Stronger and higher proportion of beneficial amino acid changing mutations in humans compared to mice and flies

Authors:
Ying Zhen[a*], Christian D. Huber[a], Robert W. Davies[b], Kirk E. Lohmueller[a,c*]

Affiliations:

[a] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095

[b] Genetics and Genomic Biology and The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada.

[c] Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

* To whom correspondence should be addressed: klohmueller@ucla.edu, zhen@g.ucla.edu

ABSTRACT

Quantifying and comparing the amount of adaptive evolution among different species is key to understanding evolutionary processes. Previous studies have shown differences in adaptive evolution across species, however their specific causes remain elusive. Here, we use improved modeling of weakly deleterious mutations and the demographic history of the outgroup species and estimate that 30-34% of nonsynonymous substitutions between humans and outgroup species have been fixed by positive selection. This estimate is much higher than previous estimates, which did not account for the population size of the outgroup species. Next, we directly estimate the proportion and selection coefficients of newly arising strongly beneficial nonsynonymous mutations in humans, mice, and *D. melanogaster* by examining patterns of polymorphism and divergence. We develop a novel composite likelihood framework to test whether these parameters differ across species. Overall, we reject a model with the same proportion and the same selection coefficients of beneficial mutations across species, and estimate that humans have a higher proportion of beneficial mutations compared to Drosophila and mice. We demonstrate that this result cannot be attributed to biased gene conversion. In summary, we find the proportion of beneficial mutations is higher in humans than in *D. melanogaster* or mice, suggesting that organismal complexity, which increases the number of steps required in adaptive walks, may be a key predictor of the amount of adaptive evolution within a species.

INTRODUCTION

Since the inception of molecular population genetics, there has been tremendous interest in quantifying the amount of adaptive evolution in different organisms. The neutral theory of molecular evolution postulated that beneficial mutations are rare, and many of the substitutions between species are neutral [1]. One early challenge to this theory originated from a comparison of polymorphism and substitutions (also known as divergence) at synonymous sites and nonsynonymous sites in Drosophila [2,3]. Under models without positive selection, the ratio of nonsynonymous to synonymous changes should remain equal when comparing polymorphisms and substitutions. In contrast to this prediction, a genome-wide excess of nonsynonymous substitutions between species was observed, a pattern indicative of an abundance of positive selection in Drosophila. More formally, Smith and Eyre-Walker (2002) proposed a statistic, alpha, which is the proportion of nonsynonymous substitutions between species that can be attributed to positive selection. Application of their approach in Drosophila has found that at least 40% of nonsynonymous substitutions have been fixed by positive selection [3].

Since the publication of the original study, alpha has been estimated from different species across the tree of life [4]. Estimates of alpha vary tremendously across species, tending to be higher in insects [5,6], but much lower in primates [6,7] and plants [8]. In these latter species, formal tests have been unable to reject that alpha is zero (*i.e.* no positive selection) [6,7,9]. It is not clear why alpha varies across species. One possibility is that alpha is higher for species with larger population sizes, which could occur if adaptation is mutation limited, and therefore species with larger population sizes would have more beneficial mutations. The fixation probability of a given beneficial mutation also would be higher in species with larger population size, but this effect is likely to only be important for very weakly beneficial mutations. Evidence indicates that, in some cases, alpha is indeed related to population size. For example, Phifer-Rixey *et al.* found that estimates of alpha were higher for species of mice that have larger population sizes compared to species with smaller population sizes [10]. Further, recent studies have found a positive correlation between alpha and population size when comparing different species of sunflowers [11] and from phylogenetically diverse taxa [12]. More recently, Galtier found a positive correlation between alpha and effective population size for 44 animal species [13]. Additional evidence that there is more positive selection in larger populations stems from recent analyses of linked selection. Corbett-Detig *et al.* found increased evidence for linked selection in species with larger population sizes, though the mechanism driving this pattern is not immediately clear [14]. Further, Nam *et al.* have suggested that across primates, species with larger population sizes have had more selective sweeps [15].

While evidence suggests that adaptation could be mutation limited and this could be driving the variation in alpha across species, it is important to note that other factors could be influencing the alpha statistic[16]. By definition, alpha is the proportion of nonsynonymous substitutions attributable to positive selection. As such, it is heavily influenced by the total number of substitutions and thus the number of substitutions attributable to the fixation of weakly deleterious mutations. For two populations with the same number of beneficial substitutions, the one with a higher number of substitutions due to weakly deleterious mutations will have a lower alpha. Indeed, because the number of weakly deleterious mutations fixed is inversely related to population size, this effect could drive the correlation between alpha and population size. In support of this prediction, Galtier found that the rate of adaptive divergence showed no correlation with population size [13]. Similar arguments have been made by Phifer-Rixey *et al.* [10]

A recent conceptual and theoretical investigation of beneficial mutations under Fisher's geometric model found that population size was a poor predictor of alpha and the rate of adaptive divergence [17]. Instead, organismal complexity (here defined as the number of phenotypes under selection) and the rate of environmental changes were better predictors of alpha. They also point out that the distribution of fitness effects (DFE) for newly arising beneficial mutations likely differs with population size. Because small populations likely have had more fixations of weakly deleterious alleles, they are further from the fitness optimum. Thus, small populations then have the potential for more new beneficial mutations compared to larger populations.

In summary, the amount of adaptive evolution in disparate species with varying population sizes remains elusive. Equally unclear is the best metric to quantify the amount of positive selection in distinct species. Boyko *et al.* found that by assuming some fraction (0 to 1.86%) of new mutations is positively selected, they could better match the frequency spectrum of polymorphisms and the counts of human-chimpanzee differences. Importantly, models with weaker selection coefficients for beneficial mutations tended to have a higher proportion of positively selected mutations than models with stronger selection [7]. Further, by fitting a DFE from Fisher's Geometric model to polymorphism data in humans and *Drosophila melanogaster*, Huber *et al.* found a higher proportion (14%) of new weakly beneficial mutations in humans compared to Drosophila (<1%) [18]. They attributed this higher estimate of the proportion of new beneficial mutations in humans as compensating for deleterious alleles that became fixed due to the small population size that in turn moved the human population away from the fitness optimum.

However, direct comparison the DFEs of beneficial mutations across species has not been performed rigorously before, and in previous work on estimating alpha, there were unjustified assumptions about the demography of the outgroup species that can have substantial impacts on the inference. Here we directly estimate the proportion and selection coefficients of newly arising strongly beneficial mutations in humans, mice, and *D. melanogaster* by examining patterns of polymorphism and divergence. We then develop a composite likelihood framework to test whether these parameters differ across species. This approach enables a more direct comparison of the amount of beneficial mutations across different species, and is also less confounded by the fixation of weakly deleterious mutations. Overall, we reject a model with the same proportion and the same selection coefficients of beneficial mutations across species, and estimate that humans have a higher proportion of beneficial mutations compared to Drosophila and mice. Using improved modeling of weakly deleterious mutations and demographic models, particularly correcting for the population size of outgroup species, we estimate that 30-34% of nonsynonymous substitutions between humans and the outgroup is driven by positive selection, much higher than previously thought. In addition, we explore the effect of biased gene conversion (BGC) on our estimates of adaptive evolution by looking at subsets of sites that are unaffected by BGC and find that while BGC influence our estimates of positive selection in a species-specific manner, it cannot account for our main findings that the proportion and strength of beneficial mutations differ across species.

RESULTS

Estimates of alpha for multiple species

We first estimated alpha from coding regions of humans, mice and *D. melanogaster*. We analyzed published genomic datasets to obtain counts of synonymous and nonsynonymous polymorphisms (P_S and P_N, respectively) as well as synonymous and nonsynonymous

substitutions between species (D_S and D_N, respectively). In total, 19.1Mb of coding sequence for human, 26.6 Mb of coding sequence for mice and 15.8Mb of coding sequence for *D. melanogaster* were used in our analysis (see Methods). For computation of alpha in humans, we used chimpanzee and macaque as outgroup species. For mice, we used rat as outgroup, and for *D. melanogaster*, we used *D. simulans* as the outgroup species (see Methods).

Alpha was first estimated using an extension of the McDonald-Kreitman (MK) test (Smith and Eyre-Walker 2002b; equation (2); table 1; supplementary table S1). To examine the effect of slightly deleterious mutations on alpha, we also filtered the data with several minor allele frequency (MAF) cutoffs (supplementary table S2). After removing low frequency polymorphisms with MAF less than 20%, the estimated alpha is close to zero for humans when using chimpanzee as outgroup species (table 1), consistent with previous estimates of alpha in humans [6,7]. However, when using macaque as the outgroup species, the estimate of alpha is -0.22 (table 1), suggesting that the choice of outgroup species could greatly influence alpha. Nevertheless, these results suggest at most only a very small proportion of nonsynonymous substitutions have been fixed by positive selection in the human lineage. In contrast, for *D. melanogaster* and mice, the estimated alpha is 49% and 40%, respectively, with MAF filter at 20% (table 1). Both of these estimates of alpha are comparable to those seen in previous studies [5,6,10]. These results suggest that the proportion of substitutions fixed by positive selection varies drastically across species, and for species with larger population sizes, like mice and *D. melanogaster*, adaptive forces may have had a greater contribution to divergence.

Model based inference of alpha

Due to these concerns regarding estimating alpha directly from MK table counts, model-based approaches to quantify alpha have been developed. Boyko *et al.* and Eyre-Walker *et al.* estimated alpha as the proportion of the observed nonsynonymous substitutions (D_{NO}) that cannot be explained by models with only neutral and deleterious mutations [6,7]. Specifically, they assume a DFE for deleterious and neutral mutations as well as a demographic model, and predicted the expected number of nonsynonymous substitutions (D_{NE}). The excess of D_{NO} compared to the D_{NE} is attributed to fixations driven by positive selection. Here, we have extended the approach of Boyko *et al.* using *prfreq* [7] and applied it to data from humans, mice and *D. melanogaster* (see Methods).

Because a wide range of divergence times were reported in previous studies for each species, we estimate divergence time that best fit our data using the observed D_S. We then use *prfreq* to predict D_N using this estimated divergence time, the gamma-distributed DFE, and the demographic model (see Methods). Applying this framework using chimpanzee as the outgroup species, we initially estimated that 15% of the nonsynonymous substitutions between humans and chimp were driven by positive selection (table 1). Our estimate is comparable to the inference by Boyko *et al.*, where authors implemented similar approach and estimated that approximately 10% of human-chimp nonsynonymous substitutions were fixed by positive selection [7].

Interestingly, when using macaque as the outgroup species, we estimate the proportion of nonsynonymous substitutions fixed by positive selection in humans is close to zero, using the aforementioned framework (table 1). While it is possible that this difference could reflect distinct evolutionary events experienced by different outgroups or different periods of history, we consider that it could be an artifact of the modeling assumptions. Specifically, one assumption of *prfreq* is that the effective population size of the outgroup species is the same as the ancestral

population size, which is not the case for the species considered here, especially for humans, chimpanzees, and macaques. The inferred human ancestral size is 7067, which is much smaller than previous estimates of the human-macaque or human-chimp ancestral population sizes in the range of tens of thousands [20–25]. Using a population size for the outgroup species that is too small likely biases estimate of alpha because more of the nonsynonymous substitutions could be attributed to the fixation of weakly deleterious mutations, causing alpha to be under-estimated.

To more accurately model the larger outgroup population size in chimpanzees and macaques, we add an additional ancient epoch where the population size is 73,000 individuals before the time of the human-chimpanzee divergence in our initial two-epoch human demographic model (supplementary table S4), which is within the range of estimated population sizes of human-chimpanzee and human-macaque common ancestors (see Methods). This is our three-epoch model. Because this added epoch is ancient, it does not affect the polymorphism pattern within humans (see Methods). Importantly, the larger ancient population size better reflects the effective population size of the outgroup species, thus yielding a more accurate estimate of the number of substitutions between species. Using the modified three-epoch model for humans, we obtain comparable estimates of alpha regardless of our outgroup choice. Specifically, we estimate approximately 33% of human-chimpanzee nonsynonymous substitutions were fixed by positive selection and approximately 30% of human-macaque nonsynonymous substitutions were fixed by positive selection (supplementary table S1). Importantly, our estimates of alpha in humans using the more realistic three-epoch demographic model are much higher than the previously reported estimates [7], implying that there is a greater contribution of positive selection to nonsynonymous substitutions than previously appreciated.

Similarly, we improved the two-epoch models of *D. melanogaster* and mice to three-epoch models by including an additional ancient epoch at the time of divergence with their outgroup species. Specifically, for *D. melanogaster* it had been inferred that *D. simulans* have slightly larger N_e [26], so we added an ancient epoch of 1.5× the current population size at the *D. melanogaster* and *D. simulans* split (supplementary table S4). For mice, a previous study estimated that the outgroup rat species has an effective population size about fivefold lower than wild house mice [27]. Thus we added an ancient epoch that was 0.2× the current population size of mice (supplementary table S4). Using the three-epoch models for *D. melanogaster* and mice, we estimated their alpha to be 68% and 38%, respectively (table 1, supplementary table S1) compared with 58% for *D. melanogaster* and 48% for mice (table 1, supplementary table S1) using the original two-epoch model. The differences between these estimates reflect the importance of accurately modeling the population size of outgroup species for calculations of alpha.

When we apply *DFE-alpha* to our three species, alpha is estimated to be 24% for humans using chimpanzee as outgroup, 2% for humans using macaque as outgroup, 71% for *D. melanogaster*, and 51% for mice (table 1). These estimates are all slightly higher compared to estimates from our two-epoch models. However, the estimates of alpha computed for humans differ significantly depending on whether the macaque or chimpanzee is used as the outgroup. We additionally estimate alpha for substitutions that occurred on the human lineage, using the human-macaque alignment to polarize substitutions between human and chimp (see Methods). We estimate that alpha equals 18.3% using the 2-epoch model and 19.7% using the 3-epoch model.

Test whether p^+ and s^+ differ across species

Thus far we have examined the proportion of nonsynonymous substitutions that have been fixed by positive selection. This is the outcome of the evolutionary process and is a function of the input of beneficial mutations as well as how they are affected by demography and natural selection. Here we take a different approach to quantify the properties of new beneficial mutations. Specifically, we estimate the DFE including new beneficial mutations. Our model of the DFE includes two additional parameters compared to the base model that only includes deleterious mutations. For each species, we estimate the proportion of new mutations that are beneficial (p^+) and their selection coefficient (s^+). We then test whether these two parameters differ across species.

The number of nonsynonymous substitutions (D_N) is Poisson distributed [28], with rate parameter equal to:

$$E[D_N] = 2N\mu[\int G(s)u(s)\,(1-p^+)ds + u(s^+)p^+] \quad (1)$$

where $G(s)$ is the DFE of deleterious and neutral mutations, $u(s)$ is the fixation probability of deleterious and neutral mutations, $u(s^+)$ is the fixation probability of beneficial mutations, and p^+ is the proportion of beneficial mutations. We then use a Poisson log-likelihood function for D_N in each species and a series of likelihood ratio tests to determine whether p^+ and s^+ differ across species (see Methods).

Using this framework, we find that the full model H1, where each species is allowed to have its own p^+ and s^+, fits D_N significantly better than the constrained null model, where p^+ and s^+ are constrained to be the same across all three species (Likelihood Ratio Test (LRT)) statistic Λ=122,724, df=4, $P<10^{-16}$; fig. 1; supplementary table S5). Taking the MLEs of p^+ and s^+ for the full model, we predicted D_N between species, which matched the observed D_N (supplementary fig. S1). When we used the MLEs for the constrained model, the predicted D_N did not match the observed D_N (supplementary fig. S1).

We estimate that humans have a higher proportion of strongly beneficial nonsynonymous mutations than *Drosophila* and mouse (supplementary table S5 and fig. 1). Specifically, we estimate that approximately 2.15% of new nonsynonymous mutations in humans are beneficial with a selection coefficient of approximately 2.45E-05 (outgroup: chimpanzee), approximately 0.0075% of new nonsynonymous mutations in *D. melanogaster* are beneficial with a selection coefficient of approximately 4.99E-05, and approximately 1.97% of new nonsynonymous mutations in mice are beneficial with a selection coefficient of 1.21E-05 (supplementary table S5). It is important to point out that models with a larger selection coefficient tend to have a lower proportion of positively selected mutations than models with weaker selection. Consequently, the likelihoods of these parameter values can be very close to the likelihoods at the MLEs.

To examine if any two species out of three share the same values of p^+ and s^+, we performed LRTs comparing each pair of species. In all pairwise tests, the model where each species has its own p^+ and s^+ fits the observed D_N significantly better than a model where p^+ and s^+ are constrained to be the same in the tested two species (supplementary table S5). These results suggest each species has their own unique values of p^+ and s^+. This result is robust regardless of outgroup species and demography (supplementary table S5).

We next investigated whether it is possible that either p^+ or s^+ is the same across species, but the other parameter varies. Specifically, we allowed p^+ to differ across species, then explored whether a model with the same s^+ could fit all species. This is shown in conditional likelihood plots, where assuming the same s^+ for all species, humans would need a higher proportion of

beneficial mutations compared to mice and *D. melanogaster* to match the observed D_N (fig. 2A). Similarly, allowing s^+ to differ across species, a model with the same p^+ across all species could fit the data. When we forced the same p^+ for all species, s^+ for beneficial mutations in humans would be larger compared to that in mice and *D. melanogaster* (fig. 2B). However, for the same p^+ value, s^+ could not be the same in all species. Similarly, for the same s^+ value, p^+ could not be the same across species.

Test whether gamma+ and p+ differ across species

Humans, *D. melanogaster*, and mice have drastically different population sizes. These different population sizes can influence the efficacy of selection within each species. Thus, we next examined whether the selection coefficient scaled by current population size ($gamma^+=2Ns^+$) and p^+ differ across species.

We find the model (Full model H1) where each species has its own different $gamma^+$ and p^+ fits the observed D_N significantly better than a model (constrained model H0) where $gamma^+$ and p^+ are constrained to be the same across all three species (LRT statistic Λ=30,061; df=4, $P<10^{-16}$, fig. 3 and supplementary table S5). Taking the MLEs of $gamma^+$ and p^+ for the full model, we predict D_N between species, which matches the observed D_N. When using the MLEs for the constrained model, the predicted D_N does not match the observed data (supplementary fig. S1).

Under the full model, we estimate that approximately 1% of new nonsynonymous mutations in humans are beneficial with $gamma^+$ of 1.38, approximately 2% of new nonsynonymous mutations in *D. melanogaster* are beneficial with $gamma^+$ of 1.95, and approximately 4% of new nonsynonymous mutations in mice are beneficial with $gamma^+$ of 4.8 (supplementary table S5). Similarly, models with stronger $gamma^+$ tended to have a lower proportion of positively selected mutations than models with weaker selection, and the likelihoods can be very close to those at the MLEs. The current population sizes are 16539, 7616700, 488948 for humans, *D. melanogaster* and mice, respectively. As a result, we are searching for p^+ within the same range of $gamma^+$ (see Methods), but markedly different ranges of s^+ for these three species. Thus, the relative ordering of the MLEs of p^+ across species considering $gamma^+$ is not necessarily the same as that for the s^+ results described above.

To examine if any two species out of three may share the same values of $gamma^+$ and p^+, we performed LRTs comparing each pair of species. We find that the model where each species has its own different $gamma^+$ and p^+ fit the observed D_N significantly better than a model where $gamma^+$ and p^+ are constrained to be the same in the tested two species (supplementary table S5). Thus each species likely has its unique $gamma^+$ and p^+. This result is generally robust regardless of outgroup species. However, when using the human two-epoch model, we cannot reject the hypothesis that human and *Drosophila* have the same $gamma^+$ and p^+ (supplementary table S5). One must note that the difference in chimpanzee population size with human ancestor population size is well established in literature [20–25], so the two-epoch model makes unrealistic assumptions and thus, this comparison is not meaningful.

We next investigated whether it is possible that either $gamma^+$ or p^+ is the same across species, while the other parameter varies. The results are shown in conditional likelihood plots, where with the same $gamma^+$ for all species, humans would need a lower proportion of beneficial mutations compared to mice and *Drosophila* to fit the observed D_N (supplementary fig. S2A). Similarly, allowing $gamma^+$ to differ across species, a model with the same p^+ could fit all species. When we force the same p^+ for all species, humans have smaller $gamma^+$ for

beneficial mutations (supplementary fig. S2B). However, for the same p^+ value, $gamma^+$ cannot be the same in all species. Similarly, for the same $gamma^+$ value, p^+ cannot be the same across species.

Effects of biased gene conversion

Biased gene conversion (BGC) is the preferred transmission of G/C alleles (S: strong alleles) at the expense of A/T alleles (W: weak alleles). This process is common in mammals, including two of our three targeted species: humans and mice. BGC could impact patterns of genetic diversity significantly [29–31] and bias estimates of rate of positive selection [32], especially when comparing between species with BGC and species without BGC, i.e. *D. melanogaster* [33].

To test whether BGC drives the observed pattern of positive selection across species, we filtered the human and mouse data to keep only strong to strong or weak to weak mutations (herein called SSWW mutations), which are not affected by BGC. For humans, the filtered SSWW polymorphisms have a similar SFS as the full dataset (supplementary fig. S3A). Thus we use the demographic and DFE parameters estimated from the full data. We used the observed number of synonymous SSWW polymorphisms to estimate the mutation rate of human SSWW mutations to be 3.14E-09, which is comparable to the previous estimates [30,34]. Following the method used on the full data, we re-estimate the human-chimpanzee divergence time that fits best to the observed SSWW D_S. We then use *prfreq* to predict the SSWW D_N using this new estimated divergence time, DFE, and demographic models (supplementary table S1). We estimate that under the two-epoch demographic model and the improved three-epoch model, approximately 17.1% or approximately 34.3% of the observed SSWW D_N in humans using chimpanzee as outgroup was driven by positive selection, respectively (table 1). These estimates of alpha from SSWW sites are slightly elevated but are comparable to the estimates from the full dataset (table 1).

For mice, however, the SFS of SSWW polymorphism has a very different shape compared to the SFS from the full dataset (supplementary fig. S3B). Thus, we re-estimated the demographic and DFE parameters for mice SSWW mutations (supplementary table S3; see Methods).

We then estimate that under the two-epoch demographic model and the improved three-epoch model, approximately 33.4% and 19.2% of the observed SSWW D_N in mice was driven by positive selection, respectively (table 1, supplementary table S1). These estimates of proportion of positive selection from SSWW mutations are much lower than those estimated from the full dataset (table 1). This suggests that biased gene conversion may account for some of the nonsynonymous substitutions between mouse and rat.

Removing the effect of BGC in humans and mice by using only the SSWW changes in these two species and the full dataset of *D. melanogaster*, we again quantify and compare the strength and proportion of new beneficial mutations across all three species. Using the same composite likelihood framework as above, we find that the model where each species has its own p^+ and s^+ fits the observed D_N significantly better than a model where p^+ and s^+ are constrained to be the same across all three species (LRT statistic Λ=3,213, df=4, $P<10^{-16}$; fig. 4; supplementary table S5). Models comparing each pair of species suggest that each species has its own unique p^+ and s^+, regardless of outgroup and demography (supplementary table S4). Allowing p^+ to differ across species, a model with the same s^+ across all species could fit the data, and *vice versa.* When we force the same s^+ for all species, humans still have the highest proportion of new beneficial mutations, *D. melanogaster* has the smallest proportions of new

beneficial mutations, and mice have an intermediate proportion of beneficial mutations (fig. 4). When we force the same p^+ for all species, humans have the largest selection coefficient, *D. melanogaster* has the weakest selection coefficient, and mice have an intermediate selection coefficient (fig. 4).

Lastly, we find that the model where each species has its own p^+ and $gamma^+$ fits the observed D_N significantly better than a model where p^+ and $gamma^+$ are constrained to be the same across all three species (LRT Λ=1,114, df=4, $P<10^{-16}$; supplementary fig. S4; supplementary table S5). For pairwise comparisons, however, we cannot reject the null hypothesis that *D. melanogaster* and mice have the same p^+ and $gamma^+$ (supplementary table S5). Only under uncorrected demographic scenarios (i.e. human two-epoch), sometimes we cannot reject humans and *D. melanogaster*, or humans and mice having the same p^+ and $gamma^+$ supplementary table S5). These results are also reflected in the conditional likelihood plots. Allowing p^+ to differ across species, a model with the same $gamma^+$ across all species could fit the data, and *vice versa*. When we force the same $gamma^+$ for all species, humans have the smallest proportion of new mutations being beneficial, while *D. melanogaster* and mice both have higher proportions of beneficial mutations than humans. But *D. melanogaster* and mice could have the same or different proportions, depending on the $gamma^+$ (supplementary fig. S4). When we force the same p^+ for all species, humans have the lowest $gamma^+$, and *D. melanogaster* and mice both have a higher $gamma^+$, but their relative strength with each other depends on p^+(supplementary fig. S4).

DISCUSSION

Here we used novel composite likelihood procedures to show that the amount of adaptive evolution and the DFE of newly arising beneficial mutations differ across species. We find that the species with smaller population size (i.e. humans) has stronger and/or more abundant new beneficial mutations than the other two species with much larger population sizes (i.e. mice and *D. melanogaster*). Our findings are consistent with predictions made in Lourenço et al [17]. Namely they used Fisher's geometric model to argue that a smaller population will have more fixations of weakly deleterious mutations pushing it further away from the fitness optimum, thus creating more opportunities of beneficial compensatory mutations. We also find that alpha varies in a complex way across species. Indeed, alpha is the result of an intricate interplay between the DFE, demography, and population size. Importantly, our model-based alpha estimate for humans is approximately 30%, which is much higher than all previous estimates [6,7], and this result is robust to the choice of different outgroups and BGC. In addition, although our estimates indicate a higher alpha in *D. melanogaster* and mice than in humans, the difference between mice and humans is small (38% vs. 33%). Interestingly, when taking BGC into account, mice have a smaller alpha (19%) than human (34%), which goes in the opposite direction seen when not controlling for BGC. After removing the potentially confounding effects of BGC, alpha is no longer correlated with population size.

One major improvement of our method over previous similar approaches from Boyko et al and *DFE-alpha* [6,7] is that we take into account the difference in outgroup population size. In Boyko *et al.*, the outgroup population size is assumed to be the same as that found in the ancestral in-group population. *DFE-alpha* allows for inference of demographic parameters of one- to three- epoch models, and the outgroup population size is assumed to be the same as the ancestral in-group population. When considering species like humans in relation to other

primates, like chimpanzee or macaque, this assumption almost certainly does not hold as primate species have population sizes at least several fold larger than the estimated ancestral human population size of approximately 8,000 individuals. The size of the outgroup population matters because it affects the fixation probability of weakly deleterious alleles[1]. As such, the amount of nonsynonymous substitutions attributed to weakly deleterious mutations is highly affected by the population size of the outgroup. Consequently, estimates of alpha are then affected as well. Our approach includes an additional population size change, such that the outgroup can have a more realistic population size. By using more realistic population sizes in the outgroup species, the alpha estimates we obtained for human are similar when using the chimpanzee or macaque as outgroup species. This is strong evidence that our method is more accurate as all the other methods give drastically different estimates of alpha using these two different outgroups. Interestingly, Eyre-Walker and Keightley also suggested that alpha in human could be as high as 0.31 if the effective population size of humans and macaques was much higher than 10,000 until very recently [6], agreeing with our current estimates. Future studies should carefully consider outgroup population size and should use statistical methods that allow for additional size changes.

Huber *et al.* found 15% of nonsynonymous mutations in humans are weakly beneficial, consistent with Fisher's geometric model. This proportion of weakly beneficial mutations is much higher than that in *D. melanogaster* that have a larger population size. Because strongly beneficial mutations are thought to become fixed rapidly, they are not observed in polymorphism data. Thus, the Huber et al. study does not include these strongly beneficial mutations. Here, taking advantage of the availability of large genome sequences of several relative species, we can estimate the proportion of strongly beneficial mutations that are fixed between species. Thus, the results presented here in terms of p^+ refer to strongly beneficial (s^+>1E-5) mutations. In our method, the weakly beneficial mutations are already accounted for using the DFE from Huber *et al.* as many weakly beneficial mutations are likely segregating as polymorphisms. Intriguingly, we find that humans have a higher proportion of strongly beneficial mutations than Drosophila. This finding is in the same direction as what was found for weakly beneficial mutations in Huber et al.

It is important to emphasize that we quantify adaptive evolution from two different perspectives. First, we estimate the DFE of newly arising beneficial mutations, i.e. p^+ and s^+. Our method and the method of Boyko *et al.* aim to understand the properties of new beneficial mutations, the beginning point where beneficial mutation appear and enter the population. Second, we estimate the proportion of adaptive nonsynonymous substitutions between species, alpha. This latter statistic is the end point where a number of factors such as demography, genetic drift, and natural selection all come into play. The results of how the DFE for the newly arising beneficial mutations varies across species could be in the opposite direction to what has been found considering fixed differences. This is expected as these two approaches measure distinct quantities and different aspects of adaptive evolution.

Our estimate of alpha for human lineage is 18.3% using 2-epoch model and 19.7% using 3-epoch model, which is comparable to the estimate of human-lineage alpha by Urrichio *et al.,* despite the use of different analytical approaches. Alpha is expected to be lower on the human lineage as compared to the chimp or macaque lineage due to the higher proportion of weakly deleterious amino acid substitutions on the human lineage to their smaller population size.

One previous explanation for varying estimates of alpha across species was that adaptation is mutation limited and there are more beneficial mutations in organisms with larger population sizes. This view was not supported by a simulation study by Lourenço *et al.* that

considered a changing DFE over time in the context of Fisher's geometric model [16]. Instead, they found that the population size only weakly related to alpha, and the rate at which the environment changed was an important predictor of the amount of adaptive evolution, as environmental shifts moved the population from the fitness optimum, creating the opportunity for new beneficial mutations. However, Connallon and Clark found that environmental heterogeneity reduces the fraction of beneficial mutations by inflating the standardized mutation size in Fisher's geometric model [35]. Lourenco *et al.* also found that organismal complexity, here defined as the number of phenotypes under selection, was a key predictor of the amount of adaptive evolution within species. Through a "cost of complexity", more complex organisms have a harder time adapting to new environmental conditions due to the additional constraints imposed by the increased number of traits under selection. As such, adaptive walks require more beneficial mutations.

Our results presented here are in broad agreement with this conceptual model. Specifically, we do not find that species with larger population sizes (i.e. *D. melanogaster*) have more beneficial mutations. Instead, we find that p^+ is higher in humans than in *D. melanogaster* or mice. Second, while it is hard to precisely define organismal complexity, previous work has found more protein-protein interactions in humans than in *D. melanogaster* [36,37], suggesting that humans may be more complex than flies. If this is the case, then our findings of a higher p^+ in humans than flies and mice supports the arguments from Lourenco *et al.* [16] that adaptive walks after an environmental shift are less efficient and require more steps (*i.e.* beneficial mutations) in more complex organisms, leading to higher p^+ in complex organisms. Lastly, while it is hard to say which species has experienced more environmental shifts, changing environments may also be contributing to the disparate estimates of p^+ across species.

METHODS

Polymorphism and divergence data sets for humans, mice, and D. melanogaster

For humans, we used polymorphism data from 112 individuals from Yoruba in Ibadan, Nigeria (YRI) from the 1000 genomes project [38]. Published genome alignments of human and chimpanzee (hg19/pantro4), and human and *Macaca mulatta* (hg19/rheMac3) were downloaded from UCSC (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/). For *D. melanogaster*, polymorphism data were from 197 African *D. melanogaster* lines from the *Drosophila* Population Genomics Project phase 3 data of samples from Zambia, Africa [39]. For divergence, *D. melanogaster* and *D. simulans* genic alignments (Dmel v5/Dsim v2) were extracted from the multi-species alignments from [40]. Only autosomal regions were used in our analysis. Human and *Drosophila* polymorphism data were filtered and down-sampled to 100 chromosomes as described in Huber *et al.* [17]

For mice, raw data (fastq) was downloaded for 10 *M. m. castaneus* individuals that were collected in the northwest Indian state of Himachal Pradesh [41,42]. Reads were mapped against mouse genome mm9 using bwa [43] and stampy [44], duplicate reads were marked using Picard, and further pre-processing was done following GATK Best Practice guidelines [45]. Variants were called using the GATK UnifiedGenotyper and filtered using the GATK VQSR using Affymetrix Mouse Diversity Genotyping Array sites [46]. We further filtered the dataset to only retain sites with a sample size of at least 16 chromosomes and down-sampled all sites with larger sample size to a sample size of 16 chromosomes using the hypergeometric probability distribution. Published genome alignments of mice and rat (mm9/rn5) were downloaded from UCSC (http://hgdownload.soe.ucsc.edu/goldenPath/mm9/vsRn5/axtNet/). For each species,

polymorphism data and divergence data were intersected, and only coding regions shared by both datasets were used in our analysis.

In total, 19.1Mb of coding sequences for human, 26.6 Mb of coding sequences for mice and 15.8Mb of coding sequences for *D. melanogaster* were included. The nonsynonymous and synonymous total sequence lengths (L_{NS}, L_S) were estimated using multipliers of $L_{NS} = 2.85 \times L_S$ in *Drosophila*, and $L_{NS} = 2.31 \times L_S$ in humans and mice from Huber et al [17]. In these filtered coding sequences, we annotated synonymous and nonsynonymous sites in both polymorphism and substitution data for each species. Human variants were annotated using the SeattleSeq Annotation pipeline (http://snp.gs.washington.edu/SeattleSeqAnnotation138/). Mice and Drosophila variants were annotated using SnpEeff v3.6 using the mice NCBIM37.66 annotation database and the *D. melanogaster* BDGP5.75 annotation database, respectively. Sites that are annotated as near-splice, or loss of function were removed. The ratio of nonsynonymous/synonymous differences between human and chimp sequences in our dataset is about 0.65, which is consistent with several previous reports from different datasets [47–49].

From the down-sampled polymorphism data, we calculated the synonymous and nonsynonymous SFS, and used the folded SFS for all further inferences to avoid misidentification of the ancestral state.

Calculation of alpha

For each species, alpha was calculated using an extension of the McDonald-Kreitman test formulated in Smith and Eyre-Walker [18]:

$$\alpha = 1 - \frac{D_S P_N}{D_N P_S} \qquad (2)$$

Here, D_S is the number of synonymous substitutions, D_N is the number of nonsynonymous substitutions, P_S is the number of synonymous polymorphisms, and P_N is the number of nonsynonymous polymorphisms.

Demographic and DFE inferences for mice

We used methods established in Huber et al to infer demography and DFE of neutral and deleterious mutations from the mouse polymorphism data [17]. In short, we first used the synonymous SFS to infer demographic parameters for a two epoch-model using ∂a∂i [28,50]. Then, we used a Poisson likelihood function to estimate the parameters of a gamma-distributed DFE of new neutral and deleterious nonsynonymous mutations using the nonsynonymous SFS, conditional on the estimated demographic parameters [7,51].

prfreq estimates of alpha

To implement the *prfreq* approach to estimate alpha, for each species, we need a demographic model and a DFE for neutral and deleterious mutations to predict the D_{NE} that is accounted for by neutral and deleterious forces. For humans and *D. melanogaster*, we use demographic and DFE parameters from Huber *et al.* (supplementary table S3) [17]. For mice, we conduct our own inference of these parameters by summarizing the polymorphism data by the folded site frequency spectrum (SFS; see Methods). For mice, using the synonymous SFS, we infer that the ancestral population size is approximately 206,500 which expanded 2.4-fold 293,000 generations ago (supplementary table S3). Conditional on this demographic model, we estimate the DFE for new nonsynonymous mutations in mice. We assume that the DFE follows a gamma distribution and estimate its shape parameter alpha to be 0.21 and scale parameter beta to be 0.083 (supplementary table S3). These estimates are within the same magnitude of previous

estimates from Huber *et al.*, which used a much smaller dataset (<0.1% of the total sites used in our study). For both the two-epoch models and the three-epoch models, we first found the demographic parameters (supplementary table S4) that fit the observed number of synonymous substitutions using *prfreq*. Here the number of synonymous substitutions equals 2 × divergence time × mutation rate. We estimated the divergence time (*tdiv*) for each model and species using this method because there is a wide range of divergence times from the literature for each species. Second, using this divergence time, demography, and DFE inferred from Huber *et al.*, or as described above for mice, we estimated the expected number nonsynonymous substitutions (D_{NE}) using *prfreq* according to Sawyer and Hartl eqn 13 (Sawyer and Hartl 1992; Boyko et al. 2008). Then, alpha is calculated as

$$\alpha = \frac{D_{NO} - D_{NE}}{D_{NO}} \quad , \qquad (3)$$

where D_{NO} is the observed number of nonsynonymous substitutions.

DFE-alpha

Data files and the program v2.15 were downloaded from the following link: http://www.homepages.ed.ac.uk/pkeightl//dfe_alpha/download-dfe-alpha.html. Folded synonymous and nonsynonymous SFS were used as input in the inferences. *est_alpha_omega* program was used to estimate the proportion of adaptive divergence, *i.e.* alpha.

Coalescent simulations to compare human two-epoch and three-epoch models

To evaluate whether patterns of neutral polymorphism would be predicted to be different under the human two-epoch and three-epoch demographic models, we conducted coalescent simulations under these models using *ms* [52]. Specifically, we simulated 1000 replicates for each scenario and calculated the mean number of synonymous segregating sites across replicates. Both models showed similar numbers of neutral segregating sites (34075 for two-epoch model and 33810 for three-epoch model), suggesting that using population sizes more appropriate for the outgroup population will not affect polymorphism data in the in-group sample.

Composite likelihood approach for testing whether p^+ and s^+ differ across species

We first used *prfreq* to numerically solve the forward diffusion equation of allele frequency change for the specified demographic model and mutation rate to generate a look-up table for the expected number of nonsynonymous substitutions for a range of s^+ (10^{-5}-10^{-2}) for each species. We focused on this range to capture strongly advantageous mutations. For each species, we then did a grid search of $\log_{10}(s^+)$ (-5 to -2) and p^+(0-7.5%). We are interested in this range of strong s^+ because weakly beneficial mutations still segregating in polymorphisms should be taken into account by the DFE being fit to the SFS. We use a Poisson log-likelihood function to calculate the log-likelihood (LL) for each combination of s^+ and p^+. We find the MLE of s^+ and p^+ for each species under each demographic model that maximizes the LL and best fit the observed D_N. This is the full model (H1) where each species is allowed to have its own s^+ and p^+. Our H0 hypothesis is the constrained model, where two or three species under certain demographic scenario have the same s^+ and p^+. The LL of the constrained model is the sum of LL for each s^+ and p^+ for each species under comparison. We then find the MLE for the constrained model and calculate the likelihood ratio between H1 and H0.

Composite likelihood approach for testing whether p^+ and gamma$^+$ differ across species

Similarly, we used *prfreq* to generate a look-up table for the expected number of nonsynonymous substitutions for a range of $gamma^+$ (1-10^3) for each species under each demographic model. For each species, we performed a grid search of $\log_{10}(gamma^+)$ (0-3) and p^+(0-7.5%). Note, because effective population sizes differ over several orders of magnitude across our three species, we are searching across drastically different ranges of s^+ as compared to our previous inference described above. We again use a Poisson log-likelihood function to calculate the LL for each combination of $gamma^+$ and p^+. We find the MLE of $gamma^+$ and p^+ for each species under each demographic model that best fit our observed nonsynonymous divergences. This is the full model (H1) where each species is allowed to have its own $gamma^+$ and p^+. Our H0 hypothesis is the constrained model, where two or three species under certain demographic scenarios have the same $gamma^+$ and p^+. The LL of the constrained model is the sum of LL for each $gamma^+$ and p^+ for each species under comparison. We then find the MLE for the constrained model and calculate the likelihood ratio between H1 and H0.

Conditional likelihoods

To examine whether all three species could have the same s^+ or p^+, we examine the conditional likelihoods. To make the conditional likelihood curve for p^+, for each s^+ value, we look for the p^+ that maximizes the likelihood as well as the p^+ values that have a likelihood within three LL of this maximum LL for this s^+ (*i.e.* $p^+ | s^+$). To make the conditional likelihood curve for s^+, for each p^+ value, we look for the s^+ that maximizes the likelihood as well as the s^+ values that have a LL within three LL units of this maximum LL for this p^+ (*i.e.* $s^+ | p^+$).

Similarly, we examine whether all three species could have the same $gamma^+$ or p^+. To make the conditional likelihood curve for p^+, for each $gamma^+$ value, we look for the p^+ that maximizes the likelihood as well as the p^+ values that have a likelihood within three LL of this maximum LL for this $gamma^+$ (*i.e.* $p^+ | gamma^+$). To make the conditional likelihood curve for $gamma^+$, for each p^+ value, we look for the $gamma^+$ that maximizes the likelihood as well as the $gamma^+$ values that have a LL within three LL units of this maximum LL for this p^+ (*i.e.* $gamma^+ | p^+$).

Estimating alpha on the human lineage

Human-chimp substitutions were polarized by the macaque sequence. Substitutions were assigned to human lineage if the bases differ between human and macaque, but were the same between chimpanzee and macaque in the pairwise genome alignments. Substitutions that are in regions where the human and macaque sequences were un-alignable and substitutions that differ among human, chimp and macaque cannot be polarized (3.8% total) and were filtered out. The total length of coding regions was scaled by this filter accordingly.

To estimate the expected number of substitutions on human lineage, we follow the method described in section *prfreq estimates of alpha,* with several modifications. Specifically, 1) we use *prfreq* to compute the expected D_S count for both human and chimp lineages; 2) we use *prfreq* to compute the expected $D_{S.outgroup}$ count for a population with constant size (i.e. 7067 for 2-epoch model, and 73,000 for 3-epoch model), with a split time equal to the one previously estimated from Ds between human-chimp, 3) we divide $D_{S.outgroup}$ by 2 to find the number of nonsynonymous substitutions fixed on the chimp lineage. Then we subtract that number from the expected D_S count from both lineages. This reminder should be the expected synonymous divergence on the human lineage. 4) we adjust *tdiv* to match the expected synonymous divergence on the human lineage with the observed synonymous divergence on the human

lineage. 5) With the adjusted *tdiv*, we similarly estimate the expected nonsynonymous divergence on the human lineage using the same approach as for synonymous divergence described in step 3, except we include the DFE of deleterious mutations. 6) alpha is calculated using equation (3).

Filtering to only include sites not affected by biased gene conversion

Full polymorphism and substitution datasets of humans and mice were filtered to keep only SSWW mutations that were not affected by BGC. This includes only A to T, T to A, C to G and G to C changes. These changes are only a small subset of all variable sites. The nonsynonymous and synonymous sequence lengths (L_{NS}, L_S) depend on the transition/transversion ratio and the CpG mutational bias. SSWW mutations are all transversions and do not included any CpG mutations, leading to a multiplier of L_{NS} = 5.21 x L_S in both humans and mice. To compute this: 1) we used numbers of 0-, 2-, 3- and 4- fold sites in human from Veeraham [53]; 2) we consider all 2-fold sites to be nonsynonymous (because SSWW mutations are all transversions); and 3) we do not consider a mutational bias of CpG sites (because CpG sites are not included in the SSWW set). In addition, because SSWW mutations are only a small subset of all mutations, mutation rates need to be scaled down to the SSWW specific mutation rate. To estimate the demographic parameters and DFE for SSWW polymorphisms in mice, first, we used the observed number of synonymous SSWW polymorphisms to estimate the mutation rate for mice SSWW mutations to be 5.99E-10. Then, using the SFS for SSWW synonymous polymorphisms, we inferred that the ancestral population size is approximately 246,256 which expanded 1.7-fold approximately 262,000 generations ago (supplementary table S3). Conditional on this demographic model, we estimated the DFE for new nonsynonymous SSWW mutations in mice. We assume that the DFE follows a gamma distribution and estimate its shape parameter alpha to be 0.21 and scale parameter beta to be 0.050 (supplementary table S3). These estimates are within the same magnitude of the estimates from the full dataset and a previous study [17]. We re-estimated the mouse-rat divergence time that fits best with the observed SSWW D_S. We then re-inferred p^+ and $s^+/gamma^+$ as done previously, using the filtered data, new mutation rates, and new values of L_{NS}.

Data availability

The datasets analyzed during the current study are available in these published reference articles 38-40 as described in Methods section.

Acknowledgements

We thank Lawrence Uricchio and David Enard for advice, discussions, and sharing their manuscript, and Tanya Phung, Jazlyn Mooney, Clare Marsden and Jesse Garcia for helpful comments on our manuscript. This work was supported by a Searle Scholars Fellowship and NIH Grant R35GM119856 (to K.E.L.). We acknowledge support from a QCB Collaboratory Postdoctoral Fellowship to Y.Z. and the QCB Collaboratory Community directed by Matteo Pellegrini.

Author contributions

K.E.L conceived of and supervised the study. Y.Z. carried out all analyses of alpha. C.D.H carried demographic and gamma-DFE inference based on the SFS. R.W.D. processed mice raw

data to genotypes. Y.Z. generated all figures. Y.Z., C.D.H., R.W.D and K.E.L. all participated in manuscript preparation.

Competing interests

The authors declare no competing interests.

References

1. Kimura, M. *The Neutral Theory of Molecular Evolution*. (Cambridge University Press, 1983).
2. Fay, J. C., Wyckoff, G. J. & Wu, C.-I. Testing the neutral theory of molecular evolution with genomic data from Drosophila. *Nature* **415,** 1024–1026 (2002).
3. Smith, N. G. C. & Eyre-Walker, A. Adaptive protein evolution in Drosophila. *Nature* **415,** 1022–1024 (2002).
4. Fay, J. C. Weighing the evidence for adaptation at the molecular level. *Trends Genet. TIG* **27,** 343–349 (2011).
5. Andolfatto, P. Adaptive evolution of non-coding DNA in Drosophila. *Nature* **437,** 1149–1152 (2005).
6. Eyre-Walker, A. & Keightley, P. D. 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).
7. Boyko, A. R. *et al.* Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* **4,** e1000083 (2008).
8. Gossmann, T. I. *et al.* Genome Wide Analyses Reveal Little Evidence for Adaptive Evolution in Many Plant Species. *Mol. Biol. Evol.* **27,** 1822–1832 (2010).
9. Foxe, J. P. *et al.* Selection on Amino Acid Substitutions in Arabidopsis. *Mol. Biol. Evol.* **25,** 1375–1383 (2008).
10. Phifer-Rixey, M. *et al.* Adaptive evolution and effective population size in wild house mice. *Mol. Biol. Evol.* **29,** 2949–2955 (2012).
11. Strasburg, J. L. *et al.* Effective population size is positively correlated with levels of adaptive divergence among annual sunflowers. *Mol. Biol. Evol.* **28,** 1569–1580 (2011).
12. Gossmann, T. I., Keightley, P. D. & Eyre-Walker, A. The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. *Genome Biol. Evol.* **4,** 658–667 (2012).
13. Galtier, N. Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis. *PLoS Genet.* **12,** e1005774 (2016).
14. Corbett-Detig, R. B., Hartl, D. L. & Sackton, T. B. Natural selection constrains neutral diversity across a wide range of species. *PLoS Biol* **13,** e1002112 (2015).
15. Nam, K. *et al.* Evidence that the rate of strong selective sweeps increases with population size in the great apes. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 1613–1618 (2017).
16. Rousselle, M., Mollion, M., Nabholz, B., Bataillon, T. & Galtier, N. Overestimation of the adaptive substitution rate in fluctuating populations. *Biol. Lett.* **14,** 20180055 (2018).
17. Lourenço, J. M., Glémin, S. & Galtier, N. The Rate of Molecular Adaptation in a Changing Environment. *Mol. Biol. Evol.* **30,** 1292–1301 (2013).
18. Huber, C. D., Kim, B. Y., Marsden, C. D. & Lohmueller, K. E. Determining the factors driving selective effects of new nonsynonymous mutations. *Proc. Natl. Acad. Sci.* **114,** 4465–4470 (2017).

19. Smith, N. G. C. & Eyre-Walker, A. Adaptive protein evolution in Drosophila. *Nature* **415,** 1022–1024 (2002).
20. Chen, F.-C. & Li, W.-H. Genomic Divergences between Humans and Other Hominoids and the Effective Population Size of the Common Ancestor of Humans and Chimpanzees. *Am. J. Hum. Genet.* **68,** 444–456 (2001).
21. Hernandez, R. D. *et al.* Demographic Histories and Patterns of Linkage Disequilibrium in Chinese and Indian Rhesus Macaques. *Science* **316,** 240–243 (2007).
22. Hobolth, A., Christensen, O. F., Mailund, T. & Schierup, M. H. Genomic Relationships and Speciation Times of Human, Chimpanzee, and Gorilla Inferred from a Coalescent Hidden Markov Model. *PLOS Genet.* **3,** e7 (2007).
23. Burgess, R. & Yang, Z. Estimation of hominoid ancestral population sizes under bayesian coalescent models incorporating mutation rate variation and sequencing errors. *Mol. Biol. Evol.* **25,** 1979–1994 (2008).
24. Prado-Martinez, J. *et al.* Great ape genetic diversity and population history. *Nature* **499,** 471–475 (2013).
25. Schrago, C. G. The Effective Population Sizes of the Anthropoid Ancestors of the Human–Chimpanzee Lineage Provide Insights on the Historical Biogeography of the Great Apes. *Mol. Biol. Evol.* **31,** 37–47 (2014).
26. Andolfatto, P., Wong, K. M. & Bachtrog, D. Effective Population Size and the Efficacy of Selection on the X Chromosomes of Two Closely Related Drosophila Species. *Genome Biol. Evol.* **3,** 114–128 (2011).
27. Ness, R. W. *et al.* Nuclear Gene Variation in Wild Brown Rats. *G3 Genes Genomes Genet.* **2,** 1661–1664 (2012).
28. Sawyer, S. A. & Hartl, D. L. Population genetics of polymorphism and divergence. *Genetics* **132,** 1161–1176 (1992).
29. Duret, L. & Galtier, N. Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes. *Annu. Rev. Genomics Hum. Genet.* **10,** 285–311 (2009).
30. Lachance, J. & Tishkoff, S. A. Biased Gene Conversion Skews Allele Frequencies in Human Populations, Increasing the Disease Burden of Recessive Alleles. *Am. J. Hum. Genet.* **95,** 408–420 (2014).
31. Bolívar, P., Mugal, C. F., Nater, A. & Ellegren, H. Recombination Rate Variation Modulates Gene Sequence Evolution Mainly via GC-Biased Gene Conversion, Not Hill–Robertson Interference, in an Avian System. *Mol. Biol. Evol.* msv214 (2015). doi:10.1093/molbev/msv214
32. Corcoran, P., Gossmann, T. I., Barton, H. J., Slate, J. & Zeng, K. Determinants of the Efficacy of Natural Selection on Coding and Noncoding Variability in Two Passerine Species. *Genome Biol. Evol.* **9,** 2987–3007 (2017).
33. Robinson, M. C., Stone, E. A. & Singh, N. D. Population genomic analysis reveals no evidence for GC-biased gene conversion in Drosophila melanogaster. *Mol. Biol. Evol.* **31,** 425–433 (2014).
34. Kong, A. *et al.* Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488,** 471 (2012).
35. Connallon, T. & Clark, A. G. The distribution of fitness effects in an uncertain world. *Evol. Int. J. Org. Evol.* **69,** 1610–1618 (2015).
36. Valentine, J. W., Collins, A. G. & Meyer, C. P. Morphological complexity increase in metazoans. *Paleobiology* **20,** 131–142 (1994).

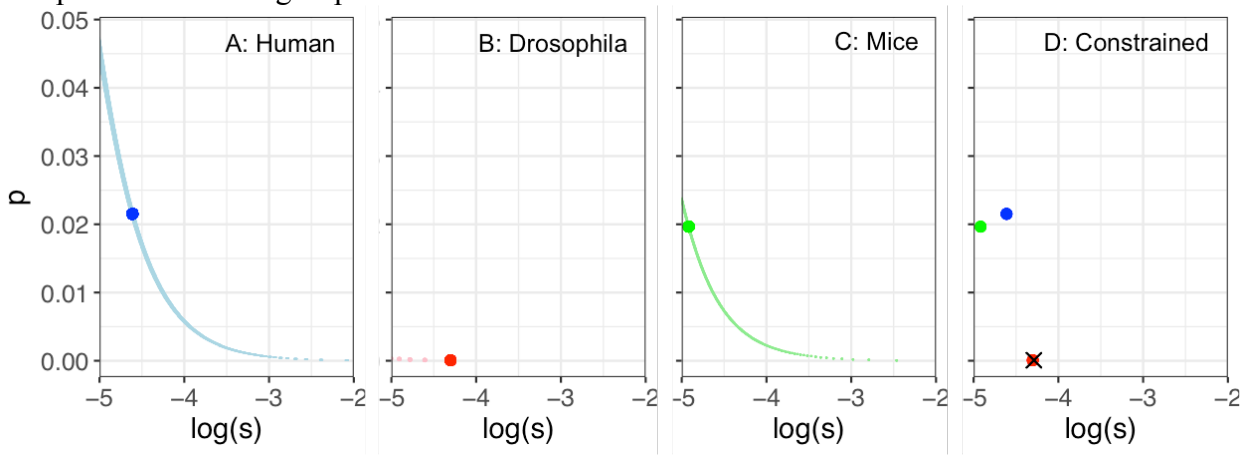
773 37. Stumpf, M. P. H. *et al.* Estimating the size of the human interactome. *Proc. Natl. Acad. Sci.*
774 **105,** 6959–6964 (2008).
775 38. Consortium, T. 1000 G. P. An integrated map of genetic variation from 1,092 human
776 genomes. *Nature* **491,** 56–65 (2012).
777 39. Lack, J. B. *et al.* The Drosophila Genome Nexus: A Population Genomic Resource of 623
778 Drosophila melanogaster Genomes, Including 197 from a Single Ancestral Range
779 Population. *Genetics* **199,** 1229–1241 (2015).
780 40. Hu, T. T., Eisen, M. B., Thornton, K. R. & Andolfatto, P. A second-generation assembly of
781 the Drosophila simulans genome provides new insights into patterns of lineage-specific
782 divergence. *Genome Res.* **23,** 89–98 (2013).
783 41. Halligan, D. L., Oliver, F., Eyre-Walker, A., Harr, B. & Keightley, P. D. Evidence for
784 Pervasive Adaptive Protein Evolution in Wild Mice. *PLOS Genet.* **6,** e1000825 (2010).
785 42. Halligan, D. L. *et al.* Contributions of Protein-Coding and Regulatory Change to Adaptive
786 Molecular Evolution in Murid Rodents. *PLOS Genet.* **9,** e1003995 (2013).
787 43. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform.
788 *Bioinforma. Oxf. Engl.* **25,** 1754–1760 (2009).
789 44. Lunter, G. & Goodson, M. Stampy: a statistical algorithm for sensitive and fast mapping of
790 Illumina sequence reads. *Genome Res.* **21,** 936–939 (2011).
791 45. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing
792 next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).
793 46. Yang, H. *et al.* A customized and versatile high-density genotyping array for the mouse. *Nat.*
794 *Methods* **6,** 663–666 (2009).
795 47. Bustamante, C. D. *et al.* Natural selection on protein-coding genes in the human genome.
796 *Nature* **437,** 1153–1157 (2005).
797 48. Torgerson, D. G. *et al.* Evolutionary Processes Acting on Candidate cis-Regulatory Regions
798 in Humans Inferred from Patterns of Polymorphism and Divergence. *PLoS Genet* **5,**
799 e1000592 (2009).
800 49. Enard, D., Messer, P. W. & Petrov, D. A. Genome-wide signals of positive selection in
801 human evolution. *Genome Res.* **24,** 885–895 (2014).
802 50. Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H. & Bustamante, C. D. Inferring the
803 Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency
804 Data. *PLoS Genet* **5,** e1000695 (2009).
805 51. Kim, B. Y., Huber, C. D. & Lohmueller, K. E. Inference of the Distribution of Selection
806 Coefficients for New Nonsynonymous Mutations Using Large Samples. *Genetics* **206,** 345–
807 361 (2017).
808 52. Hudson, R. R. Generating samples under a Wright–Fisher neutral model of genetic variation.
809 *Bioinformatics* **18,** 337–338 (2002).
810 53. Veeramah, K. R., Gutenkunst, R. N., Woerner, A. E., Watkins, J. C. & Hammer, M. F.
811 Evidence for Increased Levels of Positive and Negative Selection on the X Chromosome
812 versus Autosomes in Humans. *Mol. Biol. Evol.* **31,** 2267–2282 (2014).
813

814 Table 1. Estimates of α using different methods

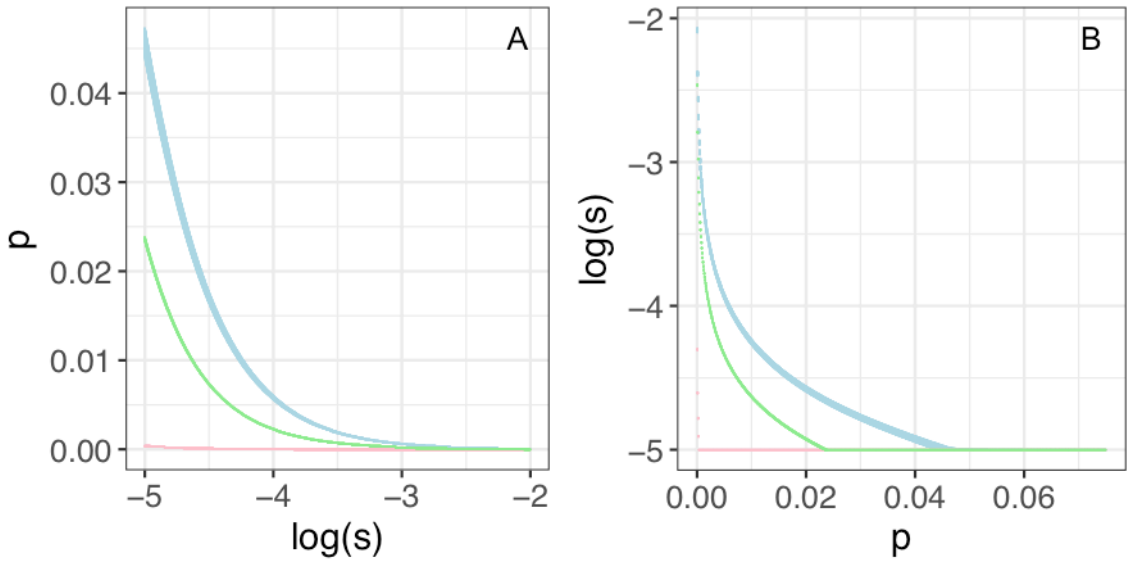
			Method of Inference				
	Species	Outgroup	original: all	original: MAF>20%	this paper: two-epoch	this paper: three-epoch	DFEα
full data	Human	Chimpanzee	-0.41	-0.01	0.15	0.33	0.24
	Human	Macaque	-0.70	-0.22	-0.04	0.30	0.02
	D. melanoga	*D. simulans*	-0.13	0.49	0.58	0.68	0.71
	Mice	Rat	0.25	0.40	0.48	0.38	0.51
SSWW only	Human	Chimpanzee	-0.37	0.08	0.17	0.34	-
	Mice	Rat	-0.14	0.08	0.33	0.19	-

815
816

fig. 1
Log-likelihood surfaces for p^+ and s^+ for different species. (A) Human. (B) Drosophila. (C) Mice. (D) The constrained model, H0, where p^+ and s^+ are constrained to be the same across all three species. Log-Likelihoods are calculated using grid search method of $\log_{10}(s)$ in the range of -5 to -2 and p^+ in the range of 0-7.5%. The large point in panels A-C represents the MLE for each species, and grid points within 3 LL of each MLE are shown by the solid lines. In panel D, MLEs for each species are represented as larger points and the black cross represents the MLE of the constrained model. Note three-epoch demographic model is used for each species and we use chimpanzee as the outgroup for humans.



828 fig. 2
829 Conditional log-likelihood surfaces. (A) Maximizing p^+ given particular values of s^+ and (B)
830 maximizing s^+ given particular values of p^+. Only grid points within 3 LL of the MLEs for each
831 parameter for each species are shown.



832

fig. 3

Log-likelihood surfaces for p^+ and $gamma^+$ for different species. (A) Human. (B) Drosophila. (C) Mice. (D) The constrained model, H0, where p^+ and $gamma^+$ are constrained to be the same across all three species. Log-likelihoods are calculated using grid search method of $\log_{10}(gamma)$ in the range of 0-3 and p^+ in the range of 0-7.5%. The large point in panels A-C represents the MLE for each species, and grid points within 3 LL of each MLE are shown by the solid lines. In panel D, MLEs for each species are represented as larger points and the black cross represents the MLE of the constrained model. Note, the three-epoch demographic model is used for each species and we use the chimpanzee as the outgroup for humans.

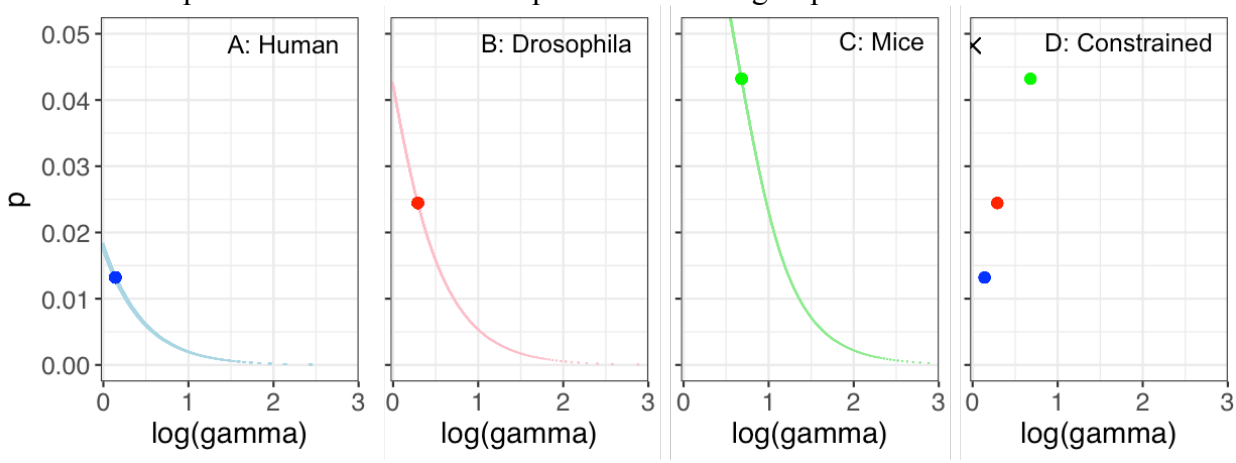


fig. 4

Log-likelihood surfaces for sites unaffected by biased gene conversion (SSWW sites in mammals). (A-C) show the log-likelihood surfaces for p^+ and s^+ for different species. (D) shows the constrained model, H0, where p^+ and s^+ are constrained to be the same across all three species. Log-Likelihoods are calculated using grid search method of $\log_{10}$(s) in the range of -5 to -2 and p^+ in the range of 0-7.5%. The large point in panels A-C represents the MLE for each species, and grid points within 3 LL of each MLE are shown by the solid lines. In panel D, MLEs for each species are represented as larger points and the black cross represents the MLE of the constrained model. Note, the three-epoch demographic model is used for each species and we use the chimpanzee as the outgroup for humans. (E) shows the conditional log-likelihood surface maximizing p^+ given particular values of s^+ and (F) shows the conditional log-likelihood surface maximizing s^+ given particular values of p^+. In panels E-F, only grid points within 3 LL of the MLEs of for each parameter for each species are shown.