Supplementary Online Information

The deleterious mutation load is insensitive to recent population history

Yuval B. Simons[1,*], Michael C. Turchin[2,*], Jonathan K. Pritchard[2,4,5†] and Guy Sella[3,6,†]

January 27, 2014

[1]Department of Ecology, Evolution, and Behavior, The Hebrew University of Jerusalem

[2]Department of Human Genetics, The University of Chicago

[3]Department of Ecology and Evolution, The University of Chicago

[4]Howard Hughes Medical Institute

[5]Departments of Biology and Genetics, Stanford University

[6]Departments of Biological Sciences, Columbia University

[*]These authors contributed equally.

[†]To whom correspondence should be addressed: pritch@stanford.edu, gsella@math.huji.ac.il.

Nature Genetics: doi:10.1038/ng.2896

List of Supplementary Tables and Figures

Nature Genetics: doi:10.1038/ng.2896

Supplementary table 1:
Changes to load under the bottleneck and growth models

			Effectively neutral	Weak		Strong
				closer to neutral	closer to strong	
Bottleneck	Semi-dominant	fixed	increase	increase	increase	—
		segregating	decrease	decrease	increase	unchanged
		total	unchanged	increase	increase	unchanged
	Recessive	fixed	increase	increase	increase	—
		segregating	decrease	decrease	increase	transient increase
		total	unchanged	increase	increase	transient increase
Growth	Semi-dominant	fixed	decrease	decrease		—
		segregating	increase	increase		unchanged
		total	unchanged	unchanged		unchanged
	Recessive	fixed	decrease	decrease		—
		segregating	increase	increase		transient decrease
		total	unchanged	unchanged		transient decrease

Supplementary Table 1: Changes to load under the bottleneck and growth models. The effects on fixed, segregating and total load are depicted by selection regime. The symbol — denotes the cases in which there is no contribution to load both before and after the change in population size.

Nature Genetics: doi:10.1038/ng.2896

Supplementary table 2:

Estimated mean frequencies in AAs and EAs at different classes of sites

Method	Chr.	Category	# SNVs	AA_{Mean}	AA_{SE}	EA_{Mean}	EA_{SE}	t-score
Non-coding	Aut	—	300209	0.034	0.00026	0.034	0.00028	0.44
Non-coding	X	—	8355	0.030	0.0015	0.028	0.0016	1.1
Synonymous	Aut	—	220391	0.033	0.00030	0.033	0.00032	0.87
Synonymous	X	—	7001	0.028	0.0016	0.029	0.0018	-0.10
Non-synonymous	Aut	—	351265	0.014	0.00015	0.014	0.00016	0.40
Non-synonymous	X	—	10293	0.012	0.00086	0.012	0.00095	0.076
PolyPhen2	Aut	D	121280	0.0078	0.00011	0.0076	0.00012	1.2
PolyPhen2	Aut	P	65400	0.012	0.00018	0.012	0.00020	0.52
PolyPhen2	Aut	B	132047	0.019	0.00024	0.019	0.00026	0.55
PolyPhen2	X	D	3205	0.0072	0.00065	0.0079	0.00078	-0.99
PolyPhen2	X	P	1957	0.013	0.0012	0.012	0.0012	0.98
PolyPhen2	X	B	3948	0.014	0.0011	0.014	0.0012	0.044
Sift	Aut	D	145986	0.0095	0.00012	0.0093	0.00013	1.6
Sift	Aut	T	180091	0.018	0.00021	0.018	0.00022	-0.13
Sift	X	D	4251	0.0099	0.00076	0.0096	0.00082	0.34
Sift	X	T	5517	0.017	0.0013	0.017	0.0015	-0.29
LRT	Aut	D	146701	0.0060	8.5e-05	0.0060	9.5e-05	-0.11
LRT	Aut	N	160179	0.020	0.00024	0.020	0.00026	0.20
LRT	Aut	U	13845	0.0066	0.00036	0.006	0.00039	2.6
LRT	X	D	3270	0.0038	0.00037	0.0034	0.00034	0.93
LRT	X	N	4548	0.017	0.0014	0.017	0.0016	-0.37
LRT	X	U	886	0.0052	0.0013	0.0046	0.0015	0.40
MutationTaster	Aut	D	155138	0.0022	2.9e-05	0.0017	3.0e-05	18
MutationTaster	Aut	A	5089	0.00089	9.5e-05	0.00056	4.8e-05	4.3
MutationTaster	Aut	N	161169	0.0062	6.8e-05	0.0047	6.7e-05	21
MutationTaster	Aut	P	9040	0.36	0.0047	0.39	0.0051	-6.5
MutationTaster	X	D	3860	0.021	0.0021	0.023	0.0023	-1.2
MutationTaster	X	A	76	0.0010	0.00058	0.00039	0.00017	1.5
MutationTaster	X	N	5566	0.0030	0.00026	0.0013	0.00022	7.0
MutationTaster	X	P	131	0.16	0.028	0.16	0.029	0.28

Nature Genetics: doi:10.1038/ng.2896

Supplementary Table 2: Comparison of mean frequencies in AAs and EAs at different classes of sites, classified according to whether the sites are on the autosomes or X, and using a variety of different functional classifications (after application of our bias-correction method). For this table, the data were subsampled down to 3852 chromosomes for AAs and EAs each, to enable X vs autosome comparisons. Note that the mean frequencies in each row are not significantly different ($|t - score| < 2$, with the sole exception of the functional classifications from MutationTaster (which are highly significant). The unusual results for MutationTaster likely arise because MutationTaster uses previously estimated population frequencies in its classification, thus introducing further biases for population genetic analysis that are not properly addressed by correction method.

Nature Genetics: doi:10.1038/ng.2896

Supplementary table 3:
Estimated mean frequencies with and without bias correction

Method	Chr.	Category	Without bias correction				With bias correction			
			AA_{Mean}	AA_{SE}	EA_{Mean}	EA_{SE}	AA_{Mean}	AA_{SE}	EA_{Mean}	EA_{SE}
Non-synonymous	Aut	—	0.014	0.00015	0.014	0.000162	0.014	0.00015	0.014	0.00016
PolyPhen2	Aut	D	0.0038	9.3E-05	0.0033	1.0E-04	0.0078	0.00011	0.0076	0.00012
PolyPhen2	Aut	P	0.0060	0.00017	0.0053	0.00019	0.012	0.00018	0.012	0.00020
PolyPhen2	Aut	B	0.026	0.00035	0.026	0.00037	0.019	0.00024	0.019	0.00026
Sift	Aut	D	0.0061	0.00013	0.0055	0.00014	0.0095	0.00012	0.0093	0.00013
Sift	Aut	T	0.020	0.00026	0.021	0.00028	0.018	0.00021	0.018	0.00022
LRT	Aut	D	0.0028	6.4E-05	0.0025	7.4E-05	0.0060	8.5e-05	0.0060	9.5e-05
LRT	Aut	N	0.023	0.00029	0.023	0.00031	0.020	0.00024	0.020	0.00026
LRT	Aut	U	0.0081	0.00048	0.0071	5.0E-04	0.0066	0.00036	0.006	0.00039
MutationTaster	Aut	D	0.0017	4.3E-05	0.0011	4.3E-05	0.0022	2.9e-05	0.0017	3.0e-05
MutationTaster	Aut	A	0.0013	0.00034	0.00099	0.00032	0.00089	9.5e-05	0.00056	4.8e-05
MutationTaster	Aut	N	0.013	0.00024	0.012	0.00025	0.0062	6.8e-05	0.0047	6.7e-05
MutationTaster	Aut	P	0.26	0.0027	0.30	0.0032	0.36	0.0047	0.39	0.0051

Supplementary Table 3: Comparison of estimated mean frequencies in samples of 3852 chromosomes, with and without bias correction of the functional annotations. Recall that we observed that all four functional prediction methods typically have low probabilities of assigned 'damaging' status to SNVs where the genome reference carries the derived allele. Notice that prior to applying the bias correction (using all SNVs), AAs tend to have higher allele frequencies at putatively damaging sites, as reported by Tennessen et al. This is likely because most of the reference genome is of non-African origin. After applying our bias correction, we observe that AAs and EAs have essentially identical allele frequencies in all functional categories (except for MutationTaster, likely for reasons discussed above).

Nature Genetics: doi:10.1038/ng.2896

Supplementary table 4:
Estimated mean frequencies using different methods for classifying sites

Category	AA_{Mean}	AA_{SE}	EA_{Mean}	EA_{SE}	T-Stat
Uncorrected (biased) PolyPhen Scores					
Prob. Damaging	0.00277	6.79e-05	0.00239	7.31e-05	5.4
Poss. Damaging	0.00452	0.00013	0.00401	0.00014	3.84
Benign	0.0208	0.000278	0.0212	0.000297	-1.34
Bias-corrected PolyPhen Scores					
Prob. Damaging	0.00593	8.11e-05	0.00582	8.76e-05	1.23
Poss. Damaging	0.00955	0.00014	0.00948	0.000151	0.488
Benign	0.0154	0.000186	0.0153	2e-04	0.527
Human-independent PolyPhen Scores					
3<PSIC	0.0056	0.0002	0.0054	0.0003	0.45
1.5<PSIC<3	0.011	0.0002	0.011	0.0002	-0.06
PSIC<1.5	0.019	0.0003	0.019	0.0003	-0.07

Supplementary Table 4: Comparison of estimated mean frequencies at autosomal nonsynonymous sites in the Fu et al data, using the full autosomal samples. The top block of data use the uncorrected (biased) PolyPhen scores, and suggest significant differences between populations. The middle block of data applies our bias correction, and shows no significant differences between populations. The bottom block of data uses an unpublished version of the PolyPhen "PSIC" scores that are calculated independent of the human reference sequence, and hence are unbiased (kindly provided by the Shamil Sunyaev lab). These too show no significant difference between populations. Note that DAFs differ between the second two blocks of data due to arbitrary choices in score cutoffs.

Nature Genetics: doi:10.1038/ng.2896

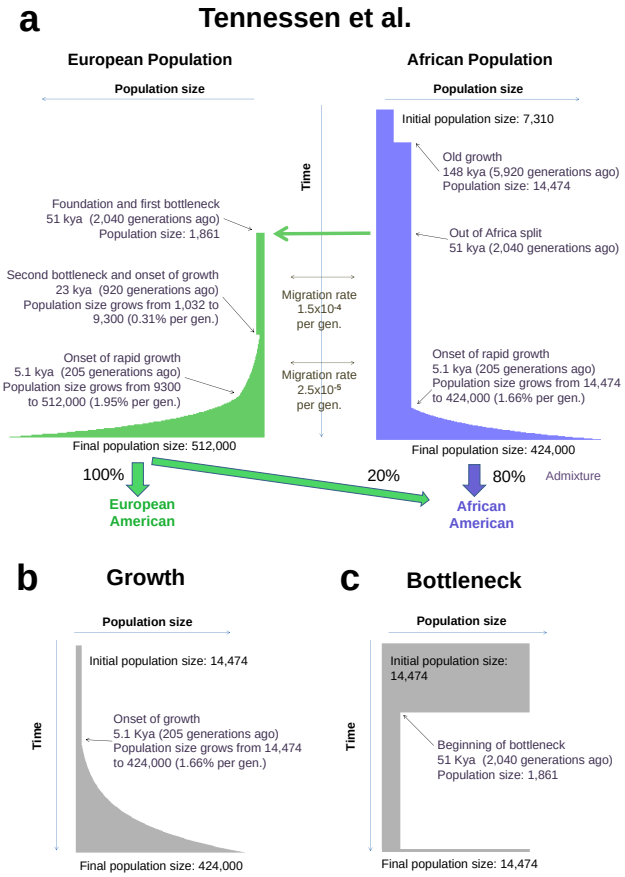
Supplementary table 5:
Summary of 1000 Genomes analysis

Category	YRI_{Mean}	YRI_{SE}	CEU_{Mean}	CEU_{SE}	P-value
Individual-Level Counts					
Synonymous	18,141	119	17,992	122	N.S.
Nonsynonymous	9903	104	9825	80	N.S.
Prob. Damaging	2153	31	2111	26	N.S.
Poss. Damaging	1851	27	1836	24	N.S.
Benign	5899	67	5878	55	N.S.

Supplementary Table 5: Summary of 1000 Genomes Analysis. This table shows the mean numbers of derived alleles per individual in the YRI and CEU populations. The functional categories (Probably/Possibly Damaging and Benign) were obtained from PolyPhen, and adjusted using our bias correction method. SEs obtained by bootstrapping across SNVs. We also obtained identical conclusions (i.e., no difference between populations) when the analysis was done in terms of DAFs, and also when we used the human-independent PolyPhen (PSIC) scores.

Nature Genetics: doi:10.1038/ng.2896

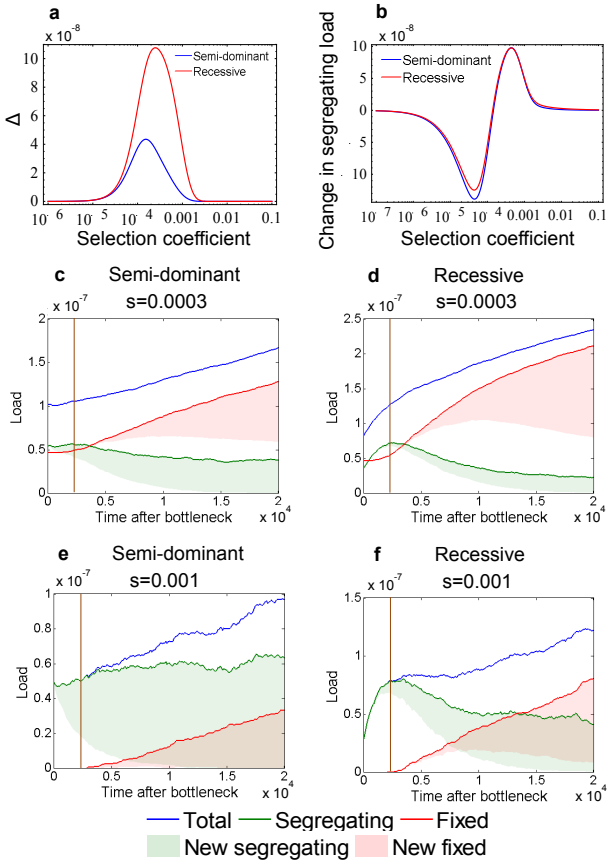
Supplementary figure 1:
Demographic scenarios



Supplementary Fig. 1: The three demographic models that we consider. A) The Out-of-Africa model estimated by Tennessen et al. [2]. C) Exponential growth.B) A population bottleneck. All population sizes are given as number of diploid individuals. In some cases, in order to study the equilibration process, we extend the growth scenario to include a priod with a constant population size after growth and the bottleneck model to include a longer period with a reduced population size.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 2:
Changes in load shortly after a bottleneck

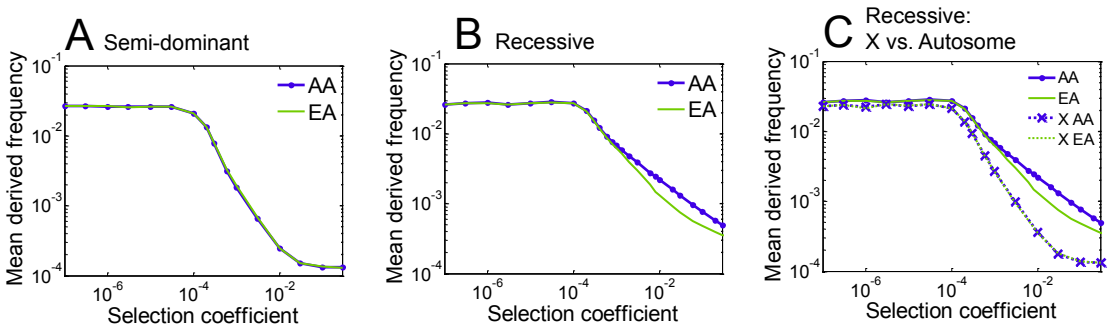


Supplementary Fig. 2: The changes in load shortly after a bottleneck. The figure shows (A) the expected change in fixed load due to mutations that segregated before the bottleneck and (B) the expected change in segregating load due to the bottleneck as a function of the selection coefficient. Shown are segregating, fixed and total load from new and all mutations as a function of time since the population size decrease. The semi-dominant (C and E) and recessive cases (D and F) are shown with a selection coefficient in the weak selection regime closer to neutral ($s = 0.0003$) and closer to strong ($s = 0.001$).

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 3:

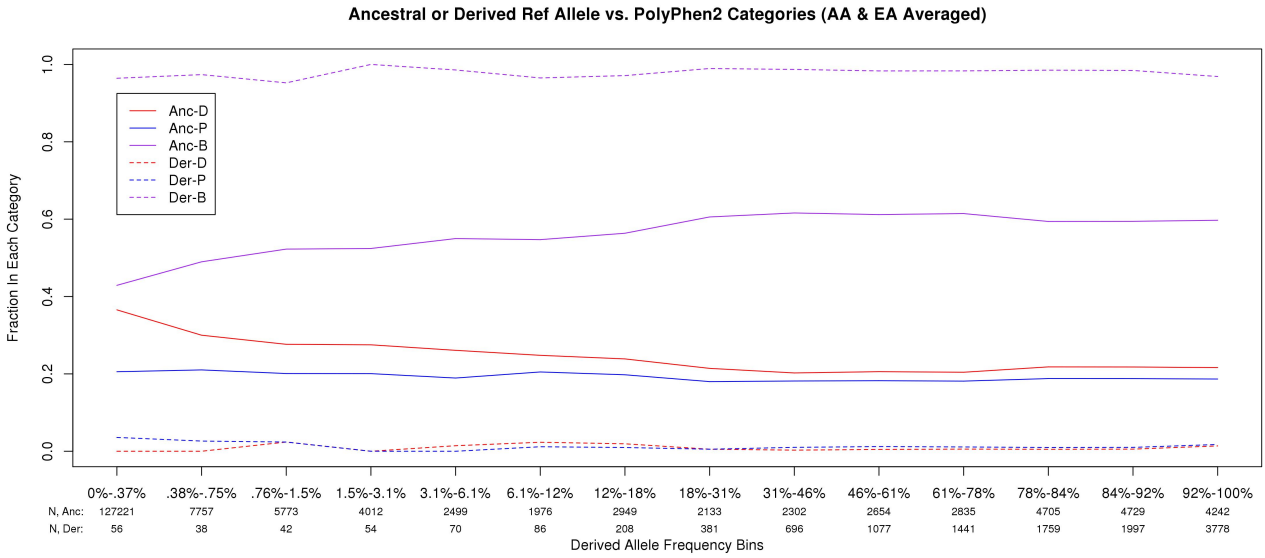
Predicted mean derived frequencies as a function of selection coefficient



Supplementary Fig. 3: Mean derived frequencies predicted as a function of selection coefficient, for the AA and EA demographies. Notice that in (**A**) we predict that for semi-dominant sites AAs and EAs should have essentially identical mean derived frequencies for all levels of selection. In (**B**) we predict a small increase in mean frequencies for AAs at recessive sites with moderate-strong selection. (**C**) provides X vs autosome comparisons under the recessive model; note that recessive alleles on the X experience selection as dominant alleles in males.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 4:
Reference bias in PolyPhen 2

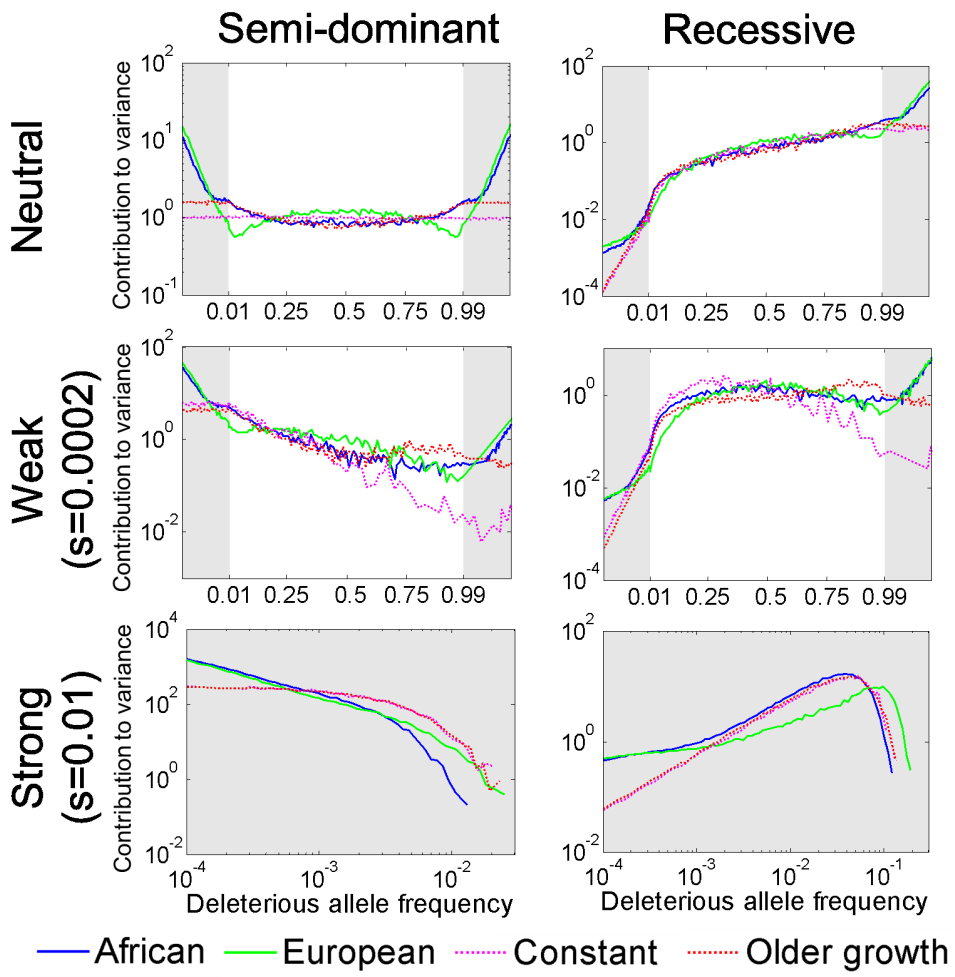


Supplementary Fig. 4: Illustration of the reference bias present in PolyPhen 2 [22]. The other functional prediction methods that we considered have a similar bias. The x-axis shows the mean population frequency of nonsynonymous SNVs in the Fu et al data (the left-most bins cover very narrow intervals of frequencies since most of the data are present in these bins). The y-axis plots the fraction of SNVs in each bin that are classified into each of the three PolyPhen categories: **B**enign, **P**ossibly damaging, Probably **D**amaging; and shown separately according to whether the genome reference sequence carries the ancestral or the derived allele. Notice that when the reference carries the ancestral allele, an SNV is classified as Damaging with a probability that ranges from nearly 40% at low frequencies to ≈20% at high frequencies (solid red line). In contrast, for SNVs where the reference carries the derived allele, the fraction of Damaging alleles is near 0% at all frequencies (dotted red line).

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 5:

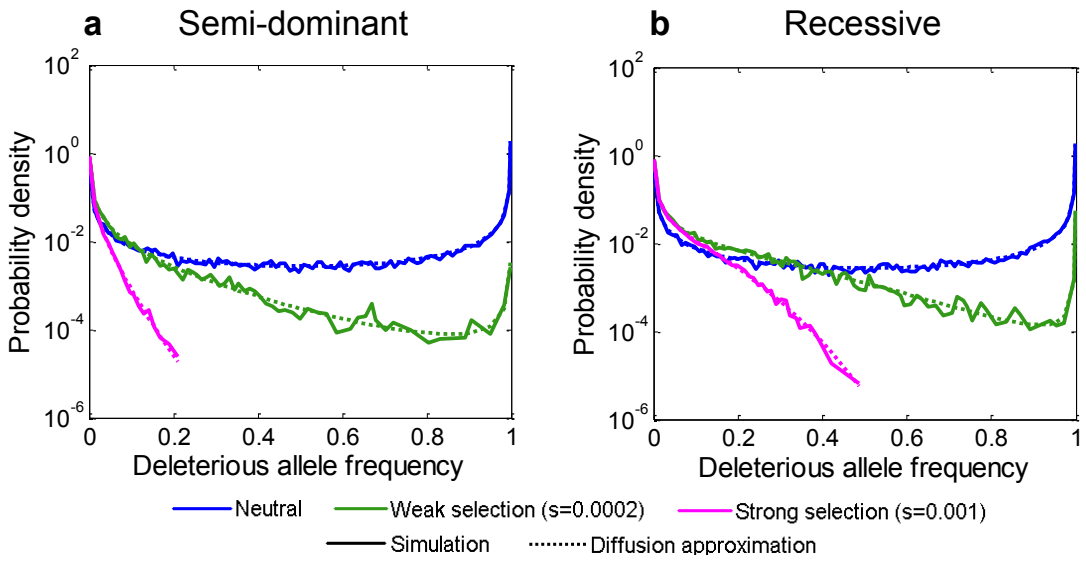
Contribution of different allele frequencies to variance in disease risk



Supplementary Fig. 5: The proportional contribution of different allele frequencies to variance in disease risk, under the Tennessen et al. model for Africans and Europeans. Shaded regions correspond to a logarithmic scale on the x-axis, which is included to show the (minor) effects of recent growth.

Nature Genetics: doi:10.1038/ng.2896

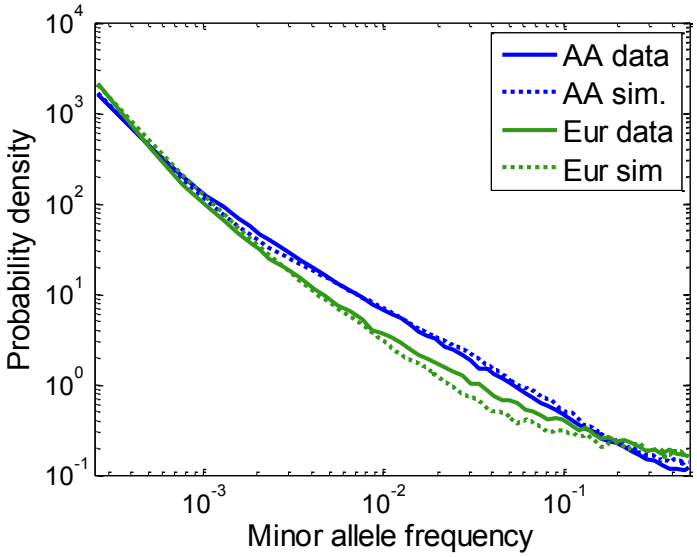
Supplementary figure 6:
Comparison of theoretical and simulated frequency spectra



Supplementary Fig. 6: Comparison of theoretical and simulated frequency spectra for a constant population size in the (A) semi-dominant and (B) recessive models. Shown are the results based on the diffusion approximation (solid) and on simulations (dashed) for several selection coefficients. The population size was taken as $N = 14,474$ and the mutation rate as $u = 2.36 \cdot 10^{-8}$ per generation per site. The number of runs for each set of parameters was 10^6.

Nature Genetics: doi:10.1038/ng.2896

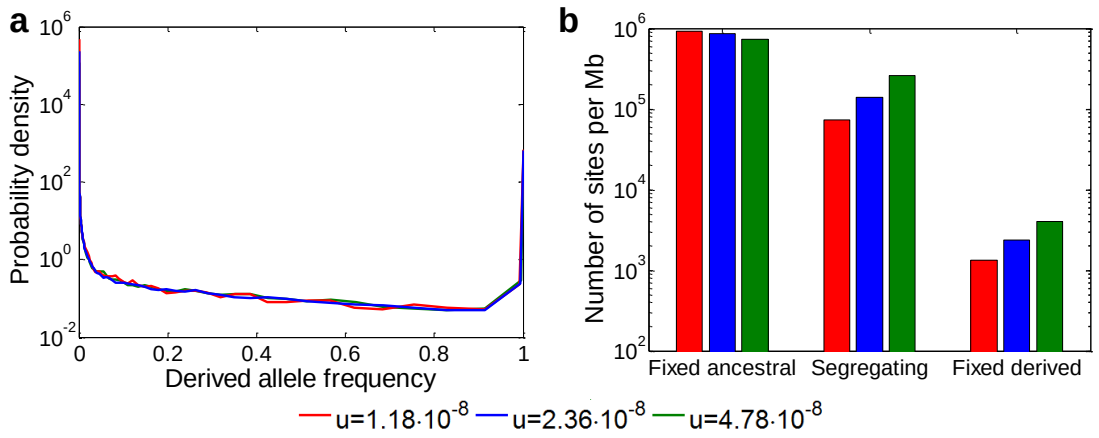
Supplementary figure 7:
Comparison of estimated and simulated frequency spectra



Supplementary Fig. 7: Comparison of the minor allele frequency spectrum in data from Fu et. al. and in simulations based on the Tennessen et al. model. The spectra are for a sample size of 3852 chromosomes in AA and EA populations, for both the data and simulations.

Nature Genetics: doi:10.1038/ng.2896

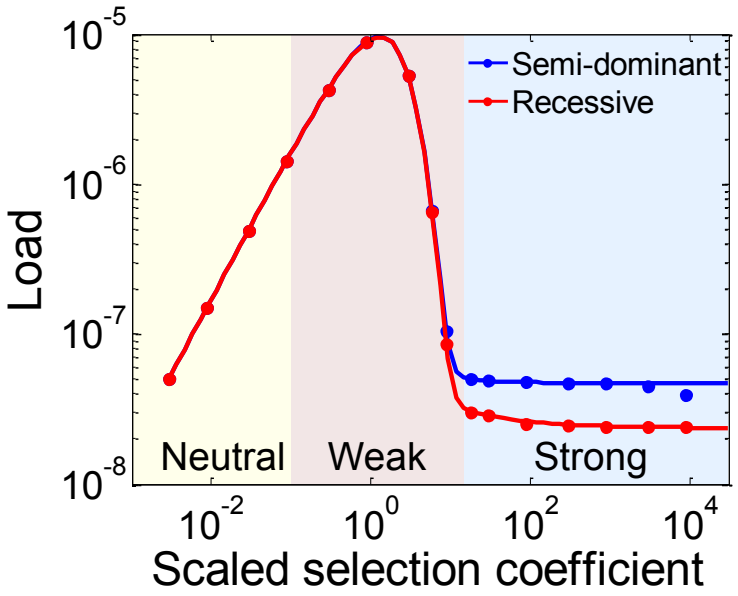
Supplementary figure 8:
Sensitivity to mutation rate



Supplementary Fig. 8: Sensitivity of (A) the frequency spectrum and (B) the number of segregating and fixed sites to the mutation rate. The results are shown for simulations of the African population but are qualitatively similar for the European population.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 9:
Load in a population of constant size

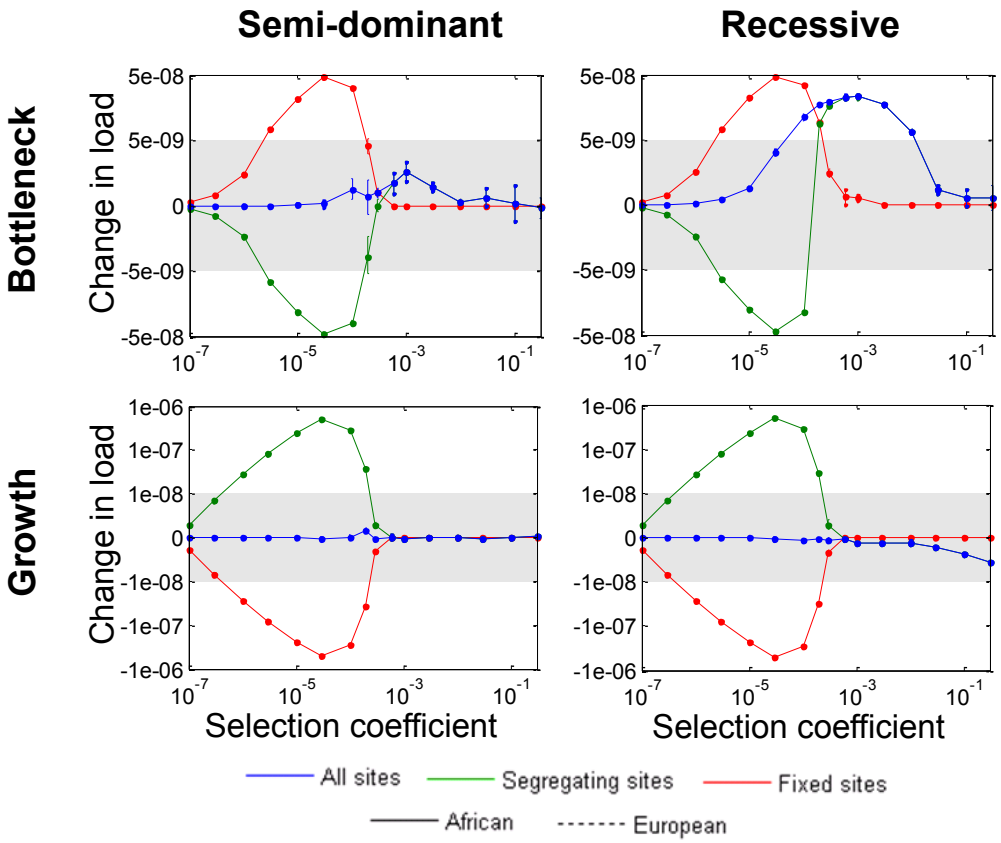


Supplementary Fig. 9: Load as a function of selection coefficient in a population of constant size. Results are shown for the semi-dominant (blue) and recessive models (red), where the diffusion approximation is shown as a solid line and simulation results as circles. The population size is $N = 14,474$.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 10:

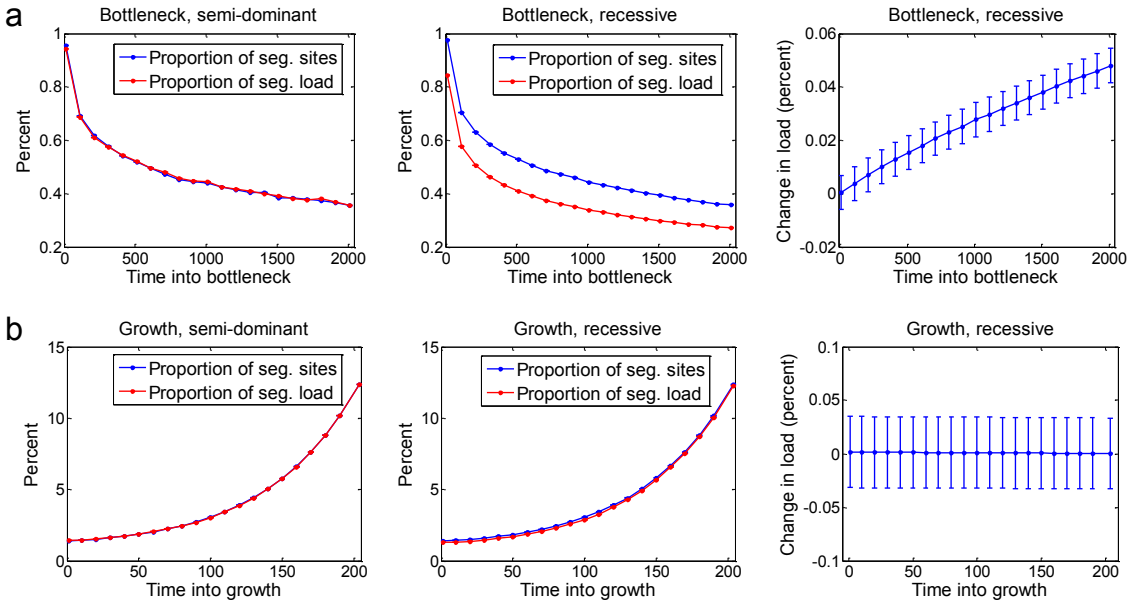
Changes to load under the bottleneck and growth models



Supplementary Fig. 10: The changes to the segregating, fixed and total load under the bottleneck and growth models. Analogous graphs for the Tennessen et al. model are presented in Figure 3 of the main text. Changes are measured by comparison to a population in which the population size has remained constant at the size that it was at the beginning of the demographic model. In the shaded areas, load is shown on linear scale; otherwise it is shown on logarithmic scale.

Nature Genetics: doi:10.1038/ng.2896

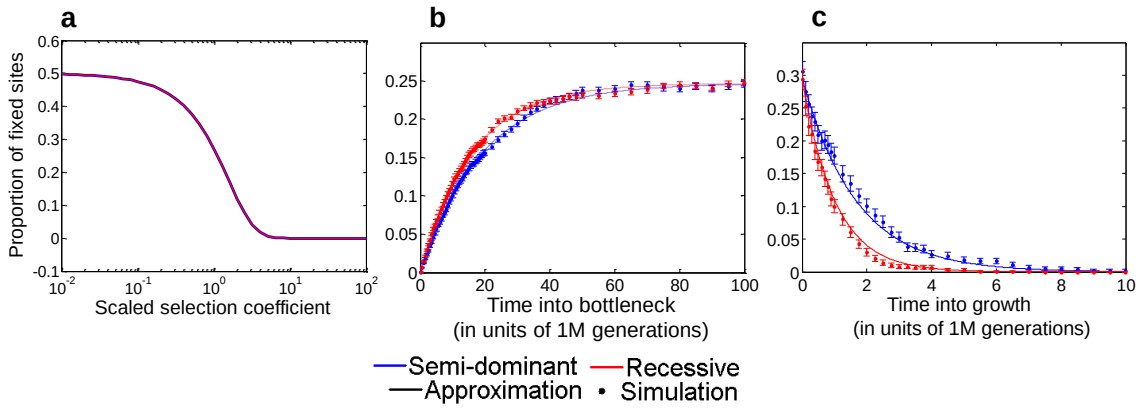
Supplementary figure 11:
Load in the effectively neutral regime



Supplementary Fig. 11: Segregating and total load in the bottleneck and growth models in the effectively neutral regime. The proportion of segregating sites, their proportional contribution to load, and the proportional change in total load are shown as a function of time (A) after the bottleneck and (B) since the onset of growth. The selection coefficient is $s = 10^{-7}$. In the semi-dominant case, the expected total load is always $s/2$ regardless of changes in population size; in the recessive case, changes to the proportion of segregating sites affect the total load, but this effect is negligibly small.

Nature Genetics: doi:10.1038/ng.2896

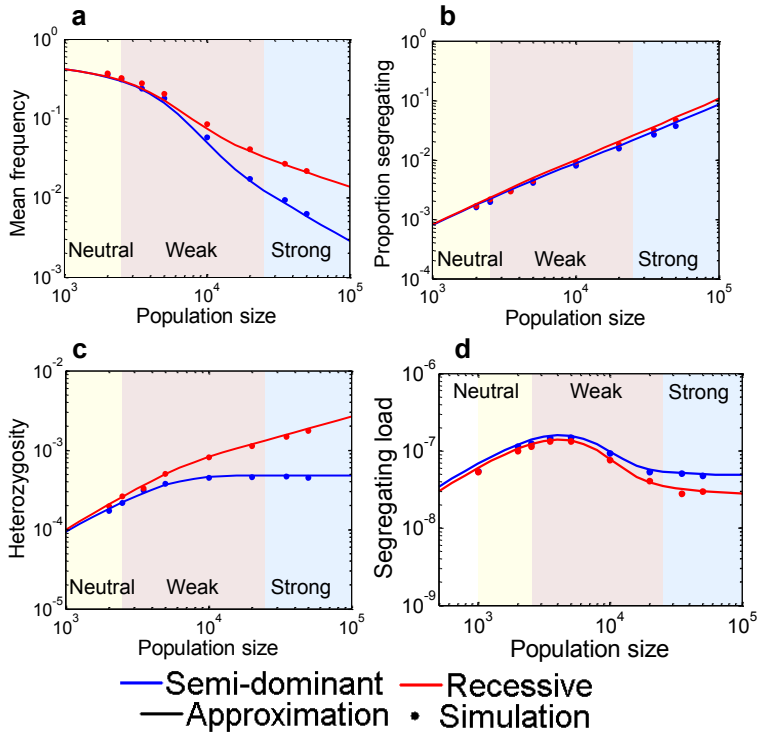
Supplementary figure 12:
Fixed sites in the weak selection regime



Supplementary Fig. 12: Proportion of sites fixed for deleterious alleles in the weak selection regime. In all graphs, the selection coefficient is $s = 10^{-4}$. (A) The equilibrium proportion as a function of the scaled selection coefficient ($\alpha = 2Ns$), where the population size was varied. (B) The proportion as a function of time after the change in population size in the bottleneck model. (C) The proportion as a function of time after the change in population size in the growth model.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 13:
Equilibrium properties of segregating sites

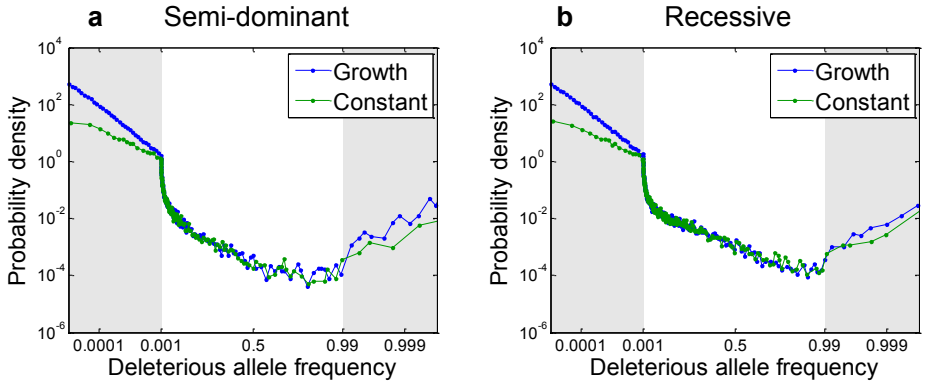


Supplementary Fig. 13: Equilibrium properties of segregating sites as a function of population size in constant population size models. In all graphs, $s = 2 \cdot 10^{-4}$. (A) The average frequency of segregating deleterious alleles. (B) The proportion of segregating sites. (C) Heterozygosity. (D) Segregating load.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 14:

Frequency spectrum of weakly deleterious sites with and without growth

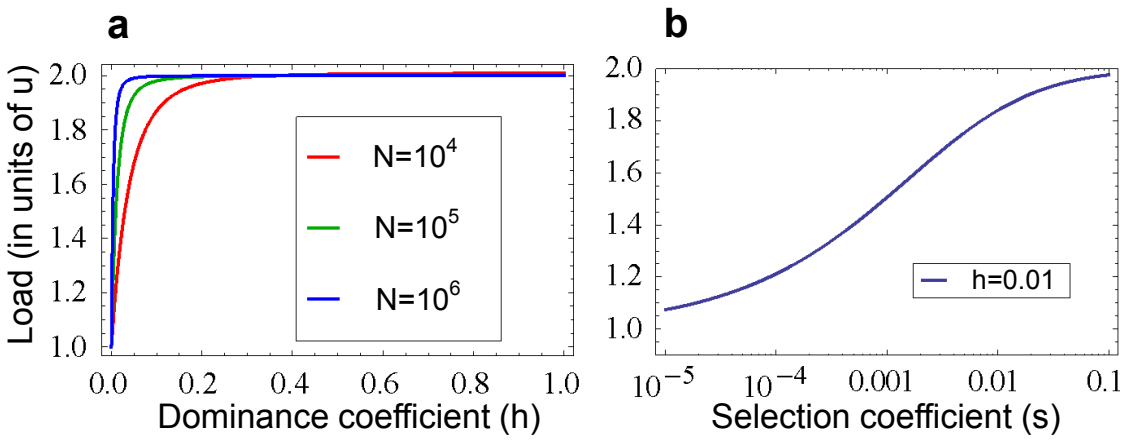


Supplementary Fig. 14: The frequency spectrum of weakly deleterious segregating sites in models with and without growth. In the shaded areas, frequency is shown on logarithmic scale; otherwise it is shown on linear scale.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 15:

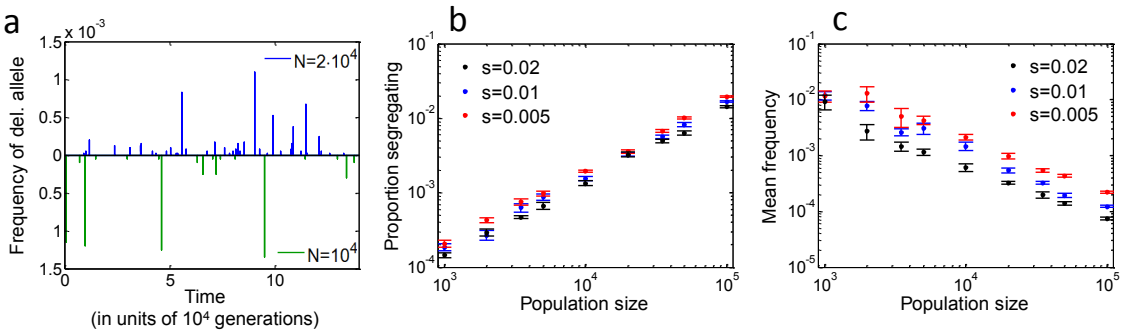
Dependence of the load on the dominance coefficient at equilibrium



Supplementary Fig. 15: The dependence of the load on the dominance coefficient at equilibrium. The graphs were generated using the diffusion approximation for the stationary distribution assuming that the deleterious allele frequency is small [3]. A) Load as a function of the dominance coefficient h, with $s = 0.01$ and population size $N = 10^4, 10^5$ and 10^6. B) Load as a function of the selection coefficient s, with $h = 0.01$ and $N = 10^6$.

Nature Genetics: doi:10.1038/ng.2896

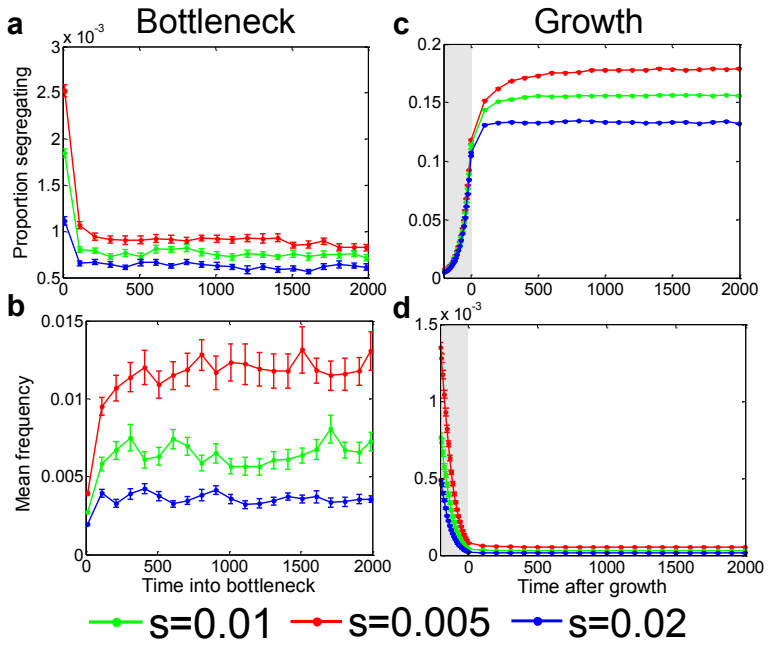
Supplementary figure 16:
Equilibrium properties of segregating sites in the quasi-dominant case



Supplementary Fig. 16: The equilibrium properties of segregating sites in the quasi-dominant case. In all graphs, $h = 0.5$ and $u = 10^{-8}$. A) Frequency of deleterious alleles as a function of time in simulations with two population sizes, corresponding to $N = 10^4$ and $2 \cdot 10^4$. In both cases, $s = 0.01$. B) The expected proportion of segregating sites as a function of population size. C) The expected frequency of deleterious alleles at segregating sites as a function of population size.

Nature Genetics: doi:10.1038/ng.2896

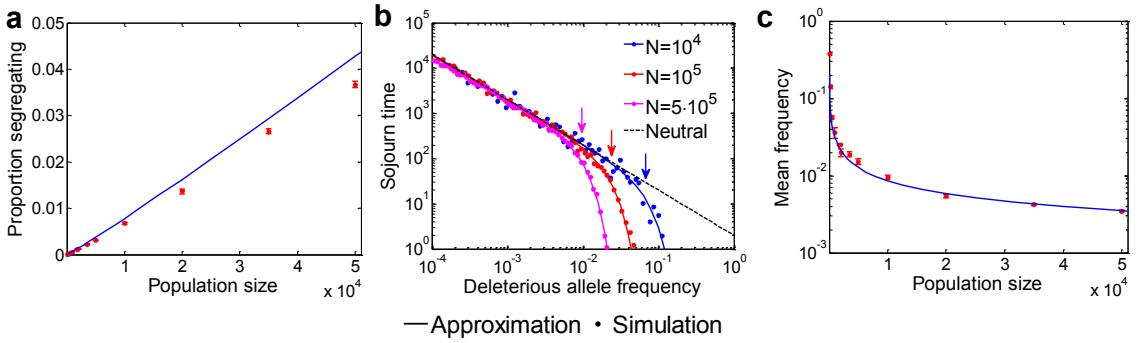
Supplementary figure 17:
Properties of segregating sites as a function of time for the quasi-dominant case



Supplementary Fig. 17: The properties of segregating sites as a function of time for the quasi-dominant case. In all graphs, $h = 0.5$. The proportion of segregating sites after (A) the reduction in population size in the bottleneck model and (C) the onset of growth. The expected frequency of deleterious alleles at segregating sites after (B) the reduction in population size in the bottleneck model and (D) after the onset of growth. The shaded region is the period of growth in the Tennessen model.

Nature Genetics: doi:10.1038/ng.2896

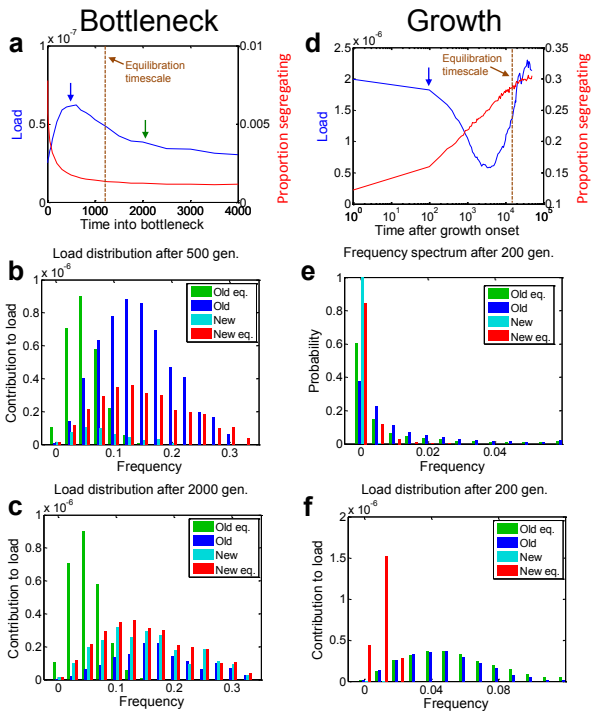
Supplementary figure 18:
Properties of segregating sites at equilibrium in the recessive case



Supplementary Fig. 18: The properties of segregating sites at equilibrium in the recessive case, as a function of population size. The selection coefficient is $s = 0.01$. (A) The proportion of segregating sites.(B) The sojourn time of deleterious alleles for different population sizes. The threshold frequency of $\frac{1}{\sqrt{2Ns}}$ for each population size is marked by an arrow with the corresponding color. (C) The average frequency of deleterious alleles.

Nature Genetics: doi:10.1038/ng.2896

Supplementary figure 19:
Load as a function of time in the recessive case

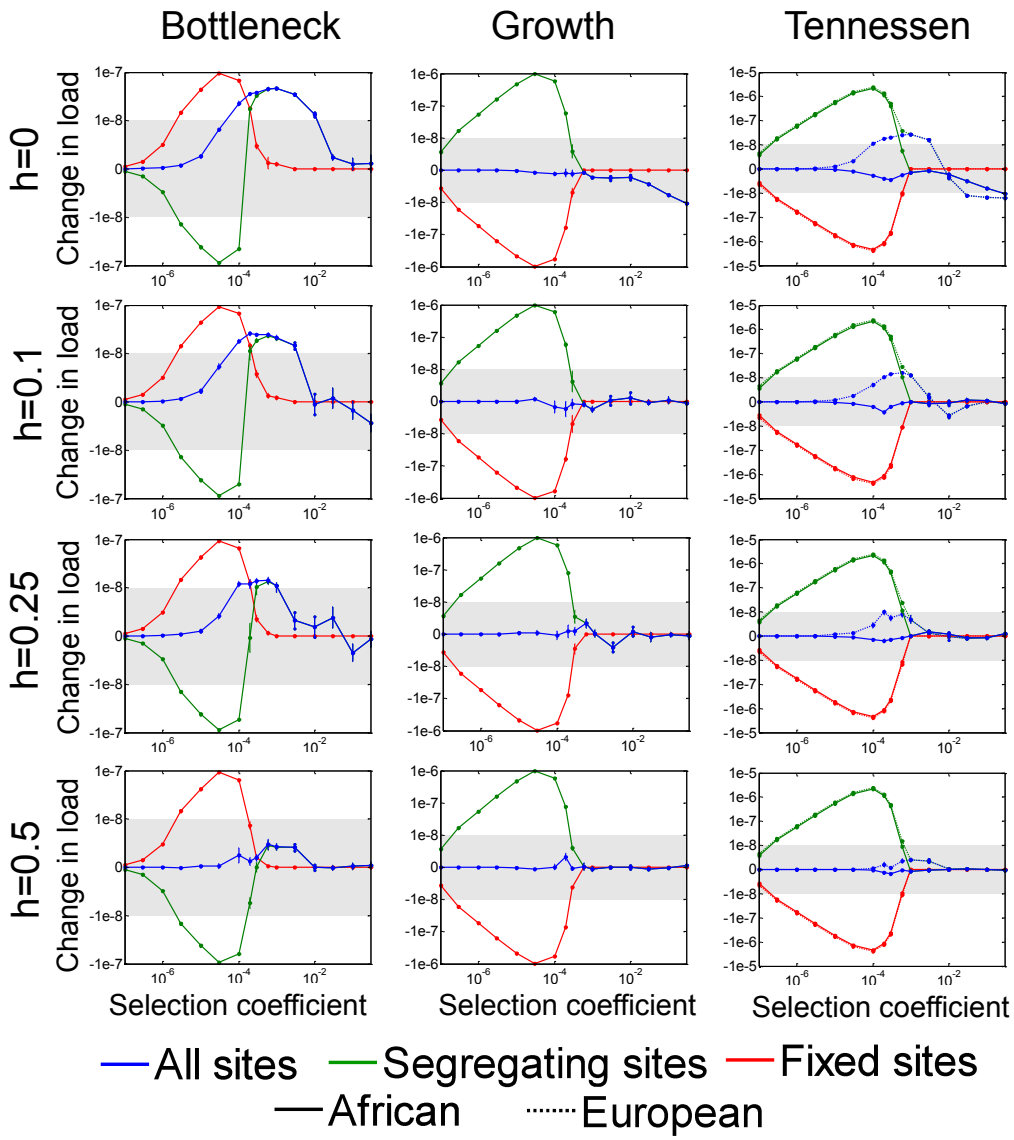


Supplementary Fig. 19: Load as a function of time in the recessive case. The selection coefficient is $s = 0.01$. A) The load and proportion of segregating sites as a function of time after the reduction in population size. B) The contribution to load of old and new mutations as a function of frequency, at the time of peak load (500 generations after the reduction in population size, indicated by a blue arrow in A). C) Same as B but for the time since the Out-of-Africa bottleneck, i.e., 50Kya (indicated by a green arrow in A). D) The load and proportion of segregating sites as a function of time after the onset of growth. E) The allele frequency distribution of old and new mutations at the end of the growth period (200 generations after onset, indicated by an arrow in D). F) The contribution to load of old and new mutations as a function of frequency at the end of the growth period.

Nature Genetics: doi:10.1038/ng.2896

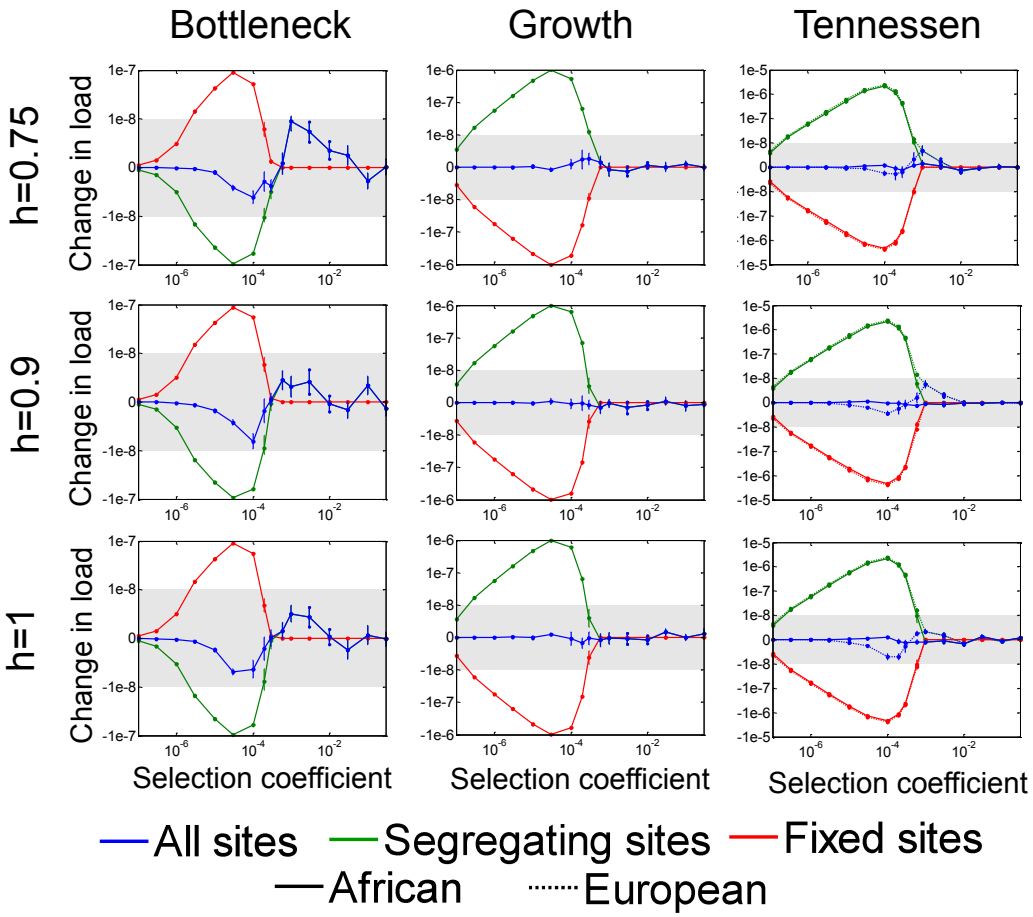
Supplementary figure 20:

Changes in load under the three demographic models with different dominance coefficients



Supplementary Fig. 20: Continued on the next page.

Nature Genetics: doi:10.1038/ng.2896



Supplementary Fig. 20: Changes in load under the three demographic models with different dominance coefficients. $h = 0$ and $1/2$ correspond to the results in Supplementary Figure 10 and are provided for comparison.

Nature Genetics: doi:10.1038/ng.2896

Supplementary note

Contents

Nature Genetics: doi:10.1038/ng.2896

1 Model and simulations

Our basic model considers selection at a single site. We use the standard bi-allelic diploid model with (in this order) two-way mutation, viability selection, drift and, in some cases, migration [1]. Specifically, we assume there are two alleles at a site: normal (N) and deleterious (D). An N allele mutates to the D allele with probability u per gamete, per generation and the reverse mutation occurs with probability v. Unless noted otherwise, we assume that mutation is symmetric, i.e., $u = v$. The absolute fitnesses of the three genotypes NN, ND and DD are 1, $1 - hs$ and $1 - s$, respectively, where $s > 0$ and $h \geq 0$. We focus on semi-dominant ($h = \frac{1}{2}$) and fully recessive ($h = 0$) selection because these two cases exhibit the full range of qualitative behaviors (with selection acting primarily on heterozygotes in one and only on homozygotes in the other), but we also consider the robustness of our findings to other dominance coefficients (section 2.4). Allele frequencies in the next generation follow from Wright-Fisher sampling with these viabilities, sometimes with migration, and the population size and migration rates vary according to the demographic scenario considered.

For each demographic scenario, we ran simulations of a single site for the semi-dominant and recessive cases and varied the selection coefficient such that selection ranges from effectively neutral to strong. For a given set of parameters, the number of runs was determined by requiring a sampling error of less than 2% in estimates of the main summaries (e.g., the mean deleterious allele frequency and squared frequency). Error bars denoting estimates of one standard deviation around the mean are provided in all the graphs based on simulations, unless they are too small to be visible. Each run begins with one of the two alleles fixed, where the proportion of runs that start with each allele is given by the expectation at equilibrium. A burn-in period of $\geq 10N$ generations with constant population size N follows in order to ensure an equilibrium distribution of segregating sites. The initial state is defined as ancestral and the other state as derived; the derived and deleterious allele frequen-

Nature Genetics: doi:10.1038/ng.2896

cies are recorded at the end of the simulation. The code is written in C++ and is available upon request.

Demographic scenarios. We consider three demographic scenarios. The most detailed is the Out-of-Africa demographic model for African-Americans (AA) and European-Americans (EA) estimated by Tennessen et al. [2] (Supplementary Figure 1A). The model includes the Out-of-Africa split of European ancestors, changes in population size before and after the split (specifically a severe bottleneck in Europeans following the split and recent rapid growth in both Europeans and Africans) and migration between the populations after the split (see Supplementary Figure 1A for details). Finally, the model includes recent admixture between the populations, which we include in our simulations only when we compare our results to data from AAs.

While the Tennessen et al. model was parameterized in a diffusion framework, i.e., in continuous time, Wright-Fisher simulations require discrete numbers of generations and individuals. We therefore divide the times by 25 years per generation (the generation time that Tennessen et al. assume) and round the number of individuals associated with any of the parameters (e.g., growth) to the nearest integer. We implement migration by sampling alleles from the local population with probability $1 - m$ and from the other population with probability m each generation.

We also study two simpler demographic scenarios. To understand the effects of recent explosive growth of human populations, we use a simple model of exponential growth with parameters matching those of the African population in the Tennessen et al. model (see Supplementary Figure 1B for details). For the purpose of analysis, this scenario is sometimes extended by adding a period with constant population size after growth ends. Similarly, to investigate the effects of the bottleneck in Europeans at the Out-of-Africa split, we consider a simple model of a bottleneck with parameters matching those of the European bottleneck in the Tennessen et al. model (see Supplementary Figure 1C for details). Here, we sometimes extend the

Nature Genetics: doi:10.1038/ng.2896

period after the reduction in population size to study longer-term equilibration to reduced population sizes.

Validating the simulation. We used two approaches to check the validity of the simulations. For a constant population size, we compared the frequency spectra from simulations with those expected under the diffusion approximation (cf. [3]) for the neutral case as well as for several semi-dominant and recessive selection coefficients (Supplementary Figure 6). We note that obtaining similar frequency spectra implies that simpler summaries, such as the number of segregating sites under neutrality or the average deleterious allele frequency at mutation-selection balance, will also be similar.

For the more elaborate Out-of-Africa demographic model, we compared the minor allele frequency spectrum from neutral simulations with the spectrum observed at non-coding sites in Fu et al. [4] We cosider non-coding sites for this purpose as these are assumed to be under the least selection (Supplementary Figure 7). In their Supplementary Figure 2A, Tennessen et al. find a close agreement between the observed spectra and a diffusion approximation under their demographic model. We find close agreement of our neutral simulations to data from both AAs and EAs and the slight differences that we do find are similar to those in their Supplementary Figure 2A [2].

Sensitivity to mutation rate. Unless noted otherwise, we follow Tennessen et al. [2] in using a mutation rate of $u = 2.36 \cdot 10^{-8}$ per bp per generation. Given that recent estimates suggest a lower mutation rate (e.g. Kong et al. [5], Sun et al. [6]), we examine here the sensitivity of our simulation results to this assumption. We find the derived allele frequency spectrum to be extremely robust, remaining essentially unchanged when we double or halve the mutation rate (Supplementary Figure 8A). As expected, the number of segregating sites and the number of sites fixed for the derived allele increase (linearly) with the mutation rate (Supplementary Figure 8B). The increase in the number of sites fixed for the derived allele follows from the in-

Nature Genetics: doi:10.1038/ng.2896

creased rate of fixation in the burn in period (akin to fixations that occur between the ancestor of humans and chimpanzees and the Out-of-Africa split). Thus, assuming a different mutation rate will affect some of our quantitative results. Notably, if the mutation rate in humans is indeed lower than the one we use, as recent estimates suggest, the proportion of segregating sites would be lower, resulting in an even smaller effect of recent demographic history on load than our analysis suggests (see section 2). Our qualitative finding of a negligible effect on load is unchanged. Moreover, our results concerning the effects of recent demography on genetic architecture derive from the frequency spectrum and therefore are unaffected.

2 The effects of demography on load

We assume that fitness is multiplicative across sites and that selected sites are at Linkage Equilibrium (LE). The absolute fitness of individual i can then be written as

$$W_i = \prod_{j=1}^{M} w_{i,j},$$

where the product is taken over the M sites contributing to fitness and $w_{i,j}$ is the contribution of site j, which depends on the genotype of the individual and on the selection and dominance coefficients at that site. Given LE, the contributions of sites to the expected fitness in the population are independent and therefore

$$E(W_i) = \prod_{j=1}^{M} E(w_{i,j}) \approx \exp(-\sum_{j=1}^{M}(2h_j s_j p_j q_j + s_j q_j^2)),$$

where p_j and q_j are the frequencies of the normal and deleterious alleles at site j. We note that the approximation applies for strong selection because the frequency q_j is small, as well as for weak selection because then the selection coefficient is small. Finally, taking an expectation over evolutionary realizations (which is equivalent to an expectation over many sites with the same parameters in a single realization)

Nature Genetics: doi:10.1038/ng.2896

yields

$$E(W) \approx \exp(-\sum_{j=1}^{M}(2h_j s_j E(p_j q_j) + s_j E(q_j^2))). \quad (1)$$

The latter expression relates the population dynamics at a site with the overall reduction in fitness.

Genetic load is defined as the relative reduction in average fitness caused by deleterious alleles, calculated as

$$L = \frac{W_{max} - \bar{W}}{W_{max}},$$

where W_{max} is the fitness of an individual without deleterious alleles and $\bar{W}$ is the average fitness [1]. Denoting the terms associated with a single site in Equation 1 by

$$l(h,s) \equiv 2hsE(pq) + sE(q^2) = s(2hE(q) + (1-2h)E(q^2)), \quad (2)$$

the fitness function can be rewritten as

$$E(W) \approx \exp(-\sum_{j=1}^{M} l(h_j, s_j)).$$

This form emphasizes that the reduction in fitness caused by a single site generally depends on the first two moments of the deleterious allele frequency. Specifically, in the semi-dominant model, it depends only on the first moment

$$l(\frac{1}{2}, s) = sE(q),$$

and in the recessive model it depends only on the second

$$l(0, s) = sE(q^2).$$

Moreover, this form shows that $l(h,s)$ provides a natural additive measure for the expected reduction in fitness caused by a site.

Throughout the manuscript we therefore use $l(h,s)$ as our measure for the contribution of a site to load. For a model with a single site, it coincides with the definition

of load, as $E(L) = l(h, s)$. For more than one site,

$$E(L) \approx 1 - \exp(-\sum_{j=1}^{M} l(h_j, s_j)).$$

Given that in our model, the load from all sites is a simple function of the sum of $l(h, s)$ across sites, for brevity, we refer to $l(h, s)$ as load.

With a constant population size, the load exhibits three standard dynamic regimes depending on the scaled selection coefficient (Supplementary Figure 9): (i) An effectively neutral regime, in which $\alpha = 2Ns \ll 1$ and the effects of selection are negligible compared to drift; (ii) a weak selection (or nearly neutral) regime, in which $\alpha = 2Ns \approx 1$ and the effects of selection and drift are comparable; (iii) a strong selection regime, in which $\alpha = 2Ns \gg 1$ and selection dominates over drift.

In what follows our analysis is divided according to these three regimes. When the population size changes, the boundaries between regimes are affected. Moreover, the rate at which the equilibrium for a new population size is attained depends on the summary of the data considered. We consider summaries for segregating sites, e.g., the proportion of segregating sites and the allele frequency at these sites, and summaries for fixed sites, e.g., the proportion of sites fixed for the deleterious allele (which we call fixed state). Specifically, we are interested in the effects of demography on the contribution of segregating and fixed sites to load, which we refer to as fixed and segregating load, and in their sum, which we refer to as total load. We consider the behavior of these statistics for the two simple demographic models, which together allow us to understand all qualitative behaviors exhibited under the more detailed Tennessen et al. model (10). For these demographic models, we primarily consider two modes of inheritance (semi-dominant and recessive).

To simplify our theoretical analysis, we make several reasonable assumptions about the parameters of the model. For brevity, we focus on the case with symmetric mutation ($u = v$) and, because we are considering human populations, we assume that the population mutation rate per site is small, i.e., that $\beta = 2Nu \ll 1$. We also

Nature Genetics: doi:10.1038/ng.2896

assume that the selection coefficient is small, i.e., $s \ll 1$. A summary of our analyses are presented in Supplementary Figure 10 and Table 1. A detailed description of the behavior in each regime follows.

2.1 The effectively neutral regime

When selection is negligible compared to drift, the behavior of deleterious alleles is well approximated by that of neutral alleles. As the properties of neutral alleles (e.g., the proportion of segregating sites and frequency spectrum) in models with constant and varying population sizes have been studied exhaustively (e.g., [9, 10, 11]), here we focus only on the implications concerning load.

First, we consider how load depends on the selection coefficient at equilibrium for a constant population size. If deleterious alleles behave like neutral ones, the first two moments of the deleterious allele frequency distribution do not depend on the selection coefficient and therefore the load is proportional to the selection coefficient (see Eq. 2). This explains the linear relationship between selection coefficient and load shown in Supplementary Figure 9.

At equilibrium, load depends negligibly on the population size. Using the diffusion approximation for the stationary deleterious allele frequency distribution [3], the expansion of the load to first order in α and β yields

$$l(h,s) = \frac{s}{2}(1 - \frac{1}{2}\alpha - 2(1-2h)\beta).$$

Thus, as long as $\beta \ll 1$ and $\alpha \ll 1$, the load is well approximated by $s/2$ regardless of the population size and dominance coefficient (hence the similarity in load for the semi-dominant and recessive cases in Supplementary Figure 9). Intuitively, this follows from the fact that the great majority of sites are fixed, and because selection is negligible, half of them are fixed for the deleterious allele ($\frac{u}{u+v}$ for asymmetric mutation).

Nature Genetics: doi:10.1038/ng.2896

The same reasoning implies that changes in population size will have a negligible effect on the total load in this regime (Supplementary Figure 11). While changes in population size affect the proportion of segregating sites and thus their contribution to load, so long as the population mutation rate remains negligibly small ($\beta \ll 1$), the segregating load will remain negligible compared to the fixed load. In the bottleneck model, the proportion of segregating sites decreases to a new equilibrium after the reduction in population size (Supplementary Figure 11A). This explains the decrease in segregating load, which is balanced by an increase in fixed load (Supplementary Figure 10). By the same token, in the growth model, the segregating load increases but is balanced by a decrease in fixed load, resulting in a negligible change to the total load (Supplementary Figure 10 and Supplementary Figure 11B). In this case, however, segregating sites are still far from their new equilibrium at present (see the next section).

2.2 The weak selection regime

In the weakly selected regime, selection and drift have comparable effects on the dynamics of deleterious alleles. As a result, at equilibrium, even moderate differences in population size can affect the balance between selection and drift. Changes in population size also shift the balance, and are followed by transient changes at fixed and segregating sites until a new equilibrium is attained. To understand these effects, we consider the behavior at equilibrium and the rate at which it is approached. For this purpose, it is helpful to use the low mutation rate (LMR) approximation in which mutant alleles at a segregating site have a single origin; in other words, we ignore mutations that arise during the sojourn of a mutant allele from the time it arises on a background fixed for the other allele to the time it reaches fixation or loss in the population.

The effect of population size on the proportion of sites fixed for the normal and deleterious alleles. At equilibrium, the rate at which deleterious alleles arise

Nature Genetics: doi:10.1038/ng.2896

and fix is equal to the rate at which normal alleles arise and fix. This balance can be written as

$$2Nup\pi(-2Ns, h, \frac{1}{2N}) = 2Nvq\pi(2Ns, 1-h, \frac{1}{2N}),$$

where π denotes the fixation probability, which depends on the scaled selection and dominance coefficients and on the initial frequency [12] (because $s \ll 1$, we ignore second order terms in s). For $s \ll 1$ and any dominance coefficient, this yields

$$\frac{q}{p} = \frac{u}{v}\frac{\pi(-2Ns, h, \frac{1}{2N})}{\pi(2Ns, 1-h, \frac{1}{2N})} \approx \frac{u}{v}e^{-2Ns}.$$

Namely, at equilibrium, the proportion of fixed deleterious sites declines exponentially with the scaled selection coefficient $\alpha = 2Ns$ (Supplementary Figure 12A). Thus, for a given selection coefficient s, the population size has a dramatic effect on the proportion of sites fixed for the deleterious allele, declining from the neutral, mutation-driven, proportions for $s \ll \frac{1}{2N}$ to approximately 0 for $s \gg \frac{1}{2N}$.

Importantly, however, when the population size changes, the new equilibrium proportion may be attained very slowly. The fractions, $p(t)$ and $q(t)$, of sites fixed for the normal and deleterious alleles t generations after a change in population size (assuming $p(t) + q(t) = 1$) are well approximated by the model

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -2N_a u\pi(-2N_a s, h, \frac{1}{2N_a}) & 2N_a v\pi(2N_a s, 1-h, \frac{1}{2N_a}) \\ 2N_a u\pi(-2N_a s, h, \frac{1}{2N_a}) & -2N_a v\pi(2N_a s, 1-h, \frac{1}{2N_a}) \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix},$$

where N_a is the population size after the change, and fixation times (on the order of $4N_a$ generations) are neglected. An additional contribution from sites that were segregating before the change is considered below. In this approximation, the change in the fraction of sites fixed for the deleterious alleles is

$$q(t) = q_a^{eq}\left(1 - e^{-\frac{t}{\tau}}\right) + q_b^{eq}e^{-\frac{t}{\tau}},$$

where q_b^{eq} and q_a^{eq} are the equilibrium fractions corresponding to the population sizes before and after the change, and

$$\tau = \left[2N_a\left(u\pi(-2N_a s, h, \frac{1}{2N_a}) + v\pi(2N_a s, 1-h, \frac{1}{2N_a})\right)\right]^{-1}$$

Nature Genetics: doi:10.1038/ng.2896

is the timescale of the exponential approach to the new equilibrium. For the semi-dominant case and $s \ll 1$, this time scale is well approximated by

$$\tau \approx \left[u\frac{\alpha}{e^{\alpha}-1} + v\frac{\alpha}{1-e^{-\alpha}} \right]^{-1},$$

demonstrating that it is mutation-limited. This is also true for other dominance coefficients. In other words, following an instantaneous change in population size, the proportion of sites fixed for the deleterious allele will change extremely slowly, at a rate that is inversely proportional to the mutation rate (Supplementary Figure 12B and C).

Because the equilibrium is reached slowly, recent demographic changes in humans should have had little effect on the proportion of sites fixed for the deleterious alleles and hence on the fixed load. The bottleneck at the Out-of-Africa split is estimated to have reduced the population size from $\sim 14,000$ to $1,800$ approximately 2000 generations ago [2]. Once a new equilibrium is reached, there will be a substantial increase in the proportion of fixed deleterious alleles; for example, for a semi-dominant deleterious allele with selection coefficient of $s = 10^{-4}$, it would increase it from 0.05 to 0.4. Yet the change over 2000 generations is minimal, increasing this proportion only by $3 \cdot 10^{-5}$. The estimated 200 generations since the onset of rapid growth in humans is similarly much too short a time period for any measurable effect on the fixed load (which in this case would decrease over large time periods).

The effects of population size on segregating sites. First we consider how the equilibrium properties of segregating sites depend on population size in models with constant population size (Supplementary Figure 13). The deleterious allele frequency at segregating sites decreases with increasing population size, because the efficacy of selection is greater in larger populations (Supplementary Figure 13A). In turn, the proportion of segregating sites increases with population size due to the (linear) increase in the number of mutations that enter the population every generation (Supplementary Figure 13B). This is true not only for the population as a whole but also for subsamples from it of any size (Supplementary Figure 13C).

Nature Genetics: doi:10.1038/ng.2896

Finally, the deleterious allele frequency and proportion of segregating sites decrease with increasing dominance coefficient, as stronger selection in heterozygotes results in stronger selection on deleterious mutations (regardless of their frequency) and thus in a shorter sojourn through the population. Thus, in larger populations or if the dominance coefficient is greater, we expect a greater proportion of segregating sites with deleterious alleles at lower frequency.

The total load decreases monotonically when the population size increases (as can be shown using the stationary distribution based on the diffusion approximation [3], for example). This is not true of the segregating load, because the increase in the mutational input can have a greater effect than the increase in the efficacy of selection (Supplementary Figure 13D). Indeed, for selection coefficients closer to neutrality, the increase in mutational input (and the proportion of segregating sites) dominates, causing the segregating load to increase with population size (akin to the behavior in the effectively neutral regime). In contrast, for selection coefficients closer to the strong selection regime, the increase in the efficacy of selection dominates, leading to a reduction in segregating load (akin to the stronger selection regime; see section 2.3).

Next we consider the effects of a change in population size. We begin by noting that, for a given population size, the expected sojourn time of deleterious and beneficial mutations that reach fixation is shorter than that for a neutral mutation and is thus on the order of $4N$ generations or less [3]. This implies that on the order of $4N_a$ generations after a change in population size, most of the *old mutations* (i.e., those that segregated before the population size changed) have been absorbed (either due to loss or fixation), and replenished by *new mutations* (that arose and spread through the population at its new size). When this turnover process is complete, new segregating sites approach their equilibrium proportions (given a background of fixed sites).

In the bottleneck model, the reduction in the efficacy of selection causes an increase

Nature Genetics: doi:10.1038/ng.2896

in total load, where the behavior of the components of load can be understood as follows (Supplementary Figure 2). Focusing first on the contribution of old mutations to the fixed load: When old mutations are absorbed, the reduction in the efficacy of selection leads more deleterious alleles to fix than would have had the population size remained constant (at the larger size), eventually resulting in an increase in fixed load. The increase can be approximated by

$$\Delta(s,h,u,N_b,N_a) = \int_0^1 \left(\pi(-2N_a s,h,x) - \pi(-2N_b s,h,x)\right) f(x;h,2N_b s,2N_b u)dx,$$

where $f(x;h,2N_b s,2N_b u)$ is the stationary distribution before the change in population size [3]. The increase is maximized for selection coefficients at which the change in population size leads selection to transit from strong to weak, and is negligible outside this range (Supplementary Figure 2A; explaining why it is more pronounced in Supplementary Figure 2C and D than in E and F, correspondingly). The increase in deleterious fixations and load is then followed by a long-term, slower increase in the fixed load due to new mutations (Supplementary Figure 2C-F). In the parameter regime where the fixation of old mutations makes a substantial contribution to load, there is also a transient increase in segregating load before the mutations fix (in Supplementary Figure 2C for example). These effects are more pronounced in the recessive case, because of the greater frequency and proportion of segregating sites. Now focusing on the segregating load (Supplementary Figure 2B): when segregating sites attain equilibrium, the reduction in population size causes a decrease in segregating load for lower selection coefficients (Supplementary Figure 2C and D) and an increase for higher selection coefficients (Supplementary Figure 2E and F). Thus, for higher selection coefficients in the weak selection range, both old and new mutations contribute to the transient increase in segregating load observed in Supplementary Figure 10. For the lower selection coefficients in this range, the segregating load decreases both in the short and long term but the fixation of old mutations still results in an overall increase to the total load (Supplementary Figure 10). Importantly, however, on the timescale estimated for the bottleneck at the Out-of-Africa split

Nature Genetics: doi:10.1038/ng.2896

(vertical line in Supplementary Figure 2), these effects amount to a tiny increase in total load (Supplementary Figure 10).

What about in the case of growth? Human population growth is thought to have started a couple hundred of generations ago, ending with an effective population size in the hundreds of thousands and starting from a size that was thirty-fold smaller [2]. Given the estimated growth parameters, there was insufficient time for the deleterious alleles that segregated before the onset of growth to change their frequencies substantially. Indeed even with the increase in the efficacy of selection as the population size increases, in this regime, selection is too weak to have caused a substantial change in allele frequency over hundreds of generations (although it could have caused the absorption of very rare or very high frequency alleles). After growth, the resulting frequency spectrum of deleterious alleles thus reflects a superposition of the spectrum of segregating sites before growth and of the spectrum at the large number of sites in which mutations were introduced after the onset of growth (Supplementary Figure 14). The many new mutations remain at low frequencies. Because of an increase in the proportion of segregating sites, the segregating load increases at the expense of fixed load, but with negligible effects on the total load, given both the low frequency of new mutations as well as the opposing contributions of normal and deleterious mutations (Supplementary Figure 10).

2.3 The strong selection regime

In this regime, purifying selection is sufficiently strong to prevent deleterious alleles from reaching high frequencies, let alone fixation. It follows that there is only segregating load. If we assume that the deleterious allele frequency is small and that the dominance coefficient is sufficiently large, then the load is well approximated by

$$l(h, s) \approx 2hsE(q).$$

Nature Genetics: doi:10.1038/ng.2896

Stated another way, when selection against heterozygotes is sufficiently strong, then deleterious homozygotes would be too rare to affect load. Under these assumptions, the diffusion approximation at equilibrium with a constant population size [3] yields

$$E(q) \approx \frac{u}{hs},$$

implying that the load is well approximated by

$$l(h, s) \approx 2u.$$

We refer to the cases where these conditions are met as quasi-dominant.

In the recessive case, the load depends on the second moment of deleterious allele frequency. Assuming once again that the deleterious allele frequency is small, the diffusion approximation at equilibrium with a constant population size [3] yields

$$E(q^2) \approx \frac{u}{s},$$

implying that the load is well approximated by

$$l(0, s) \approx u.$$

The expressions for load in both cases are identical to the classic ones for mutation-selection balance, which are derived assuming an infinite population size [12]. They imply that at equilibrium, the load depends neither on the selection coefficient (explaining the plateaus in Supplementary Figure 9) nor on the population size.

When the dominance coefficient is sufficiently small, however, the load does depend on population size (Supplementary Figure 15). This will be the case when selection against heterozygotes is weak, i.e. when $2Nhs \gg 1$ does not hold, as then both moments of deleterious allele frequency make comparable contributions to load. Holding the selection coefficient and population size constant, in this range of dominance coefficients, the load varies continuously with h between u and $2u$ (Supplementary

Nature Genetics: doi:10.1038/ng.2896

Figure 15A). In turn, holding $h \ll 1$ and $N \gg 1$ constant, increasing s also leads the load to vary from u to $2u$ (Supplementary Figure 15B).

Next, we consider the effect of changes in population size, for the quasi-dominant and then the recessive case. We show that in the quasi-dominant case, the load remains constant and is well approximated by the classic derivations for mutation-selection balance. In the recessive case, the load exhibits transient changes before it returns to its equilibrium level.

The quasi-dominant case

In the quasi-dominant case, we can assume deleterious alleles are sufficiently rare that selection against deleterious homozygotes can be ignored and selection has negligible effects on average fitness. Under these conditions, we can approximate the trajectory of a deleterious allele using a branching process (cf. [13]), in which the number of copies that a given deleterious allele gives rise to in the next generation follows a distribution that is independent on the frequency of deleterious alleles in the population.

Consider a single deleterious allele that was introduced by mutation at time $t = 0$ and denote by $Z(t)$ the number of deleterious alleles that it gives rise to at generation t. The number of mutant alleles in the next generation can then be expressed as

$$Z(t+1) = \sum_{i=1}^{Z(t)} X_i(t),$$

where $X_i(t)$ denotes the number of offspring of the ith allele at time t and $i = 1, \ldots, Z(t)$. We denote the expected number of offspring of a single allele by λ, i.e., $E(X_i(t)) = \lambda$; if we ignore mutations back to the beneficial allele then $\lambda = 1 - hs$ and if we include them then $\lambda = 1 - hs - v$. The expected number of alleles in the next generation is then

$$E(Z(t+1)) = E(\sum_{i=1}^{Z(t)} X_i(t)) = \sum_{j=1}^{\infty} Pr(Z(t) = j) j E(X_i(t)) = E(Z(t))\lambda, \qquad (3)$$

Nature Genetics: doi:10.1038/ng.2896

or

$$E(Z(t)) = \lambda^{\tau}. \tag{4}$$

Now consider the expected number of deleterious alleles at mutation-selection balance. For this purpose, we measure time backwards from the present. We denote by $Y_\tau(\tau)$ the number of mutations introduced τ generations ago and by $Y_\tau(t)$ the number of alleles that they give rise to at time t. The number of deleterious alleles at the present can then be expressed as the sum of contributions from all the mutations in the past, i.e. $\sum_{\tau=1}^{\infty} Y_\tau(0)$, where, from Equation 4,

$$E(Y_\tau(0)) = Y_\tau(\tau)\lambda^{\tau}.$$

In turn, the expected number of new mutations in a given generation is well approximated by

$$E(Y_\tau(\tau)) = 2Nu.$$

It follows that the expected deleterious allele frequency is

$$E(q) = \frac{1}{2N} E(\sum_{\tau=1}^{\infty} Y_\tau(0)) = \frac{1}{2N} \sum_{\tau=1}^{\infty} E(Y_\tau(\tau))\lambda^{\tau} = \frac{u}{hs},$$

and thus the expected contribution to load is $2u$ - well-known results for mutation-selection balance.

Next, we consider a changing population size. We denote by $N(t)$ the population size t generations in the past and by $a(t) = \frac{N(t-1)}{N(t)}$ the proportional change in one generation. Now the expected number of new mutations introduced at a given time is proportional to the population size

$$E(Y_\tau(\tau)) = 2N(\tau)u,$$

but the fraction of new mutations in the population remains constant (u). Similarly, the expected number of alleles in the next generation is affected by changes in population size

$$E(Y_\tau(t-1)) = \lambda a(t) E(Y_\tau(t)),$$

Nature Genetics: doi:10.1038/ng.2896

but their fraction is not, because their increase in number is precisely offset by the increase in population size

$$E(\frac{Y_\tau(t-1)}{2N(t-1)}) = \lambda a(t)\frac{N(t)}{N(t-1)}E(\frac{Y_\tau(t)}{2N(t)}) = \lambda E(\frac{Y_\tau(t)}{2N(t)}).$$

It follows that the proportional contribution of alleles to the present is the same as that in a constant population size:

$$E(\frac{Y_\tau(0)}{2N(0)}) = u\lambda^\tau,$$

leaving the deleterious allele frequency and the load at the present unchanged (at $\frac{u}{hs}$ and $2u$). In other words, the expected frequency of deleterious alleles and therefore the load follow the same deterministic dynamic as they do in a population of constant size, because when the population size changes, the increase (decrease) in the copy number is precisely offset by the increase (decrease) in population size.

We note that incorporating reverse mutation and migration will not change this conclusion. Reverse mutation would reduce λ, while introducing migration would be similar to both decreasing λ (due to migration of deleterious alleles out of the population) and increasing the mutational input (due to migration of deleterious mutations into the population).

Our results clarify how the expected deleterious allele frequency and proportion of segregating sites at equilibrium depend on population size. When the population mutation rate is sufficiently low, a site switches intermittently between having no deleterious alleles and having a single mutation (by origin) in the population (Supplementary Figure 16A). Under these conditions, in a larger population size, the mutational input is larger and thus the proportion of time that a site is segregating increases (Supplementary Figure 16B). Because the trajectory of a mutation in terms of numbers of copies does not depend on the population size, the frequency of the mutation is proportional to $1/N$, so the expected frequency of deleterious alleles at segregating sites scales with $1/N$ (Supplementary Figure 16C). In turn,

Nature Genetics: doi:10.1038/ng.2896

when the population mutation rate is sufficiently high, deleterious alleles are almost always present and often have several mutational origins. Under these conditions, the proportion of segregating sites approaches 1 (Supplementary Figure 16B). Given that the expected frequency at segregating sites is $x = \frac{q}{S_{2N}}$, it follows that the allele frequency asymptotes to $q = \frac{u}{hs}$ (Supplementary Figure 16C). In turn, the variance in allele frequency decreases with population size and asymptotes to 0 in the infinite population size limit.

After a change in population size, a new equilibrium is attained much more rapidly in the strong selection regime because of the rapid turnover of deleterious alleles (see Supplementary Figure 17). However, load is unaffected.

Thinking in terms of the branching process helps us to evaluate previous conjectures about the possible effects of human growth on deleterious alleles. For example, Keinan and Clark [15] suggest that "Some degree of genetic risk for complex disease may be due to this recent rapid expansion of rare variants in the human population". It is indeed the case that the expected copy number of deleterious alleles should be greater under exponential growth; specifically, for a population growing at a geometric rate γ per generation, the copy number will change at a geometric rate of $\lambda + \gamma$ per generation, which will result in an increase if $\lambda + \gamma > 1$. Moreover, population growth increases the sojourn time of a deleterious mutation and, when $\lambda + \gamma > 1$, there is a finite probability it would never go extinct [16]. Importantly, however, the expected *frequency* of quasi-dominant deleterious alleles remains constant, so human population growth has no effect on load.

The recessive case

In this case, the load at equilibrium is again insensitive to population size, but the underlying reasons are quite different than in the quasi-dominant case. In the recessive model, a deleterious allele behaves neutrally while at low frequencies. As a result, its sojourn time (i.e., the expected time that it spends at frequency x) is well approximated by that of a neutral allele (Supplementary Figure 18B). When

Nature Genetics: doi:10.1038/ng.2896

the frequency x reaches $2Nsx^2 \approx 1$, selection on homozygotes for the deleterious alleles kicks in, and the allele should spend little time above this frequency. In the low mutation rate (LMR) approximation, we can therefore approximate the sojourn time of a recessive deleterious allele as

$$\tau(x) \approx \begin{cases} \frac{2(2N-1)}{1-x} & \text{if } 0 \leq x \leq \frac{1}{2N} \\ \frac{2}{x} & \text{if } \frac{1}{2N} \leq x < \frac{1}{\sqrt{2Ns}} \\ 0 & \text{if } \frac{1}{\sqrt{2Ns}} \leq x < 1 \end{cases},$$

where the expressions for $x < 1/\sqrt{2Ns}$ are the sojourn times (in generations) for a neutral allele (Fig 18B). In this approximation, the expected contribution of a deleterious mutation to load is then

$$s\int_0^1 x^2\tau(x)\mathrm{d}x \approx s\int_0^{\frac{1}{\sqrt{2Ns}}} x^2\frac{2}{x}\mathrm{d}x = \frac{1}{2N},$$

and, given that the expected input of new mutations per generation is $2Nu$, the overall expected load is

$$l(0,s) \approx 2Nu\frac{1}{2N} = u.$$

In other words, (in the low mutation limit) for a given population size N, a recessive allele behaves neutrally up to a frequency of $N^{-\frac{1}{2}}$, resulting in an expected contribution to load that is proportional to N^{-1}. In turn, the mutational input is proportional to N, so they exactly offset.

This back of the envelope approximation also provides an intuitive explanation for the way in which the properties of segregating sites at equilibrium depend on population size (Fig 18). First, we consider the proportion of segregating sites (Fig 18A). When the population size is sufficiently small for the LMR approximation to apply, the proportion of segregating sites can be approximated by the ratio of the sojourn time of a single mutant through the population to the time between appearances of mutations, namely:

$$S_{2N} \approx \frac{\int_0^1 \tau(x)\mathrm{d}x}{\frac{1}{2Nu}} \approx 2Nu(ln(2N/s)+2).$$

Nature Genetics: doi:10.1038/ng.2896

In a larger population size and hence with a larger mutational input, mutations of different origin will overlap, resulting in a slower increase in the proportion of segregating sites with population size. When the mutational input becomes sufficiently large, this proportion asymptotes to 1. Next, we consider the frequency of deleterious alleles. In the LMR approximation, the frequency spectrum of segregating sites can be approximated using the neutral sojourn times up to the threshold frequency $\frac{1}{\sqrt{2Ns}}$ (Fig 18B), yielding an average frequency of $E(x) \approx \frac{\frac{2}{\sqrt{2Ns}}}{2+ln\frac{2N}{s}}$. As the population size increases, such that mutations of different origins overlap, the decrease in average frequency becomes slower and asymptotes to $E(x) = E(q) = \sqrt{u/s}$ (Fig 18C). Lastly, the turnover time of segregating sites for a given population size N is on the order of $2\sqrt{\frac{2N}{s}}$. As it was for other regimes, this is the time scale for the process of equilibration following a change in population size.

We now consider the implications for the bottleneck and growth models. In the bottleneck model, after the reduction in population size, there is an increase in load followed by a decrease back to the equilibrium level (Supplementary Figure 19A). The transient increase in load (blue arrow in Supplementary Figure 19A) is dominated by the contribution of mutations that segregated before the decrease in population size. The proportion of sites that segregated before was greater and their frequencies lower than after the population size reduction, and while these segregating mutations are gradually absorbed, some of them will drift to higher frequencies, generating a transient surge in load (Supplementary Figure 19B). In turn, the newly introduced mutations have yet to reach equilibrium frequencies and, given that the contribution of the lower frequencies to load is much smaller, they contribute negligibly. In the Tennessen et al. model, the time that elapsed since the bottleneck is longer and the segregating sites are therefore closer to the new equilibrium (green arrow in Supplementary Figure 19A). Correspondingly, the relative contribution of new mutations is greater and their frequency distribution is closer to equilibrium with the new population size, and yet some contribution from the older mutations remains (Supplementary Figure 19C). These considerations also explain why load exceeds above

Nature Genetics: doi:10.1038/ng.2896

equilibrium levels in the strong selection regime in Supplementary Figure 10.

In the growth scenario, we see the opposite transient effect: the load is reduced before recovering to its equilibrium level (Supplementary Figure 19D). After the growth period, the number of segregating sites is greatly increased, but the new mutations have had little time to drift to higher frequency. As a result, new mutations segregate at very low frequencies and contribute negligibly to load (Supplementary Figure 19E and F). In turn, mutations that segregated before growth have decreased in frequency due to the increased efficacy of purifying selection, and so their contribution to load declines substantially (Supplementary Figure 19E and F). The result is a transient reduction in load (seen in Supplementary Figure 10 as well as in Supplementary Figure 19D).

2.4 Models with dominance coefficients other than 0 and $\frac{1}{2}$

Here we provide summaries of simulations with dominance coefficients other than 0 and 1/2 to illustrate that the same qualitative behaviors are observed. As shown in Supplementary Figure 20, all of the observed qualitative behaviors are included in our previous analysis and summarized in Table 1, with one possible exception.

The exception is in the bottleneck model in cases with dominance coefficients $h > 1/2$, where the total load is reduced for lower selection coefficients in the weak selection regime. The reason for this reduction in load is analogous to that for the increase in load that we saw in the recessive case in the same selection regime. For dominance coefficients greater than half, the extinction of low frequency deleterious alleles that segregated before the reduction in population size decreases load more than the fixation of high frequency deleterious alleles increases it. The opposite is true for dominance coefficients smaller than half.

Nature Genetics: doi:10.1038/ng.2896

3 Data analysis and interpretation

We used data from Fu et al. (2012) [4] and from the 1000 Genomes Project [8]. Allele frequency estimates from Fu *et al.* are available from the NHLBI GO Exome Variant Server (`http://evs.gs.washington.edu/EVS/`). These provide estimates of the derived allele frequencies at exonic SNVs in European- and African-Americans (EA and AA). Variants with allele frequencies 0 or 1 in both EA and AAs were excluded.

The haploid sample sizes in Fu et al were EA Autosomal: 8596, EA X: 6717, AA Autosomal: 4434, AA X: 3852. Our primary analysis in the main paper (reported in Figure 3) uses the full sample sizes with the autosomal data. For the purpose of Table 2 we wished to compare means on the X and autosomes. Since mean allele frequencies of segregating sites are affected by total sample size, we implemented the following subsampling strategy to facilitate direct comparisons between X and autosomes. First, we converted the reported allele frequencies for each site back into allele counts (i.e., multiplying each reported frequency by the relevant haploid sample size). Next, we randomly subsampled the autosomal EA and AA variants and the X chromosome EA variant allele frequencies down to a sample size of 3852 chromosomes each, in order to match the haploid sample size for the African-American X chromosome. Subsampling was done without replacement, using the hypergeometric sampling function in R. After sub-sampling, variants whose allele frequencies were both either 0 or 1 were once again dropped. Two-sided t-tests were used to test for allele frequency differences between groups.

1000 Genomes Project vcf files (Phase 1 Version 3) were downloaded from the official 1000 Genomes public server (`ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/`). YRI and CEU individuals with (at least) exome sequencing coverage were extracted from the original .vcf files (88 YRI individuals and 81 CEU individuals). 7 YRI individuals, chosen at random, were removed to match sample sizes between YRI and CEU. Variants that were fixed for either allele in both populations were removed.

Nature Genetics: doi:10.1038/ng.2896

Any variant that was not an SNV or did not contain ancestral allele information was also dropped.

A natural measure for comparing the difference in load between two populations is to count the mean number of derived alleles per individual at SNVs segregating within the joint sample. Note that it is essential in these calculations to define SNVs using the joint sample, otherwise sites that are fixed for the derived allele in Population A but not in Population B would lead to the erroneous conclusion that there are more derived alleles in B than in A.

For our analysis, we found that it is convenient to work with the mean derived allele frequency within each functional class. This quantity allows us to compare frequencies directly between classes, and is also conveniently computed from the Fu et al frequency data. These two measures (mean derived frequency and number of derived alleles per individual) are proportional to one another and hence must yield identical conclusions about the relative load in different populations (for a given functional class: DAF multiplied by twice the number of SNVs yields the number of derived alleles per individual, assuming that missing data have been filled in appropriately). Notice also that we are dealing with mean numbers of alleles, and so these measures are unaffected by deviations from HWE or LE which affect the variance in numbers of derived alleles per individual but not the means.

Of course the number of derived alleles is not equivalent to the number of deleterious alleles, as some variants may be neutral; additionally for weakly selected sites there is a small probability at each site that the ancestral allele is deleterious. Nonetheless, the load is expected to be monotonically increasing with the number of derived alleles. As shown in Supplementary Figure 3, we predict that at semidominant sites there should be essentially no difference in mean derived frequency between AAs and EAs, regardless of selection coefficient. At recessive sites we would expect a small increase in mean frequency in AAs at moderately and strongly selected sites. The fact that we do not observe any significant difference in allele frequencies at "probably damaging"

Nature Genetics: doi:10.1038/ng.2896

sites argues that the majority of these sites are at least partially dominant.

Mean derived allele frequencies were calculated for both populations at autosomal noncoding, synonymous, and nonsynonymous sites, as well as autosomal nonsynonymous variants belonging to the different functional categories. Standard errors for each category were estimated using the standard deviation in DAF across sites, divided by the square root of the number of sites in that category. For individual-level analyses, we computed the SD in mean number of variants per individual by bootstrapping across sites. The bootstrap analysis accounts for the evolutionary sampling variance in allele frequencies.

The ANNOVAR suite of scripts [21] was used to obtain functional predictions for each SNP from each of four prediction methods: PolyPhen2 [22], SIFT [23], LRT [25] and MutationTaster [24]. Default program settings were used in each case. The functional designations for each program are as follows: PolyPhen2: D (Probably Damaging), P (Possibly Damaging), B (Benign). SIFT: D (Damaging), T (Tolerant), LRT: D (Deleterious), N (Neutral) and U (Unknown). MutationTaster: A (Disease Causing Automatic), D (Disease Causing), P (Polymorphism Automatic) and N (Polymorphism). Coding versus non-coding and synonymous versus non-synonymous designations were also determined using ANNOVAR. (Note that we also tested the SeattleSeq annotations, and found that the overall numbers were similar (though not identical) to those obtained from ANNOVAR; as with ANNOVAR we found no evidence for a difference in DAF between populations.)

We observed that a strong reference bias exists at sites for which the genome reference sequence carries the derived reference allele. This bias has also been observed by David Reich and Shamil Sunyaev (personal communication). All four functional prediction programs designate a very high proportion of these sites as being likely nonfunctional or benign, even when the reference allele is rare in the population overall. When we condition on the overall population frequency at these sites, we find that a given site is much more likely to be classified as a probably damaging

site if the reference genome carries the ancestral allele than if it carries the derived allele (Supplementary Figure 4).

To deal with this bias, we treated the functional designations at sites where the reference allele is derived as unreliable. As an alternative, we binned all SNVs into a series of allele frequency bins (i.e., the bins shown in Supplementary Figure 4). We assumed that when we condition on the population allele frequency in a very large sample (i.e., the Fu et al sample) that the identity of the genome reference allele carries essentially no further information about the likely functional properties of a variant. Thus, within a bin, the fraction of derived-reference SNVs that fall into each functional category can be predicted from the fraction of ancestral-reference SNVs in that functional category. Thus for example, if 20% of the ancestral-reference SNVs in a given bin have functional category X, then we assume that each of the derived-reference SNVs in that bin has a 20% probability of also being in functional category X. The mean frequency of all SNVs in category X is estimated by summing across all ancestral-reference SNVs in category X plus a sum of contributions from all derived-reference SNVs, weighted by the estimated probabilities that each is in X. As shown in Table 3, the bias correction makes a substantial difference to the data analysis. Prior to applying the bias correction, the mean frequency in AAs is substantially higher than in EAs (presumably because more than half of the reference genome sequence is of non-African origin (Supplement of [14], p145)), but the bias correction makes the two frequencies virtually identical as predicted for models with dominance.

We also provide supplementary results in which we made use of a new unpublished version of PolyPhen's PSIC scores that are calculated in a human-independent (i.e., unbiased) manner. (Thanks to Ivan Adzhubey and Shamil Sunyaev for access to these.) These produce results that are very similar to those from our bias-corrected version, in the sense of showing no difference between populations.

4 The effects of demography on the genetic architecture of disease risk

A great deal of interest focuses on understanding how recent demographic history has affected the genetic architecture of disease and specifically whether the recent explosive growth has increased the contribution of rare variants to disease risk [17, 15, 18, 2]. Here, we use the theory that we developed to elucidate some of these effects. *Note that while in what follows we refer to disease risk, it also applies to any other quantitative trait.*

4.1 A model relating allele frequencies to disease susceptibility

We first consider the relationship between selection on individual loci and disease risk. The few models for this relationship differ sharply in their assumptions. At one extreme, Pritchard [19] assumed that variants that increase disease susceptibility tend to be deleterious, but that otherwise there is no relationship between the strength of selection acting on these loci and the extent to which they increase disease susceptibility. In turn, Eyre-Walker [20] assumed a correlation between the strength of selection at a locus and its contribution to disease susceptibility. All else being equal, a stronger relationship between the disease risk and fitness implies that the variants that contribute more to disease risk are under stronger selection and, as a result, tend to be younger and rarer. It also follows that their frequency distribution would be more susceptible to the effects of recent demographic events. Here we consider models for the two extremes: one in which the effect sizes are independent on the selection coefficients and the other where the effect sizes are proportional to the selection coefficients.

To model how genetic variation relates to disease risk, we consider the L loci that contribute to disease risk and denote the genotype of individual i at these loci by $\mathbf{G_i} = (g_{i,1}, \ldots, g_{i,L})$. We assume that each of the loci is bi-allelic, with a normal (N) and

susceptible (S) alleles, and therefore denote the genotype at locus j ($j = 1, \ldots, L$) as $g_{i,j} = NN$, NS, or SS. We then assume that the probability of developing the disease (ignoring life-history details) takes the form

$$P(\mathbf{G}) = F(\sum_{j=1}^{L} \alpha_j(g_j)),$$

where F is a monotonically increasing function with continuous derivatives that takes values between 0 and 1 and that

$$\alpha_j(g) = \begin{cases} 0 & \text{if } g = NN \\ h_j a_j & \text{if } g = NS \\ a_j & \text{if } g = SS \end{cases},$$

where h_j and a_j denote the dominance coefficient and effect size of the contribution to susceptibility at locus j. Finally, we assume that the effect of each locus is small, such that we can approximate the variance in susceptibility by *the first term in a Taylor expansion, i.e.,*

$$V(P(G)) \approx [F'(\sum_{j=1}^{L} E(\alpha_j(g_j)))]^2 \sum_{j=1}^{L} V(\alpha_j(g_j)), \tag{5}$$

where the variances are taken over the population and

$$V(\alpha(g); x, a, h) = a^2 x(1-x) \left[(2h-1)^2 x^2 + (1-4h^2)x + 2h^2\right],$$

where x is the S-allele frequency.

Our model in which the effect sizes are independent on the selection coefficients (and similarly for dominance coefficients) follows directly. For simplicity we assume that the effect sizes and dominant coefficients are constant, as assuming a distribution yields similar results for all the quantities that we consider below. *The variance in disease susceptibility then follows from Eq. 5, where the a_j's and h_j's are constant across loci and the distribution of allele frequencies (the x's) is deterimined by the (independent) selection and dominance coefficients (for fitness) at these loci.*

Nature Genetics: doi:10.1038/ng.2896

Next, we consider the model in which the disease itself is the agent of selection. In other words that the fitness cost results entirely from the probability of developing the disease. Denoting the fitness of affected individuals by W_a and of unaffected by W_u, the relationship between fitness, W, and the probability of developing the disease then takes the form

$$W = PW_a + (1 - P)W_u.$$

In turn, in our model, the relationship between genotype and fitness is

$$W(\mathbf{G_i}) = \prod_{j=1}^{L} w_{i,j} \approx \exp\left(-\sum_{j=1}^{L} \alpha_j(g_{i,j})\right),$$

where

$$\alpha_j(g) = \begin{cases} 0 & \text{if } g = NN \\ h_j s_j & \text{if } g = ND \\ s_j & \text{if } g = DD \end{cases},$$

and we assume that $s_j \ll 1$ and therefore use an exponential approximation. Equating our two expressions for fitness leads to the following model for the relationship between disease risk and genotype

$$P(\mathbf{G}) = \frac{W_u - W(\mathbf{G})}{W_u - W_a} = \frac{W_u}{W_u - W_a} - \frac{1}{W_u - W_a} exp\left(-\sum_{j=1}^{L} \alpha_j(g_j)\right).$$

It follows that under this model, the dominance coefficient and effect size for the contribution to disease risk equal those for fitness (justifying our use of the same notation for the αs in both).

We now return to the contribution of individual loci to disease risk under this model. Assuming that each locus has a small contribution, i.e., that $\alpha_j(g) \ll 1$ (which follows from $s_j \ll 1$) for $j = 1, \ldots, L$, we can approximate the variance in disease risk by

$$V(P) \approx \exp(-2\sum_{j=1}^{L} E(\alpha_j(g_j)))\sum_{j=1}^{L} V(\alpha_j(g_j)). \quad (6)$$

Nature Genetics: doi:10.1038/ng.2896

In other words, the contribution of an individual locus to variation in disease risk is proportional to the variance in fitness at that locus. Here, we consider semi-dominant and recessive loci for which the variances are

$$V(x; s, \frac{1}{2}) = \frac{1}{2}s^2x(1-x) \tag{7}$$

and

$$V(x; s, 0) = s^2x^2(1-x^2), \tag{8}$$

correspondingly.

4.2 Demographic effects on the variance

Supplementary Figure 5 depicts how different allele frequencies at semi-dominant and recessive loci contribute to the variance in disease risk under the Tennessen et al. [2] model (expanding on Figure 4 in the main text). Because we consider only one selection coefficient at a time, the relationship between effect sizes and selection coefficient has no effect here; however, we do assume that the dominance coefficient for fitness and for disease risk are the same. The graphs can also be interpreted as the proportional contribution of different allele frequencies to the variance in fitness among individuals. To elucidate the effects of recent demographic events, we also show results for the model with a constant population size (equivalent to the one for the African population before the onset of growth) and for a population that experienced the same instantaneous increase in population size as the ancestral African population in the Tennessen et al. model but then remained constant (from $\sim 7,000$ to $\sim 14,500$ around $6,000$ generations ago, cf. Supplementary Figure 1A), which we refer to as the older growth model.

Demographic effects in the semi-dominant case. First, we consider the effectively neutral regime (Supplementary Figure 5A). In the model with constant population size, the proportional contribution is uniform across frequencies, as expected [3]. In the model of older growth, there is an increased contribution of low and

high frequency alleles to the variance (as diversity patterns did not have sufficient time to reach equilibrium yet). In the model for Africans, a similar pattern is observed, with a tiny increase in the contribution from rare alleles due to recent growth (amounting to 0.41% of variance in deleterious variants with frequency below 0.1% and 0.4% in variants above 99.9%). In the model for Europeans, the increase due to growth is also negligible (0.61% of variance in variants with frequency below 0.1% and 0.6% in variants above 99.9%). However, the bottleneck leads to an increased contribution of intermediate frequencies at the expense of moderately low and high frequency alleles (since low and high frequency alleles are quickly lost or fixed after the reduction in population size).

In the weak selection regime (Supplementary Figure 5B), selection leads to a shift towards lower frequencies and thus to an increased contribution to variance of lower frequency alleles. In turn, the effect of older growth is to increase the contribution of high frequencies: the reason being that before the increase in population size, a greater proportion of sites is fixed for the deleterious allele and at such sites, normal mutations lead to high frequency deleterious alleles. The recent growth in the model for Africans further causes a small increase in the contribution of rare alleles (amounting to 1.4% of variance in variants with frequency below 0.1% and 0.07% in variants above 99.9%). In the model for Europeans, this increase is also small (1.9% of variance in variants with frequency below 0.1% and 0.1% in variants above 99.9%), but the bottleneck again has a substantial effect, increasing the contribution of intermediate frequencies at the expense of lower and higher frequencies.

In the strong selection regime, because of the quick turnover of deleterious alleles, the older increase in population size and the bottleneck in Europeans are too far in the past to have had an effect on alleles that are currently segregating (Supplementary Figure 5C). By the same token, in the Tennessen et al. model, alleles segregating at present are young and therefore the recent growth resulted in a decrease in their frequencies (cf. section 2.3), substantially increasing the contribution of rare alleles to variance (with $\sim 70\%$ of the variance contributed by alleles at frequency below

Nature Genetics: doi:10.1038/ng.2896

0.1%).

Demographic effects in the recessive case. In this case, recent growth has little effect in all selection regimes. The contribution of low frequency alleles to variance is much smaller because their effect on load or disease risk is manifested only in homozygotes (Supplementary Figure 5D-F). As a result, the increase in the number of rare deleterious alleles caused by recent growth has a negligible effect on their contribution to the variance in disease risk under both the model for Europeans and Africans (amounting to $\sim 10^{-4}\%$ in the neutral regime, $\sim 5 \cdot 10^{-4}\%$ in the weakly selected and $\sim 0.01\%$ in the strongly selected regime, in variants with frequency below 0.1%). In turn, the increase in the number of high frequency alleles (due to normal mutants on a deleterious background) has a higher impact but it is still quite small (amounting to $\sim 1\%$ in the neutral regime and $\sim 0.2\%$ in the weakly selected regime that are due to variants with frequency above 99.9%).

In the weak and strong selection regimes, there is a peak in the contribution to variance at intermediate frequency (Supplementary Figure 5E and F). Moving from low to intermediate frequencies, the contribution to the variance of a mutant allele increases (see Equation 8). This increase is halted, however, because at higher frequencies, selection on homozygotes for the deleterious allele kicks in, leading to few alleles at high frequencies. (Specifically, for a constant population size and given a low mutation rate, the frequency spectrum of deleterious alleles is well approximated by $C\frac{e^{-\alpha x^2}}{x}$, where C is a normalizing constant [3], and thus the contribution to variance can be approximated by $De^{-\alpha x^2}x(1-x)^2$, where D is a normalizing constant.) In the model for Africans (and for older growth), this peak is at higher frequencies in the weak selection regime (Supplementary Figure 5E), because the older increase in population size led to relatively more high frequency alleles at present.

The bottleneck in the model for Europeans has a much more pronounced effect, causing a shift toward intermediate allele frequencies and a corresponding shift in the contribution to variance in all selection regimes (Supplementary Figure 5D-F). As

Nature Genetics: doi:10.1038/ng.2896

opposed to the semi-dominant case, this is also true for the strong selection regime, as recessive deleterious alleles can reach substantial allele frequencies.

Summary. Population growth increases the relative proportion of rare alleles and could therefore be expected to increase their relative contribution to the variance in disease risk. However, because rare alleles contribute less to the variance to begin with, this effect may be relatively small. Assessing the effects of growth on the genetic architecture of disease risk therefore requires quantification. Here, we have shown that, at least based on current estimates of recent growth, the effects on the variance in disease risk are expected to be negligible. The one exception is the case of strongly selected quasi-dominant alleles, which are young and therefore whose frequencies do reflect the recent population size expansion. Interestingly, in this case, while the architecture of disease risk is substantially affected by growth, the expected load (or disease prevalence) remains unchanged, i.e., the same load will be due to many more deleterious alleles that segregate at lower frequencies than had the population not grown.

In contrast to growth, the bottleneck in European populations should have increased the proportion of intermediate frequency deleterious alleles at the expense of low and high frequency ones (with the exception of strongly selected quasi-dominant alleles, because they are so young). In other words, in these populations, there will be only a small effect on load but a substantial effect on the architecture of disease, with a greater proportion of the variance in disease risk due to intermediate frequency alleles.

4.3 The contribution of rare alleles in a mixture model

In reality, we expect that the variants underlying a complex disease will have a variety of selection coefficients and effect sizes rather than a single one. Under a model with such a mixture, the expected contributions of different allele frequencies to the variance in disease risk can be derived as follows. For simplicity, assume

that mutations are semi-dominant (so the dominance coefficient is dropped from the notation). At a site with selection coefficient s, the expected contribution to the variance from deleterious alleles below frequency ω is

$$V_\omega(s) = \frac{1}{2} C E(a^2|s) \int_0^\omega f(x;s) x(1-x) dx, \tag{9}$$

where $E(a^2|s)$ is the expectation of the effect size squared for sites with selection coefficient s, $f(x;s)$ is the probability of the deleterious allele being at frequency x (here, we do not condition of the allele segregating) and the proportion coefficient C is akin to the first term in Equation 5. The overall contribution to variance of a site is $V_1(s)$ and the fraction of that contribution coming from variants below frequency ω is $\Theta_\omega(s) \equiv \frac{V_\omega(s)}{V_1(s)}$. When all sites are considered jointly, denoting the input of mutations with selection coefficient s by $\mu(s)$, the expected proportion of variance from deleterious alleles below frequency ω is then

$$\Theta_\omega = \frac{\int_s \mu(s) V_1(s) \Theta_\omega(s) ds}{\int_s \mu(s) V_1(s) ds}. \tag{10}$$

Examining the terms in Equation 10 suggests that the contribution of rare alleles depends strongly on the relationship between effect sizes and selection coefficients. Specifically, the proportional contribution of rare alleles $\Theta_{0.1\%}(s)$ becomes substantial only for strong selection coefficients (Figure 4D in the main text), as shown in section 4.2. The behavior of the overall contribution to variance $V_1(s)$, however, depends on the relationship between effect sizes and selection coefficients. If we assume that the effect sizes do not depend on the selection coefficients (or more precisely that $E(a^2|s)$ is constant) then $V_1(s)$ from weakly selected sites is much greater than from strongly selected sites (Figure 4E in the main text) and rare alleles will make an important contribution only if a very large fraction of the mutational input is at strongly selected sites. If we assume the other extreme in which the effect sizes are proportional to the selection coefficient (or more precisely that $E(a^2|s) \propto s^2$, as in the model in section 4.1) then $V_1(s)$ strongly increases with the s (Figure 4E in the main text) and rare alleles would make an important contribution unless the

Nature Genetics: doi:10.1038/ng.2896

fraction of the mutational input at strongly selected sites is very small. In reality, the outcome could be anywhere in between.

As an illustration, we consider a simple model in which we vary the correlation between selection on variants and their effect on a trait. We assume that half of the newly arising mutations have a weak selection coefficient $s_w = 0.0002$ and half have a strong selection coefficient of $s_s = 0.01$. For strongly selected mutations, the effect size on the trait, a, is chosen to be cs_s with probability $\frac{1}{2}(1+p)$ and cs_w with probability $\frac{1}{2}(1-p)$, where c is a positive constant and $0 \leq p \leq 1$; correspondingly, for weakly selected mutations the effect size is chosen to be cs_w with probability $\frac{1}{2}(1+p)$ and cs_s with probability $\frac{1}{2}(1-p)$. In this model, the marginal distributions of selection coefficients and effect sizes do not depend on p, while the correlation between them is equal to p. To obtain Figure 4F in the main text we therefore vary p between 0 and 1.

References

[1] Charlesworth, B. & Charlesworth D. *Elements of Evolutionary Genetics* (Roberts and Co., 2010).

[2] Tennessen, J. A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337**, 64-69 (2012).

[3] Ewens, W. J. *Mathematical Population Genetics*, 2nd ed. (Springer, 2004).

[4] Fu, W. *et al.* Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216-220 (2013).

[5] Kong, A. *et al.* Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471475 (2012).

Nature Genetics: doi:10.1038/ng.2896

[6] Sun, J. X. *et al.* A direct characterization of human mutation based on microsatellites. *Nature Genetics* **44**, 11611165 (2012).

[7] Gutenkunst, R. N., Hernandez R. D., Williamson S. H. & Bustamante C. D. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLoS Genet.* **5**, e1000695 (2009).

[8] The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 10611073 (2010).

[9] Tajima, F. The effect of change in population size on DNA polymorphism. *Genetics* **123**, 597601 (1989).

[10] Hudson, R. R. Gene genealogies and the coalescent process. *Oxf. Surv. Evol. Biol.* **7**, 144 (1990).

[11] Wakeley, K. *Coalescent Theory: An Introduction* (Roberts and Co., 2008).

[12] Gillespie, J. H. *Population Genetics: A Concise Guide*, 2nd ed. (Johns Hopkins University Press, 2004).

[13] Feller, W. *An Introduction to Probability Theory and its Applications* (Wiley, 1968).

[14] Green, R. E. *et al.* A Draft Sequence of the Neandertal Genome. *Science* **328**, 710-722 (2010).

[15] Keinan, A. & Clark, A. G. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* **336**, 740-743 (2012).

[16] Otto, S. P. & Whitlock, M. C. The probability of fixation in populations of changing size. *Genetics* **146**, 723-733 (1997).

[17] Coventry, A. *et al.* Deep resequencing reveals excess rare recent variants consistent with explosive population growth. *Nat. Commun.* **1**, 131 (2010).

[18] Nelson, M. R. *et al.* An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**, 100-104 (2012).

[19] Pritchard, J.K. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **69**, 124-137 (2001).

[20] Eyre-Walker, A. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proc. Natl. Acad. Sci. USA* **107**, 1752-1756 (2010).

[21] Wang, K., Li, M. &Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

[22] Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248-249 (2010).

[23] Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols* **4**, 1073-1081 (2009).

[24] Schwarz, J.M., Rödelsperger, C., Schuelke, M. &Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575-576 (2010).

[25] Chun, S. & Fay, J. C. Identification of deleterious mutations within three human genomes. *Genome Res.* **19**, 1553-1561 (2009).

Nature Genetics: doi:10.1038/ng.2896