Opinion

Cell PRESS

Weighing the evidence for adaptation at the molecular level

Justin C. Fay

Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University, St. Louis, MO, USA

The abundance of genome polymorphism and divergence data has provided unprecedented insight into how mutation, drift and natural selection shape genome evolution. Application of the McDonald–Kreitman (MK) test to such data indicates a pervasive influence of positive selection, particularly in *Drosophila* species. However, evidence for positive selection in other species ranging from yeast to humans is often weak or absent. Although evidence for positive selection could be obscured in some species, there is also reason to believe that the frequency of adaptive substitutions could be overestimated as a result of epistatic fitness effects or hitchhiking of deleterious mutations. Based on these considerations it is argued that the common assumption of independence among sites must be relaxed before abandoning the neutral theory of molecular evolution.

Estimating the frequency of positive selection using the MK test

The extent to which molecular evolution is driven by positive selection has long been debated. The neutral theory (see Glossary) holds that the vast majority of DNA sequence differences between species are neutral [1] or nearly neutral [2] with respect to fitness. However, models assuming that natural selection plays a dominant role in driving molecular evolution can also explain many features of DNA polymorphism and divergence data [3]. As a consequence, great efforts have been made to identify patterns of variation that distinguish between neutral and selective models [1,3–5]. Because of these efforts, many examples of genes evolving under positive selection have been documented. However, genome-wide estimates of the fraction of substitutions driven by positive selection are required to test the neutral theory of molecular evolution.

One test for positive selection that has been extensively applied to genome-wide polymorphism and divergence data is the MK test [6]. This test and related methods can be used to estimate the fraction of substitutions driven by positive selection by comparing observed rates of divergence to those expected based on polymorphism data (Box 1). Application of the MK test to several species has provided evidence that positive selection has driven a large portion of interspecific differences, bringing the neutral theory into question [7–11]. However, other species show little or no evidence that positive selection has driven a large proportion of interspecific differences. The disparate results of the MK test necessitate a careful consideration of what factors can explain these results, and particularly whether estimates of the rate of adaptive evolution are robust to the assumptions of the test. I discuss here one of the assumptions of the MK test that has received relatively little attention: sites evolve independently of one another. Given both theoretical and empirical evidence that this assumption is not valid, I argue that estimates of the frequency of adaptive substitution are too unreliable to reject the neutral theory of molecular evolution.

Evidence for positive selection based on the MK test

Of 38 species for which polymorphism has been compared to divergence across multiple genes, 47% show evidence that positive selection has significantly increased the nonsynonymous substitution rate based on the MK test or related tests (Table 1; Table S1 in the supplementary material online). For those species that do show significant evidence of positive selection, the estimated fraction of nonsynonymous substitutions driven by positive selection is relatively high, often around 40% (Figure 1).

Evidence for positive selection varies among taxonomic groups. The strongest evidence comes from *Drosophila*, where all six species that have been examined show

Glossary

Neutral theory: a theory that the vast majority of DNA substitutions between species is the result of neutral mutations and random drift rather than selectively driven substitutions. The neutral theory does not assert that all mutations are neutral or that there is no adaptive evolution, but instead that deleterious mutations are eliminated from a population and that positively selected mutations make only a small contribution to divergence between species.

Effective population size: the size of a randomly mating population of constant size that would effectively recapitulate patterns and levels of variation observed in a real population. The effective (rather than actual) size of a population is used to account for the fact that most real populations are not constant in size and mating is not entirely random. The effective population size of a species is relevant to estimates of α because weakly deleterious and advantageous mutations are effectively neutral when their fitness effects are less than the reciprocal of the effective size of a population.

Alpha (α): the estimated fraction of substitutions that were driven by positive selection, and for nonsynonymous substitutions can be estimated by $1-(pN/pS)/(dN/dS)$.

Selective constraint: the selective constraint on a site, or the average constraint across multiple sites, is a function of the degree to which fitness is reduced when mutated. In the absence of positive selection, pN/pS and dN/dS provide measures of the selective constraint on a site because they measure the extent to which the rate of nonsynonymous variation is reduced by negative selection relative to synonymous variation (assuming synonymous variation is neutral).

Epistasis: although epistasis is used in many different contexts, here it is used to refer to mutations with effects that depend on genotypes at other loci.

Hitchhiking: the process by which the frequency of mutations linked to an advantageous mutation are influenced by the spread of the advantageous mutation through a population.

Corresponding author: Fay, J.C. (jfay@genetics.wustl.edu).

0168-9525/$ – see front matter © 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2011.06.003 Trends in Genetics, September 2011, Vol. 27, No. 9

Box 1. The McDonald–Kreitman test

The MK test compares rates of DNA polymorphism within species to divergence between species [6]. If all nonsynonymous and synonymous variation is neutral, the ratio of the nonsynonymous to synonymous substitution rate between species (*dN*/*dS*) is expected to be equal to the ratio of nonsynonymous to synonymous polymorphism within species (*pN*/*pS*). However, positive selection can increase *dN*/*dS* above *pN*/*pS* because adaptive mutations spread quickly through a population but have a cumulative effect on divergence. In its original formulation, the MK test determines whether the ratio of nonsynonymous to synonymous fixed differences is significantly different from that of polymorphic sites in a single gene. Subsequent studies used the combined data from multiple genes across the genome to test whether *dN*/*dS* is significantly greater than *pN*/*pS* and to estimate the fraction of nonsynonymous substitutions driven by positive selection, α, from the degree to which *dN*/*dS* is elevated over *pN*/*pS*, or more specifically: $\alpha = 1 - (pN/pS)/(dN/dS)$ [74,75]. An example illustrating how α is estimated is shown in Figure I. A similar logic can be used to estimate the fraction of noncoding substitutions driven by positive selection [76].

A significant obstacle in estimating the frequency of positive selection is accounting for deleterious mutations. Deleterious polymorphism can increase *pN*/*pS* above *dN*/*dS*, even in the presence of positive selection [77]. The effect of deleterious polymorphism on *pN*/*pS* can be enhanced by a recent reduction in population size [78]. A number of methods have been developed to account for segregating deleterious polymorphism so as to estimate more accurately the frequency of positive selection [16,19,22,65,75]. However, deleterious mutations can also become fixed during historic periods of reduced population size, leading to an increase in *dN*/*dS* and spurious signals of positive selection [40,75].

	Nonsynonymous (N)	Synonymous (S)	N/S
Polymorphism	6	25	0.24
Divergence	16	40	0.40

$$\alpha = 1 - (pN/pS)/(dN/dS) = 1 - (0.24)/(0.40) = 0.40$$

TRENDS in Genetics

Figure I. An estimate of the fraction of nonsynonymous substitutions driven by positive selection (α) using the McDonald-Kreitman test. The MK test determines whether the ratio of nonsynonymous to synonymous polymorphisms (N/S) is significantly different from the same ratio obtained from divergence data using Fisher's exact test or a comparable test statistic. In this example, N/S of divergence is twice as great as that of polymorphism and leads to an estimate that 40% of nonsynonymous substitutions were driven by positive selection. Note that α can be estimated from the number of changes or rates of change because rates are measured by the number of changes divided by the number of sites surveyed, and these must be the same for polymorphism and divergence data.

evidence of positive selection. One of the six species, *Drosophila miranda*, has conflicting reports of positive selection, and this could be a consequence of low levels of polymorphism or a dependency on which genes were studied [12–16]. In vertebrates, plants, bacteria and fungi, some species show strong evidence of positive selection and others show weak or no evidence. In vertebrates there is evidence of positive selection in chickens [17] and mice [18], but evidence in humans is weak or absent [19–22]. The lack of evidence in humans could be a consequence of an obscured signal due to a recent reduction in effective population size [22] (Box 1). In bacteria, *Escherichia coli* and *Salmonella enterica* show strong evidence of positive selection [23]. Even so, the ratio of the nonsynonymous to synonymous substitution rate (*dN*/*dS*) is greater than that of the rate of polymorphism (*pN*/*pS*) for only four of 12 other bacterial species [24]. Note that the authors of the latter study interpreted their results as most probably being a consequence of purifying selection. In plants there is evidence of positive selection in four out of 13 species. Interestingly, the species showing evidence of positive selection are also those estimated to have the largest effective population size [25–27], a trend that is also reflected in other taxa [8]. Under some but not all models of molecular evolution [28], species with large effective population sizes are expected to have higher rates of adaptive substitution, and signals of positive selection are less likely to be obscured by deleterious polymorphism [29]. A valuable feature of the plant research is that evidence of positive selection does not appear to be related to the mating system because both the predominantly selfing species, *Arabidopsis thaliana*, and its outcrossing relative, *Arabidopsis lyrata*, show no evidence of positive selection [30]. Finally, two yeast species show no evidence of positive selection [31–33].

Table 1. Reported evidence of positive selection based on the MK test

Group	Species	Evidence of positive selection[a] (%)	Refs.
Fungi	2	0 (0%)	[33]
Plants[b]	13	4 (31%)	[25–27,30,79]
Bacteria	14	6 (43%)	[23,24]
Vertebrates	3	2 (67%)	[17,18,22]
Insects[b]	6	6 (100%)	[13,15,37,80–82]
All	38	18 (47%)	

[a]Evidence for positive selection is based on a significant excess of nonsynonymous divergence based on the MK test.

[b]There are both positive and negative reports of positive selection for some species.

The neutralist and selectionist interpretations

As is often the case, there are both neutral and selective interpretations of the mixed evidence for positive selection based on the MK test. The selectionist view is that positive selection is indeed pervasive, at least in some species, and that the absence of evidence in other species is a consequence of an obscured signal of selection or a lower rate of adaptive evolution. In humans and multiple plant species, the absence of a strong signal of selection could be a consequence of their small effective population sizes [7,8,10,11,18,22,25,26,29]. However, the absence of evidence in other species, particularly bacteria and yeast species, is

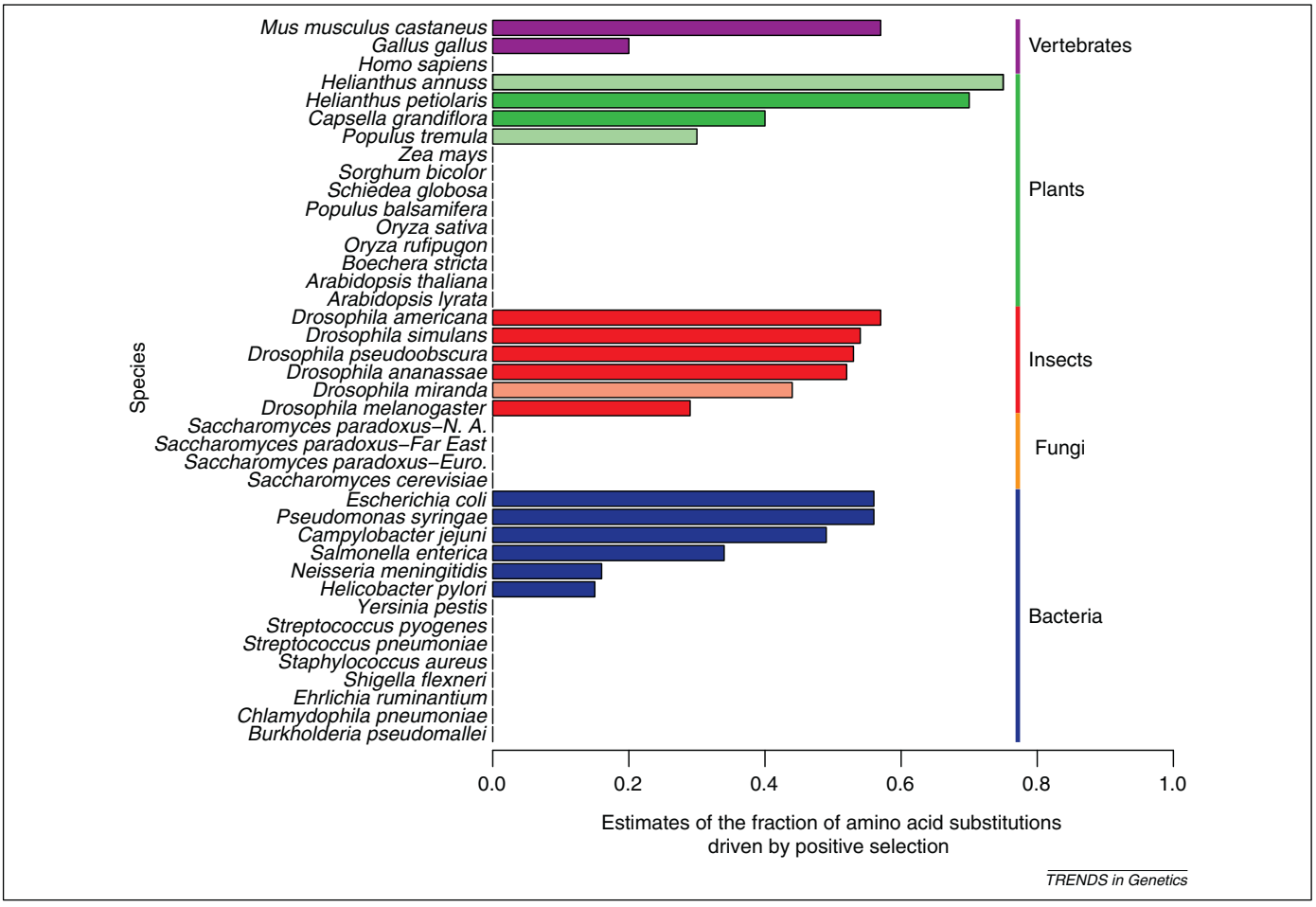

Figure 1. Estimates of the fraction of nonsynonymous substitutions driven by positive selection (α) grouped by taxa. Estimates of α are from published reports (Table S1) using a variety of different methods based on the MK test. Estimates of α are only shown for species with significant MK test results, with the exception of 12 bacterial species [24] for which α was not estimated in the original publication and for which the equation in Box 1 was used. Species with conflicting reports of positive selection are shown in lighter shades. Separate results are shown for subpopulations of *Saccharomyces paradoxus* because they are monophyletic and show evidence of reproductive isolation [82]. Together, these results indicate that there is substantial variation in the evidence for positive selection based on the MK test.

not easy to explain because they have large effective population sizes [34,35]. One potential explanation for yeast and bacteria species is that they could have a greater propensity for local adaptation, in which case pN/pS might be inflated above dN/dS even when positive selection has had a significant impact on dN/dS [8].

The neutralist view is that positive selection is rare and that factors other than positive selection generate significant MK test results. A reduction in population size can lead to fixation of deleterious mutations and spurious evidence of positive selection (Box 1). Although changes in population size could be responsible for spurious signals of positive selection in some species, it seems unlikely that all of the *Drosophila* species experienced a demographic history that would incur a similar false signal of selection. Positive selection could also be overestimated due to statistical biases related to the assumption that synonymous and nonsynonymous sites share the same genealogical history [36] and the summation over sparse contingency tables [37,38]. However, neither of these statistical considerations appear to account for the *Drosophila* data. Thus, positive selection is often [7–11] but not always [24,39] the preferred interpretation for significant MK test results.

In weighing the evidence for positive selection it is important to consider the many assumptions of the MK test [6,40]. One assumption which has received much less attention than others is that sites evolve independently of one another. Although non-independence among sites might not always generate false signals of selection, two scenarios are discussed below where non-independence could lead to overestimates of the frequency of positive selection. First, epistasis can result in changes in selective constraint. Second, positive selection can increase the substitution rate of linked deleterious mutations through hitchhiking.

Epistasis can result in changes in selective constraint

The MK test assumes that the selective constraints on a sequence remain constant over time. The selective constraint on a site is determined by the degree to which fitness is reduced when mutated, and can be measured by the reduction in the substitution rate relative to a neutral site, for example by dN relative to dS. In the absence of any fitness effect, a site is unconstrained and will evolve neutrally. Changes in effective population size influence selective constraints because populations have trouble removing deleterious mutations with small fitness effects

Box 2. Changes in selective constraint generated by epistasis

Given a non-linear phenotype–fitness function, an allele can have a positive, negative or no fitness effect depending on its genetic background. Figure I shows a simple example of a phenotype–fitness function where a mutation that causes a slight increase in the phenotype of an individual (e.g. height) is expected to increase fitness within individuals who have intermediate phenotypes but to have little or no effect on fitness within individuals who have extreme phenotypes. In this context, maximizing population fitness results in a loss of constraint and evolution of neutrality [42]. Note that if the phenotype–fitness function is bell-shaped a mutation can have positive or negative effects depending on the phenotype of the individual in which it occurs. In models that allow for epistatic fitness effects [68], species can become fixed for alternative genotypes without a change in fitness, resulting in divergence between species at sites under negative selection within species. Figure I shows an example where four out of six sites have diverged between species despite the presence of purifying selection on four of the six sites within species.

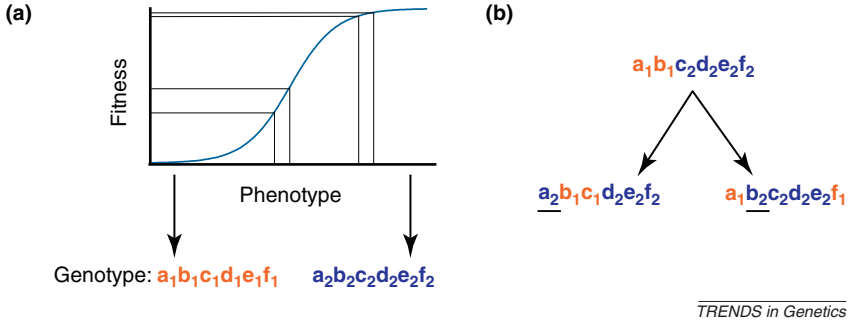

Figure I. Epistatic fitness effects under a simple evolutionary model. **(a)** Schematic of a phenotype–fitness function that generates epistatic fitness effects. Lines in the graph show how the expected fitness effect of a mutation depends on the phenotype of an individual. This is assumed be an additive function of multiple loci, six are shown for illustrative purposes. Genotypes at six loci **(a–f)** are shown for two extreme phenotypes below the graph. **(b)** A representation of species divergence that can arise as a consequence of the model shown in **(a)**. As species diverge, their genotypes can diverge via drift through neutral phenotype-space. This results in sites (underlined) that have undergone neutral divergence between species but that are subject to purifying selection within species.

[2,40]. However, the selective constraint on a site can also depend on epistatic interactions with other sites in the genome.

Epistasis is relevant to the evolution of a site whenever the fitness effect of a mutation depends on other genetic or environmental variation. In the case of genetic dependencies, epistasis is defined by fitness effects at one site that depend on the genotype present at other sites. Epistatic fitness effects can result from non-additive relationships between genotype and phenotype, or phenotype and fitness. Epistatic fitness effects could be common because they are expected to occur as a general consequence of any non-linear phenotype–fitness function, even under an additive genetic model [41] (Box 2). Several models of molecular evolution include epistatic fitness effects so as to account for the possibility that selective constraints are not static but can change over time [42,43]. Although models that incorporate epistasis have been investigated in the context of substitution rate heterogeneity [44], their predictions in the context of the MK test are at present unknown. In addition to changes in selective constraint, epistatic interactions between a mutation and its genetic background can also result in a mutation having positive effects in some individuals but negative effects in others (Box 2). This complicates the classification of evolutionary models with epistasis into the neutral regime because it appears unlikely that mutations with epistatic fitness effects would only cause changes in selective constraint without also being able to cause previously deleterious or neutral mutations to become advantageous. Interestingly, rapid temporal fluctuations between positive and negative selection can result in significant MK test results because the dynamics of sites under fluctuating selection are determined by their effects on the variance in fitness rather than the mean, which is assumed to be zero [45,46].

Empirical evidence of temporal changes in selective constraint

There is an abundance of evidence that the effects of a mutation depend on its genetic background. Empirical evidence of epistasis is frequently encountered in quantitative trait mapping [47], experimental evolution [48,49], reconstruction of substitutions that have occurred during evolution [50,51], and patterns of molecular evolution [52,53]. In many cases, large epistatic effects can be inferred from the strongly deleterious or even lethal effects of a mutation in one genetic background relative to the absence of any noticeable effects in another background [54,55]. However, the mere presence of epistasis does not imply that the frequency of positive selection has been overestimated by the MK test.

For epistasis to cause spurious signals of positive selection dN/dS must be consistently greater than pN/pS across the genome, such that a large number of sites are constrained within a species but were less constrained or unconstrained in the past. However, without an explicit model for how epistasis generates changes in constraint over time, it is hard to know whether there are also a large number of neutral sites within a species that were constrained during divergence between species.

Two observations suggest that selective constraint can systematically change over time. First, duplicated genes exhibit low levels of constraint immediately after gene duplication followed by increased levels of constraint over time, as measured by a decrease in dN/dS over time [56]. Based on this empirical pattern, pN/pS could be consistently lower than dN/dS for duplicated genes because the

high rate of divergence following gene duplication is transient and would therefore contribute to divergence between species but not to polymorphism within species. This pattern of temporal changes in selective constraint can occur as a consequence of subfunctionalization of duplicate genes [57], but can also occur as a consequence of stabilizing selection on other types of redundant genetic elements such as transcription factor binding sites within enhancers [58].

A second observation relevant to temporal changes in selective constraint is that, in both bacteria and vertebrates, *dN*/*dS* of all genes decreases as a function of the time-period over which it is measured [59,60]. Explanations for this pattern include fixation of deleterious polymorphism in ancestral populations [61] and statistical biases of the methods used to estimate *dN*/*dS* [60]. Regardless of the cause, the pattern implies that *dN*/*dS* or estimates of *dN*/*dS* are higher for closely related species compared to distantly related species. Interestingly, there is no evidence of positive selection for many of the species with high rates of divergence to their outgroup, for example bacterial and yeast species (Table S1). Nevertheless, without knowing why *dN*/*dS* changes over time, it is hard to know how *pN*/*pS* would be affected relative to *dN*/*dS*.

In summary, even though epistasis can result in changes in selective constraint, its impact on the MK test has yet to be determined. Gene duplication provides a concrete example of how epistatic fitness effects can result in changes in constraint over time. Although the genome-wide changes in *dN*/*dS* might not be related to epistasis [60], they suggests that the elevation of *dN*/*dS* over *pN*/*pS* could be an unreliable estimator of the frequency of positive selection.

Hitchhiking of deleterious mutations

The MK test assumes that any increase in *dN*/*dS* above *pN*/*pS* is due to positive selection. Although weakly deleterious mutations can become polymorphic in a population and reach fixation, they are expected to increase *pN*/*pS* more than *dN*/*dS* [40,62], assuming selective constraints have not changed and that there are no effects of selection at linked sites. However, when the effects of a deleterious mutation do not outweigh the effects of an advantageous mutation, deleterious mutations can be rapidly fixed due to hitchhiking with a linked advantageous mutation. As a consequence, positive selection can result in a dramatic increase in the deleterious substitution rate [63,64]. Although hitchhiking of deleterious mutations requires the presence of positive selection, a single strong hitchhiking event could result in the fixation of multiple slightly deleterious mutations. Without accounting for this possibility, it is hard to know the degree to which hitchhiking of deleterious mutations causes overestimation of the frequency of positive selection based on the MK test.

Weighing the evidence for adaptation based on the MK test

Application of the MK test to DNA polymorphism and divergence data has revealed two features regarding evidence of positive selection: heterogeneity across species [7,8,10] and uniformity among genes within a species [23,65,66]. Neither neutral nor selective models have yet provided compelling explanations for these observations. Although deleterious polymorphism could obscure evidence of positive selection in species with small population sizes, this cannot explain the absence of evidence for selection in species with large effective population sizes. Furthermore, it seems unlikely that positive selection would have a uniform impact across genes, unless positive selection is driven by the continual fixation of deleterious mutations [67]. By comparison, neutral models that invoke population bottlenecks could explain results from some species but do not provide a compelling explanation for the *Drosophila* data. Given the vexing patterns of heterogeneous MK test results, I argue that other factors known to affect the MK test must be considered when estimating the fraction of substitutions driven by positive selection.

One of the factors that could have a considerable influence on the MK test is the assumption that sites evolve independently of one another. Given both theoretical considerations and empirical support that sites do not evolve independently of one another [42–44,47–55,57,58,63,64,68], I suggest that estimates of positive selection based on the MK test are unreliable. Epistatic fitness effects provide one mechanism that can lead to non-independence among sites and result in changes in selective constraint over time. Any case where a site is constrained within a species but was free to diverge between species will produce spurious evidence of positive selection based on the MK test. Purifying selection operating on redundant genetic elements, such as duplicated genes or transcription factor binding sites, provides a potential but as yet unexplored explanation for positive MK test results. Furthermore, the observation that constraint or measurement of constraint consistently changes as a function of the time interval over which it is measured [59,60] raises questions over the reliability of using the elevation of *dN*/*dS* over *pN*/*pS* to estimate the frequency of positive selection. In addition to epistasis, hitchhiking also causes non-independence among sites. Hitchhiking could lead to overestimation of the frequency of positive selection due to the fixation of weakly deleterious mutations linked to a strongly advantageous mutation. Although fixation of deleterious mutations creates the opportunity for further adaptive substitutions, it makes it difficult to estimate reliably the frequency of positive selection based on the MK test.

Despite the concerns regarding the interpretation of the MK test results, many other patterns of molecular evolution are relevant to understanding how positive selection shapes genome evolution and these often point to a pervasive influence of positive selection [7,9,10]. The reduced levels of synonymous polymorphism surrounding recently fixed nonsynonymous substitutions is just one pattern that supports a dominant influence of positive selection [13,27,69–72]. Although this pattern is also expected to occur as a result of background selection, it is best explained by positive selection due to a number of patterns that are not consistent with a background selection model [13,27,69–72]. However, disentangling the effects of positive and negative selection, especially in the context of hitchhiking of deleterious mutations [71], makes it difficult

to estimate the fraction of substitutions driven by positive selection based only on patterns of hitchhiking.

Concluding remarks

Even if non-independence among sites limits our ability to distinguish between positive and negative selection based on significant MK test results, it is worth investigating the theoretical consequences and empirical evidence for more realistic models of molecular evolution that include both positive and negative selection and that do not assume sites evolve independently of one another. In this context, it might be less relevant to distinguish between models of positive and negative selection and more relevant to distinguish models of purifying selection from those of adaptive evolution, where there is a positive fitness flux [73], or where there is a change in phenotype. Regardless of the model of selection, non-independence among sites must be investigated before we conclude that positive selection is a pervasive feature of molecular evolution and dismiss the neutral theory of molecular evolution.

Acknowledgments

I would like to acknowledge participants of the Kavli Institute for Theoretic Physics conference in Population Genetics and Genomics for discussions, members of the Fay laboratory for comments, and multiple anonymous reviewers whose comments on this and a previous version of this work were extremely helpful in formulating my views in relation to the views of others in the field. This work was supported by the National Institute of General Medical Sciences (GM080669 and GM086412).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tig.2011.06.003.

References

1 Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*, Cambridge University Press
2 Ohta, T. (1973) Slightly deleterious mutant substitutions in evolution. *Nature* 246, 96–98
3 Gillespie, J.H. (1991) *The Causes of Molecular Evolution*, Oxford University Press
4 Kreitman, M. and Akashi, H. (1995) Molecular evidence for natural selection. *Annu. Rev Ecol. Syst.* 26, 403–422
5 Fay, J. and Wu, C-I. (2003) Sequence divergence, functional constraint, and selection in protein evolution. *Annu. Rev. Genomics Hum. Genet.* 4, 213–235
6 McDonald, J. and Kreitman, M. (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351, 652–654
7 Sella, G. *et al.* (2009) Pervasive natural selection in the *Drosophila* genome? *PLoS Genet.* 5, e1000495
8 Siol, M. *et al.* (2010) The population genomics of plant adaptation. *New Phytol.* 188, 313–332
9 Hahn, M.W. (2008) Toward a selection theory of molecular evolution. *Evolution* 62, 255–265
10 Wright, S. and Andolfatto, P. (2008) The impact of natural selection on the genome: emerging patterns in *Drosophila* and *Arabidopsis*. *Annu. Rev. Ecol. Syst.* 39, 193–213
11 Eyre-Walker, A. (2006) The genomic rate of adaptive evolution. *Trends Ecol. Evol.* 21, 569–575
12 Bachtrog, D. and Andolfatto, P. (2006) Selection, recombination and demographic history in *Drosophila miranda*. *Genetics* 174, 2045–2059
13 Bachtrog, D. (2008) Similar rates of protein adaptation in *Drosophila miranda* and *D. melanogaster*, two species with different current effective population sizes. *BMC Evol. Biol.* 8, 334
14 Bartolomé, C. *et al.* (2005) Patterns of selection on synonymous and nonsynonymous variants in *Drosophila miranda*. *Genetics* 169, 1495–1507
15 Haddrill, P.R. *et al.* (2010) Estimating the parameters of selection on nonsynonymous mutations in *Drosophila pseudoobscura* and *D. miranda*. *Genetics* 185, 1381–1396
16 Loewe, L. and Charlesworth, B. (2006) Inferring the distribution of mutational effects on fitness in *Drosophila*. *Biol. Lett.* 2, 426–430
17 Axelsson, E. and Ellegren, H. (2009) Quantification of adaptive evolution of genes expressed in avian brain and the population size effect on the efficacy of selection. *Mol. Biol. Evol.* 26, 1073–1079
18 Halligan, D.L. *et al.* (2010) Evidence for pervasive adaptive protein evolution in wild mice. *PLoS Genet.* 6, e1000825
19 Boyko, A.R. *et al.* (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* 4, e1000083
20 Gojobori, J. *et al.* (2007) Adaptive evolution in humans revealed by the negative correlation between the polymorphism and fixation phases of evolution. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3907–3912
21 Zhang, L. and Li, W.H. (2005) Human SNPs reveal no evidence of frequent positive selection. *Mol. Biol. Evol.* 22, 2504–2507
22 Eyre-Walker, A. and Keightley, P.D. (2009) Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Mol. Biol. Evol.* 26, 2097–2108
23 Charlesworth, J. and Eyre-Walker, A. (2006) The rate of adaptive evolution in enteric bacteria. *Mol. Biol. Evol.* 23, 1348–1356
24 Hughes, A.L. *et al.* (2008) Synonymous and nonsynonymous polymorphisms versus divergences in bacterial genomes. *Mol. Biol. Evol.* 25, 2199–2209
25 Gossmann, T.I. *et al.* (2010) Genome wide analyses reveal little evidence for adaptive evolution in many plant species. *Mol. Biol. Evol.* 27, 1822–1832
26 Slotte, T. *et al.* (2010) Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. *Mol. Biol. Evol.* 27, 1813–1821
27 Ingvarsson, P.K. (2010) Natural selection on synonymous and nonsynonymous mutations shapes patterns of polymorphism in *Populus tremula*. *Mol. Biol. Evol.* 27, 650–660
28 Gillespie, J.H. (2001) Is the population size of a species relevant to its evolution? *Evolution* 55, 2161–2169
29 Ellegren, H. (2009) A selection model of molecular evolution incorporating the effective population size. *Evolution* 63, 301–305
30 Foxe, J.P. *et al.* (2008) Selection on amino acid substitutions in *Arabidopsis*. *Mol. Biol. Evol.* 25, 1375–1383
31 Doniger, S.W. *et al.* (2008) A catalog of neutral and deleterious polymorphism in yeast. *PLoS Genet.* 4, e1000183
32 Liti, G. *et al.* (2009) Population genomics of domestic and wild yeasts. *Nature* 458, 337–341
33 Elyashiv, E. *et al.* (2010) Shifts in the intensity of purifying selection: an analysis of genome-wide polymorphism data from two closely related yeast species. *Genome Res.* 20, 1558–1573
34 Lynch, M. and Conery, J.S. (2003) The origins of genome complexity. *Science* 302, 1401–1404
35 Tsai, I.J. *et al.* (2008) Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4957–4962
36 Andolfatto, P. (2008) Controlling type-I error of the McDonald–Kreitman test in genomewide scans for selection on noncoding DNA. *Genetics* 180, 1767–1771
37 Shapiro, J.A. *et al.* (2007) Adaptive genic evolution in the *Drosophila* genomes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2271–2276
38 Stoletzki, N. and Eyre-Walker, A. (2011) Estimation of the neutrality index. *Mol. Biol. Evol.* 28, 63–70
39 Nei, M. *et al.* (2010) The neutral theory of molecular evolution in the genomic era. *Annu. Rev. Genomics Hum. Genet.* 11, 265–289
40 Fay, J. and Wu, C-I. (2001) The neutral theory in the genomic era. *Curr. Opin. Genet. Dev.* 11, 642–646
41 Brodie, E.D., 3rd (2000) Why evolutionary genetics does not always add up. In *Epistasis and the Evolutionary Process* (Wolf, J., ed.), pp. 3–19, Oxford University Press
42 Hartl, D.L. *et al.* (1985) Limits of adaptation: the evolution of selective neutrality. *Genetics* 111, 655–674
43 Takahata, N. (1991) Statistical models of the overdispersed molecular clock. *Theor. Popul. Biol.* 39, 329–344
44 Cutler, D.J. (2000) Understanding the overdispersed molecular clock. *Genetics* 154, 1403–1417

45 Huerta-Sanchez, E. *et al.* (2008) Population genetics of polymorphism and divergence under fluctuating selection. *Genetics* 178, 325–337
46 Mustonen, V. and Lässig, M. (2007) Adaptations to fluctuating selection in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2277–2282
47 Malmberg, R.L. and Mauricio, R. (2005) QTL-based evidence for the role of epistasis in evolution. *Genet. Res.* 86, 89–95
48 Sanjuán, R. *et al.* (2004) The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15376–15379
49 Pepin, K.M. and Wichman, H.A. (2007) Variable epistatic effects between mutations at host recognition sites in phiX174 bacteriophage. *Evolution* 61, 1710–1724
50 Weinreich, D.M. *et al.* (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312, 111–114
51 Lunzer, M. *et al.* (2010) Pervasive cryptic epistasis in molecular evolution. *PLoS Genet.* 6, e1001162
52 Huelsenbeck, J. (2002) Testing a covariotide model of DNA substitution. *Mol. Biol. Evol.* 19, 698–707
53 Callahan, B. *et al.* (2011) Correlated evolution of nearby residues in Drosophilid proteins. *PLoS Genet.* 7, e1001315
54 Kondrashov, A. *et al.* (2002) Dobzhansky–Muller incompatibilities in protein evolution. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14878–14883
55 Dowell, R.D. *et al.* (2010) Genotype to phenotype: a complex problem. *Science* 328, 469
56 Lynch, M. and Conery, J.S. (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151–1155
57 Force, A. *et al.* (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545
58 Bullaughey, K. (2011) Changes in selective effects over time facilitate turnover of enhancer sequences. *Genetics* 187, 567–582
59 Rocha, E.P.C. *et al.* (2006) Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J. Theor. Biol.* 239, 226–235
60 Wolf, J.B.W. *et al.* (2009) Nonlinear dynamics of nonsynonymous (dN) and synonymous (dS) substitution rates affects inference of selection. *Genome Biol. Evol.* 1, 308–319
61 Peterson, G.I. and Masel, J. (2009) Quantitative prediction of molecular clock and ka/ks at short timescales. *Mol. Biol. Evol.* 26, 2595–2603
62 Fay, J. *et al.* (2001) Positive and negative selection on the human genome. *Genetics* 158, 1227–1234
63 Bachtrog, D. and Gordo, I. (2004) Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* 58, 1403–1413
64 Hartfield, M. and Otto, S. (2011) Recombination and hitchhiking of deleterious alleles. *Evolution* DOI: 10.1111/j.1558-5646.2011.01311.x
65 Bierne, N. and Eyre-Walker, A. (2004) The genomic rate of adaptive amino acid substitution in *Drosophila*. *Mol. Biol. Evol.* 21, 1350–1360
66 Welch, J.J. (2006) Estimating the genomewide rate of adaptive protein evolution in *Drosophila*. *Genetics* 173, 821–837
67 Charlesworth, J. and Eyre-Walker, A. (2007) The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16992–16997
68 Gavrilets, S. (1997) Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* 12, 307–312
69 Sattath, S. *et al.* (2011) Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila simulans*. *PLoS Genet.* 7, e1001302
70 Macpherson, J.M. *et al.* (2007) Genomewide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in *Drosophila*. *Genetics* 177, 2083–2099
71 Andolfatto, P. (2007) Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genome Res.* 17, 1755–1762
72 Cai, J.J. *et al.* (2009) Pervasive hitchhiking at coding and regulatory sites in humans. *PLoS Genet.* 5, e1000336
73 Mustonen, V. and Lässig, M. (2009) From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* 25, 111–119
74 Smith, N. and Eyre-Walker, A. (2002) Adaptive protein evolution in *Drosophila*. *Nature* 415, 1022–1024
75 Fay, J. *et al.* (2002) Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. *Nature* 415, 1024–1026
76 Andolfatto, P. (2005) Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 437, 1149–1152
77 Charlesworth, J. and Eyre-Walker, A. (2008) The McDonald-Kreitman test and slightly deleterious mutations. *Mol. Biol. Evol.* 25, 1007–1015
78 Eyre-Walker, A. (2002) Changing effective population size and the McDonald–Kreitman test. *Genetics* 162, 2017–2024
79 Begun, D.J. *et al.* (2007) Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* 5, e310
80 Maside, X. and Charlesworth, B. (2007) Patterns of molecular variation and evolution in *Drosophila americana* and its relatives. *Genetics* 176, 2293–2305
81 Grath, S. *et al.* (2009) Molecular evolution of sex-biased genes in the *Drosophila ananassae* subgroup. *BMC Evol. Biol.* 9, 291
82 Sniegowski, P. *et al.* (2002) *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. *FEM Yeast Res.* 1, 299–306