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], and **David Reich**[1,2,4]

[1]Broad Institute of Harvard and MIT, Cambridge, MA, USA

[2]Department of Genetics, Harvard Medical School, Boston, MA

[3]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[4]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA

Abstract

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 non-synonymous 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 non-synonymous mutations faster than either modern humans or Neanderthals. To reconcile these findings with patterns that have been interpreted as evidence of 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, and instead are 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 governs the importance of the stochastic force of genetic drift. For a constant-size population in equilibrium, the effectiveness is fully determined by the product Ns[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 hundred thousand years have reduced the nucleotide diversity (the number of

differences per base pair between an individual's two chromosomes) 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 ten thousand years has led to rapid expansions.

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 (non-synonymous mutations)[5–9]. A key study measured the proportion of polymorphic positions that are non-synonymous in 20 Europeans and 15 African Americans, and showed that while both classes of sites have a reduced rate in Europeans, the reduction is proportionally less for non-synonymous sites[5]. This pattern was interpreted as being due in part to reduced effectiveness of natural selection against weakly deleterious alleles in Europeans than in West Africans due to the smaller effective population size in Europeans since separation[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 non-synonymous 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 frequency for non-synonymous alleles would have been lower than for synonymous alleles, as negative selection places downward pressure on derived allele frequencies[5]. The different frequency distributions at non-synonymous 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 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 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 from Y, determining the ancestral state based on the chimpanzee genome (*PanTro2*), and recording all the differences. We count the number of derived

mutations in genome X that are not seen in Y, and vice versa, and define a statistic $R_{X/Y}$ that is a ratio of the two. 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 (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 loads between the two populations.

We measured $R_{WestAfrica/Europe}$ in four sequencing datasets: (1) coding regions of genes (exomes) in 15 African Americans and 20 European Americans[5]; (2) exomes from 1,089 individuals in the 1000 Genomes Project (1KG)[13]; (3) exomes from 1,088 African Americans and 1,351 European Americans[6], and (4) 24 whole genomes sequenced to high coverage[14,15] (Supplementary Table 1). We inferred the ancestral allele based on comparison to chimpanzee. As expected for sites unaffected by selection and for indistinguishable differences in mutation rates in the history of the two populations, $R_{WestAfrica/Europe}(synonymous)$ is within two standard errors of 1 (Table 1 and Supplementary Table 2). However, $R_{WestAfrica/Europe}(non\text{-}synonymous)$ is 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 of ref. [16] for similar population comparisons. To extend these results to a more diverse set of populations, we computed $R_{X/Y}$ between all possible pairs of 11 populations each represented by two deep genome sequences[15], and all pairwise comparisons of 14 populations from the 1000 Genomes Project[13]. We observe no significant differences for any population pair despite sometimes profound differences in demographic history (Figure 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 show that if selection acts additively and coefficients are in the range $s \in [-0.004, -0.0004]$, $R_{WestAfrica/Europe}$ is expected to be below 0.95 and detectable given the standard errors of our measurements (Figure 2). However, if many mutations had selection coefficients outside this range, the signal could be diluted to the point of not being detectable. Indeed, when we compute the expected value of $R_{WestAfrica/Europe}$ integrating over a previously fitted distribution of selection coefficients[18], we find that $R_{WestAfrica/Europe}(non\text{-}synonymous)$ is 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 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 is opposite to that for additively acting mutations, however, reflecting the fact that recessively acting mutations that drift up in frequency due to the bottleneck can be efficiently purged through the action of selection as they are exposed in homozygous form[22] (Figure 2; Supplementary Figure 1). The difference in direction for additively and recessively acting mutations suggests that until there is a reliable model of demographic history and the 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 non-synonymous mutations, we stratified non-synonymous sites in two ways. First, we stratified by predicted functional effect. The PolyPhen-2[23] and SIFT[24] algorithms both predict function in a way that is dependent on the ancestral/derived status of allelic variants compared with 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 (Methods), and which fully corrects for the bias (Supplementary Figure 2). We detect no significant differences in the accumulation of deleterious mutations[16] (Table 1, Table 2, Supplementary Table 2, Supplementary Table 3). We caution, however, that these tests may also have limited power, as when we fit a distribution of additively acting selection coefficients to each of the three PolyPhen classes (Supplementary Note), and then simulate under these distributions, we predict ratios of $R_{WestAfrica/Europe}$ of 0.984–0.993 (Supplementary Table 5), close enough to 1 to not be able to distinguish a significant difference given the standard error 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 might enhance power as the population split between African and non-African populations occurred only in the last roughly one hundred thousand years, and the mutations that arose prior to the 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 to infer the local time since the most recent common ancestor at each location in the genome (masking out the exome to avoid circularly using the same sites for our statistic). When we pool over sample pairs to increase statistical power, and compute $R_{African/non-African}$ only for the subset of the genome with the lowest TMRCA, we still detect no significant differences (Supplementary Table 6).

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

Tests for differences in the load of non-synonymous mutations are not always null, as we found when we analyzed deeply sequenced genomes from an archaic Denisovan[14] and Neanderthal[15] and compared them to present-day humans. The ancestors of both 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 3–6 times smaller[14]. A challenge in comparing the accumulation of mutations in present-day to ancient samples is that fewer mutations are expected to have occurred in the ancient samples because they are closer in time to the common ancestor. We therefore divided the accumulation of non-synonymous mutations specific to each genome by the accumulation of synonymous sites: $R'_{X/Y}(non-synonymous\ class) = R_{X/Y}\ (non-synonymous\ class)/R_{X/Y}(synonymous)$. After removing C→T and G→A mutations in the archaic genomes that may be susceptible to degradation leading to ancient DNA errors, we infer that present-

day humans have accumulated deleterious mutations at a significantly lower rate than Denisovans since separation: $R'_{Modern/Denisova}(non\text{-}synonymous) = 0.872 \pm 0.034$ (P=0.0002) (Supplementary Table 7)[14]. However, $R'_{Modern/Neanderthal}(non\text{-}synonymous) = 1.037 \pm 0.037$ is consistent with 1, indicating that deleterious mutations accumulated faster in Denisovan than in Neanderthal ancestors since they separated (Table 2). The different patterns—despite similar inferred demographic histories—suggests that fitting models to patterns of neutral genetic variation and then simulating under these models may not produce an accurate prediction of the relative effectiveness of selection in pairs of populations (Figure 2). A corollary is that current models for the joint history of West Africans and non-Africans may not produce accurate predictions. Little is known about the duration of the out-of-Africa bottleneck: a short sharp bottleneck and a long and 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 the 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 removal of mutations across populations, we analyzed a class of sites that is much larger than the class of non-synonymous substitutions so we can 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]. Since BGC only acts on heterozygous sites, it occurs at a rate proportional to heterozygosity, or $2p(1\text{-}p)$ for a mutation of frequency p, exactly mimicking additive selection[25]. We find that $R'^{G/C \to A/T}_{West\ Africans/Europeans} = 0.995 \pm 0.002$ and $R'^{A/T \to G/C}_{West\ Africans/Europeans} = 1.000 \pm 0.002$ by using an $R'^{class1 \to class2}_{X/Y} = R^{class1 \to class2}_{X/Y}/(R^{A \leftrightarrow T}_{X/Y} + R^{C \leftrightarrow G}_{X/Y})$ statistic that normalizes by the rate of A↔T and C↔G substitutions not expected to be affected by BGC. This also corrects for possible differences in mutation rates across populations. For the comparison of diverse West Africans to diverse non-Africans, we detect 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 they split is likely to have been extremely small. However, San Bushmen from southern Africa have about 1% fewer G/C→A/T mutations than all other humans (significant at up to 8 standard errors). This is the first direct detection of a different rate of accumulation of a class of mutations in some present-day humans than in others. The San are among the most deeply diverged present-day humans, and a possible explanation for our observation is that the San's longer isolation time would have provided opportunity for differences in the effectiveness of removal of mutations across populations to accumulate to the point of having a detectable effect in the high sensitivity BGC analysis[26].

Re-interpretation 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 study of an alternative statistic: the proportion of polymorphic sites in the exome that are non-synonymous. This 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 the change in this statistic over time. While our simulations show qualitative patterns that are consistent with those reported previously[5], they also provide new insight due to a modification to the program that allows us, in every generation, to dissect how selection (versus mutation and genetic drift) contributes to the expected change in the proportion of non-synonymous sites in that generation. The simulations show that during and after a population bottleneck, the per generation change in the proportion of segregating sites that are non-synonymous is not driven by selection being less effective at reducing this ratio than it was in the ancestral population. Instead, following a short period at the start of the bottleneck when selection's effectiveness in changing this statistic is reduced, selection begins to be more effective at reducing the proportion of non-synonymous sites per generation than it was prior to the bottleneck (Figure 3). Thus, the rate at which selection reduces the value of this statistic per generation is enhanced rather than diminished by the bottleneck, which pushes the statistic in the direction opposite to that in which it actually moves. We can conclude from this that it is primarily non-selective forces that drove the dynamics of this statistic since West African and European population separation.

Intuitively, what is explaining these simulation results? Prior to the West African / European split, allele frequencies of non-synonymous polymorphisms would, on average, have been much lower due to the depletion of non-synonymous sites by selection, and the per-site density of non-synonymous segregating sites would also have been lower. A population entering the bottleneck primarily loses rare alleles, so the non-synonymous 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 re-expands, the allele frequencies for non-synonymous sites also adjust faster, because the same flux of new mutations into both classes causes a faster rate of replenishment of non-synonymous sites than synonymous sites due to an initially lower density. It is the greater proportional impact of new mutations on non-synonymous sites per generation that occurs after a bottleneck—because the class has been depleted by the bottleneck—that is driving the observed effects (Figure 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 non-synonymous 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 than in 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 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 non-synonymous and synonymous sites prior to the split.

DISCUSSION

It is tempting to interpret the indistinguishable accumulations of deleterious mutations across present-day humans 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 prevalence rates of complex disease that have an additive genetic architecture[16,19]. However, recessively or epistatically acting mutations work in combination to contribute to disease risk, and since demography affects allele frequencies, it affects the rate of co-occurrence of alleles. For example, Table 1 and Supplementary Table 9 show that the absolute count of alleles occurring in homozygous form is higher in non-Africans than in Africans for all functional site classes[5]. Thus, the relative risk for diseases with non-additive architectures may 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 will affect disease risk.

ONLINE METHODS

Data

The datasets we analyzed were published previously and are summarized here. We determined the ancestral allele at each position based on comparison to the chimpanzee genome (*PanTro2*) except in the case of the Celera dataset where we used the previously reported determination[5].

Celera—PCR amplification and Sanger sequencing was performed on 15 African American (AA) and 20 European American (EA) 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 from both AA and EA[5].

1000 Genomes (1KG)—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 dataset includes 2 samples each from 6 non-African and 5 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 dataset reported in ref. [15]. We only analyzed sites with genotype quality scores (GQ) of 45.

Mutation annotation

We annotated coding mutations using ANNOVAR[28], which classifies sites as "non-synonymous", "synonymous", "stop-gain" or "stop-loss". We sub-classified variants using a version of PolyPhen-2 that was created specifically for this study that is independent of the ancestral/derived status of the human genome reference sequence ("human-free Polyphen-2"). To guarantee independence of 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 of the scores for the two alleles. By construction, this is symmetric with respect to reference/non-reference (and also ancestral/derived, 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 are interested in the expected number of mutations in a randomly sampled haploid exome from one population that are not seen in a randomly sampled comparison exome from another population. To compute this in a situation where we have many exomes available from each population, we do 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 obtain the expected value if we performed an infinite number of random samplings. To compute this, at each variable site i, we define 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 is the count of the mutant allele in a sample of n_Y^i exomes from population Y. The expectations are obtained by summing over all sites.

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

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

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

We then defined the ratio statistics

$$R_{X/Y} = L_{X\,notY} / L_{Y\,notX} \quad R_{X/Y}^2 = L_{X\,notY}^2 / L_{Y\,notX}^2$$

Weighted Block Jackknife to estimate standard errors

We obtained standard errors using a Weighted Block Jackknife[12]. We divided the SNP datasets into 100 contiguous blocks and then recompute the statistic on all of the data except for the data from that block. The variation can be converted to a standard error using jackknife theory. We assessed significance based on the number of standard errors from the null expectation of $R_{X/Y}$=1, and compute a *P*-value using a Z-score assuming a normal distribution.

Time-stratified computation of relative accumulation of deleterious mutation

We began with data from 10 experimentally phased genomes, all processed identically[15]. These genomes consist of one each from the populations in Table 2 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 the Pairwise Sequential Markovian Analysis (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 $R_{African/non\text{-}African}$-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 (BGC)

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 treat as neutral (and use as the denominator of $R'_{X/Y}$). For BGC analyses, we use the entire genome, excluding sites in the exome.

$R'_{X/Y}$-statistic: Correcting for branch shortening and differences in mutation rates

For analyses involving the archaic Denisovan and Neanderthal samples, which are many tens of thousands of years old and thus have experienced less evolution from the common ancestor than present-day humans to which they are compared, we do not expect that $L_{archaic\,not\,modern} = L_{modern\,non\,archaic}$ even for neutral sites. For all analyses involving ancient samples, we normalize both $L_{X\,not\,Y}$ and $L_{Y\,not\,X}$ by the accumulation of mutations at sites that are expected to act neutrally (synonymous sites for coding sequences and A/T + C/G for BGC). We then define:

$$R'^{class}_{X/Y} = (L^{class}_{X\,notY} / L^{class}_{Y\,notX}) / (L^{normalization}_{X\,notY} / L^{normalization}_{Y\,notX}) = R^{class}_{X/Y} / R^{normalization}_{X/Y}$$

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 Denisova or Neanderthal[13]. In Supplementary Table 7, we document that this error process is substantial enough to bias statistics involving Denisova. We therefore restrict computation of $R'_{X/Y}$ involving the ancient samples to sites that are not C→T and G→A substitutions (for the sake of comparability, we also do not analyze these classes of sites for non-ancient samples in analyses that also involve archaic samples).

Simulations

We wrote a forward simulation program in C that implements an infinite sites model. Each mutation is 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 based on the 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 use a demographic model previously fitted to genetic data[6], as well as a simple bottleneck and expansion model (Supplementary Figure 1 reports results for four histories using simulation parameters shown in Supplementary Table 4). For comparisons of West African and archaic population history we also use a previously fitted demographic model[15]. We use a mutation rate of 2×10^{-8}/base pair/generation.

At each simulated site, we compute 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 out of a total of N individuals in the population. We compute the probability that 40 chromosomes are polymorphic at a site i as 1 minus the hypergeometric probability of 0 or 40 derived alleles:

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

We average 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 $L_{X\ not\ Y}$, $L_{Y\ not\ X}$, 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 a weighting based on the distribution of human selection coefficients for non-synonymous sites inferred in ref. [18], where the probability of a given value of -s is drawn from a gamma distribution fitted to European genetic data with α=0.206 and β=15400. For analyses of the expected value of $R_{WestAfrica/Europe}$ 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 the evolutionary dynamics into the 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 $All_{s,i,j-1,k}$ be the count of derived alleles that have selection coefficient s at nucleotide position i in the $j-1$th generation in simulation replicate k. We use our simulation machinery to sample the count of derived alleles in the subsequent generation $All_{s,i,j,k}$, assuming that the selection coefficient in the next generation is the same. We also independently sample the count of derived alleles in that generation $NonSel_{s,i,j,k}$ assuming that selection stopped in that generation (s=0). Since the count of derived alleles is always sampled based on $All_{s,i,j-1,k}$ in the previous generation (not $NonSel_{s,i,j-1,k}$), this procedure ensures that the accumulation of derived alleles at each position corresponds 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 average over A simulation replicates, and B simulated nucleotides per replicate:

$$E[All_{s,j}] = \frac{1}{A \times B} \sum_{i=1}^{A} \sum_{k=1}^{B} All_{s,i,j,k} \quad E[NonSel_{s,j}] = \frac{1}{A \times B} \sum_{i=1}^{A} \sum_{k=1}^{B} NonSel_{s,i,j,k}$$

In practice, we also want 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 non-synonymous substitution, which we empirically adjust to obtain a ratio that matches 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 use a distribution fitted to data by ref. [18]). We then obtain the expected density of non-synonymous sites by integrating over the distribution of selection coefficients, which we do in practice by performing a large number of simulations for each of a range of selection coefficients, and then grid-averaging:

$$E[All_{nonsyn,j}] = z_{nonsyn} \int E[All_{s,j}] f(s) ds \quad E[NonSel_{nonsyn,j}] = z_{nonsyn} \int E[NonSel_{s,j}] f(s) ds$$

We also define the expectation for synonymous sites:

$$E[All_{syn,j}] = E[NonSel_{syn,j}] = (1 - z_{nonsyn}) E[All_{0,j}]$$

We define the proportion of sites that are non-synonymous in a given generation as:

$$PropAll_j = \frac{All_{nonsyn,j}}{All_{nonsyn,j} + All_{syn,j}} \quad PropNonSel_j = \frac{NonSel_{nonsyn,j}}{NonSel_{nonsyn,j} + NonSel_{syn,j}}$$

The expected change in the proportion of non-synonymous sites in generation *j* is:

$\delta PropAll_j = PropAll_j - PropAll_{j-1}$	[all evolutionary forces]
$\delta PropNonSel_j = PropNonSel_j - PropAll_{j-1}$	[mutation and drift only]
$\delta PropSel_j = PropAll_j - PropNonSel_j.$	[selective forces only]

We define 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 two populations had not yet split 2500 generations ago and were in mutation-selection-drift equilibrium):

$$PropSel_j = \delta PropSel_j - \delta PropSel_{-2,500}$$

$$PropNonSel_j = \delta PropNonSel_j - \delta PropNonSel_{-2,500}$$

$$PropAll_j = \delta PropAll_j - \delta PropAll_{-2,500}$$

These statistics are positive if the effectiveness of 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

1. Charlesworth B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 2009; 10:195–205. [PubMed: 19204717]

2. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011; 475:493–496. [PubMed: 21753753]

3. Keinan A, Mullikin JC, Patterson N, Reich D. Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. Nat Genet. 2007; 39:1251–1255. [PubMed: 17828266]

4. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. Bayesian inference of ancient human demography from individual genome sequences. Nat Genet. 2011; 43:1031–1034. [PubMed: 21926973]

5. Lohmueller KE, et al. Proportionally more deleterious genetic variation in European than in African populations. Nature. 2008; 451:994–997. [PubMed: 18288194]

6. Tennessen JA, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012; 337:64–69. [PubMed: 22604720]

7. Fu W, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature. 2013; 493:216–220. [PubMed: 23201682]

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. 2013; 9:e1003815. [PubMed: 24086152]

9. Lim ET, et al. Distribution and Medical Impact of Loss-of-Function Variants in the Finnish Founder Population. PLoS Genet. 2014; 10:e1004494. [PubMed: 25078778]

10. Kidd JM, et al. Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. Am J Hum Genet. 2012; 91:660–671. [PubMed: 23040495]

11. Carmi S, et al. Sequencing an Ashkenazi reference panel supports population-targeted personal genomics and illuminates Jewish and European origins. Nat Commun. 2014; 5:4835. [PubMed: 25203624]

12. Kunsch HR. The Jackknife and the Bootstrap for General Stationary Observations. Annals of Statistics. 1989; 17:1217–1241.

13. Abecasis GR, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]

14. Meyer M, et al. A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012; 338:222–226. [PubMed: 22936568]

15. Prufer K, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature. 2014; 505:43–49. [PubMed: 24352235]

16. Simons YB, Turchin MC, Pritchard JK, Sella G. The deleterious mutation load is insensitive to recent population history. Nat Genet. 2014

17. Gravel S, et al. Demographic history and rare allele sharing among human populations. Proc Natl Acad Sci U S A. 2011; 108:11983–11988. [PubMed: 21730125]

18. Boyko AR, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet. 2008; 4:e1000083. [PubMed: 18516229]

19. Gazave E, Chang D, Clark AG, Keinan A. Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. Genetics. 2013; 195:969–978. [PubMed: 23979573]

20. Lohmueller KE. The impact of population demography and selection on the genetic architecture of complex traits. PLoS Genet. 2014; 10:e1004379. [PubMed: 24875776]

21. Lohmueller KE. The distribution of deleterious genetic variation in human populations. bioRxiv. 2014

22. Balick DJ, Do R, Reich D, Sunyaev SR. Response to a population bottleneck can be used to infer recessive selection.

23. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Chapter 7. Curr Protoc Hum Genet. 2013; (Unit7):20. [PubMed: 23315928]

24. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001; 11:863–874. [PubMed: 11337480]

25. Duret L, Galtier N. Biased gene conversion and the evolution of mammalian genomic landscapes. Annu Rev Genomics Hum Genet. 2009; 10:285–311. [PubMed: 19630562]

26. Schuster SC, et al. Complete Khoisan and Bantu genomes from southern Africa. Nature. 2010; 463:943–947. [PubMed: 20164927]

ADDITIONAL REFERENCES FOR ONLINE METHODS

27. Gnirke A, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 2009; 27:182–189. [PubMed: 19182786]

28. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]

29. Sunyaev SR, et al. PSIC: profile extraction from sequence alignments with position-specific counts of independent observations. Protein Eng. 1999; 12:387–394. [PubMed: 10360979]

	Denisova	Neanderthal	IBS (Spanish)	GBR (British)	FIN (Finns)	CEU (European)	JPT (Japanese)	CHS (Chinese)	CHB (Chinese)	PUR (Puerto Rican)	MXL (Mexican)	CLM (Colomb.)	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
				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)	ASW
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 non-synonymous 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 observed in the population in the row divided by the population in the column. ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation. Ratios for Neanderthal and Denisova 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 due to less time elapsed since divergence (all other comparisons are un-normalized). Ratios involving Neanderthal and Denisova also remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation.
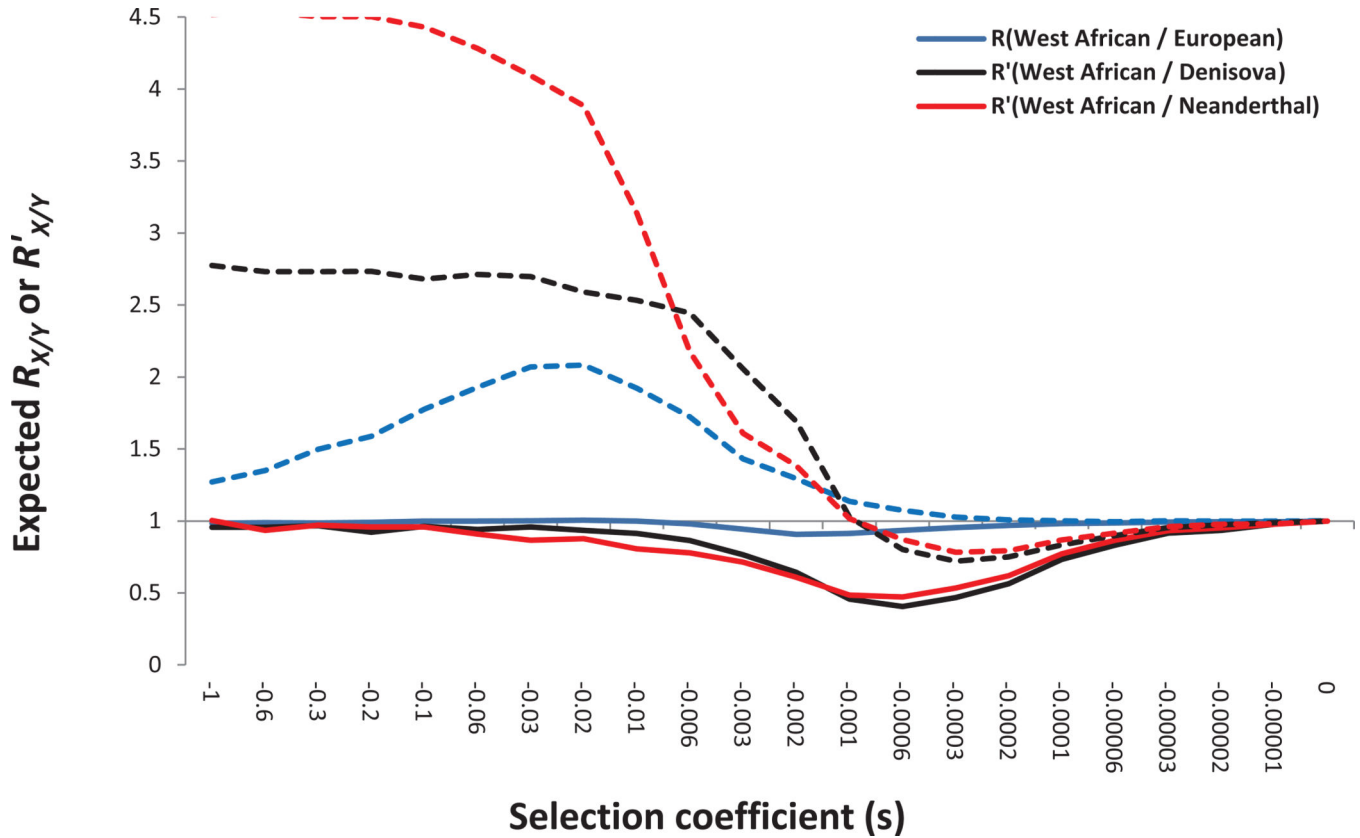
Figure 2. The effect of demographic history on the accumulation of deleterious mutations
To study the expected value of $R_{WestAfrica/Europe}$ stratified by selection coefficient, we simulated a previously published model of the joint history of West Africans and Europeans[6], for a range of selection coefficients, assuming both additive ($h=0.5$) and recessive ($h=0$) models of selection. For the additive case, $R_{WestAfrica/Europe}$ 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 thus a true signal may be difficult to detect. We also simulated a published model of the history of Denisovans and Neanderthals[15]. The simulations predict similar curves for $R'_{WestAfrica/Denisova}$ and $R'_{WestAfrica/Neanderthal}$ 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 $R'_{WestAfrica/Denisova}$ is expected to be below a detectable ratio of 0.95 for $s \in (-0.00002, -0.03)$ and that $R'_{WestAfrica/Neanderthal}$ is expected to be below 0.95 for $s \in (-0.00002, -0.09)$. For recessively acting alleles, the directionality of the effects are opposite.

a

b

c
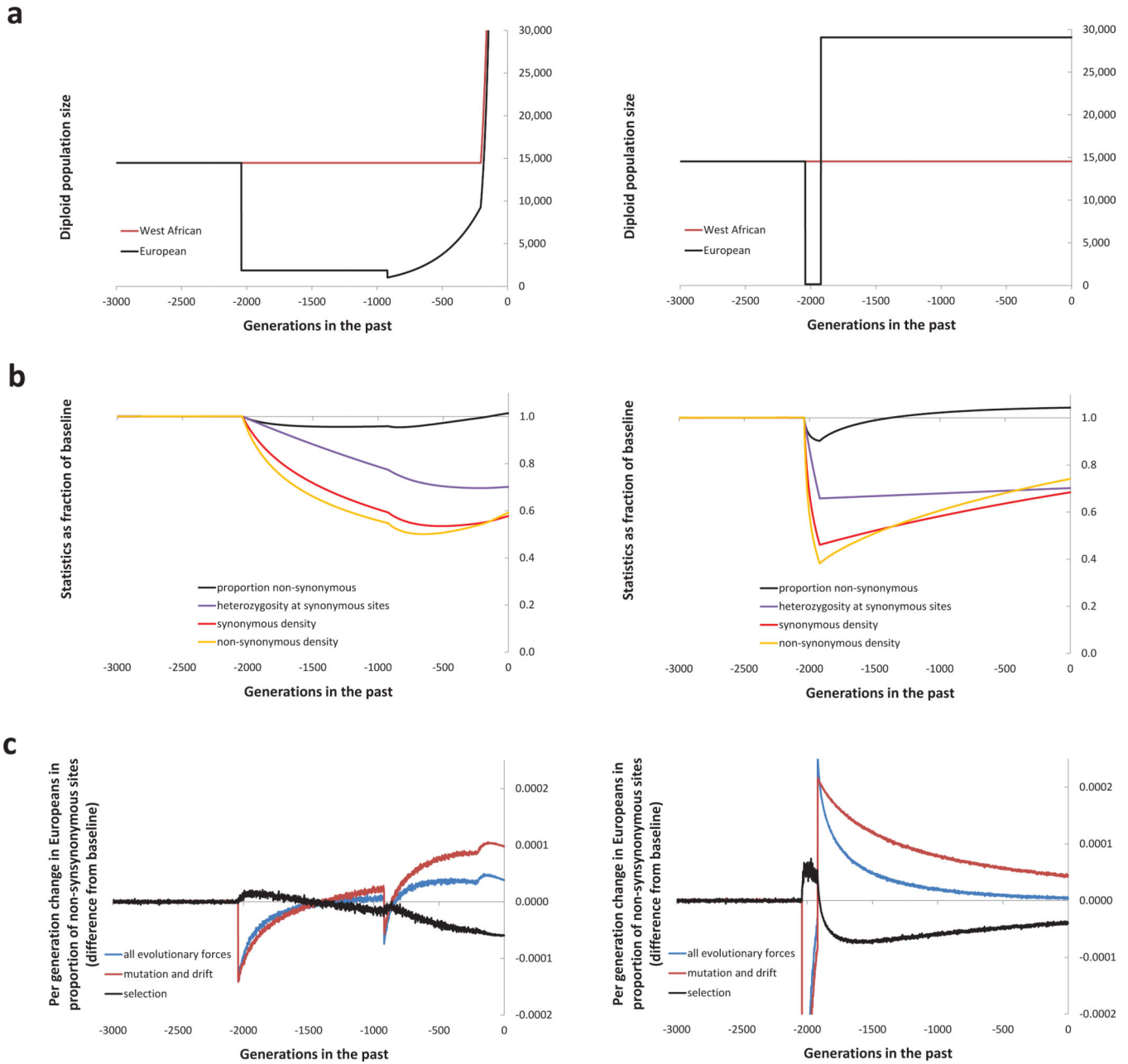
Figure 3. The rise in the proportion of non-synonymous sites in Europeans compared with West Africans is not due to a reduced effectiveness of selection in Europeans since the split

(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 non-synonymous sites in Europeans is higher than in the ancestral population (black). We also show heterozygosity at unselected sites (gray), synonymous site density (red), and non-synonymous site density (yellow). (C) Partitioning of the change in the

proportion of non-synonymous sites per generation into selective and others 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 non-synonymous mutations due to selection minus its value prior to 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 non-synonymous mutations, just that the decrease per generation due to this quantity is less than in the past.

Table 1

Accumulation of different classes of mutation in exomes of West African compared to exomes of European ancestry

Data set	West Africans	Europeans	R: Relative accumulation of mutations				R²: Relative accumulation of homozygous mutations				
			R(synonymous)	R(All non-synonymous)	R(Benign)	R(Possibly damaging)	R(Probably damaging)	R²(synonymous)	R²(All non-synonymous)	R²(Benign)	R²(Possibly damaging)	R²(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)
1KG 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)

Notes: ±1 standard errors are obtained from a Weighted Block Jackknife with 100 equally sized blocks. For the whole genomes, Yoruba+Mandenka represent West Africans, and French+Sardinian represent Europeans. For the 1000 Genomes Data (1KG), YRI represent West Africans and CEU represent Europeans. The Celera and ESP analyses represent people of West African ancestry using African Americans.

Table 2

$R_{X/Y}$(*nonsynonymous*) for all pairs of deep genome populations (bottom left) and 1000 Genomes populations (top right)

(deep row label)		IBS (Spanish)	GBR (British)	FIN (Finns)	CEU (European)	JPT (Japanese)	CHS (Chinese)	CHB (Chinese)	PUR (Puerto Rican)	MXL (Mexican)	CLM (Colomb.)	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**
Denisova	**Denisova (1)**	**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**
Neandertal	0.857 (0.043)	**Neandertal(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**
Dinka	0.849 (0.035)*	0.98 (0.042)	**Dinka (2)**	**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**
Mandenka	0.853 (0.038)	0.99 (0.04)	1.014 (0.017)	**Mandenka (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**
Mbuti	0.88 (0.035)	1.03 (0.041)	1.03 (0.016)	1.017 (0.016)	**Mbuti (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**
San	0.899 (0.037)	1.05 (0.04)	1.014 (0.016)	1.004 (0.016)	0.99 (0.017)	**San (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**
Yoruba	0.85 (0.037)*	0.993 (0.039)	0.99 (0.018)	0.982 (0.016)	0.99 (0.017)	0.975 (0.017)	**Yoruba (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**
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)**	**PUR (55)**	0.996 (0.005)	0.987 (0.003)	0.992 (0.011)	0.978 (0.01)	0.998 (0.008)	**PUR**
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)	0.997 (0.019)	**French (2)**	**MXL (64)**	0.99 (0.004)	0.995 (0.011)	0.981 (0.011)	1.001 (0.009)	**MXL**
Han	0.868 (0.036)	1.01 (0.04)	1.027 (0.018)	0.999 (0.019)	1.00 (0.019)	1.018 (0.019)	1.044 (0.02)	1.021 (0.017)	1.051 (0.019)	**Han (2)**	**CLM (60)**	1.002 (0.011)	0.989 (0.01)	1.01 (0.008)	**CLM**
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)**	**YRI (88)**	0.986 (0.004)	1.007 (0.004)	**YRI**
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)**	**LWK (96)**	1.27 (0.011)*	**LWK**
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)**	**ASW (61)**	
Deep genomes	Denisova	Neandertal	Dinka	Mandenka	Mbuti	San	Yoruba	Dai	French	Han	Karitiana	Papuan	Sardinian		

* highlighted cell (yellow) — numbers >4 standard errors from expectation.

Notes: ±1 standard errors (parentheses) are based on a Weighted Block Jackknife. We highlight numbers >4 standard errors from expectation. Ratios for Neanderthal and Denisova 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 due to less time elapsed since divergence (all other comparisons are un-normalized). Ratios involving Neanderthal and Denisova also remove C→T and G→A mutations to avoid high error rates due to ancient DNA degradation (Table S7). Ratios are based on the accumulation observed in the population in the row divided by that in the population in the column. Numbers in parentheses are sample sizes.