The Neutral Theory in Light of Natural Selection

Andrew D. Kern*[,1] and Matthew W. Hahn[2]

[1]Department of Genetics, Rutgers University, Piscataway, NJ

[2]Department of Biology and Department of Computer Science, Indiana University Bloomington, IN

***Corresponding author:** E-mail: kern@biology.rutgers.edu.

Associate editor: Sudhir Kumar

Abstract

In this perspective, we evaluate the explanatory power of the neutral theory of molecular evolution, 50 years after its introduction by Kimura. We argue that the neutral theory was supported by unreliable theoretical and empirical evidence from the beginning, and that in light of modern, genome-scale data, we can firmly reject its universality. The ubiquity of adaptive variation both within and between species means that a more comprehensive theory of molecular evolution must be sought.

Key words: **natural selection, neutral theory, population genetics.**

Perspective

Introduction

On the 50th anniversary of the neutral theory of molecular evolution, we have been charged with the task of asking: how has the neutral theory fared in light of adaptive variation within and between species? In a word, poorly. While neutral models have without doubt begat tremendous theoretical fruits, including whole conceptual structures (e.g., the coalescent), the explanatory power of the neutral theory has never been exceptional. Five decades after its proposal, in the age of cheap genome sequencing and tremendous population genomic data sets, the explanatory power of the neutral theory looks even worse. In this perspective, we argue that with modern data in hand, each of the original lines of evidence for the neutral theory are now falsified, and that genomes are shaped in prominent ways by the direct and indirect consequences of natural selection.

To begin, we should make clear what the neutral theory claims about nature. It is not simply a statement about the presence of neutral mutations, nor about the large fraction of eukaryotic genomes that are nonfunctional—neither of these assertions would be contested by current competing hypotheses. Furthermore, the neutral theory is not merely a neutral model, to be used as a null hypothesis against which more interesting hypotheses can be tested. The neutral theory instead posits a positive thesis about nature: that differences between species are due to neutral substitutions (not adaptive evolution), and that polymorphisms within species are not only neutral but also have dynamics dominated by mutation-drift equilibrium. It was these claims, and their attendant theoretical justifications, that were the original attraction of the neutral theory as an explanatory framework. However, we must also acknowledge the important roles that both Kimura (1968) and King and Jukes (1969) played in the field's acceptance of neutral mutations at all. Although we argue here that the neutral theory has not held up in light of genomic data, it is certainly the case that neutral mutations—in both functional and nonfunctional parts of the genome—are now widely recognized. The presence of neutral variation was certainly not part of the orthodoxy of the late 1960s, in which balancing selection predominated (discussed in chapter 2 of Kimura 1983).

Original Evidence for the Neutral Theory

Given the historical purpose of this issue, we wish to step back and examine the original lines of evidence that were offered to justify the neutral theory of molecular evolution. As we will argue, none of them stand up to modern scrutiny. In fact, many researchers in the field are unlikely to be aware of these original arguments—and would be even less likely to believe them—even if they are self-described neutralists.

Kimura (1968) famously used the "cost of selection" to argue that rates of protein evolution, at that point calculated from three loci in a handful of mammals, were too rapid to be compatible with natural selection driving substitutions. Unfortunately, Kimura's calculation was flawed in a number of ways. For instance, Kimura overestimates the number of protein-coding sites in the genome by two orders of magnitude (he uses 4×10^9 bp, whereas the real number is closer to 3×10^7 bp in the largest genomes). Even given the data on rates Kimura had available, fixing this number alone would remove the conflict between the rate of protein evolution at the nucleotide level and Haldane's (1957) upper limit of 300 years per adaptive substitution. However, even without the benefit of modern genome sequence data, Kimura's cost of selection argument was critiqued immediately on theoretical grounds. Maynard Smith (1968) and Sved (1968) argued that Haldane's results could be ameliorated by truncation selection and density-dependence, and therefore that Haldane's cost (and by proxy Kimura's argument) was too restrictive.

Perhaps the most compelling argument against the cost of selection being evidence for neutral evolution came from

© The Author(s) 2018. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990864
by Indiana University Libraries - Bloomington user
on 31 July 2018

Felsenstein (1971). Felsenstein rederived the expected cost of selection in two separate ways, allowing him to cast the problem as one of a population chasing a moving optimum after environmental change. He showed that, depending on the initial frequency of a beneficial allele after the environment has changed and the number of offspring per parent, the maximum rate of adaptive substitution varies over orders of magnitude. This is not to say that there are no limits to the rate of adaptation in populations (Weissman and Barton 2012), only that it is not clear to what extent such limits operate in nature. As a consequence, Kimura's central 1968 calculation—that the rate of amino acid evolution was too high given Haldane's calculated rate limit due to the cost of selection—is both technically and conceptually flawed.

Three years later, Kimura's justification for the neutral theory had shifted from limits on the rate of evolution to its constancy among lineages. As stated in Kimura and Ohta (1971): "Probably the strongest evidence for the theory is the remarkable uniformity for each protein molecule in the rate of mutant substitutions in the course of evolution." Armed with data from a handful of proteins, the constancy of the rate of amino acid substitution among disparate lineages was presumed to be due to the fact that all substitutions were neutral. Accumulating evidence from additional proteins, coupled with better analyses (Langley and Fitch 1974), soon showed that this constancy was an illusion (Gillespie 1989; Cutler 2000; Bedford and Hartl 2008).

Decades of data later, it is clear that the original pillars of the neutral theory do not hold. However, there are certainly neutral mutations and neutral substitutions, so perhaps some parts of the neutral theory can be saved when new data are brought to bear on the subject. In the next section, we examine whether genomic data on the neutrality of between-species divergence and within-species levels of polymorphism match any of the predictions of the neutral theory.

The Evidence for Selection

How much of the genome is directly or indirectly influenced by adaptive natural selection? Since the first data on variation in nucleotide sequences within a population were collected (Aquadro and Greenberg 1983; Kreitman 1983), this question has been a central focus of population genetics. Many different tests have been developed to test for the action of positive or balancing selection—often against a null model that assumes neutrality—and new genomic data have inspired new and more powerful methods for detecting selection in all its various forms.

One of the most powerful and robust tests for the action of positive selection on divergence between species was suggested by McDonald and Kreitman (1991). The so-called McDonald–Kreitman (MK) test combines polymorphism and divergence data in order to test a prediction of the neutral model that these quantities should be proportional to one another for both synonymous and nonsynonymous variants. While the MK test will not be able to detect adaptive evolution on only one or a few fixed differences, no matter the strength of selection, it is much more powerful than tests based solely on divergence (such as d_N/d_S) and much more robust to nonequilibrium demographic histories than tests based solely on polymorphism (such as Tajima's D).

Application of the MK test to data from protein-coding genes has revealed a predominant role for adaptive natural selection. The first such studies were carried out in *Drosophila melanogaster* and *D. simulans*, finding that ~50% of all amino acid substitutions have been fixed by positive selection (Fay et al. 2002; Smith and Eyre-Walker 2002; Sawyer et al. 2003; Bierne and Eyre-Walker 2004; Begun et al. 2007; Shapiro et al. 2007; Langley et al. 2012). Accumulating whole-genome data from a variety of species has continued to find a large fraction of substitutions fixed by positive selection (Charlesworth and Eyre-Walker 2006; Halligan et al. 2010; Carneiro et al. 2012; Tsagkogeorga et al. 2012; Galtier 2016). Even purported exceptions to this pattern have given way upon closer analysis. In humans, though the overall fraction of amino acid substitutions fixed by positive selection is estimated to be zero (Boyko et al. 2008; Li et al. 2008), careful functional characterization of individual proteins has revealed that a large fraction of all genes which interact with pathogens show pervasive evidence for positive selection (Enard et al. 2016; Ebel et al. 2017). Similarly, despite the lack of signal of positive selection in early studies of plants that used small numbers of loci (Gossmann et al. 2010), newer data sets with larger numbers of genes have again found strong patterns of adaptation (Williamson et al. 2014; Grivet et al. 2017).

One unavoidable charge against the MK test is that it is based on expectations of a neutral model. Although the utility of neutral models does not necessarily support the accuracy of the neutral theory as a statement about nature, there is a certain ambivalence to accepting one and not the other. But consider the alternative: if the field used a model of positive selection as the null hypothesis, failure to reject this null should of course *not* be taken as evidence for selection. In many cases there is little alternative at the moment except to use a neutral model as the null hypothesis, in order to break free of the claims of the neutral theory.

In contrast to methods for examining the effect of selection on divergence, methods for understanding how selection shapes within-species patterns of variation are highly dependent on nonequilibrium population histories. In order to account for this reliance, researchers either take a predefined fraction of loci in the tails of a distribution as the number affected by selection, or assume that a nonequilibrium history explains the majority of the data by fitting a highly embellished model of demography in order to erase all signs of outliers. While there are some promising methods to coestimate selection and demography (see below), to understand the genome-wide effects of selection we must take a different approach.

One of the most striking impacts of natural selection on genomes is the near universal correlation between rates of recombination and levels of polymorphism (Hahn 2008; Cutter and Payseur 2013; Corbett-Detig et al. 2015; see fig. 1). Under neutrality, no relationship between levels of polymorphism and recombination is expected, as the number and frequency of neutral mutations is unaffected by

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990884
by Indiana University Libraries - Bloomington user
on 31 July 2018

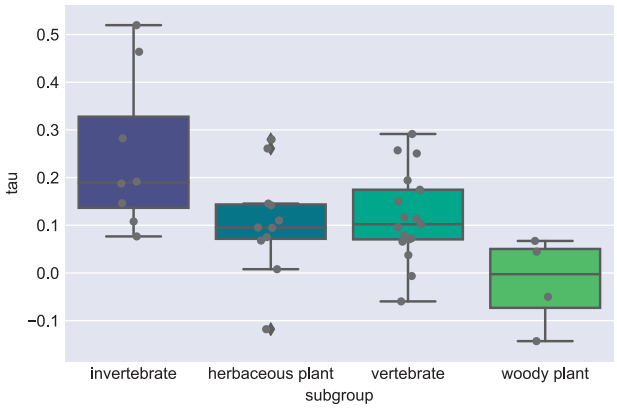

FIG. 1. Correlation coefficients ("tau") between levels of polymorphism and recombination rate from 40 genomes belonging to various multicellular subgroups (data from Corbett-Detig et al. 2015).

recombination (Hudson 1983). In the presence of selection, however, levels of polymorphism are reduced by an amount proportional to the strength of selection and the recombination rate (Maynard Smith and Haigh 1974; Kaplan et al. 1989; Charlesworth et al. 1993; Barton 1998). As such, there will be less polymorphism in regions of lower recombination, and more polymorphism in regions of higher recombination. The correlation between recombination and polymorphism could formally have a neutral explanation, if, for instance, recombination were mutagenic. Begun and Aquadro (1992) tested for such an effect by looking at the correlation between recombination and divergence, but found no relationship between the two. Additional alternative neutral explanations for this relationship have also been excluded (McGaugh et al. 2012; Pease and Hahn 2013). Thus, at the whole-genome scale it is readily apparent that selective forces need to be invoked to adequately explain gross features of population genetic variation. However, it is less clear to what extent linked positive versus linked negative selection is predominant, and the effect of each may differ across species.

The positive correlation between polymorphism and recombination across many plant and animal species is striking for a number of reasons. First, these results imply that almost no loci are free from the effects of selection, in any organism. Far from being limited to only the regions of lowest recombination, published 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 (Hahn 2008; Sella et al. 2009). Second, in the absence of other forces, the reduction in variation caused by linked selection will rebound to equilibrium levels relatively rapidly (Simonsen et al. 1995; Barton 1998). The fact that polymorphism is correlated with recombination implies that in almost every species examined, at almost every locus, there has recently been a selected allele nearby (whether advantageous or deleterious), such that levels of polymorphism are not at mutation-drift equilibrium. An equilibrium between mutation and drift is a central tenet of the neutral theory (Kimura and Ohta 1971); therefore, current data appear to be fundamentally incompatible with the neutral theory.

In addition to settling existing arguments, genome-scale studies have uncovered challenges to the neutral theory unimagined 50 years ago. From the exquisite detail on local adaptation from even species with low effective population sizes (such as humans; Fan et al. 2016), to broad patterns gathered from across the tree of life, increased sequencing has further marginalized the neutral theory. We review a few of these advances in what follows.

Over the past decade, both empirical data and theoretical advances have sufficiently accumulated to suggest that adaptive evolution is not mutation-limited in natural populations. Instead, selection from standing variation may be the typical response to an environmental shift (Gillespie 1991; Hermisson and Pennings 2005; Messer and Petrov 2013; Garud et al. 2015; Sheehan and Song 2016; Schrider and Kern 2017). The abundance of these "soft" selective sweeps means that even if drift plays an important role in some portion of the sojourn of an allele, the influence of natural selection can still dominate the evolutionary trajectory at other points. While selection from standing variation within a focal population is a potent source of adaptive variation, yet another source is beneficial mutations from other populations or species. Adaptive introgression, while long hypothesized to be an important source of variation (Anderson and Stebbins 1954), has only recently been shown to be common in nature (reviewed in Hedrick 2013). Indeed, recent examples of adaptive introgression include a wide swath of organismal diversity including plants (Bechsgaard et al. 2017), fungi (Cheeseman et al. 2014), insects (Salazar et al. 2010; Fontaine et al. 2015), and even our own distant ancestors (Huerta-Sánchez et al. 2014). Thus, the ubiquity of adaptive introgression provides another route toward adaptation, and additional sources of potentially adaptive variation. Taken together, modern evidence for soft sweeps and adaptive introgression suggest that the supply of beneficial mutations will not be a major limiting factor over evolutionary time.

While the search for selective sweeps of any stripe has been a dominant theme in population genetics, there is good theoretical reason to believe that phenotypes that are highly polygenic (i.e., that result from genetic contributions at many loci) might not be associated with fixation of advantageous alleles at all (Pritchard et al. 2010; Jain and Stephan 2017). This implies that, for a large number of evolutionarily important phenotypes, searching for selective sweeps might be an effort made in vain. The signals of selection will be much more subtle and possibly much more pervasive—the GWAS revolution in humans over the past decade has revealed that many phenotypes are polygenic. In response, a growing number of researchers have focused on devising methods that might be able to detect the signatures of polygenic selection in the genome. The most intuitive approaches combine information from GWAS with population genetic information on allele frequencies, asking whether a specific phenotypic difference between populations is associated with increased differentiation of the specific alleles known to affect the trait (Turchin et al. 2012; Berg and Coop 2014), or if such trait-associated SNPs are associated with signals of linked selection (Field et al. 2016). The knowledge of functional alleles

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990884
by Indiana University Libraries - Bloomington user
on 31 July 2018

across species will enable similar analyses in many more systems, especially in species in which we can examine loci that are known to directly affect fitness (Agren et al. 2013).

It also must be stressed that the evidence for selection summarized above has come from across sequenced genomes—in coding and noncoding regions—due to all different types of mutations—not just single nucleotide differences. Many of the strongest signals of selective sweeps are found in noncoding regions, possibly affecting RNA genes or the *cis*-regulatory apparatus of nearby protein-coding genes (Wang et al. 1999; Tishkoff et al. 2007). Extensions of the MK test to such regulatory sequences have revealed a large fraction of substitutions in these regions fixed due to positive selection (Jenkins et al. 1995; Ludwig and Kreitman 1995; Crawford et al. 1999; Kohn et al. 2004; Andolfatto 2005; MacDonald and Long 2005; Holloway et al. 2007; Jeong et al. 2008; Torgerson et al. 2009). We now also appreciate the wide range of different types of mutations that may be underlying adaptation, not just single nucleotide substitutions. Changes to gene copy-number (Perry et al. 2007; Schrider et al. 2013), the insertion of transposable elements (Daborn et al. 2002; Schlenke and Begun 2004; González et al. 2008), and even large inversions (Stefansson et al. 2005; Kolaczkowski et al. 2011; Cheng et al. 2012; Kirkpatrick and Kern 2012; Reinhardt et al. 2014) have all been involved in adaptive natural selection.

The Way Forward

We have presented accumulated evidence from the past 50 years that natural selection has played the predominant role in shaping within- and between-species genetic variation. As a consequence, we believe that the neutral theory has been overwhelmingly rejected, and that as a field we must continue to develop alternate theories of molecular evolution.

How will such a change in view affect how we make inferences from sequence data? Rejecting the neutral theory does not mean embracing adaptive storytelling, nor does it mean that we must forsake all models that assume neutrality. But we must recognize that assuming a neutral model for the sake of statistical convenience can positively mislead our inferences. One area where this problem is especially dire is in the estimation of demographic histories. While most populations almost certainly have a nonequilibrium history, attempting to infer the details of these histories without accounting for selective forces can mislead us in multiple ways. For instance, methods may infer migration between populations when none has occurred (Mathew and Jensen 2015; Roux et al. 2016), or they may infer nonequilibrium dynamics even in equilibrium populations (Ewing and Jensen 2016; Schrider et al. 2016). Meanwhile, nonadaptive storytelling in the form of overly fit demographic models can mask all signs of natural selection (Hahn 2008). Recent methods for coestimating selection and demography (Li and Stephan 2006; Sheehan and Song 2016) are moving us one important step forward: the ability to estimate demography without assuming neutrality. In parallel, newer methods for detecting selection that are suitably robust to demographic misspecification (Schrider and Kern 2016) provide the ability to detect all of the signals of selection even in the presence of nonequilibrium demography.

In order to more completely remove the lingering misapprehensions of the neutral theory, we must of course replace it with an explanatory theory of greater value. A more sufficient model of genetic variation would at minimum have to account for the direct and indirect effects of selective sweeps (Maynard Smith and Haigh 1974; Gillespie 2000) and the direct and indirect effects of purifying selection (Charlesworth et al. 1993; Hudson and Kaplan 1995), while simultaneously accounting for variation in population size and population structure. If this already sounds like a difficult task to accomplish, we can raise the stakes and add that population genetic models that operate in continuous space, that is, those that reflect the basic realities of geography, are still only in their infancy. Coupled with increasing amounts of data from new types of population samples—for example, those including noncontemporaneous individuals (such as from ancient DNA) or from very large pedigrees—future theories of molecular evolution will have to be able to service an ever-widening set of approaches. Thus, 50 years after the birth of the neutral theory, we wish to both celebrate its history and move on to more productive efforts.

Acknowledgment

ADK was supported by National Institute of General Medical Sciences award no. R01GM117241.

References

Agren J, Oakley CG, McKay JK, Lovell JT, Schemske DW. 2013. Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A.* 110(52):21077–21082.

Anderson E, Stebbins GL. 1954. Hybridization as an evolutionary stimulus. *Evolution* 8(4):378–388.

Andolfatto P. 2005. Adaptive evolution of non-coding DNA in Drosophila. *Nature* 437(7062):1149–1152.

Aquadro CF, Greenberg BD. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103(2):287–312.

Barton NH. 1998. The effect of hitch-hiking on neutral genealogies. *Genet Res.* 72(2):123–133.

Bechsgaard J, Jorgensen TH, Schierup MH. 2017. Evidence for adaptive introgression of disease resistance genes among closely related *Arabidopsis* species. *G3 (Bethesda)* 7(8):2677–2683.

Bedford T, Hartl DL. 2008. Overdispersion of the molecular clock: temporal variation of gene-specific substitution rates in *Drosophila*. *Mol Biol Evol.* 25(8):1631–1638.

Begun DJ, Aquadro CF. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356(6369):519.

Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh Y-P, Hahn MW, Nista PM, Jones CD, Kern AD, Dewey CN, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* 5(11):e310.

Berg JJ, Coop G. 2014. A population genetic signal of polygenic adaptation. *PLoS Genet.* 10(8):e1004412.

Bierne N, Eyre-Walker A. 2004. The genomic rate of adaptive amino acid substitution in *Drosophila*. *Mol Biol Evol.* 21(7):1350–1360.

Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, Lohmueller KE, Adams MD, Schmidt S, Sninsky JJ, Sunyaev SR, et al.

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990884
by Indiana University Libraries - Bloomington user
on 31 July 2018

2008. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* 4(5):e1000083.
Carneiro M, Albert FW, Melo-Ferreira J, Galtier N, Gayral P, Blanco-Aguiar JA, Villafuerte R, Nachman MW, Ferrand N. 2012. Evidence for widespread positive and purifying selection across the European rabbit (*Oryctolagus cuniculus*) genome. *Mol Biol Evol.* 29(7):1837–1849.
Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134(4):1289–1303.
Charlesworth J, Eyre-Walker A. 2006. The rate of adaptive evolution in enteric bacteria. *Mol Biol Evol.* 23(7):1348–1356.
Cheeseman K, Ropars J, Renault P, Dupont J, Gouzy J, Branca A, Abraham A-L, Ceppi M, Conseiller E, Debuchy R. 2014. Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nat Commun.* 5:2876.
Cheng C, White BJ, Kamdem C, Mockaitis K, Costantini C, Hahn MW, Besansky NJ. 2012. Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics* 190(4):1417–1432.
Corbett-Detig RB, Hartl DL, Sackton TB. 2015. Natural selection constrains neutral diversity across a wide range of species. *PLoS Biol.* 13(4):e1002112.
Crawford DL, Segal JA, Barnett JL. 1999. Evolutionary analysis of TATA-less proximal promoter function. *Mol Biol Evol.* 16(2):194–207.
Cutler DJ. 2000. Understanding the overdispersed molecular clock. *Genetics* 154(3):1403–1417.
Cutter AD, Payseur BA. 2013. Genomic signatures of selection at linked sites: unifying the disparity among species. *Nat Rev Genet.* 14(4):262–274.
Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N. 2002. A single P450 allele associated with insecticide resistance in *Drosophila*. *Science* 297(5590):2253–2256.
Ebel ER, Telis N, Venkataram S, Petrov DA, Enard D. 2017. High rate of adaptation of mammalian proteins that interact with *Plasmodium* and related parasites. *PLoS Genet.* 13(9):e1007023.
Enard D, Cai L, Gwennap C, Petrov DA. 2016. Viruses are a dominant driver of protein adaptation in mammals. *eLife* 5:e12469.
Ewing GB, Jensen JD. 2016. The consequences of not accounting for background selection in demographic inference. *Mol Ecol.* 25(1):135–141.
Fan S, Hansen MEB, Lo Y, Tishkoff SA. 2016. Going global by adapting local: a review of recent human adaptation. *Science* 354(6308):54–59.
Fay JC, Wyckoff GJ, Wu C-I. 2002. Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. *Nature* 415(6875):1024–1026.
Felsenstein J. 1971. On the biological significance of the cost of gene substitution. *Am Nat.* 105(941):1–11.
Field Y, Boyle EA, Telis N, Gao Z, Gaulton KJ, Golan D, Yengo L, Rocheleau G, Froguel P, McCarthy MI, Pritchard JK. 2016. Detection of human adaptation during the past 2000 years. *Science* 354(6313):760–764.
Fontaine MC, Pease JB, Steele A, Waterhouse RM, Neafsey DE, Sharakhov IV, Jiang X, Hall AB, Catteruccia F, Kakani E, et al. 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* 347(6217):1258524.
Galtier N. 2016. Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genet.* 12(1):e1005774.
Garud NR, Messer PW, Buzbas EO, Petrov DA. 2015. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet.* 11(2):e1005004.
Gillespie JH. 1989. Lineage effects and the index of dispersion of molecular evolution. *Mol Biol Evol.* 6(6):636–647.
Gillespie JH. 1991. The causes of molecular evolution. Oxford: Oxford University Press.
Gillespie JH. 2000. Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* 155(2):909–919.
González J, Lenkov K, Lipatov M, Michael Macpherson J, Petrov DA. 2008. High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biol.* 6(10):e251.
Gossmann TI, Song B-H, Windsor AJ, Mitchell-Olds T, Dixon CJ, Kapralov MV, Filatov DA, Eyre-Walker A. 2010. Genome wide analyses reveal little evidence for adaptive evolution in many plant species. *Mol Biol Evol.* 27(8):1822–1832.
Grivet D, Avia K, Vaattovaara A, Eckert AJ, Neale DB, Savolainen O, González-Martínez SC. 2017. High rate of adaptive evolution in two widespread European pines. *Mol Ecol.* 26(24):6857–6870.
Hahn MW. 2008. Toward a selection theory of molecular evolution. *Evolution* 62(2):255–265.
Haldane JBS. 1957. The cost of natural selection. *J Genet.* 55(3):511.
Halligan DL, Oliver F, Eyre-Walker A, Harr B, Keightley PD. 2010. Evidence for pervasive adaptive protein evolution in wild mice. *PLoS Genet.* 6(1):e1000825.
Hedrick PW. 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol.* 22(18):4606–4618.
Hermisson J, Pennings PS. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169(4):2335–2352.
Holloway AK, Lawniczak MKN, Mezey JG, Begun DJ, Jones CD. 2007. Adaptive gene expression divergence inferred from population genomics. *PLoS Genet.* 3(10):e187.
Hudson RR. 1983. Properties of a neutral allele model with intragenic recombination. *Theor Popul Biol.* 23(2):183–201.
Hudson RR, Kaplan NL. 1995. Deleterious background selection with recombination. *Genetics* 141(4):1605–1617.
Huerta-Sánchez E, Jin X, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X. 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512(7513):194–197.
Jain K, Stephan W. 2017. Rapid adaptation of a polygenic trait after a sudden environmental shift. *Genetics* 206(1):389–406.
Jenkins DL, Ortori CA, Brookfield JFY. 1995. A test for adaptive change in DNA sequences controlling transcription. *Proc R Soc Lond B* 261(1361):203–207.
Jeong S, Rebeiz M, Andolfatto P, Werner T, True J, Carroll SB. 2008. The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132(5):783–793.
Kaplan NL, Hudson RR, Langley CH. 1989. The "hitchhiking effect" revisited. *Genetics* 123(4):887–899.
Kimura M. 1968. Evolutionary rate at the molecular level. *Nature* 217(5129):624–626.
Kimura M. 1983. The neutral theory of molecular evolution. Cambridge: Cambridge University Press.
Kimura M, Ohta T. 1971. Protein polymorphism as a phase of molecular evolution. *Nature* 229(5285):467–469.
King JL, Jukes TH. 1969. Non-Darwinian evolution. *Science* 164(3881):788–798.
Kirkpatrick M, Kern A. 2012. Where's the money? Inversions, genes, and the hunt for genomic targets of selection. *Genetics* 190(4):1153–1155.
Kohn MH, Fang S, Wu C-I. 2004. Inference of positive and negative selection on the 5′ regulatory regions of *Drosophila* genes. *Mol Biol Evol.* 21(2):374–383.
Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. *Genetics* 187(1):245–260.
Kreitman M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304(5925):412–417.
Langley CH, Fitch WM. 1974. An examination of the constancy of the rate of molecular evolution. *J Mol Evol.* 3(3):161–177.
Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB, Kolaczkowski B. 2012. Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics* 192(2):533–598.
Li H, Stephan W. 2006. Inferring the demographic history and rate of adaptive substitution in *Drosophila*. *PLoS Genet.* 2(10):e166.
Li YF, Costello JC, Holloway AK, Hahn MW. 2008. "Reverse ecology" and the power of population genomics. *Evolution* 62(12):2984–2994.

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990884
by Indiana University Libraries - Bloomington user
on 31 July 2018

Ludwig MZ, Kreitman M. 1995. Evolutionary dynamics of the enhancer region of *even-skipped* in *Drosophila*. *Mol Biol Evol.* 12(6):1002–1011.
MacDonald SJ, Long AD. 2005. Prospects for identifying functional variation across the genome. *Proc Natl Acad Sci U S A*.102(Suppl 1):6614–6621.
Mathew LA, Jensen JD. 2015. Evaluating the ability of the pairwise joint site frequency spectrum to co-estimate selection and demography. *Front Genet.* 6:268.
McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351(6328):652–654.
McGaugh SE, Heil CSS, Manzano-Winkler B, Loewe L, Goldstein S, Himmel TL, Noor MAF. 2012. Recombination modulates how selection affects linked sites in *Drosophila*. *PLoS Biol.* 10(11):e1001422.
Messer PW, Petrov DA. 2013. Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol Evol.* 28(11):659–669.
Pease JB, Hahn MW. 2013. More accurate phylogenies inferred from low-recombination regions in the presence of incomplete lineage sorting. *Evolution* 67(8):2376–2384.
Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, et al. 2007. Diet and the evolution of human amylase gene copy number variation. *Nat Genet.* 39(10):1256–1260.
Pritchard JK, Pickrell JK, Coop G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol.* 20(4):R208–R215.
Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel geographic variation in *Drosophila melanogaster*. *Genetics* 197(1):361–373.
Roux C, Fraïsse C, Romiguier J, Anciaux Y, Galtier N, Bierne N. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* 14(12):e2000234.
Salazar C, Baxter SW, Pardo-Diaz C, Wu G, Surridge A, Linares M, Bermingham E, Jiggins CD. 2010. Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. *PLoS Genet.* 6(4):e1000930.
Sawyer SA, Kulathinal RJ, Bustamante CD, Hartl DL. 2003. Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. *J Mol Evol.* 57(0):S154–S164.
Schlenke TA, Begun DJ. 2004. Strong selective sweep associated with a transposon insertion in *Drosophila simulans*. *Proc Natl Acad Sci U S A.* 101(6):1626–1631.
Schrider DR, Kern AD. 2016. S/HIC: robust identification of soft and hard sweeps using machine learning. *PLoS Genet.* 12(3):e1005928.
Schrider DR, Kern AD. 2017. Soft sweeps are the dominant mode of adaptation in the human genome. *Mol Biol Evol.* 34(8):1863–1877.
Schrider DR, Navarro FCP, Galante PAF, Parmigiani RB, Camargo AA, Hahn MW, de Souza SJ. 2013. Gene copy-number polymorphism caused by retrotransposition in humans. *PLoS Genet.* 9(1):e1003242.
Schrider DR, Shanku AG, Kern AD. 2016. Effects of linked selective sweeps on demographic inference and model selection. *Genetics* 204(3):1207–1223.
Sella G, Petrov DA, Przeworski M, Andolfatto P. 2009. Pervasive natural selection in the *Drosophila* genome. *PLoS Genet.* 5(6):e1000495.
Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, Turissini DA, Fang S, Wang H-Y, Hudson RR, Nielsen R, et al. 2007. Adaptive genic evolution in the *Drosophila* genomes. *Proc Natl Acad Sci U S A.* 104(7):2271–2276.
Sheehan S, Song YS. 2016. Deep learning for population genetic inference. *PLoS Comput Biol.* 12(3):e1004845.
Simonsen KL, Churchill GA, Aquadro CF. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* 141(1):413–429.
Smith JM. 1968. "Haldane's dilemma" and the rate of evolution. *Nature* 219(5159):1114–1116.
Smith JM, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genetical Res.* 23(1):23–35.
Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in *Drosophila*. *Nature* 415(6875):1022–1024.
Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, Baker A, Jonasdottir A, Ingason A, Gudnadottir VG, et al. 2005. A common inversion under selection in Europeans. *Nat Genet.* 37(2):129–137.
Sved JA. 1968. Possible rates of gene substitution in evolution. *Am Nat.* 102(925):283–293.
Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, et al. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet.* 39(1):31–40.
Torgerson DG, Boyko AR, Hernandez RD, Indap A, Hu X, White TJ, Sninsky JJ, Cargill M, Adams MD, Bustamante CD, et al. 2009. Evolutionary processes acting on candidate cis-regulatory regions in humans inferred from patterns of polymorphism and divergence. *PLoS Genet.* 5(8):e1000592.
Tsagkogeorga G, Cahais V, Galtier N. 2012. The population genomics of a fast evolver: high levels of diversity, functional constraint, and molecular adaptation in the tunicate *Ciona intestinalis*. *Genome Biol Evol.* 4(8):852–861.
Turchin MC, Chiang CWK, Palmer CD, Sankararaman S, Reich D, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Hirschhorn JN. 2012. Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat Genet.* 44(9):1015–1019.
Wang R-L, Stec A, Hey J, Lukens L, Doebley J. 1999. The limits of selection during maize domestication. *Nature* 398(6724):236–239.
Weissman DB, Barton NH. 2012. Limits to the rate of adaptive substitution in sexual populations. *PLoS Genet.* 8(6):e1002740.
Williamson RJ, Josephs EB, Platts AE, Hazzouri KM, Haudry A, Blanchette M, Wright SI. 2014. Evidence for widespread positive and negative selection in coding and conserved noncoding regions of *Capsella grandiflora*. *PLoS Genet.* 10(9):e1004622.

Downloaded from https://academic.oup.com/mbe/article-abstract/35/6/1366/4990884
by Indiana University Libraries - Bloomington user
on 31 July 2018