nature genetics

No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans

Ron Do[1,2], Daniel Balick[1,3], Heng Li[1,2], Ivan Adzhubei[3], Shamil Sunyaev[1,3] & David Reich[1,2,4]

Non-African populations have experienced size reductions in the time since their split from West Africans, leading to the hypothesis that natural selection to remove weakly deleterious mutations has been less effective in the history of non-Africans. To test this hypothesis, we measured the per-genome accumulation of nonsynonymous substitutions across diverse pairs of populations. We find no evidence for a higher load of deleterious mutations in non-Africans. However, we detect significant differences among more divergent populations, as archaic Denisovans have accumulated nonsynonymous mutations faster than either modern humans or Neanderthals. To reconcile these findings with patterns that have been interpreted as evidence of the less effective removal of deleterious mutations in non-Africans than in West Africans, we use simulations to show that the observed patterns are not likely to reflect changes in the effectiveness of selection after the populations split but are instead likely to be driven by other population genetic factors.

The effectiveness with which natural selection removes deleterious mutations from a population depends not only on the selection coefficient (s) of a mutation but also on the population size (N), which determines the magnitude of the stochastic force of genetic drift. For a constant-size population in equilibrium, effectiveness is fully determined by the product Ns (ref. 1). Because of the dependence on N, the rate at which deleterious alleles are removed from a population depends, in theory, on demographic history. Demographic differences across human populations are well documented. For example, founder events in the last 100,000 years have reduced nucleotide diversity (the number of differences per base pair between paired chromosomes in an individual) in non-Africans by at least 20% relative to West Africans[2–4], reflecting times when the ancestors of non-Africans had relatively smaller population sizes. Similarly, the advent of agriculture in the last 10,000 years has led to rapid population expansions.

[1]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [2]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [3]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [4]Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to S.S. (ssunyaev@rics.bwh.harvard.edu) or D.R. (reich@genetics.med.harvard.edu).

Received 19 February 2014; accepted 9 December 2014; published online 12 January 2015; doi:10.1038/ng.3186

To investigate whether selection has differed in its effectiveness across populations, some studies have contrasted mutation classes thought to be subject to little selection (synonymous mutations in genes) to those potentially subject to purifying selection (nonsynonymous mutations)[5–9]. A key study measured the proportion of polymorphic positions that were nonsynonymous in 20 Europeans and 15 African Americans and showed that, whereas both classes of sites had a reduced rate in Europeans, the reduction was proportionally less for nonsynonymous sites[5]. This pattern was interpreted as being due, in part, to a reduced effectiveness of natural selection against weakly deleterious alleles in Europeans in comparison to West Africans due to the smaller effective population size in Europeans since the separation of these populations[5]. Subsequent studies have confirmed the observation[6,8,10] and have often given a similar interpretation[7–9,11]. What these studies have shown is that there has been an interaction between the forces of natural selection and demographic history that has affected the total number of nonsynonymous polymorphisms[5–11]. However, it does not follow that there have been differences in the effectiveness of selection after the population split. In the common ancestral population of Europeans and West Africans, the average derived allele frequency for nonsynonymous sites would have been lower than for synonymous sites, as negative selection places downward pressure on derived allele frequencies[5]. The different frequency distributions at nonsynonymous and synonymous sites would have responded differently to the bottleneck that then occurred in European populations, simply because the different initial shapes of the distributions would be distorted in different ways by the bottleneck. Thus, the empirically observed differences between Europeans and West Africans could arise independently of differences in the effectiveness of selection after the population split.

RESULTS

No significant differences in the load of deleterious mutations across human populations

The most direct way to contrast the effectiveness of selection between two populations is to sample a single haploid genome from each population, count all the differences and measure which of the two carries an excess. Any genomes that are compared in this manner are by definition separated by the same amount of time since their most recent common ancestor at each chromosomal location. In the absence of selection and assuming no differences in the mutation rate between the two populations, both genomes are expected to harbor the same number of genome-specific mutations. In the presence of selection, however, mutations are removed from the ancestral populations of

Table 1 Accumulation of different classes of mutation in exomes of West African ancestry in comparison to exomes of European ancestry

			R (relative accumulation of mutations)				R^2 (relative accumulation of homozygous mutations)				
Data set	West Africans	Europeans	R (synonymous)	R (all nonsynonymous)	R (benign)	R (possibly damaging)	R (probably damaging)	R^2 (synonymous)	R^2 (all nonsynonymous)	R^2 (benign)	R^2 (possibly damaging)	R^2 (probably damaging)
24 deep genomes	4	4	1.022 (0.012)	1.008 (0.015)	1.002 (0.018)	1.007 (0.040)	1.031 (0.038)	0.652 (0.014)	0.629 (0.019)	0.628 (0.021)	0.602 (0.043)	0.660 (0.047)
Celera exomes	15	20	0.982 (0.011)	1.010 (0.019)	1.011 (0.022)	1.019 (0.043)	0.992 (0.039)	0.610 (0.011)	0.586 (0.047)	0.583 (0.021)	0.616 (0.053)	0.586 (0.047)
1000 Genomes Project exomes	88	85	1.019 (0.010)	0.994 (0.012)	0.999 (0.015)	0.955 (0.028)	1.011 (0.026)	0.655 (0.012)	0.656 (0.014)	0.639 (0.018)	0.599 (0.032)	0.642 (0.032)
ESP exomes	1,088	1,351	1.004 (0.009)	1.001 (0.011)	0.993 (0.013)	1.001 (0.021)	1.037 (0.029)	0.605 (0.010)	0.608 (0.011)	0.598 (0.015)	0.578 (0.025)	0.630 (0.034)

We obtained ±1 standard errors from a weighted-block jackknife with 100 equally sized blocks. For the whole genomes, Yoruba and Mandenka samples represent West Africans and French and Sardinian samples represent Europeans. For the 1000 Genomes Project data, YRI represents West Africans and CEU represents Europeans. The Celera and Exome Sequencing Project (ESP) analyses represent people of West African ancestry using data from African Americans.

each of the two genomes at a rate that depends on the product Ns. Thus, differences in the effectiveness with which selection removes mutations from the two populations can be inferred from a detected asymmetry in the mutation counts between the two genomes. Here we test for differences in the accumulation of mutations between two genomes by sampling one genome from population X and one genome from population Y, determining the ancestral state on the basis of the chimpanzee genome (PanTro2) and recording all the differences. We count the number of derived mutations in genome X that are not seen in genome Y and vice versa, and we define a statistic $R_{X/Y}$ that is a ratio of the two counts. We average over all possible pairs of genomes from the two populations and compute a standard error using a weighted-block jackknife to correct for correlation among neighboring sites (Online Methods)[12]. If selection has been equally effective and mutation rates have been the same since the population split, $R_{X/Y}$ is expected to equal 1. This statistic is monotonically related to the difference in mutation load between the two populations.

We measured $R_{\text{WestAfrican/European}}$ in four sequencing data sets: (i) the coding regions of genes (exomes) in 15 African Americans and 20 European Americans[5]; (ii) exomes from 1,089 individuals in the 1000 Genomes Project[13]; (iii) exomes from 1,088 African Americans and 1,351 European Americans[6]; and (iv) 24 whole genomes sequenced to high coverage[14,15] (**Supplementary Table 1**). We inferred the ancestral allele on the basis of comparison to chimpanzee. As expected for sites unaffected by selection and for indistinguishable differences in mutation rate in the history of the two populations, $R_{\text{WestAfrican/European}}$ (synonymous) was within two standard errors of 1 (**Table 1** and **Supplementary Table 2**). However, $R_{\text{WestAfrican/European}}$ (nonsynonymous) was also indistinguishable from 1. Thus, our data provide no evidence that purging of weakly deleterious mutations has been less effective in Europeans than in West Africans, similar to the finding in ref. 16 for similar population comparisons. To extend these results to a more diverse set of populations, we computed $R_{X/Y}$ for all possible pairs of 11 populations, each represented by 2 deep genome sequences[15], and all pairwise comparisons of 14 populations from the 1000 Genomes Project[13]. We observed no significant differences for any population pair, despite there sometimes being profound differences in demographic history (**Fig. 1** and **Supplementary Table 3**).

To contextualize these null findings, we carried out simulations using fitted models of the histories of West African and European populations[5,6,17] (**Supplementary Table 4**). The simulations showed that, if selection acts additively and coefficients are in the range $s \in [-0.004, -0.0004]$, $R_{\text{WestAfrican/European}}$ is expected to be below 0.95 and detectable given the standard errors of our measurements (**Fig. 2**). However, if many mutations have selection coefficients outside this range, the signal could be diluted to the point of not being detectable. Indeed, when we computed the expected value of $R_{\text{WestAfrican/European}}$ integrating over a previously fitted distribution of selection coefficients[18], we found that $R_{\text{WestAfrican/European}}$ (nonsynonymous) was expected to be 0.987, too close to 1 to be reliably detected given the standard errors of our measurements (**Table 1**). This is consistent with other studies that have concluded that, assuming additively acting mutations, the mutational load in West Africans and Europeans is expected to be indistinguishable when measured on a per-genome basis[16,19–21]. We also simulated recessively acting mutations and in this case predict a stronger difference across populations. The direction of the difference was opposite to that for additively acting mutations, however, reflecting the fact that recessively acting mutations that drift up in frequency owing to a bottleneck can be efficiently purged through the action of selection, as they are exposed in homozygous form[22] (**Fig. 2** and **Supplementary Fig. 1**). The difference in direction for

			IBS (Spanish)	GBR (British)	FIN (Finn)	CEU (European)	JPT (Japanese)	CHS (Chinese)	CHB (Chinese)	PUR (Puerto Rican)	MXL (Mexican)	CLM (Colombian)	YRI (Nigerian)	LWK (Kenyan)	ASW (Afr. Am.)	1KG
		TSI (98)	1.026 (0.005)	1.003 (0.003)	1.003 (0.004)	1 (0.003)	0.998 (0.01)	1.005 (0.011)	1.001 (0.011)	1.017 (0.004)	1.014 (0.006)	1.004 (0.005)	1.005 (0.012)	0.992 (0.011)	1.013 (0.01)	TSI
			IBS (14)	0.978 (0.005)	0.977 (0.005)	0.974 (0.005)	0.974 (0.011)	0.981 (0.011)	0.978 (0.011)	0.993 (0.006)	0.989 (0.008)	0.979 (0.006)	0.986 (0.012)	0.972 (0.012)	0.992 (0.01)	IBS
	Denisova (1)			GBR (89)	0.999 (0.003)	0.996 (0.002)	0.995 (0.01)	1.002 (0.011)	0.998 (0.01)	1.014 (0.005)	1.011 (0.006)	1.001 (0.005)	1.003 (0.012)	0.989 (0.011)	1.01 (0.01)	GBR
Neanderthal	0.857 (0.043)	Neanderthal (1)			FIN (93)	0.997 (0.003)	0.995 (0.01)	1.003 (0.011)	0.999 (0.011)	1.015 (0.005)	1.011 (0.007)	1.001 (0.005)	1.003 (0.013)	0.99 (0.012)	1.011 (0.01)	FIN
Dinka	0.849 (0.035)	0.98 (0.042)	Dinka (2)			CEU (85)	0.998 (0.011)	1.005 (0.011)	1.002 (0.011)	1.018 (0.005)	1.014 (0.007)	1.004 (0.005)	1.006 (0.013)	0.992 (0.012)	1.013 (0.01)	CEU
Mandenka	0.853 (0.038)	0.99 (0.04)	1.014 (0.017)	Mandenka (2)			JPT (89)	1.008 (0.004)	1.004 (0.003)	1.019 (0.009)	1.016 (0.008)	1.006 (0.009)	1.007 (0.013)	0.993 (0.012)	1.014 (0.011)	JPT
Mbuti	0.88 (0.035)	1.03 (0.041)	1.03 (0.016)	1.017 (0.016)	Mbuti (2)			CHS (100)	0.996 (0.002)	1.012 (0.009)	1.009 (0.008)	0.999 (0.009)	1.001 (0.013)	0.988 (0.012)	1.008 (0.011)	CHS
San	0.899 (0.037)	1.05 (0.04)	1.014 (0.016)	1.004 (0.016)	0.99 (0.017)	San (2)			CHB (97)	1.015 (0.01)	1.012 (0.008)	1.002 (0.009)	1.004 (0.013)	0.99 (0.012)	1.011 (0.011)	CHB
Yoruba	0.85 (0.037)	0.993 (0.039)	0.99 (0.018)	0.982 (0.016)	0.965 (0.016)	0.975 (0.017)	Yoruba (2)			PUR (55)	0.996 (0.005)	0.987 (0.003)	0.992 (0.011)	0.978 (0.01)	0.998 (0.008)	PUR
Dai	0.86 (0.036)	1.017 (0.041)	1.012 (0.019)	1.001 (0.021)	0.983 (0.018)	0.999 (0.019)	1.024 (0.019)	Dai (2)			MXL (64)	0.99 (0.004)	0.995 (0.011)	0.981 (0.011)	1.001 (0.009)	MXL
French	0.832 (0.036)	0.971 (0.037)	0.992 (0.019)	0.978 (0.02)	0.961 (0.019)	0.973 (0.019)	0.997 (0.019)	0.972 (0.018)	French (2)			CLM (60)	1.002 (0.011)	0.989 (0.01)	1.01 (0.008)	CLM
Han	0.868 (0.036)	1.01 (0.04)	1.027 (0.018)	1.016 (0.021)	1.00 (0.019)	1.018 (0.019)	1.044 (0.02)	1.021 (0.017)	1.051 (0.019)	Han (2)			YRI (88)	0.986 (0.004)	1.007 (0.004)	YRI
Karitiana	0.833 (0.037)	0.967 (0.034)	0.97 (0.018)	0.961 (0.019)	0.943 (0.018)	0.954 (0.018)	0.975 (0.019)	0.941 (0.019)	0.97 (0.019)	0.923 (0.018)	Karitiana (2)			LWK (96)	1.27 (0.011)	LWK
Papuan	0.85 (0.037)	0.991 (0.039)	1.012 (0.021)	0.999 (0.019)	0.982 (0.02)	0.997 (0.018)	1.021 (0.02)	1.002 (0.02)	1.022 (0.018)	0.976 (0.02)	1.057 (0.021)	Papuan (2)			ASW (61)	
Sardinian	0.859 (0.038)	0.988 (0.038)	0.997 (0.018)	0.985 (0.019)	0.967 (0.018)	0.977 (0.018)	1.006 (0.018)	0.978 (0.017)	1.004 (0.017)	0.958 (0.017)	1.035 (0.022)	0.987 (0.018)	Sardinian (2)			
Deep genomes	Denisova	Neanderthal	Dinka	Mandenka	Mbuti	San	Yoruba	Dai	French	Han	Karitiana	Papuan				

Figure 1 Relative load of nonsynonymous mutations $R_{X/Y}$ for diverse pairs of populations. Results for the deep genomes are given at the bottom left, and results for 1000 Genomes Project populations are given at the top right. Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations for the population in the column. Standard errors (±1; in parentheses) are based on a weighted-block jackknife. We highlight numbers >4 standard errors from expectation. Ratios for Neanderthal and Denisovan samples are normalized by the number of synonymous sites specific to each genome, to adjust for the expectation of fewer mutations in the ancient samples than in the present-day human samples owing to less time elapsed since divergence (all other comparisons are not normalized). Ratios involving Neanderthal and Denisovan samples also remove C→T and G→A substitutions to avoid high error rates due to ancient DNA degradation.

additively and recessively acting mutations suggests that, until there is a reliable model of demographic history and a joint distribution of dominance and selection coefficients in humans, it will be impossible to make a reliable theoretical prediction about whether West Africans or Europeans carry a higher per-genome load.

To boost statistical power to detect differences in the load of nonsynonymous mutations, we stratified nonsynonymous sites in two ways. First, we stratified by predicted functional effect. The PolyPhen-2 (ref. 23) and SIFT[24] algorithms both predict function in a way that is dependent on the ancestral/derived status of allelic variants relative to the human reference genome, which has a particular ancestry at every segment that can bias measurements. We therefore implemented a version of PolyPhen-2 that is independent of the allelic status of the human reference genome (Online Methods) and that fully corrects for the bias (**Supplementary Fig. 2**). We detected no significant differences in the accumulation of deleterious mutations[16] (**Table 1**, **Figure 1** and **Supplementary Tables 2** and **3**). We caution, however, that these tests might also have limited power; for example, when we fit a distribution of additively acting selection coefficients to each of the three PolyPhen functional classes (**Supplementary Note**) and then simulated under these distributions, we predicted ratios of $R_{\text{WestAfrican/European}}$ of 0.984–0.993 (**Supplementary Table 5**), close enough to 1 to not be able to distinguish a significant difference given the standard errors of our measurements. The second way we attempted to boost power was by restricting analysis to locations where pairs of African and non-African individuals share relatively recent common ancestors. We reasoned that this approach might enhance power, as the population split between African and non-African populations occurred only in the last roughly 100,000 years and the mutations that arose before population divergence would be expected to contribute equally to the descendant populations and thus dilute any true signal. We carried out this analysis using four experimentally phased African and six experimentally phased non-African genomes[15]. For each pair of populations, we used the pairwise sequential Markovian coalescent (PSMC) to infer the local time since the most recent common ancestor at each location in the genome (masking the exome to avoid circularly using the same sites for our statistic). When we pooled over sample pairs to increase statistical power and computed $R_{\text{African/non-African}}$ only for the subset of the genome with the shortest times to the most recent common ancestor, we still detected no significant differences between the populations (**Supplementary Table 6**).

A greater load of deleterious mutations in archaic Denisovans than in present-day humans

Tests for differences in the load of nonsynonymous mutations are not always null, as we found when we analyzed the deeply sequenced genomes from an archaic Denisovan[14] and Neanderthal[15] and compared

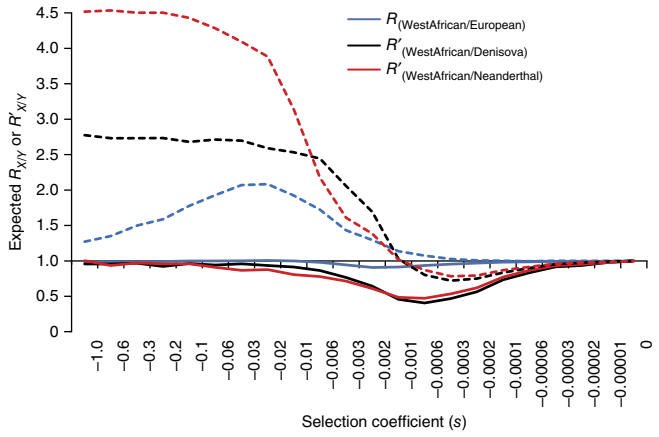


Figure 2 The effect of demographic history on the accumulation of deleterious mutations. To study the expected value of $R_{\text{WestAfrican/European}}$ stratified by selection coefficient, we simulated a previously published model of the joint history of West Africans and Europeans[6], for a range of selection coefficients, assuming both additive ($h = 0.5$) and recessive ($h = 0$) models of selection. For the additive case (solid lines), $R_{\text{WestAfrican/European}}$ dips below a confidently detectable ratio of 0.95 (given the standard errors of our empirical measurements) for $s \in (-0.0004, -0.004)$. Real distributions of selection coefficients may include a large fraction of their density outside this range, and a true signal may thus be difficult to detect. We also simulated a published model of the history of Denisovans and Neanderthals[15]. The simulations predict similar curves for $R'_{\text{WestAfrican/Denisovan}}$ and $R'_{\text{WestAfrican/Neanderthal}}$, reflecting their similar inferred demographic histories (we use a normalized R' statistic to correct for the effects of branch shortening in these ancient genomes). The simulations show that $R'_{\text{WestAfrican/Denisovan}}$ is expected to be below a detectable ratio of 0.95 for $s \in (-0.00002, -0.03)$ and that $R'_{\text{WestAfrican/Neanderthal}}$ is expected to be below 0.95 for $s \in (-0.00002, -0.09)$. For recessively acting alleles (dashed lines), the directionalities of the effects are often opposite.

them to genomes from present-day humans. The ancestors of both archaic individuals are inferred to have maintained relatively small effective population sizes since their main separation from present-day humans, consistent with their levels of genetic diversity having been three to six times smaller[14]. A challenge in comparing the accumulation of mutations in present-day samples to that in ancient samples is that fewer mutations are expected to have occurred in the ancient samples, as they are closer in time to the common ancestor. We therefore divided the accumulation of nonsynonymous mutations specific to each genome by the accumulation of synonymous sites: $R'_{X/Y}$ (nonsynonymous) = $R_{X/Y}$ (nonsynonymous)/$R_{X/Y}$ (synonymous). After removing C→T and G→A mutations in the archaic genomes that might be susceptible to degradation leading to errors in ancient DNA sequences, we inferred that present-day humans had accumulated deleterious mutations at a significantly lower rate than Denisovans since their separation: $R'_{\text{modern/Denisovan}}$ (nonsynonymous) = 0.872 ± 0.034 ($P = 0.0002$) (**Supplementary Table 7**)[14]. However, $R'_{\text{modern/Neanderthal}}$ (nonsynonymous) = 1.037 ± 0.037 is consistent with 1, indicating that deleterious mutations accumulated faster in Denisovan than in Neanderthal ancestors since they separated (**Fig. 1**). The different patterns—despite similar inferred demographic histories—suggest that fitting models to patterns of neutral genetic variation and then simulating under these models might not produce an accurate prediction of the relative effectiveness of selection in pairs of populations (**Fig. 2**). A corollary is that current models for the joint history of West Africans and non-Africans might not produce accurate predictions. Little is known about the duration of the out-of-Africa bottleneck: a short, sharp bottleneck and a long, drawn-out bottleneck are both consistent with most analyses. The primary influence on the cumulative effectiveness of selection is the duration of the bottleneck, and so the uncertainty about its duration is important.

Differences in the effectiveness of biased gene conversion across populations

As a final way to boost power to detect any differences in the effectiveness of the removal of mutations across populations, we analyzed a class of sites that is much larger than the class of nonsynonymous substitutions so that we could make measurements with far smaller standard errors. Specifically, we focused on the class of sites affected by biased gene conversion (BGC), a process in which DNA repair acting on heterozygous sites in gene conversion tracts favors GC over AT alleles[25]. Because BGC only acts on heterozygous sites, it occurs at a rate proportional to heterozygosity, or $2p(1 - p)$ for a mutation of frequency p, exactly mimicking additive selection[25]. We found that $R'^{\text{GC}\to\text{AT}}_{\text{WestAfrican/European}} = 0.995 \pm 0.002$ and $R'^{\text{AT}\to\text{GC}}_{\text{WestAfrican/European}} = 1.00 \pm 0.002$ by using a statistic $R'^{\text{class1}\to\text{class2}}_{X/Y} = R^{\text{class1}\to\text{class2}}_{X/Y} / \left(R^{\text{A}\leftrightarrow\text{T}}_{X/Y} + R^{\text{C}\leftrightarrow\text{G}}_{X/Y}\right)$ that normalizes by the rates of A↔T and C↔G substitutions that are not expected to be affected by BGC. This approach also corrects for possible differences in mutation rate across populations. For the comparison of diverse West Africans to diverse non-Africans, we detected no significant differences (multiple-hypothesis testing–adjusted $P > 0.05$) (**Supplementary Table 8**). The very small standard errors allow us to state with confidence that any differences in the effectiveness of selection in the history of non-Africans and West Africans since the split is likely to have been extremely small. However, San Bushmen from southern Africa had about 1% fewer G/C→A/T mutations than all other humans (significant at up to eight standard errors). To our knowledge, this is the first direct detection of a different rate of accumulation of a class of mutations in one present-day human population in comparison to others. The San are among the most deeply diverged present-day humans, and a possible explanation for our observation is that the longer isolation time for the San would have provided an opportunity for differences in the effectiveness of mutation removal across populations to accumulate to the point of having a detectable effect in the high-sensitivity BGC analysis[26].

Reinterpretation of previous evidence for less effective selection in Europeans

Previous suggestions that weakly deleterious mutations have been removed less effectively in Europeans than in West Africans were largely based on the study of an alternative statistic: the proportion of polymorphic sites in the exome that are nonsynonymous. This statistic is significantly higher in Europeans than in West Africans[5]. We investigated the population genetic forces shaping this statistic by carrying out simulations that allowed us to study the dynamics of this statistic over time. While our simulations showed qualitative patterns that were consistent with those reported previously[5], they also provided new insight owing to a modification to the program that allowed us, in every generation, to dissect how selection (versus mutation and genetic drift) contributed to the expected change in the proportion of nonsynonymous sites in that generation. The simulations showed that, during and after a population bottleneck, the per-generation change in the proportion of segregating sites that were nonsynonymous was not driven by selection being less effective at reducing this ratio than

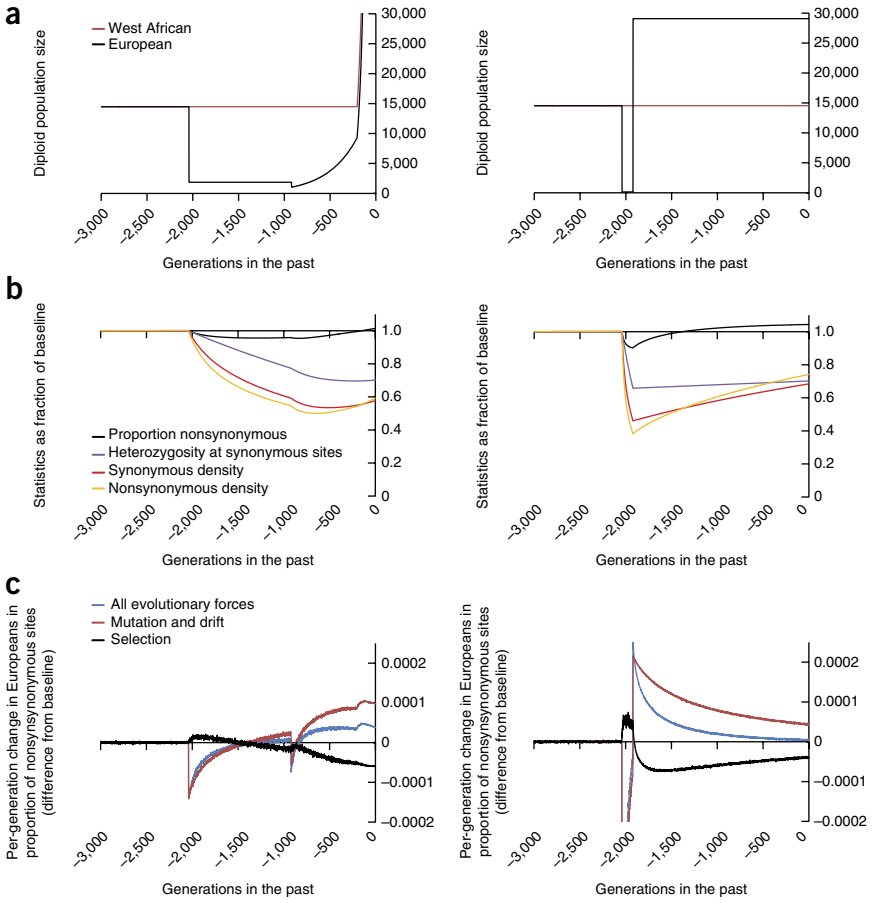


Figure 3 The rise in the proportion of nonsynonymous sites in Europeans in comparison to West Africans is not due to reduced effectiveness of selection in Europeans since the population split. (**a**) The West African and European diploid population sizes for the two simulated models (left (ref. 6) and right, a bottleneck followed by expansion), both of which specify a population split 2,040 generations ago. Subsequent panels are restricted to Europeans, as the West African population size does not fluctuate enough to cause statistics to deviate substantially from the baseline. (**b**) Key statistics as a fraction of the baseline. The present proportion of nonsynonymous sites in Europeans is higher than in the ancestral population (black). We also show heterozygosity at unselected sites (blue), synonymous site density (red) and nonsynonymous site density (yellow). (**c**) Partitioning of the change in the proportion of nonsynonymous sites per generation into selective and other forces. For both models, the temporal dynamics are driven by the forces of mutation and stochastic changes in allele frequency (the curves are positively correlated) and not by negative selective forces (negatively correlated). We plot the per-generation change in the proportion of nonsynonymous mutations due to selection minus its value before the West African–European population split, used as a baseline. A positive value does not mean that selection is working to increase the proportion of nonsynonymous mutations, just that the decrease per generation due to this quantity is less than in the past.

it was in the ancestral population. Instead, after a short period at the start of the bottleneck when the effectiveness of selection in changing this statistic was reduced, selection began to be more effective at reducing the proportion of nonsynonymous sites per generation than it was before the bottleneck (**Fig. 3**). Thus, the rate at which selection reduced the value of this statistic per generation was enhanced rather than diminished by the bottleneck, which for much of the history means that selection pushed the statistic in the direction opposite to that in which it actually moved. We can conclude from this observation that it is primarily non-selective forces that drove the dynamics of this statistic since the separation of West African and European populations.

Intuitively, what explains these simulation results? Before the West African–European split, allele frequencies of nonsynonymous polymorphisms would, on average, have been much lower owing to the depletion of nonsynonymous sites by selection, and the per-site density of nonsynonymous segregating sites would also have been lower. A population entering a bottleneck primarily loses rare alleles, so the nonsynonymous distribution is predicted to be affected more strongly in each generation by the constant flux of new mutations than the synonymous site distribution, as our simulations show. Once the population expands again, the allele frequencies for nonsynonymous sites also adjust faster because the same flux of new mutations into both classes causes a faster rate of replenishment of nonsynonymous sites than synonymous sites, owing to an initially lower density for this class. It is the greater proportional impact of new mutations on nonsynonymous sites per generation that occurs after a bottleneck—because the class has been depleted by the bottleneck—that is driving the observed effects (**Fig. 3**). Putting this another way, we agree with previous reports that have suggested that interactions between the effects of demographic history and natural selection are responsible for the empirically observed differences in the proportion of nonsynonymous segregating sites across human populations[5–11]. However, we differ in the interpretation. Our simulations show that the observed patterns are not driven by a reduced effectiveness of selection at removing slightly deleterious alleles in some human populations in comparison to others since they separated, as has been hypothesized to explain the patterns observed in comparisons of West Africans to non-Africans[5,7], as well as in comparisons of French Canadians[8], Finns[9] and Ashkenazi Jews[11] to European populations that have not experienced recent bottlenecks. Instead, the patterns are driven primarily by new mutation and drift, acting on the different distributions that existed at nonsynonymous and synonymous sites before the population split.

DISCUSSION

It is tempting to interpret the indistinguishable accumulations of deleterious mutations across present-day human populations as implying that the overall genetic burden of disease should be similar for diverse populations. To the extent that mutations act additively, this is correct, as it implies that the complex demographic events of the past are not expected to lead to substantial population differences in the prevalence rates of complex diseases that have an additive genetic architecture[16,19]. However, recessively or epistatically acting mutations work in combination to contribute to disease risk, and, because demography affects allele frequencies, it affects the rate of co-occurrence of alleles. For example, the absolute count of alleles occurring in homozygous form is higher in non-Africans than in Africans for all functional site classes (**Table 1** and **Supplementary Table 9**)[5]. Thus,

the relative risk for diseases with non-additive architectures might be influenced by demography. It will be important to determine the extent to which mutations contributing to phenotypes act non-additively, which will largely determine the extent to which demographic differences among human population affect disease risk.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank J. Akey, D. Altshuler, C. Bustamante, S. Castellano, C. de Filippo, A. Keinan, A. Kondrashov, E. Lander, K. Lohmueller, S. Mallick, S. Pääbo, N. Patterson, J. Pritchard, M. Przeworski, J. Schaiber, G. Sella and M. Slatkin for valuable discussions. R.D. was supported by a Banting fellowship from the Canadian Institutes of Health Research. S.S. was supported by US National Institutes of Health grants R01GM078598 and R01MH101244. D.R. was supported by US National Institutes of Health grants GM100233 and HG006399 and US National Science Foundation grant 1032255 and is an investigator of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

R.D., D.B., H.L., I.A., S.S. and D.R. performed analyses. S.S. and D.R. supervised the research. R.D., D.B., S.S. and D.R. wrote the manuscript with the assistance of all coauthors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Charlesworth, B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* **10**, 195–205 (2009).
2. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
3. Keinan, A., Mullikin, J.C., Patterson, N. & Reich, D. Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat. Genet.* **39**, 1251–1255 (2007).
4. Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G. & Siepel, A. Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034 (2011).
5. Lohmueller, K.E. *et al.* Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451**, 994–997 (2008).
6. Tennessen, J.A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337**, 64–69 (2012).
7. Fu, W. *et al.* Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216–220 (2013).
8. 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).
9. Lim, E.T. *et al.* Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet.* **10**, e1004494 (2014).
10. Kidd, J.M. *et al.* Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. *Am. J. Hum. Genet.* **91**, 660–671 (2012).
11. Carmi, S. *et al.* Sequencing an Ashkenazi reference panel supports population-targeted personal genomics and illuminates Jewish and European origins. *Nat. Commun.* **5**, 4835 (2014).
12. Kunsch, H.R. The jackknife and the bootstrap for general stationary observations. *Ann. Stat.* **17**, 1217–1241 (1989).
13. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
14. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
15. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
16. Simons, Y.B., Turchin, M.C., Pritchard, J.K. & Sella, G. The deleterious mutation load is insensitive to recent population history. *Nat. Genet.* **46**, 220–224 (2014).
17. Gravel, S. *et al.* Demographic history and rare allele sharing among human populations. *Proc. Natl. Acad. Sci. USA* **108**, 11983–11988 (2011).
18. Boyko, A.R. *et al.* Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* **4**, e1000083 (2008).
19. 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).
20. Lohmueller, K.E. The impact of population demography and selection on the genetic architecture of complex traits. *PLoS Genet.* **10**, e1004379 (2014).
21. Lohmueller, K.E. The distribution of deleterious genetic variation in human populations. *bioRxiv* doi:10.1101/005330 (2014).
22. Balick, D.J., Do, R., Reich, D. & Sunyaev, S.R. Response to a population bottleneck can be used to infer recessive selection. *bioRxiv* doi:10.1101/003491 (2014).
23. Adzhubei, I., Jordan, D.M. & Sunyaev, S.R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* Chapter 7 Unit 7.20 (2013).
24. Ng, P.C. & Henikoff, S. Predicting deleterious amino acid substitutions. *Genome Res.* **11**, 863–874 (2001).
25. Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).
26. Schuster, S.C. *et al.* Complete Khoisan and Bantu genomes from southern Africa. *Nature* **463**, 943–947 (2010).

ONLINE METHODS

Data. The data sets we analyzed were published previously and are summarized here. We determined the ancestral allele at each position on the basis of comparison to the chimpanzee genome (PanTro2), except in the case of the Celera data set where we used the previously reported determination[5].

Celera. PCR amplification and Sanger sequencing were performed on 15 African-American and 20 European-American samples over the coding sequences of 10,150 genes. We downloaded ancestral and derived allele counts for 39,440 autosomal SNPs from the supplementary materials of the original study, restricting to sites with genotypes available for both African Americans and European Americans[5].

1000 Genomes Project. A total of 1,089 samples from 14 populations were analyzed in Phase 1 of the 1000 Genomes Project. Illumina-based exome sequencing[13] was performed to ~100× average coverage after solution hybrid capture of the exome[27].

ESP. A total of 1,088 African Americans and 1,351 European Americans were sequenced as part of the National Heart, Lung, and Blood Institute Exome Sequencing Project. Illumina-based exome sequencing was performed to ~100× average coverage after solution hybrid capture of the exome[6].

24 Genomes. This data set included two samples each from six non-African and five sub-Saharan African genomes, an archaic human from Denisova Cave in Siberia sequenced to 31× coverage and an archaic Neanderthal from Denisova Cave in Siberia sequenced to 52× coverage. All sequencing data are based on Illumina technology. We used the version of this data set reported in ref. 15. We only analyzed sites with genotype quality (GQ) scores of ≥45.

Mutation annotation. We annotated coding mutations using ANNOVAR[28], which classifies sites as 'nonsynonymous', 'synonymous', 'stop gain' or 'stop loss'. We subclassified variants using a version of PolyPhen-2 that was created specifically for this study where annotation is independent of the ancestral/derived status of the human genome reference sequence (human-free PolyPhen-2). To guarantee the independence of the PolyPhen-2 predictions from the human genome reference sequence, we modified PolyPhen-2 to rely solely on the multi-species conservation score used in this method[29]. This score reflects the likelihood of observing a given amino acid at a site conditional on the observed pattern of amino acid changes in the phylogeny and is the most informative feature of PolyPhen-2. The predictions in our simplified PolyPhen-2 method are based on the absolute value of the difference in the scores for the two alleles. By construction, this is symmetric with respect to reference/non-reference (and also ancestral/derived and major/minor) allele status. This procedure is similar to the original version of PolyPhen but relies on the PolyPhen-2 homology search and alignment pipeline.

Statistics. We were interested in the expected number of mutations in a randomly sampled haploid exome from one population that were not seen in a randomly sampled comparison exome from another population. To compute this in a situation where we had many exomes available from each population, we did not wish to literally randomly choose a single exome from each population, as this would reduce the sample size in our analysis, resulting in decreased precision of our estimates. Instead, we obtained the expected value if we were to perform an infinite number of random samplings. To compute this value, at each variable site i, we defined d_X^i as the count of the mutant allele at that site in a sample of n_X^i exomes from population X. Similarly, d_Y^i was the count of the mutant allele in a sample of n_Y^i exomes from population Y. The expectation values were obtained by summing over all sites:

$$L_{X,\text{not }Y} = \sum_i (d_X^i / n_X^i)(1 - d_Y^i / n_Y^i)$$

For some analyses, we wished to compute the relative probability that a population was homozygous for a derived allele whereas the other population was not. Thus, we defined an additional statistic, now imposing a correction for limited sample size (because we needed to sample two alleles from each population, we needed to sample without replacement):

$$L^2_{X,\text{not }Y} = \sum_i \frac{2d_X^i\left(n_X^i - d_X^i\right)}{n_X^i\left(n_X^i - 1\right)}\left(1 - \frac{2d_Y^i\left(n_Y^i - d_Y^i\right)}{n_Y^i\left(n_Y^i - 1\right)}\right)$$

We then defined the ratio statistics as follows:

$$R_{X/Y} = L_{X,\text{not }Y}/L_{Y,\text{not }X}$$

$$R^2_{X/Y} = L^2_{X,\text{not }Y}/L^2_{Y,\text{not }X}$$

Weighted-block jackknife to estimate standard errors. We obtained standard errors using a weighted-block jackknife[12]. We divided the SNP data sets into 100 contiguous blocks and then recomputed the statistic on all of the data except for the data from that block. The variation could be converted to a standard error using jackknife theory. We assessed significance on the basis of the number of standard errors from the null expectation of $R_{X/Y} = 1$ and computed a P value using a z score assuming a normal distribution.

Time-stratified computation of the relative accumulation of deleterious mutation. We began with data from ten experimentally phased genomes, all processed identically[15]. These genomes consisted of one each from the populations in **Figure 1** except for the Dinka. We then combined the haploid genomes from 4 African and 6 non-African individuals in all possible pairs to make 96 = (2 × 4) × (2 × 6) pseudo-diploid individuals. We masked the data from the exome and ran PSMC[2] on the data to estimate the time since the most recent common ancestor of the two phased genomes at each location in the genome. We stratified the data into three subsets of inferred time depth and then computed the $R_{\text{African/non-African}}$ statistic within each time-stratified subset (using exomic sites that had been masked from the PSMC analysis so we could independently use these data for downstream analysis).

Analysis of sites susceptible to biased gene conversion. We computed the accumulation of mutations susceptible to BGC for three different substitution classes: G/C→A/T (G→A, G→T, C→A or C→T) mimicking negative selection, A/T→G/C (A→G, T→G, A→C or T→C) mimicking positive selection and A↔T or G↔C (A→T, T→A, C→G or G→C), which we treated as neutral (and used as the denominator of $R_{X/Y}$). For BGC analyses, we used the entire genome, excluding sites in the exome.

The $R'_{X/Y}$ statistic: correcting for branch shortening and differences in mutation rate. For analyses involving samples from the archaic Denisovan and Neanderthal populations, which are many tens of thousands of years old and thus have experienced less evolution from the common ancestor than the present-day humans to whom they were compared, we did not expect $L_{\text{archaic, not-modern}}$ to equal $L_{\text{modern, not-archaic}}$, even for neutral sites. For all analyses involving ancient samples, we normalize both $L_{X,\text{not }Y}$ and $L_{Y,\text{not }X}$ by the accumulation of mutations at sites that were expected to act neutrally (synonymous sites for coding sequences and sites with A/T + C/G for BGC). We then defined:

$$R'^{\text{class}}_{X/Y} = \left(L^{\text{class}}_{X,\text{not }Y} / L^{\text{class}}_{Y,\text{not }X}\right) / \left(L^{\text{normalization}}_{X,\text{not }Y} / L^{\text{normalization}}_{Y,\text{not }X}\right) = R^{\text{class}}_{X/Y} / R^{\text{normalization}}_{X/Y}$$

This $R'_{X/Y}$ statistic not only corrects for branch shortening in the ancient samples but also has the benefit of correcting for any differences in mutation rate that might have arisen in one population or the other since they separated.

Avoiding the confounder of ancient DNA degradation. Ancient DNA data are known to have a high rate of C→T and G→A errors, which persist at a measurable rate even in high-coverage genomes such as those from Denisovan or Neanderthal individuals[13]. In **Supplementary Table 7**, we document that this error process is substantial enough to bias statistics involving Denisovans. We therefore restricted the computation of $R_{X/Y}$ involving the ancient samples to sites that were not C→T and G→A substitutions (for the sake of comparability, we also did not analyze these classes of sites for non-ancient samples in analyses that also involved archaic samples).

Simulations. We wrote a forward simulation program in C that implemented an infinite-sites model. Each mutation was assumed to occur at an unlinked site.

There is an initial burn-in period of 250,000 generations to generate an equilibrium allele frequency spectrum. The simulator samples the allele counts in the current generation on the basis of frequencies in the previous generation, the selection coefficient s, the dominance coefficient h (usually set to additive or $h = 0.5$) and the current population size.

For modeling West African and European history in the main text, we used a demographic model previously fitted to genetic data[6], as well as a simple bottleneck and expansion model (**Supplementary Fig. 1** reports the results for four histories using the simulation parameters shown in **Supplementary Table 4**). For comparisons of West African and archaic population histories, we also used a previously fitted demographic model[15]. We used a mutation rate of 2×10^{-8} mutations per base pair per generation.

At each simulated site, we computed the probability of it being discovered as polymorphic in a sample of size 40, assuming that K_i is the total number of derived alleles for a total of N individuals in the population. We computed the probability that 40 chromosomes were polymorphic at a site i as 1 minus the hypergeometric probability of 0 or 40 derived alleles:

$$\text{Probability that site } i \text{ is segregating} = 1 - \frac{\binom{K_i}{0}\binom{N-K_i}{40}}{\binom{N}{40}} - \frac{\binom{K_i}{40}\binom{N-K_i}{0}}{\binom{N}{40}}$$

We averaged this probability over all simulated positions to obtain the density of segregating sites.

Code availability. The code in C and Perl that was used for the simulations is available on request from D.R.

Integrating over distributions of selection coefficients. For some statistics, we wished to obtain an expected value integrating over distributions of selection coefficients. To achieve this, we carried out simulations for different coefficients. For **Figure 3**, we simulated each of 19 values: $s = \{-1 \times 10^{-6}, -2 \times 10^{-6}, -5 \times 10^{-6}, -1 \times 10^{-5}, -2 \times 10^{-5}, -5 \times 10^{-5}, -1 \times 10^{-4}, -2 \times 10^{-4}, -5 \times 10^{-4}, -1 \times 10^{-3}, -2 \times 10^{-3}, -5 \times 10^{-3}, -0.01, -0.02, -0.05, -0.1, -0.2, -0.5, -1\}$. To compute expected values for $L_{X,\text{ not }Y}$, $L_{Y,\text{ not }X}$ and the density of segregating sites per base pair in a fixed sample size of 40 chromosomes, we used a weighted average of the values of the simulated single selection coefficient statistics. For most analyses, we used weighting based on the distribution of human selection coefficients for the nonsynonymous sites inferred in ref. 18, where the probability of a given value of $-s$ was drawn from a gamma distribution fitted to European genetic data with $\alpha = 0.206$ and $\beta = 15{,}400$. For analyses of the expected value of $R_{\text{WestAfrican/European}}$ stratified by PolyPhen-2 functional class, we used the values inferred in the **Supplementary Note**. Further details of the integration over selection coefficients are given in the **Supplementary Note**.

Partitioning evolutionary dynamics into effects due to selection, mutation and drift. We modified the simulations to sample derived alleles in the next generation at each simulated nucleotide under two alternative assumptions: assuming that all evolutionary forces are operating, and assuming that only non-selective forces are operating.

Let $\text{All}_{s,i,j-1,k}$ be the count of derived alleles that have selection coefficient s at nucleotide position i in the $j-1$ generation in simulation replicate k. We used our simulation machinery to sample the count of derived alleles in the subsequent generation $\text{All}_{s,i,j,k}$, assuming that the selection coefficient in the next generation was the same. We also independently sampled the count of derived alleles in that generation, $\text{NonSel}_{s,i,j,k}$, assuming that selection stopped in that generation ($s = 0$). Because the count of derived alleles was always sampled on the basis of $\text{All}_{s,i,j-1,k}$ in the previous generation (not $\text{NonSel}_{s,i,j-1,k}$), this procedure ensured that the accumulation of derived alleles at each position corresponded to what is expected for an unchanging selection coefficient over time.

To compute the expected value of counts of segregating sites in generation j assuming that all analyzed nucleotides when mutated produce an allele of selection coefficient s, we averaged over A simulation replicates and B simulated nucleotides per replicate:

$$E[\text{All}_{s,j}] = \frac{1}{A \times B}\sum_{i=1}^{A}\sum_{k=1}^{B}\text{All}_{s,i,j,k}$$

$$E[\text{NonSel}_{s,j}] = \frac{1}{A \times B}\sum_{i=1}^{A}\sum_{k=1}^{B}\text{NonSel}_{s,i,j,k}$$

In practice, we also wanted to integrate over a distribution of selection coefficients. Let z_{nonsyn} be the fraction of nucleotides in the genome than when mutated result in a nonsynonymous substitution, which we empirically adjusted to obtain a ratio that matched the data in West Africans (**Supplementary Note**). Let $f(s)$ be the distribution of selection coefficients for *de novo* substitutions (in many of our simulations, we used a distribution fitted to data by ref. 18). We then obtained the expected density of nonsynonymous sites by integrating over the distribution of selection coefficients, which we did in practice by performing a large number of simulations for each of a range of selection coefficients and then grid averaging:

$$E[\text{All}_{\text{nonsyn},j}] = z_{\text{nonsyn}}\int E[\text{All}_{s,j}]f(s)ds$$

$$E[\text{NonSel}_{\text{nonsyn},j}] = z_{\text{nonsyn}}\int E[\text{NonSel}_{s,j}]f(s)ds$$

We also defined the expectation for synonymous sites:

$$E[\text{All}_{\text{syn},j}] = E[\text{NonSel}_{\text{syn},j}] = (1 - z_{\text{nonsyn}})E[\text{All}_{0,j}]$$

We defined the proportion of sites that were nonsynonymous in a given generation as follows:

$$\text{PropAll}_j = \frac{\text{All}_{\text{nonsyn},j}}{\text{All}_{\text{nonsyn},j} + \text{All}_{\text{syn},j}}$$

$$\text{PropNonSel}_j = \frac{\text{NonSel}_{\text{nonsyn},j}}{\text{NonSel}_{\text{nonsyn},j} + \text{NonSel}_{\text{syn},j}}$$

The expected change in the proportion of nonsynonymous sites in generation j is:

$$\delta\text{PropAll}_j = \text{PropAll}_j - \text{PropAll}_{j-1} \text{ (all evolutionary forces)}$$

$$\delta\text{PropNonSel}_j = \text{PropNonSel}_j - \text{PropAll}_{j-1} \text{ (mutation and drift only)}$$

$$\delta\text{PropSel}_j = \text{PropAll}_j - \text{PropNonSel}_j \text{ (selective forces only)}$$

We defined the effectiveness of an evolutionary force in a generation—measured as the magnitude of its effect in that generation on a statistic of interest—by comparing it to the baseline when the population was constant in size (we call this 2,500 generations ago for convenience, as for both the demographic histories we simulated, the 2 populations had not yet split 2,500 generations ago and were in mutation-selection-drift equilibrium):

$$\Delta\text{PropSel}_j = \delta\text{PropSel}_j - \delta\text{PropSel}_{-2,500}$$

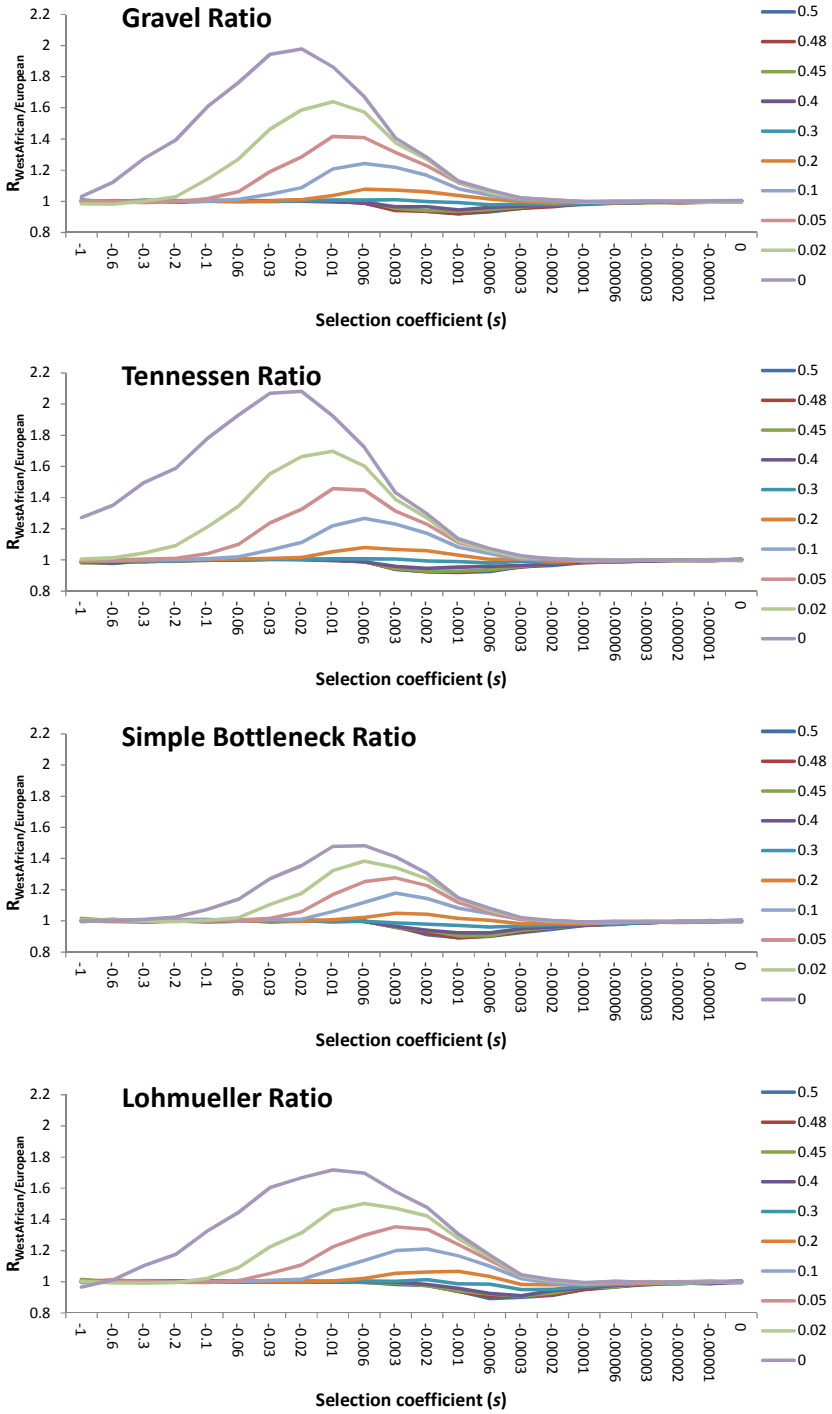
$$\Delta\text{PropNonSel}_j = \delta\text{PropNonSel}_j - \delta\text{PropNonSel}_{-2,500}$$

$$\Delta\text{PropAll}_j = \delta\text{PropAll}_j - \delta\text{PropAll}_{-2,500}$$

These statistics are positive if the effectiveness of the removal of mutations due to an evolutionary force is less than in the ancestral population and negative if the effectiveness is greater than in the ancestral population.

27. Gnirke, A. *et al.* Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **27**, 182–189 (2009).
28. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
29. Sunyaev, S.R. *et al.* PSIC: profile extraction from sequence alignments with position-specific counts of independent observations. *Protein Eng.* **12**, 387–394 (1999).

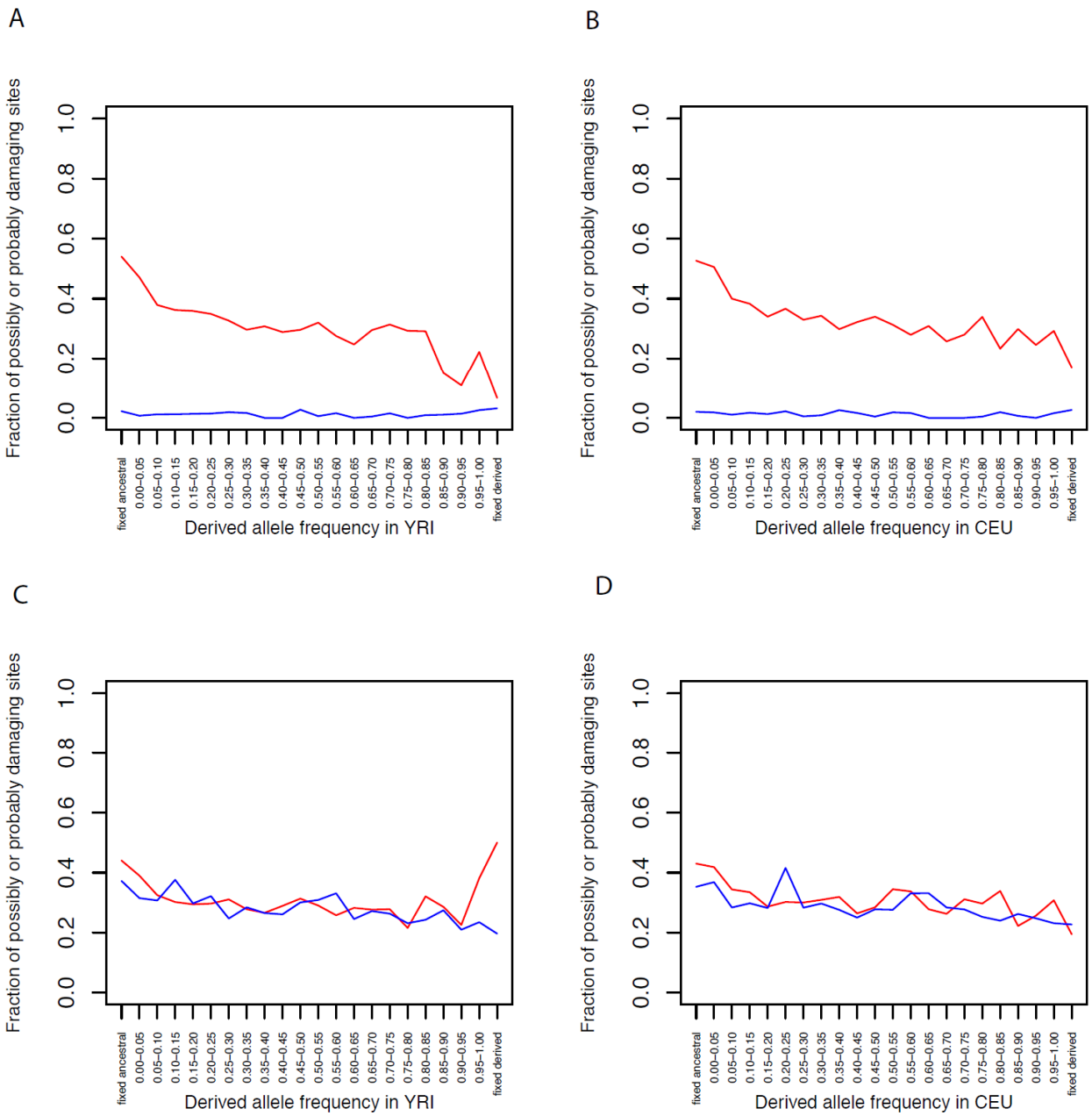
doi:10.1038/ng.3186



Supplementary Figure 1

***R*WestAfrican/European for four demographic histories (simulations).**

We show the expected accumulation of deleterious mutation in West Africans compared with Europeans at the present. We explore a range of selection coefficients *s* and dominance coefficients *h*, for the four models of demographic history specified in **Supplementary Table 4**. We observe a greater accumulation of deleterious mutations in West Africans for recessively acting mutations ($h = 0$) and a greater accumulation in Europeans for additively acting mutations ($h = 0.5$).



Supplementary Figure 2

Our modified version of PolyPhen-2 has no reference bias.
In each of the panels, the *y* axis shows the fraction of nonsynonymous segregating sites in 1000 Genomes Project data that are labeled by PolyPhen-2 as being 'possibly damaging' or 'probably damaging', and the *x* axis shows the derived allele frequency in 1000 Genomes Project European Americans (CEU) or Yoruba Nigerians (YRI). The data are stratified into sites where the human reference sequence allele is ancestral (red) or derived (blue). (**a**,**b**) Standard PolyPhen-2. The probability of being labeled as likely to be damaging is strongly dependent on the status of the human reference sequence. (**c**,**d**) Reference-free PolyPhen-2 has no such dependence.

Supplementary Information for: "No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans"

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 1: Sample sizes in each dataset

Dataset	Population	N
24 diverse genomes[1]	Denisova	1
	Neanderthal	1
	Mbuti	2
	San	2
	Mandenka	2
	Yoruba	2
	Dinka	2
	Papuan	2
	Sardinian	2
	Dai	2
	Karitiana	2
	Han	2
	French	2
Lohmueller[2]	African American	15
	European American	20
1000 Genomes[3]	ASW	61
	CEU	85
	CHB	97
	CHS	100
	CLM	60
	FIN	93
	GBR	89
	IBS	14
	JPT	89
	LWK	96
	MXL	64
	PUR	55
	TSI	98
	YRI	88
Exome Sequencing Project[4]	African American	1,088
	European American	1,351

ASW: African Ancestry in Southwest US; CEU: Utah residents (CEPH) with Northern and Western European ancestry; CHB: Han Chinese in Beijing, China; CHS: Han Chinese South; CLM: Colombian in Medellin, Colombia; FIN: Finnish from Finland; GBR: British from England and Scotland (GBR); IBS: Iberian populations in Spain; JPT: Japanese in Tokyo, Japan; LWK: Luhya in Webuye, Kenya; MXL: Mexican Ancestry in Los Angeles, CA; MXL: Mexican Ancestry in Los Angeles, CA; PUR: Puerto Rican in Puerto Rico; TSI: Toscani in Italia: YRI: Yoruba in Ibadan, Nigeria.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 2: Version of Table 1 for sites with a consistent allele in apes

Data set	West Africans	Europeans	*R:* Relative accumulation of mutations					R^2: Relative accumulation of homozygous mutations				
			R(synonymous)	*R(All non-synonymous)*	*R(Benign)*	*R(Possibly damaging)*	*R(Probably damaging)*	*R^2(synonymous)*	*R^2 (All non-synonymous)*	*R^2 (Benign)*	*R^2 (Possibly damaging)*	*R^2 (Probably damaging)*
24 deep genomes	4	4	1.011 (0.014)	1.015 (0.015)	1.015 (0.019)	0.991 (0.039)	1.038 (0.038)	0.628 (0.015)	0.626 (0.018)	0.630 (0.021)	0.570 (0.043)	0.661 (0.052)
Celera exomes	15	20	0.990 (0.012)	1.012 (0.020)	1.018 (0.023)	1.009 (0.044)	0.991 (0.043)	0.599 (0.011)	0.572 (0.051)	0.585 (0.024)	0.604 (0.054)	0.572 (0.051)
1KG exomes	88	85	1.001 (0.012)	0.987 (0.013)	0.992 (0.016)	0.948 (0.029)	1.003 (0.028)	0.624 (0.013)	0.616 (0.014)	0.613 (0.017)	0.575 (0.031)	0.612 (0.035)
ESP exomes	1,088	1,351	1.007 (0.011)	0.999 (0.012)	0.993 (0.014)	0.984 (0.028)	1.040 (0.030)	0.603 (0.011)	0.594 (0.014)	0.585 (0.016)	0.551 (0.025)	0.628 (0.038)

Notes: This is the same analysis as Table 1, restricting to sites where chimpanzee and at least one of gorilla and orangutan have an allele call and all of the great apes are consistent (data from the EPO six-way primate alignment). ±1 standard errors are from a Weighted Block Jackknife. For the whole genomes, 2 Yoruba + 2 Mandenka represent West Africans, and 2 French + 2 Sardinian represent Europeans. For the 1000 Genomes Data (1KG), YRI represent West Africans and CEU Europeans. The Celera and ESP datasets use African Americans to represent people with West African ancestry.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 3: Expansion of Table 2 into PolyPhen2 classes

Supplementary Table 3A – Synonymous mutations for all pairs of 24 deep genomes (bottom left) and 1000 Genomes populations (top right)

			IBS (Spanish)	GBR (British)	FIN (Finnish)	CEU (European)	JPT (Japanese)	CHS (Chinese)	CHB (Chinese)	PUR (Pu.Ric.)	MXL (Mexican)	CLM (Colom.)	YRI (Nigerian)	LWK (Kenyan)	ASW (Afr.Am.)	1KG
		TSI (98)	1.015 (0.004)	1.002 (0.003)	0.997 (0.004)	0.999 (0.003)	0.988 (0.008)	0.994 (0.009)	0.991 (0.008)	1.002 (0.003)	0.99 (0.005)	0.991 (0.004)	0.981 (0.009)	0.973 (0.008)	0.987 (0.007)	**TSI (Italian)**
			IBS (14)	0.987 (0.004)	0.982 (0.004)	0.984 (0.004)	0.974 (0.008)	0.98 (0.008)	0.977 (0.008)	0.988 (0.005)	0.976 (0.006)	0.976 (0.005)	0.97 (0.009)	0.962 (0.009)	0.975 (0.008)	**IBS (Spanish)**
	Denis-ova (1)			**GBR (89)**	0.995 (0.003)	0.997 (0.002)	0.986 (0.008)	0.992 (0.008)	0.989 (0.008)	1 (0.003)	0.988 (0.005)	0.989 (0.004)	0.979 (0.009)	0.972 (0.008)	0.985 (0.007)	**GBR (British)**
Neand-erthal	n/a	**Neand. (1)**			**FIN (93)**	1.002 (0.003)	0.991 (0.008)	0.997 (0.008)	0.994 (0.008)	1.005 (0.004)	0.993 (0.005)	0.994 (0.004)	0.983 (0.009)	0.976 (0.009)	0.99 (0.007)	**FIN (Finnish)**
Dinka	n/a	n/a	**Dinka (2)**			**CEU (85)**	0.989 (0.008)	0.994 (0.008)	0.992 (0.008)	1.003 (0.004)	0.991 (0.005)	0.991 (0.004)	0.981 (0.009)	0.974 (0.009)	0.988 (0.008)	**CEU (Eur.)**
Mand-enka	n/a	n/a	1.001 (0.013)	**Mand-enka (2)**			**JPT (89)**	1.007 (0.003)	1.004 (0.003)	1.014 (0.007)	1.003 (0.007)	1.003 (0.007)	0.99 (0.01)	0.983 (0.009)	0.997 (0.009)	**JPT (Japan.)**
Mbuti	n/a	n/a	0.99 (0.013)	0.993 (0.012)	**Mbuti (2)**			**CHS (100)**	0.997 (0.002)	1.008 (0.008)	0.997 (0.007)	0.997 (0.007)	0.986 (0.01)	0.978 (0.01)	0.992 (0.009)	**CHS (Chinese)**
San	n/a	n/a	0.979 (0.014)	0.975 (0.014)	0.982 (0.014)	**San (2)**			**CHB (97)**	1.01 (0.007)	0.999 (0.007)	1 (0.007)	0.988 (0.01)	0.98 (0.009)	0.994 (0.008)	**CHB (Chinese)**
Yoruba	n/a	n/a	0.981 (0.013)	0.981 (0.012)	0.99 (0.011)	1.004 (0.014)	**Yoruba (2)**			**PUR (55)**	0.989 (0.004)	0.989 (0.003)	0.979 (0.008)	0.972 (0.008)	0.986 (0.007)	**PUR (Pu.Ric.)**
Dai	n/a	n/a	0.969 (0.015)	0.971 (0.014)	0.978 (0.014)	0.994 (0.014)	0.988 (0.013)	**Dai (2)**			**MXL (64)**	1 (0.004)	0.988 (0.009)	0.981 (0.008)	0.995 (0.008)	**MXL (Mexican)**
French	n/a	n/a	0.966 (0.013)	0.971 (0.014)	0.977 (0.014)	0.991 (0.014)	0.984 (0.012)	0.995 (0.016)	**French (2)**			**CLM (60)**	0.988 (0.008)	0.98 (0.008)	0.995 (0.007)	**CLM (Colomb.)**
Han	n/a	n/a	0.99 (0.017)	0.992 (0.014)	0.996 (0.016)	1.013 (0.015)	1.009 (0.014)	1.029 (0.015)	1.028 (0.017)	**Han (2)**			**YRI (88)**	0.992 (0.003)	1.007 (0.003)	**YRI (Nigerian)**
Karit-iana	n/a	n/a	0.983 (0.017)	0.983 (0.015)	0.986 (0.016)	1.001 (0.016)	0.994 (0.014)	1.012 (0.015)	1.02 (0.017)	0.987 (0.018)	**Karitiana (2)**			**LWK (96)**	1.015 (0.003)	**LWK (Kenyan)**
Papuan	n/a	n/a	0.958 (0.015)	0.962 (0.015)	0.97 (0.015)	0.982 (0.015)	0.977 (0.013)	0.985 (0.016)	0.991 (0.016)	0.959 (0.016)	0.97 (0.016)	**Papuan (2)**			**ASW (61)**	
Sard-inian	n/a	n/a	0.976 (0.013)	0.972 (0.015)	0.978 (0.015)	0.996 (0.013)	0.987 (0.013)	1.004 (0.015)	1.01 (0.013)	0.978 (0.014)	0.991 (0.016)	1.018 (0.015)	**Sardinian (2)**			
Deep genomes	**Denis-ova**	**Neande r-thal**	**Dinka**	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karitiana**	**Papuan**				

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation.

* $R_{X/Y}$ ratios involving the ancient Denisova and Neanderthal samples are not shown as fewer mutations are expected for these genomes than for present-day human genomes since divergence. Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population in the column. The number in parentheses indicates the number of samples per population.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 3B – PolyPhen2 "Benign" mutations for all pairs of 24 deep genomes (bottom left) and 1000 Genomes populations (top right)

			IBS (Spanish)	GBR (British)	FIN (Finn)	CEU (Eur. Am.)	JPT (Jap.)	CHS (Chinese)	CHB (Chinese)	PUR (Pu.Ric.)	MXL (Mex)	CLM (Colom.)	YRI (Nigerian)	LWK (Kenyan)	ASW (Afr. Am.)	1KG
		TSI (98)	1.02 (0.006)	0.998 (0.004)	0.995 (0.005)	0.994 (0.004)	1 (0.013)	1.003 (0.014)	1 (0.014)	1.009 (0.005)	1.007 (0.008)	0.999 (0.006)	0.996 (0.015)	0.987 (0.014)	1.006 (0.012)	**TSI (Italian)**
			IBS (14)	0.979 (0.006)	0.976 (0.007)	0.974 (0.007)	0.983 (0.014)	0.985 (0.014)	0.982 (0.014)	0.991 (0.008)	0.988 (0.01)	0.981 (0.008)	0.982 (0.015)	0.972 (0.014)	0.991 (0.012)	**IBS (Spanish)**
	Denis-ova (1)			**GBR (89)**	0.996 (0.004)	0.995 (0.003)	1.002 (0.014)	1.005 (0.014)	1.001 (0.014)	1.011 (0.006)	1.008 (0.009)	1.001 (0.006)	0.997 (0.015)	0.988 (0.014)	1.008 (0.012)	**GBR (British)**
Neand-erthal	0.88 (0.053)	**Neand. (1)**			**FIN (93)**	0.999 (0.004)	1.005 (0.014)	1.008 (0.014)	1.004 (0.014)	1.014 (0.007)	1.012 (0.009)	1.004 (0.007)	1 (0.016)	0.991 (0.015)	1.011 (0.013)	**FIN (Finnish)**
Dinka	0.858 (0.039)	0.975 (0.044)	**Dinka (2)**			**CEU (85)**	1.006 (0.014)	1.009 (0.015)	1.006 (0.014)	1.015 (0.006)	1.013 (0.009)	1.005 (0.007)	1.001 (0.015)	0.992 (0.014)	1.012 (0.012)	**CEU (European)**
Mand-enka	0.85 (0.043)	0.97 (0.045)	0.996 (0.02)	**Mandenka (2)**			**JPT (89)**	1.003 (0.005)	0.999 (0.004)	1.008 (0.012)	1.006 (0.01)	0.999 (0.011)	0.996 (0.016)	0.986 (0.015)	1.006 (0.014)	**JPT (Japanese)**
Mbuti	0.882 (0.039)	1.016 (0.045)	1.024 (0.021)	1.022 (0.021)	**Mbuti (2)**			**CHS (100)**	0.996 (0.003)	1.005 (0.012)	1.003 (0.01)	0.995 (0.011)	0.993 (0.016)	0.984 (0.015)	1.003 (0.014)	**CHS (Chinese)**
San	0.897 (0.041)	1.03 (0.045)	1.006 (0.019)	1.009 (0.02)	0.988 (0.018)	**San (2)**			**CHB (97)**	1.009 (0.012)	1.007 (0.01)	0.999 (0.011)	0.996 (0.016)	0.987 (0.015)	1.007 (0.014)	**CHB (Chinese)**
Yoruba	0.84 (0.042)	0.963 (0.042)	0.98 (0.021)	0.987 (0.02)	0.963 (0.019)	0.976 (0.02)	**Yoruba (2)**			**PUR (55)**	0.998 (0.006)	0.99 (0.004)	0.989 (0.013)	0.98 (0.012)	0.999 (0.01)	**PUR (Pu.Ric.)**
Dai	0.851 (0.042)	0.982 (0.044)	0.984 (0.022)	0.984 (0.024)	0.965 (0.022)	0.976 (0.023)	1.001 (0.022)	**Dai (2)**			**MXL (64)**	0.992 (0.005)	0.991 (0.014)	0.982 (0.013)	1.001 (0.011)	**MXL (Mexican)**
French	0.844 (0.042)	0.967 (0.039)	0.986 (0.023)	0.987 (0.023)	0.964 (0.024)	0.978 (0.024)	1.004 (0.023)	0.998 (0.023)	**French (2)**			**CLM (60)**	0.997 (0.013)	0.988 (0.012)	1.007 (0.01)	**CLM (Colomb.)**
Han	0.857 (0.041)	0.963 (0.042)	0.998 (0.021)	1 (0.024)	0.981 (0.024)	0.998 (0.023)	1.022 (0.024)	1.022 (0.02)	1.021 (0.024)	**Han (2)**			**YRI (88)**	0.99 (0.004)	1.01 (0.005)	**YRI (Nigerian)**
Karitiana	0.836 (0.043)	0.945 (0.038)	0.95 (0.023)	0.954 (0.022)	0.931 (0.022)	0.942 (0.021)	0.965 (0.023)	0.954 (0.025)	0.955 (0.024)	0.935 (0.022)	**Karit-iana (2)**			**LWK (96)**	1.02 (0.005)	**LWK (Kenyan)**
Papuan	0.858 (0.042)	0.979 (0.042)	1 (0.024)	1.002 (0.023)	0.978 (0.024)	0.992 (0.022)	1.021 (0.023)	1.023 (0.023)	1.012 (0.022)	0.999 (0.025)	1.07 (0.028)	**Papuan (2)**			**ASW (61)**	
Sard-inian	0.864 (0.044)	0.97 (0.041)	0.983 (0.022)	0.992 (0.023)	0.969 (0.023)	0.979 (0.024)	1.009 (0.022)	1.003 (0.023)	1.001 (0.022)	0.987 (0.024)	1.047 (0.029)	0.991 (0.022)	**Sardinian (2)**			
Deep genomes	**Denis-ova**	**Neand-erthal**	**Dinka**	**Mandenka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karit-iana**	**Papuan**				

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation.

* *R*-ratios computed using Denisova and Neanderthal are normalized by the number of synonymous sites on each lineage, to adjust for the fewer mutations in the ancient sample than on present-day human lineages since divergence (the *R´* statistic described in the main text). Ratios involving Neanderthal and Denisova also remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Supplementary Table 7). Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population shown in the column. The number in parentheses indicates the number of samples per population.

Nature Genetics: doi:10.1038/ng.3186

Supp. Table 3C – PolyPhen2 "Possibly damaging" mutations for all pairs of 24 deep genomes (bottom left) and 1000 Genomes pops. (top right)

		IBS (Spanish)	GBR (British)	FIN (Finnish)	CEU European	JPT Japanese	CHS Chinese	CHB Chinese	PUR Pu.Ric.	MXL Mexican	CLM Colom.	YRI Nigerian	LWK Kenyan	ASW Afr. Am.	1KG
	TSI (98)	1.056 (0.013)	1.024 (0.008)	1.009 (0.012)	1.013 (0.008)	0.969 (0.025)	0.985 (0.027)	0.985 (0.026)	1.045 (0.012)	1.015 (0.018)	1.006 (0.012)	1.057 (0.031)	1.026 (0.028)	1.047 (0.024)	**TSI (Italian)**
		IBS (14)	0.969 (0.012)	0.955 (0.013)	0.959 (0.012)	0.921 (0.025)	0.936 (0.026)	0.935 (0.025)	0.991 (0.013)	0.963 (0.018)	0.954 (0.013)	1.014 (0.031)	0.984 (0.028)	1.002 (0.025)	**IBS (Spanish)**
			GBR (89)	0.985 (0.009)	0.989 (0.006)	0.948 (0.025)	0.963 (0.026)	0.963 (0.025)	1.021 (0.011)	0.992 (0.018)	0.983 (0.011)	1.038 (0.031)	1.007 (0.028)	1.027 (0.024)	**GBR (British)**
				FIN (93)	1.005 (0.01)	0.961 (0.024)	0.977 (0.026)	0.977 (0.025)	1.036 (0.012)	1.007 (0.018)	0.998 (0.012)	1.05 (0.031)	1.019 (0.028)	1.04 (0.025)	**FIN (Finnish)**
					CEU (85)	0.958 (0.026)	0.973 (0.028)	0.973 (0.027)	1.031 (0.012)	1.002 (0.018)	0.993 (0.012)	1.046 (0.031)	1.015 (0.028)	1.036 (0.024)	**CEU (European)**
						JPT (89)	1.019 (0.01)	1.018 (0.009)	1.075 (0.025)	1.048 (0.023)	1.037 (0.022)	1.084 (0.033)	1.052 (0.03)	1.075 (0.029)	**JPT (Japanese)**
							CHS (100)	0.999 (0.007)	1.058 (0.025)	1.03 (0.024)	1.02 (0.022)	1.069 (0.033)	1.038 (0.03)	1.06 (0.029)	**CHS (Chinese)**
								CHB (97)	1.058 (0.025)	1.031 (0.023)	1.021 (0.022)	1.07 (0.033)	1.038 (0.031)	1.06 (0.029)	**CHB (Chinese)**
									PUR (55)	0.972 (0.014)	0.964 (0.009)	1.021 (0.026)	0.99 (0.024)	1.009 (0.019)	**PUR (Pu.Ric.)**
										MXL (64)	0.991 (0.011)	1.044 (0.029)	1.014 (0.027)	1.034 (0.023)	**MXL (Mexican)**
											CLM (60)	1.052 (0.028)	1.021 (0.025)	1.041 (0.022)	**CLM (Colomb.)**
												YRI 88)	0.97 (0.008)	0.988 (0.01)	**YRI (Nigerian)**
													LWK (96)	1.018 (0.011)	**LWK (Kenyan)**
														ASW (61)	

	Denisova (1)													
Neander-thal	0.769 (0.07)	**Neand-erthal(1)**												
Dinka	0.749 (0.06)	0.936 (0.071)	**Dinka (2)**											
Mand-enka	0.788 (0.066)	0.982 (0.074)	1.075 (0.047)	**Mand-enka (2)**										
Mbuti	0.812 (0.064)	1.025 (0.079)	1.043 (0.05)	0.983 (0.044)	**Mbuti (2)**									
San	0.832 (0.064)	1.054 (0.078)	1.043 (0.044)	0.982 (0.038)	1.005 (0.042)	**San (2)**								
Yoruba	0.793 (0.063)	1.006 (0.073)	1.017 (0.039)	0.954 (0.036)	0.979 (0.041)	0.966 (0.035)	**Yoruba (2)**							
Dai	0.807 (0.066)	1.063 (0.086)	1.075 (0.048)	1.022 (0.047)	1.028 (0.052)	1.049 (0.051)	1.059 (0.048)	**Dai (2)**						
French	0.72 (0.056)	0.925 (0.071)	1.013 (0.048)	0.954 (0.044)	0.968 (0.045)	0.966 (0.044)	0.994 (0.048)	0.931 (0.042)	**French (2)**					
Han	0.815 (0.066)	1.054 (0.088)	1.126 (0.049)	1.055 (0.049)	1.071 (0.05)	1.078 (0.049)	1.102 (0.051)	1.061 (0.048)	1.124 (0.051)	**Han (2)**				
Karitiana	0.763 (0.066)	0.967 (0.082)	1.068 (0.052)	1.013 (0.048)	1.008 (0.051)	1.022 (0.049)	1.034 (0.048)	0.976 (0.048)	1.04 (0.053)	0.925 (0.051)	**Karit-iana (2)**			
Papuan	0.759 (0.067)	0.981 (0.086)	1.071 (0.05)	1.008 (0.046)	1.019 (0.047)	1.037 (0.05)	1.05 (0.05)	0.998 (0.049)	1.068 (0.049)	0.952 (0.045)	1.012 (0.054)	**Papuan (2)**		
Sardinian	0.767 (0.064)	0.985 (0.076)	1.031 (0.051)	0.968 (0.042)	0.965 (0.043)	0.975 (0.043)	1.013 (0.046)	0.937 (0.042)	1.003 (0.037)	0.881 (0.041)	0.961 (0.05)	0.946 (0.043)	**Sardinian (2)**	
Deep genomes	**Denisova**	**Neand-erthal**	**Dinka**	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karit-iana**	**Papuan**		

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation.

* *R*-ratios computed using Denisova and Neanderthal are normalized by the number of synonymous sites on each lineage, to adjust for the fewer mutations in the ancient sample than on present-day human lineages since divergence (the *R´* statistic described in the main text). Ratios involving Neanderthal and Denisova also remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Supplementary Table 7). Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population shown in the column. The number in parentheses indicates the number of samples per population.

Nature Genetics: doi:10.1038/ng.3186

Supp. Table 3D – PolyPhen2 "Probably damaging" mutations for all pairs of 24 deep genomes (bottom left) and 1000 Genomes pops. (top right)

			IBS (Spanish)	GBR (British)	FIN (Finnish)	CEU European	JPT Japanese	CHS Chinese	CHB Chinese	PUR Pu.Ric.	MXL Mexican	CLM Colomb.	YRI Nigerian	LWK Kenyan	ASW Afr. Am.	1KG
		TSI (98)	1.025 (0.013)	1.004 (0.007)	1.03 (0.009)	1.012 (0.007)	1.014 (0.025)	1.029 (0.025)	1.024 (0.025)	1.026 (0.009)	1.042 (0.018)	1.022 (0.012)	0.998 (0.026)	0.981 (0.024)	1.008 (0.021)	**TSI (Italian)**
			IBS (14)	0.979 (0.013)	1.005 (0.014)	0.987 (0.013)	0.991 (0.028)	1.006 (0.028)	1.001 (0.028)	1.002 (0.014)	1.018 (0.022)	0.998 (0.016)	0.978 (0.026)	0.962 (0.025)	0.987 (0.021)	**IBS (Spanish)**
	Denis-ova (1)			**GBR (89)**	1.026 (0.008)	1.008 (0.007)	1.01 (0.026)	1.026 (0.026)	1.021 (0.026)	1.023 (0.01)	1.039 (0.019)	1.018 (0.013)	0.995 (0.027)	0.978 (0.025)	1.005 (0.022)	**GBR (British)**
Neander-thal	0.841 (0.099)	**Neand-ertal(1)**			**FIN (93)**	0.983 (0.008)	0.986 (0.024)	1.001 (0.025)	0.996 (0.025)	0.998 (0.011)	1.014 (0.017)	0.993 (0.012)	0.975 (0.026)	0.958 (0.024)	0.984 (0.021)	**FIN (Finnish)**
Dinka	0.918 (0.07)	1.048 (0.091)	**Dinka (2)**			**CEU (85)**	1.002 (0.025)	1.018 (0.025)	1.013 (0.025)	1.015 (0.01)	1.03 (0.018)	1.01 (0.013)	0.988 (0.026)	0.971 (0.024)	0.998 (0.021)	**CEU (European)**
Mandenka	0.94 (0.072)	1.083 (0.086)	1.029 (0.042)	**Mand-enka (2)**			**JPT (89)**	1.017 (0.009)	1.012 (0.009)	1.011 (0.023)	1.028 (0.02)	1.007 (0.022)	0.986 (0.028)	0.969 (0.026)	0.995 (0.024)	**JPT (Japanese)**
Mbuti	0.95 (0.075)	1.100 (0.088)	1.041 (0.041)	1.026 (0.04)	**Mbuti (2)**			**CHS (100)**	0.994 (0.007)	0.996 (0.022)	1.012 (0.02)	0.991 (0.021)	0.973 (0.028)	0.957 (0.026)	0.982 (0.024)	**CHS (Chinese)**
San	0.988 (0.071)	1.131 (0.089)	1.024 (0.039)	1.005 (0.04)	0.98 (0.038)	**San (2)**			**CHB (97)**	1.001 (0.022)	1.017 (0.021)	0.996 (0.021)	0.978 (0.027)	0.961 (0.026)	0.986 (0.024)	**CHB (Chinese)**
Yoruba	0.972 (0.076)	1.117 (0.099)	1.007 (0.035)	0.985 (0.039)	0.961 (0.036)	0.975 (0.036)	**Yoruba (2)**			**PUR (55)**	1.015 (0.014)	0.995 (0.008)	0.977 (0.022)	0.96 (0.021)	0.986 (0.017)	**PUR (Pu.Ric.)**
Dai	0.964 (0.079)	1.126 (0.101)	1.075 (0.046)	1.047 (0.046)	1.013 (0.044)	1.051 (0.04)	1.085 (0.041)	**Dai (2)**			**MXL (64)**	0.98 (0.012)	0.964 (0.027)	0.948 (0.026)	0.972 (0.022)	**MXL (Mexican)**
French	0.902 (0.076)	1.029 (0.089)	0.994 (0.045)	0.959 (0.044)	0.937 (0.044)	0.957 (0.04)	0.972 (0.039)	0.908 (0.039)	**French (2)**			**CLM (60)**	0.98 (0.024)	0.963 (0.023)	0.989 (0.019)	**CLM (Colomb.)**
Han	0.983 (0.079)	1.18 (0.097)	1.061 (0.047)	1.048 (0.048)	1.013 (0.039)	1.043 (0.041)	1.08 (0.043)	0.984 (0.041)	1.11 (0.053)	**Han (2)**			**YRI (88)**	0.983 (0.009)	1.01 (0.009)	**YRI (Nigerian)**
Karitiana	0.9 (0.081)	1.063 (0.093)	0.967 (0.047)	0.938 (0.047)	0.929 (0.045)	0.938 (0.044)	0.96 (0.044)	0.861 (0.041)	0.966 (0.046)	0.875 (0.046)	**Karit-iana (2)**			**LWK (96)**	1.027 (0.011)	**LWK (Kenyan)**
Papuan	0.913 (0.074)	1.051 (0.097)	1.006 (0.052)	0.974 (0.046)	0.964 (0.044)	0.979 (0.044)	0.995 (0.048)	0.924 (0.041)	1.02 (0.048)	0.911 (0.042)	1.046 (0.053)	**Papuan (2)**			**ASW (61)**	
Sardinian	0.931 (0.075)	1.069 (0.099)	1.02 (0.046)	0.969 (0.043)	0.958 (0.038)	0.969 (0.043)	0.988 (0.044)	0.919 (0.041)	1.014 (0.045)	0.917 (0.04)	1.052 (0.055)	1.008 (0.048)	**Sardinian (2)**			
Deep genomes	**Denisova**	**Neander-thal**	**Dinka**	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karit-iana**	**Papuan**				

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation.

* *R*-ratios computed using Denisova and Neanderthal are normalized by the number of synonymous sites on each lineage, to adjust for the fewer mutations in the ancient sample than on present-day human lineages since divergence (the *R´* statistic described in the main text). Ratios involving Neanderthal and Denisova also remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Supplementary Table 7). Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population shown in the column. The number in parentheses indicates the number of samples per population.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 4: Parameters of simulated demographic models

Gravel[5] [π_{Eur}/π_{Afr}=0.72]

Time in gens.	$2N_{Afr}$	$2N_{Eur}$
$-300000 \le t < 3880$	28948	
$3880 \le t < 5000$	28948	3,722
$5000 \le t \le 5921$	28948	$2064e^{.003858(t-5000)}$

Simple bottleneck [π_{Eur}/π_{Afr}=0.69]

Time in gens.	$2N_{Afr}$	$2N_{Eur}$
$-300000 \le t < 3880$	28948	
$3880 \le t < 4080$	28948	500
$4080 \le t \le 5921$	28948	28948

Tennessen[4] [π_{Eur}/π_{Afr}=0.70]

Time in gens.	$2N_{Afr}$	$2N_{Eur}$
$-300000 \le t < 3880$	28948	
$3880 \le t < 5000$	28,948	3,722
$5000 \le t < 5716$	28,948	$2064e^{.00307(t-5000)}$
$5716 \le t \le 5921$	$28948e^{.00166(t-5716)}$	$18900e^{.00195(t-5716)}$

Lohmueller[2] [π_{Eur}/π_{Afr}=0.70]

Time in gens.	$2N_{Afr}$	$2N_{Eur}$
$-300000 \le t < 100002$	15672	
$100002 \le t < 101772$	15556	11398
$101772 \le t < 107706$	51272	11398
$107706 \le t \le 108580$	51272	60060

Notes: All simulations use $\mu = 2\times10^{-8}$ and burn in from generation -250,000 to 0. The switch from sampling every 100 to every 1 generations occurs at generation +1000 for the three models that end at time 5,921, and at +99,000 for the Lohmueller model[2]. Summary statistics at the end of the simulation are shown.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 5: Expected $R_{WestAfrica/Europe}$ for different models of demography

	$R(Non-synonymous)$	$R(Benign)$	$R(Possibly\ damaging)$	$R(Probably\ damaging)$
Estimated percentage of sites in each of three selective coefficient bins				
Percent of sites that are neutral	19%	27%	16%	9%
Percent of sites with weak selection coefficients: $s = -10^{-3}$	47%	60%	54%	27%
Percent of sites with strong selection coefficients: $s = -10^{-2}$	33%	13%	29%	64%
Model of history simulated				
Tennessen[4]	0.989	0.990	0.985	0.988
Gravel[5]	0.987	0.988	0.984	0.986
Lohmueller[2]	0.992	0.993	0.989	0.991

Notes: As described in the Supplementary Note, we assume that selective coefficients take on only one of three values: $s = 0$ ("neutral"), -10^{-3} ("weak"), and -10^{-2} ("strong"), and then fit the density in each of these bins using site frequency spectrum data under the assumption of mutations all acting additively with no epistasis. In the bottom section of the table, we show the value of $R_{WestAfrica/Europe}$ expected for each demographic model and distribution of selective coefficients. The expected values are less than two standard errors from 1 (using the empirically estimated Weighted Block Jackknife standard errors from Table 1), indicating that we do not expect there to be a clearly detectable difference in the accumulation of deleterious mutations comparing Europeans to West Africans.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 6: $R_{African/Non\text{-}African}$-statistic stratified by coalescent time depth

PSMC point estimate of time depth (posterior decoding)*	***R(synonymous)***	***R(All non-synonymous)***	***R(Benign)***	***R(Possibly damaging)***	***R(Probably damaging)***
≥0.0008 expected divergence ≥800,000 years	1.046 (0.014)	1.050 (0.017)	1.063 (0.019)	0.976 (0.059)	1.025 (0.053)
0.0004-0.0008 expected divergence 400,000-800,000 years	1.018 (0.022)	1.026 (0.028)	1.005 (0.046)	1.028 (0.101)	1.012 (0.057)
<0.0004 expected divergence <400,000 years	1.008 (0.033)	1.009 (0.044)	0.931 (0.061)	1.017 (0.145)	1.146 (0.093)

Notes: ±1 standard errors are from a Weighted Block Jackknife. For this analysis, we compare 8=4×2 sub-Saharan African to 12=6×2 non-African phased genomes. We restrict to sites with a GATK genotype quality of ≥70, and that have a consistent genotype between GATK and samtools.

* We present two coalescent time estimates based on the PSMC. The first is the expected per-base pair sequence divergence for genealogies of this time depth. The second is a calibration to years, based on making the further assumption of a mutation rate of 0.5×10^{-9} / bp / year. (The dates would halve using another conventional mutation rate assumption of 1.0×10^{-9} / bp / year.)

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 7: Key statistics as a function of allelic substitution patterns

Substitution type	Benign	Possibly damaging	Probably damaging	Non-synonymous
$R_{WestAfrica/Europe}$				
C→T or G→A	0.981 (0.028)	0.942 (0.059)	0.978 (0.053)	0.976 (0.023)
T→C or A→G	1.035 (0.034)	0.995 (0.082)	0.941 (0.102)	1.023 (0.029)
A→C or T→G	1.016 (0.054)	1.127 (0.112)	1.183 (0.089)	1.082 (0.042)
C→A or G→T	1.001 (0.060)	1.213 (0.127)	1.113 (0.114)	1.065 (0.051)
A→T or T→A	1.017 (0.100)	0.995 (0.188)	0.927 (0.136)	1.001 (0.076)
C→G or G→C	0.971 (0.049)	0.957 (0.082)	1.078 (0.093)	0.989 (0.038)
All but C→T or G→A	1.018 (0.021)	1.053 (0.051)	1.073 (0.052)	1.033 (0.017)
All sites	1.002 (0.018)	1.007 (0.040)	1.031 (0.038)	1.008 (0.015)
$R'_{WestAfrica/Europe}$				
C→T or G→A	0.956 (0.030)	0.918 (0.058)	0.953 (0.049)	0.951 (0.025)
T→C or A→G	1.005 (0.043)	0.966 (0.084)	0.914 (0.099)	0.993 (0.038)
A→C or T→G	1.009 (0.080)	1.118 (0.134)	1.173 (0.134)	1.074 (0.084)
C→A or G→T	1.055 (0.087)	1.278 (0.146)	1.172 (0.145)	1.122 (0.084)
A→T or T→A	1.060 (0.143)	1.039 (0.209)	0.968 (0.158)	1.044 (0.117)
C→G or G→C	0.945 (0.064)	0.931 (0.096)	1.049 (0.106)	0.962 (0.060)
All but C→T or G→A	1.004 (0.029)	1.038 (0.052)	1.059 (0.051)	1.019 (0.024)
All sites	0.981 (0.022)	0.986 (0.041)	1.010 (0.037)	0.987 (0.019)
$R'_{AllModern/Denisova}$				
C→T or G→A	0.898 (0.035)	0.828 (0.065)	0.565 (0.033)	0.812 (0.027)
T→C or A→G	0.860 (0.046)	0.904 (0.120)	0.697 (0.091)	0.851 (0.042)
A→C or T→G	0.830 (0.119)	0.725 (0.143)	1.475 (0.313)	0.908 (0.117)
C→A or G→T	1.085 (0.127)	0.789 (0.123)	1.080 (0.189)	1.026 (0.103)
A→T or T→A	0.848 (0.162)	0.919 (0.228)	0.791 (0.227)	0.857 (0.149)
C→G or G→C	0.766 (0.075)	0.799 (0.118)	1.244 (0.207)	0.833 (0.072)
All but C→T or G→A	0.865 (0.039)	0.810 (0.058)	0.985 (0.070)	0.872 (0.034)
All sites	0.929 (0.029)	0.870 (0.046)	0.760 (0.035)	0.889 (0.024)
$R'_{AllModern/Neanderthal}$				
C→T or G→A	0.953 (0.040)	0.994 (0.094)	0.909 (0.059)	0.953 (0.033)
T→C or A→G	1.046 (0.056)	1.053 (0.126)	0.928 (0.156)	1.038 (0.051)
A→C or T→G	0.937 (0.105)	1.132 (0.197)	1.462 (0.283)	1.085 (0.105)
C→A or G→T	1.073 (0.125)	0.991 (0.160)	1.188 (0.201)	1.086 (0.112)
A→T or T→A	0.988 (0.214)	0.786 (0.208)	0.928 (0.224)	0.944 (0.168)
C→G or G→C	0.816 (0.070)	1.099 (0.173)	1.169 (0.185)	0.919 (0.077)
All but C→T or G→A	0.997 (0.037)	1.074 (0.079)	1.181 (0.092)	1.037 (0.036)
All sites	0.993 (0.026)	1.065 (0.062)	1.063 (0.054)	1.015 (0.025)
$R'_{Denisova/Neanderthal}$				
C→T or G→A	1.089 (0.060)	1.210 (0.147)	1.695 (0.160)	1.215 (0.055)
T→C or A→G	1.250 (0.098)	1.072 (0.182)	1.409 (0.277)	1.244 (0.087)
A→C or T→G	1.049 (0.184)	1.594 (0.447)	0.687 (0.201)	1.083 (0.170)
C→A or G→T	0.849 (0.128)	1.188 (0.266)	1.133 (0.251)	0.966 (0.136)
A→T or T→A	1.195 (0.275)	0.913 (0.272)	1.141 (0.387)	1.135 (0.218)
C→G or G→C	1.027 (0.138)	1.368 (0.323)	1.029 (0.204)	1.091 (0.126)
All but C→T or G→A	1.132 (0.068)	1.290 (0.118)	1.172 (0.140)	1.164 (0.059)
All sites	1.064 (0.044)	1.198 (0.090)	1.433 (0.096)	1.141 (0.040)

Notes: ±1 standard errors are from a Weighted Block Jackknife. Statistics computed using Denisova and Neanderthal are normalized by the number of synonymous sites on each lineage (the R' statistic). Red highlighting indicates a nominal $P < 0.001$ for $R < 1$ or $R' < 1$, and green highlighting indicates a nominal $P < 0.001$ for $R > 1$ or $R' > 1$. These results document a significantly higher rate of accumulation of deleterious mutations in Denisova than in present-day humans whether the analysis is performed over all sites, or excluding C→T and G→A sites which are known to be subject to high rates of error in ancient DNA. There is no significant evidence of a higher rate if accumulation of deleterious mutations in Neanderthals compared with present-day humans.

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 8: Biased gene conversion analysis for all population pairs

Supplementary Table 8A: Unnormalized $R_{X/Y}$ statistics: bottom left G/C→A/T, top right A/T→G/C

	Neand-erthal (1)	Denisova (1)	Dinka (2)	Mand-enka (2)	Mbuti (2)	San (2)	Yoruba (2)	Dai (2)	French (2)	Han (2)	Karitiana (2)	Papuan (2)	Sardin-ian (2)	
Denisova	1.06 (0.006)													**Sardinian**
Dinka	1.133 (0.004)	1.064 (0.004)												**Papuan**
Mandenka	1.132 (0.005)	1.063 (0.004)	1.01 (0.001)											**Karitiana**
Mbuti	1.14 (0.005)	1.07 (0.004)	1.013 (0.002)	1.004 (0.001)										**Han**
San	1.137 (0.004)	1.065 (0.004)	1.002 (0.001)	0.994 (0.001)	0.99 (0.001)									**French**
Yoruba	1.125 (0.004)	1.056 (0.004)	0.994 (0.001)	0.985 (0.001)	0.981 (0.001)	0.992 (0.001)								**Dai**
Dai	1.142 (0.005)	1.071 (0.004)	0.992 (0.002)	0.984 (0.002)	0.98 (0.002)	0.991 (0.002)	0.999 (0.002)							**Yoruba**
French	1.11 (0.004)	1.045 (0.004)	0.986 (0.001)	0.978 (0.001)	0.975 (0.002)	0.985 (0.002)	0.992 (0.001)	0.992 (0.002)						**Mbuti**
Han	1.16 (0.005)	1.085 (0.004)	0.998 (0.001)	0.99 (0.001)	0.986 (0.002)	0.997 (0.002)	1.006 (0.001)	1.008 (0.001)	1.016 (0.001)					**San**
Karitiana	1.112 (0.005)	1.047 (0.004)	0.986 (0.002)	0.977 (0.002)	0.974 (0.002)	0.984 (0.002)	0.992 (0.002)	0.991 (0.002)	0.999 (0.002)	0.983 (0.002)				**Dinka**
Papuan	1.124 (0.005)	1.057 (0.004)	0.992 (0.002)	0.984 (0.002)	0.98 (0.002)	0.99 (0.001)	0.998 (0.001)	1 (0.002)	1.007 (0.002)	0.992 (0.002)	1.008 (0.002)			**Mandenka**
Sardinian	1.114 (0.004)	1.048 (0.004)	0.992 (0.002)	0.984 (0.001)	0.981 (0.002)	0.99 (0.002)	0.998 (0.001)	1 (0.002)	1.008 (0.001)	0.992 (0.001)	1.008 (0.002)	0.999 (0.002)		**Neander-thal**
	Neand-erthal	**Denisova**	**Dinka**	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karitiana**	**Papuan**		

	Papuan	Karit-iana	Han	French	Dai	Yoruba	San	Mbuti	Mand-enka	Dinka	Denisova	Neander-thal	
Sardin-ian (2)	0.994 (0.002)	1.002 (0.002)	0.984 (0.002)	1.004 (0.001)	0.99 (0.002)	0.987 (0.002)	0.975 (0.002)	0.974 (0.002)	0.979 (0.002)	0.989 (0.002)	1.104 (0.003)	1.100 (0.003)	**Sardinian**
Papuan (2)	1.008 (0.002)	0.991 (0.002)	1.01 (0.002)	0.996 (0.002)	0.992 (0.002)	0.979 (0.002)	0.979 (0.002)	0.985 (0.002)	0.994 (0.002)	1.110 (0.003)	1.106 (0.003)	**Papuan**
Kariti-ana (2)	0.981 (0.002)	1.002 (0.002)	0.987 (0.002)	0.986 (0.002)	0.973 (0.002)	0.972 (0.002)	0.978 (0.002)	0.987 (0.002)	1.103 (0.003)	1.099 (0.003)	**Karitiana**
Han (2)	1.02 (0.002)	1.006 (0.001)	1 (0.002)	0.988 (0.002)	0.985 (0.002)	0.992 (0.002)	1.001 (0.002)	1.112 (0.003)	1.112 (0.003)	**Han**
French (2)	0.986 (0.002)	0.984 (0.002)	0.972 (0.002)	0.971 (0.002)	0.977 (0.002)	0.986 (0.002)	1.102 (0.003)	1.098 (0.003)	**French**
Dai (2)	0.996 (0.002)	0.983 (0.002)	0.981 (0.002)	0.987 (0.002)	0.997 (0.002)	1.109 (0.003)	1.108 (0.003)	**Dai**
Yoruba (2)	0.986 (0.001)	0.986 (0.001)	0.992 (0.001)	1.002 (0.001)	1.116 (0.003)	1.112 (0.002)	**Yoruba**
San (2)	1 (0.002)	1.006 (0.002)	1.015 (0.002)	1.128 (0.003)	1.125 (0.002)	**San**
Mbuti (2)	1.007 (0.001)	1.016 (0.001)	1.127 (0.003)	1.123 (0.003)	**Mbuti**
Mand-enka (2)	0.989 (0.002)	1.121 (0.003)	1.116 (0.003)	**Mandenka**
Dinka (2)	1.112 (0.003)	1.109 (0.003)	**Dinka**
Denisova (1)	0.998 (0.003)	**Denisova**
Neander-thal (1)	**Neander-thal**
	Sardinian	**Papuan**	**Karitiana**	**Han**	**French**	**Dai**	**Yoruba**	**San**	**Mbuti**	**Mandenka**	**Dinka**	**Denisova**	

Notes: Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population in the column. The number in parentheses indicates the number of samples per population. ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation. We observe significant deviations from one in most pairwise comparisons. This could be due to different sequence error rates across samples which are small but significant given the small standard errors, or different mutation rates across samples. We therefore correct for such systematic differences across samples in Supplementary Table 6B by normalizing by the substitution rate differences at *G/C* and *A/T* sites, which are not subject to biased gene conversion. Ratios involving Neanderthal and Denisova remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Supplementary Table 7).

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 8B: Normalized $R'_{X/Y}$ statistics: bottom left G/C→A/T, top right A/T→G/C

	Neander-thal (1)														
Denisova	1.004 (0.007)	**Denisova (1)**													
Dinka	1.034 (0.005)	1.014 (0.005)	**Dinka (2)**												
Mand-enka	1.024 (0.005)	1.005 (0.005)	0.998 (0.002)	**Mand-enka (2)**											
Mbuti	1.028 (0.005)	1.009 (0.005)	0.997 (0.002)	1 (0.003)	**Mbuti (2)**										
San	1.023 (0.004)	1.002 (0.005)	0.987 (0.002)	0.99 (0.002)	0.989 (0.003)	**San (2)**									
Yoruba	1.025 (0.005)	1.006 (0.005)	0.995 (0.002)	0.998 (0.002)	0.997 (0.002)	1.009 (0.002)	**Yoruba (2)**								
Dai	1.041 (0.005)	1.023 (0.005)	0.995 (0.002)	0.998 (0.003)	0.997 (0.003)	1.009 (0.002)	1.000 (0.002)	**Dai (2)**							
French	1.022 (0.004)	1.005 (0.005)	1.002 (0.002)	1.005 (0.002)	1.004 (0.003)	1.015 (0.002)	1.007 (0.002)	1.009 (0.003)	**French (2)**						
Han	1.054 (0.005)	1.033 (0.005)	0.997 (0.002)	1 (0.002)	1 (0.003)	1.01 (0.002)	1.003 (0.003)	1.003 (0.002)	0.995 (0.003)	**Han (2)**					
Karitiana	1.02 (0.005)	1.003 (0.005)	0.993 (0.003)	0.997 (0.003)	0.996 (0.003)	1.007 (0.003)	0.999 (0.003)	0.998 (0.003)	0.99 (0.003)	0.995 (0.003)	**Karitiana (2)**				
Papuan	1.025 (0.005)	1.009 (0.005)	0.995 (0.002)	0.997 (0.003)	0.997 (0.003)	1.008 (0.003)	1.000 (0.003)	0.999 (0.002)	0.992 (0.003)	0.997 (0.003)	1.001 (0.003)	**Papuan (2)**			
Sardinian	1.021 (0.005)	1.005 (0.005)	1.001 (0.002)	1.003 (0.002)	1.003 (0.002)	1.014 (0.002)	1.006 (0.002)	1.007 (0.003)	0.998 (0.003)	1.004 (0.003)	1.008 (0.003)	1.007 (0.003)	**Sardin-ian (2)**		
	Neander-thal	**Denisova**	**Dinka**	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karitiana**	**Papuan**			

	Papuan	Karit-iana	Han	French	Dai	Yoruba	San	Mbuti	Mand-enka	Dinka	Denisova	Neander-thal	
Sardin-ian (2)	1.001 (0.003)	1.002 (0.003)	0.996 (0.003)	0.994 (0.003)	0.998 (0.003)	0.994 (0.003)	0.998 (0.002)	0.996 (0.002)	0.999 (0.003)	0.997 (0.002)	1.058 (0.003)	1.009 (0.003)	**Sardinian**
Papuan (2)		1.001 (0.003)	0.995 (0.003)	0.994 (0.003)	0.996 (0.003)	0.994 (0.003)	0.997 (0.003)	0.996 (0.003)	0.998 (0.003)	0.996 (0.003)	1.059 (0.004)	1.008 (0.003)	**Papuan**
	Karit-iana (2)		0.993 (0.003)	0.992 (0.003)	0.994 (0.003)	0.993 (0.003)	0.996 (0.003)	0.994 (0.003)	0.998 (0.003)	0.995 (0.002)	1.057 (0.004)	1.008 (0.003)	**Karitiana**
			Han (2)	0.999 (0.003)	1.001 (0.003)	0.998 (0.003)	1.001 (0.003)	1 (0.003)	1.002 (0.003)	1 (0.002)	1.059 (0.004)	1.011 (0.003)	**Han**
				French (2)	1.003 (0.003)	0.999 (0.003)	1.002 (0.002)	1 (0.002)	1.004 (0.002)	1.001 (0.002)	1.06 (0.003)	1.01 (0.003)	**French**
					Dai (2)	0.997 (0.003)	1.001 (0.003)	0.998 (0.002)	1.001 (0.003)	0.999 (0.002)	1.059 (0.004)	1.009 (0.003)	**Dai**
						Yoruba (2)	1.003 (0.003)	1.002 (0.002)	1.005 (0.002)	1.004 (0.002)	1.063 (0.003)	1.013 (0.003)	**Yoruba**
							San (2)	0.999 (0.002)	1.002 (0.003)	1 (0.002)	1.061 (0.003)	1.013 (0.003)	**San**
								Mbuti (2)	1.003 (0.002)	1.001 (0.002)	1.063 (0.003)	1.013 (0.003)	**Mbuti**
									Mand-enka (2)	0.998 (0.002)	1.06 (0.003)	1.01 (0.003)	**Mandenka**
										Dinka (2)	1.06 (0.003)	1.012 (0.003)	**Dinka**
											Denisova (1)	0.945 (0.004)	**Denisova**
												Neander-thal (1)	

Notes: Ratios are based on the accumulation of mutations observed in the population in the row divided by the accumulation of mutations observed in the population in the column. The number in parentheses indicates the number of samples per population. ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation. Ratios are normalized by the sum of $A \rightarrow T$, $T \rightarrow A$, $C \rightarrow G$ and $G \rightarrow C$ mutations on each lineage, producing an R' statistic that adjusts for differences in the rates of accumulations of mutations on different lineages since divergence, as well as the different ages of the samples. By normalizing, we highlight any differences in rates that are above and beyond these processes. Ratios involving Neanderthal and Denisova remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Supplementary Table 7).

Nature Genetics: doi:10.1038/ng.3186

Supplementary Table 9: R^2-statistic for all population pairs

	Dinka (2)										
Mand-enka	0.927 (0.038)	**Mand-enka (2)**									
Mbuti	0.956 (0.036)	1.025 (0.038)	**Mbuti (2)**								
San	0.907 (0.032)	0.962 (0.036)	0.949 (0.036)	**San (2)**							
Yoruba	0.92 (0.041)	0.99 (0.04)	0.969 (0.035)	1.029 (0.038)	**Yoruba (2)**						
Dai	1.601 (0.066)	1.684 (0.064)	1.609 (0.065)	1.681 (0.065)	1.713 (0.071)	**Dai (2)**					
French	1.448 (0.06)	1.528 (0.066)	1.454 (0.06)	1.513 (0.064)	1.541 (0.066)	0.879 (0.034)	**French (2)**				
Han	1.608 (0.059)	1.677 (0.07)	1.605 (0.064)	1.691 (0.068)	1.702 (0.072)	1.009 (0.036)	1.142 (0.045)	**Han (2)**			
Karitiana	2.141 (0.078)	2.228 (0.082)	2.091 (0.077)	2.147 (0.079)	2.239 (0.085)	1.468 (0.054)	1.619 (0.06)	1.455 (0.052)	**Karitiana (2)**		
Papuan	1.879 (0.07)	1.963 (0.078)	1.867 (0.058)	1.925 (0.064)	1.975 (0.067)	1.214 (0.048)	1.358 (0.048)	1.203 (0.049)	0.866 (0.032)	**Papuan (2)**	
Sardinian	1.524 (0.057)	1.603 (0.063)	1.522 (0.059)	1.579 (0.059)	1.612 (0.058)	0.936 (0.034)	1.064 (0.038)	0.933 (0.034)	0.655 (0.025)	0.787 (0.029)	**Sardinian (2)**
	Dinka	**Mand-enka**	**Mbuti**	**San**	**Yoruba**	**Dai**	**French**	**Han**	**Karitiana**	**Papuan**	

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation.

* For all population pairs, we show the $R^2_{X/Y}$ statistic. Ratios are based on the expected rate in the population in the row divided by the expected rate in the population in the column. Number in parentheses indicates the samples per population.

Nature Genetics: doi:10.1038/ng.3186

SUPPLEMENTARY NOTE

1 Inferred distributions of selection coefficients for PolyPhen-2 classes

1.1 Abstract

This note details the empirical fitting of the site frequency spectrum (SFS) from 1000 genomes data to determine the underlying distribution of fitness effects (DFE) for new mutations. Of particular interest is the DFE for PolyPhen-2 classes.

1.2 Aims and goals

Here we describe the technique used to analyze the distribution of selective effects of de novo mutations that form the distribution of fitness effects (DFE). Our primary aim is to infer this distribution from the site frequency spectrum of polymorphic non-synonymous alleles in the context of a given demographic history and total mutation rate. The de novo DFE in humans is in principle independent of population history and other demographic differences between individuals, allowing us to infer the distribution from a single fixed demography without loss of generality, provided the demographic inference is accurate.

1.3 Site frequency spectra

We use coding sequences from the 1000 genomes Yoruban (YRI) and Northern Europeans from Utah (CEU) populations to create a site frequency spectrum (SFS) in the form of a minor allele frequency (MAF) spectrum for both synonymous and non-synonymous sites. Additionally, we stratify the non-synonymous SFS by predicted PolyPhen-2 classes, labeled benign, possibly damaging, and probably damaging in order of increased predicted effect.

1.3.1 Simulated MAFs

Using the demographic inferences given in ref. 5, we simulate a genome of length 100Mb through the inferred demographic histories of European and African populations for a range of selective effects. In particular, the simulator tracks the derived allele frequencies of 10^8 independently evolving sites, in the infinite recombination limit with no linkage. Mutations

Nature Genetics: doi:10.1038/ng.3186

are introduced at a rate $\mu = 2 \times 10^{-8}$ per site per individual per generation. The population size is time dependent and reflects the demography associated with the population of interest. After completing roughly 5000 generations of recent demographic history, the allele frequencies are subsampled to the sample size of the associated 1000 genomes population sample, 88 for YRI and 85 for CEU. The results of this simulation provide expectations for the MAF for alleles with a single selective coefficient s. We simulate separately for $s = \{0, -10^{-3}, -10^{-2}\}$, which we consider to be neutral, weakly deleterious, and strongly deleterious, respectively. These selective coefficients are chosen to represent the range of realistic selective effects expected to be segregating in the human population. Alleles of stronger selective effect are likely to be absent in all but the largest population samples, and will be incorporated into the $s = -10^{-2}$ fitness class in our fit. These simulated MAFs provide the basis for our fit, as we will estimate the coefficients of their linear combination to determine the DFE.

1.4 Overall scale and target size

The number of bases simulated clearly overestimates the length of the human coding genome. The total coding genome is thought to be roughly $30Mb$ long, accounting for about 1% of the whole genome. Since estimates of both the mutation rate and target size are known to be relatively imprecise, we use the synonymous MAF to determine the overall rescaling for fitting our simulations to 1000 genomes data. Additionally, this method accounts for coverage issues, etc., assuming the same fraction of synonymous and non-synonymous sites are affected.

1.4.1 Scale factor for synonymous sites

Assuming synonymous sites are selectively neutral, we use a maximum likelihood fit with a single parameter to determine the scale factor for synonymous sites. The log likelihood is calculated as follows.

$$\log \mathcal{L} = \sum_{i=1}^{N} (D_i \log[F_i] - F_i) \tag{1}$$

Here D_i represents the i^{th} bin of the MAF from data, where $i \in [1, N]$ corresponds to allele count in the sample ranging from singletons at frequency $x = i/2N = 1/2N$ to alleles present in half of the haploid individuals at $x = N/2N = 1/2$. Similarly, F_i corresponds to counts in the fit to simulation, and is a function of fit parameters ϵ_k. For the present purposes, we are interested in determining the maximum likelihood for the following form of $F_i(\epsilon)$.

$$F_i(\epsilon) = \epsilon \; S_0^i \tag{2}$$

S_i^0 represents the i^{th} count of the MAF for the appropriately down-sampled neutral simulation with $s = 0$. The maximum log likelihood is given by the following expression.

$$\max[\log \mathcal{L}(\epsilon^{syn})] = \max \left[\sum_{i=1}^{N} (D_i \log[\epsilon^{syn} S_0^i] - \epsilon^{syn} S_0^i) \right] \tag{3}$$

Nature Genetics: doi:10.1038/ng.3186

We use the YRI synonymous MAF D_i^{YRIsyn} and the simulated YRI MAF for $s = 0$ to determine ϵ^{YRIsyn} numerically. The synonymous scale factor for YRI is determined by the maximum log likelihood value at $\epsilon^{YRIsyn} = 0.093$. Analogously, the synonymous scale factor for CEU has a maximum log likelihood value of $\epsilon^{CEUsyn} = 0.097$.

1.4.2 Scale factor for non-synonymous sites

Kruykov et al.[6] estimates the synonymous and non-synonymous fractions of the coding genome to be 0.32 and 0.68, respectively. This can be used to determine the appropriate scale factor for non-synonymous sites. The scale factor is simply the ratio of the total mutation rate in the target to the total simulated mutation rate.

$$\epsilon^{syn} = \frac{U_{syn}^{data}}{U^{sim}} = \frac{(\mu L_{syn})}{U^{sim}} \tag{4}$$

This can be solved for μ and substituted in to the non-synonymous expression to determine the non-synonymous scale factor.

$$\begin{aligned} \epsilon^{nonsyn} &= \frac{U_{nonsyn}^{data}}{U^{sim}} = \frac{(\mu L_{nonsyn})}{U^{sim}} \\ &= \frac{L_{nonsyn}}{L_{syn}} \epsilon^{syn} = \left(\frac{68}{32}\right) \epsilon^{syn} \end{aligned} \tag{5}$$

We find the following scale factors for the YRI and CEU simulated data.

ϵ_{nonsyn}^{YRI}	ϵ_{nonsyn}^{CEU}
0.198	0.207

1.4.3 Scale factors for PolyPhen-2 classes

The PolyPhen-2 software provides functional predictions that can be stratified into 3 classes: benign, possibly damaging, and probably damaging. One can compute the target size of these classes as a fraction of the total non-synonymous coding genome. This is accomplished by enumerating all possible point mutations from the hg19 human reference genome and classifying each mutation. We use the context dependent 64×4 weight matrix of single point mutations from a given triplet to all others[7]. Each of the 4^3 possible triplets has an associated matrix. Using HumVar, we compute approximate fractions for PolyPhen-2 classes found in the following table.

prediction	fraction (%)
benign	50.0
possibly damaging	16.7
probably damaging	33.3
unknown*	$\ll 1$

To confirm that this estimate is not biased by ancestry or recent demography, we stratify the human reference genome by predicted ancestry and find no substantial difference from these approximate values. From these fractions, we compute the appropriate scale factors for our fitting procedure.

Nature Genetics: doi:10.1038/ng.3186

ϵ^{YRI}_{benign}	$\epsilon^{YRI}_{possibly}$	$\epsilon^{YRI}_{probably}$
0.099	0.033	0.066

ϵ^{CEU}_{benign}	$\epsilon^{CEU}_{possibly}$	$\epsilon^{CEU}_{probably}$
0.104	0.035	0.069

1.5 Maximum Likelihood fit

Using the scale factors determined in the previous section, we compute the maximum log likelihood for a linear combination of selective effects. For simplicity, we choose to represent the DFE as a sum of several single s effect classes, rather than using a continuous functional form. We acknowledge that this three point mass model is a simplification of the true distribution of selection coefficients, but believe that it is useful for the purpose of obtaining a rough prediction of the expected value of the R-statistic for specific PolyPhen-2 classes.

$$\log \mathcal{L}(\{\alpha_k\}) = \sum_{i=1}^{N} \left(D_i \log[F_i(\{\alpha_k\})] - F_i(\{\alpha_k\})\right) \tag{6}$$

We use the following form for the fit function $F(\{\alpha_k\})$.

$$F_i(\{\alpha_k\}) = \epsilon^{nonsyn} \sum_k \alpha_k \; S_k^i = \epsilon^{nonsyn} \left(\alpha_0 \; S_0^i + \alpha_3 \; S_3^i + \alpha_2 \; S_2^i\right) \tag{7}$$

We employ the notation $k = 0$ for the simulated $s = 0$ MAF, $k = 3$ for the simulated $s = -10^{-3}$ MAF, and $k = 2$ for the simulated $s = -10^{-2}$ MAF. In this form, S_i^3 represents the MAF for the weakly selected sites, and α_3 is the fraction of the DFE that falls into this category. By estimating the maximum likelihood we can re-assemble the DFE in a rudimentary form as a fraction of mutations that fall into the category of neutral, weakly deleterious, and strongly deleterious. Since the overall scale factor is fixed, the α_k coefficients must be normalized with the following constraint.

$$\sum_k \alpha_k = 1 \tag{8}$$

This restricts the fit function as follows.

$$F_i(\{\alpha_k\}) = \epsilon^{nonsyn} \left(\alpha_0 \; S_0^i + \alpha_3 \; S_3^i + (1 - \alpha_0 - \alpha_3) \; S_2^i\right) \tag{9}$$

Note that for the present purposes, we have chosen 2 free parameters to fit, such that $\{\alpha_k\} = \{\alpha_0, \alpha_3\}$. For a 3 parameter fit with an additional nearly neutral class at $s = -10^{-4}$, for example, we simply introduce α_4 and S_4^i and modify the constraint $(\alpha_0 + \alpha_3 + \alpha_3 + \alpha_2) = 1$, with free parameters $\{\alpha_k\} = \{\alpha_0, \alpha_4, \alpha_3\}$. This method can be easily extended to fit an arbitrary number of parameters by including additional S_k^i for various selective effects. We have found this unnecessary for the present purposes, as it results in the effective overfitting of the DFE.

The maximum likelihood fit for 2 parameters is given simply by the following equations.

$$\max\left[\log \mathcal{L}(\alpha_0, \alpha_3)\right] = \max\left[\sum_{i=1}^{N} \left(D_i \log[F_i(\{\alpha_k\})] - F_i(\{\alpha_k\})\right)\right] \tag{10}$$

$$F_i(\{\alpha_k\}) = \epsilon^{nonsyn} \left(\alpha_0 \; S_0^i + \alpha_3 \; S_3^i + (1 - \alpha_0 - \alpha_3) \; S_2^i\right) \tag{11}$$

Nature Genetics: doi:10.1038/ng.3186

1.6 Results

Using the method outlined above, we compute the maximum likelihood fits for various PolyPhen-2 classes using YRI, CEU, and a joint measure that is the sum of the log likelihood functions of both YRI and CEU. Since the DFE should in principle be independent of demographic history, one can use the overlap of the independent measures in YRI and CEU in the form of the joint log likelihood (defined as a sum of the two log likelihoods) to produce a fit that is less sensitive to demographic errors in either of the two populations individually. The maximum log likelihood fit is summarized in the tables below. Errors are given for the joint fit, as this will be used in our subsequent analysis.

2 parameter fit (YRI)	neutral ($s = 0$)	weak ($s = -10^{-3}$)	strong ($s = -10^{-2}$)
all non-synonymous	0.20	0.44	0.36
benign	0.28	0.56	0.16
possibly damaging	0.17	0.50	0.34
probably damaging	0.09	0.25	0.66

2 parameter fit (CEU)	neutral ($s = 0$)	weak ($s = -10^{-3}$)	strong ($s = -10^{-2}$)
all non-synonymous	0.18	0.55	0.27
benign	0.26	0.68	0.06
possibly damaging	0.15	0.63	0.21
probably damaging	0.08	0.32	0.60

2 parameter fit (Joint)	neutral ($s = 0$)	weak ($s = -10^{-3}$)	strong ($s = -10^{-2}$)
all non-synonymous	0.19 $\pm$0.01	0.47 $\pm$0.04	0.33 $\pm$0.05
benign	0.27 $\pm$0.02	0.60 $\pm$0.07	0.13 $\pm$0.07
possibly damaging	0.16 $\pm$0.03	0.54 $\pm$0.11	0.29 $\pm$0.11
probably damaging	0.09 $\pm$0.01	0.27 $\pm$0.06	0.64 $\pm$0.06

1.6.1 Log Likelihood plots

The log likelihood surface for the two parameter fit can be visualized in a contour plot shown in Figure A. We note that the normalization condition $\sum_k \alpha_k = 1$ determines the strongly deleterious class uniquely. Figure B plots log likelihood contours for the benign, possibly damaging, and probably damaging PolyPhen-2 classes. We note a trend in the location of the maximum towards smaller values with increased predicted effect. All of the mass that vanishes in this process contributes to enhancing the weight of the strongly deleterious class. This is consistent with the stratification by PolyPhen-2 score, reinforcing our results.

Nature Genetics: doi:10.1038/ng.3186

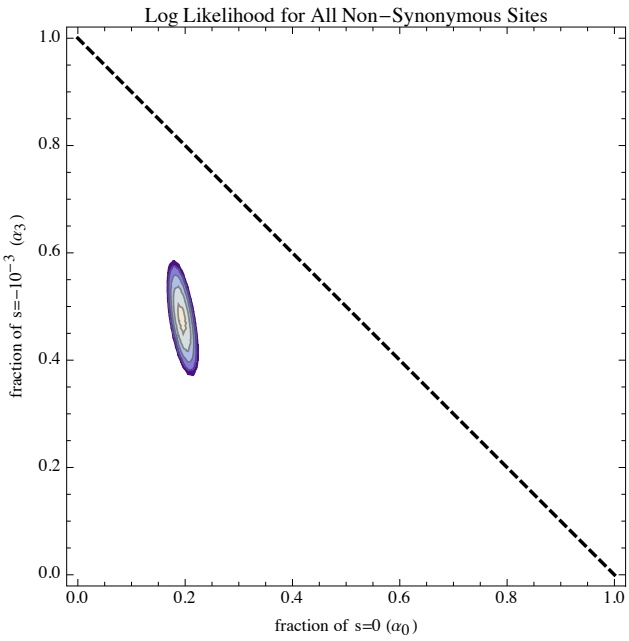


Figure A: The log likelihood plot for the joint inference from YRI and CEU data for all non-synonymous sites is shown for a two parameter fit. Contours are plotted representing two standard deviations from the peak. The coefficients of $s = 0$ and $s = -10^{-3}$, represented as (α_0,α_3), are plotted on the x and y axes, respectively. The fraction of strongly deleterious ($s = -10^{-2}$) sites in the DFE is constrained by the equation $\alpha_0 + \alpha_3 + \alpha_2 = 1$. This constraint restricts allowed values to below the dashed line. The maximum likelihood fit is located at $\{\alpha_0, \alpha_3, \alpha_2\} = \{0.19, 0.47, 0.33\}$.

Nature Genetics: doi:10.1038/ng.3186

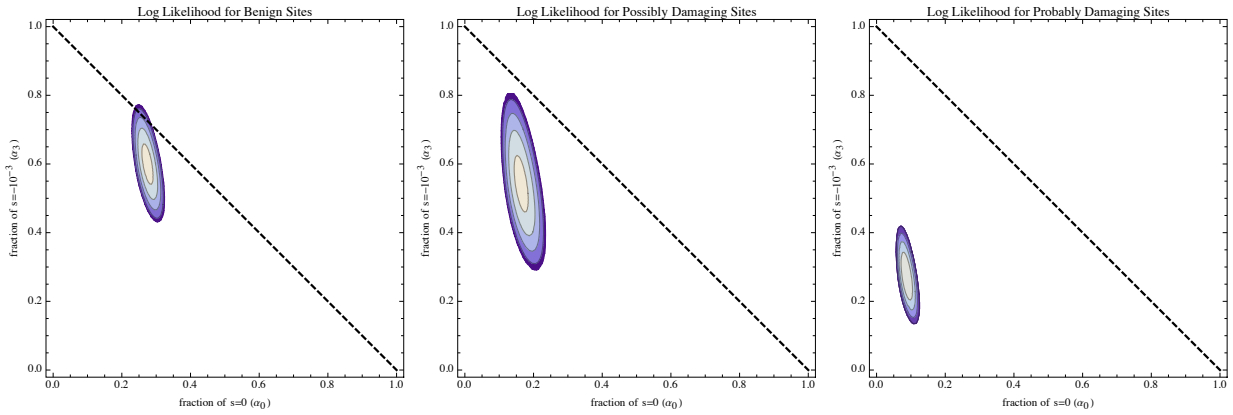


Figure B: Log Likelihood plots for the 2 parameter fit from the joint inference of YRI and CEU data are plotted for PolyPhen-2 classes. LEFT: Benign sites. MIDDLE: Possibly damaging sites. RIGHT: Probably damaging sites. All plots have axes (α_0,α_3) corresponding to neutral and weakly deleterious alleles and display two standard deviations from the maximum. The constraint $\alpha_0 + \alpha_3 + \alpha_2 = 1$ is satisfied, and only values below the dashed line are allowed. Note that the fit favors smaller fractions of neutral and weakly deleterious sites in favor of strongly deleterious sites with increasing PolyPhen-2 score, consistent with prediction.

1.7 Using the DFE to appropriately weight R

Here we use the inferred distribution of fitness effects, $\rho(s)$, to define an expected value $\langle R \rangle$ corresponding to the value of R that we expect to observe in population data. The appropriately weighted mutation load $\langle L \rangle$ for a given population is given by convoluting the load at different s values over the DFE.

$$\langle L \rangle = \int ds\ \rho(s)\ L(s) \tag{12}$$

This is true for both populations independently, since the DFE is roughly the same, allowing us to compute the expected $\langle R \rangle$ as follows.

$$\langle R \rangle = \frac{\langle L \rangle_{pop0}}{\langle L \rangle_{pop1}} = \frac{\int ds\ \rho(s)\ L^{pop0}(s)}{\int ds\ \rho(s)\ L^{pop1}(s)} \tag{13}$$

For the discretization of the DFE into neutral, weakly deleterious, and strongly deleterious components, this can be rewritten as the following sum.

$$\begin{aligned} \langle R \rangle &= \frac{\sum_k \alpha_k L^{pop0}(s_k)}{\sum_k \alpha_k L^{pop1}(s_k)} \\ &= \frac{\alpha_0 L^{pop0}(s=0) + \alpha_3 L^{pop0}(s=-10^{-3}) + \alpha_2 L^{pop0}(s=-10^{-2})}{\alpha_0 L^{pop1}(s=0) + \alpha_3 L^{pop1}(s=-10^{-3}) + \alpha_2 L^{pop1}(s=-10^{-2})} \end{aligned} \tag{14}$$

Here the α_k correspond to the fractions given in the results table above, and can represent appropriate values for all non-synonymous sites, or those for any of the PolyPhen-2 classes.

Nature Genetics: doi:10.1038/ng.3186

1.7.1 Computing $\langle R \rangle$, the weighted R statistic

Here, we calculate a weighted mutation load for population 0 (African) and population 1 (European) using fractions obtained from the maximum likelihood fits from the inferred distribution of fitness effects from Section 1.6 and from simulated mutation loads for average selection coefficients $s = \{0, -0.001, -0.01\}$. We calculated the weighted R statistic, denoted $\langle R \rangle$, as the ratio of the weighted mutation loads corresponding to population 0 and population 1. We calculate $\langle R \rangle$ for all non-synonymous sites, in addition to Polyphen classes, including benign, possibly damaging, and probably damaging sites (see tables below).

We calculated the expected $\langle R \rangle$ from simulations for four demographic models: Tennessen[4], Gravel[5], Lohmueller[2], and a simple bottleneck without exponential growth. We compare $\langle R \rangle$ from simulations with the R statistic observed in African Americans/European Americans from the Exome Sequencing Project (ESP) to assess the validity of different demographic models. Using this approach, we are unable to reject the Tennessen, Gravel and Lohmueller models, since $\langle R \rangle$ from these models are all within the 95% confidence intervals of R from ESP for all classes. The square bottleneck prediction is 2.09 standard errors from the empirical observation from the ESP measurement which is weakly suggestive that this model is not consistent with the data. These results suggest that this approach, the accumulation of deleterious mutations in two populations, along with the inferred DFE, can be a useful tool to evaluate the validity of different demographic models.

all non-synonymous sites	$\langle L \rangle_{pop0}$	$\langle L \rangle_{pop1}$	$\langle R \rangle$
Tennessen	0.000139	0.000140	0.989
Gravel	0.000138	0.000140	0.987
Lohmueller	0.000113	0.000114	0.992
Simple Bottleneck	0.000138	0.000141	0.978

benign	$\langle L \rangle_{pop0}$	$\langle L \rangle_{pop1}$	$\langle R \rangle$
Tennessen	0.000193	0.000195	0.990
Gravel	0.000192	0.000194	0.988
Lohmueller	0.000157	0.000158	0.993
Simple Bottleneck	0.000192	0.000196	0.979

possibly damaging	$\langle L \rangle_{pop0}$	$\langle L \rangle_{pop1}$	$\langle R \rangle$
Tennessen	0.000123	0.000125	0.985
Gravel	0.000123	0.000125	0.984
Lohmueller	0.000101	0.000103	0.989
Simple Bottleneck	0.000123	0.000126	0.973

Nature Genetics: doi:10.1038/ng.3186

probably damaging	$\langle L \rangle_{pop0}$	$\langle L \rangle_{pop1}$	$\langle R \rangle$
Tennessen	0.00006909	0.00006995	0.988
Gravel	0.00006887	0.00006982	0.986
Lohmueller	0.00005698	0.00005748	0.991
Simple Bottleneck	0.00006885	0.00007048	0.977

Nature Genetics: doi:10.1038/ng.3186

2 The proportion of non-synonymous sites is driven by neutral demographic history

2.1 Aims and Goals

We performed computer simulations of two models of demographic history (shown in Figure 3A) that differ qualitatively with regard to the history after the population split: (1) "Tennessen et al. 2012"[4], and (2) "Bottleneck and growth" . We tuned the parameters of the "Bottleneck and growth" model to match the Tennessen et al. 2012 (ref. 4) model both for the population split time (2,040 generations ago) and the final predicted heterozygosities at synonymous sites in both West Africans and Europeans.

2.2 Qualitative differences between the two models of demographic history

There is an important difference between the two models. For Tennessen et al. 2012 (ref. 4), West African populations are larger than European populations for most of the history since their split, and thus selection against weakly deleterious mutations would be expected to operate less effectively in European history. For the Bottleneck and Growth model, the opposite is the case.

Weighting of selection coefficients for Boyko et al. 2008				**Tennessen et al. 2012 demographic model results[4]**				**Bottleneck & Growth demographic model results**			
s	gamma density	bin width	Weight: proportional to (gamma) x (bin width)	segregating sites /bp	$L_{Afr\text{-}not\text{-}Eur}$	$L_{Eur\text{-}not\text{-}Afr}$	$R_{Afr/Eur}$	segregating sites /bp	$L_{Afr\text{-}not\text{-}Eur}$	$L_{Eur\text{-}not\text{-}Afr}$	$R_{Afr/Eur}$
-0.000001	1.8×10^{-1}	1.4×10^{-6}	0.029	2.8×10^{-3}	6.1×10^{-4}	6.1×10^{-4}	1	3.4×10^{-3}	6.2×10^{-4}	6.2×10^{-4}	1
-0.000002	1.1×10^{-1}	1.7×10^{-6}	0.021	2.8×10^{-3}	6.1×10^{-4}	6.1×10^{-4}	1	3.3×10^{-3}	6.1×10^{-4}	6.2×10^{-4}	1
-0.000005	5.1×10^{-2}	3.9×10^{-6}	0.022	2.8×10^{-3}	6.0×10^{-4}	6.0×10^{-4}	1	3.3×10^{-3}	6.1×10^{-4}	6.1×10^{-4}	1
-0.00001	2.9×10^{-2}	7.1×10^{-6}	0.023	2.7×10^{-3}	5.9×10^{-4}	5.9×10^{-4}	1	3.3×10^{-3}	5.9×10^{-4}	5.9×10^{-4}	1
-0.00002	1.7×10^{-2}	1.7×10^{-5}	0.033	2.6×10^{-3}	5.6×10^{-4}	5.6×10^{-4}	1	3.2×10^{-3}	5.6×10^{-4}	5.6×10^{-4}	0.99
-0.00005	8.2×10^{-3}	3.9×10^{-5}	0.036	2.3×10^{-3}	4.7×10^{-4}	4.8×10^{-4}	0.99	2.9×10^{-3}	4.7×10^{-4}	4.8×10^{-4}	0.99
-0.0001	4.7×10^{-3}	7.1×10^{-5}	0.037	1.9×10^{-3}	3.5×10^{-4}	3.6×10^{-4}	0.98	2.4×10^{-3}	3.5×10^{-4}	3.6×10^{-4}	0.97
-0.0002	2.7×10^{-3}	1.7×10^{-4}	0.053	1.4×10^{-3}	2.1×10^{-4}	2.1×10^{-4}	0.97	1.9×10^{-3}	2.1×10^{-4}	2.2×10^{-4}	0.95
-0.0005	1.3×10^{-3}	3.9×10^{-4}	0.057	8.7×10^{-4}	8.4×10^{-5}	8.9×10^{-5}	0.94	1.3×10^{-3}	8.3×10^{-5}	9.1×10^{-5}	0.91
-0.001	7.6×10^{-4}	7.1×10^{-4}	0.059	6.5×10^{-4}	4.1×10^{-5}	4.5×10^{-5}	0.92	9.8×10^{-4}	4.1×10^{-5}	4.6×10^{-5}	0.90
-0.002	4.3×10^{-4}	1.7×10^{-3}	0.084	4.9×10^{-4}	2.0×10^{-5}	2.2×10^{-5}	0.92	6.6×10^{-4}	2.0×10^{-5}	2.2×10^{-5}	0.94
-0.005	2.0×10^{-4}	3.9×10^{-3}	0.089	2.8×10^{-4}	8.0×10^{-6}	8.3×10^{-6}	0.97	3.0×10^{-4}	8.1×10^{-6}	8.1×10^{-6}	1
-0.01	1.1×10^{-4}	7.1×10^{-3}	0.090	1.6×10^{-4}	4.0×10^{-6}	4.0×10^{-6}	0.99	1.6×10^{-4}	4.0×10^{-6}	4.0×10^{-6}	1
-0.02	6.1×10^{-5}	1.7×10^{-2}	0.120	8.0×10^{-5}	2.0×10^{-6}	2.0×10^{-6}	1	7.9×10^{-5}	2.0×10^{-6}	2.0×10^{-6}	1
-0.05	2.4×10^{-5}	3.9×10^{-2}	0.105	3.2×10^{-5}	8.0×10^{-7}	8.0×10^{-7}	1	3.2×10^{-5}	8.0×10^{-7}	8.0×10^{-7}	1
-0.1	9.9×10^{-6}	7.1×10^{-2}	0.078	1.6×10^{-5}	4.0×10^{-7}	4.0×10^{-7}	1	1.6×10^{-5}	4.0×10^{-7}	4.0×10^{-7}	1
-0.2	2.9×10^{-6}	1.7×10^{-1}	0.056	8.0×10^{-6}	2.0×10^{-7}	2.0×10^{-7}	1	8.0×10^{-6}	2.0×10^{-7}	2.0×10^{-7}	1
-0.5	1.8×10^{-7}	3.9×10^{-1}	0.0077	3.2×10^{-6}	8.0×10^{-8}	8.0×10^{-8}	1	3.2×10^{-6}	8.0×10^{-8}	8.0×10^{-8}	1
-1	3.3×10^{-9}	2.9×10^{-1}	0.00011	1.6×10^{-6}	4.0×10^{-8}	4.0×10^{-8}	1	1.6×10^{-6}	4.0×10^{-8}	4.0×10^{-8}	1
Non-synonymous	N/A	1	1	7.6×10^{-4}	1.27×10^{-4}	1.29×10^{-4}	0.987	9.6×10^{-4}	1.28×10^{-4}	1.30×10^{-4}	0.983
Synonymous	N/A	N/A	N/A	2.8×10^{-3}	6.17×10^{-4}	6.17×10^{-4}	0.999	9.6×10^{-4}	6.21×10^{-4}	6.22×10^{-4}	0.998

Nature Genetics: doi:10.1038/ng.3186

2.3 Simulations

We simulated 10 billion base pairs for a selection coefficient of $s = 0$ ("synonymous sites") and 1 billion base pairs for each of 19 negative selection coefficients ("non-synonymous"), as shown in the table. For the total number of segregating non-synonymous sites, we weighted each of the 19 selection coefficients based on an inferred distribution of human selection coefficients from Boyko et al. 2008 (ref. 8) (the fit to European genetic variation data in which $-s$ follows a gamma distribution with α=0.206 and β=15,400). Specifically, we took the gamma density ("gamma density" in Table 1 of that paper), and multiplied it by the range of selection coefficients represented by that bin ("bin width"). We renormalized the products so that they summed to one.

For each simulation, we tabulated the number of segregating sites per generation over the last 3,000 generations of history assuming a sample size of 40 for both West Africans and 40 Europeans (we used a hypergeometric distribution to obtain the expected probability of each polymorphic site being heterozygous given that sample size).

2.4 Matching the empirical ratio of non-synonymous to synonymous sites

To obtain numbers for all non-synonymous sites we used the weights shown in the table to compute the expected rate of segregating sites per base pair in both West Africans and Europeans in each generation.

To compute the proportion of non-synonymous sites in each generation, we used the following equation, with the factor of 3.42 chosen to be what was needed for the proportion to equal the empirical value in ref. 2 (0.479). We note that this is only a scaling factor. Similar behaviors would be obtained with other scaling factors so the exact choice of scaling factor is not important for our qualitative results.

$$proportion\ nonsynonymous = \frac{3.42 \times nonsynonymous}{3.42 \times nonsynonymous + synonymous}$$

2.5 Detailed discussion of the simulation results

Figure 3B shows the temporal dynamics of the density of non-synonymous sites (per base pair), the density of synonymous sites, and the proportion of all sites that are non-synonymous. We only show results here for Europeans (in both simulated models, the West African population size changes very little and so the statistics hardly change, at least compared with Europeans).

Both simulated demographic models show the same qualitative feature as the simulations presented in ref. 2. Non-synonymous and synonymous segregating site densities are initially decreased in Europeans by the bottleneck for all classes of selection coefficients, with the proportional effect being larger for non-synonymous sites. In the recovery period, however, non-synonymous segregating site densities increase faster than synonymous ones. Thus, the total proportion of sites that are non-synonymous also increases in this period. In both simulated demographic models, the proportion of non-synonymous sites thus has a non-trivial behavior of initially falling and then rising, eventually passing the baseline as first observed by ref. 2.

Nature Genetics: doi:10.1038/ng.3186

Lohmueller et al. 2008 (ref. 2) argued that the observation of an elevated rate of non-synonymous sites in present-day Europeans (compared with West Africans as a baseline) is evidence of less effective natural selection to remove weakly deleterious mutations in European than in West African populations since their separation. If this is the case, it is surprising that the Bottleneck and Growth model where population sizes have been larger in European than in West Africans populations for most of their history shows the same qualitative effect.

What then is driving the rise in the proportion of non-synonymous sites in Europeans that begins within hundreds of generations after the bottleneck for both simulated models, if it is not reduced effectiveness of deleterious mutation in Europeans?

A key intuition that is helpful for understanding this behavior is that prior to the West African / European population split, the density of non-synonymous segregating sites is expected to have been much lower than the density of synonymous segregating sites due to the action of natural selection. This pattern would only have been intensified by the preferential loss of segregating sites for the non-synonymous class due to the out-of-Africa bottleneck, as our simulations show (Figure 3B).

Once the population began expanding, genetic drift would have been reduced and equilibrium would have favored a higher density of segregating sites both for non-synonymous and synonymous classes. The non-synonymous site class approaches its equilibrium relatively more quickly than the synonymous site class once the population grows, as the non-synonymous class experiences the same flux of new mutations as synonymous sites (actually an even higher flux, as the target size is larger). Since non-synonymous sites start out with a lower baseline density, the proportional rate of their approach to equilibrium is faster than for synonymous segregating sites, explaining our observation. Selected classes of mutations turn over more quickly, and thus approach equilibrium more quickly[9].

Nature Genetics: doi:10.1038/ng.3186

Supplementary References

[1] Prüfer, K. et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505,** 43-9 (2014).

[2] Lohmueller, K.E. *et al.* Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451,** 994-7 (2008).

[3] Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56-65 (2012).

[4] Tennessen, J.A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337,** 64-9 (2012).

[5] Gravel, S. *et al.* Demographic history and rare allele sharing among human populations. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 1198-8 (2011).

[6] Kryukov, G.V. *et al.* Power of deep, all-exon resequencing for discovery of human trait genes. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 3871-3876 (2009).

[7] Asthana, S. *et al.* Analysis of Sequence Conservation at Nucleotide Resolution. *PLoS Comput. Biol.* **3,** e254 (2007).

[8] Boyko, A.R. *et al.* Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* **4,** e1000083 (2008).

[9] Reich, D.E. & Lander, E.S. On the allelic spectrum of human disease. *Trends Genet.* **17,** 502-10 (2001).

Nature Genetics: doi:10.1038/ng.3186