Sequences

Part II

How to detect adaptation?

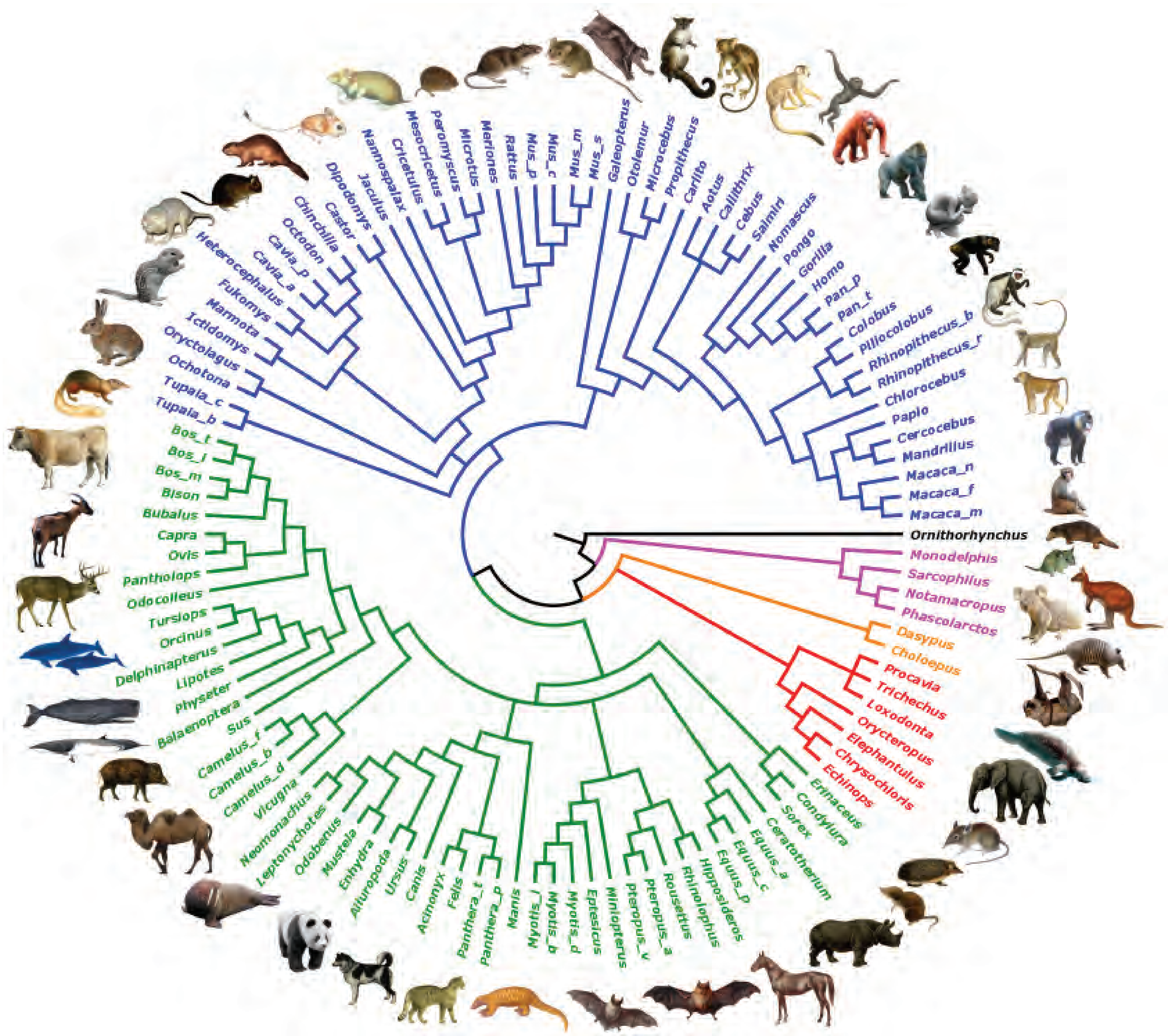
Is adaptation predictable across evolutionary scales?

Is adaptation predictable across evolutionary scales?

We first need to test for adaptation at different scales

Protein coding DNA alignment for 14.509 conserved genes across mammals

Phylogenetic scale



Population scale

Equus caballus (1 population - 6 individuals)

Bos taurus (2 populations - 34 individuals)

Capra hircus/aegagrus (6 populations - 217 individuals)

Ovis aries/orientalis/vignei (5 populations - 278 individuals)

Chlorocebus sabaeus (9 populations - 157 individuals)

Homo sapiens (5 populations - 2504 individuals)

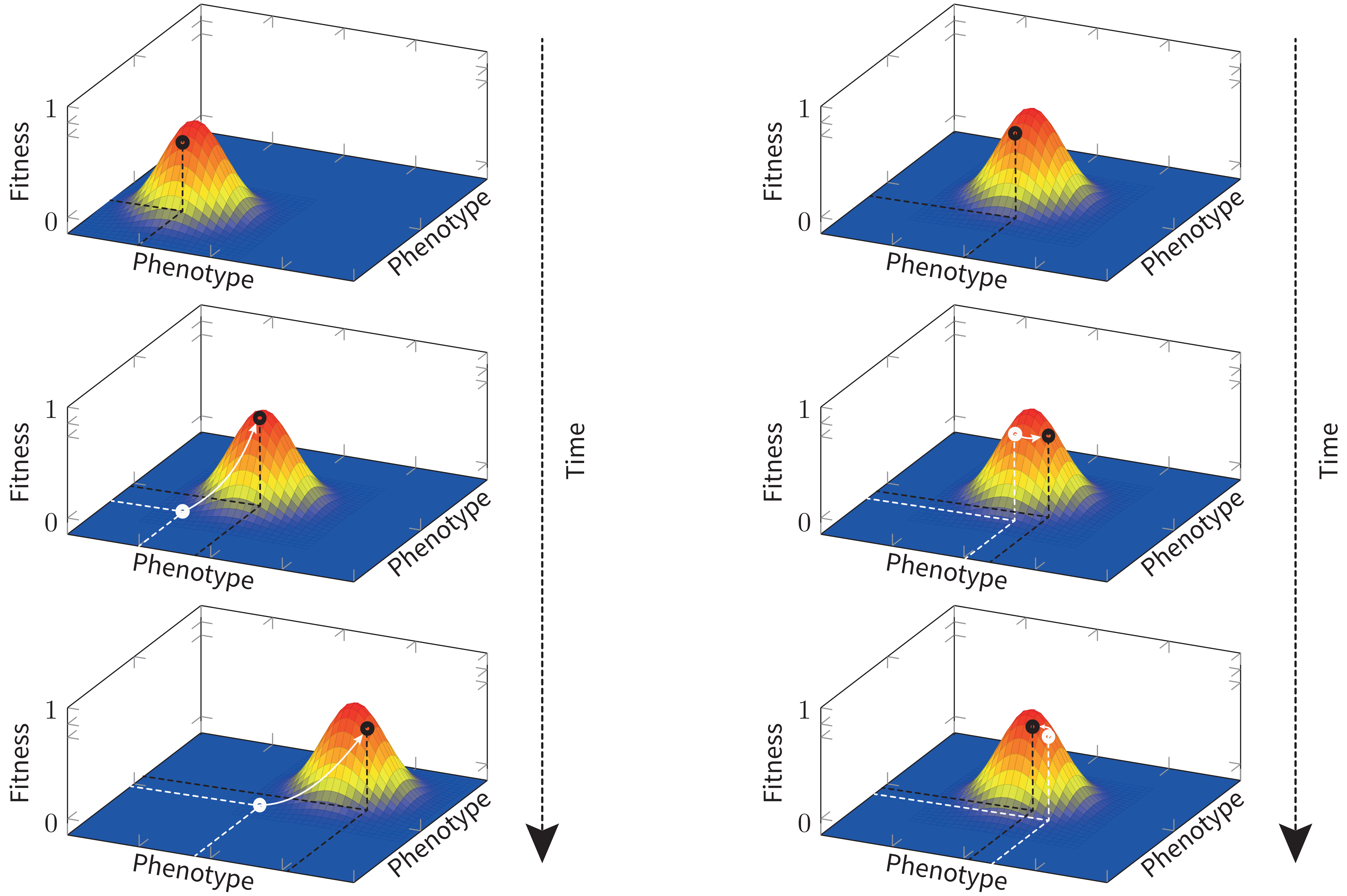
Scornavacca *et al* (2019); Howe *et al* (2021)

What is adaptation?

Adaptation occurs on a changing fitness landscape

Traits

Changing fitness landscape **Stable fitness landscape**



- **The optimal state is a moving target.**
- **Environmental changes that are external.**

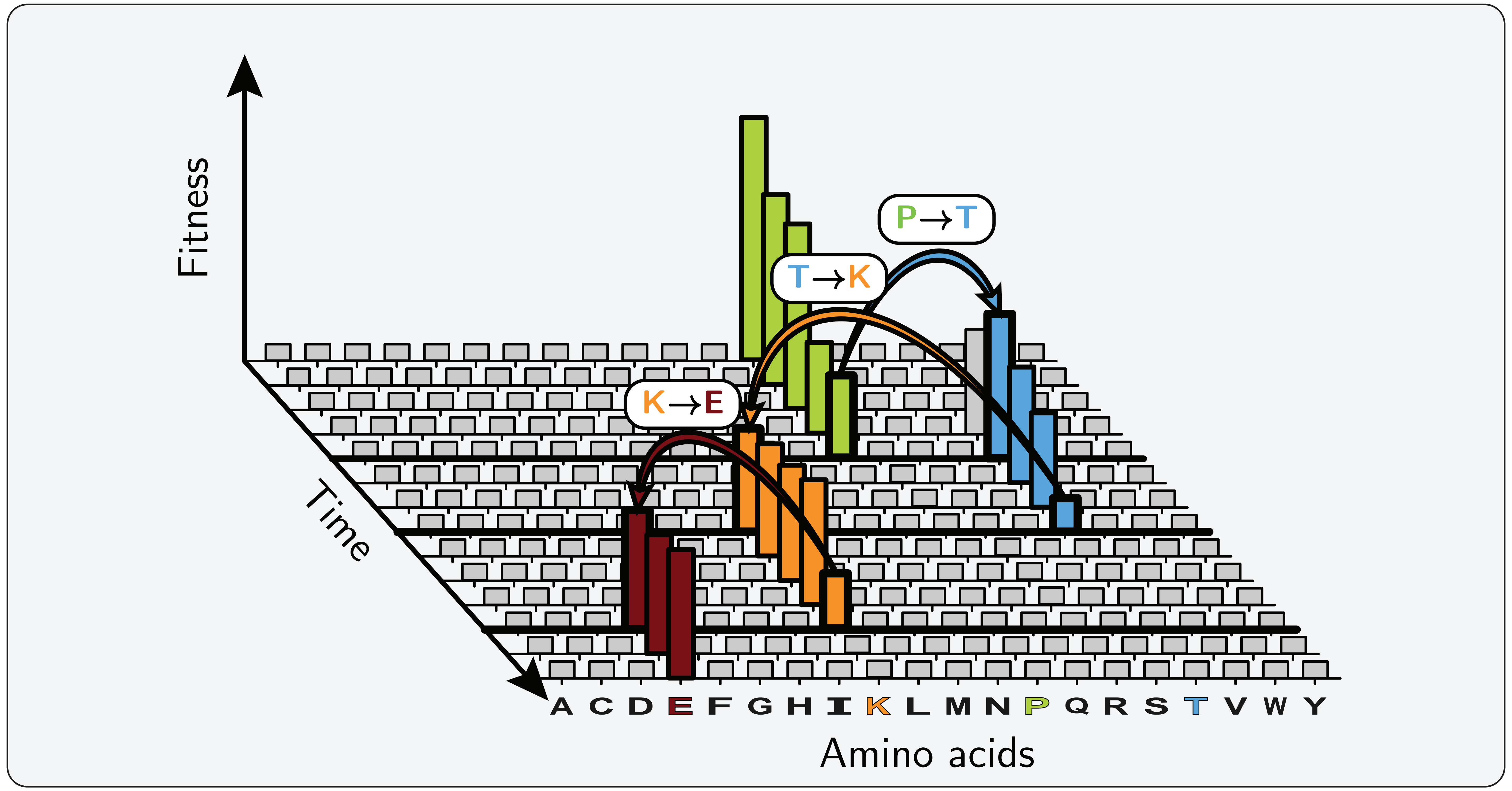
Sella & Hirsh (2005); Mustonen & Lässig (2009)

What is adaptation? (protein coding DNA case)

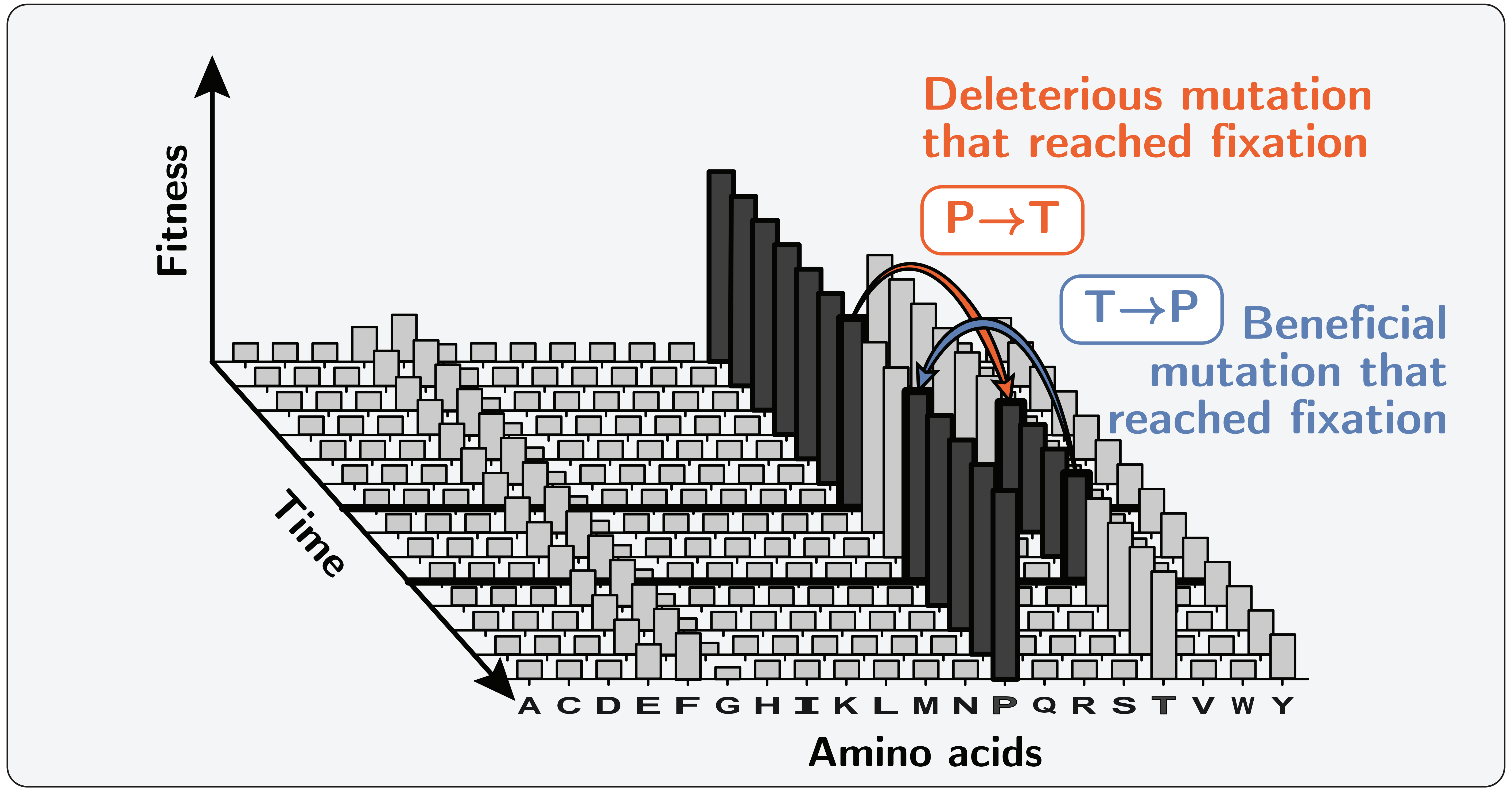
Adaptation occurs on a changing fitness landscape

Sequences

Changing fitness landscape



Stable fitness landscape



- The optimal state is a moving target.
- Environmental changes that are external.

- A stable fitness landscape is a null model of evolution without adaptation.

Sella & Hirsh (2005); Mustonen & Lässig (2009)

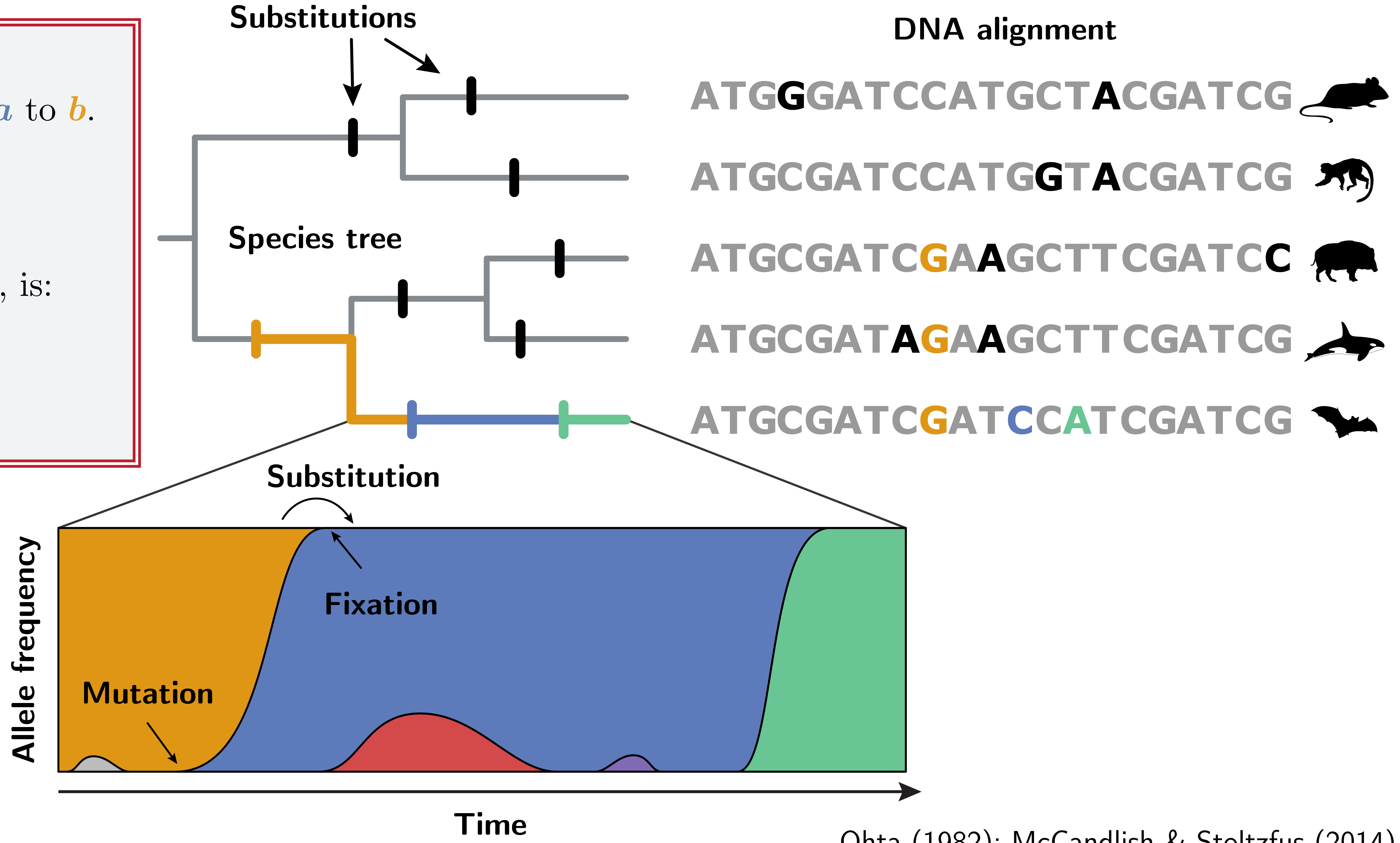
Can we estimate fitnesses from the pattern of substitutions?

Whether a mutation reached fixation depends on its fitness effect.

- $\mu_{a \to b}$ the mutational rate from state a to b.
- F_a the scaled fitness of state a.
- F_b the scaled fitness of state b.

The substitution rate from a to b, $q_{a \to b}$, is:

$$q_{a \to b} = \mu_{a \to b} \frac{F_b - F_a}{1 - e^{F_a - F_b}}.$$



Ohta (1982); McCandlish & Stoltzfus (2014)

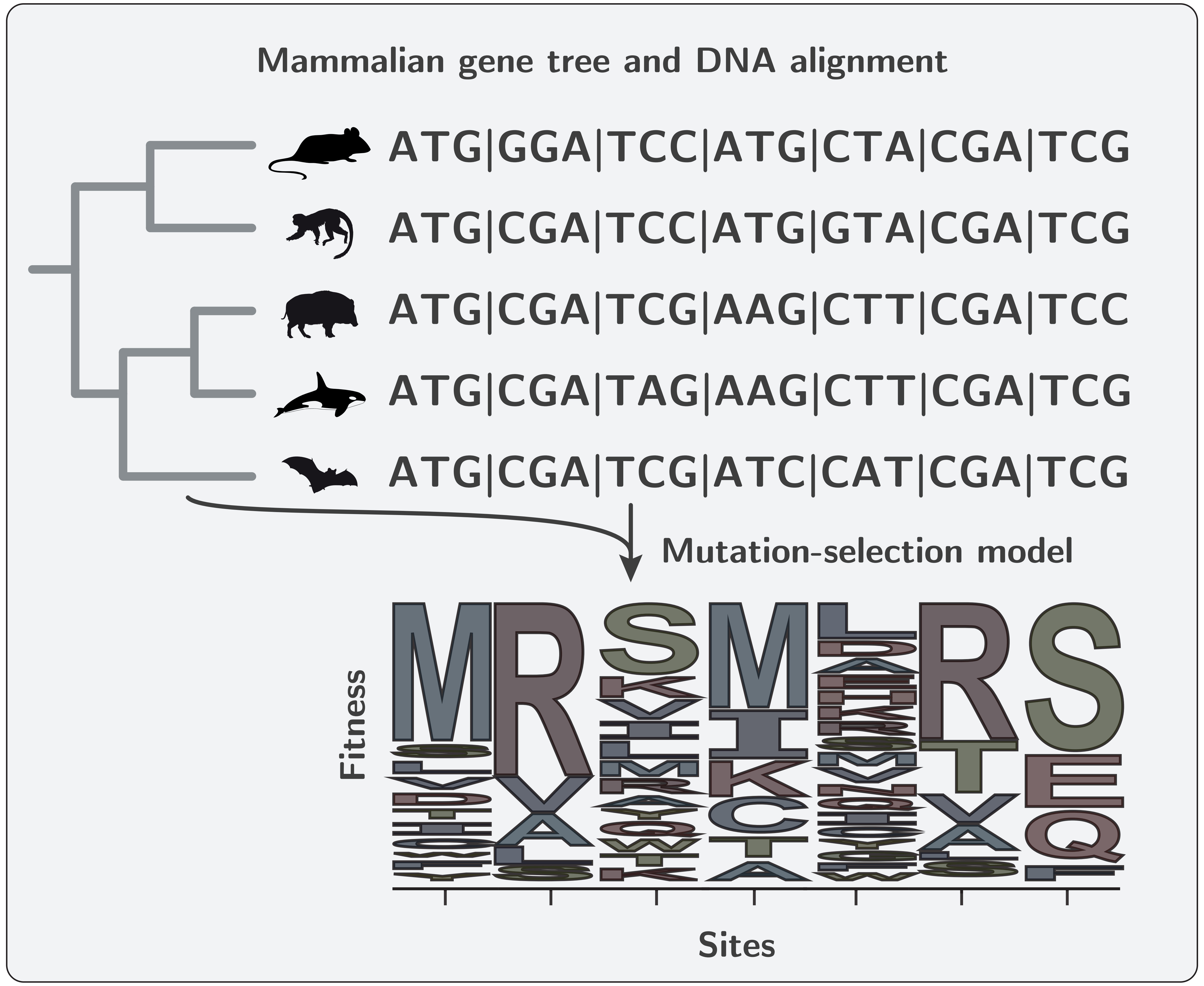
How to estimate the contribution of amino acids to fitness?

By fitting a mutation-selection model at the phylogenetic scale.

Mutation-selection model

- $\mu_{a \to b}$: mutation rate from codon a to b.
- $q_{a \to b}$: substitution rate from codon a to b.
- F_a: scaled fitness of the amino-acid encoded by codon a (F_b for codon b).

$$\begin{cases} q_{a \to b} = \mu_{a \to b} & \text{if synonymous,} \\ q_{a \to b} = \mu_{a \to b} \times \dfrac{F_b - F_a}{1 - \mathrm{e}^{F_a - F_b}} & \text{if non-synonymous.} \end{cases}$$


Mammalian gene tree and DNA alignment

ATG|GGA|TCC|ATG|CTA|CGA|TCG
ATG|CGA|TCC|ATG|GTA|CGA|TCG
ATG|CGA|TCG|AAG|CTT|CGA|TCC
ATG|CGA|TAG|AAG|CTT|CGA|TCG
ATG|CGA|TCG|ATC|CAT|CGA|TCG

- **Input: alignment of protein-coding DNA sequences and phylogenetic tree.**
- **Output: amino-acid fitness profiles estimated by mutation-selection models.**

Halpern & Bruno (1998); Tamuri & Goldstein (2012); Rodrigue & Lartillot (2017); Rodrigue *et al* (2021)

Can we predict the rate of protein evolution while assuming no adaptation?

We first fit a stable first landscape then predict the rate of protein evolution.

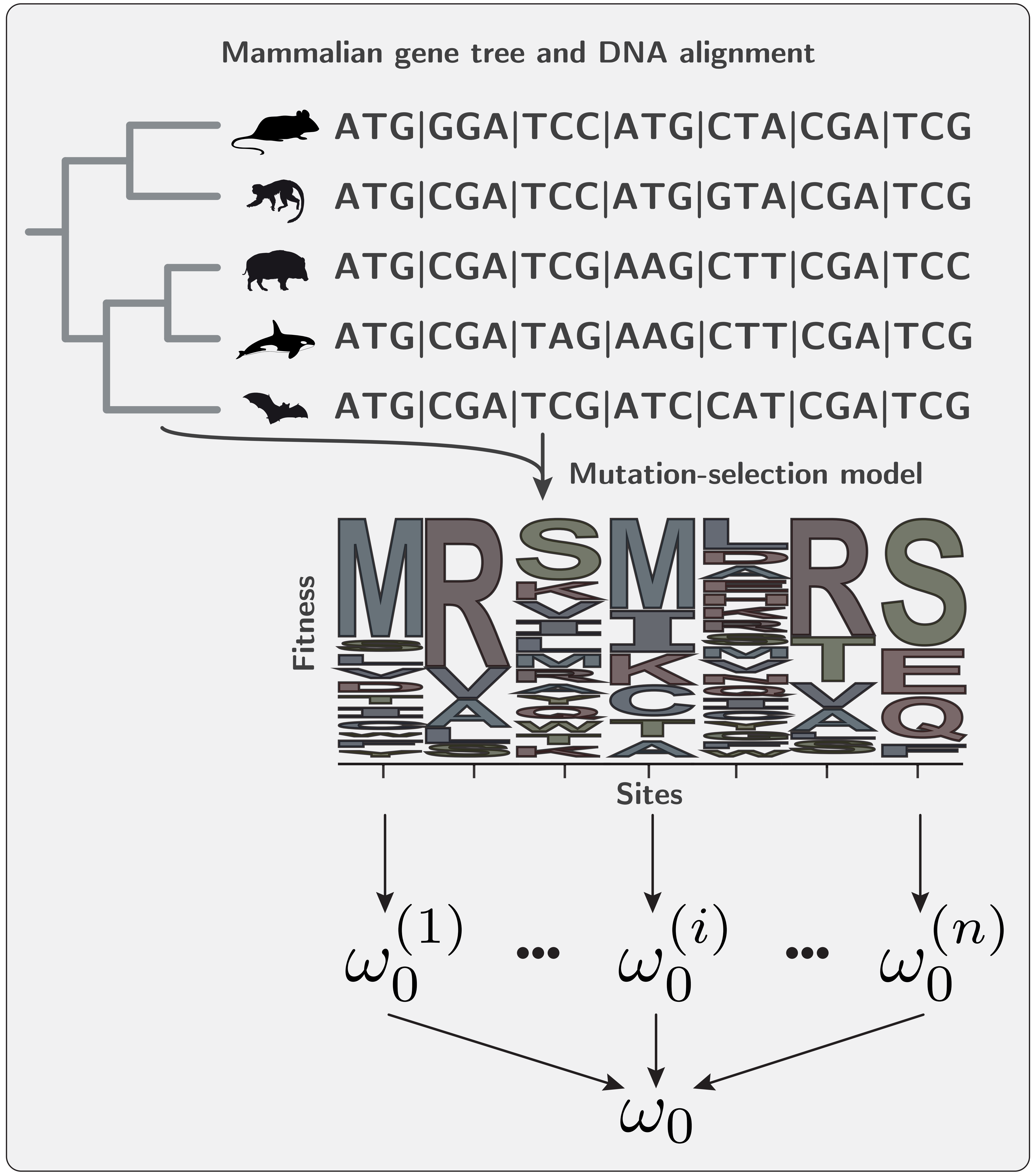
Mutation-selection model

- $\mu_{a \to b}$: mutation rate from codon a to b.
- $q_{a \to b}^{(i)}$: substitution rate from codon a to b at site i.
- $\pi_a^{(i)}$: equilibrium frequency of codon a at site i.

$$\omega_0^{(i)} = \frac{\left\langle \pi_a^{(i)} q_{a \to b}^{(i)} \right\rangle}{\left\langle \pi_a^{(i)} \mu_{a \to b} \right\rangle},$$

$$\Rightarrow \omega_0 = \frac{1}{n} \sum_{i=1}^{n} \omega_0^{(i)}.$$

- $\langle \cdot \rangle$ is the average over all pairs of non-synonymous codons.
- n: number of codon sites in the DNA alignment.

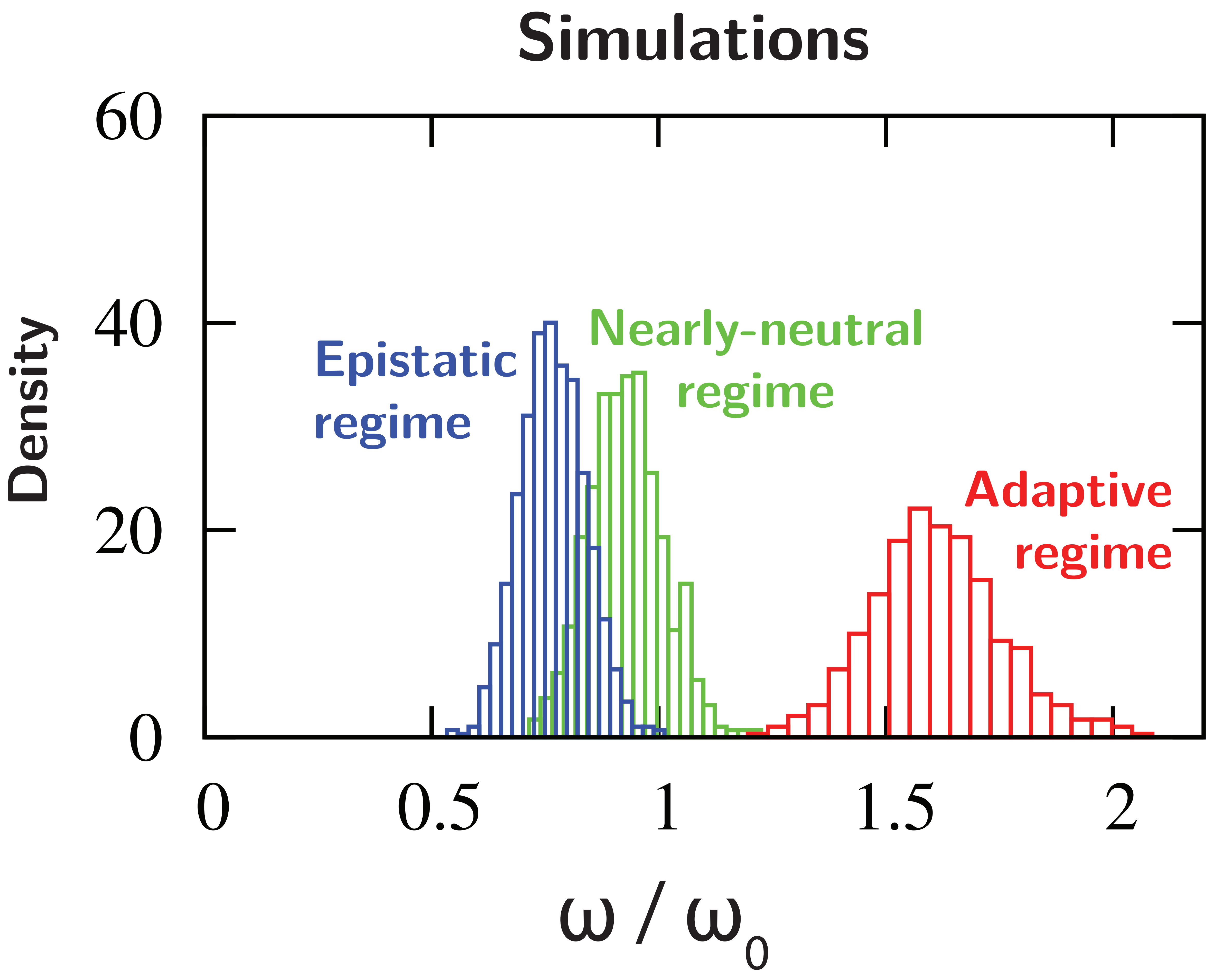


Spielman & Wilke (2015); Dos Reis (2015); Rodrigue & Lartillot (2017)

How to use the mutation-selection model to detect adaptation?

Contrasting ω and ω_0 to detect a changing fitness landscape.

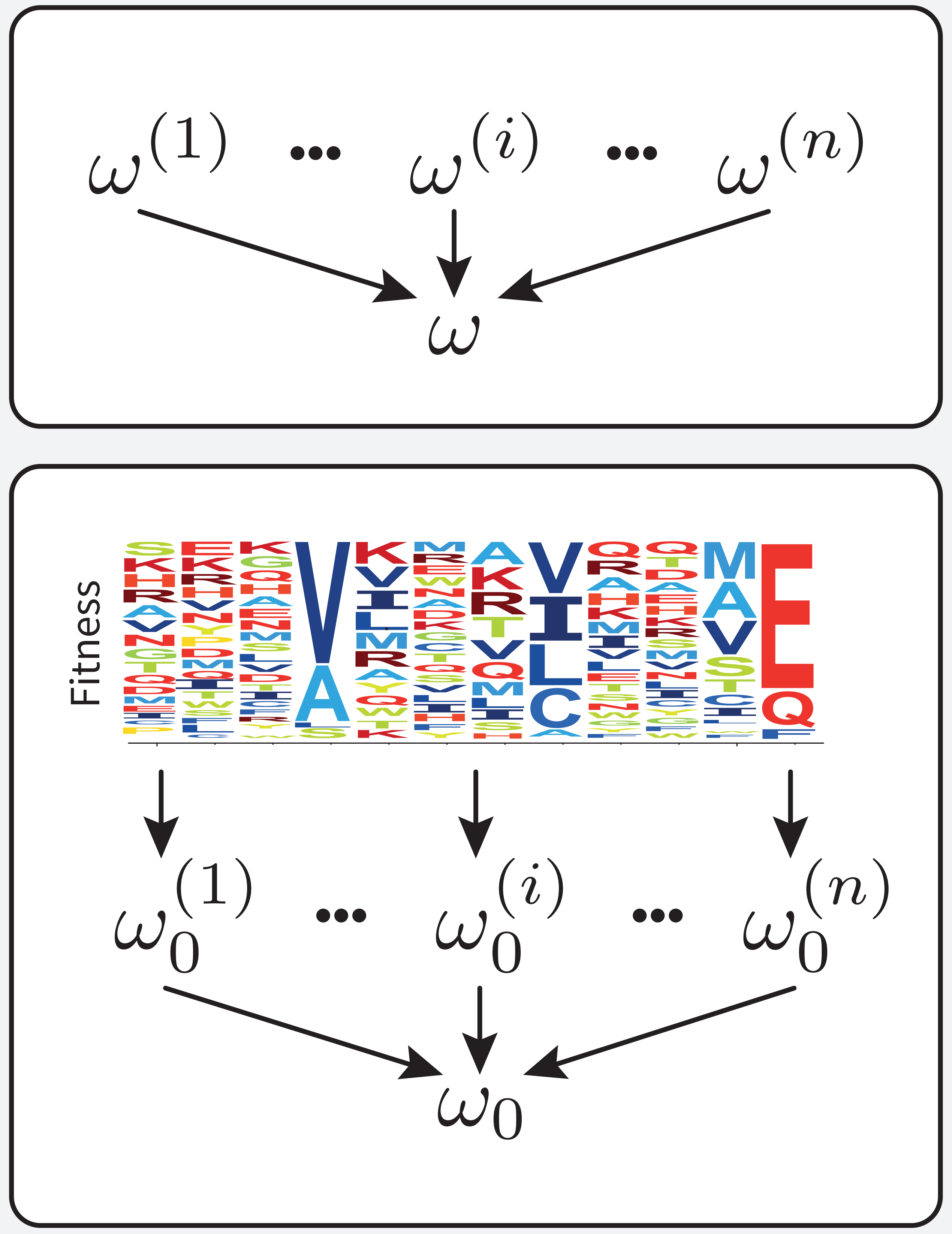
- ω: estimated rate of evolution under classical codon model.
- ω_0: predicted rate of evolution under the mutation-selection model.



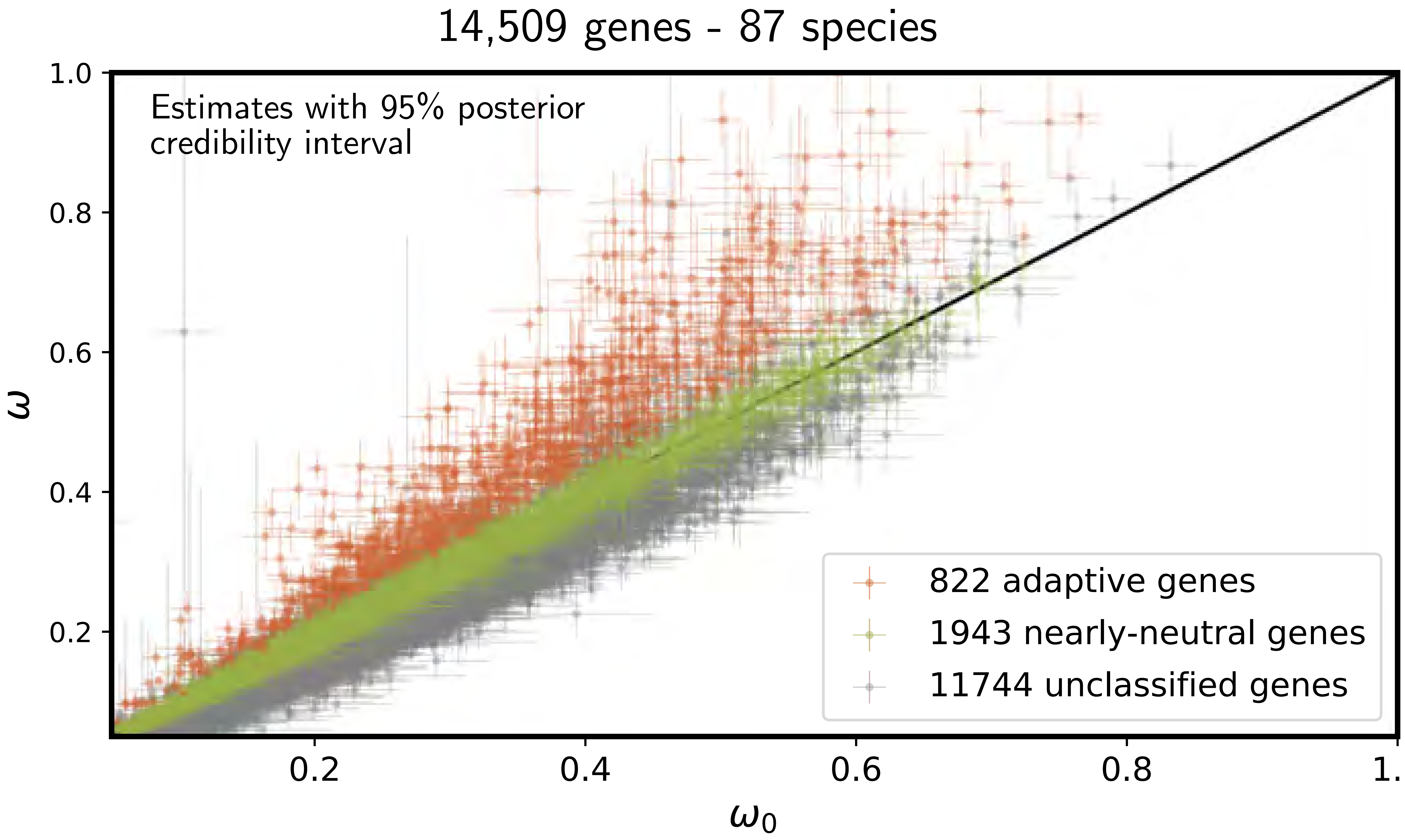
Rodrigue & Lartillot (2017)

Can the mutation-selection model detect adaptation?

We can detect genes with $\omega>\omega_0$ across mammals while they still have $\omega<1$.



Phylogenetic codon models on protein-coding DNA alignments

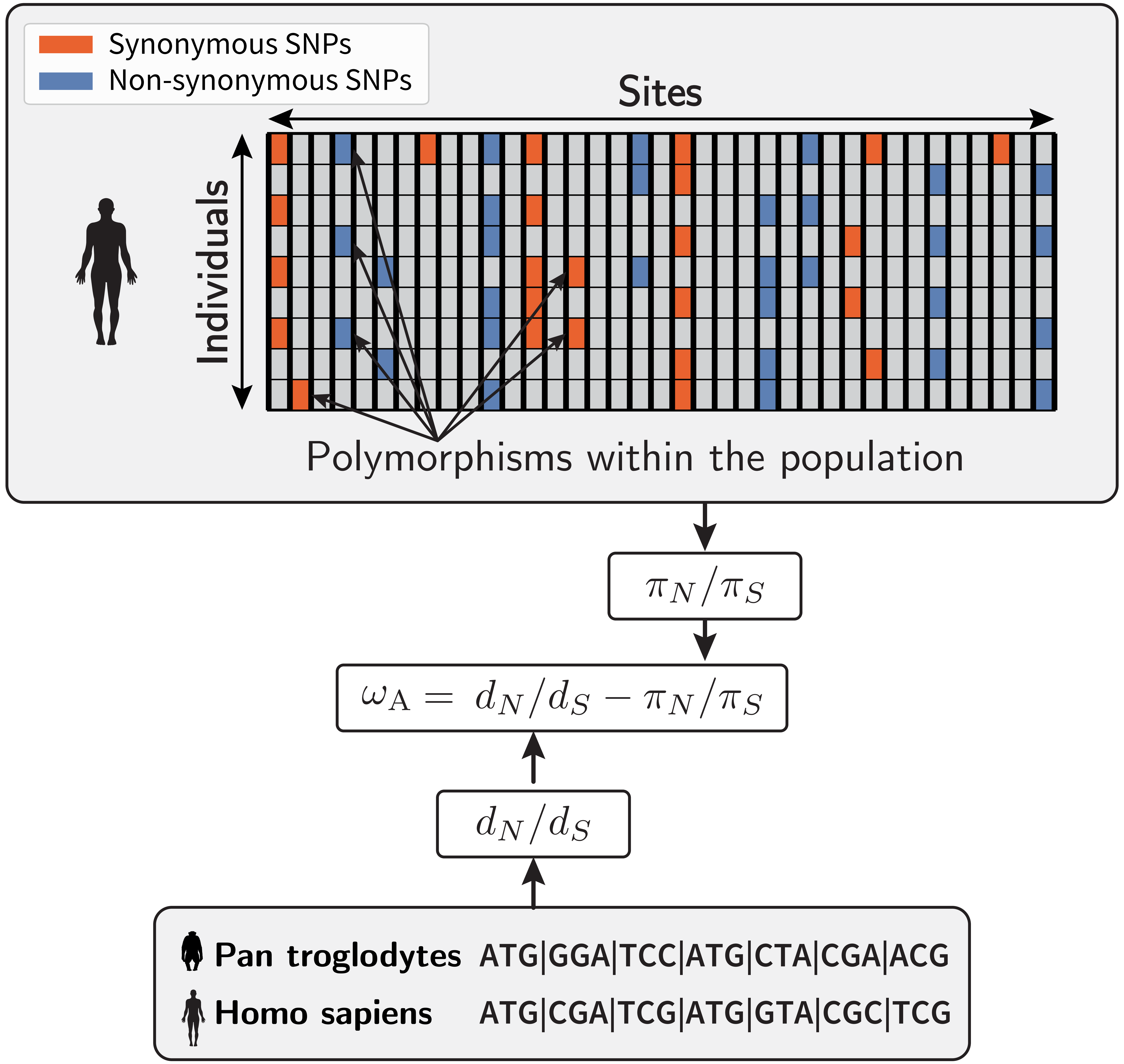


- **Genes predicted to be under adaptation at the phylogenetic scale are enriched in ontologies related to immunity, response to virus and external membrane.**

Latrille et al. (2023)

How to test for adaptation in a lineage?

Contrasting substitutions to polymorphism with the McDonald & Kreitman test.



🟥 Synonymous SNPs
🟦 Non-synonymous SNPs

Sites

Individuals

Polymorphisms within the population

π_N/π_S

$\omega_{\mathrm{A}} = d_N/d_S - \pi_N/\pi_S$

d_N/d_S

Pan troglodytes ATG|GGA|TCC|ATG|CTA|CGA|ACG
Homo sapiens ATG|CGA|TCG|ATG|GTA|CGC|TCG

McDonald & Kreitman (1991), Messer & Petrov (2013), Tataru et al. (2017)

Is adaptation at different evolutionary scale comparable?

Adaptation at the phylogenetic scale predicts adaptation in a terminal lineage.

14,509 genes - 87 species

$\omega^{(1)} \cdots \omega^{(i)} \cdots \omega^{(n)} \rightarrow \omega$

$\omega_0^{(1)} \cdots \omega_0^{(i)} \cdots \omega_0^{(n)} \rightarrow \omega_0$

Phylogenetic codon models on protein-coding DNA alignments

$\omega > \omega_0$

Genes classified as adaptive

Concatenate genes classified as adaptive

$\omega \simeq \omega_0$

Genes classified as nearly-neutral

Subsampling genes classified as nearly-neutral and concatenate

Polymorphisms within the population

π_N/π_S

$\omega_{\mathrm{A}} = d_N/d_S - \pi_N/\pi_S$

d_N/d_S

Pan troglodytes ATG|GGA|TCC|ATG|CTA|CGA|ACG

Homo sapiens ATG|CGA|TCG|ATG|GTA|CGC|TCG

Population-genetics rate of adaptive evolution

- **Genes predicted to be under adaptation at the phylogenetic scale are under adaptation at the population-genetic scale.**

Latrille et al. (2023)

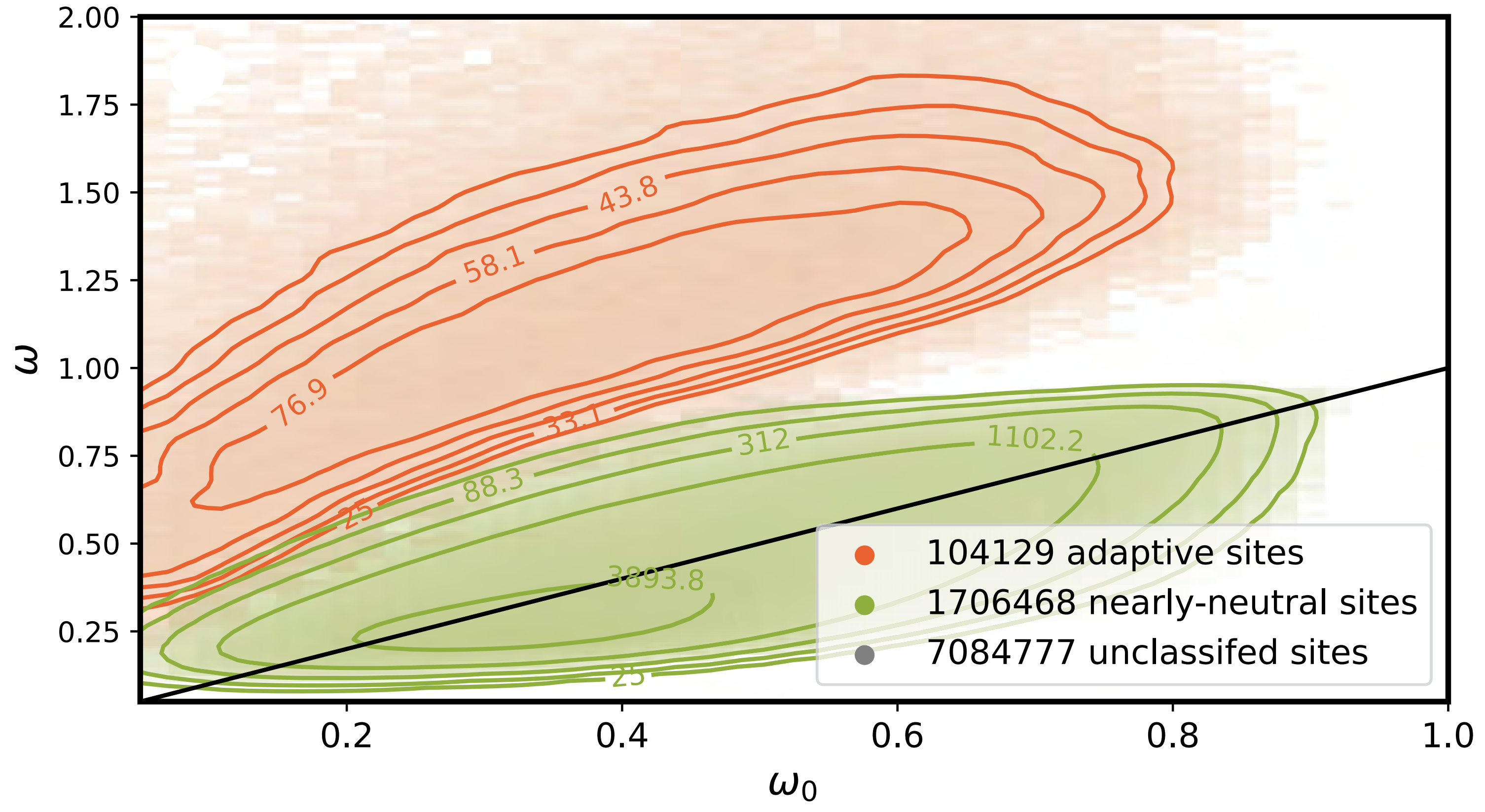
Are the results replicable?

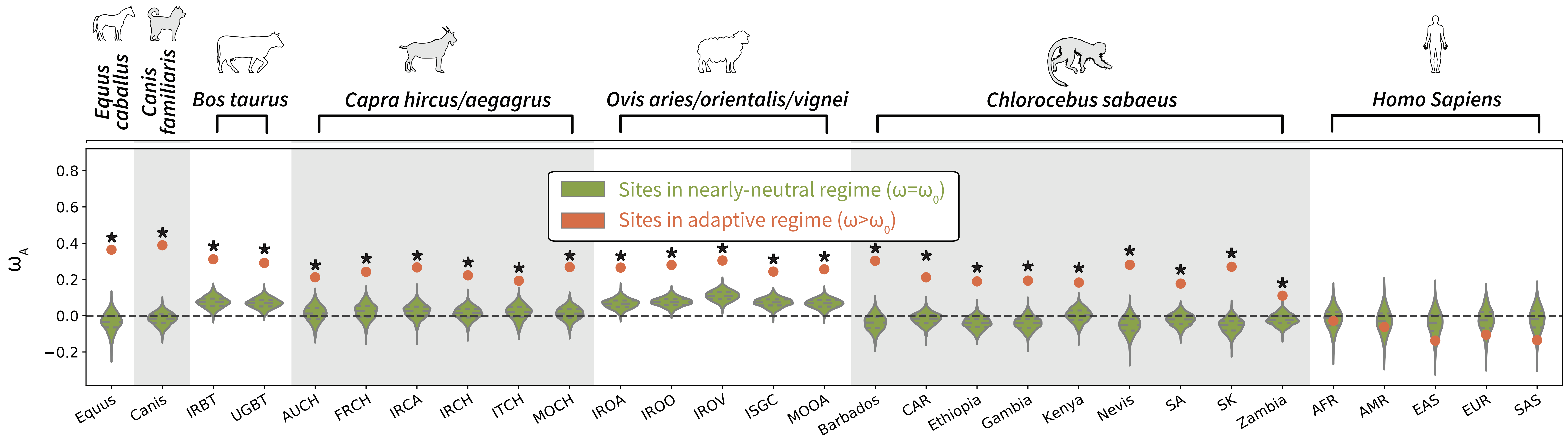
Adaptation at the phylogenetic scale predicts adaptation in terminal lineages.

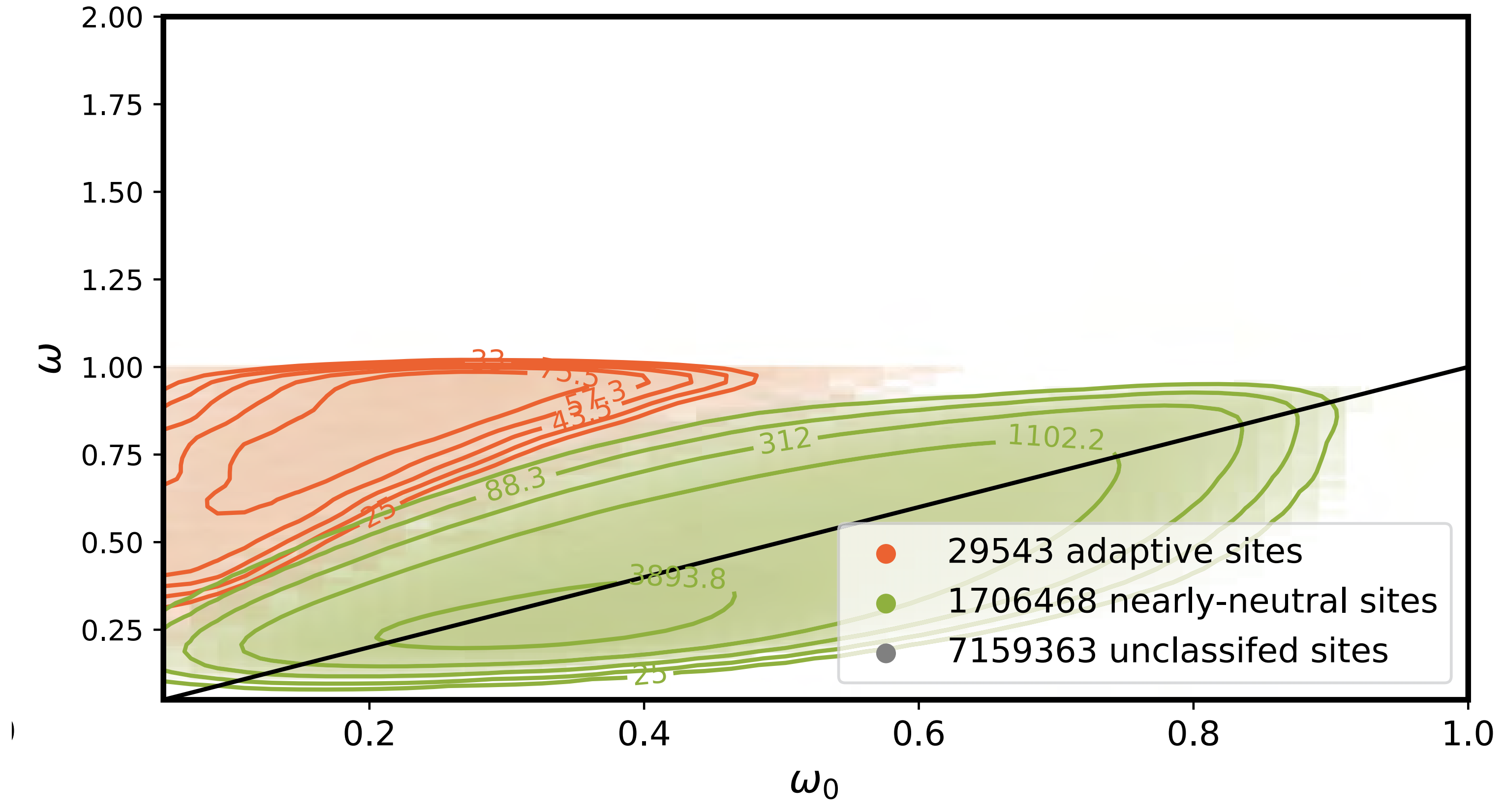
Latrille et al. (2023)

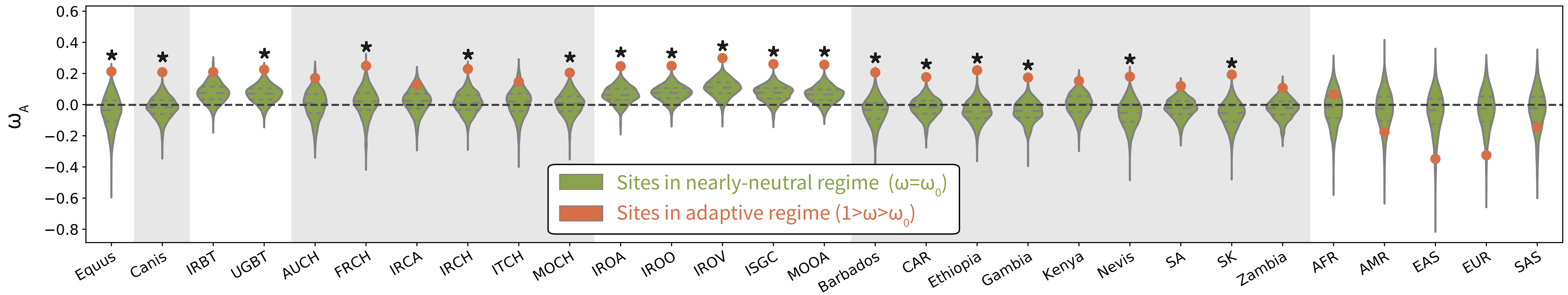
Are the results generalisable to sites instead of sites?

Sites under adaptation across mammals are under adaptation in terminal lineages.





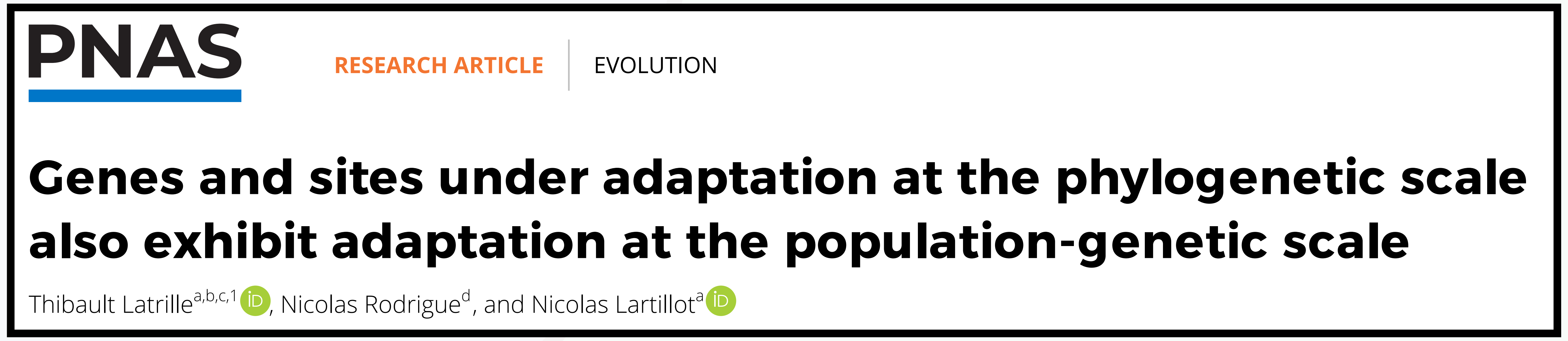




Latrille et al. (2023)

Part II
How to detect adaptation?
Is adaptation predictable across evolutionary scales?

- A stable fitness lanscape is a null model of evolution.
- Adaptation as deviation from this null model.
- Adaptation at the phylogenetic scale predicts adaptation in terminal lineages and populations.



PNAS RESEARCH ARTICLE | EVOLUTION

Genes and sites under adaptation at the phylogenetic scale also exhibit adaptation at the population-genetic scale

Thibault Latrille[a,b,c,1], Nicolas Rodrigue[d], and Nicolas Lartillot[a]