www.sciencemag.org/content/vol/6337/539/suppl/DC1

Supplementary Materials for

Negative selection in humans and fruit flies involves synergistic epistasis

Mashaal Sohail, Olga A. Vakhrusheva, Jae Hoon Sul, Sara L. Pulit, Laurent C. Francioli, Genome of the Netherlands Consortium, Alzheimer's Disease Neuroimaging Initiative, Leonard H. van den Berg, Jan H. Veldink, Paul I. W. de Bakker, Georgii A. Bazykin, Alexey S. Kondrashov,* Shamil R. Sunyaev*

*Corresponding author. Email: ssunyaev@rics.bwh.harvard.edu (S.R.S.); kondrash@umich.edu (A.S.K.)

Published 5 May 2017, *Science* **356**, 539 (2017)
DOI: 10.1126/science.aah5238

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:
(available at www.sciencemag.org/content/vol/6337/539/suppl/DC1)

aah5238_Supplementary_Tables (Excel). A tab-delineated file containing all supplementary tables that did not fit into the supplementary materials file.

aah5238_Datafiles (Excel). A tab-delineated file containing DatafileS1 (a complete list of all candidate LoF singletons in six human data sets that were used for mutation burden analysis), DatafileS2 (a complete list of all samples in Drosophila Population Genomics Project (DPGP3) and The Drosophila Genetic Reference Panel (DGRP) with additional annotation), and DatafileS3 (a complete list of all candidate rare LoF variants with minor allele count ≤ 5 in two fruit fly datasets that were used for mutation burden analysis).

TABLE OF CONTENTS

Materials and Methods

Human datasets

Analysis of genome-wide mutation burden was performed on six independent datasets. The ideal dataset consists of high diversity, low-admixture, unrelated individuals from a single randomly mating population. We used three high quality European whole genome sequencing (WGS) datasets for this study - Genome of the Netherlands (GoNL)(*15*), Alzheimer's Disease Neuroimaging Initiative (ADNI), and Dutch controls from Project MinE, a amyotrophic lateral sclerosis study. We also analyzed data from three non-European populations from the 1000 genomes Phase I project, one African population (YRI) and two East Asian populations (JPT, CHS)(*18*).

The GoNL dataset (http://www.nlgenome.nl/) consists of phased whole genome-sequences of 250 Dutch parent-child trios sequenced at ~13x average coverage. Sequence data from the parent generation was used for this study.

The ADNI dataset (http://adni.loni.usc.edu/) consists of 808 whole genome-sequences sequenced at ~30x average coverage. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD (Department of Radiology, UCSF School of Medicine, San Francisco, CA, USA). The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The largest cohort has European ancestry but the dataset also includes samples of other ancestry. Only individuals of European ancestry were selected for this analysis.

Project MinE (http://www.projectmine.com) consists of ~4500 samples from different ancestries, and includes 1806 samples of Dutch ancestry post quality control, with 617 controls and 1189 cases. The average coverage is ~40x for this dataset. Only Dutch controls were analyzed for this study.

The 1000 Genomes Phase I Project (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/) provides sequence data for 1,092 individuals. We selected 3 different non-European populations for our analysis - 88 sub-Saharan Africans (YRI), 89 Japanese samples from Tokyo Japan (JPT), and 100 Southern Han Chinese samples (CHS). The Yoruba population is particularly suited to our analysis, as it has not gone through the out-of-Africa bottleneck, and is therefore, high diversity. Moreover, African populations are undergoing the second demographic transition currently; therefore, selection due to pre-reproductive mortality is not as relaxed in these populations as it is in industrialized European populations that already underwent the second demographic transition (*28*, *29*). For the 1000 Genomes cohorts, only SNPs that were discovered using Exome sequencing were used for our analysis.

For all analyses, ancestral alleles were distinguished from derived alleles based on EPO multiple-sequence alignments (available from the 1000 Genomes project). Only SNPs with high confidence on predicted ancestral alleles were analyzed.

Fly datasets

Analysis of genome-wide mutation burden was also performed on two independent fly datasets. We analyzed whole-genome polymorphism data for two *D. melanogaster* populations – African flies from Phase 3 of the Drosophila Population Genomics Project (DPGP3)(*16*) and North American flies from Freeze 2.0 of the The Drosophila Genetic Reference Panel (DGRP*)*(*19*, *30*). Polymorphism sequence data for both datasets was obtained in the format of pseudochromosome assemblies from the Drosophila Genome Nexus website (http://johnpool.net/genomes.html)(*16*). The same two-round mapping and SNP calling strategy was applied to both DPGP3 and DGRP datasets provided by the Drosophila Genome Nexus project (*16*).This makes variant calling methodology consistent between DGRP and DPGP3 datasets. However Drosophila Genome Nexus variant calls for the DGRP dataset (*16*) may differ from the SNPs available from The Drosophila Genetic Reference Panel website (http://dgrp2.gnets.ncsu.edu)(*30*).

While DPGP3 dataset consists of haploid embryo genomes (*16*), DGRP provides genomic sequences of inbred lines (*19*). Residual tracts of heterozygosity present in the genomes of the DGRP dataset after 20 generations of full-sib mating are already masked in files with DGRP pseudochromosome assemblies available on the Drosophila Genome Nexus website. Each genome in the DGRP dataset has only one sequence for each chromosome arm restricting analysis to homozygous regions of the genome. Thus DPGP3 genomes are truly haploid, while DGRP genomes could be considered effectively haploid.

Genomes belonging to DPGP3 and DGRP datasets have different fractions of bases masked due to true heterozygosity, missing data, pseudoheterozygosity or other technical artifacts (*16*). On average DGRP genomes have 22.57 % of the bases masked (with the median being equal to 16.11%), while DPGP3 genomes have only 6.91 % of the bases masked (the median = 6.87%)(Data file S1).

For *D. melanogaster*, the DPGP3 population has low levels of admixture and less variation in genomic coverage between samples as compared to the DGRP dataset. Another strong point of the DPGP3 dataset is high level of genetic diversity that in contrast to North-American DGRP population has not been affected by out-of-Africa bottleneck (*31*).

Human data analysis

Sample quality control

Outliers were detected and removed based on inbreeding coefficient, ethnicity using principal component analysis (PCA), contamination, and relatedness. Total numbers of singletons and SNPs were computed to detect and remove outliers that were more than 3 standard deviations from the genome wide mean for the dataset (direction of results was not affected by outlier removal). The number of samples removed in each filtering step for all 6 datasets is listed in Table S9.

The GoNL Consortium sequenced 500 unrelated individuals from 250 trios. The consortium removed two samples after quality control due to contamination. We further detected outlier samples based on number of singletons per genome. Only 3 individuals were flagged as outliers. 495 samples were retained for further population genetic analysis.

The ADNI study sequenced 808 individuals. Plink was used to run identity-by-descent analysis and 5 related pairs were detected (Pi_hat > 0.4). One sample for each related pair was chosen randomly and removed. The remaining 803 individuals consisted of both samples with self-reported European and non-European ancestry (2 Indian/Alaskan, 10 Asian, 27 African, 9 multi-ethnic, 2 unknown, 2 Hawaiian, and 751 European). We performed principal component analysis on the 803 unrelated individuals using EIGENSTRAT (*32*) to detect outliers (non-European samples). After outlier removal (5 outlier removal iterations on 10 principal components, with outliers defined as outside 6 standard deviations), we retained 744 of the 751 European samples for further analysis. Finally, we detected and removed outliers based on the numbers of singletons and SNPs per genome. Thirty individuals were flagged as outliers and removed. 714 samples were retained for further population genetic analysis.

Project MinE includes 1806 Dutch samples post quality control, with 617 controls and 1189 cases. We further removed 8 related samples (twins) from the controls. Finally, we detected 4 outliers by genome-wide singletons, and 4 additional outliers by genome-wide SNPs. We retained 601 samples for further population genetic analysis.

The 1000 genomes project has an unexpectedly high level of inbreeding and relatedness (*33*). We chose proposed subsets of unrelated and outbred individuals using Fsuite (*33*) for our analysis.

Variant quality control

Single nucleotide polymorphisms (SNPs) residing within protein-coding genes were the only type of genetic variation considered for the analysis of the human datasets. All SNPs were called against the GRCh37 human reference assembly. Only bi-allelic SNPs were analyzed. We used various quality control measures to remove putative false positives before performing population genetic analyses (Table S8). For all six datasets, only SNPs with no missing data were analyzed, and SNPs out of Hardy-Weinberg equilibrium (P-value < 10^{-6}) were removed. Genetic variation in all human datasets except Project MinE was called using the Broad Institute's GATK pipeline (*34*)(*35*). We only considered SNPs that passed GATK's Variant Quality Score Recalibration (VQSR) filter for further analysis.

In the GoNL dataset, the average sequencing coverage is ~13x. To remove poorly sequenced genomic regions from consideration, all SNPs with less than half or greater than twice the mean coverage were flagged as the inaccessible genome (*15*) and removed from further analysis.

In the ADNI dataset, the average sequencing coverage is ~30x. For this dataset, as suggested by the GATK group, we removed poorly sequenced genomic regions by treating all genotypes with low quality (GQ < 20) as missing.

To obtain access to genetic variation in African and other non-European populations, we downloaded release 3 of the 1000 genomes Phase 1 project. No additional filtering on coverage or genotype quality was performed on SNPs in the 1000 genomes populations, as the VCF files do not provide any sequencing depth or quality information.

In the MinE dataset, the average sequencing coverage is ~40x. Genetic variation in the MinE dataset was called using Illumina's Isaac pipeline (*36*), which generates sample-level gVCF files. These files were merged using Illumina's agg tool to obtain

genetic variation across all samples. For each SNP, agg returns a PF variable describing the proportion of samples with a "PASS" variant at that genomic position. Sequencing depth (DP) and PF were binned by Ti/Tv and het/hom-non-ref ratio to determine quality control thresholds to eliminate putative false positives. SNPs that were "PASS" variants in only a small proportion of samples (PF < 0.6) or were in inaccessible parts of the genome (DP < 10x per sample or DP > 40x per sample) were removed from further analysis. We also removed all sites with QUAL < 30 and treated all genotypes with GQ < 10 as missing.

Variant annotation

Functional consequences of genetic variants were annotated using Ensembl Variant Effect Predictor version 82. One transcript was chosen per variant using an ordered set of criteria (canonical status of transcript, APRIS isoform annotation, transcript support level, biotype of transcript with protein coding preferred, CCDS status of transcript, variant consequence rank in order of severity from more severe to less severe, and transcript or feature length with longer preferred). Alleles were classified as synonymous, missense (nonsynonymous), nonsense (stop gain mutations resulting in a premature stop codon leading to a shortened transcript, and stop loss mutations where at least one base of the terminator stop codon is changed resulting in an elongated transcript), and splice-disrupting (splice acceptor mutations that change the 2 base region at the 3' end of an intron and splice donor mutations that change the 2 base region at 5' end of an intron). Loss-of-Function (LoF) alleles were defined as the joint set of nonsense and splice-disrupting alleles. Any synonymous allele that is also within the splice site region (within 1-3 bases of the exon or 3-8 bases of the intron) was removed from further analysis. Any transcripts marked for nonsense-mediated decay were also removed from further analysis.

Mutation burden for each human was calculated as the number of derived alleles in the genome of a given human for each functional type of mutations. We calculated the mutation burden for singletons (Table 1) and for different derived allele frequency cutoffs (Table S2).

Gene filters

To focus our analysis only on truly deleterious protein-damaging mutations we discarded pseudogenes and genes belonging to taste receptor and olfactory receptor families, as it is likely that protein-damaging mutations in these rapidly evolving non-essential genes will have only minor fitness effects. For this filtering step, we downloaded a list of genic positions and descriptions for the GRCh37 human reference assembly from Ensembl Biomart.

To remove confounding effects of known long-range LD, we also removed the MHC region on chromosome 6 (28Mbp – 35Mbp), inversion on chromosome 8 (6 Mbps – 15 Mbps) and long LD stretch on chromosome 17 (40 Mbps – 45 Mbps) from all analyses.

The number of SNPs and singletons left after above-mentioned filtering steps are shown in Table S8.

Analysis of the crucial genome

We used genic selection coefficients estimated using Exome sequencing data from 60,706 individuals to determine a set of genes with crucial functions (*37*). Briefly, selection coefficients against heterozygous loss of gene function were estimated using rare protein truncating variants (PTVs), and Bayesian estimates were obtained for individual genes. Genes found under strongest selection were shown to be enriched in embryonic lethal mouse knockouts, putatively cell-essential genes inferred from human tumor cells, Mendelian disease genes, and regulators of transcription (*37*). We used this set of 1599 crucial genes with heterozygous selection coefficients exceeding 0.2 for our crucial genome analysis (Fig. 2b, Fig. S8, Table S22).

When only their crucial genome was considered, humans (Figs. 2B, S8) showed an underdispersion in their missense mutation burden. In contrast, synonymous alleles remained overdispersed, though to a lesser degree. This is likely because overdispersion scales not only with selection strength, but also with the number of alleles. The effect of confounders increases variance proportionally to the number of alleles, as every pair of alleles contributes to the excess positive LD (the excess variance is a sum of pairwise LDs, see pg. 14 of supplement for further details).

Given that synonymous alleles are equal or fewer in number than missense alleles in human genes, we cannot explicitly obtain P-values by resampling. However, analysis in synonymous alleles at the same allele frequency does not show a signal of negative LD, but rather a signal of weak positive LD as discussed above. We also obtained P-values for σ^2/V_A by permuting functional consequences across variants (Table S22).

Analysis of simple insertions and deletions

A set of high quality simple insertions and deletions (indels) were generated as part of the GoNL structural variants (SV) dataset (*38*). Briefly, 250 parent-offspring families (769 individuals) from the Dutch population were used to generate a high-quality SV-integrated, haplotype-resolved reference panel, using 12 different variant detection tools representing 4 algorithmic approaches (gapped alignment and split-read mapping, discordant read pair, read depth and de novo genome assembly). The results from the different detection tools were combined into a consensus set containing 9 different forms of SVs and indels. The GoNL study further selected a representative set of variant candidates for validation using PCR amplification of breakpoint junctions, and subsequent sequencing of the PCR products via Sanger or MiSeq sequencing (98% confirmation rate for simple indels was observed)(*38*).

We obtained access to resulting genotype calls for 646,011 short insertions (1–20 bp) and 1,093,289 short deletions (1–20 bp). Before performing population genetic analyses, we further removed variants out of Hardy-Weinberg equilibrium ($p < 10^{-3}$), and only kept bi-allelic variants with a low proportion of missing data (only variants with greater than 99% genotypes present were retained). Functional consequences of simple indels were annotated using Ensembl Variant Effect Predictor version 82. Analogously to SNPs, we removed known genomic regions with long-range LD for this analysis as well, and attempted to focus our analysis only on truly deleterious protein-damaging indels by discarding pseudogenes and genes belonging to taste receptor and olfactory receptor families. We report results for the joint mutation burden category of LoF SNPs and frameshift indels (Table S19). P-values were computed by resampling coding

synonymous SNPs and intronic indels at matching allele frequency as LoF SNPs and frameshift indels respectively.

Fly data analysis

Reference genome and annotation data

D. melanogaster reference genome sequence and annotation data were downloaded from the UCSC Genome Browser site (https://genome.ucsc.edu)(*39*). We used dm3 version of *D. melanogaster* genome assembly from the UCSC database corresponding to BDGP release 5 of *D. melanogaster* genome reference sequence. Our analysis is based on FlyBase v. 5.12 protein-coding gene annotations for *D. melanogaster* genome (*40*). Only canonical isoforms of the genes were considered in the analysis. Genes residing outside of 5 *D. melanogaster* euchromatic chromosome arms (2L, 2R, 3L, 3R and X) were not taken into consideration.

Gene filtering

Annotation errors

To focus our analysis only on truly deleterious protein-damaging mutations we discarded gene models with putative annotation errors. A total of 48 gene models were excluded after this step.

Genes were excluded from further analysis if any of the following conditions were true:

1) If the CDS carried a premature termination codon in the reference genome assembly.

2) If the CDS was lacking canonical termination stop codon.

3) If the CDS length was not a multiple of three.

Chemoreceptor genes

We additionally excluded from all analyses sequences belonging to chemoreceptor gene and odorant-binding protein gene families due to the fact that these genes have previously been shown to undergo frequent pseudogenization (*41*, *42*) and are enriched for nonsense alleles (*30*, *43*). These two large gene families include multiple rapidly evolving paralogs that are frequently lost in the course of evolution (*44*).

Thus it is likely that protein-damaging mutations in these rapidly evolving non-essential genes will have only minor fitness effects. Lists of genes constituting families of chemoreceptors and odorant-binding proteins were downloaded from FlyBase (http://flybase.org). There were a total of 121 and 52 genes belonging to *D. melanogaster* chemoreceptor and odorant binding protein gene families respectively. flyBaseToCG.txt table available on the UCSC site was used to map CGID gene symbols to transcript IDs. Additional 164 gene models were removed after this step.

Inversions

Polymorphic inversions have been previously shown to account for the majority of population structure in fly populations (*30*, *45*). Assuming that population structure is expected to lead to overdispersion in mutation burden we separately analyzed SNPs residing in inversion-free regions of *D. melanogaster* genome. For this purpose we retrieved genomic coordinates of the inversions known to segregate in fly populations from published studies (*30*, *45*) and obtained a list of genes that do not overlap with

known inversions. The original number of FlyBase canonical gene models and the numbers of gene models left after above-mentioned filtering steps are shown in Table S10.

Sample quality control

For both fly datasets we detected outlier samples with extremely large or small numbers of SNPs or extremely large or small numbers of genomic bases masked to N. A sample was flagged as an outlier if the number of SNPs or the number of masked genomic bases in a sample was more than 3 standard deviations away from its corresponding dataset mean. We identified 7 and 6 outliers in the DGRP and DPGP3 datasets respectively and excluded them from further analysis.

Sample ZI382 belonging to the DPGP3 dataset was removed from consideration due to chromosome X missing from the assembly. We additionally removed from the analysis all samples from the DGRP dataset that had > 20% of the genomic sequence masked. After these steps there were 191 out of 197 flies left in the DPGP3 dataset and 125 out of 205 flies left in the DGRP dataset. The samples retained for the analysis are listed in Data File S1.

Minor variant identification, variant annotation and quality control

Single nucleotide polymorphisms residing within protein-coding genes were the only type of genetic variation considered for the analysis of the fly datasets. Specifically we analyzed SNPs falling within protein coding regions of the genes and splice site SNPs. A consensus sequence for the coding portion of each gene was constructed separately for two populations by picking the nucleotide that occurs most frequently at each position. After that we scanned resulting consensus CDS sequences in a codon-wise manner and discarded from the analysis codons containing ambiguous nucleotides (N) in the consensus. Analogously we constructed consensus sequences for each splice site of each gene. Only splice sites with canonical dinucleotides (GT for donor and AG for acceptor sites) in the consensus and in the reference *D. melanogaster* genome were retained for further analysis.

We searched for segregating sites in each dataset independently and calculated allele frequencies for all available alleles at each site. We excluded from the analysis SNPs in codons carrying more than one SNP at least in one fly as the effects of such SNPs could not be evaluated independently and such cases are likely to be enriched with double mutations (Table S11). Analogously splice sites carrying more than one SNP at least in one fly were removed from the dataset (Table S12). SNPs were classified as synonymous, missense (nonsynonymous), nonsense (stop gain mutations resulting in a premature stop codon leading to a shortened transcript) or splice-disrupting according to Flybase annotation of *D. melanogaster* genome with respect to the major (consensus) variant. Stop loss mutations were not included in nonsense alleles for fly genomes as they are not expected to be under strong negative selection (*43*). All mutations falling in splice sites were labeled as splice-disrupting except for the SNPs in donor splice sites resulting in a weak variant of donor site (GT-> GC). Loss-of-Function (LoF) alleles were defined as the joint set of nonsense and splice-disrupting alleles.

Mutation burden for each fly was calculated as the number of minor alleles in the genome of a given fly for each functional type of mutations. We calculated the mutation

burden for all minor alleles (with minor allele frequency < 50%) and for different minor allele frequency cutoffs. Missing genotypes are expected to inflate variance of the mutation burden, thus only sites without missing data were included in the calculation of the mutation burden in the DPGP3 dataset. We used all available segregating sites when calculating the mutation burden in the DGRP dataset as only 11% of the segregating codon sites and 13% of the segregating splice sites were left after exclusion of the sites with missing genotypes due to a large fraction of the genomic bases masked from individual DGRP genomes (Table S11, Table S12).

Due to high levels of polymorphism in the DPGP3 population a significant fraction of codons have more than 2 alleles (8% of the variable codons are multi-allelic in the DPGP3 dataset). To account for the effects of presence of multi-allelic sites on our results we computed mutational burden for each functional class of alleles (nonsense, nonsynonymous, synonymous) after discarding codons with more than one minor allele belonging to a given functional class (Table S13). Analogously multi-allelic splice sites were also removed at this step (Table S13). The direction of the effect remained unchanged after exclusion of multi-allelic sites.

Analysis of the essential genome

The list of essential *D. melanogaster* genes was downloaded from the DEG database (http://tubic.tju.edu.cn/deg/)(*46*). There are 339 genes listed as essential in *D. melanogaster* in the DEG database. We were able to unambiguosly map 267 genes out of this list to the genes from the filtered list we used in the analysis. We calculated σ^2/V_A for the number of missense and synonymous alleles residing within the essential genes in both *D. melanogaster* datasets. All missense and synonymous alleles with minor allele frequency up to 50% were considered for this analysis. Low numbers of LoF alleles in the essential genes do not allow separate analysis of this type of variant (there are 6 and 3 LoF alleles in the essential genes in the DPGP3 and DGRP datasets respectively). Missense alleles residing within the essential genes show underdispersion in the DPGP3 dataset ($\sigma^2/V_A = 0.947$, Fig. S9A) but not in the DGRP dataset ($\sigma^2/V_A = 2.729$).

To assess the significance of the underdispersion signal for the missense alleles from essential genes in the DPGP3 dataset we resampled synonymous and missense alleles at the population frequencies matching the frequency distribution of the missense alleles residing within the essential genes and calculated σ^2/V_A in the resampled datasets. In addition we separately resampled synonymous alleles residing within the essential genes. Missense alleles residing in the essential genes are significantly underdispersed compared to the resampled distributions of synonymous alleles randomly picked in the genome ($p = 0.002$) as well as compared to the resampled distributions of synonymous alleles randomly picked from the essential genes ($p < 10^{-3}$).
In addition, only 20.9% of the resampled missense datasets show σ^2/V_A less than or equal to the corresponding ratio for the missense alleles residing within the essential genes (Fig. S9B).

We realized that the resampled sets of synonymous alleles picked from the essential genes show overdispersion as compared to the resampled sets of synonymous alleles with the matching population frequencies picked from the random genomic locations (Fig. S9B). As long as there are only 267 essential genes in the dataset the average physical distance between SNPs restricted to the essential genes is expected to be

much smaller as compared to SNPs chosen from random genomic positions as many variants fall into the same genes, and this in turn is supposed to lead to stronger LD between such SNPs and overdispersion in the mutation burden for such alleles. To control for the effects of physical linkage between SNPs from the essential genes we generated 1000 random sets of synonymous and missense SNPs restricting the maximum number of genes in each set to 400. Indeed, the distribution of σ^2/V_A in the resampled synonymous datasets obtained this way is much closer to the corresponding distribution in the resampled datasets of the synonymous alleles restricted to the essential genes (Fig. S9C). The extent of underdispersion for missense alleles from the essential genes in the DPGP3 population is even more pronounced when compared to the resampled datasets of synonymous ($p < 10^{-3}$) and missense ($p = 0.076$) alleles if the gene number is controlled for (Fig. S9C).

dN/dS analysis

We obtained dN/dS values for *D. melanogaster* genes possessing only a single ortholog in the *melanogaster* subgroup from the published study (*47*) available at ftp://ftp.flybase.net/genomes/12_species_analysis/clark_eisen/paml. This dataset contained dN/dS values for 8510 genes, 7888 out of these genes could be unambiguously mapped to the genes from the filtered list we used in our study. Single dN/dS ratio estimates for each gene from model M0 were used in the subsequent analyses. Genes with total tree length for dS > 4 or values of dN/dS > 0.7 were excluded from further consideration leaving a total of 7836 genes.

Remaining genes were subdivided into five equal-sized bins according to the dN/dS ratio where bin 1 contains the most slowly evolving genes and bin 5 contains the most rapidly evolving genes. The median values of dN/dS for the bins 1-5 are as follows: 0.018, 0.041, 0.063, 0.097, and 0.184. σ^2/V_A for the number of common LoF, missense and synonymous alleles was computed separately for each bin (Table S20). Missense alleles residing within the most slowly evolving genes (dN/dS bin 1) show underdispersion ($\sigma^2/V_A = 0.98$) in the DPGP3 dataset, which is in line with the underdispersion of the missense alleles in the essential *D. melanogaster* genes in this dataset. LoF alleles residing within relatively slowly evolving genes are underdispersed in both fly datasets (dN/dS bins 1-3 in the DPGP3 dataset and dN/dS bins 1-2 in the DGRP dataset).

More generally σ^2/V_A for the number of LoF and missense alleles tends to increase with the relaxation of selective constraint on a gene in both fly datasets while synonymous alleles show no such tendency (Table S20). However this trend could be attributed not only to epistatic interactions between deleterious alleles but also to the increase in the mean number of the deleterious alleles occurring with the increasing dN/dS ratio.

To explicitly control for the mean number of alleles as well as for allele frequencies we resampled missense alleles in dN/dS bins 2,3,4,5 and synonymous alleles in dN/dS bins 1,2,3,4,5 at the population frequencies matching the frequency distribution of the missense alleles in the most conserved dN/dS bin 1 (Fig. S18). We calculated median values and 95% confidence intervals for σ^2/V_A in 1000 resamplings for each bin of dN/dS. This analysis was performed separately for resampled synonymous and missense alleles.

After controlling for the allele frequencies, σ^2/V_A for the number of missense alleles shows an obvious decrease with the increase in the strength of purifying selection in both fly datasets while no similar trend was observed for synonymous alleles (Fig. S18). The dependency of the variance of the number of missense alleles on the degree of a gene's evolutionary constraint in the absence of such dependency for synonymous alleles points to selective forces as the main drivers of the signal.

Computing variance-based estimator of net LD

For each individual in a population of size N, X_i is a discrete random variable that represents the number of derived alleles present at locus i and can take values 0, 1 or 2. The mutation burden for individual k with L polymorphic loci is defined as,

$$B^k = \sum_{i=1}^{L} X_i$$

Under multiplicative selection, the variance of the mutations burden σ^2 is equal to the additive variance V_A computed as $\sum_i 2p_i\,(1-p_i)$ for all loci i with mutant allele frequency p_i in the genome (*10*). Note that V_A is also mathematically equivalent to the genome-wide nucleotide diversity **π** (*48*). For the mutation burden distribution,

$$\mu = \frac{\sum_{k=1}^{N} B^k}{N}$$

$$\sigma^2 = VAR\left(\sum_{i=1}^{L} X_i\right) = \frac{1}{N-1}\sum_{k=1}^{N}(B^k - \mu)^2$$

$$V_A = \sum_{i=1}^{L} VAR(X_i) = \sum_{i=1}^{L}\left(\frac{1}{N-1}\sum_{k=1}^{N}(X_i^k - \bar{X}_i)^2\right) \;\; where\; \bar{X}_i = \frac{\sum_{k=1}^{N} X_i^k}{N}$$

For a diploid population,

$$Net\ LD = \sigma^2 - V_A = VAR\left(\sum_{i=1}^{L} X_i\right) - \sum_{i=1}^{L} VAR(X_i) = 2\sum_{i,j} COV(X_i, X_j) = 4\sum_{i,j} D_{i,j}$$

In the case of independent loci, or when negative disequilibrium exactly cancels out positive disequilibrium in a population,

$$VAR\left(\sum_{i=1}^{L} X_i\right) = \sum_{i=1}^{L} VAR(X_i)$$

$$Net\ LD = \sigma^2 - V_A = 2\sum_{i,j} COV(X_i, X_j) = 4\sum_{i,j} D_{i,j} = 0$$

Genome-wide net LD for each dataset was computed using scripts written in the R programming language. Variance estimations were computed on X_i matrices using the matrixStats CRAN package. Net LD is normalized in two ways: per pair of derived alleles (divided by the square of the mean mutation burden μ) and per pair of loci (Table 1). This allows for comparison between different functional classes of variants (mutation burden for synonymous and missense alleles has higher μ than mutation burden for LoF alleles).

For our primary human (GoNL, ADNI, MinE) and fly (DPGP3) datasets, the distribution of rare LoF alleles is underdispersed (Table 1); nonsense alleles in the MinE dataset, if considered separately, are the exception, although underdispersion was also observed for stop gain alleles in this dataset at a slightly higher allele frequency threshold (Table S2).

One-sided P-values for σ^2/V_A were obtained by permuting functional consequences across variants (1000 permutations) for each human dataset (Table S1). A joint P-value for all human datasets was computed by meta-analysis using Stouffer's method weighted by each dataset's sample size (*49*).

One-sided P-values for σ^2/V_A of LoF alleles in both human and fruit fly datasets were also obtained by resampling synonymous alleles at matched allele frequency in each dataset (Fig. 3). For each LoF allele, a matching synonymous allele at the same allele frequency was picked to generate a set of synonymous alleles with the same μ as the LoF alleles in the dataset. σ^2/V_A was computed on the set of resampled synonymous alleles. This process was repeated 1000 times to generate an empirical null distribution for σ^2/V_A for each dataset. A joint P-value for all human datasets and all fly datasets respectively was computed by meta-analysis using Stouffer's method weighted by each dataset's sample size (*49*). Analogously we also resampled nonsynonymous alleles at matched allele frequencies as LoF alleles for each dataset and obtained a σ^2/V_A distribution for the sets of resampled nonsynonymous alleles (Fig. 3).

If the p-value for a variant type in a dataset was less than 0.001 by permutation or resampling, we used p = 0.001 as its value for the meta-analysis.

By meta-analysis, the underdispersion signal in rare LoF alleles is more significant in flies compared to humans (Fig. 3) which may be due to weaker recombination in flies compared to humans (*50*, *51*). Recombination opposes the reduction in genetic variance caused by negative LD. Also, human populations have recently experienced relaxed selection (*28*, *29*). Without selection, recombination would rapidly destroy linkage disequilibrium between deleterious alleles.

Furthermore, through regression analysis and resampling experiments, we showed that the underdispersion signal persists after correcting for potential confounders of population structure and batch processing (Table S5), variable coverage across the genome (Table S18), and that it is not driven by a small set of outliers (Table S17).

Computing $D_{i,j}$ based estimator of net LD

In the GoNL dataset, Plink was used to compute pairwise correlation coefficient r^2 between all pairs of loci for each functionally annotated class of variants. The ${r_{i,j}}^2$ values for each pair of SNPs i and j were used to compute $D_{i,j}$ values using the formula,

$$r_{i,j}{}^2 = \frac{D_{i,j}{}^2}{p_A p_a p_B p_b}$$

where p_A and p_a are the major and minor allele frequencies respectively at polymorphic locus i and p_B and p_b are the major and minor allele frequencies respectively at polymorphic locus j.

We summed $D_{i,j}$ values for each intra-chromosomal pair of SNPs and each inter-chromosomal pair of SNPs to partition net LD by linkage (Table S6). We also summed $D_{i,j}$ values by chromosome for every intra-chromosomal pair of SNPs on that chromosome (Table S7). Finally, we plotted net LD binned by physical distance between pairs of SNPs (Fig. S6, Fig. S7).

Identifying positive disequilibria sources in rare mutation burden

Even for a set of independent alleles, overdispersion in the mutation burden is observed if genome-wide positive LD is present due to population structure, which can also be seen as deviations from Hardy-Weinberg equilibrium (Wahlund effect) for the entire genome (*17*, *52*). Overdispersion may also be caused by DNA samples sequenced or processed in different batches, which can introduce heterogeneity with a clustering effect similar to that of geographic structure.

We computed Pearson's r in rare mutation burden between spouses in the GoNL dataset (r = 0.31). We removed all samples that were not in a spousal pair (5 samples) for this analysis. Rare mutation burden was computed using coding synonymous singletons. All remaining GoNL samples were divided into 3 broad geographic regions (north, central, south)(*15*) and average rare mutation burden (μ) was computed for each subpopulation (Fig. S3, $\mu_{north} = 26.88$, $\mu_{central} = 31.98$, $\mu_{south} = 33.74$). We observed that rare mutation burden for synonymous alleles decreases in a south-to-north gradient. This is consistent with the pattern observed for common variants in the GoNL dataset, which can be explained by sequential bottlenecks as the population moved northwards (*15*). We also observed a positive correlation between rare mutation burden and the first principal component (Fig. S4). We conducted a multivariate regression analysis to study sources of overdispersion for the genome-wide rare mutation burden computed using synonymous singletons. Genome-wide mutation burden was regressed under the following model:

$$Burden = \beta_0 + \beta_1 batch\ 1 + .. + \beta_5 batch\ 5 + \beta_6 region\ North + \beta_7 region\ Central + \beta_8 region\ South + \beta_9 PC\ 1 + .. + \beta_{19} PC\ 10$$

where $batch$ refers to the sequencing batch (our GoNL samples were sequenced in 5 batches), $region$ refers to the geographic region where the sample originated (The Netherlands is divided into 3 broad regions - north, central, and south), and PC refers to principal component.

Principal components in the GoNL dataset were computed using EIGENSTRAT (*32*). To perform PCA analysis, we removed SNPs with greater than 5% missing data, sites out of Hardy-Weinberg (P-value < 0.001), SNPs residing in the inaccessible genome of GoNL, and retained only common SNPs (minor allele frequency > 0.05). Before

computing principal components, we removed regions with long range LD (*53*), and performed two-step LD-pruning using PLINK to obtain a set of independent SNPs. In step one of LD pruning, we used a window of 200 SNPs with a step size of 5 and a VIF threshold of 1.03. In step two of LD pruning, we performed pairwise pruning using an r^2 threshold of 0.1. Table S4 lists the coefficients and P-values for all covariates of population and technical structure in the multivariate regression model. We repeated the regression analysis for missense alleles and LoF alleles (Tables S15, S16).

We regressed mutation burden at each locus X_i under the same model as above. We used the residuals for burden at each locus to compute the residualized genome-wide mutation burden. We observed no correlation between the residualized burden and the first principal component (Fig. S5). We computed σ^2 and V_A for the residualized mutation burden (Table S5). Positive disequilibria due to population structure and other sources of correlations between samples, such that sub-populations show different values of the mean mutation burden μ, leads to overdispersion in the rare mutation burden ($\sigma^2 > V_A$). Reduced overdispersion in the residualized mutation burden compared to the raw mutation burden shows that a proportion of the overdispersion can be explained by geographic and technical covariates in our regression model.

Simulations for a finite population with realistic demography

We used SLiM 2.0 (*54*) to conduct forward population genetics simulations with realistic demography. We ran 100 replicates each of African and European populations modeled as per the demography published in Tennessen et al (*55*). Each replicate was started with a population (N = 14474) that had been burned in for 40,000 generations, after which the Tennessen et al demography was applied as follows:

1. Out of Africa bottleneck starting 2040 generations ago, shrinking European population size N to 1861.
2. Second European bottleneck starting 920 generations ago shrinking N to 1032, followed immediately by exponential growth in Europeans at 0.307%.
3. Explosive growth at 1.95% in Europeans and 1.66% in Africans starting 204 generations ago.

Africans (N = 11,754) and Europeans (N = 68,858) were sampled separately for mutation burden analysis. All simulations had a length of 1 Mb, mutation rate of 10^{-8} per generation per base pair, and recombination rate of 10^{-5} per generation per base pair. The high recombination rate was chosen to simulate largely unlinked sites. Strength of selection acting on deleterious alleles was varied between -10^{-1}, -10^{-2}, -10^{-3}, -10^{-4}, and -10^{-5}. Alleles were assumed to be additive (h = 0.5).

Mutation burden was computed on singletons for each selection coefficient, and σ^2/V_A was calculated in African and European samples separately (Fig. S10). Population structure leading to inbreeding ('heterogeneous demography') was modeled by combining the African and European samples together before performing mutation burden analysis (Fig. 2a).

The generation at which each mutation arose was used to stratify alleles by age and the analysis was repeated in two separate age brackets (Fig. S11). We verified in

simulations that, although synonymous alleles may be older than LoF alleles at the same allele frequency, our statistic gives almost identical values for old and new alleles if analyzed separately (Fig. S11). Unsurprisingly, there is larger variance for older alleles but the effect is quantitatively almost negligible.

Analytics for an infinite population with synergistic epistasis

We used theoretical estimates derived by Charlesworth (*13*) for σ^2 and μ of the rare mutation burden in an infinite population under multiplicative and epistatic selection. These are reproduced below for $V = \sigma^2$ and μ in the case where $w(x) = e^{-\alpha x - \frac{1}{2}\beta x^2}$ is the fitness function for an individual carrying x mutations, with $\alpha = hs$ and $\beta = 2h^2 b$, for selection coefficient (s), dominance coefficient (h) and pairwise epistasis coefficient (b). Given the genomic deleterious mutation rate per generation U,

$$\mu = \frac{U - V(\alpha - \beta U)}{\beta V}$$

$$f(V) = V^3\beta^2\left(1 + \frac{Z}{2}\right) + V^2\beta(1 + \alpha - \beta U) - V(2\beta U - \alpha) - U = 0$$

where $Z = E(\frac{1}{r_{i,j} + 2hs})$ and $r_{i,j} = recombination\ frequency\ between\ locus\ i\ and\ j$

We solved these equations in Matlab for $r_{i,j} = 0.5$ (free recombination) and plotted σ^2/V_A ($\mu \approx V_A$ for rare mutations) at the mutation-selection balance, under multiplicative selection ($\beta = 0$, Fig S1), and epistatic selection for a range of α and β values (Fig. S2). We also note that the quadratic model is only one possible model of synergistic epistasis (another model, for example, is truncation selection) and we present Figure S2 to provide intuition rather than as a way to estimate parameters.

Truncation selection represents the extreme mode of synergistic epistasis (*4*) and leads to the smallest ρ, where ρ describes the factor reduction in the variance of the mutation burden due to dependencies between independent alleles. For example, if 50% of individuals with above average numbers of mutations would produce no offspring, ρ would be 0.36 under a normal approximation if the average genomic number of mutations is high. Because free recombination halves LD within a single generation, at the mutation-selection equilibrium, we should expect $\sigma^2 = V_A/(2 - \rho)$, where V_A is the variance of the mutation burden under linkage equilibrium. Thus our observed reduction in variance ($\sigma^2/V_A \sim 0.9$) is consistent with that calculated ($\rho = 0.89$) for a truncation of less than 2% of the population.

Supplementary Figures

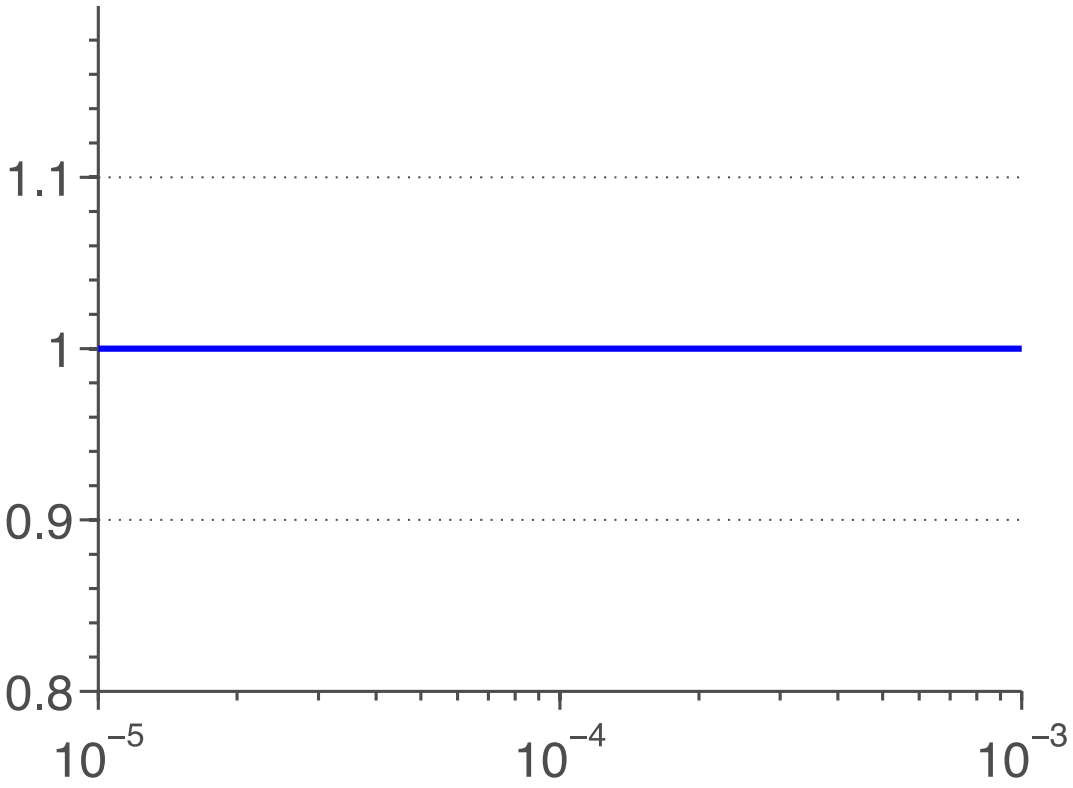


Fig. S1. Predictions for rare mutation burden under multiplicative selection.

Values of σ^2/V_A are shown as a function of the strength of selection (s) under free recombination (r = 0.5). Analytical solutions (*blue*) for σ^2 and μ $(\approx V_A)$ of the rare mutation burden in an infinite population under multiplicative selection are derived in previous work (*16*) for the fitness function $w(x) = e^{-\alpha x - \frac{1}{2}\beta x^2}$ for an individual carrying x mutations with $\alpha = hs$ and $\beta = 2h^2 b$ given selection coefficient (s), dominance coefficient (h) and pairwise epistasis coefficient (b). Here, alleles are assumed to be additive $(h = 0.5)$, and to have negligible epistasis $(b = 10^{-15})$.

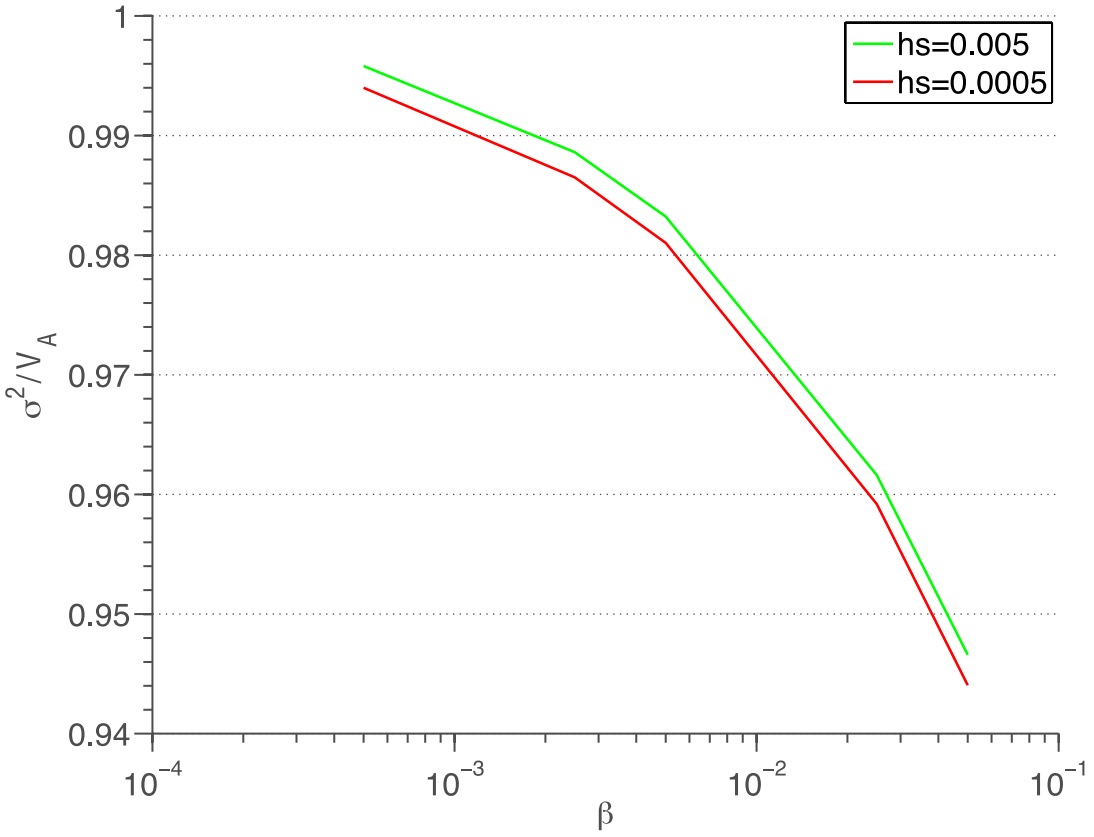


Fig. S2. Predictions for rare mutation burden under a quadratic model of synergistic epistasis.
Values of σ^2/V_A are shown as a function of the extent of epistasis (β) under free recombination (r = 0.5) for different selection strengths (hs). Analytical solutions for σ^2 and μ ($\approx V_A$) of the rare mutation burden in an infinite population under a quadratic model of synergistic epistasis are derived in previous work (*16*) for the fitness function $w(x) = e^{-\alpha x - \frac{1}{2}\beta x^2}$ for an individual carrying x mutations with $\alpha = hs$ and $\beta = 2h^2b$ given selection coefficient (s), dominance coefficient (h) and pairwise epistasis coefficient (b). These are solved here for U = 0.08 per genome per generation for the class of LoF (nonsense and splice-disrupting) mutations (*56*). Alleles are assumed to be additive ($h = 0.5$). We note that the quadratic model is only one possible model of synergistic epistasis (another model, for example, is truncation selection) and we present this plot to provide intuition rather than as a way to estimate parameters.

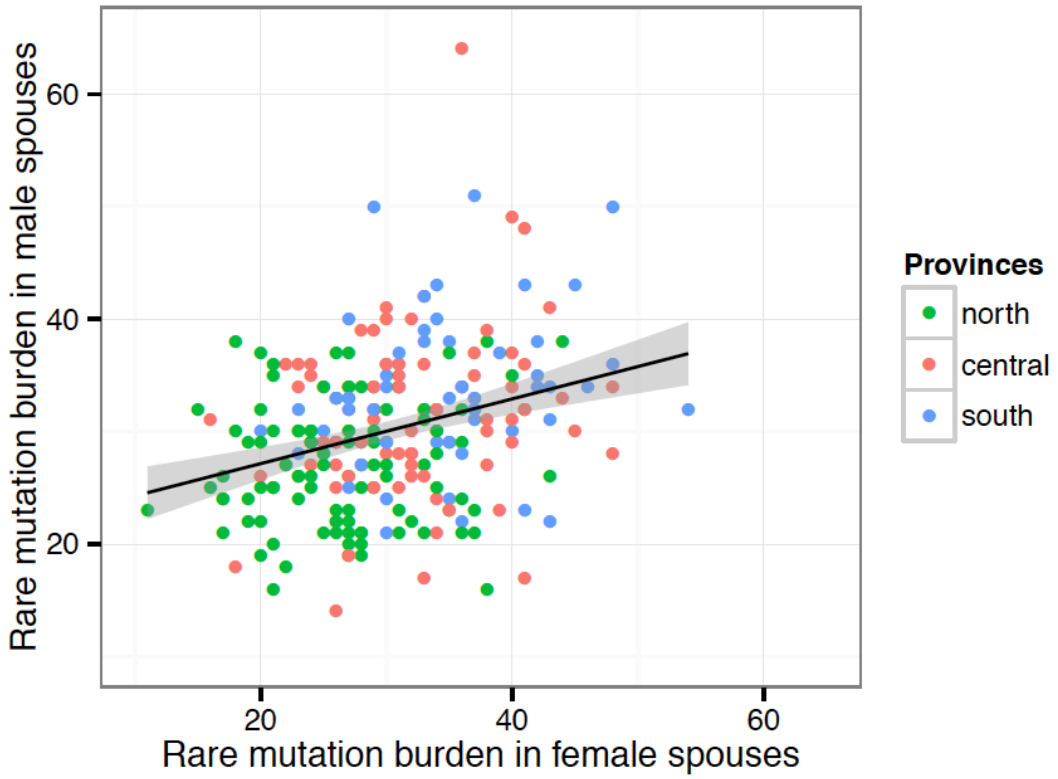


Fig. S3. Rare mutation burden for spousal pairs in the GoNL dataset.

Synonymous singletons were used to compute mutation burden for every sample in the GoNL dataset. Spouses show a positive correlation in rare mutation burden (Pearson's r = 0.31). Dutch provinces are divided into 3 regions – north, central and south (*15*). Each point is colored by the region of origin of the male spouse (only 1 pair has the two spouses originating from different regions). Mean rare mutation burden varies between the three regions ($\mu_{north} = 26.88$, $\mu_{central} = 31.98$, $\mu_{south} = 33.74$).

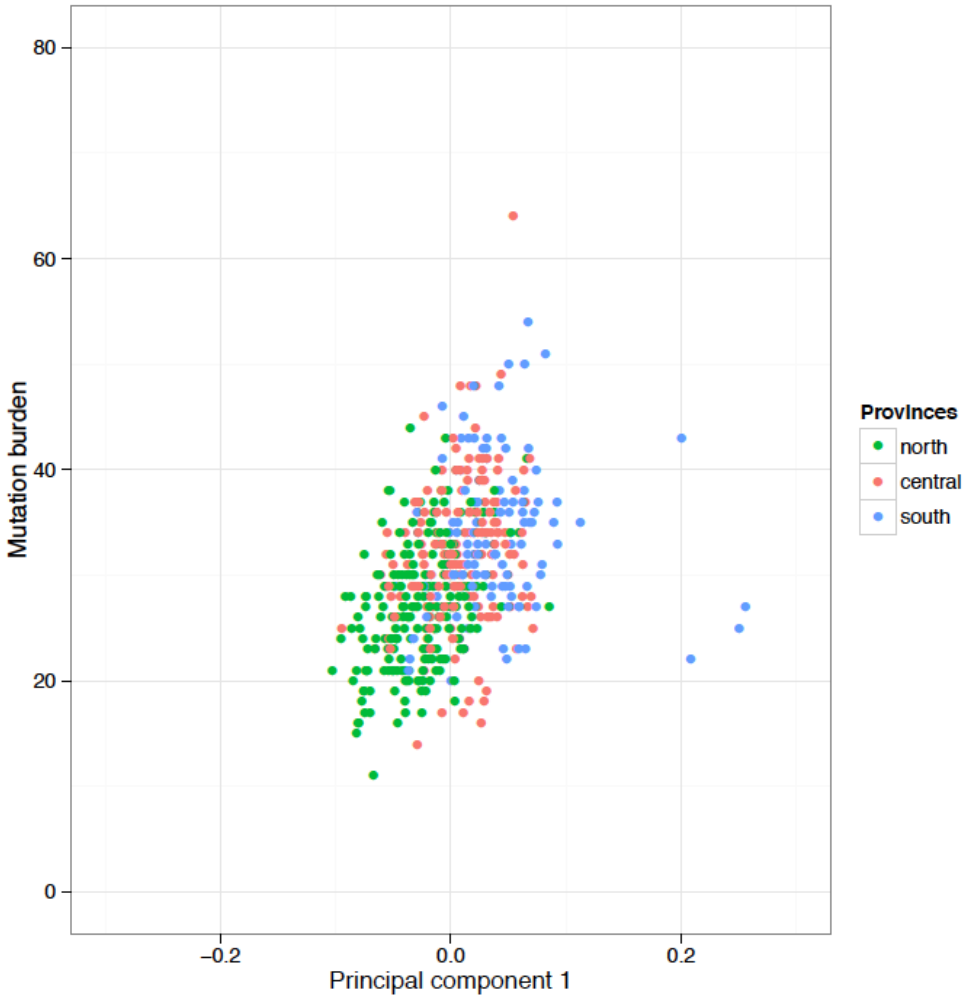


Fig. S4. Rare mutation burden in the GoNL dataset reflects geographic structure.

Synonymous singletons were used to compute mutation burden for every sample in the GoNL dataset. Mutation burden shows a positive correlation, along a south-north cline, with the first principal component (see methods for details) computed on the GoNL dataset (Pearson's r = 0.4).

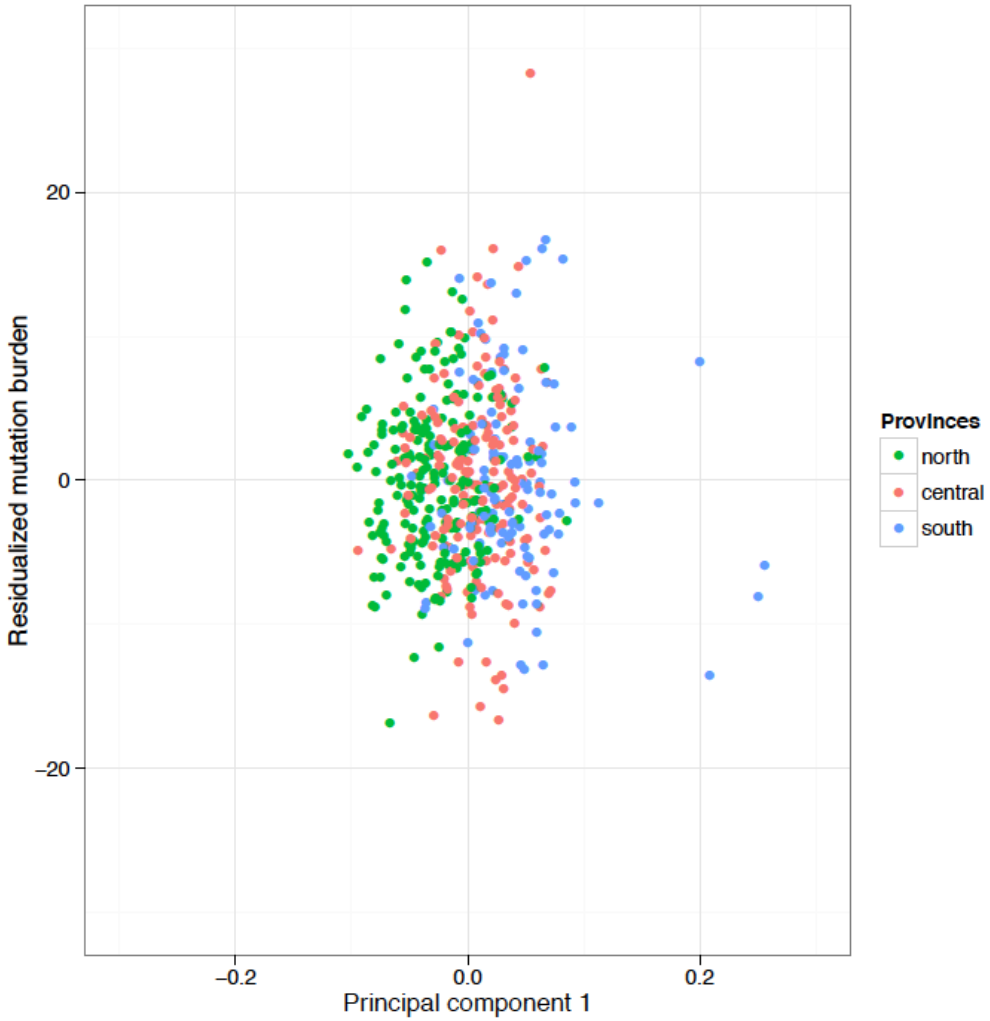
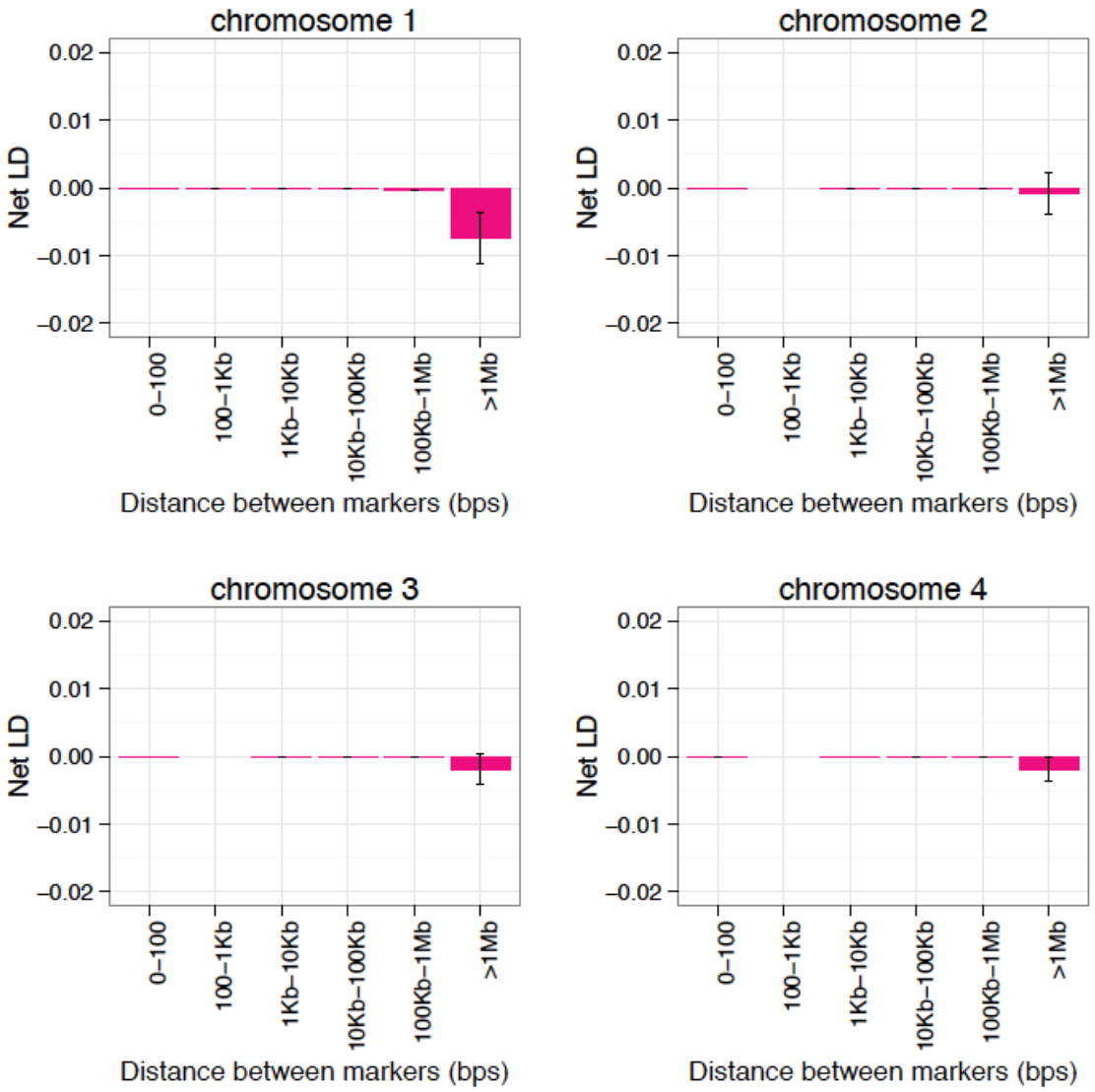
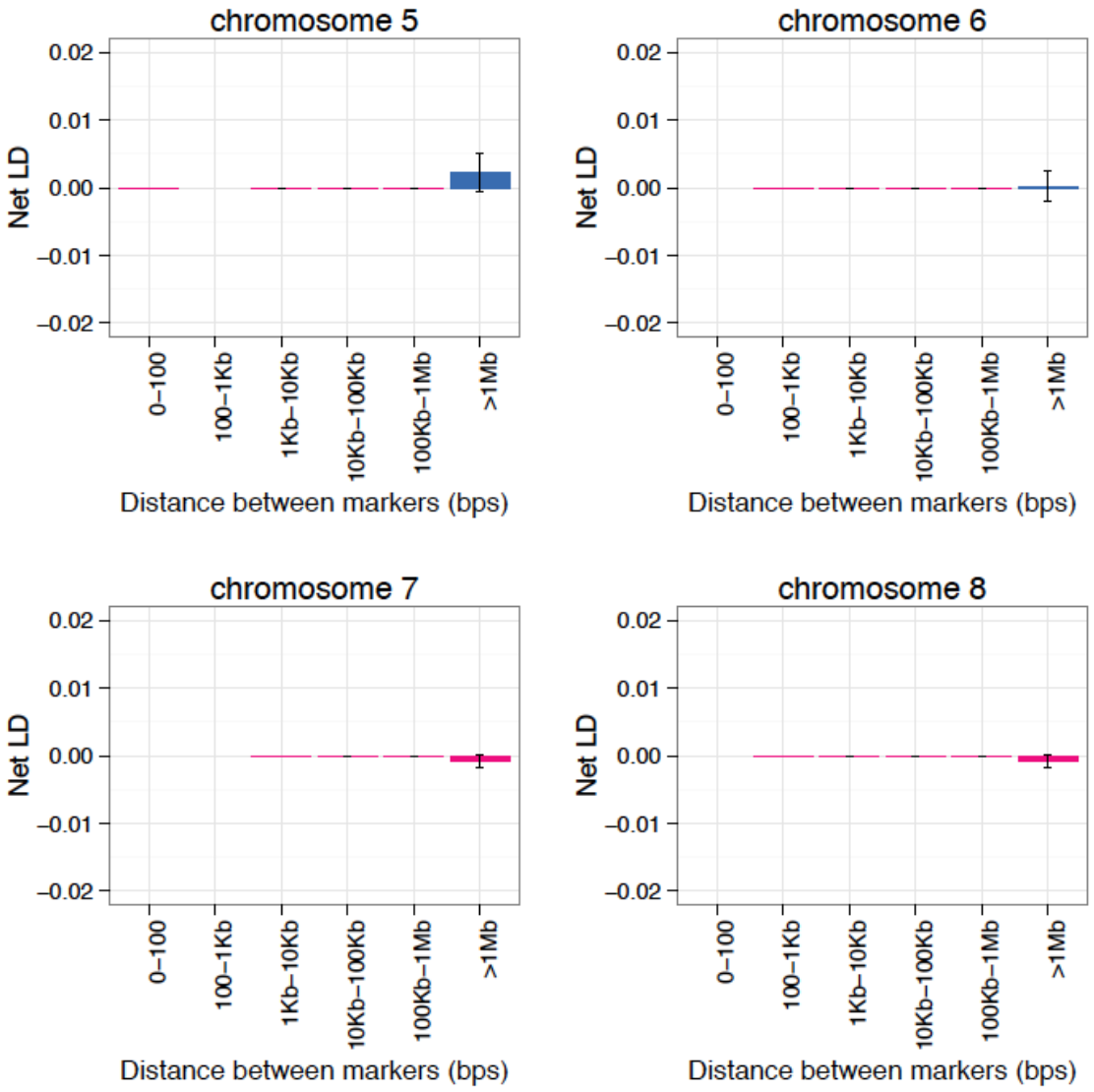
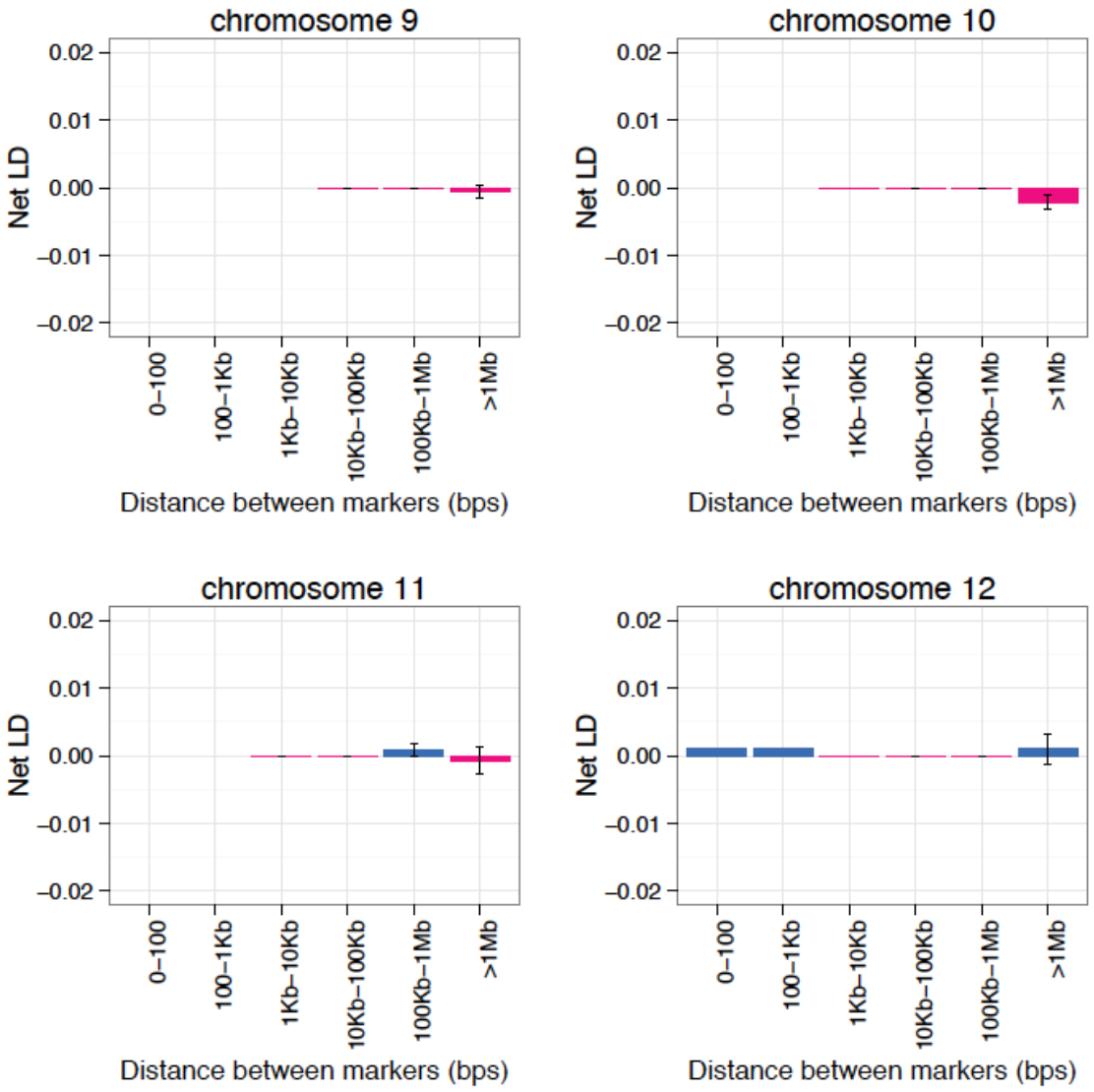


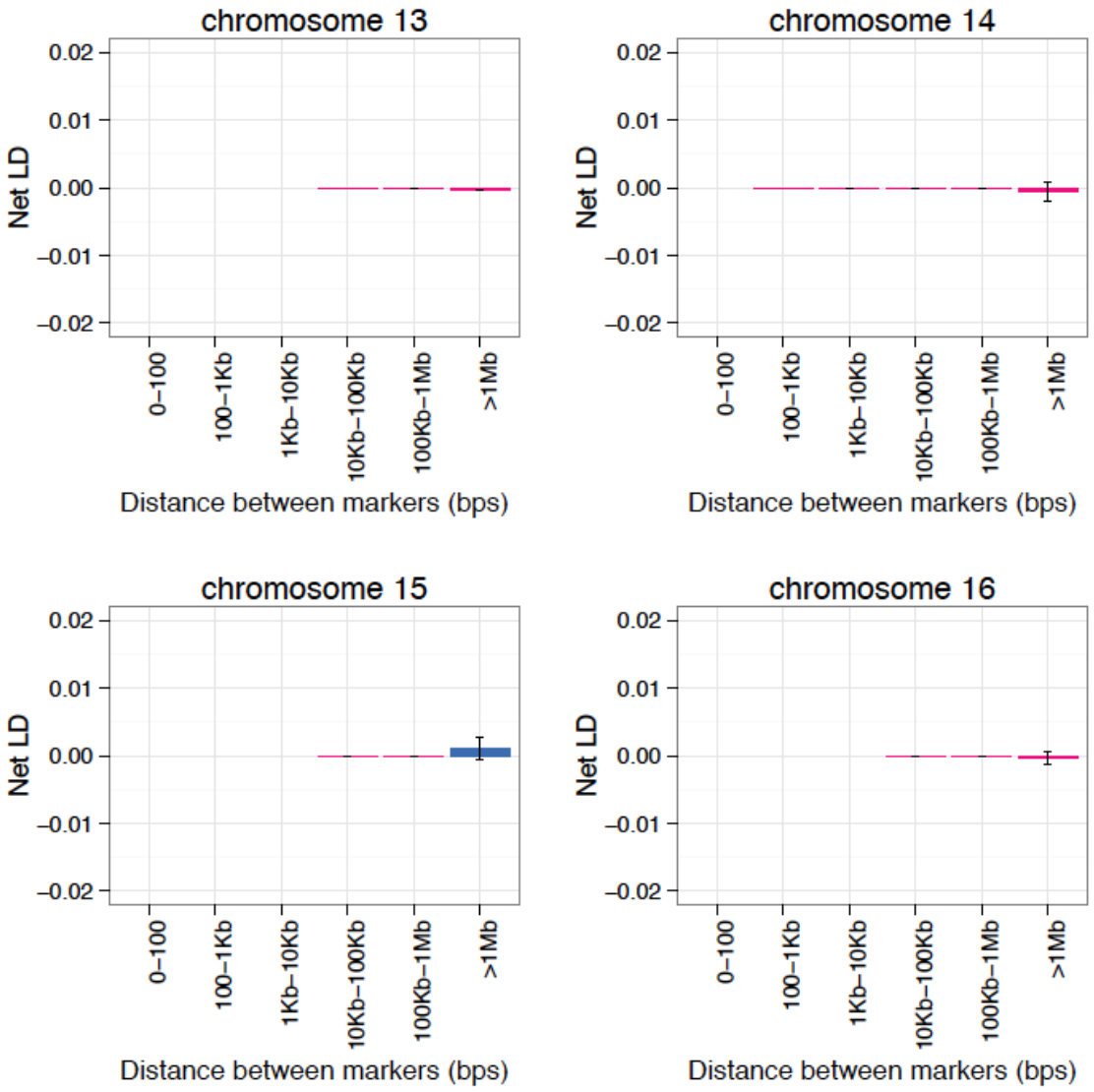
Fig. S5. Residuals for rare mutation burden in the GoNL dataset.

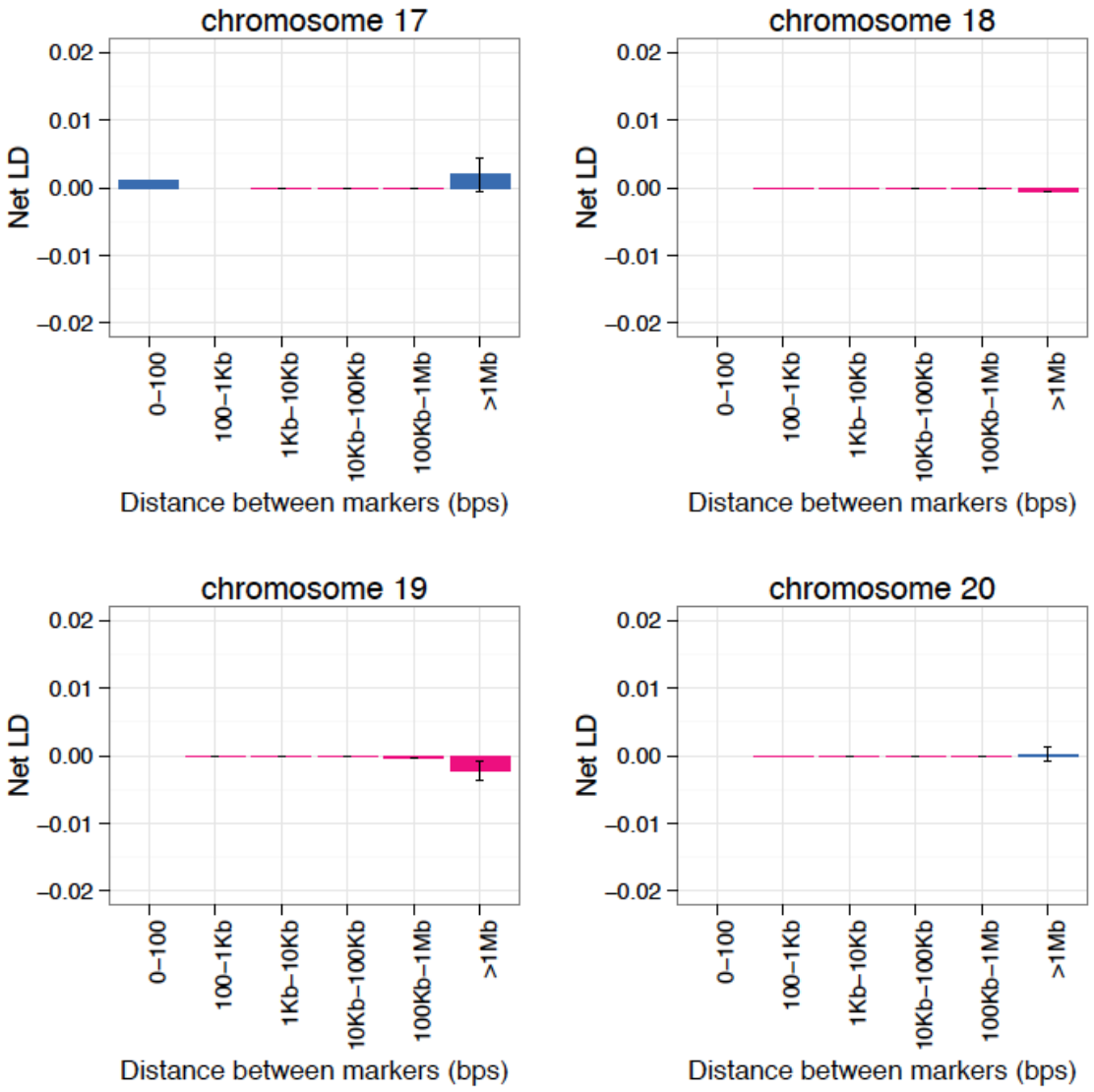
Synonymous singletons were used to compute mutation burden for every sample in the GoNL dataset. Mutation burden was residualized under a generalized linear model consisting of 10 principal components and other covariates for geographic structure (see methods for model details). Residualized mutation burden does not show a positive correlation with the first principal component computed on the GoNL dataset (Pearson's r = -5.59×10^{-17}).


chromosome 1
chromosome 2
chromosome 3
chromosome 4
Net LD
0.02
0.01
0.00
−0.01
−0.02
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 5
chromosome 6
chromosome 7
chromosome 8
Net LD
0.02
0.01
0.00
−0.01
−0.02
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 9
chromosome 10
chromosome 11
chromosome 12
Net LD
0.02
0.01
0.00
−0.01
−0.02
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 13
chromosome 14
chromosome 15
chromosome 16
Net LD
0.02
0.01
0.00
−0.01
−0.02
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 17
chromosome 18
chromosome 19
chromosome 20
Net LD
0.02
0.01
0.00
−0.01
−0.02
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)

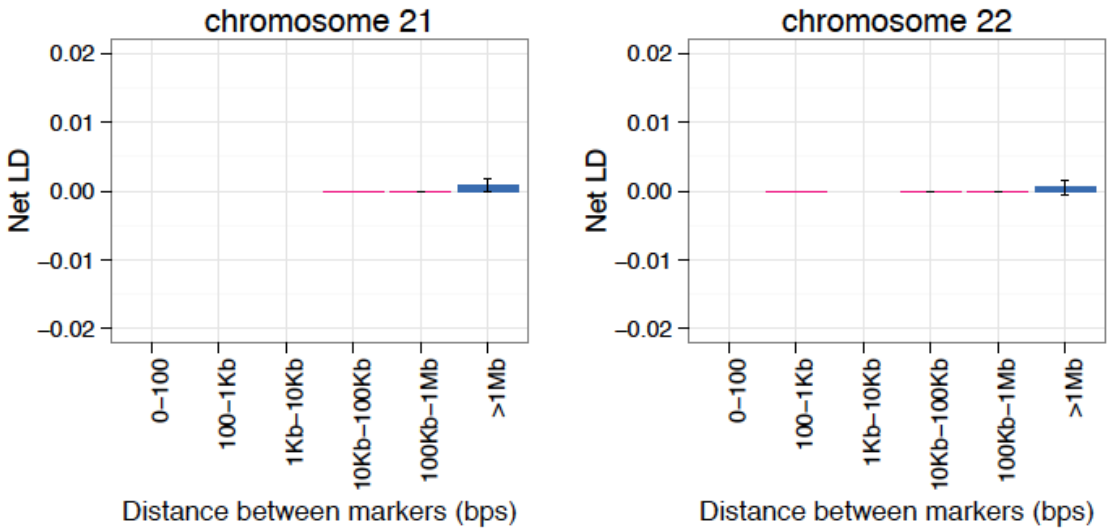
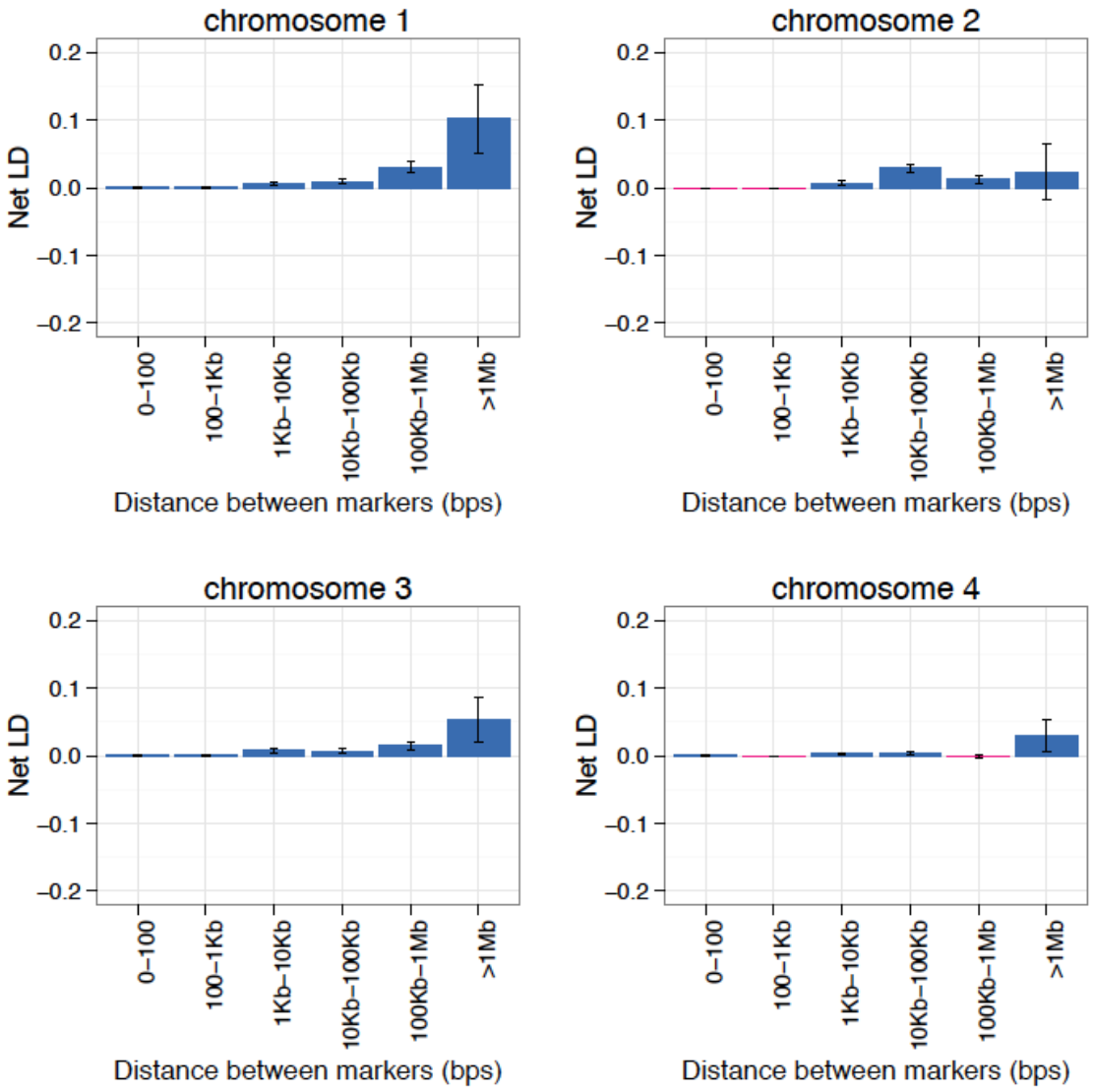
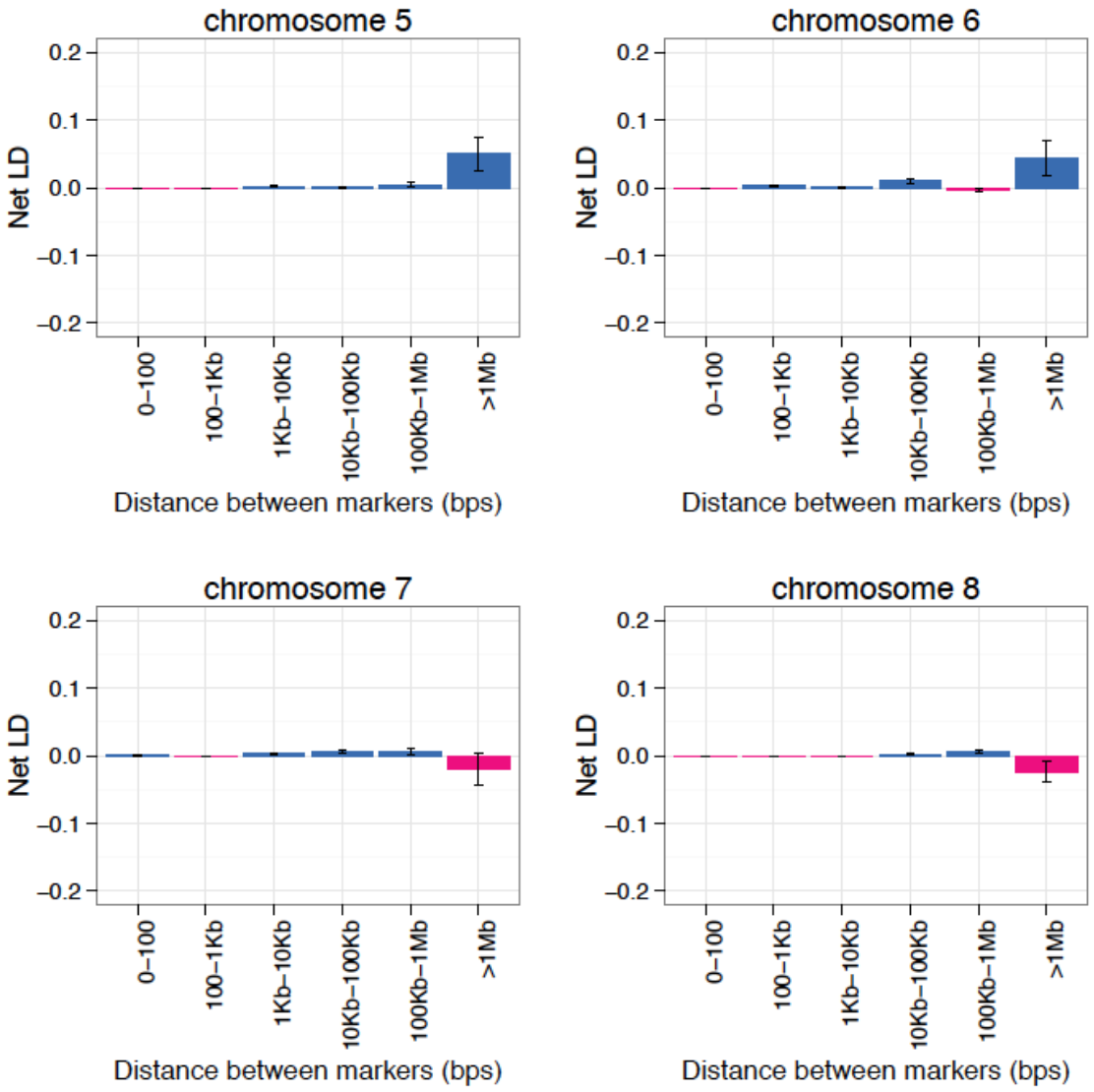
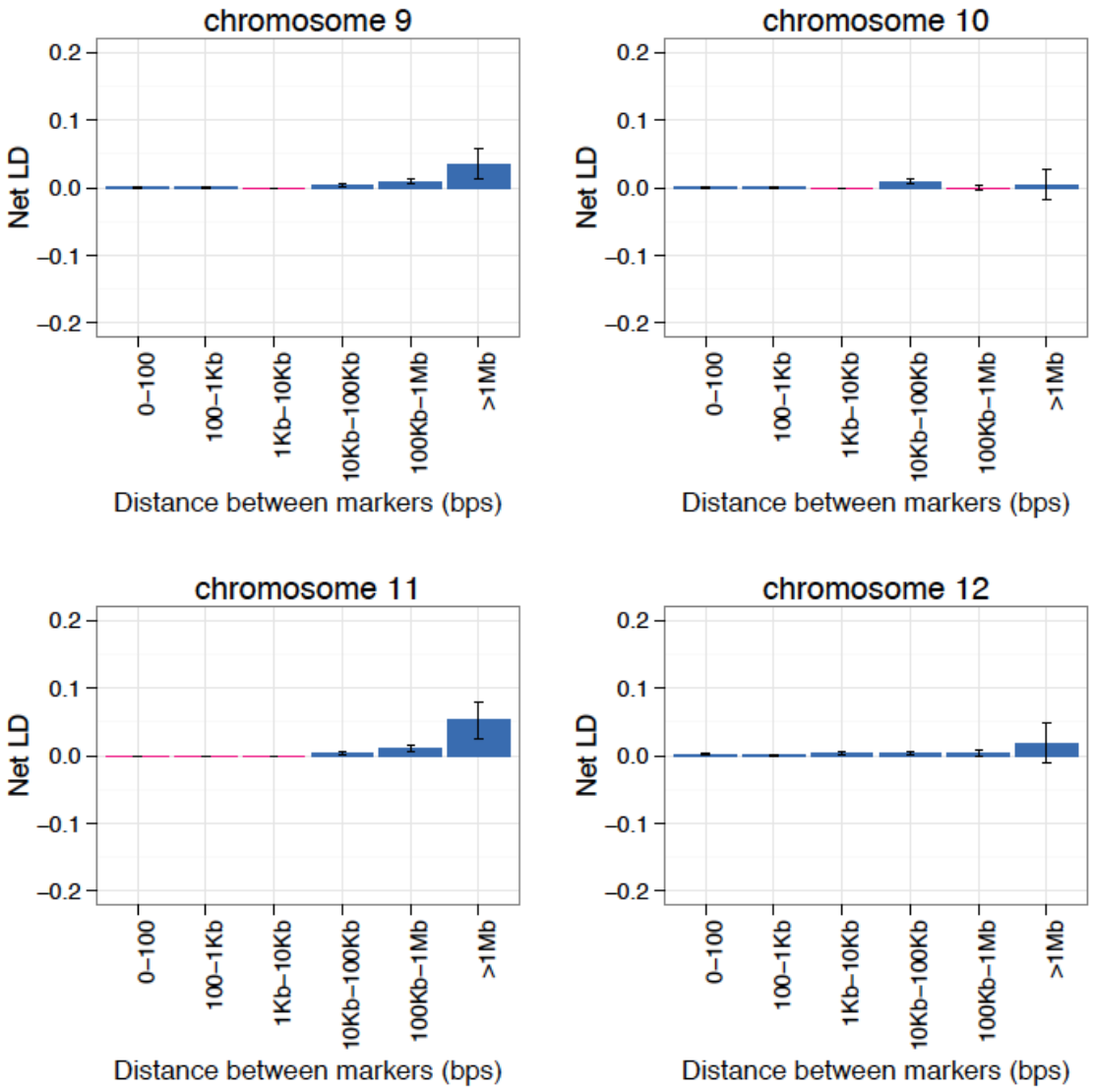


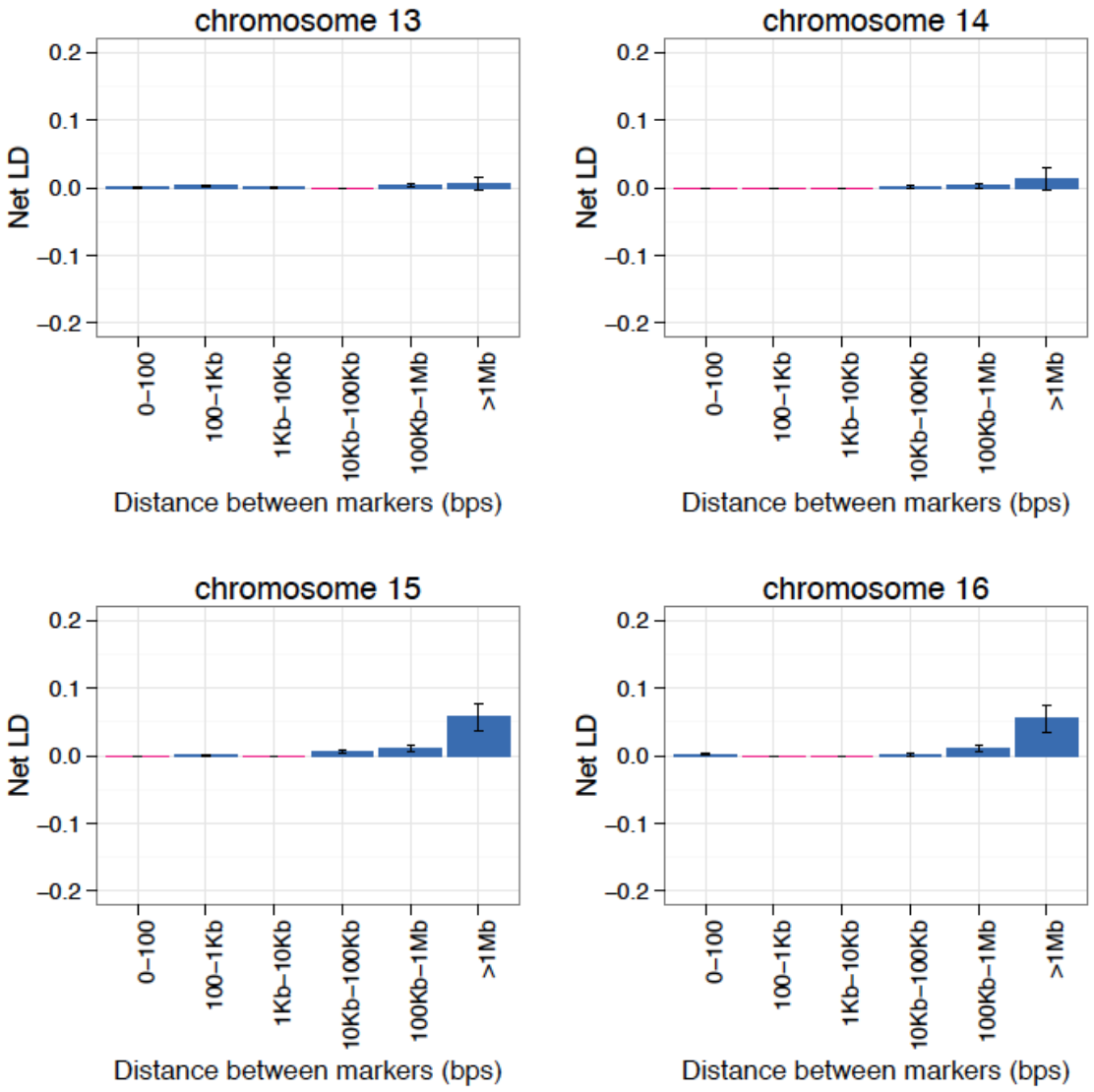
Fig. S6. Net LD as a function of physical distance between rare deleterious alleles.

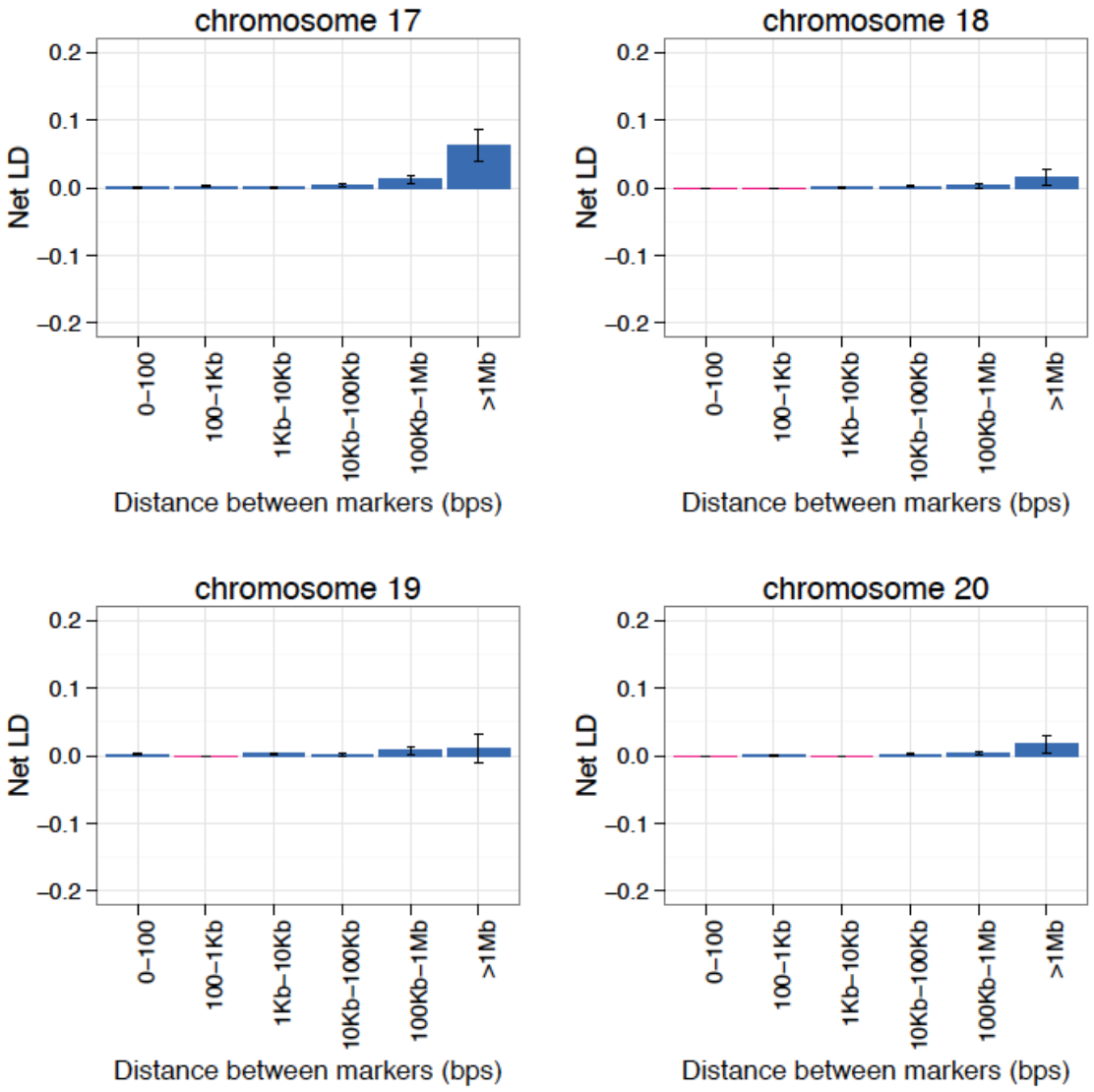
Data are shown for intra-chromosomal interactions between splice-disrupting and stop gain singletons in twenty-two autosomal chromosomes in the GoNL dataset. Net LD values are shown binned by physical distance between loci. Net LD in a given bin can be positive (*blue)* or negative (*pink).* Error bars show standard errors.


chromosome 1
chromosome 2
chromosome 3
chromosome 4
Net LD
0.2
0.1
0.0
−0.1
−0.2
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 5
chromosome 6
chromosome 7
chromosome 8
Net LD
0.2
0.1
0.0
−0.1
−0.2
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 9
chromosome 10
chromosome 11
chromosome 12
Net LD
0.2
0.1
0.0
−0.1
−0.2
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 13
chromosome 14
chromosome 15
chromosome 16
Net LD
0.2
0.1
0.0
−0.1
−0.2
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)


chromosome 17
chromosome 18
chromosome 19
chromosome 20
Net LD
0.2
0.1
0.0
−0.1
−0.2
0–100
100–1Kb
1Kb–10Kb
10Kb–100Kb
100Kb–1Mb
>1Mb
Distance between markers (bps)

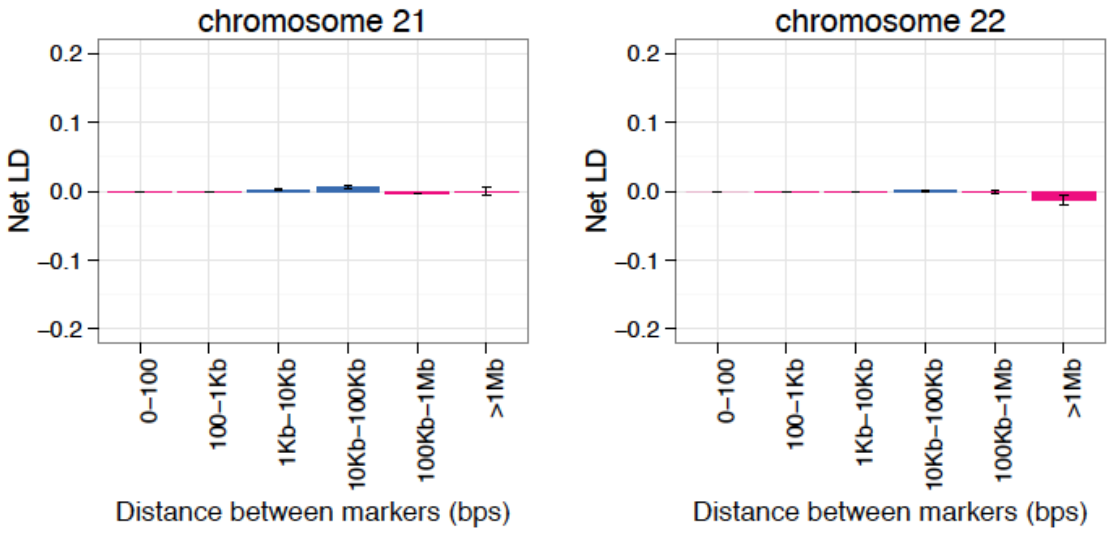


Fig. S7. Net LD as a function of physical distance between rare synonymous alleles.

Data are shown for intra-chromosomal interactions between synonymous singletons in twenty-two autosomal chromosomes in the GoNL dataset. Net LD values are shown binned by physical distance between loci. Net LD in a given bin can be positive (*blue)* or negative (*pink).* Error bars show standard errors.

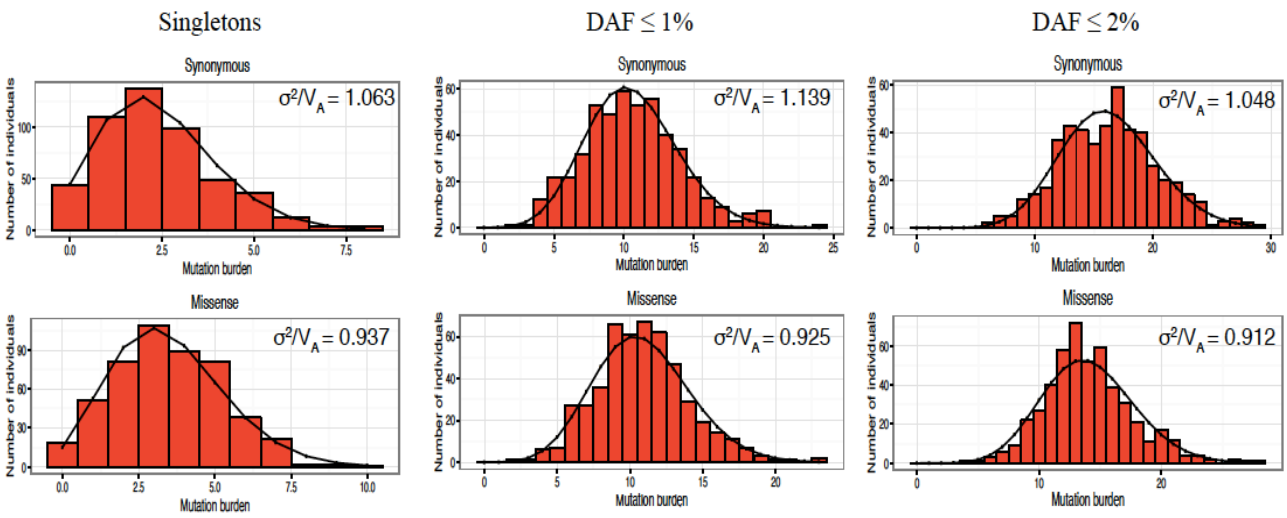


Fig. S8. Mutation burden in the crucial genome in humans.

Mutation burden was computed on synonymous and missense singletons and alleles with derived allele frequency up to 1% and 2% in the GoNL dataset. Only most selectively constrained genes were used for this analysis (estimated selection coefficient against heterozygous protein truncating variants exceeding 0.2)(*37*). Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*). P-values for σ^2/V_A were computed by permuting functional consequences across variants (see Table S22).

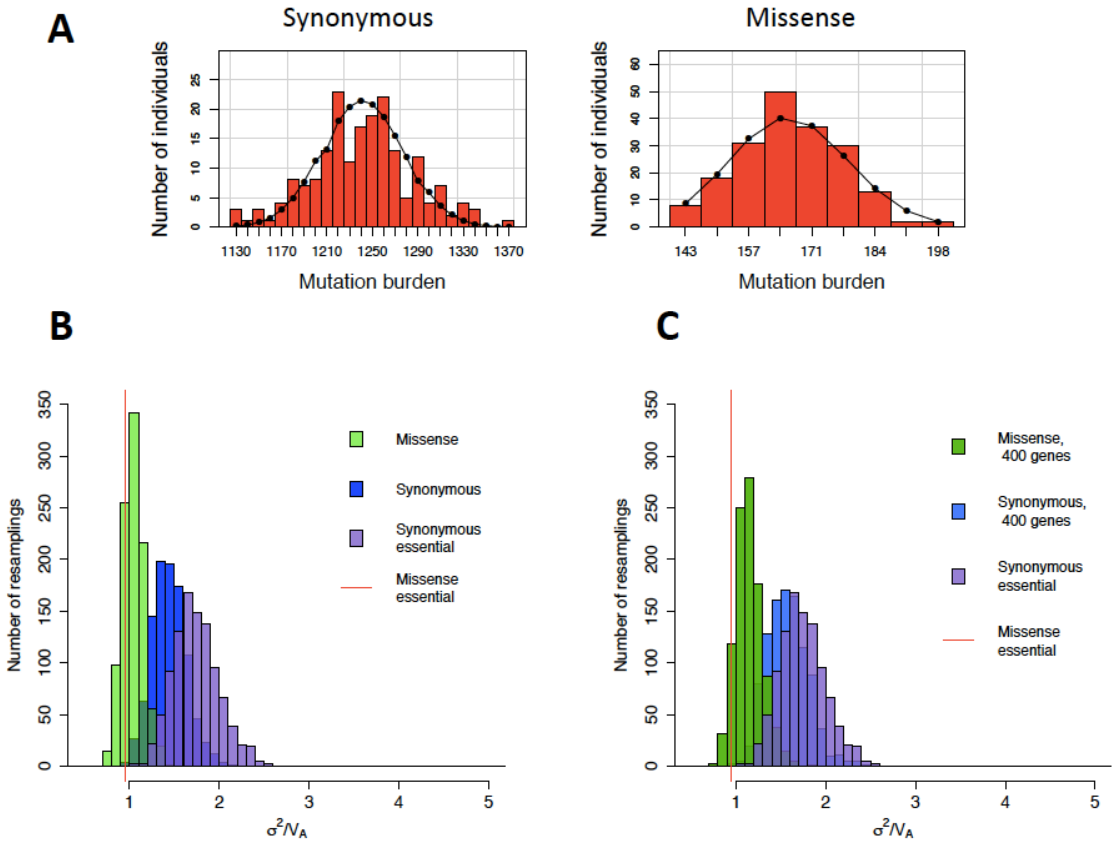


Fig. S9. Mutation burden in the essential genome in *D. melanogaster*.

Mutation burden was computed on alleles with minor allele frequency up to 50% in the DPGP3 dataset. Only set of 267 genes considered essential from the DEG database were used for this analysis. (**A**) Synonymous and missense rare mutation burden (*red*) in the *D. melanogaster* essential genome, overlaid with Poisson distributions (*black*) having identical means. (**B**, **C**) Resampling distributions of σ^2/V_A for missense alleles in the *D. melanogaster* essential genome. Synonymous and missense alleles were resampled at population frequencies matching the frequency distribution of missense alleles residing within the essential genes. Resampled datasets were used to obtain null distributions for σ^2/V_A. We generated 1000 resampled sets of synonymous and missense SNPs picking alleles at random in the *D. melanogaster* genome (**B**), and restricting the maximum number of genes in each set to 400 (**C**) to control for the effects of physical linkage between SNPs from the essential genes. Analogously we obtained resampled distributions of synonymous alleles randomly picked from the *D. melanogaster* essential genes (**B**, **C**). Underdispersion for missense alleles in the essential genome remained significant when compared to the resampled sets of synonymous alleles with ($p < 10^{-3}$) or without ($p = 0.002$) controlling for the gene number.

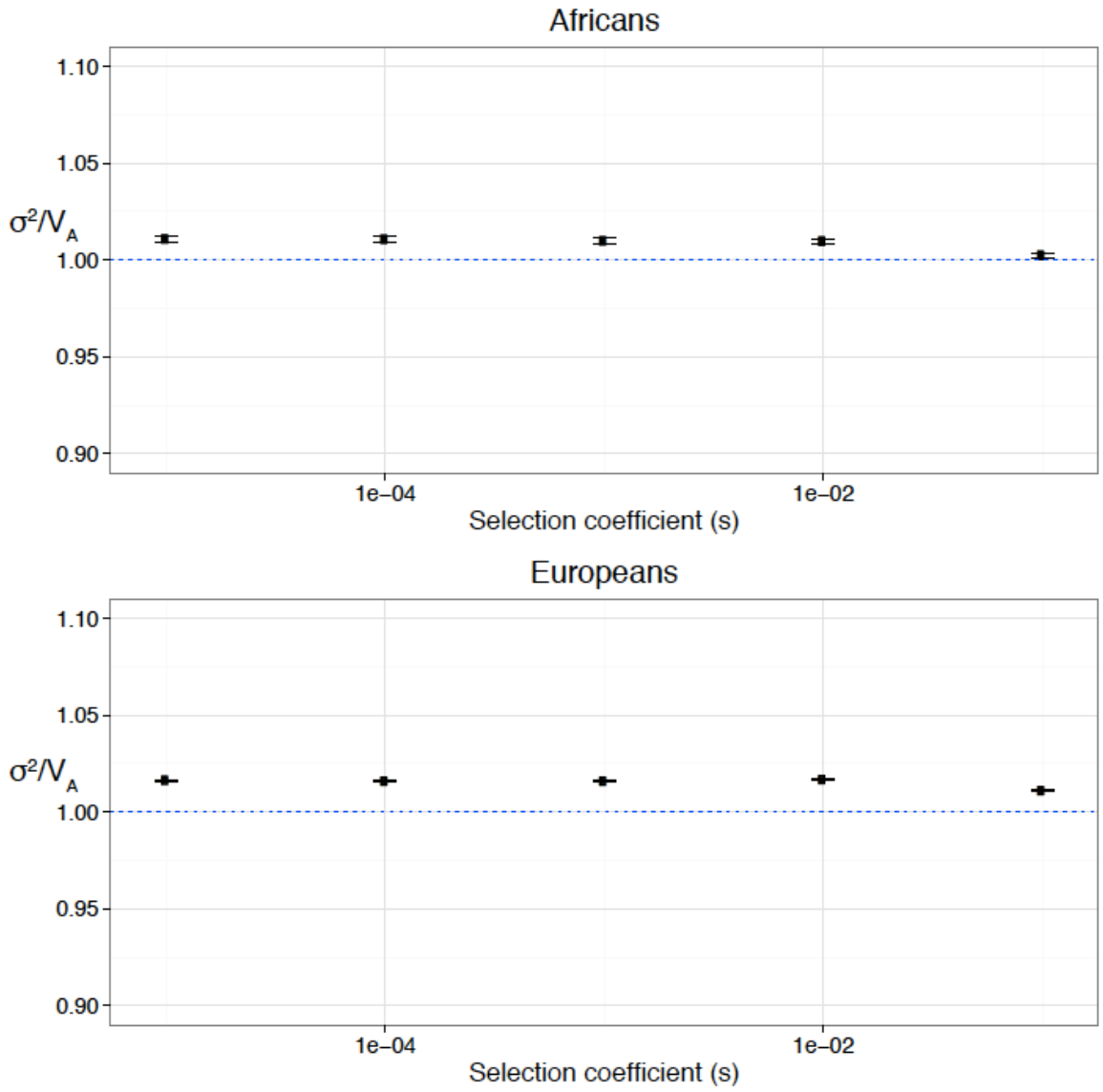


Fig. S10. Simulated mutation burden in African and European populations.

SLiM 2.0 was used to simulate a population with demography from Tennessen et al (*55*). σ^2/V_A of the mutation burden computed on singletons was calculated for Africans (*top*) and Europeans (*bottom*). A (*blue*) dotted line is drawn at the null expectation for a randomly mating population at equilibrium. All simulations had a length of 1 Mb, mutation rate of 10^{-8} per generation per base pair, and recombination rate of 10^{-5} per generation per base pair. The high recombination rate was chosen to simulate largely unlinked sites. Strength of selection acting on deleterious alleles was varied between -10^{-1}, -10^{-2}, -10^{-3}, -10^{-4}, and -10^{-5}. Alleles were assumed to be additive (h = 0.5).

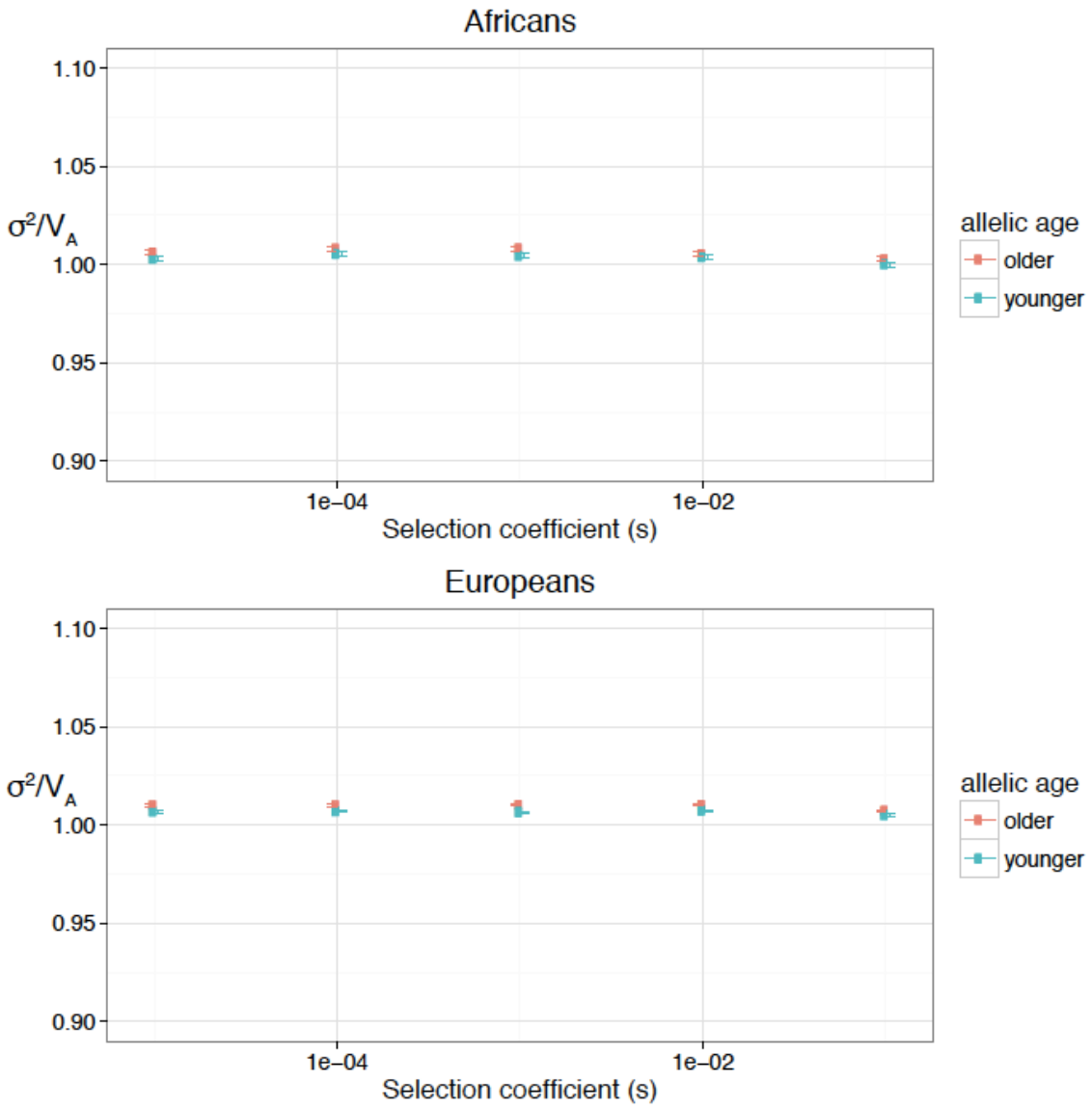


Fig. S11. Simulated mutation burden for alleles of different ages in African and European populations.
SLiM 2.0 was used to simulate a population with demography from Tennessen et al (*55*). σ^2/V_A of the mutation burden computed on singletons was calculated for alleles segregated by their age into two groups (older, younger) in Africans (*top*) and Europeans (*bottom*). All simulations had a length of 1 Mb, mutation rate of 10^{-8} per generation per base pair, and recombination rate of 10^{-5} per generation per base pair. The high recombination rate was chosen to simulate largely unlinked sites. Strength of selection acting on deleterious alleles was varied between -10^{-1}, -10^{-2}, -10^{-3}, -10^{-4}, and -10^{-5}. Alleles were assumed to be additive (h = 0.5).

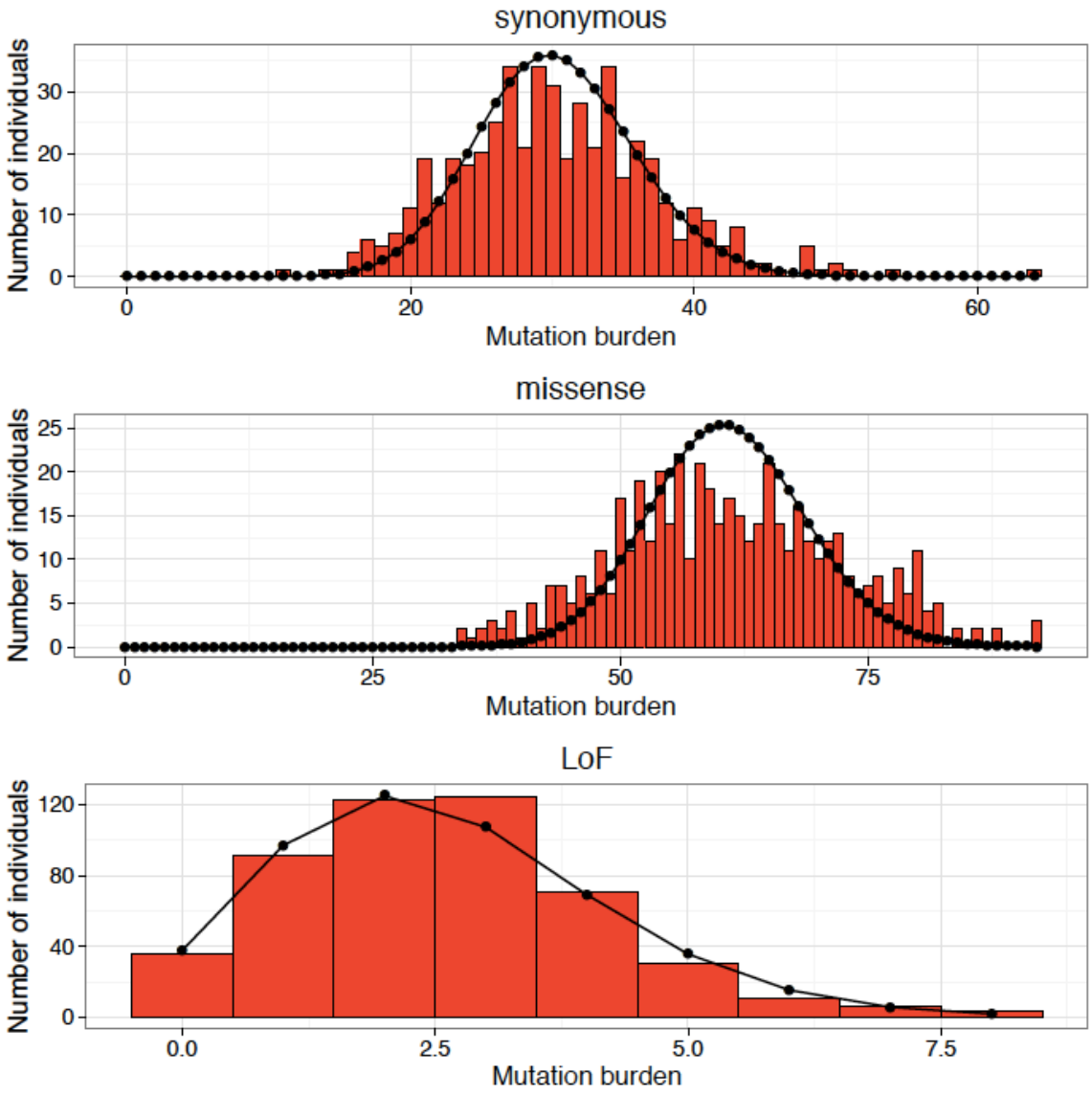


Fig. S12. Mutation burden in the GoNL dataset overlaid with Poisson distributions having identical means.
Mutation burden was computed on synonymous, missense and LoF singletons in the GoNL dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*).

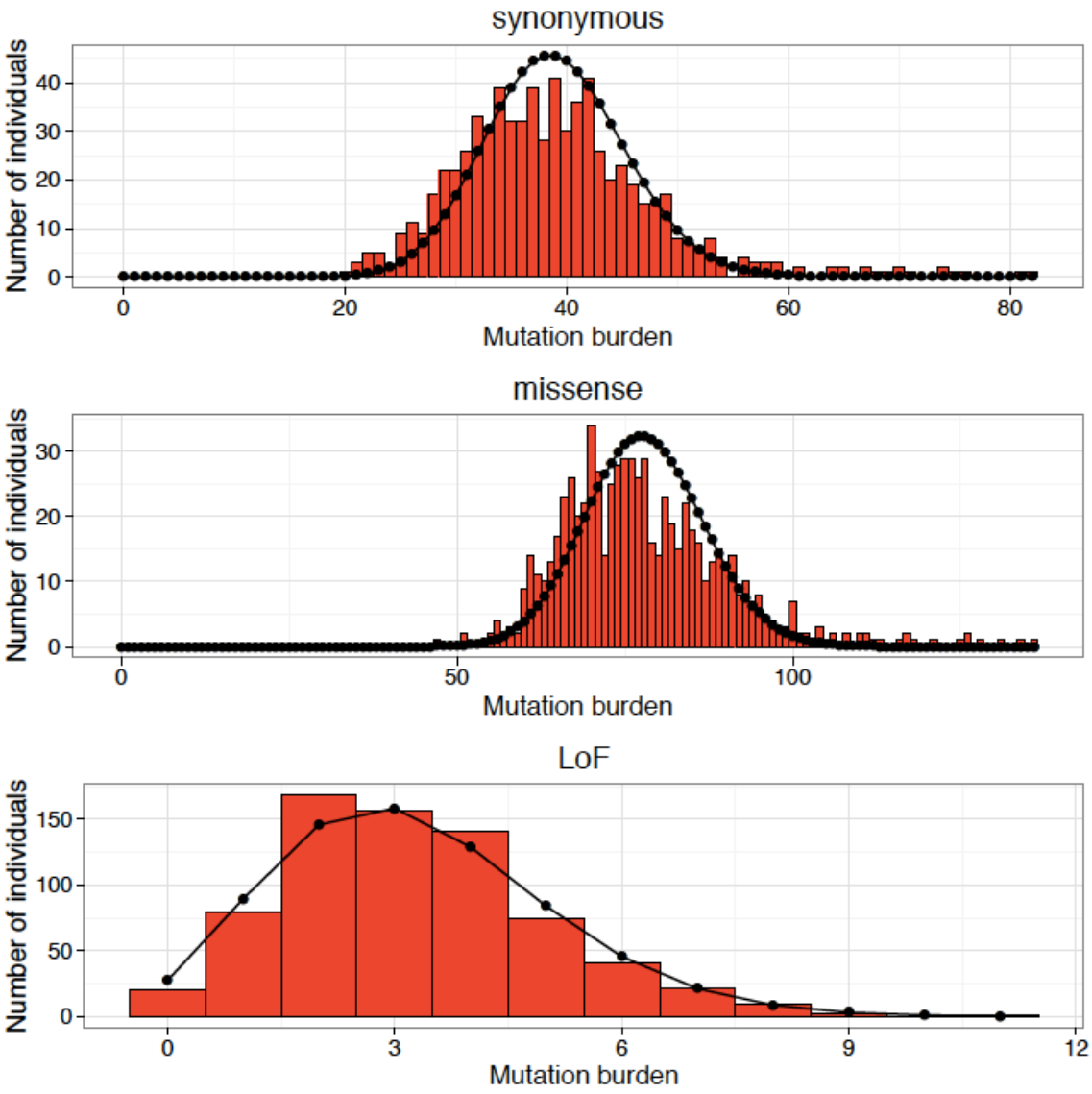


Fig. S13. Mutation burden in the ADNI dataset overlaid with Poisson distributions having identical means.
Mutation burden was computed on synonymous, missense and LoF singletons in the ADNI dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*).

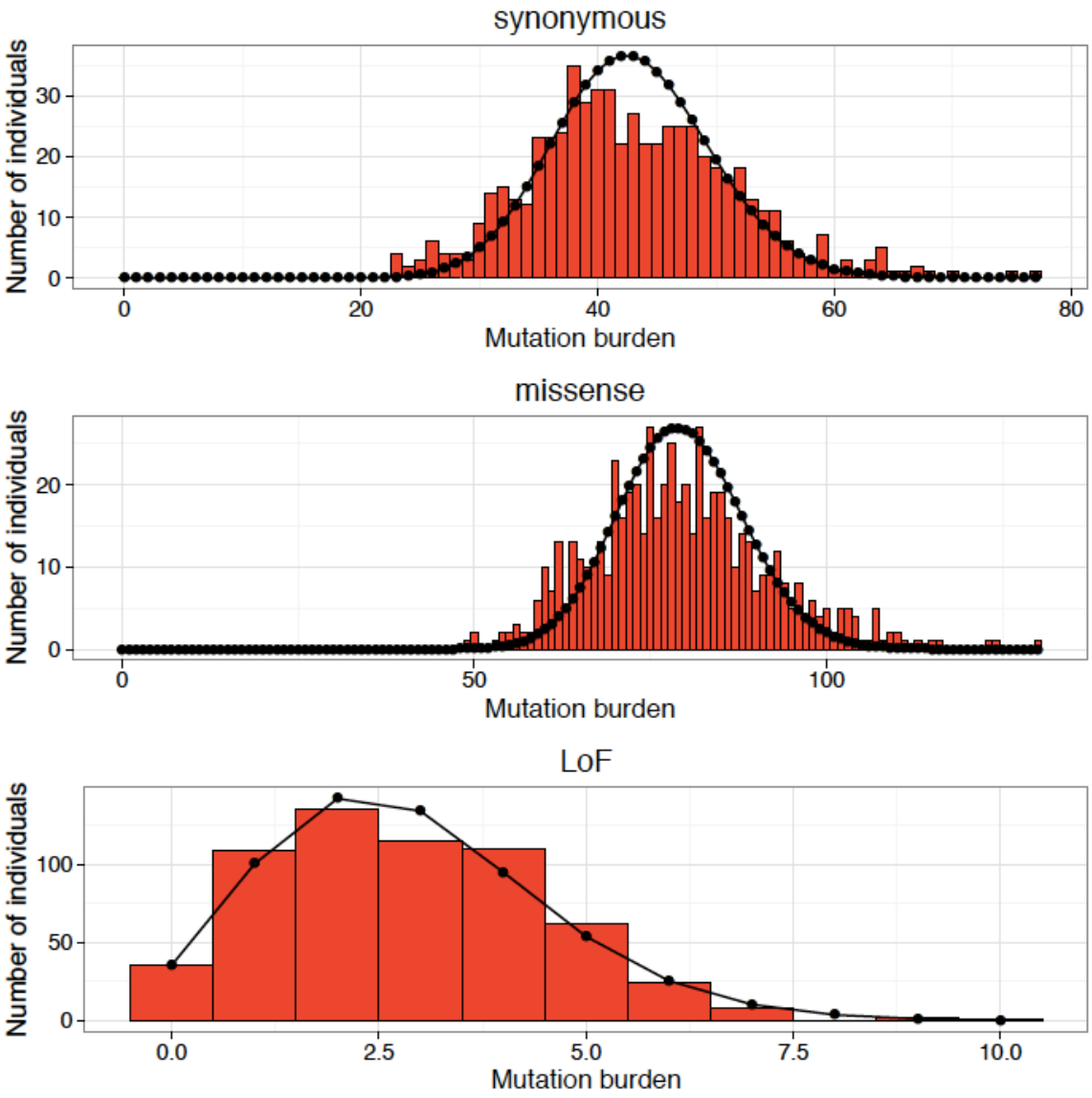


Fig. S14. Mutation burden in the MinE dataset overlaid with Poisson distributions having identical means.
Mutation burden was computed on synonymous, missense and LoF singletons in the MinE dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*).

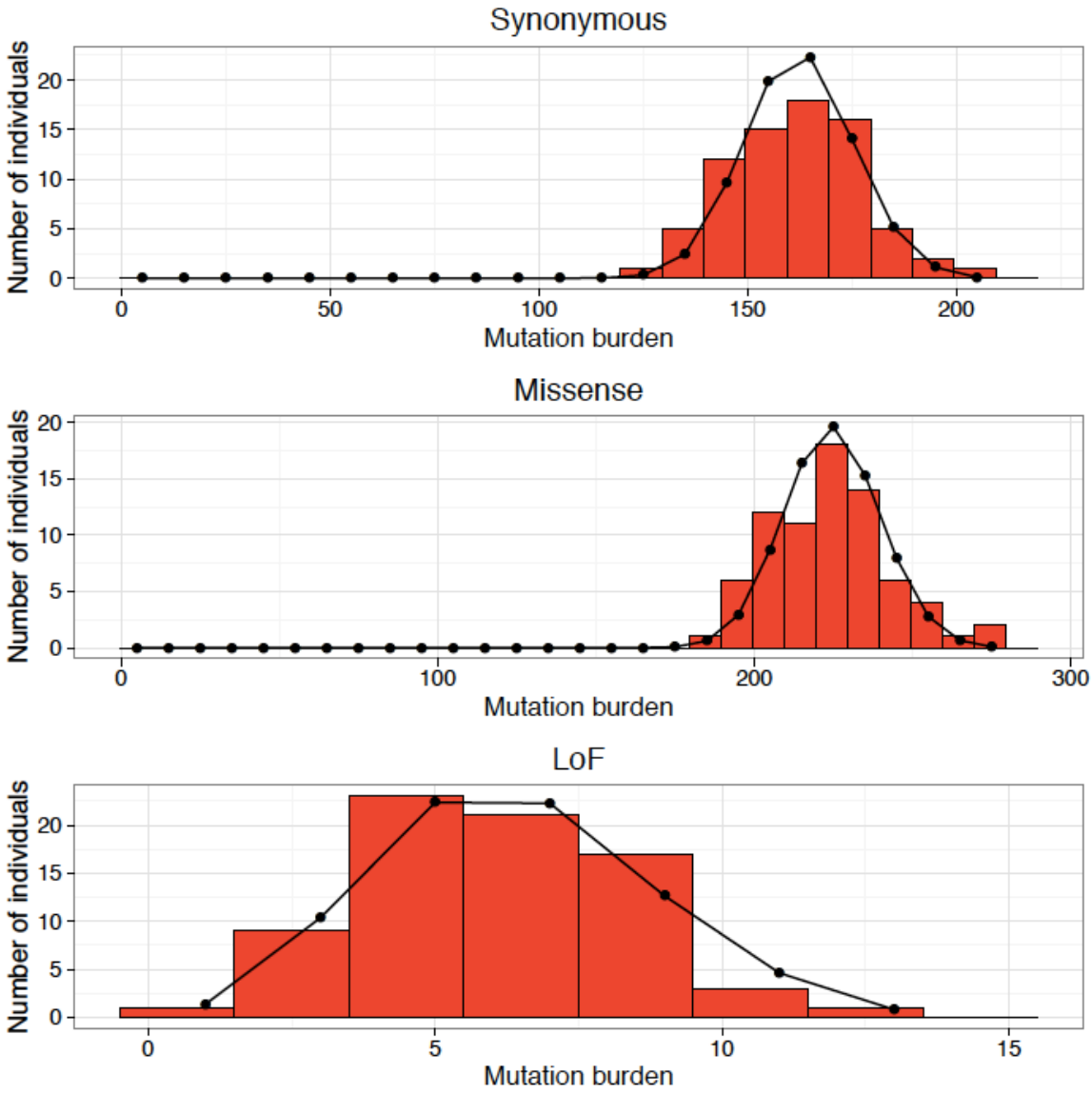


Fig. S15. Mutation burden in the 1000 Genomes YRI dataset overlaid with Poisson distributions having identical means.

Mutation burden was computed on synonymous, missense and LoF singletons in the 1000 Genomes YRI dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*).

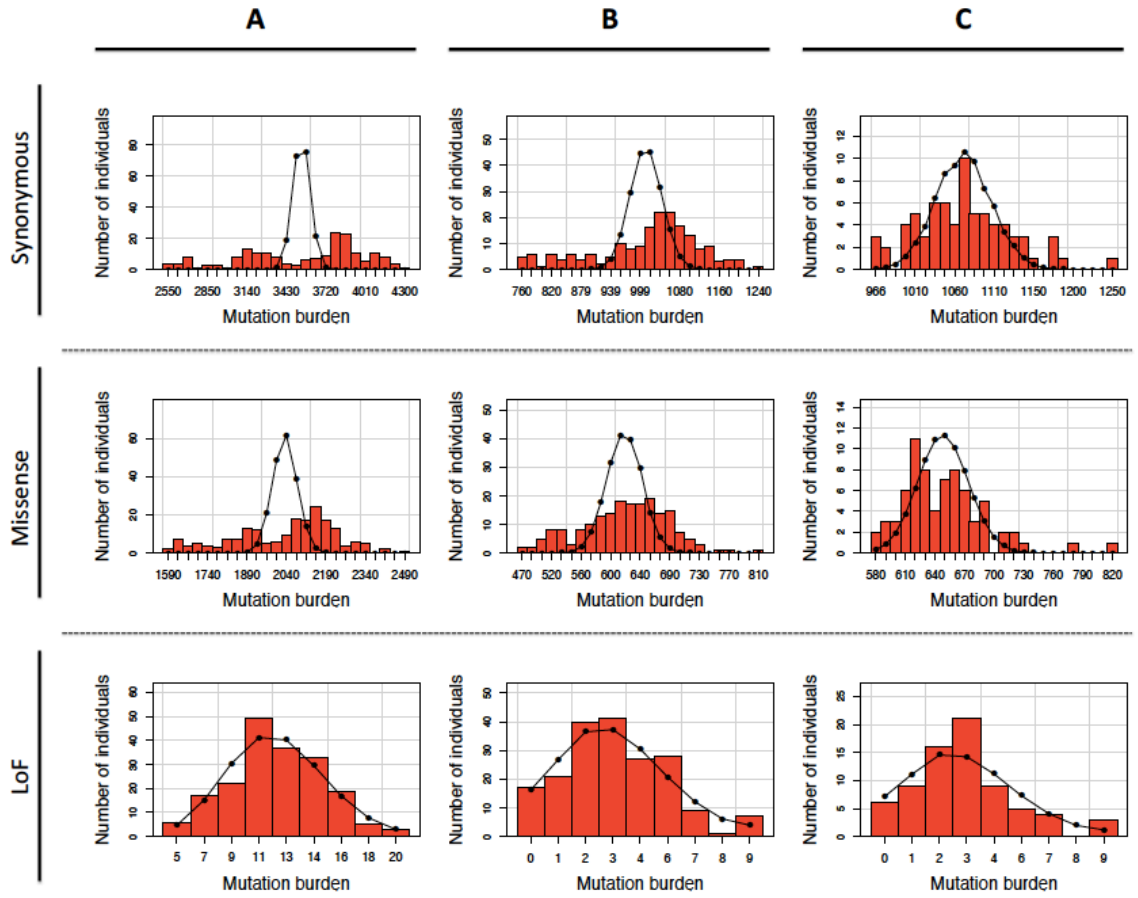


Fig. S16. Mutation burden in the DPGP3 dataset overlaid with Poisson distributions having identical means.
Mutation burden was computed on synonymous, missense and LoF alleles up to a minor allele count of 5 in the DPGP3 dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*). **(A)** Mutation burden was computed for all available rare alleles. **(B)** Mutation burden was computed only for alleles residing within regions of the *D. melanogaster* genome devoid of inversion polymorphisms. **(C)** Mutation burden was computed only in inversion-free individuals for alleles residing within the regions of the *D. melanogaster* genome devoid of inversion polymorphisms.

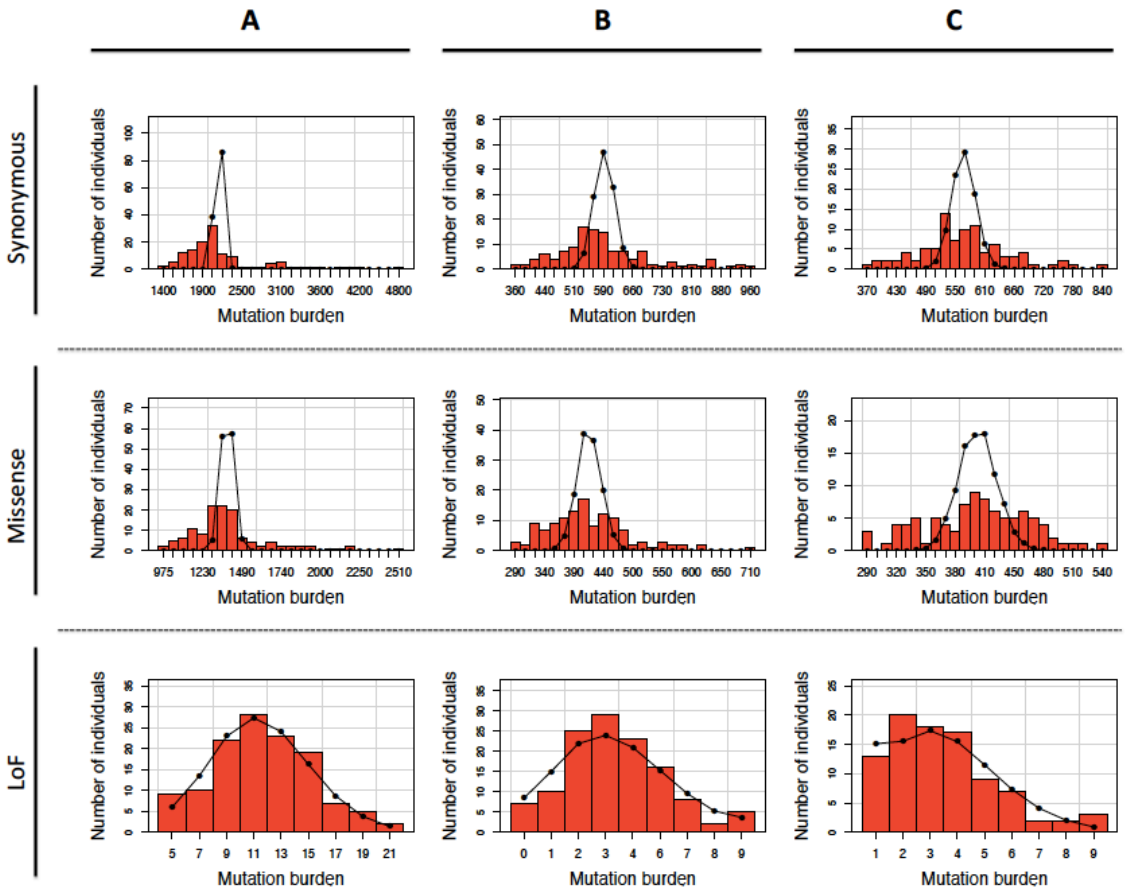


Fig. S17. Mutation burden in the DGRP dataset overlaid with Poisson distributions having identical means. Mutation burden was computed on synonymous, missense and LoF alleles up to a minor allele count of 5 in the DGRP dataset. Black curve shows the expectation under Poisson distribution with the maximum likelihood estimator of its single parameter θ derived from the data (*red*). **(A)** Mutation burden was computed for all available rare alleles. **(B)** Mutation burden was computed only for the alleles residing within regions of the *D. melanogaster* genome devoid of inversion polymorphisms. **(C)** Mutation burden was computed only in inversion-free individuals for alleles residing within the regions of the *D. melanogaster* genome devoid of inversion polymorphisms.

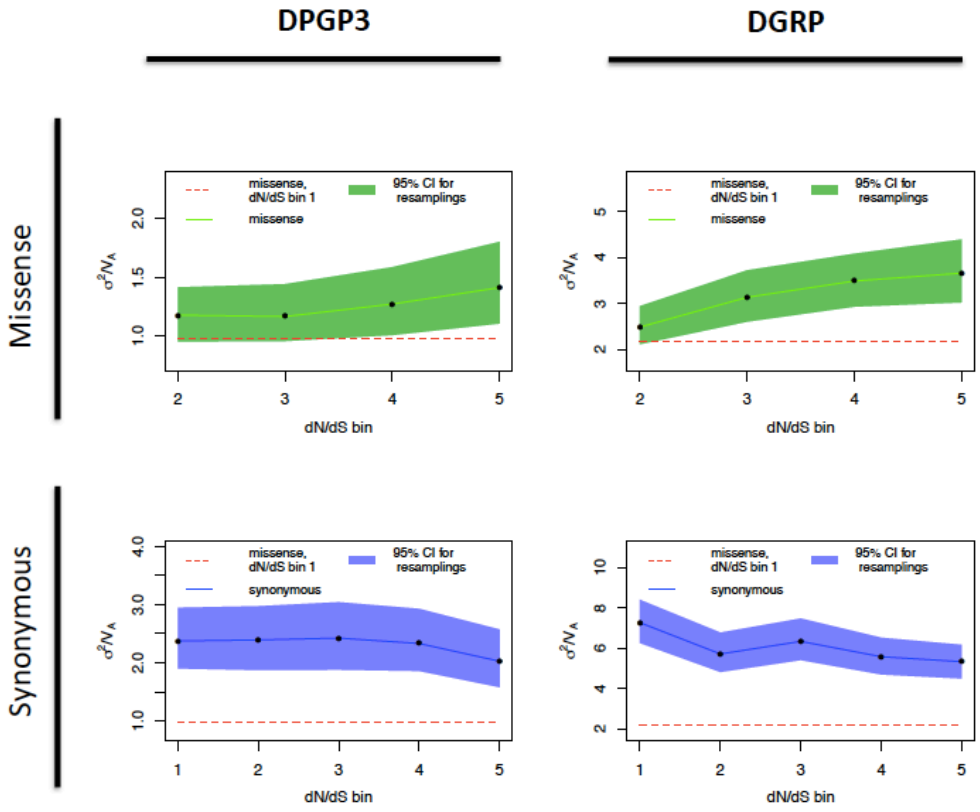


Fig. S18. Mutation burden for common missense and synonymous alleles residing within genes evolving at different rates in two *D. melanogaster* datasets.

Mutation burden was computed on synonymous and missense alleles with minor allele frequency up to 50% in the DPGP3 and DGRP datasets. Genes were subdivided into five equal-sized bins according to the dN/dS ratio where bin 1 contains the most slowly evolving genes and bin 5 contains the most rapidly evolving genes. Missense alleles in bins 2,3,4,5 (*top*) and synonymous alleles in bins 1,2,3,4,5 (*bottom*) were resampled at population frequencies matching the population frequencies of missense alleles in dN/dS bin 1. The value of σ^2/V_A for missense alleles in dN/dS bin 1 (dashed line) is shown for reference. Black points connected with solid lines and the colored area represent median values and 95% confidence intervals for σ^2/V_A calculated in 1000 resamplings for each dN/dS bin.

Supplementary Tables

Table S1. Mutation burden for singletons in six human datasets.

Mutation burden was computed on synonymous, missense and LoF singletons in European, African and East Asian human populations. P-values for σ^2/V_A were computed by two methods; by resampling synonymous alleles at matched allele frequency as LoF alleles, and by permuting functional consequences across variants. A joint P-value for each functional variant type was computed by weighted meta-analysis using Stouffer's method. Values of $\sigma^2/V_A < 1$ and P-values < 0.05 are highlighted.
In a separate excel file.

Table S2. Mutation burden for rare and common alleles in six human datasets.

Mutation burden was computed on synonymous, missense, stop gain and splice-disrupting singletons and alleles with derived allele frequency up to 0.5%, 1%, 2% and 50% in European, African and East Asian human populations. P-values for σ^2/V_A were computed by resampling synonymous alleles at matched allele frequency as LoF alleles. A joint P-value for each deleterious variant type was computed by weighted meta-analysis using Stouffer's method. Values of $\sigma^2/V_A < 1$ and P-values < 0.05 are highlighted.
In a separate excel file.

Table S3. Mutation burden for rare and common alleles in two *D. melanogaster* datasets.

Mutation burden was computed on synonymous, missense and LoF alleles in African and American fruit fly datasets. For rare variants, alleles with a minor allele count up to 1, 2 or 5 were included. For common variants, alleles up to 50% minor allele frequency were included. P-values for σ^2/V_A were computed by resampling synonymous alleles at matched allele frequency as LoF alleles. A joint P-value for each deleterious variant type was computed by weighted meta-analysis using Stouffer's method on genome-wide burden including inversions. Values of $\sigma^2/V_A < 1$ and P-values < 0.05 are highlighted.
In a separate excel file.

Variable	Coefficient	Std.Error	P-value
Intercept	28.486	1.294	$< 2x10^{-16}$
Sequencing batch 2	3.401	1.469	0.021
Sequencing batch 3	2.998	1.304	0.022
Sequencing batch 4	2.294	1.316	0.082
Sequencing batch 5	2.755	1.347	0.041
Region - north	-2.335	0.758	0.002
Region - south	0.859	0.863	0.320
Principal component 1	49.641	7.989	$1.13x10^{-9}$
Principal component 2	-26.924	7.120	$1.76x10^{-4}$
Principal component 3	14.704	6.336	0.021
Principal component 4	-7.264	6.523	0.266
Principal component 5	4.849	6.381	0.448
Principal component 6	2.882	6.266	0.646
Principal component 7	-6.358	6.247	0.309
Principal component 8	-6.052	6.230	0.332
Principal component 9	-7.730	6.343	0.224
Principal component 10	7.040	6.279	0.263

Table S4. Multivariate regression analysis for rare synonymous mutation burden.

Synonymous singletons were used to compute mutation burden for every sample in the GoNL dataset. Mutation burden was residualized under a generalized linear model consisting of 10 principal components and other covariates for geographic structure (see methods for model details).

Table S5. Properties of residualized mutation burden.

Raw mutation burden was computed on synonymous, missense and LoF singletons in the GoNL dataset. *Residualized* mutation burden was computed by correcting for geographic and technical covariates in a multivariate regression model (see methods for details). Values of $\sigma^2/V_A < 1$ are highlighted.

In a separate excel file.

	Net LD			*Mean LD per pair of loci*		
Variant type	**Genome-wide**	**Inter-chromosomal**	**Intra-chromosomal**	**Genome-wide**	**Inter-chromosomal**	**Intra-chromosomal**
Synonymous	20.409	16.782	3.627	4.549×10^{-8}	3.959×10^{-8}	1.471×10^{-7}
Missense	65.489	57.770	7.720	3.605×10^{-8}	3.370×10^{-8}	7.543×10^{-8}
Stop gain	-0.106	-0.080	-0.026	-8.186×10^{-8}	-6.561×10^{-8}	-3.570×10^{-7}
Splice	-0.042	-0.039	-0.004	-1.068×10^{-7}	-1.042×10^{-7}	-1.490×10^{-7}
Splice and stop gain	-0.182	-0.135	-0.047	-5.832×10^{-8}	-4.584×10^{-8}	-2.654×10^{-7}

Table S6. Net LD for rare alleles partitioned into intra-chromosomal and inter-chromosomal components.

Net LD was computed for synonymous, missense, stop gain and splice-disrupting singletons in the GoNL dataset, and partitioned by summing $D_{i,j}$ values for all intra-chromosomal pairs of SNPs and all inter-chromosomal pairs of SNPs separately (see methods for details). Net LD normalized per pair of loci is also shown.

Chromosome	Net intra-chromosomal LD by chromosome		
	Synonymous	Missense	Splice and stop gain
1	0.5927	0.9401	-0.0311
2	0.2880	1.3311	-0.0040
3	0.3382	0.7028	-0.0083
4	0.1454	0.4624	-0.0082
5	0.2267	0.3676	0.0084
6	0.2117	0.2443	0.0005
7	-0.0189	0.1969	-0.0033
8	-0.0648	-0.1567	-0.0037
9	0.2008	0.6133	-0.0027
10	0.0589	0.3451	-0.0090
11	0.2614	0.5889	0.0005
12	0.1288	0.3945	0.0118
13	0.0534	0.0286	-0.0013
14	0.0711	0.3316	-0.0032
15	0.2966	0.1723	0.0040
16	0.2734	0.4084	-0.0017
17	0.3252	0.3772	0.0113
18	0.0847	0.1265	-0.0029
19	0.0963	0.2889	-0.0100
20	0.0920	-0.1830	0.0005
21	0.0216	0.0362	0.0032
22	-0.0569	0.1030	0.0018

Table S7. Net LD for rare alleles by chromosome.

Net LD was computed for synonymous, missense, stop gain and splice-disrupting singletons in the GoNL dataset, and partitioned by summing $D_{i,j}$ values for all intra-chromosomal pairs of SNPs in each chromosome separately (see methods for details).

Table S8. Variant quality control in human datasets.

Variants were filtered based on various quality control criteria (see methods for details). Raw number of variants, filtering steps and final number of variants used for population genetics analysis are shown for six human datasets.
In a separate excel file.

Table S9. Sample quality control in human datasets.

Samples were filtered based on various quality control criteria (see methods for details). Raw number of samples, filtering steps and final number of samples used for population genetics analysis are shown for six human datasets.
In a separate excel file.

Filter	The total number of canonical *D. melanogaster* FlyBase gene models retained after filtering
Unfiltered (total)	13300
Pseudogenes, putative annotation errors	13252
Pseudogenes, putative annotation errors, chemoreceptor genes	13088
Pseudogenes, putative annotation errors, chemoreceptor genes, genes residing in known inversions	4881

Table S10. The numbers of *D. melanogaster* FlyBase canonical gene models retained after various filtering steps. Only genes on 5 euchromatic chromosome arms are considered.

Dataset	Total number of segregating codon sites	Total number of segregating codon sites after excluding codons with putative double mutations	Total number of segregating codon sites after excluding codons with putative double mutations and codons with missing genotypes
DPGP3	1147021	1128927	858135
DGRP	471696	463690	51608

Table S11. Total numbers of segregating codon sites in *D. melanogaster* datasets.

Dataset	Total number of segregating splice sites	Total number of segregating splice sites after excluding splice sites with putative double mutations	Total number of segregating splice sites after excluding splice sites with putative double mutations and splice sites with missing genotypes
DPGP3	712	703	516
DGRP	297	294	38

Table S12. Total numbers of segregating splice sites in *D. melanogaster* datasets.

Table S13. Mutation burden for rare and common alleles in two *D. melanogaster* datasets after exclusion of multi-allelic splice sites and codon sites with more than one minor allele belonging to the same functional class.

Mutation burden was computed on synonymous, missense and LoF alleles in African and American fruit fly datasets. For rare variants, alleles with a minor allele count up to 1, 2 or 5 were included. For common variants, alleles up to 50% minor allele frequency were included. P-values for σ^2/V_A were computed by resampling synonymous alleles at matched allele frequency as LoF alleles. A joint P-value for each deleterious variant type was computed by weighted meta-analysis using Stouffer's method on genome-wide burden including inversions. Values of $\sigma^2/V_A < 1$ and P-values < 0.05 are highlighted.
In a separate excel file.

Table S14. Missense mutation burden for rare and common alleles in two *D. melanogaster* datasets.

Mutation burden was computed on missense alleles in African and American fruit fly datasets. For rare variants, alleles with a minor allele count up to 5 were included. For common variants, alleles up to 50% minor allele frequency were included. P-values for the σ^2/V_A were computed by permuting functional consequences across variants. A joint P-value for each functional variant type was computed by weighted meta-analysis using Stouffer's method. Values of $\sigma^2/V_A < 1$ and P-values < 0.05 are highlighted.
In a separate excel file.

Variable	Coefficient	Std. Error	P-value
Intercept	57.518	1.974	$< 2x10^{-16}$
Sequencing batch 2	4.540	2.242	0.043
Sequencing batch 3	4.570	1.990	0.022
Sequencing batch 4	4.510	2.009	0.025
Sequencing batch 5	3.691	2.056	0.073
Region - north	-3.274	1.157	0.005
Region - south	3.088	1.317	0.019
Principal component 1	73.729	12.191	$2.96x10^{-9}$
Principal component 2	-61.995	10.866	$2.04x10^{-8}$
Principal component 3	-2.089	9.669	0.829
Principal component 4	1.751	9.954	0.860
Principal component 5	5.510	9.737	0.572
Principal component 6	-3.281	9.562	0.732
Principal component 7	-2.186	9.533	0.819
Principal component 8	1.374	9.507	0.885
Principal component 9	-3.042	9.679	0.753
Principal component 10	-0.975	9.583	0.919

Table S15. Multivariate regression analysis for rare missense mutation burden.

Missense singletons were used to compute mutation burden for every sample in the GoNL dataset. Mutation burden was residualized under a generalized linear model consisting of 10 principal components and other covariates for geographic structure (see methods for model details).

Variable	Coefficient	Std. Error	P-value
Intercept	2.533	0.323	2.99×10^{-14}
Sequencing batch 2	0.041	0.367	0.912
Sequencing batch 3	-0.003	0.326	0.993
Sequencing batch 4	0.206	0.329	0.531
Sequencing batch 5	0.002	0.337	0.995
Region - north	-0.196	0.189	0.302
Region - south	0.246	0.216	0.254
Principal component 1	-0.738	1.996	0.712
Principal component 2	0.139	1.779	0.938
Principal component 3	-1.246	1.583	0.432
Principal component 4	-0.132	1.629	0.935
Principal component 5	0.129	1.594	0.936
Principal component 6	-1.838	1.565	0.241
Principal component 7	-2.910	1.561	0.063
Principal component 8	2.057	1.556	0.187
Principal component 9	-2.229	1.585	0.160
Principal component 10	-0.117	1.569	0.941

Table S16. Multivariate regression analysis for rare LoF mutation burden.

LoF singletons were used to compute mutation burden for every sample in the GoNL dataset. Mutation burden was residualized under a generalized linear model consisting of 10 principal components and other covariates for geographic structure (see methods for model details).

	Mean (μ)	Additive variance (V_A)	Variance (σ^2)	σ^2/V_A	Net LD per pair of alleles	P-value
sample size = 495						
Synonymous	30.257	30.257	50.685	1.675	0.022	
Missense	60.883	60.883	126.436	2.077	0.018	
LoF	2.576	2.576	2.395	0.930	-0.027	0.025
sample size = 247						
Synonymous	51.279	51.279	84.641	1.651	0.013	
Missense	97.449	97.449	166.387	1.707	0.007	
LoF	3.891	3.891	3.464	0.890	-0.028	0.037
sample size = 123						
Synonymous	80.098	80.098	137.810	1.721	0.009	
Missense	145.358	145.358	214.281	1.474	0.003	
LoF	5.325	5.325	4.680	0.879	-0.023	0.096
sample size = 82						
Synonymous	103.146	103.146	216.349	2.097	0.011	
Missense	179.805	179.805	275.270	1.531	0.003	
LoF	5.988	5.988	5.025	0.839	-0.027	0.090

Table S17. Subsampling experiments to test sensitivity of underdispersion signal to changes in sample size.
Mutation burden was computed on synonymous, missense and LoF singletons in the GoNL dataset, and in the dataset down-sampled to half, quarter and one-sixth its original size**.** P-values for σ^2/V_A were computed by resampling synonymous alleles at matched allele frequency as LoF alleles in each subsampled dataset.

	Mean (μ)	Additive variance (V_A)	Variance (σ^2)	σ^2/V_A	Net LD per pair of alleles	P-value
Singletons						
Synonymous	7.582	7.582	8.697	1.147	0.019	
Missense	15.107	15.107	18.157	1.202	0.013	
LoF	0.653	0.653	0.616	0.944	-0.086	0.146
DAF ≤ 1%						
Synonymous	36.501	36.916	39.732	1.076	0.002	
Missense	58.861	59.448	59.647	1.003	0.000	
LoF	1.741	1.742	1.593	0.915	-0.049	0.060
DAF ≤ 2%						
Synonymous	55.917	56.357	64.007	1.136	0.002	
Missense	82.669	83.197	88.222	1.060	0.001	
LoF	2.390	2.393	2.194	0.917	-0.035	0.077

Table S18. Mutation burden analysis restricted only to sites with coverage close to the dataset mean.

Mutation burden was computed on synonymous, missense, and LoF singletons and alleles with derived allele frequency up to 1% and 2% in the GoNL dataset. Only sites with an average coverage between 12x and 14x were considered (~13x is the dataset average). P-values for σ^2/V_A were computed by resampling synonymous alleles at matched allele frequency as LoF alleles.

	Mean (μ)	Additive variance (V_A)	Variance (σ^2)	σ^2/V_A	Net LD per pair of alleles	P-value
Singletons						
Synonymous	30.257	30.257	50.685	1.675	0.022	
Missense	60.883	60.883	126.436	2.077	0.018	
LoF and frameshift	4.469	4.469	4.719	1.056	0.013	0.218
DAF ≤ 1%						
Synonymous	130.867	132.087	200.549	1.518	0.004	
Missense	220.156	221.911	273.727	1.233	0.001	
LoF and frameshift	11.800	11.881	11.132	0.937	-0.005	0.023
DAF ≤ 2%						
Synonymous	202.099	203.366	317.312	1.560	0.003	
Missense	309.760	311.282	364.090	1.170	0.001	
LoF and frameshift	15.325	15.405	14.507	0.942	-0.004	0.043

Table S19. Mutation burden for rare SNPs and indels.

Mutation burden was computed on synonymous, missense, and LoF SNPs for singletons and alleles with derived allele frequency up to 1% and 2%, and on frameshift indels for singletons and alleles with minor allele frequency up to 1% and 2% in the GoNL dataset. P-values for σ^2/V_A were computed by resampling synonymous alleles and intronic indels at matched allele frequency as LoF alleles and frameshift indels respectively.

Table S20. Mutation burden for common alleles versus evolutionary rate of genes in two *D. melanogaster* datasets.

Mutation burden was computed on missense alleles with minor allele frequency up to 50% in African and American fruit fly datasets. Genes were subdivided into five equal-sized bins according to the dN/dS ratio where bin 1 contains the most slowly evolving genes and bin 5 contains the most rapidly evolving genes. Mutation burden was computed separately for each bin. Values of $\sigma^2/V_A < 1$ are highlighted.
In a separate excel file.

Table S21. Mutation burden for rare alleles versus evolutionary rate of genes in two *D. melanogaster* datasets.

Mutation burden was computed on missense alleles with minor allele count up to 5 in African and American fruit fly datasets. Genes were subdivided into five equal-sized bins according to the dN/dS ratio where bin 1 contains the most slowly evolving genes and bin 5 contains the most rapidly evolving genes. Mutation burden was computed separately for each bin. Values of $\sigma^2/V_A < 1$ are highlighted.
In a separate excel file.

	Mean (μ)	Additive variance (V_A)	Variance (σ^2)	σ^2/V_A	Net LD per pair of alleles	P-value
Singletons						
Synonymous	2.414	2.414	2.567	1.063	0.026	0.911
Missense	3.489	3.489	3.271	0.937	-0.018	0.219
LoF	0.032	0.032	0.031	0.970	-0.939	0.617
DAF ≤ 1%						
Synonymous	10.602	10.716	12.208	1.139	0.013	0.979
Missense	10.851	10.940	10.123	0.925	-0.007	0.030
LoF	0.044	0.044	0.043	0.958	-0.944	0.289
DAF ≤ 2%						
Synonymous	16.269	16.380	17.165	1.048	0.003	0.847
Missense	14.059	14.141	12.901	0.912	-0.006	0.063
LoF	0.044	0.044	0.043	0.958	-0.944	0.285

Table S22. Mutation burden in the crucial genome in humans.

Mutation burden was computed on synonymous, missense and LoF singletons and alleles with derived allele frequency up to 1% and 2% in the GoNL dataset. Only most selectively constrained genes were used for this analysis (estimated selection coefficient against heterozygous protein truncating variants exceeding 0.2)(*37*). P-values for σ^2/V_A were computed by permuting functional consequences across variants.

Consortia

This study/database makes use of data generated by the Genome of the Netherlands Project. A full list of the investigators is available from http://www.nlgenome.nl. Funding for the project was provided by the Netherlands Organization for Scientific Research under award number 184021007, dated July 9, 2009 and made available as a Rainbow Project of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL). The sequencing was carried out in collaboration with the Beijing Institute for Genomics (BGI).

The members of the Genome of the Netherlands Consortium are Laurent C. Francioli, Androniki Menelaou, Sara L. Pulit, Freerk van Dijk, Pier Francesco Palamara, Clara C. Elbers, Pieter B.T. Neerincx, Kai Ye, Victor Guryev, Wigard P. Kloosterman, Patrick Deelen, Abdel Abdellaoui, Elisabeth M. van Leeuwen, Mannis van Oven, Martijn Vermaat, Mingkun Li, Jeroen F.J. Laros, Lennart C. Karssen, Alexandros Kanterakis, Najaf Amin, Jouke Jan Hottenga, Eric-Wubbo Lameijer, Mathijs Kattenberg, Martijn Dijkstra, Heorhiy Byelas, Jessica van Setten, Barbera D.C. van Schaik, Jan Bot, Isac J. Nijman, Ivo Renkens, Tobias Marschall, Alexander Schnhuth, Jayne Y. Hehir-Kwa, Robert E Handsaker, Paz Polak, Mashaal Sohail, Dana Vuzman, Fereydoun Hormozdiari, David van Enckevort, Hailiang Mei, Vyacheslav Koval, Matthijs H. Moed, K. Joeri van der Velde, Fernando Rivadeneira, Karol Estrada, Carolina Medina-Gomez, Aaron Isaacs, Steven A. McCarroll, Marian Beekman, Anton J.M. de Craen, H. Eka D. Suchiman, Albert Hofman, Ben Oostra, Andr G. Uitterlinden, Gonneke Willemsen, LifeLines Cohort Study, Mathieu Platteel, Jan H. Veldink, Leonard H. van den Berg, Steven J. Pitts, Shobha Potluri, Purnima Sundar, David R. Cox, Shamil R. Sunyaev, Johan T. den Dunnen, Mark Stoneking, Peter de Knijff, Manfred Kayser, Qibin Li, Yingrui Li, Yuanping Du, Ruoyan Chen, Hongzhi Cao, Ning Li, Sujie Cao, Jun Wang, Jasper A. Bovenberg, Itsik Pe'er, P. Eline Slagboom, Cornelia M. van Duijn, Dorret I. Boomsma, Gert-Jan B van Ommen, Paul I.W. de Bakker, Morris A. Swertz, and Cisca Wijmenga.

Part of the data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Part of the data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.;

Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References and Notes

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

2. E. E. Shnol, A. S. Kondrashov, The effect of selection on the phenotypic variance. *Genetics* **134**, 995–996 (1993). Medline

3. M. Kimura, T. Maruyama, The mutational load with epistatic gene interactions in fitness. *Genetics* **54**, 1337–1351 (1966). Medline

4. J. F. Crow, M. Kimura, Efficiency of truncation selection. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 396–399 (1979). doi:10.1073/pnas.76.1.396 Medline

5. A. S. Kondrashov, Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988). doi:10.1038/336435a0 Medline

6. J. A. G. M. de Visser, S. F. Elena, The evolution of sex: Empirical insights into the roles of epistasis and drift. *Nat. Rev. Genet.* **8**, 139–149 (2007). doi:10.1038/nrg1985 Medline

7. H.-C. Chiu, C. J. Marx, D. Segrè, Epistasis from functional dependence of fitness on underlying traits. *Proc. Biol. Sci.* **279**, 4156–4164 (2012). doi:10.1098/rspb.2012.1449 Medline

8. C. Bank, R. T. Hietpas, J. D. Jensen, D. N. A. Bolon, A systematic survey of an intragenic epistatic landscape. *Mol. Biol. Evol.* **32**, 229–238 (2015). doi:10.1093/molbev/msu301 Medline

9. O. Puchta, B. Cseke, H. Czaja, D. Tollervey, G. Sanguinetti, G. Kudla, Network of epistatic interactions within a yeast snoRNA. *Science* **352**, 840–844 (2016). doi:10.1126/science.aaf0965 Medline

10. M. Bulmer, *The Mathematical Theory of Quantitative Genetics* (Clarendon, 1980).

11. Materials and methods are available as supporting materials.

12. A. S. Kondrashov, Dynamics of unconditionally deleterious mutations: Gaussian approximation and soft selection. *Genet. Res.* **65**, 113–121 (1995). doi:10.1017/S0016672300033139 Medline

13. B. Charlesworth, Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**, 199–221 (1990). doi:10.1017/S0016672300025532 Medline

14. M. G. Bulmer, The effect of selection on genetic variability: A simulation study. *Genet. Res.* **28**, 101–117 (1976). doi:10.1017/S0016672300016797 Medline

15. L. C. Francioli, A. Menelaou, S. L. Pulit, F. van Dijk, P. F. Palamara, C. C. Elbers, P. B. T. Neerincx, K. Ye, V. Guryev, W. P. Kloosterman, P. Deelen, A. Abdellaoui, E. M. van Leeuwen, M. van Oven, M. Vermaat, M. Li, J. F. J. Laros, L. C. Karssen, A. Kanterakis, N. Amin, J. J. Hottenga, E.-W. Lameijer, M. Kattenberg, M. Dijkstra, H. Byelas, J. van Setten, B. D. C. van Schaik, J. Bot, I. J. Nijman, I. Renkens, T. Marschall, A. Schönhuth, J. Y. Hehir-Kwa, R. E. Handsaker, P. Polak, M. Sohail, D. Vuzman, F. Hormozdiari, D. van Enckevort, H. Mei, V. Koval, M. H. Moed, K. J. van der Velde, F. Rivadeneira, K. Estrada, C. Medina-Gomez, A. Isaacs, S. A. McCarroll, M. Beekman, A. J. M. de Craen, H. E. D. Suchiman, A. Hofman, B. Oostra, A. G. Uitterlinden, G. Willemsen, L. L. C.

Study, M. Platteel, J. H. Veldink, L. H. van den Berg, S. J. Pitts, S. Potluri, P. Sundar, D. R. Cox, S. R. Sunyaev, J. T. Dunnen, M. Stoneking, P. de Knijff, M. Kayser, Q. Li, Y. Li, Y. Du, R. Chen, H. Cao, N. Li, S. Cao, J. Wang, J. A. Bovenberg, I. Pe'er, P. E. Slagboom, C. M. van Duijn, D. I. Boomsma, G.-J. B. van Ommen, P. I. W. de Bakker, M. A. Swertz, C. Wijmenga; Genome of the Netherlands Consortium, Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat. Genet.* **46**, 818–825 (2014). doi:10.1038/ng.3021 Medline

16. J. B. Lack, C. M. Cardeno, M. W. Crepeau, W. Taylor, R. B. Corbett-Detig, K. A. Stevens, C. H. Langley, J. E. Pool, The *Drosophila* genome nexus: A population genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. *Genetics* **199**, 1229–1241 (2015). doi:10.1534/genetics.115.174664 Medline

17. R. C. Yang, Zygotic associations and multilocus statistics in a nonequilibrium diploid population. *Genetics* **155**, 1449–1458 (2000). Medline

18. G. R. Abecasis, A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin, R. E. Handsaker, H. M. Kang, G. T. Marth, G. A. McVean; 1000 Genomes Project Consortium, An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012). doi:10.1038/nature11632 Medline

19. T. F. C. Mackay, S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles, D. Zhu, S. Casillas, Y. Han, M. M. Magwire, J. M. Cridland, M. F. Richardson, R. R. H. Anholt, M. Barrón, C. Bess, K. P. Blankenburg, M. A. Carbone, D. Castellano, L. Chaboub, L. Duncan, Z. Harris, M. Javaid, J. C. Jayaseelan, S. N. Jhangiani, K. W. Jordan, F. Lara, F. Lawrence, S. L. Lee, P. Librado, R. S. Linheiro, R. F. Lyman, A. J. Mackey, M. Munidasa, D. M. Muzny, L. Nazareth, I. Newsham, L. Perales, L.-L. Pu, C. Qu, M. Ràmia, J. G. Reid, S. M. Rollmann, J. Rozas, N. Saada, L. Turlapati, K. C. Worley, Y.-Q. Wu, A. Yamamoto, Y. Zhu, C. M. Bergman, K. R. Thornton, D. Mittelman, R. A. Gibbs, The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482**, 173–178 (2012). doi:10.1038/nature10811 Medline

20. B. Charlesworth, Why we are not dead one hundred times over. *Evolution.* **67**, 3354–3361 (2013). doi:10.1111/evo.12195 Medline

21. J. Felsenstein, The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974). Medline

22. N. H. Barton, S. P. Otto, Evolution of recombination due to random drift. *Genetics* **169**, 2353–2370 (2005). doi:10.1534/genetics.104.032821 Medline

23. S. Besenbacher, S. Liu, J. M. G. Izarzugaza, J. Grove, K. Belling, J. Bork-Jensen, S. Huang, T. D. Als, S. Li, R. Yadav, A. Rubio-García, F. Lescai, D. Demontis, J. Rao, W. Ye, T. Mailund, R. M. Friborg, C. N. S. Pedersen, R. Xu, J. Sun, H. Liu, O. Wang, X. Cheng, D. Flores, E. Rydza, K. Rapacki, J. Damm Sørensen, P. Chmura, D. Westergaard, P. Dworzynski, T. I. A. Sørensen, O. Lund, T. Hansen, X. Xu, N. Li, L. Bolund, O. Pedersen, H. Eiberg, A. Krogh, A. D. Børglum, S. Brunak, K. Kristiansen, M. H. Schierup, J. Wang, R. Gupta, P. Villesen, S. Rasmussen, Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. *Nat. Commun.* **6**, 5969 (2015). doi:10.1038/ncomms6969 Medline

24. C. M. Rands, S. Meader, C. P. Ponting, G. Lunter, 8.2% of the Human genome is constrained: Variation in rates of turnover across functional element classes in the human lineage. *PLOS Genet.* **10**, e1004525 (2014). doi:10.1371/journal.pgen.1004525 Medline

25. Y. Lesecque, P. D. Keightley, A. Eyre-Walker, A resolution of the mutation load paradox in humans. *Genetics* **191**, 1321–1330 (2012). doi:10.1534/genetics.112.140343 Medline

26. S. C. Stearns, S. G. Byars, D. R. Govindaraju, D. Ewbank, Measuring selection in contemporary human populations. *Nat. Rev. Genet.* **11**, 611–622 (2010). Medline

27. E. C. Larsen, O. B. Christiansen, A. M. Kolte, N. Macklon, New insights into mechanisms behind miscarriage. *BMC Med.* **11**, 154 (2013). doi:10.1186/1741-7015-11-154 Medline

28. H. M. E. Hed, Trends in opportunity for natural selection in the Swedish population during the period 1650-1980. *Hum. Biol.* **59**, 785–797 (1987). Medline

29. A. Courtiol, I. J. Rickard, V. Lummaa, A. M. Prentice, A. J. C. Fulford, S. C. Stearns, The demographic transition influences variance in fitness and selection on height and BMI in rural Gambia. *Curr. Biol.* **23**, 884–889 (2013). doi:10.1016/j.cub.2013.04.006 Medline

30. W. Huang, A. Massouras, Y. Inoue, J. Peiffer, M. Ràmia, A. M. Tarone, L. Turlapati, T. Zichner, D. Zhu, R. F. Lyman, M. M. Magwire, K. Blankenburg, M. A. Carbone, K. Chang, L. L. Ellis, S. Fernandez, Y. Han, G. Highnam, C. E. Hjelmen, J. R. Jack, M. Javaid, J. Jayaseelan, D. Kalra, S. Lee, L. Lewis, M. Munidasa, F. Ongeri, S. Patel, L. Perales, A. Perez, L. Pu, S. M. Rollmann, R. Ruth, N. Saada, C. Warner, A. Williams, Y.-Q. Wu, A. Yamamoto, Y. Zhang, Y. Zhu, R. R. H. Anholt, J. O. Korbel, D. Mittelman, D. M. Muzny, R. A. Gibbs, A. Barbadilla, J. S. Johnston, E. A. Stone, S. Richards, B. Deplancke, T. F. C. Mackay, Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. *Genome Res.* **24**, 1193–1208 (2014). doi:10.1101/gr.171546.113 Medline

31. P. Duchen, D. Živković, S. Hutter, W. Stephan, S. Laurent, Demographic inference reveals African and European admixture in the North American *Drosophila melanogaster* population. *Genetics* **193**, 291–301 (2013). doi:10.1534/genetics.112.145912 Medline

32. A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006). doi:10.1038/ng1847 Medline

33. S. Gazal, M. Sahbatou, M.-C. Babron, E. Génin, A.-L. Leutenegger, High level of inbreeding in final phase of 1000 Genomes Project. *Sci. Rep.* **5**, 17453 (2015). doi:10.1038/srep17453 Medline

34. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi:10.1101/gr.107524.110 Medline

35. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). doi:10.1038/ng.806 Medline

36. C. Raczy, R. Petrovski, C. T. Saunders, I. Chorny, S. Kruglyak, E. H. Margulies, H.-Y. Chuang, M. Källberg, S. A. Kumar, A. Liao, K. M. Little, M. P. Strömberg, S. W. Tanner, Isaac: Ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics* **29**, 2041–2043 (2013). doi:10.1093/bioinformatics/btt314 Medline

37. C. A. Cassa, D. Weghorn, D. J. Balick, D. M. Jordan, D. Nusinow, K. E. Samocha, A. O'Donnell-Luria, D. G. MacArthur, M. J. Daly, D. R. Beier, S. R. Sunyaev, Estimating the selective effects of heterozygous protein-truncating variants from human exome data. *Nat. Genet.* (2017). doi:10.1038/ng.3831 Medline

38. J. Y. Hehir-Kwa, T. Marschall, W. P. Kloosterman, L. C. Francioli, J. A. Baaijens, L. J. Dijkstra, A. Abdellaoui, V. Koval, D. T. Thung, R. Wardenaar, I. Renkens, B. P. Coe, P. Deelen, J. de Ligt, E.-W. Lameijer, F. van Dijk, F. Hormozdiari, A. G. Uitterlinden, C. M. van Duijn, E. E. Eichler, P. I. de Bakker, M. A. Swertz, C. Wijmenga, G. B. van Ommen, P. E. Slagboom, D. I. Boomsma, A. Schönhuth, K. Ye, V. Guryev, J. S. Cao, P. B. T. Neerincx, M. Dijkstra, G. Byelas, A. Kanterakis, J. Bot, M. Vermaat, J. F. J. Laros, J. T. den Dunnen, P. de Knijff, L. C. Karssen, E. M. van Leeuwen, N. Amin, F. Rivadeneira, K. Estrada, J.-J. Hottenga, V. M. Kattenberg, D. van Enckevort, H. Mei, M. Santcroos, B. D. C. van Schaik, R. E. Handsaker, S. A. McCarroll, A. Ko, P. Sudmant, I. J. Nijman, A. G. Uitterlinden, C. M. van Duijn, E. E. Eichler, P. I. W. de Bakker, M. A. Swertz, C. Wijmenga, G.-J. B. van Ommen, P. E. Slagboom, D. I. Boomsma, A. Schönhuth, K. Ye, V. Guryev; Genome of the Netherlands Consortium, A high-quality human reference panel reveals the complexity and distribution of genomic structural variants. *Nat. Commun.* **7**, 12989 (2016). doi:10.1038/ncomms12989 Medline

39. W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, D. Haussler, The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002). doi:10.1101/gr.229102 Medline

40. M. A. Crosby, J. L. Goodman, V. B. Strelets, P. Zhang, W. M. Gelbart; FlyBase Consortium, FlyBase: Genomes by the dozen. *Nucleic Acids Res.* **35** (Database), D486–D491 (2007). doi:10.1093/nar/gkl827 Medline

41. C. S. McBride, Rapid evolution of smell and taste receptor genes during host specialization in Drosophila sechellia. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4996–5001 (2007). doi:10.1073/pnas.0608424104 Medline

42. A. Gardiner, D. Barker, R. K. Butlin, W. C. Jordan, M. G. Ritchie, Drosophila chemoreceptor gene evolution: Selection, specialization and genome size. *Mol. Ecol.* **17**, 1648–1657 (2008). doi:10.1111/j.1365-294X.2008.03713.x Medline

43. Y. C. G. Lee, J. A. Reinhardt, Widespread polymorphism in the positions of stop codons in Drosophila melanogaster. *Genome Biol. Evol.* **4**, 533–549 (2012). doi:10.1093/gbe/evr113 Medline

44. C. S. McBride, J. R. Arguello, B. C. O'Meara, Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics* **177**, 1395–1416 (2007). doi:10.1534/genetics.107.078683 Medline

45. R. B. Corbett-Detig, D. L. Hartl, Population genomics of inversion polymorphisms in Drosophila melanogaster. *PLOS Genet.* **8**, e1003056 (2012). doi:10.1371/journal.pgen.1003056 Medline

46. H. Luo, Y. Lin, F. Gao, C. T. Zhang, R. Zhang, DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res.* **42**, D574–D580 (2014). doi:10.1093/nar/gkt1131 Medline

47. A. M. Larracuente, T. B. Sackton, A. J. Greenberg, A. Wong, N. D. Singh, D. Sturgill, Y. Zhang, B. Oliver, A. G. Clark, Evolution of protein-coding genes in Drosophila. *Trends Genet.* **24**, 114–123 (2008). doi:10.1016/j.tig.2007.12.001 Medline

48. W.-H. Li, *Molecular Evolution* (Sinauer Associates, 1997).

49. P. Liptak, On the combination of independent tests. *Magyar Tud. Akad. Mat. Kutato Int. Kozl*. **3**, 171–197 (1958).

50. B. Charlesworth, D. Charlesworth, *Elements of Evolutionary Genetics* (Freeman, 2010).

51. M. G. Bulmer, Linkage disequilibrium and genetic variability. *Genet. Res.* **23**, 281–289 (1974). doi:10.1017/S0016672300014920 Medline

52. R.-C. Yang, Gametic and zygotic associations. *Genetics* **165**, 447–450 (2003). Medline

53. A. L. Price, M. E. Weale, N. Patterson, S. R. Myers, A. C. Need, K. V. Shianna, D. Ge, J. I. Rotter, E. Torres, K. D. Taylor, D. B. Goldstein, D. Reich, Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* **83**, 132–135, author reply 135–139 (2008). doi:10.1016/j.ajhg.2008.06.005 Medline

54. P. W. Messer, SLiM: Simulating evolution with selection and linkage. *Genetics* **194**, 1037–1039 (2013). doi:10.1534/genetics.113.152181 Medline

55. J. A. Tennessen, A. W. Bigham, T. D. O'Connor, W. Fu, E. E. Kenny, S. Gravel, S. McGee, R. Do, X. Liu, G. Jun, H. M. Kang, D. Jordan, S. M. Leal, S. Gabriel, M. J. Rieder, G. Abecasis, D. Altshuler, D. A. Nickerson, E. Boerwinkle, S. Sunyaev, C. D. Bustamante, M. J. Bamshad, J. M. Akey; Broad GO; Seattle GO; NHLBI Exome Sequencing Project, Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337**, 64–69 (2012). doi:10.1126/science.1219240 Medline

56. G. V. Kryukov, L. A. Pennacchio, S. R. Sunyaev, Most rare missense alleles are deleterious in humans: Implications for complex disease and association studies. *Am. J. Hum. Genet.* **80**, 727–739 (2007). doi:10.1086/513473 Medline