Mammalian protein-coding genes exhibit widespread beneficial mutations that are not adaptive

T. Latrille[1†], **J. Joseph**[2†], **D. A. Hartasánchez**[1], **N. Salamin**[1]

[1]Department of Computational Biology, Université de Lausanne, Lausanne, Switzerland

[2]Laboratoire de Biométrie et Biologie Evolutive, UMR5558, Université Lyon 1, Villeurbanne, France

[†]These authors contributed equally to this work

thibault.latrille@ens-lyon.org

May 3, 2023

Abstract

Mutations can be beneficial by bringing innovation to their bearer, allowing them to adapt to environmental change. However, mutations can also be beneficial because they are reverting previous deleterious changes, simply restoring pre-existing functions. By integrating phylogenomic and population data in mammals, we estimated the contribution of beneficial back-mutations to molecular evolution, which had remained widely overlooked so far. Applying a mutation-selection model, we estimated amino acid fitness landscapes for mammalian protein-coding genes. Assuming these fitness landscapes do not change, any beneficial mutation is necessarily a back-mutation. We confirmed that these beneficial back-mutations are positively selected in extant populations and demonstrated that a substantial part of ongoing positive selection is not driven by adaptation to environmental change.

***K*eywords** Adaptation · Back-mutations · Phylogenetics · Population-genetics · Codon models

Adaptation is one of the main processes shaping the diversity of forms and functions across the tree of life [1]. Evolutionary adaptation is tightly linked to environmental change and species responding to this change [2, 3]. For adaptation to occur, there must be variation within populations, which mostly appears via mutations in the DNA sequence. While neutral mutations will not impact an individual's fitness, deleterious mutations have a negative effect, and beneficial mutations improve their bearer's fitness. A beneficial mutation is thus more likely than a neutral mutation to invade the population and reach fixation, resulting in a substitution at the species level. Upon environmental change, because adaptive beneficial mutations toward new fitness optima are more likely, the number of substitutions also increases (fig. 1A). An increased substitution rate is thus commonly interpreted as a sign of adaptation [4–6]. The availability of large-scale genomic data and the development of theoretical models have enabled the detection and quantification of substitution rate changes across genes and lineages [7–9]. These approaches, now common practice in evolutionary biology, helped better understand the processes underpinning the rates of molecular evolution.

However, a collateral effect has been conflating beneficial mutations with adaptive evolution when adaptive evolution is not the only process that can lead to beneficial mutations [10, 11].

In a constant environment, a deleterious mutation can reach fixation by genetic drift [12]. A new mutation restoring the ancestral fitness will thus be beneficial (fig. 1B), even though the environment has not changed [11, 13–15]. The restoration of the ancestral fitness can either happen through a mutation at a different locus – called a compensatory mutation [11, 13], or at the locus of the initial mutation – called a beneficial back-mutation [10, 16]. While compensatory mutations change the sequence and thus induce genetic diversification, beneficial back-mutations reduce genetic diversity and do not contribute to genetic innovation. Although Tomoko Ohta considered beneficial back-mutations negligible in her nearly-neutral theory [12], their importance has now been acknowledged for expanding populations [10]. However, differentiating between an adaptive mutation and a beneficial back-mutation remains challenging [17]. Indeed, an adaptive mutation responding to a change in the environment and a beneficial back-mutation have equivalent consequences for their bearer [10]. Similarly, at the population level, both types of mutations will result in a positive transmission bias of the beneficial allele. However, at the macro-evolutionary scale, the consequences of these two types of mutations are fundamentally different. While adaptive mutations promote phenotype diversification (fig. 1C), beneficial back-mutations promote phenotype stability and may help preserve well-established biological systems (fig. 1D). Additionally, the direction of adaptive evolution is unpredictable because it is caused by an unforeseen change in the environment and, hence, in the underlying fitness landscape [18]. On the other hand, beneficial back-mutations are predictable if the underlying stable fitness landscape is known because changes from non-optimal amino acids to optimal ones will be expected.

Fitness landscape reconstruction

The mutation-selection framework permits to link the patterns of substitution along a phylogeny with the underlying fitness landscape [19, 20]. Such mutation-selection models applied to protein-coding DNA sequences allow us to estimate relative fitnesses for all amino acids for each site of the sequence, explicitly assuming that the underlying fitness landscape is stable along the phylogenetic tree [21–23]. Importantly, because mutation-selection models at the phylogenetic scale are based on population-genetics equations, their estimates of selection coefficients are directly interpretable as fitness effects for an individual within a population.

Theoretically, we can thus use large-scale genomic data to assess whether the fitness estimated at the phylogenetic scale predict the fitness effect at the population scale for both deleterious and beneficial back-mutations. The placental mammals represent an excellent study system to perform such analysis. Having originated ∼102 million years ago, they diversified quickly [24]. Additionally, polymorphism data are available for many species [25], as are high quality protein-coding DNA alignments across the genome [26, 27]. By performing our analysis on 14,509 orthologous protein-coding genes across 87 species, we focus on sequences under functional constraints across mammals. For each gene, fitting the mutation-selection model to the multi-species sequence alignment, assuming that the underlying fitness landscape is stable along the phylogenetic tree, allows us to obtain relative fitnesses for all amino acids for each site of the alignment (fig. 2A). From the fitness landscape, we can calculate the scaled selection coefficient (S_0) for any possible amino acid

change at each position. According to its S_0 value, we can classify any mutation as either a deleterious mutation toward a less fit amino acid ($\mathcal{D}_0 := S_0 < -1$), a nearly-neutral mutation ($\mathcal{N}_0 := -1 < S_0 < 1$) or a beneficial back-mutation toward a known fitter amino acid ($\mathcal{B}_0 := S_0 > 1$). Having identified which potential DNA changes represent beneficial back-mutations (fig. 2B), we retrieved polymorphism data from 28 populations belonging to 6 genera (*Equus*, *Bos*, *Capra*, *Ovis*, *Chlorocebus*, and *Homo*) to assess the presence of beneficial back-mutations at the population scale. We focused on mutations currently segregating within populations and substitutions in the terminal branches, and checked if any of these observed changes were beneficial back-mutations (fig. 2C-E). In this study, by integrating large-scale genomic datasets at both phylogenetic and population scales, we have been able to estimate the number of beneficial back-mutations across the entire exome of the six genera (figs. 2F-G).

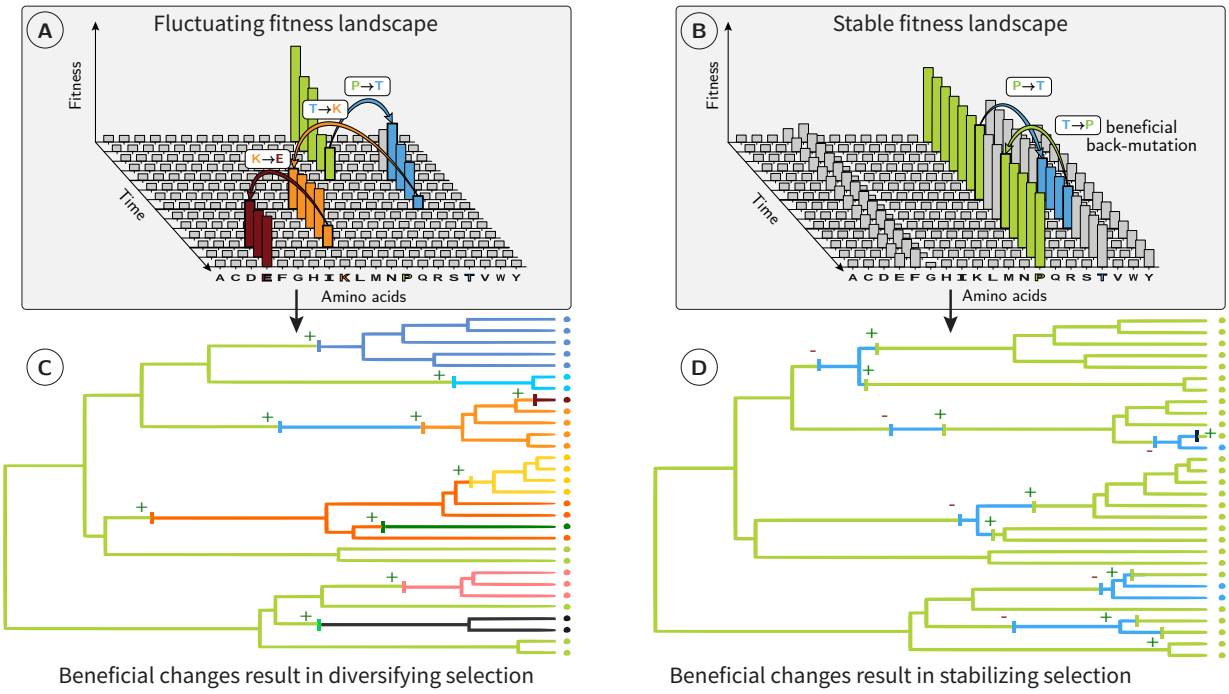

Figure 1: (A & B) For a given codon position of a protein-coding DNA sequence, amino acids (x-axis) have different fitness values (y-axis). Under a fluctuating fitness landscape (A), these fitnesses change with time. The protein sequence follows the moving target defined by the amino-acid fitnesses. Since substitutions are preferentially accepted if they are in the direction of this target, substitutions are, on average, adaptive. At the phylogenetic scale (C), beneficial substitutions are common (positive signs), promoting phenotype diversification across species. Under a stable fitness landscape (B), most mutations reaching fixation are either slightly deleterious reaching fixation due to drift or are beneficial back-mutations restoring a more optimal amino acid. At the phylogenetic scale (D), deleterious substitutions (negative signs) are often reverted via beneficial back-mutations (positive signs), promoting phenotype stability and preserving well-established biological systems. Even though, individually, any back-mutation might have a small beneficial effect on its bearer, we expect beneficial back-mutations to be scattered across the genome and the genome-wide signature of beneficial back-mutations to be detectable and quantifiable.

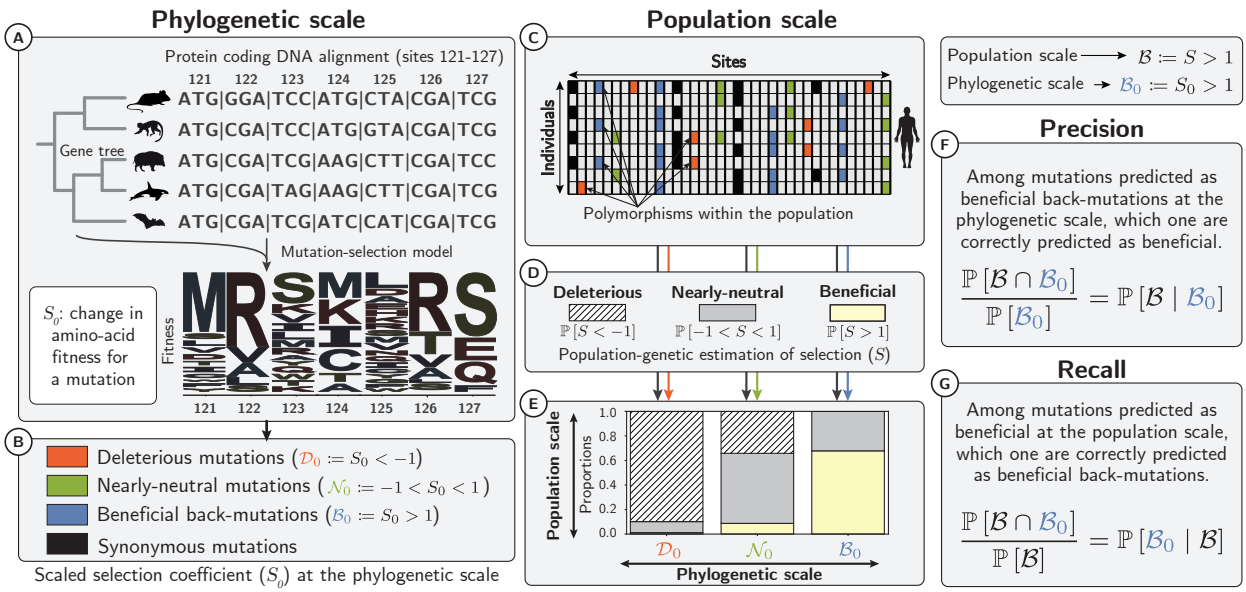

Figure 2: Selection coefficients at the phylogenetic and population scales. At the phylogenetic scale (A), we estimated the amino-acid fitness for each site from protein-coding DNA alignments using mutation-selection codon models. For every possible mutation, the difference in amino acid fitness before and after the mutation allows us to compute the selection coefficient at the phylogenetic scale (S_0). Depending on S_0 (B) mutations can be predicted as deleterious ($\mathcal{D}_0$), nearly-neutral ($\mathcal{N}_0$) or beneficial back-mutations ($\mathcal{B}_0$) toward a fitter amino acid and repairing existing functions. At the population scale, each observed single nucleotide polymorphism (SNP) segregating in the population can also be classified according to its S_0 value (C). Occurrence and frequency in the population of non-synonymous polymorphisms, contrasted to synonymous polymorphisms (deemed neutral) is used to estimate selection coefficients (D-E) at the population scale (S), for each class of selection ($\mathcal{D}_0$, $\mathcal{N}_0$, $\mathcal{B}_0$). We can thus assess whether S_0 predicts S and compute *precision* (F) and *recall* (G) for each class. The *recall* value for class $\mathcal{B}_0$ is the probability of back-mutations among all beneficial ones (G).

Selection along the terminal branches

First, we assessed whether fitness effects derived from the mutation-selection model at the phylogenetic scale predict selection occurring in terminal branches. We recovered the mutations that reached fixation in the terminal branches of the six genera. We only considered mutations fixed in a population as substitutions in the corresponding branch and discarded segregating mutations. We classified each substitution identified in the terminal branches as either $\mathcal{D}_0$, $\mathcal{N}_0$, or $\mathcal{B}_0$ depending on its S_0 value obtained at the phylogenetic scale (fig. 2A-C). Importantly, the mammalian alignment used to estimate the amino acid fitness landscape did not include the genera for which we reconstructed the substitutions to avoid fallacy and circular reasoning. Because S_0 was based on estimating fitness effects at the mammalian scale, mutations with $S_0 > 1$ were considered beneficial back-mutations ($\mathcal{B}_0$) toward a fitter amino acid. Among all the substitutions found in each terminal branch, between 10 and 13% were $\mathcal{B}_0$ (table S1 and fig. 3B for humans). In principle, beneficial back-mutations are bound to reach fixation more often than neutral mutations. Hence, we calculated the

d_N/d_S ratio of non-synonymous over synonymous divergence for all terminal lineages, focusing on the non-synonymous changes predicted as beneficial back-mutations ($d_N(\mathcal{B}_0)/d_S$). We obtained values between 1.10 and 1.61 in the different lineages (table S1), implying i) that $\mathcal{B}_0$ mutations reached fixation more frequently than synonymous mutations that are supposed to be neutral, and ii) that these back-mutations are effectively beneficial.

This result further indicated that using d_N/d_S as an estimate of adaptation is biased due to the presence of beneficial back-mutations among the non-synonymous substitutions. By discarding all beneficial back-mutations we can obtain an estimate of d_N/d_S which is not inflated. By comparing these two ways of calculating d_N/d_S (see section 1.5 in Materials & Methods), we calculated that beneficial back-mutations inflate d_N/d_S values by between 9 and 11% across genera while only representing between 1.2 and 1.4% of non-synonymous mutations (tables S2-3 and fig. 3A for humans).

Selection in populations

Second, we assessed whether our calculated S_0 values are also indicative of the selective forces exerted at the population level. We retrieved single nucleotide polymorphisms (SNPs) segregating in 28 mammalian populations. To determine if SNPs were ancestral of derived, we reconstructed the ancestral genome of each population. We then classified every non-synonymous SNP as either $\mathcal{D}_0$, $\mathcal{N}_0$, or $\mathcal{B}_0$ according its S_0 value (fig. 2B-C).

In humans, some SNPs have been associated with specific clinical prognosis terms obtained by clinical evaluation of the impact of variants on human Mendelian disorders [28]. Thus, they provide an orthogonal estimate of the effect of a mutation on human health, allowing us to evaluate whether this effect follows its predicted fitness effect. Therefore, we investigated whether the non-synonymous SNPs classified as $\mathcal{D}_0$ or $\mathcal{B}_0$ showed enrichment in specific clinical terms compared to SNPs classified as $\mathcal{N}_0$. Our results show that SNPs predicted as deleterious are associated with clinical terms such as *Likely Pathogenic* and *Pathogenic*, implying that, in general, the selective pressure of a mutation exerted across mammals is also predictive of its clinical effect in humans (table S4) [29]. Conversely, back-mutations are associated with clinical terms such as *Benign* and *Likely Benign*, which shows that back-mutations are less likely to be functionally damaging (table S5).

In addition to clinical prognosis, frequencies at which SNPs are segregating within populations provide information on their selective effects. For instance, deleterious SNPs usually segregate at lower frequencies because of purifying selection, which tends to remove them from the population (fig. 3C for humans). By gathering information across many SNPs, it is possible to estimate the distribution of fitness effects (DFE) at the population scale, taking synonymous SNPs as a neutral expectation [30–33]. From the estimated DFE, we can derive the proportion of beneficial mutations ($\mathbb{P}[\mathcal{B}]$), nearly-neutral mutations ($\mathbb{P}[\mathcal{N}]$) and deleterious mutations ($\mathbb{P}[\mathcal{D}]$) at the population scale (see section 1.6 in Materials & Methods). These approaches offer a unique opportunity to contrast selection coefficients estimated at the population scale (S) and at the phylogenetic scale (S_0).

Across our selection classes ($\mathcal{D}_0$, $\mathcal{N}_0$ and $\mathcal{B}_0$), one can ultimately estimate the proportion of correct and incorrect predictions, leading to an estimation of *precision* and *recall* (fig. 2F-G and section 1.7 in Materials & Methods). Across 28 populations of different mammal species, mutations predicted to be deleterious at the phylogenetic scale ($\mathbb{P}[\mathcal{D}]$) were indeed purged at the population scale, with a *precision* in the range of 89–96% (table 1 and fig. 3D for humans). Conversely, a *recall* in the range of 95–100% implies that mutations found to be deleterious at the population scale were most likely also predicted to be deleterious at the phylogenetic scale (table 1). Altogether, purifying selection is largely predictable and amino acids with negative fitness across mammals have been effectively purged away in each population.

Mutations predicted as $\mathcal{N}_0$ were effectively composed of a mix of neutral and selected mutations with varying *precision* (36–73%) and *recall* (30–45%) across the different populations (table 1, fig. 3D for humans). The variable proportions between populations can be explained by the effective number of individuals in the population (N_e), a major driver of selection efficacy. Since N_e is not directly measurable, we used Watterson's θ_S as a proxy to order populations by their levels of synonymous diversity. Higher synonymous diversity was associated with a smaller proportion of nearly-neutral mutations (fig. 4A). This result follows the prediction of the nearly-neutral theory and suggests that in populations with higher diversity (e.g., *Bos* or *Ovis*), discrimination between beneficial and deleterious mutations is more likely to occur (figs. S1-S6). Conversely, many more mutations are effectively neutral in populations with lower diversity (e.g., *Homo*).

Finally, mutations predicted to be beneficial back-mutations based on the selection coefficients at the phylogenetic scale ($\mathcal{B}_0$) were indeed beneficial for individuals bearing them, with a *precision* (fig. 2F) in the range of 13–86% (table 1 and fig. 3D for humans). This result confirms that selection towards amino acids restoring existing functions is ongoing in these populations. Importantly, the *recall* value in this case, computed as $\mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}]$, is the probability for a beneficial mutation at the population scale to be a beneficial back-mutation, i.e., going toward a fitter amino acid (fig. 2G). In other words, the *recall* value quantifies the number of beneficial mutations restoring damaged genomes instead of creating adaptive innovations. Across the 28 populations, this proportion is in the range of 11–81% (table 1).

Discussion

Beneficial mutations are not necessarily adaptive

This study represents an essential step toward integrating the different evolutionary scales necessary to understand the combined effects of mutation, selection, and drift on genome evolution. In particular, we have been able to quantify the proportion of beneficial back-mutations among all beneficial mutations, which has only been achievable by combining genome-wide data from both phylogenetic and population scales. At the phylogenetic scale, codon diversity at each site of a protein-coding DNA alignment allows for reconstructing an amino acid fitness landscape, assuming that this landscape is stable along the phylogeny. These amino-acid fitness landscapes allow us to predict any mutation's selection coefficient (S_0) along a protein-coding sequence. We can compare these selective effects to observations at the population level (figs. S1-3). By doing so, we confirmed that mutations predicted to be deleterious ($\mathcal{D}_0 \coloneqq S_0 < -1$) are purified away in extant populations. Our results concur with previous studies showing that SIFT scores [34,

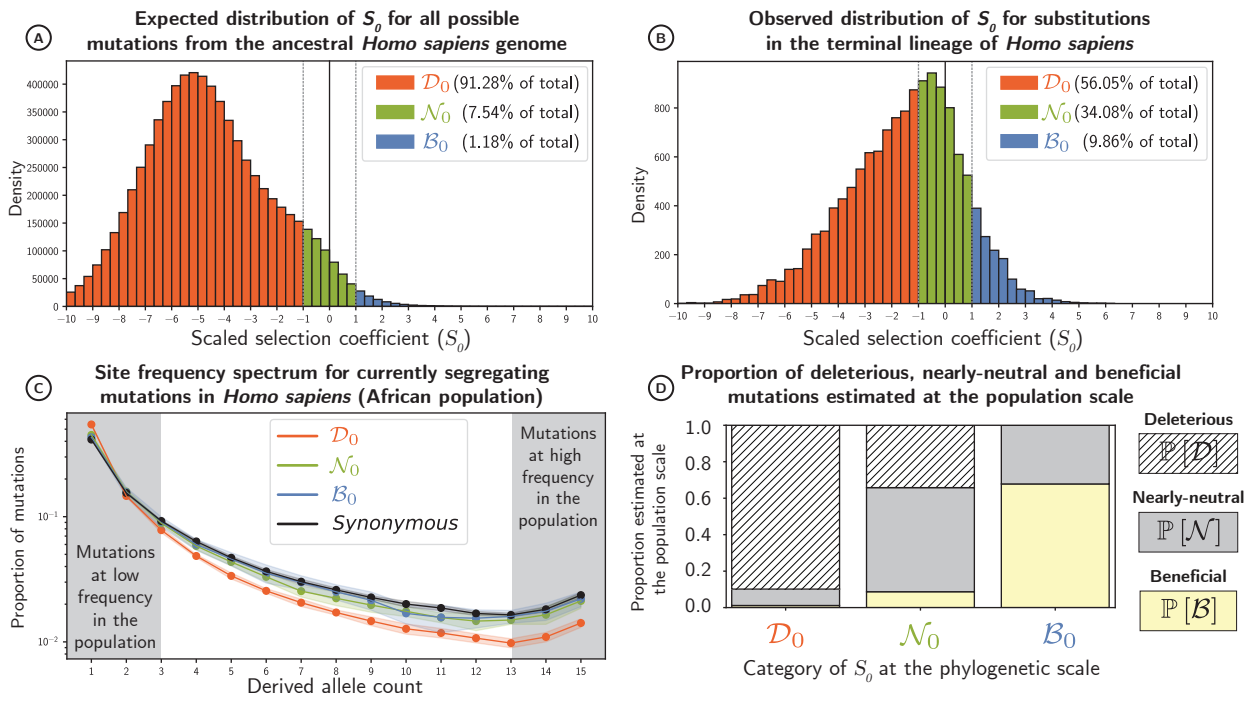

Figure 3: (A) Distribution of scaled selection coefficients (S_0), predicted for all possible mutations away from the ancestral human genome. Mutations are divided into three classes of selection: deleterious ($\mathcal{D}_0$), nearly-neutral ($\mathcal{N}_0$) and beneficial ($\mathcal{B}_0$, supposedly beneficial back-mutations). (B) Distribution of scaled selection coefficients (S_0) for all observed substitutions along the *Homo* branch after the *Homo-Pan* split. If there are fewer substitutions than expected, this class is thus undergoing purifying selection, as is the case for $\mathcal{D}_0$. (C) The site-frequency spectrum (SFS) in humans of African descent for a random sample of 16 alleles (means in solid lines and standard deviations in color shades) for each class of selection and for synonymous mutations, supposedly neutral (black). The SFS represents the proportion of mutations (y-axis) with a given number of derived alleles in the population (x-axis). At high frequencies, deleterious mutations are underrepresented. (D) Proportion of beneficial $\mathbb{P}[\mathcal{D}]$, nearly-neutral $\mathbb{P}[\mathcal{N}]$, and deleterious mutations $\mathbb{P}[\mathcal{B}]$ estimated at the population scale for each class of selection at the phylogenetic scale. Proportions depicted here are not weighted by their mutational opportunities.

35], based on amino acid alignments across species, also inform on the deleterious fitness effects exerted at the population scale [36]. However – contrary to SIFT scores – our mutation-selection model is parameterized by a fitness function such that changes are directly interpretable as fitness effects. In this regard, an interesting prediction of our model is that back-mutations are beneficial because they revert previous deleterious changes. We have tested this hypothesis and found that these back-mutations ($\mathcal{B}_0 := S_0 > 1$) are indeed beneficial in extant populations. We estimated that between 11 and 81% of all beneficial mutations in mammalian populations have not been driven by adaptation but instead have reversed deleterious substitutions which have accumulated along the phylogeny. More specifically, in 21 out of 28 populations analyzed, the percentage of beneficial mutations estimated to be back-mutations falls between 20 and 40%. These results suggest that many beneficial mutations are non-adaptive but rather restore ancestral states of higher fitness. Hence, we

			Deleterious mutations $\mathcal{D} \coloneqq S < -1$ $\mathcal{D}_0 \coloneqq S_0 < -1$		**Nearly-neutral mutations** $\mathcal{N} \coloneqq -1 < S < 1$ $\mathcal{N}_0 \coloneqq -1 < S_0 < 1$		**Beneficial mutations** $\mathcal{B} \coloneqq S > 1$ $\mathcal{B}_0 \coloneqq S_0 > 1$	
Population	**Species**	θ_S	**Precision** $\mathbb{P}[\mathcal{D} \mid \mathcal{D}_0]$	**Recall** $\mathbb{P}[\mathcal{D}_0 \mid \mathcal{D}]$	**Precision** $\mathbb{P}[\mathcal{N} \mid \mathcal{N}_0]$	**Recall** $\mathbb{P}[\mathcal{N}_0 \mid \mathcal{N}]$	**Precision** $\mathbb{P}[\mathcal{B} \mid \mathcal{B}_0]$	**Recall** $\mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}]$
Equus c.	Equus caballus	0.002	0.908	0.976	0.730	0.383	0.580	0.809
Iran	Bos taurus	0.002	0.898	1.000	0.632	0.340	0.857	0.277
Uganda	Bos taurus	0.005	0.942	0.963	0.497	0.393	0.488	0.483
Australia	Capra hircus	0.003	0.934	0.966	0.524	0.425	0.368	0.196
France	Capra hircus	0.004	0.936	0.965	0.506	0.412	0.368	0.211
Iran (C. aegagrus)	Capra hircus	0.004	0.940	0.961	0.478	0.438	0.221	0.114
Iran	Capra hircus	0.004	0.945	0.960	0.419	0.400	0.368	0.211
Italy	Capra hircus	0.004	0.937	0.963	0.548	0.426	0.270	0.220
Morocco	Capra hircus	0.004	0.942	0.964	0.522	0.428	0.368	0.276
Iran	Ovis aries	0.007	0.953	0.954	0.450	0.406	0.169	0.392
Iran (O. orientalis)	Ovis aries	0.009	0.958	0.953	0.417	0.431	0.153	0.180
Iran (O. vignei)	Ovis aries	0.007	0.960	0.951	0.360	0.449	0.195	0.133
Various	Ovis aries	0.008	0.955	0.954	0.435	0.426	0.190	0.228
Morocco	Ovis aries	0.008	0.956	0.955	0.449	0.410	0.136	0.353
Barbados	Chlorocebus sabaeus	0.004	0.925	0.969	0.536	0.377	0.565	0.306
Central Afr. Rep.	Chlorocebus sabaeus	0.006	0.940	0.964	0.495	0.404	0.429	0.274
Ethiopia	Chlorocebus sabaeus	0.005	0.925	0.969	0.560	0.400	0.462	0.239
Gambia	Chlorocebus sabaeus	0.005	0.934	0.969	0.610	0.409	0.486	0.774
Kenya	Chlorocebus sabaeus	0.005	0.938	0.965	0.519	0.434	0.478	0.243
Nevis	Chlorocebus sabaeus	0.003	0.923	0.970	0.609	0.393	0.516	0.393
South Africa	Chlorocebus sabaeus	0.006	0.935	0.965	0.542	0.409	0.490	0.360
Saint Kitts	Chlorocebus sabaeus	0.004	0.926	0.969	0.562	0.377	0.515	0.367
Zambia	Chlorocebus sabaeus	0.006	0.937	0.965	0.494	0.413	0.499	0.249
African	Homo sapiens	0.002	0.899	0.970	0.571	0.334	0.678	0.319
Admixed American	Homo sapiens	0.002	0.888	0.972	0.581	0.301	0.635	0.361
East Asian	Homo sapiens	0.002	0.894	0.973	0.583	0.335	0.611	0.233
European	Homo sapiens	0.002	0.896	0.972	0.581	0.350	0.619	0.217
South Asian	Homo sapiens	0.002	0.898	0.972	0.580	0.354	0.628	0.222

Table 1: *Precision* and *recall* for estimated selection coefficient of mutations given by mutation-selection models (S_0). *Precision* is the estimation of the selection coefficient at population scale (S) given that S_0 is known. Conversely, *recall* is the estimation of S_0 given selection coefficient at the population scale (S) is known. *Recall* for beneficial mutations ($\mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}]$) is thus the proportion of beneficial back-mutations among all beneficial mutations. Watterson's θ_S is the observed genetic diversity calculated for synonymous changes.

can correctly estimate the extent of adaptive evolution only if we account for the number of beneficial back-mutations [37, 38]. Altogether, we argue that we should dissociate positive selection from adaptive evolution and limit the use of adaptive mutations to those that are associated with adaptation to environmental change as such [10, 11].

Interpreting the proportion of beneficial back-mutations

Across the genome, beneficial-back mutations and deleterious mutations reaching fixation create a balance in which genomes are constantly damaged and restored simultaneously at different loci. Since the probability of fixation of mutations depends on the effective population size (N_e), the history of N_e plays a crucial role in setting the number of beneficial back-mutations compensating for deleterious mutations [39]. For example, a population size expansion will increase the efficacy of selection, and a larger proportion of mutations will be beneficial (otherwise effectively neutral), thus increasing the number of beneficial back-mutations. On the other hand, a population that has experienced a high N_e throughout its history should be closer to an optimal state under a stable fitness landscape, having suffered fewer fixations of deleterious mutations and therefore decreasing the probability of beneficial back-mutations [40]. Overall, we expect the proportion of beneficial back-mutations to be more dependent on N_e's long-term expansions and contractions than on the short-term ones [10, 40]. Indeed the proportion of back-mutations among all beneficial mutations did

not show a clear relationship with the synonymous diversity within a population (Watterson's θ_S) (fig. 4B), which should, in theory, reflect the population's short-term N_e.

The exact estimation of the contribution of beneficial back-mutations to positive selection relies on some hypotheses at both the phylogenetic and population scales and is sensitive to methodological limitations. Indeed, to be conservative, we considered that any mutation under positive selection at the population scale ($\mathcal{B}$) but not predicted as such at the phylogenetic scale (not $\mathcal{B}_0$) is potentially an adaptation. However, adaptation is not the only factor hindering the detection of beneficial back-mutations; data quality and potentially inadequate modeling choices of both the fitness landscape and the DFE might also lead to missed predictions. Given that our amino-acid fitness landscapes are site-specific and also independent of one another, missed prediction can be also due to epistasis where the fitness landscape at a given site depends on the residue present at other sites [41, 42]. Moreover, a mutation inferred as nearly-neutral at the phylogenetic scale ($\mathcal{N}_0$) but positively selected at the population scale ($\mathcal{B}$) will also be considered a missed prediction, even if this discrepancy could just be due to a population size expansion [43–45]. Given this inflation of missed predictions, our estimated proportion of adaptive mutations among beneficial ones is likely to be an overestimation. In practice, advantageous mutations can be depleted in the $\mathcal{N}_0$ and $\mathcal{D}_0$ class in some species but pervasive in others, depending on the model used to infer the DFE at the population scale (tables S6-9). It appears that our estimation can be quite sensitive to model misspecification and overall, while we provide an order of magnitude for the contribution of beneficial back-mutations to positive selection, further methodological development on the estimation of the distribution of fitness effects is needed to increase the confidence on this value.

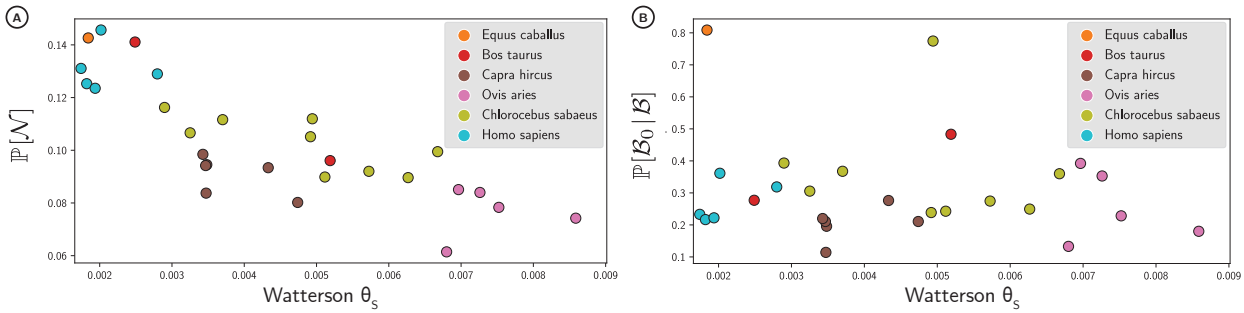

Figure 4: (A) Proportion of nearly-neutral mutations at the population scale ($\mathbb{P}[\mathcal{N}]$ in the y-axis), shown as a function of the synonymous diversity (Watterson's θ_S in the x-axis). (B) Proportion of beneficial back-mutations among all beneficial mutations ($\mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}]$ in the y-axis), shown as a function of the synonymous diversity (Watterson's θ_S in the x-axis).

Detecting adaptation above the nearly-neutral background

A long-standing debate in molecular evolution is whether the variations we observe between species in protein-coding genes are primarily due to nearly-neutral mutations reaching fixation by drift or primarily due to adaptation [12, 46–48]. Here we provide evidence that in mammalian orthologs, many substitutions occur

through fixation of both deleterious mutations and beneficial back-mutations. However, detecting adaptation above this background of nearly-neutral substitutions remains a central question [46, 49].

One first strategy is precisely to use a nearly-neutral substitution model as a null model of evolution. Under a strictly neutral evolution of protein-coding sequence, we expect the ratio of non-synonymous over synonymous substitutions (d_N/d_S) to be equal to one. Deviations from this neutral expectation, such as $d_N/d_S > 1$, which can be generated by an excess of non-synonymous substitutions, is generally interpreted as a sign of adaptation. However, as shown in this study, a $d_N/d_S > 1$ is not necessarily a signature of adaptation but can be due to beneficial back-mutations. So, by relaxing the strict neutrality and assuming a stable fitness landscape instead, one can predict the expected rate of evolution, called ω_0 [50, 51]. Adaptation can thus be considered as evolution under a changing fitness landscape and tested as such by searching for the signature of $d_N/d_S > \omega_0$ [15, 23, 52]. Using a stable fitness landscape as a null model of evolution, thus accounting for selective constraints exerted on the different amino acids, increased the statistical power in testing for adaptation [53]. Instead of relying solely on summary statistics (such as d_N/d_S or ω_0), another strategy to detect adaptation is to include changes in the fitness landscapes inherently within the mutation-selection framework [54]. Such mechanistic models could be more general than site-specific fitness landscapes, including epistasis and changing fitness landscapes [55, 56].

Here, we have provided empirical evidence that an evolutionary model assuming a stable fitness landscape at the mammalian scale allows us to predict the fitness effects of mutations in extant populations and individuals, acknowledging the balance between deleterious and beneficial back-mutations. We argue that such a model represents a null expectation for the evolution of protein-coding genes in the absence of adaptation. In that sense, to avoid conflating beneficial mutations with adaptive evolution, the term “adaptation“ should retain its original meaning associated with a change in the underlying fitness landscape and be modelled as such.

Acknowledgements

We gratefully acknowledge the help of Mélodie Bastian, Nicolas Lartillot, Carina Farah Mugal, Laurent Duret, Alexandre Reymond, Daniele Silvestro and Nicolas Gambardella for their advice and reviews concerning this manuscript. This work was performed using the computing facilities of the CC LBBE/PRABI. This study makes use of data generated by the NextGen Consortium. **Funding:** Université de Lausanne; Agence Nationale de la Recherche, Grant ANR-19-CE12-0019 / HotRec. The European Union’s Seventh Framework Programme (FP7/2010-2014) provided funding for the project under grant agreement no 244356 - “NextGen”. **Author contributions:** Original idea: T.L. and J.J.; Model conception: T.L., J.J. and N.S.; Code: T.L.; Data analyses: T.L. and J.J.; Interpretation: T.L., J.J., D.A.H. and N.S.; First draft: T.L. and J.J.; Editing and revisions: T.L., J.J., D.A.H. and N.S. Project management and funding: N.S. **Competing interests:** The authors declare no conflicts of interest. **Data and materials availability:** The data underlying this article are available at `https://doi.org/10.5281/zenodo.7878954`. Snakemake pipeline, analysis scripts and documentation are available at github.com/ThibaultLatrille/SelCoeff.

1 Materials & Methods

1.1 Phylogenetic dataset

Protein-coding DNA sequence alignments in placental mammals and their corresponding gene trees come from the OrthoMaM database and were processed as in Latrille *et al.* [53]. OrthoMaM contains a total of 116 mammalian reference sequences in v10c [26, 27, 57].

Genes located on the X and Y chromosomes and on the mitochondrial genome were discarded from the analysis because the level of polymorphism – which is necessary for population-based analyses – is expected to be different in these three regions compared to the autosomal genome. Sequences of species for which we used population-level polymorphism (see section 1.3) and their sister species, were removed from the analysis to ensure independence between the data used in the phylogenetic and population scales. Sites in the alignment containing more than 10% of gaps across the species were discarded. Altogether, our genome-wide dataset contains $14,509$ protein-coding DNA sequences in 87 placental mammals.

1.2 Selection coefficient (S_0) in a phylogeny-based method

We analyzed the phylogenetic-level data using mutation-selection models. These models assume the protein-coding sequences are at mutation-selection balance under a fixed fitness landscape characterized by a fitness vector over the 20 amino acids at each site [19, 21, 58]. Mathematically, the rate of non-synonymous substitution from codon a to codon b ($q^{(i)}_{a\mapsto b}$) at site i of the sequence is equal to the rate of mutation of the underlying nucleotide change ($\mu_{a\mapsto b}$) multiplied by the scaled probability of mutation fixation ($\mathbb{P}^{(i)}_{a\mapsto b}$). The probability of fixation depends on the difference between the scaled fitness of the amino acid encoded by the mutated codon ($F^{(i)}_b$) and the amino acid encoded by the original codon ($F^{(i)}_a$) at site i [59, 60].

The rate of substitution from codon a to b at a site i is thus:

$$\begin{cases} q^{(i)}_{a\mapsto b} &= 0 \text{ if codons } a \text{ and } b \text{ are more than one mutation away,} \\ q^{(i)}_{a\mapsto b} &= \mu_{a\mapsto b} \text{ if codons } a \text{ and } b \text{ are synonymous, and} \\ q^{(i)}_{a\mapsto b} &= \mu_{a\mapsto b}\dfrac{F^{(i)}_b - F^{(i)}_a}{1-\mathrm{e}^{F^{(i)}_a - F^{(i)}_b}} \text{ if codons } a \text{ and } b \text{ are non-synonymous.} \end{cases} \tag{1}$$

Fitting the mutation-selection model on a multi-species sequence alignment leads to an estimation of the gene-wide 4×4 nucleotide mutation rate matrix ($\boldsymbol{\mu}$) as well as the 20 amino-acid fitness landscape ($\boldsymbol{F^{(i)}}$) at each site i. The selection coefficient for a mutation from codon a to codon b at site i is defined as:

$$S^{(i)}_0(a\mapsto b) = \Delta F^{(i)} = F^{(i)}_b - F^{(i)}_a. \tag{2}$$

In our subsequent derivation the source (a) and target (b) codons as well as the site (i) are implicit and thus never explicitly written. We used the Bayesian software *BayesCode* (`https://github.com/ThibaultLatrille/bayescode`) to estimate the selection coefficients for each protein-coding gene in the mammalian dataset. We ran the Markov chain Monte-Carlo (MCMC) algorithm implemented in BayesCode for $2,000$ generations as described in Latrille *et al.* [53]. For each gene, after discarding a burn-in period of $1,000$ generations of MCMC, we obtained posterior mean estimates (over the $1,000$ generations left of MCMC) of the mutation rate matrix ($\boldsymbol{\mu}$) as well as the 20 amino acid fitness landscape ($\boldsymbol{F^{(i)}}$) at each site i.

1.3 Polymorphism dataset

The genetic variants representing the population level polymorphisms were obtained from the following species and their available datasets: *Equus caballus* (EquCab2 assembly in the EVA study PRJEB9799 [61]), *Bos taurus* (UMD3.1 assembly in the NextGen project: https://projects.ensembl.org/nextgen/), *Ovis aries* (Oar_v3.1 assembly in the NextGen project), *Capra hircus* (CHIR1 assembly in the NextGen project), converted to ARS1 assembly with dbSNP identifiers [62]), *Chlorocebus sabaeus* (ChlSab1.1 assembly in the EVA project PRJEB22989 [63]), *Homo sapiens* (GRCh38 assembly in the 1000 Genomes Project [64]). In total, we analyzed 28 populations across the 6 different species with polymorphism data. The data was processed as described in Latrille *et al.* [53].

Only bi-allelic single nucleotide polymorphisms (SNPs) found within a gene were in our polymorphism dataset, while nonsense variants and indels were discarded. To construct the dataset, we first recovered the location of each SNP (represented by its chromosome, position, and strand) in the focal species and matched it to its corresponding position in the coding sequence (CDS) using gene annotation files (GTF format) downloaded from Ensembl (ensembl.org). We then verified that the SNP downloaded from Ensembl matched the reference in the CDS in FASTA format. Next, the position in the CDS was converted to the corresponding position in the multi-species sequence alignment (containing gaps) from the OrthoMaM database (see section 1.2) for the corresponding gene by doing a global pairwise alignment (Biopython function pairwise2). This conversion from genomic position to alignment position was only possible when the assembly used for SNP-calling was the same as the one used in the OrthoMaM alignment, the GTF annotations, and the FASTA sequences. SNPs were polarized using the three closest outgroups found in the OrthoMaM alignment with est-usfs v2.04 [65], and alleles with a probability of being derived lower than 0.99 were discarded.

1.4 Mutational opportunities

The mutational opportunities of any new mutation refer to its likelihood of falling into a specific category (synonymous, deleterious, nearly-neutral, or beneficial). Deriving such opportunities is necessary to estimate the strength of selection exerted at the population scale since different categories might have different mutational opportunities, and thus polymorphism and divergence need to be corrected accordingly (see sections 1.5, 1.6, and 1.7). To calculate mutational opportunities, we reconstructed the ancestral genome of each of the 28 populations, which differs from the corresponding species' reference genome since it accounts for the variability present in the specific population. So, for each population, we used the polarized SNPs to reconstruct the ancestral DNA sequence, containing only substitutions and no segregating polymorphisms. If a SNP was still segregating in the population, we used its ancestral version.

From the reconstructed ancestral genome, all possible mutations were computed, weighted by the instantaneous rate of change between nucleotides obtained from the mutation rate matrix ($\boldsymbol{\mu}$, see section 1.2), summing to μ_{tot} across the whole genome, and to μ_{syn} when restricted to synonymous mutations. Finally, the mutational opportunities for synonymous mutations were computed as the total number of sites across the genome (L_{tot}) weighted by the proportion of synonymous mutations among all possible mutations as:

$$L_{\text{syn}} = L_{\text{tot}} \frac{\mu_{\text{syn}}}{\mu_{\text{tot}}}. \tag{3}$$

Similarly, for non-synonymous mutations, the total mutation rate for each class of selection $x \in \{\mathcal{D}_0, \mathcal{N}_0, \mathcal{B}_0\}$, called $\mu(x)$, was estimated as the sum across all non-synonymous mutations if their selection coefficient at the phylogenetic scale is in the class $S_0 \in x$. Accordingly, the mutational opportunities ($L(x)$) for each class of selection coefficient (x) was finally computed as the total number of sites across the genome (L_{tot}) weighted by the ratio of the aggregated mutations rates falling in the class $\mu(x)$:

$$L(x) = L_{\text{tot}} \frac{\mu(x)}{\mu_{\text{tot}}}. \tag{4}$$

Finally, $\mathbb{P}[x]$ is the probability for a non-synonymous mutation to be in the class x, thus computed as:

$$\mathbb{P}[x] = \frac{L(x)}{\sum_{y \in \{\mathcal{D}_0, \mathcal{N}_0, \mathcal{B}_0\}} L(y)}. \tag{5}$$

1.5 Substitution mapping and d_N/d_S in the terminal branch

For each gene and each population with polymorphism data available, the genome at the base of the population genealogy was reconstructed (see section 1.4). We used the ancestral genome reconstructed at the base of the population genealogy and the three closest outgroups found in the OrthoMaM alignment to infer the ancestral protein-coding DNA sequences for each node of the 4-taxa tree by applying the M5 codon model (gamma site rate variation) as implemented in FastML.v3.11 [66]. All polarized codon substitutions were then obtained by comparing the ancestral protein-coding DNA sequences before the split from the sister species (closest outgroup) to the genome at the base of the population genealogy. We considered *Ceratotherium simum simum* as *Equus caballus'* sister species; *Bison bison bison* as *Bos taurus'* sister species; *Pantholops hodgsonii* as *Ovis aries'* sister species; *Pantholops hodgsonii* as *Capra hircus'* sister species; *Macaca mulatta* as *Chlorocebus sabaeus'* sister species and finally, we considered *Pan troglodytes* as *Homo sapiens'* sister species. The selection coefficient (S_0) of each substitution was then obtained by comparing the difference in amino acid fitnesses for each polarized non-synonymous codon substitution, by referring to the site-specific fitness profile obtained in the phylogeny-based method (see section 1.2). Finally, the rate of non-synonymous over synonymous substitutions for a given class of selection coefficient ($x \in \{\mathcal{D}_0, \mathcal{N}_0, \mathcal{B}_0\}$) was computed as:

$$\begin{cases} d_N(x) & = \dfrac{D(x)}{L(x)}, \\ d_S & = \dfrac{D_{\text{syn}}}{L_{\text{syn}}}, \end{cases} \tag{6}$$

where $D(x)$ was the number of non-synonymous substitutions in class x, D_{syn} was the number of synonymous substitutions across the genome, while $L(x)$ and L_{syn} were the numbers of non-synonymous and synonymous mutational opportunities, respectively, as defined in section 1.4. $\delta(d_N/d_S)$ was computed as the difference between d_N/d_S computed over all substitutions and d_N/d_S when we removed beneficial back-mutations $d_N(S_0 < 1)/d_S$, normalized by d_N/d_S. Note that the quantities $\delta(d_N/d_S)$ and $\delta(d_N)$ are equivalent due to the simplification of the factor d_S:

$$\delta(d_N/d_S) = \frac{d_N/d_S - d_N(S_0 < 1)/d_S}{d_N/d_S} = \frac{d_N - d_N(S_0 < 1)}{d_N} = \delta(d_N). \tag{7}$$

1.6 Scaled selection coefficients (S) in a population-based method

To obtain a quantitative estimate of the distribution of selection coefficients for each category of SNPs, we used the polyDFE model [33, 67]. This model uses the count of derived alleles to infer the distribution of fitness effects (DFE). The probability of sampling an allele at a given frequency (before fixation or extinction) is informative of its scaled selection coefficient at the population scale (S). Therefore, pooled across many sites, the site-frequency spectrum (SFS) provides information on the underlying S of mutations. However, estimating a single S for all sampled mutations is biologically unrealistic, and a DFE of mutations is usually assumed [30, 31]. The polyDFE[33, 67] software implements a mixture of a Γ and exponential distributions to model the DFE of non-synonymous mutations, while synonymous mutations are considered neutral. The model estimates the parameters β_d, b, p_b and β_b for non-synonymous mutations as:

$$\phi\left(S;\beta_d,b,p_b,\beta_b\right)=\begin{cases}(1-p_b)\,f_\Gamma(-S;-\beta_d,b) & \text{if } S\leq 0,\\ p_b f_e(S;\beta_b) & \text{if } S>0,\end{cases} \tag{8}$$

where $\beta_d \leq -1$ is the estimated mean of the DFE for $S \leq 0$; $b \geq 0.2$ is the estimated shape of the Γ distribution; $0 \leq p_b \leq 1$ is the estimated probability that $S > 0$; $\beta_b \geq 1$ is the estimated mean of the DFE for $S > 0$; and $f_\Gamma(S; m, b)$ is the density of the Γ distribution with mean m and shape b, while $f_e(S; m)$ is the density of the exponential distribution with mean m.

PolyDFE requires one SFS for non-synonymous mutations and one for synonymous mutations (neutral expectation), as well as the number of sites on which each SFS was sampled. For populations containing more than 8 individuals, the SFS was subsampled down to 16 chromosomes (8 diploid individuals) without replacement (hyper-geometric distribution) to alleviate the effect of different sampling depths in the 28 populations. Altogether, for each class of selection ($x \in \{\mathcal{D}_0, \mathcal{N}_0, \mathcal{B}_0\}$) of non-synonymous SNPs, we aggregated all the SNPs in the selection class x as an SFS. The number of sites on which each SFS was sampled is given by $L(x)$ for the non-synonymous SFS and L_{syn} for the synonymous SFS respectively. For each class of selection x, once fitted to the data using maximum likelihood with polyDFE, the parameters of the DFE $(\beta_d, b, p_b, \beta_b)$ were used to compute $\mathbb{P}[\mathcal{D} \mid x]$, $\mathbb{P}[\mathcal{N} \mid x]$, and $\mathbb{P}[\mathcal{B} \mid x]$ as:

$$\mathbb{P}[\mathcal{D}\mid x]=\mathbb{P}[S<-1\mid x]=(1-p_b)\int_1^{+\infty} f_\Gamma(S;-\beta_d,b)\mathrm{d}S, \tag{9}$$

$$\mathbb{P}[\mathcal{N}\mid x]=\mathbb{P}[-1<S<1\mid x]=p_b\int_0^1 f_e(S;\beta_b)\mathrm{d}S+(1-p_b)\int_0^1 f_\Gamma(S;-\beta_d,b)\mathrm{d}S, \tag{10}$$

$$\mathbb{P}[\mathcal{B}\mid x]=\mathbb{P}[S>1\mid x]=p_b\int_1^{+\infty} f_e(S;\beta_b)\mathrm{d}S. \tag{11}$$

1.7 *Precision* and *recall*

For readability, we give here *precision* and *recall* for beneficial mutations ($\mathcal{B}_0$ and $\mathcal{B}$), but it can be obtained using the same derivation for the deleterious mutations ($\mathcal{D}_0$ and $\mathcal{D}$) and nearly-neutral mutations ($\mathcal{N}_0$ and $\mathcal{N}$).

Precision is the proportion of mutations correctly predicted as beneficial ($\mathbb{P}[\mathcal{B} \cap \mathcal{B}_0]$) out of all predicted as beneficial back-mutations ($\mathbb{P}[\mathcal{B}_0]$), which can be written as a conditional probability:

$$\frac{\mathbb{P}[\mathcal{B} \cap \mathcal{B}_0]}{\mathbb{P}[\mathcal{B}_0]} = \mathbb{P}[\mathcal{B} \mid \mathcal{B}_0]. \tag{12}$$

Namely, *precision* corresponds to the probability for a back-mutation ($\mathcal{B}_0$) to be effectively beneficial at the population level ($\mathcal{B}$). This probability, computed from eq. 11, is obtained by restricting our analysis to SNPs that are predicted to be beneficial back-mutations (yellow fill for the category $\mathcal{B}_0$ in fig. 3D).

Recall is the proportion of mutations correctly predicted as beneficial ($\mathbb{P}[\mathcal{B} \cap \mathcal{B}_0]$) out of all beneficial mutations ($\mathbb{P}[\mathcal{B}]$), which can be written as a conditional probability:

$$\frac{\mathbb{P}[\mathcal{B} \cap \mathcal{B}_0]}{\mathbb{P}[\mathcal{B}]} = \mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}]. \tag{13}$$

Namely, *recall* corresponds to the probability for a beneficial mutation at the population level ($\mathcal{B}$) to be a beneficial back-mutation ($\mathcal{B}_0$). Using Bayes theorem, *recall* can be re-written as:

$$\mathbb{P}[\mathcal{B}_0 \mid \mathcal{B}] = \frac{\mathbb{P}[\mathcal{B} \mid \mathcal{B}_0] \times \mathbb{P}[\mathcal{B}_0]}{\mathbb{P}[\mathcal{B}]}, \tag{14}$$

where $\mathbb{P}[\mathcal{B} \mid \mathcal{B}_0]$ and $\mathbb{P}[\mathcal{B}_0]$ can be calculated using equations 12 and 5, respectively, and $\mathbb{P}[\mathcal{B}]$ is the probability of a mutation to be beneficial at the level of the population, which can be computed from the law of total probabilities as:

$$\mathbb{P}[\mathcal{B}] = \sum_{x \in \{\mathcal{D}_0, \mathcal{N}_0, \mathcal{B}_0\}} \mathbb{P}[\mathcal{B} \mid x] \times \mathbb{P}[x]. \tag{15}$$

References

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Merrell, D. J. *The Adaptive Seascape: The Mechanism of Evolution* 284 pp. (U of Minnesota Press, 1994).
3. Gavrilets, S. & Losos, J. B. Adaptive Radiation: Contrasting Theory with Data. *Science* **323,** 732–737 (2009).
4. McDonald, J. H. & Kreitman, M. Adaptative Protein Evolution at Adh Locus in Drosophila. *Nature* **351,** 652–654 (1991).
5. Smith, N. G. & Eyre-Walker, A. Adaptive Protein Evolution in Drosophila. *Nature* **415,** 1022–1024 (2002).
6. Welch, J. J. Estimating the Genomewide Rate of Adaptive Protein Evolution in Drosophila. *Genetics* **173,** 821–837 (2006).
7. Yang, Z. & Bielawski, J. R. Statistical Methods for Detecting Molecular Adaptation. *Trends in Ecology and Evolution* **15,** 496–503 (2000).

8. Eyre-Walker, A. The Genomic Rate of Adaptive Evolution. *Trends in Ecology & Evolution* **21,** 569–575 (2006).

9. Moutinho, A. F., Bataillon, T. & Dutheil, J. Y. Variation of the Adaptive Substitution Rate between Species and within Genomes. *Evolutionary Ecology* **34,** 315–338 (2019).

10. Charlesworth, J. & Eyre-Walker, A. The Other Side of the Nearly Neutral Theory, Evidence of Slightly Advantageous Back-Mutations. *Proceedings of the National Academy of Sciences* **104,** 16992–16997 (2007).

11. Mustonen, V. & Lässig, M. From Fitness Landscapes to Seascapes: Non-Equilibrium Dynamics of Selection and Adaptation. *Trends in genetics* **25,** 111–119 (2009).

12. Ohta, T. The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics* **23,** 263–286 (1992).

13. Hartl, D. L. & Taubes, C. H. Compensatory Nearly Neutral Mutations: Selection without Adaptation. *Journal of Theoretical Biology* **182,** 303–309 (1996).

14. Sella, G. & Hirsh, A. E. The Application of Statistical Physics to Evolutionary Biology. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 9541–9546 (2005).

15. Cvijović, I., Good, B. H., Jerison, E. R. & Desai, M. M. Fate of a Mutation in a Fluctuating Environment. *Proceedings of the National Academy of Sciences* **112,** E5021–E5028 (2015).

16. Piganeau, G. & Eyre-Walker, A. Estimating the Distribution of Fitness Effects from DNA Sequence Data: Implications for the Molecular Clock. *Proceedings of the National Academy of Sciences of the United States of America* **100,** 10335–10340 (2003).

17. Chi, P. B., Kosater, W. M. & Liberles, D. A. Detecting Signatures of Positive Selection against a Backdrop of Compensatory Processes. *Molecular Biology and Evolution* **37,** 3353–3362 (2020).

18. Bazykin, G. A. Changing Preferences: Deformation of Single Position Amino Acid Fitness Landscapes and Evolution of Proteins. *Biology letters* **11,** 20150315 (2015).

19. Halpern, A. L. & Bruno, W. J. Evolutionary Distances for Protein-Coding Sequences: Modeling Site-Specific Residue Frequencies. *Molecular Biology and Evolution* **15,** 910–917 (1998).

20. McCandlish, D. M. & Stoltzfus, A. Modeling Evolution Using the Probability of Fixation: History and Implications. *Quarterly Review of Biology* **89,** 225–252 (2014).

21. Rodrigue, N. & Philippe, H. Mechanistic Revisions of Phenomenological Modeling Strategies in Molecular Evolution. *Trends in Genetics* **26,** 248–252 (2010).

22. Tamuri, A. U. & Goldstein, R. A. Estimating the Distribution of Selection Coefficients from Phylogenetic Data Using Sitewise Mutation-Selection Models. *Genetics* **190,** 1101–1115 (2012).

23. Rodrigue, N. & Lartillot, N. Detecting Adaptation in Protein-Coding Genes Using a Bayesian Site-Heterogeneous Mutation-Selection Codon Substitution Model. *Molecular Biology and Evolution* **34,** 204–214 (2017).

24. Foley, N. M. *et al.* A Genomic Timescale for Placental Mammal Evolution. *Science* **380,** eabl8189 (2023).

25. Howe, K. L. *et al.* Ensembl 2021. *Nucleic Acids Research* **49,** D884–D891 (2021).

26. Ranwez, V. *et al.* OrthoMaM: A Database of Orthologous Genomic Markers for Placental Mammal Phylogenetics. *BMC Evolutionary Biology* **7,** 1–12 (2007).
27. Scornavacca, C. *et al.* OrthoMaM V10: Scaling-up Orthologous Coding Sequence and Exon Alignments with More than One Hundred Mammalian Genomes. *Molecular Biology and Evolution* **36** (ed Tamura, K.) 861–862 (2019).
28. Landrum, M. J. *et al.* ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. *Nucleic Acids Research* **46,** D1062–D1067 (2018).
29. Sullivan, P. F. *et al.* Leveraging Base-Pair Mammalian Constraint to Understand Genetic Variation and Human Disease. *Science* **380,** eabn2937 (2023).
30. Eyre-Walker, A., Woolfit, M. & Phelps, T. The Distribution of Fitness Effects of New Deleterious Amino Acid Mutations in Humans. *Genetics* **173,** 891–900 (2006).
31. Eyre-Walker, A. & Keightley, P. D. Estimating the Rate of Adaptive Molecular Evolution in the Presence of Slightly Deleterious Mutations and Population Size Change. *Molecular Biology and Evolution* **26,** 2097–2108 (2009).
32. Galtier, N. Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis. *PLoS Genetics* **12,** e1005774 (2016).
33. Tataru, P., Mollion, M., Glémin, S. & Bataillon, T. Inference of Distribution of Fitness Effects and Proportion of Adaptive Substitutions from Polymorphism Data. *Genetics* **207,** 1103–1119 (2017).
34. Ng, P. C. & Henikoff, S. SIFT: Predicting Amino Acid Changes That Affect Protein Function. *Nucleic Acids Research* **31,** 3812–3814 (2003).
35. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. & Ng, P. C. SIFT Missense Predictions for Genomes. *Nature Protocols* **11,** 1–9 (1 2016).
36. Chen, J., Bataillon, T., Glémin, S. & Lascoux, M. Hunting for Beneficial Mutations: Conditioning on SIFT Scores When Estimating the Distribution of Fitness Effect of New Mutations. *Genome Biology and Evolution* (2021).
37. Keightley, P. D. & Eyre-Walker, A. What Can We Learn about the Distribution of Fitness Effects of New Mutations from DNA Sequence Data? *Philosophical Transactions of the Royal Society B: Biological Sciences* **365,** 1187–1193 (2010).
38. Rice, D. P., Good, B. H. & Desai, M. M. The Evolutionarily Stable Distribution of Fitness Effects. *Genetics* **200,** 321–329 (2015).
39. Latrille, T., Lanore, V. & Lartillot, N. Inferring Long-Term Effective Population Size with Mutation–Selection Models. *Molecular Biology and Evolution* **38,** 4573–4587 (2021).
40. Huber, C. D., Kim, B. Y., Marsden, C. D. & Lohmueller, K. E. Determining the Factors Driving Selective Effects of New Nonsynonymous Mutations. *Proceedings of the National Academy of Sciences* **114,** 4465–4470 (2017).
41. Starr, T. N. & Thornton, J. W. Epistasis in Protein Evolution. *Protein Science* **25,** 1204–1218 (2016).
42. Park, Y., Metzger, B. P. H. & Thornton, J. W. Epistatic Drift Causes Gradual Decay of Predictability in Protein Evolution. *Science* **376,** 823–830 (2022).

43. Lanfear, R., Kokko, H. & Eyre-Walker, A. Population Size and the Rate of Evolution. *Trends in Ecology and Evolution* **29,** 33–41 (2014).

44. Jones, C. T., Youssef, N., Susko, E. & Bielawski, J. P. Shifting Balance on a Static Mutation–Selection Landscape: A Novel Scenario of Positive Selection. *Molecular Biology and Evolution* **34,** 391–407 (2017).

45. Platt, A., Weber, C. C. & Liberles, D. A. Protein Evolution Depends on Multiple Distinct Population Size Parameters. *BMC Evolutionary Biology* **18,** 1–9 (2018).

46. Kimura, M. Evolutionary Rate at the Molecular Level. *Nature* **217,** 624–626 (1968).

47. Jensen, J. D. *et al.* The Importance of the Neutral Theory in 1968 and 50 Years on: A Response to Kern and Hahn 2018. *Evolution* **73,** 111–114 (2019).

48. Gillespie, J. H. Substitution Processes in Molecular Evolution. III. Deleterious Alleles. *Genetics* **138,** 943–952 (1994).

49. Ohta, T. & Gillespie, J. H. Development of Neutral and Nearly Neutral Theories. *Theoretical Population Biology* **49,** 128–142 (1996).

50. Spielman, S. J. & Wilke, C. O. The Relationship between dN/dS and Scaled Selection Coefficients. *Molecular biology and evolution* **32,** 1097–1108 (2015).

51. Dos Reis, M. How to Calculate the Non-Synonymous to Synonymous Rate Ratio of Protein-Coding Genes under the Fisher-Wright Mutation-Selection Framework. *Biology Letters* **11,** 20141031 (2015).

52. Rodrigue, N., Latrille, T. & Lartillot, N. A Bayesian Mutation-Selection Framework for Detecting Site-Specific Adaptive Evolution in Protein-Coding Genes. *Molecular Biology and Evolution* **38,** 1199–1208 (2021).

53. Latrille, T., Rodrigue, N. & Lartillot, N. Genes and Sites under Adaptation at the Phylogenetic Scale Also Exhibit Adaptation at the Population-Genetic Scale. *Proceedings of the National Academy of Sciences of the United States of America* **120,** e2214977120 (2023).

54. Tamuri, A. U. & dos Reis, M. A Mutation-Selection Model of Protein Evolution under Persistent Positive Selection. *Molecular Biology and Evolution* (2021).

55. Goldstein, R. A. & Pollock, D. D. Sequence Entropy of Folding and the Absolute Rate of Amino Acid Substitutions. *Nature Ecology & Evolution* **1,** 1923–1930 (2017).

56. Stolyarova, A. V. *et al.* Senescence and Entrenchment in Evolution of Amino Acid Sites. *Nature Communications* **11,** 4603 (1 2020).

57. Douzery, E. J. *et al.* OrthoMaM v8: A Database of Orthologous Exons and Coding Sequences for Comparative Genomics in Mammals. *Molecular Biology and Evolution* **31,** 1923–1928 (2014).

58. Yang, Z. & Nielsen, R. Mutation-Selection Models of Codon Substitution and Their Use to Estimate Selective Strengths on Codon Usage. *Molecular Biology and Evolution* **25,** 568–579 (2008).

59. Wright, S. Evolution in Mendelian Populations. *Genetics* **16,** 97–159 (1931).

60. Fisher, R. A. *The Genetical Theory of Natural Selection* (The Clarendon Press, 1930).

61. Al Abri, M. A., Holl, H. M., Kalla, S. E., Sutter, N. B. & Brooks, S. A. Whole Genome Detection of Sequence and Structural Polymorphism in Six Diverse Horses. *PLoS ONE* **15,** e0230899 (2020).

62. Sherry, S. T. *et al.* dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Research* **29,** 308–311 (2001).

63. Svardal, H. *et al.* Ancient Hybridization and Strong Adaptation to Viruses across African Vervet Monkey Populations. *Nature Genetics* **49,** 1705–1713 (2017).

64. Zheng-Bradley, X. *et al.* Alignment of 1000 Genomes Project Reads to Reference Assembly GRCh38. *GigaScience* **6,** gix038 (2017).

65. Keightley, P. D. & Jackson, B. C. Inferring the Probability of the Derived vs the Ancestral Allelic State at a Polymorphic Site. *Genetics* **209,** 897–906 (2018).

66. Ashkenazy, H. *et al.* FastML: A Web Server for Probabilistic Reconstruction of Ancestral Sequences. *Nucleic Acids Research* **40,** W580–W584 (2012).

67. Tataru, P. & Bataillon, T. *polyDFE: Inferring the Distribution of Fitness Effects and Properties of Beneficial Mutations from Polymorphism Data* in *Methods in Molecular Biology* 125–146 (Humana Press Inc., 2020).