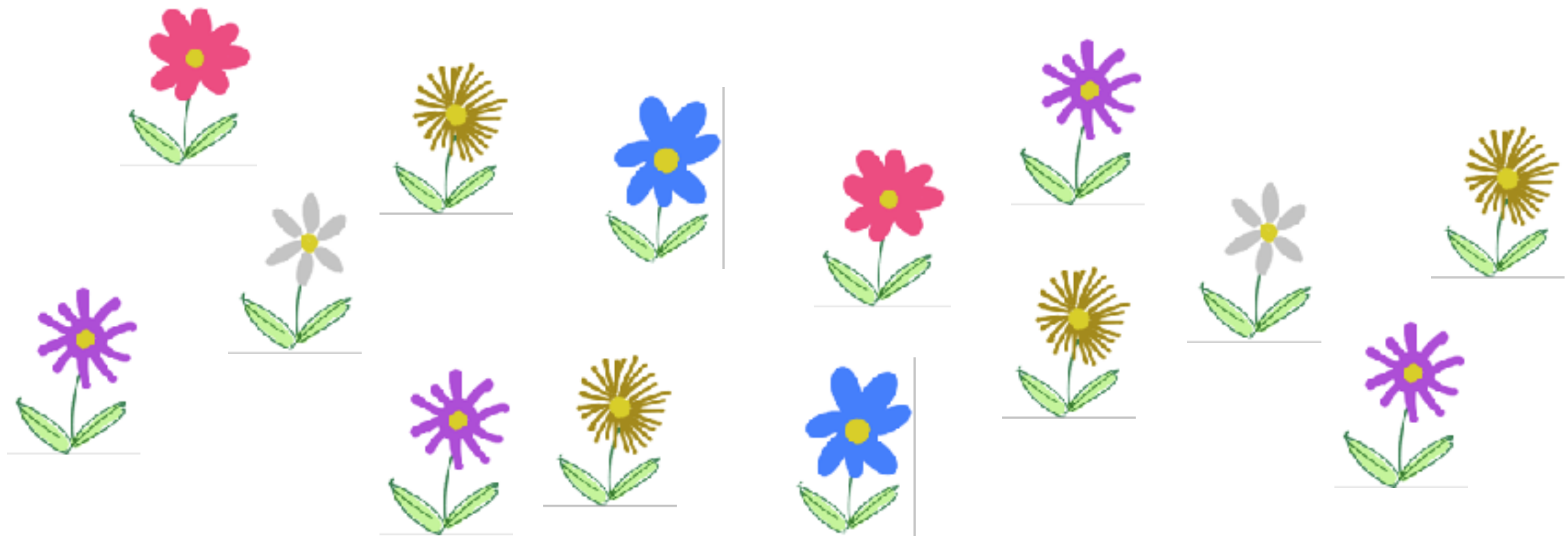




Practice of Population genetics



Lab. Crop Evolution

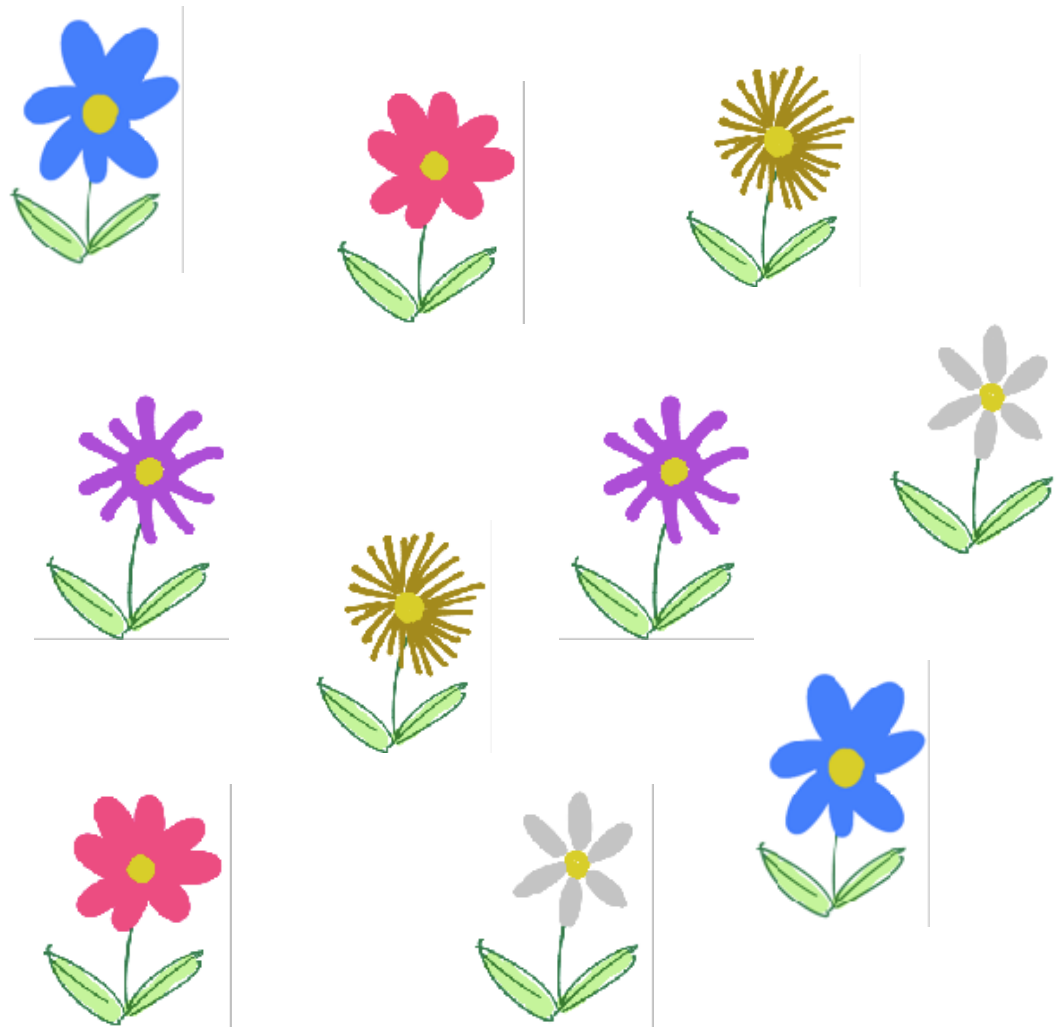


What is a diversity ?

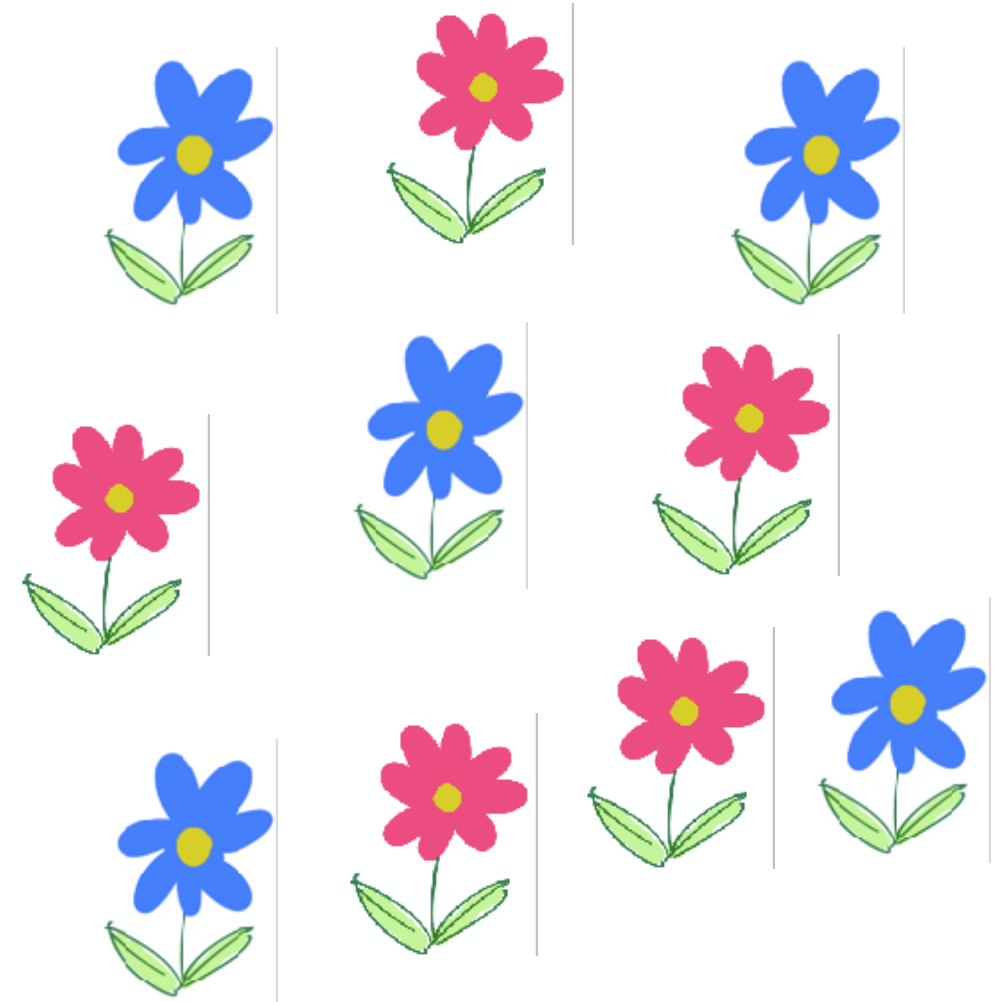


Which is more diverse ?

A

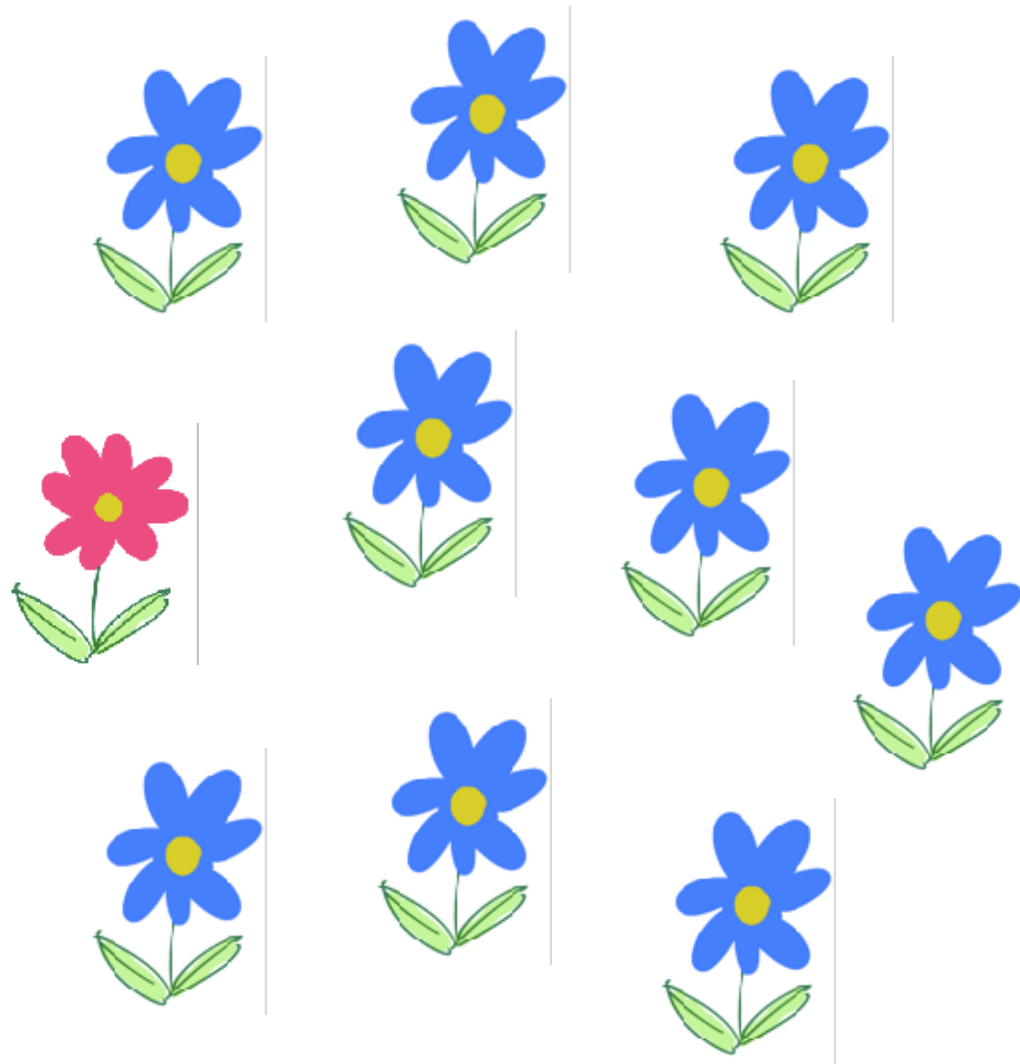


B

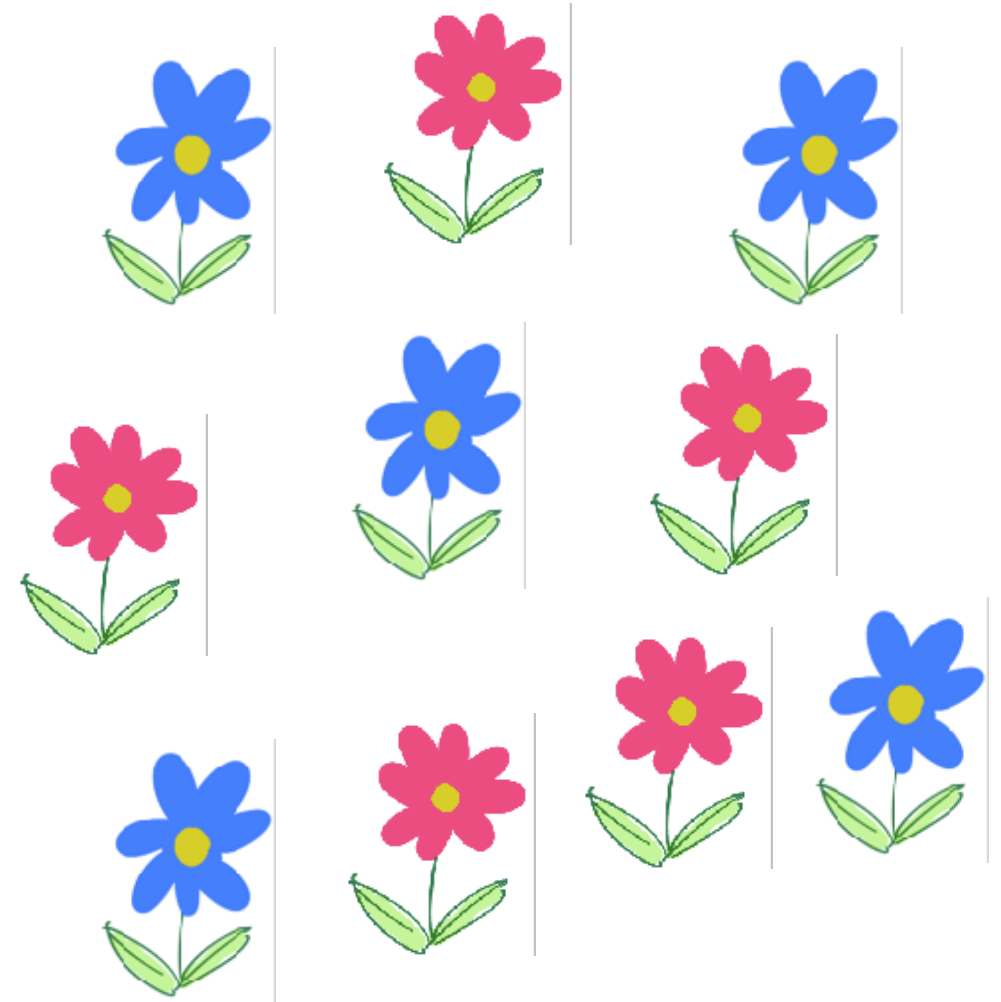


Which is more diverse ?

A

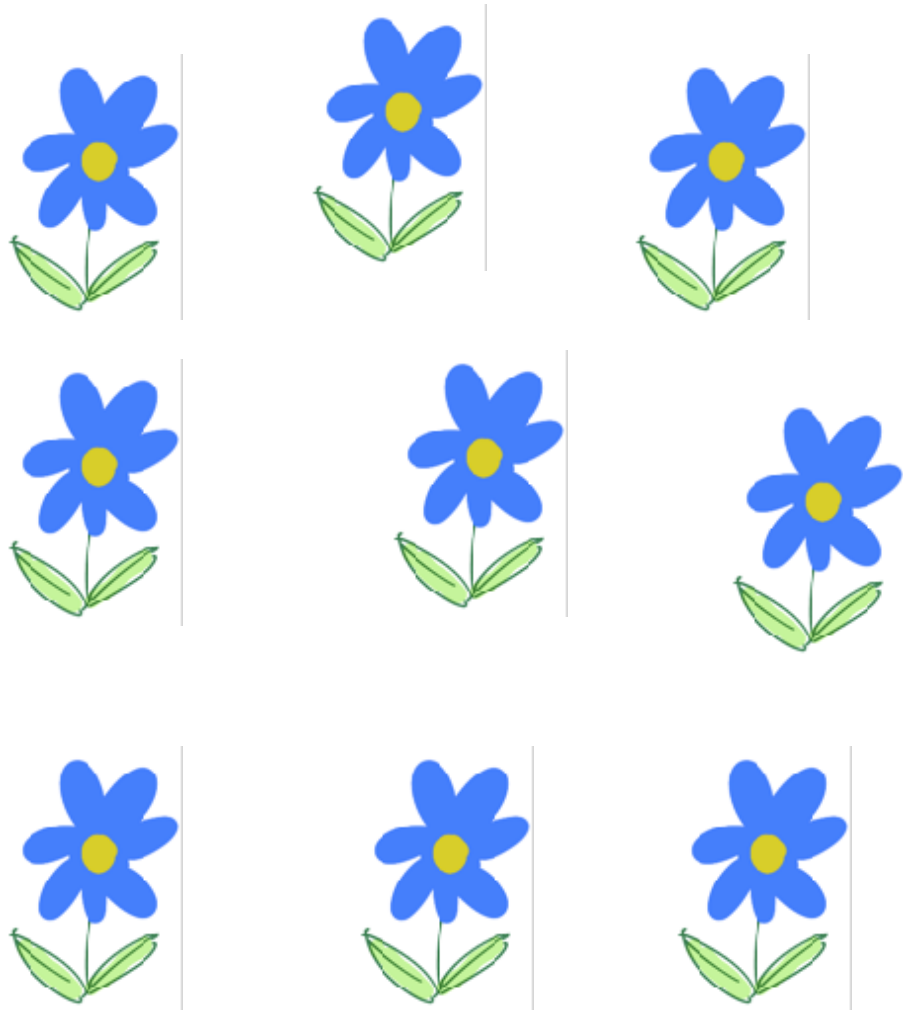


B



Which is more diverse ?

A

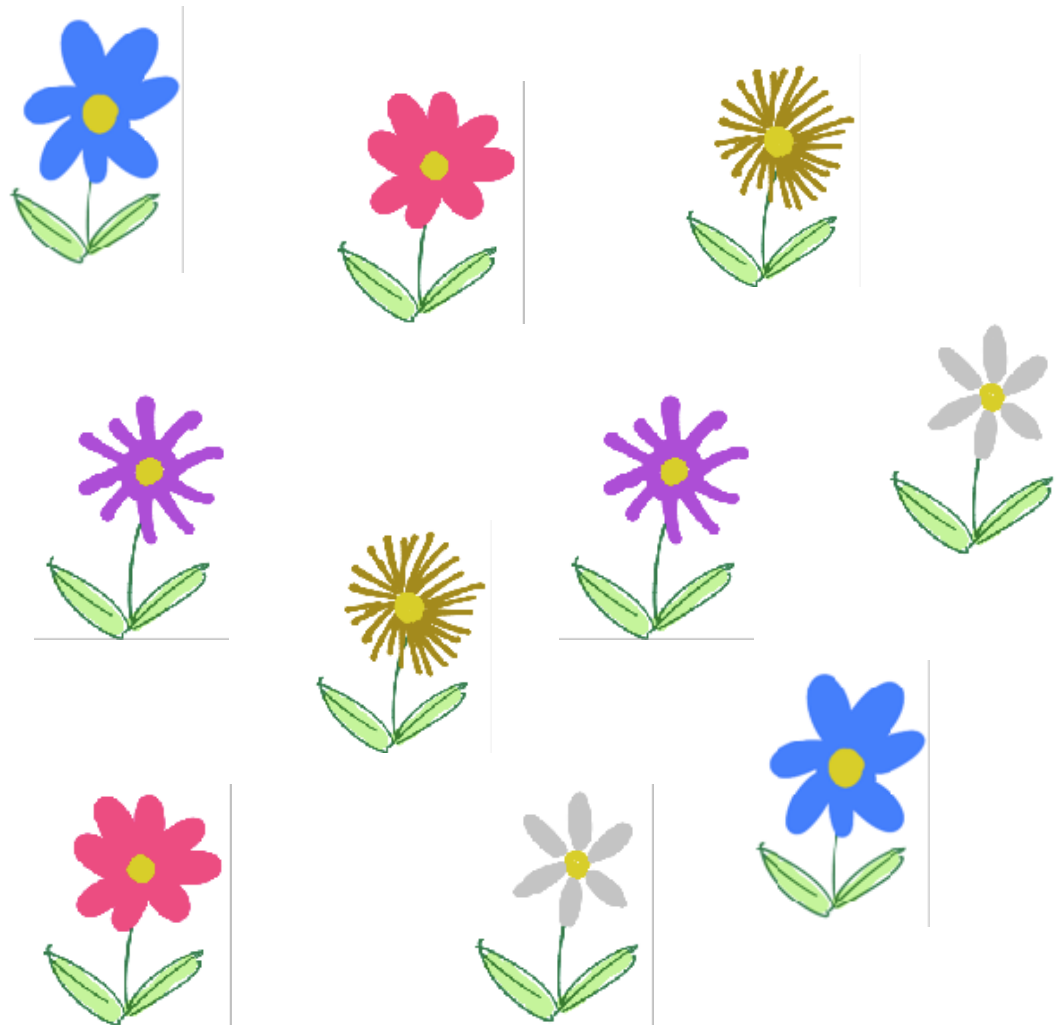


B

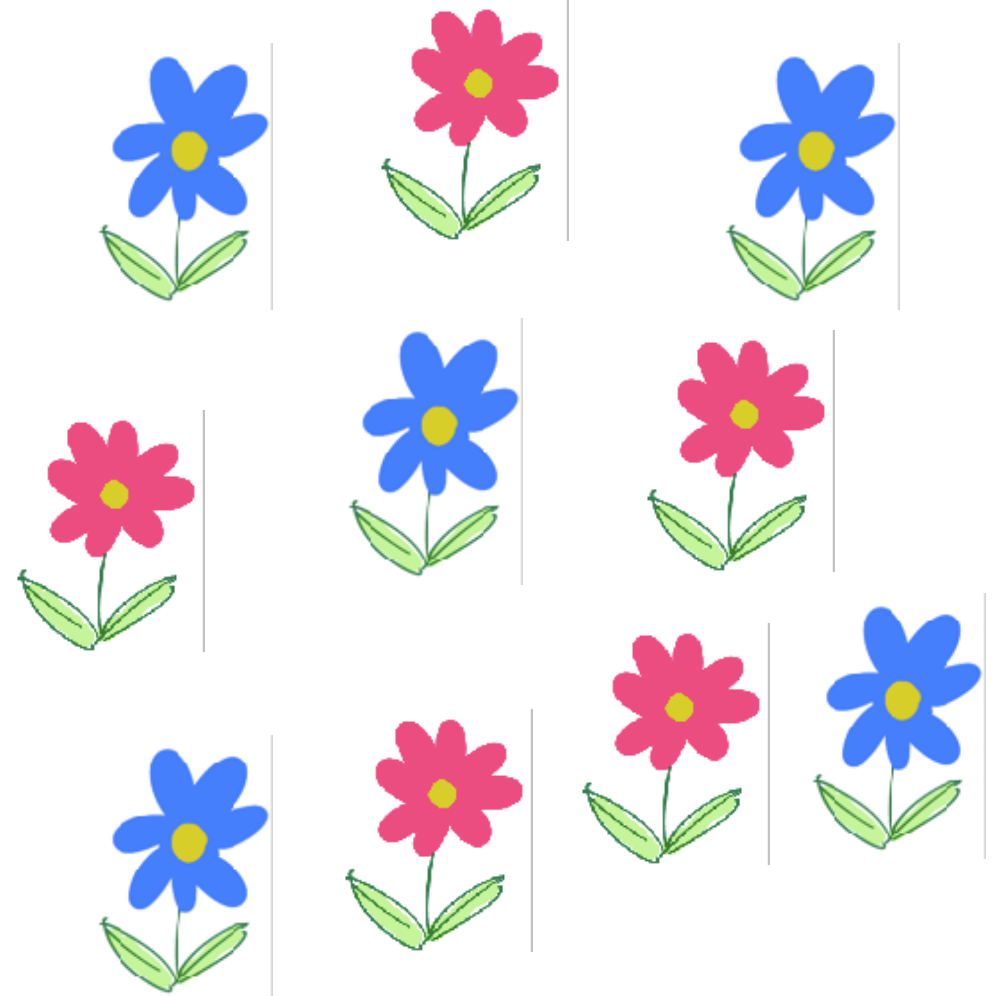


Which is more diverse ?

A



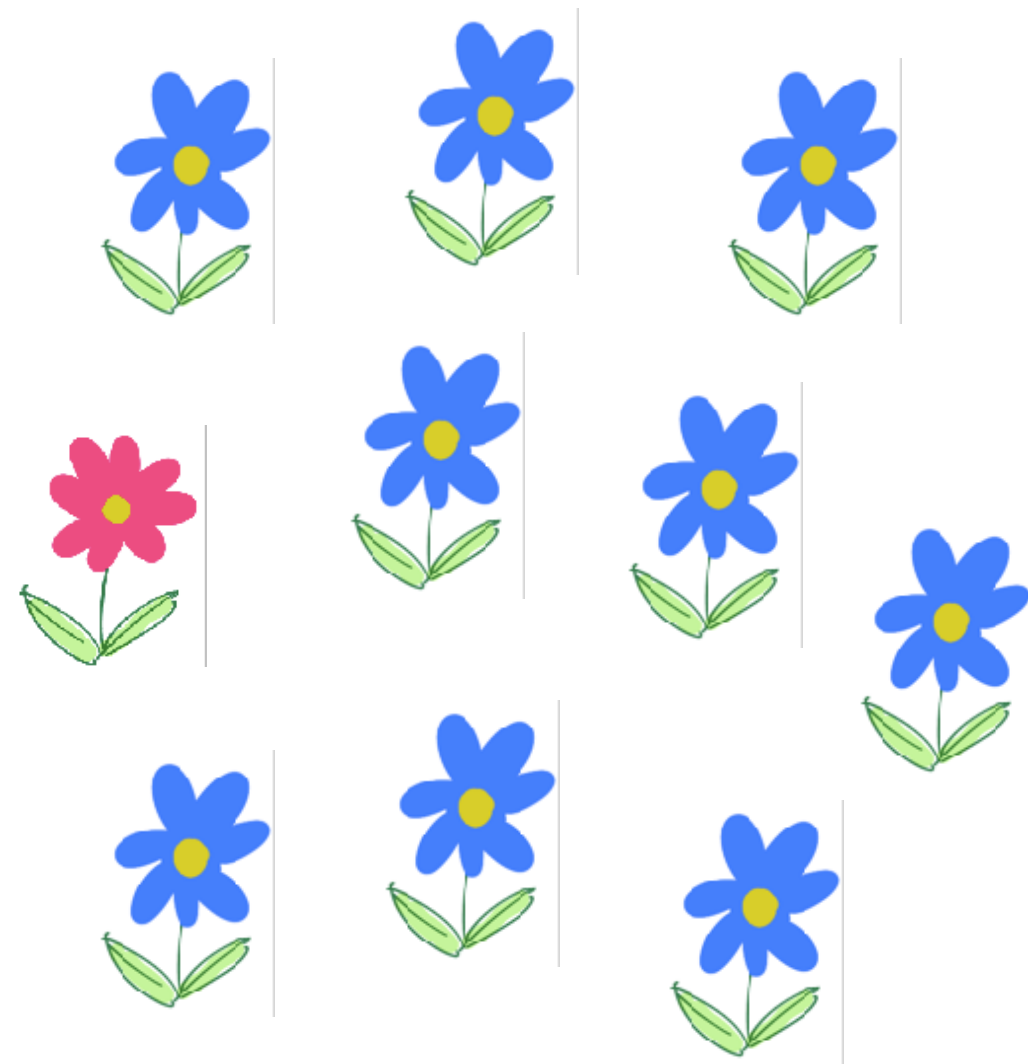
B



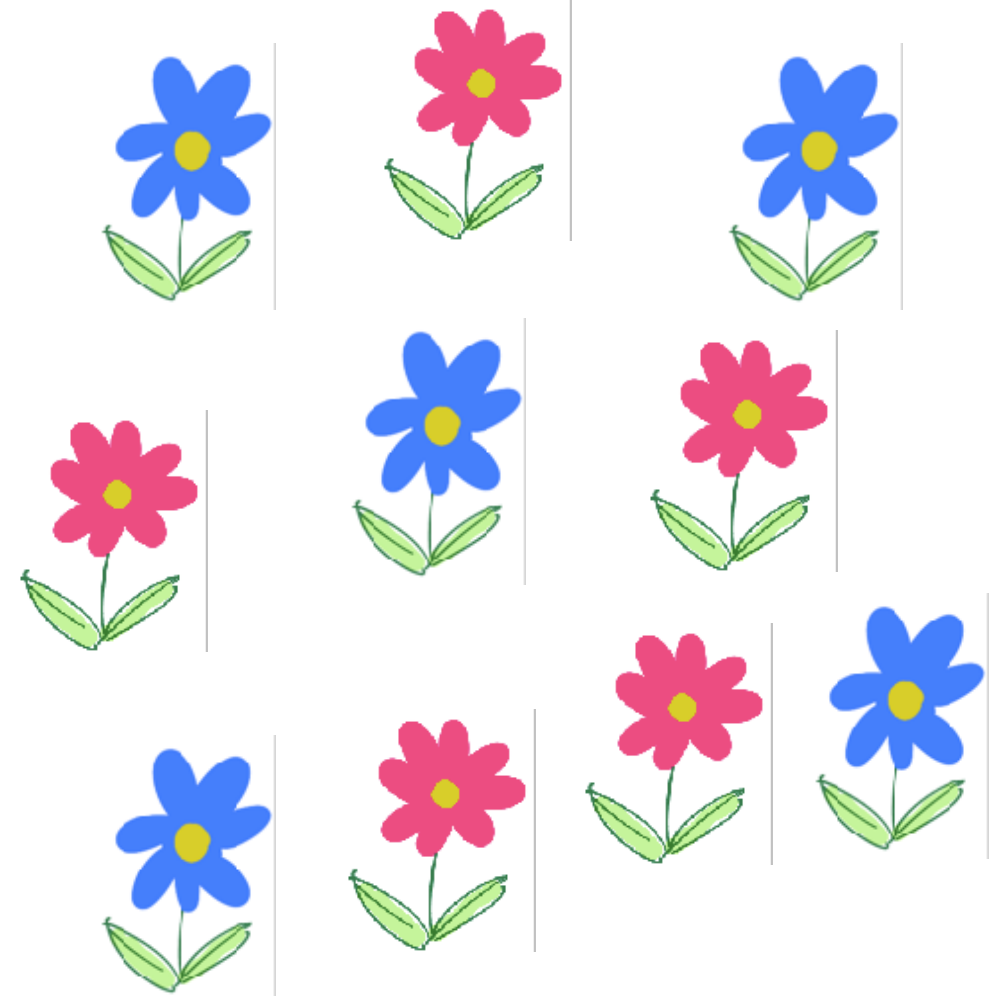
Diversity is related to variety.

Which is more diverse ?

A



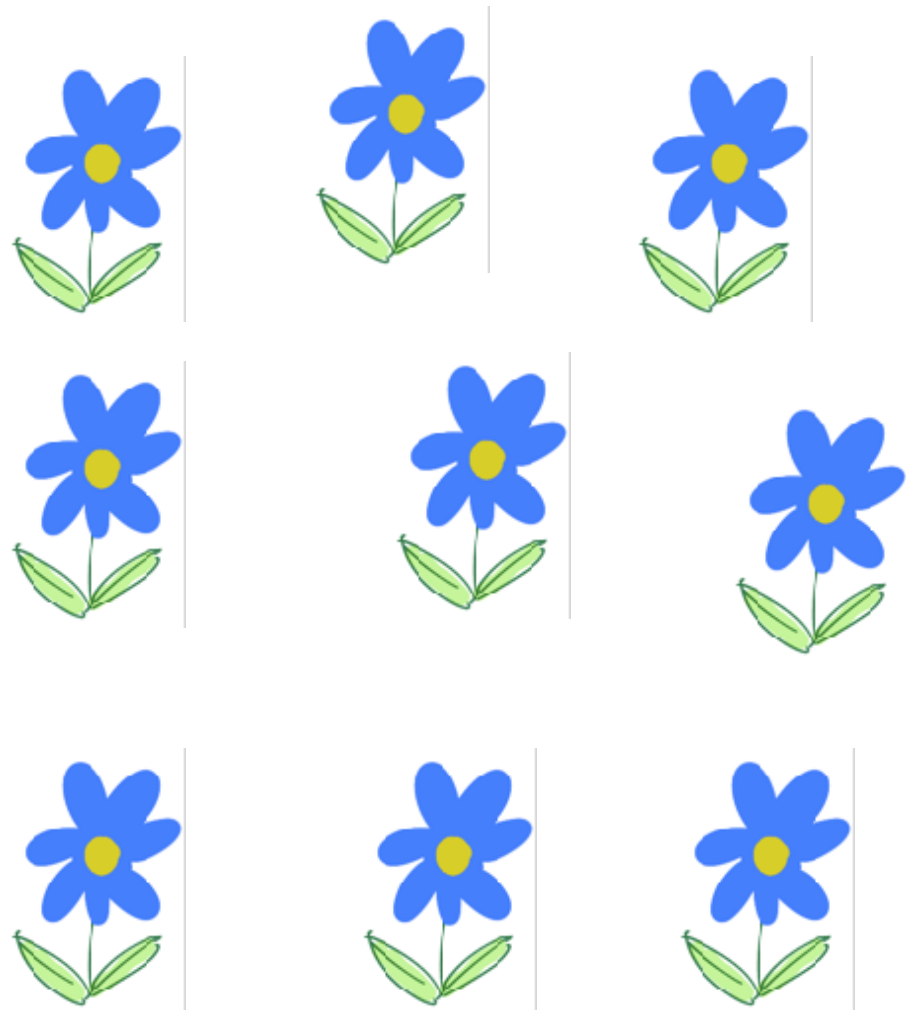
B



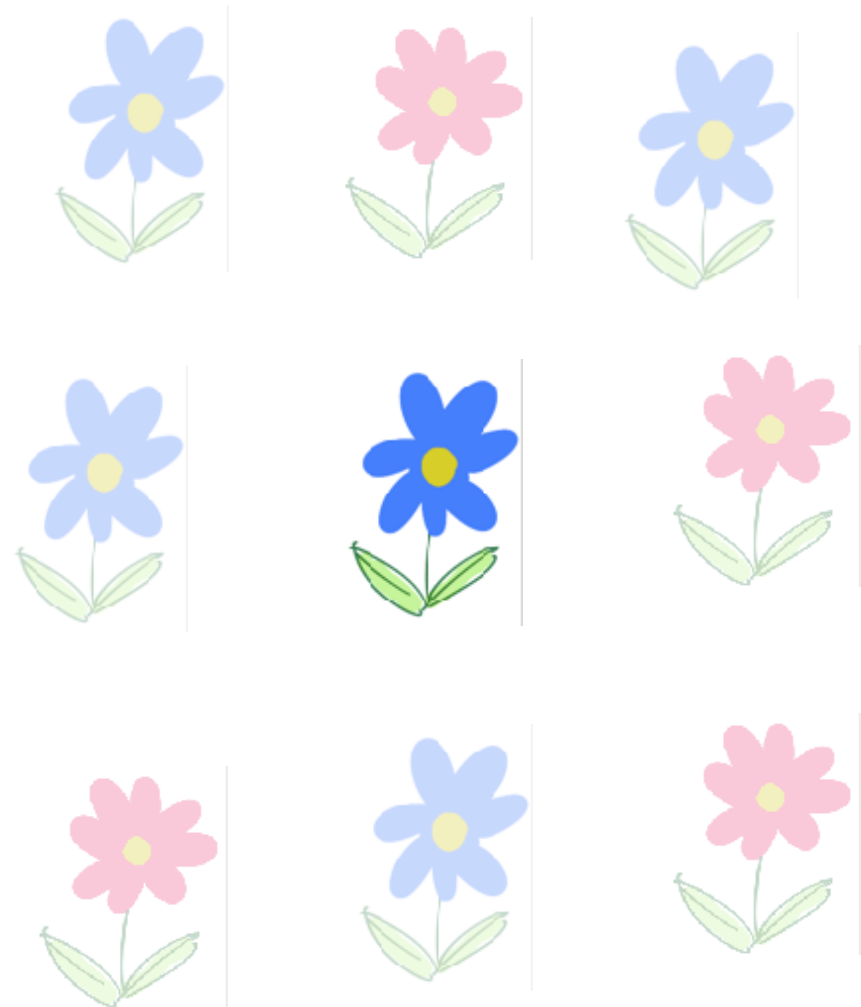
Diversity is related to frequency.

Which is more diverse ?

A



B



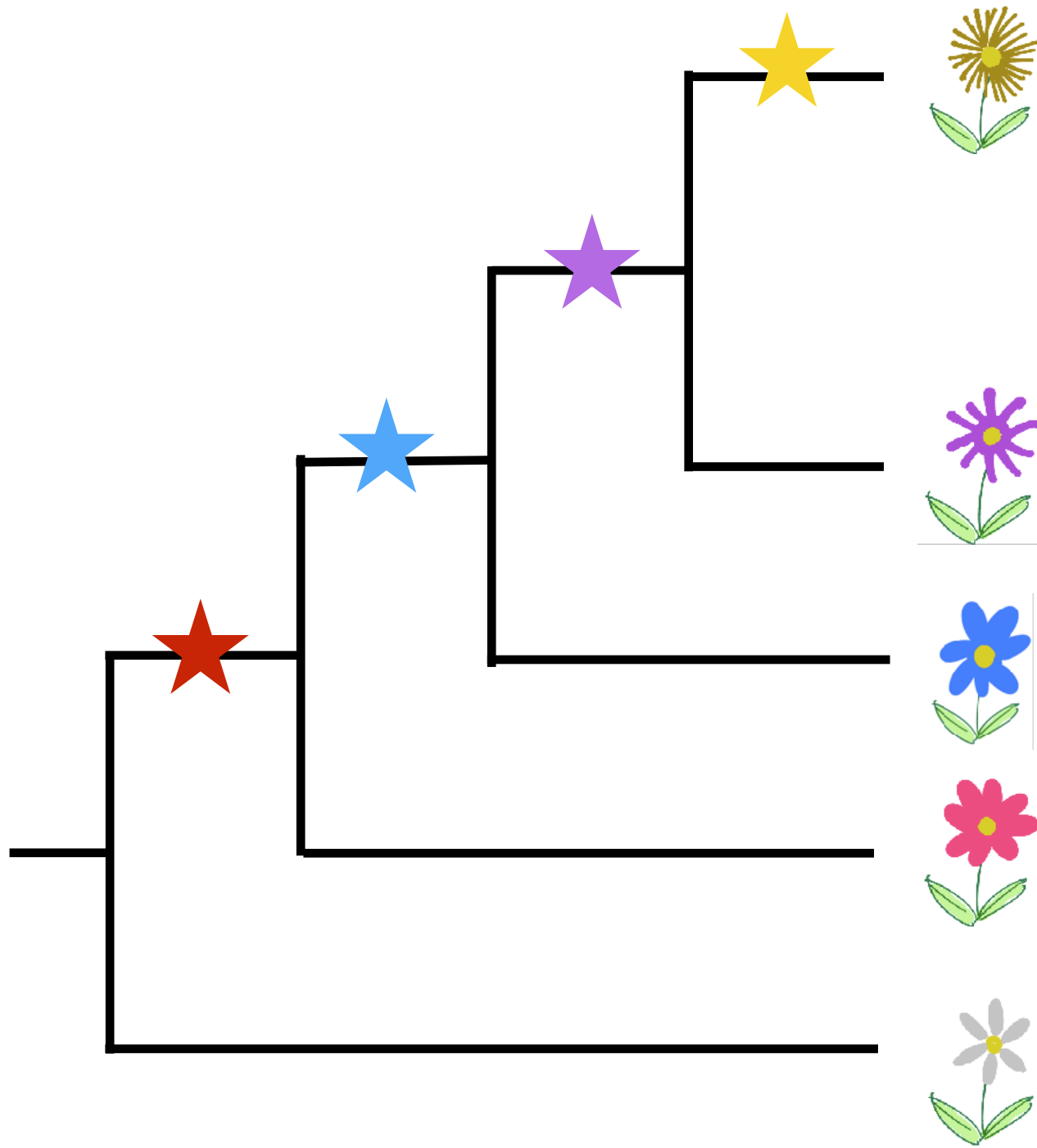
No. of samples is related to accuracy.

Diversity is related to variety and frequency.

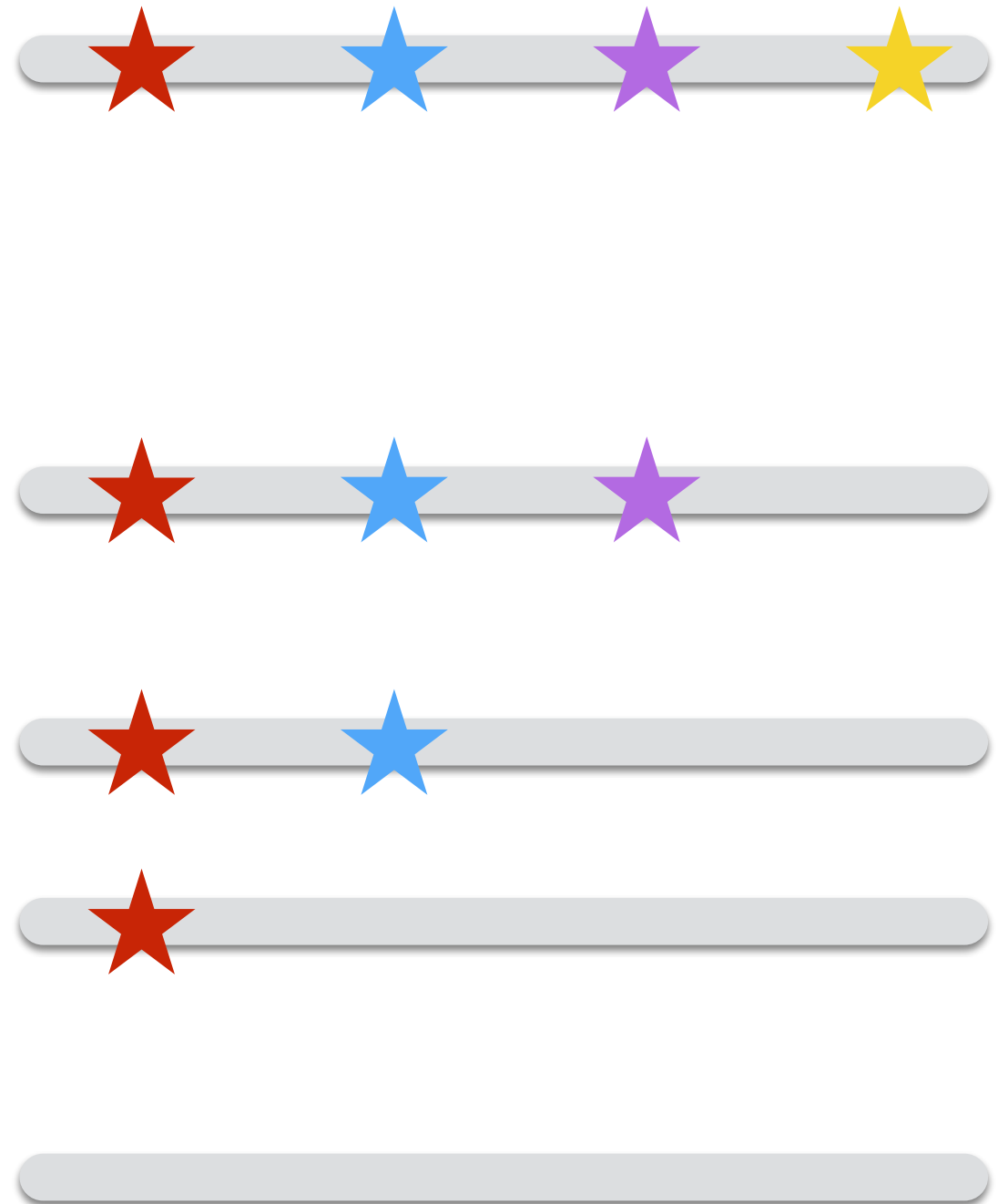
However, how can we calculate diversity from genome.



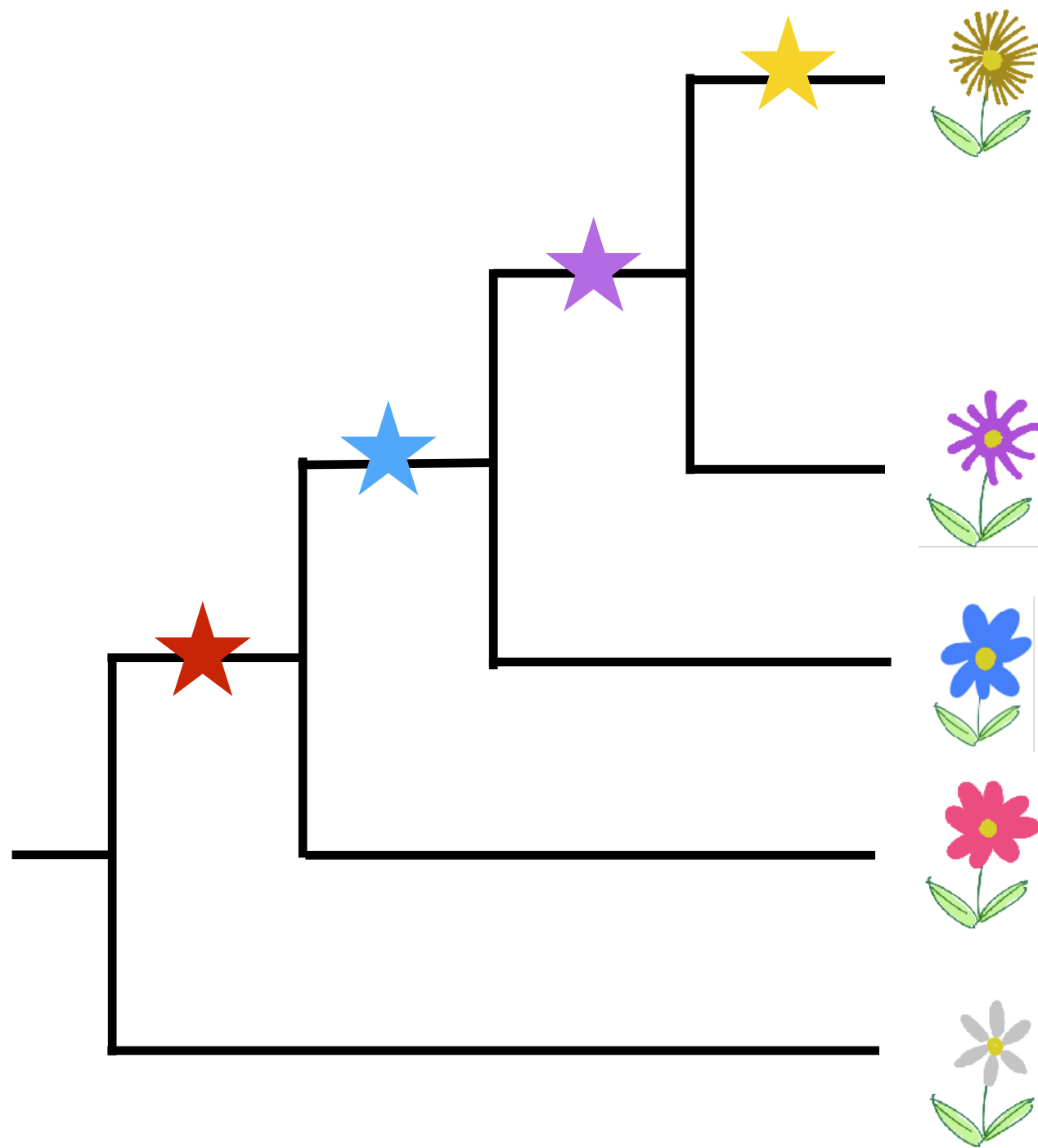
Phylogenetic tree



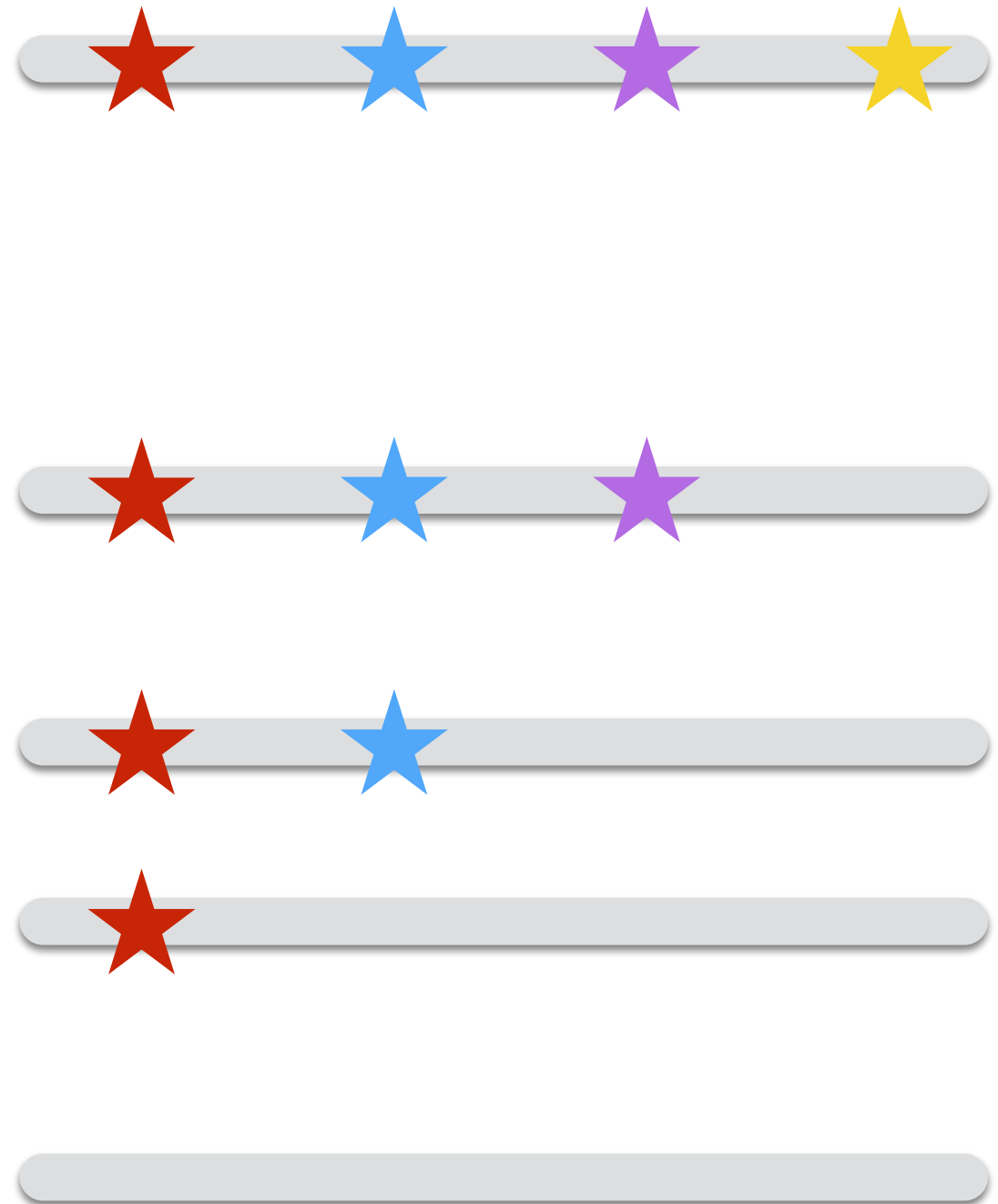
Genome sequence



Phylogenetic tree

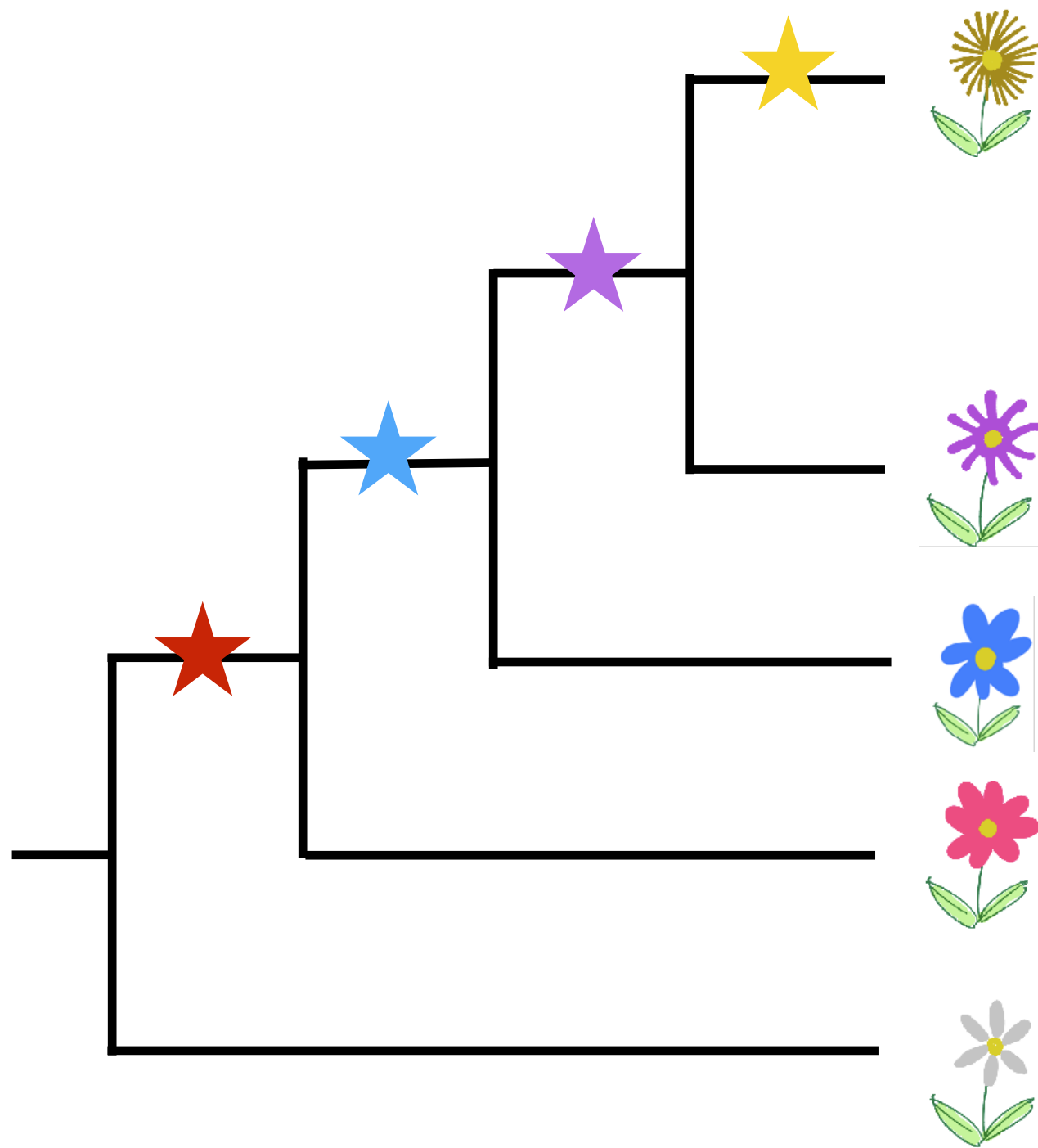


Genome sequence

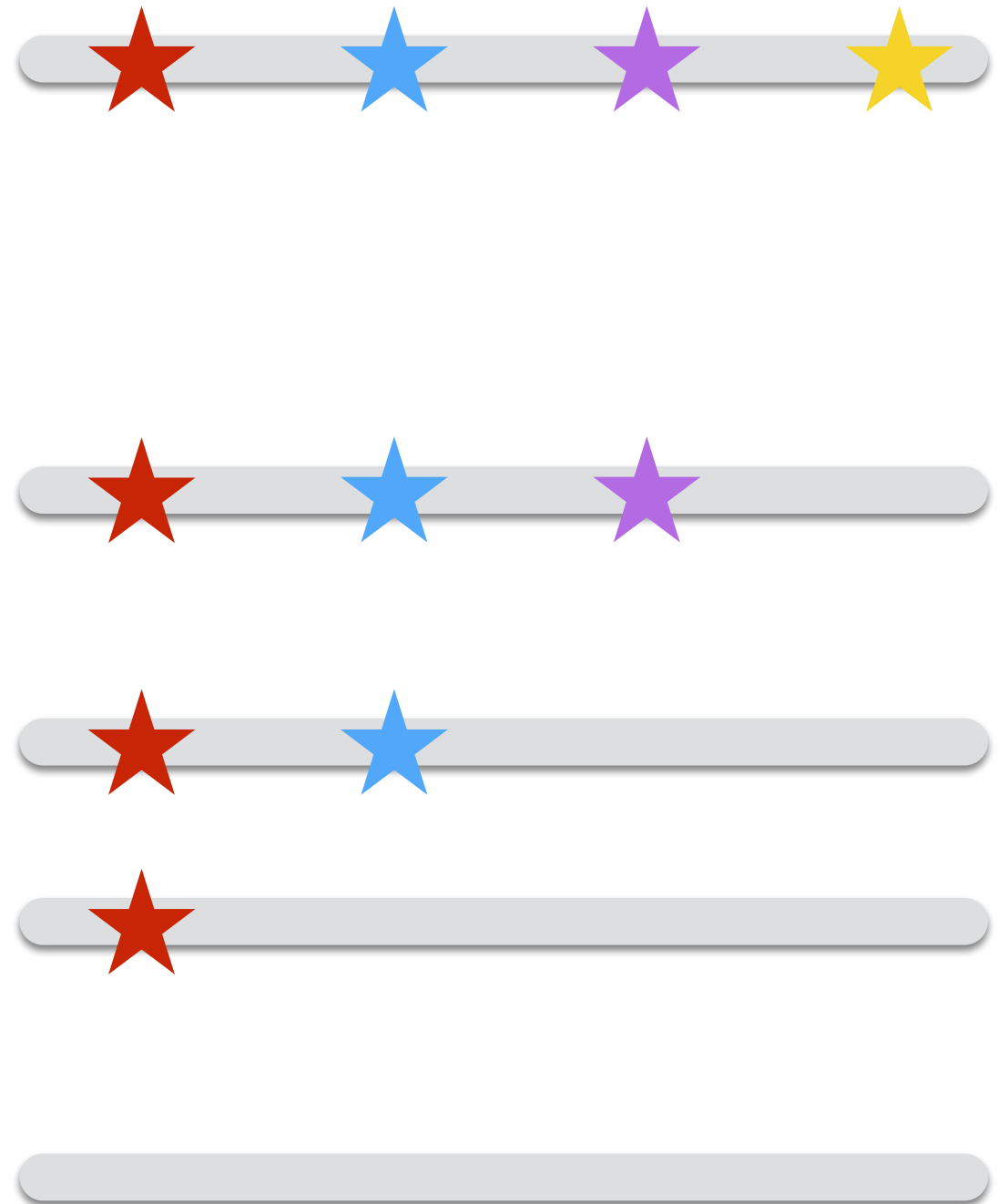


Q. Which  vs.  and  vs.  is more distant?

Phylogenetic tree



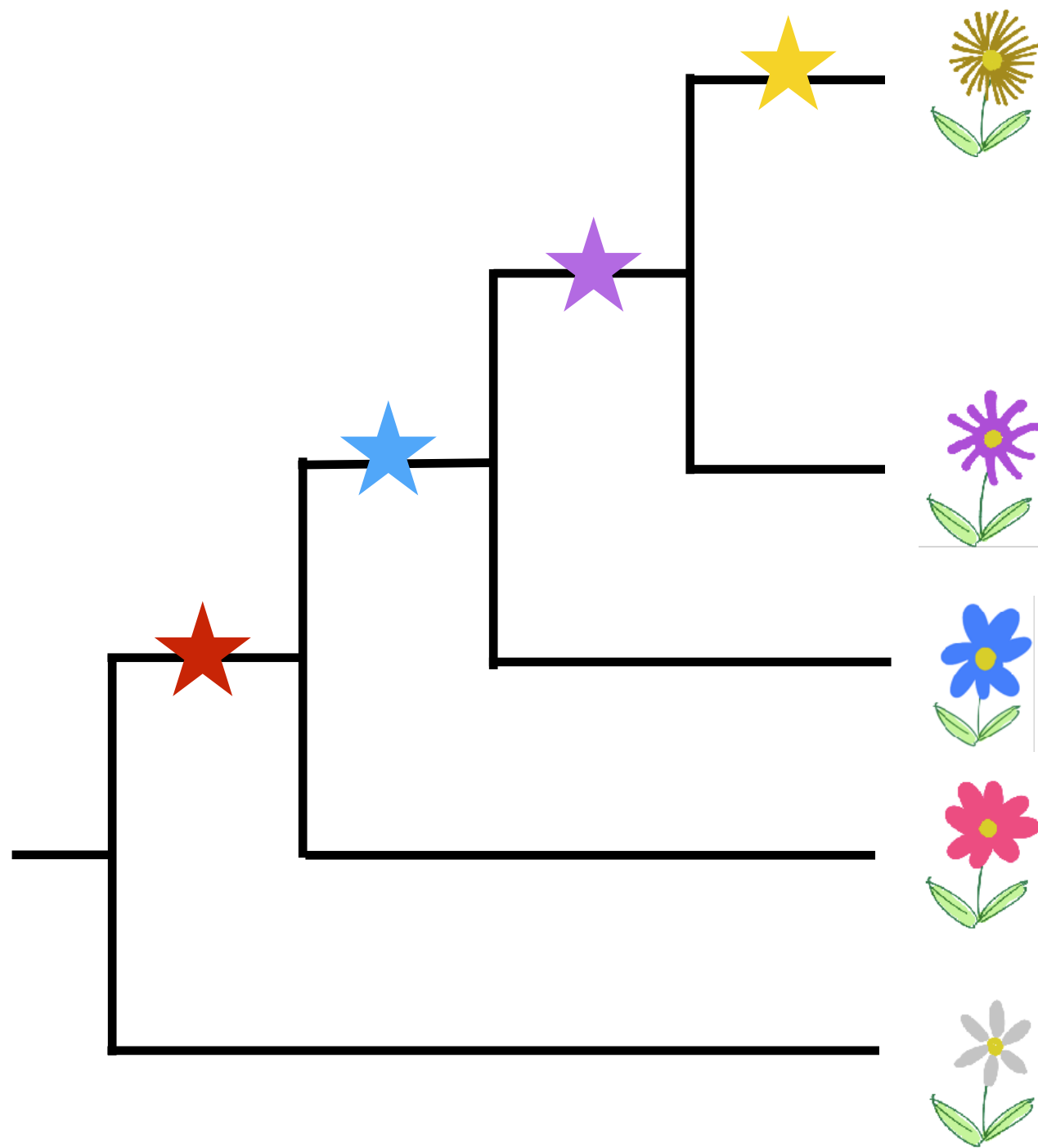
Genome sequence



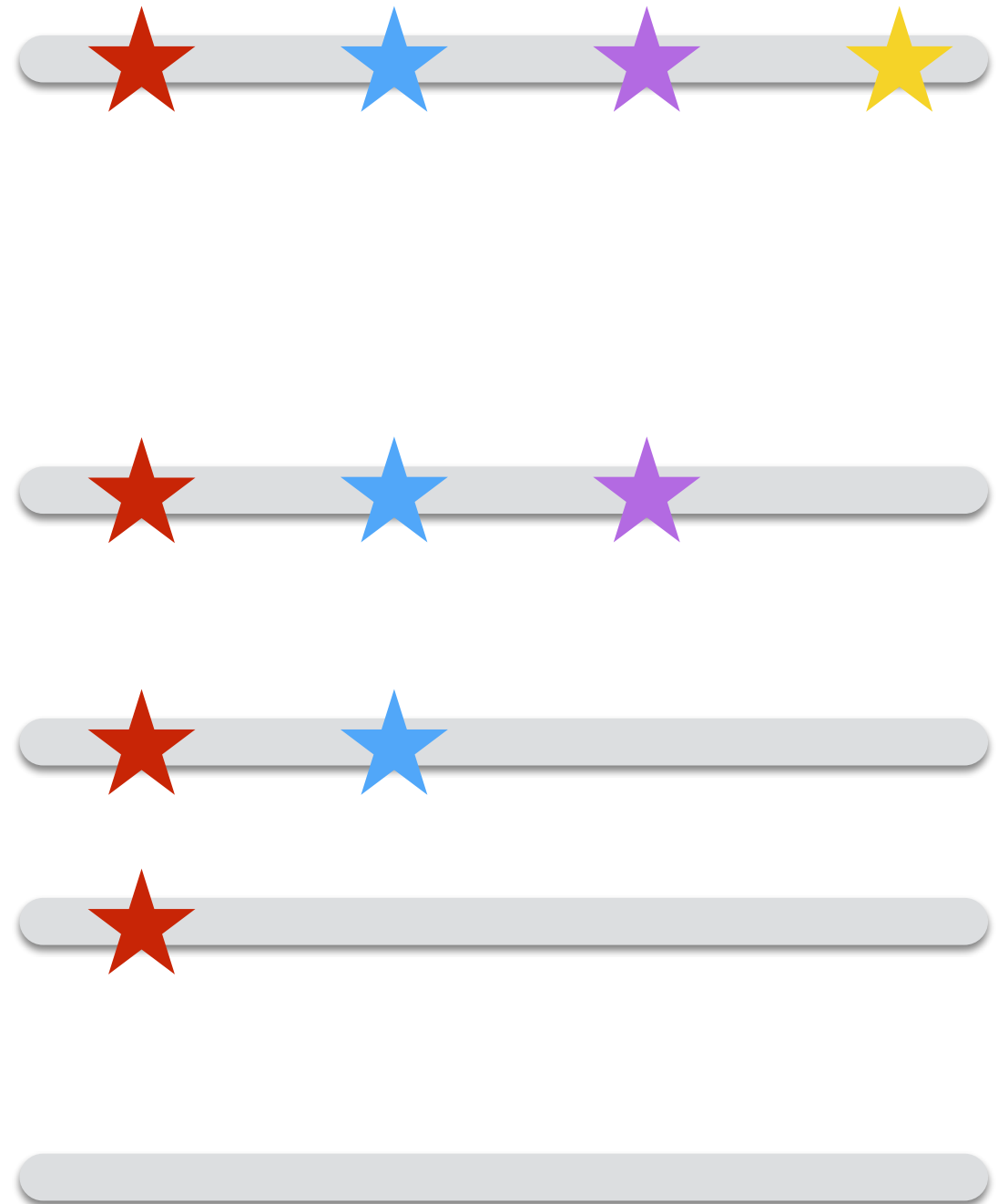
Q. Which  vs.  and  vs.  is more distant?

A.  vs. 

Phylogenetic tree

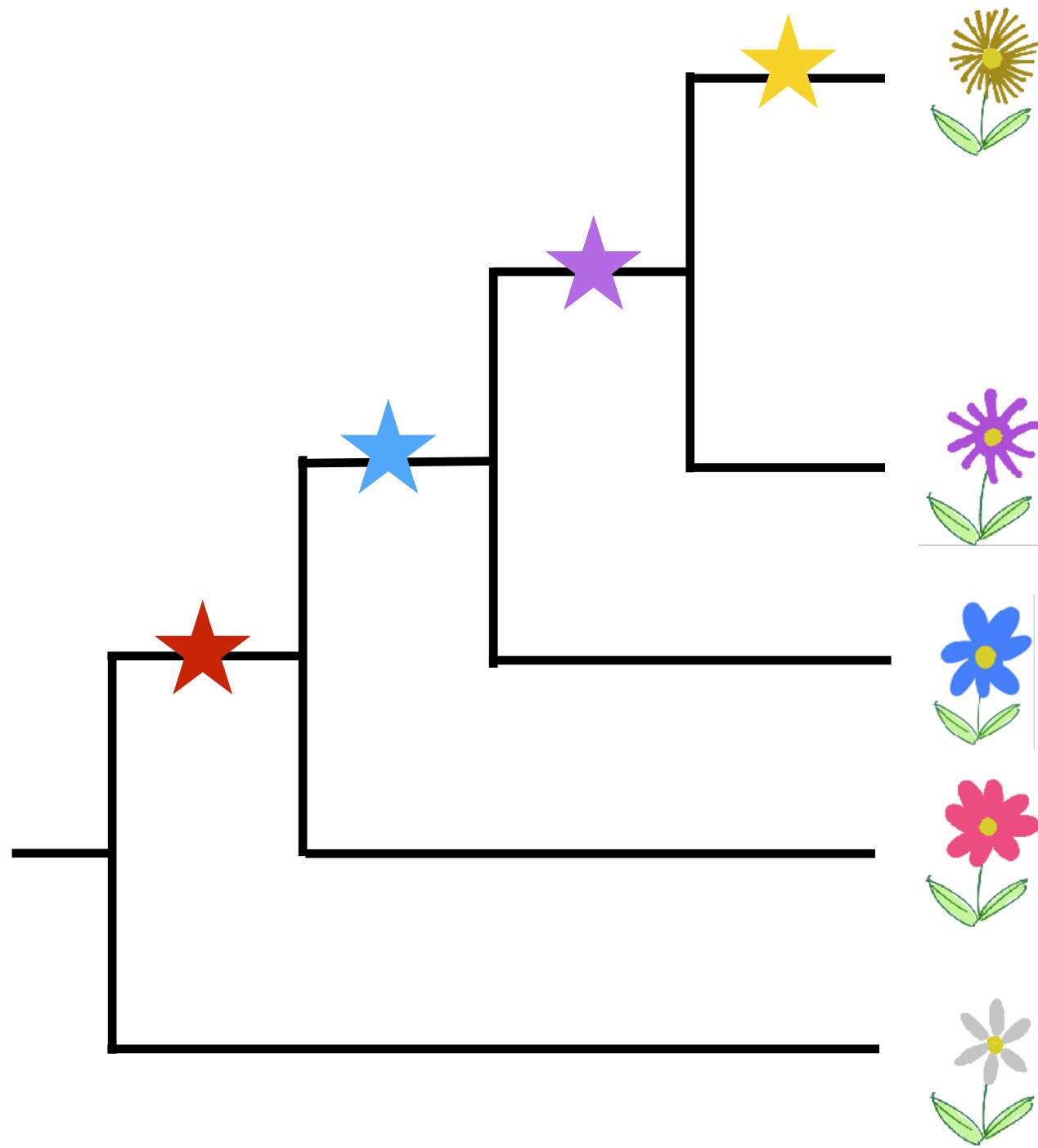


Genome sequence

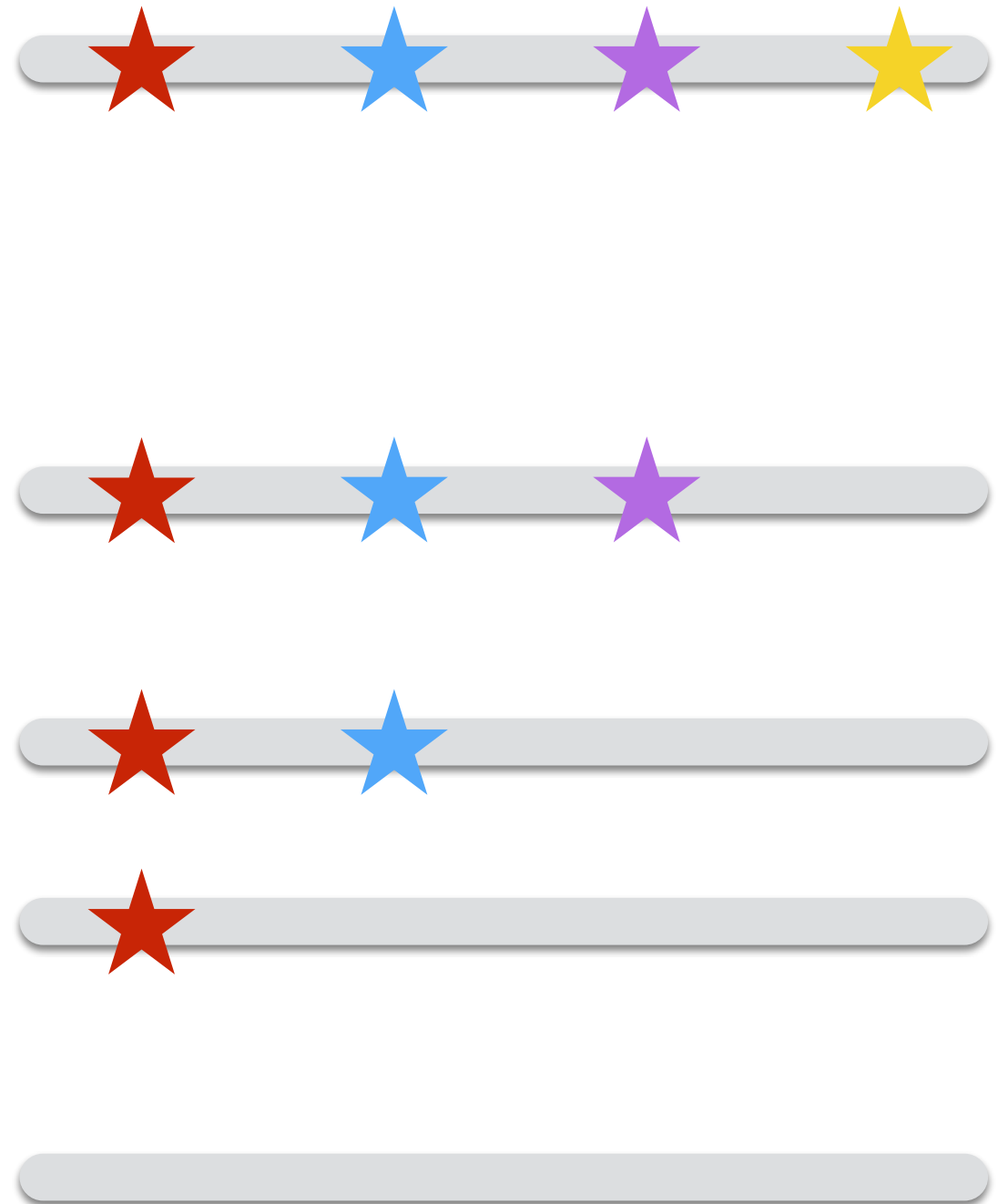


Q. How did you calculate genetic distance?

Phylogenetic tree



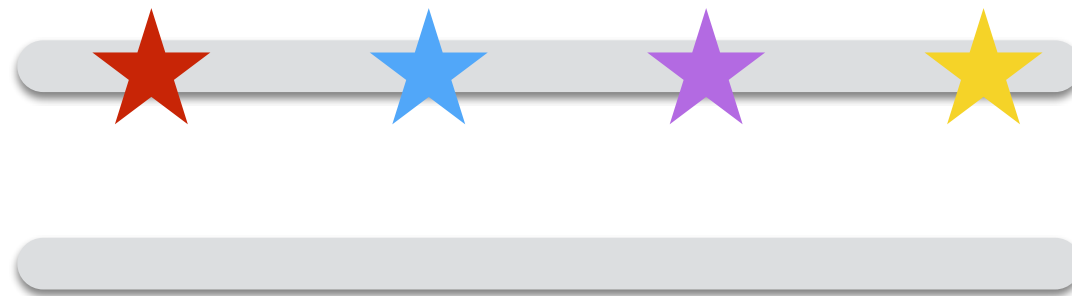
Genome sequence



Q. How did you calculate genetic distance?

A. segregating site

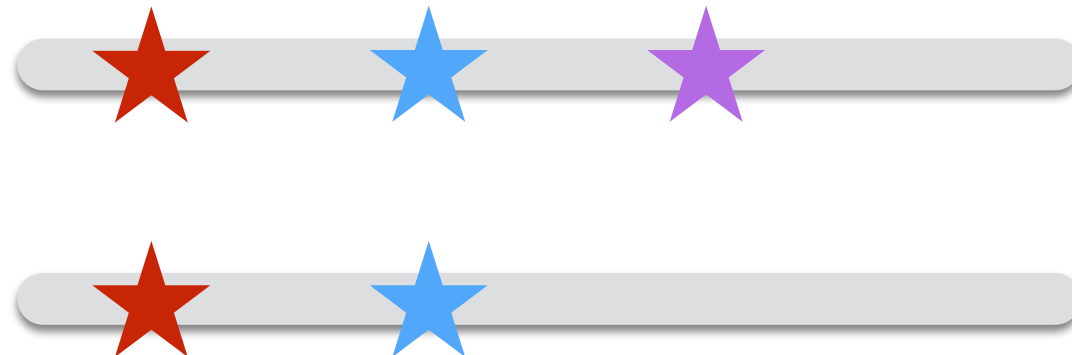
Example 1



segregating site

$$S = 4$$

Example 2



$$S = 1$$

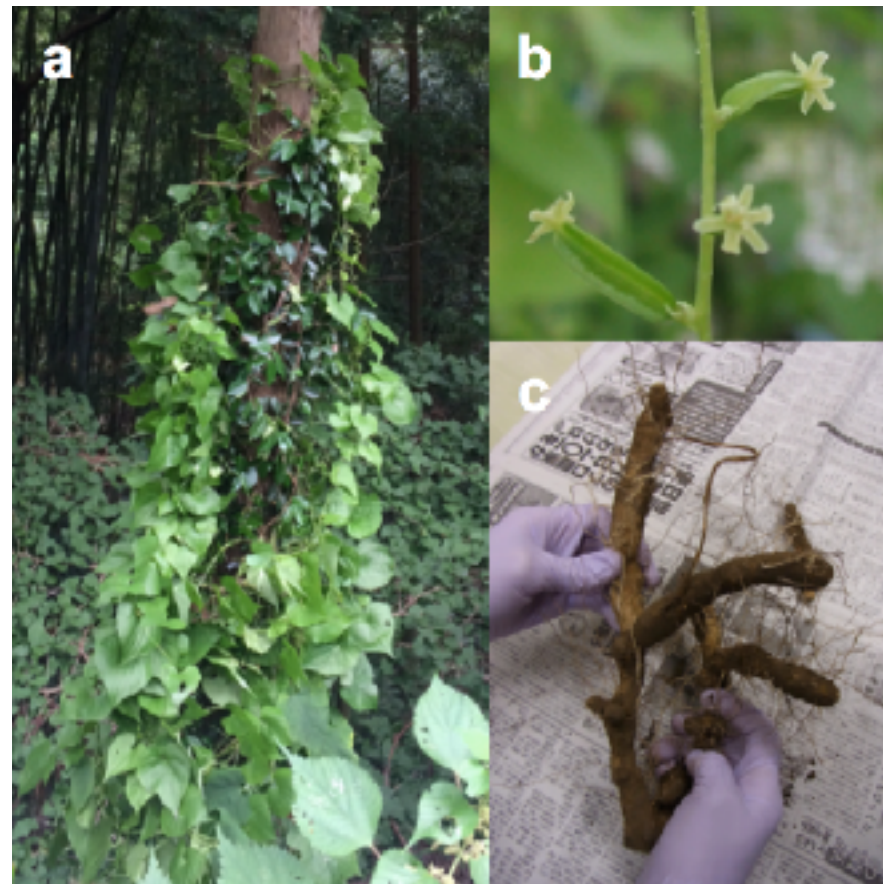
Example 3



$$S = 0$$

Let's count segregating site!

Dioscorea tokoro



S2 = ?

Dioscorea sylvatica



LLIFLE www.llifle.com

S1 = ?

Dioscorea elephantipes



Wikipedia commons

S3 = ?

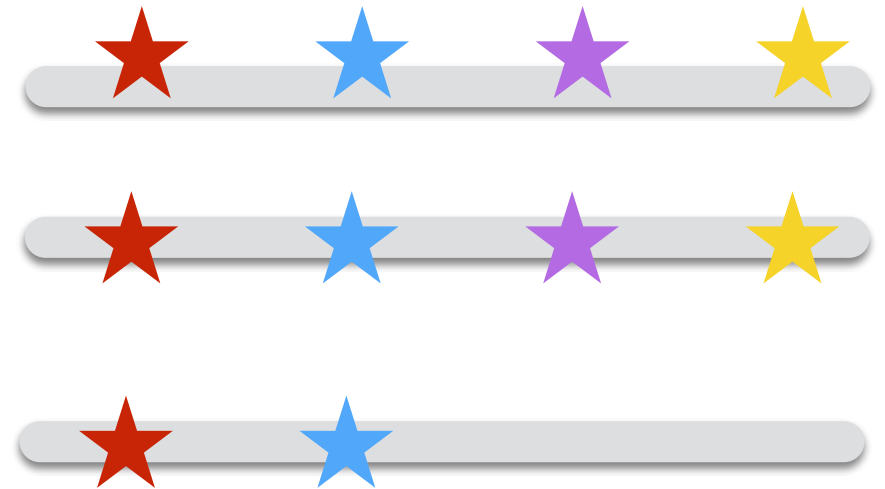


Which is more diverse?

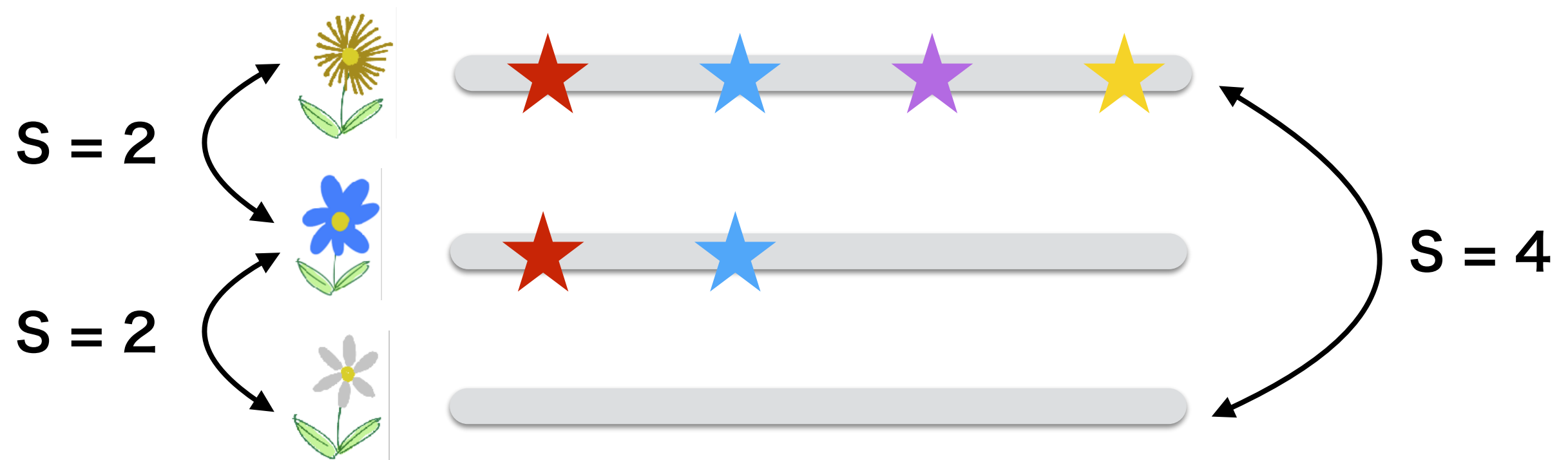
Population A



Population B



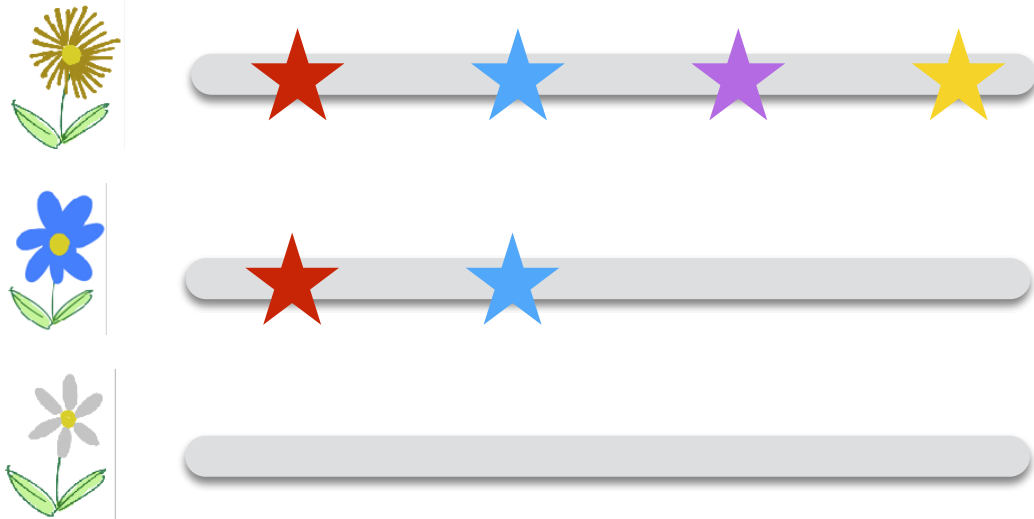
Measurement of Diversity (method 1)



$$\pi = \frac{(\text{Sum of all combinations of } S)}{(\text{No. of combination of sequences})} = \frac{2 + 2 + 4}{3} = \frac{8}{3}$$

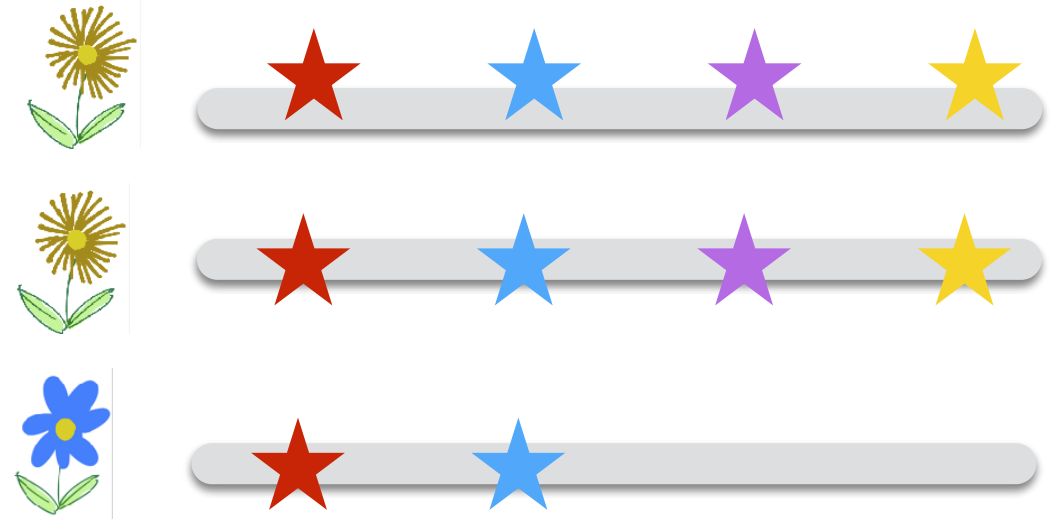
Which is more diverse?

A



$$\frac{2 + 2 + 4}{3} = \frac{8}{3}$$

B

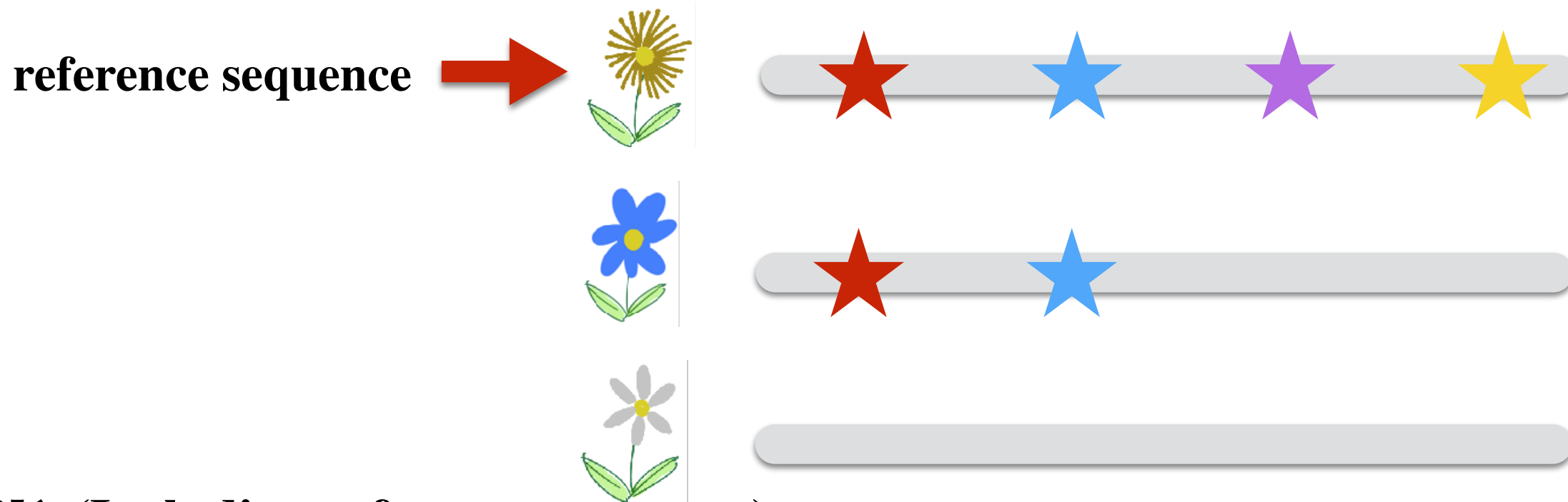


$$\frac{0 + 2 + 2}{3} = \frac{4}{3}$$

>

Measurement of Diversity (method 1)

✱ Method 1 and method 2 show the same result.



(1) (Including reference sequence)

No. of base same to reference

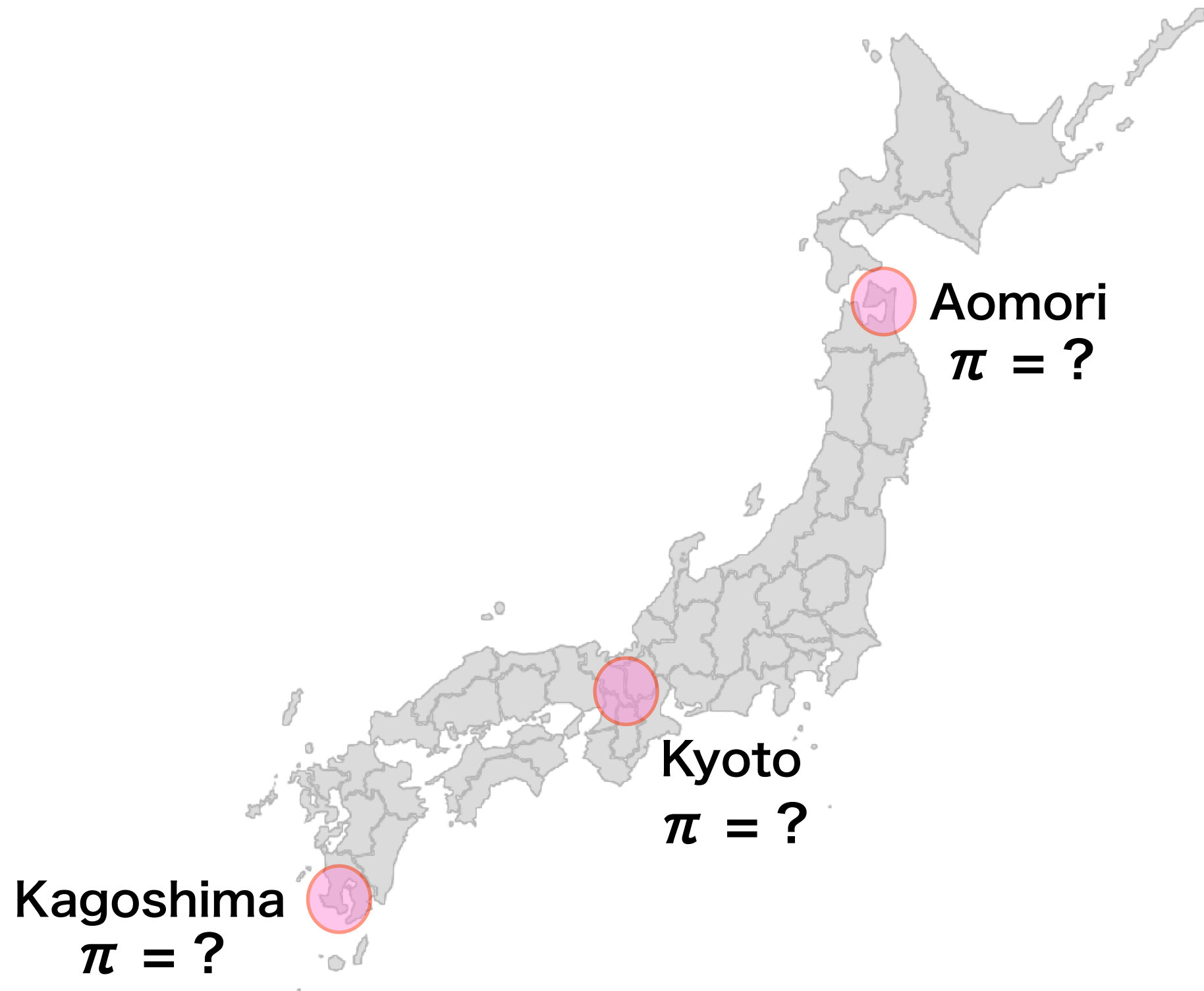
No. of base different from reference

(2) (same) × (different)

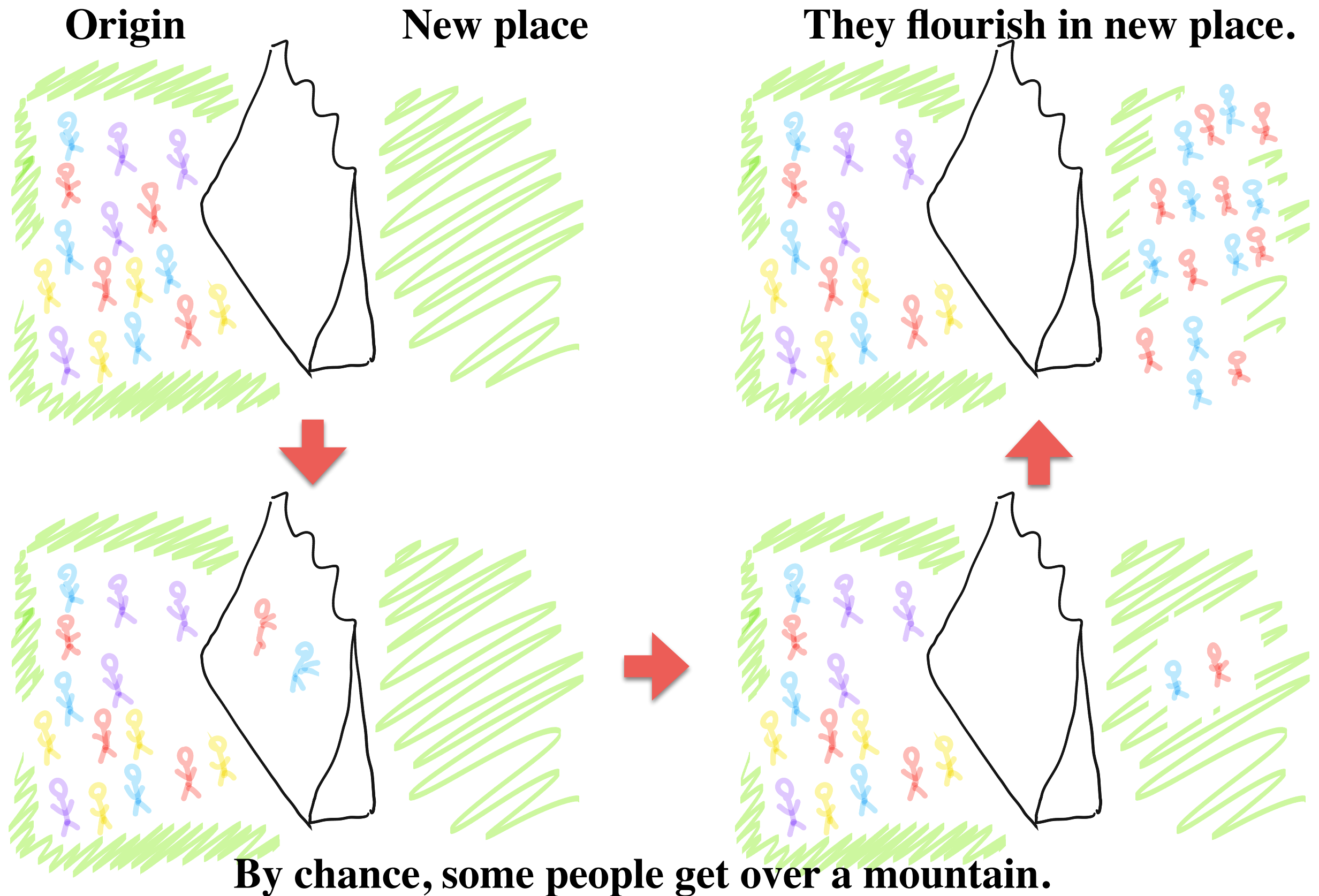
	2	2	1	1	
	×	×	×	×	
	1	1	2	2	
	+	+	+	+	
	2	2	2	2	
	=				8

(3) divide by No. of all combinations of sequences $8 / {}_3C_2 = 8/3$

Let's calculate π !



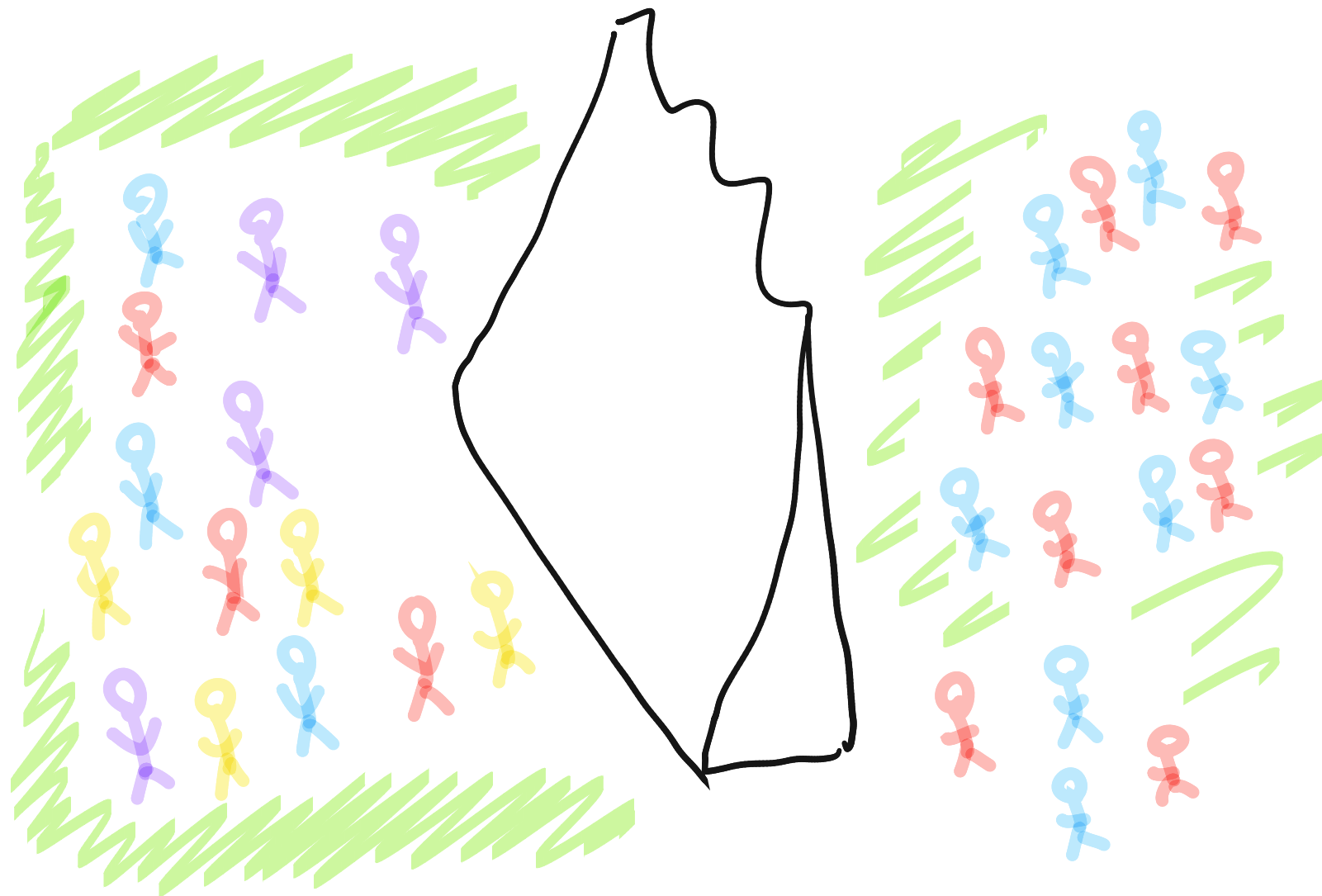
What is the effect of bottleneck?



Then, origin place is more diverse.

High diversity

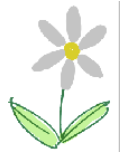
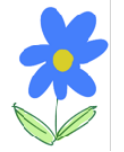
Low diversity



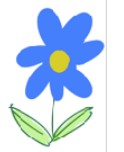
So, we can consider where is origin of a species using π .

Which is more distant between 2 populations?

A



vs.



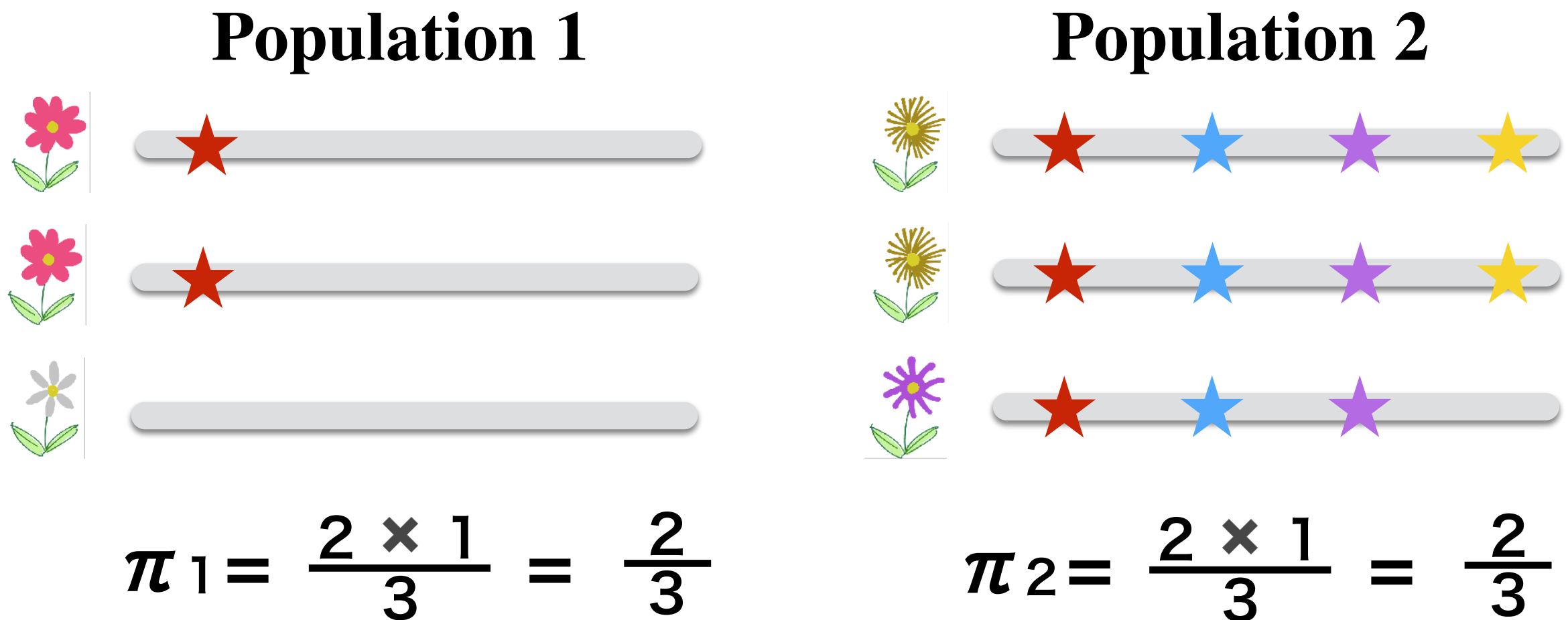
B



vs.



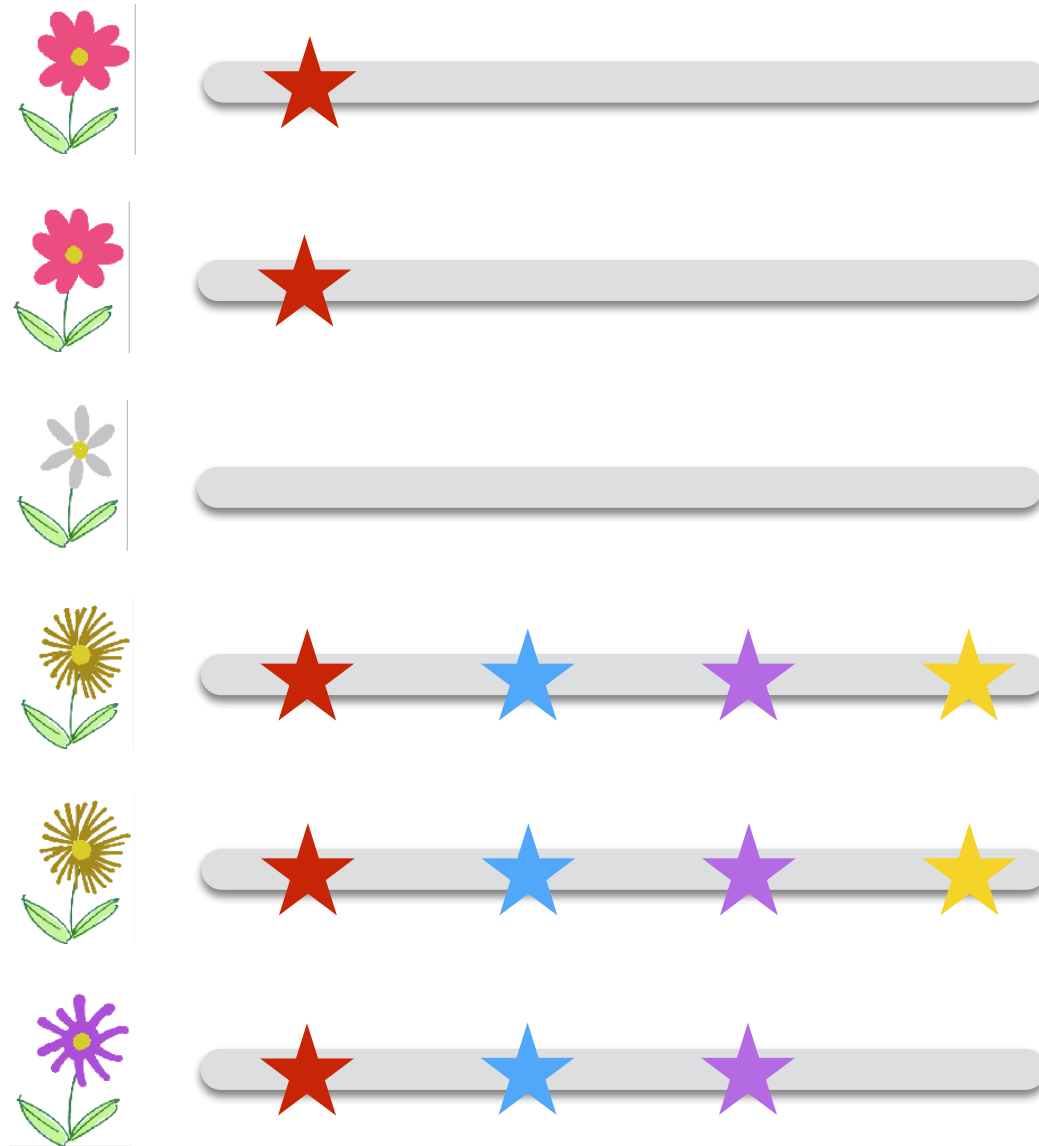
Measurement of genetic distance between two populations



(1) Calculate π in each populations.

Measurement of genetic distance between two populations

(2) assume two population as one population.



(3) calculate π in that population.

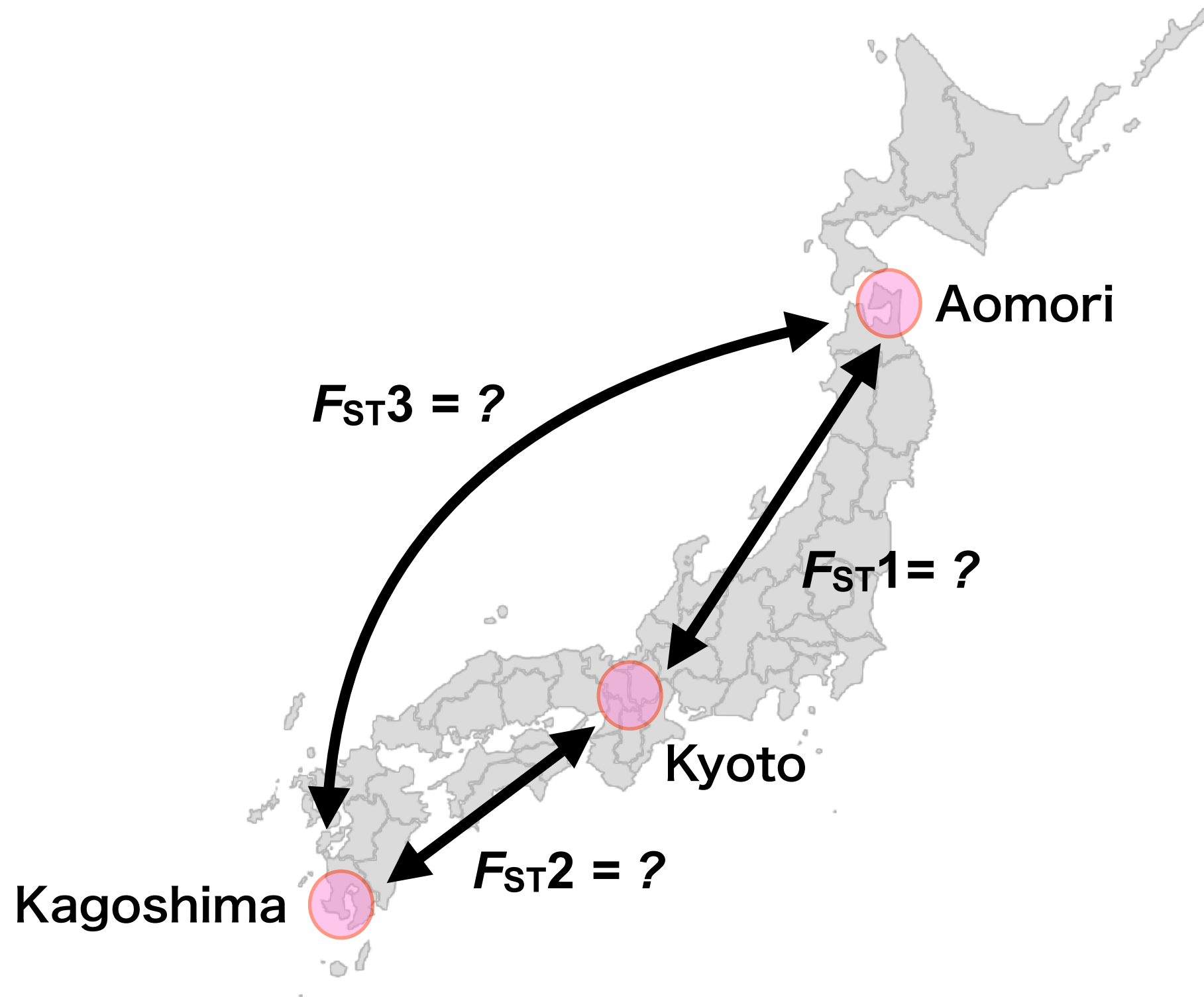
$$\pi_{12} = \frac{5 \times 1 + 3 \times 3 + 3 \times 3 + 4 \times 2}{6C_2} = \frac{31}{15}$$

Measurement of genetic distance between two populations

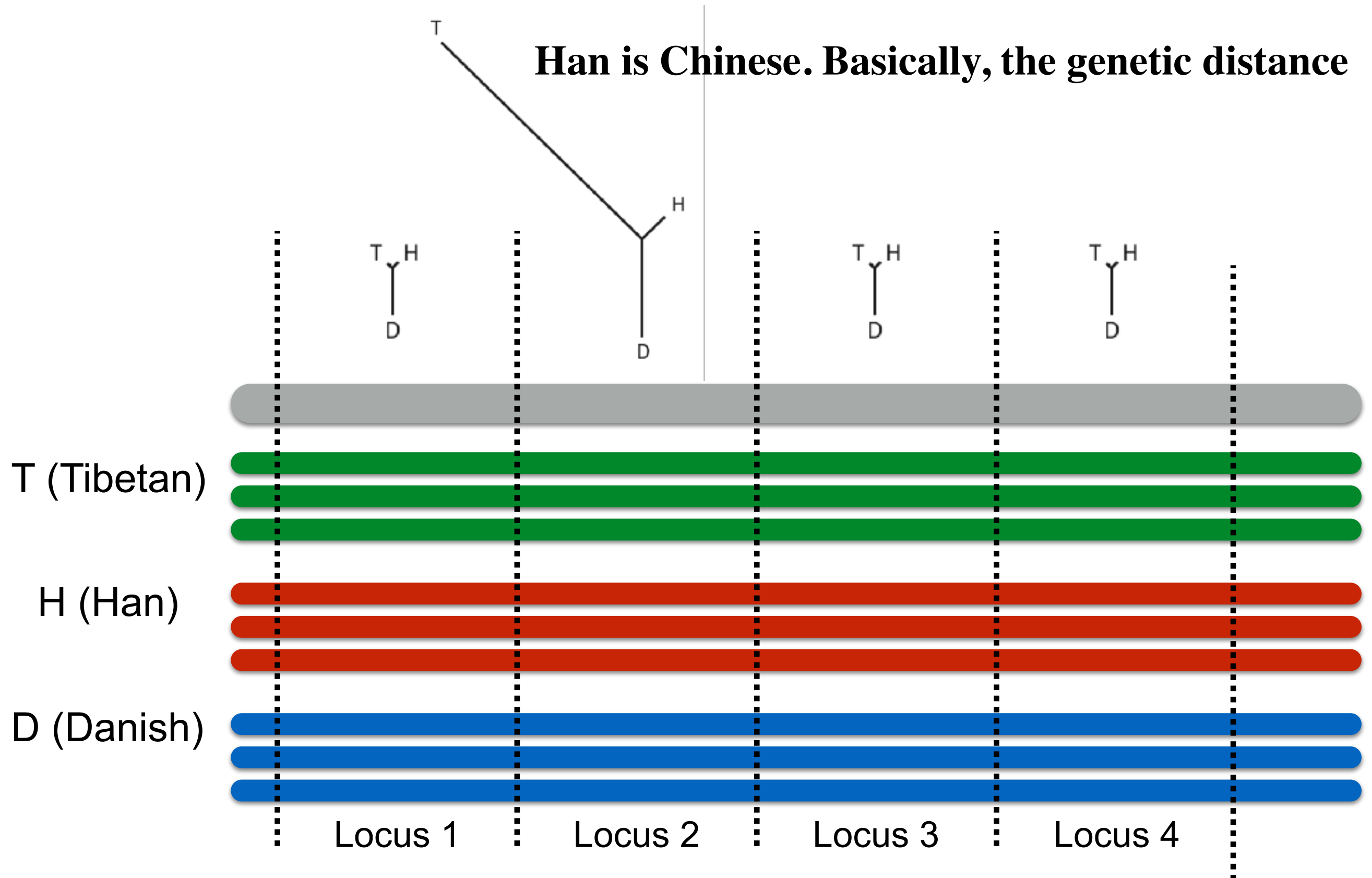
$$\begin{aligned} F_{ST} &= \frac{\pi_{12} - (\pi_1 + \pi_2) / 2}{\pi_{12}} \\ &= \frac{\frac{31}{15} - \left(\frac{2}{3} + \frac{2}{3} \right) / 2}{\frac{31}{15}} \\ &= \frac{21}{31} \end{aligned}$$

F_{ST} is the measure between 0 and 1. Usually, F_{ST} over 0.4 is very high.

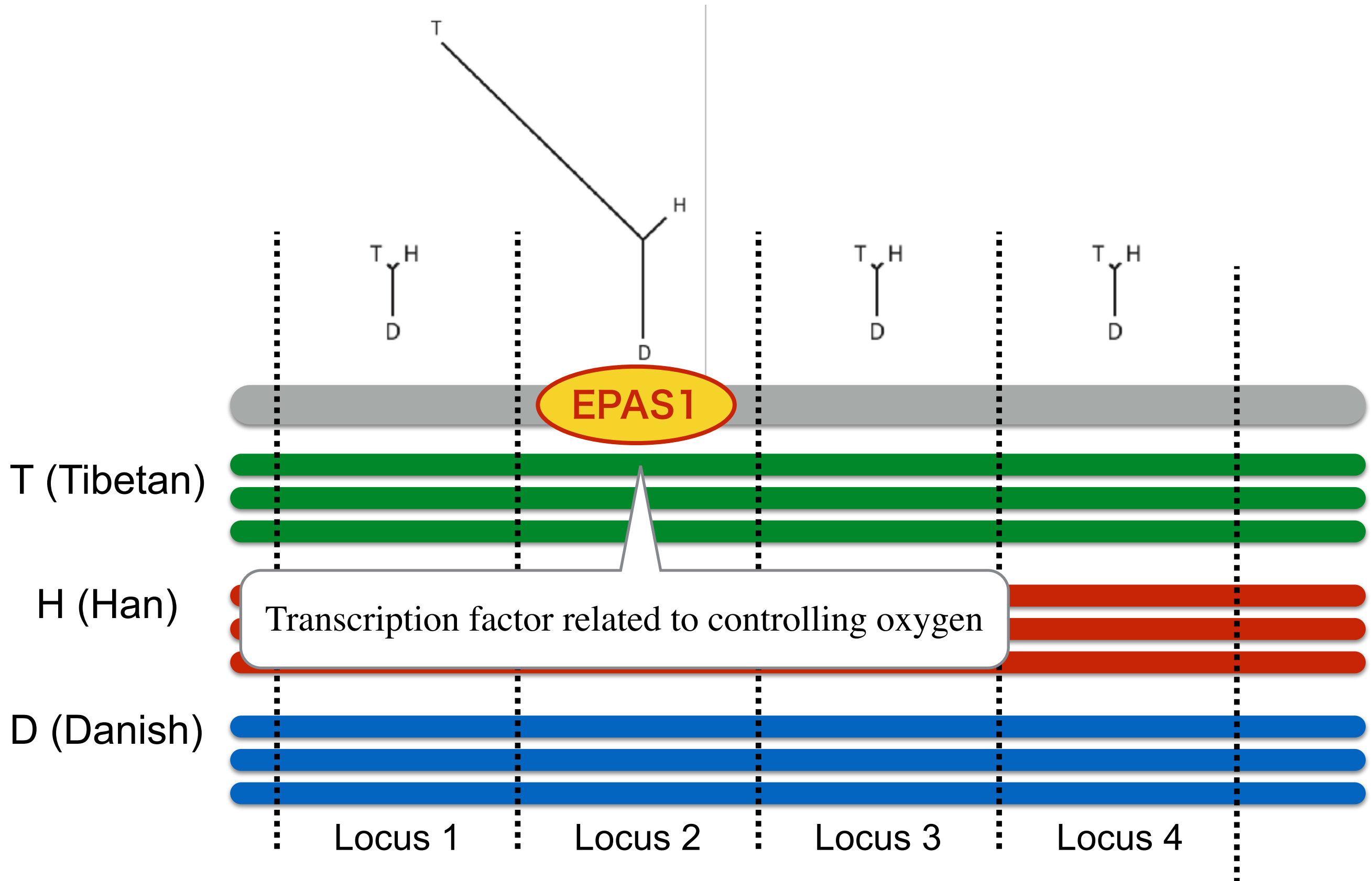
Let's calculate F_{ST} !



Genome scan using F_{ST}



Genome scan using F_{ST}



Han is Chinese. Basically, the genetic distance between Chinese and Tibetan is closer than Danish. However, some Tibetan loci is distant from Chinese due to local adaptation. For example, *EPAS1* is so. This gene is for controlling oxygen and related to hypoxia. Probably, this is a local adaptation to high altitude land in Tibetan.