Estimating the mutation load in human genomes

Brenna M. Henn[1], Laura R. Botigué[1], Carlos D. Bustamante[2], Andrew G. Clark[3] and Simon Gravel[4]

Abstract | Next-generation sequencing technology has facilitated the discovery of millions of genetic variants in human genomes. A sizeable fraction of these variants are predicted to be deleterious. Here, we review the pattern of deleterious alleles as ascertained in genome sequencing data sets and ask whether human populations differ in their predicted burden of deleterious alleles — a phenomenon known as mutation load. We discuss three demographic models that are predicted to affect mutation load and relate these models to the evidence (or the lack thereof) for variation in the efficacy of purifying selection in diverse human genomes. We also emphasize why accurate estimation of mutation load depends on assumptions regarding the distribution of dominance and selection coefficients — quantities that remain poorly characterized for current genomic data sets.

[1]*Department of Ecology and Evolution, Stony Brook University, 650 Life Sciences Building, Stony Brook, New York 11794–5245, USA.*
[2]*Stanford University School of Medicine, Department of Genetics, 291 Campus Drive, Stanford, California 94305, USA.*
[3]*Cornell University, Department of Molecular Biology and Genetics, 526 Campus Road, Ithaca, New York 14853–2703, USA.*
[4]*McGill University, Department of Human Genetics and Genome Quebec Innovation Centre, 740 Dr Penfield Avenue, Montreal, Quebec H3A 0G1, Canada.*
Correspondence to B.M.H.
e-mail: brenna.henn@stonybrook.edu
doi:10.1038/nrg3931
Published online 12 May 2015

The process of mutation constantly creates deleterious variation in a population. These mutations can persist for some time, depending on the intensity of drift and purifying selection. The burden of deleterious variants carried by a population was the subject of classical work in population genetics during the mid-twentieth century and was termed mutation load[1,2]. This mathematical theory described the expected mutation load under idealized genetic models whereby deleterious mutations reduce the reproductive success of carriers compared to a hypothetical genotype with no such deleterious variation. As mutations occur over time, populations accumulate a mutation load compared to a hypothetical population with only the fittest genotypes. A key finding was that very deleterious variants, despite their large potential for damage, tend to be quickly eliminated and rarely rise to high frequencies. By contrast, variants of weaker effect may reach appreciable frequencies owing to random drift and can contribute significantly to mutation load because they affect more individuals in the population[1–3].

The role of genetic drift in these models raises the possibility that different human populations may have varying mutational burden, given the varied patterns of population growth and decline that have characterized different human groups since their initial divergence more than 100,000 years ago[4–7]. Although the theory of genetic load generated strong interest in the 1950s and 1960s, there has been limited opportunity to test these models in the context of human genomics.

In this Review, we synthesize recent work characterizing the frequency of deleterious variants in the human genome and the behaviour of these variants under different demographic models. What are the characteristics of deleterious variants that have been discovered in large-scale sequencing experiments? Do demographic simulations predict differences in mutation load among populations, and how realistic are these models of human demographic history? What other important parameters — such as dominance, epistasis or interaction with the environment — should be considered when calculating the burden of deleterious alleles in each population? Although there have been substantial advances in quantifying how mutation load may vary among human populations, a complete understanding will remain elusive until we can better characterize the relative roles of local adaption and purifying selection for mutations in the human genome.

Models of mutation load

Neutral theory emerged in the context of empirical and theoretical work on genetic load in the mid-twentieth century. At that time, genetic polymorphisms were typically considered to be functional[8]. However, as new protein polymorphism data were generated, much more genetic variation was discovered within and among species than had been previously appreciated. The rate of amino acid substitutions across species phylogenies, estimated at one substitution per genome every 2 years

Genetic load
Reduction in population fitness compared to a theoretical 'perfectly adapted' genotype. This reduction can be caused by the constant influx of new, deleterious mutations (the mutation load), but the genetic load also encompasses reduction in fitness caused by other phenomena, such as inbreeding and changing environment.

Neutral theory
A theory stating that the variation observed within and between species is largely determined by neutral mutation and genetic drift, and not by natural selection. Neutral theory became the basis for many additional population genetic models.

in mammals, was deemed to be too rapid for plausible models of selective evolution[8]; such a rapid rate of adaptation could only be accomplished through the selective deaths of an exceedingly high number of less fit individuals. This substitution load would lead to population decline. Motoo Kimura recognized the significance of the estimated evolutionary rates for genetic loci and instead proposed that the vast majority of polymorphisms were in fact neutral with regard to fitness[1–3]. This shift in worldview to one where neutrality is the dominant factor driving allele frequency change recasts population genetic models in terms of two evolutionary forces: the neutral mutation rate and genetic drift. In this setting, genetic polymorphism is simply "a transient phase of molecular evolution" (REF. 9). As purifying selection quickly eliminates highly deleterious mutations, rates of heterozygosity (θ) simply reflect the product of the mutation rate (μ), the fraction of mutations that are neutral (f), and the effective population size (N_e): that is, $\theta = 4N_e\mu f$. This result only holds if positive selection or weak negative selection is rare. Kimura's focus on genetic drift in a finite population led to an examination of the interaction between genetic drift, natural selection and mutation in determining the accumulation of deleterious alleles in a population.

Tomoko Ohta and Kimura extended the principles of the neutral theory to argue that mutations with very small fitness effects behave effectively as if they were neutral[1]. If a mutation induces a fractional change (s) in the expected number of offspring of carriers, it is effectively neutral if $|s| << 1/N_e$. The evolution of such loci can be accurately modelled using equations involving only drift and mutation. Nearly neutral mutations were defined as a related class of loci with selection coefficients s approximately equal to $1/N_e$. A given mutation with a selection coefficient s that is effectively neutral in a small population can behave nearly neutrally in an intermediate-size population but is eliminated in a large population in which drift is considerably reduced. Its effect on average fitness therefore depends on the population demography. Genetic drift occasionally drives nearly neutral mutations to fixation[10], leading to a decrease in fitness of the entire population, sometimes referred to as drift load[11].

To compare the overall effect of demography on fitness across human populations, we turn to the definition of genetic load (L), the reduction in fitness in a population or species attributable to the presence of alleles that are detrimental in comparison with the genotype that has the maximum fitness (BOX 1): $L = (W_{max} - W_{mean})/W_{max}$. W_{max} is the maximum possible fitness, and W_{mean} is the average fitness of all genotypes in the population. W_{max} is often assigned a value of 1 for algebraic convenience. This definition applies wherever genotypic fitness can be measured or inferred. However, most theoretical results are established under simplifying assumptions of time-independent fitness across generations[12], environmental uniformity and assumptions of additive or multiplicative effects across loci. Even though the maximum fitness W_{max} is easy to identify in idealized models, it is much more challenging to arrive at meaningful empirical estimates in real populations.

Mutation load is the component of genetic load that is attributable to the reduction in fitness caused by recent deleterious mutations. Other components of genetic load include the segregation load, the inbreeding load and the transitory load. Segregation load occurs when a heterozygous genotype has a higher fitness than either of the homozygotes (that is, a heterozygote advantage or overdominance). Inbreeding load occurs when the frequency of homozygous recessive deleterious alleles is increased beyond Hardy–Weinberg expectations as a result of inbreeding[13]. Assuming no selection, the proportion of homozygotes is $\Pr(AA) = Fp + (1-F)p^2$ under inbreeding versus $\Pr(AA) = p^2$ under Hardy–Weinberg equilibrium, where p is the allele frequency and F is the rate of autozygosity. Transitory load occurs while populations adapt to a new fitness landscape and the previously optimal genotype becomes suboptimal. Multiple classes of effects potentially contribute to genetic load in humans, but we focus here on the mutation load and the inbreeding load[14]. In an infinite population, the classical mutation–selection balance

Box 1 | **Summary statistics for mutation load**

A variety of summary statistics have been used to quantify differences in mutation load between human populations. Some studies estimate mutation load by comparing the estimated load per individual in a population. The total number of derived deleterious alleles present in a single individual's genome is a straightforward statistic if an unbiased ancestral genome is available. Under this metric, derived homozygotes are counted twice and heterozygotes are counted once. Under neutrality, each individual is expected to carry the same number of mutations, with some stochasticity owing to the finite genome. There is little evidence of substantive differences between populations in the mean number of deleterious alleles per individual[49,51].

There are several alternative approaches that consider more general statistics of the allele frequency distribution, such as the average frequency among all deleterious alleles, or the proportion of nonsynonymous to synonymous variants. It is fairly straightforward to identify differences across populations in these more general statistics. For example, FIG. 2a shows the site frequency spectrum (SFS) for variants predicted to have a large deleterious effect in four populations: the western African Yoruba (YRI) have a notable excess of low-frequency variants, whereas populations with Out-of-Africa ancestry such as the Japanese (JPT), Tuscans (TSI) and Mexicans (MXL) have an excess of fixed variants. These statistics measure the interaction between selective forces and drift. The analysis finds that the frequency distributions of deleterious variants are different across populations, which has important consequences for the genetic architecture of disease across populations.

Strikingly, we find no published estimates of the mutation load L_T as it is classically defined[2] using human genome sequence data (see above); we do so here for four populations (FIG. 4). Only slight differences are detected across human populations when we consider an additive model (with a dominance coefficient (h) of 0.5). As reflected in the SFS, there are more deleterious variants in the YRI population than in the JPT population (20,672 versus 13,392, respectively), but these rare deleterious variants occur, on average, at lower frequencies. The contribution of large-effect variants to mutation load is slightly higher in the Out-of-Africa population (2.40 in JPT versus 2.37 in YRI) because more of these variants are at higher frequencies in the JPT population. We assume all of the large-effect mutations (genomic evolutionary rate profiling (GERP) score 4:6) to be equally damaging. However, strong differences emerge under a recessive model ($h = 0$) (FIG. 4). In summary, if the mutation load is calculated according to classical models and a distribution of fitness effect (DFE), then differences depend largely on dominance mode. These calculations are laden with multiple assumptions (see main text) and are not suitable for estimating disease prevalence within current populations for public health considerations.

indicates that the expected mutation load at a single site is bounded between μ and 2μ, depending on the level of dominance at that site. Importantly, it does not depend on the selection coefficient at that site: the increased cost per damaging allele is cancelled exactly by the reduction in frequency due to selection. However, most populations, including human populations, are neither infinite nor in mutation–selection balance. The equilibrium results still hold approximately true in finite populations for very deleterious variants, for which mutation and selection are the largest effects. For weakly deleterious (that is, nearly neutral) variants, however, drift becomes more significant and can act to increase the average load.

Dominance

The proportion of deleterious mutations that are recessive, additive or dominant is an open question in human genetics, but characterizing dominance is crucial for evolutionary and medical genetics. The effect of dominance on fitness is quantified by the parameter h, where the fitness of example genotypes *AA*, *Aa* and *aa* are 1, $1-hs$ and $1-s$, respectively. Across loci, there is a distribution of h, with $h=0$ for recessive alleles, $h=0.5$ for additive alleles, and all levels of partial dominance, including outside this (0,1) range. Dominance is perhaps the most important quantity that has not been estimated from genome-wide data. For a large population at equilibrium, a classical prediction is that the load per new deleterious mutation is greater under an additive model than under a recessive model. This is because, by definition, additive mutations exhibit some penetrance, whereas recessive mutations do not. Dominance can lead to substantial differences in load across populations because differences in population history can have a strong impact on the proportion of homozygotes[15,16]. By contrast, the effect of drift on load under an additive model is much weaker. The effect of dominance on load also depends on the frequency of deleterious variants (FIG. 1). New variants are almost exclusively found in heterozygous form, so rare recessive mutations have little impact on load.

Hints about the distribution of dominance come from a range of experimental systems. Mutation-accumulation experiments in model organisms indicate that there is an inverse relationship between dominance and the severity of mutations[17–19]: the more severe a mutation, the lower its dominance coefficient h (that is, the more recessive it is). The average dominance of mildly deleterious mutations across a variety of studied non-human organisms is partially recessive: $h \sim 0.25$ (REF. 20). We also know that there are many recessive mutations that have strong effects in humans and that are particularly evident in consanguineous unions or endogamous populations[21–23], sometimes referred to as inbreeding depression[14]. A recent population genetic approach by Szpiech *et al.*[24] considered long runs of homozygosity (ROH) across human populations and looked at the enrichment of deleterious variants in runs of different length. Long ROH contain more deleterious variants on average than short ROH or homozygotes not in ROH, indicating that the deleterious variants are more likely to lie on long, recent haplotypes. They suggest that long ROH represent recent inbreeding, demonstrating how recent nonrandom mating can exacerbate deleterious effects for recessive loci.

Even mutations of weak effect demonstrate inbreeding depression for human height in European populations[25], which is indicative of a recessive or partially recessive model for the majority of deleterious mutations. Well-validated recessive disease mutations, such as those used in newborn disease screening panels, have been shown to be regularly observed in standard sequencing data sets; as many as 45% of individuals carried at least one recessive allele of strong effect in their exome[26,27]. The X chromosome carries a larger proportion of rare deleterious variation than the autosomes, potentially because recessive alleles are exposed to selection in hemizygous males and are thus kept at lower average frequency than recessive alleles in the autosome[28]. In an attempt to quantify recessive versus dominant diseases, Erickson and Mitchison[29] surveyed

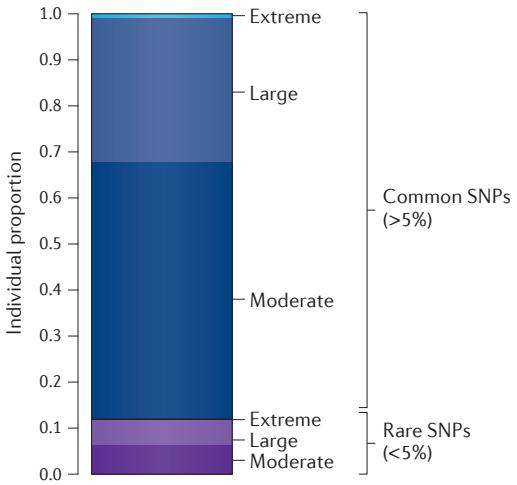

Figure 1 | **Proportion of deleterious variants found in an individual's genome classified by their frequency in the population (common versus rare).** We wanted to ascertain whether the deleterious portion of an individual's genome is mostly represented by rare or common variants. For the Yoruba (YRI) population in the 1000 Genomes Project, variants were assigned to three selection regimes (moderate, large and extreme), according to genomic evolutionary rate profiling (GERP) score categories in increasing order of phylogenetic conservation: 2:4, 4:6 and >6. The more conserved a site is, the more likely it is that a new allele is deleterious (BOX 2). Deleterious single-nucleotide polymorphisms (SNPs) with a derived allele frequency lower than 5% within the population (shown in purple) are classified as 'rare' and the rest as 'common' (shown in blue). Almost 70% of the deleterious SNPs found in an individual genome are common, and most of them have a small predicted effect ('moderate'). Half of the rare SNPs also have a moderate effect, and half of them have a large effect, demonstrating how low-frequency, large-effect variants have not yet been purged by purifying selection.

Substitution load
The difference between optimal fitness and mean fitness in a population undergoing a selective sweep. A locus undergoing a selective sweep will result in some individuals with lower fitness. If multiple loci are under adaptation, the number of individuals who will not reproduce in a generation becomes too large to realistically maintain a stable population. This substitution load puts a limit on the rate of adaptation and is sometimes referred to as Haldane's cost of selection.

Nearly neutral mutations
Variants of relatively weak selective effects that can be accounted for by an extension of neutral theory. Mutations that are slightly deleterious or slightly beneficial will behave as neutral depending on the relationship between population size and the selection coefficient. Because they can reach high frequency, nearly neutral variants can have a large impact on the genetic load, and their evolution is sensitive to fluctuations in population size.

Inbreeding load
Reduction in fitness caused by an excess of recessive homozygotes following inbreeding within a population. The inbreeding load measures the difference between the average fitness of individuals in a population and the fitness of a hypothetical randomly mating population with the same allele frequencies.

Mutation–selection balance
An equilibrium model in which the frequency of an allele is determined by recurrent mutation and the selection coefficient against the allele.

Consanguineous union
In clinical genetics, a union between two individuals who are related as second cousins or closer, with the inbreeding coefficient (F) equal to or higher than 0.0156.

Box 2 | **Properties of private versus shared variants**

A large proportion of rare variants are found to be private to a population, whereas common variants are more likely to be shared across populations (see Supplementary information S3 (figure)). For example, Gravel *et al.*[57] considered the probability that two mutant alleles drawn at random from the global population come from different subpopulations, and they have defined the sharing ratio as the reduction in this probability compared to random mating. This sharing ratio across continental human populations is 0.8 for variants at 30–50% frequency but only 0.1 for variants at 5% frequency. The large number of shared common variants is expected from the relatively low degree of genetic differentiation seen among human populations. The large number of private and rare variants can be explained by neutral forces: rare variants are likely to have occurred during or after population divergence. These variants will be population-specific and found at very low frequencies. This can be compounded by the effect of natural selection, which tends to keep deleterious variants at low frequency and may act differentially across populations. Lohmueller *et al.*[48] and Peischl *et al.*[66] show that the proportion of deleterious alleles is higher among rare variants than among common variants, and that rare variants that are also population-specific are even more likely to be deleterious[45].

Variants that are shared across populations are also typically older and are therefore more likely to be found at higher frequencies at a global scale. Again, this pattern can be largely explained by neutral forces: if these mutations are benign or neutral, they can be maintained over long periods of time and in multiple populations. If these mutations are slightly or even moderately deleterious, they have probably been driven to higher frequencies and spread across populations owing to the increased effect of genetic drift during range expansions in very early human dispersals. Peischl *et al.*[66] have performed simulations that show an increase in frequency of deleterious mutations under a range expansion model and have observed that 10% of common variants shared between Africans and non-Africans are predicted to be deleterious. A similar proportion (14%) of large-effect variants are found to be shared between the eastern African Luhya population and the Finns from northern Europe in the 1000 Genomes Project Phase 1 data[34]. More interestingly, however, these variants are generally found at very high frequencies and actually represent 86% of the total number of large-effect variants in the data set. This scenario is consistent with most of the large-effect variants being private and occurring at low frequencies, and a smaller proportion of variants being shared but common — a view that is compatible with the range expansion model. However, shared large-effect variants can also be driven to high frequencies in cases where they have a beneficial impact on the fitness of the population.

14 diseases for which a variety of genetic susceptibilities have been well characterized and found that disease-associated genes on the autosomes were more likely to be classified as recessive by the Online Mendelian Inheritance in Man (OMIM) database.

Rare variants and deleterious variants

The rise of low-cost, large-scale next-generation sequencing has empowered the study of human genetic variation in ever-larger samples at a genome-wide scale. Most newly discovered genomic variants are found in fewer than 1 in 1,000 people. These rare variants tend to be geographically restricted[30,31] or even restricted to an individual or a family (BOX 2). Rare variants also tend to have emerged more recently than common variants[32]. Compared to common variants, these rare variants are more likely to affect protein composition and to do so in a more disruptive manner, and they are more likely to occur at predicted functional sites[28,33–38]. Furthermore, the lower the frequency of a variant in a sample, the more likely it is to be annotated as deleterious using a variety of variant effect prediction algorithms (see Supplementary information S1 (box)).

Variant effect prediction algorithms attempt to combine available information to predict the effect of a mutation on function (its impact on protein structure and function, expression, degradation and so on) or on evolutionary fitness (that is, the expected number of offspring that a carrier leaves). These effects are distinct but sometimes related: some mutations are strongly evolutionarily deleterious precisely because they have an impact on the protein structure or function of important genes. For example, loss-of-function mutations[33,34,39] disrupt the generation of a fully functional protein either by the introduction of a stop codon or by truncating the reading frame of the protein, and are thus selected against if the gene product is essential. Mutations that affect non-essential genes or that slightly alter protein function or expression are likely to be less evolutionarily deleterious than loss-of-function variants.

Rare, deleterious variants may lead to early-onset diseases and may also inflate an individual's susceptibility to common, complex diseases[37,40–43]. For example, approximately 70% of nonsynonymous variants identified within 200 genes encoding drug targets were estimated to have sufficiently large negative selection coefficients such that these variants would be unlikely to reach even 5% frequency within the current European population[2,36,44]. This may reflect partial ascertainment bias because genes encoding drug targets seem to be under stronger purifying selection than the average gene. However, even in whole-exome sequence data, it is predicted that 47% of single-nucleotide variants detected in large population samples (>3,000)[37] are deleterious. The concordance of predictions of deleterious effects across leading effect prediction algorithms is modest (see Supplementary information S1 (box)); thus, there is still substantial uncertainty regarding the true number of functional or deleterious variants[2,33,37]. Despite this uncertainty, variants of large effect are enriched among rare variants in several populations from the 1000 Genomes Project Phase 1 data set — a conclusion that is independent of the prediction algorithm[37,45] (FIG. 2).

Empirical estimates of load in humans

The distribution and evolution of deleterious mutations are fundamental to understanding the genetic architecture of human disease. Recent studies have asked whether diseases are caused by common variants that are shared across populations, or by rare variants that are specific to a population or family[36,42] (BOX 2). The relative proportion of rare versus common variations contributing to human disease may differ by population, depending on their unique historical modes of population growth or bottlenecks[37,46]. Mutation load provides a framework for quantifying and summarizing the overall effect of population-specific history on deleterious variation. Perhaps more importantly, these recent modelling studies of mutation load highlight the complexity of understanding genetic disease risk using genomic data alone.

All humans carry many deleterious mutations in their genome sequence (FIG. 1). New mutations that enter the gene pool have widely varying impacts on fitness, and we

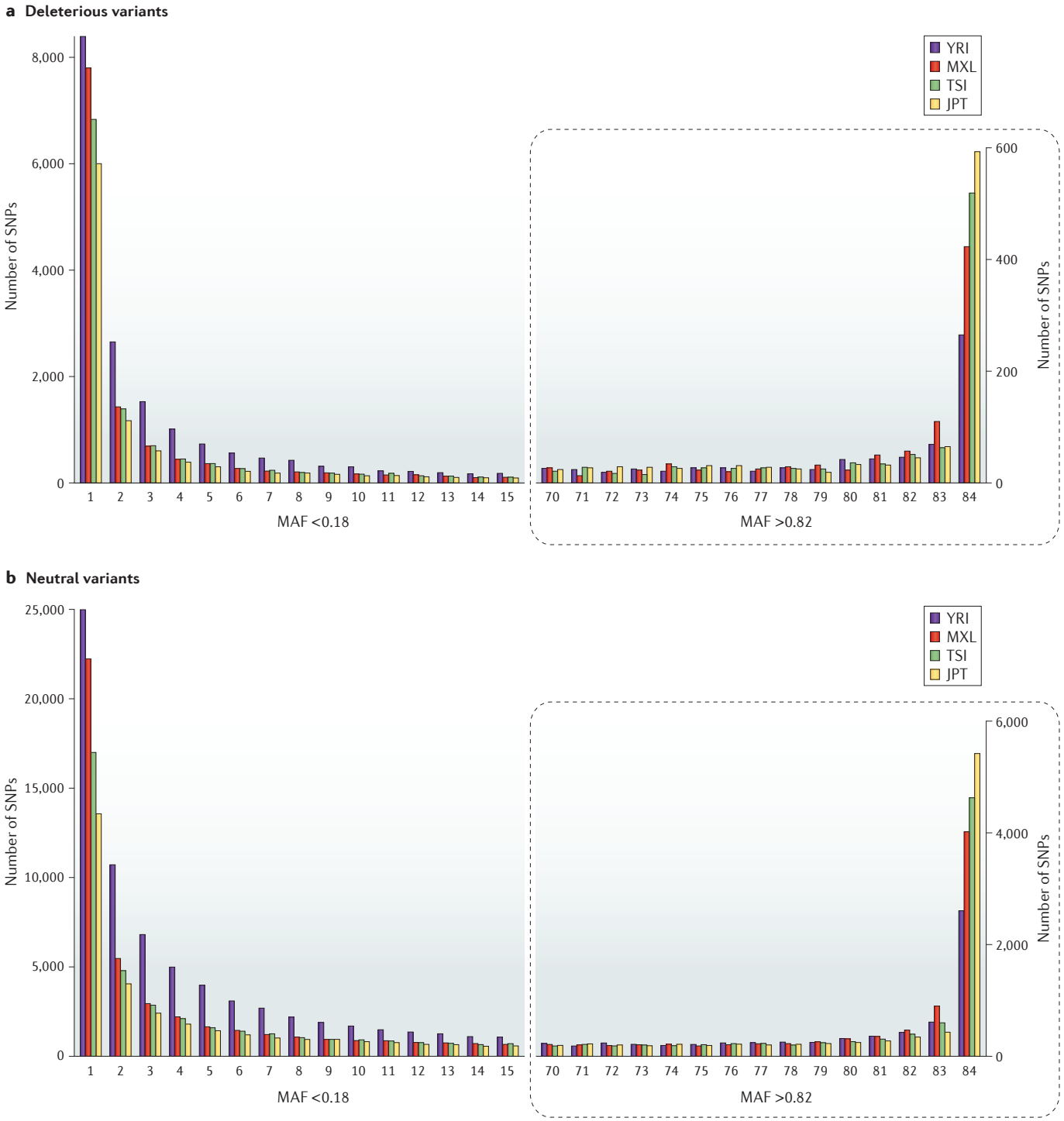

Figure 2 | **Differences in the site frequency spectrum across populations for deleterious and neutral variants.** The site frequency spectrum (SFS) can be a powerful method for summarizing genomic data. The figure shows the SFSs for four populations, focusing on both low-frequency variants (minor allele frequency (MAF) <0.18; left panels) and nearly fixed variants (MAF >0.82; right panels). Derived variants were annotated with genomic evolutionary rate profiling (GERP) scores (see Supplementary information S1 (box)). In part **a**, we plot single-nucleotide polymorphisms (SNPs) that are predicted to have a 'large' deleterious effect (GERP >4). In part **b**, we plot SNPs that are predicted to have a 'neutral' effect (GERP <2). Using 1000 Genomes Project Phase 1 exome data[34], we sampled 42 individuals from the Yoruba (YRI, Nigeria), Mexican (MXL, Mexico), Tuscan (TSI, Italy) and Japanese (JPT, Japan) populations. Only individuals sequenced on the same Agilent exome platform were compared here to avoid biases in target capture between platforms. Demography results in different SFS for each population. Neutral variants provide a null demographic model. The African YRI population have the highest number of rare deleterious variants, although the JPT and TSI populations have many more deleterious fixed variants, possibly owing to ancient founder effects resulting in the fixation by strong drift (also noted in REF. 48). By comparing the difference between the neutral and deleterious SFS (see Supplementary information S2 (figure)), one can infer the impact of purifying selection. For example, non-African populations have a larger proportion of deleterious variants that are fixed than that seen neutrally.

Distribution of fitness effects
(DFE). The distribution of selection coefficients associated with newly arising mutations in a population.

Out-of-Africa
A model by which a small group of modern humans exited Africa approximately 50,000 years ago and dispersed into the Eurasian continents. This movement was accompanied by a severe, possibly tenfold, population bottleneck.

can think of them as being drawn from an underlying distribution of fitness effects (DFE)[47]. An early paper by Morton and colleagues[21] used consanguineous marriages to measure the inbreeding load: the total mutational damage in humans caused by the excess proportion of recessive homozygotes. Total mutational damage was defined as the average number of mutations that would be lethal if they occurred in a homozygous state. The estimated mutational damage was 3–5 lethal equivalent mutations per zygote, and this number was likely to be an underestimate given that this inference relied only on stillbirths and other major pre-reproductive abnormalities but not, for example, on infertility.

Recently, with increased availability of human whole-genome sequences, studies have been able to assay the number of deleterious mutations in larger numbers of samples and across populations, reviving interest in characterizing the human mutation load. However, empirical work has been limited and has primarily considered populations of European and African-American ancestry[26,27,48–50]. In one of the first studies to revisit this topic, Lohmueller *et al.*[48] aimed to address whether human populations carried different numbers of deleterious mutations. Using a set of ~10,000 genes, the authors quantified the total number of damaging single-nucleotide polymorphisms (SNPs) in two population samples ($n = 15$ African-Americans; $n = 20$ European-Americans) and also the per-genome rate in heterozygous versus derived homozygous states[48]. They found a significantly higher proportion of putatively deleterious alleles in the European-American sample than in the African-American sample. Among SNPs that were specific to each population (that is, SNPs only segregating in European-Americans or only segregating in African-Americans), the proportion of predicted damaging mutations was significantly higher in European-Americans (16%) than in African-Americans (12%). However, the total number of deleterious variants was greater in African-Americans because African-Americans carried vastly more neutral variants as well. Under an Out-of-Africa bottleneck (discussed below), Lohmueller *et al.*[48] showed via forward simulations that a severe bottleneck coupled with subsequent population growth could result in the considerable differences in the proportion of deleterious mutations.

Whereas Lohmueller *et al.*[48] hypothesized that this result might be due to reduced efficacy of selection (BOX 3) after the Out-of-Africa bottleneck, this view has been recently contested[51]. Simons *et al.*[49] revisited this question using a larger data set of recently generated exome sequences from individuals of European and African-American ancestry[52]. They contrasted the number of deleterious alleles per individual in each population (regardless of zygosity), annotated either as nonsynonymous or predicted to be damaging by PolyPhen (see Supplementary information S1 (box)). Under this summary of the data, European and African-American individuals carried, on average, similar numbers of deleterious mutations, and there was no significant difference in the average frequency of deleterious mutations. Simons *et al.*[49] concluded that the differences observed by Lohmueller *et al.*[48] did not indicate differences in mutation load. Rather, they suggested that the data can be explained by assuming that each population has the same amount of deleterious variation but that populations differ in how many of these deleterious variants are common and how many are rare. These results were replicated in a smaller sample from the 1000 Genomes Project Phase 1 data; data from Yoruba (Nigeria) and European-American populations (Utah residents, CEU)[34] demonstrated that the lack of difference was not due to recent European admixture in African-Americans.

The difference between the original paper by Lohmueller *et al.*[48] and the paper by Simons *et al.*[49] is largely one of interpretation: despite distinct data sets, the two studies found comparable differences in the distribution of allele frequencies across populations[53]. The disagreement is about whether these differences are informative about the efficacy of selection[54]. However, Fu *et al.*[53] re-analysed the same exome sequence data as Simons *et al.*[49] but instead annotated deleterious variants with a conservation-based algorithm, PhyloP (Supplementary information S1 (box)). Fu *et al.*[53] found significantly more deleterious alleles in Europeans than in African-Americans; these mutations were typically mildly deleterious and many of them were fixed in the European population. Thus, the choice of the functional prediction algorithm can have a large impact on the final interpretation. Further efforts to functionally characterize variants through high-throughput mutagenesis, and the resulting improvement of functional algorithms that incorporate such experimental evidence, will be a huge asset in resolving these long-standing questions.

Recent population history also demonstrates how some deleterious alleles can reach high frequency following a severe bottleneck. Casals *et al.*[50] examined the effect of a strong bottleneck in a French-Canadian population descended from French migrants who settled in the Quebec region in the beginning of the seventeenth

Box 3 | **Efficacy of purifying selection**

Lohmueller *et al.*[48] proposed that different patterns of deleterious variation across populations might be due to differences in the efficacy of selection (specifically, the higher proportion of nonsynonymous to synonymous variants among Europeans), but how can selection be more efficient if the mean number of deleterious mutations per individual, such as between Europeans and African-Americans, is not different (see also REF. 80)? Lohmueller *et al.*[48] estimate the efficacy of selection by comparing it to the effect of drift at a given locus. Given selection coefficient (s), negative selection is more efficient in larger populations because drift is reduced. This definition is inspired by the nearly neutral theory, which proposes that the fixation of deleterious alleles depends crucially on the ability of drift to overcome negative selection at individual sites. However, for rare variation and over short periods of time, this efficacy may have little to do with mean fitness decrease in a population: copies of a recent deleterious allele evolve almost independently from each other and of the population size. If we define the efficiency of selection as its ability to purge deleterious alleles globally, we may not see any appreciable difference between human populations: for short timescales, we can have an equal number of deleterious alleles across populations but differences in drift (that is, in the changes in frequency of these variants). Measuring the efficacy of selection by its effect on load, and by its relative strength versus drift, can lead to markedly different conclusions[54].

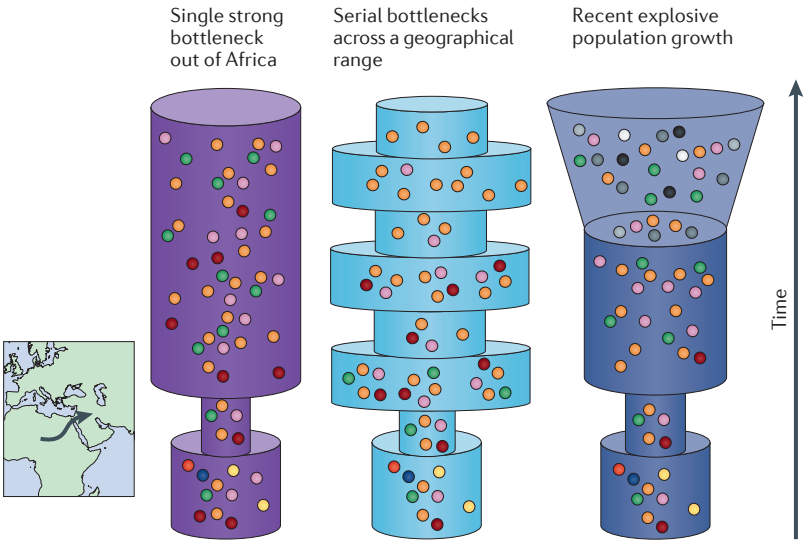

Figure 3 | **Schematic of different demographic models for the Out-of-Africa dispersal.** Three demographic models have been discussed in the context of changes in genetic load due to extreme genetic drift across different human populations. All three models allow for a severe Out-of-Africa bottleneck and recovery but with varying degrees of subsequent changes in population size. Coloured dots indicate allelic diversity; the width of the column is proportional to the effective population size (N_e). The bottom tube represents the ancestral African population size, with later events occurring in temporal sequence towards the top of the figure.

century. Analysis of more than 100 exome sequences identified reduced heterozygosity in French-Canadians compared to the source French population (19% and 12.5%, respectively), as expected under a founder effect. French-Canadians also carry proportionally more missense alleles both at low frequencies and fixed, whereas the French carry proportionally more missense alleles at intermediate frequencies. Casals *et al.*[50] applied the genomic evolutionary rate profiling (GERP) score statistic (see Supplementary information S1 (box)) and found that missense and nonsense mutations in the French-Canadian populations have larger GERP scores, suggesting that they have a more functionally deleterious effect than their counterparts in the French population. In other words, the deleterious variants observed in the French-Canadians show, on average, larger negative selection coefficients. Similarly, the Finnish population also carries a higher proportion of low-frequency loss-of-function variants (which are expected to be highly damaging) than their counterparts in other European countries as a result of a recent population bottleneck[55,56].

Demographic simulations of mutation load

Empirical data from the analysis of human genome sequences seem to support discordant theories of genetic load. Do human populations have differing levels of mutation load? Has purifying selection acted more efficiently in some human populations than in others? Some of this confusion is due to different reported summary statistics (BOX 1), but the debate also centres on simulated results obtained under several idealized demographic models. Multiple simulation efforts have considered three major demographic effects: an Out-of-Africa bottleneck markedly reducing variation in non-African populations; serial founder effects across a geographical range whereby drift is increased during the founding of new populations; and rapid population growth due to recent agricultural and technological changes (FIG. 3). These models are by no means mutually exclusive, and many simulations include bottleneck and growth periods. However, these idealized models can help us to build intuition about the effect of different events on patterns of diversity and examine their effect over time. As we are dealing with a dynamic system rather than a population at equilibrium, some effects are short-lived and others take many generations to evolve before a strong difference is detectable.

Bottleneck. A scenario that is commonly simulated is a classical bottleneck, in which a population experiences a drastic reduction in size before recovering. The most readily observed effect of a bottleneck is an increase in genetic drift, which in turn reduces heterozygosity. The amount of drift depends on the bottleneck intensity: $I = T/(2N_B)$, where T is the duration of the bottleneck and N_B is the effective population size during the bottleneck. A single Out-of-Africa bottleneck model captures the reduction in genomic diversity in populations currently residing outside Africa[16] and is arguably the most noticeable genomic consequence of varying demographic histories in human populations[16,57]. Many studies support a severe bottleneck during the initial colonization of the Eurasian continents, reducing the N_e of the founders to fewer than 1,000–2,500 individuals[4,5,36,46]. Recent coalescent analysis[58,59] based on whole-genome sequences also supports an Out-of-Africa bottleneck with a nearly 15-fold reduction in N_e to only about 1,000 individuals, leading to a higher proportion of recent common ancestry among non-African individuals[4,60]. Interestingly, the coalescent method also suggests a bottleneck of varying magnitudes in African populations at approximately the same time as the Out-of-Africa dispersal. It remains to be determined whether this is due to a reduction in substructure across African populations, parallel bottlenecks during the marine isotope stage 4 (MIS 4) and MIS 5 glacial periods, founder effects during the expansion throughout Africa[61] or other demographic possibilities.

Below, we consider in detail published simulations of the Out-of-Africa bottleneck on load. The single bottleneck models simulated by Lohmueller[48,62] and Simons *et al.*[49] vary in the length of bottleneck considered (T), from instantaneous to 7,700 generations. The Out-of-Africa bottleneck length inferred from genetic data varies from a few hundred generations to 50,000 years[57,60]. Immediately following a deep bottleneck, the number of deleterious polymorphisms decreases, reflecting the fixation of low- and high-frequency variants. However, as the population recovers, the now larger population accumulates rare deleterious variants with DFE that is more similar to that of new mutations. These new variants will tend to have more deleterious effects than the variants

Serial founder effects
Serial expansion of a small ancestral population into a new geographical range; each new deme is created by sampling a small number of individuals to colonize the next location, resulting in a reduced effective population size.

in the pre-bottleneck population. Importantly, during the bottleneck, the reduced population size inflates the role of random genetic drift and allows some deleterious mutations to drift to intermediate and high frequencies. This ultimately leads to more high-frequency and fixed deleterious variants in a population that has undergone a bottleneck (FIG. 2); however, this can be a slow process. The simulations by Simons *et al.*[49] under a single-step bottleneck model illustrate how these two opposing factors can interact: the decrease in the number of deleterious variants is balanced by the increased frequency of the remaining variants. These opposing forces cancel out exactly immediately after the bottleneck. The cancellation is maintained over time for very deleterious variants (as simulated in Simons *et al.*[49]) or very weakly deleterious ones, but alleles that are more moderately deleterious can be more influenced by differential purifying selection[54]. Assuming weaker overall variant effects, Fu *et al.*[53] found that a tenfold bottleneck would increase mutation load by 4.5% relative to a constant-size population, primarily owing to common and fixed variants of weak effect.

Serial founder effects and range expansion. The second strong signature in genetic data from non-African populations is the continuous trend of decreasing genetic diversity (for example, heterozygosity) with increasing distance from eastern Africa[63,64]. This observation can be modelled by serial founder effects in which a small founder population buds off from the ancestral group and contains only a subset of the original diversity as it colonizes a new uninhabited geographical area[33,35]. When described using geographically explicit simulation models, the effect of genetic drift is further exacerbated by the sampling of demes from the wave front of the population as it expands in space (see Moreau *et al.*[65] for a historically documented example). For populations towards the end of the range expansion (for example, Native Americans), demographic history is characterized by many, perhaps hundreds, of bottlenecks followed by population recoveries. Although it is computationally intensive to simulate, the serial founder effect model probably best describes the long period of human population history after the dramatic Out-of-Africa event[33].

Simulations using the serial founder effect show that varying demographic details can result in large differences in genetic load. Peischl *et al.*[66] used spatially explicit forward simulations to examine the effect of extreme drift at the expansion wave front on the pattern of deleterious alleles. These wave front expansions will affect both new mutations and standing variation where drift is especially extreme and alleles can 'surf' to high frequencies rapidly[44]. A range expansion can thus increase the mutation load of a population at the edge of an expansion relative to one in the geographical centre, assuming similar environments and selection coefficients. This effect is particularly pronounced for small selection coefficients and mutations that newly occur either during or after the Out-of-Africa bottleneck. Expansion load in these simulations was particularly sensitive to the carrying capacity (K) — the population size reached before founding a new deme and allowing for migration — with large carrying capacities reducing the probability of fixation[67]. Simulations of the Out-of-Africa serial founder effect, which model K as approximately 1,000, indicate that moderate rates of migration are a good fit to current human heterozygosity[33]. These results would involve a high probability of local fixation for new deleterious mutations, even with selection coefficients as strong as –0.01.

Population growth. We know that the global human population has grown at a prodigious rate; this has been well documented from historical and archaeological records from the past few thousand years (see Supplementary information S2 (figure)). However, what has been appreciated only recently is the magnitude of the impact of this recent growth on the pattern of genetic variation in humans. It is only with large sample sizes that surveys of population variation reveal this impact. The studies of Coventry *et al.*[68], Nelson *et al.*[36], Tennessen *et al.*[37] and Fu *et al.*[52] examined either a selected set of genes or exome sequences in thousands of individuals, and all studies give a consistent picture: larger sample sizes identify many more rare variants than expected under a constant population model. This excess of rare variation, reflected by the nearly fivefold excess of singletons in the sample of >10,000 individuals of Coventry *et al.*[68], is consistent with realistic models of recent population growth. Demographers report global human growth rates on the order of 1–2% per year for the past century or longer; although a lower rate is obtained from genetic data, this is probably because genetic growth rates reflect the N_e.

Population growth stretches back more than just a few thousand years in many regions of the world. Analysis of mitochondrial DNA (mtDNA) lineages shows a strong increase in female effective population size (N_{ef}) during the Holocene in Africa and Eurasia[69,70] compared to Upper Paleolithic lineages. Zheng *et al.*[70] analysed >300 East Asian mitochondrial genomes and found that major mtDNA lineages underwent expansions starting around 13 thousand years ago (kya) and lasting until 4 kya, with changes in N_e from a few hundreds of thousands to millions. Western African populations, far from being the constant population size often portrayed in simulations, have experienced tremendous growth, particularly over the past 5,000 years as Bantu-speaking populations and other groups adopted agriculture[4,37,69]. Europeans also probably experienced a notable bottleneck during the last glacial period starting 21 kya as northern and central Europe became glaciated[71], but populations recovered after 15 kya just before the widespread adoption of agriculture resulted in sustained growth.

Population expansion has a complex effect on the fate of deleterious variation. For example, growth increases the mean survival time of all new mutations in the population, including deleterious ones. However, a longer survival time does not necessarily mean a larger effect on genetic load; variants survive longer largely because they are allowed to exist at lower frequencies after the population has grown. Similarly, population

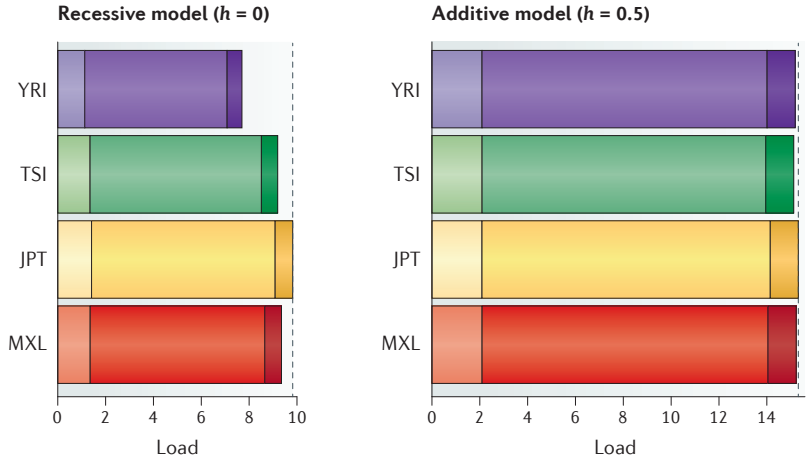

Figure 4 | **Mutation load under an additive and a recessive model.** Using the same data set as in FIG. 2, we computed the total mutation load[2] for each population. Genomic evolutionary rate profiling (GERP) scores were annotated for whole-exome data. Variants were grouped into three categories according their GERP score (2:4, 4:6 and >6), corresponding to different biological functional effects. The more phylogenetically conserved a site is, the more likely it is that a new allele is deleterious and has a high GERP score (see Supplementary information S1 (box)). Within each category, three selection coefficients were assigned, using the inferred *s* coefficients in Boyko *et al.*[47]: $s = -4.5 \times 10^{-4}$, $s = -4.5 \times 10^{-3}$ and $s = -1 \times 10^{-2}$. The total mutation load is the sum of load for each locus[2]. The mutation load under an additive model is higher than the mutation load under a recessive model because the phenotypic effect of a variant is masked in the recessive homozygous state. Although only slight differences exist between populations for an additive model of dominance (~1.5%), strong differences occur under a recessive model because of the differential number of derived homozygotes among populations. JPT, Japanese (Japan); MXL, Mexican (Mexico); TSI, Tuscan (Italy); YRI, Yoruba (Nigeria).

expansions increase the proportion of recent (and rare) mutations in a population, and these rare variants tend to be more deleterious than common variants because of differences in the efficacy of purifying selection between rare and common variants. However, this may not have a large effect on the genetic load because rare variants contribute a small fraction to an individual's overall mutation load[49,72] (FIG. 1).

Future directions

With a few exceptions, the differences in genetic load identified across human populations in recent studies are expected to have little bearing on the health and reproduction of present-day populations, as the current fitness landscape is very different from what it was over the past 100,000 years. Exceptions include populations that have undergone particularly strong bottlenecks and experienced a temporary increase in recessive deleterious variants. Models of mutation load that attempt to accurately quantify differences across populations or to test the predictions of specific theoretical models should consider several other interconnected issues: variant prediction accuracy, spatial variation in selection coefficients and local adaptation.

As noted in Supplementary information S1 (box), almost all the variants considered in these studies are predicted to be deleterious, and most estimates rely on bioinformatic heuristics that are informative but far from perfect. There is substantial discrepancy among methods. One potential issue is the radical misassignment of selection coefficients for adaptive variants. For example, the EDAR-V370A missense mutation is computationally predicted to have a strong effect on a downstream signal transducer, and yet this mutation may be locally adaptive in East Asia for an increased number of eccrine sweat glands[73]. Even if local adaptation is not considered to be a pervasive force in recent human evolution[74], small numbers of adaptive alleles under a selective sweep model will reach high frequency in the population and contribute significantly to the mutation load if they are erroneously annotated as deleterious. One potential way to overcome the complications of local adaptation would be to disregard alleles found at high frequencies, although that would not be appropriate for all non-African populations, as severe genetic drift led to bona fide high-frequency deleterious variants[50] (FIG. 2). Even for alleles that have a more global distribution, the notion that alleles can be assigned an absolute fitness coefficient that is shared among all human groups is untenable. Alleles that are deleterious in some human groups have been shown, in some cases, to be beneficial in others. As an example, it has been shown that children with anaemia have a fourfold higher risk of pneumonia at high altitudes than at lower altitudes[75]. Glucose-6-phosphate dehydrogenase (G6PD) deficiency alleles may have a negative consequence of haemolytic anaemia but nevertheless carry a fitness advantage in areas where malaria is endemic. The genome-wide dependence of fitness coefficients on time and place is largely unknown.

Finally, the above models assume that fitness effects are additive over all loci in the genome, meaning that they ignore the possibility of epistatic interactions. However, in support of a role for epistatic interactions in model organisms, experimental mutation-accumulation experiments often see diminishing returns (that is, negative) epistasis[76,77]. Recent studies have also searched for epistastic interactions in humans in the context of genetic association studies for common diseases, but most of these studies have remained severely underpowered for detecting these effects and would require larger sample sizes. The detection of epistasis is affected by the 'explosion' in the number of tests (there are $n(n-1)/2$ pairwise epistasis tests across n SNPs) and by the fact that these tests require contrasts of phenotypes across multiple genotypic classes and therefore very large sample sizes to recover rare double homozygous genotypes. In short, the absence of numerous reports of epistasis in the genetic architecture of human disease has not provided evidence for a lack of epistasis[78].

Because of such issues, there may be insufficient information in the genomes of all human individuals to accurately infer all the unknown parameters governing human evolution. To answer these questions, we will need to complement genome sequence data with direct experiments measuring the cellular impact of mutations through high-throughput mutagenesis and with experimental evolution in model organisms.

Conclusions

Current human genome sequence data sets and simulations provide conflicting evidence for differences in mutation load across human populations. Various statistics have been used to summarize the distribution of deleterious polymorphisms within populations, and this has contributed to the confusion. Although recent work has emphasized an abundance of deleterious rare variants, rare variants have only a small effect on differences in mutation load between populations (FIG. 1). Moreover, several deleterious mutations may also exist in the non-coding portion of the genome[79], meaning that studies focusing on the analysis of exome sequence data have only studied a small portion of the actual mutation load that may be found in the human genome[38].

Rather than focusing on rare variants, we believe that assessment and simulation of different dominance models is key to understanding the distribution of mutation load across populations. If a fraction of deleterious alleles are recessive, as predicted from disease and model organism studies, then the mutation load observed in bottlenecked Out-of-Africa populations is predicted to be higher than that in African populations owing to the presence of expressed homozygotes (FIG. 4).

In addition to the question of dominance, there are three other areas that require extensive research to understand the phenomenon of mutation load. First, epistasis has a key role in determining complex-trait phenotypes in model organisms and is likely to have a similar role in humans. At the same time, we know little of the underlying mechanisms driving epistatic interactions, especially for rare variants, or the degree to which epistasis in fitness effects influences allele frequency dynamics. Second, the role of local adaptation in human populations from different environments and cultures must be described in order to discriminate between frequent deleterious alleles and frequent beneficial ones. Third, research should focus on integrating bioinformatic and experimental approaches to validate predicted variant effects on the phenotype. An improved partitioning of variants into a DFE would increase the utility of evolutionary and disease models, and is key to improving our understanding of the differences in the architecture of complex diseases across human populations.

1. Ohta, T. & Gillespie, J. Development of neutral and nearly neutral theories. *Theor. Popul. Biol.* **49**, 128–142 (1996).
This paper reviews the development of the neutral and nearly neutral theories by key contributors to the field of population genetics.
2. Kimura, M., Maruyama, T. & Crow, J. F. The mutation load in small populations. *Genetics* **48**, 1303–1312 (1963).
This is a foundational paper on the effect of drift on mutation load in finite populations, demonstrating that mildly deleterious alleles can contribute more to load than strongly deleterious alleles.
3. King, J. L. & Jukes, T. H. Non-Darwinian evolution. *Science* **164**, 788–798 (1969).
4. Marth, G. T., Czabarka, E., Murvai, J. & Sherry, S. T. The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics* **166**, 351–372 (2004).
5. Laval, G., Patin, E., Barreiro, L. B. & Quintana-Murci, L. Formulating a historical and demographic model of recent human evolution based on resequencing data from noncoding regions. *PLoS ONE* **5**, e10284 (2010).
6. Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G. & Siepel, A. Bayesian inference of ancient human demography from individual genome sequences. *Nature Genet.* **43**, 1031–1034 (2011).
7. Veeramah, K. R. *et al.* An early divergence of KhoeSan ancestors from those of other modern humans is supported by an ABC-based analysis of autosomal resequencing data. *Mol. Bio Evol.* **29**, 617–630 (2012).
8. Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).
9. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1985).
10. Ohta, T. Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96–98 (1973).
11. Crow, J. F. Genetic loads and the cost of natural selection. *Math. Top. Popul. Genet.* **1**, 128–177 (1970).
12. Agrawal, A. F. & Whitlock, M. C. Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annu. Rev. Ecol. Evol. Syst.* **43**, 115–135 (2012).
13. Crow, J. F. 2. The concept of genetic load: a reply. *Am. J. Hum. Genet.* **15**, 310–315 (1963).
14. Charlesworth, D. & Willis, J. H. Fundamental concepts in genetics: the genetics of inbreeding depression. *Nature Rev. Genet.* **10**, 783–796 (2009).
This is a broad review of inbreeding depression and heterosis, fitness phenomena that are caused by the presence of deleterious recessive mutations in populations.
15. Li, J. Z. *et al.* Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**, 1100–1104 (2008).
16. Henn, B. M., Cavalli-Sforza, L. L. & Feldman, M. W. The great human expansion. *Proc. Natl Acad. Sci. USA* **109**, 17758–17764 (2012).
17. Agrawal, A. F. & Whitlock, M. C. Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics* **187**, 553–566 (2011).
The distribution of dominance coefficients is directly measured from yeast knockout experiments, showing that large-effect mutations tend to be more recessive than weak-effect mutations.
18. Mukai, T., Chigusa, S. I., Mettler, L. E. & Crow, J. F. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 335–355 (1972).
19. Houle, D., Hughes, K. A., Assimacopoulos, S. & Charlesworth, B. The effects of spontaneous mutation on quantitative traits. II. Dominance of mutations with effects on life-history traits. *Genet. Res.* **70**, 27–34 (1997).
20. Manna, F., Martin, G. & Lenormand, T. Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* **189**, 923–937 (2011).
21. Morton, N. E., Crow, J. F. & Muller, H. J. An estimate of the mutational damage in man from data on consanguienous marriages. *Proc. Natl Acad. Sci. USA* **42**, 855–863 (1956).
This is among the earliest work to empirically measure the mutation load in humans by considering the reduction in fitness due to recessive mutations in consanguineous unions.
22. Reich, D. E. & Lander, E. S. On the allelic spectrum of human disease. *Trends Genet.* **17**, 502–510 (2001).
23. Bittles, A. H. & Black, M. L. Consanguinity, human evolution, and complex diseases. *Proc. Natl Acad. Sci. USA* **107**, 1779–1786 (2010).
24. Szpiech, Z. A. *et al.* Long runs of homozygosity are enriched for deleterious variation. *Am. J. Hum. Genet.* **93**, 90–102 (2013).
25. McQuillan, R. *et al.* Evidence of inbreeding depression on human height. *PLoS Genet.* **8**, e1002655 (2012).
26. Tabor, H. K. *et al.* Pathogenic variants for Mendelian and complex traits in exomes of 6,517 European and African Americans: implications for the return of incidental results. *Am. J. Hum. Genet.* **95**, 183–193 (2014).
Based on analysis of the exome sequences of >6,500 individuals, this study shows that nearly 45% of individuals carry a known variant associated with severe Mendelian diseases.
27. Xue, Y. *et al.* Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am. J. Hum. Genet.* **91**, 1022–1032 (2012).
28. Li, Y. *et al.* Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nature Genet.* **42**, 969–972 (2010).
29. Erickson, R. P. & Mitchison, N. A. The low frequency of recessive disease: insights from ENU mutagenesis, severity of disease phenotype, GWAS associations, and demography: an analytical review. *J. Appl. Genet.* **55**, 319–327 (2014).
30. De la Cruz, O. & Raska, P. Population structure at different minor allele frequency levels. *BMC Proc.* **8**, S55 (2014).
31. Henn, B. M., Gravel, S., Moreno-Estrada, A., Acevedo-Acevedo, S. & Bustamante, C. D. Fine-scale population structure and the era of next-generation sequencing. *Hum. Mol. Genet.* **19**, R221–R226 (2010).
32. Mathieson, I. & McVean, G. Demography and the age of rare variants. *PLoS Genet.* **10**, e1004528 (2014).
33. Deshpande, O., Batzoglou, S., Feldman, M. W. & Luca Cavalli-Sforza, L. A serial founder effect model for human settlement out of Africa. *Proc. Biol. Sci.* **276**, 291–300 (2009).
34. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **490**, 56–65 (2013).
35. DeGiorgio, M., Jakobsson, M. & Rosenberg, N. A. Explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proc. Natl Acad. Sci. USA* **106**, 16057–16062 (2009).
36. Nelson, M. R. *et al.* An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**, 100–104 (2012).
37. Tennessen, J. A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337**, 64–69 (2012).
38. Goode, D. L. *et al.* Evolutionary constraint facilitates interpretation of genetic variation in resequenced human genomes. *Genome Res.* **20**, 301–310 (2010).
39. MacArthur, D. G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).
40. Pritchard, J. K. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **69**, 124–137 (2001).
41. Agarwala, V., Flannick, J., Sunyaev, S., GoT2D Consortium & Altshuler, D. Evaluating empirical bounds on complex disease genetic architecture. *Nature Genet.* **45**, 1418–1427 (2013).

42. Gibson, G. Rare and common variants: twenty arguments. *Nature Rev. Genet.* **13**, 135–145 (2012).
43. Maher, M. C., Uricchio, L. H., Torgerson, D. G. & Hernandez, R. D. Population genetics of rare variants and complex diseases. *Hum. Hered.* **74**, 118–128 (2012).
44. Klopfstein, S. The fate of mutations surfing on the wave of a range expansion. *Mol. Bio. Evol.* **23**, 482–490 (2005).
45. Marth, G. T. *et al.* The functional spectrum of low-frequency coding variation. *Genome Biol.* **12**, R84 (2011).
46. Keinan, A. & Clark, A. G. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* **336**, 740–743 (2012).
47. Boyko, A. R. *et al.* Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* **4**, e1000083 (2008).
This paper estimates selection coefficients for alleles with different predicted deleterious effects in humans and includes a discussion of methods to infer the DFE via site frequency spectra.
48. Lohmueller, K. E. *et al.* Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451**, 994–997 (2008).
This is a formative paper considering the proportion of deleterious mutations in European-Americans compared to African-Americans based on analysis of an early genome sequencing data set. The higher proportion of deleterious variants in European-Americans was ascribed to increased genetic drift during the Out-of-Africa bottleneck.
49. Simons, Y. B., Turchin, M. C., Pritchard, J. K. & Sella, G. The deleterious mutation load is insensitive to recent population history. *Nature Genet.* **46**, 220–224 (2014).
This paper challenges the earlier studies (for example, reference 48) by demonstrating, via simulation, that the average number of deleterious mutations per individual under an additive model should be the same across populations for different human demographic histories.
50. Casals, F. *et al.* Whole-exome sequencing reveals a rapid change in the frequency of rare functional variants in a founding population of humans. *PLoS Genet.* **9**, e1003815 (2013).
51. Do, R. *et al.* No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nature Genet.* **47**, 126–131 (2015).
52. Fu, W. *et al.* Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216–220 (2013).
53. Fu, W., Gittelman, R. M., Bamshad, M. J. & Akey, J. M. Characteristics of neutral and deleterious protein-coding variation among individuals and populations. *Am. J. Hum. Genet.* **95**, 421–436 (2014).
This paper shows that European-American individuals carry slightly more deleterious derived alleles in their genome sequences, on average, than African-Americans under a conservation-based framework to predict variant function; this is consistent with Out-of-Africa bottleneck simulations.
54. Gravel, S. When is selection effective? *bioRXiv* http://dx.doi.org/10.1101/010934 (2014).
55. Lim, E. T. *et al.* Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet.* **10**, e1004494 (2014).
56. Sajantila, A. *et al.* Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *Proc. Natl Acad. Sci. USA* **93**, 12035–12039 (1996).
57. Gravel, S. *et al.* Demographic history and rare allele sharing among human populations. *Proc. Natl Acad. Sci. USA* **108**, 11983–11988 (2011).
58. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
59. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
60. Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H. & Bustamante, C. D. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).
61. Henn, B. M. *et al.* Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc. Natl Acad. Sci.* **108**, 5154–5162 (2011).
62. Lohmueller, K. E. The impact of population demography and selection on the genetic architecture of complex traits. *PLoS Genet.* **10**, e1004379 (2014).
Reprising his earlier work in reference 48, this paper focuses on patterns of deleterious variants over time, given different demographic scenarios of expansion, bottleneck and combinations of demographic events.
63. Ramachandran, S. *et al.* Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc. Natl Acad. Sci. USA* **102**, 15942–15947 (2005).
64. Sousa, V., Peischl, S. & Excoffier, L. Impact of range expansions on current human genomic diversity. *Curr. Opin. Genet. Dev.* **29**, 22–30 (2014).
65. Moreau, C. *et al.* Deep human genealogies reveal a selective advantage to be on an expanding wave front. *Science* **334**, 1148–1150 (2011).
66. Peischl, S., Dupanloup, I., Kirkpatrick, M. & Excoffier, L. On the accumulation of deleterious mutations during range expansions. *Mol. Ecol.* **22**, 5972–5982 (2013).
This is a complex simulation study showing that range expansion can result in expansion load from deleterious mutations that rise to a high frequency on a geographical wave front.
67. Flaxman, S. M. Surfing downhill: when should population range expansion be characterized by reductions in fitness? *Mol. Ecol.* **22**, 5963–5965 (2013).
68. Coventry, A. *et al.* Deep resequencing reveals excess rare recent variants consistent with explosive population growth. *Nature Commun.* **1**, 131–136 (2010).
69. Gignoux, C. R., Henn, B. M. & Mountain, J. L. Rapid, global demographic expansions after the origins of agriculture. *Proc. Natl Acad. Sci. USA* **108**, 6044–6049 (2011).
70. Zheng, H.-X., Yan, S., Qin, Z.-D. & Jin, L. MtDNA analysis of global populations support that major population expansions began before Neolithic time. *Sci. Rep.* **2**, 745 (2012).
71. Forster, P. Ice ages and the mitochondrial DNA chronology of human dispersals: a review. *Phil. Trans. R. Soc. Lond. B.* **359**, 255–264 (2004).
72. Gazave, E., Chang, D., Clark, A. G. & Keinan, A. Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. *Genetics* **195**, 969–978 (2013).
73. Kamberov, Y. G. *et al.* Modeling recent human evolution in mice by expression of a selected *EDAR* variant. *Cell* **152**, 691–702 (2013).
74. Hernandez, R. D. *et al.* Classic selective sweeps were rare in recent human evolution. *Science* **331**, 920–924 (2011).
75. Moschovis, P. P. *et al.* Childhood anemia at high altitude: risk factors for poor outcomes in severe pneumonia. *Pediatrics* **132**, e1156–e1162 (2013).
76. Whitlock, M. C. & Bourguet, D. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution* **54**, 1654–1660 (2000).
77. Fry, J. D. On the rate and linearity of viability declines in *Drosophila* mutation-accumulation experiments: genomic mutation rates and synergistic epistasis revisited. *Genetics* **166**, 797–806 (2004).
78. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl Acad. Sci. USA* **109**, 1193–1198 (2012).
79. Arbiza, L. *et al.* Genome-wide inference of natural selection on human transcription factor binding sites. *Nature Genet.* **45**, 723–729 (2013).
80. Lohmueller, K. E. The distribution of deleterious genetic variation in human populations. *Curr. Opin. Genet. Dev.* **29**, 139–146 (2014).

Acknowledgements
The authors thank K. Lohmueller, S. Peischl and L. Excoffier for discussion regarding these topics.

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
OMIM: http://omim.org

SUPPLEMENTARY INFORMATION
See online article: S1 (box) | S2 (figure) | S3 (figure)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF