

## REVIEW

# Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level

AL Hughes

*Department of Biological Sciences, University of South Carolina, Columbia, SC, USA*

Recent years have seen an explosion of interest in evidence for positive Darwinian selection at the molecular level. This quest has been hampered by the use of statistical methods that fail adequately to rule out alternative hypotheses, particularly the relaxation of purifying selection and the effects of population bottlenecks, during which the effectiveness of purifying selection is reduced. A further problem has been the assumption that positive selection will generally involve repeated amino-acid changes to a single protein. This model was derived from the case of the vertebrate major histocompatibility complex (MHC), but

the MHC proteins are unusual in being involved in protein–protein recognition and in a co-evolutionary process of pathogens. There is no reason to suppose that repeated amino-acid changes to a single protein are involved in selectively advantageous phenotypes in general. Rather adaptive phenotypes are much more likely to result from other causes, including single amino-acid changes; deletion or silencing of genes or changes in the pattern of gene expression.

*Heredity* (2007) **99**, 364–373; doi:10.1038/sj.hdy.6801031; published online 11 July 2007

**Keywords:** positive selection; nearly neutral theory; nucleotide substitution

## Introduction

Natural selection has long been one of the central organizing principles of biology; and in recent decades, it has become possible to study natural selection at the most fundamental level, that of the genome (Hughes, 1999). Evolutionary biologists typically distinguish two main types of natural selection: (1) purifying selection, which acts to eliminate deleterious mutations and (2) positive (Darwinian) selection, which favors advantageous mutations. Positive selection can, in turn, be further subdivided into directional selection, which tends toward fixation of an advantageous allele, and balancing selection, which maintains a polymorphism. The neutral theory of molecular evolution (Kimura, 1983) predicts that purifying selection is ubiquitous, but that both forms of positive selection are rare, whereas not denying the importance of positive selection in the origin of adaptations.

By means of nucleotide sequencing, we now have access to new kinds of evidence regarding the past and ongoing action of natural selection. Unfortunately, the task of deciphering this evidence has been hampered by the widespread use of inappropriate statistical methods and by a flawed philosophical framework, inherited from the past, which inhibits a true understanding of the nature and action of natural selection. The purpose of this paper is to provide a review of the current state of

research on natural selection at the molecular level with emphasis on the faulty methods and unwarranted assumptions that currently impede progress in our understanding of evolution.

In particular, I examine the recent emphasis on positive Darwinian selection favoring repeated amino-acid changes at a limited set of sites in a given protein. This type of selection at the molecular level is the most easily studied statistically, because one can make use of comparisons between synonymous and nonsynonymous substitutions; and there are a few well-studied examples of this type of selection, notably the genes of the vertebrate major histocompatibility complex (MHC; Hughes and Nei, 1988, 1989). However, I argue that this type of selection is likely to be very rare. Moreover, because this type of selection occurs mainly in molecules involved in protein–protein recognition and typically involves a co-evolutionary process, it is unlikely to be involved in the evolution of major morphological and developmental adaptations.

## Historical background

### Neo-Darwinism

The modern concept of natural selection dates from the Neo-Darwinian synthesis of the 1920–1930's, when the original insight of Darwin and Wallace was combined with Mendelian genetics to model evolution as the change in gene frequencies in populations. The importance of these pioneering studies to evolutionary biology should not be minimized, but at the same time, it is necessary to realize that there were limitations to the original Neo-Darwinists' understanding of the

Correspondence: Dr AL Hughes, Department of Biological Sciences, University of South Carolina, Coker Life Sciences Bldg, 700 Sumter St, Columbia, SC 29208, USA.

E-mail: austin@biol.sc.edu

Received 10 April 2007; revised 4 June 2007; accepted 6 June 2007; published online 11 July 2007

evolutionary process. The original Neo-Darwinists knew nothing about the physical nature of the gene and little of how genes actually affect phenotypic traits. Also, at that time, there was little knowledge regarding such simple matters as the census numbers of natural populations and the frequency of population bottlenecks.

As a result of these gaps in knowledge, certain unrealistic views became part and parcel of the Neo-Darwinist account of natural selection, and unfortunately, many of these are still with us, at least implicitly. The following are three major areas of misconception among the Neo-Darwinists:

- (1) It was assumed that most natural populations are very large. Because of this assumption, when Sewall Wright first proposed an important role for genetic drift (that is, changes in allele frequency due to chance effects in a finite population), his models were dismissed as rarely applicable to the real world (Fisher and Ford, 1950). Neo-Darwinist population genetics relied on deterministic models that assumed infinite population sizes; and this assumption was justified by the claim (supported by no real evidence) that most natural populations are very large. Dismissing the importance of genetic drift, the Neo-Darwinists argued that most polymorphisms are maintained by balancing selection. Moreover, it was believed that any beneficial mutation, no matter how slight the selective advantage it conferred, would eventually reach fixation.
- (2) Artificial selection on quantitative traits was taken as a model of the evolutionary process. It was easily shown, in agriculture or in the laboratory, that populations of most organisms contain sufficient additive genetic variance to obtain a response to selection on quantitative traits, such as measures of body size or increased yield of agriculturally valuable products such as milk in dairy cattle or grain size in food plants. Generalizing from this experience, it was assumed that natural populations are endowed with essentially unlimited additive genetic variance, implying that any sort of selection imposed by environmental changes will encounter abundant genetic variation on which to act. Moreover, this model was extended to evolutionary time as well as ecological time. This way of thinking ignored the substantial evidence from selection experiments that the response to selection on any trait essentially comes to a halt after a number of generations as the genetic variance for the trait in question is depleted; thereafter, further progress depends on the introduction of new variants either through outcrossing or new mutations (Falconer, 1981).
- (3) There was a tendency to denigrate the importance of mutation in the evolutionary process. Darwin himself was aware of 'sports,' in domestic animals, which we would call mutations. But in his later works, Darwin dismissed the importance of 'sports' in the evolutionary process in favor of his bizarre, essentially Lamarckian theory of pangenesis (Larson, 2004). According to Darwin's theory, the environment itself caused changes to the germ plasm, ensuring that the appropriate selectable variation would always be present. Although the Neo-Darwinists rejected the Lamarckian aspects of Darwin's

theory, certain presuppositions lingered, the most obvious being the assumption mentioned above that the selectable variation in natural populations is essentially unlimited. In addition, evolutionary biologists retained Darwin's 'gradualism'. Thus, it was imagined that phenotypes change in a continuous fashion over evolutionary time, by small almost imperceptible increments rather than in a saltational fashion.

#### The neutral theory of molecular evolution

The dawn of the molecular era in biology also saw the first serious challenges to the Neo-Darwinist worldview, in Motoo Kimura's neutral theory of molecular evolution (Kimura, 1968, 1983). Before proposing the neutral theory, Kimura had devoted over a decade to the study of evolutionary dynamics in finite populations, a study for which he adapted mathematical tools (such as the diffusion approximation) that were new to population genetics (Kimura, 1955, 1957, 1964). In developing a sophisticated understanding of the role of population size in the evolutionary change of gene frequencies, Kimura made a contribution to evolutionary biology that is arguably second only to Darwin's. However, many population geneticists viewed Kimura's work (if they understood it at all) as a mere mathematical curiosity, since they were wedded to the Neo-Darwinist assumption of infinite population size.

In the neutral theory, Kimura (1968) advanced the radical hypothesis that most evolutionary changes at the molecular level are due to the chance fixation of selectively neutral mutations. This hypothesis stimulated vigorous debates in the 1970s between 'selectionists' and 'neutralists.' In these debates, the selectionist camp often seriously misunderstood Kimura's theory and thus aimed their rhetorical fire at straw men. Many of these misunderstandings persist in the biological literature to this day.

The neutral theory predicts both (1) that most polymorphisms are selectively neutral and are maintained by genetic drift; and (2) that most changes at the molecular level that are fixed over evolutionary time are selectively neutral and are fixed by drift. Thus, the neutral theory provides a conceptual framework uniting ecological and evolutionary time frames. It is often stated even today that the neutral theory predicts that most mutations are selectively neutral. But this is not a prediction of the neutral theory. The neutral theory predicts that the majority of mutations that are fixed over evolutionary time are selectively neutral. When the neutral theory was first proposed, the extent of noncoding DNA in the genomes of eukaryotes was not known. Now, given the evidence that a substantial majority of the nucleotides in a typical mammalian genome may be nonfunctional, we may hypothesize that most mutations occurring in nonfunctional regions are selectively neutral. And, given the abundance of nonfunctional regions, it follows that a majority of mutations in such genomes are probably selectively neutral. But this is not a consequence of the neutral theory *per se*.

In fact, as regard to the coding regions, the neutral theory predicts that most mutations are not selectively neutral. Rather, because most mutations in coding regions are nonsynonymous (amino-acid-altering) and

thus disrupt protein structure, most mutations in coding regions are selectively deleterious. One of the most important predictions of the neutral theory is thus that purifying selection will predominate in coding regions (and in other functionally important regions as well).

It is by now such a truism of molecular biology that functionally important sequences are 'conserved' or 'functionally constrained' (that is, evolve slowly), that many biologists probably do not realize that this generalization is a prediction of the neutral theory. Moreover, few remember that the selectionists made the opposite prediction; that the most functionally important regions of proteins should evolve rapidly. Some of the most basic techniques of bioinformatics depend on the fact that the neutralists were right in this case. Homology searches and sequence alignments depend on the fact that functionally important sequences are conserved over evolutionary time. If the selectionists had been right, these everyday tools of modern biology would be impossible.

Kimura's associate, Tomoko Ohta, developed certain aspects of the neutral theory in the so-called 'nearly neutral theory' (Ohta, 1973, 1976). It is worth emphasizing that the nearly neutral theory is not an alternative theory, competing with the neutral theory. Rather, the nearly neutral theory is really a corollary of the neutral theory that focuses on the issue of slightly deleterious mutations. An important prediction of the neutral theory is that, when the selection coefficient in favor of an advantageous mutant or against a deleterious mutant is less than the reciprocal of twice the effective population size, that mutant becomes effectively neutral and is not exposed to selection. Thus, in a species with small effective population size, slightly deleterious mutations can increase in frequency as a result of genetic drift. The same would of course be true of slightly advantageous mutations, but the latter are expected to be less common than slightly deleterious mutations. Thus, one characteristic of a species that undergoes an extended population bottleneck will be an increase in frequency of slightly deleterious alleles, leading to fixation of some such alleles. If the effective population size increases after a population bottleneck, natural selection again becomes efficient in removing slightly deleterious alleles.

There is evidence that such a process has occurred in the human species, which is known to have undergone a bottleneck (Harpending *et al.*, 1998). The human genome includes numerous nonsynonymous single-nucleotide polymorphisms (SNPs) that show lower gene diversities than either synonymous SNPs or intronic SNPs in the same loci (Freudenberg-Hua *et al.*, 2003; Hughes *et al.*, 2003; Zhao *et al.*, 2003). The reduction in gene diversity is particularly striking in the case of nonsynonymous mutations that cause radical amino-acid changes (Hughes *et al.*, 2003). The reduced frequency of these nonsynonymous mutations is evidence that they are subject to ongoing purifying selection. A similar pattern is seen at many bacterial and viral protein-coding loci (Hughes, 2005, 2007; Hughes *et al.*, 2007). These results provide strong support for the nearly neutral theory and evidence against theories that positive selection on protein-coding genes is widespread.

As mentioned, nearly neutral mutations include slightly advantageous as well as slightly deleterious mutations (Ohta, 2002). Sawyer *et al.* (2007) argue that

slightly advantageous mutations are likely to be as frequent as slightly deleterious mutations, using an argument from Fisher (1930), based on analogy with the process of adjusting a mechanical contrivance such as a microscope. This analogy appears reasonable if applied to continuously varying phenotypic traits such as body size. For example, in a population variants that increase body size above the population mean may be roughly as common as those that decrease body size below the population mean, as assumed in standard models of quantitative genetics (Falconer, 1981). But the analogy seems less certain when applied to changes at the amino-acid sequence level. For instance, there is evidence that the strength of selection against a given amino-acid replacement increases as a function of the chemical distance from the amino acid replaced (Yampolsky *et al.*, 2005); yet, for any amino acid, the possible replacements causing substantial chemical dissimilarity far outnumber those with similar chemical properties.

#### Testing for positive selection: the MHC example

Just as nucleotide sequence data were starting to become available, Kimura (1977) pointed out that the neutral theory predicts that, in protein-coding genes, the number of synonymous nucleotide substitutions per synonymous site ( $d_S$ ) typically exceeds the number of nonsynonymous substitutions per nonsynonymous site ( $d_N$ ). This pattern is predicted because the neutral theory predicts that most nonsynonymous mutations are harmful to protein function and thus will tend to be eliminated by purifying selection, whereas synonymous mutations are much more likely to be selectively neutral or nearly so. The fact that the vast majority of genes show this pattern provides strong support for the neutral theory (Li *et al.*, 1985; Endo *et al.*, 1996; Hughes and Friedman, 2005; Hughes and French, 2007).

In fact, the mere comparison of  $d_S$  and  $d_N$  may underestimate the extent of purifying selection on coding sequences. Subramanian and Kumar (2006) point out that the highly mutable CpG positions occur at a much higher frequency in replacement sites in codons than they do in introns. As a consequence, the intensity of purifying selection on coding sequences is probably greater than previously inferred.

The logic behind the comparison of  $d_S$  and  $d_N$  led to another prediction: if natural selection acts to favor repeated changes at the amino-acid level, we would expect to see a pattern of  $d_N > d_S$ . This prediction has been supported in a small number of cases. One of the first such cases involved the genes of the vertebrate MHC, and it is instructive to review briefly the distinctive features of the MHC that may be unfamiliar to many biologists.

The MHC is a multi-gene family encoding cell-surface glycoproteins that present peptides to T cells (Klein, 1987). All jawed vertebrates possess two subfamilies of MHC molecules known, respectively, as the class I MHC (which present peptides to CD8+ or cytotoxic T cells) and the class II MHC (which present peptides to CD4+ or helper T cells). The class I MHC genes were the first to be discovered, and their discovery was due to the role that the class I molecules (also known as 'major transplantation antigens') play in controlling transplant rejection. By transplantation experiments in mice, it was

discovered that the class I MHC loci are highly polymorphic; and the same high level of polymorphism was discovered in the MHC genes of humans and other vertebrate species. This polymorphism was mysterious because, at this time, the natural function of the MHC genes was not known.

Zinkernagel and Doherty (1974) discovered that the class I MHC molecules present foreign antigens to cytotoxic T cells (CTL), thereby triggering the killing of cells infected by intracellular pathogens. The same authors also had data indicating that different class I allelic products present different antigens; and this fact suggested a mechanism to explain the maintenance of MHC polymorphism. Doherty and Zinkernagel (1975) argued that, in a population exposed to a variety of pathogens, a heterozygote at class I MHC loci will have an advantage, because the ability to present a broader array of foreign antigens than a homozygote will confer broader immune surveillance.

Doherty and Zinkernagel's (1975) hypothesis was the first to explain MHC polymorphism with reference to the actual biological function of these molecules, but it proved very difficult to test. The usual way of testing a hypothesis about heterozygote advantage would be to compare fitnesses of homozygotes and heterozygotes in natural populations—a daunting task in the case of the MHC. Statistical analysis of DNA sequences provided an alternative approach. A crucial step was the availability of the first crystal structure of a class I MHC molecule (Bjorkman *et al.*, 1987). This structure showed revealed the region of the molecule (the peptide-binding region or PBR) where peptides are bound by the class I MHC, thereby suggesting a testable prediction: if MHC polymorphism is maintained by a form of balancing selection that relates to peptide-binding (and therefore to disease resistance), this selection should be focused on the codons encoding the PBR (Hughes and Nei, 1988).

In the late 1980's, only a few class I MHC sequences were available to test this prediction. But even with a small number of sequences, comparison of synonymous and nonsynonymous substitutions provided a powerful method of testing Doherty and Zinkernagel's (1975) hypothesis. Hughes and Nei (1988) reasoned that, if class I MHC polymorphism is maintained by balancing selection focused on the PBR, then we should observe  $d_N > d_S$  in the PBR codons, whereas the rest of the gene should show the pattern ( $d_S > d_N$ ) seen in most protein-coding genes. This pattern is exactly what was observed (Hughes and Nei, 1988). A similar pattern was later observed in the case of class II MHC molecules as well (Hughes *et al.*, 1994).

It is worthwhile to emphasize several points in which the MHC case differs from a number of subsequent studies. First, in the case of the MHC, the work of Doherty and Zinkernagel (1975) and the availability of a crystal structure (Bjorkman *et al.*, 1987) provided an *a priori* hypothesis, based on biological reasoning, regarding the codons upon which we expect positive selection to be focused. This is in marked contrast with numerous recent studies where it is claimed that positive selection acts on a given gene or certain codons within a given gene, but no biological hypothesis for the nature of this selection is provided (Hughes *et al.*, 2006).

Another important aspect of the MHC case is that the comparison of  $d_S$  and  $d_N$  revealed positive selection in

this case because a limited set of codons encoding the PBR residues have been subject over time to recurring positive selection favoring one amino-acid change after another (Hughes and Hughes, 1995). This occurs because these codons are involved in a co-evolutionary process with pathogens that exposes them to continuing selection. However, there is no reason to suppose that such selection—favoring continued amino-acid change at a limited set of codons—is characteristic of positive selection in general. Yet biologists continue to expect that the pattern seen in the MHC is a 'signature of positive selection' that will be seen in other genes as well.

## Flawed statistical approaches

Reasoning that evidence for positive selection can be obtained from the relative numbers of synonymous and nonsynonymous substitutions, biologists developed a number of new tests for positive selection. Unfortunately, the most widely used such approaches are fundamentally flawed as tests for positive selection, because they do not effectively rule out alternative interpretations (Hughes *et al.*, 2006). Here, I briefly discuss the conceptual flaws underlying some widely used methods.

### The McDonald–Kreitman test

The McDonald–Kreitman (MK) test compares polymorphism and divergence at synonymous sites in a protein-coding gene (McDonald and Kreitman, 1991). Two species are compared, which are ideally so closely related that 'multiple hits' (unobserved nucleotide changes) are not a problem, but not so closely related that there is a chance of shared ('trans-species') polymorphism at the locus under study (Graur and Li, 1991; Whittam and Nei, 1991). The results are in the form of a contingency table: counts of polymorphic synonymous ( $P_s$ ) and nonsynonymous ( $P_n$ ) sites vs counts of synonymous ( $D_s$ ) and nonsynonymous ( $D_n$ ) divergence; that is, sites at which a change has occurred between the two species. The test is based on the expectation that  $P_n:P_s$  will equal  $D_n:D_s$  under conditions of strict neutrality. This expectation can be tested by a simple  $2 \times 2$  test of independence.

Several problems with the MK test have been noted. For example, changes over time in the rate of synonymous substitution (for example, through changes in mutation rate) can cause changes in the ratio of  $P_n:P_s$  to  $D_n:D_s$  independent of selection on amino-acid sequences (Gerber *et al.*, 2001). Moreover, problems may arise when, as a result of recombination within the gene or set of genes analyzed, all sites do not share the same evolutionary history (Shapiro *et al.*, 2007).

A perhaps more serious problem with the MK test is that it cannot distinguish between positive Darwinian selection and any factor that causes purifying selection to become relaxed or to become less efficient. For example, the nearly neutral theory predicts that during a population bottleneck, slightly deleterious mutations may no longer be effectively removed by purifying selection, and thus a certain number of such mutations may drift to fixation (Ohta, 1993; Eyre-Walker, 2002; Hughes *et al.*, 2006). As a result,  $D_n:D_s$  will exceed  $P_n:P_s$ . Since the MK test compares species and bottlenecks are a frequent occurrence in speciation, this phenomenon is likely to

lead to a high rate of false detection of positive selection by this test.

McDonald and Kreitman's (1991) original paper applied the MK test to the alcohol dehydrogenase (*Adh*) gene of *Drosophila melanogaster* and found greater  $Dn:Ds$  than  $Pn:Ps$ . They considered the hypothesis that this result might be explained by the fixation of slightly deleterious alleles during a bottleneck, but they preferred the hypothesis of adaptive evolution as a 'simpler' explanation. However, Ohta (1993) showed that  $Dn:Ds$  is much greater in *Adh* of Hawaiian *Drosophila* than in either the *melanogaster* or *obscura* species groups. Because the Hawaiian species are known to have undergone population bottlenecks in speciation (DeSalle and Templeton, 1988), this observation strongly supports the 'nearly neutral' hypothesis of fixation of slightly deleterious alleles during population bottlenecks, rather than the hypothesis of positive selection (Ohta, 1993). In spite of this result, numerous subsequent studies have been published which have failed to consider the alternative 'nearly neutral' explanation of a positive result of the MK test.

A further problem with the MK test relates to the way that deals with polymorphism within species. The MK test simply counts the numbers of synonymous and nonsynonymous polymorphism sites within a species. However, as noted by Tajima (1989), there are two aspects to polymorphism in DNA sequence data: (1) the mean number of pairwise difference between sequences; and (2) the number of polymorphic (or 'segregating') sites. The former will tend to be relatively high when most polymorphisms are at intermediate frequencies, while the latter will tend to be high when there is an excess of rare polymorphisms. Because the MK test only incorporates the latter aspect of within-species polymorphism (in the counts of  $Ps$  and  $Pn$ ), the test is unduly influenced by rare polymorphisms, many of which are likely to be slightly deleterious nonsynonymous polymorphisms in the process of being eliminated by ongoing purifying selection (Hughes *et al.*, 2003; Hughes, 2005, 2007).

The way the MK test handles within-species polymorphism can explain some anomalous results of recent studies using this method, results that are not easily explained in a selectionist framework (Eyre-Walker, 2006). Several studies using the MK test show a high level of 'positive selection' on between-species amino-acid differences in comparisons between *D. melanogaster* and other *Drosophila* species (Smith and Eyre-Walker, 2002; Sawyer *et al.*, 2007; Shapiro *et al.*, 2007), whereas in human–chimpanzee comparisons, a much lower proportion of genes show evidence of 'positive selection' (Bustamante *et al.*, 2005; Gojobori *et al.*, 2007).

The high rate of 'positive selection' detected by the MK test in *Drosophila* can be explained by fixation of slightly deleterious mutations during a bottleneck in the process of speciation (Ohta, 1993). The level of nucleotide diversity in *D. melanogaster* is at least five times as great as that in the human species, indicating a much larger long-term effective population size in the former than in the latter (Li and Sadler, 1991). With an origin in Sub-Saharan Africa, this species was largely unaffected by Pleistocene glaciation, a major cause of bottlenecks in species of the North Temperate zones (Hughes and Hughes, 2007). Given a large effective population size for

a long time, the nearly neutral theory predicts that slightly deleterious mutations will have a good chance of being purged by purifying selection. Thus, the highly effective purifying selection within *D. melanogaster*, by lowering  $Pn$ , causes  $Dn$  to appear large by comparison.

The human species, by contrast, underwent a bottleneck of long duration early in the origin of modern humans (Harpending *et al.*, 1998). As mentioned previously, the human population shows evidence of an excess of rare nonsynonymous polymorphisms, as expected if many of these polymorphisms represent slightly deleterious mutations that increased in frequency during the bottleneck and now are in the process of being eliminated by purifying selection (Hughes *et al.*, 2003). Consistent with this interpretation, Bustamante *et al.* (2005) reported a genome-wide  $Dn:Ds$  ratio of 0.60 in human–chimpanzee comparisons, but a  $Pn:Ps$  ratio of 0.91 within the human species. By contrast, a recent analysis of data from *D. melanogaster* showed a  $Dn:Ds$  ratio of 0.37 and a  $Pn:Ps$  ratio of 0.31 (Shapiro *et al.*, 2007). Thus, the explanation for the lack of evidence of 'positive selection' in humans is evidently that numerous slightly deleterious polymorphisms serve to increase the relative value of  $Pn$ . Similar reasoning can also explain why the MK test fails to find evidence of positive selection in *Arabidopsis thaliana* (Bustamante *et al.*, 2002). The reduced recombination due to a mating system based on partial self-fertilization in this plant species causes a reduced efficiency of purifying selection against slightly deleterious mutations, resulting in a relatively high value of  $Pn$ .

Andolfatto (2005) used a modification of the MK test in a study that claimed evidence of positive selection on noncoding regions in *Drosophila* species. In this case, noncoding sites, rather than nonsynonymous sites, were compared with synonymous sites and a greater relative abundance of divergence than of polymorphism was reported at noncoding sites. These results may reflect fixation of slightly deleterious mutations in certain noncoding regions, including regulatory sequences and micro-RNA sequences. On the other hand, the results of this study may be simply an artifact of the greater difficulty of aligning noncoding regions between species than within species, leading to an overestimate of the rate of divergence between species in noncoding regions.

#### Positive selection at individual codons

A number of methods make use of phylogenies to reconstruct the pattern of nucleotide change at individual codons; according to the developers of these methods, codons with a pattern of  $d_N > d_S$  can be considered subject to positive selection (Suzuki and Gojobori, 1999; Yang *et al.*, 2000). One problem with these methods is that they assume that the phylogeny is known with 100% accuracy. This may be true in some cases, but in the case of genes subject to positive selection, the phylogeny is often difficult or impossible to reconstruct. Positive selection often leads to parallel/convergent sequence changes, as similar amino-acid substitutions are favored in phylogenetically distant lineages; and abundance parallel/convergent change ('homoplasy') misleads phylogenetic reconstruction. Moreover, genes subject to positive selection are often subject to a high rate of recombination, which renders

meaningless the very idea of a phylogeny of the gene, since different parts of the same cistronic sequence have different phylogenies.

The MHC genes provide a number of striking examples of parallel/convergent amino-acid changes (Yeager and Hughes, 1999). Parallel/convergent change is particularly well documented in the case of experimental SIV infections in rhesus monkeys, where the same CTL escape mutants independently occurred in virus infecting different monkeys bearing the same class I MHC molecules (Hughes *et al.*, 2001). Certain loci in both the class I and class II MHC are known for very high levels of interallelic recombination; such recombination includes both recombination of exons from different alleles as well as exchange of short sequence motifs ('cassettes'), presumably by a gene conversion-like mechanism (Yeager and Hughes, 1999). These examples suggest that parallel/convergent evolution and recombination are of frequent occurrence in genes subject to positive selection, rendering the real-world applicability of phylogeny-based methods questionable.

Another problem with these methods is the assumption that a codon at which  $d_N > d_S$  must be subject to positive selection. Given the stochastic nature of mutation, it is to be expected that  $d_N > d_S$  will frequently occur at certain codons just by chance (Hughes and Friedman, 2005). In fact, when nucleotide substitutions are relatively rare, synonymous and nonsynonymous will tend to be negatively correlated (Hughes and Friedman, 2005). Codon-based tests for positive selection thus fail adequately to rule out alternatives to the hypothesis of positive selection (Hughes *et al.*, 2006). This leads to an overly nonconservative test, particularly when likelihood methods are used (Suzuki and Nei, 2004; Friedman and Hughes, 2007).

A final problem with codon-based methods is that they are applied in the absence of any *a priori* hypothesis. This situation is very different from the MHC case, in which Hughes and Nei (1988, 1989) tested a prediction of Doherty and Zinkernagel's (1975) hypothesis regarding natural selection on MHC loci. The output of a study using codon-based methods includes a list of codons at which positive selection is alleged to operate. But the analysis itself provides no information regarding the biological factors that might be responsible for the alleged selection. This situation is particularly problematic in the case of likelihood methods, which are prone to false positives, because the set of codons at which positive selection is said to operate is likely to include codons at which purifying selection is relaxed and/or codons at which synonymous substitutions are absent simply by chance.

#### Positive selection or relaxation of purifying selection?

In addition to the statistical tests described above, a number of recent studies comparing  $d_S$  and  $d_N$  have assumed that positive selection is occurring if  $d_N$  exceeds  $d_S$  even by a small amount. However, in such cases, it is difficult to distinguish between, on the one hand, positive selection and, on the other hand, relaxation of purifying selection. For reasons described above, the MK test and the codon-based methods also fail to distinguish clearly between positive selection and the relaxation of purifying selection (or population factors that render

purifying selection less efficient). Thus, we are faced with the paradoxical situation where a large proportion of cases in which positive selection has been claimed may really be cases in which purifying selection either is relaxed or has experienced a period of reduced efficiency.

This is important because authors have frequently used statistical evidence of positive selection to support the argument that a specific gene has played a role in adaptive evolution. For example, it has been claimed that changes to the microcephalin protein in the primate lineage have played some role in the evolution of increased brain size of hominids (Wang and Su, 2004). But the fact that the methods used do not clearly differentiate positive selection from the relaxation of purifying selection means that the same data might just as plausibly be interpreted as evidence of a reduction in the importance of microcephalin in the higher primates. Relaxation of purifying selection on microcephalin in higher primates might have occurred because changes in the pattern of brain development in the primates have rendered some previously important role of microcephalin no longer essential. Of course, it is not possible at the present time to decide between these two hypotheses. But the prevalence of a selectionist mentality and the use of nonconservative methods places biologists in a very undesirable position, where there are no clear-cut criteria for deciding whether a given protein is particularly important or particularly unimportant for a given biological process.

#### An inappropriate model of selection

Even aside from the methodological problems described above, there is a further conceptual difficulty with many recent studies of positive selection at the molecular level; namely, the assumption that selection will act as it does on the PBR of MHC molecules to favor repeated amino-acid changes in a limited set of codons within a given gene. There is no biological reason to suppose that selection will show this pattern in molecules that are neither involved in a co-evolutionary process nor in protein-protein recognition.

In fact, there is good reason to believe that natural selection favoring major phenotypic adaptations generally follows a very different pattern. For example, a single amino-acid replacement, rather than a series of amino-acid replacements, may often give rise to a new phenotype that is adaptive. A striking example is provided by a single amino-acid replacement in the melanocortin-1 receptor that plays a major role in a change in coat color in Florida beach mice *Peromyscus polionotus*, believed to be adaptive because it enhances concealment in the beach environment (Hoekstra *et al.*, 2006).

Note that comparisons of  $d_S$  and  $d_N$ ; the MK test; and codon-based tests would all be powerless to detect selection in a case like that of the beach mouse, involving a single nonsynonymous change at a single codon. But there is reason to expect that many adaptive phenotypes must involve single amino-acid replacements. One reason for this expectation is that, if a phenotypic change required a series of amino-acid replacements in a given gene, natural selection alone could not produce it. Unless each of the amino-acid replacements in the series is at

least slightly advantageous, there would be no reason for natural selection to favor the intermediate stages of such a series.

Whole or partial gene deletions provide another example of a kind of mutational event that may also contribute to important phenotypic adaptations yet is not detected by methods concentrating on synonymous and nonsynonymous substitutions. For example, in cavefish (genus *Astyanax*), different deletions that knocked out a pigmentation gene have occurred independently in different populations (Protas *et al.*, 2006). Comparison of complete genome sequences suggests that gene loss is an important factor in evolution (Hughes and Friedman, 2004), and patterns of selective gene loss may be responsible for many phenotypic adaptations.

Moreover, many important changes at the phenotypic level are not caused by changes in protein sequences but by changes in the patterns of gene expression. Examples include the loss of eyes in cavefish (Jeffrey, 2005) and the changes in bill morphology in Galápagos finches (Abzhanov *et al.*, 2004). Over three decades ago, King and Wilson (1975) hypothesized that the major phenotypic differences between human and chimpanzee are likely to have arisen from differences in gene expression. This hypothesis is plausible given that a more prolonged period of brain growth in human than in chimpanzee is likely to underlie several of the major phenotypic differences between the two species. Yet curiously, numerous studies have searched for genes with accelerated rates of amino-acid replacement between human and chimpanzee as candidates for playing a major role in the evolution of uniquely human adaptations. Studies that expect to find the basis for unique human adaptations in repeated amino-acid replacements in a set of proteins are almost certain to uncover nothing of interest.

One might ask why a biologist has persisted in applying the model of repeated amino-acid replacements as a hallmark of positive selection even if it is unlikely to be applicable in most cases. One possibility is that this is a case of what Atlan (1999) has called *l'effet réverbère* (named for the proverbial drunk who searches for his lost keys under a streetlamp, not because that is where he lost them but because the light is better there); in other words, researchers test for this particular type of positive selection because of the availability of statistical methods that will allegedly detect this type of selection. By contrast, there are no generally applicable statistical tests that can be used to search for adaptive evolution involving a single nonsynonymous substitution or a change in expression pattern. Another possibility is that the fondness of biologists for the model of repeated amino-acid changes represents an atavistic holdover of 'Darwinian' gradualism.

Whether or not phenotypic change over evolutionary time is saltational ('punctuated') or gradual has been much debated (Gould, 2002). However, it is worth remarking that even when phenotypic change is gradual (or at least proceeds by relatively modest saltations), gradual accumulation of amino-acid changes in one or more proteins will not in most cases give rise to gradual change at the phenotypic level. Rather, the example of beak size and shape in birds suggests that gradual morphological change is likely to be caused by gradual changes in the timing and level of expression of major

developmental switch genes (Abzhanov *et al.*, 2004; Wu *et al.*, 2004).

Hoekstra and Coyne (2007) have recently reviewed the literature on the genetic basis of phenotypic adaptations and found relatively few reported cases where changes in *cis*-regulatory elements are responsible for phenotypic changes. These authors did not discuss certain known cases (such as that of avian bill shape Abzhanov *et al.*, 2004) where change in the expression pattern of a transcription factor causes morphological change, presumably because the role of *cis*-regulatory elements was not determined in these cases. On the other hand, Hoekstra and Coyne (2007) cited 16 cases where mutations in protein-coding genes are associated with morphological change. It is worth noting that all but one of these cases involved pigmentation, and 10 cases involved independent mutations in different lineages to a single gene: the *Mcl1r* gene encoding the melanocortin-1 receptor. Moreover, many of the mutations to protein-coding genes involved loss of function, including loss of phosphorylation sites (1 case) and deletions or frame-shifts (5 cases). In fact, most or all of the changes to the *Mcl1r* gene probably involve some degree of loss of functionality. For example, in the beach mouse case cited above, there is evidence that the amino-acid replacement responsible for coat color change causes reduced ligand binding (Hoekstra *et al.*, 2006).

In addition to examples of morphological change, Hoekstra and Coyne (2007) list 16 cases involving plant life history traits, altitudinal physiology, insecticide resistance, and visual pigments. These 16 cases included 5 where loss of protein function was clearly implicated. In all categories of phenotypic change, there were 20 cases where amino-acid replacements were implicated in the functional change. In 14 of these 20 cases (70%), the phenotypic change could be attributed to a single amino-acid replacement. Where there were multiple amino-acid replacements, the role of the individual replacements in the phenotypic change was not clear.

Thus, the survey of Hoekstra and Coyne (2007) suggests that phenotypic changes caused by changes in protein-coding genes often involve losses of function and may be particularly likely to involve aspects of morphology such as pigmentation where a loss of function in one component of a complex pathway will produce a marked phenotypic difference. Moreover, in most cases where amino-acid replacements are involved, the phenotypic change can be explained by a single mutation.

Hoekstra and Coyne's (2007) data may include a few cases in which a series of amino-acid changes has given rise to a new phenotype in a gradual fashion, but such cases are most likely to involve ligand binding. For example, there are four amino-acid changes in the melanocortin-1 receptor that are associated with melanism in lava-dwelling pocket mice *Chaetodipus intermedius*, and these amino-acid replacements probably involved in interactions with other proteins (Nachman *et al.*, 2003). Consider a case where decreased binding of a given ligand by a given receptor gives rise to a selectively favored phenotype (such as cryptic coloration). In those circumstances, a single amino-acid replacement causing a slight decrease in binding would be selectively favored, as would additional replacements further decreasing binding. However, once binding is effectively eliminated, no further changes will be favored. A similar process

would occur if increased binding between a given ligand and receptor were favored; once the appropriate degree of binding is achieved, further amino-acid replacements would not be favored.

Thus, selection favoring either decreased or increased ligand binding will ordinarily result in a pattern very unlike that seen in the PBR of MHC genes, where repeated nonsynonymous changes are favored in the same set of codons over a long period of time. Yet, the latter is the type of selection that widely used tests (including both MK and codon-based methods) are designed to detect. We are faced with a paradoxical situation indeed where the most commonly used statistical methods are designed to detect a form of selection that is almost never is of importance in context to which the methods are applied. The fact that these methods are overly nonconservative and fail adequately to rule out alternative hypotheses (Hughes *et al.*, 2006) only aggravates the situation.

In the effort to understand the molecular basis of adaptations, comparison of gene expression patterns is likely to provide important insights that are lacking if we focus on coding sequence changes alone (Khaitovich *et al.*, 2006). Some authors have attempted to formulate models of neutral changes in gene expression pattern over evolutionary time with the goal of providing a null hypothesis against which the hypothesis of adaptive change in gene expression might be tested (Khaitovich *et al.*, 2005). However, if such tests are to be done, it will be important to consider that a sudden acceleration in the rate of change of gene expression pattern may be caused by the relaxation of purifying selection as well as by positive selection (Gilad *et al.*, 2006). Developing statistical methods that can decide between these two alternative hypotheses will be a formidable challenge.

Indeed, biologists need to consider whether the statistical analysis of sequence or expression data alone really provides an adequate test of the hypothesis of adaptive evolution in most cases. Consider again the example of coat color in beach mice mentioned previously. Hoekstra *et al.* (2006) used sequence analysis to uncover the major mutational change underlying the coat color change in beach mice. But, in follow-up studies, the most straightforward way to test the hypothesis that this change is adaptive would be to conduct experiments in behavioral ecology, to test the prediction that the color change reduces detection of the mice by predators in their natural habitat. Likewise, when the genetic basis of other possibly adaptive traits is discovered, the next logical step will often be to conduct appropriate biochemical, physiological or ecological experiments to test of the hypothesis that the trait confers a fitness advantage. Such empirical testing provides far more convincing evidence of a trait's adaptive value than can any statistical argument.

The same reasoning applies in the case of gene expression data. Given experimental evidence that a particular change in gene expression underlies a particular phenotype, follow-up studies should be designed to test experimentally the hypothesis that the phenotype in question confers an adaptive advantage in a natural setting. Such an experimental approach is much more likely to yield insights of value than any statistical analysis of the patterns of variation in gene expression.

## Conclusion: does it matter?

For the past 20 years, there has been a tendency on the part of journal editors and reviewers to assume that every case of alleged statistical evidence for positive selection is worthy of publication, even in the absence of a plausible biological mechanism underlying the alleged selection. One unintentionally beneficial effect of the widespread use of overly nonconservative tests for positive selection has been a saturation of the literature with such claims. It is to be hoped that this saturation will in turn lead editors and reviewers to question the value of yet another statistically based claim of evidence for positive selection divorced from any biological mechanism.

Indeed, it is worth asking what can be learned from the conclusion that positive selection has acted—even if that conclusion is true—in cases where we have no knowledge of the biological basis of that selection. In some cases, reporting evidence of positive selection, especially when the evidence is based on conservative tests, may serve to stimulate further studies that will examine the function of the gene involved and suggest possible mechanisms for the apparent positive selection (Hughes and French, 2007). On the other hand, it may prove difficult to test certain claims of positive selection that have been proposed on the basis of statistical arguments alone. For example, a recent paper by Mustonen and Lässig (2007) applied a model of fluctuating selection to noncoding regions of the *Drosophila* genome. Even if the authors' hypothesis is correct, it will not be easy to reconstruct the environmental factors that might have given rise to this kind of selection.

It has sometimes been argued that searching for positive selection on the human genome can be justified, because such studies yield insights into the basis of human complex diseases (Kelley *et al.*, 2007). The rationale behind such an expectation is rarely clarified. There may be a few cases where positive selection is associated with genetic disease, but surely there is no reason to expect a relationship to hold in general. Sickle-cell anemia represents the one well-established case where a genetic disease results from adaptive evolution. But this is an isolated example; indeed, one might characterize the sickle-cell gene as a very poor sort of adaptation, since it imposes such a severe cost in the loss of homozygotes. It seems reasonable to predict that such costly adaptations will be few, since, were a new mutant to arise that provides an equivalent benefit without the high cost, it would spread quickly at the expense of the costly adaptation.

On the other hand, there is strong evidence for the presence of abundant slightly deleterious variants in the human population. Given that many of these variants are known to be subject to ongoing purifying selection, it seems much more plausible to examine these variants as candidates for a role in complex disease than to search for positively selected variants (Hughes *et al.*, 2003; Yampolsky *et al.*, 2005).

As regard to the infectious disease, the detection of positive selection at the molecular level has been a powerful tool for examining the role of host immune recognition in shaping the evolution of pathogens (for example, Allen *et al.*, 2000). But the use of nonconservative methods has sometimes suggested the occurrence of



positive selection on pathogen proteins not known to interact with the host immune system or indeed to play a direct role in the process of infection (for example, Perez-Losada *et al.*, 2005). Of course in some cases positive selection may indeed be occurring as a result of previously unknown processes, but there is a real danger that the reporting of false positives will obscure our understanding of host–pathogen interactions, overwhelming the genuine cases of immune-driven selection in an avalanche of ill-founded claims. In our ongoing efforts to understand the biology of major human pathogens, the unsupported assumption that positive selection is ubiquitous can cause real mischief, providing a false picture of the nature of the host–parasite interaction that can retard progress toward effective therapies.

## Acknowledgements

This research was supported by Grant GM43940 from the National Institutes of Health.

## References

- Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004). *Bmp4* and morphological variation of beaks in Darwin's finches. *Science* **305**: 1462–1465.
- Allen TM, O'Connor DH, Jing P, Dzuris JL, Mothé BR, Vogel TU *et al.* (2000). Tat-specific CTL select for SIV escape variants during resolution of primary viremia. *Nature* **407**: 386–390.
- Atlan H (1999). *La fin du 'tout génétique'?*. Institut National de la Recherche Agronomique: Paris.
- Andolfatto P (2005). Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**: 1149–1152.
- Bjorkman PJ, Saper MA, Samraoui B, Bennet WAS, Strominger JL, Wiley DC (1987). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**: 512–518.
- Bustamante CD, Fledel-Alon A, Williamson S, Nielsen R, Hubisz MT, Glanowski S *et al.* (2005). Natural selection on protein-coding genes in the human genome. *Nature* **437**: 1153–1157.
- Bustamante CD, Nielsen R, Sawyer SA, Olsen KM, Purugganan MD, Hartl DL (2002). The cost of inbreeding in *Arabidopsis*. *Nature* **416**: 531–534.
- Desalle R, Templeton AR (1988). Founder effects and the rate of mitochondrial DNA evolution in Hawaiian *Drosophila*. *Evolution* **42**: 1076–1084.
- Doherty PC, Zinkernagel RM (1975). Enhanced immunologic surveillance in mice heterozygous at the *H-2* complex. *Nature* **256**: 50–52.
- Endo T, Ikeo K, Gojobori T (1996). Large-scale search for genes on which positive selection may operate. *Mol Biol Evol* **13**: 685–690.
- Eyre-Walker A (2002). Changing effective population size and the McDonald–Kreitman test. *Genetics* **162**: 2017–2024.
- Falconer DS (1981). *Introduction to Quantitative Genetics*, 2nd edn. Longman: London.
- Fisher RA (1930). *The Genetical Theory of Natural Selection*. Oxford University Press: Oxford.
- Fisher RA, Ford EB (1950). The 'Sewall Wright' effect. *Heredity* **4**: 117–119.
- Freudenberg-Hua Y, Freudenberg J, Kluck N, Cichon S, Propping P, Nöthen MM (2003). Single nucleotide variation analysis in 65 candidate genes for CNS disorders in a representative sample of the European population. *Genome Res* **13**: 2271–2276.
- Friedman R, Hughes AL (2007). Likelihood-ratio tests for positive selection of human and mouse duplicate genes reveal nonconservative and anomalous properties of widely used methods. *Mol Phyl Evol* **42**: 388–393.
- Gerber AS, Loggins R, Kumar S, Dowling TE (2001). Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? *Annu Rev Genet* **35**: 539–566.
- Gilad Y, Oshlack A, Rifkin SA (2006). Natural selection on gene expression. *Trends Genet* **22**: 456–461.
- Gojobori J, Tang H, Akey JM, Wu C-I (2007). Adaptive evolution in humans revealed by the negative correlation between the polymorphism and fixation phases of evolution. *Proc Natl Acad Sci USA* **104**: 3907–3912.
- Gould SJ (2002). *The Structure of Evolutionary Theory*. Belknap Press: Cambridge, MA.
- Graur D, Li W-H (1991). Neutral mutation hypothesis test. *Nature* **354**: 114–115.
- Harpending HC, Batzer MA, Gurven M, Jorde LB, Rogers AR, Sherry ST (1998). Genetic traces of ancient demography. *Proc Natl Acad Sci USA* **95**: 1961–1967.
- Hoekstra HE, Coyne JA (2007). The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**: 995–1016.
- Hoekstra HE, Hirschmann RJ, Bundley RA, Insel PA, Crossland JP (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**: 101–104.
- Hughes AL (1999). *Adaptive Evolution of Genes and Genomes*. Oxford University Press: New York.
- Hughes AL (2005). Evidence for abundant slightly deleterious polymorphisms in bacterial populations. *Genetics* **169**: 553–558.
- Hughes AL (2007). Micro-scale signature of purifying selection in Marburg virus genomes. *Gene* **392**: 266–272.
- Hughes AL, French JO (2007). Homologous recombination and the pattern of nucleotide substitution in *Ehrlichia ruminantium*. *Gene* **387**: 31–37.
- Hughes AL, Friedman R (2004). Shedding genomic ballast: extensive parallel loss of ancestral gene families in animals. *J Mol Evol* **59**: 827–833.
- Hughes AL, Friedman R (2005). Variation in the pattern of synonymous and nonsynonymous difference between two fungal genomes. *Mol Biol Evol* **22**: 1320–1324.
- Hughes AL, Friedman R, Glenn NL (2006). The future of data analysis in evolutionary genomics. *Curr Genomics* **7**: 227–234.
- Hughes AL, Hughes MA (2007). Coding sequence polymorphism in avian mitochondrial genomes reflects population histories. *Mol Ecol* **16**: 1369–1376.
- Hughes AL, Hughes MA, Friedman R (2007). Variable intensity of purifying selection on cytotoxic T-lymphocyte epitopes in hepatitis C virus. *Virus Res* **123**: 147–153.
- Hughes AL, Hughes MK (1995). Natural selection on the peptide-binding regions of major histocompatibility complex molecules. *Immunogenetics* **42**: 233–243.
- Hughes AL, Hughes MK, Howell CY, Nei M (1994). Natural selection at the class II major histocompatibility complex loci of mammals. *Phil Trans R Soc Lond B* **346**: 359–367.
- Hughes AL, Nei M (1988). Pattern of nucleotide substitution at MHC class I loci reveals overdominant selection. *Nature* **335**: 167–170.
- Hughes AL, Nei M (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA* **86**: 958–962.
- Hughes AL, Packer B, Welch R, Bergen AW, Chanock SJ, Yeager M (2003). Widespread purifying selection at polymorphic sites in human protein-coding loci. *Proc Natl Acad Sci USA* **100**: 15754–15757.
- Hughes AL, Westover K, da Silva J, O'Connor DH, Watkins DI (2001). Simultaneous positive and purifying selection on overlapping reading frames of the tat and vpr genes of Simian Immunodeficiency virus. *J Virol* **75**: 7966–7972.
- Jeffrey WR (2005). Adaptive evolution of eye degeneration in the Mexican blind cavefish. *J Hered* **96**: 185–196.

- Kelley JL, Madeoy J, Calhoun JC, Swanson W, Akey JM (2007). Genomic signatures of positive selection in humans and the limits of outlier approaches. *Genome Res* **16**: 980–989.
- Khaitovich P, Enard W, Lachmann M, Pääbo S (2006). Evolution of primate gene expression. *Nat Rev Genet* **7**: 693–702.
- Khaitovich P, Pääbo S, Weiss G (2005). Toward a neutral evolutionary model of gene expression. *Genetics* **170**: 929–939.
- Kimura M (1955). Solution of a process of random genetic drift with a continuous model. *Proc Natl Acad Sci USA* **41**: 144–150.
- Kimura M (1957). Some problems of stochastic processes in genetics. *Ann Math Stat* **28**: 882–891.
- Kimura M (1964). Diffusion models in population genetics. *J Appl Prob* **1**: 177–232.
- Kimura M (1968). Evolutionary rate at the molecular level. *Nature* **217**: 624–626.
- Kimura M (1977). Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **267**: 275–276.
- Kimura M (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press: Cambridge.
- King MC, Wilson AC (1975). Evolution at two levels in humans and chimpanzees. *Science* **188**: 107–116.
- Klein J (1987). *Natural History of the Major Histocompatibility Complex*. Wiley: New York.
- Larson EJ (2004). *Evolution: the Remarkable History of a Scientific Theory*. Modern Library: New York.
- Li W-H, Sadler LA (1991). Low nucleotide diversity in man. *Genetics* **129**: 513–523.
- Li W-H, Wu C-I, Luo C-C (1985). A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol* **2**: 150–174.
- McDonald JH, Kreitman M (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 114–116.
- Mustonen V, Lässig M (2007). Adaptations to fluctuating selection in *Drosophila*. *Proc Natl Acad Sci USA* **104**: 2277–2282.
- Nachman MW, Hoekstra HE, D'Agostino SL (2003). The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci USA* **100**: 5268–5273.
- Ohta T (1973). Slightly deleterious mutant substitutions in evolution. *Nature* **246**: 96–98.
- Ohta T (1976). Role of very slightly deleterious mutations in molecular evolution and polymorphism. *Theor Pop Biol* **10**: 254–275.
- Ohta T (1993). Amino acid substitution at the *Adh* locus of *Drosophila* is facilitated by small population size. *Proc Natl Acad Sci USA* **90**: 4548–4551.
- Ohta T (2002). Near-neutrality in evolution of genes and gene regulation. *Proc Natl Acad Sci USA* **99**: 16134–16137.
- Perez-Losada M, Viscidi RP, Demma JC, Zenilman J, Crandall KA (2005). Population genetics of *Neisseria gonorrhoeae* in a high-prevalence community using a hypervariable outer membrane *porB* and 13 slowly evolving housekeeping genes. *Mol Biol Evol* **22**: 1887–1902.
- Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffrey WR *et al.* (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genet* **38**: 107–111.
- Sawyer SA, Parsch J, Zhang Z, Hartl DL (2007). Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila*. *Proc Natl Acad Sci USA* **104**: 6504–6510.
- Shapiro JA, Huang W, Zhang C, Hubisz MJ, Lu J, Turissini DA *et al.* (2007). Adaptive evolution in the *Drosophila* genome. *Proc Natl Acad Sci USA* **104**: 2271–2276.
- Smith NG, Eyre-Walker A (2002). Adaptive protein evolution in *Drosophila*. *Nature* **415**: 1022–1024.
- Subramanian S, Kumar S (2006). Higher intensity of purifying selection on >90% of the human genes revealed by the intrinsic replacement mutation rates. *Mol Biol Evol* **23**: 2283–2287.
- Suzuki Y, Gojobori T (1999). A method for detecting positive selection at single amino acid sites. *Mol Biol Evol* **16**: 1315–1328.
- Suzuki Y, Nei M (2004). False-positive selection identified by ML-based methods: examples from the *Sig1* gene of the diatom *Thalassiosira weissflogii* and the *tax* gene of a human T-cell lymphotropic virus. *Mol Biol Evol* **21**: 914–921.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Wang Y, Su B (2004). Molecular evolution of microcephalin, a gene determining human brain size. *Hum Mol Genet* **13**: 1131–1137.
- Whittam TS, Nei M (1991). Neutral mutation hypothesis test. *Nature* **354**: 115–116.
- Wu P, Jiang T-X, Suksaweang S, Widelitz RB, Chuong C-M (2004). Molecular shaping of the beak. *Science* **305**: 1465–1466.
- Yampolsky LY, Kondrashov FA, Kondrashov AS (2005). Distribution of the strength of selection against amino acid replacements in human proteins. *Hum Mol Genet* **14**: 3191–3201.
- Yang Z, Nielsen R, Goldman N, Pedersen AM (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431–449.
- Yeager M, Hughes AL (1999). Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. *Immunol Rev* **167**: 45–58.
- Zhao Z, Fu Y-X, Hewett-Emmett D, Boerwinkle E (2003). Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. *Gene* **312**: 207–213.
- Zinkernagel RM, Doherty PC (1974). Immunological surveillance against altered self components by sensitized T lymphocytes in lymphocytic choriomeningitis. *Nature* **251**: 547–548.