doi:10.1111/j.1558-5646.2007.00308.x

TOWARD A SELECTION THEORY OF MOLECULAR EVOLUTION

Matthew W. Hahn

Department of Biology and School of Informatics, 1001 E. 3rd Street, Indiana University, Bloomington, IN 47405

E-mail: mwh@indiana.edu

Received November 27, 2007

Accepted November 28, 2007

KEY WORDS: Adaptation, neutral theory, population genetics, positive selection.

> "When the facts change, I change my mind. What do you do sir?"
>
> —John Maynard Keynes

The Neutral Theory

Models describing the dynamics of genetic variants with no effect on fitness—so-called neutral models—have been around almost as long as the field of population genetics (Fisher 1922; Wright 1931). Decades after the first models were introduced Motoo Kimura gave a complete description of the dynamics of neutral mutations in finite populations, using mathematical tools borrowed from particle physics (Kimura 1955). Although the elegance of this and other results from Kimura and colleagues were uncontested, their applicability to data seemed remote until experiments revealed enormous amounts of molecular genetic variation, both within and between species (Zuckerkandl and Pauling 1965; Harris 1966; Lewontin and Hubby 1966). The observed levels of variation appeared inconsistent with models that proposed selective effects for all or most mutations, and what has become known as the Neutral Theory of Molecular Evolution was born (Kimura 1968; King and Jukes 1969; Kimura and Ohta 1971).

Despite contentious argument over the validity of the Neutral Theory (Kimura 1983; Gillespie 1991), it has become the predominant framework for research in population genetics and molecular evolution for almost 40 years. Increasingly complex models describing the expected patterns of variation within and between species allow researchers to ask about the evolutionary processes acting in nature, both at single loci and in increasingly large datasets encompassing all or most genes in a genome. The Neutral Theory provides a theoretical basis for understanding DNA variation with clear, testable hypotheses and an array of statistical tools that distinguish natural selection from random genetic drift (Kreitman 2000; Nielsen 2001; Hahn 2007).

However, the recent paper by Begun and colleagues (Begun et al. 2007) should finally begin to change people's view of this scientific paradigm. Although results inconsistent with the Neutral Theory have been mounting for some time (see below), the field has continued to use it as a foundation for understanding the molecular world. As the first true "population genomic" dataset, the results of Begun et al. force us to see that the central predictions of the Neutral Theory do not hold in natural populations. Far from just the half-caught glimpses of nonneutral evolution afforded by studies of limited numbers of loci, by sequencing the whole genomes of multiple lines of *Drosophila simulans* this work should cause a major shift in how we interpret DNA variation within populations and among species. As the conclusions of this article are appropriately cautious with respect to the implications of the work, I will use this essay to provide a wider view of the importance of these results and a synthesis with previous results. To do this I will address the two major tenets of the Neutral Theory, and how increasing amounts of data are showing that these claims and their attendant predictions do not hold for the vast majority of genes and species. I also argue that the implications of our continued use of neutral models are dire—at

least if we hope to truly understand the evolutionary forces that shape genomes—and require the development of a new Selection Theory of molecular evolution.

THE DIRECT SELECTION CLAIM

The Direct Selection claim of the Neutral Theory is that the vast majority of polymorphisms within species and fixed differences between species have no effect on fitness—that is, that there is no direct selection on them, and that they are neutral. This does not mean that all possible mutations are neutral, only that the observed mutations are neutral. Strongly deleterious variants are rarely segregating in populations and have an even smaller chance of being fixed, and so are not observed; likewise, adaptive mutations might make up a small fraction of differences between species but are fixed so quickly that they are not sampled when polymorphic. (Ohta and Kimura [1971] put an upper limit of ~8% on the proportion of advantageous substitutions consistent with the Neutral Theory.) That synonymous changes in coding sequences or changes in nonfunctioning noncoding sequences are neutral is a trivial extension of this claim (but see Akashi 1994; Resch et al. 2007), and most disagreements revolve around the neutrality of mutations that have the capacity to change the phenotype.

The Direct Selection claim of the Neutral Theory is also one of the most misunderstood ideas in molecular evolution, as the term "neutral" is often conflated with "unconstrained" (see Kimura 1983, chapter 3). The Neutral Theory is perfectly compatible with strong selective constraint on a sequence as long as all of the observed changes—no matter how few—are neutral. DNA sequences are either constrained or unconstrained, but this distinction does not tell us whether they are evolving neutrally or not (although all mutations in unconstrained sequences are assumed to be neutral). To conclude that a sequence is evolving "nonneutrally" in this context means either that segregating polymorphisms are not neutral with respect to fitness, or that fixed differences were not simply fixed by drift but by adaptive natural selection.

Regardless, the Direct Selection claim provides the fundamental foundation for many current tests of selection, by providing the null (neutral) hypothesis against which instances of selection can be evaluated. Unfortunately, it is now clear that data collected from a large number of loci are inconsistent with this claim. The most direct tests of this claim are carried out by comparing the number and frequency of functionally relevant mutations (either coding or regulatory) to those without an effect on function (usually synonymous or intronic). Polymorphism and divergence data from these classes of sequences can be combined in the McDonald–Kreitman (MK) test (McDonald and Kreitman 1991). The prediction of the neutral model is that the ratio of functional:nonfunctional polymorphisms (e.g., nonsynonymous:synonymous) will be equal to the same ratio among fixed differences. An excess of nonsynonymous or regulatory fixed differences relative to synonymous changes is interpreted as evidence for adaptive substitutions, whereas an excess of nonsynonymous or regulatory polymorphisms is interpreted as evidence that segregating variation is either mildly deleterious or under strong balancing selection (McDonald and Kreitman 1991; Weinreich and Rand 2000).

Previous studies using the MK test on hundreds of genes in *Drosophila melanogaster* and *D. simulans* have concluded that anywhere between 30% and 94% of all amino acid substitutions were fixed by adaptive natural selection (Fay et al. 2002; Smith and Eyre-Walker 2002; Sawyer et al. 2003; Bierne and Eyre-Walker 2004; Shapiro et al. 2007). Similar studies on the untranslated and regulatory regions upstream of *Drosophila* genes revealed similar estimates of adaptive divergence (Kohn et al. 2004; Andolfatto 2005), as have studies of protein-coding genes in *Escherichia coli* (Charlesworth and Eyre-Walker 2006). In contrast, studies of both humans and *Arabidopsis thaliana* have revealed an excess of nonsynonymous polymorphisms consistent with the nonneutral evolution of segregating variation (Bustamante et al. 2002, 2005). As both humans and *A. thaliana* have much smaller population sizes than either *Drosophila* or *E. coli*, it is believed that the observed excess of variation is largely due to mildly deleterious mutations that are not purged from small populations but do not contribute to interspecific divergence. Alternative methods that compare the allele frequencies of nonsynonymous polymorphisms and synonymous polymorphisms are also consistent with a large amount of segregating deleterious polymorphism in both humans and *Arabidopsis* (Williamson et al. 2005; Eyre-Walker et al. 2006; Kim et al. 2007). It should also be noted that even though the MK test is relatively powerful compared to other tests of molecular evolution, it is still very conservative and likely misses a large proportion of nonneutral evolution (Akashi 1999); it is likely to be even more conservative if many nonsynonymous polymorphisms are actually advantageous mutations and on their way to fixation.

The genome-wide data presented by Begun et al. (2007) confirm these previous patterns and reveal a new pattern using the MK test. As with previous studies in *Drosophila*, Begun et al. find a large number of substitutions with evidence for adaptive evolution. Approximately 54% of all nonsynonymous fixed differences, ~35% of fixations in untranslated regions, and ~7% of intergenic fixations have been fixed by positive selection along the lineage leading to *D. simulans* (all of these estimates are significantly different than zero). Curiously, Begun et al. also find a significant excess of segregating polymorphisms, but only for mutations that do not lie in coding regions. Taken together, the results of Begun et al. and other large studies provide overwhelming evidence for direct selection on both coding and regulatory mutations, manifest as excess polymorphism and divergence in different species, and are associated with different types of mutations. Neutral substitutions

clearly still occur, but given our current power to detect selection it is likely that they represent a minority of all changes. Although large datasets of this kind come from only a handful of model organisms, one must conclude that the preponderance of evidence to date does not support the Direct Selection claim of the Neutral Theory.

THE LINKED SELECTION CLAIM

The Linked Selection claim of the Neutral Theory is that linked selection does not affect a vast majority of loci, and therefore that variation in nature reflects the predictions of neutral models. Population genetics theory provides quantitative expectations of the level and frequencies of polymorphisms under neutrality. However, if selection acts on even a small fraction of mutations, patterns of variation at linked loci will be affected as neutral polymorphisms are dragged along with selected ones. The effect of selection on mutations linked to neutral polymorphisms, whether advantageous ("hitchhiking"; Maynard Smith and Haigh 1974) or deleterious ("background selection"; Charlesworth et al. 1993), is to lower the level of variation and skew the frequency spectrum of mutations relative to neutral expectations. An alternative way to state this claim is that most loci are expected to be at mutation–drift equilibrium—where the number of mutations entering a population is equal to the number being lost due to drift—although recent changes in population size may disturb this equilibrium.

To better understand the Linked Selection claim and its consequences, we need to understand two simple neutral expectations concerning the amount of variation within a species and the magnitude of divergence between species. At mutation–drift equilibrium the average number of polymorphic sites between two sequences within the same species, denoted π, is determined by the expression $4N\mu_0$, where N is the population size and μ_0 is the neutral mutation rate (Tajima 1983). This means that as either the population size or the mutation rate goes up, so does the amount of variation within a species, and vice versa. Likewise, when all mutations are neutral the average number of nucleotide differences measured between two sequences from different species, denoted d, is determined by $2T\mu_0$, where T is the time back to the ancestor of the two sequences and mutations can occur on both lineages. Note that d is unaffected by population size (Kimura 1968) and is determined solely by the amount of time that has elapsed and the mutation rate, which also affects levels of polymorphism. In addition, it has been shown theoretically that linked selection does not affect the level of neutral divergence (Birky and Walsh 1988), although the selected mutation itself obviously contributes to the number of differences between species. Because the number of differences between single representative sequences from two different species includes both the fixed substitutions that have accumulated and the derived polymorphisms present in those individual samples, it is common to use a corrected distance that represents only fixed divergence: $d^* = d - \pi$ (Nei 1987). The simple relationships presented in this paragraph represent the most basic expectations of the Neutral Theory, and provide us with testable hypotheses that can be applied to population genetic datasets. Below I discuss four empirical observations that are fundamentally inconsistent with neutral predictions.

The paradox of variation

One of the first challenges to the Neutral Theory was leveled against the initial prediction outlined above that π and other measures of genetic diversity should be linearly proportional to population size. The observation that diversity was only weakly correlated with apparent population size was called the "paradox of variation" (Lewontin 1974) and still remains. Measurements of nucleotide variation from hundreds of species across the tree of life continue to show that even though population sizes vary across many orders of magnitude (from ubiquitous bacteria to exceedingly rare vertebrates), the mean difference in nucleotide diversity between prokaryotes and vertebrates only spans two orders of magnitude (Lynch 2006). Among eukaryotes, levels of variation in mitochondrial DNA show no correlation with population size (Bazin et al. 2006) and there is only a weak relationship between nuclear genes and population size (Lynch 2006), even though the Neutral Theory clearly predicts a linear relationship. The two most complete models of linked selection either predict no relationship between diversity and population size (hitchhiking; Gillespie 2001) or a weakly positive relationship (background selection; Charlesworth et al. 1993). It is clear from the data collected thus far that comparisons of levels of diversity across species are not consistent with the Neutral Theory.

Negative relationship between polymorphism and divergence

A second major prediction of the Linked Selection claim is that there should be a strong positive correlation between polymorphism and divergence. These two measures should be tightly correlated under neutrality because levels of both polymorphism and divergence are the result of the magnitude of the neutral mutation rate. However, comparing polymorphism within *D. simulans* (π) to divergence between *D. simulans* and *D. melanogaster* (d^*) reveals that there is actually a negative correlation between the two across the genome (average across all chromosome arms: $r = -0.39$). To control for possible differences in neutral mutation rate between *D. simulans* and *D. melanogaster*, we can also compare polymorphism to the divergence that has occurred only on the *D. simulans* lineage by using *D. yakuba* to polarize changes. This comparison also shows a negative correlation between π and d^* ($r = -0.17$), completely inconsistent with the predictions of the Neutral Theory. Coalescent simulations show that negative correlations between polymorphism and divergence are not expected under neutral conditions (N. Nista and M. W. Hahn, unpubl. data).

If neutral models do not predict the observed patterns, it is worthwhile asking whether there are selection models that do. The answer appears to be that a range of selection models do predict the negative correlation, as the following example shows. If mutations are largely advantageous (i.e., the hitchhiking model), then we would expect that higher fixation rates of adaptive mutations would lead to decreased levels of polymorphism as the effect of hitchhiking becomes more pronounced. Indeed, a selection–mutation–drift equilibrium expectation for the amount of variation at a neutral locus linked to a constant influx of adaptive mutations is $\pi = 4N\mu_0/(1 + 2N\rho y^2)$, where ρ is the rate of fixation of advantageous mutations and y^2 is proportional to the effect of linkage (Gillespie 2000). As we can see, increasing the rate of adaptive change results in decreasing levels of variation at linked loci. (Because ρ is actually a function of N, π does not scale linearly with population size in this model [Gillespie 2001].) Likewise, the amount of divergence expected with adaptive mutations is given by $d = 2T \times 4Ns\mu_T f_A$, where s is the selective advantage of the mutant allele, μ_T is the total mutation rate, and f_A is the fraction of all mutations that are advantageous (Kimura 1983). In this case, divergence increases as the rate and strength of adaptive evolution increases, which results in a net negative correlation between polymorphism and divergence. Consistent with this view, Begun et al. found that genes in *D. simulans* that had experienced the greatest increase in the rate of nonsynonymous substitution also showed dramatically decreased levels of polymorphism.

Positive relationship between polymorphism and recombination

One of the most striking results to challenge the Neutral Theory was the discovery of a correlation where none was expected. Under neutrality, no relationship between levels of polymorphism and recombination is expected, as the number and frequency of neutral mutations should be unaffected by the recombinational environment (Hudson 1983). However, low levels of polymorphism in several *D. melanogaster* genes located in regions of low crossing-over prompted Begun and Aquadro (1992) to demonstrate that in fact there was a positive correlation between polymorphism and recombination across a range of recombination rates. This result now appears to be one of the most universal patterns of population genetics, with similar relationships found in every species examined: human (Nachman et al. 1998; Przeworski et al. 2000), mouse (Nachman 1997; Takahashi et al. 2004), *C. elegans* (Cutter and Payseur 2003), mosquito (Stump et al. 2005; Slotman et al. 2006), *A. thaliana* (Kim et al. 2007), tomato (Stephan and Langley 1998; Roselius et al. 2005), sea beet (Kraft et al. 1998), maize (Tenaillon et al. 2001), and goatgrass (Dvorak et al. 1998). Begun et al. provide one of the most complete studies to date, and show that there is a very strong correlation between recombination and π on the *D. simulans* X chromosome ($r = 0.45$; Fig. 1). As the article points out, this result is especially striking given that the estimates of crossing-over come from *D. melanogaster*; nonetheless, the fact that such a strong relationship is observed argues that recombination rates between the two species are quite conserved.

There are two clear hypotheses—one neutral and one selective—for why recombination and polymorphism should be correlated. If recombination itself is mutagenic then regions of high recombination will have more mutations, resulting in higher levels of polymorphism. This neutral model also therefore predicts that divergence (d^*) should be positively correlated with recombination if this process is mutagenic (Begun and Aquadro 1992), which is not found in the *D. simulans* data ($r = 0.03$).

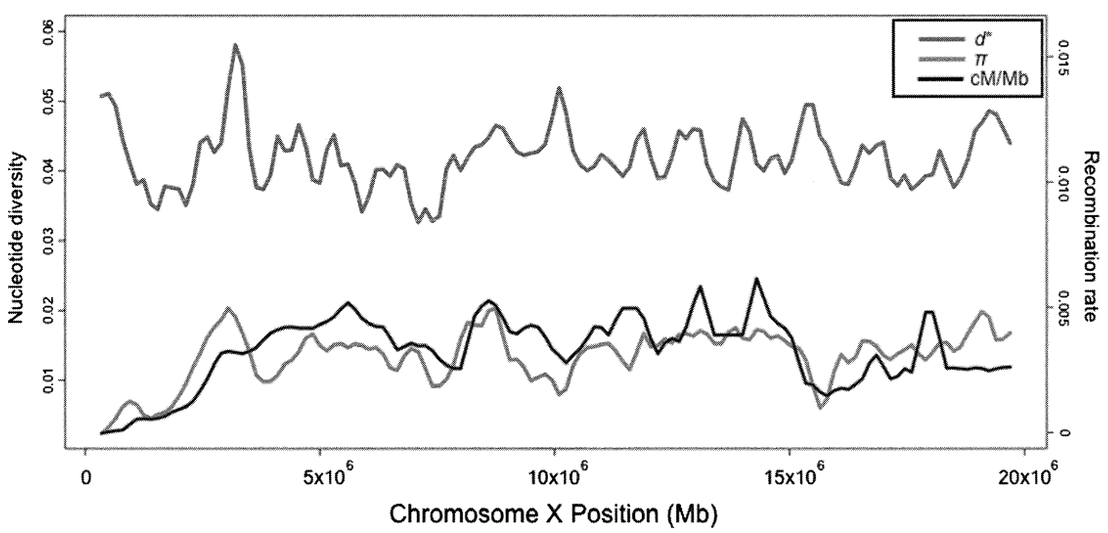

Figure 1. Polymorphism, divergence, and recombination across the *D. simulans* X chromosome. Measures of polymorphism (π), divergence (d^*), and recombination (measured in centiMorgans per megabase [cM/MB]) are shown (all from Begun et al. 2007). The raw data have been Loess smoothed for ease of visualization.

However, this effect may explain some of the variation in humans (Lercher and Hurst 2002; Hellmann et al. 2003; Hellmann et al. 2005). The alternative hypothesis is that some form of linked selection is acting across the *D. simulans* genome such that loci in regions of higher recombination are more likely to escape the effects of nearby selection, whether advantageous or deleterious (Aquadro et al. 1994). If, for instance, adaptive fixations are occurring continuously across the genome then levels of polymorphism are reduced by an amount proportional to the strength of selection and the recombination rate, although divergence levels are unaffected (Birky and Walsh 1988). As shown above the equilibrium level of variation under a hitchhiking model is $\pi = 4N\mu_0/(1 + 2N\rho y^2)$, which is expected to show a positive correlation with the rate of recombination (because y^2 is inversely proportional to recombination). Alternative models of linked selection also predict this positive relationship (Wiehe and Stephan 1993; Gillespie 1994; Hudson and Kaplan 1995; Payseur and Nachman 2002). In addition, the correlation between recombination and the allele frequency of mutations found in both flies (Andolfatto and Przeworski 2001) and humans (Stajich and Hahn 2005) can only be explained by linked selection.

The positive correlation between polymorphism and recombination across many species is astounding for a number of reasons. First, these results imply that almost no loci are free from the effects of linked selection. Even after trimming 2.5 megabases from the proximal and distal ends of the X chromosome—where recombination is the lowest—there is a significant correlation between π and recombination in *D. simulans* ($r = 0.26$). Far from being limited to only the regions of lowest recombination these patterns suggest that all loci but those with the highest rates of recombination are affected, and even these loci may simply show the least effects of linked selection. Second, in the absence of other forces the reduction in variation caused by linked selection will rebound to neutral-equilibrium levels relatively rapidly (Simonsen et al. 1995). The fact that polymorphism is correlated with recombination implies that in every species examined at almost every locus there has been a recent selective event (whether advantageous or deleterious), such that levels of polymorphism are not at equilibrium. These data are therefore fundamentally incompatible with the expectations of the Neutral Theory.

Regional similarity in levels of polymorphism

A novel example of the effects of linked selection is also afforded by the results of Begun and colleagues. Previous studies have shown that levels of divergence among mammals are locally similar at a megabase scale (Lercher et al. 2001; Webster et al. 2004; Gaffney and Keightley 2005; Hellmann et al. 2005). Begun et al. show that the same is true for *Drosophila*, with significant correlations in levels of divergence at loci hundreds of kilobases apart (see also Hahn 2006). If this regional similarity in divergence is simply due to differences in neutral mutation rates, then levels of polymorphism in the *D. simulans* genome should also show such a pattern. Begun et al. do find regional similarities in polymorphism, although they seem to extend almost three times as far as similarities in divergence. This same difference was found when comparing regional similarity in polymorphism and divergence in humans (Smith and Lercher 2002).

Because levels of polymorphism can be locally similar due to shared genealogical histories or population bottlenecks—processes that do not affect levels of divergence—it is possible that these neutral processes have caused similarities in polymorphism to stretch over greater distances than in divergence. However, simulations show that neither linkage nor demographic history (nor both together) is sufficient to explain the observed patterns (P. Nista and M.W. Hahn unpubl. data). Instead, it appears that again some form of linked selection must be invoked. One possibility is that hitchhiking events create "islands" of relatively homogeneous levels of polymorphism extending over very long distances, and that the action of many such events along a single chromosome are enough to create the observed patterns. A second possibility, more in keeping with the apparent relationship between recombination and polymorphism, is that selection and recombination interact to cause regional similarities in polymorphism. This pattern would arise because recombination rates themselves are regionally similar and are correlated over long distances (Kong et al. 2002). If the effects of selection on linked variation are limited by recombination rate, then a uniform distribution of selective events would lead to levels of polymorphism that are similar over the same scale as recombination. Whichever scenario turns out to be the correct explanation, the results of Begun et al. are clearly not consistent with neutrality.

Implications

Since the proposal of the Neutral Theory, every few years has seen the publication of a paper summarizing data that challenge its preeminence (e.g., Gillespie 1984; Kreitman 1996; Fay and Wu 2001). However—and despite the mounting evidence of natural selection that each successive author has been able to draw upon—the general conclusion has always been that even though we do not necessarily believe the Neutral Theory, neutral models are easier to parameterize and provide a clear null model. The title of Kreitman's 1996 review sums up this feeling: "The neutral theory is dead. Long live the neutral theory." In the following I outline some of the reasons why neutral models have had such staying power, but why continued use of the Neutral Theory as a guiding framework can positively mislead researchers and skew our understanding of nature.

To see the advantage of using neutral models, consider a situation in which researchers eventually come to agree that 50%

of all nonsynonymous differences in *D. melanogaster* were fixed by adaptive evolution. With this proportion as a starting point we must now specify a large number of additional parameters and distributions to model genetic variation. For example, we must now answer the following questions: What is the distribution of selection coefficients of these nonsynonymous differences? What are their dominance coefficients? Has selection acted on standing polymorphism or only on newly arising mutations? Is there epistasis among mutations, and if so, what form does it take? What does the fine-scale recombination map look like across the genome? And of course all of these need to be specified again for each new species considered. None of this information is needed if all mutations are neutral; therefore simple scientific expediency has tended to win out.

Some of the reluctance to move away from neutral models is also likely to be a continued reaction to rampant pan-selectionism and adaptationist storytelling (cf. Gould and Lewontin 1979). It is certainly true that a mature field of evolutionary biology needs to consider both adaptationist and nonadaptationist explanations for natural phenomena, and it can be forcefully argued that nonadaptationist hypotheses are in general more parsimonious given equal evidence (Lynch 2007). However, the overwhelming evidence from studies of molecular variation does not support the Neutral Theory, and therefore neutral explanations are arguably not more parsimonious given all of the evidence. The consequence of this is that we have tied ourselves into philosophical knots by using null models no one believes but are easily parameterized. Below I describe one widespread example of the biased view forced on us by neutral models and how a move away from such models may help us to better understand molecular variation.

Concomitant with the development of the Neutral Theory, Cavalli-Sforza (1966) and Lewontin and Krakauer (1973) proposed what was to become one of the main axioms of modern molecular population genetics: "While natural selection will operate differently for each locus and each allele at a locus, the effect of breeding structure is uniform over all loci and all alleles" (Lewontin and Krakauer 1973 [italics in original]). Put another way, this says that the action of natural selection affects only a small region of the genome whereas demographic history affects the whole genome. Many authors have used this logic to try to disentangle the effects of selection and demography in large datasets (reviewed in Thornton et al. 2007). The common mode of inference employed by these studies assumes that the majority of genes provide information about the demographic history of a population, whereas the genes in the tails of the distribution for some statistic (e.g., π or Tajima's D) are the most likely to be under selection (see Fig. 2A). Because all distributions have tails, a number of refinements to these methods have been made to provide statistical evidence of selection—by simulating either a wide range of possible demographic histories (Akey et al. 2004), realistic demographic models known from other data (Stajich and Hahn 2005), or demographic models estimated from the "background" data of the dataset under study (Nielsen et al. 2005).

Although the Cavalli-Sforza/Lewontin/Krakauer axiom is obviously true for any single gene under selection, it implicitly assumes that most genes are unaffected by natural selection. However, if most loci are affected by linked selection, then patterns of variation at these loci are influenced by a combination of demography and selection (Fig. 2B). Consider data collected from a population at demographic equilibrium but with high rates of adaptive natural selection: because hitchhiking results in an excess of low-frequency mutations, standard approaches would lead us to infer that a population expansion or bottleneck had occurred as most loci show this excess. To find the targets of selection we would then "recenter" the distribution of test statistics by simulating neutral data under the inferred demographic history, whether or not there is any independent (nonmolecular) data about the validity of such a history. As might be imagined, this procedure will cause us to miss many or most of the genes undergoing adaptive natural selection, or even to reject those with strong evidence simply because they are in the middle of the distribution. It can therefore be considered an extremely low-powered method for detecting selection. For example, Mekel-Brobov et al. (2005) found multiple signatures of natural selection on the *ASPM* gene in non-African humans. Although there has been considerable debate over the effect of this gene on intelligence and brain size among humans, the more relevant responses to this article have focused on showing that the patterns of variation seen are not due to selection. These responses have generally followed the Cavalli-Sforza/Lewontin/Krakauer axiom, in two different ways: (1) there is a demographic model that can explain the data (Currat et al. 2006), or (2) the gene is uninteresting because it looks like many other genes in the human genome (Yu et al. 2007). Regardless of the validity of any of these criticisms (see Mekel-Bobrov et al. 2006 and Mekel-Bobrov and Lahn 2007 for the authors' responses), it is clear that they follow the logic of an idealized neutral world.

The above discussion begs the question of whether the true demographic history of a population can ever be inferred from molecular data given the pervasive effects of linked selection. It is clear that the common use in phylogeographic studies of "neutral" markers such as the mitochondrial D-loop or microsatellites are not neutral if they are either linked to loci under selection (Bazin et al. 2006) or are subject to direct selection themselves (Rockman and Wray 2002). Even when studies have attempted to infer population histories from loci that do not overlap coding regions—so as to minimize the effect of linked selection—there can still be strong effects of linked selection. From the *D. simulans* data presented above it is clear that using noncoding loci

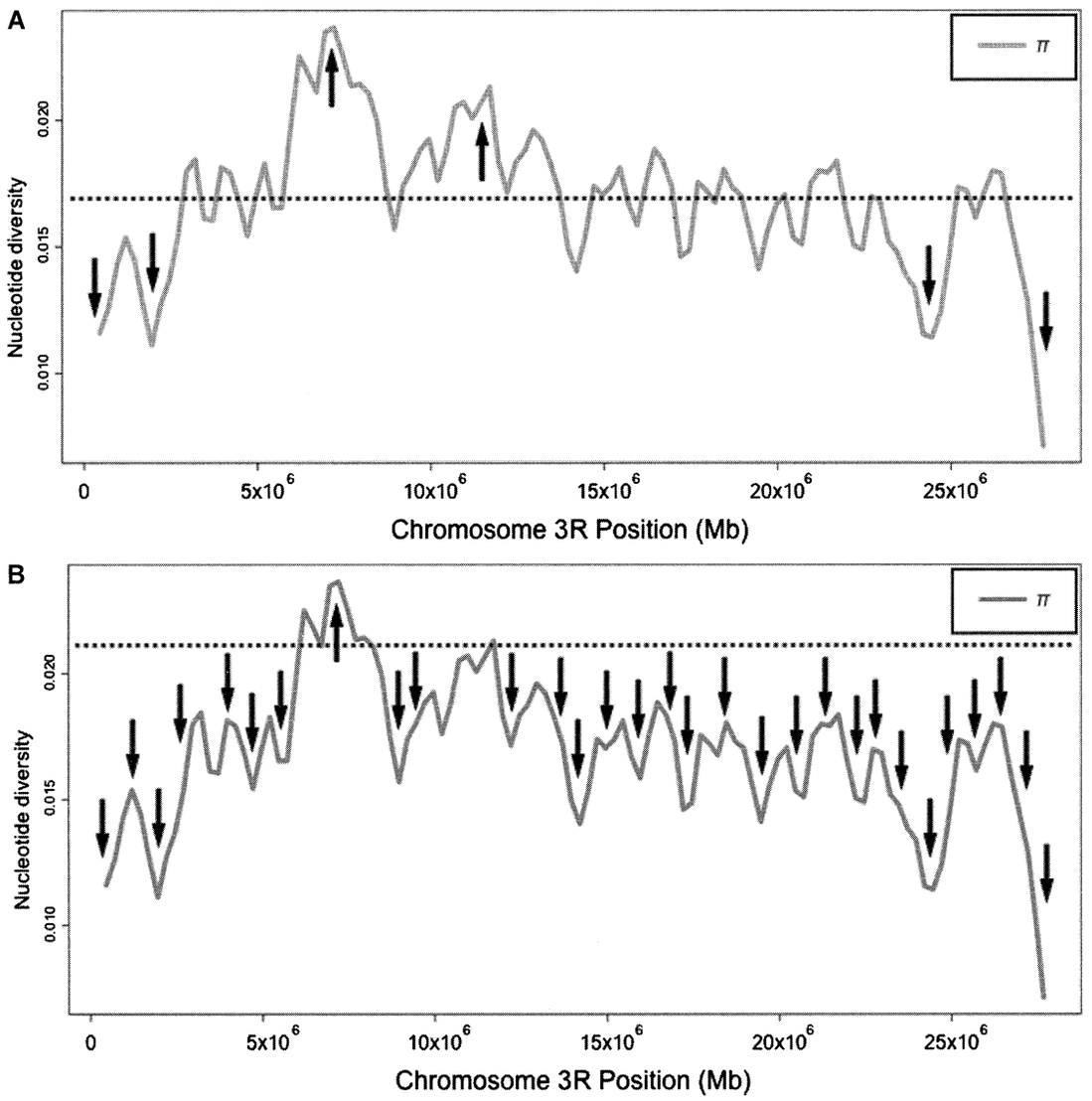

Figure 2. (A) The neutralist interpretation, and (B) a selectionist interpretation of variation in levels of polymorphism. The red line in both shows measured values of polymorphism (π) across *D. simulans* chromosome 3R (from Begun et al. 2007). Arrows indicate hypothetical effects of linked selection in raising or lowering levels of polymorphism. The dashed line represents the expected mutation–drift equilibrium level of polymorphism under neutrality, given as the average value of π on chromosome 3R in panel A and a hypothetical value unaffected by linked selection in panel B.

to infer demographic histories (e.g., Haddrill et al. 2005) does not obviate the problem of linked selection, especially if there is direct adaptive natural selection on noncoding loci in the same species (Andolfatto 2005). It is possible, however, that in species with genomes larger than those of *Drosophila* it might be possible to find truly neutral markers—multiple studies in humans have attempted to use such loci (e.g., Frisse et al. 2001; Rockman et al. 2003). Unfortunately, it still appears as though there is a strong correlation between levels of variation and recombination rates at these loci. For example, reanalysis of the data from Frisse et al. (2001) finds a correlation of $r = 0.64$ between π and map-based estimates of recombination among the Hausa people of Africa, and $r = 0.47$ between Tajima's D statistic (a measure of the frequency spectrum of mutations) and recombination in the same population (M. W. Hahn, unpubl. data).

Taken together the examples presented above illustrate one of the most misleading and intellectually disingenuous aspects of neutralist interpretations—the excess of low-frequency mutations expected under widespread natural selection can be explained away as the result of demography simply because most loci show this pattern. There are rarely independent data on demographic changes, and when there is—as with human migration out of Africa—disagreements of tens of thousands of years on the date of such changes are explained as estimation error or inconclusive archaeological dates rather than the joint effects of demography and selection. Although populations migrating to new environments

are surely subject to many new selective agents, evidence for adaptive evolution is often cast aside because a nonequilibrium model fits the data equally well. It would be very surprising if the high rates of adaptive fixations found across coding and noncoding regions of the genome did not occur in these populations. Even when data are collected from populations that are thought a priori to represent equilibrium histories (e.g., African populations of humans and flies), patterns consistent with rampant linked selection are interpreted as population expansions without any corroborating evidence. The joint effects of linked selection and demography may also explain why seemingly so few species are inferred to have had population contractions: the excess of intermediate-frequency mutations expected under a contraction may be cancelled out by patterns of linked selection, resulting in apparently "equilibrium" populations.

If we begin to change our view of the forces that shape variation, we begin to see that there may be very few loci in any genome that have escaped the effects of linked selection. This means that there are few genes that provide an unbiased view of demographic history and therefore that inferences of selection based on inferred histories are hugely conservative. It also means that π will rarely be a good estimator of $4N\mu_0$. Figure 2 attempts to show this major difference between neutral and selective views of molecular variation. The neutral view is shown in Figure 2A, where the Cavalli-Sforza/Lewontin/Krakauer axiom leads us to believe that the mean value of π represents the neutral expectation—whether or not the population is in demographic equilibrium—and only the most extreme deviations from this mean indicate either balancing selection (high values) or positive selection (low values). A selective view (Fig. 2B) acknowledges that most loci have levels of polymorphism much lower than are expected under neutrality, as linked selection is affecting almost the entire genome. Contrary to the neutral view, genes with the highest levels of polymorphism may simply be the only ones to have escaped linked selection, although balancing selection surely acts on some loci. These views provide vastly different conclusions about the amount of selection acting in nature, as well as about the relative effects of disparate evolutionary forces in shaping polymorphism.

Toward a Selection Theory of Molecular Evolution

If a case is to be made that the Neutral Theory is no longer an appropriate description of molecular variation, then we must replace it with a theory that includes a much larger role for natural selection. In addition, use of a "Selection Theory" will likely necessitate a change in the way that many inferences in population genetics are made: assuming we did agree that 50% (or some other number) of nonsynonymous changes were fixed by positive selection, this would then have to be used as the null model to test alternatives against. Alternatively—and in contrast to the statistical beliefs of many population geneticists—it may be that a Selection Theory requires much wider use of estimation methods (such as Bayesian approaches) rather than standard inferences driven by testing alternatives to an unrealistic null model.

The most challenging task in developing a Selection Theory is likely to be the absence of a single model that describes every species. Each species' specific biology will dictate whether deleterious or advantageous mutations are most commonly found, as suggested by the different patterns of variation found in humans and *Arabidopsis* versus *Drosophila* discussed above. For example, Reed et al. (2005) found that a background selection model fit human polymorphism data very well, both in terms of the level and frequency of polymorphisms. This suggests that much of the linked selection in humans will be driven by deleterious mutations. In contrast, the Begun et al. study of *D. simulans* found an excess of high-frequency mutations, consistent with a model of recurrent adaptive evolution. For now it is unclear which form of linked selection will predominate, or whether both will be found equally. We do not have enough data from enough species with varying demographic, life-history, and mating systems to make clear generalizations about when to expect one type of selective model or the other. What is clear is that instead of comparing simple selective models to complex demographic models (e.g., Wall et al. 2002) we should be attempting to distinguish among more realistic selective models (e.g., Reed et al. 2005). The theoretical tools necessary to do this are available, and the dissemination of more simulation programs that enable all researchers to investigate these alternatives are becoming available (Spencer and Coop 2004).

It should be stressed that by arguing for a turn away from the Neutral Theory I am not making the case for rampant adaptation or pan-selectionism in its widest meaning. Rather, the patterns apparent from multiple species at multiple loci make the case for rampant nonneutrality. These widespread deviations from neutrality neither distinguish between advantageous or deleterious mutations as their cause, nor do they tell us whether they are the result of direct selection or are merely the spandrels of linked selection. The results also do not say that all fixed differences between species are due to adaptive natural selection. In fact, whatever the proximate causes of deviations from neutrality, the ultimate results are likely to be the retardation of adaptation and the fixation of mildly deleterious mutations (Hill and Robertson 1966). Whatever the general conclusions drawn, it is clear that adherence to the Neutral Theory in the face of mounting evidence for selection is unwarranted, despite the intellectual effort required to shift our view. It is simply inconsistent to claim both that there are high rates of adaptive evolution and that the Neutral Theory is an adequate description of nature. As I began this essay with a quote from a dead economist, I will also end with one that seems to

sum up 40 years of research into molecular evolution: "It is a far, far better thing to have a firm anchor in nonsense than to put out on the troubled sea of thought [John Kenneth Galbraith]." Insert "neutrality" and "selection" as needed.

ACKNOWLEDGMENTS

I am grateful to P. Nista for discussion and for gathering many of the data cited here. D. Begun, L. Moyle, P. Nista, and B. Payseur also made many helpful comments on the manuscript, although this should not imply that they condone any apostate opinions held by the author. MWH is supported by grants from the NSF and NIH.

LITERATURE CITED

Akashi, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. Genetics 136:927–935.

———. 1999. Inferring the fitness effects of DNA mutations from polymorphism and divergence data: statistical power to detect directional selection under stationarity and free recombination. Genetics 151:221–238.

Akey, J. M., M. A. Eberle, M. J. Rieder, C. S. Carlson, M. D. Shriver, D. A. Nickerson, and L. Kruglyak. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol. 2:1591–1599.

Andolfatto, P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437:1149–1152.

Andolfatto, P., and M. Przeworski. 2001. Regions of lower crossing over harbor more rare variants in African populations of *Drosophila melanogaster*. Genetics 158:657–665.

Aquadro, C. F., D. J. Begun, and E. C. Kindahl. 1994. Selection, recombination, and DNA polymorphism in *Drosophila*. Pp. 46–56 *in* B. Golding, ed. Non-neutral evolution. Chapman and Hall, New York.

Bazin, E., S. Glemin, and N. Galtier. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312:570–572.

Begun, D. J., and C. F. Aquadro. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. Nature 356:519–520.

Begun, D. J., A. K. Holloway, K. Stephens, L. W. Hillier, Y.-P. Poh, M. W. Hahn, P. M. Nista, C. D. Jones, A. D. Kern, C. Dewey, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. 5:e310.

Bierne, N., and A. Eyre-Walker. 2004. The genomic rate of adaptive amino acid substitution in *Drosophila*. Mol. Biol. Evol. 21:1350–1360.

Birky, C. W. Jr., and J. B. Walsh. 1988. Effects of linkage on rates of molecular evolution. Proc. Natl. Acad. Sci. USA 85:6414–6418.

Bustamante, C. D., R. Nielsen, S. A. Sawyer, K. M. Olsen, M. D. Purugganan, and D. L. Hartl. 2002. The cost of inbreeding in *Arabidopsis*. Nature 416:531–534.

Bustamante, C. D., A. Fledel-Alon, S. Williamson, R. Nielsen, M. T. Hubisz, S. Glanowski, D. M. Tanenbaum, T. J. White, J. J. Sninsky, R. D. Hernandez et al. 2005. Natural selection on protein-coding genes in the human genome. Nature 437:1153–1157.

Cavalli-Sforza, L. L. 1966. Population structure and human evolution. Proc. R. Soc. Lond. B 164:362–379.

Charlesworth, J., and A. Eyre-Walker. 2006. The rate of adaptive evolution in enteric bacteria. Mol. Biol. Evol. 23:1348–1356.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303.

Currat, M., L. Excoffier, W. Maddison, S. P. Otto, N. Ray, M. C. Whitlock, and S. Yeaman. 2006. Comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in homo sapiens" and "microcephalin, a gene regulating brain size, continues to evolve adaptively in humans." Science 313:172a.

Cutter, A. D., and B. A. Payseur. 2003. Selection at linked sites in the partial selfer *Caenorhabditis elegans*. Mol. Biol. Evol. 20:665–673.

Dvorak, J., M. C. Luo, and Z. L. Yang. 1998. Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. Genetics 148:423–434.

Eyre-Walker, A., M. Woolfit, and T. Phelps. 2006. The distribution of fitness effects of new deleterious amino acid mutations in humans. Genetics 173:891–900.

Fay, J. C., and C. I. Wu. 2001. The neutral theory in the genomic era. Curr. Opin. Genet. Develop. 11:642–646.

Fay, J. C., G. J. Wyckoff, and C. I. Wu. 2002. Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. Nature 415:1024–1026.

Fisher, R. A. 1922. On the dominance ratio. Proc. R. Soc. Edinbourgh 42:321–341.

Frisse, L., R. R. Hudson, A. Bartoszewicz, J. D. Wall, J. Donfack, and A. Di Rienzo. 2001. Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. Am. J. Hum. Genet. 69:831–843.

Gaffney, D. J., and P. D. Keightley. 2005. The scale of mutational variation in the murid genome. Genome Res. 15:1086–1094.

Gillespie, J. H. 1984. The status of the neutral theory. Science 224:732–733.

———. 1991. The causes of molecular evolution. Oxford Univ. Press, New York.

———. 1994. Alternatives to the neutral theory. Pp. 1–17 *in* B. Golding, ed. Non-neutral evolution. Chapman and Hall, New York.

———. 2000. Genetic drift in an infinite population: The pseudohitchhiking model. Genetics 155:909–919.

———. 2001. Is the population size of a species relevant to its evolution? Evolution 55:2161–2169.

Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc. Lond. B 205:581–598.

Haddrill, P. R., K. R. Thornton, B. Charlesworth, and P. Andolfatto. 2005. Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. Genome Res. 15:790–799.

Hahn, M. W. 2006. Accurate inference and estimation in population genomics. Mol. Biol. Evol. 23:911–918.

———. 2007. Detecting natural selection on *cis*-regulatory DNA. Genetica 129:7–18.

Harris, H. 1966. Enzyme polymorphisms in man. Proc. R. Soc. Lond. B 164:298–310.

Hellmann, I., I. Ebersberger, S. E. Ptak, S. Paabo, and M. Przeworski. 2003. A neutral explanation for the correlation of diversity with recombination rates in humans. Am. J. Hum. Genet. 72:1527–1535.

Hellmann, I., K. Prufer, H. K. Ji, M. C. Zody, S. Paabo, and S. E. Ptak. 2005. Why do human diversity levels vary at a megabase scale? Genome Res. 15:1222–1231.

Hill, W. G., and A. Robertson. 1966. Effect of linkage on limits to artificial selection. Genetical Res. 8:269–294.

Hudson, R. R. 1983. Properties of a neutral allele model with intragenic recombination. Theor. Popul. Biol. 23:183–201.

Hudson, R. R., and N. L. Kaplan. 1995. Deleterious background selection with recombination. Genetics 141:1605–1617.

Kim, S., V. Plagnol, T. T. Hu, C. Toomajian, R. M. Clark, S. Ossowski, J. R. Ecker, D. Weigel, and M. Nordborg. 2007. Recombination and

linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. 39:1151–1155.

Kimura, M. 1955. Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. USA 41:144–150.

———. 1968. Evolutionary rate at the molecular level. Nature 217:624–626.

———. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge.

Kimura, M., and T. Ohta. 1971. Protein polymorphism as a phase of molecular evolution. Nature 229:467–469.

King, J. L., and T. H. Jukes. 1969. Non-Darwinian evolution. Science 164:788–798.

Kohn, M. H., S. Fang, and C.-I. Wu. 2004. Inference of positive and negative selection on the 5' regulatory regions of *Drosophila* genes. Mol. Biol. Evol. 21:374–383.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, et al. 2002. A high-resolution recombination map of the human genome. Nat. Genet. 31:241–247.

Kraft, T., T. Sall, I. Magnusson-Rading, N. O. Nilsson, and C. Hallden. 1998. Positive correlation between recombination rates and levels of genetic variation in natural populations of sea beet (*Beta vulgaris* subsp. *maritima*). Genetics 150:1239–1244.

Kreitman, M. 1996. The neutral theory is dead. Long live the neutral theory. Bioessays 18:678–683.

———. 2000. Methods to detect selection in populations with applications to the human. Annu. Rev. Genom. Hum. Genet. 1:539–559.

Lercher, M. J., and L. D. Hurst. 2002. Human SNP variability and mutation rate are higher in regions of high recombination. Trends Genet. 18:337–340.

Lercher, M. J., E. J. B. Williams, and L. D. Hurst. 2001. Local similarity in evolutionary rates extends over whole chromosomes in human-rodent and mouse-rat comparisons: implications for understanding the mechanistic basis of the male mutation bias. Mol. Biol. Evol. 18:2032–2039.

Lewontin, R. C. 1974. The genetic basis for evolutionary change. Columbia Univ. Press, New York, NY.

Lewontin, R. C., and J. L. Hubby. 1966. A molecular approach to the study of genetic heterozygosity in natural populations. II. amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595–609.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene-frequency as a test of the theory of the selective neutralism of polymorphisms. Genetics 74:175–195.

Lynch, M. 2006. The origins of eukaryotic gene structure. Mol. Biol. Evol. 23:450–468.

———. 2007. The origins of genome architecture. Sinauer Associates, Inc., Sunderland, MA.

Maynard Smith, J., and J. Haigh. 1974. The hitch-hiking effect of a favorable gene. Genet. Res. 23:23–35.

McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351:652–654.

Mekel-Bobrov, N., and B. T. Lahn. 2007. Response to Comments by Timpson et al. and Yu et al. Science 317:1036b.

Mekel-Bobrov, N., S. L. Gilbert, P. D. Evans, E. J. Vallender, J. R. Anderson, R. R. Hudson, S. A. Tishkoff, and B. T. Lahn. 2005. Ongoing adaptive evolution of ASPM, a brain size determinant in *Homo sapiens*. Science 309:1720–1722.

Mekel-Bobrov, N., P. D. Evans, S. L. Gilbert, E. J. Vallender, R. R. Hudson, and B. T. Lahn. 2006. Response to comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in homo sapiens" and "microcephalin, a gene regulating brain size, continues to evolve adaptively in humans." Science 313:172b.

Nachman, M. W. 1997. Patterns of DNA variability at X-linked loci in *Mus domesticus*. Genetics 147:1303–1316.

Nachman, M. W., V. L. Bauer, S. L. Crowell, and C. F. Aquadro. 1998. DNA variability and recombination rates at X-linked loci in humans. Genetics 150:1133–1141.

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. Heredity 86:641–647.

Nielsen, R., S. Williamson, Y. Kim, M. J. Hubisz, A. G. Clark, and C. Bustamante. 2005. Genomic scans for selective sweeps using SNP data. Genome Res. 15:1566–1575.

Ohta, T., and M. Kimura. 1971. On the constancy of the evolutionary rate of cistrons. J. Mol. Evol. 1:18–25.

Payseur, B. A., and M. W. Nachman. 2002. Natural selection at linked sites in humans. Gene 300:31–42.

Przeworski, M., R. R. Hudson, and A. Di rienzo. 2000. Adjusting the focus on human variation. Trends Genet. 16:296–302.

Reed, F. A., J. M. Akey, and C. F. Aquadro. 2005. Fitting background-selection predictions to levels of nucleotide variation and divergence along the human autosomes. Genome Res. 15:1211–1221.

Resch, A. M., L. Carmel, L. Marino-Ramirez, A. Y. Ogurtsov, S. A. Shabalina, I. B. Rogozin, and E. V. Koonin. 2007. Widespread positive selection in synonymous sites of mammalian genes. Mol. Biol. Evol. 24:1821–1831.

Rockman, M. V., and G. A. Wray. 2002. Abundant raw material for cis-regulatory evolution in humans. Mol. Biol. Evol. 19:1991–2004.

Rockman, M. V., M. W. Hahn, N. Soranzo, D. B. Goldstein, and G. A. Wray. 2003. Positive selection on a human-specific transcription factor binding site regulating *ILA* expression. Curr. Biol. 13:2118–2123.

Roselius, K., W. Stephan, and T. Stadler. 2005. The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. Genetics 171:753–763.

Sawyer, S. A., R. J. Kulathinal, C. D. Bustamante, and D. L. Hartl. 2003. Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. J. Mol. Evol. 57:S154–S164.

Shapiro, J. A., W. Huang, C. H. Zhang, M. J. Hubisz, J. Lu, D. A. Turissini, S. Fang, H. Y. Wang, R. R. Hudson, R. Nielsen, et al. 2007. Adaptive genic evolution in the *Drosophila* genomes. Proc. Natl. Acad. Sci. USA 104:2271–2276.

Simonsen, K. L., G. A. Churchill, and C. F. Aquadro. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics 141:413–429.

Slotman, M. A., L. J. Reimer, T. Thiemann, G. Dolo, E. Fondjo, and G. C. Lanzaro. 2006. Reduced recombination rate and genetic differentiation between the M and S forms of *Anopheles gambiae* s.s. Genetics 174:2081–2093.

Smith, N. G. C., and A. Eyre-Walker. 2002. Adaptive protein evolution in *Drosophila*. Nature 415:1022–1024.

Smith, N. G. C., and M. J. Lercher. 2002. Regional similarities in polymorphism in the human genome extend over many megabases. Trends Genet. 18:281–283.

Spencer, C. C. A., and G. Coop. 2004. SelSim: a program to simulate population genetic data with natural selection and recombination. Bioinformatics 20:3673–3675.

Stajich, J. E., and M. W. Hahn. 2005. Disentangling the effects of demography and selection in human history. Mol. Biol. Evol. 22:63–73.

Stephan, W., and C. H. Langley. 1998. DNA polymorphism in *Lycopersicon* and crossing-over per physical length. Genetics 150:1585–1593.

Stump, A. D., M. C. Fitzpatrick, N. F. Lobo, S. Traore, N. F. Sagnon, C. Costantini, F. H. Collins, and N. J. Besansky. 2005. Centromere-proximal

differentiation and speciation in ***Anopheles gambiae***. Proc. Natl. Acad. Sci. USA 102:15930–15935.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460.

Takahashi, A., Y. H. Liu, and N. Saitou. 2004. Genetic variation versus recombination rate in a structured population of mice. Mol. Biol. Evol. 21:404–409.

Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley, and B. S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays* L.). Proc. Natl. Acad. Sci. USA 98:9161–9166.

Thornton, K. R., J. D. Jensen, C. Becquet, and P. Andolfatto. 2007. Progress and prospects in mapping recent selection in the genome. Heredity 98:340–348.

Wall, J. D., P. Andolfatto, and M. Przeworski. 2002. Testing models of selection and demography in *Drosophila simulans*. Genetics 162:203–216.

Webster, M. T., N. G. C. Smith, M. J. Lercher, and H. Ellegren. 2004. Gene expression, synteny, and local similarity in human noncoding mutation rates. Mol. Biol. Evol. 21:1820–1830.

Weinreich, D. M., and D. M. Rand. 2000. Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. Genetics 156:385–399.

Wiehe, T. H. E., and W. Stephan. 1993. Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from *Drosophila melanogaster*. Mol. Biol. Evol. 10:842–854.

Williamson, S. H., R. Hernandez, A. Fledel-Alon, L. Zhu, R. Nielsen, and C. D. Bustamante. 2005. Simultaneous inference of selection and population growth from patterns of variation in the human genome. Proc. Natl. Acad. Sci. USA 102:7882–7887.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Yu, F. L., R. S. Hill, S. F. Schaffner, P. C. Sabeti, E. T. Wang, A. A. Mignault, R. J. Ferland, R. K. Moyzis, C. A. Walsh, and D. Reich. 2007. Comment on "Ongoing adaptive evolution of ASPM, a brain size determinant in Homo sapiens." Science 316:370b.

Zuckerkandl, E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Pp. 97–166 *in* V. Bryson and H. J. Vogel, eds. Evolving genes and proteins. Academic Press, New York.

Associate Editor: M. Rausher