Genetic diversity in humans and non-human primates and its evolutionary consequences

Naoki Osada[1,2,3*]

[1]*Department of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan*
[2]*Department of Genetics, SOKENDAI (The Graduate University for Advanced Studies),*
1111 Yata, Mishima, Shizuoka 411-8540, Japan
[3]*Division of Bioengineering and Bioinformatics, Graduate School of Information Science and Technology,*
Hokkaido University, Sapporo, Hokkaido 060-0814, Japan

Genetic diversity is a key parameter in population genetics and is important for understanding the process of evolution and for the development of appropriate conservation strategies. Recent advances in sequencing technology have enabled the measurement of genetic diversity of various organisms at the nucleotide level and on a genome-wide scale, yielding more precise estimates than were previously achievable. In this review, I have compiled and summarized the estimates of genetic diversity in humans and non-human primates based on recent genome-wide studies. Although studies on population genetics demonstrated fluctuations in population sizes over time, general patterns have emerged. As shown previously, genetic diversity in humans is one of the lowest among primates; however, certain other primate species exhibit genetic diversity that is comparable to or even lower than that in humans. There exists greater than 10-fold variation in genetic diversity among primate species, and I found weak correlation with species fecundity but not with body or propagule size. I further discuss the potential evolutionary consequences of population size decline on the evolution of primate species. The level of genetic diversity negatively correlates with the ratio of non-synonymous to synonymous polymorphisms in a population, suggesting that proportionally greater numbers of slightly deleterious mutations segregate in small rather than large populations. Although population size decline is likely to promote the fixation of slightly deleterious mutations, there are molecular mechanisms, such as compensatory mutations at various molecular levels, which may prevent fitness decline at the population level. The effects of slightly deleterious mutations from theoretical and empirical studies and their relevance to conservation biology are also discussed in this review.

Key words: genetic diversity, nucleotide diversity, primates, deleterious mutations, conservation

INTRODUCTION

Genetic diversity is a measure of the genetic differences among individuals within a population and is an important research parameter in several areas of biology. For example, genetic diversity, generated as a consequence of mutations, is the ultimate source of variation for the long-term evolution of organisms; therefore, it is an important parameter in evolutionary biology. Its importance in biomedical research is attributable to the genetic variability within humans or experimental animal models in the reaction to drugs or in disease susceptibility. Furthermore, genetic diversity is an important criterion in endangered species conservation. Genetic diversity among humans provides insight not only into the demographic history of human populations but also into the long-term evolution of the human lineage. Comparative studies on human evolution chiefly focus on two divergent aspects, the similarities and differences of humans to other organisms, which can be addressed by a comparison of humans to other non-human primates. However, until recently, our knowledge of genetic diversity among non-human primates was limited owing to technical drawbacks.

A wide range of genetic diversity exists among the different taxa of living organisms, with seemingly simple

---

* Corresponding author. E-mail: nosada@ist.hokudai.ac.jp

organisms such as bacteria exhibiting greater genetic diversity than complex multicellular organisms such as vertebrates (Lynch, 2006). Recent studies based on large-scale sequencing revealed that different families of animals exhibit different levels of genetic diversity, which are highly correlated with life history traits of the animals, such as body size, longevity, dispersal ability, fecundity and propagule size (Romiguier et al., 2014). In general, species with a larger body size, longer life span, restricted niche and low fecundity have lower genetic diversity. The existence of this universal trend is surprising given that the population size of organisms, which is a major determinant of genetic diversity, can fluctuate as a consequence of environmental changes, and that closely related species can have significantly different population sizes. Although the underlying biological reasons are not clear, this correlation was attributed to the different ecological strategies, termed *r/K* selection strategies, typically employed by organisms (MacArthur and Wilson, 1967; Pianka, 1970). Despite the controversy surrounding this issue (Reznick et al., 2002), these terms remain useful as labels for the life history traits of different organisms. Organisms whose life history is subject to *r*-selection (*r*-strategists) and *K*-selection (*K*-strategists) invest more resources toward the production and survival of offspring, respectively. Here, in the logistic growth model, *r* and *K* are parameters that represent population growth rate and carrying capacity, respectively. In general, *K*-strategists tend to exhibit lower genetic diversity than *r*-strategists, although *r*- and *K*-strategists are not always clearly distinguishable.

The theory of population genetics advocates that the effective population sizes of organisms exert a large influence on organismal evolution. The efficacy of natural selection correlates positively with effective population sizes (e.g., Kimura, 1964; see below for more detail); therefore, a reduction in genetic diversity could result in the enhanced fixation of slightly deleterious mutations, and eventually in decreased average fitness of the population. Therefore, genetic diversity is an important parameter in both evolutionary and conservation biology.

In this review, the current status of research on genetic diversity in humans and non-human primates is summarized, with a particular focus on genome-wide single nucleotide polymorphisms (SNPs). This review deals with four major topics: 1) the unbiased measurement of genetic diversity in different organisms. In the current genomic era, nucleotide diversity (heterozygosity at the level of single nucleotides) on a genome-wide scale provides the best statistical measure of genetic diversity; the methods employed for obtaining such statistical measures are briefly introduced in this review. 2) Previous studies have reported genetic diversity on a genome-wide level in several humans and non-human primates. Their results are summarized in this review to provide a reference for future studies. 3) The aforementioned data have been employed for conducting meta-analysis to clarify the correlation between data on genetic diversity and life history traits of primates, as observed at the level of animal families. 4) Finally, the effects of changes in population size on evolution and its relevance to conservation biology are discussed.

HOW IS GENETIC DIVERSITY QUANTIFIED?

Effective population size and nucleotide diversity
Early molecular studies investigated polymorphisms at the protein level to estimate genetic diversity within populations (Harris, 1966; Lewontin, 1967). Heterozygosity, defined as the probability of obtaining different alleles upon random sampling in a population, was frequently employed to quantify the level of genetic diversity. However, there are several problems associated with the use of protein heterozygosity as an indicator of neutral genetic diversity. For example, only mutations that have a significant effect on protein mobility in electrophoresis are detectable, and a large fraction of cryptic genetic variants exerting subtle effects on protein structure are liable to be overlooked. Moreover, some of the protein polymorphisms may not be selectively neutral; different proteins are under different levels of purifying (negative) selection, rendering it difficult to compare levels of genetic diversity estimated using different proteins.

The relatively easy determination of nucleotide sequence using DNA sequencing technology has overcome the above-mentioned problems. Heterozygosity in repeat length and nucleotide sequences is widely employed in current studies of genetic diversity. Nucleotide diversity (π), which is equivalent to heterozygosity, is defined as the average number of nucleotide differences per site between two randomly sampled sequences and is among the most frequently employed statistical measures for estimating genetic diversity within populations using SNP data (Nei and Li, 1979; Tajima, 1983). Assuming the infinite-sites model in Wright–Fisher populations of constant size,

$$\pi = 4N_e\mu,$$

where N_e is the effective population size and μ is the mutation rate per site per generation (see more detail in Charlesworth, 2009). According to this equation, the estimation of π and μ would allow the calculation of N_e. Although the accurate estimation of μ is difficult and controversial, μ in human nuclear genomes has been estimated to fall within the range between 1.0×10^{-8} and 2.5×10^{-8} per site per generation (Scally and Durbin, 2012). Therefore, π allows a rough estimation of the effective population size of a species (Lynch, 2010; Campbell and Eichler, 2013). For example, under the assumption that in

humans, $\pi = 1.0 \times 10^{-3}$ and $\mu = 2.5 \times 10^{-8}$, N_e is estimated to be 10,000.

Statistical properties of nucleotide diversity Because π represents the average difference between two randomly chosen DNA sequences, this statistical measure is less sensitive to low-frequency (rare) polymorphisms. Therefore, the estimates of π are robust against recent changes in population size, the segregation of deleterious mutations, and sequencing errors, which typically occur at low frequencies in samples. In other words, π is an indicator of long-term effective population size. There are several other statistical measures for quantifying genetic diversity within populations, such as those proposed by Watterson (1975) and Ewens (1972); however, those statistics are more sensitive to the presence of low-frequency polymorphisms, which could be caused by recent population size changes. In contrast to these measurements of historical N_e, an alternative approach for the estimation of N_e by monitoring changes in allele frequency per generation could measure contemporary effective population sizes of small populations (Crow and Morton, 1955; Krimbas and Tsakas, 1971; Nei and Tajima, 1981). The statistics sensitive to recent population size changes are not explained in the current review.

Analysis of the nucleotide sequences of protein-coding genes enables the separate estimation of π for non-synonymous (π_N) and synonymous (π_S) sites. Assuming that synonymous sites are selectively neutral, π_S can be considered an indicator of genetic diversity in populations. In addition, the ratio of π_N to π_S reflects the strength of natural selection on the segregation of amino acid mutations within populations. Thus, the use of nucleotide sequences for the estimation of genetic diversity allows the effects of natural selection and neutral genetic drift to be disentangled. For example, at a genome-wide level, if the level of π_S is identical in two species but one has lower π_N than the other, it can be deduced that the effect of purifying selection has been stronger in the species with lower π_N. However, the assumption that π_S is unaffected by selection is not satisfied in certain species (Chamary et al., 2006). Previous studies have proposed that synonymous sites are subject to weak selection for ensuring the efficiency of translation (Ikemura, 1981; Akashi, 1995). Codon usage bias appears to be weak in mammals (Subramanian, 2008); nevertheless, the non-neutral evolution of synonymous sites could have occurred in the highly expressed genes that are likely to be over-represented in transcriptome data. In addition, proteins of highly expressed genes are generally under stronger purifying selection and are more conserved than the other genes both between and within species (Liu et al., 2008). Therefore, the action of natural selection on highly expressed genes potentially decreases the estimated values of π_N and π_S from background genomic values.

Advantage of genome-wide data for estimating genetic diversity Another important caveat in the use of heterozygosity and nucleotide diversity for the estimation of genetic diversity is their relatively large variance as a consequence of random genetic drift. In the absence of recombination, the expected difference between two randomly sampled sequences from a population is proportional to the coalescence time (the time elapsed since the last common ancestral allele) between the two sequences. Under the conditions of constant population size, the expected coalescence time and variance are $2N_e$ generations and $4N_e^2$, respectively (Tajima, 1983). The accuracy of estimation is improved by increasing the number of samples, but the extent of the improvement is lower than that obtained upon increasing the number of analyzed loci. This is because samples from a population are not statistically independent (Tajima, 1983; Pluzhnikov and Donnelly, 1996; Felsenstein, 2006). Mitochondrial DNA sequences, which have been widely employed in genetic research on non-human primates, provide precise estimates of the effect of recent demographic changes, when the sample size is very large (Maruvka et al., 2011). However, single locus analysis fails to provide a reliable estimate of the effects of long-term demographic changes on genomes. In addition, mutation rate heterogeneity among loci may also affect the estimation of genetic diversity when a small number of loci are analyzed. In particular, the mutation rate of microsatellites is highly variable among loci, making the accurate estimation of mutation rate difficult (Payseur et al., 2011). Despite the mutation rate heterogeneity among loci, if we assume that the magnitude of variation does not differ greatly among different species, comparing genome-wide averages of nucleotide diversity is a straightforward way to contrast nucleotide diversities between different species. Therefore, estimation of nucleotide diversity (or a related statistical measure) from data obtained from multiple loci, and ideally from whole-genome sequencing, is desirable for comparing genetic diversity between different taxa. Of note, the levels of genetic diversity in autosomal and sex-linked chromosomal loci are not directly comparable, because the effective population sizes of sex-linked loci (for example, the mitochondrial genome and sex chromosomes) differ from those of autosomes. In addition, the levels of natural selection operating on these loci can be different, and recent demographic changes exert different effects on genetic diversity in sex-linked chromosomes and in autosomes (Pool and Nielsen, 2007; Hammer et al., 2010; Osada et al., 2013).

The recent use of second-generation DNA sequencers has allowed large-scale genome and transcriptome sequencing in several non-model organisms. Such approaches permit the estimation of genetic diversity in populations with relatively small sample sizes (Lynch, 2008; Haubold et al., 2010; Gayral et al., 2013; Osada,

2014), given the absence of any strong inbreeding effect (for example, in captive-born animals) or cryptic population structure. This review deals primarily with data obtained from nuclear gene resequencing, particularly from genome resequencing and transcriptome sequencing studies carried out in primates, with lesser emphasis on studies of mitochondrial DNA and microsatellites.

Genetic diversity mirrors long-term effective population size; however, population size fluctuates over time as a consequence of stochastic and environmental changes. Recent studies employing whole-genome sequencing data for estimating past demographic changes in species have revealed interesting population histories for primates (Li and Durbin, 2011; Higashino et al., 2012; Prado-Martinez et al., 2013; Carbone et al., 2014; Fan et al., 2014; Prufer et al., 2014; Zhou et al., 2014), as well as drastic and often multiple changes in effective population sizes. In theory, estimates of N_e represent the harmonic mean of N_e over past generations, which implies that population size contraction exerts a significant impact on genetic diversity in populations. Some of the estimated changes in population size are similar among different primate species that do not share recent demographic history, suggesting that common environmental factors (for instance, global temperature decline during the glacial period) are responsible for common patterns in population size changes.

GENETIC DIVERSITY IN THE NUCLEAR GENOMES OF PRIMATE SPECIES

Great apes (*Hominidae*) The genetic diversity in primate species is summarized in Fig. 1. One of the earliest studies on genetic diversity at the nucleotide level revealed very low levels of genetic diversity in humans
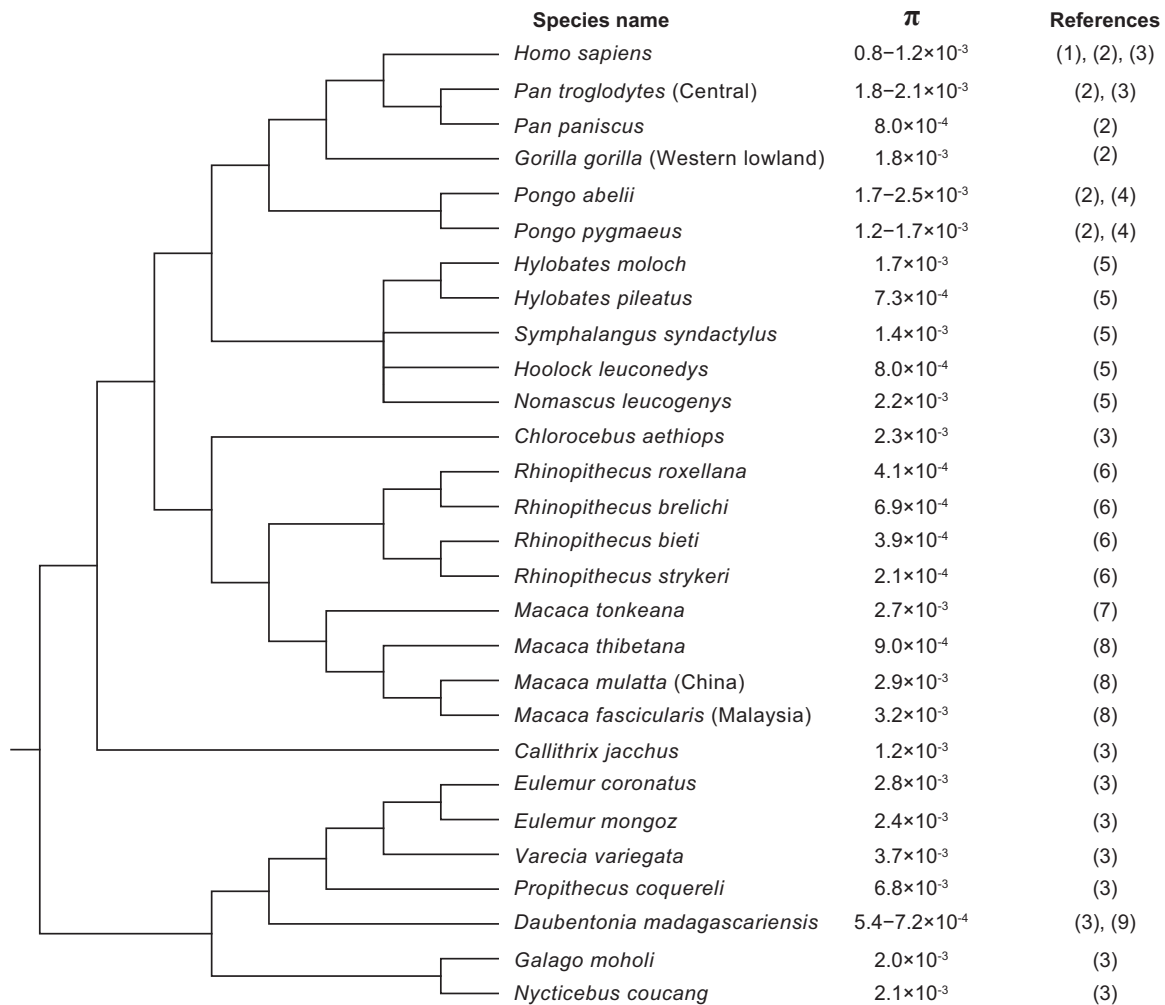
Species name	π	References
Homo sapiens	0.8–1.2×10^{-3}	(1), (2), (3)
Pan troglodytes (Central)	1.8–2.1×10^{-3}	(2), (3)
Pan paniscus	8.0×10^{-4}	(2)
Gorilla gorilla (Western lowland)	1.8×10^{-3}	(2)
Pongo abelii	1.7–2.5×10^{-3}	(2), (4)
Pongo pygmaeus	1.2–1.7×10^{-3}	(2), (4)
Hylobates moloch	1.7×10^{-3}	(5)
Hylobates pileatus	7.3×10^{-4}	(5)
Symphalangus syndactylus	1.4×10^{-3}	(5)
Hoolock leuconedys	8.0×10^{-4}	(5)
Nomascus leucogenys	2.2×10^{-3}	(5)
Chlorocebus aethiops	2.3×10^{-3}	(3)
Rhinopithecus roxellana	4.1×10^{-4}	(6)
Rhinopithecus brelichi	6.9×10^{-4}	(6)
Rhinopithecus bieti	3.9×10^{-4}	(6)
Rhinopithecus strykeri	2.1×10^{-4}	(6)
Macaca tonkeana	2.7×10^{-3}	(7)
Macaca thibetana	9.0×10^{-4}	(8)
Macaca mulatta (China)	2.9×10^{-3}	(8)
Macaca fascicularis (Malaysia)	3.2×10^{-3}	(8)
Callithrix jacchus	1.2×10^{-3}	(3)
Eulemur coronatus	2.8×10^{-3}	(3)
Eulemur mongoz	2.4×10^{-3}	(3)
Varecia variegata	3.7×10^{-3}	(3)
Propithecus coquereli	6.8×10^{-3}	(3)
Daubentonia madagascariensis	5.4–7.2×10^{-4}	(3), (9)
Galago moholi	2.0×10^{-3}	(3)
Nycticebus coucang	2.1×10^{-3}	(3)

Fig. 1. Phylogenetic tree and estimates of π in primates. Note that the branch length is not proportional to time. The tree has been extracted from previous studies (Perelman et al., 2011; Zhou et al., 2014), except for the star-like phylogeny of gibbons in the present study, attributable to the extreme complexity of the speciation process (Chan et al. 2013; Carbone et al. 2014). References: (1) Arbiza et al. (2014), (2) Prado-Martinez et al. (2013), (3) Perry et al. (2012), (4) Locke et al. (2011), (5) Carbone et al. (2014), (6) Zhou et al. (2014), (7) Evans et al. (2014), (8) Fan et al. (2014), (9) Perry et al. (2013).

compared with *Drosophila* (Li and Sadler, 1991). In recent genome-wide studies, π has been estimated to be approximately $1.0–1.2 \times 10^{-3}$ and 8.0×10^{-4} in African and non-African human populations, respectively (Perry et al., 2012; Prado-Martinez et al., 2013; Arbiza et al., 2014). Whether the other great apes, such as chimpanzees and gorillas, harbor similar levels of genetic diversity to human populations remained a subject of controversy (Morin et al., 1994; Wise et al., 1997; Cooper et al., 1998; Kaessmann et al., 1999, 2001; Yu et al., 2003, 2004; Fischer et al., 2011). Recent genome resequencing studies revealed a strong genetic structure at the subspecies level in chimpanzees and gorillas, but a high level of variability in genetic diversity within particular subspecies (Hvilsom et al., 2012; Perry et al., 2012; Prado-Martinez et al., 2013). This discrepancy is attributable to a scarcity of analyzed loci and genetic heterogeneity among subspecies.

Among chimpanzees, π in central chimpanzees (*Pan troglodytes troglodytes*) is approximately twice that in humans ($1.8–2.1 \times 10^{-3}$), whereas the genetic diversity of the western chimpanzee (*P. troglodytes verus*) is lower than that of African humans and similar to non-African humans. The other two subspecies of chimpanzee (eastern chimpanzee or *P. troglodytes schweinfurthii* and Nigeria–Cameroon chimpanzee or *P. troglodytes ellioti*) exhibit genetic diversity intermediate to those of the African humans and central chimpanzees. On the other hand, bonobo (*P. paniscus*), a sister species of the common chimpanzee (*P. troglodytes*), exhibits low genetic diversity, which is comparable to that of non-African humans ($\pi = 8.0 \times 10^{-4}$).

Gorillas exhibit a pattern similar to that of chimpanzees, with the western lowland gorillas (*Gorilla gorilla gorilla*) displaying the highest genetic diversity among gorillas and the other two subspecies, eastern lowland and cross river gorillas (*G. beringei graueri* and *G. gorilla diehli*, respectively), displaying slightly higher genetic diversity than African humans.

Despite their current small census size, orangutans display the highest genetic diversity among the great apes. The Bornean orangutans (*Pongo pygmaeus*) exhibit a similar level of genetic diversity to central chimpanzees and western lowland gorillas ($\pi = 1.7 \times 10^{-3}$). The Sumatran orangutans (*P. abelii*), on the other hand, display the highest genetic diversity among the great apes ($\pi = 2.5 \times 10^{-3}$); however, their census population size was estimated to be approximately 6,000–7,000 (Meijaard and Wich, 2007; Nater et al., 2013), which is attributable to a recent decline in population size. The use of different samples and methods has yielded slightly lower estimates of genetic diversity in both species (Locke et al., 2011).

In addition to extant primate species, recent advances in DNA sequencing technology using archaic samples has enabled the study of genetic diversity in extinct populations. Genome sequencing of archaic humans such as the Denisovans and Neanderthals revealed surprisingly low genetic diversity in these extinct populations. After correcting for the effects of strong inbreeding, the estimated nucleotide diversities of the Altai Neanderthal and Denisovan individuals were found to be $2.1–2.2 \times 10^{-4}$, which is just 30% of that found in extant human populations (Prufer et al., 2014). Studies on genetic diversity in extinct primate species are likely to have interesting implications for conservation biology.

Lesser apes (*Hylobatidae*) In contrast to the great apes, few studies have investigated genetic diversity in gibbons using multiple nuclear loci (Kim et al., 2011; Chan et al., 2013; Carbone et al., 2014). The most recent study, employing whole-genome resequencing, revealed great variability in the genetic diversity of gibbons within genera, with π values ranging from 7.3×10^{-4} in *Hylobates pileatus* to 2.2×10^{-3} in *Nomascus leucogenys* (Carbone et al., 2014). Sequencing of multiple genes revealed the highest π value ($4.4–4.7 \times 10^{-3}$) in *H. muelleri* (Kim et al., 2011; Chan et al., 2013). These studies agreed that the ancestral population size of gibbons was large, that they diverged over a relatively short time period approximately 5 million years ago, and that they experienced continuous gene flow following speciation, which complicates the evolutionary history of gibbons.

Old World monkeys (*Catarrhini*) The genomic features of the Old World monkeys (*Cercopithecidae*) have been relatively well studied because of their importance for biomedical research. Cynomolgus (crab-eating) macaques (*Macaca fascicularis*) were classified into four major groups based on their genetic characteristics: Indonesian–Malaysian, Indochinese, Philippine and Mauritian. The Indonesian–Malaysian population exhibits the highest genetic diversity among the *fascicularis* group of macaques ($\pi = 3.2 \times 10^{-3}$) (Osada et al., 2010; Higashino et al., 2012; Fan et al., 2014). The cynomolgus macaque population on the island of Mauritius, which originated from the Indonesian–Malaysian population in the 16th century, shows about 20% less genetic diversity than the Indonesian–Malaysian population (Osada et al., 2015). Rhesus macaques (*M. mulatta*), a sister species of the cynomolgus macaques, are subdivided into Indian and Chinese populations, and exhibit slightly lower genetic diversity at the species level ($\pi \approx 2.0–2.9 \times 10^{-3}$) (Hernandez et al., 2007; Osada et al., 2010; Yan et al., 2011; Perry et al., 2012; Fan et al., 2014). In contrast to these widely distributed species, the habitat of the Tibetan macaques (*M. thibetana*) is restricted to a limited region of China. As a result, they exhibit markedly low genetic diversity ($\pi = 9.0 \times 10^{-4}$), comparable to that of non-African humans (Fan et al., 2014). A study employing restriction site-associated DNA sequencing (RAD-seq) estimated the π

value of Tonkean macaque (*M. tonkeana*) as approximately 2.7×10^{-3}, using regions distal from the annotated genes (Evans et al., 2014).

Among the species in the genus *Chlorocebus*, relatively high genetic diversity ($\pi = 2.3 \times 10^{-3}$) was estimated in vervet (grivet) monkeys (*Chlorocebus aethiops*) using transcriptome data (Perry et al., 2012). Genome sequencing of the Vero cell line, which was established from *C. sabaeus*, revealed higher heterozygosity than humans, although direct comparison of this result with other studies is not feasible because of loss of heterozygosity and *de novo* mutations in the immortalized cell line (Osada et al., 2014).

Among the other groups of Old World monkeys, the sequencing of 12 autosomal loci in baboons (*Papio*) revealed a range of π from 4.2×10^{-4} to 1.9×10^{-3} and the existence of rampant gene flow between species (Boissinot et al., 2014). A recent study employing whole-genome sequencing in snub-nosed monkeys (genus *Rhinopithecus*) revealed extremely low genetic diversity in this group: 2.1×10^{-4} in *Rhinopithecus strykeri*, 3.9×10^{-4} in *R. bieti*, 4.1×10^{-4} in *R. roxellana* and 6.9×10^{-4} in *R. brelichi* (Zhou et al., 2014). Despite the strikingly low genetic diversity in certain species of this genus, caution should be exercised in interpretation, as these values were estimated from single individuals, and could be strongly influenced by the population structure of *Rhinopithecus* spp., most of which are endangered and live in fragmented habitats.

New World monkeys (*Platyrrhini*) Very few studies employing multiple nuclear gene sequences have been conducted to estimate genetic diversity within New World monkey species (Hiwatashi et al., 2010). The common marmoset (*Callithrix jacchus*), with its small body size, short life cycle and high fecundity, is widely used in biomedical research, and is one of the best-studied species of New World monkeys in the context of genomics research (The Marmoset Genome Sequencing and Analysis Consortium, 2014). A transcriptome study revealed π in the common marmoset to be as low as in African humans (1.2×10^{-3}).

Prosimians (*Lemuriformes* and *Tarsiiformes*) Prosimians are a paraphyletic group that includes the infraorders *Lemuriformes* and *Tarsiiformes*. Studies on genetic diversity in this group are summarized in this section. The group includes many endangered species, and surveying genetic diversity in this group is crucial to their conservation. However, studies conducted to date have largely been restricted to mitochondrial DNA and microsatellites (e.g., Blair et al., 2014; Brown et al., 2014). Transcriptome data revealed relatively high genetic diversity in this group compared to Simian primates, with the exception of aye-ayes (*Daubentonia madagascariensis*; $\pi \approx 7.2 \times 10^{-4}$), whose genetic diversity was found to be similar to or lower than non-African humans (Perry et

al., 2012). The estimate of genetic diversity in aye-ayes using transcriptome data is consistent with estimates obtained using data from whole-genome resequencing (Perry et al., 2013). The following estimates of π have been made for the other members of this group: 2.8×10^{-3} in crowned lemur (*Eulemur coronatus*), 2.4×10^{-3} in mongoose lemur (*E. mongoz*), 3.7×10^{-3} in black-and-white ruffed lemur (*Varecia variegata*), 2.1×10^{-3} in slow loris (*Nycticebus coucang*), 2.0×10^{-3} in mohol bushbaby (*Galago moholi*), and 6.8×10^{-3} in Coquerel's sifaka (*Propithecus verreauxi coquereli*) (Perry et al., 2012). Unfortunately, genome-wide studies on genetic diversity in *Tarsiiformes* have not been published to date.

THE CAUSE OF VARIATION IN GENETIC DIVERSITY

General trends in primate genetic diversity Existing data reveal three important trends in the genetic diversity of primates. First, lower levels of genetic diversity are often observed in primates compared to other mammals with small body sizes, including wild mice (*Mus musculus castaneus*; $\pi = 7.9 \times 10^{-3}$) (Halligan et al., 2010) and rabbits (*Oryctolagus cuniculus algirus* and *O. cuniculus cuniculus*; $\pi \approx 7.1–8.2 \times 10^{-3}$) (Carneiro et al., 2012). The levels are comparable to those found in domesticated animals such as dogs (*Canis lupus familiaris*) and cows (*Bos taurus*) (The Bovine HapMap Consortium, 2009; Gray et al., 2009), as well as mammals with large body sizes such as giant pandas (*Ailuropoda melanoleuca*) and brown or polar bears (*Ursus arctos* or *U. maritimus*) (Hailer et al., 2012; Zhao et al., 2013). The only exception is Coquerel's sifaka, which has a similar level of genetic diversity to wild mice. Second, the level of genetic diversity is quite variable even within genera. Closely related sister species sometimes exhibit considerably different levels of genetic diversity. For example, of the members of the genus *Macaca*, genetic diversity is approximately 3–4 times lower in Tibetan macaques than in cynomolgus macaques (Fan et al., 2014). In addition, several primate species display moderate to strong levels of population structure; such populations are often classified into different subspecies, and certain subspecies may exhibit lower genetic diversity than the other subspecies. Third, although there is a certain level of fluctuation within genera, historical genetic diversity appears to have decreased along the lineage to humans.

Does primate genetic diversity correlate with life history traits? Given that genetic diversity among primates varies more than 10-fold (Fig. 1), examining the possibility of predicting genetic diversity in these animals on the basis of their life history traits is an interesting proposition. Toward this end, propagule size (body size of juvenile individuals at the time of weaning) and fecun-

dity, two life history traits that show extremely strong correlation with genetic diversity across a wide range of animal taxa, were compiled for the species presented in Fig. 1, and examined for statistically significant correlation with π. Given the possibility of a phylogenetic correlation among propagule size, fecundity and π, the analysis was carried out using the method of phylogenetically independent contrasts proposed by Felsenstein (1985), which attempts to minimize the phylogenetic correlation of traits under the assumption of Brownian motion in trait evolution. The results are shown in Fig. 2. Fecundity, which was approximated to the inverse of reproduction interval time (Romiguier et al., 2014), showed a statistically significant positive correlation with genetic diversity ($P = 0.029$). On the other hand, propagule size, which displayed the highest correlation with genetic diversity across a broad spectrum of animal taxa, failed to exhibit significant correlation with genetic diversity in primates ($P = 0.400$).

The above-mentioned results indicate that despite certain significant trends, life history traits cannot be employed as general predictors of genetic diversity within primates. The low predictability is attributable to the relatively narrow range of trait values in primates. In addition, as mentioned earlier, most of the primate species, including humans, have experienced very recent fluctuations in population size, with different patterns of fluctuation among closely related species (Li and Durbin, 2011; Higashino et al., 2012; Fan et al., 2014; Zhou et al., 2014). Evolutionary changes in life history traits appear to be much slower than changes in effective population sizes caused by environmental factors, which could explain the weak association between genetic diversity and life history traits among primates.

EVOLUTIONARY CONSEQUENCES OF LOW GENETIC DIVERSITY

Impact of small effective population size on slightly deleterious mutations As summarized in the previous sections, certain primate lineages, including the modern human lineage, display a trend towards decreasing genetic diversity. The theory of population genetics predicts a reduction in the efficacy of natural selection as the effective population size decreases, because genetic drift is expected to be stronger in small populations (Kimura, 1968; Ohta, 1973). Figure 3 shows the predicted frequency spectrum of deleterious mutations relative to neutral mutations. The strength of natural selection on mutations is represented in this case by
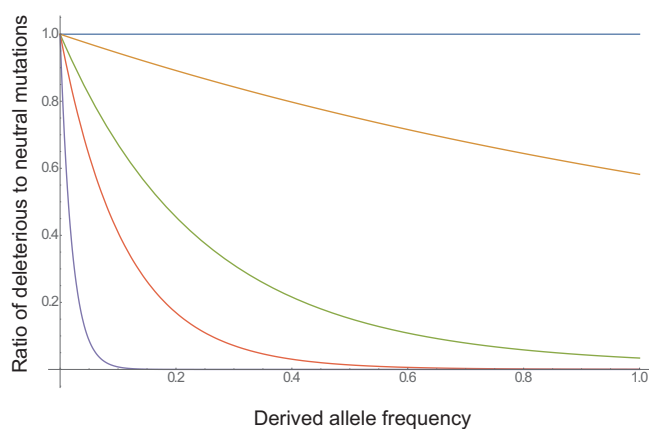
Fig. 3. Predicted frequency spectrum of deleterious mutations relative to neutral mutations. Blue, orange, green, red and purple lines represent scaled selection coefficient (S) values of 0 (neutral), −1, −5, −10 and −50, respectively.
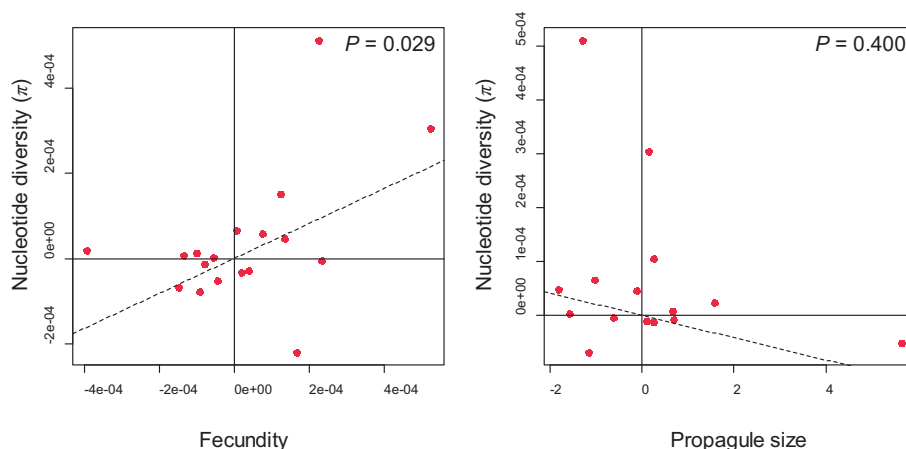
Fig. 2. Correlation between nucleotide diversity (π) and life history traits. To eliminate phylogenetic dependence, standardized contrasts or phylogenetically independent contrasts (PIC; Felsenstein, 1985) were employed. PIC analysis was performed using the APE package in R (Paradis et al., 2004). Branch length and life history traits were retrieved from Perelman et al. (2011) and Tacutu et al. (2013), respectively. A) Correlation between π and fecundity (inverse of the averaged inter-litter interval in females), B) correlation between π and propagule size (average volume of infants at the time of weaning).

$S = 2N_es$, where s is the selective advantage or disadvantage of mutations in heterozygous individuals. As shown in Fig. 3, mutations with $S < -5$ are unlikely to become fixed in a population but will continue to segregate at a low frequency, while mutations with $S < -50$ would rarely be encountered at allele frequencies of $> 10\%$. A reduction in N_e has essentially the same effect as a decrease in s, and is expected to result in the segregation of a greater number of slightly deleterious mutations within a population of small effective population size. For example, a 10-fold decrease in population size is expected to result in 10-fold stronger effects of deleterious mutations at the population level.

The ratio of nucleotide diversity at non-synonymous to synonymous sites (π_N/π_S) is an indicator of the degree of purifying selection within a population, assuming the absence of linkage between the synonymous and non-synonymous sites. The effect of linkage has been investigated using computer simulations and shown to be negligible when the selection is not strong (Boyko et al., 2008). Given that the values of π_N and π_S differ among genes, and that fluctuation of the ratio is possible for each gene, the ratio of averaged π_N to averaged π_S across genes or sites is usually considered. Figure 4 shows a reanalysis of the results obtained by Perry et al. (2012), where π_N and π_S were investigated across a wide range of primate species using transcriptome sequencing. Similar to the plot in Fig. 2, the data were analyzed considering phylogenetic independence. A statistically significant negative correlation ($P = 0.008$) was obtained between the level of genetic diversity (π_S) and the indicator of purifying selection within populations (π_N/π_S), confirming the segregation of an increased number of slightly deleterious mutations in small populations. As shown in Fig. 3, a small number of the segregating deleterious mutations
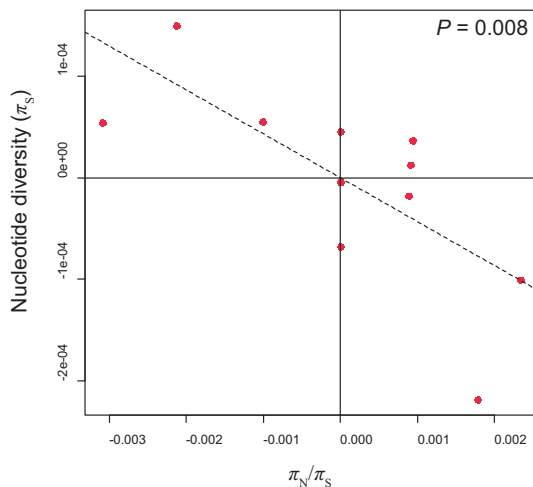

Fig. 4. Plot of π_S and π_N/π_S in 12 primate species, using data from Perry et al. (2012). Data were generated as described in the legend to Fig. 2.

could become fixed, potentially leading to a decline in the long-term fitness of organisms.

Notably, although humans are among the primate species with the lowest genetic diversity, other primates such as common marmosets and aye-ayes also possess low π_S and high π_N/π_S, indicating that the reduction in the efficacy of natural selection and the acceleration of genetic drift occurred in parallel in different primate lineages. Thus, in this context, humans do not present a unique case.

Influence of non-adaptive forces of evolution on the human genome The evolutionary process at the population level is governed by four fundamental forces: natural selection, mutation, genetic drift and recombination. With the exception of natural (Darwinian) selection, which results in rapid fixation of beneficial mutations in a population, the other three are categorized as non-adaptive forces. Several molecular studies have highlighted the importance of non-adaptive forces in the evolution of the human genome, as thoroughly reviewed by Harris (2010) and strongly advocated by Saitou (2014). Evidence supporting the dominating role of non-adaptive forces in shaping the evolution of the human genome is briefly summarized in this section; however, it should be noted that examining the signature of adaptive evolution in human genomic sequences is also important. First, as described in the previous section, the value of π_N/π_S in the human population is higher than in the majority of investigated species, suggesting that the effect of purifying selection is weaker in human populations than in other species with larger effective population sizes. Second, the ratio of non-synonymous to synonymous substitutions (d_N/d_S) is higher in the human lineage than in other mammalian lineages such as the mouse and rat lineages (Ohta, 1995; Kosiol et al., 2008). The ratio is even higher after the divergence from the last common ancestor of humans and chimpanzees (Wang et al., 2007). Similarly, a higher proportion of radical amino acid changes (as opposed to conservative ones) have become fixed between humans and macaques than between mice and rats (Hughes and Friedman, 2009). Finally, a reduction in the level of polymorphism near fixed mutations, which is a hallmark of the recent and rapid fixation of beneficial mutations (the so-called selective sweep), was observed in *Drosophila melanogaster* and *D. simulans* (Sattath et al., 2011) but not in humans (Hernandez et al., 2011; Lohmueller et al., 2011). Taken together, these patterns support, at least at the molecular level, the absence of a genome-wide signature of positive adaptive evolution in the human genome, with the exception of certain outlier loci (Sabeti et al., 2002, 2006). However, the adaptive evolution of the human genome remains an issue of controversy (e.g., Akey, 2009; Cai et al., 2009; Enard et al., 2014).

Genetic diversity and conservation genetics

Whether the segregation and fixation of deleterious mutations decreases organismal fitness is an important question. The decline in the average fitness of a population is termed genetic load, and is of great importance to public health in humans and to the conservation of non-human primates. The genetic load attributable to mutations in particular is referred to as the mutation load (Muller, 1950). While somewhat counterintuitive, recent studies using computer simulations predict that the mutation load caused by segregating mutations is unlikely to be strongly affected by recent demographic changes; this prediction is supported by the pattern of polymorphisms in humans (Lohmueller, 2014b; Simons et al., 2014; Do et al., 2015). This insensitivity to recent demographic changes is attributable to particular characteristics of the frequency of deleterious mutations and their strength. The frequency of deleterious mutations in populations is skewed towards rare (Fig. 3). For example, a decrease in population size results in the rapid purging of several deleterious mutations that occur at low frequencies, which in turn is balanced by increased frequencies of other deleterious mutations due to stronger genetic drift. Therefore, a drastic change in genetic load does not occur as a consequence of recent demographic changes, and a new equilibrium is gradually attained through the fixation and segregation of slightly deleterious mutations. Notably, although genetic load is insensitive to recent demographic changes, the frequency distribution of deleterious mutations in a population is altered after the change in population size (Lohmueller, 2014a).

Despite the theoretical predictions, the general consensus is that low genetic diversity, and therefore small population size, increases the extinction risk of populations. Several field studies have discovered a positive correlation between population size and the individual fitness of animals and plants (Reed and Frankham, 2003; Leimu et al., 2006). This discrepancy is attributable to a complex interaction between deleterious mutations and individual fitness (for example, epistatic interactions among deleterious mutations), or the simplification inherent in commonly employed models of population genetics. Alternatively, despite the fluctuations in population size over time, the long-term effective population sizes are in turn determined by life history traits (Romiguier et al., 2014), which could affect fitness through the fixation of slightly deleterious mutations.

In the context of long-term evolution, the fixation of slightly deleterious mutations could result in a significant decline in organismal fitness and lead organisms to extinction (Lande, 1994). In sexual organisms, recombination could function to shield populations from the catastrophic effects of mutation load by eliminating deleterious mutations (Lynch and Gabriel, 1990). In addition, compensatory mutations capable of ameliorating the deleterious effects of other mutations potentially contribute towards preventing organisms from continuous fitness decline (Akashi et al., 2012). A large body of evidence supports the possible existence of several mechanisms for compensatory evolution at the molecular level, including RNA structure (Chen et al., 1999; Meer et al., 2010), protein structure (Yanofsky et al., 1964; DePristo et al., 2005), protein–protein interaction (Osada and Akashi, 2012), nucleosome binding sites (Kenigsberg et al., 2010; Langley et al., 2014), transcription factor binding sites (Doniger and Fay, 2007; Weirauch and Hughes, 2010), codon usage bias (Li, 1987), splice sites (Denisov et al., 2014) and gene loss (Harcombe et al., 2009). Such molecular mechanisms could partially arrest fitness decline in organisms with small population sizes.

The study conducted by Perry et al. (2012) showed that genetic diversity is not always a good predictor of the endangered state of species. This is also the case with the Sumatran and Bornean orangutans (Nater et al., 2013), where the more abundant Bornean population was found to exhibit lower genetic diversity, which is attributable to the very recent destruction of the habitats of the Bornean population by humans. Therefore, genetic diversity alone cannot be employed as the sole criterion for deciding whether a species is endangered. However, a decline in genetic diversity does increase the risk of extinction, and knowledge of the level of genetic diversity facilitates the determination of appropriate conservation strategies for a particular species. This is because species with lower levels of genetic diversity are likely to be more severely affected by inbreeding depression in breeding programs. Investigations of genetic diversity in several wild-living primates have become increasingly feasible with the dramatic reduction in sequencing costs. Although the data I have reviewed in this article are limited to a dozen primate species, the availability of large quantities of information in the future should greatly benefit studies on evolutionary and conservation biology of all living organisms, including primates.

REFERENCES

Akashi, H. (1995) Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. Genetics **139**, 1067–1076.

Akashi, H., Osada, N., and Ohta, T. (2012) Weak selection and protein evolution. Genetics **192**, 15–31.

Akey, J. M. (2009) Constructing genomic maps of positive selection in humans: Where do we go from here? Genome Res. **19**, 711–722.

Arbiza, L., Gottipati, S., Siepel, A., and Keinan, A. (2014) Contrasting X-linked and autosomal diversity across 14

human populations. Am. J. Hum. Genet. **94**, 827–844.

Blair, C., Heckman, K., Russell, A., and Yoder, A. (2014) Multilocus coalescent analyses reveal the demographic history and speciation patterns of mouse lemur sister species. BMC Evol. Biol. **14**, 57.

Boissinot, S., Alvarez, L., Giraldo-Ramirez, J., and Tollis, M. (2014) Neutral nuclear variation in baboons (genus *Papio*) provides insights into their evolutionary and demographic histories. Am. J. Phys. Anthropol. **155**, 621–634.

Bovine HapMap Consortium (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science **324**, 528–532.

Boyko, A. R., Williamson, S. H., Indap, A. R., Degenhardt, J. D., Hernandez, R. D., Lohmueller, K. E., Adams, M. D., Schmidt, S., Sninsky, J. J., Sunyaev, S. R., et al. (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet. **4**, e1000083.

Brown, R. M., Weghorst, J. A., Olson, K. V., Duya, M. R. M., Barley, A. J., Duya, M. V., Shekelle, M., Neri-Arboleda, I., Esselstyn, J. A., Dominy, N. J., et al. (2014) Conservation genetics of the Philippine tarsier: Cryptic genetic variation restructures conservation priorities for an island archipelago primate. PLoS ONE **9**, e104340.

Cai, J. J., Macpherson, J. M., Sella, G., and Petrov, D. A. (2009) Pervasive hitchhiking at coding and regulatory sites in humans. PLoS Genet. **5**, e1000336.

Campbell, C. D., and Eichler, E. E. (2013) Properties and rates of germline mutations in humans. Trends Genet. **29**, 575–584.

Carbone, L., Alan Harris, R., Gnerre, S., Veeramah, K. R., Lorente-Galdos, B., Huddleston, J., Meyer, T. J., Herrero, J., Roos, C., Aken, B., et al. (2014) Gibbon genome and the fast karyotype evolution of small apes. Nature **513**, 195–201.

Carneiro, M., Albert, F. W., Melo-Ferreira, J., Galtier, N., Gayral, P., Blanco-Aguiar, J. A., Villafuerte, R., Nachman, M. W., and Ferrand, N. (2012) Evidence for widespread positive and purifying selection across the European rabbit (*Oryctolagus cuniculus*) genome. Mol. Biol. Evol. **29**, 1837–1849.

Chamary, J. V., Parmley, J. L., and Hurst, L. D. (2006) Hearing silence: Non-neutral evolution at synonymous sites in mammals. Nat. Rev. Genet. **7**, 98–108.

Chan, Y.-C., Roos, C., Inoue-Murayama, M., Inoue, E., Shih, C.-C., Pei, K. J.-C., and Vigilant, L. (2013) Inferring the evolutionary histories of divergences in *Hylobates* and *Nomascus* gibbons through multilocus sequence data. BMC Evol. Biol. **13**, 82.

Charlesworth, B. (2009) Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. **10**, 195–205.

Chen, Y., Carlini, D. B., Baines, J. F., Parsch, J., Braverman, J. M., Tanda, S., and Stephan, W. (1999) RNA secondary structure and compensatory evolution. Genes Genet. Syst. **74**, 271–286.

Cooper, G., Rubinsztein, D. C., and Amos, W. (1998) Ascertainment bias cannot entirely account for human microsatellites being longer than their chimpanzee homologues. Hum. Mol. Genet. **7**, 1425–1429.

Crow, J. F., and Morton, N. E. (1955) Measurement of gene frequency drift in small populations. Evolution **9**, 202–214.

Denisov, S. V., Bazykin, G. A., Sutormin, R., Favorov, A. V., Mironov, A. A., Gelfand, M. S., and Kondrashov, A. S. (2014) Weak negative and positive selection and the drift load at splice sites. Genome Biol. Evol. **6**, 1437–1447.

DePristo, M. A., Weinreich, D. M., and Hartl, D. L. (2005) Mis-

sense meanderings in sequence space: A biophysical view of protein evolution. Nat. Rev. Genet. **6**, 678–687.

Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., and Reich, D. (2015) No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nat. Genet. **47**, 126–131.

Doniger, S. W., and Fay, J. C. (2007) Frequent gain and loss of functional transcription factor binding sites. PLoS Comput. Biol. **3**, e99.

Enard, D., Messer, P. W., and Petrov, D. A. (2014) Genome-wide signals of positive selection in human evolution. Genome Res. **24**, 885–895.

Evans, B. J., Zeng, K., Esselstyn, J. A., Charlesworth, B., and Melnick, D. J. (2014) Reduced representation genome sequencing suggests low diversity on the sex chromosomes of tonkean macaque monkeys. Mol. Biol. Evol. **31**, 2425–2440.

Ewens, W. J. (1972) The sampling theory of selectively neutral alleles. Theor. Popul. Biol. **3**, 87–112.

Fan, Z., Zhao, G., Li, P., Osada, N., Xing, J., Yi, Y., Du, L., Silva, P., Wang, H., Sakate, R., et al. (2014) Whole-genome sequencing of Tibetan macaque (*Macaca thibetana*) provides new insight into the macaque evolutionary history. Mol. Biol. Evol. **31**, 1475–1489.

Felsenstein, J. (1985) Phylogenies and the comparative method. Am. Nat. **125**, 1–15.

Felsenstein, J. (2006) Accuracy of coalescent likelihood estimates: Do we need more sites, more sequences, or more loci? Mol. Biol. Evol. **23**, 691–700.

Fischer, A., Prüfer, K., Good, J. M., Halbwax, M., Wiebe, V., André, C., Atencia, R., Mugisha, L., Ptak, S. E., and Pääbo, S. (2011) Bonobos fall within the genomic variation of chimpanzees. PLoS ONE **6**, e21605.

Gayral, P., Melo-Ferreira, J., Glémin, S., Bierne, N., Carneiro, M., Nabholz, B., Lourenco, J. M., Alves, P. C., Ballenghien, M., Faivre, N., et al. (2013) Reference-free population genomics from next-generation transcriptome data and the vertebrate–invertebrate gap. PLoS Genet. **9**, e1003457.

Gray, M. M., Granka, J. M., Bustamante, C. D., Sutter, N. B., Boyko, A. R., Zhu, L., Ostrander, E. A., and Wayne, R. K. (2009) Linkage disequilibrium and demographic history of wild and domestic canids. Genetics **181**, 1493–1505.

Hailer, F., Kutschera, V. E., Hallström, B. M., Klassert, D., Fain, S. R., Leonard, J. A., Arnason, U., and Janke, A. (2012) Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science **336**, 344–347.

Halligan, D. L., Oliver, F., Eyre-Walker, A., Harr, B., and Keightley, P. D. (2010) Evidence for pervasive adaptive protein evolution in wild mice. PLoS Genet. **6**, e1000825.

Hammer, M. F., Woerner, A. E., Mendez, F. L., Watkins, J. C., Cox, M. P., and Wall, J. D. (2010) The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. Nat. Genet. **42**, 830–831.

Harcombe, W., Springman, R., and Bull, J. (2009) Compensatory evolution for a gene deletion is not limited to its immediate functional network. BMC Evol. Biol. **9**, 106.

Harris, E. E. (2010) Nonadaptive processes in primate and human evolution. Am. J. Phys. Anthropol. **143**, **suppl. 51**, 13–45.

Harris, H. (1966) Enzyme polymorphisms in man. Proc. R. Soc. Lond. B Biol. Sci. **164**, 298–310.

Haubold, B., Pfaffelhuber, P., and Lynch, M. (2010) Mlrho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. Mol. Ecol. **19**, **Suppl. 1**, 277–284.

Hernandez, R. D., Hubisz, M. J., Wheeler, D. A., Smith, D. G., Ferguson, B., Rogers, J., Nazareth, L., Indap, A., Bourquin, T., McPherson, J., et al. (2007) Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques. Science 316, 240–243.

Hernandez, R. D., Kelley, J. L., Elyashiv, E., Melton, S. C., Auton, A., McVean, G., 1000 Genomes Project, Sella, G., and Przeworski, M. (2011) Classic selective sweeps were rare in recent human evolution. Science 331, 920–924.

Higashino, A., Sakate, R., Kameoka, Y., Takahashi, I., Hirata, M., Tanuma, R., Masui, T., Yasutomi, Y., and Osada, N. (2012) Whole-genome sequencing and analysis of the Malaysian cynomolgus macaque (*Macaca fascicularis*) genome. Genome Biol. 13, R58.

Hiwatashi, T., Okabe, Y., Tsutsui, T., Hiramatsu, C., Melin, A. D., Oota, H., Schaffner, C. M., Aureli, F., Fedigan, L. M., Innan, H., et al. (2010) An explicit signature of balancing selection for color-vision variation in New World monkeys. Mol. Biol. Evol. 27, 453–464.

Hughes, A. L., and Friedman, R. (2009) More radical amino acid replacements in primates than in rodents: Support for the evolutionary role of effective population size. Gene 440, 50–56.

Hvilsom, C., Qian, Y., Bataillon, T., Li, Y., Mailund, T., Sallé, B., Carlsen, F., Li, R., Zheng, H., Jiang, T., et al. (2012) Extensive X-linked adaptive evolution in central chimpanzees. Proc. Natl. Acad. Sci. USA 109, 2054–2059.

Ikemura, T. (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J. Mol. Biol. 151, 389–409.

Kaessmann, H., Wiebe, V., and Pääbo, S. (1999) Extensive nuclear DNA sequence diversity among chimpanzees. Science. 286, 1159–1162.

Kaessmann, H., Wiebe, V., Weiss, G., and Pääbo, S. (2001) Great ape DNA sequences reveal a reduced diversity and an expansion in humans. Nat. Genet. 27, 155–156.

Kenigsberg, E., Bar, A., Segal, E., and Tanay, A. (2010) Widespread compensatory evolution conserves DNA-encoded nucleosome organization in yeast. PLoS Comput. Biol. 6, e1001039.

Kim, S. K., Carbone, L., Becquet, C., Mootnick, A. R., Li, D. J., de Jong, P. J., and Wall, J. D. (2011) Patterns of genetic variation within and between gibbon species. Mol. Biol. Evol. 28, 2211–2218.

Kimura, M. (1964) Difusion models in population genetics. J. Appl. Probab. 1, 177–232.

Kimura, M. (1968) Evolutionary rate at the molecular level. Nature 217, 624–626.

Kosiol, C., Vinar, T., da Fonseca, R. R., Hubisz, M. J., Bustamante, C. D., Nielsen, R., and Siepel, A. (2008) Patterns of positive selection in six mammalian genomes. PLoS Genet. 4, e1000144.

Krimbas, C. B., and Tsakas, S. (1971) The genetics of *Dacus oliae*. V. Changes of esterase polymorphism in a natural population following insecticide control-selection or drift? Evolution 25, 454–460.

Lande, R. (1994) Risk of population extinction from fixation of new deleterious mutations. Evolution 48, 1460–1469.

Langley, S. A., Karpen, G. H., and Langley, C. H. (2014) Nucleosomes shape DNA polymorphism and divergence. PLoS Genet. 10, e1004457.

Leimu, R., Mutikainen, P. I. A., Koricheva, J., and Fischer, M. (2006) How general are positive relationships between plant population size, fitness and genetic variation? J. Ecol. 94, 942–952.

Lewontin, R. C. (1967) An estimate of average heterozygosity in man. Am. J. Hum. Genet. 19, 681–685.

Li, H., and Durbin, R. (2011) Inference of human population history from individual whole-genome sequences. Nature 475, 493–496.

Li, W.-H. (1987) Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. J. Mol. Evol. 24, 337–345.

Li, W. H., and Sadler, L. A. (1991) Low nucleotide diversity in man. Genetics 129, 513–523.

Liu, J., Zhang, Y., Lei, X., and Zhang, Z. (2008) Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective. Genome Biol. 9, R69.

Locke, D. P., Hillier, L. W., Warren, W. C., Worley, K. C., Nazareth, L. V., Muzny, D. M., Yang, S. P., Wang, Z., Chinwalla, A. T., Minx, P., et al. (2011) Comparative and demographic analysis of orang-utan genomes. Nature 469, 529–533.

Lohmueller, K. E. (2014a) The distribution of deleterious genetic variation in human populations. Curr. Opin. Genet. Dev. 29, 139–146.

Lohmueller, K. E. (2014b) The impact of population demography and selection on the genetic architecture of complex traits. PLoS Genet. 10, e1004379.

Lohmueller, K. E., Albrechtsen, A., Li, Y., Kim, S. Y., Korneliussen, T., Vinckenbosch, N., Tian, G., Huerta-Sanchez, E., Feder, A. F., Grarup, N., et al. (2011) Natural selection affects multiple aspects of genetic variation at putatively neutral sites across the human genome. PLoS Genet. 7, e1002326.

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

Lynch, M. (2008) Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. Mol. Biol. Evol. 25, 2409–2419.

Lynch, M. (2010) Evolution of the mutation rate. Trends Genet. 26, 345–352.

Lynch, M., and Gabriel, W. (1990) Mutation load and the survival of small populations. Evolution 44, 1725–1737.

MacArthur, R. H., and Wilson, E. O. (1967) The theory of island biogeography. Princeton University Press, Princeton.

Marmoset Genome Sequencing and Analysis Consortium (2014) The common marmoset genome provides insight into primate biology and evolution. Nat. Genet. 46, 850–857.

Maruvka, Y. E., Shnerb, N. M., Bar-Yam, Y., and Wakeley, J. (2011) Recovering population parameters from a single gene genealogy: An unbiased estimator of the growth rate. Mol. Biol. Evol. 28, 1617–1631.

Meer, M. V., Kondrashov, A. S., Artzy-Randrup, Y., and Kondrashov, F. A. (2010) Compensatory evolution in mitochondrial tRNAs navigates valleys of low fitness. Nature 464, 279–282.

Meijaard, E., and Wich, S. (2007) Putting orang-utan population trends into perspective. Curr. Biol. 17, R540.

Morin, P., Moore, J., Chakraborty, R., Jin, L., Goodall, J., and Woodruff, D. (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265, 1193–1201.

Muller, H. J. (1950) Our load of mutations. Am. J. Hum. Genet. 2, 111–176.

Nater, A., Arora, N., Greminger, M. P., van Schaik, C. P., Singleton, I., Wich, S. A., Fredriksson, G., Perwitasari-

Farajallah, D., Pamungkas, J., and Krützen, M. (2013) Marked population structure and recent migration in the critically endangered Sumatran orangutan (*Pongo abelii*). J. Hered. **104**, 2–13.

Nei, M., and Li, W. H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76**, 5269–5273.

Nei, M., and Tajima, F. (1981) Genetic drift and estimation of effective population size. Genetics **98**, 625–640.

Ohta, T. (1973) Slightly deleterious mutant substitutions in evolution. Nature **246**, 96–98.

Ohta, T. (1995) Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. J. Mol. Evol. **40**, 56–63.

Osada, N. (2014) Extracting population genetics information from a diploid genome sequence. Front. Ecol. Evol. **2**, doi:10.3389/fevo.2014.00007.

Osada, N., and Akashi, H. (2012) Mitochondrial–nuclear interactions and accelerated compensatory evolution: Evidence from the primate cytochrome *c* oxidase complex. Mol. Biol. Evol. **29**, 337–346.

Osada, N., Uno, Y., Mineta, K., Kameoka, Y., Takahashi, I., and Terao, K. (2010) Ancient genome-wide admixture extends beyond the current hybrid zone between *Macaca fascicularis* and *M. mulatta*. Mol. Ecol. **19**, 2884–2895.

Osada, N., Nakagome, S., Mano, S., Kameoka, Y., Takahashi, I., and Terao, K. (2013) Finding the factors of reduced genetic diversity on X chromosomes of *Macaca fascicularis*: Male-driven evolution, demography, and natural selection. Genetics **195**, 1027–1035.

Osada, N., Kohara, A., Yamaji, T., Hirayama, N., Kasai, F., Sekizuka, T., Kuroda, M., and Hanada, K. (2014) The genome landscape of the African green monkey kidney-derived Vero cell line. DNA Res. **21**, 673–683.

Osada, N., Hettiarachchi, N., Adeyemi Babarinde, I., Saitou, N., and Blancher, A. (2015) Whole-genome sequencing of six Mauritian cynomolgus macaques (*Macaca fascicularis*) reveals a genome-wide pattern of polymorphisms under extreme population bottleneck. Genome Biol. Evol. **7**, 821–830.

Paradis, E., Claude, J., and Strimmer, K. (2004) APE: Analyses of phylogenetics and evolution in R language. Bioinformatics **20**, 289–290.

Payseur, B. A., Jing, P., and Haasl, R. J. (2011) A genomic portrait of human microsatellite variation. Mol. Biol. Evol. **28**, 303–312.

Perelman, P., Johnson, W. E., Roos, C., Seuánez, H. N., Horvath, J. E., Moreira, M. A. M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., et al. (2011) A molecular phylogeny of living primates. PLoS Genet. **7**, e1001342.

Perry, G. H., Melsted, P., Marioni, J. C., Wang, Y., Bainer, R., Pickrell, J. K., Michelini, K., Zehr, S., Yoder, A. D., Stephens, M., et al. (2012) Comparative RNA sequencing reveals substantial genetic variation in endangered primates. Genome Res. **22**, 602–610.

Perry, G. H., Louis, E. E., Ratan, A., Bedoya-Reina, O. C., Burhans, R. C., Lei, R., Johnson, S. E., Schuster, S. C., and Miller, W. (2013) Aye-aye population genomic analyses highlight an important center of endemism in northern Madagascar. Proc. Natl. Acad. Sci. USA **110**, 5823–5828.

Pianka, E. R. (1970) On *r*- and *K*-selection. Am. Nat. **104**, 592–597.

Pluzhnikov, A., and Donnelly, P. (1996) Optimal sequencing strategies for surveying molecular genetic diversity. Genetics **144**, 1247–1262.

Pool, J. E., and Nielsen, R. (2007) Population size changes reshape genomic patterns of diversity. Evolution **61**, 3001–3006.

Prado-Martinez, J., Sudmant, P. H., Kidd, J. M., Li, H., Kelley, J. L., Lorente-Galdos, B., Veeramah, K. R., Woerner, A. E., O'Connor, T. D., Santpere, G., et al. (2013) Great ape genetic diversity and population history. Nature **499**, 471–475.

Prufer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P. H., de Filippo, C., et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. Nature **505**, 43–49.

Reed, D. H., and Frankham, R. (2003) Correlation between fitness and genetic diversity. Conserv. Biol. **17**, 230–237.

Reznick, D., Bryant, M. J., and Bashey, F. (2002) *r*- and *K*-selection revisited: The role of population regulation in life-history evolution. Ecology **83**, 1509–1520.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., et al. (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature **515**, 261–263.

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., et al. (2002) Detecting recent positive selection in the human genome from haplotype structure. Nature **419**, 832–837.

Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T. S., Altshuler, D., and Lander, E. S. (2006) Positive natural selection in the human lineage. Science **312**, 1614–1620.

Saitou, N. (2014) *Homo sapiens* under neutral evolution. Genes and Environment **36**, 99–102.

Sattath, S., Elyashiv, E., Kolodny, O., Rinott, Y., and Sella, G. (2011) Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila simulans*. PLoS Genet. **7**, e1001302.

Scally, A., and Durbin, R. (2012) Revising the human mutation rate: Implications for understanding human evolution. Nat. Rev. Genet. **13**, 745–753.

Simons, Y. B., Turchin, M. C., Pritchard, J. K., and Sella, G. (2014) The deleterious mutation load is insensitive to recent population history. Nat. Genet. **46**, 220–224.

Subramanian, S. (2008) Nearly neutrality and the evolution of codon usage bias in eukaryotic genomes. Genetics **178**, 2429–2432.

Tacutu, R., Craig, T., Budovsky, A., Wuttke, D., Lehmann, G., Taranukha, D., Costa, J., Fraifeld, V. E., and de Magalhaes, J. P. (2013) Human ageing genomic resources: Integrated databases and tools for the biology and genetics of ageing. Nucleic Acids Res. **41**, D1027–D1033.

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

Wang, H. Y., Chien, H. C., Osada, N., Hashimoto, K., Sugano, S., Gojobori, T., Chou, C. K., Tsai, S. F., Wu, C. I., and Shen, C. K. (2007) Rate of evolution in brain-expressed genes in humans and other primates. PLoS Biol. **5**, e13.

Watterson, G. A. (1975) On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7**, 256–276.

Weirauch, M. T., and Hughes, T. R. (2010) Conserved expression without conserved regulatory sequence: The more things change, the more they stay the same. Trends Genet. **26**, 66–74.

Wise, C. A., Sraml, M., Rubinsztein, D. C., and Easteal, S. (1997) Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees. Mol. Biol. Evol. **14**, 707–716.

Yan, G., Zhang, G., Fang, X., Zhang, Y., Li, C., Ling, F., Cooper, D. N., Li, Q., Li, Y., van Gool, A. J., et al. (2011) Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. Nat. Biotech. **29**, 1019–1023.

Yanofsky, C., Horn, V., and Thorpe, D. (1964) Protein structure relationships revealed by mutational analysis. Science **146**, 1593–1594.

Yu, N., Jensen-Seaman, M. I., Chemnick, L., Kidd, J. R., Deinard, A. S., Ryder, O., Kidd, K. K., and Li, W. H. (2003) Low nucleotide diversity in chimpanzees and bonobos. Genetics **164**, 1511–1518.

Yu, N., Jensen-Seaman, M. I., Chemnick, L., Ryder, O., and Li, W. H. (2004) Nucleotide diversity in gorillas. Genetics **166**, 1375–1383.

Zhao, S., Zheng, P., Dong, S., Zhan, X., Wu, Q., Guo, X., Hu, Y., He, W., Zhang, S., Fan, W., et al. (2013) Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nat. Genet. **45**, 67–71.

Zhou, X., Wang, B., Pan, Q., Zhang, J., Kumar, S., Sun, X., Liu, Z., Pan, H., Lin, Y., Liu, G., et al. (2014) Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. Nat. Genet. **46**, 1303–1310.