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

Thibault Latrille, Nicolas Rodrigue, Nicolas Lartillot

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]

Edited by Rasmus Nielsen, University of California Berkeley, Berkeley, CA; received September 26, 2022; accepted February 11, 2023

Adaptation in protein-coding sequences can be detected from multiple sequence alignments across species or alternatively by leveraging polymorphism data within a population. Across species, quantification of the adaptive rate relies on phylogenetic codon models, classically formulated in terms of the ratio of nonsynonymous over synonymous substitution rates. Evidence of an accelerated nonsynonymous substitution rate is considered a signature of pervasive adaptation. However, because of the background of purifying selection, these models are potentially limited in their sensitivity. Recent developments have led to more sophisticated mutation–selection codon models aimed at making a more detailed quantitative assessment of the interplay between mutation, purifying, and positive selection. In this study, we conducted a large-scale exome-wide analysis of placental mammals with mutation–selection models, assessing their performance at detecting proteins and sites under adaptation. Importantly, mutation–selection codon models are based on a population-genetic formalism and thus are directly comparable to the McDonald and Kreitman test at the population level to quantify adaptation. Taking advantage of this relationship between phylogenetic and population genetics analyses, we integrated divergence and polymorphism data across the entire exome for 29 populations across 7 genera and showed that proteins and sites detected to be under adaptation at the phylogenetic scale are also under adaptation at the population-genetic scale. Altogether, our exome-wide analysis shows that phylogenetic mutation–selection codon models and the population-genetic test of adaptation can be reconciled and are congruent, paving the way for integrative models and analyses across individuals and populations.

adaptation | phylogenetic | population genetics | codon models

Significance

Detecting genes under adaptation represents a key step in the decoding of genomes. Several methods have been proposed, focusing either on the short time scale (population genetics, e.g., human populations) or on the long time scale (phylogenetics, e.g., across mammals). However, the accuracy of these methods is still under debate, and it is still unclear whether the signatures of adaptation are congruent across evolutionary scales. In this study, using phylogenetic methods and gathering genome data across and within species, we show that the signatures of adaptation at the phylogenetic and population-genetic scales can be reconciled. While providing a mutual confirmation of the two approaches, our work paves the way for further methodological integration between microevolutionary and macroevolutionary genomics.

Present-day genetic sequences are informative of populations' past evolutionary history and can carry signatures of selection at different scales. One main goal in the study of molecular evolution is to disentangle and quantify the intensity of neutral, adaptive, and purifying evolution acting on sequences, leveraging variations in sequences between and within species. Theoretically, in order to detect adaptive evolution, one must have data where part of the sequence is known to be under a neutral regime, which can be used as a null model. In the case of protein-coding DNA sequences, synonymous sites are usually taken as proxies for neutral sites, although there are instances where they are indeed under selection (1–3). Nonsynonymous mutations, on the other hand, might be under a mixture of varying degrees of adaptive and purifying selection. Contrasting synonymous and nonsynonymous changes, two different types of methods have emerged to quantify both positive and purifying selection acting on protein-coding sequences. One method, stemming from phylogeny, uses a multiple sequence alignment comprising genes from different species and relies on codon models to deduce the selective regime from the patterns of nonsynonymous versus synonymous substitutions (4, 5). Starting with the work of McDonald and Kreitman (6), another method, stemming from population genetics, contrasts polymorphism within a population and divergence to a closely related species.

At the population-genetic scale, one of the most widely used tests for adaptation relies on the substitutions between two closely related species and polymorphism within one population (6). Under a strict neutral model (i.e., assuming that nonsynonymous mutations are either neutral or strongly selected), the ratios of nonsynonymous polymorphisms over synonymous polymorphisms (π_N/π_S) and of nonsynonymous substitutions over synonymous substitutions (d_N/d_S) are expected to be equal. If, in addition, strongly advantageous mutations occur, they are fixed rapidly in the population, thus contributing solely to divergence but not to polymorphism. As a result, the positive difference between d_N/d_S and π_N/π_S gives an estimate of the adaptive rate $\omega_A = d_N/d_S - \pi_N/\pi_S$ (7). This simple argument is not strictly valid in the presence of moderately deleterious

Author contributions: T.L., N.R., and N.L. designed research; T.L. performed research; T.L. contributed new analytic tools; T.L. analyzed data; and T.L., N.R., and N.L. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: thibault.latrille@ens-lyon.org.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2214977120/-/DCSupplemental.

Published March 10, 2023.

nonsynonymous mutations, which can segregate at a substantial frequency in the population without reaching fixation, thus contributing solely to polymorphism, and not to divergence, potentially resulting in an underestimation of the rate of adaptive evolution (6, 8). Subsequent developments have tried to correct for this effect by relying on an explicit nearly neutral model (9, 10), so as to estimate the rate of evolution expected in the absence of adaptation (called ω_0) based on polymorphism and then to compare it with the rate of evolution, $\omega = d_N/d_S$, to get an estimate of the rate of adaptation as $\omega_A = \omega - \omega_0$.

In their current formulation, phylogeny-based methods rely on the ratio of nonsynonymous substitutions over synonymous substitutions, called ω (4, 5). Since this ratio is a contrast between the rates before and after the action of selection on the protein, it thus provides an estimate of the strength of selection exerted at the amino acid level, whether synonymous changes are driven solely by mutation or by both mutation and selection (11). In this context, an excess in nonsynonymous substitutions, leading to $\omega > 1$, indicates that the protein is undergoing recurrent positive selection for amino acid changes, meaning that the protein is putatively under adaptive evolution. Conversely, a deficit in nonsynonymous substitutions, leading to $\omega < 1$, means the protein is under purifying selection. In practice, proteins are typically under a mix of adaptive and purifying selection dominated by the latter, thus typically leading to an $\omega < 1$ even in the presence of positive selection. At a finer scale, site models can detect a specific site (i) of the sequence with a $\omega^{(i)} > 1$ (12, 13). Site models have the advantage of greater sensitivity and the ability to pinpoint where positive selection acts on the protein. However, even at the level of a single site under recurrent adaptation, not all amino acids are expected to be adaptive, leading to $\omega^{(i)}$ capturing a mix of adaptive and purifying selection, reducing the sensitivity of the test. An alternative approach to detect adaptation would be to rely on an explicit nearly neutral model as the null against which to detect deviations, similarly to the McDonald and Kreitman test. As a recent development in this direction, the so-called phylogenetic mutation–selection models provide a null model by estimating the fitness landscape over amino acid sequences, for each site of the sequence (11, 14, 15). At the mutation–selection balance, the probability for a specific codon to be fixed in the population is proportional to its fitness, and a mutation from a high-fitness amino acid toward a low-fitness amino acid will have a small probability of fixation, genuinely accounting for purifying selection. Conversely, only nearly neutral mutations between high-fitness amino acids will tend to be permitted by the model, allowing for the explicit calculation of the nearly neutral rate of nonsynonymous substitutions at mutation–selection balance, called ω_0 (16, 17). By contrasting ω estimated by ω-based codon models and ω_0 calculated from mutation–selection models, one can hope to extract the rate of adaptation $\omega_A^{phy} = \omega - \omega_0$.

Interestingly, the rate of adaptation is directly comparable between phylogenetic and population-genetic methods since both seek a deviation of ω from a nearly neutral null model, estimated with mutation–selection models in a phylogenetic context (ω_0) or from standing polymorphism in a population-genetic context (π_N/π_S). This raises the question of whether the two signals of adaptation are correlated, thus representing a unique opportunity to confront phylogeny-based and population-based methods. These two methods work over very different time scales; for that reason, they might be capturing different signals: long-term evolutionary Red-Queen for phylogeny-based methods versus events of adaptation in specific lineages for population-based methods. Nonetheless, we expect sites and proteins under long-term evolutionary Red-Queen regimes to maintain their signal of adaptation in several independent lineages for which the McDonald and Kreitman test is applied.

Accordingly, in this study, we first applied ω-based and mutation–selection codon models to whole exome data from placental mammals, so as to quantify the rate ω_A^{phy} for each site and protein and detect signatures of adaptive evolution at the phylogenetic scale. Then, we developed a pipeline integrating (and aligning) divergence and polymorphism data across the entire exome for 29 populations across 7 genera, namely, *Equus*, *Canis*, *Bos*, *Capra*, *Ovis*, *Chlorocebus*, and *Homo*. Finally, using this pipeline, we assessed the congruence between the phylogeny-based and population-based approaches, by testing whether the group of sequences detected with a high rate of adaptation in the phylogeny-based method also displays a high rate of adaptation according to the population-based method.

Results

Detecting Genes and Sites under Adaptation. We derived a two-step approach (*Methods*), which we applied to a set of alignments of orthologous genes at the scale of placental mammals. The d_N/d_S estimated by the site model (ω) is plotted against the d_N/d_S predicted by the nearly neutral mutation–selection model (ω_0) for genes (scatter plot in Fig. 1*A*) and sites (density plot in Fig. 1*B*). An excess of ω relative to ω_0 is a typical signature of ongoing positive selection (17, 18). For such comparison to be valid, both models estimating ω and ω_0 should have assumptions as similar as possible regarding the underlying mutation process, modeled as a Muse and Gaut (4, 16). As a control, we validated our Bayesian estimates of ω against their maximum-likelihood counterpart (*Methods* and *SI Appendix*, Fig. S1).

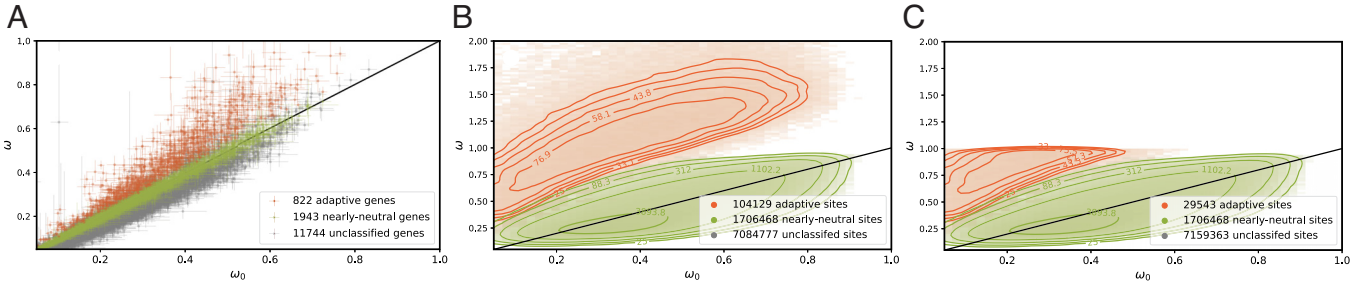

Fig. 1. Detection of protein-coding sequences ongoing adaptation at the phylogenetic scale. ω estimated by the site model against ω_0 calculated by the mutation–selection model. Scatter plot of 14,509 genes in panel *A*, with a 95% Bayesian credible interval ($\alpha = 0.05$). Density plot of sites in panels *B* and *C*. Genes and sites are then classified as adaptive ($\omega > \omega_0$ in red) or nearly neutral ($\omega \simeq \omega_0$ in green). In panel *C*, the set of sites detected exclusively by mutation–selection codon models have a mean $\omega < 1$.

Given these validations, an excess of ω compared to ω_0 can be retrieved as a ratio ($\omega/\omega_0 > 1$) or as a difference ($\omega - \omega_0 > 0$). However, since ω_0 can be close to 0 for genes and sites strongly constrained (Fig. 1), ratios can be particularly large and mathematically undefined, ultimately leading to artifacts. Moreover, comparing the posterior credibility intervals of ω and ω_0 as a difference rather than a ratio is more intuitive and visual (Fig. 1*A*). Altogether, genes or sites were considered to be under an adaptive regime (in red) if the value of their ω is higher than that of their ω_0, with nonoverlapping 95% posterior credibility intervals. This procedure retrieved 822 out of 14,509 genes, which are putatively under a long-term evolutionary Red-Queen regime. At the site level, the nearly neutral assumption appears to be rejected for 104,129 out of 8,895,374 sites. Of note, this selection procedure is not meant as a routine statistical test, but only as an enrichment procedure, for the needs of the subsequent analysis shown below. In practice, this selection is likely to be conservative and to have a rate of false discovery of the order of 1% at the gene level and 5% at the site level (*Methods*).

Remarkably, selection based on $\omega > \omega_0$ is more sensitive than based on the commonly used criterion of $\omega > 1$ since ω_0 is always lower than 1 by definition (16). Thus, we can uncover sites under adaptation ($\omega > \omega_0$) with a mean ω lower than 1 (29,543 sites in Fig. 1*C*). These sites could not have been detected by ω-based codon models relying on the criterion that $\omega > 1$. At the gene level, only two genes have an estimated $\omega > 1$ such that this distinction is not relevant.

Ontology Enrichment Tests. Next, we investigated whether the genes classified as adaptive ($\omega > \omega_0$) showed enrichment in specific ontology terms. Thus, we performed 775 instances of Fisher's exact test to estimate ontology enrichment by contrasting with genes in the control group, not classified as adaptive. Forty-two ontologies are observed with a p-value (p_v) corrected for multiple comparison (Holm–Bonferroni correction, p_v^{adj}) lower than the risk $\alpha = 0.05$ (*SI Appendix*, Table S1). At a finer scale, we weighted genes by their proportion of sites considered under adaptation with a ω-based site model ($\omega > 1$, *SI Appendix*, Table S2) or with a mutation–selection model ($\omega > \omega_0$, *SI Appendix*, Table S3). For each ontology, the proportion of sites under adaptation is compared between the set of genes sharing this given ontology and the rest of the genes (Mann–Whitney U test). The statistical test based on the first criterion ($\omega > 1$) is correlated with ontologies related to immune processes, while the statistical test based on the second criterion ($\omega > \omega_0$) is also correlated with ontologies related to the external membrane and cellular adhesion.

Congruence Between Phylogeny- and Population-Based Methods. Finally, we investigated whether the phylogeny-based and the population-based methods give congruent results in terms of detection of adaptive evolution (Fig. 2). To do so, population genomic data were collected for 29 populations across 7 genera (*SI Appendix*, Fig. S2). For each population, ω_A based on the McDonald and Kreitman (MK) test (6) was computed on the concatenated sequence of the 822 genes classified as adaptive by the phylogeny-based method (red dots in Figs. 2 and 3). This result was compared to a null distribution obtained by computing ω_A over sets of 822 genes that were randomly sampled (1,000 replicates) among the genes classified as nearly neutral according to the mutation–selection model (green violins in Figs. 2 and 3). Importantly, the terminal lineages over which the population-genetic method was applied were not included in the phylogenetic analysis so as to avoid fallacy and circularity in the estimation of ω_A and ω_A^{phy}. As a result, the two methods are working on entirely nonoverlapping compartments of the evolutionary history across mammals. For all 29 populations, the ω_A estimated by the population-genetic method was significantly higher for the putatively adaptive gene set than for the putatively nearly neutral gene sets of the same size (at a risk $\alpha = 0.05$ corrected for multiple testing, Holm–Bonferroni correction). There is thus a good qualitative agreement between the two methods as to what they capture and interpret as positive selection at the gene level.

The same procedure was applied at a finer scale with sites instead of genes. For each population, ω_A was computed on the concatenated sequence of the 104,129 sites classified as

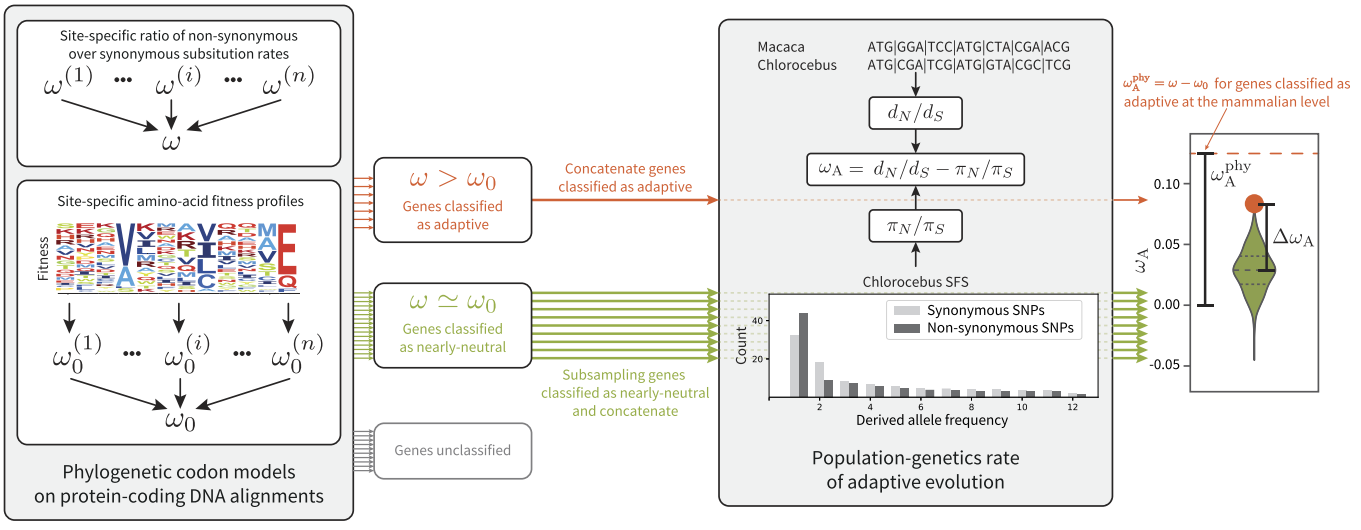

Fig. 2. Integrating divergence and polymorphism for the detection of adaptation. At the phylogenetic level, ω (classical codon models) and ω_0 (mutation–selection codon models) are computed from protein-coding DNA alignments, allowing us to classify genes into adaptive (in red) and nearly neutral (in green) regime. At the population-genetic level, for each population, ω_A is computed on the concatenated sequence of genes classified as under adaptation. The result is compared to the empirical null distribution of ω_A in each population, obtained by randomly sampling (1,000 replicates) a subset under a nearly neutral regime.

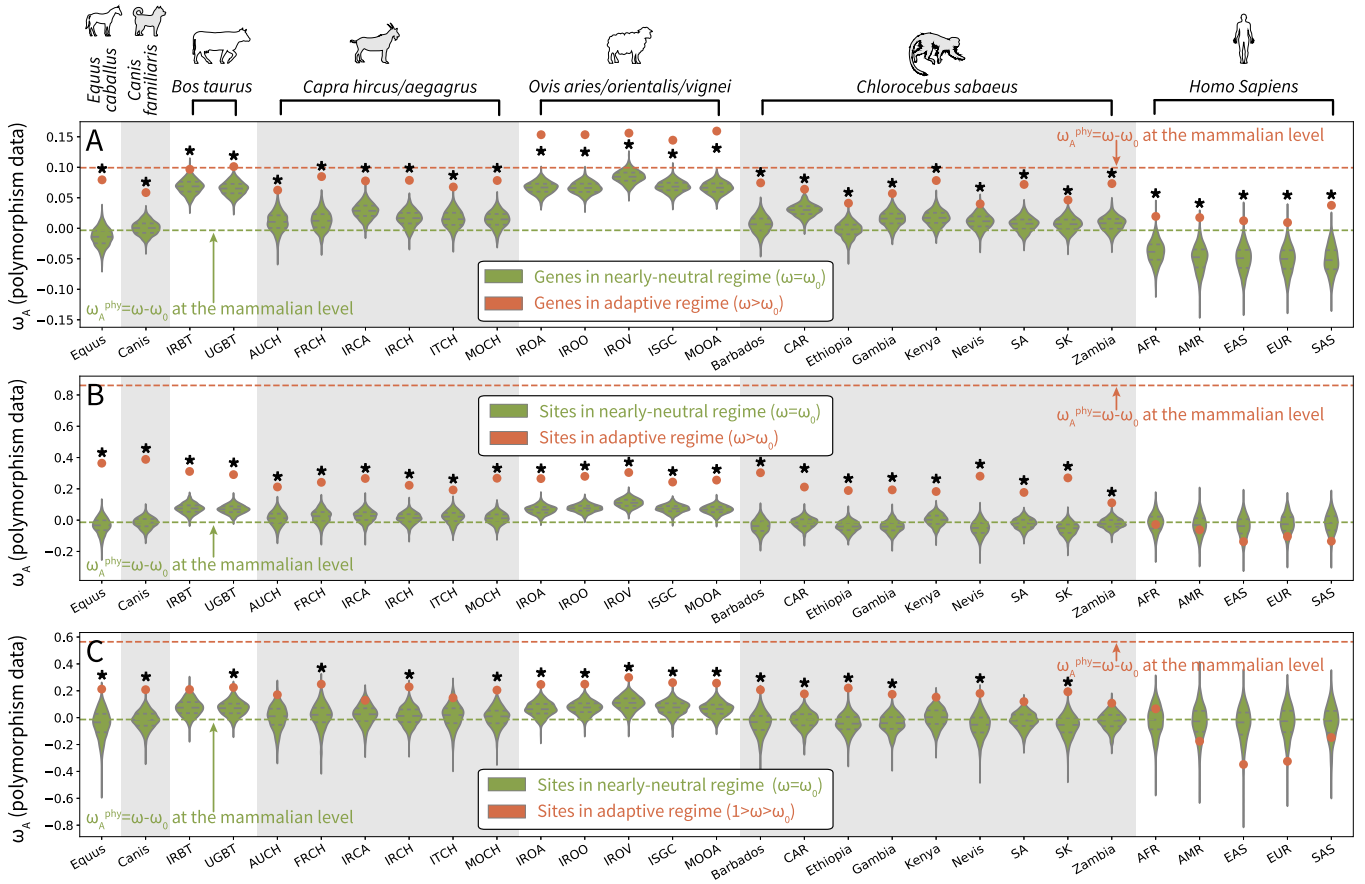

Fig. 3. Enrichment of adaptation at the population-genetic scale for 29 populations across 7 genera at the gene (panel *A*) and site (panels *B* and *C*) level. For each population, ω_A is computed on 822 genes (*A*) and 104,129 sites (*B*) having a high rate of adaptation at the phylogenetic scale ($\omega > \omega_0$ in red). In panel *C*, the set of 29,543 sites are detected exclusively by mutation–selection codon models with a mean $\omega < 1$. The result is compared to the empirical null distribution of ω_A, obtained by randomly sampling (1,000 replicates) a subset of genes and sites under a nearly neutral regime (violin plot in green). * signify that the p_V corrected for multiple comparison (Holm–Bonferroni correction) is lower than the risk $\alpha = 0.05$. The acronym of populations and the quantitative value of ω_A and p_V are shown in Table 1.

adaptive by the phylogeny-based method and compared to the empirical null distribution (Fig. 3*B*) and Table 1. Out of 29 populations, 24 have an ω_A estimated by the population-genetic method significantly higher for the putatively adaptive site-set than for the putatively nearly neutral site-sets of the same size taken at random (at a risk $\alpha = 0.05$ corrected for multiple testing, Holm–Bonferroni correction). In particular, the five populations for which the test is not significant are the human populations. Coincidentally, this is the species with the lowest genetic diversity, resulting in underpowered statistical tests due to fewer polymorphisms available per loci. This statistical effect is combined with the biological effect of increased genetic drift, owing to the small effective population size in ancestral human populations, resulting in less efficient selection. Altogether, we argue that finding evidence for positive selection in human populations is statistically more challenging due to both less efficient selection and few data available.

Except for *Equus* and *Humans*, on average, the ω_A returned by MK is positive even for the putatively nearly neutral replicates and significantly so for *Bos* (ω_A in the range 0.65 to 0.68 for genes and sites) and *Ovis* (ω_A in the range 0.66 to 0.84 for genes and sites). This suggests the presence of a background of positive selection captured by MK methods but not by phylogenetic methods. This background signal could correspond either to adaptation specifically present in the terminal lineages on which the MK method is applied and absent over the rest of the mammalian tree or to low-intensity recurrent positive selection, present over the tree but nevertheless missed by phylogenetic methods, owing to a lack of sensitivity. Alternatively, part of it could be an artifact of MK methods due, for example, to a recent demographic expansion (*Bos* and *Ovis* are the two among those analyzed by the population-genetic approach showing the highest levels of synonymous diversity) or to a more general mismatch between short- and long-term effective population size (N_e) (19).

Regardless of its exact cause, subtracting this background may give a more meaningful basis for a quantitative comparison between phylogenetic and population-genetic approaches (Fig. 2). To do so, we can compare the $\Delta\omega_A$ between the putatively adaptive set and the control replicates to the $\omega_A^{phy} = \omega - \omega_0$ returned by the phylogenetic method. Of note, across all analyses shown in Fig. 3 *A* and *B*, this population-genetic $\Delta\omega_A$ is always smaller than the phylogenetic ω_A^{phy}. This asymmetry is expected as a result of a selection bias: The genes of the test set were selected precisely for their high phylogenetic signal, while keeping a blind eye to their population-genetic signal. From this perspective, the ratio $\Delta\omega_A/\omega_A^{phy}$ can be interpreted as an estimate of the fraction of the total signal captured by the phylogenetic enrichment procedure that is confirmed by MK statistics. This ratio, hereafter called the confirmation rate, is indicated in Table 1.

At the gene level, the confirmation rate is relatively high, ranging from 30% to 90%. At the site level, the confirmation rate

Table 1. Across 29 populations (rows), table of quantitative value of $\Delta\omega_A$ between the set classified as adaptive and nearly neutral shown in Fig. 3

			Genes (822)			Sites (104,129)			Sites ($\omega < 1$) (29,543)		
Population	Species	π_S	$\Delta\omega_A$	p_v^{adj}	$\frac{\Delta\omega_A}{\omega_A^{phy}}$	$\Delta\omega_A$	p_v^{adj}	$\frac{\Delta\omega_A}{\omega_A^{phy}}$	$\Delta\omega_A$	p_v^{adj}	$\frac{\Delta\omega_A}{\omega_A^{phy}}$
Diverse (Equus)	Equus caballus	0.002	0.094	**0.0***	0.928	0.399	**0.0***	0.459	0.258	**0.0***	0.446
Diverse (Canis)	Canis familiaris	0.004	0.058	**0.0***	0.557	0.406	**0.0***	0.463	0.227	**0.0***	0.392
Iran (IRBT)	Bos taurus	0.008	0.028	**0.020***	0.278	0.237	**0.0***	0.272	0.134	0.150	0.231
Uganda (UGBT)	Bos taurus	0.008	0.036	**0.0***	0.355	0.222	**0.0***	0.254	0.156	**0.017***	0.270
Australia (AUCH)	Capra hircus	0.003	0.052	**0.0***	0.506	0.202	**0.0***	0.230	0.168	0.143	0.290
France (FRCH)	Capra hircus	0.003	0.073	**0.0***	0.709	0.220	**0.0***	0.250	0.236	**0.039***	0.407
Iran (IRCA)	Capra aegagrus	0.004	0.049	**0.0***	0.482	0.242	**0.0***	0.275	0.108	0.396	0.186
Iran (IRCH)	Capra hircus	0.004	0.062	**0.0***	0.610	0.210	**0.0***	0.239	0.217	**0.017***	0.375
Italy (ITCH)	Capra hircus	0.003	0.052	**0.0***	0.511	0.174	**0.0***	0.199	0.134	0.308	0.232
Morocco (MOCH)	Capra hircus	0.004	0.064	**0.0***	0.626	0.256	**0.0***	0.292	0.201	**0.0***	0.347
Iran (IROA)	Ovis aries	0.007	0.087	**0.0***	0.847	0.199	**0.0***	0.228	0.183	**0.017***	0.316
Iran (IROO)	Ovis orientalis	0.009	0.087	**0.0***	0.848	0.204	**0.0***	0.233	0.176	**0.0***	0.304
Iran (IROV)	Ovis vignei	0.005	0.072	**0.0***	0.697	0.194	**0.0***	0.222	0.192	**0.0***	0.332
Various (ISGC)	Ovis aries	0.008	0.076	**0.0***	0.742	0.171	**0.0***	0.195	0.189	**0.0***	0.326
Morocco (MOOA)	Ovis aries	0.008	0.093	**0.0***	0.905	0.189	**0.0***	0.216	0.193	**0.0***	0.333
Barbados	Chlorocebus sabaeus	0.003	0.068	**0.0***	0.665	0.341	**0.0***	0.390	0.248	**0.0***	0.430
Central African Republic (CAR)	Chlorocebus sabaeus	0.006	0.034	**0.0***	0.334	0.229	**0.0***	0.262	0.195	**0.0***	0.338
Ethiopia	Chlorocebus sabaeus	0.005	0.044	**0.0***	0.425	0.231	**0.0***	0.264	0.264	**0.0***	0.457
Gambia	Chlorocebus sabaeus	0.005	0.041	**0.0***	0.403	0.236	**0.0***	0.270	0.217	**0.0***	0.375
Kenya	Chlorocebus sabaeus	0.004	0.061	**0.0***	0.598	0.181	**0.0***	0.207	0.152	0.150	0.264
Nevis	Chlorocebus sabaeus	0.003	0.029	**0.020***	0.279	0.332	**0.0***	0.380	0.237	**0.017***	0.410
South Africa (SA)	Chlorocebus sabaeus	0.006	0.065	**0.0***	0.633	0.199	**0.0***	0.228	0.142	0.108	0.246
Saint Kitts (SK)	Chlorocebus sabaeus	0.004	0.040	**0.0***	0.388	0.324	**0.0***	0.371	0.253	**0.0***	0.439
Zambia	Chlorocebus sabaeus	0.006	0.066	**0.0***	0.642	0.132	**0.0***	0.151	0.131	0.150	0.227
African (AFR)	Homo sapiens	0.002	0.059	**0.012***	0.568	−0.010	1.000	−0.012	0.089	1.000	0.155
Ad Mixed American (AMR)	Homo sapiens	0.002	0.067	**0.006***	0.647	−0.029	1.000	−0.034	−0.141	1.000	−0.244
East Asian (EAS)	Homo sapiens	0.002	0.063	**0.006***	0.610	−0.096	1.000	−0.111	−0.296	1.000	−0.513
European (EUR)	Homo sapiens	0.002	0.061	**0.015***	0.590	−0.078	1.000	−0.089	−0.289	1.000	−0.500
South Asian (SAS)	Homo sapiens	0.002	0.089	**0.0***	0.866	−0.113	1.000	−0.130	−0.111	1.000	−0.193

p_v^{adj} associated to the test are corrected for multiple comparison (Holm–Bonferroni correction, bold numbers and * for $p_v^{adj} < 0.05$). $\frac{\Delta\omega_A}{\omega_A^{phy}}$ is the ratio of $\Delta\omega_A$ at the population-genetic level and the phylogenetic level. π_S is the observed genetic diversity (number of single nucleotide polymorphisms per site) counted over synonymous sites.

is lower (30% on average), which could betray a higher rate of false discovery at the site level, or could be the result of subtle molecular evolutionary processes, such as intermittent adaptation (on some but not on all branches) or within-gene turnover (ongoing adaptation targeting different sites on different branches).

After discarding sites with a mean $\omega > 1$, the remaining 29,543 sites classified as being under an adaptive regime have $1 > \omega > \omega_0$ and are specifically discovered by the mutation–selection approach. Since their ω is less than 1, they could not be detected by classical codon models. This raises the question of the empirical value of these findings. Indeed, while mutation–selection methods are more sophisticated and may therefore have a greater sensitivity, they may also be more prone to producing false positives. The phylogenetic/population-genetic contrast developed here can be used to assess this important point. As shown in (Fig. 3*C*) and Table 1, out of 29 populations, for 17 out of the 29 populations that have been analyzed, the confirmation rate is significantly positive ($\alpha = 0.05$, Holm–Bonferroni correction) and of the order of 10% on average. This importantly suggests the presence of a background of low-intensity positive selection, which is missed by classical codon models, but partially detected by mutation–selection models. In other words, the approach can detect a long-term evolutionary Red-Queen even for a site with $\omega < 1$ that is still under adaptation at the population-genetic scale.

Because genes and sites classified as adaptive have a higher ω than genes/sites classified as nearly neutral, ω_A could simply be higher for genes with higher ω due to this confounding factor. Thus, we performed additional experiments where ω is controlled to be the same in the nearly neutral replicate and the adaptive set of genes (*SI Appendix*, Figs. S3–S7 and Tables S5–S9). Additionally, we performed the same experiments with a more stringent risk $\alpha = 0.005$ (10 times greater) to classify genes and sites as adaptive (*SI Appendix*, Figs. S8–S10 and Tables S9–S10). Our results are robust to both controlling for ω and with a different threshold to classify genes and sites as adaptive. Finally, we computed ω_A using the software polyDFE (20), which relies on the synonymous and nonsynonymous unfolded site-frequency spectra (SFS) to estimate the distribution of fitness effects of mutations (DFE) and the rate of adaptation (*SI Appendix*, Fig. S11–S18 and Tables S11–S18). Depending on the underlying assumptions for the shape of the DFE and the definition of ω_A, we observed a wide range of ω_A both for the set of adaptive and nearly neutral genes/sites. However, the

statistical test for the enrichment of ω_A between the set of adaptive and nearly neutral genes/sites gives results in the same direction whether computed by polyDFE or as McDonald and Kreitman statistic (6), although the confirmation rate and the associated p_v are different.

Discussion

Quantifying the rate of adaptation assumes that we can measure the rate of evolution and, more importantly, its deviation from a null model of evolution disallowing adaptation. For phylogenetic codon models, this null model of evolution is usually assumed to be neutral evolution and the rate of evolution computed as the ratio of nonsynonymous over synonymous substitution rates (ω) is thus compared to 1. We first showed that, at the phylogenetic scale, ω can be compared to its expectation under the mutation–selection model (ω_0), a nearly neutral model instead of a neutral model of evolution, giving a quantitative estimate of the rate of adaptation as $\omega_A^{phy} = \omega - \omega_0$. Moreover, because ω_0 is by definition lower than 1 (16), detecting adaptation as $\omega > \omega_0$ instead of $\omega > 1$ is always a more sensitive statistical test. We thus argue that mutation–selection provides both a theoretical and practical improvement over the current method of testing to detect adaptation. The application of this approach exome-wide across placental mammals suggests that 822 out of 14,509 proteins are under a long-term evolutionary Red-Queen, with ontology terms related to immune processes and the external membrane of cells. Enrichment in ontologies related to immune processes is expected, as found in previous studies (13). However, we also detect enrichment in ontologies related to the external membrane and cell adhesion. Genes pertaining to these ontologies might be under adaptation as a consequence of being the target of viruses and parasites (21, 22). Altogether, the mutation–selection method effectively detects adaptation regardless of the background of purifying selection and returns reasonable candidates for adaptive evolution. Of note, in its current implementation, and unlike classical codon models (23, 24), the mutation–selection approach does not yet provide a proper and well-calibrated statistical test for calling genes or sites under adaptation with a well-controlled frequentist risk. This was not a problem in the enrichment analysis conducted in this article, which relies on downstream controls based on random permutations. Nevertheless, the encouraging results obtained here give motivation for developing such a test, which should then have an increased power to detect adaptation, compared to classical codon models relying on the $\omega > 1$ criterion.

At the population-genetic scale, the availability of approaches to detect adaptation (6, 25) raises the question of whether the rate of adaptation calculated at the phylogenetic scale as ω_A^{phy} is congruent with the rate calculated at the population genetics scale by McDonald and Kreitman (MK) (6) as $\omega_A = d_N/d_S - \pi_N/\pi_S$. In this light, the set of genes and sites detected to be under adaptation at the phylogenetic scale showed a significant increase in ω_A such as inferred by the population-based method (29 populations across 7 genera). Quantitatively, about 30 to 90% of the signal detected by the phylogeny-based approach is confirmed by the population-based approach. This result is in stark contrast with studies comparing ω-based codon models at the gene level with MK methods, which found that the set of genes detected at different scales does not seem to overlap beyond random expectations (26). The reasons for this discrepancy are not totally clear. The use of different codon modeling strategies could play a role here. More fundamentally however, our study relies on a large and densely sampled phylogeny with $\simeq$100 taxa across placental mammals, versus 5 *Drosophila* and 5 *Brassicaceae* in Chen et al. (26). As a result, the phylogenetic aspect of our analysis benefits from an increased power while being also inherently more focused on genes characterized by recurrent adaptation over a very large evolutionary scale (i.e., long-term evolutionary Red Queens), for which population-genetic signals of adaptation may be more easily recovered. We thus showed empirically that the mutation–selection codon model provides a null (nearly neutral) model from which we can disentangle purifying and adaptive evolution. However, our procedure still has some limitations.

First, synonymous mutations are considered neutral in our implementation of mutation–selection codon models. In other words, we assumed no synonymous codon usage bias (CUB) or a weak effect as it is suggested in mammals (27). However, a strong CUB can influence the comparison between ω and ω_0 (16). In our mutation–selection codon model, in addition to modeling selection for amino acids at the site level, modeling CUB is possible by introducing another layer of selection for synonymous codons at the gene level (11, 15). Such a model could be used to detect adaptation in a more general context for which CUB cannot be overlooked (e.g., *Drosophila*). Further work will thus be required to relax the assumption that synonymous mutations are neutral and to assess the sensitivity of ω_A^{phy} to CUB. Second, mutation–selection codon models assume a constant effective population size, while it has been established that its fluctuations have a major effect on selection dynamics (28, 29). Estimating changes in effective population size in a mutation–selection framework is possible (30), although too computationally intensive in its current implementation to be performed genome-wide. Third, epistasis is not modeled while it can have a large effect on the response of the rate of evolution with change in population size (31). More generally, pervasive epistasis generates an entrenchment of the amino acids (32–34), resulting in slowing down of the rate of evolution (17, 35) or a standstill (36). Consequently, our estimation of the predicted rate of evolution computed at mutation–selection balance (ω_0) is overestimated given that epistasis is not taken into account in our model. Such an overestimated ω_0 results in an underestimated rate of adaptation $\omega_A^{phy} = \omega - \omega_0$, ultimately leading to a conservative test of adaptation which could theoretically be improved. Accurately modeling the underlying fitness landscape is thus a direction for future research in mutation–selection models to improve the estimation of ω_A^{phy}, by means of relaxing model assumptions (34) or by leveraging empirically determined fitness landscapes (18, 37).

On the other hand, at the population-genetic scale, the greatest limitation to detecting adaptation is the lack of power determined by the genetic diversity since polymorphisms are rare and estimation of π_N/π_S requires pooling many sites for which variations are available. Since the effects of mild purifying selection are more pronounced on longer time scales (i.e., mildly deleterious mutations contribute disproportionately to polymorphism, compared to divergence), ω_A as computed by MK can be biased by moderately deleterious mutations (8, 38) and by the change in population size through time (39). To overcome this bias, model-based approaches relying on the synonymous and nonsynonymous site-frequency spectra (SFS) to estimate the distribution of fitness effects of mutations (DFE), so as to account for the contribution of mild selective effects to standing polymorphism, have been developed (9, 20) and are often used (10, 40). However, the broad range of ω_A estimated on sets of genes/sites classified as nearly neutral suggests that

these models are lacking power, even more than the MK statistic, because of the sparsity of the SFS. Besides changes in population size biasing the estimation (19), we argue that inferring ω_A using an underlying DFE model is also highly sensitive to assumptions for the shape of the DFE and the definition of ω_A. For example, the value of ω_A is computed as an integral, where the bounds of this integral are debated by different authors (10, 41). It is thus relatively easy to change the definition of ω_A (*SI Appendix*, Figs. S13–S16 and Tables S13–S16) or to constrain the underlying DFE (*SI Appendix*, Figs. S13–S18 and Tables S13–S18) to obtain a wide range of ω_A on the same dataset. Taken together, we argue that comparing ω_A to 0 is not a robust test for adaptation. Instead, ω_A for a particular genomic region of interest should be compared to other genomic regions for which the nearly neutral evolution is not rejected, and the difference $\Delta\omega_A$ should be compared to 0, as done in this study. More generally, our empirical analysis emphasizes the limitations of, and the difficulties raised by, the model-based population-genetic approaches. In this respect, further exploring the congruence (or lack thereof) between phylogenetic and population-genetic approaches will represent a useful asset to clarify those delicate problems, given that similar benefits are also expected on the side of phylogenetic approaches, which are far from immune from methodological limitations.

More broadly on a theoretical level, this work leverages a specific overlap between phylogenetic and population genetics, namely that the rate of adaptation ω_A^{phy} in phylogenetic codon models and ω_A in the MK test should theoretically be directly comparable. Based on this theoretical relationship, our study is paving the way for studies and methods augmenting molecular polymorphism data within species with information about divergence data between species (42) and by assessing empirically the relationship between phylogenetic and population genetics analyses (43). In this light, mutation–selection models at the phylogenetic scale can play a dual role: pinpointing genes and sites under adaptation ($\omega_A^{phy} > 0$) and also seeking the genomic region for which the nearly neutral theory is not rejected ($\omega_A^{phy} \simeq 0$).

Materials and Methods

Phylogenetic Dataset. Protein-coding DNA sequences alignments in placental mammals and their corresponding gene trees were extracted from the OrthoMaM database, containing 116 mammalian reference sequences in v10c (44–46). Genes located on the X, Y, and mitochondrial chromosomes were discarded from the analysis since the number of polymorphisms, necessary in the population-based method, is expected to be different on these sequences. Additionally, sequences from the species for which polymorphism are available as well as their sister species have been discarded from the analysis to ensure independence between the data used in the phylogenetic and population-genetic method, so as to avoid fallacy and circularity in the estimation of ω_A and ω_A^{phy}. Altogether, we analyzed 14,509 protein-coding DNA sequences alignment containing at most 87 reference sequences of placental mammals.

Adaptation in Phylogeny-Based Method. Classical codon models estimate a parameter $\omega = d_N/d_S$, namely, the ratio of the nonsynonymous over the synonymous substitution rates (4, 5). In the so-called site models, ω is allowed to vary across sites (12, 47). In *Bayescode*, site-specific $\omega^{(i)}$ (Fig. 1*B*, y-axis) are modeled as independent and identically distributed random effects from a gamma distribution (48). Therefore, this corresponds to a continuous version of the M5 model of Yang et al. (12). In a second step, the average over sites is calculated, giving estimates of ω for each protein-coding sequence (Fig. 1*A*, y-axis). To assess the sensitivity of our estimate of ω to the prior as implemented in *BayesCode*, we ran the *CODEML* site model (F1x4 MG, 8 discrete categories of ω) from the PAML software (24) on a sample of 100 random genes. Both ω estimates from *BayesCode* and *CODEML* are strongly correlated at the site level ($r^2 = 0.985$, slope of 0.92 in *SI Appendix*, Fig. S1*A*) and at the gene level ($r^2 = 0.991$, slope of 0.9 in *SI Appendix*, Fig. S1*B*).

In contrast to ω-based codon models, mutation–selection models assume that the protein-coding sequence is at mutation–selection balance under a fixed fitness landscape, which is itself characterized by a fitness vector over the 20 amino acids at each site (11, 14, 15). Mathematically, the rate of nonsynonymous 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 from the underlying DNA change ($\mu_{a\mapsto b}$) multiplied by the scaled probability of fixation of the mutation ($\mathbb{P}^{(i)}_{a\mapsto b}$). Crucially, the probability of fixation depends on the difference of scaled fitness between the amino acid encoded by the mutated codon ($F^{(i)}_b$) and the fitness of the amino acid encoded by the original codon ($F^{(i)}_a$) at a given site i (49–51). Altogether, the rate of substitution from codon a to b at a given site i is

$$q^{(i)}_{a\mapsto b} = \mu_{a\mapsto b}\mathbb{P}^{(i)}_{a\mapsto b} = \mu_{a\mapsto b}\frac{F^{(i)}_b - F^{(i)}_a}{1 - e^{F^{(i)}_a - F^{(i)}_b}}. \quad [1]$$

Fitting the mutation–selection model on a sequence alignment leads to an estimation of the mutation rate matrix ($\boldsymbol{\mu}$) as well as the 20 amino acid fitness landscape ($\boldsymbol{F}^{(i)}$) at each site i. From these parameters, one can compute $\omega_0^{(i)}$ (Fig. 1*B*, x-axis), the site-specific rate of nonsynonymous over synonymous substitution at the mutation–selection balance:

$$\omega_0^{(i)} = \frac{\sum_{a\in\mathcal{C}}\sum_{b\in\mathcal{N}_a}\pi_a^{(i)}q^{(i)}_{a\mapsto b}}{\sum_{a\in\mathcal{C}}\sum_{b\in\mathcal{N}_a}\pi_a^{(i)}\mu_{a\mapsto b}}, \quad [2]$$

where $\mathcal{C}$ is the set of all the possible codons (61 by discarding stop codons), $\pi_a^{(i)}$ is the equilibrium frequency of codon a at site i, and $\mathcal{N}_a$ is the set of codons that are nonsynonymous to a (16, 17). The equilibrium frequency of codon a at site i is a function of the nucleotide frequencies at its three positions and the amino acid scaled fitness ($F^{(i)}_a$):

$$\pi_a^{(i)} = \frac{\sigma_{a[1]}\sigma_{a[2]}\sigma_{a[3]}e^{F^{(i)}_a}}{\sum_{b\in\mathcal{C}}\sigma_{b[1]}\sigma_{b[2]}\sigma_{b[3]}e^{F^{(i)}_b}}, \quad [3]$$

where $\sigma_{a[j]}(a[j] \in \{A, T, C, G\})$ is the equilibrium frequency (given by the mutational matrix) of the nucleotide at position $j \in \{1, 2, 3\}$ of codon a. In a second step, the average over sites is calculated, giving estimates of ω_0 for each protein-coding sequence (Fig. 1*A*, x-axis). Under the assumption that the protein is under a nearly neutral regime, the calculated ω_0 (mutation–selection model) and the estimated ω (site model) should be the same (16).

We ran the Bayesian software *BayesCode* (https://github.com/ThibaultLatrille/bayescode) on each protein-coding DNA alignment (52). Each Monte Carlo Markov chain (MCMC) is run during 2,000 points, with a burn-in of 1,000 points, to obtain the posterior mean of ω and ω_0 across the MCMC as well as the 95% posterior credibility interval for genes and sites. Genes and sites classified under an adaptive regime (in red) are rejecting the nearly neutral assumption such that the lower bound for the credible interval of ω ($\alpha = 0.05$) is above the upper bound of the credible interval of ω_0 ($\alpha = 0.05$), meaning that the value of their ω is higher than that of their ω_0. Because this is a unilateral test ($\omega > \omega_0$) and the two credible intervals are independent, the risk is $(\alpha/2)^2 = 0.025^2 = 0.000625$ for each test. Empirically, the nearly neutral assumption appears to be rejected for 822 out of 14,509 genes, while $0.000625 \times 14{,}509 \simeq 9$ genes are expected due to the multiple testing, suggesting a $9/822 \simeq 1\%$ rate of false positive at the gene level. At the site level, the nearly neutral assumption appears to be rejected for 104,129 out of 8,895,374 sites, while $0.000625 \times 8{,}895{,}374 \simeq 5{,}560$ are expected due to the multiple testing, suggesting a $5{,}560/104{,}129 \simeq 5\%$ rate of false positive

at the site level. Genes and sites are classified under a nearly neutral regime (in green) if the average ω is within the credible interval of the ω_0, and respectively, the average ω_0 is also within the credible interval of ω, meaning $\omega = \omega_0$. Additionally, the set of sites detected exclusively by mutation–selection codon models have a mean $\omega < 1$. Genes and sites that do not fall under any of these categories are considered unclassified.

Polymorphism Dataset. Each Single Nucleotide Polymorphism (SNP) (chromosome, position, strand) in the focal species was matched to its relative position (chromosome, position, strand) in the protein-coding DNA alignment by first converting the genomic positions to relative position in the coding sequence (CDS) using gene annotation files (GTF format) downloaded from Ensembl (ensembl.org). We then verified that the SNPs downloaded from Ensembl were matching the reference in the CDS (FASTA format). Second, the relative position in the CDS was converted to position in the multiple sequence alignment (containing gaps) from the OrthoMaM database (44–46) by doing a global pairwise alignment, using the Biopython function pairwise2, between CDS FASTA and the sequence found in the alignment. This conversion from genomic position to position in the alignment is possible only if the assembly used for SNP calling is the same as the one used in the alignment, the GTF annotations, and the FASTA sequences.

We retrieved the genetic variants representing the population level polymorphism from the following species and respective available datasets: *Equus caballus* EquCab2 assembly in the EVA study PRJEB9799 (53), *Canis familiars* CanFam3.1 assembly in the EVA study PRJEB24066 (54), *Bos taurus* (UMD3.1 assembly in the NextGen project), *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 (55), *Chlorocebus sabaeus* ChlSab1.1 assembly in the EVA project PRJEB22989 (56), and *Homo sapiens* GRCh38 assembly from the 1,000-genome project (57, 58).

Variants not inside genes are discarded at the beginning of the analysis. Insertions and deletions are not analyzed, and only single nucleotide polymorphisms (SNPs) with only one mutant allele are considered. Stop codon mutants are also discarded. For populations containing more than 8 sampled individuals, the site-frequency spectrum (SFS) is subsampled down to 16 chromosomes (8 diploid individuals) without replacement to alleviate the effect of different sampling depths in the 29 populations. Moreover, subsampling mitigates the impact of moderately deleterious mutations segregating at a low frequency on π_N/π_S since they are more likely to be discarded than polymorphism segregating at a higher frequency. The Snakemake pipeline for integrating polymorphism and divergence data uses custom scripts written in Python 3.9.

Rate of Adaption in Population-Based Method. The genes and sites classified as under adaptation are concatenated. For each population, π_N/π_S is computed as the sum of nonsynonymous over synonymous polymorphism on the concatenated SFS. d_N/d_S is computed on the concatenated pairwise alignment between focal and sister species extracted from OrthoMaM; the d_N/d_S count is performed by *yn00*. We considered *Ceratotherium simum simum* as *Equus caballus* sister species; *Ursus maritimus* as *Canis familiars* 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. Altogether, $\omega_A = d_N/d_S - \pi_N/\pi_S$ is thus computed for each population on genes and sites classified as under adaptation. The result is compared to the empirical null distribution of ω_A, obtained by randomly sampling (1,000 sampling replicates) a subset of genes/sites classified as nearly neutral.

Other methods to compute ω_A such as polyDFE (20) are also used (*SI Appendix*, Eqs. 3–20), which relies on the synonymous and nonsynonymous unfolded site-frequency spectra (SFS) to estimate the distribution of fitness effects of mutations (DFE) and the rate of adaptation. In polyDFE, the GammaExpo model assumes that the fitness effect of weakly deleterious nonsynonymous mutations are distributed according to a negative Gamma, and the fitness effects of weakly advantageous mutations are distributed exponentially. This method is an extension of the methods introduced by Eyre-Walker and collaborators (9, 59). Unfolded SFSs are obtained by polarizing SNPs using the 3 closest outgroups found in the OrthoMam alignment with est-usfs v2.04 (60).

Data, Materials, and Software Availability. The data underlying this article are available at https://doi.org/10.5281/zenodo.7107233. Scripts and instructions necessary to reproduce the empirical experiments on the original dataset or with user-specified datasets are available at https://github.com/ThibaultLatrille/AdaptaPop. Archive data have been deposited in Genes and sites under adaptation at the phylogenetic scale also exhibit adaptation at the population-genetic scale, https://zenodo.org/record/7543458. Previously published data were used for this work. This study makes use of data generated by the NextGen Consortium.

ACKNOWLEDGMENTS. We gratefully also acknowledge the help of Nicolas Galtier and Julien Joseph for their advice and review 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. The European Union's Seventh Framework Programme (FP7/2010-2014) provided funding for the project under grant agreement no. 244356, "NextGen." Funding: Université de Lausanne; Agence Nationale de la Recherche, Grant ANR-15-CE12-0010-01/DASIRE; Agence Nationale de la Recherche, Grant ANR-19-CE12-0019/HotRec; Grant ANR-20-CE02-0008/NEGA; Natural Sciences and Engineering Research Council of Canada.

Author affiliations: [a]Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, Laboratoire de Biométrie et Biologie Evolutive, UMR5558, 69100 Villeurbanne, France; [b]École Normale Supérieure de Lyon, Université de Lyon, 69342 Lyon, France; [c]Department of Computational Biology, Université de Lausanne, 1015 Lausanne, Switzerland; and [d]Department of Biology, Institute of Biochemistry, and School of Mathematics and Statistics, Carleton University, K1S 5B6 Ottawa, Canada

1. L. Duret, D. Mouchiroud, Expression pattern and surprisingly, gene length shape codon usage in *Caenorhabditis, Drosophila*, and *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4482–4487 (1999).
2. L. Duret, Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* **12**, 640–649 (2002).
3. N. Galtier *et al.*, Codon usage bias in animals: Disentangling the effects of natural selection, effective population size, and GC-biased gene conversion. *Mol. Biol. Evol.* **35**, 1092–1103 (2018).
4. S. V. Muse, B. S. Gaut, A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Mol. Biol. Evol.* **1**, 715–724 (1994).
5. N. Goldman, Z. Yang, A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736 (1994).
6. J. H. McDonald, M. Kreitman, Adaptative protein evolution at Adh locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
7. N. G. Smith, A. Eyre-Walker, Adaptive protein evolution in *Drosophila*. *Nature* **415**, 1022–1024 (2002).
8. A. Eyre-Walker, P. D. Keightley, N. G. C. Smith, D. Gaffney, Quantifying the slightly deleterious mutation model of molecular evolution. *Mol. Biol. Evol.* **19**, 2142–2149 (2002).
9. A. Eyre-Walker, P. D. Keightley, Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol. Biol. Evol.* **26**, 2097–2108 (2009).
10. N. Galtier, Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genet.* **12**, e1005774 (2016).
11. Z. Yang, R. Nielsen, Mutation-selection models of codon substitution and their use to estimate selective strengths on codon usage. *Mol. Biol. Evol.* **25**, 568–579 (2008).
12. Z. Yang, R. Nielsen, N. Goldman, AmK Pedersen, Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**, 431–449 (2000).
13. C. Kosiol *et al.*, Patterns of positive selection in six mammalian genomes. *PLOS Genet.* **4**, e1000144 (2008).
14. A. L. Halpern, W. J. Bruno, Evolutionary distances for protein-coding sequences: Modeling site-specific residue frequencies. *Mol. Biol. Evol.* **15**, 910–917 (1998).
15. N. Rodrigue, H. Philippe, Mechanistic revisions of phenomenological modeling strategies in molecular evolution. *Trends Genet.* **26**, 248–252 (2010).
16. S. J. Spielman, C. O. Wilke, The relationship between dN/dS and scaled selection coefficients. *Mol. Biol. Evol.* **32**, 1097–1108 (2015).
17. N. Rodrigue, N. Lartillot, Detecting adaptation in protein-coding genes using a Bayesian site-heterogeneous mutation-selection codon substitution model. *Mol. Biol. Evol.* **34**, 204–214 (2017).
18. J. D. Bloom, Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. *Biol. Direct* **12**, 1–24 (2017).
19. M. Rousselle, M. Mollion, B. Nabholz, T. Bataillon, N. Galtier, Overestimation of the adaptive substitution rate in fluctuating populations. *Biol. Lett.* **14**, 20180055 (2018).
20. P. Tataru, T. Bataillon, "polyDFE: Inferring the distribution of fitness effects and properties of beneficial mutations from polymorphism data" in *Methods in Molecular Biology* (Humana Press Inc., 2020), vol. 2090, pp. 125–146.

21. D. Enard, L. Cai, C. Gwennap, D. A. Petrov, Viruses are a dominant driver of protein adaptation in mammals. *eLife* **5**, e12469 (2016).
22. E. R. Ebel, N. Telis, S. Venkataram, D. A. Petrov, D. Enard, High rate of adaptation of mammalian proteins that interact with Plasmodium and related parasites. *PLOS Genet.* **13**, e1007023 (2017).
23. W. S. W. Wong, Z. Yang, N. Goldman, R. Nielsen, Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* **168**, 1041–1051 (2004).
24. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
25. P. W. Messer, D. A. Petrov, Frequent adaptation and the McDonald–Kreitman test. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8615–8620 (2013).
26. Q. Chen *et al*., Two decades of suspect evidence for adaptive molecular evolution – Negative selection confounding positive selection signals. *Natl. Sci. Rev.* **9**, nwab217 (2021).
27. J. B. Plotkin, G. Kudla, Synonymous but not the same: The causes and consequences of codon bias. *Nat. Rev. Genet.* **12**, 32–42 (2011).
28. R. Lanfear, H. Kokko, A. Eyre-Walker, Population size and the rate of evolution. *Trends Ecol. Evol.* **29**, 33–41 (2014).
29. A. Platt, C. C. Weber, D. A. Liberles, Protein evolution depends on multiple distinct population size parameters. *BMC Evol. Biol.* **18**, 1–9 (2018).
30. T. Latrille, V. Lanore, N. Lartillot, Inferring long-term effective population size with mutation–selection models. *Mol. Biol. Evol.* **38**, 4573–4587 (2021).
31. T. Latrille, N. Lartillot, Quantifying the impact of changes in effective population size and expression level on the rate of coding sequence evolution. *Theor. Popul. Biol.* **142**, 57–66 (2021).
32. R. A. Goldstein, Evolutionary perspectives on protein thermo dynamics. *Int. Conf. Comput. Sci.* **3039**, 718–727 (2004).
33. R. A. Goldstein, S. T. Pollard, S. D. Shah, D. D. Pollock, Nonadaptive amino acid convergence rates decrease over time. *Mol. Biol. Evol.* **32**, 1373–1381 (2015).
34. R. A. Goldstein, D. D. Pollock, Sequence entropy of folding and the absolute rate of amino acid substitutions. *Nat. Ecol. Evol.* **1**, 1923–1930 (2017).
35. R. Patel, V. Carnevale, S. Kumar, Epistasis creates invariant sites and modulates the rate of molecular evolution. *Mol. Biol. Evol.* **39**, msac106 (2022).
36. N. Youssef, E. Susko, A. J. Roger, J. P. Bielawski, Evolution of amino acid propensities under stability-mediated epistasis. *Mol. Biol. Evol.* **39**, msac030 (2022).
37. D. Schrempf, N. Lartillot, G. Szöllősi, Scalable empirical mixture models that account for across-site compositional heterogeneity. *Mol. Biol. Evol.* **37**, 3616–3631 (2020).
38. S. Y. W. Ho, M. J. Phillips, A. Cooper, A. J. Drummond, Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* **22**, 1561–1568 (2005).
39. A. Eyre-Walker, Changing effective population size and the McDonald-Kreitman test. *Genetics* **162**, 2017–2024 (2002).
40. A. F. Moutinho, T. Bataillon, J. Y. Dutheil, Variation of the adaptive substitution rate between species and within genomes. *Evol. Ecol.* **34**, 315–338 (2019).
41. P. Tataru, T. Bataillon, polyDFEv2.0: Testing for invariance of the distribution of fitness effects within and across species. *Bioinformatics* **35**, 2868–2869 (2019).
42. J. Chen, T. Bataillon, S. Glémin, M. Lascoux, Hunting for beneficial mutations: Conditioning on SIFT scores when estimating the distribution of fitness effect of new mutations. *Gen. Biol. Evol.* **14**, evab151 (2022).
43. J. L. Thorne, N. Lartillot, N. Rodrigue, S. C. Choi, "Codon models as a vehicle for reconciling population genetics with inter-specific sequence data" in *Codon Evolution: Mechanisms and Models* (Oxford University Press, 2012), pp. 97–110.
44. V. Ranwez *et al*., OrthoMaM: A database of orthologous genomic markers for placental mammal phylogenetics. *BMC Evol. Biol.* **7**, 1–12 (2007).
45. E. J. Douzery *et al*., OrthoMaM v8: A database of orthologous exons and coding sequences for comparative genomics in mammals. *Mol. Biol. Evol.* **31**, 1923–1928 (2014).
46. C. Scornavacca *et al*., OrthoMaM v10: Scaling-up orthologous coding sequence and exon alignments with more than one hundred mammalian genomes. *Mol. Biol. Evol.* **36**, 861–862 (2019).
47. J. P. Huelsenbeck, S. Jain, S. W. D. Frost, S. L. Kosakovsky Pond, A Dirichlet process model for detecting positive selection in protein-coding DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6263–6268 (2006).
48. N. Lartillot, N. Rodrigue, D. Stubbs, J. Richer, PhyloBayes MPI. Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615 (2013).
49. R. A. Fisher, The distribution of gene ratios for rare mutations. *Proc. R. Soc. Edinburgh* **50**, 205–220 (1930).
50. S. Wright, Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931).
51. D. M. McCandlish, A. Stoltzfus, Modeling evolution using the probability of fixation: History and implications. *Q. Rev. Biol.* **89**, 225–252 (2014).
52. N. Rodrigue, T. Latrille, N. Lartillot, A Bayesian mutation-selection framework for detecting site-specific adaptive evolution in protein-coding genes. *Mol. Biol. Evol.* **38**, 1199–1208 (2021).
53. H. M. Holl, S. E. Kalla, N. B. Sutter, S. A. Brooks, Whole genome detection of sequence and structural polymorphism in six diverse horses. *PLoS One* **15**, e0230899 (2020).
54. V. Jagannathan, C. Drögemüller, T. Leeb, DBVD Consortium (DBVDC), A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. *Animal Genet.* **50**, 695–704 (2019).
55. S. T. Sherry *et al*., dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
56. H. Svardal *et al*., Ancient hybridization and strong adaptation to viruses across African vervet monkey populations. *Nat. Genet.* **49**, 1705–1713 (2017).
57. TGP Consortium, An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
58. The 1000 Genomes Project Consortium, A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
59. A. Eyre-Walker, M. Woolfit, T. Phelps, The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* **173**, 891–900 (2006).
60. P. D. Keightley, B. C. Jackson, Inferring the probability of the derived vs the ancestral allelic state at a polymorphic site. *Genetics* **209**, 897–906 (2018).