The impact of purifying and background selection on the inference of population history: problems and prospects

Parul Johri[*, ‡], Kellen Riall[*], Hannes Becher[†], Brian Charlesworth[†], and Jeffrey D. Jensen[*, ‡]

[*]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA

[†]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, EH9 3FL, United Kingdom

[‡]Correspondence can be addressed to these authors

Corresponding authors:

Parul Johri

pjohri1@asu.edu

Jeffrey D. Jensen

Jeffrey.D.Jensen@asu.edu

ABSTRACT

Current procedures for inferring population history are generally performed under the assumption of complete neutrality - that is, by neglecting both direct selection and the effects of selection on linked sites. We here examine how the presence of direct purifying and background selection may bias demographic inference by evaluating two commonly-used methods (MSMC and *fastsimcoal2*), specifically studying how the underlying shape of the distribution of fitness effects (DFE) and the fraction of directly selected sites interact with demographic parameter estimation. The results show that, even after masking functional genomic regions, background selection effects may result in the mis-inference of population growth under models of both constant population size as well as decline. This effect is amplified as the strength of purifying selection and the density of directly selected sites increases, as indicated by the distortion of the site frequency spectrum and levels of nucleotide diversity at linked neutral sites. We also show how simulated changes in background selection effects caused by population size changes can be predicted analytically. We propose a potential method for correcting for the mis-inference of population growth caused by selection. By treating the DFE as a nuisance parameter and averaging across all potential realizations, we demonstrate that even directly selected sites may be used to infer demographic histories with reasonable accuracy.

Keywords: demographic inference, background selection, distribution of fitness effects, MSMC, *fastsimcoal2*, approximate Bayesian computation (ABC)

Running title: Demographic inference with selection

INTRODUCTION

The characterization of past population size change is a central goal of population genomic analysis - with applications ranging from anthropological to agricultural to clinical (see review by Beichman *et al.* 2018). Furthermore, use of an appropriate demographic model provides a necessary null model for assessing the impact of selection across the genome (*e.g.*, Teshima *et al.* 2006; Thornton and Jensen 2007; Jensen *et al.* 2019). Multiple strategies have been proposed for performing demographic inference, utilizing expectations related to levels of variation, the site frequency spectrum, linkage disequilibrium, and within- and between-population relatedness (*e.g.*, Gutenkunst *et al.* 2009; Li and Durbin 2011; Lukic and Hey 2012; Harris and Nielsen 2013; Excoffier *et al.* 2013; Bhaskar *et al.* 2015; Sheehan and Song 2016; Ragsdale and Gutenkunst 2017; Steinrücken *et al.* 2019; Kelleher *et al.* 2019; Speidel *et al.* 2019).

Although many methods perform well when evaluated under the standard assumption of neutrality, it is difficult in practice to assure that the nucleotide sites used in empirical analyses experience neither direct selection nor the effects of selection at linked sites. For example, inference is often performed using intergenic, 4-fold degenerate, or intronic sites. While there is evidence for weak direct selection in all of these categories in multiple organisms (*e.g.*, Haddrill *et al.* 2005; Chamary and Hurst 2005; Andolfatto 2005; Lynch 2007; Zeng and Charlesworth 2010; Choi and Aquadro 2016; Jackson *et al.* 2017), it is also clear that such sites near or in coding regions will also experience background selection (BGS; Charlesworth *et al.* 1993; Charlesworth 2013), and may periodically be affected by selective sweeps as well (Messer and Petrov 2013; Schrider et al. 2016). These effects are known to affect the local underlying effective population size, and alter both the levels and patterns of variation and linkage disequilibrium (Charlesworth et al. 1993; Kaiser and Charlesworth 2009; O'Fallon et al. 2010; Charlesworth 2013; Nicolaisen and Desai 2013; Ewing and Jensen 2016; Johri et al. 2020).

However, commonly-used approaches for performing demographic inference that assume complete neutrality, including *fastsimcoal2* (Excoffier et al. 2013) and MSMC/PSMC (Li and Durbin 2011; Schiffels and Durbin 2014), have yet to be thoroughly evaluated in the light of this assumption, which is likely to be violated in practice. There are, however, some exceptions. Rather than investigating existing software, Ewing and Jensen (2016) implemented an approximate Bayesian (ABC) approach to quantify the impact of BGS effects, demonstrating that weak purifying selection can generate a skew towards rare alleles that would be mis-

interpreted as population growth. Under certain scenarios, this resulted in a many-fold mis-inference of population size change. However, the effects of the density of directly selected sites and the shape of the distribution of fitness effects (DFE), which are probably of great importance, have yet to be considered. Spanning the range of these potential parameter values is important for understanding the implications for empirical application. For example, the proportion of the genome experiencing direct purifying selection can vary greatly between species, with estimates ranging from ~3-8% in humans, to ~12% in rice, to 37-53% in *D. melanogaster*, to 47-68% in *S. cerevisiae* (Siepel et al. 2005; Liang et al. 2018). Further, many organisms have highly compact genomes, with ~88% of the *E. coli* genome (Blattner et al. 1997) and effectively all of many virus genomes, being functional (*e.g.*, >95% of the SARS-CoV-2 genome, Wu *et al.* 2020).

While such estimates allow us to approximate the effects of BGS in some model organisms, in which recombination and mutation rates are well known, it is difficult to predict these effects in the vast majority of study systems. Moreover, while the genome-wide mean of *B*, a widely-used measure of BGS effects that measures the level of variability relative to neutral expectation, can range from ~0.45 in *D. melanogaster* to ~0.94 in humans (Charlesworth 2013; but see Pouyet *et al.* 2018), existing demographic inference approaches are usually applied across organisms without considering this important source of differences in levels of bias. Here, we examine the effects of the DFE shape and functional density on two common demographic inference approaches - the multiple sequentially Markovian coalescent (MSMC) and *fastsimcoal2*. Finally, we propose an extension within the approximate Bayesian computation (ABC) framework to address this issue, treating the DFE as a nuisance parameter and demonstrating greatly improved demographic inference even when using directly selected sites alone.

RESULTS and DISCUSSION

Effects of SNP numbers and genome size on inference under neutral equilibrium

The accuracy and performance of demographic inference was evaluated using two popular methods, MSMC (Schiffels and Durbin 2014) and *fastsimcoal2* (Excoffier et al. 2013). In order

125 to assess performance, it was first necessary to determine how much genomic information is
126 required to make accurate inference when the assumptions of neutrality are met. Chromosomal
127 segments of varying sizes (10 Mb, 50 Mb, 200 Mb, and 1 Gb) were simulated under neutrality
128 and demographic equilibrium (*i.e.,* a constant population size of 5000 diploid individuals) with
129 100 independent replicates each. For each replicate this amounted to the mean [SD] number of
130 segregating sites for each diploid individual being 1,944 [283], 9,996 [418], 40,046 [957] and
131 200,245 [1887]; for 50 diploid individuals, these values were 10,354 [225], 51,863 [567],
132 207,118 [1139] and 1,035,393 [2476] for 10 Mb, 50 Mb, 200 Mb and 1 Gb, respectively. Use of
133 MSMC resulted in incorrect inferences for all segments smaller than 1 Gb (Supp Figure 1).
134 Specifically, very strong recent growth was inferred instead of demographic equilibrium,
135 although ancestral population sizes were correctly estimated. In addition, when 2 or 4 diploid
136 genomes were used for inference, MSMC again inferred a recent many-fold growth for all
137 segment sizes even when the true model was equilibrium, but performed well when using 1
138 diploid genome for large segments (Supp Figure 1).
139 When using *fastsimcoal2* to perform demographic inference, parameters were accurately
140 estimated for all chromosomal segment sizes when the correct model (*i.e*., equilibrium) was
141 specified (left panel of Supp Figure 2). However, when model selection was performed using a
142 choice of three models (equilibrium, instantaneous size change, and exponential size change),
143 64% of replicates of the 10 Mb chromosomal segment incorrectly inferred a size change (Supp
144 Figure 3a), although the current population sizes were estimated accurately (middle panel of
145 Supp Figure 2). When the instantaneous bottleneck model was added to the set of competing
146 models, the true model (equilibrium) was correctly inferred in only 11-33% of replicates (right
147 panel of Supp Figure 2; Supp Figure 3b). These results suggest caution when performing
148 inferences on smaller regions or genomes, specifically when the number of SNPs are less than
149 ~200,000 per single diploid individual, or less than 1 million per 50 diploid individuals. Extra
150 caution should be used when interpreting population size changes inferred by MSMC when
151 using more than 1 diploid individual.
152 Given this performance, all further analyses were restricted to characterizing
153 demographic inference on data that roughly matched the structure and size of the human genome
154 - for every diploid individual, 22 chromosomes (autosomes) of size 150 Mb each were simulated,
155 which amounted to roughly 3 Gb of total sequence. Ten independent replicates of each parameter

156 combination were performed throughout, and inference utilized 1 and 50 diploid individuals for
157 MSMC and *fastsimcoal2,* respectively.
158
159 **Effect of the strength of purifying selection on demographic inference**
160 In order to test demographic inference in the presence of BGS, all 22 chromosomes were
161 simulated with exons of size 350 bp each, with varying sizes of introns and intergenic regions
162 (see Methods) being used in order to vary the fraction (5%, 10% and 20%) of the genome under
163 selection. Because the strength of selection acting on deleterious mutations affects how far the
164 effects of BGS extend, demographic inference was evaluated for various DFEs (Table 1). The
165 DFE was modelled as a discrete distribution with four fixed classes: $0 \leq 2N_e s < 1$, $1 \leq 2N_e s \leq$
166 10, $10 < 2N_e s \leq 100$ and $100 < 2N_e s \leq 2N_e$, where N_e is the effective population size and s is the
167 reduction in fitness of the homozygous mutant relative to wildtype. The fitness effects of
168 mutations were uniformly distributed within each bin, assumed to be semi-dominant, and the
169 DFE shape was altered by varying the proportion of mutations belonging to each fixed class,
170 given by f_0, f_1, f_2, and f_3, respectively (see Methods). Three DFEs highly skewed towards a
171 particular class were initially used to assess the impact of the strength of selection on
172 demographic inference (with the remaining mutations equally distributed amongst the other three
173 classes):
174 DFE1: a DFE in which 70% of mutations have weakly deleterious fitness effects (*i.e.*, $f_1 = 0.7$),
175 DFE2: a DFE in which 70% of mutations have moderately deleterious fitness effects (*i.e.*, $f_2 =$
176 0.7), and DFE3: a DFE in which 70% of mutations have strongly deleterious fitness effects (*i.e.*,
177 $f_3 = 0.7$).
178 In order to understand the effects of BGS, exonic sites were masked, and only linked
179 neutral intergenic sites were used for demographic inference. The three demographic models
180 examined were (1) demographic equilibrium, (2) a 30-fold exponential growth, mimicking the
181 recent growth experienced by European human populations, and (3) ~6-fold instantaneous
182 decline, mimicking the out-of-Africa bottleneck in human populations (Figure 1a). Although
183 these models were parameterized using previous estimates of human demographic history (Supp
184 Table 1; Gutenkunst *et al.* 2009), these basic demographic scenarios are applicable to many
185 organisms, although the magnitudes of population size changes in this case may represent an
186 extreme. Although for the 30-fold exponential growth model the ancestral population size was

consistently over-estimated when using *fastsimcoal2*, even under neutrality, the site frequency spectrum (SFS) obtained from the test simulations and from the best-fit model inferred by *fastsimcoal2* are virtually indistinguishable (Supp Table 2). In other words, the model fitting is working correctly, but the models themselves result in nearly identical expected frequency spectra (and see Rosen et al. 2018). In order to better characterize this mis-inference, we simulated 2-fold, 5-fold, 10-fold, and 20-fold exponential growth for comparison. Generally speaking, *fastsimcoal2* performs better under models of instantaneous size change compared to those involving exponential change (Supp Table 3).

Under demographic equilibrium, when 20% of the genome experiences direct selection (with masking of the directly selected sites), we found the true population size to be underestimated, as well recent population growth mis-inferred (Figure 1), with a larger bias in MSMC than *fastsimcoal2*. Stronger growth is inferred with larger fitness effects of mutations. Conversely, when the true demographic model is characterized by recent 30-fold growth, demographic inference is accurate and performs equally well for both MSMC and *fastsimcoal2*, with the exception of a slight underestimation of the ancestral population size for all DFE types. When the true model is population decline, weakly deleterious mutations alone did not affect inference drastically with either method, and it was possible to recover the true model by *fastsimcoal2*. However, moderately and strongly deleterious mutations resulted in an underestimation of population size and an inference of strong growth, to the extent that population decline would be misinterpreted as growth. We further tested the effect of BGS on demographic inference when changes in population size were less severe, namely, when population growth and decline was only 2-fold, with qualitatively similar results (Supp Figure 4).

Finally, given the strong evidence that most organisms have a bi-modal DFE with a significant proportion of strongly deleterious or lethal mutations (Sanjuán 2010; Jacquier et al. 2013; Kousathanas and Keightley 2013; Bank et al. 2014; Charlesworth 2015), we investigated the effect of this strongly deleterious class further. Thus, for comparison with the above, we simulated a rather extreme case in which 30% or 50% of exonic mutations are strongly deleterious with fitness effects uniformly sampled between $100 \leq 2N_{\text{anc}}s \leq 2N_{\text{anc}}$, with the remaining mutations being neutral (*i.e.*, DFE5 and DFE6; see Table 1). As with the above results, both equilibrium and decline models were falsely inferred as growth, with an order of magnitude underestimation of the true population size (Figure 2).

In sum, neglecting BGS frequently results in the inference of population growth, almost regardless of the true underlying demographic model.

Effects of density and inclusion/exclusion of directly selected sites on inference

Although we have shown that the presence of purifying selection biases demographic inference, the extent of mis-inference necessarily depends on the fraction of the genome experiencing direct selection. We therefore compared models in which 5%, 10% or 20% of the genome was functional. For this comparison, equal proportions of mutations in each DFE bin were assumed (*i.e.,* $f_0 = f_1 = f_2 = f_3 = 0.25$) corresponding to DFE4 (Table 1). As before, when the true model was growth, inference was unbiased, with a slight underestimation of ancestral population size when 20% of the genome experiences selection (Figure 3). Population decline was inferred reasonably well if less than 10% of the genome experienced direct selection, but could be mis-inferred as growth with greater functional density, as shown in Figure 3. Similarly, the extent to which population size is under-estimated at demographic equilibrium increases with the fraction of the genome under selection. Finally, it is noteworthy that many changes in population size that were falsely inferred were greater than 2-fold in size, suggesting the need for great caution when inferring such changes from real data.

Importantly, the results presented do not significantly differ between inference performed while including directly selected sites (*i.e.*, no masking of functional regions; Supp Figure 5) versus inference performed using linked neutral sites (*i.e.*, masking functional regions; Figures 1-3). These results suggest that the exclusion of exonic sites, which is often assumed to provide a sufficiently neutral dataset to enable accurate demographic inference, is not necessarily a satisfactory solution unless gene density is low.

Effect of BGS on model selection and inferred time of size change using *fastsimcoal2*

In order to quantify the effects of BGS on model selection, four competing models were used for inference: equilibrium, instantaneous size change (growth/decline), exponential size change (growth/decline), and an instantaneous bottleneck. Although demographic equilibrium was almost always inferred as instantaneous size change (70-100% of replicates), when there existed a substantial proportion of highly deleterious mutations (DFE3 and DFE5) exponential growth was generally inferred instead. However, when there was a true size change, *fastsimcoal2*

249 performed well in distinguishing between exponential *vs.* instantaneous change models even in
250 the presence of BGS (Supp Figure 6), provided that the magnitude of size change was large.
251 When size changes were on the order of 2-fold, exponential growth was consistently inferred to
252 be instantaneous.

253 With respect to model choice between growth and decline in the presence of BGS
254 (irrespective of instantaneous *vs.* exponential change), as the density of selected sites and
255 strength of purifying selection increased, both equilibrium and decline models were more likely
256 to be inferred as growth (Supp Figure 7), while true growth models were generally chosen
257 correctly. It should be added that with such large chromosome sizes (3 Gb of total sequence
258 data), model selection was not observed to vary between replicates using *fastsimcoal2* for any
259 given parameter combination. Thus, in the presence of BGS, high-confidence calls of an
260 incorrect underlying demographic model appear likely.

261 With regards to the time of inferred size change, when the true model was exponential
262 growth, the model was always correctly identified and inference of the time of change was only
263 slightly under-estimated in the presence of BGS (Supp Figure 8 and 9), consistent with the fact
264 that BGS will further skew the site frequency spectrum towards rare alleles. When the true model
265 was decline, and the model was correctly identified as such, the time of change was modestly
266 over-estimated (Supp Figure 10, 11) – up to 1.3-fold for 6× growth and 2.9-fold for 2× growth
267 (when 20% of the genome is exonic). However, when the decline model was incorrectly
268 identified as growth, the timing of size change was severely under-estimated (Supp Figure 10,
269 11).

270

271 **Effect of heterogeneity in recombination rates, mutation rates, and repeat masking**

272 Variation in recombination and mutation rates, as well as the masking of repeat regions, may
273 also affect demographic inference procedures. We evaluated this issue by simulating
274 heterogeneity in both mutation and recombination rates (based on estimated human genome
275 maps, as described in the Methods section), and masking 10% of each simulated segment
276 drawing from the empirical distribution of repeat lengths in the human genome (Supp Figure 12).
277 In general, inferences under neutrality (Supp Figures 13-15) as well as under BGS (Supp Figures
278 16-18) were unaffected under all demographic models. These investigations suggest that such
279 demographic inference is robust to genome-wide variation in rates of recombination and

mutation, as well as the presence of repeat elements, at least for the parameters considered here. Thus, serious mis-inference is more likely to be caused by selection. These observations also suggest that simulations performed with mean rates of recombination and mutation, as in this study, are sufficient to evaluate biases caused by BGS.

Effects of BGS on diversity and the SFS under various demographic models: theoretical expectations versus simulation results

To better understand how BGS can lead to different biases in the inference of population history, we investigated the extent of BGS effects under all three demographic models, with respect to both the reduced diversity relative to neutrality, as well as the shape of the SFS at linked neutral sites. First, we found that B, the diversity relative to the neutral expectation, differed among demographic scenarios, disproportionately amplifying mis-inference under equilibrium and decline (Figure 4). After a population decline, B was lower than that before the size change; while after population expansion, B increased relative to that in the ancestral population, sometimes approaching 1 (Figure 4). This may seem paradoxical, given that the magnitude of the scaled selection coefficient ($2N_es$) decreases with decreasing N_e (*i.e.*, the efficacy of purifying selection decreases, and could thus be expected to result in larger values of B under population decline). Conversely, with increasing N_e, B may be reduced.

However, these expectations apply only once a population has maintained a given N_e for sufficient time such that equilibrium has been approached. During the initial stages of population size change, and shortly afterwards, the dynamics of B tend to show a trend opposite to this long-term expectation (see also Figure 5 of Torres *et al.* 2020). This is because differences in N_e caused by different initial levels of BGS cause differences in the rates of response to changes in population size – a small value of N_e (corresponding to low B) results in a faster response compared with a high value (Fay and Wu 1999; Hey and Harris 1999; Pool and Nielsen 2007; Pool and Nielsen 2009; Campos et al. 2014; Torres et al. 2020). The relative diversity values observed with different initial equilibrium B values after a short period of population size change may thus be very different from both the initial and final equilibrium values, so that the apparent B values estimated by comparing diversities with and without BGS differ from the equilibrium values. The overall effect is that there is an apparent increase in B immediately following a population decline, and a decrease immediately following an expansion. Analytical models

311 describing these effects are presented in the Appendix. These models used the simulated values
312 of *B* at equilibrium before the population size changes to predict the apparent *B* values at the
313 ends of the periods of size change (see the Methods and Appendix). It can be seen from Figure 4
314 that there is good agreement between these predictions and the simulation results.
315 Because several demographic estimation methods are based on fitting a demographic
316 model to the SFS, it is also of interest to understand whether BGS can skew the SFS to different
317 extents under different demographic models. Although it is well known that BGS causes a skew
318 of the SFS towards rare variants under equilibrium models (Charlesworth et al. 1995; Nicolaisen
319 and Desai 2013), the effect of BGS on the SFS with population size change has not been much
320 explored (but see Johri *et al.* 2020; Torres *et al.* 2020). As shown in Figure 5, with a population
321 size decline, the SFS of derived alleles is more skewed towards rare variants when BGS is
322 operating, especially when *B* is initially small, since the effects of BGS work in opposition to the
323 effects of the population size reduction. This difference in the left skew of the SFS with and
324 without BGS is much less noticeable in the case of population expansion, since here the effects
325 of BGS and the expansion act in a similar direction.
326 As with the estimates of the apparent *B* values discussed above, analytical predictions of
327 the expected SFS after an instantaneous / exponential change in population size can be made,
328 using the values of *B* at equilibrium before the population size change using the formulae of
329 Polanski and Kimmel (2003) and Polanski *et al.* (2003) for the purely neutral case, as described
330 in the Methods section. Although the shape of the SFS is affected by both demography and BGS,
331 the impact of BGS is often comparatively small. Figure 5 shows that the overall shape of the SFS
332 is predicted reasonably well by the analytical results, although deviations are to be expected for
333 the rare allele classes, which are the most sensitive to demographic change and selection.
334 Overall, the results imply that BGS is more likely to bias demographic inference post-decline
335 compared with post-increase, consistent with the performance of the methods described above.
336 However, it should be noted that the exact patterns observed will depend on the timing of
337 population size changes relative to the time of sampling, as well as the value of *B* prior to the
338 size change. The patterns described here thus represent only a small subset of the possibilities. In
339 addition, all else being equal, in populations with large long-term population sizes (and thus low
340 values of *B*), BGS would be expected to result in even larger biases than those observed here.
341

342 **A potential solution: averaging across all possible DFEs**

343 As shown above, demographic inference can be strongly affected by uncorrected for BGS
344 effects, as well as by direct purifying selection. A potential solution is thus to correct for these
345 effects when performing inference of population history. A widely-used approach to estimating
346 direct selection effects, DFE-alpha, takes a stepwise approach to inferring demography, by using
347 a presumed neutral class (synonymous sites); conditional on that demography, it then estimates
348 the parameters of the DFE (Keightley and Eyre-Walker 2007; Eyre-Walker and Keightley 2009;
349 Schneider et al. 2011; Keightley and Jackson 2018). However, this approach does not include the
350 possibility of effects of selection at linked sites, which can result an over-estimate of population
351 growth, and while the DFE may not be mis-inferred strongly (Kim et al. 2017), there is
352 substantial mis-inference of the DFE if synonymous sites experience direct selection (Johri et al.
353 2020).

354 Building on this idea, Johri et al. (2020) recently proposed an approach that includes both
355 direct and background effects of purifying selection, simultaneously inferring the deleterious
356 DFE and demography. By utilizing the decay of BGS effects around functional regions, they
357 demonstrated high accuracy under the simple demographic models examined. Moreover, the
358 method makes no assumptions about the neutrality of synonymous sites, and can thus be used to
359 estimate selection acting on these sites, as well as in non-coding functional elements. However,
360 this computationally-intensive approach is specifically concerned with jointly inferring the DFE
361 and demographic parameters. As such, if an unbiased characterization of the population history
362 is the sole aim, this procedure may be needlessly involved. We thus here examine the possibility
363 of instead treating the DFE as an unknown nuisance parameter, averaging across all possible
364 DFE shapes, in order to assess whether demographic inference may be improved simply by
365 correcting for these selection effects without inferring their underlying parameter values. This
366 approach utilizes functional (*i.e.*, directly selected) regions, a potential advantage in populations
367 for which only coding data may be available (*e.g.*, exome-capture data; see Jones and Good
368 2016), or more generally in organisms with largely functional genomes.

369 In order to illustrate this approach, a functional genomic element was simulated under
370 demographic equilibrium, 2-fold exponential population growth and 2-fold exponential
371 population decline with four different DFE shapes (as described previously, and shown in Figure
372 6). A number of summary statistics were calculated (see Methods) for the entire functional

region. Inference was first performed assuming strict neutrality, and inferring a one-epoch size change (thus estimating the ancestral (N_{anc}) and current population sizes (N_{cur})). As was found with the other inference approaches examined, population sizes were underestimated and a false inference of population growth was observed in almost all cases when selective effects are ignored (Figure 6).

Next, the assumption of neutrality was relaxed, and mutations were simulated with fitness effects characterized by a discrete DFE, with the same fitness classes given above (f_0, f_1, f_2, f_3). Values for f_i were drawn from a uniform prior between 0 and 1, such that $\sum f_i = 1$. Note that no assumptions were made about which sites in the genomic region were functionally important, or regarding the presence/absence of a neutral class. These directly selected sites were then used to infer demographic parameters. We found that, by varying the shape of the DFE, averaging across all realizations, and estimating only parameters related to population history, highly accurate inference of modern and ancestral population sizes is possible (Figure 6). These results demonstrate that, even if the true DFE of a population is unknown (as will always be the case in reality), it is possible to infer demographic history with reasonable accuracy by approximately correcting for these selective effects.

This proposed method is most applicable to organisms in which recombination rates are reasonably well known. If the assumed recombination rate is 2-fold lower than the true rate, the ABC approach infers growth by over-estimating the current population size; correspondingly, if the assumed recombination rate is higher than the true rate, the current population size is under-estimated (Supp Figure 19). Interestingly, in both cases the ancestral population sizes are correctly inferred, consistent with previous results (Johri *et al*. 2020).

CONCLUSION

While commonly used approaches for inferring demography assume neutrality and independence among segregating sites, it is very difficult to verify those assumptions empirically. In addition, there is considerable evidence for wide-spread effects of selection on linked sites in many commonly studied organisms (Hernandez et al. 2011; Cutter and Payseur 2013; Williamson et al. 2014; Elyashiv et al. 2016; Campos et al. 2017; Booker and Keightley 2018; Pouyet et al. 2018; Torres et al. 2018; Castellano et al. 2020). Accordingly, we have explored how violations of the

404 assumption of neutrality may affect demographic inference, particularly with regard to the
405 underlying strength of purifying selection and the genomic density of directly selected sites.
406 Generally speaking, the neglect of these effects (*i.e.*, background selection) results in an
407 inference of population growth, with the severity of the growth model roughly scaling with
408 selection strength and density. Thus, when the true underlying model is in fact growth,
409 demographic mis-inference is not particularly severe; when the true underlying model is constant
410 size or decline, the mis-inference can be extreme, with a many-fold underestimation of
411 population size.

412 It is important to note that BGS effects extend to genomic distances in a way that is
413 positively related to the strength of purifying selection. For instance, strongly and moderately
414 deleterious mutations affect patterns of diversity at large genomic distances, whereas mildly
415 deleterious mutations primarily skew allele frequencies at adjacent sites. Thus, if intergenic
416 regions further away from exons are used to perform demographic inference, it is predominantly
417 strongly deleterious mutations that are likely to bias inferences. In contrast, if synonymous sites
418 are used to infer demographic history, mildly deleterious mutations will be most important. Thus,
419 as we here focused on relatively sparsely-coding genomes (resembling human-like gene
420 densities) and used intergenic sites for inference, moderately and strongly deleterious mutations
421 resulted in more severe mis-inference. In addition, the effect of decay of BGS due to mildly
422 deleterious mutations depends on multiple parameters. For instance, with an exon of length 500
423 bp and *Drosophila*-like parameters (*e.g.,* $N_e = 10^6$; recombination rate = mutation rate = 10^{-8} /
424 site / generation), B will increase from 0.53 (at 10 bases from the end of the exon) to 0.94 at a
425 distance of 1000 bases. On the other hand, with human-like parameters ($N_e = 10^4$; recombination
426 rate = mutation rate = 10^{-8} / site / generation) the corresponding change in B is only 0.981 to
427 0.982 (Supp Table 4). Thus, mildly deleterious mutations will have drastically different effects in
428 this regard, depending on the underlying population parameters. Hence, a large effect of
429 selection is expected when such approaches are applied to synonymous sites in smaller and more
430 compact genomes with higher gene densities, with a larger input of deleterious mutations at
431 closely linked sites.

432 Comparing the two inference methods investigated here, it appears that *fastsimcoal2* is
433 less prone to inferring false fluctuations in population size. However, both methods falsely infer
434 growth in the presence of BGS, with increasing severity as the density of coding regions

435 increases. The times of population growth inferred by both methods appear to be affected in
436 unpredictable ways when the inferred model is incorrect. When the general model is correctly
437 identified, BGS leads to inference of more recent growth, and more ancient decline, than the
438 reality. Further, we observed little difference in performance when using all SNPs relative to
439 when SNPs are thinned to be separated by 5kb (Supp Figure 20), presumably because the models
440 investigated do not generate strong LD. Overall, the degree of mis-inference caused by a neglect
441 of BGS is largely similar between the two methods. Although variation in mutation and
442 recombination rates across the genome alone did not strongly affect demographic inference by
443 either method, our evaluations in the current study are restricted to a specific parameter space
444 resembling those of human populations. The effects of this variation on organisms with more
445 extreme rate fluctuations remain in need of investigation.
446 It is noteworthy that, even when all sites are strictly neutral or only 5% of the genome
447 experiences direct selection, demographic equilibrium is mis-estimated by MSMC as a series of
448 size changes. The pattern of these erroneous size changes lend a characteristic shape to the
449 MSMC curve (*i.e.*, ancient decline and recent growth) which appears to resemble the
450 demographic history previously inferred for the Yoruban population (Schiffels and Durbin
451 2014), including the time at which changes in population size occurred (Supp Figure 21).
452 Previous work has demonstrated that the resulting demographic model does not in fact fit the
453 observed SFS in the Yoruban population (Beichman et al. 2017; Lapierre et al. 2017). A similar
454 shape has also been inferred in the vervet subspecies (Warren *et al.* 2015; Figure 4), in passenger
455 pigeons (Hung *et al.* 2014; Figure 2), in elephants (Palkopoulou *et al.* 2018; Figure 4), in
456 *Arabidopsis* (Fulgione *et al.* 2018; Figure 3), and in grapevines (Zhou *et al.* 2017; Figure 2A).
457 Although the inferred population size fluctuations under simulated neutrality are only
458 ~1.2-fold, in most empirical applications the fluctuations are of a somewhat larger magnitude (~
459 2-fold in pigeons, *Arabidopsis*, and grapevines). Nonetheless, this neutral performance of
460 MSMC under demographic equilibrium is concerning, and adds to the other previously published
461 cautions concerning the interpretation of MSMC results. For example, Mazet *et al*. (2016) and
462 Chikhi *et al*. (2018) demonstrated that, under constant population size with hidden structure,
463 MSMC may suggest false size changes (see also Orozco-terWengel 2016). In addition, MSMC
464 has been reported to falsely infer growth prior to instantaneous bottlenecks (Bunnefeld et al.
465 2015). In addition, we observed that if insufficient genomic data is utilized, or more than one

466 diploid genome is used to perform inference, MSMC falsely infers recent growth of varying
467 magnitudes (as previously observed by Beichman et al. 2017).
468 In sum, we find that the effects of purifying and background selection result in similar
469 demographic mis-inference across approaches, and that masking functional sites does not yield
470 accurate parameter estimates. In order to side-step many of these difficulties, our proposed
471 approach of inferring demography by averaging selection effects across all possible DFE shapes
472 within an ABC framework appears to be promising. Utilizing only functional regions, we found
473 a great improvement in accuracy, without making any assumptions regarding the true underlying
474 shape of the DFE or the neutrality of particular classes of sites. As such, this approach represents
475 a more computationally efficient avenue if only demographic parameters are of interest, and
476 ought to be particularly useful in the great majority of organisms in which unlinked neutral sites
477 either do not exist, or are difficult to identify and verify.
478
479
480 **METHODS**
481 ***Simulations***: Simulations were performed using SLiM 3.1 (Haller and Messer 2019) with 10
482 replicates per evolutionary scenario. For every replicate, 22 chromosomes of 150Mb each were
483 simulated, totaling ~3 Gb of information per individual genome (similar to the amount of
484 information in a human genome). Within each chromosome, 3 different types of regions were
485 simulated, representing non-coding intergenic, intronic, and exonic regions. Based on the NCBI
486 RefSeq human genome annotation, downloaded from the UCSC genome browser for hg19
487 (http://genome.ucsc.edu/; Kent *et al.* 2002), mean values of exon sizes and intron numbers per
488 gene were calculated. To represent mean values for the human genome (Lander et al. 2001), each
489 gene was comprised of 8 exons and 7 introns, and exon lengths were fixed at 350 bp. By varying
490 the lengths of the intergenic and intronic regions, three different genomic configurations with
491 varying densities of functional elements were simulated and compared - with 5%, 10% and 20%
492 of the genome being under direct selection - hereafter referred to as genome5, genome10, and
493 genome20, respectively. Genome5 was comprised of introns of 3000 bp and intergenic sequence
494 of 31000 bp, genome10 of introns of 1500 bp and intergenic sequence of 15750 bp, while
495 genome20 was comprised of introns of 600 bp and intergenic sequence of 6300 bp. The total
496 chromosome sizes of these genomes were approximately 150 Mb (150,018,599 bp, 150,029,949

497 bp, 150,003,699 bp in genome5, 10, and 20, respectively) with 2737, 5164, and 11278 genes per
498 chromosome in genome5, 10 and 20, respectively. In order to be conservative with respect to the
499 performance of existing demographic estimators, intronic and intergenic regions were assumed
500 to be neutral.

501 Recombination and mutation rates were assumed to be equal at 1×10^{-8} /site / generation.
502 Neither crossover interference (see the discussion in Campos and Charlesworth 2019) nor gene
503 conversion were modeled. Exonic regions in the genomes experienced direct purifying selection
504 given by a discrete DFE comprised of 4 fixed classes (Johri et al. 2020), whose frequencies are
505 denoted by f_i: f_0, with $0 \leq 2N_e s < 1$ (*i.e.*, effectively neutral mutations), f_1, with $1 \leq 2N_e s < 10$
506 (*i.e.*, weakly deleterious mutations), f_2, with $10 \leq 2N_e s < 100$ (*i.e.*, moderately deleterious
507 mutations), and f_3, with $100 \leq 2N_e s < 2N_e$ (*i.e.*, strongly deleterious mutations), where N_e is the
508 effective population size and s is the reduction in fitness of the mutant homozygote relative to
509 wild-type. Within each bin, the distribution of s was assumed to be uniform. All mutations were
510 assumed to be semi-dominant. In all cases, the N_e corresponding to the DFE refers to the
511 ancestral effective population size. Six different types of DFE were simulated: DFE1: a DFE
512 skewed largely towards mildly deleterious mutations, given by f_0=0.1, f_1=0.7, f_2=0.1, f_3=0.1;
513 DFE2: a DFE skewed towards moderately deleterious mutations, f_0=0.1, f_1=0.1, f_2=0.7, f_3=0.1;
514 DFE3: a DFE skewed towards strongly deleterious mutations, f_0=0.1, f_1=0.1, f_2=0.1, f_3=0.7;
515 DFE4: a DFE with equal proportions of all mutations, f_0=0.25, f_1=0.25, f_2=0.25, f_3=0.25;
516 DFE5: a DFE with equal proportions of neutral and strongly deleterious mutations, f_0=0.5,
517 f_1=0.0, f_2=0.0, f_3=0.5; and DFE6: a DFE with a majority of neutral mutations and a minority of
518 strongly deleterious mutations, f_0=0.7, f_1=0.0, f_2=0.0, f_3=0.3.

519 Three different demographic models were tested for each of these DFEs (Supp Table 1):
520 1) demographic equilibrium, 2) recent exponential 30-fold growth, resembling that estimated for
521 the human CEU population (Gutenkunst et al. 2009), and 3) ~6-fold instantaneous decline,
522 resembling the out-of-Africa bottleneck in humans (Gutenkunst et al. 2009). For simulations of
523 demographic equilibrium and decline, population sizes and time of change were scaled down by
524 a factor of 10 (with corresponding scaling of the recombination rate, mutation rate, and selection
525 coefficients), while simulations of growth were not scaled.
526

527 ***Running MSMC***: In order to quantify the effect of purifying selection on demographic
528 inference, we used entire chromosomes generated by SLiM to generate input files for MSMC.
529 For comparison, and in order to quantify the effect of BGS alone on demographic inference, we
530 masked the exonic regions to generate input files. For all parameters, MSMC was performed on
531 a single diploid genome, as the results for this case were the most accurate (Supp Figure 1). Input
532 files were made using the script ms2multihetsep.py provided in the msmc-tools-Repository
533 downloaded from https://github.com/stschiff/msmc-tools. MSMC1 and 2 were run as follows:
534 *msmc_1.1.0_linux64bit -t 5 -r 1.0 -o output_genomeID input_chr1.tab input_chr2.tab …*
535 *input_chr22.tab.* Population sizes obtained from MSMC were plotted up to the maximum
536 number of generations obtained from MSMC, and the final value of the ancestral population size
537 was extended indefinitely as a dashed line.
538
539 ***Running Fastsimcoal2***: In order to minimize the effects of linkage disequilibrium (LD), only
540 SNPs separated by 5 kb were used for inference, following Excoffier et al. (2013). Inference was
541 also performed by including all SNPs in order to assess the impact of violating the assumption of
542 no LD on inference. Site frequency spectra (SFS) were obtained for both sets of SNPs for all 10
543 replicates of every combination of demographic history and DFE. SNPs from all 22
544 chromosomes were pooled together to calculate the SFS. *Fastsimcoal2* was used to fit each SFS
545 to 4 distinct models: (a) equilibrium, which estimates only a single population size parameter
546 (N); (b) instantaneous size change (decline/growth), which fits 3 parameters - ancestral
547 population size (N_{anc}), current population size (N_{cur}), and time of change (T); (c) exponential size
548 change (decline/growth), which also estimates 3 parameters - N_{anc}, N_{cur} and T; and (d) an
549 instantaneous bottleneck model with 3 parameters – N_{anc}, intensity, and time of bottleneck. The
550 parameter search ranges for both ancestral and current population sizes in all cases were
551 specified to be uniformly distributed between 100-500000 individuals, while the parameter range
552 for time of change was specified to be uniform between 100-10000 generations in all models.
553 The intensity of the bottleneck, as specified by the population size during the reduction, was also
554 sampled uniformly from the range 100-500000 individuals. The following command line was
555 used to run *fastsimcoal2*:
556 *fsc26 -t demographic_model.tpl -n 150000 -d -e demographic_model.est -M -L 50 -q,*

557 Model selection was performed as recommended by Excoffier et al. (2013). For each
558 demographic model, the maximum of maximum likelihoods from all replicates was used to
559 calculate the Akaike Information Criterion (AIC) = 2 × number of parameters – 2 ×
560 ln(likelihood). The relative likelihoods (Akaike's weight of evidence) in favor of the ith model
561 was then calculated by:

562 $$w(i) = \frac{e^{-0.5\Delta_i}}{\sum_{j=1}^{4} e^{-0.5\Delta_j}}$$

563 where $\Delta_i = AIC_i - AIC_{min}$. The model with the highest relative likelihood was selected as the
564 best model, and the parameters estimated using that model were used to plot the final inferred
565 demography.
566
567 ***Simulations of variable recombination and mutation rates, and repeat masking***: In order to
568 simulate variation in recombination and mutation rates, all 22 chromosomes were simulated by
569 mimicking chromosome 6 (~171Mb) of the human genome. Recombination rates (HapMap)
570 obtained from Yoruban populations (McVean et al. 2004; Myers et al. 2005) were obtained from
571 the UCSC genome browser, while the mutation rate map
572 (https://molgenis26.target.rug.nl/downloads/gonl_public/mutation_rate_map/release2/) was
573 assumed to correspond to estimates obtained from *de novo* mutations (Francioli et al. 2015), as in
574 Castellano *et al.* (2020). Absolute values of mutation rates were normalized in order to maintain
575 the mean mutation rate across the genome at ~ 1.0×10^{-8} per site per generation. Recombination
576 and mutation rate estimates were taken from positions of approximately 10 Mb to 160 Mb, with
577 the recombination map starting at 10010063 bp and the mutation map starting at 10010001 bp.
578 Regions with missing data for either of the two estimates were simulated with rates
579 corresponding to the previous window, except for the case of centromeres in which no
580 recombination was assumed. In order to understand the effect of excluding centromeric regions
581 in empirical studies, the 4Mb region corresponding to the centromere was masked,
582 corresponding to 48.5 to 52.5 Mb of the simulated 150Mb chromosomes. In order to evaluate the
583 effect of masking repeat regions, random segments comprising 10% of each chromosome were
584 masked. The lengths of these segments were drawn from the lengths of repeat regions found in
585 the human genome (Supp Figure 12), as obtained from the repeat regions in the *hg19* assembly
586 of the human genome from the UCSC genome browser.

587

588 ***Performing inference by approximate Bayesian computation (ABC)***: ABC was performed
589 using the R package "abc" (Csilléry et al. 2010), and non-linear regression aided by a neural net
590 (used with default parameters as provided by the package) was used to correct for the
591 relationship between parameters and statistics (Johri et al. 2020). To infer posterior estimates, a
592 tolerance of 0.1 was applied (*i.e.*, 10% of the total number of simulations were accepted by ABC
593 in order to estimate the posterior probability of each parameter). The weighted medians of the
594 posterior estimates for each parameter were used as point estimates. ABC inference was
595 performed under two conditions: (1) complete neutrality, or (2) the presence of direct purifying
596 selection. In both cases only 2 parameters were inferred - ancestral (N_{anc}) and current (N_{cur})
597 population sizes. However, in scenario 2, the shape of the DFE was also varied. Specifically, the
598 parameters f_0, f_1, f_2, and f_3 were treated as nuisance parameters and were sampled such that $0 \leq f_i$
599 ≤ 1, and $\Sigma_i f_i = 1$, for $i = 0$ to 3. In addition, in order to limit the computational complexity
600 involved in the ABC framework, values of f_i were restricted to multiples of 0.05 (*i.e.*, $f_i \in \{0.0,$
601 $0.05, 0.10, \ldots, 0.95, 1.0\}\ \forall\ i$), which allowed us to sample 1,771 different DFE realizations.
602 Simulations were performed with functional genomic regions, and the demographic model was
603 characterized by 1-epoch changes in which the population either grows or declines exponentially
604 from ancestral to current size, beginning at a fixed time in the past.

605 For the purpose of illustration, and for a contrast with the human-like parameter set
606 above, parameters for ABC testing were selected to resemble those of *D. melanogaster* African
607 populations. Priors on ancestral and current population sizes were drawn from a uniform
608 distribution between 10^5-10^7 diploid individuals, while the time of change was fixed at 10^6
609 ($\sim N_e$) generations. In order to simulate functional regions, 94 single-exon genes, as described in
610 Johri *et al*. (2020) and provided in
611 https://github.com/paruljohri/BGS_Demography_DFE/blob/master/DPGP3_data.zip, were
612 simulated with recombination rates specific to those exons (https://petrov.stanford.edu/cgi-
613 bin/recombination-rates_updateR5.pl) (Fiston-Lavier et al. 2010; Comeron et al. 2012). Mutation
614 rates were assumed to be fixed at 3×10^{-9} per site per generation (Keightley et al. 2009;
615 Keightley et al. 2014).

616 All parameters were scaled by the factor 320 in order to decrease computational time,
617 using the principle first described by Hill and Robertson (1966), and subsequently employed by

618 others (Comeron and Kreitman 2002; Hoggart et al. 2007; Kaiser and Charlesworth 2009; Kim
619 and Wiehe 2009; Uricchio and Hernandez 2014; Campos and Charlesworth 2019). The scaled
620 population sizes thus ranged between ~300–30000 and were reported as scaled values in the
621 main text. One thousand replicate simulations were performed for every parameter combination
622 (N_{anc}, N_{cur}, f_0, f_1, f_2, f_3); for performing ABC inference, 50 diploid genomes were randomly
623 sampled without replacement, and summary statistics were calculated using pylibseq 0.2.3
624 (Thornton 2003). The following summary statistics were calculated across the entire exonic
625 region for every exon: nucleotide site diversity (π), Watterson's θ, Tajima's D, Fay and Wu's H
626 (both absolute and normalized), number of singletons, haplotype diversity, LD-based statistics
627 (r^2, D, D'), and divergence (*i.e.*, number of fixed mutations per site per generation after the burn-
628 in period). Means and variances (between exons) of all of the above (a total of 22) were used as
629 final summary statistics to perform ABC. As opposed to the above examples, in this inference
630 scheme only exonic data (*i.e.*, directly selected sites) were utilized. Test datasets were generated
631 in exactly the same fashion as described above.
632
633 ***Analytical expectations for the relative site frequencies***: To compute the expected relative
634 frequencies of site frequency classes, the approach of Polanski and Kimmel (2003) was
635 followed. They describe a method for computing the "probability that a SNP has *b* mutant
636 bases", which is equivalent to the expected site frequency spectrum (SFS) of derived variants.
637 This method (their equations 3-10) allows for the specification of arbitrary population size
638 histories and sample sizes. For reasons of computational precision, a sample size of 10 diploid
639 genomes was chosen. The demographic scenarios detailed in Supplementary Table 1 were
640 implemented as piecewise functions of the effective population size (counting haploid genomes),
641 and the effect of BGS was included by scaling these functions by values of B before population
642 size change as obtained from the forwards-in-time simulations described above. A Mathematica
643 notebook detailing these results, will be made available upon publication (see data availability
644 statement). In addition, analytical expressions can be obtained for pairwise diversity values when
645 there are step changes or exponential growth in population size, as described in the Appendix
646 and in an example program that calculates diversity values after exponential growth.
647

Data availability: The following data will be made publicly available upon acceptance of the manuscript: (1) Scripts used to perform simulations; (2) Input files used to run *fastsimcoal2*; (3) Scripts used for plotting; (4) Plotted results of MSMC and *fastsimcoal2* for all models and scenarios tested in this work; (5) A Mathematica (version 12.1) notebook detailing calculations of analytical expectations for the relative SFS; (6) An example program (Fortran script) demonstrating how to obtain analytical expressions for values of B after exponential growth. All supplemental files will be made publicly available at https://github.com/paruljohri/demographic_inference_with_selection.

Acknowledgements

We would like to thank Susanne Pfeifer for helpful discussions related to this project, and for feedback on the manuscript. This research was conducted using resources provided by Research Computing at Arizona State University (http://www.researchcomputing.asu.edu) and the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. This work was funded by National Institutes of Health grant R01GM135899 to JDJ.

APPENDIX

There are two simple scenarios for population size change for which simple explicit expressions for the expected pairwise coalescent time or diversity can be obtained, without using the methodology of Polanski and Kimmel (2003) and Polanski *et al.* (2003) – a step change in N or an exponential growth in N. First consider the coalescent process for a step change, where the current and initial effective population sizes are denoted by N_{e1} and N_{e0}, respectively. Let B be the background selection parameter at the start of the process of change, corresponding to effective size N_{e0}. For convenience, time is scaled in units of $2N_{e1}$ generations, and the time of the change in population size on this scale is denoted by T_0, counting back from the present time, $T = 0$. T_0 is assumed to be sufficiently small that B remains approximately constant during the period since the change in size. Denote the ratio N_{e0}/N_{e1} by

675 R. The derivation for the case of a step change in population size is similar to that given by Pool and
676 Nielsen (2009) for the purpose of comparing X chromosomes and autosomes.
677 Between times T and T_0, coalescence occurs at a rate B^{-1} on the chosen timescale, so that the
678 contribution from this period to the net coalescent time for a pair of alleles sampled at $T = 0$ is:
679

680 $$B^{-1}\int_0^{T_0} T\exp(-B^{-1}T)\,\mathrm{d}T = B - B\exp(-B^{-1}T_0) - T_0\exp(-B^{-1}T_0)$$

681
682 There is a probability of $\exp(-B^{-1}T_0)$ that there is no coalescence when T lies between 0 and T_0,
683 after which coalescence occurs at a rate $1/BR$, giving a net contribution to the coalescence time of:
684

685 $$(BR + T_0)\exp\,(-B^{-1}T_0)$$

686
687
688 The net coalescence time for the stepwise change with BGS is given by the sum of these
689 two expressions:
690

691 $$B[1 + (R - 1)\exp(-B^{-1}T_0)] \qquad (1a)$$

692
693 If this expression is compared to the corresponding equation with $B = 1$, the apparent value of B
694 at the time of sampling of the pair of alleles is given by:
695

696 $$B_s = \frac{B[1+(R-1)\exp(-B^{-1}T_0)]}{[1+(R-1)\exp(-T_0)]} \qquad (1b)$$

697
698 Next, consider a process of exponential change in population size, starting at an initial effective
699 size of N_{e0} at t_0 generations in the past and ending at size N_{e1}, such that the instantaneous growth rate r
700 per generation is r=equal to $\ln(N_{e1}/N_{e0})/t_0$. The effective population size at time t in the past is $N_e(t)$ =
701 $N_{e1}\exp(-rt)$; with BGS, the rate of coalescence at time t is $1/BN_e(t)$. As before, the BGS parameter is
702 assumed to remain constant over the period of population size change. It follows that the probability of
703 no coalescence by generation t in the past is:

$$P_{nc}(t) = \exp\left[-\int_0^t (2BN_{e1})^{-1} \exp(rt)\,\mathrm{d}t\right] = \exp[c^{-1}(1 - e^{rt})] \qquad (2)$$

where $c = 2BNe_1r$.

The pre-growth period with $t > t_0$ contributes an expected coalescent time of $(2BN_{e0} + t_0)P_{nc}(t_0)$, on the scale of generations.

Following Slatkin and Hudson (1991), to obtain the contribution from the period with $t > t_0$, it is convenient to measure time as $\tau = rt$. The probability of coalescence between τ and $\tau + \mathrm{d}\tau$ is then given by:

$$P_c(\tau) = c^{-1}e^{\tau} \exp[c^{-1}(1 - e^{\tau})]\,\mathrm{d}\tau \qquad (3)$$

The contribution from this period to the expected coalescent time is given by the integral of $\tau P_c(\tau)$ between 0 and τ_0. Following Slatkin and Hudson (1991), by transforming to $u = \exp(\tau)$, this contribution can be expressed as the following integral:

$$\bar{\tau}_1 = c^{-1}e^{c^{-1}} \int_1^{u_0} \ln(u) e^{-uc^{-1}} du \qquad (4)$$

This integral can easily be evaluated numerically. The corresponding mean coalescent time on the scale of generations is obtained by division by r, and the result can be added to $(2BN_{e0} + t_0)P_{nc}(t_0)$, yielding the net expected coalescent time. By dividing the resulting expression by the corresponding expression with $B = 1$, the apparent BGS effect at the time of sampling can be obtained, in the same way as for the step change model.

FIGURES AND TABLES

Table 1: Proportion (f_i) of mutations in each class of the discrete distribution of fitness effects (DFE) simulated in this study.

	f_0	f_1	f_2	f_3
DFE1	0.1	0.7	0.1	0.1
DFE2	0.1	0.1	0.7	0.1
DFE3	0.1	0.1	0.1	0.7
DFE4	0.25	0.25	0.25	0.25
DFE5	0.5	0.0	0.0	0.5
DFE6	0.7	0.0	0.0	0.3

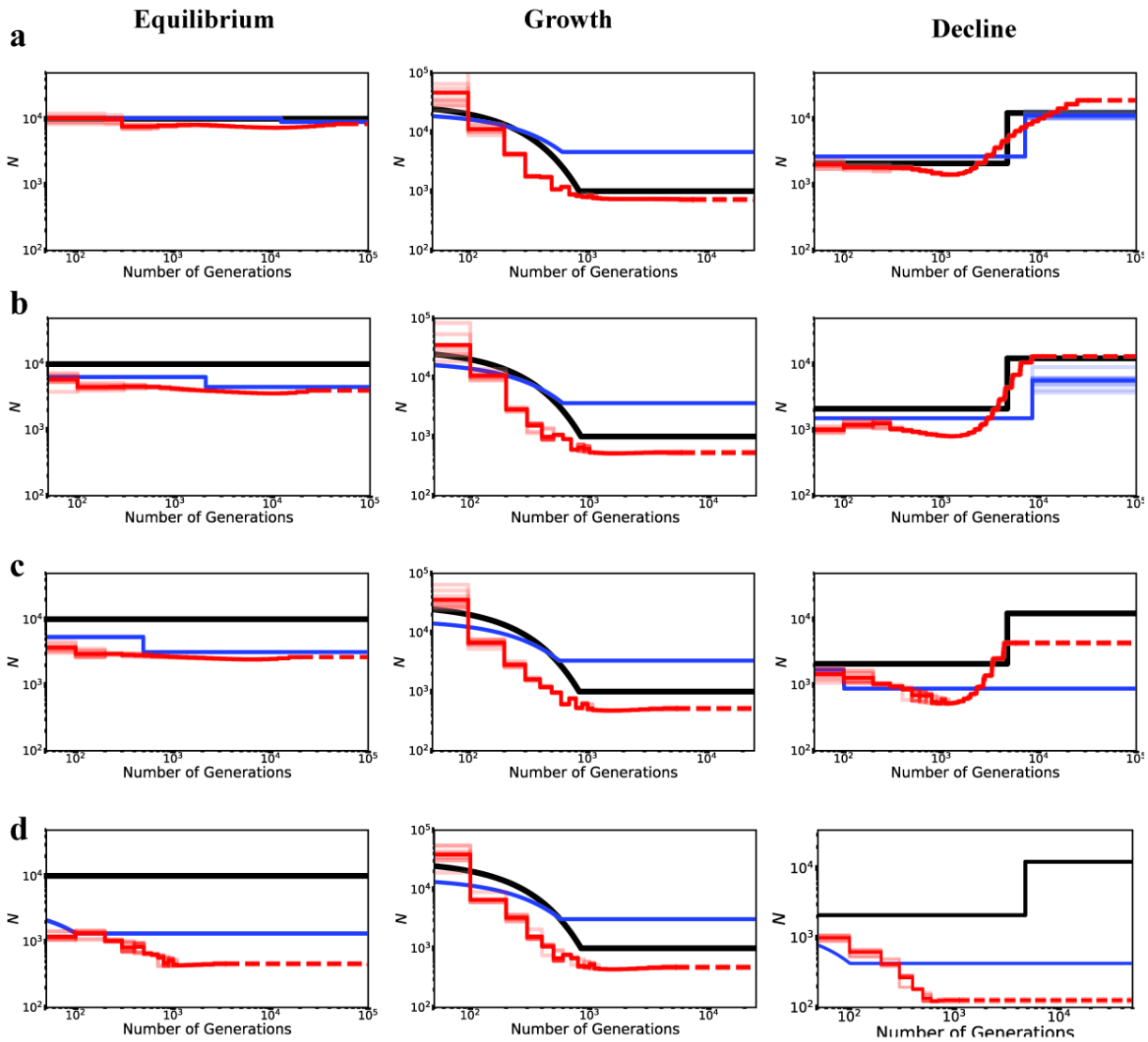


Figure 1: Inference of demography by MSMC (red lines; 10 replicates) and *fastsimcoal2* (blue lines; 10 replicates) with and without BGS, under demographic equilibrium (left column), 30-fold exponential growth (middle column), and ~6-fold instantaneous decline (right column). The true demographic models are depicted as black lines, with the x-axis origin representing the present day. (a) All genomic sites are strictly neutral. (b) Exonic sites experience purifying selection specified by a DFE comprised largely of weakly deleterious mutations (DFE1: $f_0 = 0.1$, $f_1 = 0.7$, $f_2 = 0.1$, $f_3 = 0.1$). (c) Exonic sites experience purifying selection specified by a DFE comprised largely of moderately deleterious mutations (DFE2: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.7$, $f_3 = 0.1$). (d) Exonic sites experience purifying selection specified by a DFE comprised largely of strongly deleterious mutations (DFE3: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.1$, $f_3 = 0.7$). Exons represent 20% of the genome, and exonic sites were masked/excluded when performing demographic inference, quantifying the effects of BGS alone. The dashed lines represent indefinite extensions of the ancestral population sizes.

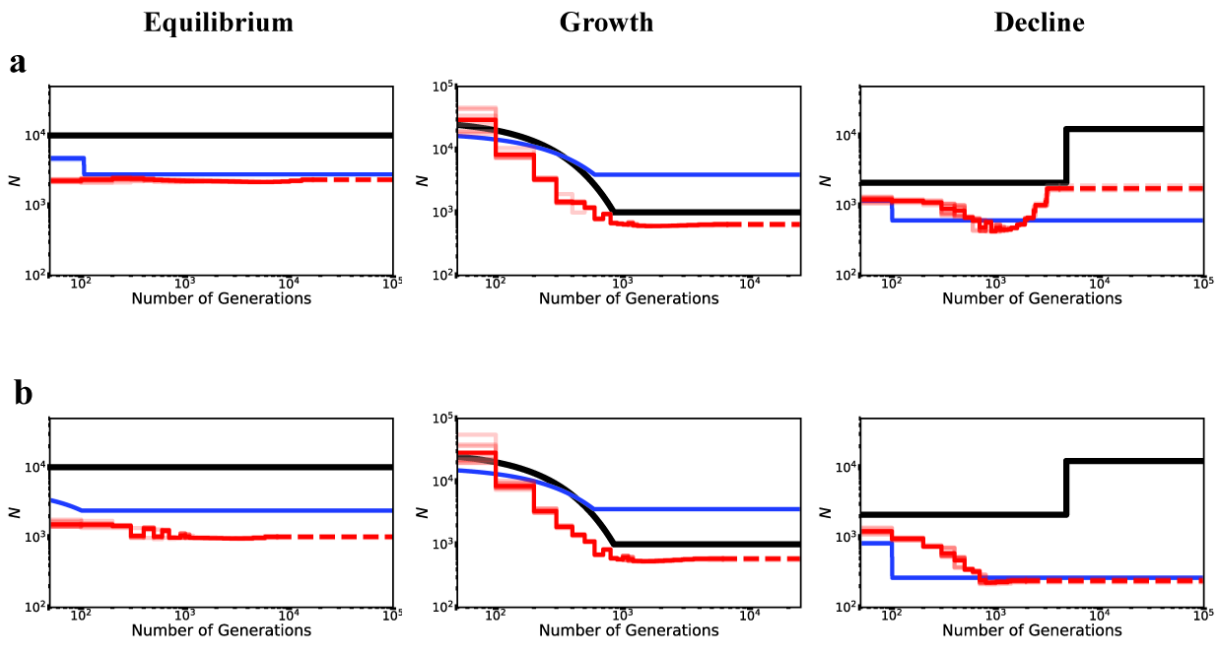


Figure 2: Inference of demography by MSMC (red lines; 10 replicates) and *fastsimcoal2* (blue lines; 10 replicates) in the presence of BGS generated by strongly deleterious mutations. Directly selected sites comprised 20% of the genome and were masked when performing demographic inference. (a) Exons experience purifying selection, with 30% of new mutations being strongly deleterious, and the remaining being neutral (specified by DFE6). (b) Exons experience purifying selection, with 50% of new mutations being strongly deleterious, and the remainder being neutral (specified by DFE5). The true demographic models are given as black lines, with the x-axis origin representing the present day. The dashed lines represent indefinite extensions of the ancestral population sizes.

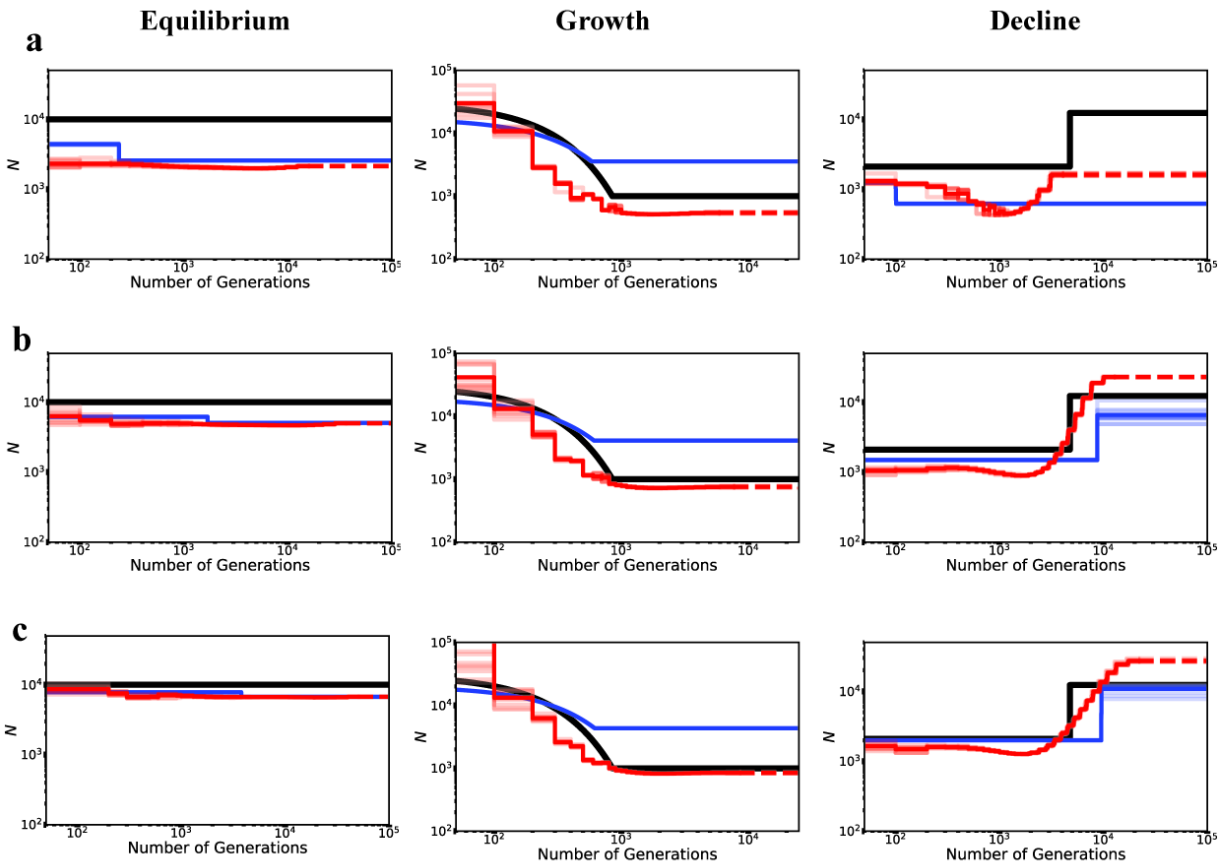


Figure 3: Inference of demography by MSMC (red lines; 10 replicates) and *fastsimcoal2* (blue lines; 10 replicates) in the presence of BGS with varying proportions of the genome under selection, for demographic equilibrium (left column), exponential growth (middle column), and instantaneous decline (right column). Exonic sites were simulated with purifying selection with all f_i values equal to 0.25 (DFE4), and were masked when performing inference. Directly selected sites comprise (a) 20% of the simulated genome, (b) 10% of the simulated genome, and (c) 5% of the simulated genome. The true demographic models are given by the black lines, with the x-axis origin representing the present day. The dashed lines represent indefinite extensions of the ancestral population sizes.

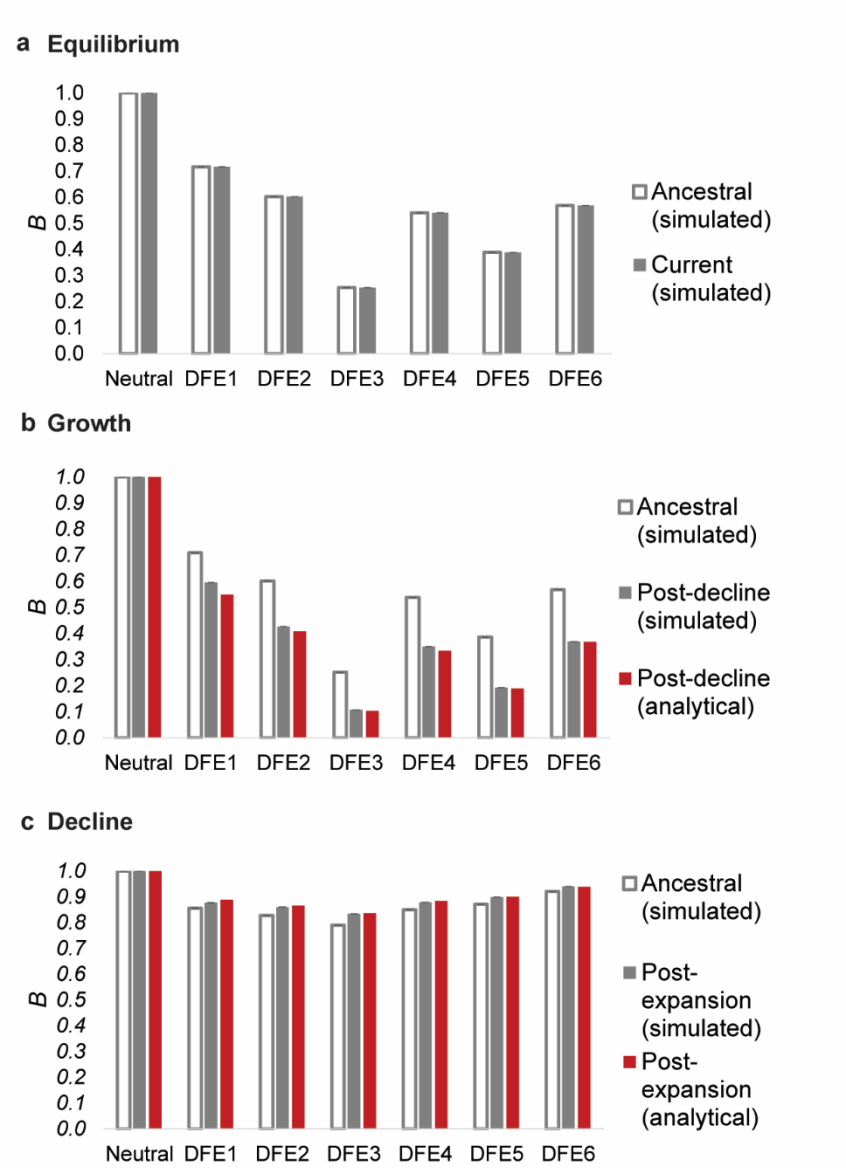


Figure 4: B is the nucleotide site diversity (π) with BGS relative to its purely neutral expectation (π_0). The results are shown for (a) demographic equilibrium, (b) population growth, and (c) population decline. All cases refer to size changes forward in time, the ancestral B (*i.e.*, B pre-change in population size) is shown in white bars, B post-change in population size is shown in solid gray bars, and the analytical expectations for the post-size change B is shown as red bars. Various DFEs were used: **DFE1**: $f_0 = 0.1$, $f_1 = 0.7$, $f_2 = 0.1$, $f_3 = 0.1$. **DFE2**: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.7$, $f_3 = 0.1$. **DFE3**: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.1$, $f_3 = 0.7$. **DFE4**: $f_0 = f_1 = f_2 = f_3 = 0.25$. **DFE5:** $f_0 = 0.5$, $f_1 = 0.0$, $f_2 = 0.0$, $f_3 = 0.5$. **DFE6:** $f_0 = 0.7$, $f_1 = 0.0$, $f_2 = 0.0$, $f_3 = 0.3$. Exonic sites comprised ~10% of the genome, roughly mimicking the density of the human genome.

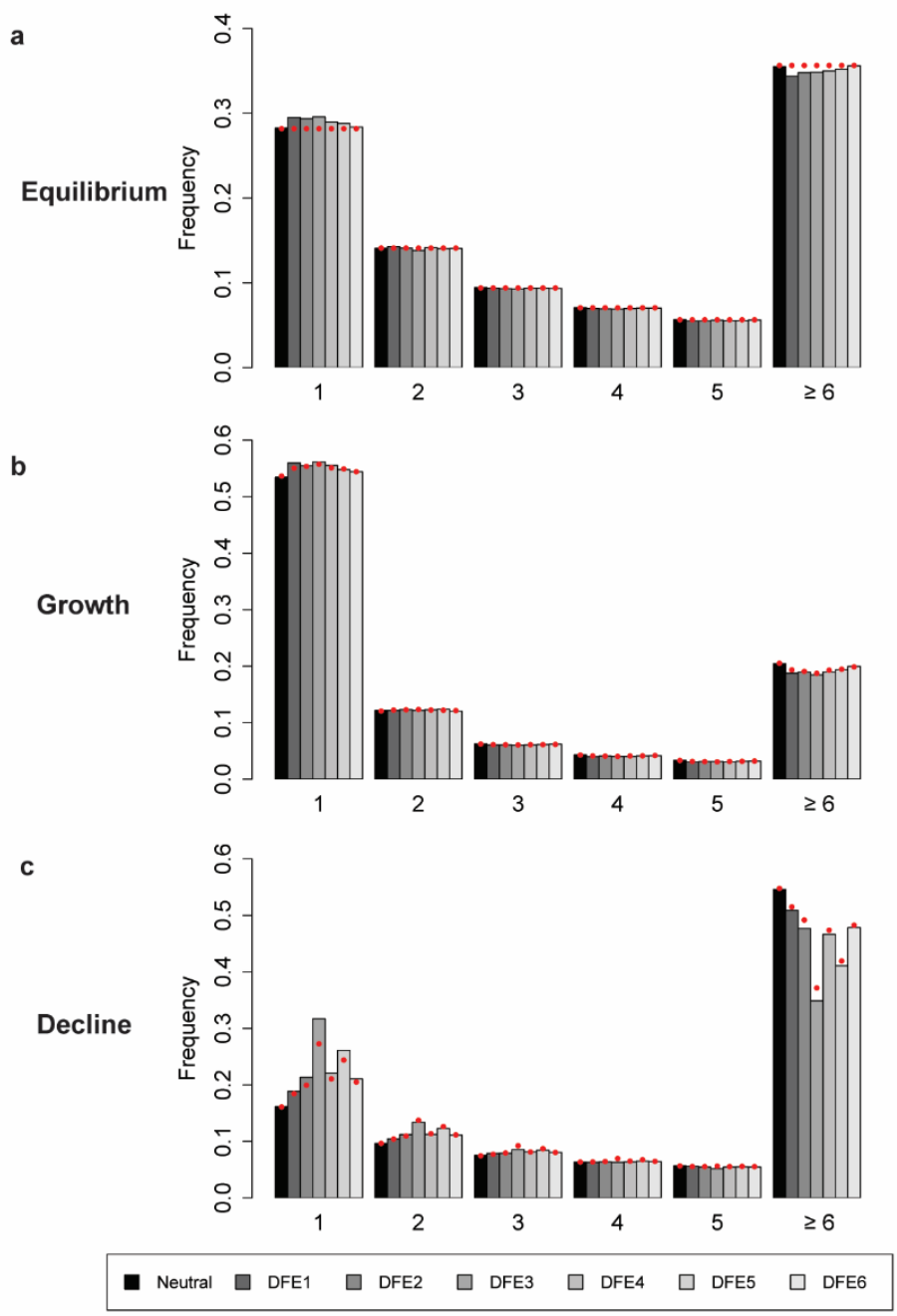


Figure 5: The site frequency spectrum (SFS) of derived allele frequencies at neutral sites from 10 diploid genomes under (a) demographic equilibrium, (b) population growth, and (c) population decline, under the same DFEs as shown in Figure 4. The x-axis indicates the number of sample alleles (out of 20) carrying the derived variant. Exonic sites comprised ~10% of the genome, roughly mimicking the density of the human genome. The red solid circles give the values predicted analytically with a purely neutral model, but correcting for BGS by using the *B* values of the ancestral population (*i.e.*, pre-change in population size) obtained from simulations, in order to quantify the effective population size.

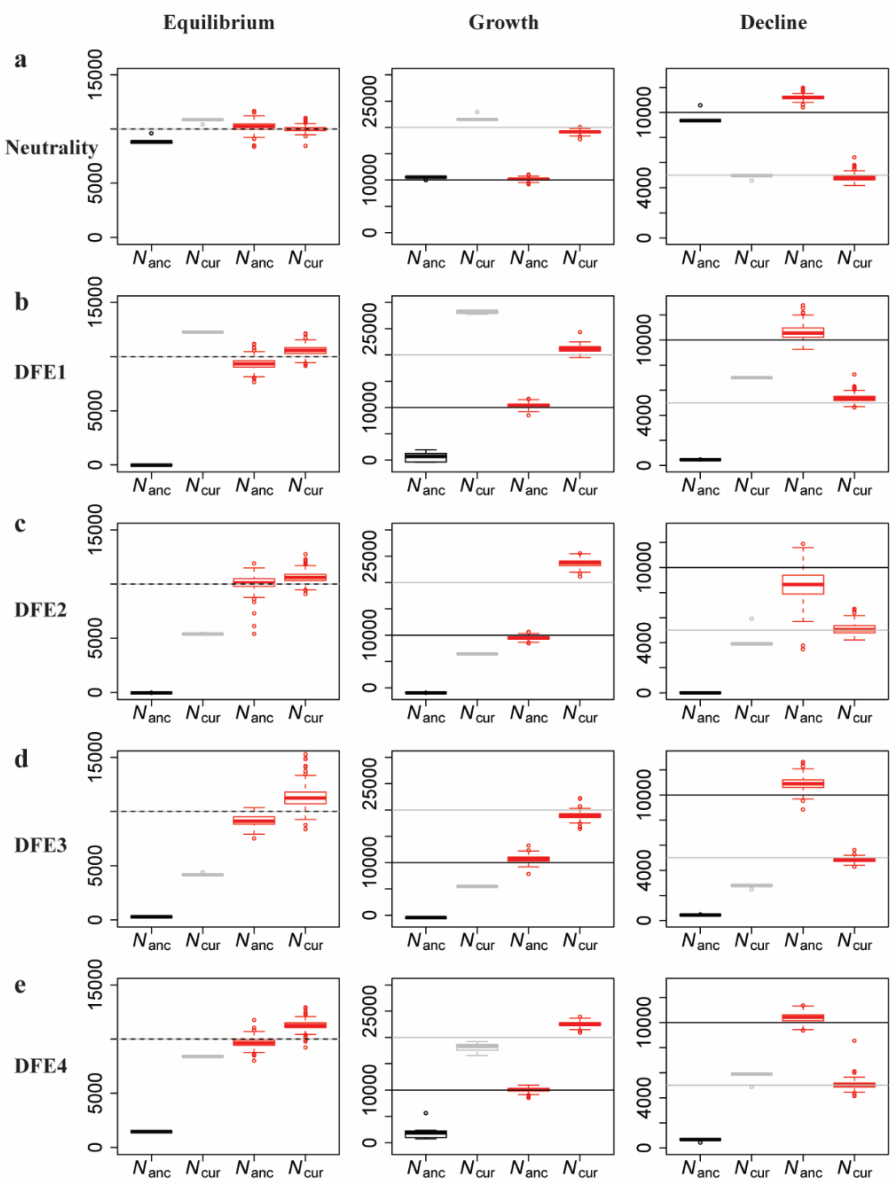


Figure 6: Comparison of estimates of ancestral (N_{anc}) and current (N_{cur}) population sizes when assuming neutrality *vs* when varying the DFE shape as a nuisance parameter, using an ABC framework. Inference is shown for demographic equilibrium (left column), 2-fold exponential growth (middle column), and 2-fold population decline (right column), for five separate DFE shapes that define the extent of direct purifying selection acting on the genomic segment for which demographic inference is performed: (a) neutrality ($f_0 = 1$, $f_1 = 0$, $f_2 = 0$, $f_3 = 0$), (b) weak purifying selection (DFE1: $f_0 = 0.1$, $f_1 = 0.7$, $f_2 = 0.1$, $f_3 = 0.1$), (c) moderately strong purifying selection (DFE2: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.7$, $f_3 = 0.1$), (d) strong purifying selection (DFE3: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.1$, $f_3 = 0.7$), and (e) a DFE in which all classes of mutations are equally frequent (DFE4: $f_0 = f_1 = f_2 = f_3 = 0.25$). In each, the horizontal lines give the true values (black for N_{anc}; and gray for N_{cur}) and the box-plots give the estimated values. Black and gray boxes represent estimates when assuming neutrality, while red boxes represent estimates when the DFE is treated as a nuisance parameter.

818 **REFERENCES**

819

820 Andolfatto P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 437:1149–
821 1152.

822 Bank C, Ewing GB, Ferrer-Admettla A, Foll M, Jensen JD. 2014. Thinking too positive?
823 Revisiting current methods of population genetic selection inference. *Trends Genet.*
824 30:540–546.

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

827 Beichman AC, Phung TN, Lohmueller KE. 2017. Comparison of single genome and allele
828 frequency data reveals discordant demographic histories. *G3* 7:3605–3620.

829 Bhaskar A, Wang YXR, Song YS. 2015. Efficient inference of population size histories and
830 locus-specific mutation rates from large-sample genomic variation data. *Genome Res.*
831 25:268–279.

832 Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner
833 JD, Rode CK, Mayhew GF, et al. 1997. The complete genome sequence of *Escherichia*
834 *coli* K-12. *Science* 277:1453–1462.

835 Booker TR, Keightley PD. 2018. Understanding the factors that shape patterns of nucleotide
836 diversity in the house mouse genome. *Mol. Biol. Evol.* 35:2971–2988.

837 Bunnefeld L, Frantz LAF, Lohse K. 2015. Inferring bottlenecks from genome-wide samples of
838 short sequence blocks. *Genetics* 201:1157–1169.

839 Campos JL, Charlesworth B. 2019. The effects on neutral variability of recurrent selective
840 sweeps and background selection. *Genetics* 212:287–303.

841 Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014. The relation between
842 recombination rate and patterns of molecular evolution and variation in *Drosophila*
843 *melanogaster*. *Mol. Biol. Evol.* 31:1010–1028.

844 Campos JL, Zhao L, Charlesworth B. 2017. Estimating the parameters of background selection
845 and selective sweeps in *Drosophila* in the presence of gene conversion. *Proc. Natl. Acad.*
846 *Sci. U.S.A.* 114:E4762–E4771.

847 Castellano D, Eyre-Walker A, Munch K. 2020. Impact of mutation rate and selection at linked
848 sites on DNA variation across the genomes of humans and other Homininae. *Genome*
849 *Biol. Evol.* 12:3550–3561.

850 Chamary J, Hurst LD. 2005. Evidence for selection on synonymous mutations affecting stability
851 of mRNA secondary structure in mammals. *Genome Biol.* 6:R75.

852 Charlesworth B. 2013. Background selection 20 years on. The Wilhelmine E. Key 2012
853 invitational lecture. *J. Hered.* 104:161–171.

854 Charlesworth B. 2015. Causes of natural variation in fitness: evidence from studies of
855 *Drosophila* populations. *Proc. Natl. Acad. Sci. U.S.A.* 112:1662–1669.

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

858 Charlesworth D, Charlesworth B, Morgan MT. 1995. The pattern of neutral molecular variation
859 under the background selection model. *Genetics* 141:1619–1632.

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

863 Choi JY, Aquadro CF. 2016. Recent and long term selection across synonymous sites in
864 *Drosophila ananassae*. *J. Mol. Evol.* 83:50–60.

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

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

869 Csilléry K, Blum MGB, Gaggiotti OE, François O. 2010. Approximate Bayesian Computation
870 (ABC) in practice. *Trends Ecol. Evol.* 25:410–418.

871 Cutter AD, Payseur BA. 2013. Genomic signatures of selection at linked sites: unifying the
872 disparity among species. *Nat. Rev. Genet.* 14:262–274.

873 Elyashiv E, Sattath S, Hu TT, Strutsovsky A, McVicker G, Andolfatto P, Coop G, Sella G. 2016.
874 A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet.*
875 12:e1006130.

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

878 Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013. Robust demographic
879 inference from genomic and SNP data. *PLoS Genet.* 9:e1003905.

880 Eyre-Walker A, Keightley PD. 2009. Estimating the rate of adaptive molecular evolution in the
881 presence of slightly deleterious mutations and population size change. *Mol. Biol. Evol.*
882 26:2097–2108.

883 Fay JC, Wu CI. 1999. A human population bottleneck can account for the discordance between
884 patterns of mitochondrial versus nuclear DNA variation. *Mol. Biol. Evol.* 16:1003–1005.

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

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

890 Fulgione A, Koornneef M, Roux F, Hermisson J, Hancock AM. 2018. Madeiran Arabidopsis
891 thaliana reveals ancient long-range colonization and clarifies demography in Eurasia.
892 *Mol. Biol. Evol.* 35:564–574.

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

896 Haddrill PR, Thornton KR, Charlesworth B, Andolfatto P. 2005. Multilocus patterns of
897 nucleotide variability and the demographic and selection history of *Drosophila*
898 *melanogaster* populations. *Genome Res.* 15:790–799.

899 Haller BC, Messer PW. 2019. SLiM 3: Forward genetic simulations beyond the Wright–Fisher
900 model. *Mol. Biol. Evol.* 36:632–637.

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

903 Hernandez RD, Kelley JL, Elyashiv E, Melton SC, Auton A, McVean G, Project 1000 Genomes,
904 Sella G, Przeworski M. 2011. Classic selective sweeps were rare in recent human
905 evolution. *Science* 331:920–924.

906 Hey J, Harris E. 1999. Population bottlenecks and patterns of human polymorphism. *Mol Biol*
907 *Evol* 16:1423–1426.

908 Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.*
909 8:269–294.

910 Hoggart CJ, Chadeau-Hyam M, Clark TG, Lampariello R, Whittaker JC, Iorio MD, Balding DJ.
911 2007. Sequence-level population simulations over large genomic regions. *Genetics*
912 177:1725–1731.

913 Hung C-M, Shaner P-JL, Zink RM, Liu W-C, Chu T-C, Huang W-S, Li S-H. 2014. Drastic
914 population fluctuations explain the rapid extinction of the passenger pigeon. *Proc. Nat.*
915 *Acad. Sci. U.S.A.* 111:10636–10641.

916 Jackson BC, Campos JL, Haddrill PR, Charlesworth B, Zeng K. 2017. Variation in the intensity
917 of selection on codon bias over time causes contrasting patterns of base composition
918 evolution in *Drosophila*. *Genome Biol. Evol.* 9:102–123.

919 Jacquier H, Birgy A, Nagard HL, Mechulam Y, Schmitt E, Glodt J, Bercot B, Petit E, Poulain J,
920 Barnaud G, et al. 2013. Capturing the mutational landscape of the beta-lactamase TEM-1.
921 *Proc. Natl. Acad. Sci. U.S.A.* 110:13067–13072.

922 Jensen JD, Payseur BA, Stephan W, Aquadro CF, Lynch M, Charlesworth D, Charlesworth B.
923 2019. The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern
924 and Hahn 2018. *Evolution* 73:111–114.

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

927 Jones MR, Good JM. 2016. Targeted capture in evolutionary and ecological genomics. *Mol.*
928 *Ecol.* 25:185–202.

929 Kaiser VB, Charlesworth B. 2009. The effects of deleterious mutations on evolution in non-
930 recombining genomes. *Trends Genet.* 25:9–12.

931 Keightley PD, Eyre-Walker A. 2007. Joint inference of the distribution of fitness effects of
932 deleterious mutations and population demography based on nucleotide polymorphism
933 frequencies. *Genetics* 177:2251–2261.

934 Keightley PD, Jackson BC. 2018. Inferring the probability of the derived vs. the ancestral allelic
935 state at a polymorphic site. *Genetics* 209:897–906.

936 Keightley PD, Ness RW, Halligan DL, Haddrill PR. 2014. Estimation of the spontaneous
937 mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics*
938 196:313–320.

939 Keightley PD, Trivedi U, Thomson M, Oliver F, Kumar S, Blaxter ML. 2009. Analysis of the
940 genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation
941 lines. *Genome Res.* 19:1195–1201.

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

944 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler and D. 2002.
945 The Human Genome Browser at UCSC. *Genome Res.* 12:996–1006.

946 Kim BY, Huber CD, Lohmueller KE. 2017. Inference of the distribution of selection coefficients
947 for new nonsynonymous mutations using large samples. *Genetics* 206:345–361.

948 Kim Y, Wiehe T. 2009. Simulation of DNA sequence evolution under models of recent
949 directional selection. *Brief. Bioinformatics* 10:84–96.

950 Kousathanas A, Keightley PD. 2013. A comparison of models to infer the distribution of fitness
951 effects of new mutations. *Genetics* 193:1197–1208.

952 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle
953 M, FitzHugh W, et al. 2001. Initial sequencing and analysis of the human genome.
954 *Nature* 409:860–921.

955 Lapierre M, Lambert A, Achaz G. 2017. Accuracy of demographic inferences from the site
956 frequency spectrum: the case of the Yoruba population. *Genetics* 206:439–449.

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

959 Liang P, Saqib HSA, Zhang X, Zhang L, Tang H. 2018. Single-base resolution map of
960 evolutionary constraints and annotation of conserved elements across major grass
961 genomes. *Genome Biol. Evol.* 10:473–488.

962 Lukic S, Hey J. 2012. Demographic inference using spectral methods on SNP data, with an
963 analysis of the human out-of-Africa expansion. *Genetics* 192:619–639.

964 Lynch M. 2007. The Origins of Genome Architecture. Sunderland, Massachusetts: Sinauer
965 Associates

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

969 McVean GAT, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P. 2004. The fine-scale
970 structure of recombination rate variation in the human genome. *Science* 304:581–584.

971 Messer PW, Petrov DA. 2013. Frequent adaptation and the McDonald–Kreitman test. *Proc. Natl.*
972 *Acad. Sci. U.S.A.* 110:8615–8620.

973 Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of
974 recombination rates and hotspots across the human genome. *Science* 310:321–324.

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

977 O'Fallon BD, Seger J, Adler FR. 2010. A continuous-state coalescent and the impact of weak
978 selection on the structure of gene genealogies. *Mol. Biol. Evol.* 27:1162–1172.

979 Orozco-terWengel P. 2016. The devil is in the details: the effect of population structure on
980 demographic inference. *Heredity* 116:349–350.

981 Palkopoulou E, Lipson M, Mallick S, Nielsen S, Rohland N, Baleka S, Karpinski E, Ivancevic
982 AM, To T-H, Kortschak RD, et al. 2018. A comprehensive genomic history of extinct
983 and living elephants. *Proc. Nat. Acad. Sci. U.S.A.* 115:E2566–E2574.

984 Polanski A, Bobrowski A, Kimmel M. 2003. A note on distributions of times to coalescence,
985 under time-dependent population size. *Theor Popul Biol* 63:33–40.

986 Polanski A, Kimmel M. 2003. New explicit expressions for relative frequencies of single-
987 nucleotide polymorphisms with application to statistical inference on population growth.
988 *Genetics* 165:427–436.

989 Pool JE, Nielsen R. 2007. Population size changes reshape genomic patterns of diversity.
990 *Evolution* 61:3001–3006.

991 Pool JE, Nielsen R. 2009. Correction for Pool and Nielsen (2007). *Evolution* 63:1671.

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

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

997 Rosen Z, Bhaskar A, Roch S, Song YS. 2018. Geometry of the sample frequency spectrum and
998 the perils of demographic inference. *Genetics* 210:665–682.

999 Sanjuán R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common
1000 patterns revealed by site-directed mutagenesis studies. *Phil. Trans. R. Soc. B.* 365:1975–
1001 1982.

1002 Schiffels S, Durbin R. 2014. Inferring human population size and separation history from
1003 multiple genome sequences. *Nat. Genet.* 46:919–925.

1004 Schneider A, Charlesworth B, Eyre-Walker A, Keightley PD. 2011. A method for inferring the
1005 rate of occurrence and fitness effects of advantageous mutations. *Genetics* 189:1427–
1006 1437.

1007 Schrider DR, Shanku AG, Kern AD. 2016. Effects of linked selective sweeps on demographic
1008 inference and model selection. *Genetics* 204:1207–1223.

1009 Sheehan S, Song YS. 2016. Deep learning for population genetic inference. *PLoS Comput. Biol.*
1010 12:e1004845.

1011 Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J,
1012 Hillier LW, Richards S, et al. 2005. Evolutionarily conserved elements in vertebrate,
1013 insect, worm, and yeast genomes. *Genome Res.* 15:1034–1050.

1014 Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable
1015 and exponentially growing populations. *Genetics* 129:555–562.

1016 Speidel L, Forest M, Shi S, Myers SR. 2019. A method for genome-wide genealogy estimation
1017 for thousands of samples. *Nat. Genet.* 51:1321–1329.

1018 Steinrücken M, Kamm J, Spence JP, Song YS. 2019. Inference of complex population histories
1019 using whole-genome sequences from multiple populations. *Proc. Natl. Acad. Sci. U.S.A.*
1020 116:17115–17120.

1021 Teshima KM, Coop G, Przeworski M. 2006. How reliable are empirical genomic scans for
1022 selective sweeps? *Genome Res* 16:702–712.

1023 Thornton K. 2003. Libsequence: a C++ class library for evolutionary genetic analysis.
1024 *Bioinformatics* 19:2325–2327.

1025 Thornton KR, Jensen JD. 2007. Controlling the false-positive rate in multilocus genome scans
1026 for selection. *Genetics* 175:737–750.

1027 Torres R, Stetter MG, Hernandez RD, Ross-Ibarra J. 2020. The temporal dynamics of
1028 background selection in nonequilibrium populations. *Genetics* 214:1019–1030.

1029 Torres R, Szpiech ZA, Hernandez RD. 2018. Human demographic history has amplified the
1030 effects of background selection across the genome. *PLoS Genet.* 14:e1007387.

1031 Uricchio LH, Hernandez RD. 2014. Robust forward simulations of recurrent hitchhiking.
1032 *Genetics* 197:221–236.

1033 Warren WC, Jasinska AJ, García-Pérez R, Svardal H, Tomlinson C, Rocchi M, Archidiacono N,
1034 Capozzi O, Minx P, Montague MJ, et al. 2015. The genome of the vervet (*Chlorocebus*
1035 *aethiops sabaeus*). *Genome Res.* 25:1921–1933.

1036 Williamson RJ, Josephs EB, Platts AE, Hazzouri KM, Haudry A, Blanchette M, Wright SI.
1037 2014. Evidence for widespread positive and negative selection in coding and conserved
1038 noncoding regions of *Capsella grandiflora*. *PLoS Genet.* 10:e1004622.

1039 Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y, et al.
1040 2020. A new coronavirus associated with human respiratory disease in China. *Nature*
1041 579:265–269.

1042 Zeng K, Charlesworth B. 2010. Studying patterns of recent evolution at synonymous sites and
1043 intronic sites in *Drosophila melanogaster*. *J. Mol. Evol.* 70:116–128.

1044 Zhou Y, Massonnet M, Sanjak JS, Cantu D, Gaut BS. 2017. Evolutionary genomics of grape
1045 (*Vitis vinifera ssp. vinifera*) domestication. *Proc. Nat. Acad. Sci. U.S.A.* 114:11715–
1046 11720.

1047
1048
1049

Supplementary Information

Supp Table 1: Parameters underlying the human-like demographic models considered.

	Demographic models	**Ancestral population size**	**Current population size**	**Time of change in generations**
1	Equilibrium	10,000	10,000	NA
2	Exponential growth	1000	30,000	850
3	Instantaneous decline	12,300	2,100	4,750

Supp Table 2: The site frequency spectrum (SFS) obtained from 50 diploid individuals from the simulated data (which was provided to *fastsimcoal2* for inference) and the SFS corresponding to the best model inferred by *fastsimcoal2*, in the case of 30-fold exponential growth under neutrality.

	Fastsimcoal2 best model		Simulated data
Frequency of SNP (out of 100)	Mean	SD	Mean
1	0.529851	0.007443	0.533701
2	0.144710	0.005779	0.135345
3	0.062517	0.005029	0.059647
4	0.034133	0.003670	0.034636
5	0.021766	0.002589	0.023198
6	0.015518	0.001808	0.017067
7	0.011996	0.001255	0.013212
8	0.009808	0.000862	0.010737
9	0.008347	0.000576	0.009145
10	0.007304	0.000369	0.007903
11	0.006519	0.000214	0.006843
12	0.005909	0.000105	0.006222
13	0.005408	0.000039	0.005552
14	0.004991	0.000064	0.005050
15	0.004640	0.000109	0.004844
16	0.004329	0.000141	0.004410
17	0.004065	0.000172	0.004148
18	0.003838	0.000190	0.003860
19	0.003627	0.000209	0.003663
20	0.003441	0.000221	0.003489
21	0.003273	0.000230	0.003236
22	0.003122	0.000236	0.003155
23	0.002984	0.000240	0.003042
24	0.002854	0.000242	0.002901
25	0.002736	0.000244	0.002807
26	0.002625	0.000243	0.002701
27	0.002536	0.000245	0.002560
28	0.002437	0.000244	0.002451
29	0.002356	0.000244	0.002338
30	0.002273	0.000241	0.002308
31	0.002199	0.000239	0.002150
32	0.002127	0.000237	0.002133
33	0.002067	0.000236	0.002146
34	0.002004	0.000233	0.002032
35	0.001944	0.000230	0.001900

36	0.001891	0.000227	0.001876
37	0.001839	0.000226	0.001894
38	0.001790	0.000221	0.001761
39	0.001741	0.000219	0.001758
40	0.001698	0.000216	0.001709
41	0.001657	0.000213	0.001663
42	0.001616	0.000212	0.001645
43	0.001580	0.000210	0.001583
44	0.001543	0.000205	0.001493
45	0.001508	0.000203	0.001524
46	0.001475	0.000201	0.001444
47	0.001439	0.000197	0.001480
48	0.001414	0.000196	0.001398
49	0.001386	0.000193	0.001367
50	0.001353	0.000191	0.001323
51	0.001330	0.000189	0.001380
52	0.001302	0.000186	0.001284
53	0.001274	0.000182	0.001237
54	0.001254	0.000181	0.001213
55	0.001232	0.000178	0.001239
56	0.001210	0.000177	0.001180
57	0.001189	0.000175	0.001181
58	0.001167	0.000172	0.001176
59	0.001147	0.000170	0.001131
60	0.001127	0.000169	0.001135
61	0.001108	0.000167	0.001097
62	0.001092	0.000165	0.001115
63	0.001075	0.000163	0.001135
64	0.001057	0.000161	0.000993
65	0.001039	0.000160	0.001078
66	0.001027	0.000158	0.001054
67	0.001011	0.000156	0.001010
68	0.000991	0.000154	0.000995
69	0.000978	0.000152	0.000986
70	0.000964	0.000150	0.000946
71	0.000952	0.000150	0.000927
72	0.000937	0.000147	0.000922
73	0.000926	0.000146	0.000966
74	0.000910	0.000144	0.000898
75	0.000897	0.000143	0.000916
76	0.000888	0.000141	0.000910
77	0.000877	0.000140	0.000844
78	0.000866	0.000139	0.000866

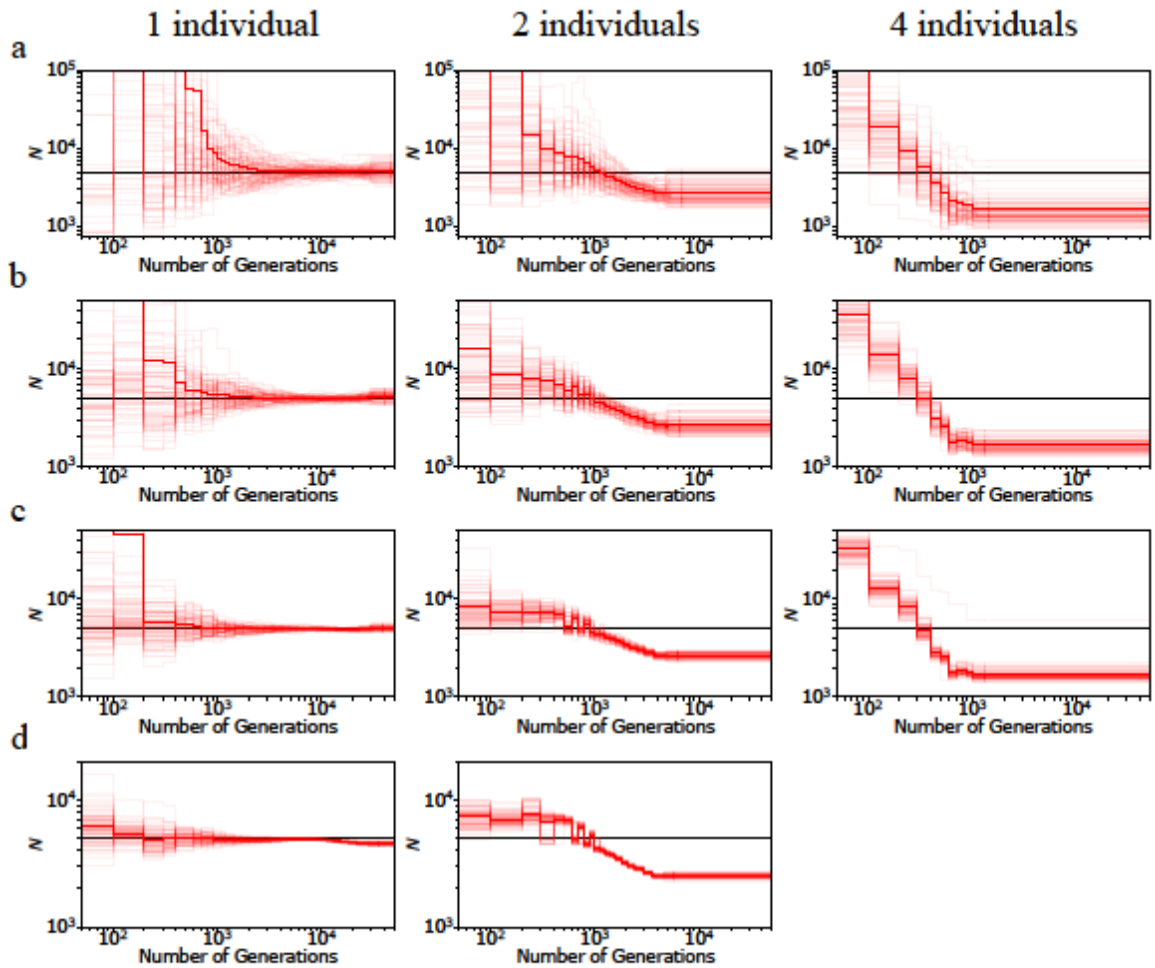
79	0.000853	0.000137	0.000869
80	0.000845	0.000136	0.000810
81	0.000831	0.000135	0.000819
82	0.000824	0.000134	0.000792
83	0.000812	0.000132	0.000753
84	0.000801	0.000131	0.000735
85	0.000793	0.000130	0.000803
86	0.000783	0.000128	0.000790
87	0.000775	0.000128	0.000767
88	0.000768	0.000127	0.000754
89	0.000757	0.000125	0.000771
90	0.000749	0.000124	0.000776
91	0.000742	0.000124	0.000737
92	0.000734	0.000123	0.000713
93	0.000727	0.000122	0.000724
94	0.000717	0.000120	0.000724
95	0.000709	0.000119	0.000708
96	0.000702	0.000118	0.000687
97	0.000696	0.000117	0.000674
98	0.000688	0.000116	0.000653
99	0.000682	0.000116	0.000679

Supp Table 3: Inference of ancestral population size (Nanc), current population size (Ncur), and time of change, using instantaneous and exponential size change models in *fastsimcoal2*.

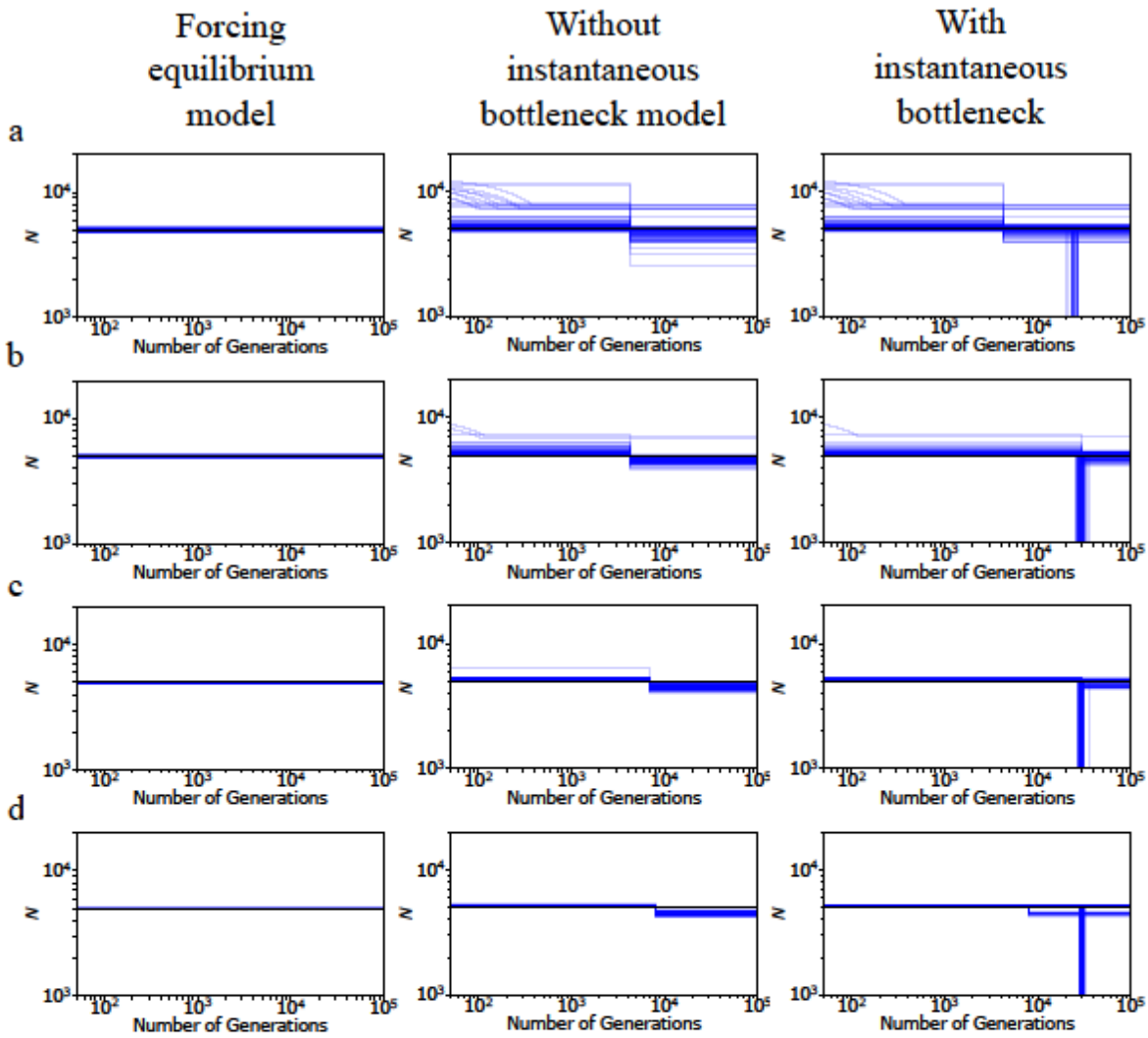
	Exponential Model			Instantaneous Model			True Values		
Fold Change	Ncur	Nanc	Time	Ncur	Nanc	Time	Ncur	Nanc	Time
2x	4414	2134	222	1844	964	485	2000	1000	850
5x	8057	2883	425	3990	1005	488	5000	1000	850
10x	11569	3425	530	6704	1020	522	10000	1000	850
20x	16670	4104	585	11018	1037	544	20000	1000	850
30X	21149	4624	611	14272	1033	567	30000	1000	850

Supp Table 4: Nucleotide diversity in the presence of BGS relative to that under neutrality (B), calculated for a neutral site distance y bases from the end of a gene/exon of length 500 bp. The exon experiences purifying selection with strength $2N_e s=10$, where N_e is the effective population size and s is the reduction in fitness. Shown below is $t=hs$ where h is the dominance coefficient (assumed to be 0.5 here), r is the recombination rate per site per generation, and u is the mutation rate per site per generation. B was calculated using Equation 2 of Johri *et al*. (2020).

N_e	$t\,(2N_e s=10)$	r	u	$B\,(y=1)$	$B\,(y=10)$	$B\,(y=1000)$
10^4	0.00025	1.00×10^{-8}	1.00×10^{-8}	0.9805	0.9806	0.9820
		1.00×10^{-6}	1.00×10^{-6}	0.5152	0.5312	0.9444
10^6	2.5×10^{-6}	1.00×10^{-8}	1.00×10^{-8}	0.5152	0.5312	0.9445
		1.00×10^{-6}	1.00×10^{-6}	0.4920	0.8227	0.9992



Supp Figure 1: Performance of MSMC under neutrality and demographic equilibrium when using 1, 2, and 4 diploid individuals for inference, for varying chromosome sizes: (a) 10 Mb, (b) 50 Mb, (c) 200 Mb, (d) 1 Gb. The numbers of replicates for each panel varied slightly around 100. MSMC runs with 4 diploid individuals could not be obtained due to the long computational times required. As can be seen, MSMC has a tendency to take a common shape, often falsely indicating recent population growth.



Supp Figure 2: Performance of *fastsimcoal2* under neutrality and demographic equilibrium with varying chromosome sizes: (a) 10 Mb, (b) 50 Mb, (c) 200 Mb, (d) 1 Gb, when different model choices are provided: *Left column*: the correct model of constant population size is specified; *Middle column*: model selection was performed with 3 models – equilibrium, instantaneous size change, and exponential size change; *Right column*: model selection was performed with 4 models – equilibrium, instantaneous size change, exponential size change, and instantaneous bottleneck. Inference was performed using 50 diploid individuals and the inferred population size estimates of the best model are plotted (blue lines). The numbers of replicates for each panel varied slightly around 100. The true model is shown in black.

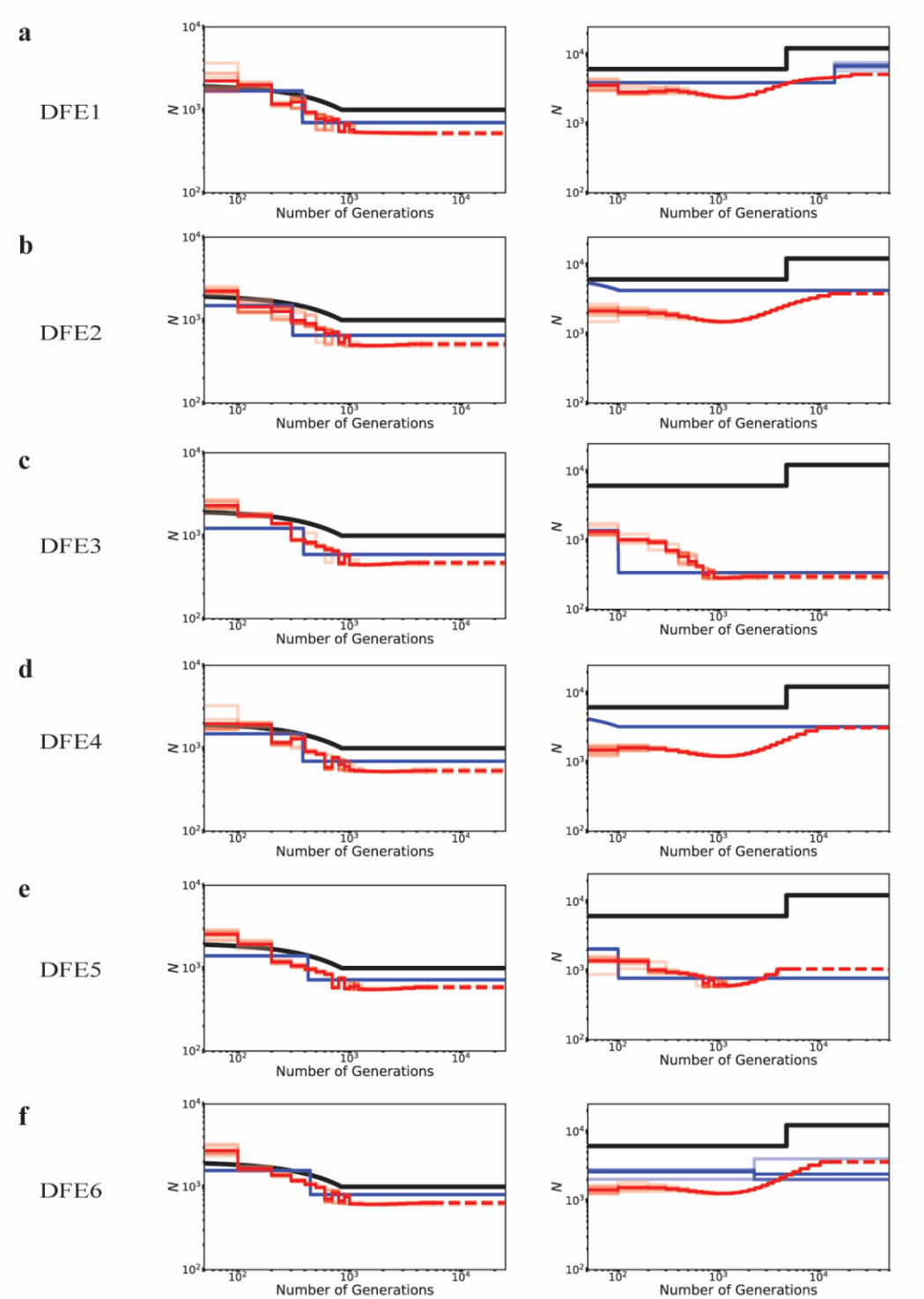
a

Size \ Test Model	Equilibrium	Instantaneous Change	Exponential Change
10 Mb	38	56	6
50 Mb	34	62	4
200 Mb	24	72	4
1 Gb	19	75	6

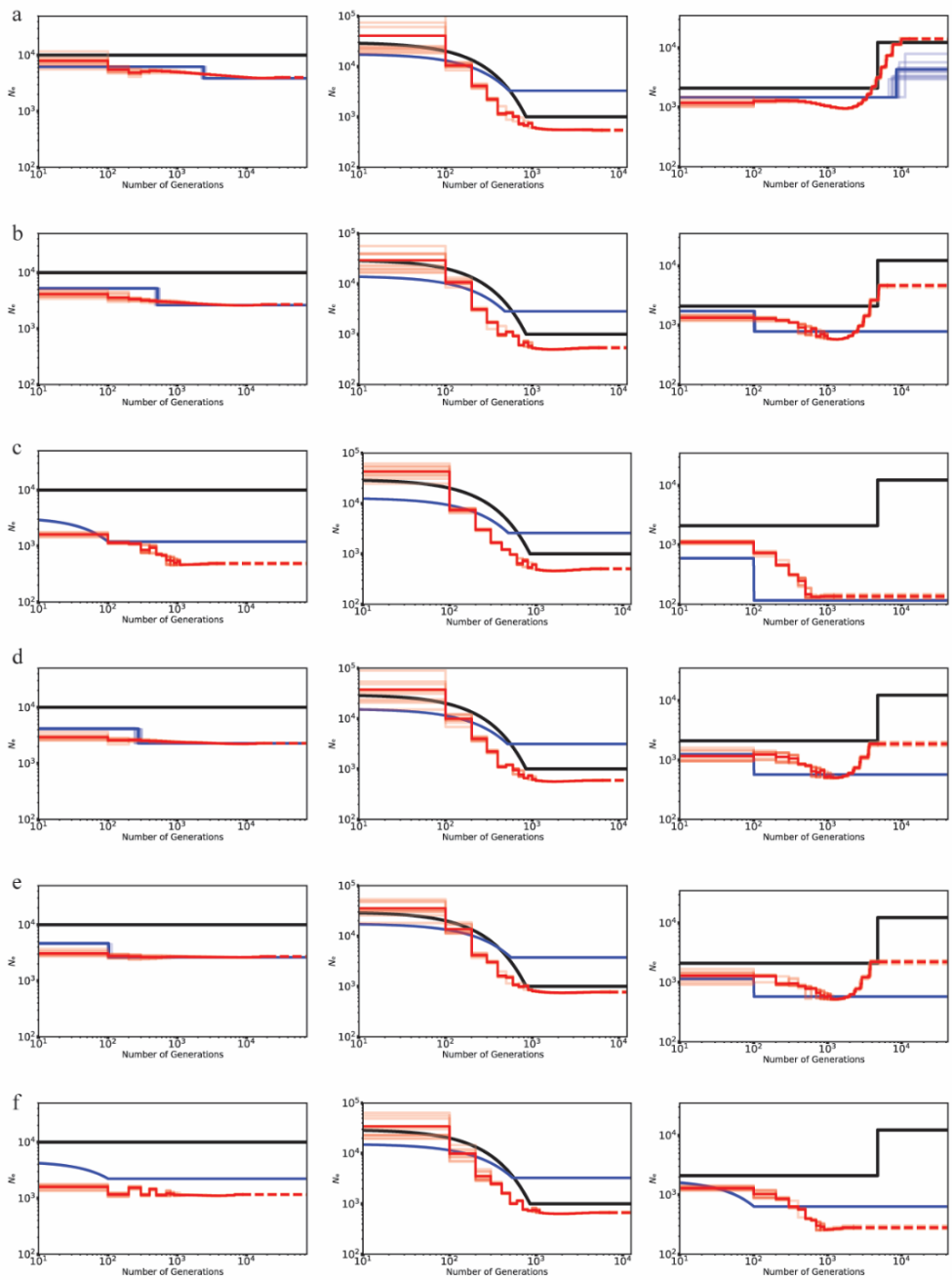
b

Size \ Test Model	Equilibrium	Instantaneous Change	Exponential Change	Instant Bottleneck
10 Mb	33	46	5	16
50 Mb	29	48	2	21
200 Mb	19	51	2	28
1 Gb	11	47	5	37

Supp Figure 3: Model selection for 100 replicates (corresponding to Supp Figure 2) in which the true model was neutral equilibrium. (a) Model selection was performed with 3 models: equilibrium, instantaneous size change, and exponential size change; (b) Model selection was performed with 4 models: equilibrium, instantaneous size change, exponential size change, and instantaneous bottleneck.



Supp Figure 4: Inferred demography from MSMC (red lines) and *fastsimcoal2* (blue lines) in the presence of background selection, with the true DFE shown to the left of the panel, for 2-fold instantaneous decline (right column) and 2-fold exponential growth (left column). In this case, 20% of the genome experienced direct selection. The true demographic models are shown as black lines.



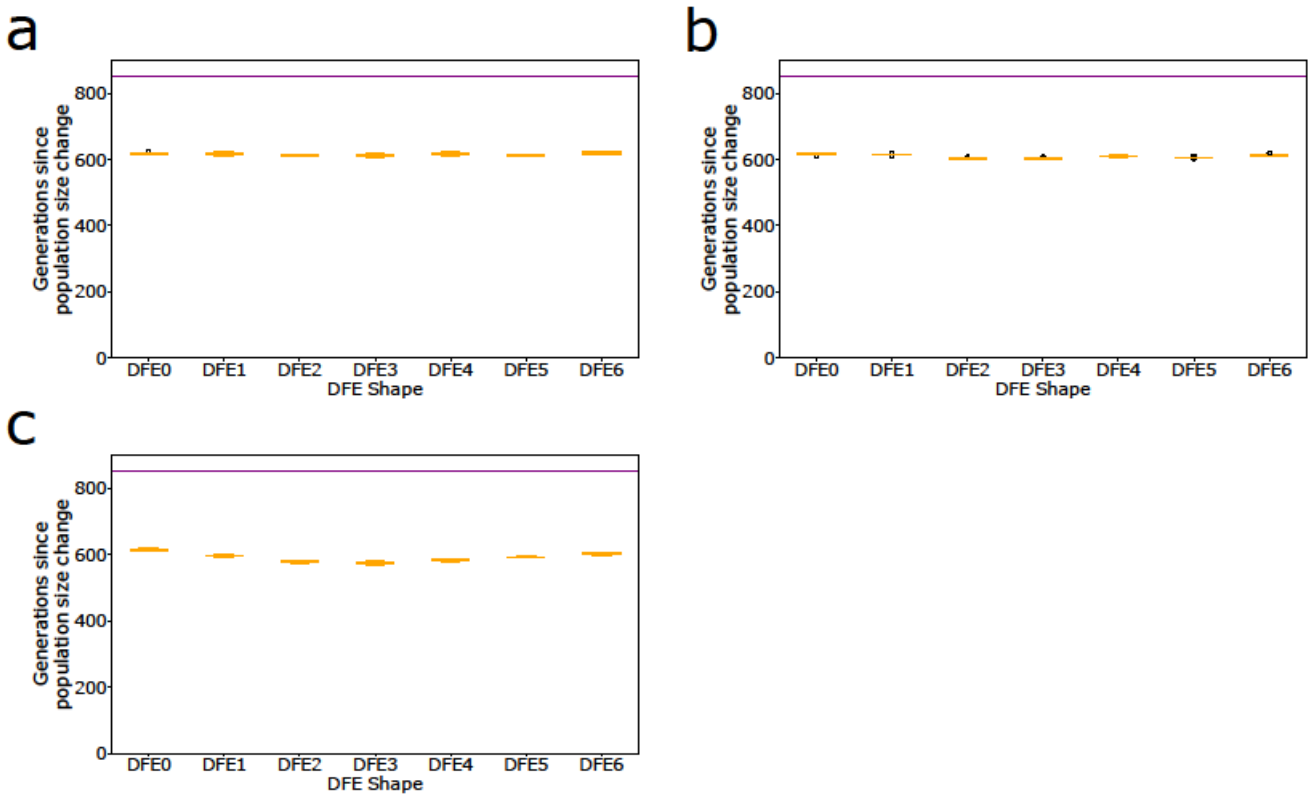
Supp Figure 5: Inference of demography by MSMC (red lines; 10 replicates) and *fastsimcoal2* (blue lines; 10 replicates) under demographic equilibrium (left column), 30-fold exponential growth (middle column), and ~6-fold instantaneous decline (right column) in the presence of direct purifying selection (*i.e.*, directly selected sites are not masked). The true demographic model is depicted in black lines. Exonic sites experience purifying selection specified by the following DFEs: (a) DFE1: $f_0 = 0.1$, $f_1 = 0.7$, $f_2 = 0.1$, $f_3 = 0.1$, (b) DFE2: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.7$, $f_3 = 0.1$, (c) DFE3: $f_0 = 0.1$, $f_1 = 0.1$, $f_2 = 0.1$, $f_3 = 0.7$, (d) DFE4: $f_0 = 0.25$, $f_1 = 0.25$, $f_2 = 0.25$, $f_3 = 0.25$, (e) DFE5: $f_0 = 0.7$, $f_1 = 0.0$, $f_2 = 0.0$, $f_3 = 0.3$, (f) DFE6: $f_0 = 0.5$, $f_1 = 0.0$, $f_2 = 0.0$, $f_3 = 0.5$. In this case, 20% of the genome was under selection.

	Functional Density	5%				10%				20%			
True Model	Fast Model / DFE Shape	Equilibrium	Instantaneous Change	Exponential Change	Instant Bottleneck	Equilibrium	Instantaneous Change	Exponential Change	Instant Bottleneck	Equilibrium	Instantaneous change	Exponential Change	Instant Bottleneck
Equilibrium	DFE0	0	10	0	0	0	10	0	0	0	10	0	0
	DFE1	0	7	3	0	0	8	2	0	0	10	0	0
	DFE2	0	9	1	0	0	10	0	0	0	10	0	0
	DFE3	0	9	1	0	0	10	0	0	0	0	10	0
	DFE4	0	10	0	0	0	9	1	0	0	10	0	0
	DFE5	0	10	0	0	0	10	0	0	0	0	10	0
	DFE6	0	7	3	0	0	10	0	0	0	9	1	0
Instantaneous Decline (6x)	DFE0	0	10	0	0	0	10	0	0	0	10	0	0
	DFE1	0	10	0	0	0	10	0	0	0	10	0	0
	DFE2	0	10	0	0	0	10	0	0	0	10	0	0
	DFE3	0	10	0	0	0	10	0	0	0	9	1	0
	DFE4	0	10	0	0	0	10	0	0	0	10	0	0
	DFE5	0	10	0	0	0	10	0	0	0	10	0	0
	DFE6	0	10	0	0	0	10	0	0	0	10	0	0
Exponential Growth (30x)	DFE0	0	0	10	0	0	0	10	0	0	0	10	0
	DFE1	0	0	10	0	0	0	10	0	0	0	10	0
	DFE2	0	0	10	0	0	0	10	0	0	0	10	0
	DFE3	0	0	10	0	0	0	10	0	0	0	10	0
	DFE4	0	0	10	0	0	0	10	0	0	0	10	0
	DFE5	0	0	10	0	0	0	10	0	0	0	10	0
	DFE6	0	0	10	0	0	0	10	0	0	0	10	0
Instantaneous Decline (2x)	DFE0	0	10	0	0	0	10	0	0	0	10	0	0
	DFE1	0	10	0	0	0	10	0	0	0	10	0	0
	DFE2	0	10	0	0	0	10	0	0	0	4	6	0
	DFE3	0	10	0	0	0	10	0	0	0	5	5	0
	DFE4	0	10	0	0	0	10	0	0	0	7	3	0
	DFE5	0	10	0	0	0	10	0	0	0	10	0	0
	DFE6	0	10	0	0	0	10	0	0	0	8	2	0
Exponential Growth (2x)	DFE0	0	10	0	0	0	10	0	0	0	10	0	0
	DFE1	0	10	0	0	0	10	0	0	0	10	0	0
	DFE2	0	10	0	0	0	10	0	0	0	10	0	0
	DFE3	0	10	0	0	0	10	0	0	0	10	0	0
	DFE4	0	10	0	0	0	10	0	0	0	10	0	0
	DFE5	0	10	0	0	0	10	0	0	0	10	0	0
	DFE6	0	10	0	0	0	10	0	0	0	10	0	0

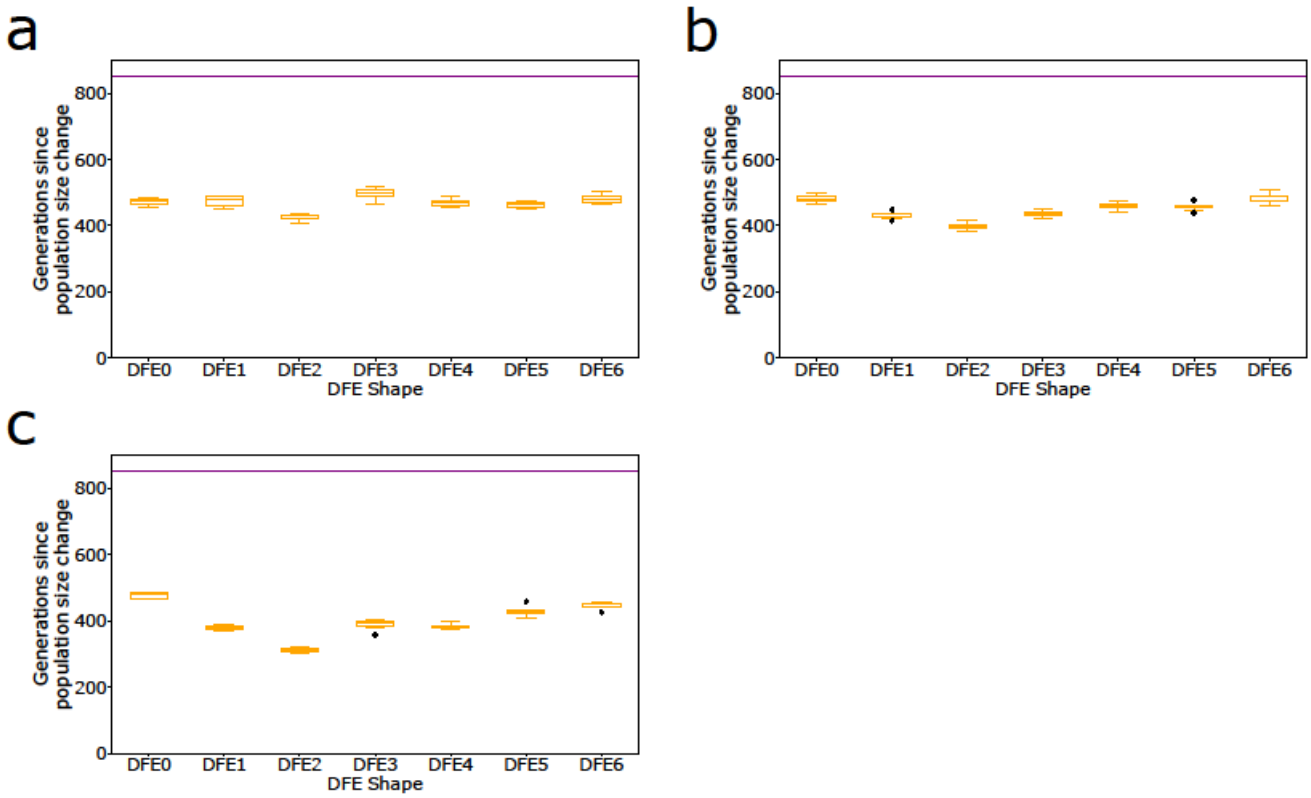
Supp Figure 6: Model selection by *fastsimcoal2* in the presence of BGS, when chosen from four possible models: equilibrium, instantaneous size change, exponential size change, and instantaneous bottleneck. Results are shown when SNPs separated at a distance of 5 kb were used for inference, and when 5%, 10% or 20% of the genomes were exonic.

	Functional Density	5%		10%		20%	
True Model	Behavior / DFE Shape	Growth	Decline	Growth	Decline	Growth	Decline
Equilibrium	DFE0	10	0	10	0	10	0
	DFE1	10	0	10	0	10	0
	DFE2	10	0	10	0	10	0
	DFE3	10	0	10	0	10	0
	DFE4	10	0	10	0	10	0
	DFE5	10	0	10	0	10	0
	DFE6	10	0	10	0	10	0
Instantaneous Decline (6x)	DFE0	0	10	0	10	0	10
	DFE1	0	10	0	10	0	10
	DFE2	0	10	0	10	10	0
	DFE3	0	10	10	0	10	0
	DFE4	0	10	0	10	10	0
	DFE5	0	10	10	0	10	0
	DFE6	0	10	0	10	10	0
Exponential Growth (30x)	DFE0	10	0	10	0	10	0
	DFE1	10	0	10	0	10	0
	DFE2	10	0	10	0	10	0
	DFE3	10	0	10	0	10	0
	DFE4	10	0	10	0	10	0
	DFE5	10	0	10	0	10	0
	DFE6	10	0	10	0	10	0
Instantaneous Decline (2x)	DFE0	0	10	0	10	0	10
	DFE1	0	10	0	10	0	10
	DFE2	0	10	0	10	10	0
	DFE3	0	10	10	0	10	0
	DFE4	0	10	0	10	10	0
	DFE5	0	10	0	10	10	0
	DFE6	0	10	0	10	10	0
Exponential Growth (2x)	DFE0	10	0	10	0	10	0
	DFE1	10	0	10	0	10	0
	DFE2	10	0	10	0	10	0
	DFE3	10	0	10	0	10	0
	DFE4	10	0	10	0	10	0
	DFE5	10	0	10	0	10	0
	DFE6	10	0	10	0	10	0

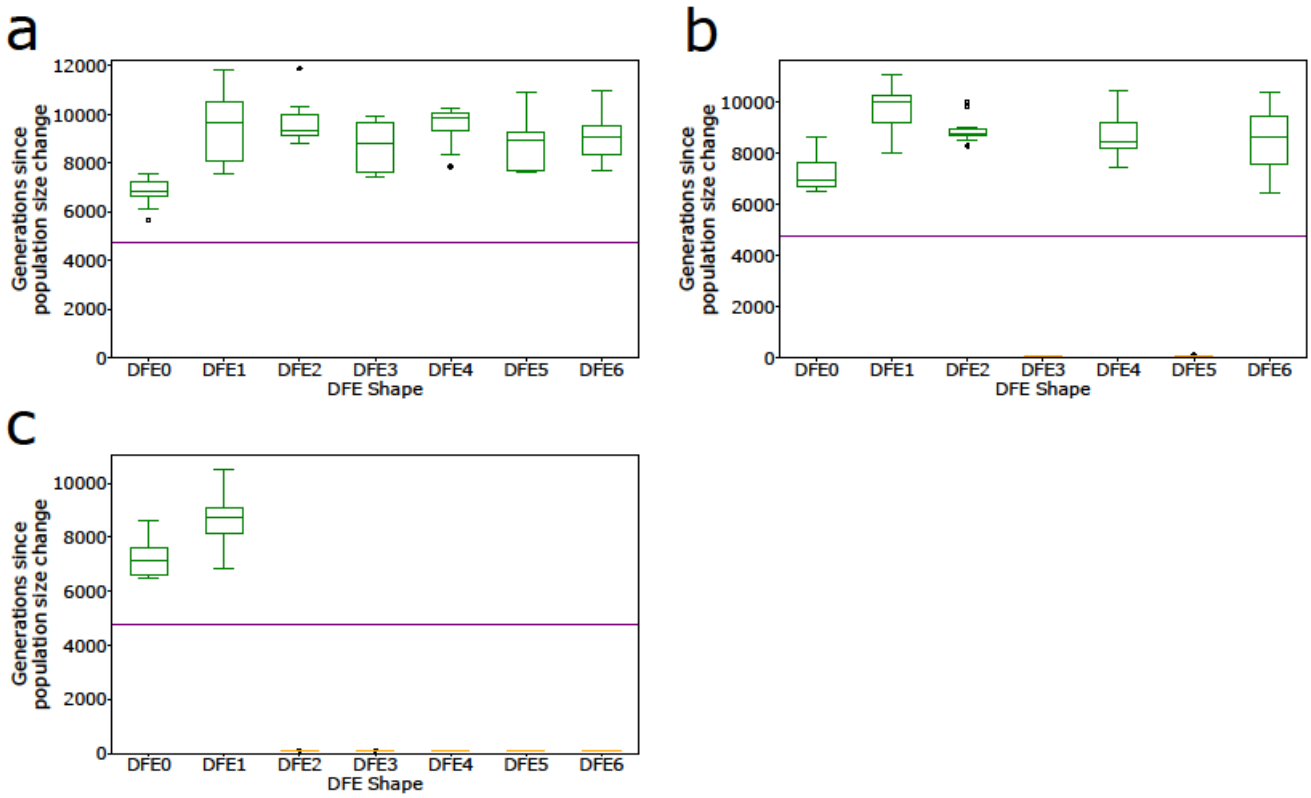
Supp Figure 7: Effects of BGS on inference of growth or decline by *fastsimcoal2*. Results are shown when SNPs separated by a distance of 5 kb were used for inference, and when 5%, 10% and 20% of the genomes were exonic.



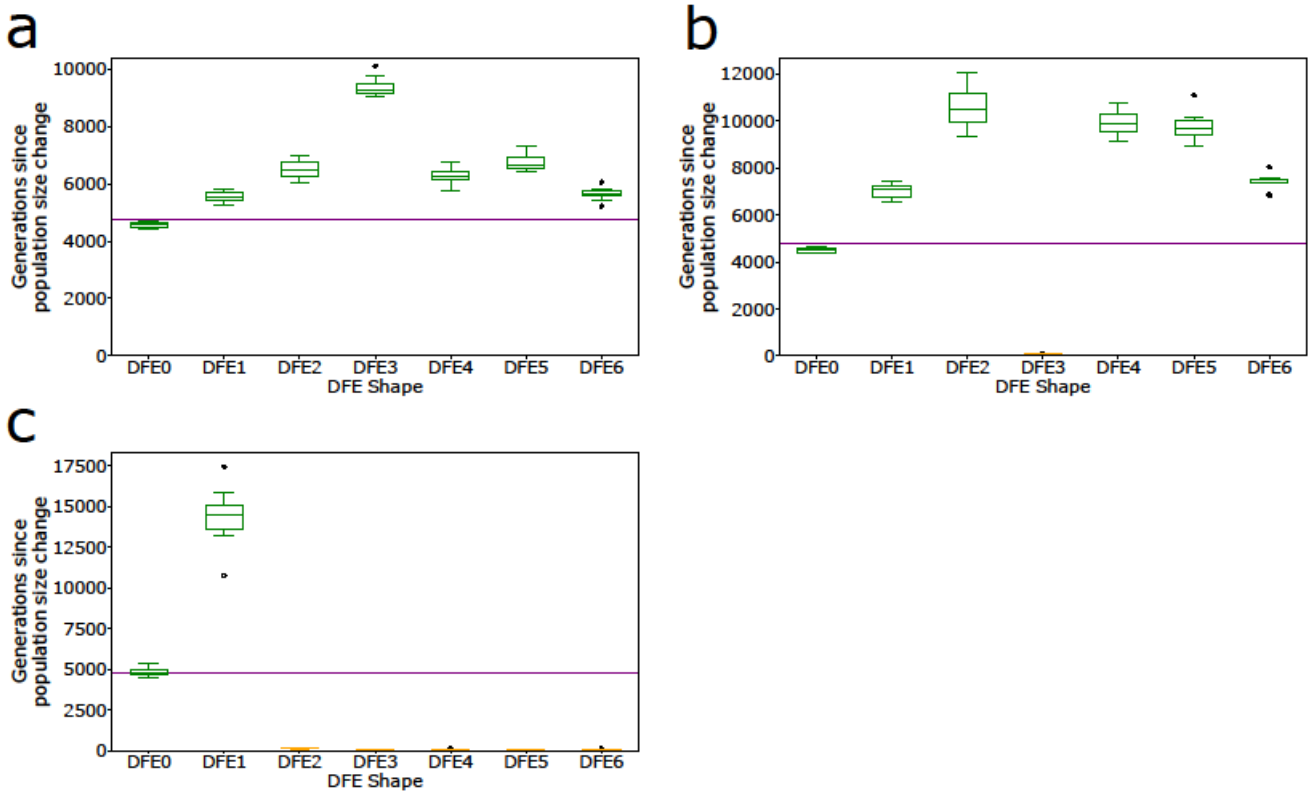
Supp Figure 8: Inference of the timing of 30-fold exponential growth by *fastsimcoal2* in the presence of BGS, when (a) 5%, (b) 10%, and (c) 20% of the genome is exonic (compared to neutrality (DFE0)). The purple line represents the true timing.



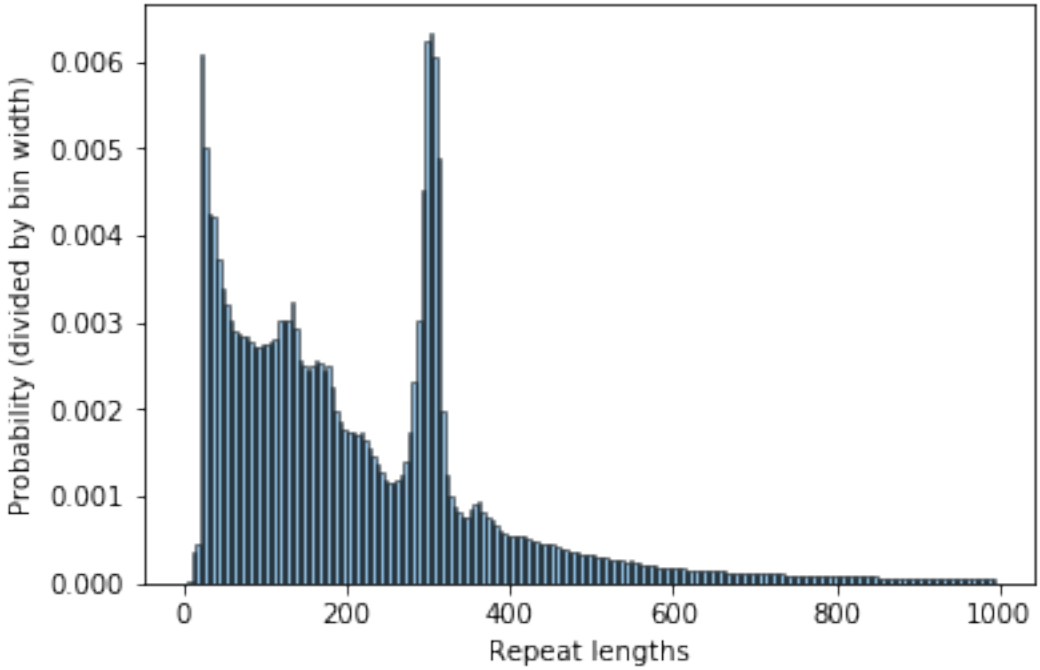
Supp Figure 9: Inference of the timing of 2-fold exponential growth by *fastsimcoal2* in the presence of BGS, when (a) 5%, (b) 10%, and (c) 20% of the genome is exonic (compared to neutrality (DFE0)). The purple line represents the true timing.



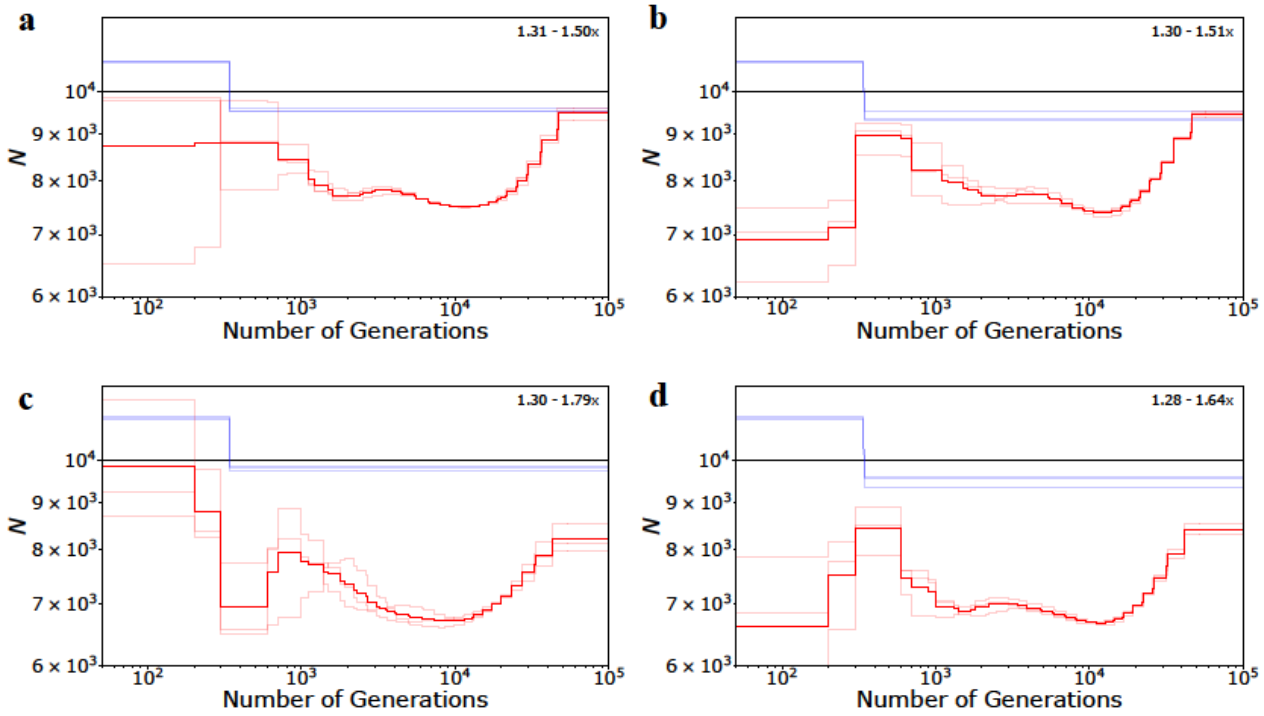
Supp Figure 10: Inference of the timing of 6-fold instantaneous decline by *fastsimcoal2* in the presence of BGS, when (a) 5%, (b) 10%, and (c) 20% of the genome is exonic (compared to neutrality (DFE0)). The purple line represents the true timing. Boxplots are presented in green if decline was inferred, and in yellow if growth was inferred.



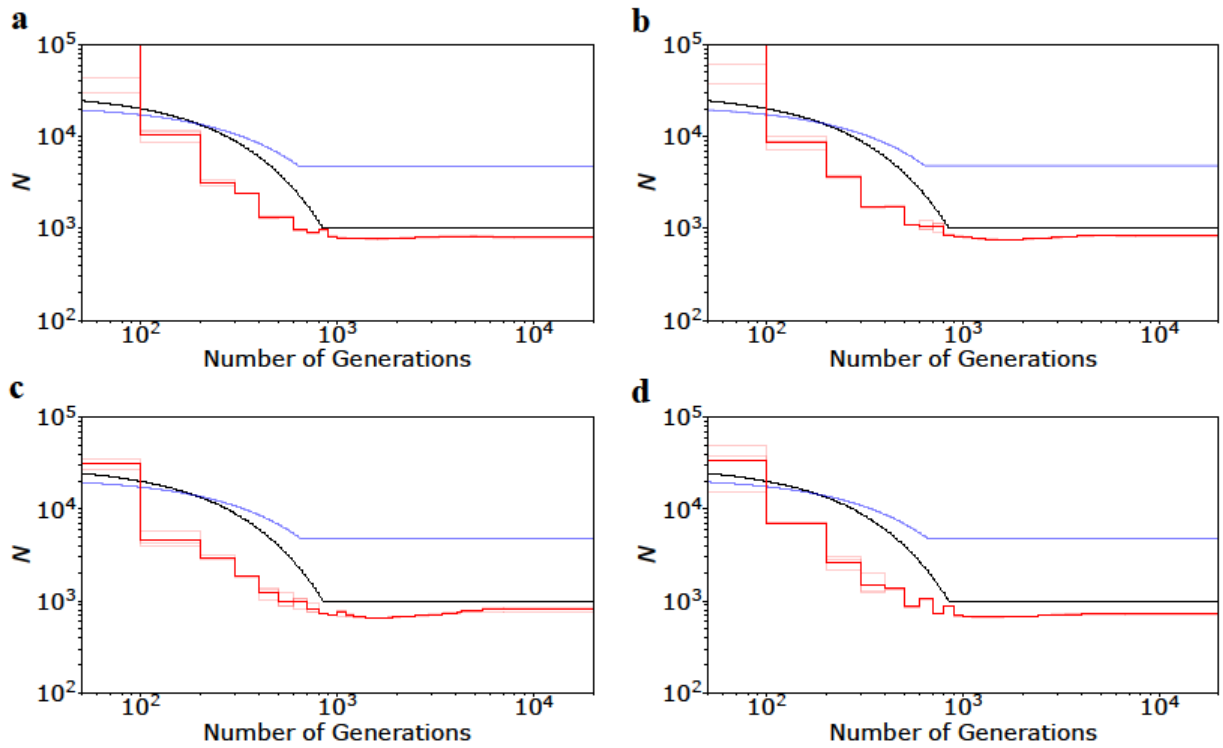
Supp Figure 11: Inference of the timing of 2-fold instantaneous decline by *fastsimcoal2* in the presence of BGS, when (a) 5%, (b) 10%, and (c) 20% of the genome is exonic (compared to neutrality (DFE0)). The purple line represents the true timing. Boxplots are presented in green if decline was inferred, and in yellow if growth was inferred.



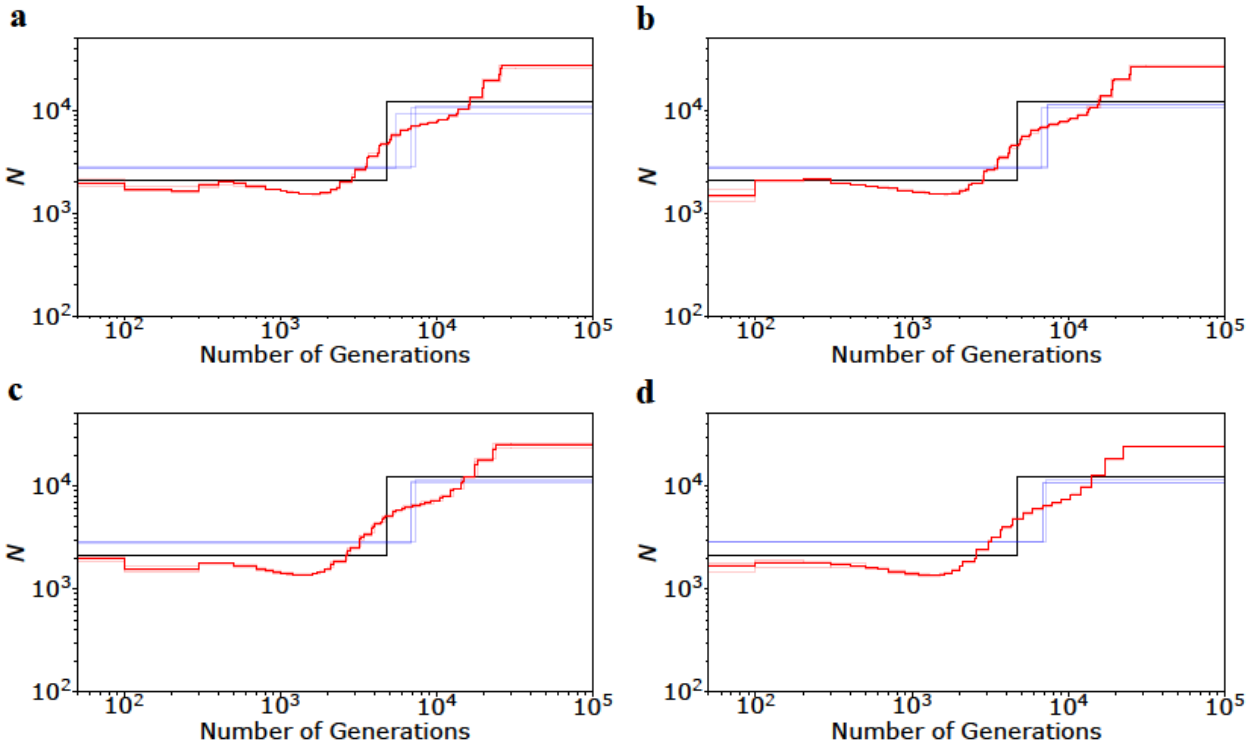
Supp Figure 12: Distribution of lengths of repeat regions in the human genomes (*hg19*). Shown above is the distribution of lengths up to 1000 bp, although lengths of repeat regions range between 6 to 160602 bp.



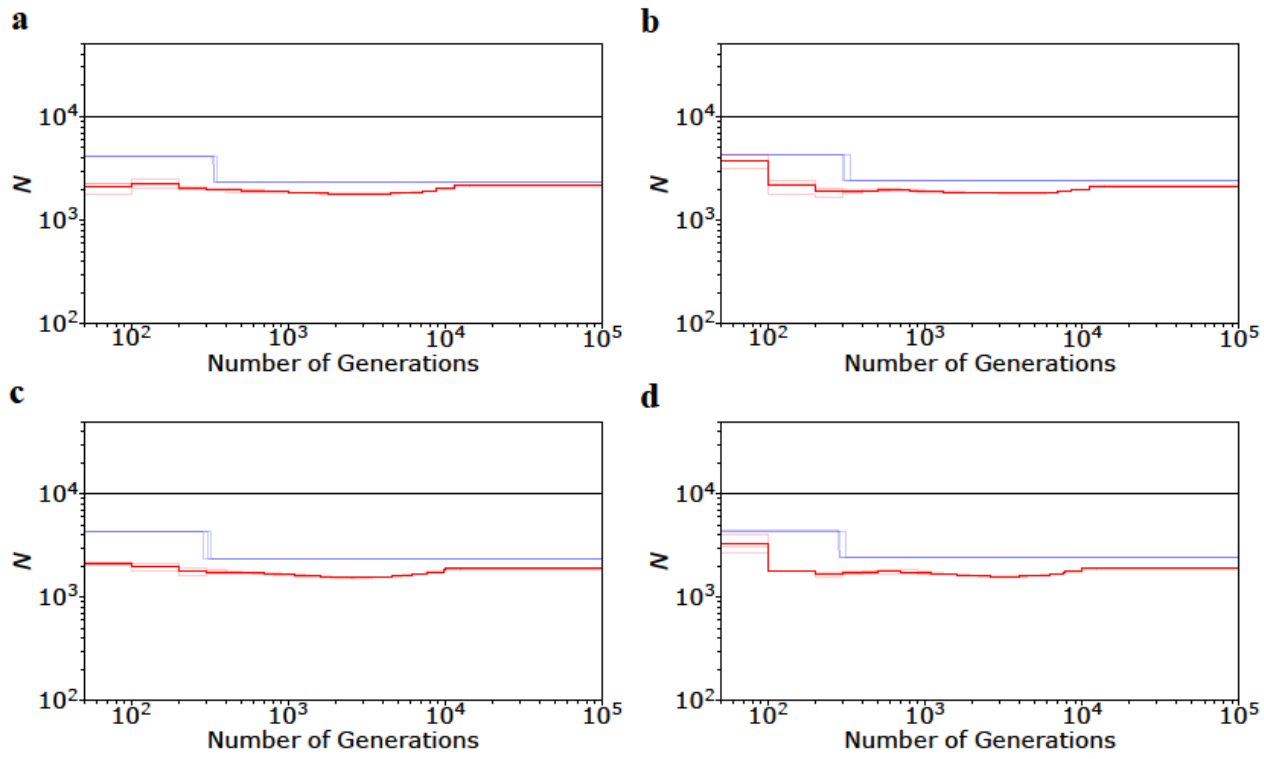
Supp Figure 13: Performance of demographic inference by MSMC (red lines) and *fastsimcoal2* (blue lines) under different scenarios of neutrality, when the true model is equilibrium: (a) there is variation in recombination and mutation rates, (b) there is variation in recombination and mutation rates and the centromeric region is masked, (c) there is variation in recombination and mutation rates, and short regions resembling repeats (comprising 10% of each chromosome) are randomly masked across the genome, and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome. The maximum and minimum fold change detected in every scenario is indicated on the upper right corner.



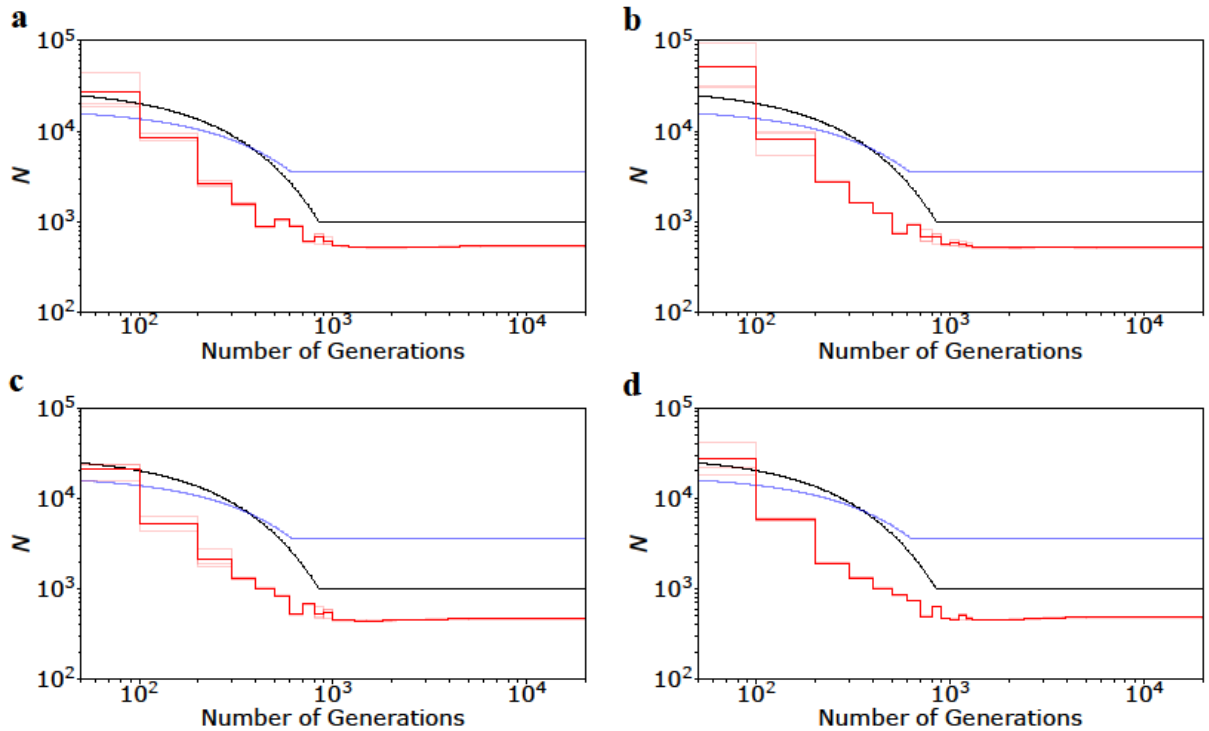
Supp Figure 14 : Performance of demographic inference by MSMC (red lines) and *fastsimcoal2* (blue lines) under different scenarios of neutrality, when the true model is 30-fold exponential growth: (a) there is variation in recombination and mutation rates across the genome, (b) there is variation in recombination and mutation rates, and the centromeric region is masked, (c) there is variation in recombination and mutation rates, and short regions resembling repeats are randomly masked across the genome (comprising of 10% of each chromosome), and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome.



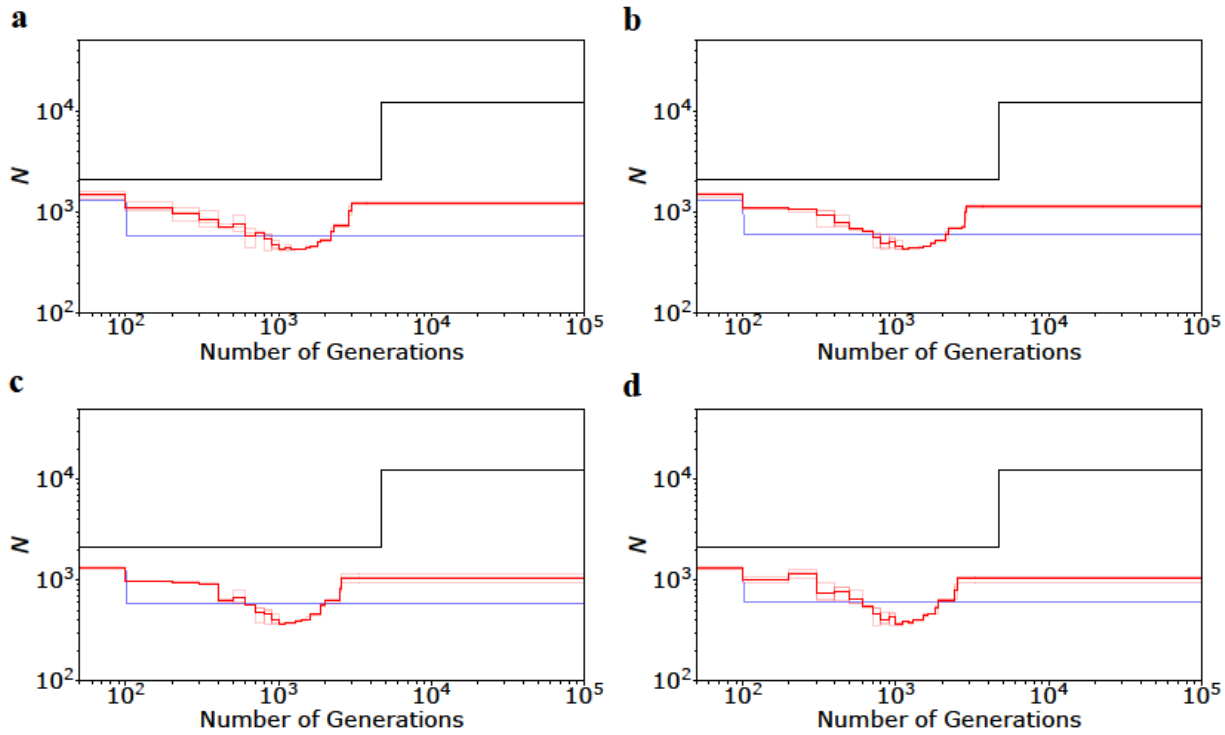
Supp Figure 15: Performance of demographic inference by MSMC under different scenarios of neutrality, when the true model is 6-fold instantaneous decline: (a) there is variation in recombination and mutation rates across the genome, (b) there is variation in recombination and mutation rates, and the centromeric region is masked, (c) there is variation in recombination sand mutation rates, and short regions resembling repeats are randomly masked across the genome (comprising of 10% of each chromosome), and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome.



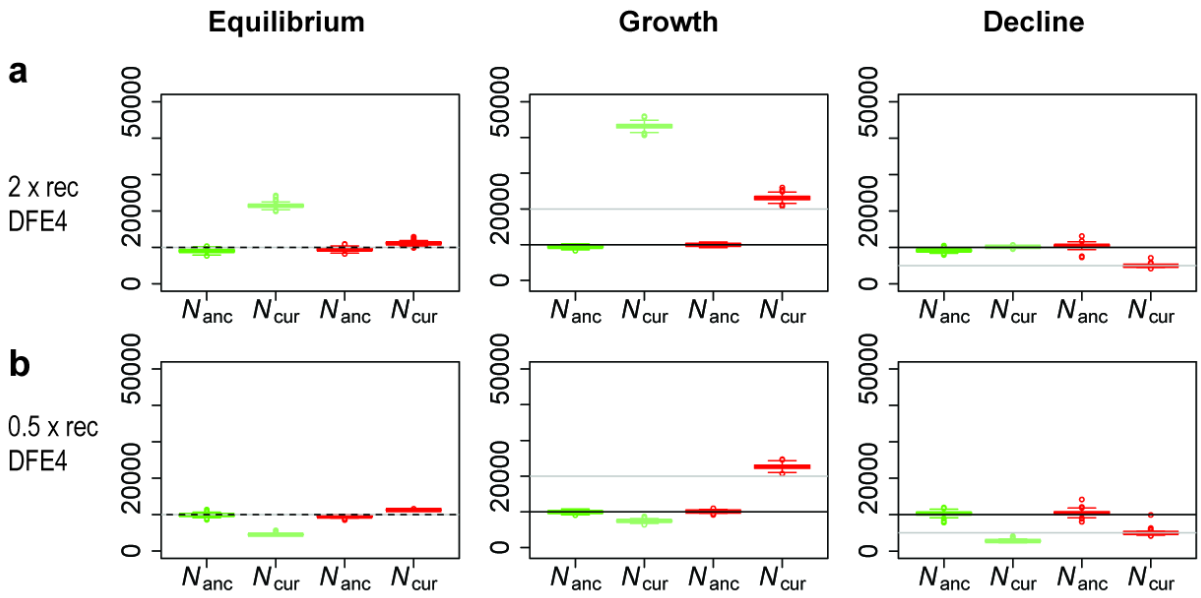
Supp Figure 16: Performance of demographic inference by MSMC (red lines) and *fastsimcoal2* (blue lines) in the presence of background selection, under different scenarios when the true model is equilibrium: (a) there is variation in recombination and mutation rates, (b) there is variation in recombination and mutation rates and the centromeric region is masked, (c) there is variation in recombination and mutation rates, and short regions resembling repeats (comprising 10% of each chromosome) are randomly masked across the genome, and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome. Exons comprise of 20% of the genome, experience purifying selection given by DFE4 ($f_0 = f_1 = f_2 = f_3 = 0.25$), and are masked when performing inference.



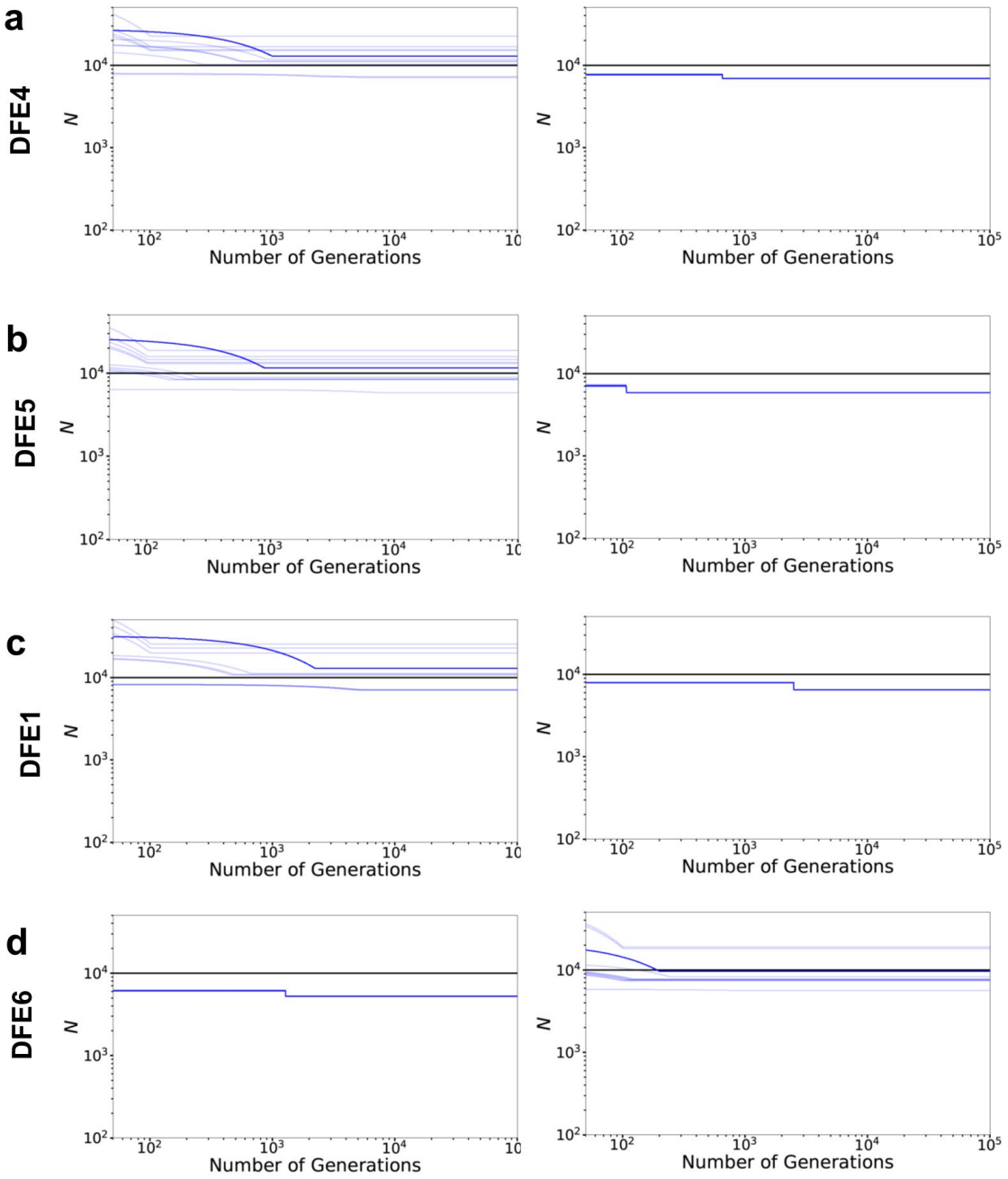
Supp Figure 17: Performance of demographic inference by MSMC (red lines) and *fastsimcoal2* (blue lines) in the presence of background selection, under different scenarios when the true model is 30-fold exponential growth: (a) there is variation in recombination and mutation rates, (b) there is variation in recombination and mutation rates and the centromeric region is masked, (c) there is variation in recombination and mutation rates, and short regions resembling repeats (comprising 10% of each chromosome) are randomly masked across the genome, and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome. Exons comprise of 20% of the genome, experience purifying selection given by DFE4 ($f_0 = f_1 = f_2 = f_3 = 0.25$), and are masked when performing inference.

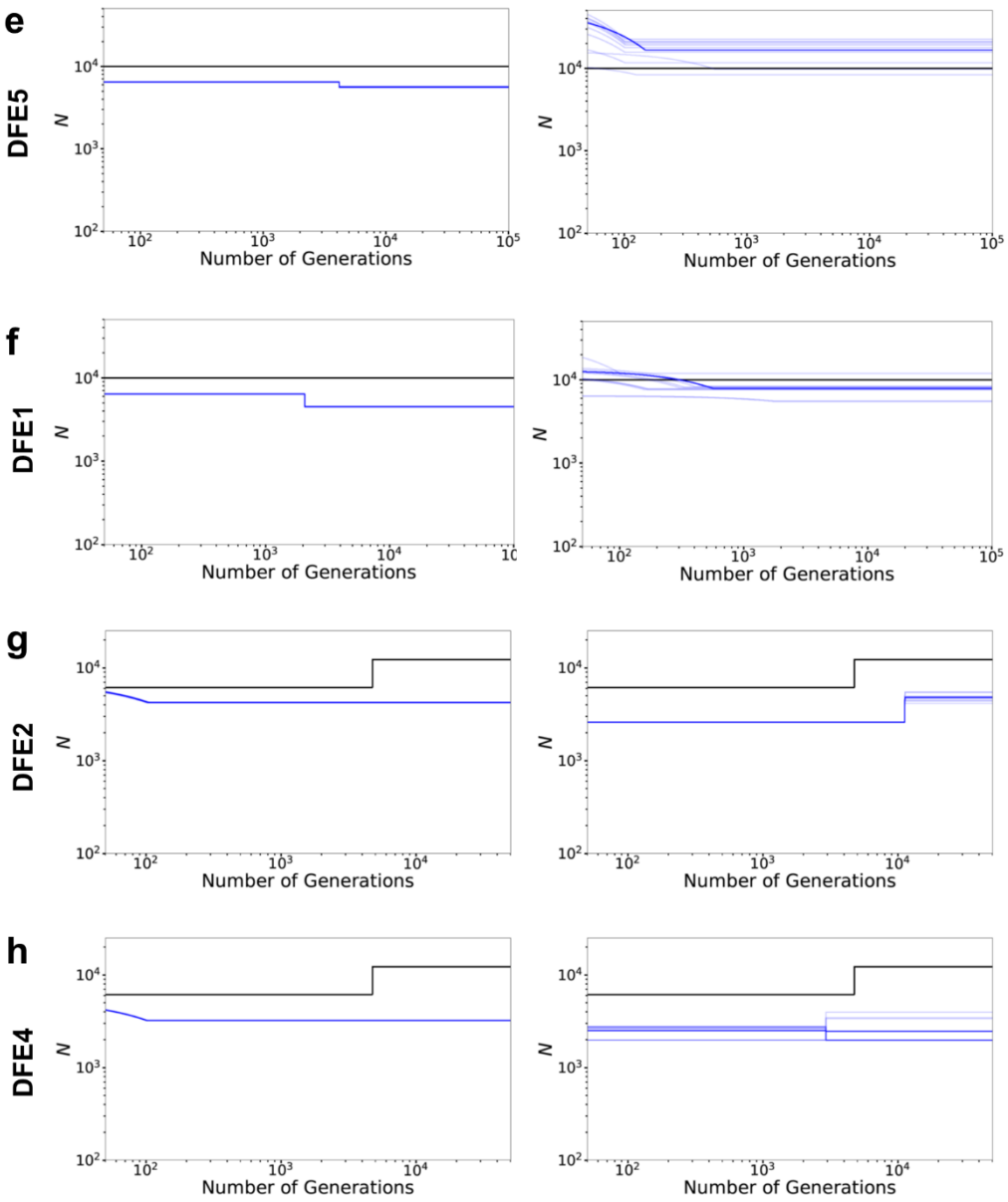


Supp Figure 18: Performance of demographic inference by MSMC (red lines) and *fastsimcoal2* (blue lines) in the presence of background selection, under different scenarios when the true model is a 6-fold instantaneous decline: (a) there is variation in recombination and mutation rates, (b) there is variation in recombination and mutation rates and the centromeric region is masked, (c) there is variation in recombination and mutation rates, and short regions resembling repeats (comprising 10% of each chromosome) are randomly masked across the genome, and (d) there is variation in recombination and mutation rates, and the centromere as well as small-sized repeats are randomly masked across the genome. Exons comprise of 20% of the genome, experience purifying selection given by DFE4 ($f_0 = f_1 = f_2 = f_3 = 0.25$), and are masked when performing inference.

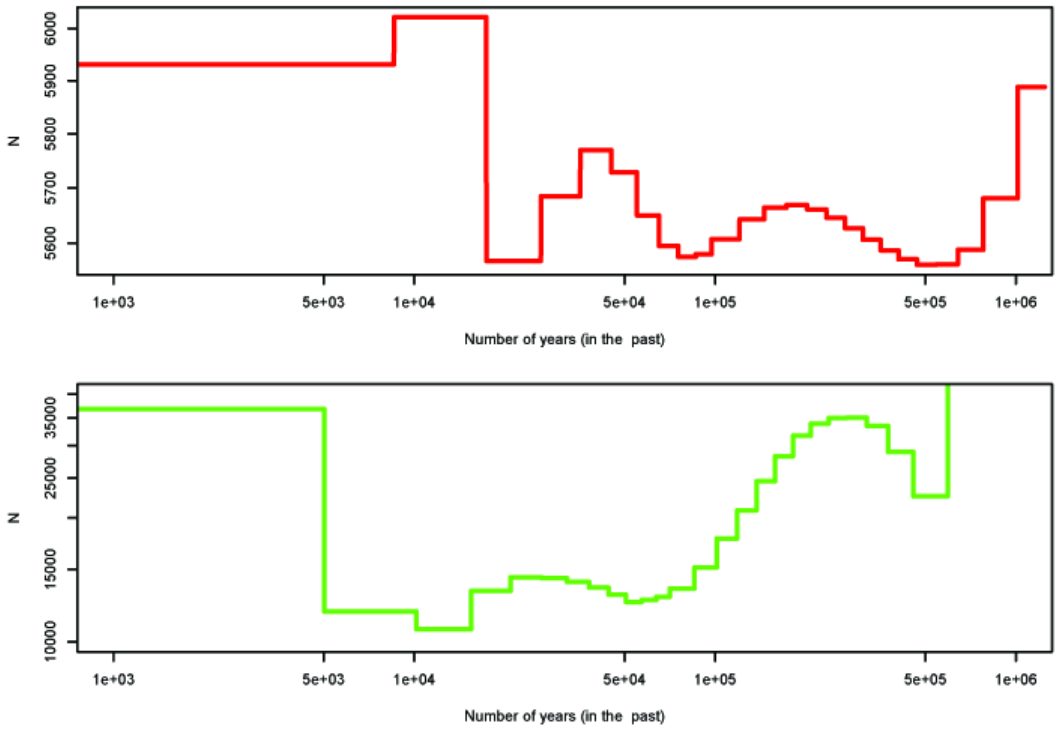


Supp Figure 19: Performance of the ABC method when recombination rate is mis-specified. (a) The true recombination rate is 2-fold higher than that assumed, and (b) the true recombination rate is 2-fold lower than that assumed. Boxplots in green show the posterior estimates when the recombination rate is higher or lower than assumed. For comparison, boxplots in red show the posterior inferred when the corresponding recombination rate is correctly specified. The black line displays the true ancestral population size (N_{anc}) and the gray line represents the true current population size (N_{cur}).


a
DFE4
b
DFE5
c
DFE1
d
DFE6
N
Number of Generations



Supp Figure 20: Scenarios in which demographic inference by *fastsimcoal2* resulted in different best models when performed using all SNPs (right column) and when SNPs were thinned to be separated by 5 kb (left column). The best model was defined as the model corresponding to the lowest AIC over all ten replicates. The DFEs are indicated on the left. (a) Direct purifying selection under DFE4 (exonic sites are not masked) in 5% of the genome; (b) Direct purifying selection under DFE5 (exonic sites are not masked) in 5% of the genome; (c) Direct purifying selection under DFE1 (exonic sites are not masked) in 10% of the genome; (d) Direct purifying selection under DFE6 (exonic sites are not masked) in 10% of the genome; (e) BGS under DFE5 (*i.e.*, exonic sites are masked) in which 5% of the genome is exonic; (f) BGS under DFE1 (*i.e.*, exonic sites are masked) in which 20% of the genome is exonic; (g) BGS under DFE2 (*i.e.*, exonic sites are masked) in which 20% of the genome is exonic; (h) BGS under DFE4 (*i.e.*, exonic sites are masked) in which 20% of the genome is exonic.



Supp Figure 21: Inference of demographic history by MSMC. Top panel / red line: simulations in which the true model is constant population size, and 50% of new mutations in exons are strongly deleterious with the remainder being neutral, where exons comprise 5% of the genome. Bottom panel / green line: the empirical estimate of population history of the YRI population inferred with MSMC by Schiffels and Durbin (2014). The x-axis is in years (assuming a generation time of 30 years). Note that the y-axes are on different scales, and the magnitude of change observed in the empirical data is considerably larger in the simulated data. Thus, this comparison is only meant to illustrate this common shape taken in MSMC plots (and see similar shapes in, for example, vervets (Warren *et al.* 2015; Figure 4) and passenger pigeons (Hung *et al.* 2014; Figure 2)).