Biases in ARG-based inference of historical population size in populations experiencing selection

Jacob I. Marsh[1,*], Parul Johri[1,2,3,*]

[1] Department of Biology, University of North Carolina, Chapel Hill, NC 27599

[2] Department of Genetics, University of North Carolina, Chapel Hill, NC 27599

[3] Integrative Program for Biological and Genome Sciences, University of North Carolina, Chapel Hill, NC 27599

*Corresponding author contact: jacob.marsh@unc.edu, pjohri@unc.edu

Abstract

Inferring the demographic history of populations provides fundamental insights into species dynamics and is essential for developing a null model to accurately study selective processes. However, background selection and selective sweeps can produce genomic signatures at linked sites that mimic or mask signals associated with historical population size change. While the theoretical biases introduced by the linked effects of selection have been well established, it is unclear whether ARG-based approaches to demographic inference in typical empirical analyses are susceptible to mis-inference due to these effects. To address this, we developed highly realistic forward simulations of human and *Drosophila melanogaster* populations, including empirically estimated variability of gene density, mutation rates, recombination rates, purifying and positive selection, across different historical demographic scenarios, to broadly assess the impact of selection on demographic inference using a genealogy-based approach. Our results indicate that the linked effects of selection minimally impact demographic inference for human populations, though it could cause mis-inference in populations with similar genome architecture and population parameters experiencing more frequent recurrent sweeps. We found that accurate demographic inference of *D. melanogaster* populations by ARG-based methods is compromised by the presence of pervasive background selection alone, leading to spurious inferences of recent population expansion which may be further worsened by recurrent sweeps, depending on the proportion and strength of beneficial mutations. Caution and additional testing with species-specific simulations are needed when inferring population

history with non-human populations using ARG-based approaches to avoid mis-inference due to the linked effects of selection.

Keywords

Demographic inference, ancestral recombination graph, background selection, selective sweeps, human population genomics, *Drosophila melanogaster*

Running title

Biases in ARG-based demographic inference due to selection

Introduction

Inferring the demographic histories of natural populations using population-genetic approaches provides fundamental insights into both evolutionary and ecological species dynamics. Demographic inference methods predict the effective size of ancestral populations at different time points (epochs) in the past which best explains observed levels and patterns of neutral diversity among sequenced individuals (reviewed in Beichman et al. 2018; Marchi et al. 2021). Inferred population history can be correlated with archaeological and anthropological findings to, for example, uncover the date and scale of population bottlenecks and migration events (Mellars 2006; Jouganous et al. 2017; Bergström et al. 2020; Hu et al. 2023). Additionally, demographic history is needed to establish a null model for the inference of selection across the genome (Teshima et al. 2006; Caicedo et al. 2007; Nielsen et al. 2007; Thornton and Jensen 2007; Jensen et al. 2019). As we accumulate increasingly rich genomic resources and apply advanced methods for inferring evolutionary parameters of populations, it is essential that we understand the interactive effects between selection and historic demography on contemporary variation, including the ways in which selection can impact empirical inference of demography using modern methods.

Demographic history has traditionally been inferred using statistical summaries of genomic variation such as the site frequency spectrum (SFS; Gutenkunst et al. 2009; Kamm et al. 2020; Excoffier et al. 2021), linkage disequilibrium (LD; Hayes et al. 2003; Harris and Nielsen 2013; Ragsdale and Gutenkunst 2017), by approximate Bayesian computation approaches that use multiple summary statistics to perform inference (Beaumont et al. 2002; Duchen et al. 2013; Kapopoulou et al. 2018; Johri et al. 2020), or by modelling local

genealogies (Li and Durbin 2011; Schiffels and Durbin 2014; Terhorst et al. 2017; Schiffels and Wang 2020). SFS-based approaches such as *∂a∂i* (Gutenkunst et al. 2009) and *fastsimcoal2* (Excoffier et al. 2021) fit the observed SFS at putatively neutral sites to the expected SFS under a range of pre-determined demographic scenarios using diffusion theory and coalescent simulations respectively. However, SFS-based methods assume independence between segregating sites, which is violated when using dense variant data due to linkage disequilibrium between proximal loci. In contrast, sequentially Markovian coalescent (SMC) methods typically employ Hidden Markov Models (HMM) to generate local genealogies between historical recombination breakpoints across the genome, allowing for the estimation of genome-wide coalescence rates correlating to a specific branch-length across the genealogies (Li and Durbin 2011; Schiffels and Durbin 2014; Schiffels and Wang 2020). The coalescence rate across the genome is inversely proportional to the number of recombining individuals at a given time point in a panmictic population, thus the inverse instantaneous coalescent rate (IICR) can be used as a scaling factor to estimate relative effective population size even if it is also influenced by population structure and migration rate and should thus be interpreted with care (Mazet et al. 2016; Chikhi et al. 2018). SMC-based methods are limited to very few individuals (2-16 haplotypes) which limits analysis of recent demography. *SMC++* (Terhorst et al. 2017) was subsequently developed to address this limitation, accommodating the analysis of hundreds of samples to improve inference of recent demography. Furthermore, emerging methods for capturing detailed genealogies across larger populations present additional promising opportunities to improve demographic inference by harnessing richer sources of genetic information (Kelleher et al. 2019; Speidel et al. 2019; Zhang et al. 2023).

Ancestral recombination graphs (ARGs) capture detailed genealogical relationships between individuals by recording all ancestral recombination/coalescence events throughout the genome (Hudson 1991; Griffiths and Marjoram 1997), typically employing HMM algorithms similarly to SMC-based methods (Li and Stephens 2003). Among recently developed scalable ARG approximation tools, *Relate* has emerged as the most accessible and widely used for demographic inference as of yet (*e.g.* Speidel et al. 2019; Almarri et al. 2021; Minadakis et al. 2023; Moorjani and Hellenthal 2023; Teterina et al. 2023; Ishigohoka and Liedvogel 2024). To approximate an ARG, *Relate* estimates topological features of local genealogies (trees) based on position-specific distance matrices calculated between

haplotypes. Following the estimation of local tree topology, *Relate* then estimates branch lengths using a Markov Chain Monte Carlo (MCMC) algorithm, initially assuming constant effective population size. Estimated branch lengths are then used to re-estimate historical population sizes from the inverse of the averaged coalescence rate at different epochs, allowing in turn for branch lengths to be re-estimated under the inferred demography. This process is iteratively repeated to co-estimate branch lengths and historical population size. The authors presented simulations of human data suggesting that *Relate* provides improved inference when compared to traditional local genealogy methods *MSMC* and *SMC++,* particularly for estimates of population size in the recent past (Speidel et al. 2019).

Although most demographic inference approaches perform accurately when tested with simulations assuming strict neutrality, indirect evolutionary forces, notably background selection (BGS) due to purifying selection (Charlesworth et al. 1993) and selective sweeps caused by positive selection (Maynard Smith and Haigh 1974), can produce genomic signatures that mimic or mask signals associated with historical population size change (Nicolaisen and Desai 2013), confounding inference of demography (Messer and Petrov 2013; Ewing and Jensen 2016; Schrider et al. 2016; Johri et al. 2021; Boitard et al. 2022; Smith and Hahn 2023). For example, a population bottleneck caused by non-selective factors (*e.g.*, founder effects following migration) will result in a population having fewer immediate ancestors in subsequent generations; a hard selective sweep for a specific allele similarly leads to fewer immediate ancestors in subsequent generations due to increased fitness for a small proportion of the population (Teshima et al. 2006). Accurate inference of historical population sizes therefore becomes a pre-requisite for, and pre-requires, accurate inference of selection - presenting a circular dilemma (Johri, Aquadro, et al. 2022).

While a few recent simulation-based approaches have developed methods to jointly infer selection and demography while accounting for the linked effects of selection (Sheehan and Song 2016; Johri et al. 2020; Johri et al. 2023), the majority of widely-used demographic inference approaches assume strict neutrality, *i.e.*, ignore the effects of direct and indirect selection, which is typically attempted by masking functional regions (*e.g.*, conserved elements, protein-coding exons) from the analysis. However, LD with selected loci often extends beyond functional regions, violating assumptions of neutrality by impacting genetic diversity at putatively neutral sites. Findings by Pouyet *et al.* (2018) suggest that the majority of the human genome is influenced by BGS and biased gene conversion which in turn biases

the SFS and demographic inference. Furthermore, Johri *et al.* (2021) demonstrated that BGS biases demographic inference when purifying selection is sufficiently pervasive, and spurious inference of recent population expansion is observed under population decline and equilibrium scenarios when using *fastsimcoal2* (Excoffier et al. 2021) and *MSMC* (Schiffels and Wang 2020). In contrast, Boitard *et al.* (2022), who explored the relative decrease in effective population size due to BGS and selective sweeps, suggest that the heterogeneity in the rate of coalescence across the genome, as generated by the linked effects of selection, can instead lead to spurious apparent decreases in effective population size. While previous studies clearly outline the theoretical biases introduced by BGS and selective sweeps to demographic inference, further analysis is needed to determine the extent to which the linked effects of selection impact empirical population genomic inference using ARG-based approaches where complex genome architecture and realistic effects of selection are present. Moreover, because we do not have precise estimates of the fraction and strength of new beneficial mutations in natural populations, it has been unclear if a reasonable extent of recurrent positive selection will significantly impact linked neutral variation in human and *D. melanogaster*-like populations.

Here, we test how demographic inference using an ARG-based approach is affected by selection in complex genomes using highly realistic simulations relevant to human and *Drosophila melanogaster*-like populations. To do so, we developed forward simulations incorporating variable gene density, recombination rates, and mutation rates, using realistic parameters for human and *D. melanogaster*-like chromosomes. We assess the effects on inference of different simulated demographic histories, with the effects of background selection and recurrent sweeps, parameterized by previously published estimates. Our findings provide insights into pitfalls that arise when ARG-based demographic inference approaches are applied to human populations and other species. Moreover, we collate previous estimates of positive selection parameters inferred from human and *D. melanogaster*-like populations, providing a potentially useful resource to other researchers (Table 1).

Results

Neutrality with variable recombination/mutation rates in human-like populations

To assess the performance of *Relate* with realistic human data we performed detailed forward simulations of a whole chromosome, informed by empirical estimates, under five demographic histories: constant population size, recent expansion, recent contraction, bottleneck/recovery, and repeated bottleneck/recovery (Figure 1). Mutation and recombination rate heterogeneity was modelled across the chromosome according to published maps (The International HapMap Consortium 2005; Francioli et al. 2015) for all simulations (except where stated), rather than simulating a constant rate. Under equilibrium, (*i.e.*, constant population size N), *Relate* inferred the correct demographic history and only slightly underestimated the population size up to $4N$ generations prior to sampling (Figure 1A). *Relate* effectively identified the presence of historical population expansion and bottleneck/recovery, however the precise timing and speed of the bottleneck and initial population growth/recovery was less accurate (Figure 1B, 1D). The presence, timing, and scale of population contraction was correctly identified under neutrality (Figure 1C), though a very recent brief increase in population size approximately 1000 generations from present was spuriously inferred prior to a sharp decrease toward the present (Figure S1); *i.e.*, *Relate* correctly estimates the direction of population size change and the approximate scale of the overall decline in population size over N_{anc} (the ancestral population size) generations, though overestimates the rapidity of the decline in very recent generations for human-like populations. For the repeated bottleneck/recovery scenario *Relate* identified the presence of the most recent bottleneck/recovery, although the distant bottleneck/recovery was not inferred and elevated variation between replicates was seen along with mis-inference prior to the most recent bottleneck/recovery (Figure 1E). Our results suggest that mutation and recombination rate heterogeneity across the chromosome by itself had a negligible effect on simulated diversity and subsequent demographic inference for human simulations (Figure 1).

Background selection with variable recombination/mutation rates in human-like populations

BGS effects were modelled along with variable recombination and mutation rates. The simulated 138.6 Mb genome consisted of an exon-intron-intergenic structure, with the first and last exons of all genes flanked by UTR regions; realistic gene density heterogeneity was modelled by varying the distance between adjacent genes (see Methods). All coding and UTR regions experienced selection, while 4.4% and 3.6% of all new mutations in introns and intergenic regions respectively experienced selection, matching empirical estimates (Siepel et al. 2005; Dukler et al. 2022). This resulted in 3.15% of genome-wide sites experiencing selection in simulations of human-like populations, consistent with recent estimates of 3.2% of the human genome experiencing selection (Dukler et al. 2022). Purifying selection was modelled by simulating the distribution of fitness effects (DFE) estimated by Huber *et al.* (2017) and Johri *et al.* (2023), as discrete uniform distributions describing proportions of new mutations drawn from four classes of selective effects. While both estimates predict similar proportions of effectively neutral mutations ($0 \leq 2N_{anc}s_d < 1$), the DFE estimated by Johri *et al.* (2023) comprises more mild ($1 \leq 2N_{anc}s_d \leq 10$) and moderately ($10 \leq 2N_{anc}s_d < 100$) deleterious mutations, and no strongly deleterious new mutations ($2N_{anc}s_d \geq 100$), compared to the Huber *et al*. (2017) DFE which predicts 22% of new mutations at conserved sites have strongly deleterious effects (see Methods).

The addition of purifying selection using the Huber *et al*. (2017) DFE reduced genetic diversity in neutral regions by 25-38% (Figure 1; Table S1, S2), with greater reductions for the population contraction scenario (38%) and lesser reductions for population expansion (25%). In comparison, we found the reduction of diversity under the Huber *et al.* (2017) DFE was greater than the 10-16% reduction under the DFE estimated by Johri *et al.* (2023; Figure 1; Table S1, S2). While the introduction of BGS led to an overall reduction in estimates of historical N inferred by *Relate* across all scenarios, as expected, the effect was relatively minor for human-like genomes and population parameters (Figure 1). While the true demographic history was largely recovered accurately in the presence of BGS in human-like populations, under the repeated bottleneck/recovery scenario there was considerable stochasticity between replicates in the more distant past as generations from present approached $4N_{anc}$ (Figure 1E).

Effects of recurrent selective sweeps on demographic inference in human-like populations

Selective sweeps were modelled in addition to variable recombination and mutation rates and BGS effects by introducing beneficial mutations as an additional class in the DFE describing the 3.15% of chromosome-wide sites experiencing selection. To model the DFE of new beneficial mutations, we first performed a literature survey and collated previous estimates of relevant parameters from human populations (Table 1). While several studies have estimated the fraction of beneficial substitutions (α) in human populations, there are only two recent studies that have obtained estimates of the proportion of new mutations that are beneficial (f_{pos}) and the distribution of selection coefficients of beneficial mutations (s_a). However, these estimates vary widely because the two studies use very different data. Moreover, Zhen *et al.* (2020) demonstrated that a wide range of f_{pos} (proportion of new beneficial mutations at selected sites) and $2N_{anc}\overline{s_a}$ (mean population scaled selection coefficient of beneficial mutations) combinations could explain the observed substitution rates at nonsynonymous sites in humans across a likelihood surface. Thus, to explore a wide parameter space that could be relevant for human and great ape populations (see Table 1), we tested a range of f_{pos} and $2N_{anc}\overline{s_a}$ combinations in a grid-like fashion (Table 2). Beneficial mutations were introduced in simulations with one of two different proportions of new mutations (f_{pos}= 0.01, f_{pos}= 0.001); with each proportion, we tested population-scaled selection coefficients ($2N_{anc}s_a$) drawn from an exponential distribution with mean values of 10, 100, and 500. Note that our parameter combination $f_{pos} = 0.01$ and $2N_{anc}\overline{s_a} = 10$ closely matches the empirical estimate provided by Zhen *et al*. (2020; Table 1).

The introduction of selective sweeps due to relatively rare beneficial mutations (f_{pos}= 0.001), alongside BGS and variable recombination/mutation rates, resulted in effects on genetic diversity and demographic inference ranging from negligible to severe, depending on both the selection parameters and simulated demographic scenario (Figure 2; Table S1, S2). When the strength of positive selection was mild ($2N_{anc}\overline{s_a} = 10$, f_{pos}= 0.001), negligible changes in nucleotide diversity and demographic inference were observed (Figure 2; Table S1, S2). Moderate, rare positive selection ($2N_{anc}\overline{s_a}$ = 100, f_{pos}= 0.001), resulted in moderate

reductions in nucleotide diversity relative to strict neutrality ($\frac{\pi_{neu}}{\pi_0}$ = 0.53-0.71; Figure 2; Table S1, S2), though demographic inference appeared largely unaffected.

Strong, rare positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.001) resulted in large reductions in nucleotide diversity ($\frac{\pi_{neu}}{\pi_0}$ = 0.17-0.49; Figure 2; Table S1, S2) and led to demographic mis-inferences using *Relate*. For all demographic scenarios, strong, rare positive selection ($2N_{anc}\bar{s_a}$ = 500; f_{pos}= 0.001) resulted in gradual population expansion (approximately five-fold) from $4N_{anc}$ to around $1.5N_{anc}$ generations in the past, with increased stochasticity between replicates (Figure 2). In addition, it caused recent population expansion to be spuriously inferred under constant population size (Figure 2A), and constant population size to be incorrectly inferred under population contraction (Figure 2C). The presence of a recent population expansion was recovered under strong, rare positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.001) as was a recent bottleneck, though large reductions in diversity and estimates of *N* were apparent, especially as generations from present approached $4N_{anc}$ (Figure 2; Table S1, S2).

The introduction of selective sweeps due to relatively common beneficial mutations (f_{pos}= 0.01), resulted in greater effects on genetic diversity and demographic inference than those produced under rarer occurrences of new beneficial mutations (f_{pos}= 0.001). Mild, common positive selection ($2N_{anc}\bar{s_a}$ = 10, f_{pos}= 0.01) led to only slight reductions in neutral diversity (π_{neu}) and did not impact demographic inference (Figure 3; Table S1, S2). Moderate strength, common positive selection ($2N_{anc}\bar{s_a}$ = 100, f_{pos}= 0.01) led to large reductions in π_{neu} and led to poorly inferred demographic histories, including high stochasticity as generations from present approach $4N_{anc}$ and failure to recover a recent population contraction (Figure 3C), similar to effects seen for populations experiencing strong, rare positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.001) parameter combination (Figure 2C). Strong, common positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.01) resulted in extreme reductions of π_{neu} and failure to recover demographic histories, with severely underestimated *N* and spuriously inferred growth regardless of the simulated scenario (Figure 3). In summary, population history is mis-inferred by *Relate* when using neutral sites from human-like genomes if the combination of frequency (f_{pos}) and strength of beneficial mutations ($2N_{anc}\bar{s_a}$) is sufficiently high ($2N_{anc}\bar{s_a}f_{pos} \geq 1$).

Relevance to human populations

While we have demonstrated that there exist parameter combinations where demographic inference will be biased in human populations in the presence of recurrent sweeps, it's important to evaluate whether they are relevant to human populations. To assess the tenability of positive selection parameters for human simulations, we used the ratio of divergence at coding and neutral sites ($\frac{div_{coding}}{div_{neu}}$), as used by (Soni et al. 2023), and the fraction of beneficial substitutions at coding sites, corrected to reflect values at nonsynonymous sites (α; Table 2; Table S3-S6), to compare our simulations with empirical findings (Table 1). In particular, our simulations indicated three combinations of parameters ($2N_{anc}\bar{s_a}$ = 10, f_{pos}= 0.001; $2N_{anc}\bar{s_a}$ = 100, f_{pos}= 0.001; $2N_{anc}\bar{s_a}$ = 10, f_{pos}= 0.01) that result in α values within the published range for humans (0-0.25; Table 2) as well as plausible ratios of divergence between coding and neutral sites (Table 2). Thus, positive selection causing classic hard sweeps in humans is likely to be relatively rare and not strong, consistent with previous literature. Importantly, simulations with the parameter combinations that exhibited plausible summary statistics for human populations (Table 2) did not impact demographic inference with *Relate* (Figure 2, 3).

In contrast, when common positive selection was moderate or strong ($2N_{anc}\bar{s_a}$ = 100, f_{pos}= 0.01; $2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.01), parameter combinations that caused mis-inference with *Relate* (Figure 3), resulted in higher divergence in coding regions experiencing selection compared to neutral intergenic regions ($\frac{div_{coding}}{div_{neu}}$ > 1; Table 2), indicating that it is unlikely that these parameter combinations are relevant to human populations. Simulations with strong, rare positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.001), which also caused mis-inference with *Relate* (Figure 2), exhibited plausible $\frac{div_{coding}}{div_{neu}}$ values (Table 2), with α values (0.40-0.47) exceeding published findings for humans and great apes, however this parameter combination may be relevant to rapidly evolving distant mammalian species with higher reported α values, such as mouse (Table 1).

316 ## Effects of selection on demographic inference in *Drosophila melanogaster*-like
317 genomes
318
319 To gauge the effects of BGS and selective sweeps on demographic inference with
320 more compact genomes, we used similar approaches as described above for humans, with
321 variable recombination/mutation rates, though simulated for *D. melanogaster*-like parameters
322 (Table S7). The *D. melanogaster* genome architecture was simulated with 48.7% of genome-
323 wide sites under selection, again distributed in exons, UTRs, introns (43% sites under
324 selection), and intergenic (35% sites under selection) regions, consistent with estimates by
325 Siepel *et al.* (2005). We found that providing an accurate mask input to *Relate* for
326 recalculation of genetic distance was essential for accurate demographic inference for
327 compact genomes where variants at a substantial proportion of total sites are removed from
328 the analysis (48.7%; see Supplementary Note S1; Figure S2).
329
330 Introducing recombination rate and mutation rate heterogeneity led to a moderate
331 improvement in demographic inference performance with *Relate* for *D. melanogaster*
332 genomes across all demographic scenarios; N was relatively under-estimated by 2-fold to 10-
333 fold in the distant past as generations from present approached $4N_{anc}$ for simulated
334 populations with constant recombination and mutation rate when compared to simulated
335 populations with realistic heterogeneous rates (Figure S3). For neutral simulations with
336 realistic heterogeneity in mutation and recombination rates, *Relate* recovered the correct
337 demographic histories under equilibrium (constant population size), population expansion,
338 contraction, and the most recent bottleneck/recovery (Figure 4). However, approximately 2-
339 fold variation in population size was inferred across the $4N$ period under equilibrium, as well
340 as spurious inference of population expansion from ~2500 generations ago until present with
341 high stochasticity between replicates (Figure S4). In addition, an additional bottleneck further
342 in the past was not recovered (Figure 4).
343
344 Introducing BGS effects, modelled by applying the DFE estimated for deleterious
345 mutations by Johri *et al.* (2020) to conserved sites (48.7% region-wide), reduced nucleotide
346 diversity in neutral regions by 46-70% (Figure 4; Table S8, S9) consistent with theoretically
347 predicted values by Charlesworth (2013). As expected, BGS led to substantially under-
348 estimated N and increased stochasticity across replicates in all scenarios as generations from

present approached $4N_{anc}$, with additional spurious variation in historical population size identified, even under equilibrium (Figure 4A). *Relate* recovered the recent population expansion and recent contraction, albeit with high stochasticity between replicates and underestimation of N beyond $1.5N$ generations ago for the population contraction scenario (Figure 4B, 4C). BGS caused recent bottlenecks to be entirely mis-inferred as recent expansions with population sizes under-estimated by up to two orders of magnitude prior to the bottleneck (Figure 4D, 4E).

Selective sweeps were modelled by introducing beneficial mutations at conserved sites with varying frequency (f_{pos}) and mean positive selection coefficient ($2N_{anc}\overline{s_a}$), informed by widely varying empirically estimated values (Table 1), with BGS present. In contrast to human populations, several previous studies have estimated the relevant parameters in *D. melanogaster* populations (Table 1) though estimates vary widely, possibly due to methodological differences. We therefore simulated all estimated parameter combinations. The parameter combination (f_{pos}= 0.005, $2N_{anc}\overline{s_a}$ = 20) simulating common, weakly beneficial mutations reported by Keightley *et al.* (2016) had minimal additional effects on reducing diversity and impacting demographic inference, when compared to simulations with BGS only (Figure 4). The Campos *et al.* (2017) parameter combination (f_{pos}= 0.0002, $2N_{anc}\overline{s_a}$ = 250) simulating infrequent, moderately beneficial mutations had a larger effect on reducing diversity, particularly in the population expansion scenario (Figure 4B), though demographic inference was only minorly impacted by the addition of sweeps. The Jensen *et al.* (2008) parameters (f_{pos}= 1×10^{-6}, $2N_{anc}\overline{s_a}$ = 10000) simulating very rare, very strongly beneficial mutations had only moderate effects on reducing diversity and impacting demographic inference, though stochasticity between replicates was considerably higher in all scenarios, particularly when a recent population bottleneck occurred $1N$ generations ago (Figure 1D, 1E). However, simulated data from all three empirically estimated combinations of positive selection parameters produced summary statistics that did not match well with empirical findings: simulated $\frac{\pi_{neu}}{\pi_0}$ ranges from 0.54-0.27 (Figure 4, Table S8, S9), while Elyashiv et al. (2016) estimated $\frac{\pi_{neu}}{\pi_0}$ = 0.11-0.23; simulated $\alpha \leq 0.19$ (Table 2; Table S10) while published estimates report α = 0.40-0.71 (Table 1); and finally, while the simulated $\frac{div_{coding}}{div_{neu}} \leq 0.52$ (Table 2, S11-13), which is plausible in real populations

experiencing selection, Johri *et al.* (2020) empirically estimated a higher ratio for *D. melanogaster* ($\frac{div_{coding}}{div_{neu}}$ = 0.87).

The parameter combination of f_{pos} = 0.005 (Keightley et al. 2016) and $2N_{anc}\overline{s_a}$ = 250 (Campos et al. 2017) simulating common, moderately beneficial mutations, which was never inferred by any one previous study, produced summary statistics (α, $\frac{div_{coding}}{div_{neu}}$, $\frac{\pi_{neu}}{\pi_0}$; Table S8-S13) that best matched empirical estimates, which could potentially suggest that it most accurately models the effects of selection in real *D. melanogaster* populations. This parameter combination resulted in total reduction of neutral diversity ($\frac{\pi_{neu}}{\pi_0}$ = 0.09-0.28; Figure 4, Table S8, S9) consistent with findings by Elyashiv *et al.* (2016), as well as α (0.40-0.47; Table 2, S10) and $\frac{div_{coding}}{div_{neu}}$ (0.91-0.98; Table 1, S11-S13) best matching empirical estimates (Keightley et al. 2016; Campos et al. 2017; Lange and Pool 2018; Johri et al. 2020; Zhen et al. 2021) of all simulated *D. melanogaster* parameter combinations. Selective sweeps introduced by this parameter combination (f_{pos}= 0.005, $2N_{anc}\overline{s_a}$ = 250), with BGS present, led to failure by *Relate* to effectively infer demographic histories, with recent population expansion spuriously inferred in all simulated scenarios, severe underestimates of *N*, and extremely high stochasticity between replicates, even under equilibrium (Figure 4).

Discussion

Implications for human-like populations

Our analysis of realistic simulated populations (not accounting for structure) indicated that signatures of the linked effects of selection have the potential to cause mis-inference of historical population size using *Relate* when purifying and/or positive selection are sufficiently strong and pervasive. However, demographic inference with empirically supported parameters for human populations was shown to not be strongly impacted by selection under most tested demographic models. While BGS introduced by the DFEs estimated by Huber et al. (2017) and Johri et al. (2023) had moderate effects reducing genomic variation for simulated human populations ($B \approx$ 0.7-0.9; Table S1, S2), consistent with previous estimates (B = 0.83; Murphy et al. 2023), our findings suggest the effect of BGS on demographic inference for humans is minimal, inducing only slight under-estimation of historical population size without causing spurious inferences. Selective sweeps from

411 recurrent positive selection similarly had little effect on demographic inference with *Relate*
412 when restricted to empirically supported parameters for human populations (Eyre-Walker and
413 Keightley 2009; Galtier 2016; Uricchio et al. 2019; Zhen et al. 2021), only potentially
414 impacting the inferred history for populations that experienced repeated bottlenecks.
415
416 For simulated populations experiencing recurrent selective sweeps with parameters
417 ($2N_{anc}\bar{s_a}$, f_{pos}, α) consistent with findings for human and other great apes with similar
418 selection (McManus et al. 2015; Cagan et al. 2016; Castellano et al. 2019) and population-
419 genetic dynamics (Besenbacher et al. 2019; Schmidt et al. 2019), we found the greatest driver
420 of population history mis-inference using *Relate* was not selection but rather repeated
421 fluctuations in historical population size. We found that when a simulated population
422 experienced repeated bottlenecks *Relate* was only able to detect the most recent bottleneck,
423 and exhibited poor inference further in the past, raising concerns of its accuracy when applied
424 to many real populations. Great ape taxa including orangutans, gorillas, chimpanzees, and
425 bonobos have been reported by some to experience a recent bottleneck (McManus et al.
426 2015) and repeated population bottlenecks (Prado-Martinez et al. 2013), suggesting that
427 perhaps other demographic inference methods may be better suited at recovering repeated
428 bottlenecks. Previous archaeological evidence suggests that we expect natural human
429 populations underwent multiple bottlenecks (reviewed in Hawks et al. 2000; Nielsen et al.
430 2017). Our results suggest that while recent bottlenecks humans experienced may be
431 observable (*e.g.,* Gravel et al. 2011; Liu and Fu 2015; Terhorst et al. 2017; Almarri et al.
432 2021; Ragsdale et al. 2023), *Relate* and similar methods may not effectively identify
433 additional distant bottlenecks that may have taken place (Hawks et al. 2000).
434
435 While recurrent selective sweeps using parameters supported by findings for humans
436 ($2N_{anc}\bar{s_a}$ = 10, f_{pos}= 0.001; $2N_{anc}\bar{s_a}$ = 100, f_{pos}= 0.001; $2N_{anc}\bar{s_a}$ = 10, f_{pos}= 0.01) are not
437 strong and pervasive enough to meaningfully impact demographic inference with *Relate*, we
438 found that stronger positive selection ($2N_{anc}\bar{s_a}$ = 500, f_{pos}= 0.001) could lead to mis-
439 inference, which may be relevant to other mammalian species. Lemur (*Propithecus*
440 *coquereli*; $\alpha = 0.75$; Galtier 2016), mouse (*Mus musculus;* α = 0.41, 0.51; Halligan et al.
441 2013; Zhen et al. 2021), and rabbit populations (*Lepus granatensis, Oryctolagus cuniculus;*
442 Carneiro et al. 2012; Galtier 2016) have been estimated to have a high fraction of beneficial
443 relative to neutral substitutions at or above levels which caused *Relate* to produce spurious

444 population expansions with substantial underestimations of historical population size (α =
445 0.40-0.47; Figure 4, Table S13). Mice and rabbits have higher gene densities than humans
446 (Carneiro et al. 2014; Frankish et al. 2019) and larger effective population sizes (Carneiro et
447 al. 2011; Phifer-Rixey et al. 2012), which would exacerbate the impact of both BGS and
448 sweeps on demographic inference, though they also exhibit higher rates of recombination
449 relative to mutation rates (Jensen-Seaman et al. 2004; Booker et al. 2017; Lindsay et al.
450 2019). Overall, our results indicate that selection has the potential to cause mis-inference of
451 population history with *Relate* in some mammalian populations, though additional
452 simulations modelling alternative genomes/populations are needed to identify the precise
453 effects for non-human species. In cases where ARG-based inference using all
454 intergenic/intronic sites are expected to be impacted due to the linked effects of selection,
455 masking only selected regions may be insufficient and instead it may also be necessary to
456 mask adjacent regions most affected by BGS and selective sweeps.
457

458 ## Implications for *Drosophila*-like species with compact genomes

459

460 Our results suggest that demographic inference with *Relate* is susceptible to mis-
461 inference for species with compact genomes and pervasive selection, as modelled by our *D.*
462 *melanogaster* simulations. BGS due to purifying selection introduced according to the DFE
463 estimated by Johri et al. (2020) substantially impacted demographic inference and decreased
464 historical population size estimates by approximately ten-fold $1N_{anc}$ generations ago.
465 Interestingly, the recent approximately two-fold expansion in the most recent $1N_{anc}$
466 generations inferred by *Relate* for a population of constant size experiencing BGS (Figure
467 4A) matches well empirical estimates of two- to four-fold growth in the Zambian population
468 of *D. melanogaster* (Ragsdale and Gutenkunst 2017; Kapopoulou et al. 2018), raising
469 questions of whether these findings of population growth may be equally explained by a
470 population of constant size experiencing the linked effects of selection. In addition, the
471 effects of BGS in *D. melanogaster* resulted in the failure of *Relate* to recover some
472 demographic shifts including population bottlenecks. Various non-Zambian *D. melanogaster*
473 populations experienced severe bottlenecks, such as during transcontinental colonization
474 events (Thornton and Andolfatto 2006; Laurent et al. 2011; Duchen et al. 2013; Kapopoulou
475 et al. 2018; Arguello et al. 2019). Our results indicate that BGS alone is sufficient to entirely

476 obscure the presence of population bottleneck when inferred by *Relate*, instead spuriously
477 inferring a recent population expansion.
478
479 Selective sweeps will likely also exacerbate mis-inference in *D. melanogaster*,
480 however, the extent to which this occurs may be unclear, as our simulated parameters
481 matching empirical estimates varied widely in their frequency/intensity of selection and their
482 effect on demographic inference. The mismatch we found between summary statistics (α,
483 $\frac{div_{coding}}{div_{neu}}, \frac{\pi_{neu}}{\pi_0}$; Table S8-S13) from our simulations using widely varying published f_{pos} and
484 $2N_{anc}s_a$ combinations (Jensen et al. 2008; Keightley et al. 2016; Campos et al. 2017; Zhen et
485 al. 2021) compared to the same summary statistics reported in the literature (Elyashiv et al.
486 2016; Keightley et al. 2016; Campos et al. 2017; Johri et al. 2020; Zhen et al. 2021) raises
487 concerns of consensus regarding which parameters best describe real populations. This is
488 likely in part due to the likelihood surface as presented in Zhen *et al.* (2021), which
489 highlighted the difficulty in accurately inferring the fraction of new beneficial mutations and
490 their strength jointly. Another possibility is that the beneficial DFE in real *Drosophila*
491 populations may deviate from the expected exponential distribution (Fisher 1930; Gillespie
492 1984; Orr 2003) modelled in our simulations and inferred by previous studies, and may
493 instead be better fit by alternative distribution domains (Joyce et al. 2008; Seetharaman and
494 Jain 2014). Similarly, while we assumed that all new mutations are semidominant, increased
495 recessivity or dominance of selected mutations could result in different effects on genomic
496 variation (Teshima and Przeworski 2006). Thirdly, it is possible that the large rescaling of *D.*
497 *melanogaster*-like simulated populations that was necessary for computational tractability
498 (see Methods; Table S14) inadvertently changed some dynamics of recombination by under-
499 modelling crossover interference which increases the probability of introducing more than
500 one crossover events in a single progeny; in such a case, increased linkage effects between
501 distal sites may occur which could in turn influence selection dynamics. Generally, increased
502 linkage and a higher mutation rate per diploid genome is likely to cause interference between
503 selected alleles, leading to a possible decrease in the efficacy of positive selection.
504 Alternatively, it is possible that diminished effects of recombination could lead to increased
505 effects of background selection, resulting in a lower effective population size, and therefore
506 reduced efficacy of positive selection (*i.e.*, lower values of $N_e s_a$). While we optimized our
507 simulations based on the best fit to empirical findings within the realms of tractability, the
508 precise impacts of rescaling are not fully known and further research is needed to clarify its

effects and how best to preserve expected population genetics dynamics. Finally, positive selection in *D. melanogaster* and other populations may be primarily due to polygenic selection or selection on standing variation, in which case our assumed model of recurrent classic sweeps might not be sufficient to fit the empirical observations. Because positive selection on new mutations causes the strongest hitchhiking effect, it is unlikely that other models of positive selection will result in more severe mis-inference than that simulated in our study.

ARG-based demographic inference methods including *Relate* were initially developed to investigate human ancestry, however, these approaches have recently been applied to other organisms including insects, nematodes, and plants. Pope *et al.* (2023) reported population histories using ARG information approximated by *Relate*, inferring recent expansions in squash bee (*Eucera pruinosa*) populations following population splits coinciding with historical events. Our *D. melanogaster* results under the population bottleneck scenario (Figure 4D) suggest mis-inference due to BGS effects could partially explain some of their findings, if their methods for masking sites putatively linked to selected loci were too permissive (Pope et al. 2023; see Supplementary Note S2 for details). Beyond insects, *Relate* has recently been used to infer the demographic histories of self-fertilizing (selfing) and outcrossing *Caenorhabditis* nematodes (Teterina et al. 2023), and a grass species (*Brachypodium distachyon*) with high rates of selfing (Minadakis et al. 2023). Selfing increases the effects of background selection (Charlesworth et al. 1993; Charlesworth 2003) and selective sweeps (Hedrick 1980; Hartfield and Bataillon 2020), dramatically affects heterozygosity (Charlesworth and Charlesworth 1987) and linkage (Wright et al. 2013; Noble et al. 2021), and ultimately, changes the genealogical patterns from which *Relate* derives its demographic estimates (Nordborg and Donnelly 1997; Nordborg 2000; Strütt et al. 2023). Our results underline the need for caution when applying methods, such as *Relate*, that were designed for and tested with human parameters, to species with differing genome architecture experiencing disparate population dynamics.

Demographic inference with complex genomic architecture

While a moderate improvement in inference was seen by effectively modelling recombination and mutation rate variability across *D. melanogaster* genomes (Figure S3),

542 simulating recombination and mutation rate heterogeneity did not have a large impact on
543 demographic inference with *Relate* across human genomes, unlike observations in some
544 previous studies (Sellinger et al. 2021; Boitard et al. 2022). Sellinger *et al.* (2021) found that
545 local variation in mutation rate could lead to demographic mis-inference and spurious
546 apparent population decline using SMC methods. Realistic variation in local mutation and
547 recombination rates has been shown to impact demographic inference when using SFS-based
548 methods (Soni et al. 2024). In addition, extreme rates of recombination (cM/Mb ≤ 0.01,
549 cM/Mb ≥ 5) have been shown to result in mis-inference of historical population size as
550 calculated from nucleotide diversity using *Relate* (Novo et al. 2022). However, Upadhya and
551 Steinrücken (2022) suggested that demographic inference performance with *Relate* improves
552 when recombination rates are variable due to the presence of low recombination rate regions
553 from which exact local genealogies can be estimated, which could in part explain the
554 moderate improvement provided by modelling mutation and recombination rate
555 heterogeneity in *D. melanogaster* simulated populations (Figure S3). It is unclear whether
556 recombination and mutation rate heterogeneity may have more substantial effects in different
557 species, as recent reports suggest masking low-recombination regions in songbird species
558 may improve demographic inference with *Relate* (Ishigohoka and Liedvogel 2024).
559
560 The feasibility of finding neutral sites for demographic inference, which have not
561 been substantially affected by nearby linked sites under selection, is essential for accurate
562 genetic inference. Messer and Petrov (2013) found a 5-fold recent expansion could be mis-
563 inferred by DFE-α using synonymous sites under equilibrium due to BGS alone, with more
564 severe mis-inference (8.8-fold expansion) when selective sweeps were also modelled, with
565 parameters ($2N_{anc}\bar{s_a}$ = 20, α = 0.18) similar to our empirically supported human parameters
566 that did not lead to demographic mis-inference with *Relate* (Table 2; Figure 3A). Similarly,
567 Schrider *et al.* (2016) revealed that when substantial fractions of neutral sites used for
568 demographic inference are affected by proximal strong selective sweeps, population history
569 can be severely impaired, concluding that it is essential to identify neutral sites sufficiently
570 unlinked from recent selective sweeps to use for demographic inference. We found that
571 neutral sites in intronic and intergenic regions are indeed sufficiently unlinked from recent
572 selective sweeps to not affect inference in humans when simulating empirically supported
573 positive selection parameters (Table 2; Figure 2, 3). However, our results suggest that the
574 linked effects of selection (notably BGS) in compact genomes such as *D. melanogaster*
575 impact demographic inference when using sites in neutral regions (Figure 4). Thus, for

populations with compact genomes experiencing selection such as *D. melanogaster*, or in more extreme cases such as viruses where neutral sites unlinked to selected loci do not exist (Howell et al. 2023; Terbot et al. 2023), population history may be best inferred using joint inference approaches (*e.g.* Johri et al. 2020) that account for the linked effects of selection and demography simultaneously.

The effect of selection on gene genealogies

Our results did not reveal that the effects of selection at linked sites lead to spurious inference of recent population decline, as reported by Boitard *et al.* (2022), when tested in realistic populations. The authors suggested the conflict between their findings (Boitard et al. 2022) and findings by Johri *et al.* (2021) was due to modelling of genomic regions without variable gene density or recombination and mutation rate variability; however, the simulated populations in our current analysis addresses these concerns and we largely corroborate findings by Johri *et al.* (2021). This is because Boitard *et al*. (2022) modelled the heterogeneity in effective population size across the genome while accounting for only the rescaling effects of N_e due to background selection and sweeps. However, both processes lead to not only a decrease in N_e, but also affect the underlying gene genealogy, which mimics the genealogy of an expanding population and skews the SFS (Kim 2006; Nicolaisen and Desai 2013), and this is not accounted for by Boitard *et al*. (2022). Similarly, while Cousins *et al.* (2024) have recently implemented a *PSMC* method that accounts for the relative decrease in N_e across the genome caused by the linked effects of selection, such a method is unlikely to account for impacts of selection on local genealogies and thus may be ineffective when selection is pervasive, especially in species with compact genomes or those with little/no recombination. A final caveat of our analysis is the assumption of panmixia that may not reflect the reliability of ARG-based inference in the presence of population structure, which Boitard *et al.* (2020) showed could overshadow the impact of linked effects of selection on the IICR. Given reports that humans, for example, have experienced considerable population structure throughout their demographic history (Schiffels and Durbin 2014; Schlebusch et al. 2017; Ragsdale and Gravel 2019; Ragsdale et al. 2023), the effect of selection on inference in structured populations may be a promising avenue for additional research.

Following our analysis, the question arises: how generalizable are our findings using *Relate* to analysis with alternative ARG-based demographic inference approaches? Systematic ARG estimation using exhaustive likelihood approaches is computationally impractical (McVean and Cardin 2005), and as a result, *Relate* and similar methods use heuristics and HMMs to minimise redundant calculations (Li and Stephens 2003), greatly improving scalability with large sample sizes at the cost of precision. *Relate* (Speidel et al. 2019) and *tsinfer/tsdate* (Kelleher et al. 2019; Wohns et al. 2022) share biases when estimating recent and ancient coalescence times (Brandt et al. 2022; Upadhya and Steinrücken 2022), where accuracy drops when compared to *e.g*. the less scalable, more precise *ARGweaver* (Rasmussen et al. 2014), likely due to factors such as under-sampling of marginal trees (Deng et al. 2021). *Relate* has been shown in other studies to produce spurious inferences in very recent generation times, suggested to be due to a lack of sufficient mutations accumulating in short time intervals, as is needed for accurate inference using coalescent approaches, which alternate methods may improve upon (Santiago et al. 2020). Among the scalable approaches, *Relate* is relatively more affected by under-sampling of marginal trees (Deng et al. 2021), and preliminary reports indicated that demographic inference with *Relate* is less accurate than *tsinfer*/*tsdate* for the most recent ~1000 generations, though *Relate* outperforms *tsinfer*/*tsdate* further in the past (Fan et al. 2023) while having improved coalescence time estimation than the recently developed *ARGneedle* (Ignatieva et al. 2023; Zhang et al. 2023). Thus, while minor differences in mis-inference due to effects of selection may occur across ARG-based methods, as the genealogy at linked neutral sites itself is altered by selection (Zeng and Charlesworth 2011; Nicolaisen and Desai 2013) we expect similar biases when performing demographic inference using branch lengths from alternative ARG approximations.

Conclusions

ARG-based inference of population history remains a promising avenue for developing new methods (Brandt et al. 2024; Lewanski et al. 2024), however, *Relate* appears to be similarly impacted by the linked effects of selection as traditional SMC and SFS-based approaches (Johri et al. 2021). As a result, sufficiently pervasive positive and purifying selection leads to mis-inference of historical population size, most notably when moderately severe bottlenecks have taken place. While the effects of selection on demographic inference of human populations appears minimal, caution and additional testing with specific

642 simulations are needed when analysing compact genomes with pervasive background
643 selection and/or sufficiently high rates of positive selection, such as *D. melanogaster*, and
644 populations with complex demographic histories.
645
646 Quantifying expected interactive effects between demography and selection on
647 population variation remains necessary through the development of theoretical advances
648 (Friedlander and Steinrücken 2022) and joint-inference approaches that account for the
649 effects of selection at linked sites (Sheehan and Song 2016; Johri et al. 2020; Johri, Eyre-
650 Walker, et al. 2022;). In addition, novel scalable ARG approximation with uncertainty
651 estimates for coalescence times has recently been theorized (Wong et al. 2023), which would
652 enable the development of improved methods that can account for temporal uncertainty
653 (Brandt et al. 2024). While ARG-based demographic inference methods rapidly advance in
654 terms of scalability and sophistication, it is essential we develop in tandem an understanding
655 of their limits when applied to complex and varied empirical populations.

656 Methods

657 Simulation of chromosome structure

658
659 Simulations were performed using *SLiM* 4.0.1 (Haller and Messer 2023) with ten
660 replicates for each evolutionary scenario. Simulations of human-like populations were
661 performed assuming an ancestral population size (N_{anc}) of 20,000 diploid individuals (Sherry
662 et al. 1997; Huff et al. 2010), a mean mutation rate of 1.25×10^{-8} (Kong et al. 2012), and a
663 mean recombination rate of 1×10^{-8} per site/generation (Dumont and Payseur 2008; Kong
664 et al. 2010). All *SLiM* simulations were scaled appropriately (population sizes and the
665 number of generations were reduced, by a factor of 10 for human, while mutation and
666 recombination rates were multiplied by the same factor) resulting in a scaled N_{anc} equal to
667 2000 (Comeron and Kreitman 2002; Hoggart et al. 2007; Kim and Wiehe 2008; Uricchio and
668 Hernandez 2014). All simulations, human and *Drosophila*, included a burn-in period of
669 $14N_{anc}$ generations, with an output of 50 diploid individuals (unless specified) in the VCF
670 format. The number of output diploid individuals (50) used for demographic inference with
671 *Relate* was chosen following testing with neutral simulations with human-like parameters
672 which indicated that using 10 or fewer samples led to minor mis-inference in recent
673 generations, while increasing the number of individuals beyond 50 up to 250 did not lead to

substantial improvements in inference (Figure S5, S6). Note, we attempted to run *Relate* with 1000 samples, though results are not shown as demographic inference did not finish within 86 hours.

Drosophila: Simulations of *Drosophila melanogaster*-like populations were performed similarly, assuming $N_{anc} = 1 \times 10^6$ (Keightley et al. 2014), a mean mutation rate of 3×10^{-9} (Keightley et al. 2014), a mean recombination rate of 1×10^{-8} (Comeron et al. 2012). Given computational feasibility constraints of modelling *D. melanogaster* populations using forward-simulations with selection, we tested of a range of scaling factor and genome-size combinations (Table S14); as a result, we scaled simulations by a factor of 300 for a 5 Mb euchromatic region, which maximised the length of the region and resulted in empirically supported background selection effects (modelled as described below) in neutral regions: $B = 0.4$ simulated under constant population size ($B = 0.45$; estimated by Charlesworth 2013).

Simulating the genome architecture

To approximate the genome architecture of the average human chromosome, we interspersed 797 genes with size and features approximately matching the mean values obtained across human protein-coding genes (Mignone et al. 2002; Mayr 2016; Piovesan et al. 2019). In simulated human genomes, all genes had 11 exons of 311 bp, 10 introns of 6938 bp, a 200 bp 5′ UTR, and an 800 bp 3′ UTR, with a mean intergenic length across the chromosome of 100 Kb. The 5′ UTR was placed immediately before the first exon in each gene; the 3′ UTR was placed immediately after the final exon in each gene; all genes had forward orientation.

To capture realistic variability of gene density, human gene features for chromosome 7 were downloaded from NCBI (https://www.ncbi.nlm.nih.gov; Sayers et al. 2023) using the following command: "*7[Chr] AND "Homo sapiens"[Organism] AND ("has ccds"[Properties] AND alive[prop])*". The chromosome was divided into 1 Mb bins within which the number of genes were counted. The mean (5.47) and standard deviation (6.07) for the gene count across the 1 Mb bins were used to produce a normal distribution (limited to avoid negative gene counts and counts resulting in total genic lengths exceeding 1 Mb). We

707 developed a script that recursively sampled 138 integers from this normal distribution
708 (setSeed = 1), corresponding to the number of genes per 1 Mb bin, until the first occurrence
709 of a cumulative sum (total gene count) of exactly 797 was output and saved to generate gene
710 densities in simulations (see
711 https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/gene_density_gener
712 ation.slim). Within *SLiM*, the saved 138 integers were used to generate corresponding counts
713 of gene structures in 1 Mb bins, by varying the intergenic lengths within each bin (Figure S7),
714 with 300 Kb intergenic regions added to each end of the chromosome (*n.b.* the intergenic
715 lengths are the only genomic elements that vary in length across the chromosome, and the
716 chromosome structure was identical across all replicates and conditions). This produced a
717 chromosome with a size of 138.6 Mb and genic regions (introns and exons) spanning 58 Mb
718 (41.9%), which is consistent with the 1.27 Gb estimate of total protein coding gene length
719 including introns (Piovesan et al. 2019), which spans 40.9% of a 3.12 Gb human genome
720 (Nurk et al. 2022; Rhie et al. 2023).
721
722 *Drosophila:* The genome architecture of *Drosophila melanogaster* was simulated
723 similarly, however, due to computational feasibility constraints, a 5 Mb euchromatic region
724 was modelled based on features from chromosome arm 2R rather than a full chromosome.
725 All 565 genes had 5 exons of 314 bp each, 4 introns of 999 bp each, a 266 bp 5′ UTR and a
726 382 bp 3′ UTR, with a mean intergenic length across the region of 2654 bp (Long and
727 Deutsch 1999; Dos Santos et al. 2015; Table S7). The mean (14.69) and standard deviation
728 (7.54) for gene count across 100 Kb bins was used to generate a normal distribution of gene
729 counts in 50 variably dense 100 Kb bins totalling 565 genes over 5 Mb (see above and
730 https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/gene_density_gener
731 ation.slim), with 20 kb intergenic regions added to the ends of the region. This produced a
732 5.04 Mb region comprised of 17.70% coding sites, matching the chromosome 2R annotation
733 for the BDGP6 assembly (17.68%; Dos Santos et al. 2015).
734

735 ## Modelling heterogeneity in recombination and mutation rates

736
737 The sex-averaged mutation rate map for human chromosome 7 generated by Francioli
738 *et al.* (2015) was used as input for the mutation rates, scaled so that the average mutation rate
739 (μ) across the chromosome was 1.25×10^{-8} per site/generation (1.25×10^{-7} after scaling;

740 see Figure S8 for distribution of mutation rates). As Francioli *et al*. (2015) provided the rates
741 for each base pair transition (*i.e*., A to T, T to G, etc), the nucleotide composition for each
742 corresponding 1 Mb interval was calculated from the GRCh37 reference (BioProject
743 PRJNA31257) and multiplied by the sum of the rates for each given nucleotide. Finally, the
744 rates for each nucleotide were combined to produce μ (see
745 https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/mut_map.sh). The
746 population-specific recombination rate map for chromosome 7 generated for the YRI
747 population in the HapMap project (The International HapMap Consortium 2005) was used as
748 input for the recombination rates. The shared length (146.04 Mb) reported in both the
749 chromosome 7 recombination rate map and mutation rate map (7.01 Mb to 153.05 Mb) was
750 extracted, and all interval lengths were reduced by approximately 5.1% to have a cumulative
751 total length of 138.6 Mb rather than 146.04 Mb, so that the maximal region for which
752 recombination rates and mutation rates are reported for chromosome 7 could be modelled in
753 our simulations.
754
755 *Drosophila*: The mean mutation rate simulated was 3×10^{-9} per site/generation
756 (Keightley et al. 2014) prior to scaling and was varied across the region in 50 kb bins. The
757 mutation rates in each 50 kb bin were drawn from a normal distribution with a coefficient of
758 variation of 0.479, corresponding to the coefficient of variation for divergence at 4-fold
759 degenerate sites (d_S) between *D. simulans* and *D. melanogaster* in the middle section of the
760 2R chromosome arm (Mackay et al. 2012). The mean sex-averaged recombination rate
761 simulated (crossover probability per bp = 1×10^{-8}) was modelled on a genome-wide
762 estimate (Comeron et al. 2012), though rates were varied in 50 kb bins. The recombination
763 rates in each 50 kb bin were taken from the rates in 100 Kb bins from 10 Mb to 20 Mb
764 reported by Comeron *et al.* (2012) on the *Drosophila melanogaster* recombination rate
765 calculator v2.30 (Fiston-Lavier et al. 2010).
766

767 ## Simulating mutational fitness effects

768
769 Functional regions, coding sequences and UTRs, experienced purifying selection
770 modelled by a distribution of fitness effects (DFE) comprised of four non-overlapping
771 uniform distributions representing the effectively neutral ($0 \leq 2N_{anc}s_d < 1$), weakly
772 deleterious ($1 \leq 2N_{anc}s_d < 10$), moderately deleterious ($10 \leq 2N_{anc}s_d < 100$), and

773 strongly deleterious ($100 \leq 2N_{anc}s_d < 2N_{anc}$) class of mutations, such that f_0, f_1, f_2, and f_3
774 proportion of all new mutations belonged to each class respectively. Here, s_d represents the
775 selective disadvantage of mutant homozygotes relative to the wildtype, and all mutations
776 were assumed to be semidominant. When simulating recurrent sweeps, an f_{pos} proportion of
777 all new mutations in functional elements was assumed to be beneficial, while $1 - f_{pos}$
778 proportion was effectively neutral and deleterious. Beneficial mutations were assumed to be
779 semidominant and to follow an exponential DFE with the mean selective advantage of mutant
780 homozygotes relative to wildtypes denoted by $\overline{s_a}$. Semidominance of all mutations was
781 assumed because previous estimates used to parameterize selection in our simulations were
782 inferred under the same assumption.
783
784 The deleterious DFE estimated by Huber *et al.* (2017) for the Yoruba population was
785 used to inform the proportions of each mutation class in human simulations where BGS was
786 present: f_0 = 0.51; f_1 = 0.14; f_2 = 0.14; f_3 = 0.21; the alternative Johri *et al.* (2023) DFE
787 tested (Figure 1) was: f_0 = 0.50; f_1 = 0.20; f_2 = 0.30; f_3 = 0. Simulations that modelled
788 recurrent selective sweeps included beneficial mutations with varying values of f_{pos}, and the
789 population-scaled selection coefficients of advantageous mutations ($2N_{anc}\overline{s_a}$), as provided in
790 Table 2.
791
792 To model functional elements like regulatory elements present in intergenic and
793 intronic regions, we assumed that conserved introns and intergenic regions experience
794 selection. Previous findings report the functional proportion of the human genome
795 experiencing purifying selection is approximately between 3.2%-8.2% (Siepel et al. 2005;
796 Rands et al. 2014; Dukler et al. 2022). Approximately 60% of mutations with a deleterious
797 fitness effect across the genome have been found to be located in intronic and intergenic
798 regions (Siepel et al. 2005; Dukler et al. 2022). According to Table 2 in Dukler *et al.* (2022),
799 3.6% of intergenic regions, and 4.4% of intronic regions are conserved (i.e., containing
800 elevated proportions of selected sites). Thus, instead of modelling the precise locations of
801 regulatory elements, we assumed that 4.4% and 3.6% of all new mutations in intronic and
802 intergenic regions were selected, following the same DFE as in coding and UTR regions,
803 while the remaining majority of intergenic (96.4%) and intronic (95.6%) mutations were kept
804 neutral (*n.b.* we did not explicitly model conserved non-coding element structures within
805 introns and intergenic regions). Combined with the exon and UTR regions (2.54% of the

genome) for which all new mutations possessed selection coefficients sampled from the DFE, the mean total proportion of new mutations across the genome with selection coefficients sampled from the DFE was 6.43%; given an effectively neutral (f_0) proportion of 0.51, the conserved proportion of the genome ($p_{conserved}$) was 3.15%.

Drosophila: The structure of functional elements for *Drosophila* simulations was modelled as above, with constrained proportions of intron (43%) and intergenic (35%) sites, constrained proportion of all sites (48.7%) matching previous estimates (Andolfatto 2005; Siepel et al. 2005; Table S7). The deleterious DFE estimated by Johri *et al.* (2020) for the Zambian population was used for *Drosophila* simulations: f_0 = 0.25; f_1 = 0.49; f_2 = 0.04; f_3 = 0.22. In total, 64.9% of new mutations across the genome sampled fitness effects from the DFE and $p_{conserved}$ was 48.7%, matching previous findings of 39%-53% (Siepel et al. 2005).

Simulation of population history

Four demographic models were tested to assess the robustness of demographic inference results under a range of historical population scenarios following the burn-in period of $14N_{anc}$, at which time $N = N_{anc}$ = 2000 for human simulations (3333 for *Drosophila*) after scaling. Recent expansion was modelled by exponential growth leading to a five-fold increase in population size that started $1N_{anc}$ generations ago. Recent contraction was modelled by exponential decline leading to a five-fold decrease in population size that started $1N_{anc}$ generations ago. Bottleneck and recovery were modelled by an instantaneous ten-fold decline in population size ($1N_{anc}$ generations ago), which remained constant for 10 generations, before exponentially increasing to recover to N_{anc}. Repeated bottlenecks were simulated identically to the bottleneck-and-recovery scenario, however following recovery, population size was kept constant for $1N_{anc}$ generations, before repeating the scenario.

Calculating summary statistics

Mean nucleotide diversity in neutral regions (π_{neu}) was calculated using the calcHeterozygosity function within *SLiM*, applied to only neutral mutations in the intergenic and intronic regions, multiplied by the total proportion of neutral sites in the genome. The

839 decrease in nucleotide diversity in neutral regions due to the shared effects of background
840 selection and selective sweeps under different selection parameters was estimated by $\frac{\pi_{neu}}{\pi_0}$,
841 where π_0 was obtained from otherwise identical simulations without selection.
842
843 Divergence at neutral sites in (div_{neu}) was calculated as the number of neutral
844 substitutions in intergenic and intronic regions that occurred after $10N_{anc}$ generations (*n.b.* all
845 population size change scenarios took place after $14N_{anc}$ burn-in) divided by the total number
846 of neutral sites in intergenic and intronic regions (*i.e.*, 95.65% of intronic and 96.40% of
847 intergenic sites for human). Divergence in coding regions (div_{coding}) was similarly
848 calculated by dividing all exonic substitutions, neutral and selected, by the cumulative length
849 of exon regions in the simulated chromosome.
850
851 The proportion of nonsynonymous substitutions in coding regions that were positively
852 selected (α) was determined assuming ~30% of mutations in coding regions are synonymous,
853 and all synonymous mutations are neutral. Given ~50% of new mutations in coding regions
854 are neutral according to the human DFE provided by Huber *et al.* (2017), this suggests that
855 the remaining ~40% of neutral mutations in coding regions (~20% of all exonic mutations)
856 refer to non-synonymous sites. Therefore, α for simulated populations of humans was
857 calculated by the number of substitutions observed in different classes of mutations (sub_{class})
858 that occurred after $10N_{anc}$ using the following equation:
859
860 $$\alpha_{human} = \frac{sub_{pos}}{0.4sub_{neutral} + sub_{del} + sub_{pos}} \quad (1)$$
861
862
863 *Drosophila*: we assumed that all mutations in the effectively neutral class (comprising
864 25% of all mutations) following the Johri *et al.* (2020) DFE represent all synonymous
865 mutations, and no nonsynonymous mutations are strictly neutral. Therefore, α for *Drosophila*
866 was calculated as above, though without including any neutral substitutions:
867
868 $$\alpha_{drosophila} = \frac{sub_{pos}}{sub_{del} + sub_{pos}} \quad (2)$$
869

Scripts used to generate summary statistics are available at: https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/pi_div_alpha.sh.

Demographic inference with *Relate*

The VCF for each replicate was filtered to only contain neutral mutations in intergenic and intronic regions, before being converted to the hap/sample format using the *RelateFileFormats* script, with *–mode ConvertFromVcf*. During testing, we found that when masking selected sites from the VCF used as input for *Relate*, *i.e.*, removing SNPs in selected regions, demographic inference could be majorly impacted if a mask file was not provided to *Relate* to enable recalculation of genetic distances between sites (see Supplementary Note S1 for details). As a result, we generated a genomic mask file for each individual replicate which masks conserved regions (exons/UTRs) as well as conserved sites and a randomly chosen proportion of invariant sites in introns/intergenic regions, to account for excluding/masking functional elements for demographic inference (see https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/new_build_mask.sh). The haps, sample, and mask files were then passed to *Relate* 1.1 (Speidel et al. 2019) to perform genome-wide genealogy estimation with mutation rate (μ) = 1.25×10^{-8} for human (3×10^{-9} for *Drosophila*). The estimated current haplotype population size ($2N_{cur}$) parameter, '-*N*', required by *Relate*, was calculated as may be done in empirical analysis, rather than using the true values from simulations, by the following:

$$2N_{cur} = \frac{\pi_{neu}}{2\mu} \quad (3)$$

Where μ is the unscaled genome-wide mutation rate, and mean nucleotide diversity in neutral regions (π_{neu}) is calculated as previously described. The genealogy results were passed through *Relate*'s EstimatePopulationSize module, with '--years_per_generation 1'. Historical effective population sizes for an epoch *i* generations ago (N_i) were estimated from the coalescence rates output using the same method built into *Relate*:

$$N_i = \frac{0.5}{coal_i} \quad (4)$$

where *coal* is the mean genome-wide instantaneous coalescence rate *i* generations ago obtained from the *.coal* file output by *Relate*. The genetic maps were calculated from the recombination rate maps assuming linearity (see https://github.com/JohriLab/ARG_BGS_Demography/blob/main/scripts/rec_map2gen_map.sh); Haldane's correction (Haldane 1919) was deemed unnecessary as mean interval length between rate recalculation < 1.1Mb. *n.b.* despite the *SLiM* simulations being scaled down, the *Relate* parameters are unscaled, thus producing unscaled demographic inference results.

Figure generation was performed using *R* v4.1.0 and *Python* v3.11.3. ChatGPT 3.5, (https://chat.openai.com; accessed September 2023) was used to build two minor data processing scripts (print_nums.sh and exon_mutations.py at https://github.com/JohriLab/ARG_Selection_Demography/blob/main/scripts/), though was not used for any text generation for the manuscript.

Acknowledgements

The research in this study was conducted using computational resources provided by ITS Research Computing at the University of North Carolina at Chapel Hill. The authors declare no conflicts of interest. We thank Jeff Jensen for providing helpful comments on the manuscript.

Data availability

All relevant code and scripts generated for this study are available at https://github.com/JohriLab/ARG_BGS_Demography. The mutation and recombination rate maps used for analysis are available at https://github.com/JohriLab/ARG_BGS_Demography/tree/main/misc.

References

932 Almarri MA, Haber M, Lootah RA, Hallast P, Al Turki S, Martin HC, Xue Y, Tyler-Smith C.
933 2021. The genomic history of the Middle East. *Cell* 184:4612–4625.

934 Andolfatto P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. *Nature*
935 437:1149–1152.

936 Arguello JR, Laurent S, Clark AG. 2019. Demographic history of the human commensal
937 *Drosophila melanogaster*. *Genome Biol Evol* 11:844–854.

938 Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in
939 population genetics. *Genetics* 162:2025–2035.

940 Beichman AC, Huerta-Sanchez E, Lohmueller KE. 2018. Using genomic data to infer historic
941 population dynamics of nonmodel organisms. *Annu Rev Ecol Evol Syst* 49:433–456.

942 Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, Chen Y, Felkel S,
943 Hallast P, Kamm J, et al. 2020. Insights into human genetic variation and population
944 history from 929 diverse genomes. *Science* 367:eaay5012.

945 Besenbacher S, Hvilsom C, Marques-Bonet T, Mailund T, Schierup MH. 2019. Direct
946 estimation of mutations in great apes reconciles phylogenetic dating. *Nat Ecol Evol*
947 3:286–292.

948 Boitard S, Arredondo A, Chikhi L, Mazet O. 2022. Heterogeneity in effective size across the
949 genome: effects on the inverse instantaneous coalescence rate (IICR) and implications
950 for demographic inference under linked selection. *Genetics* 220:iyac008.

951 Booker TR, Jackson BC, Keightley PD. 2017. Detecting positive selection in the genome.
952 *BMC Biol* 15:98.

953 Brandt D, Huber CD, Chiang CWK, Ortega-Del Vecchyo D. 2024. The promise of inferring
954 the past using the ancestral recombination graph. *Genome Biol Evol* 16:evae005.

955 Brandt D, Wei X, Deng Y, Vaughn AH, Nielsen R. 2022. Evaluation of methods for
956 estimating coalescence times using ancestral recombination graphs. *Genetics*
957 221:iyac044.

958 Cagan A, Theunert C, Laayouni H, Santpere G, Pybus M, Casals F, Prüfer K, Navarro A,
959 Marques-Bonet T, Bertranpetit J, et al. 2016. Natural selection in the great apes. *Mol*
960 *Biol Evol* 33:3268–3283.

961 Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR,
962 Olsen KM, Nielsen R, McCouch SR, et al. 2007. Genome-wide patterns of nucleotide
963 polymorphism in domesticated rice. *PLoS Genet* 3:e163.

964 Campos JL, Zhao L, Charlesworth B. 2017. Estimating the parameters of background
965 selection and selective sweeps in *Drosophila* in the presence of gene conversion. *Proc*
966 *Natl Acad Sci USA* 114:E4762–E4771.

967 Carneiro M, Afonso S, Geraldes A, Garreau H, Bolet G, Boucher S, Tircazes A, Queney G,
968 Nachman MW, Ferrand N. 2011. The genetic structure of domestic rabbits. *Mol Biol*
969 *Evol* 28:1801–1816.

970 Carneiro M, Albert FW, Melo-Ferreira J, Galtier N, Gayral P, Blanco-Aguiar JA, Villafuerte
971 R, Nachman MW, Ferrand N. 2012. Evidence for widespread positive and purifying
972 selection across the European rabbit (*Oryctolagus cuniculus*) genome. *Mol Biol Evol*
973 29:1837–1849.

974 Carneiro M, Rubin C-J, Di Palma F, Albert FW, Alföldi J, Barrio AM, Pielberg G, Rafati N,
975 Sayyab S, Turner-Maier J, et al. 2014. Rabbit genome analysis reveals a polygenic
976 basis for phenotypic change during domestication. *Science* 345:1074–1079.

977 Castellano D, Macià MC, Tataru P, Bataillon T, Munch K. 2019. Comparison of the full
978 distribution of fitness effects of new amino acid mutations across great apes. *Genetics*
979 213:953–966.

980 Charlesworth B. 2013. Background selection 20 years on. *J Hered* 104:161–171.

981 Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on
982 neutral molecular variation. *Genetics* 134:1289–1303.

983 Charlesworth D. 2003. Effects of inbreeding on the genetic diversity of populations. *Philos*
984 *Trans R Soc B* 358:1051–1070.

985 Charlesworth D, Charlesworth B. 1987. Inbreeding depression and its evolutionary
986 consequences. *Annu Rev Ecol Syst* 18:237–268.

987 Chikhi L, Rodríguez W, Grusea S, Santos P, Boitard S, Mazet O. 2018. The IICR (inverse
988 instantaneous coalescence rate) as a summary of genomic diversity: insights into
989 demographic inference and model choice. *Heredity* 120:13–24.

990 Comeron JM, Kreitman M. 2002. Population, evolutionary and genomic consequences of
991 interference selection. *Genetics* 161:389–410.

992 Comeron JM, Ratnappan R, Bailin S. 2012. The many landscapes of recombination in
993 *Drosophila melanogaster*. *PLoS Genet* 8:e1002905.

994 Cousins T, Tabin D, Patterson N, Reich D, Durvasula A. 2024. Accurate inference of
995 population history in the presence of background selection. Available from:
996 http://biorxiv.org/lookup/doi/10.1101/2024.01.18.576291

997 Deng Y, Song YS, Nielsen R. 2021. The distribution of waiting distances in ancestral
998 recombination graphs. *Theor Pop Biol* 141:34–43.

999 Duchen P, Živković D, Hutter S, Stephan W, Laurent S. 2013. Demographic inference
1000 reveals African and European admixture in the North American *Drosophila*
1001 *melanogaster* population. *Genetics* 193:291–301.

1002 Dukler N, Mughal MR, Ramani R, Huang Y-F, Siepel A. 2022. Extreme purifying selection
1003 against point mutations in the human genome. *Nat Commun* 13:4312.

1004 Dumont BL, Payseur BA. 2008. Evolution of the genomic rate of recombination in mammals.
1005 *Evolution* 62:276–294.

1006 Elyashiv E, Sattath S, Hu TT, Strutsovsky A, McVicker G, Andolfatto P, Coop G, Sella G.
1007 2016. A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet*
1008 12:e1006130.

1009 Ewing GB, Jensen JD. 2016. The consequences of not accounting for background selection in
1010 demographic inference. *Mol Ecol* 25:135–141.

1011 Excoffier L, Marchi N, Marques DA, Matthey-Doret R, Gouy A, Sousa VC. 2021.
1012 *fastsimcoal2* : demographic inference under complex evolutionary scenarios.
1013 *Bioinformatics* 37:4882–4885.

1014 Eyre-Walker A, Keightley PD. 2009. Estimating the rate of adaptive molecular evolution in
1015 the presence of slightly deleterious mutations and population size change. *Mol Biol*
1016 *Evol* 26:2097–2108.

1017 Fan C, Cahoon JL, Dinh BL, Ortega-Del Vecchyo D, Huber C, Edge MD, Mancuso N,
1018 Chiang CWK. 2023. A likelihood-based framework for demographic inference from
1019 genealogical trees. Available from:
1020 http://biorxiv.org/lookup/doi/10.1101/2023.10.10.561787

1021 Fisher RA. 1930. The genetical theory of natural selection. Oxford, UK: Clarendon Press

1022 Fiston-Lavier A-S, Singh ND, Lipatov M, Petrov DA. 2010. *Drosophila melanogaster*
1023 recombination rate calculator. *Gene* 463:18–20.

1024 Francioli LC, Polak PP, Koren A, Menelaou A, Chun S, Renkens I, Van Duijn CM, Swertz
1025 M, Wijmenga C, Van Ommen G, et al. 2015. Genome-wide patterns and properties of
1026 *de novo* mutations in humans. *Nat Genet* 47:822–826.

1027 Frankish A, Diekhans M, Ferreira A-M, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu
1028 C, Wright J, Armstrong J, et al. 2019. GENCODE reference annotation for the human
1029 and mouse genomes. *Nucleic Acids Res* 47:766–773.

1030 Friedlander E, Steinrücken M. 2022. A numerical framework for genetic hitchhiking in
1031 populations of variable size. *Genetics* 220:iyac012.

1032 Galtier N. 2016. Adaptive protein evolution in animals and the effective population size
1033 hypothesis. *PLoS Genet* 12:e1005774.

1034 Gillespie JH. 1984. Molecular evolution over the mutational landscape. *Evolution* 38:1116.

1035 Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, Yu F, Gibbs RA, The
1036 1000 Genomes Project, Bustamante CD, et al. 2011. Demographic history and rare
1037 allele sharing among human populations. *Proc Natl Acad Sci USA* 108:11983–11988.

1038 Griffiths RC, Marjoram P. 1997. An ancestral recombination graph. In: Progress in
1039 population genetics and human evolution. Vol. 87. New York, NY: Springer New
1040 York. p. 257–270.

1041 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint
1042 demographic history of multiple populations from multidimensional SNP frequency
1043 data. *PLoS Genet* 5:e1000695.

1044 Haldane JBS. 1919. The combination of linkage values, and the calculation of distances
1045 between the loci of linked factors. *J Genet* 8:299–309.

1046 Haller BC, Messer PW. 2017. asymptoticMK: a web-based tool for the asymptotic
1047 McDonald–Kreitman test. *G3* 7:1569–1575.

1048 Haller BC, Messer PW. 2023. SLiM 4: multispecies eco-evolutionary modeling. *Am Nat*
1049 201:127–139.

1050 Halligan DL, Kousathanas A, Ness RW, Harr B, Eöry L, Keane TM, Adams DJ, Keightley
1051 PD. 2013. Contributions of protein-coding and regulatory change to adaptive
1052 molecular evolution in murid rodents. *PLoS Genet* 9:e1003995.

1053 Harris K, Nielsen R. 2013. Inferring demographic history from a spectrum of shared
1054 haplotype lengths. *PLoS Genet* 9:e1003521.

1055 Hartfield M, Bataillon T. 2020. Selective sweeps under dominance and inbreeding. *G3*
1056 10:1063–1075.

1057 Hawks J, Hunley K, Lee S-H, Wolpoff M. 2000. Population bottlenecks and Pleistocene
1058 human evolution. *Mol Biol Evol* 17:2–22.

1059 Hayes BJ, Visscher PM, McPartlan HC, Goddard ME. 2003. Novel multilocus measure of
1060 linkage disequilibrium to estimate past effective population size. *Genome Res*
1061 13:635–643.

1062 Hedrick PW. 1980. Hitchhiking: a comparison of linkage and partial selfing. *Genetics*
1063 94:791–808.

1064 Hoggart CJ, Chadeau-Hyam M, Clark TG, Lampariello R, Whittaker JC, De Iorio M, Balding
1065 DJ. 2007. Sequence-level population simulations over large genomic regions.
1066 *Genetics* 177:1725–1731.

1067 Howell AA, Terbot JW, Soni V, Johri P, Jensen JD, Pfeifer SP. 2023. Developing an
1068 appropriate evolutionary baseline model for the study of human cytomegalovirus.
1069 *Genome Biol Evol* 15:evad059.

1070 Hu W, Hao Z, Du P, Di Vincenzo F, Manzi G, Cui J, Fu Y-X, Pan Y-H, Li H. 2023. Genomic
1071 inference of a severe human bottleneck during the Early to Middle Pleistocene
1072 transition. *Science* 381:979–984.

1073 Hudson RR. 1991. Gene genealogies and the coalescent process. In: Oxford Surveys in
1074 Evolutionary Biology. Vol. 7. New York, NY: Oxford University Press. p. 1–44.

1075 Huff CD, Xing J, Rogers AR, Witherspoon D, Jorde LB. 2010. Mobile elements reveal small
1076 population size in the ancient ancestors of *Homo sapiens*. *Proc Natl Acad Sci USA*
1077 107:2147–2152.

1078 Ignatieva A, Favero M, Koskela J, Sant J, Myers SR. 2023. The distribution of branch
1079 duration and detection of inversions in ancestral recombination graphs. Available
1080 from: http://biorxiv.org/lookup/doi/10.1101/2023.07.11.548567

1081 Ishigohoka J, Liedvogel M. 2024. High-recombining genomic regions affect demography
1082 inference. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.02.05.579015

1083 Jensen JD, Payseur BA, Stephan W, Aquadro CF, Lynch M, Charlesworth D, Charlesworth
1084 B. 2019. The importance of the Neutral Theory in 1968 and 50 years on: A response
1085 to Kern and Hahn 2018. *Evolution* 73:111–114.

1086 Jensen JD, Thornton KR, Andolfatto P. 2008. An approximate Bayesian estimator suggests
1087 strong, recurrent selective sweeps in *Drosophila*. *PLoS Genet* 4:e1000198.

1088 Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen C-F, Thomas MA,
1089 Haussler D, Jacob HJ. 2004. Comparative recombination rates in the rat, mouse, and
1090 human genomes. *Genome Res* 14:528–538.

1091 Johri P, Aquadro CF, Beaumont M, Charlesworth B, Excoffier L, Eyre-Walker A, Keightley
1092 PD, Lynch M, McVean G, Payseur BA, et al. 2022. Recommendations for improving
1093 statistical inference in population genomics. *PLoS Biol* 20:e3001669.

1094 Johri P, Charlesworth B, Jensen JD. 2020. Toward an evolutionarily appropriate null model:
1095 jointly inferring demography and purifying selection. *Genetics* 215:173–192.

1096 Johri P, Eyre-Walker A, Gutenkunst RN, Lohmueller KE, Jensen JD. 2022. On the prospect
1097 of achieving accurate joint estimation of selection with population history. *Genome*
1098 *Biol Evol* 14:evac088.

1099 Johri P, Pfeifer SP, Jensen JD. 2023. Developing an evolutionary baseline model for humans:
1100 jointly inferring purifying selection with population history. *Mol Biol Evol*
1101 40:msad100.

1102 Johri P, Riall K, Becher H, Excoffier L, Charlesworth B, Jensen JD. 2021. The impact of
1103 purifying and background selection on the inference of population history: problems
1104 and prospects. *Mol Biol Evol* 38:2986–3003.

1105 Jouganous J, Long W, Ragsdale AP, Gravel S. 2017. Inferring the joint demographic history
1106 of multiple populations: beyond the diffusion approximation. *Genetics* 206:1549–
1107 1567.

1108 Joyce P, Rokyta DR, Beisel CJ, Orr HA. 2008. A general extreme value theory model for the
1109 adaptation of DNA sequences under strong selection and weak mutation. *Genetics*
1110 180:1627–1643.

1111 Kamm J, Terhorst J, Durbin R, Song YS. 2020. Efficiently inferring the demographic history
1112 of many populations with allele count data. *J Am Stat Assoc* 115:1472–1487.

1113 Kapopoulou A, Pfeifer SP, Jensen JD, Laurent S. 2018. The demographic history of African
1114 *Drosophila melanogaster*. *Genome Biol Evol* 10:2338–2342.

1115 Keightley PD, Campos JL, Booker TR, Charlesworth B. 2016. Inferring the frequency
1116 spectrum of derived variants to quantify adaptive molecular evolution in protein-
1117 coding genes of *Drosophila melanogaster*. *Genetics* 203:975–984.

1118 Keightley PD, Ness RW, Halligan DL, Haddrill PR. 2014. Estimation of the spontaneous
1119 mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family.
1120 *Genetics* 196:313–320.

1121 Kelleher J, Wong Y, Wohns AW, Fadil C, Albers PK, McVean G. 2019. Inferring whole-
1122 genome histories in large population datasets. *Nat Genet* 51:1330–1338.

1123 Kim Y. 2006. Allele frequency distribution under recurrent selective sweeps. *Genetics*
1124 172:1967–1978.

1125 Kim Y, Wiehe T. 2008. Simulation of DNA sequence evolution under models of recent
1126 directional selection. *Brief Bioinformatics* 10:84–96.

1127 Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA,
1128 Sigurdsson A, Jonasdottir Aslaug, Jonasdottir Adalbjorg, et al. 2012. Rate of *de novo*
1129 mutations and the importance of father's age to disease risk. *Nature* 488:471–475.

1130 Kong A, Thorleifsson G, Gudbjartsson DF, Masson G, Sigurdsson A, Jonasdottir Aslaug,
1131 Walters GB, Jonasdottir Adalbjorg, Gylfason A, Kristinsson KTh, et al. 2010. Fine-
1132 scale recombination rate differences between sexes, populations and individuals.
1133 *Nature* 467:1099–1103.

1134 Lange JD, Pool JE. 2018. Impacts of recurrent hitchhiking on divergence and demographic
1135 inference in *Drosophila*. *Genome Biol Evol* 10:1882–1891.

1136 Laurent SJY, Werzner A, Excoffier L, Stephan W. 2011. Approximate Bayesian analysis of
1137 *Drosophila melanogaster* polymorphism data reveals a recent colonization of
1138 southeast Asia. *Mol Biol Evol* 28:2041–2051.

1139 Lewanski AL, Grundler MC, Bradburd GS. 2024. The era of the ARG: An introduction to
1140 ancestral recombination graphs and their significance in empirical evolutionary
1141 genomics. *PLoS Genet* 20:e1011110.

1142 Li H, Durbin R. 2011. Inference of human population history from individual whole-genome
1143 sequences. *Nature* 475:493–496.

1144 Li N, Stephens M. 2003. Modeling linkage disequilibrium and identifying recombination
1145 hotspots using single-nucleotide polymorphism data. *Genetics* 165:2213–2233.

1146 Lindsay SJ, Rahbari R, Kaplanis J, Keane T, Hurles ME. 2019. Similarities and differences in
1147 patterns of germline mutation between mice and humans. *Nat Commun* 10:4053.

1148 Liu X, Fu Y-X. 2015. Exploring population size changes using SNP frequency spectra. *Nat*
1149 *Genet* 47:555–559.

1150 Long M, Deutsch M. 1999. Intron—exon structures of eukaryotic model organisms. *Nucleic*
1151 *Acids Res* 27:3219–3228.

1152 Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y,
1153 Magwire MM, Cridland JM, et al. 2012. The *Drosophila melanogaster* genetic
1154 reference panel. *Nature* 482:173–178.

1155 Mao Y, Catacchio CR, Hillier LW, Porubsky D, Li R, Sulovari A, Fernandes JD, Montinaro
1156 F, Gordon DS, Storer JM, et al. 2021. A high-quality bonobo genome refines the
1157 analysis of hominid evolution. *Nature* 594:77–81.

1158 Marchi N, Schlichta F, Excoffier L. 2021. Demographic inference. *Current Biol* 31:R276–
1159 R279.

1160 Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genet Res*
1161 23:23–35.

1162 Mayr C. 2016. Evolution and biological roles of alternative 3′UTRs. *Trends Cell Biol*
1163 26:227–237.

1164 Mazet O, Rodríguez W, Grusea S, Boitard S, Chikhi L. 2016. On the importance of being
1165 structured: instantaneous coalescence rates and human evolution—lessons for
1166 ancestral population size inference. *Heredity* 116:362–371.

1167 McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*.
1168 *Nature* 351:652–654.

1169 McManus KF, Kelley JL, Song S, Veeramah KR, Woerner AE, Stevison LS, Ryder OA, Ape
1170 Genome Project G, Kidd JM, Wall JD, et al. 2015. Inference of gorilla demographic
1171 and selective history from whole-genome sequence data. *Mol Biol Evol* 32:600–612.

1172 McVean GAT, Cardin NJ. 2005. Approximating the coalescent with recombination. *Philos*
1173 *Trans R Soc B* 360:1387–1393.

1174 Mellars P. 2006. Why did modern human populations disperse from Africa *ca.* 60,000 years
1175 ago: a new model. *Proc Natl Acad Sci USA* 103:9381–9386.

1176 Messer PW, Petrov DA. 2013. Frequent adaptation and the McDonald–Kreitman test. *Proc*
1177 *Natl Acad Sci USA* 110:8615–8620.

1178 Mignone F, Gissi C, Liuni S, Pesole G. 2002. Untranslated regions of mRNAs. *Genome Biol*
1179 3:1–10.

1180 Minadakis N, Williams H, Horvath R, Caković D, Stritt C, Thieme M, Bourgeois Y, Roulin
1181 AC. 2023. The demographic history of the wild crop relative *Brachypodium*
1182 *distachyon* is shaped by distinct past and present ecological niches. *Peer Community J*
1183 3:e84.

1184 Moorjani P, Hellenthal G. 2023. Methods for assessing population relationships and history
1185 using genomic data. *Annu Rev Genomics Hum Genet* 24:305–332.

1186 Murphy DA, Elyashiv E, Amster G, Sella G. 2023. Broad-scale variation in human genetic
1187 diversity levels is predicted by purifying selection on coding and non-coding
1188 elements. *eLife* 12:e76065.

1189 Nater A, Greminger MP, Arora N, Van Schaik CP, Goossens B, Singleton I, Verschoor EJ,
1190 Warren KS, Krützen M. 2015. Reconstructing the demographic history of orangutans
1191 using approximate Bayesian computation. *Mol Ecol* 24:310–327.

1192 Nicolaisen LE, Desai MM. 2013. Distortions in genealogies due to purifying selection and
1193 recombination. *Genetics* 195:221–230.

1194 Nielsen R, Akey JM, Jakobsson M, Pritchard JK, Tishkoff S, Willerslev E. 2017. Tracing the
1195 peopling of the world through genomics. *Nature* 541:302–310.

1196 Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. 2007. Recent and ongoing
1197 selection in the human genome. *Nat Rev Genet* 8:857–868.

1198 Noble LM, Yuen J, Stevens L, Moya N, Persaud R, Moscatelli M, Jackson JL, Zhang G,
1199 Chitrakar R, Baugh LR, et al. 2021. Selfing is the safest sex for *Caenorhabditis*
1200 *tropicalis*. *eLife* 10:e62587.

1201 Nordborg M. 2000. Linkage disequilibrium, gene trees and selfing: an ancestral
1202 recombination graph with partial self-fertilization. *Genetics* 154:923–929.

1203 Nordborg M, Donnelly P. 1997. The coalescent process with selfing. *Genetics* 146:1185–
1204 1195.

1205 Novo I, Santiago E, Caballero A. 2022. The estimates of effective population size based on
1206 linkage disequilibrium are virtually unaffected by natural selection. *PLoS Genet*
1207 18:e1009764.

1208 Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR,
1209 Altemose N, Uralsky L, Gershman A, et al. 2022. The complete sequence of a human
1210 genome. *Science* 376:44–53.

1211 Orr HA. 2003. The distribution of fitness effects among beneficial mutations. *Genetics*
1212 163:1519–1526.

1213 Phifer-Rixey M, Bonhomme F, Boursot P, Churchill GA, Pialek J, Tucker PK, Nachman
1214 MW. 2012. Adaptive evolution and effective population size in wild house mice. *Mol*
1215 *Biol Evol* 29:2949–2955.

1216 Piovesan A, Antonaros F, Vitale L, Strippoli P, Pelleri MC, Caracausi M. 2019. Human
1217 protein-coding genes and gene feature statistics in 2019. *BMC Res Notes* 12:315.

1218 Pope NS, Singh A, Childers AK, Kapheim KM, Evans JD, López-Uribe MM. 2023. The
1219 expansion of agriculture has shaped the recent evolutionary history of a specialized
1220 squash pollinator. *Proc Natl Acad Sci USA* 120:e2208116120.

1221 Pouyet F, Aeschbacher S, Thiéry A, Excoffier L. 2018. Background selection and biased
1222 gene conversion affect more than 95% of the human genome and bias demographic
1223 inferences. *eLife* 7:e36317.

1224 Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, Veeramah KR,
1225 Woerner AE, O'Connor TD, Santpere G, et al. 2013. Great ape genetic diversity and
1226 population history. *Nature* 499:471–475.

1227 Ragsdale AP, Gravel S. 2019. Models of archaic admixture and recent history from two-locus
1228 statistics. *PLoS Genet* 15:e1008204.

1229 Ragsdale AP, Gutenkunst RN. 2017. Inferring demographic history using two-locus statistics.
1230 *Genetics* 206:1037–1048.

1231 Ragsdale AP, Weaver TD, Atkinson EG, Hoal EG, Möller M, Henn BM, Gravel S. 2023. A
1232 weakly structured stem for human origins in Africa. *Nature* 617:755–763.

1233 Rands CM, Meader S, Ponting CP, Lunter G. 2014. 8.2% of the human genome is
1234 constrained: variation in rates of turnover across functional element classes in the
1235 human lineage. *PLoS Genet* 10:e1004525.

1236 Rasmussen MD, Hubisz MJ, Gronau I, Siepel A. 2014. Genome-wide inference of ancestral
1237 recombination graphs. *PLoS Genet* 10:e1004342.

1238 Rhie A, Nurk S, Cechova M, Hoyt SJ, Taylor DJ, Altemose N, Hook PW, Koren S,
1239 Rautiainen M, Alexandrov IA, et al. 2023. The complete sequence of a human Y
1240 chromosome. *Nature* 621:344–354.

1241 Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A. 2020. Recent demographic
1242 history inferred by high-resolution analysis of linkage disequilibrium. *Mol Biol Evol*
1243 37:3642–3653.

1244 dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, Emmert
1245 DB, Gelbart WM, the FlyBase Consortium. 2015. FlyBase: introduction of the
1246 *Drosophila melanogaster* Release 6 reference genome assembly and large-scale
1247 migration of genome annotations. *Nucleic Acids Res* 43:690–697.

1248 Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Farrell CM, Feldgarden
1249 M, Fine AM, Funk K, et al. 2023. Database resources of the National Center for
1250 Biotechnology Information in 2023. *Nucleic Acids Res* 51:D29–D38.

1251 Schiffels S, Durbin R. 2014. Inferring human population size and separation history from
1252 multiple genome sequences. *Nat Genet* 46:919–925.

1253 Schiffels S, Wang K. 2020. MSMC and MSMC2: the multiple sequentially Markovian
1254 coalescent. In: Statistical Population Genomics. Vol. 2090. Methods in Molecular
1255 Biology. New York, NY: Springer US. p. 147–166.

1256 Schlebusch CM, Malmström H, Günther T, Sjödin P, Coutinho A, Edlund H, Munters AR,
1257 Vicente M, Steyn M, Soodyall H, et al. 2017. Southern African ancient genomes
1258 estimate modern human divergence to 350,000 to 260,000 years ago. *Science*
1259 358:652–655.

1260 Schmidt JM, De Manuel M, Marques-Bonet T, Castellano S, Andrés AM. 2019. The impact
1261 of genetic adaptation on chimpanzee subspecies differentiation. *PLoS Genet*
1262 15:e1008485.

1263 Schrider DR, Shanku AG, Kern AD. 2016. Effects of linked selective sweeps on
1264 demographic inference and model selection. *Genetics* 204:1207–1223.

1265 Seetharaman S, Jain K. 2014. Adaptive walks and distribution of beneficial fitness effects.
1266 *Evolution* 68:965–975.

1267 Sellinger TPP, Abu-Awad D, Tellier A. 2021. Limits and convergence properties of the
1268 sequentially Markovian coalescent. *Mol Ecol Res* 21:2231–2248.

1269 Sheehan S, Song YS. 2016. Deep learning for population genetic inference. *PLoS Comput*
1270 *Biol* 12:e1004845.

1271 Sherry ST, Harpending HC, Batzer MA, Stoneking M. 1997. *Alu* Evolution in human
1272 populations: using the coalescent to estimate effective population size. *Genetics*
1273 147:1977–1982.

1274 Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth
1275 J, Hillier LW, Richards S, et al. 2005. Evolutionarily conserved elements in
1276 vertebrate, insect, worm, and yeast genomes. *Genome Res* 15:1034–1050.

1277 Smith ML, Hahn MW. 2023. Selection leads to false inferences of introgression using
1278 popular methods. Available from:
1279 http://biorxiv.org/lookup/doi/10.1101/2023.10.27.564394

1280 Soni V, Johri P, Jensen JD. 2023. Evaluating power to detect recurrent selective sweeps
1281 under increasingly realistic evolutionary null models. *Evolution* 77:2113–2127.

1282 Soni V, Pfeifer SP, Jensen JD. 2024. The effects of mutation and recombination rate
1283 heterogeneity on the inference of demography and the distribution of fitness effects.
1284 *Genome Biol Evol* 16:evae004.

1285 Speidel L, Forest M, Shi S, Myers SR. 2019. A method for genome-wide genealogy
1286 estimation for thousands of samples. *Nat Genet* 51:1321–1329.

1287 Strütt S, Sellinger T, Glémin S, Tellier A, Laurent S. 2023. Joint inference of evolutionary
1288 transitions to self-fertilization and demographic history using whole-genome
1289 sequences. *eLife* 12:e82384.

1290 Tataru P, Bataillon T. 2019. polyDFEv2.0: testing for invariance of the distribution of fitness
1291 effects within and across species. *Bioinformatics* 35:2868–2869.

1292 Terbot JW, Johri P, Liphardt SW, Soni V, Pfeifer SP, Cooper BS, Good JM, Jensen JD. 2023.
1293 Developing an appropriate evolutionary baseline model for the study of SARS-CoV-2
1294 patient samples.Hobman TC, editor. *PLoS Pathog* 19:e1011265.

1295 Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history
1296 from hundreds of unphased whole genomes. *Nat Genet* 49:303–309.

1297 Teshima KM, Coop G, Przeworski M. 2006. How reliable are empirical genomic scans for
1298 selective sweeps? *Genome Res* 16:702–712.

1299 Teshima KM, Przeworski M. 2006. Directional positive selection on an allele of arbitrary
1300 dominance. *Genetics* 172:713–718.

1301 Teterina AA, Willis JH, Lukac M, Jovelin R, Cutter AD, Phillips PC. 2023. Genomic
1302 diversity landscapes in outcrossing and selfing *Caenorhabditis nematodes*. *PLoS*
1303 *Genet* 19:e1010879.

The International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* 437:1299–1320.

Thornton K, Andolfatto P. 2006. Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. *Genetics* 172:1607–1619.

Thornton KR, Jensen JD. 2007. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* 175:737–750.

Upadhya G, Steinrücken M. 2022. Robust inference of population size histories from genomic sequencing data. *PLoS Comput Biol* 18:e1010419.

Uricchio LH, Hernandez RD. 2014. Robust forward simulations of recurrent hitchhiking. *Genetics* 197:221–236.

Uricchio LH, Petrov DA, Enard D. 2019. Exploiting selection at linked sites to infer the rate and strength of adaptation. *Nat Ecol Evol* 3:977–984.

Wohns AW, Wong Y, Jeffery B, Akbari A, Mallick S, Pinhasi R, Patterson N, Reich D, Kelleher J, McVean G. 2022. A unified genealogy of modern and ancient genomes. *Science* 375:eabi8264.

Wong Y, Ignatieva A, Koskela J, Gorjanc G, Wohns AW, Kelleher J. 2023. A general and efficient representation of ancestral recombination graphs. Available from: http://biorxiv.org/lookup/doi/10.1101/2023.11.03.565466

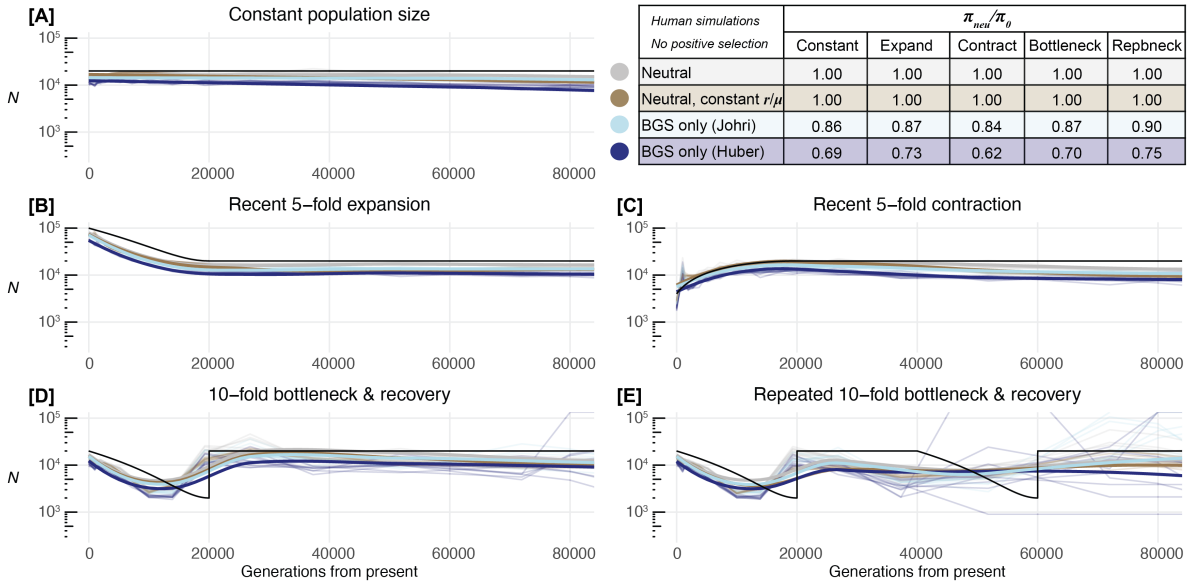
Wright SI, Kalisz S, Slotte T. 2013. Evolutionary consequences of self-fertilization in plants. *Philos Trans R Soc B* 280:20130133.

Zeng K, Charlesworth B. 2011. The joint effects of background selection and genetic recombination on local gene genealogies. *Genetics* 189:251–266.

Zhang BC, Biddanda A, Gunnarsson ÁF, Cooper F, Palamara PF. 2023. Biobank-scale inference of ancestral recombination graphs enables genealogical analysis of complex traits. *Nat Genet* 55:768–776.

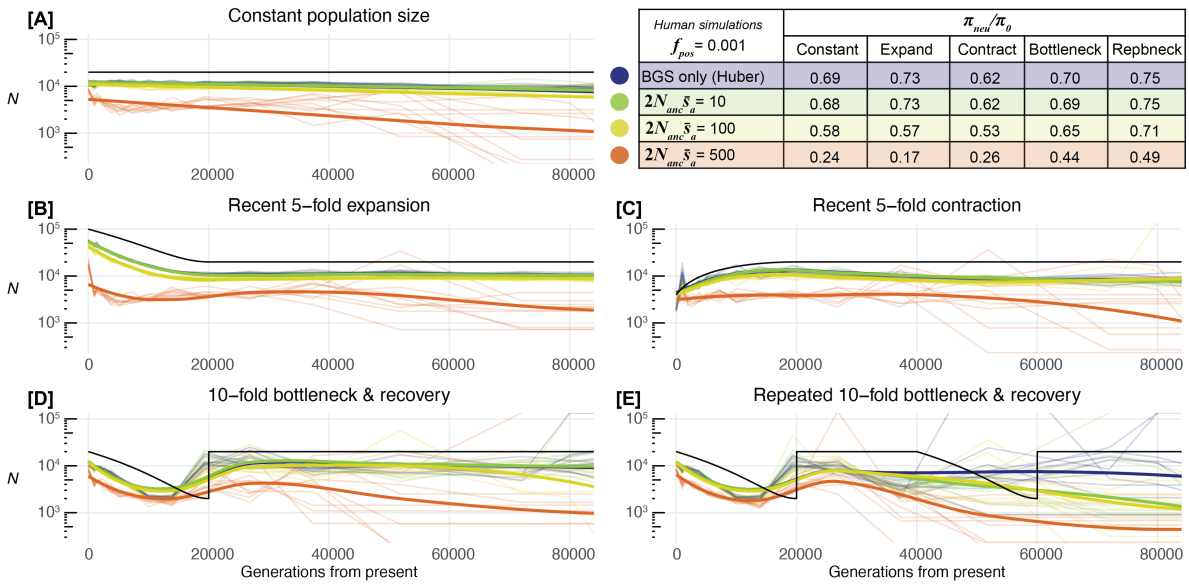
Zhen Y, Huber CD, Davies RW, Lohmueller KE. 2021. Greater strength of selection and higher proportion of beneficial amino acid changing mutations in humans compared with mice and *Drosophila melanogaster*. *Genome Res* 31:110–120.

Figures and Tables



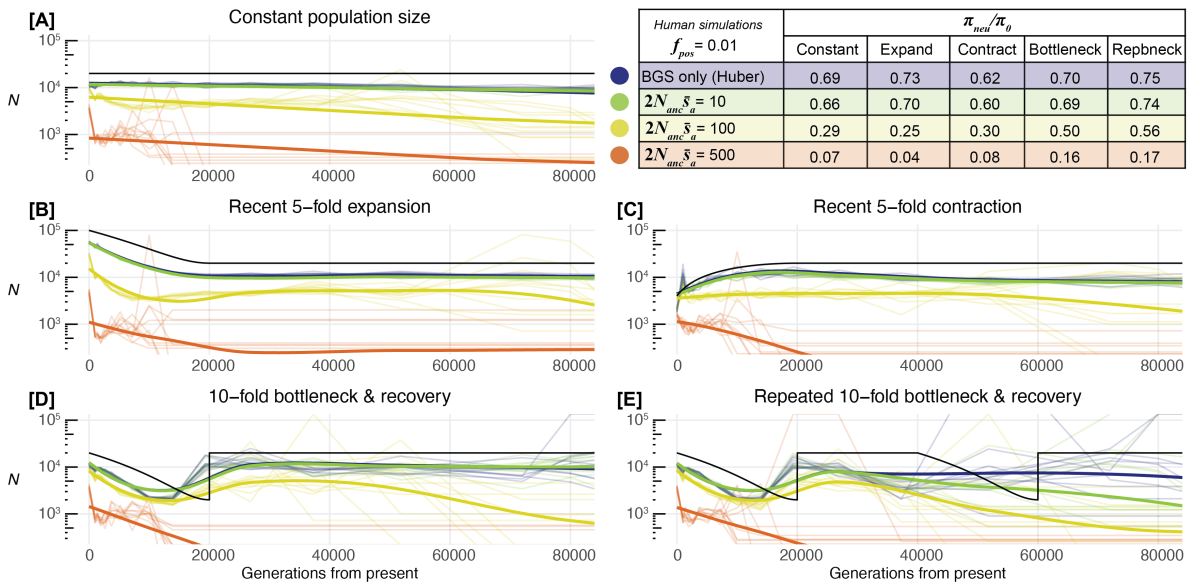
Human simulations No positive selection	π_{neu}/π_0				
	Constant	Expand	Contract	Bottleneck	Repbneck
Neutral	1.00	1.00	1.00	1.00	1.00
Neutral, constant r/μ	1.00	1.00	1.00	1.00	1.00
BGS only (Johri)	0.86	0.87	0.84	0.87	0.90
BGS only (Huber)	0.69	0.73	0.62	0.70	0.75

Figure 1. Historical population size inferred by *Relate* for human-like parameters under five demographic scenarios under neutrality and in the presence of BGS. The black line represents the true simulated population size (N) for each demographic scenario (A-E); coloured lines represent simulations without selection with varying recombination/mutation rates (grey), with constant recombination/mutation rates (gold), with BGS from the DFE reported by Johri *et al.* (2023; light blue) and with BGS from the DFE reported by Huber *et al.* (2017; dark blue). Thin coloured lines represent each of the 10 replicates per evolutionary scenario, thick coloured lines represent the moving regression (LOESS) across all replicates for a given condition. Nucleotide diversity at neutral intron/intergenic sites for simulations with selection relative to identical simulations under neutrality is presented as $\frac{\pi_{neu}}{\pi_0}$ in the table. Note, population size is presented on a $\log_{10}$ scale.



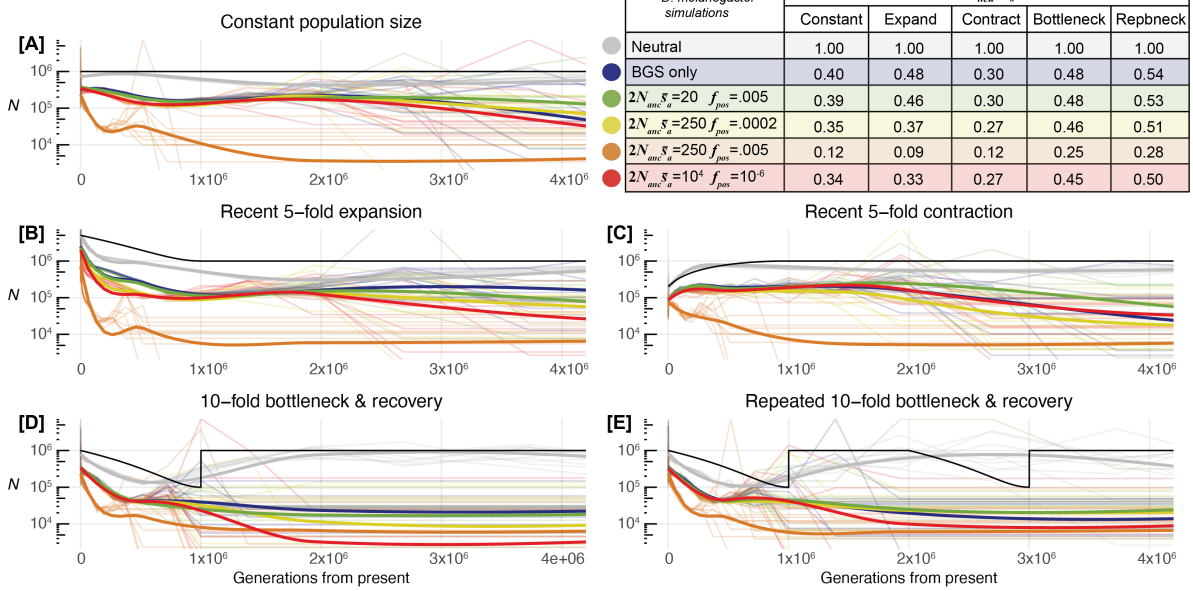
Human simulations f_{pos} = 0.001	π_{neu}/π_0				
	Constant	Expand	Contract	Bottleneck	Repbneck
BGS only (Huber)	0.69	0.73	0.62	0.70	0.75
$2N_{anc}\bar{s}_a$ = 10	0.68	0.73	0.62	0.69	0.75
$2N_{anc}\bar{s}_a$ = 100	0.58	0.57	0.53	0.65	0.71
$2N_{anc}\bar{s}_a$ = 500	0.24	0.17	0.26	0.44	0.49

Figure 2. Historical population size inferred by *Relate* for human-like parameters under five demographic scenarios experiencing selective sweeps from relatively low frequency positive selection (f_{pos}= 0.001) with variable recombination/mutation rates and BGS (DFE from Huber et al. 2017). The black line represents the true simulated population size (N) for each demographic scenario (A-E), coloured lines represent results for BGS only (blue), and different simulated mean positive selection coefficients ($2N_{anc}\bar{s_a}$). Thin coloured lines represent results for each of the 10 replicates per evolutionary scenario, thick coloured lines represent the moving regression (LOESS) across all replicates for a given condition. Nucleotide diversity at neutral intron/intergenic sites relative to identical simulations under neutrality is presented as $\frac{\pi_{neu}}{\pi_0}$ in the table. Note, population size is presented on a $\log_{10}$ scale.



Human simulations f_{pos} = 0.01	π_{neu}/π_0				
	Constant	Expand	Contract	Bottleneck	Repbneck
BGS only (Huber)	0.69	0.73	0.62	0.70	0.75
$2N_{anc}\bar{s}_a$ = 10	0.66	0.70	0.60	0.69	0.74
$2N_{anc}\bar{s}_a$ = 100	0.29	0.25	0.30	0.50	0.56
$2N_{anc}\bar{s}_a$ = 500	0.07	0.04	0.08	0.16	0.17

Figure 3. Historical population size inferred by *Relate* for human-like parameters under five demographic scenarios experiencing selective sweeps from relatively high frequency positive selection (f_{pos}= 0.01) with variable recombination/mutation rates and BGS (DFE from Huber et al. 2017). The black line represents the true simulated population size (N) for each demographic scenario (A-E), coloured lines represent results for BGS only (blue), and different simulated mean positive selection coefficients ($2N_{anc}\bar{s_a}$). Thin coloured lines represent results for each of the 10 replicates per evolutionary scenario, thick coloured lines represent the moving regression (LOESS) across all replicates for a given condition. Nucleotide diversity at neutral intron/intergenic sites relative to identical simulations under neutrality is presented as $\frac{\pi_{neu}}{\pi_0}$ in the table. Note, population size is presented on a $\log_{10}$ scale.



D. melanogaster simulations	π_{neu}/π_0				
	Constant	Expand	Contract	Bottleneck	Repbneck
Neutral	1.00	1.00	1.00	1.00	1.00
BGS only	0.40	0.48	0.30	0.48	0.54
$2N_{anc}s_a$=20 f_{pos}=.005	0.39	0.46	0.30	0.48	0.53
$2N_{anc}s_a$=250 f_{pos}=.0002	0.35	0.37	0.27	0.46	0.51
$2N_{anc}s_a$=250 f_{pos}=.005	0.12	0.09	0.12	0.25	0.28
$2N_{anc}s_a$=10^4 f_{pos}=10^{-6}	0.34	0.33	0.27	0.45	0.50

Figure 4. Historical population size inferred by *Relate* for *Drosophila melanogaster*-like parameters under five demographic scenarios, with variable recombination/mutation rates. The black line represents the true simulated population size (N) for each demographic scenario (A-E); coloured lines represent simulations under strict neutrality (grey), with BGS from the DFE reported by Johri *et al.* (2020; blue), and simulations experiencing BGS as well as sweeps from positive selection introduced with different parameter combinations. Thin coloured lines represent results for each of the 10 replicates per evolutionary scenario, thick coloured lines represent the moving regression (LOESS) across all replicates for a given condition. Nucleotide diversity at neutral intron/intergenic sites for simulations with selection relative to identical simulations under neutrality is presented as $\frac{\pi_{neu}}{\pi_0}$ in the table. Note, population size is presented on a $\log_{10}$ scale.

1385 **Table 1**. Summary of empirically estimated positive selection parameters for different animal populations, including the selection inference
1386 method used. Confidence interval values provided where available in parentheses. Annotation type of selected sites used for analysis is provided:
1387 "CDS" is protein coding sites, "Nonsyn", nonsynymous, "0-fold", 0-fold degenerate coding sites.

Taxon	Population	f_{pos}	$2N_{anc}\overline{s_a}$	α	Reference	Selected sites	Selection inference method
Human	African			0.135	Uricchio et al. 2018	CDS	ABC-MK
	YRI	0.01	6.05	0.25 (0.24-0.252)	Zhen et al. 2021	Nonsyn	Composite likelihood approach
	YRI			0.24	Zhen et al. 2021	Nonsyn	DFE-alpha (Eyre-Walker and Keightley 2009)
	Mixed			0.24 (0.18-0.29)	Galtier 2016	Nonsyn	Grapes
	Mixed			0 (-0.30-0.24)	Eyre-Walker & Keightley 2009	0-fold	DFE-alpha
	Mixed	0.000023 (0.0000022-0.0076)	0.0064 (0.0007-0.1084)		Castellano et al. 2019	CDS	polyDFE v2.0 (Tataru and Bataillon 2019) using only the SFS
Chimpanzee	Western			-0.21 (-0.46-0)	Cagan et al. 2016	0-fold	MK test (McDonald and Kreitman 1991)
		0.000029 (0.0000032-0.01)	0.0081 (0.0007-0.0911)		Castellano et al. 2019	CDS	polyDFE v2.0
Bonobo	Congo			-0.04 (-0.37-0.15)	Cagan et al. 2016	0-fold	MK test
	Congo	0.017 (0.000073-0.12)	0.076 (0.0008, 27.8)		Castellano et al. 2019	CDS	polyDFE v2.0
Gorilla	Western lowland			0.03 (-0.16-0.21)	Cagan et al. 2016	0-fold	MK test
		0.0013 (0.0000071-0.03)	0.0091 (0.0008, 0.2291)		Castellano et al. 2019	CDS	polyDFE v2.0
				0.014 (-0.116-0.110)	McManus et al. 2015	Nonsyn	DFE-alpha
Lemur	Coquerel's sifaka			0.75 (0.7- 0.78)	Galtier 2016	Nonsyn	Grapes
Mouse	Castaneus (mixed)			0.32 (0.28-0.35)	Halligan et al. 2013	0-fold	DFE-alpha
		0.000179 (0.0000865-0.0002715)	170.6 (89.7-251.55)		Booker et al. 2021	CDS	Maximum likelihood
		0.003 (0.0019-0.0048)	14.54 (9.24-23.4)		Booker et al. 2018	0-fold	DFE-alpha

	Castaneus (NW India)			0.51	Zhen et al. 2021	Nonsyn	DFE-alpha
		0.04	3.76	0.41 (0.411-0.413)	Zhen et al. 2021	Nonsyn	Composite likelihood approach
Rabbit	*Lepus granatensis*			0.77 (0.68-0.86)	Galtier 2016	Nonsyn	Grapes
	Oryctolagus cuniculus			0.64 (0.53-0.71)	Carneiro et al 2012	Nonsyn	DFE-alpha
Drosophila melanogaster	Rwandan	0.0045 (0-0.0165)	23 (10-36)	0.57	Keightley et al. 2016	Nonsyn	DFE-alpha
	Yakuba	0.00022; 0.000841 (high gene conversion rates)	249; 508 (high gene conversion rates)		Campos et al. 2017	Nonsyn	DFE-alpha
	Zambian			0.71	Zhen et al. 2021	Nonsyn	DFE-alpha
		0.04	4.92	0.60 (0.601-0.604)	Zhen et al. 2021	Nonsyn	Composite likelihood approach
				0.59	Lange & Pool 2018	Nonsyn	asymptoticMK (Haller and Messer 2017)
	East African	0.00001	10000		Jensen et al. 2008	Nonsyn	ABC
	Raleigh, NC			0.40	Elyashiv et al. 2016	Nonsyn	Maximum composite likelihood

1388

Table 2. Proportion of positively selected "nonsynonymous" substitutions at coding sites, corrected to reflect values at nonsynonymous sites (α) and the ratio of divergence in coding and neutral regions $\frac{div_{coding}}{div_{neu}}$ for human and *D. melanogaster* simulations experiencing different combinations of positive selection parameters: proportion of advantageous (f_{pos}) and mean population scaled selection coefficient of advantageous mutations ($2N_{anc}\overline{s_a}$) following an exponential distribution. Minimum and maximum values are shown, displaying the range across the five demographic scenarios tested for each parameter combination (Table S3-S6). Cells shaded green indicate simulated parameter combinations exhibiting higher divergence in neutral compared to coding regions, as expected in real populations, and α consistent with empirical observations in human and *D. melanogaster* populations (Table 1).

Human simulations		$2N_{anc}\overline{s_a}$		
		10	100	500
f_{pos}	0.001	$\frac{div_{coding}}{div_{neu}}$ = 0.54-0.55	$\frac{div_{coding}}{div_{neu}}$ = 0.59-0.63	$\frac{div_{coding}}{div_{neu}}$ = 0.80-0.89
		α = 0.02-0.03	α = 0.16-0.24	α = 0.40-0.47
	0.01	$\frac{div_{coding}}{div_{neu}}$ = 0.59-0.61	$\frac{div_{coding}}{div_{neu}}$ = 0.99-1.13	$\frac{div_{coding}}{div_{neu}}$ = 1.58-1.70
		α = 0.16-0.23	α = 0.58-0.65	α = 0.71-0.74

D. melanogaster simulations		$2N_{anc}\overline{s_a}$		
		20	250	10000
f_{pos}	0.005	$\frac{div_{coding}}{div_{neu}}$ = 0.48-0.52	$\frac{div_{coding}}{div_{neu}}$ = 0.91-0.98	
		α = 0.11-0.19	α = 0.40-0.47	
	0.0002	$\frac{div_{coding}}{div_{neu}}$ = 0.43-0.49	$\frac{div_{coding}}{div_{neu}}$ = 0.48-0.52	
		α = 0.01	α = 0.06-0.11	
	0.000001			$\frac{div_{coding}}{div_{neu}}$ = 0.46-0.51
				α = 0.01