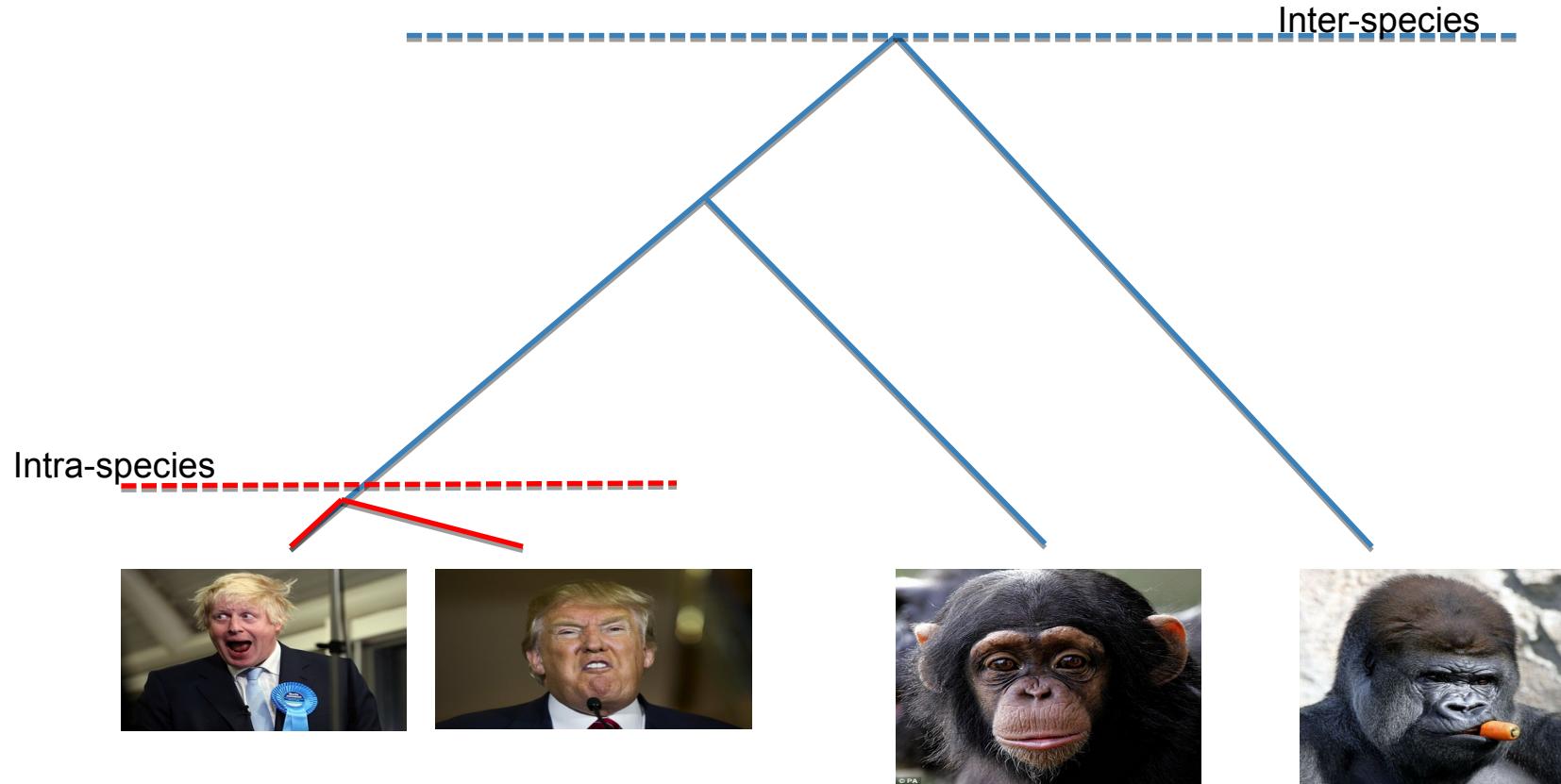


Detecting signature of natural selection from genomic data (part 2)

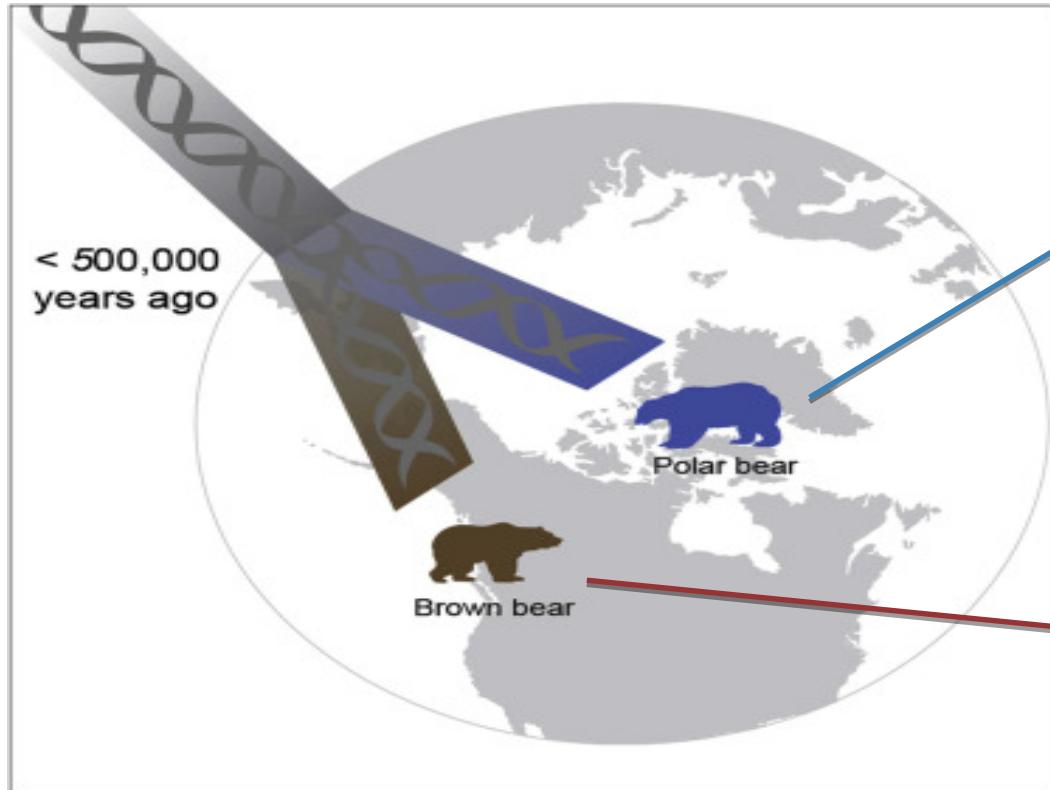
Matteo Fumagalli
2019

Inter-species variation



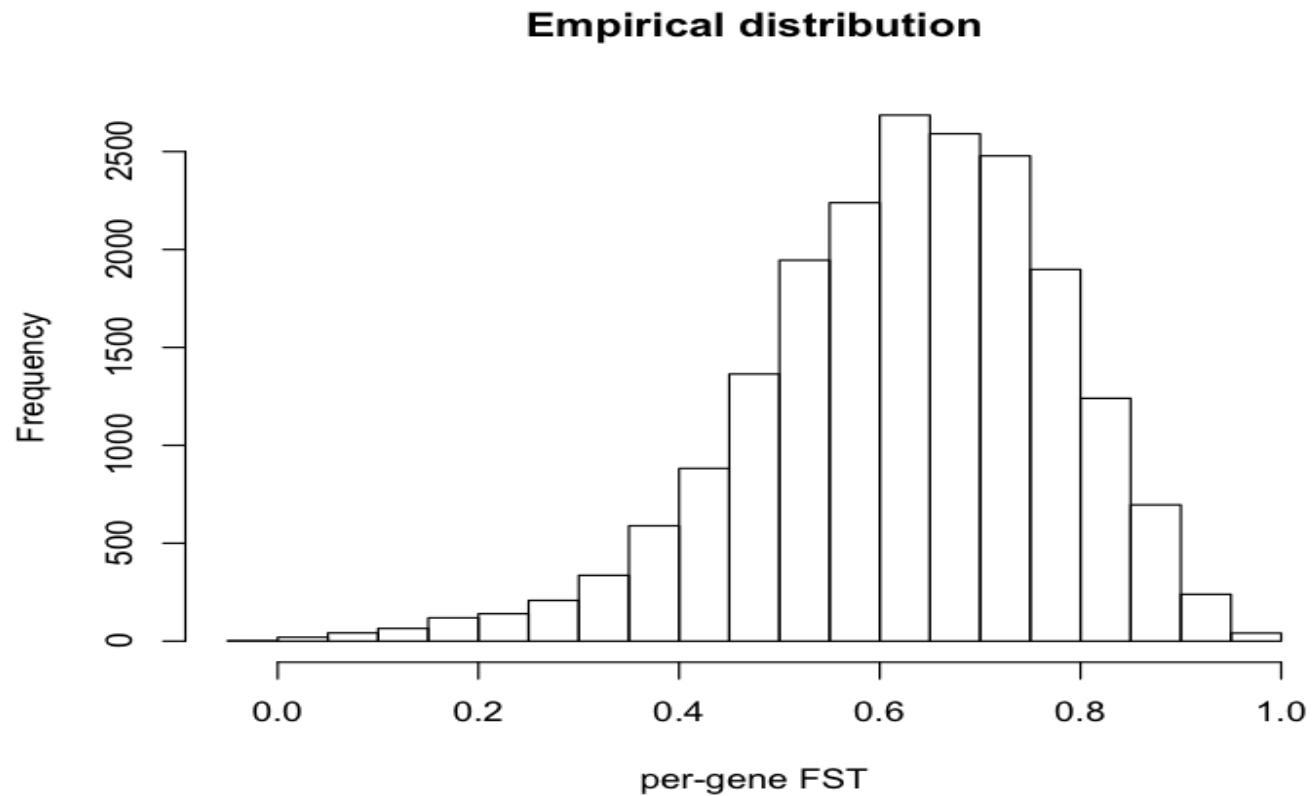
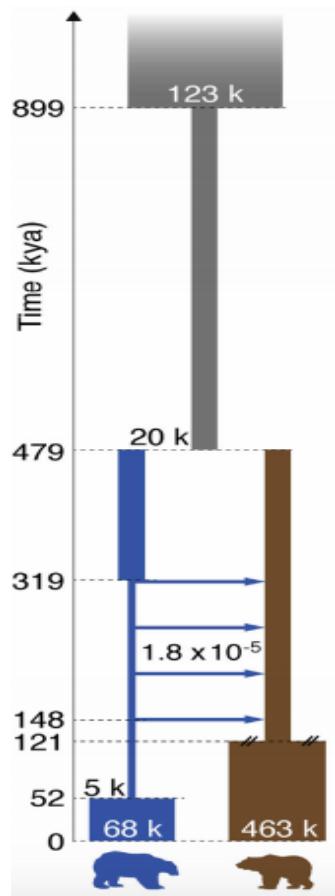
events in the deep past, macro-evolutionary trends, selection between species

Polar bears vs. Brown bears

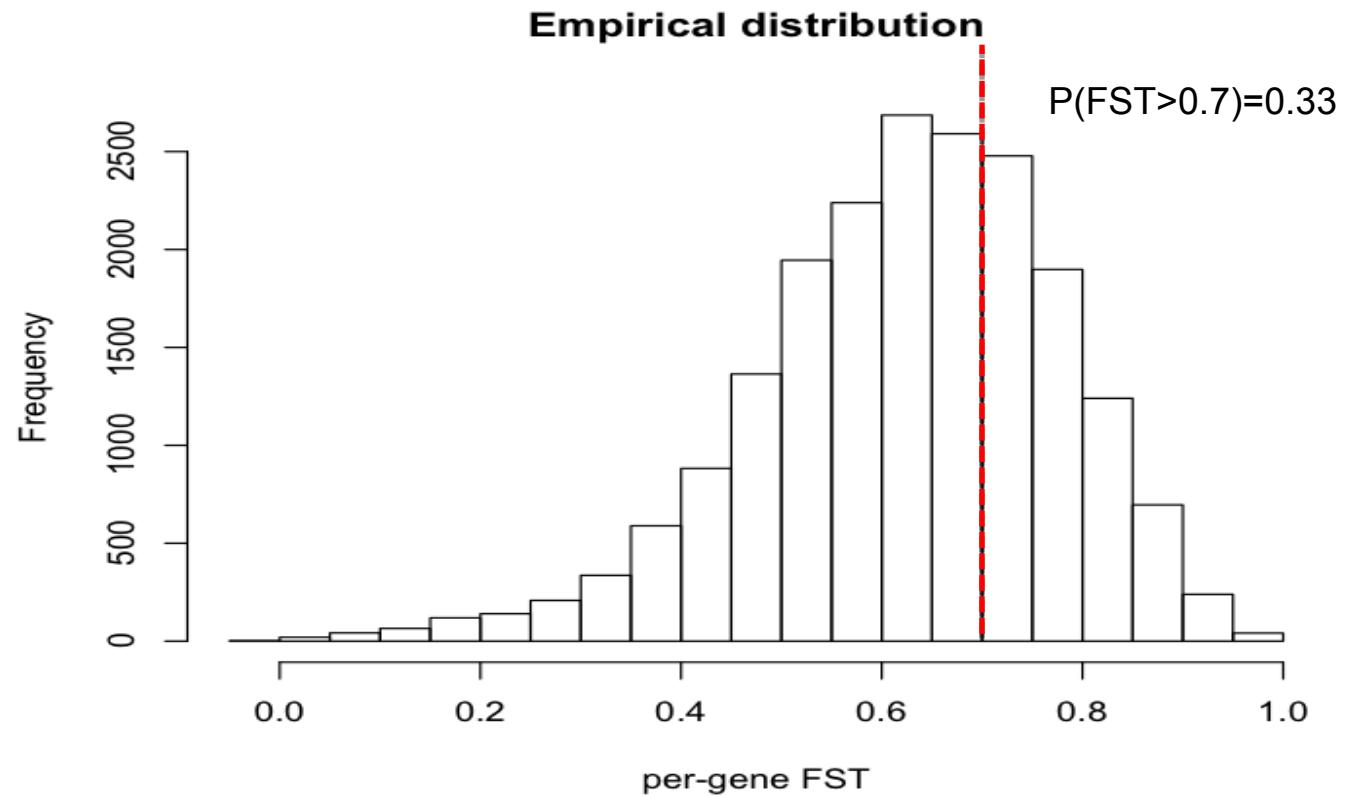
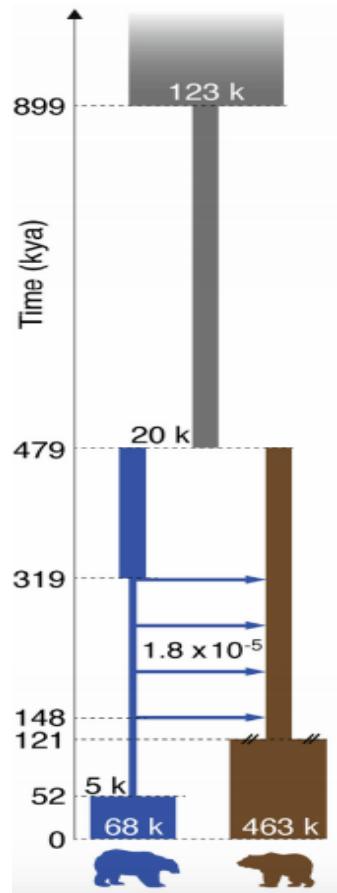


Question: what are the genetic signatures of polar bears' adaptation to the Arctic environment?

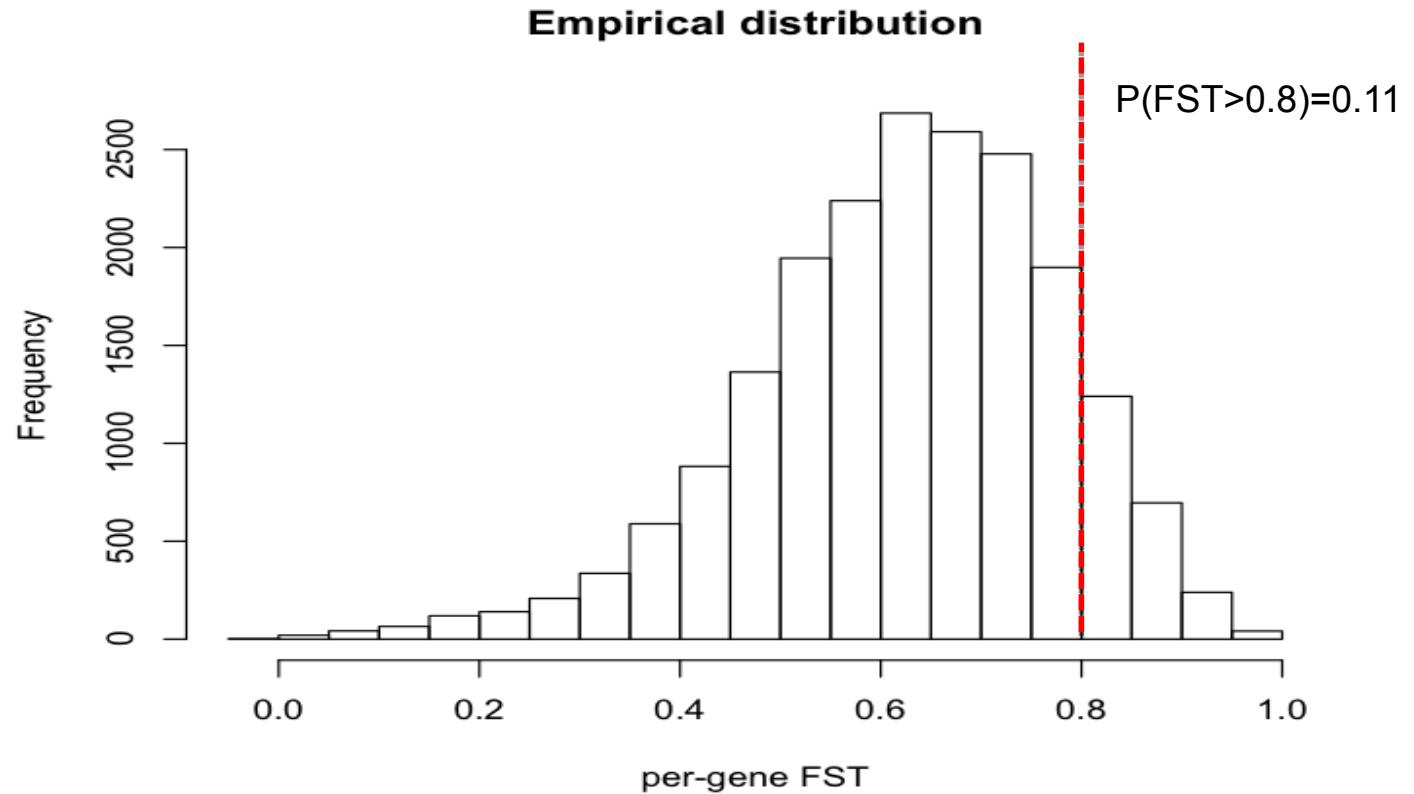
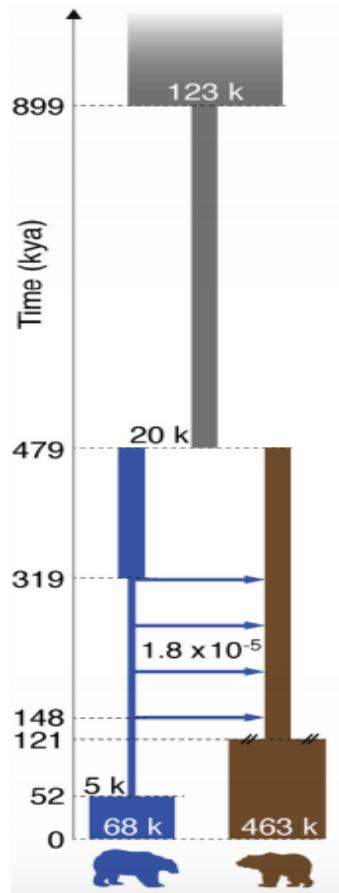
Expected genetic differentiation



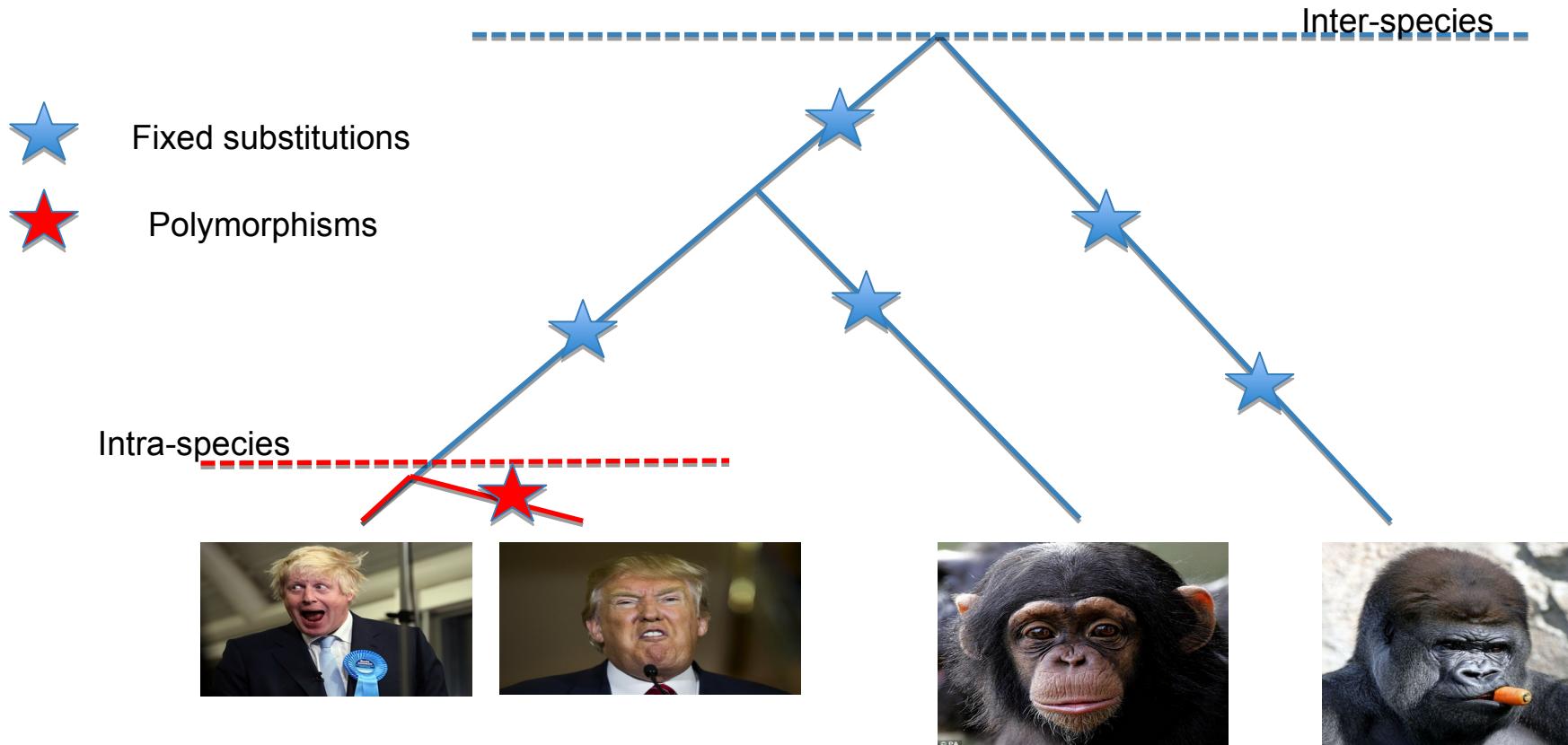
Expected genetic differentiation



Expected genetic differentiation



Inferring inter-species selection

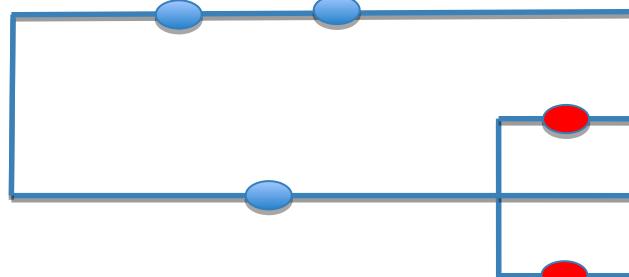


Polymorphisms and divergence

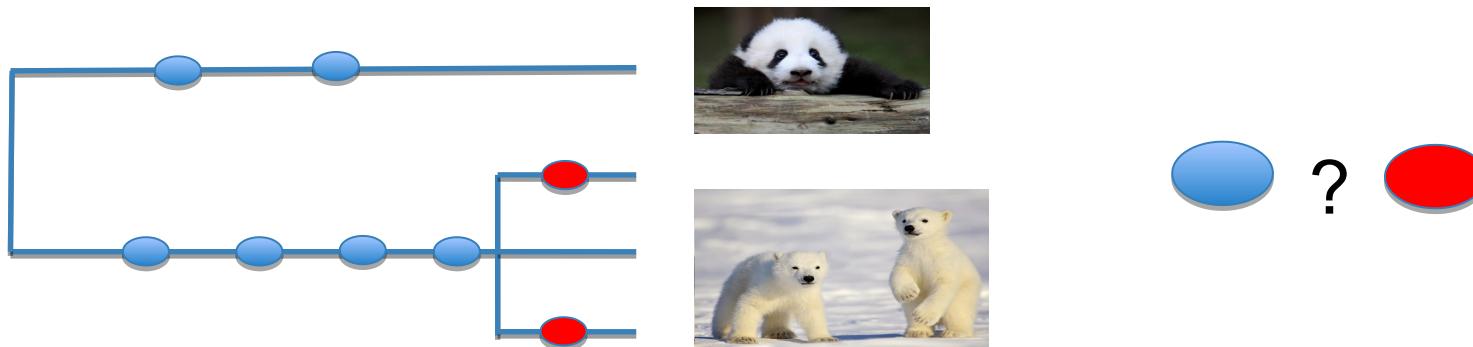
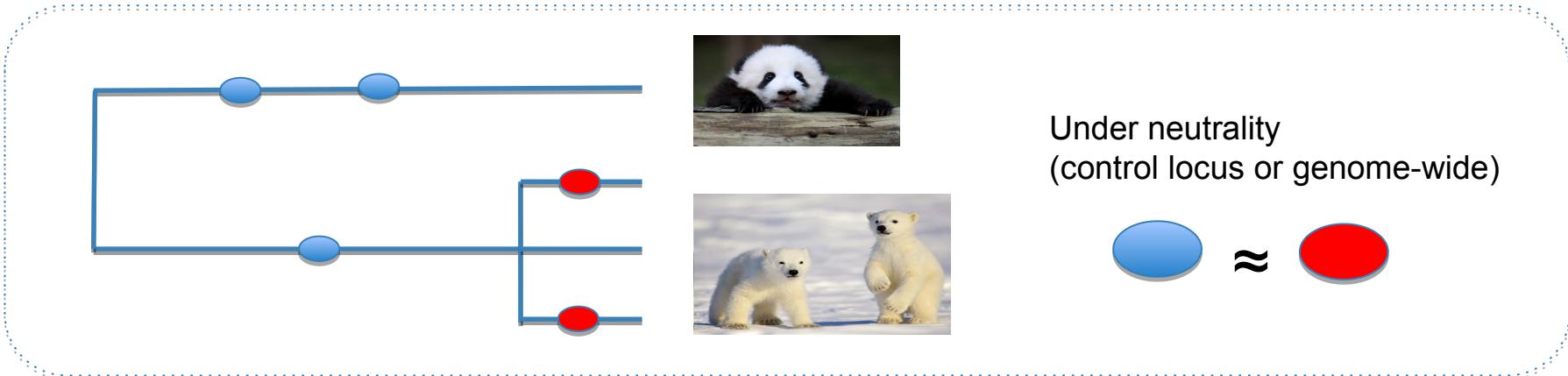
As both depend on mutation rates,
proportional.



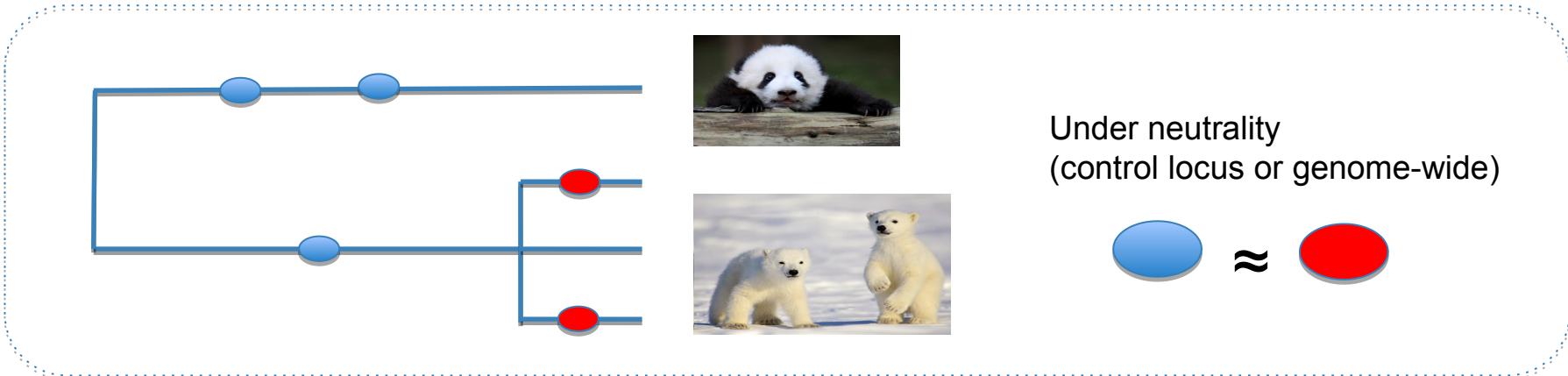
Levels are expected to be



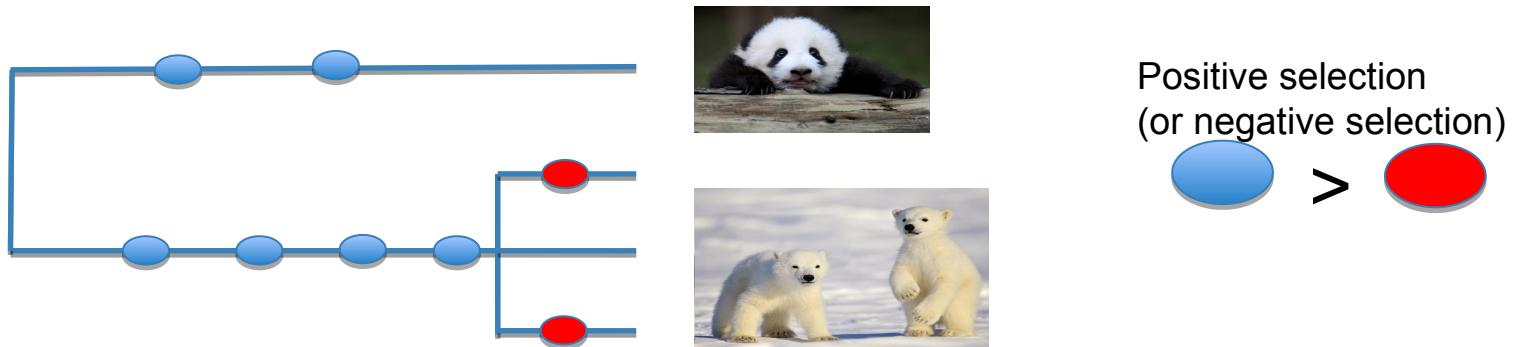
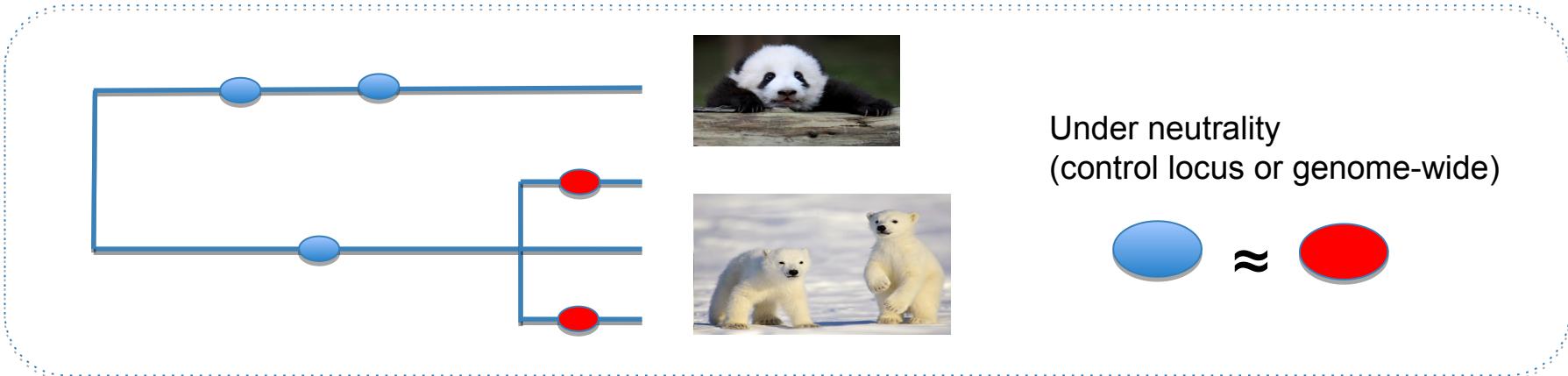
Polymorphisms and divergence



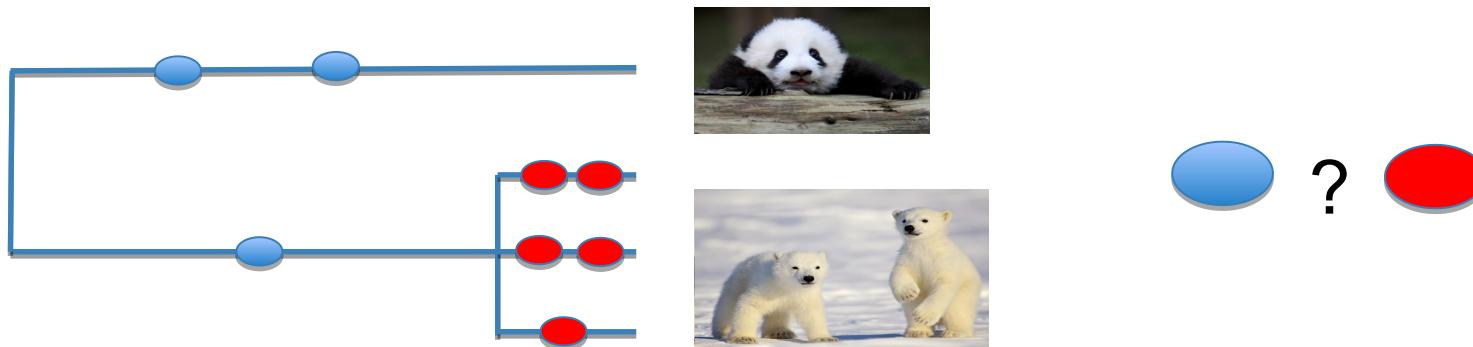
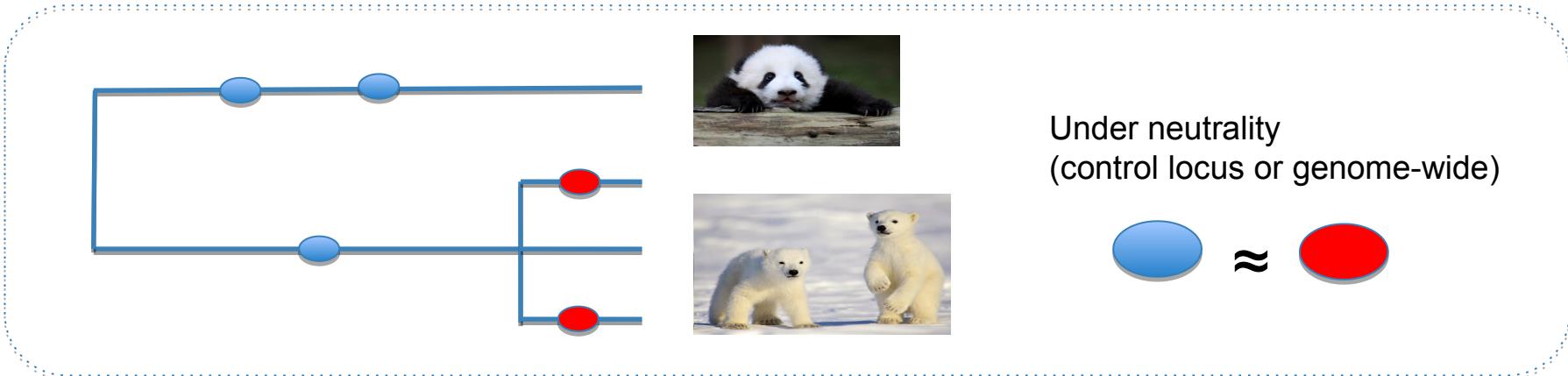
Polymorphisms and divergence



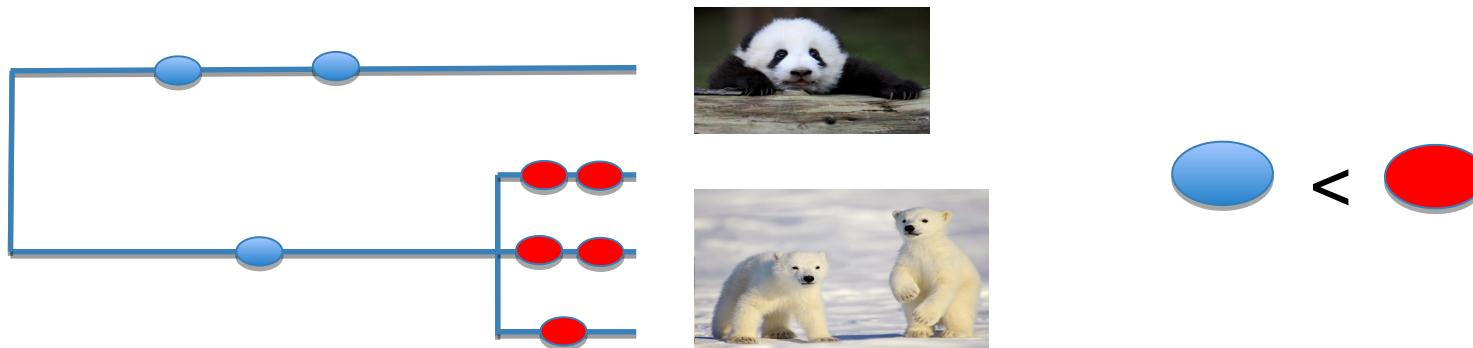
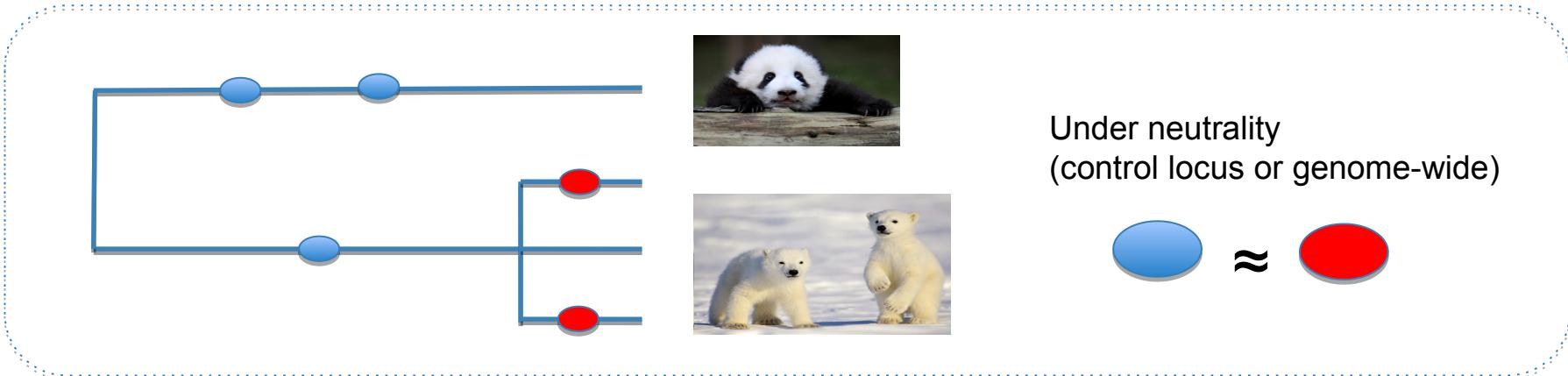
Polymorphisms and divergence



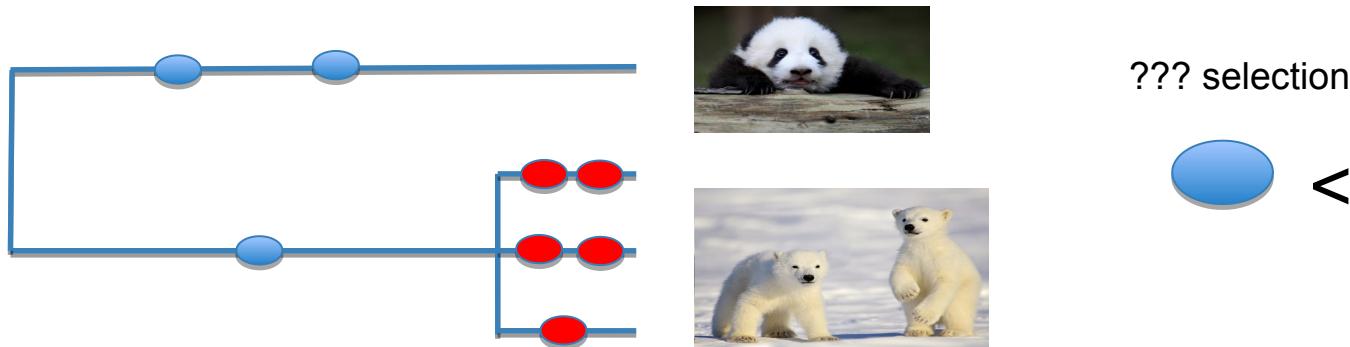
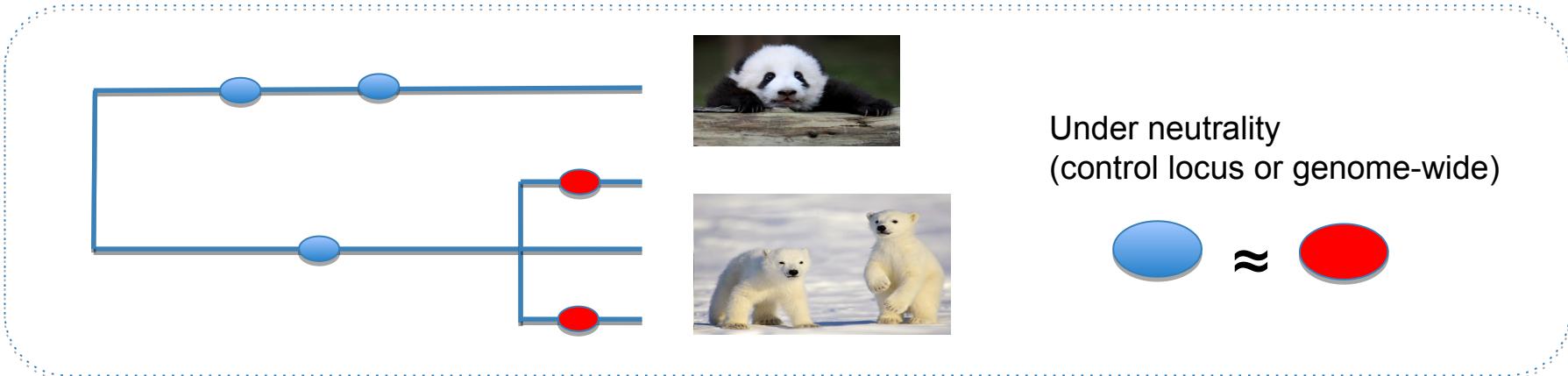
Polymorphisms and divergence



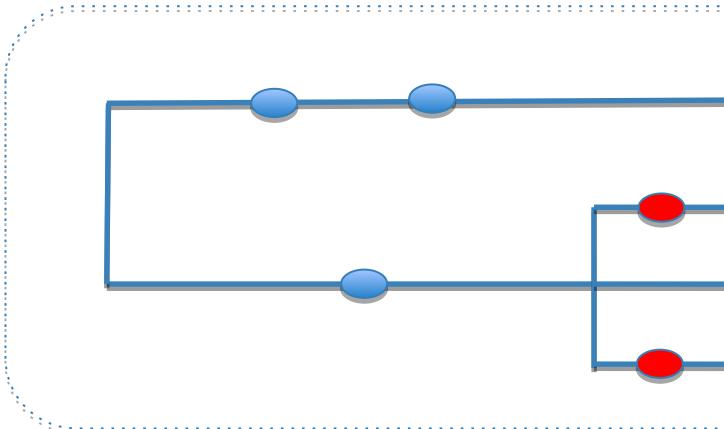
Polymorphisms and divergence



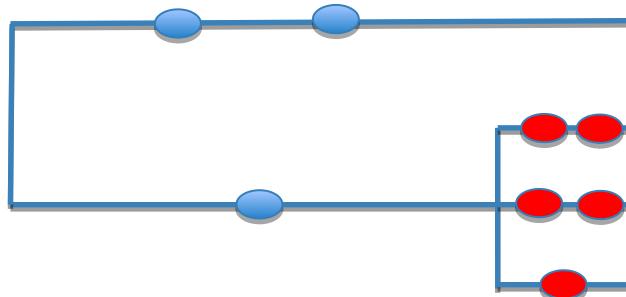
Polymorphisms and divergence



Polymorphisms and divergence



Under neutrality
(control locus or genome-wide)

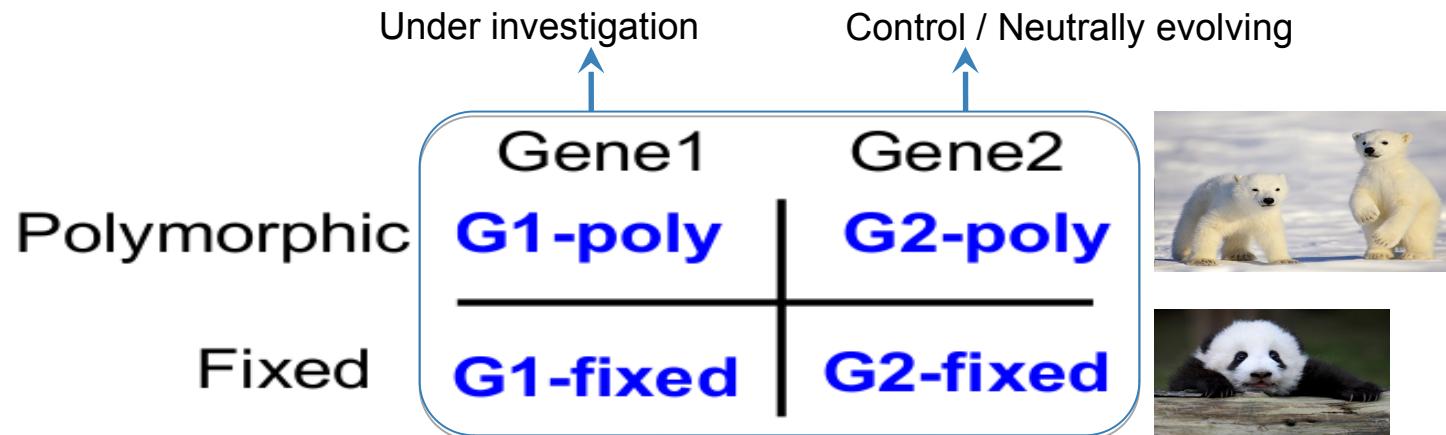


Balancing selection
(or relaxation of functional constraints)



HKA test

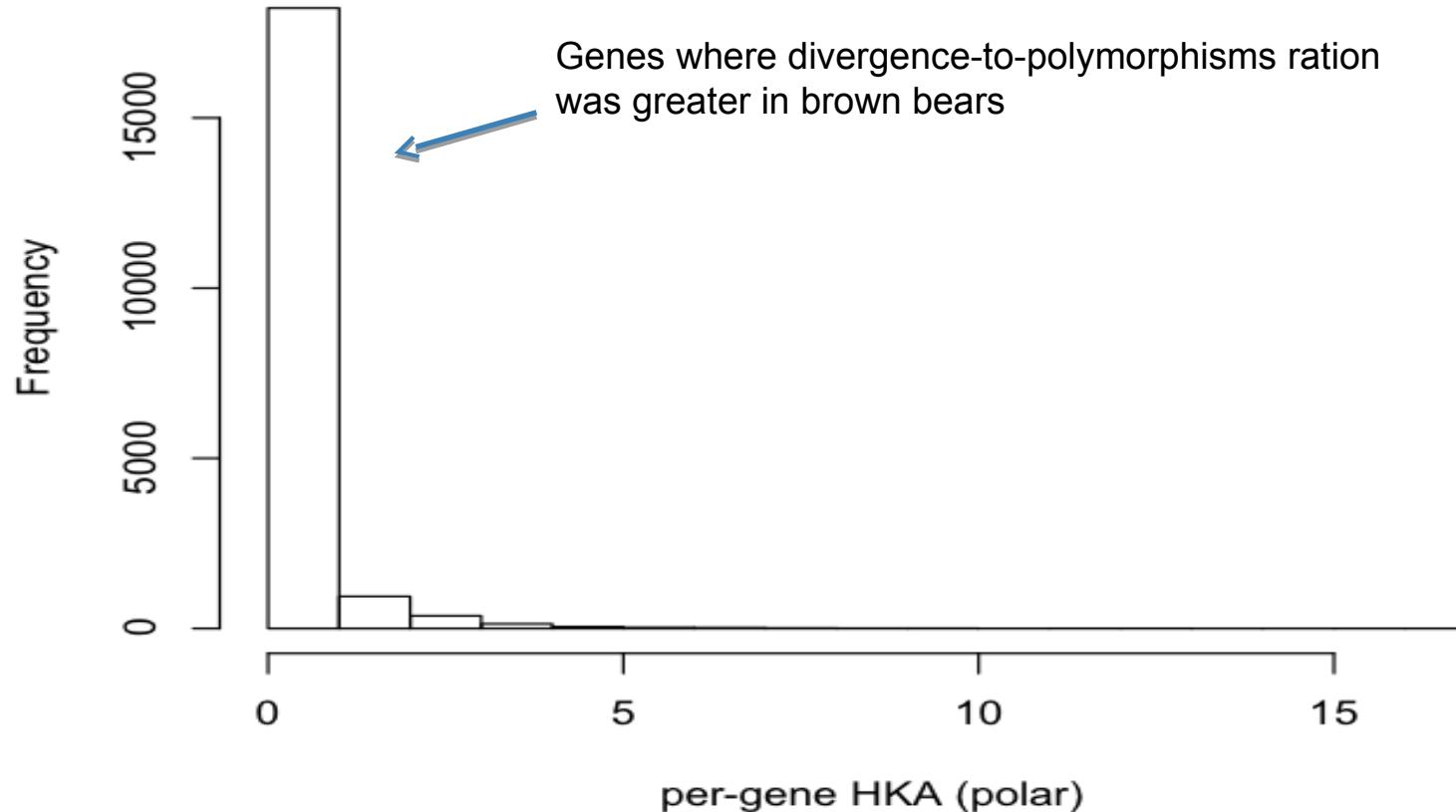
Hudson-Kreitman-Aguadè (HKA, Hudson et al. 1987) test



Contingency table 2x2: chi-square test

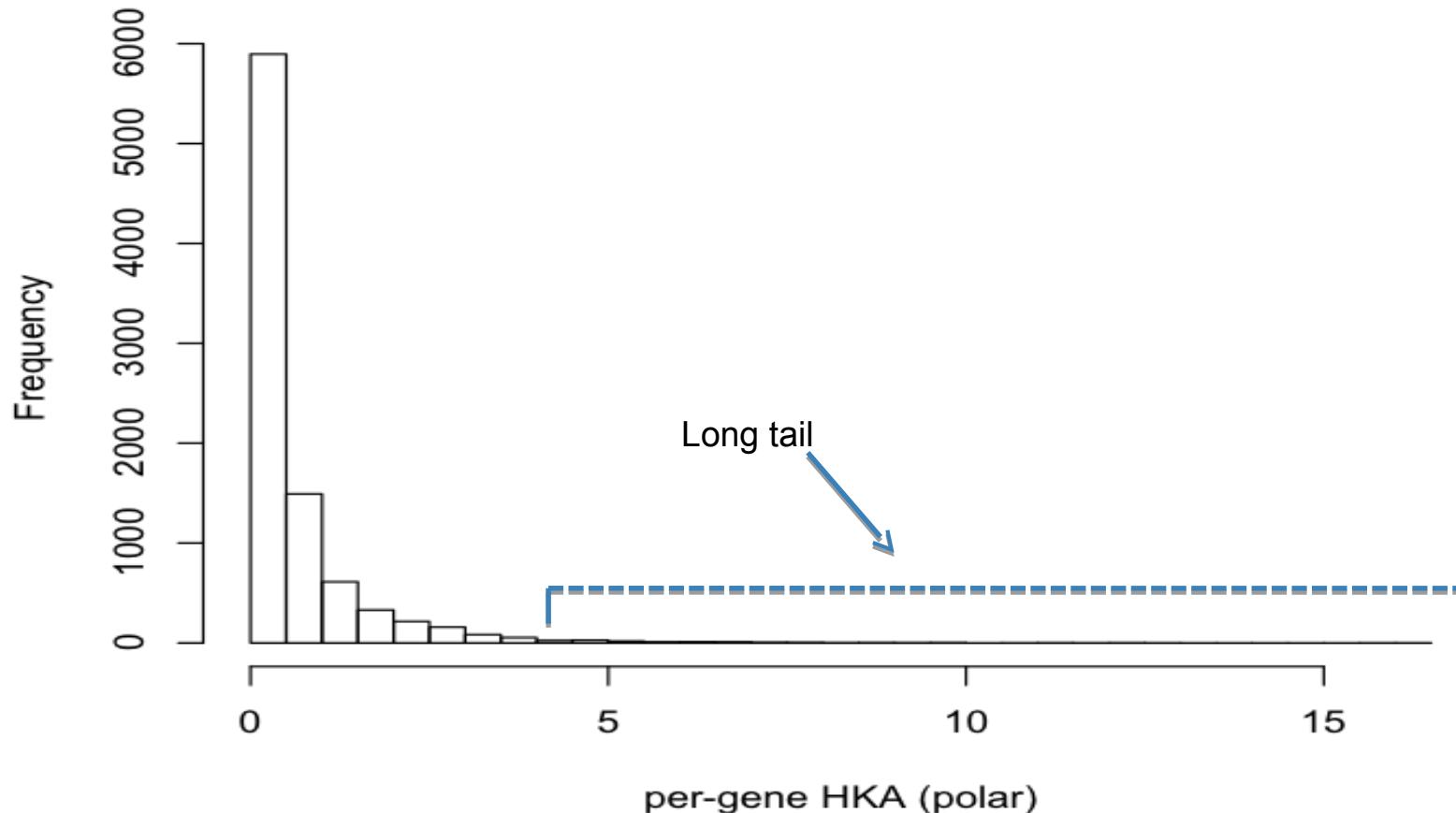
HKA test

Empirical distribution (-log10)



HKA test

Empirical distribution (-log10)



Coding polymorphisms

Substitutions in the coding regions may alter (non-synonymous) or not (synonymous) the protein sequence:

ATC CAA ...
I Q

Ancestral sequence



ATA GAA ...
I K

Species 1



ATG CAG ...
M Q

Species 2



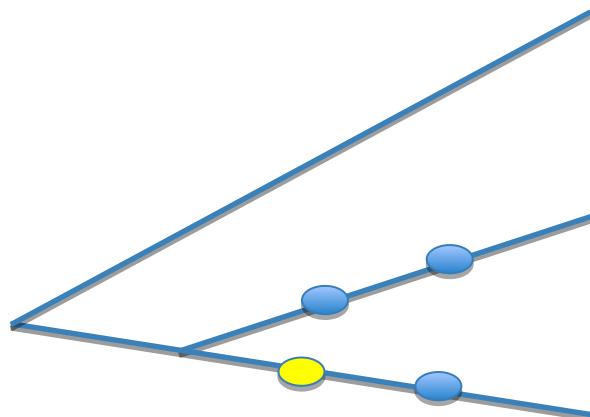
Non-synonymous



Synonymous

dN/ds

Comparison between rate of **nonsynonymous** substitutions and rate of **synonymous** substitutions (also called K_a/K_s or ω)



Non-synonymous



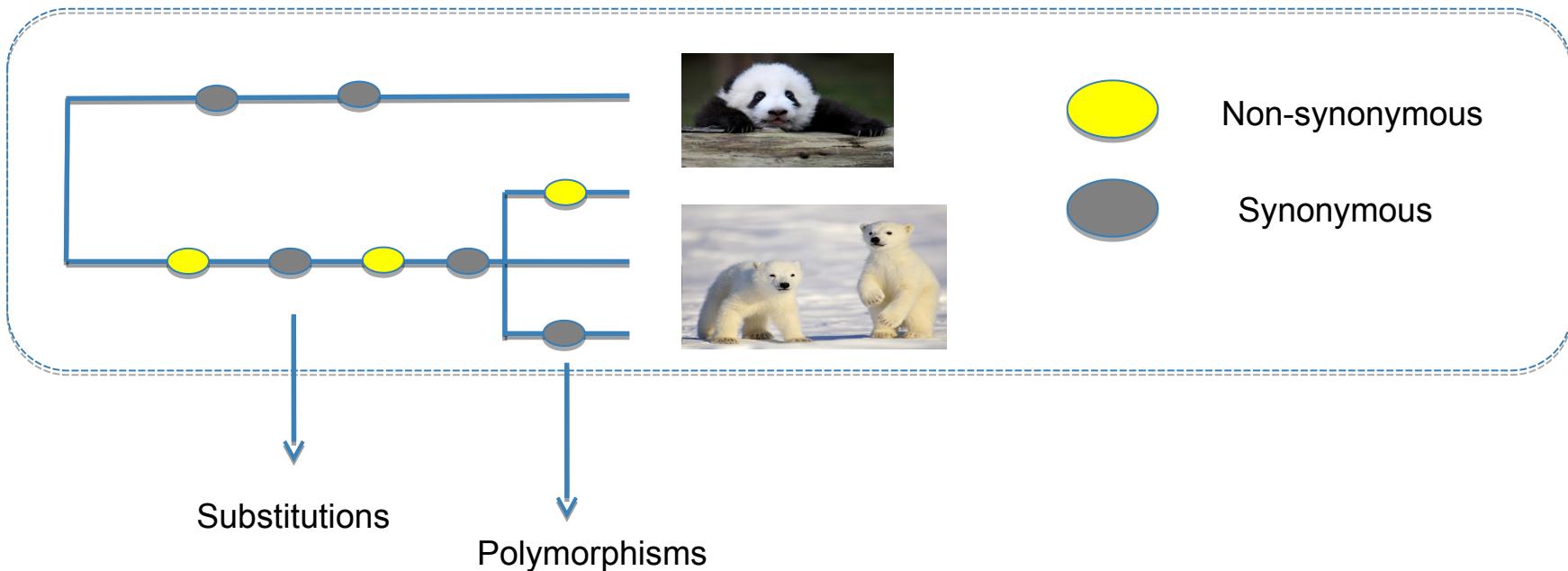
Synonymous

Positive selection
(novel proteins)
 $K_a/K > 1$

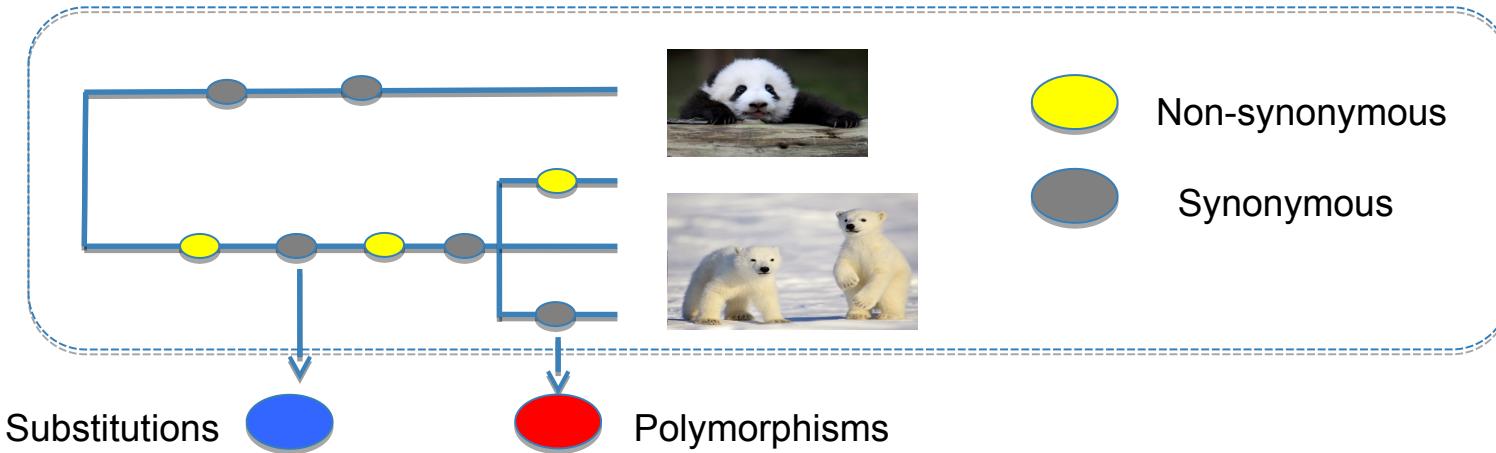
Negative selection
(against deleterious mutations)
 $K_a/K < 1$

Coding polymorphisms and divergence

Compare the amount of variation within a species to the divergence between species at non-synonymous and synonymous variants.

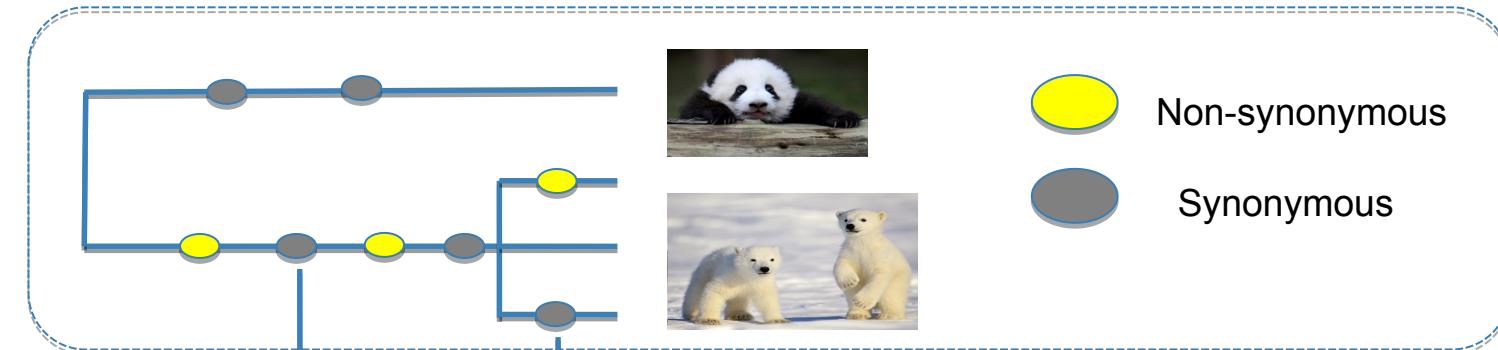


MK (McDonald–Kreitman) test



	Non-synonymous	Synonymous
Divergence		
Polymorphisms		

MK (McDonald–Kreitman) test



Substitutions

Polymorphisms

Divergence

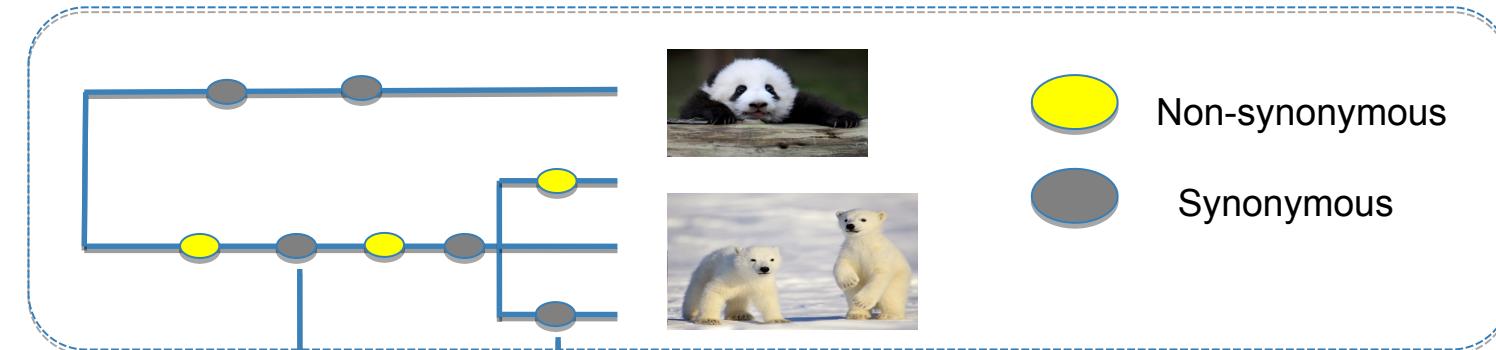
Non-synonymous

Synonymous

Polymorphisms

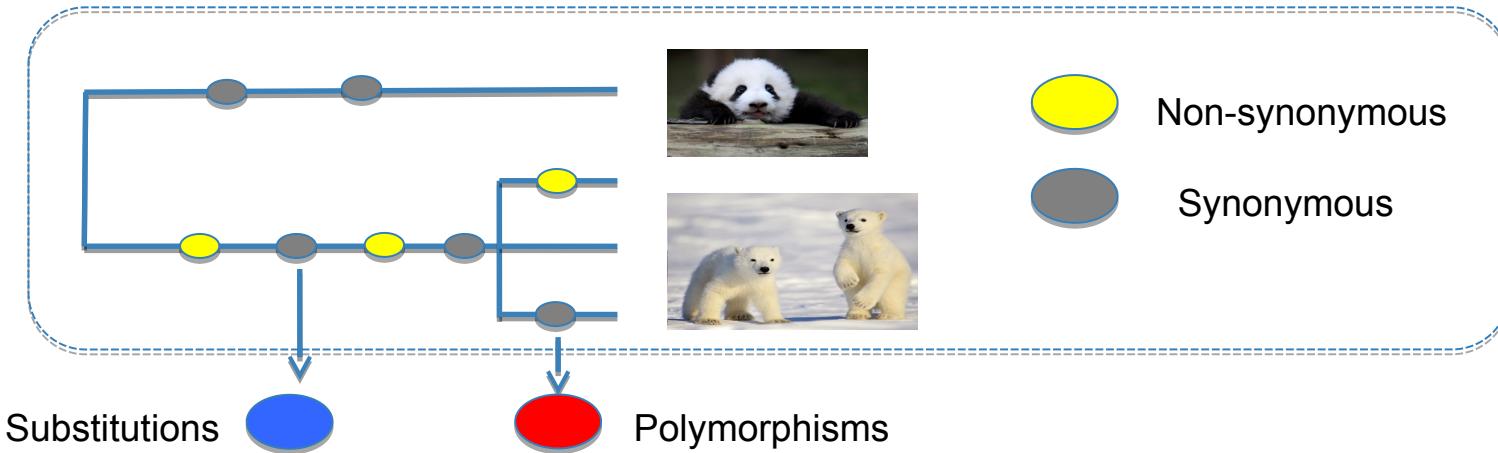
dN/dS between species

MK (McDonald–Kreitman) test



	Non-synonymous	Synonymous
Divergence	Blue circle (Substitution)	Yellow circle (Substitution)
Polymorphisms	Red circle (Polymorphism)	Yellow circle (Polymorphism)
		dN/dS within species

MK (McDonald–Kreitman) test



	Non-synonymous	Synonymous
Divergence	Substitutions (Blue)	Polymorphisms (Red)
Polymorphisms	Substitutions (Yellow)	Polymorphisms (Grey)

2x2 contingency table