

Molecular evolution in large steps - Codon substitutions under positive selection

Qingjian Chen¹, Ziwen He¹, Ao Lan¹, Haijun Wen^{1*} and Chung-I Wu^{1,2,3*}

¹State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China.

²CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China.

³Department of Ecology and Evolution, University of Chicago, Chicago Illinois 60637, USA

*Corresponding author:

Chung-I Wu, ciwu@uchicago.edu

Haijun Wen, wenhj5@mail.sysu.edu.cn

Abstract

Molecular evolution is believed to proceed in small steps. The step size can be defined by a distance reflecting physico-chemical disparities between amino acid (AA) pairs that can be exchanged by single 1 bp mutations. We show that AA substitution rates are strongly and negatively correlated with this distance but only when positive selection is relatively weak. We use the McDonald and Kreitman (MK) test to separate the influences of positive and negative selection. While negative selection is indeed stronger on AA substitutions generating larger changes in chemical properties of amino acids, positive selection operates by different rules. For 65 of the 75 possible pairs, positive selection is comparable in strength regardless of AA distance. However, the 10 pairs under the strongest positive selection all exhibit large leaps in chemical properties. Five of the 10 pairs are shared between hominoids and *Drosophila*, thus hinting at a common but modest biochemical basis of adaptation across taxa. The hypothesis that adaptive changes often take large functional steps will need to be extensively tested. If validated, molecular models will need to better integrate positive and negative selection in the search for adaptive signal.

Introduction

It is generally accepted that natural selection favors incremental small-step changes. In the accompanying paper ([Chen and Wu](#)), we use physico-chemical distances between amino acids (AAs) as a measure of step size in evolution. When negative selection is the main driving force, similar AAs are indeed more likely to be exchanged. After all, a mutant must not be too different from the wild type in order to avoid elimination. This intuition is a key rule of neutral molecular evolution ([Kimura 1983](#)).

On the other hand, it is not obvious that positive selection should also favor small-step changes. The most common reference is Fisher's geometric model (FGM) ([Fisher 1930](#)), whereby small-step changes would have a better chance of being advantageous than large-step ones. Nevertheless, a key element of FGM is still the avoidance of negative selection. In FGM, each mutation is assumed to be highly pleiotropic and large-step changes are likely to be deleterious for some phenotypes ([Wagner and Zhang 2011](#)).

An opposite argument for large-step evolution can be stated as follows: Natural selection, either positive or negative, can “discern” large-step changes better than small-step ones. For example, replacing isoleucine (Ile) with the chemically similar Valine (Val) would not alter the protein structure as much as its replacement by Arginine (Arg), which is very different chemically. In this view, an Ile → Arg replacement may be either much worse, or much better, than the Ile → Val replacement. Therefore, while negative selection would accept small-step changes (e.g., Ile → Val), positive selection might in fact favor large-step ones (e.g., Ile → Arg). The conventional wisdom is a postulate, not a fact.

Clearly, the arguments must be resolved by empirical means. The companion study ([Chen and Wu](#)) has shown that AA substitutions start to deviate from the small-step rule when negative selection becomes weaker and/or positive selection becomes stronger. In the same vein, this study aims to separate the two effects on the rate of AA substitutions.

Results

I. Amino acid distance in relation to the action of negative selection

We partitioned the conventional K_a (or D_n) measurement (number of non-synonymous substitutions per site) into 75 classes of substitutions, denoted K_i ($i = 1, 75$; (Tang, et al. 2004)). These 75 classes are AA substitutions that require only a 1 bp change. K_s (or D_s), the number of synonymous substitutions per site, is a separate class. It is reported that the rank order of K_i is highly correlated ($R > 0.9$) across taxa ranging from plants and invertebrates to mammals and primates (Chen and Wu). The main reason for this nearly universal correlation is that K_i is strongly dependent on the physicochemical properties of AAs. We hence define evolutionary AA distance of the i -th pair by $\Delta_U(i) = (U_1 - U_i)/(U_1 - U_{75})$ where U_i is the “universal exchangeability” given in Tang et al. (2004) (Tang, et al. 2004) (see Chen et al. (Chen and Wu) for an update). By this measure, $\Delta_U(1)$ is 0 for the closest pair of AAs (Ser-Thr) and $\Delta_U(75)$ is 1 for the most distant pair (Asp-Tyr).

Fig. 1 here

To see how well Δ_U may account for evolutionary rates, we separated genes into two groups: the slow group consists of the top 3 gene groups in Table 1 (0% - 60%) and the fast group consists of the last gene group (80% - 100%). *Drosophila* (*D. melanogaster* vs. *D. simulans*) and Hominoids (human vs. chimpanzee) were chosen to represent the slowest and fastest evolving taxa in our collection (Fig. 1). We found that K_i of the slow group is highly correlated with Δ_U : the R^2 is 0.904 in *Drosophila* and 0.706 in Hominoids (green lines in Fig. 1A-B). The R^2 values in all other taxa are also > 0.75 (Chen and Wu). Hence, for 60% of the genes, there is a nearly universal relationship between $\Delta_U(i)$ and K_i (Fig. 1D-E). For the fastest evolving 20% of genes, the correlation decreases sharply. R^2 among these loci is 0.798 in *Drosophila* and 0.262 in Hominoids (red lines in Fig. 1A-B).

To further understand the underlying processes, we developed a model with variable negative selection, but without positive selection (see Methods). Fig. 1C shows simulated patterns for the relaxation of negative selection. Note that the correlation remains quite high ($R^2 > 0.8$) as K_a/K_s exceeds 0.5. It is only when K_a/K_s approaches 1 with negative selection becoming fully relaxed, does the correlation breaks down (blue points of Fig. 1C). The observed pattern in Fig. 1A appears reasonably close to the simulated pattern in Fig. 1C, which shows the reduction in R when K_a/K_s increases. The pattern in Fig. 1B between human and chimpanzee, however, is very different from the corresponding simulations in Fig. 1C. The main reason appears to be that, in the fast group of genes, the points are often > 1 and scattered widely.

The correlations in all sampled taxa are given in Fig. 1D-F, where genes are separated into five bins by the rank order of K_a/K_s values with equal number of non-synonymous changes. The X-axis shows mean K_a/K_s ratios of each bin and the Y-axis is the R^2 value of their K_i against Δ_U . Simulated R^2 values are depicted by solid lines (Fig. 1D-E; see Methods). Note that the simulations assume only negative selection in the absence of positive selection. While the simulated R^2 agree reasonably well with the observed values in Fig. 1D (in *Drosophila* and *Arabidopsis*), the agreements in other taxa in Fig. 1E-F are much worse. Fig. 1E, which presents the five comparisons between vertebrate species (a pair of reptiles, a pair of birds, and three pairs of mammals), shows lower than expected correlations as K_a/K_s increases. Fig. 1F on the human-chimpanzee comparison shows the weakest correlations of all taxa.

Overall, the high correlation between the observed K_i and Δ_U starts to break down when the K_a/K_s ratio rises above 0.3 and especially above 0.5 (Fig. 1D-F). Loci with high K_a/K_s ratios are presumably less influenced by negative selection, but also possibly undergo adaptive evolution. In the next section, we attempt to separate the two effects in order to analyze the action of positive selection.

II. Separating positive and negative selection on AA substitutions

A commonly used approach to separating positive and negative selection is the McDonald and Kreitman (MK) test (McDonald and Kreitman 1991). The MK test compares the K_a/K_s ratio for between-species divergence with the P_a/P_s ratio for within-species polymorphism (McDonald and Kreitman 1991;

[Sawyer and Hartl 1992](#); [Fay, et al. 2002](#); [Smith and Eyre-Walker 2002](#); [Bustamante, et al. 2005](#); [Shapiro, et al. 2007](#)). The fitness advantage of nonsynonymous substitutions, F_A , is defined as

$$F_A = \frac{\text{Obs}[Ka/Ks]}{\text{Exp}_{\text{neutral}}[Ka/Ks]} \sim \frac{\text{Obs}[Ka/Ks]}{\text{Obs}[Pa/Ps]} \quad \text{Eq. (1.1)}$$

F_A is the observed nonsynonymous substitution relative to the expected, calibrated by K_s . $\text{Exp}_{\text{neutral}}[Ka/Ks]$ can then be replaced by $\text{Obs}[Pa/Ps]$ in Eq. (1.1) if there are few advantageous variants within species. Thus, $F_A = 1$ means that nonsynonymous substitutions are not driven by positive selection. When $F_A > 1$, we interpret F_A as the selective advantage averaged across nonsynonymous substitutions. $F_A < 1$, which is occasionally recorded, implies weaker negative selection in the polymorphism than in the divergence data (see Discussion).

Table 1 here

We applied the MK test to *Drosophila* and hominoids (see Fig. 1A and B). The results (Table 1) show $F_A = 1.98$ and 1.02 for the two taxa (see also Table S1, with a different frequency cutoff). When the genes are divided into five bins according to the Ka/Ks ratio, F_A increases steadily as Ka/Ks gets larger. Although the classification by Ka/Ks can bias the calculation of F_A (see Eq. 1.1), simulations show that the observed increase is faster than the expectation accounting for bias. The trend supports, albeit only indirectly, the suggestion that these fast-evolving genes are under positive selection, reducing the high correlation between K_i and Δ_U .

We then used the MK test on each individual class of AA substitutions that differ by only one bp. As there are 75 such classes, we conducted 75 separate MK tests. For each test, K_a , P_a , and F_A are replaced by K_i , P_i , and F_i for $i = 1, 75$ (see Methods). The fitness advantage of the i -th type of AA substitution, F_i , is

$$F_i \sim \text{Obs}[K_i/K_s] / \text{Obs}[P_i/P_s] \quad \text{Eq.(1.2)}$$

In parallel, we define the intensity of negative selection on the i -th class of AA changes as

$$G_i = 1 - \text{Obs}[P_i/P_s]. \quad \text{Eq.(1.3)}$$

G_i is the proportion of the i -th class of codon mutations eliminated by negative selection. Because low frequency polymorphisms may contain deleterious mutations not yet eliminated ([Shapiro, et al. 2007](#)) and the high frequency portion of the spectrum may harbor advantageous variants ([Fay and Wu 2000](#); [Wang, et al. 2017](#)), we followed the common practice of eliminating loci with low- and high-frequency derived alleles (see Fig. S2 and Methods for details).

Fig. 2 and Table 2 here

With the intensity of positive and negative selection defined by F_i and G_i , we first corroborate the expectation that negative selection tolerates small-step changes. The plots of G_i against Δ_U in Fig. 2A-B indeed show a strong correlation with step size (see also Fig. S1 A-B). In *Drosophila*, $R^2 = 0.89$. The R^2 is lower in hominoids, but still highly significant at 0.56. The two taxa differ substantially in the strength of negative selection as can be seen on the Y-axis: G_i ranges between 0.8 and 0.99 with a mean of 0.93 in *Drosophila* but is much weaker between hominoids (as low as 0.2, with a mean of 0.68). Since the strength of negative selection determines the correlation with Δ_U , the smaller R^2 in hominoids is not surprising.

We now address the strength of positive selection in relation to step size, Δ_U . In Fig. 2C-D (see also Fig. S1 C-D), the correlation can be seen as the composite of two groups of dots: Group I consists of the red dots representing the 10 highest F_i values and Group II consists of the remaining 65 green dots. The regression line for the 65 green dots is nearly flat for both *Drosophila* and hominoids. The small slopes suggest that the strength of positive selection does not depend strongly on Δ_U . Any of the 65 AA substitutions can have a similar likelihood of being advantageous. Red dots, on the other hand, are much often favored by positive

selection and represent AA substitutions with large Δ_U . The analysis suggests that positive selection does not favor small-step changes and the strongest adaptive signals almost always come from large-distance substitutions.

In summary, both positive and negative selection can distinguish, and act on, big-step changes better than small-step modifications. Nevertheless, there is an important difference. While negative selection appears to follow a nearly universal rule (Chen and Wu; Tang, et al. 2004), such consistency across taxa is not expected of positive selection because adaptive changes should be highly dependent on the organisms and their environments. For that reason, the similarity between Fig. 2C and D is somewhat surprising. Table 2 (and Dataset S1, see also Table S2) shows the highest 10 pairs in the F_i ranking. Five of the top 10 pairs are shared between the two taxa ($p = 0.0027$ with the overlap of 1.33 expected from the hypergeometric distribution). It will require extensive analyses beyond *Drosophila* and hominoids to determine if positive selection indeed favors the small subset of AAs given in Table 2.

III. Molecular evolution in small vs. large steps – A model

We have shown that the correlation between the observed K_i/K_s and $\Delta_U(i)$ becomes progressively weaker as the overall K_a/K_s gets higher (Fig. 1A-B). Obviously, the many K_i/K_s values above 1 require incorporation of positive selection (cf. Fig. 1B and C). An expanded model that considers the opposing dynamics of positive and negative selection is necessary. Negative selection works against large $\Delta_U(i)$ changes and positive selection tends to favor them. The model starts with K_a , expressed as

$$K_a = \sum_{i=1}^{75} w_i K_i \quad \text{Eq. (2.1)}$$

where w_i is the weight reflecting the number of sites available for AA exchanges of the i -th pair. The distribution of w_i in the *Drosophila* genome is given in Fig. 3A (top panel) where i is shown as $\Delta_U(i)$.

Fig. 3 here

We now formulate K_i ($i = 1, 75$) of Eq. (2.1), which is the average rate across all sites of a gene or genes.

$$\begin{aligned} R_i &= K_i/K_s = (1 - p_i - q_i) + p_i[N\mu f(N, s_{1i})]/\mu + q_i[N\mu f(N, s_{2i})]/\mu \\ &= 1 + p_i[Nf(N, s_{1i}) - 1] - q_i[1 - Nf(N, s_{2i})] \\ &= 1 + A_i - B_i \end{aligned} \quad \text{Eq. (2.2)}$$

where p_i and q_i are the proportion of advantageous and deleterious mutations, respectively (Ohta and Gillespie 1996; Hartl and Clark 1997; Li 1997; Chen, et al. 2018). μ is the mutation rate and will be canceled out. In the presentation, we drop the subscript i to s_1 and s_2 . We also use s for s_1 and s_2 when they are interchangeable. Then, $f(N, s) = (1 - e^{-2s}) / (1 - e^{-2Ns})$ is the fixation probability of a mutation with a selective coefficient s , where s can be > 0 (denoted by s_1) or < 0 (s_2) and N is the effective population size. $A_i = p_i[Nf(N, s_1) - 1]$ is the positive selection term and $B_i = q_i[1 - Nf(N, s_2)]$ is the negative selection term. The fixation probability of an advantageous mutation is approximately $2s_1 (> 1/N)$ and that of a deleterious mutation is $f(N, s_2) = \varepsilon \sim 0$. Hence, R_i can be simplified as

$$R_i = K_i/K_s \sim 1 + p_i(2Ns_1 - 1) - q_i = 1 + A_i - B_i \quad \text{Eq. (2.3).}$$

and

$$R_a = K_a/K_s = \sum_{i=1}^{75} w_i R_i \quad \text{Eq. (2.4).}$$

Given $p_i(2Ns_1 - 1) > 0$, it is obvious that $q_i \geq 1 - R_i$ and $p_i \leq R_i$. Note that $R_a = 0.3$ is almost at the top of R_a values seen across sequenced genomes (Jordan, et al. 2002; *Drosophila 12 Genomes*, et al. 2007; Chen, et al. 2018) and, for $R_a < 0.3$, $q > 0.7$. This means that at least 70% of mutations are deleterious, allowing negative selection to dominate the overall molecular evolution.

Using Eqs. (2.3-2.4), we then quantified the relationship between R_i and $\Delta_U(i)$ for any given R_a ($=K_a/K_s$). In Fig. 3B-C, the Y-axis shows the strength of positive selection (A_i term of Eq. 2.3; red dots above the dashed neutral line, taken from the *Drosophila* data in Fig. 2C) and negative selection (B_i term; green dots below the dashed line [see the legend]). R_i is the joint product, $1 + A_i - B_i$, shown by blue dots. The X-axis shows the step size of evolution (Δ_U).

Fig. 3B with $R_a = 0.3$ portrays the observed limit of a fast-evolving eukaryotic genome. Although the effects of positive and negative selection are opposite, the overall trend of R_i resembles that of negative selection (blue vs. green dots). It can thus be concluded that the signature of negative selection overwhelms that of adaptive evolution in most eukaryotic genomes. For that reason, the conventional view of small-step evolution (blue dots) is hardly surprising because positive selection fails to offset the impact of negative selection. In rare cases when the strength of both selection regimes is approximately equal $R_a = 1$ (Fig. 3C), the overall pattern still tilts toward small-step changes. Only in the extreme case of $R_a = 2$ (Fig. 3D), where $A_i > B_i$, does the overall R_i begin to show a weak positive correlation with step size.

IV. Rate of adaptive evolution as a function of step size

So far we have established the efficacy of positive selection on mutations of different step sizes. The rate of adaptive evolution, however, is the product of the mutation rate and the efficacy of positive selection. Since estimation of this rate of adaptive evolution demands accounting for the influence of negative selection, it is rarely attempted (see below). The available mutational inputs of all $\Delta_U(i)$ in *Drosophila* are shown in the top panel of Fig. 3A, as discussed above. Large Δ_U changes indeed occur much less frequently than small-step mutations, a well-known property of the code table itself. In the code table, neighboring amino acids tend to be physico-chemically similar ([Haig and Hurst 1999](#)). The middle panel in Fig. 3A reproduces the effect of positive selection from Fig. 2C. Combining the top and middle panels of Fig. 3A, the rate of evolution in the bottom panel appears to tilt toward small-step changes. Clearly, mutational input outweighs selection efficacy in determining the overall rate. A recent study ([Bergman and Eyre-Walker 2018](#)) on the rate of adaptive evolution presents a pattern that is similar to (but not identical with) the bottom panel of Fig. 3A. In comparison, this study aims to show the effect of selection (middle panel), independent of the mutational input (top panel).

Discussion

The conclusion that sets our study apart from the convention is that positive selection does not favor small-step changes (Fig. 2C-D). Although we use $\Delta_U(i)$ as a proxy for functional changes, the analysis should have broader implications since any measure of phenotypic change can be adopted as the step size. For example, we may consider the evolved level of gene expression as the step size when studying mutations occurring in enhancers or promoters. In this sense, the small-step hypothesis for positive selection is a curious concept. When the environment changes (say, woody plants invading the intertidal zone ([Xu, et al. 2017](#); [He, Li, et al. 2018](#)), it may require large jumps in the expression of many genes. The genome has to respond adaptively and, importantly, small-effect mutations may not be the best solution. Similarly, if the environmental changes demand modifications in the functionally important part of a protein, mutations have to fall in that part even if such alternations would have a large phenotypic effect.

In short, large- vs. small- step adaption is dictated by the environment. Small-step changes can be a good strategy for fine-tuning the phenotype, especially in an unchanging environment where the avoidance of deleterious effects is paramount. Positive selection in a new or changing environment would operate very differently. It has indeed been reported that in extreme environments or under artificial selection there is often an excess of radical amino acid substitutions ([Lu, et al. 2006](#); [Luo, et al. 2017](#); [Xu, et al. 2017](#)).

The view of small-step evolution prevails mainly because the signature of negative selection almost always dominates (see Fig. 3). Detection of the less-frequent positive selection would require filtering out the effects of negative selection. To that end, we applied the MK test to the 75 classes of AA substitutions. It should be noted that this study focuses on the *relative* F_i among the 75 classes. Hence, factors affecting the

absolute magnitude of F_i 's are of a lesser concern ([Fay, et al. 2002](#); [Charlesworth and Eyre-Walker 2008](#); [Eyre-Walker and Keightley 2009](#); [He, Chen, et al. 2018](#)).

The prevalence of negative over positive selection may often have a confounding effect on the interpretation of results ([Wang, et al. 2017](#)). For example, previous studies have compared radical (K_r , or large-step changes) with conservative (K_c ; small-step changes) AA substitutions. K_r/K_c has been shown to be higher than the overall K_a/K_s ratios ([Hughes, et al. 1990](#); [Zhang 2000](#); [Hanada, et al. 2007](#)). The trend, however, can reflect either positive selection elevating K_r or negative selection reducing K_c . In this context, Tang et al.'s two-fold rule ([Tang and Wu 2006](#)) is another attempt at resolving this issue. They noted that the average of the top 10 K_i classes is nearly twice the value of K_a . The accompanying study shows that the rule is correct only when $K_a/K_s < 0.5$ ([Chen and Wu](#)). This rule is hence not applicable for interpreting positive selection, which requires $K_a/K_s > 1$.

The concept that “nature abhors big changes” has been part of both selectionism and neutralism ([Fisher 1930](#); [Kimura 1983](#)). This concept is only half right since it only applies to negative selection. While negative selection follows tractable common rules, the pattern of positive selection is much more taxon- and environment- dependent. The interpretation of positive selection will remain uncertain until negative selection is fully and rigorously analyzed ([He, Chen, et al. 2018](#)). The latter will be an immediate challenge for the study of molecular evolution.

Materials and Methods

The definition of Δ_U

Functional distances between AA pairs are always based on the physicochemical properties, such as the Miyata distance ([Miyata, et al. 1979](#)). However, AA distances based on some physicochemical properties cannot account for all of the evolutionary exchangeability (EE) variance among amino acids. For example, AA volume differences has a prominent effect on AA properties but only accounts for 27% of the EE variance. Even the first principal component, a composite of 48 physicochemical properties, can only account for 60% of the exchangeability variance among amino acids ([Chen and Wu](#)).

A simple way to define a distance between AA pairs is directly based on their EE difference from genome data. These EE differences are highly conserved across a wide taxonomic range. We hence define an AA distance based on U_i , where U_i is the “universal exchangeability” given by Tang et al. (2004) ([Tang, et al. 2004](#)). Here U_i is the evolutionary exchangeability between each AA pair based on genome data. The evolutionary AA distance Δ_U is defined by $\Delta_U(i) = (U_1 - U_i) / (U_1 - U_{75})$, where U_1 is the most exchangeable AA pair (Ser-Thr) and U_{75} is the least exchangeable (Asp-Tyr). Hence $\Delta_U(1) = 0$ for the most similar pair (Ser-Thr) and $\Delta_U(75) = 1$ for the most dissimilar pair (Asp-Tyr). The $\Delta_U(i)$ is positively correlated with the functional distance.

The definition of negative and positive selection.

A widely-used method for detecting positive selection in coding regions is the MK test, which compares K_a/K_s relative to P_a/P_s ([McDonald and Kreitman 1991](#)). The strength of positive selection is defined as $F_A = \text{Obs}[K_a/K_s] / \text{Obs}[P_a/P_s]$. $F_A > 1$ indicates the existence of positive selection. In contrast, $F_A < 1$ means there is a relaxation of negative selection in polymorphism relative to divergence. Here, we do the MK test for each 75 amino acid pairs by dividing non-synonymous changes into 75 parts.

$$F_i \sim \text{Obs}[K_i/K_s] / \text{Obs}[P_i/P_s]$$

Strength of negative selection for each Δ_U is estimated from the polymorphism data. In the polymorphism, deleterious mutations come into the population at lower allele frequency and are then removed by negative selection. The A/S ratios of polymorphism are very high at lower allele frequencies and drop at allele prevalence increases (see Fig. S2). This drop tapers off at a certain point. We use the frequency range that has uniform P_a/P_s values in our analyses. Negative selection is then defined by

$$G_i = 1 - P_i/P_s.$$

G_i is the proportion of mutations that are eliminated by negative selection.

Multiple-alignment Data

There are eight pairs of species in Fig. 1D-E, including a pair of *Arabidopsis*, *Drosophila*, reptile, birds, and four pairs of mammals. Details about how to deal with the genome data can be found in our accompanying paper ([Chen and Wu](#)). The *Mus-Rattus* comparison was removed out as the U_i values were deduced from the comparison between *Mus-Rattus* and a pair of yeasts ([Tang, et al. 2004](#)).

Polymorphism Data

Sequences of 334 *Drosophila melanogaster* lines were collected from the *Drosophila* Genome Nexus ([Lack, et al. 2015](#)), including 137 lines from the *Drosophila* Population Genomics Project Phase 2 (DPGP2) ([Pool, et al. 2012](#)) and 197 lines from DPGP3 ([Lack, et al. 2015](#)). After masking identical by descent and admixed regions, we assembled 1,360,772 SNPs in *Drosophila melanogaster*. Human SNPs were downloaded from the 1000 Genomes Project ([Consortium, et al. 2015](#)).

To estimate derived allele frequencies, adjacent species were assigned as the ancestral state. The chimpanzee and *Drosophila simulans* genomes were used as the ancestor state for human and *Drosophila melanogaster*. If the states were unknown in the adjacent species, the higher frequency alleles were assigned as ancestral states. Divergence comparisons were based on *D. melanogaster* vs. *D. simulans* and human vs. chimpanzee. The divergence data of the two paired species were extracted from multiple alignment data above. There are 9,710 orthologous genes in *D. melanogaster* vs. *D. simulans* comparison and 11,571 in the human vs. chimpanzee comparison. Polymorphic mutations that were outside the orthologous genes above were removed. Especially, we masked CpG related mutations in both human polymorphism and divergence (CG => TG, CG => CA were removed).

Populations can tolerate deleterious mutations at a relatively low frequency. We removed deleterious mutations by setting a cutoff so that the A/S ratios were stable at the higher frequencies (see Fig. S2) (Shapiro, et al. 2007). We set two different frequency cutoff to test the influence of this choice on the results. One group is 0.2-0.8 for Hominoids and 0.15-0.85 for *Drosophila* (Results in main text). The other group is 0.2-0.95 for Hominoids and 0.15-0.95 for *Drosophila* (Results in the Supplementary Materials). Very high frequency mutations (0.95-1) were removed from further analyses because there were too many such mutations, hitting at possible positive selection in this class (Fay and Wu 2000; Wang, et al. 2017).

Relaxation of negative selection models.

Fitness effects of new deleterious mutations were drawn from the exponential distribution. Most deleterious mutations are under strong negative selection, thus the fitness is nearly 0. A small fraction of mutations are under weak negative selection with fitness far greater than 0.

$$f(x) = \lambda e^{-\lambda x}$$

Here, $f(x)$ is the frequency of mutations. The x is the fitness of new deleterious mutations. The fixation probability of such mutations is the cumulative fitness from the least fit to current fitness, thus

$$F(x) = \int_0^x f(x)dx = 1 - e^{-\lambda x}$$

Ka/Ks is the fixation probability of new non-neutral relative to neutral mutations. Thus $F(x) = Ka/Ks$ in the case where only deleterious mutations are under consideration.

$$Ka/Ks = 1 - e^{-\lambda x}$$

Here, we connect the relationship between the fitness of a mutation x and the Ka/Ks ratio. As the fitness of a new mutation is in the inverse relationship with the strength of negative selection, mutations under strong negative selection undergo significant fitness reduction. The equation above thus links the strength of negative selection and the Ka/Ks ratio. When mutations are under strong negative selection, $x \sim 0$ and Ka/Ks is nearly 0. When constraint is absent, $x \sim \text{infinite}$ and Ka/Ks is 1.

Thus we define the fitness of new deleterious mutations x and Ki/Ks for different $\Delta_U (i=1:75)$ as follows:

$$Ki/Ks = 1 - e^{-\lambda_i x} \quad \text{Eq.(3.1)}$$

We simulated relaxation of negative selection based on Eq. (3.1). First, we set x as a constant value and obtained $\lambda_i (i=1:75)$ for each Δ_U , under the lowest total Ka/Ks ratio situation from the real data. We get the Ki/Ks values for each Δ_U of the slowest evolving genes accounting for 20% of all non-synonymous changes and calculate the corresponding $\lambda_i (i=1:75)$ from Eq. (3.1). Relaxation of negative selection was simulated by predicting Ki/Ks changes as the total Ka/Ks ratio increased. We simply increased the fitness x and calculated the corresponding theoretical Ki/Ks values ($i=1:75$) based on Eq. (3.1). By changing the magnitude of x , the theoretical Ki/Ks values under every possible total Ka/Ks ratio were calculated. In the

complete absence of constraint, $\lambda_i x$ values were nearly *infinite* and theoretical K_i/K_s reached their upper limit, $K_i/K_s \sim 1$ for most Δ_U .

Acknowledgements

We would like to thank Pei Lin and members of Wu Lab for discussions and advices. This work was supported by the National Natural Science Foundation of China (91731301 and 31730046), the 985 Project (33000-18831107), and National Key Basic Research Program of China (2014CB542006).

References

- Bergman J, Eyre-Walker A. 2018. Does Adaptive Protein Evolution Proceed by Large or Small Steps at the Amino Acid Level? bioRxiv: <https://doi.org/10.1101/379073>.
- Bustamante CD, Fledel-Alon A, Williamson S, Nielsen R, Hubisz MT, Glanowski S, Tanenbaum DM, White TJ, Sninsky JJ, Hernandez RD, et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* 437:1153-1157.
- Charlesworth J, Eyre-Walker A. 2008. The McDonald-Kreitman test and slightly deleterious mutations. *Mol Biol Evol* 25:1007-1015.
- Chen B, Shi Z, Chen Q, Shibata D, Wen H, Wu C-I. 2018. Quasi-neutral molecular evolution - When positive and negative selection cancel out. bioRxiv: <https://doi.org/10.1101/330811>.
- Chen Q, Wu C-I. Evolution in small steps - A nearly-universal rule of negative selection on codons. unpublished.
- Consortium GP, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. 2015. A global reference for human genetic variation. *Nature* 526:68-74.
- Drosophila 12 Genomes C, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, et al. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* 450:203-218.
- Eyre-Walker A, Keightley PD. 2009. Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol Biol Evol* 26:2097-2108.
- Fay JC, Wu CI. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155:1405-1413.
- Fay JC, Wyckoff GJ, Wu CI. 2002. Testing the neutral theory of molecular evolution with genomic data from Drosophila. *Nature* 415:1024-1026.
- Fisher RA. 1930. The genetical theory of natural selection. Oxford,: The Clarendon press.
- Haig D, Hurst LD. 1999. A quantitative measure of error minimization in the genetic code. *J Mol Evol* 49:708.
- Hanada K, Shiu SH, Li WH. 2007. The nonsynonymous/synonymous substitution rate ratio versus the radical/conservative replacement rate ratio in the evolution of mammalian genes. *Mol Biol Evol* 24:2235-2241.
- Hartl DL, Clark AG. 1997. Principles of population genetics: Sinauer associates Sunderland.
- He Z, Chen Q, Yang H, Chen Q, Shi S, Wu C-I. 2018. Skepticism toward adaptive signals in DNA sequence comparisons-Is the neutral theory dead yet? bioRxiv: <https://doi.org/10.1101/417717>.
- He Z, Li X, Yang M, Wang X, Zhong C, Duke NC, Shi S, Wu C-I. 2018. Speciation with gene flow via cycles of isolation and migration: Insights from multiple mangrove taxa. *National Science Review*: <https://doi.org/10.1093/nsr/nwy1078>.
- Hughes AL, Ota T, Nei M. 1990. Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. *Mol Biol Evol* 7:515-524.
- Jordan IK, Rogozin IB, Wolf YI, Koonin EV. 2002. Microevolutionary Genomics of Bacteria. *Theoretical population biology* 61:435-447.
- Kimura M. 1983. The neutral theory of molecular evolution: Cambridge University Press.
- Lack JB, Cardeno CM, Crepeau MW, Taylor W, Corbett-Detig RB, Stevens KA, Langley CH, Pool JE. 2015. The Drosophila genome nexus: a population genomic resource of 623 Drosophila melanogaster genomes, including 197 from a single ancestral range population. *Genetics* 199:1229-1241.
- Li W. 1997. Molecular evolution: Sinauer Associates Incorporated.
- Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends Genet* 22:126-131.
- Luo H, Huang Y, Stepanauskas R, Tang J. 2017. Excess of non-conservative amino acid changes in marine bacterioplankton lineages with reduced genomes. *Nature Microbiology* 2:17091.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. *Nature* 351:652-654.
- Miyata T, Miyazawa S, Yasunaga T. 1979. Two types of amino acid substitutions in protein evolution. *J Mol Evol* 12:219-236.
- Ohta T, Gillespie JH. 1996. Development of neutral and nearly neutral theories. *Theoretical population biology* 49:128-142.

- Pool JE, Corbett-Detig RB, Sugino RP, Stevens KA, Cardeno CM, Crepeau MW, Duchon P, Emerson JJ, Saelao P, Begun DJ, et al. 2012. Population Genomics of sub-saharan *Drosophila melanogaster*: African diversity and non-African admixture. *PLoS Genet* 8:e1003080.
- Sawyer SA, Hartl DL. 1992. Population genetics of polymorphism and divergence. *Genetics* 132:1161-1176.
- Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, Turissini DA, Fang S, Wang HY, Hudson RR, Nielsen R, et al. 2007. Adaptive genic evolution in the *Drosophila* genomes. *Proc Natl Acad Sci U S A* 104:2271-2276.
- Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in *Drosophila*. *Nature* 415:1022-1024.
- Tang H, Wu CI. 2006. A new method for estimating nonsynonymous substitutions and its applications to detecting positive selection. *Mol Biol Evol* 23:372-379.
- Tang H, Wyckoff GJ, Lu J, Wu CI. 2004. A universal evolutionary index for amino acid changes. *Mol Biol Evol* 21:1548-1556.
- Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet* 12:204-213.
- Wang H-Y, Chen Y, Tong D, Ling S, Hu Z, Tao Y, Lu X, Wu C-I. 2017. Is the evolution in tumors Darwinian or non-Darwinian? *National Science Review* 5:15-17.
- Xu S, He Z, Zhang Z, Guo Z, Guo W, Lyu H, Li J, Yang M, Du Z, Huang Y, et al. 2017. The origin, diversification and adaptation of a major mangrove clade (*Rhizophoreae*) revealed by whole-genome sequencing. *National Science Review* 4:721-734.
- Zhang J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J Mol Evol* 50:56-68.

Figures and Tables

Table 1 – The MK test for each of the five gene groups as Ka/Ks increases

Species	Gene Group	Polymorphism		Divergence		A/S ratio		^a F _A
		A	S	A	S	Polym	Div	
<i>Drosophila</i> (<i>D. melanogaster</i> vs. <i>D. simulans</i>)	0-20%	5,659	57,094	20,206	133,647	0.099	0.151	1.525**
	20-40%	4,711	18,577	20,553	52,675	0.254	0.390	1.539**
	40-60%	4,234	10,309	21,960	32,818	0.411	0.669	1.629**
	60-80%	3,421	55,63	22,290	21,032	0.615	1.060	1.723**
	80-100%	2,449	2,492	23,126	11,090	0.983	2.085	2.122**
	ALL	20,474	94,035	108,135	251,262	0.218	0.430	1.977**
Hominoids (human vs. chimpanzee)	0-20%	780	1,708	5,783	19,412	0.457	0.298	0.652
	20-40%	557	576	5,810	6,368	0.967	0.912	0.943
	40-60%	483	414	5,884	4,199	1.167	1.401	1.201*
	60-80%	488	282	5,781	2,702	1.730	2.140	1.236*
	80-100%	372	170	6,079	1,116	2.188	5.447	2.489**
	ALL	2,680	3,150	29,337	33,797	0.851	0.868	1.020

Genes are divided into five groups in the ascending order of their Ka/Ks ratio. There is approximately the same number of non-synonymous changes in each group. The allele frequency cutoffs are 0.15-0.85 for *Drosophila* and 0.2-0.8 for Hominoids. In the supplementary Table S1, different cutoffs are used.

Polym – polymorphism; Div - divergence.

^a one tailed Fisher's exact test. * p-value < 0.05, ** p-value < 0.001. Here, we use Div(A/S) / Poly(A/S) for F_A in lieu of Eq. (1.1) since the two expressions are nearly identical.

Table 2 – Ten AA substitutions with the highest Fi in *Drosophila* and their ranks in hominoids

AA1	AA2	Δ_U (rank)	<i>Drosophila</i>				Hominoids	
			Polym	Div	Fi	Fi rank	Fi	Fi rank
Ser	Trp	0.940 (10)	9	94	3.909	1	1.678	10
Arg	Ile	0.889 (17)	17	161	3.544	2	1.150	29
Lys	Ile	0.801 (34)	37	334	3.378	3	Inf	1
Arg	Cys	0.937 (12)	38	336	3.309	4	2.190	6
Ser	Ile	0.867 (22)	72	588	3.056	5	1.270	19
Tyr	Phe	0.374 (68)	212	1727	3.049	6	1.161	28
Gly	Cys	0.923 (14)	71	550	2.899	7	0.749	70
Cys	Phe	0.964 (7)	40	308	2.882	8	2.563	5
Lys	Met	0.859 (29)	72	532	2.765	9	3.821	2
Arg	Pro	0.859 (28)	58	426	2.749	10	1.181	27
Synonymous changes			94035	251262				

The 10 pairs with the highest Fi values in *Drosophila* are given. Five of the 10 pairs are also in the top 10 in hominoids, shown in boldface (the expected overlap is 1.33; P = 0.0027). Fi = Obs[Ki/Ks]/Obs[Pi/Ps]. The complete table is given in Dataset S1. In the supplementary Table S2, different cutoffs are used.

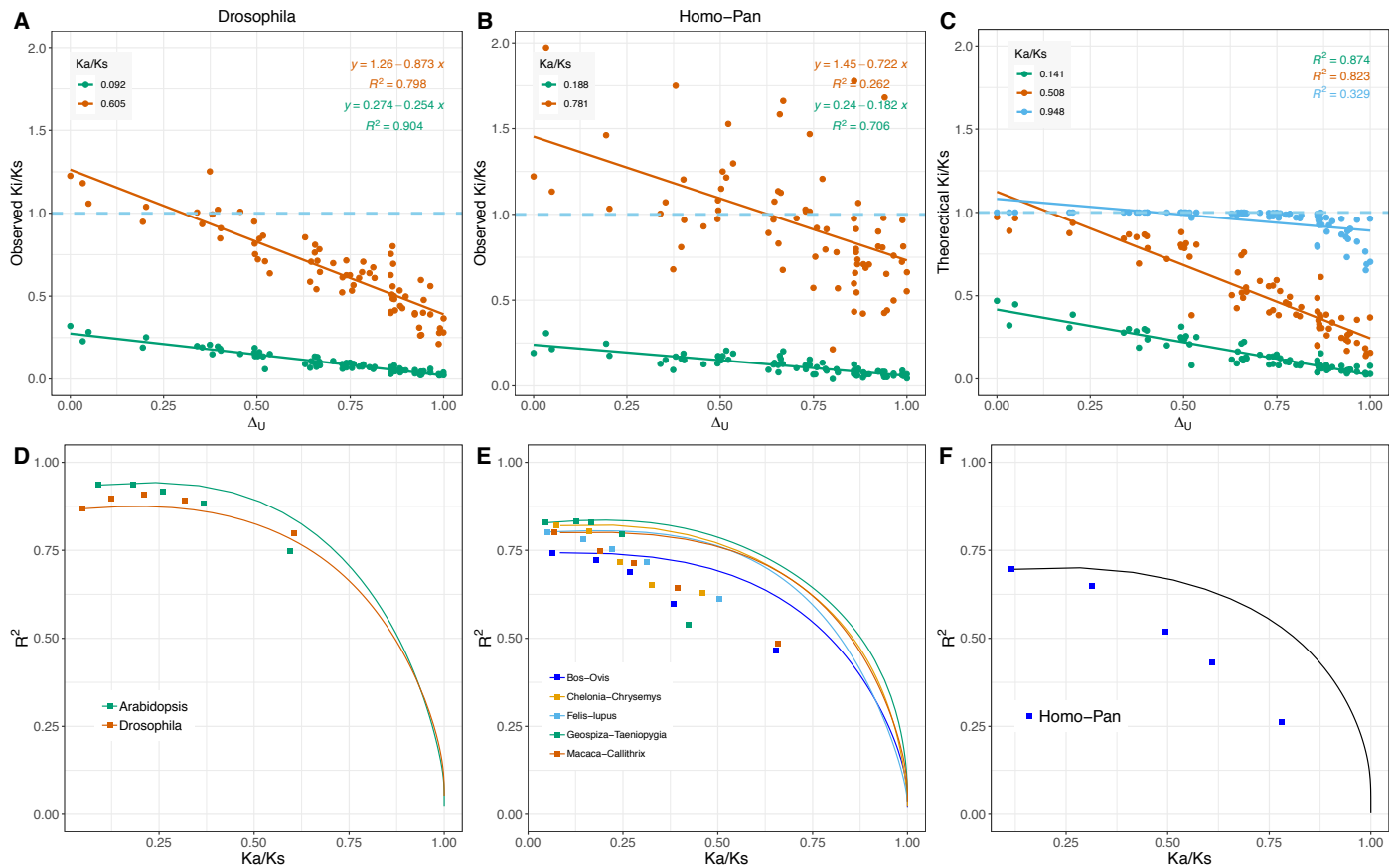


Figure 1 | Ki/Ks vs. Δ_U and their correlations. (A-B) The relationship between Δ_U and Ki/Ks from rapidly and slowly evolving genes for *Drosophila* and *homo-pan*. Here, genes are first divided into five groups in the ascending order of their Ka/Ks ratios with equal number of non-synonymous changes in each group. Rapidly evolving genes account for the top 20% of non-synonymous changes (red points). Slowly evolving genes account for bottom 0%-60% (green points). The Ka/Ks rate ratios in the legend are the average Ka/Ks ratio for each gene group. (C) The theoretical changes of Ki/Ks and Δ_U under a model with negative selection only. As the Ka/Ks ratio increases, some AA pairs ascend to their upper limit, with Ki/Ks values nearly 1 (blue points). (D-F) The theoretical (solid lines) and observed changes (box-shape points) of total Ka/Ks ratio and R^2 values of their Ki against Δ_U . Genes were again divided into five groups as above. The theoretical changes are based on a model with varying levels of negative selection only.

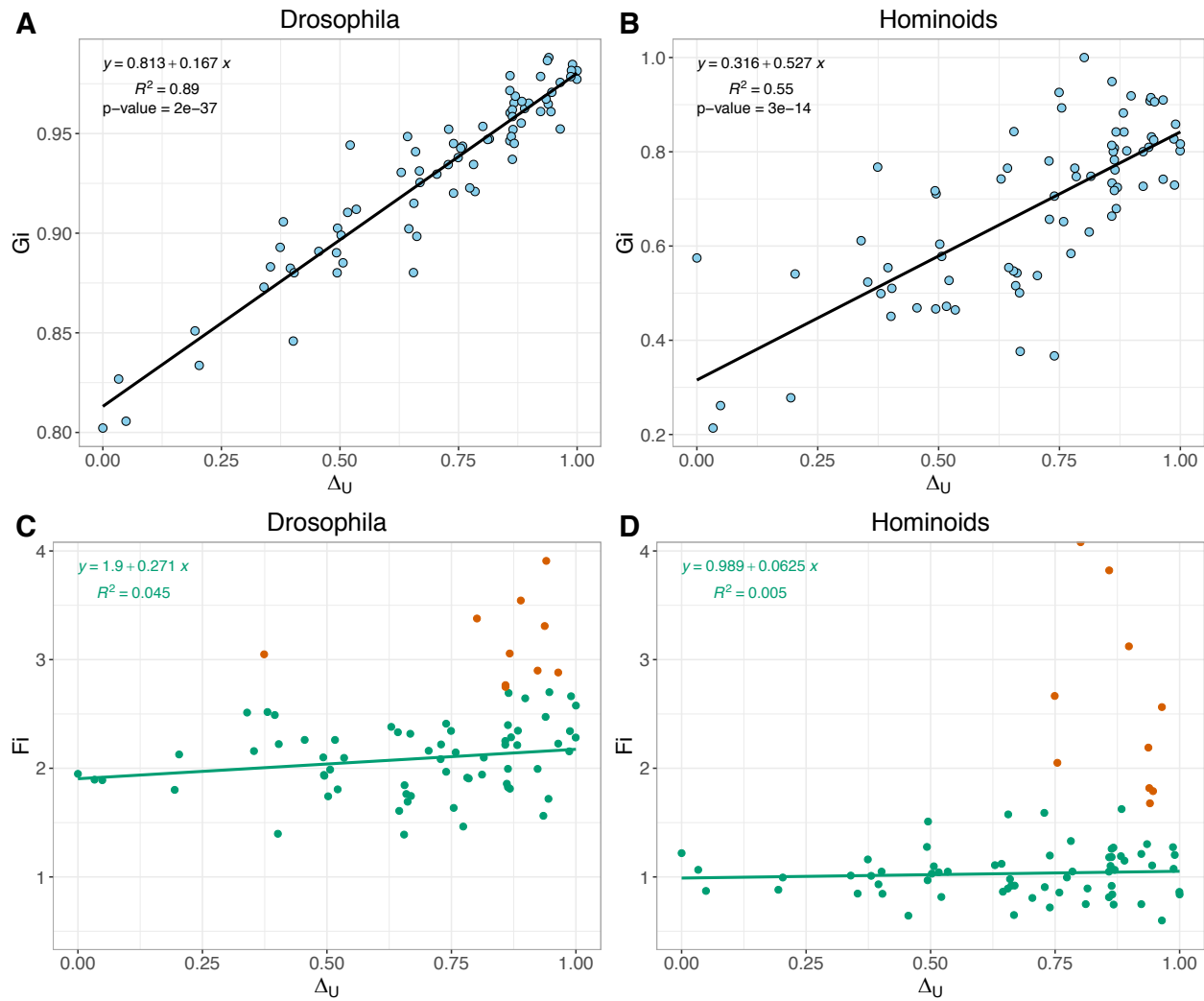


Figure 2 | The relationship between evolutionary AA distance (Δ_U) and selection intensity. (A-B) The intensity of negative selection in *Drosophila* and Hominoids. **(C-D)** positive selection intensity in *Drosophila* and Hominoids. The highest 10 F_i values are labeled by red, and the rests are green. The allele frequency cutoffs are 0.15-0.85 for *Drosophila* and 0.2-0.8 for Hominoids. In the supplementary Figure S1, different cutoffs are applied.

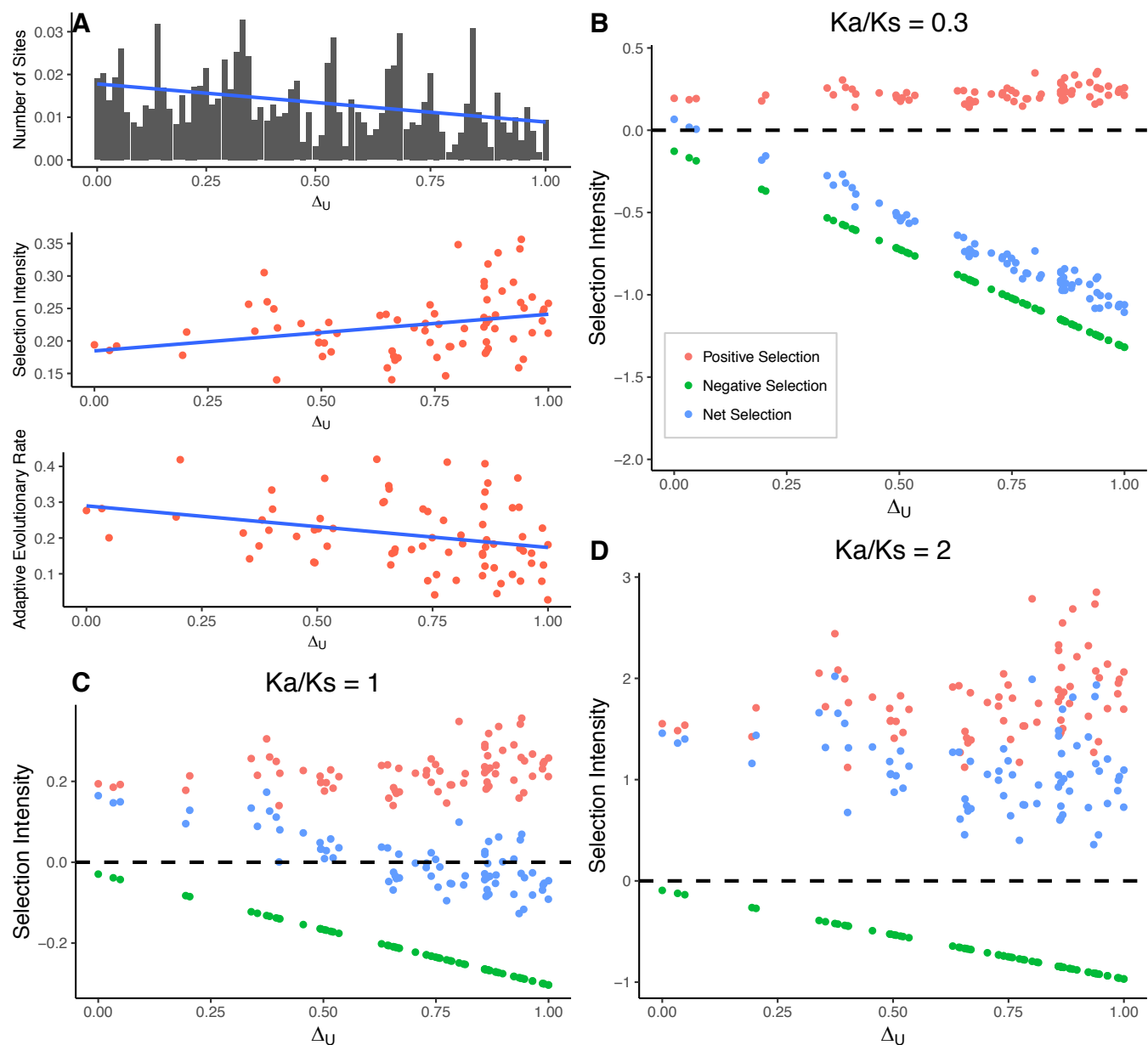


Figure 3 | A model of selection intensity in relation to functional differences (Δ_U). (A) A model with positive selection. The relative frequencies of available number of sites for each Δ_U (w_i) are given in the top panel. The strength of positive selection (A_i) relative to Δ_U is based on *Drosophila* data in Figure 2 (middle panel), but the adaptive evolutionary rate (the bottom panel) still depends on small-step changes due to the difference of available number of sites (w_i) among AAs (top panel). The blue lines are the linear regression lines. (B-D) A model of positive selection, negative selection, and their confounding effect. There are three Ka/Ks ratios, where $Ka/Ks = 0.3$ (Panel B), $Ka/Ks = 1$ (Panel C), and $Ka/Ks = 2$ (Panel D). The dashed line denotes complete neutrality, with $Ka/Ks = 1$. Above the dashed neutral line is the strength of positive selection (A_i) (red points) and below the line is negative selection (B_i , green points). The strength of negative selection (B_i) changes linearly with functional difference (Δ_U). The strength of positive selection (A_i) shown in Panel A is also shared between Panel B-C, but the relative strength is increased in Panel D to achieve a higher Ka/Ks ratio. The changes of Ki/Ks relative to Δ_U are the result of a confounding effect between negative and positive selection ($R_i = 1 + A_i - B_i$) (blue points).