The temporal dynamics of background selection in non-equilibrium populations

Raul Torres[1,*], **Markus G Stetter**[2,*], **Ryan D Hernandez**[3,4,†] **and Jeffrey Ross-Ibarra**[5,†]

[1]Biomedical Sciences Graduate Program, University of California San Francisco, San Francisco, CA, USA, [2]Botanical institute, University of Cologne, Cologne, Germany, [3]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA, [4]Department of Human Genetics and Genome Quebec Innovation Center, McGill University, Montreal, Canada, [5]Department of Evolution and Ecology, Genome Center, and Center for Population Biology, University of California, Davis, CA, USA

ABSTRACT Neutral genetic diversity across the genome is determined by the complex interplay of mutation, demographic history, and natural selection. While the direct action of natural selection is limited to functional loci across the genome, its impact can have effects on nearby neutral loci due to genetic linkage. These effects of selection at linked sites, referred to as genetic hitchhiking and background selection (BGS), are pervasive across natural populations. However, only recently has there been a focus on the joint consequences of demography and selection at linked sites, and empirical studies have sometimes come to apparently contradictory conclusions. In order to understand the relationship between demography and linked selection, we conducted an extensive forward simulation study of BGS under a range of demographic models. We found that levels of diversity compared to an equilibrium population vary over time and that the initial dynamics after a population size change are often in the opposite direction of the long-term expected trajectory. Our detailed observations of the temporal dynamics of neutral diversity in the context of selection at linked sites in nonequilibrium populations provides new intuition about why patterns of diversity under BGS vary through time in natural populations and help reconcile previously contradictory observations. Most notably, our results highlight that classical models of BGS are poorly suited for predicting diversity in nonequilibrium populations.

KEYWORDS demography, background selection, linked selection

Introduction

The effects of natural selection and demography on neutral genetic diversity within populations have long been of interest in evolutionary and population genetics. Recent efforts in sequencing tens of thousands of genomes across a multitude of species have yielded new and valuable insights into how these two forces of evolution have shaped extant patterns of genomic variation. Yet, while the theoretical underpinnings of the effects of natural selection and demography on genetic diversity have been investigated for decades (Smith and Haigh 1974; Nei *et al.* 1975; Maruyama and Fuerst 1984, 1985; Kaplan *et al.* 1989; Charlesworth *et al.* 1993; Nordborg *et al.* 1996; Hudson and Kaplan 1995; Tajima 1989), detailed investigation into how they jointly act to create patterns of diversity in different populations remains lacking.

Both theory and empirical observation have long shown that patterns of neutral genetic variation can vary regionally across the genome as a function of recombination rate (Smith and Haigh 1974; Begun and Aquadro 1992). This is because natural selection operating on selected sites not only decreases genetic variation at the focal site but can also lead to decreases in nearby neutral genetic diversity due to genetic linkage (Cutter and Payseur 2013). These effects, known as genetic hitchhiking (Smith and Haigh 1974) (in which neutral variants rise to high frequency with adaptive variants) and background selection (Charlesworth *et al.* 1993) (BGS; in which neutral variants are removed along with deleterious variants) can be widespread across the genome (Elyashiv *et al.* 2016). Evidence for selection at linked sites has been found across an array of species, including *Drosophila melanogaster* (Begun and Aquadro 1992; Comeron 2014; Charlesworth 1996; Andolfatto 2007; Sella *et al.* 2009; Elyashiv *et al.* 2016), mice (Keightley and Booker 2018), wild and domesticated rice (Flowers *et al.* 2011; Xu *et al.* 2012), *Capsella* (Williamson *et al.* 2014), monkeyflowers (Stankowski *et al.* 2018), maize (Beissinger *et al.* 2016), and humans (Sabeti *et al.* 2002; Reed *et al.* 2005; Voight *et al.* 2006; McVicker *et al.* 2009; Cai *et al.* 2009; Hernandez *et al.* 2011; Lohmueller *et al.* 2011).

Demographic change can also impact patterns of diversity across the genome. For example, neutral theory predicts that the amount of genetic diversity is proportional to a population's effective population size (N_e), such that changes in N_e should result in concomitant changes to diversity (Kimura 1983). One of the most common forms of a population size change is a population bottleneck, whereby populations suffer a large decrease followed by an expansion. Bottlenecks can occur via domestication events (Doebley *et al.* 2006; Tang *et al.* 2010; Wiener and Wilkinson 2011; Gaut *et al.* 2018), seasonal or cyclical fluctuations in population size (Elton 1924; Ives 1970; Itoh *et al.* 2009; Norén and Angerbjörn 2014), and founder events (David and Capy 1988; Dlugosch and Parker 2008; Henn *et al.* 2012). Notably,

† Department of Evolution and Ecology, University of California, Davis, CA, USA E-mail: rossibarra@ucdavis.edu; Department of Human Genetics, McGill University, Montreal, Canada E-mail: ryan.hernandez@me.com

* Authors contributed equally.

while the rate of loss of diversity in response to a population contraction is quite fast, the recovery of diversity following a population increase can be quite slow (Charlesworth 2009). As a result, large contemporary populations may still exhibit patterns of low average genetic diversity if their population size was much smaller in the recent past. In humans, this is clearly evident in European and Asian populations due to the out-of-Africa bottleneck (Auton *et al.* 2015).

Because selection at linked sites and demography are both pervasive forces across a multitude of species, the characterization of how these two forces interact with one another is necessary in order to develop a full picture on the determinants of neutral genetic diversity. The efficiency of natural selection scales proportionally with N_e and the impact of selection at linked sites on neutral diversity is likely to be greater in larger populations (Kaplan *et al.* 1989; Cutter and Payseur 2013; Corbett-Detig *et al.* 2015) (but see Gillespie (2001); Santiago and Caballero (2016)). Further, demographic changes can also increase (in the case of bottlenecks) or decrease (in the case of expansions) the rate of drift. It is therefore plausible that the rate at which diversity at a neutral locus is perturbed by selection at linked sites could be highly dependent on both the current as well as long-term N_e of the population. This competition between the strength of selection at linked sites (which increases with the census size N) and genetic drift (which decreases with census N) may be a key contributor to the limited range of diversity observed among species despite much larger observed differences in census size (Gillespie 2001; Corbett-Detig *et al.* 2015; Santiago and Caballero 2016; Lewontin 1974; Leffler *et al.* 2012). However, selection at linked sites alone may not be sufficient to explain the discrepancy between observed diversity and census populations sizes (Coop 2016), and the action of both demography and selection at linked sites in concert may provide a better model. Moreover, the heterogeneous structure of selection at linked sites across the genome may yield different responses to demography and population splits through time and their resulting effects on patterns of differentiation and divergence also remain largely unexplored (Burri 2017).

Many models of selection at linked sites were also formulated with the assumption that the population is large enough (or selection strong enough) such that mutation-selection balance is maintained (Charlesworth *et al.* 1993; Zeng 2013; Nicolaisen and Desai 2013). However, non-equilibrium demographic change may break such assumptions and forces other than selection may drive patterns of variation in regions experiencing selection at linked sites. For example, during the course of a population bottleneck, genetic drift may transiently dominate the effects of selection at many sites such that traditional models of selection will poorly predict patterns of genetic diversity. Additionally, in regions affected by selection at linked sites, the impact of genetic drift may also be exacerbated, resulting in greater losses to diversity than expected by the action of demography alone. A recent review by Comeron (2017) included a cursory investigation into the impact of demography on diversity in regions under BGS and suggested a dependency on demographic history. Recent empirical work in maize and humans has also demonstrated a strong interaction between demography and selection at linked sites (Beissinger *et al.* 2016; Torres *et al.* 2018). But these papers also demonstrate the need for a deeper understanding of the interaction between these two forces, as the two studies come to opposing conclusions about the impact of population bottlenecks and expansions on patterns of diversity in regions affected by selection at linked sites.

In order to more fully explore the joint consequences of demography and selection at linked sites, in this study we conducted extensive simulations of different demographic models jointly with the effects of BGS. We find that the time span removed from demographic events is critical for populations experiencing non-equilibrium demography and can yield contrasting patterns of diversity that reconciles apparently contradicting results (Beissinger *et al.* 2016; Torres *et al.* 2018). Additionally, the sensitivity of genetic diversity to demography is dependent on the frequency of the alleles being measured, with rare variants experiencing more dynamic changes through time.

Our results demonstrate that traditional models of selection at linked sites may be poorly suited for predicting patterns of diversity for populations experiencing recent demographic change, and that the predicted forces of BGS become apparent only after populations begin to approach equilibrium. Importantly, even simple intuition about the effect of selection at linked sites may lead to erroneous conclusions if populations are assumed to be at equilibrium. These results should motivate further research into this area and support the use of models that incorporate the joint effects of both demography and selection at linked sites.

Materials and Methods

Simulation model

We simulated a diploid and randomly mating population using fwdpy11 v1.2a (*https://github.com/molpopgen/fwdpy11*), a Python package using the fwdpp library (Thornton 2014). Selection parameters for simulating BGS followed those of Torres *et al.* (2018), with deleterious variation occurring at 20% of sites across a 2 Mb locus and the selection coefficient, s, drawn from two distributions of fitness effects (DFE). Specifically, 13% of deleterious sites were drawn from a gamma distribution (parameters: mean = α/β, variance = α/β^2) parameterized $\Gamma(\alpha = 0.0415, \beta = 80.11)$ and seven percent from a distribution parameterized $\Gamma(\alpha = 0.184, \beta = 6.25)$. These distributions mimic the DFEs inferred across non-coding and coding sites within the human genome (Torgerson *et al.* 2009; Boyko *et al.* 2008). Fitness followed a purely additive model in which the fitness effect of an allele was 0, $0.5s$, and s for homozygous ancestral, heterozygous, and homozygous derived genotypes, respectively. Per base pair mutation and recombination rates also followed those of Torres *et al.* (2018) and were 1.66×10^{-8} and 8.2×10^{-10}, respectively. We included a 200 kb neutral locus directly flanking the 2 Mb deleterious locus in order to observe the effects of BGS on neutral diversity. For all simulations, we simulated a burn-in period for 10N generations with an initial population size of 20,000 individuals before simulating under 12 specific demographic models. The demographic models included one demographic model of a constant sized population (model 1) and eleven non-equilibrium demographic models incorporating both bottlenecks and expansions (models 2-12; Figures S1-S2; Table S1). For each demographic model, we also conducted an identical set of neutral simulations without BGS by simulating only the 200 kb neutral locus. Each model scenario was simulated 5,000 times.

Diversity statistics and bootstrapping

After the burn-in period, we measured genetic diversity (π) and singleton density (ξ; the number of singletons observed within a locus) within 10 kb windows across the 200 kb neutral locus

every 50 generations using a random sample of 400 chromosomes. We measured π and ξ for each demographic model by taking the mean of these values across each set of 5,000 replicate simulations. For neutral simulations, we annotated π and ξ as π_0 and ξ_0, respectively. We took the ratio of these statistics (i.e., π/π_0 and ξ/ξ_0) in order to measure the relative impact of BGS within each demographic model. We bootstrapped the diversity statistics by sampling with replacement the 5,000 simulated replicates of each demographic model to generate a new set of 5,000 simulations, taking the mean of π and ξ across each new bootstrapped set. We conducted 10,000 bootstrap iterations and generated confidence intervals from the middle 95% of the resulting bootstrapped distribution.

Calculations of expected BGS

To calculate the predicted equilibrium π/π_0 given the instantaneous N_e at each time point for each demographic model, we used equation 14 of Nordborg *et al.* (1996), but modified it accordingly to incorporate two gamma distributions of fitness effects. Additionally, in order to properly model our simulations, we only calculated the effects of BGS on one side of the selected locus. This resulted in the following modified equation:

$$\frac{N_e}{N} \equiv \frac{\pi}{\pi_0} = f(U_T, \alpha_T, \beta_T) \times f(U_B, \alpha_B, \beta_B)$$

The first term of the equation ($f(U_T, \alpha_T, \beta_T)$) models the effects of BGS due to selection at non-coding sites according to the gamma DFE inferred by Torgerson *et al.* (2009), and the second term of the equation ($f(U_b, \alpha_B, \beta_B)$) models the effects of BGS due to selection on coding sites according to the gamma DFE inferred by Boyko *et al.* (2008). Each of these is modeled following:

$$f(U, \alpha, \beta) = exp\left(-\frac{U}{2R}\int_C^\infty \frac{1}{s}\left\{\int_0^R \frac{dz}{[1 + r(z)(1-s)/s]^2}\right\}\Gamma(s, \alpha, \beta)ds\right)$$

Here, R is the total length of the selected locus in bp, U is the total deleterious mutation rate across the selected locus, $r(z)$ is the genetic map distance between a neutral site and a deleterious mutation, and s is the selection coefficient of a deleterious mutation.

Because N_e is not explicitly included in this model of BGS, we followed previous work (Charlesworth 2012; Comeron 2014) in truncating selection at some value C (represented in the integral $\int_C^\infty$). Here, C represents the minimum selection coefficient (s) that is treated as deleterious for the model. This step effectively excludes neutral mutations from the model that should not contribute to BGS and can be modulated to mimic small or large populations (by increasing or decreasing C, respectively; Figure S3). This truncation step also affects the values used for U in the above equation, resulting in specific values of U for each DFE. We simulated different population sizes to equilibrium under our BGS simulation model to see how well the modified version of the classic model fit populations of different N_e for different values of C (Figure S3). We used $2N_eC = 0.15$ because this provided the best estimate of π/π_0 for the starting N_e of our demographic models (i.e., $N_e = 20,000$). While this value provides a coarse estimate for the effects of BGS on π/π_0 for a particular N_e, it will overestimate the effects of BGS for smaller N_e and potentially underestimate the effects of BGS for larger N_e (Figure S3).

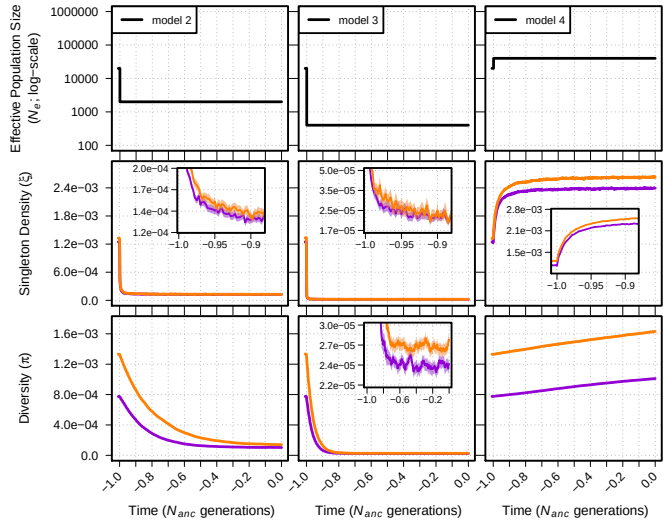


Figure 1 Singleton density (ξ per site) and diversity (π per site) for models 2-4. The top panel shows each demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Diverstiy statistics are shown for neutral simulations (orange lines) and simulations with BGS (violet lines). Insets show diversity using a log scale for improved detail. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

Results

Background selection under instantaneous population size change

We first present the joint effects of demography and BGS under simple demographic models with a single instantaneous change in size (models 2-4; Figure S1). While our simulations incorporated a 200 kb neutral region, we first focused on patterns of diversity generated within the 10 kb window nearest to the 2 Mb locus experiencing purifying selection, as this is where BGS is strongest. Doing so allowed us to observe any change in the dynamics of π and ξ as they approached new population equilibria resulting from a change in size. In the simple bottleneck models (models 2-3) we observed the expected strong decrease in ξ and π following population contraction in models of both BGS and neutrality (Figure 1). Similarly, we observed the expected rapid increase in ξ compared to π in our model of a simple population expansion (model 4; Figure 1). In all cases, values of ξ and π were lower in models with BGS and their relative values changed more rapidly than in the neutral case (Figure S4).

To examine the interaction of demography and selection observed in empirical data (Beissinger *et al.* 2016; Torres *et al.* 2018), we normalize π and ξ in models of BGS by their equivalent statistics generated under the same demographic model in the absence of any selection (π_0 and ξ_0). We observed that π/π_0 and ξ/ξ_0 were dynamic through time in response to demography, with changes occurring to both their magnitude and direction (Figure 2). Moreover, changes to ξ/ξ_0 occurred more rapidly through time compared to π/π_0. For example, in model 2 we ob-

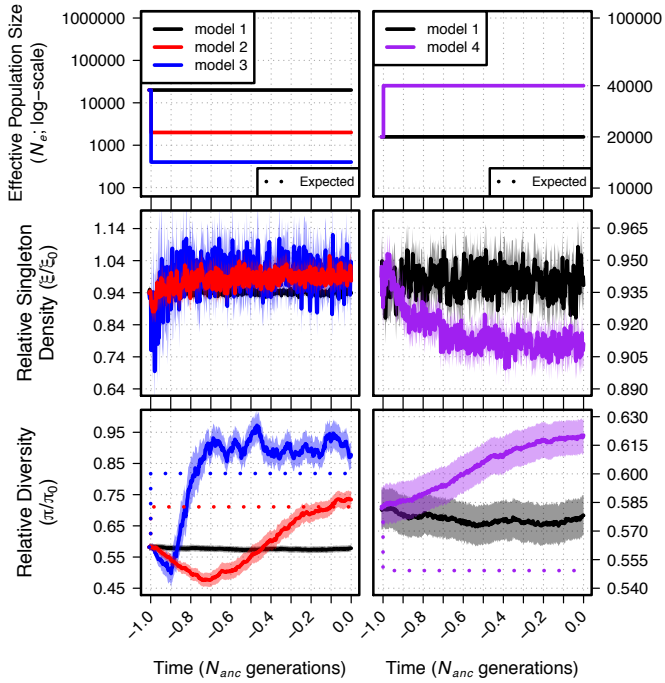


Figure 2 Relative singleton density (ξ/ξ_0) and relative diversity (π/π_0) across time for demographic models 1-4. The top panel shows each demographic model as in Figure 1. Black lines show ξ/ξ_0 and π/π_0 from simulations of a constant sized population (model 1). Dotted lines in the bottom panel show the equilibrium expectation of π/π_0 from Nordborg *et al.* (1996) given the specific selection parameters and the instantaneous N_e at each time point. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

served a dip and rise in the ξ/ξ_0 statistic relative to equilibrium (model 1) within the first $\approx 0.1N_{anc}$ generations (N_{anc} refers to the N_e of the ancestral population prior to any demographic change). Yet, for the same model, π/π_0 remained depressed for over 0.5 N_{anc} generations (Figure 2). Similar patterns were observed for model 3, which experienced a greater reduction in size, although the pattern is less clear because of the greater sampling variance of ξ/ξ_0 due to the overall lower number of singletons. In both population contraction models, π/π_0 and ξ/ξ_0 appeared to plateau at levels above that of the equilibrium model (model 1). In contrast, we observed markedly different dynamics in our model of a simple population expansion (model 4). This included a sustained increase in π/π_0 but only a transient increase in ξ/ξ_0 and within the first $\approx 0.1N_{anc}$ generations, ξ/ξ_0 declined to levels below that of the equilibrium model.

Changes in population size should lead to changes in the rate of genetic drift and the efficacy of natural selection and, thus, changes in the magnitude of BGS over time. Indeed, under equilibrium conditions (and if mutations that are effectively neutral can be ignored) the classic model of BGS (Nordborg *et al.* 1996) predicts weaker BGS (with higher π/π_0) for smaller populations and stronger BGS (with lower π/π_0) for larger populations. To compare these predictions to those of our simple demographic models, we calculated the predicted equilibrium π/π_0 under the classic model given the instantaneous N_e at each generation. In all three simple demographic models we observed that changes in π/π_0 over the short term differed qualitatively from the classic model (Figure 2; bottom panel). The classic model predicts a higher value for π/π_0 in a smaller population, yet we observed a transient drop in π/π_0 directly after a contraction (models 2 and 3). Similarly, while the classic model predicts a decrease in π/π_0 in larger populations, we observed an increase in π/π_0 with a population expansion (model 4). The trajectory of π/π_0 changed in our bottleneck models, eventually approaching the expectation predicted by the classic model, while π/π_0 in the expansion model continued to increase over the entire course of the simulation. To test if π/π_0 for the expansion model eventually reaches the expectation predicted by the classic model, we ran a limited set of simulations (2,000 total) for 11 N_{anc} generations. We found that, indeed, π/π_0 plateaued then decreased relative to its starting value, eventually approaching the expectation of the classic model after ≈ 10 N_{anc} generations (Figure S10). This was because, in the absence of selection, π plateaus more slowly when compared to π under BGS (Figure S11), and only once π under neutrality begins to approach equilibrium did we begin to observe the prediction of the classic model.

Background selection under bottleneck-expansion models

We built upon the simple two epoch demographic models to test more complex scenarios and better understand the relative effects of different events on patterns of diversity under BGS. Specifically, we simulated a population undergoing a contraction similar in size to models 2 and 3, but with a subsequent expansion to 400,000 individuals by the final generation (Figure S2). These bottleneck-expansion models included both ancient (1.0 N_{anc} generations in the past; models 5-8) and recent (0.1 N_{anc} generations in the past; models 9-12) bottlenecks with either an instantaneous expansion (models 5-6, 9-10) or a sustained bottleneck (models 7-8, 11-12). These models recapitulated several patterns observed in our simple bottleneck models, but with added dynamics. In all cases, diversity in models with BGS was both lower (Figures S5-S6) and changed more rapidly (Figures S7-S8) than in neutral simulations. Changes in diversity also occurred more quickly in models with a stronger or sustained bottleneck, and ξ again exhibited more rapid dynamics than did π. Mirroring results from our simple bottleneck scenarios, models with an ancient bottleneck (models 5-8) showed a transient decrease in ξ/ξ_0 and π/π_0 followed by an increase to higher values (Figure 3). Again changes in π/π_0 contrast with the expectations of the classic model, where BGS is expected to become more efficient in larger populations, thus resulting in an expected decrease in π/π_0 through time (Figure 3, dotted line). But while both π/π_0 and ξ/ξ_0 remain elevated in our simple bottleneck models, ξ/ξ_0 in the bottleneck-expansion models shifts direction during the course of the expansion and begins to decline, eventually reaching values below that of the equilibrium population. Though the trajectories of π/π_0 and ξ/ξ_0 were truncated for models in which the bottleneck occurred in the recent past (models 9-12; 0.1 N_{anc} generations), they nonetheless appeared to behave qualitatively similar to ancient bottleneck models (Figure S9). Notably, for these models, the ending values of ξ/ξ_0 and π/π_0 were in the opposite direction relative to model 1 when compared to the models with longer demographic histories.

Patterns of diversity across the 200 kb neutral region

We also measured patterns of π/π_0 across time for the entire 200 kb neutral region. Doing so showed the characteristic "trough"

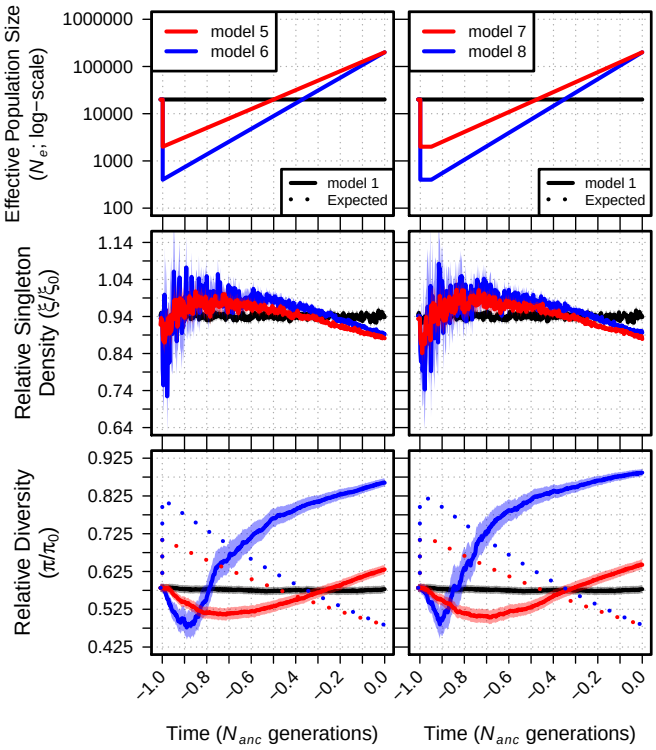


Figure 3 Relative singleton density (ξ/ξ_0) and relative diversity (π/π_0) across time for demographic models 1 and 5-8. The top panel shows each demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Black lines show ξ/ξ_0 and π/π_0 from simulations of a constant sized population (model 1). Dotted lines in the bottom panel show the equilibrium expectation of π/π_0 from Nordborg *et al.* (1996) given the specific selection parameters and the instantaneous N_e at each time point. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

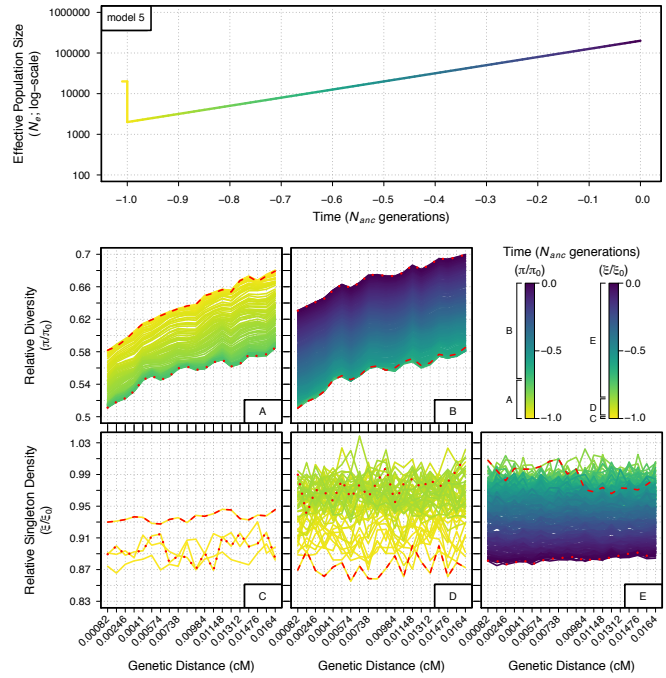


Figure 4 Temporal and spatial dynamics of relative diversity (π/π_0) and singleton density (ξ/ξ_0) under demographic model 5 across a neutral 200 kb region. The genetic distance of each 10 kb bin from the selected locus is indicated on the x-axis of the bottom panels. Each line measuring π/π_0 and ξ/ξ_0 in the bottom panels represent one of the 401 discrete generations sampled from the demographic model; colors follow the demographic model in the top panel (time is scaled as in Figures 1-3) and in the figure legend. Multiple plots are given in order to prevent overlap of the measurements between generations (see legend for specific generations covered in each plot). Red dashed lines and red dotted lines indicate the first and last generation measured within each plot (respectively).

structure of increasing relative diversity as a function of genetic distance from the deleterious locus (model 5 is shown in Figure 4, see Figure S12 for all models). Change in π/π_0 over time generally followed patterns observed in the neutral window closest to the selected region. In all of our ancient bottleneck models (models 2-3, 5-8), for example, we see a decline in π/π_0 across the entire region followed by an increase to levels higher than in the ancestral population. For recent bottlenecks (models 9-12) we see a consistent decline with no recovery and in our simple expansion model (model 4) π/π_0 increases monotonically through time.

Yet, these general patterns obscure more subtle changes in the slope of π/π_0 with increasing distance from the selected region. In models with a stronger bottleneck (models 3, 6 and 8), where we expect the efficacy of selection to be most affected, we see that the slope π/π_0 flattens over time, completely erasing the trough of diversity in the most extreme case without a recovery (model 3).

Finally, while ξ/ξ_0 across the region largely followed patterns seen in the neutral window most proximal to the selected locus, a closer look across the 200 kb regions of most models yielded no clear patterns. Troughs were slightly apparent for the final generations of some models (models 5 and 7), but the stochasticity among 10 kb windows for ξ/ξ_0 swamped any other patterns that might otherwise be evident.

Discussion

General patterns of diversity

A long history of both theoretical (Nei *et al.* 1975; Maruyama and Fuerst 1984, 1985) and empirical (Begun and Aquadro 1992; Cavalli-Sforza *et al.* 1994; Eyre-Walker *et al.* 1998) population genetics work has provided a clear picture of the impacts of demographic change on patterns of diversity in the genome. We know, for example, the impact of simple bottleneck and growth models such as those simulated here on the allele frequency spectrum (Tajima 1989; Slatkin and Hudson 1991; Griffiths and Tavare 1994). Theory also offers clear direction on the long-term effects of decreases in effective population size on the efficacy of natural selection (Kimura 1983). Likewise, classical theory on background selection provides a solid theoretical expectation for the effects of selection at linked sites on diversity in populations at demographic equilibrium (Nordborg *et al.* 1996).

Despite these efforts, there have been surprisingly few investigations addressing the expected patterns from the interaction

of demography and selection at linked sites in the context of BGS (Zeng 2013; Nicolaisen and Desai 2013; Comeron 2017). A thorough investigation into the joint effects of selection at linked sites and demography across a range of demographic parameters encountered in nature remains lacking. In addition, our parameter space explores the consequences of demography on selection at linked sites when selection becomes weak, breaking the assumptions of previously developed models (Nordborg *et al.* 1996; Nicolaisen and Desai 2013). For our DFE, $\approx 75\%$ of deleterious mutations will have a $s \leq 10^{-3}$, equivalent of $2N_{anc}s = 20$. There also remains substantial confusion in empirical population genetic analyses, where authors often equate long-term predictions of changes in effective populations size on the efficacy of natural selection to short-term responses under non-equilibrium demography (Brandvain and Wright 2016). Here, we use exhaustive simulations and analysis of different demographic models with and without BGS to show that predictions from such equilibrium models generally fail to hold up over shorter time scales, and that the predicted impacts of the combined effects of demography and selection at linked sites depend strongly on the details of the demographic model as well as the timing of sampling.

In each of our models, the initial effects observed are dominated by demography alone and fit well with theoretical expectations. Loss of diversity in the first few generations occurs equally across the entire region, independent of the distance from the selected region (Figure S12). While equilibrium models predict that the effects of BGS should be attenuated in populations with lower N_e due to the decreased efficacy of purifying selection, we instead observed a drop in π/π_0 after the bottleneck and a more rapid loss in models with a stronger bottleneck (Figures 2 and 3). These observations make it clear that effects of BGS on π/π_0 immediately following a reduction in N_e were not driven by a change in the efficacy of natural selection from population decline, but rather by increased sensitivity to allelic loss within these regions. Similarly, while theory predicts a decrease in π/π_0 in expanding populations, we see the initial response is instead an increase (Figure 2), and does not approach equilibrium expectations for $\sim 10N_{anc}$ generations (Figure S10).

Although the initial changes in diversity are dominated by the impacts of demography, as population size shifts, the efficacy of natural selection begins to change as well. In our simple bottleneck models, π/π_0 stops declining and begins to increase, eventually reaching higher values as expected under equilibrium (Figure 2). This change reflects the inability of a smaller population to effectively select against new deleterious mutations, rendering these alleles effectively neutral and decreasing the effects of BGS. These effects are countered in larger, growing populations, which is presumably why we see the rate of increase of π/π_0 slow and eventually plateau in models incorporating both bottlenecks and growth (Figure 3). Changes in the efficacy of selection are also readily observed in comparisons of diversity across neutral windows varying in recombination distance from the selected region (Figure S12). Diversity in the ancestral population increases with distance from the selected regions as expected under classical models of BGS at equilibrium (Nordborg *et al.* 1996) and observed in previous studies (Hernandez *et al.* 2011; Beissinger *et al.* 2016). But while the slope of this relationship remains constant in the generations initially following population size change in our simple bottleneck models, it begins to flatten through time, reflecting a lowered effective population size and concomitant weakened efficacy of natural selection (Figure S12; models 2-3).

The diversity-reducing effects of BGS have often been modeled as a reduction in N_e (Charlesworth *et al.* 1993), though we caution that the effects of BGS on the SFS cannot be simplified to this extent (Cvijović *et al.* 2018). Like a reduction in N_e, however, BGS exacerbates the stochastic process of drift. Because the relevant timescale for allele frequency evolution is scaled by the rate of drift (Crow and Kimura 1970), both the reduction and recovery of π/π_0 to equilibrium levels happen over fewer generations in populations with stronger bottlenecks and in regions impacted by BGS. We see this borne out in comparisons of models with stronger (Figure 2) or more sustained (Figure 3) bottlenecks, as well as comparisons of models with BGS to their equivalent neutral scenarios (Figures S4, S7, and S11). This framework also helps explain the slower changes observed in expanding populations (Figure 3), as increases in the effective size attenuate the rate of drift.

Because singleton variants represent very recent mutations, changes in ξ/ξ_0 responded quickly to changes in N_e. In our simple expansion model, for example, while π/π_0 never stops increasing, we see a relatively rapid increase in ξ/ξ_0, followed by a decrease as a larger population size increases the efficacy of selection against new deleterious mutants. And while theoretical predictions for ξ/ξ_0 are not as straightforward because of the dependency of sample size on distortions to the site-frequency spectrum (Cvijović *et al.* 2018), singleton diversity in the simple expansion model quickly plateaus at a new value below that of the ancestral population, consistent with having reached a new equilibrium value. However, signals using rare frequency bins such as ξ are inherently more difficult to capture, partly because it is less affected than π since BGS perturbs common frequency bins of the SFS more than rare ones (Cvijović *et al.* 2018). In addition, we observe much higher variance for ξ/ξ_0 compared to π/π_0 due to the smaller number of sites contributing to ξ.

Overall, then, we find that theoretical models of population demographic change or selection alone are insufficient to predict changes in diversity when both processes are at play. Both π/π_0 and ξ/ξ_0 show initial changes that often conflict with predictions from equilibrium models of BGS, reflecting instead the effects of rapid demographic change. But the rate at which diversity patterns begin to exhibit the impact changes in the efficacy of natural selection varies, differing between π/π_0 and ξ/ξ_0 and depending on the effects of demography and BGS on the rate of drift and thus the timescale over which changes can be observed. Although we have simulated under a mixture distribution of selection coefficients for new mutations (see Methods), this distribution will also play an important role in determining the threshold above which new mutations are efficiently removed by selection. In practice, it will be difficult to know any of these features in empirical data, and thus simple predictions — for example that a bottleneck will reduce the efficacy of natural selection and lead to increases in π/π_0 — are likely to be difficult to make with much accuracy.

Comparisons to empirical data

One of the motivations for the work presented here is the fact that empirical analyses evaluating the impact of demography on selection at linked sites have come to conflicting conclusions (Torres *et al.* 2018; Beissinger *et al.* 2016). Cultivated maize, for example, is thought to have undergone a bottleneck during the process of domestication (Eyre-Walker *et al.* 1998; Tenaillon *et al.* 2004; Wright *et al.* 2005), followed by a substantial expansion to

a modern size several orders of magnitude larger than its wild ancestor teosinte (Beissinger *et al.* 2016; Bellon *et al.* 2018). A recent analysis of selection at linked sites in maize and teosinte found differences, and interpreted them to be due to the timing of these demographic events. Beissinger *et al.* (2016) found that π/π_0 exhibited a greater trough around selected sites in teosinte, consistent with this taxon's larger long-term N_e. However, their analysis of ξ/ξ_0 found the opposite pattern — stronger effects in maize than teosinte — a result the authors interpreted as a reflection of the post-domestication expansion and increased efficacy of selection in maize in the recent past.

Human demographic history appears at least qualitatively similar to that of maize, with a bottleneck associated with migration out of Africa followed by considerable recent population expansion (Tennessen *et al.* 2012). In spite of this, however, analysis of selection at linked sites in humans produced different results (Torres *et al.* 2018). The bottlenecked non-African populations were found to have lower π/π_0 but higher ξ/ξ_0 than African populations. The authors interpreted this as the effects of a demographic bottleneck being exacerbated in regions of BGS, leading to greater losses of diversity (π) in Europeans. However, the authors also concluded that the lowered N_e resulting from the European bottleneck led to weaker effects of BGS overall, which were manifested in the higher observed values of ξ/ξ_0.

While there are a number of challenges associated with accurate estimation of historical demography in natural and domesticated populations (Myers *et al.* 2008; Pool *et al.* 2010; Bhaskar and Song 2014; Terhorst and Song 2015; Schraiber and Akey 2015; Beichman *et al.* 2018), our models highlight the fact that results seen in both maize and humans are plausible under even relatively simple models that are broadly consistent with the dynamics of these two systems. For example, under the bottleneck-expansion models of Figure 3, sampling in the present (generation 0) returns results similar to those seen in maize, with π/π_0 higher than a constant-size reference population but ξ/ξ_0 showing lower values due to the increased efficacy of selection in the expanding population. Yet, if a sample was taken $\approx -0.8N_{anc}$ - $-0.6N_{anc}$ generations in the past for model 5 or 7, they would reveal the observed pattern in humans, with a lower π/π_0 but a higher ξ/ξ_0, reflecting the impacts of a recent bottleneck. Indeed, in our simulations of demographic models with a short time span (models 9-12), we also observe these results when sampling in the present (Figure S9).

Conclusions

Genetic diversity across the genome is determined by the complex interplay of mutation, demographic history, and the effects of both direct and linked natural selection. While each of these processes is understood to a degree on its own, in many cases we lack either theory or sufficient empirical data to capture the effects of their interaction.

Selection at linked sites, in particular, is increasingly recognized as perhaps the primary determinant of patterns of diversity along a chromosome (Stankowski *et al.* 2018), but our ability to infer its impact is often complicated by changes in population size. Indeed, many studies make the simplifying assumption that selection at linked sites in such non-equilibrium populations can be effectively modeled using classic theory and scaling of the effective population size. Our extensive simulations show that this is not the case.

We find that the relationship between selection at linked sites and demographic change is complex, with short-term dynamics often qualitatively different from predictions under classic models. These results suggest that inferring the impact of population size change on selection at linked sites should be undertaken with caution, and is only really possible with a thorough understanding of the demographic history of the populations of interest.

Acknowledgments

MGS and JR-I would like to acknowledge funding from NSF Plant Genome Grant 1238014 and funding from the Deutsche Forschungsgemeinschaft (DFG) grant STE 2654/1-1 to MGS. RDH was partially supported by grant R01HG007644 from the National Institutes of Health, and RT was supported by a Diversity Supplement under this award. We would also like to thank Felix Andrews for statistical advice, although we did not follow it.

References

Andolfatto, P., 2007 Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. Genome Research **17**: 1755–1762.

Auton, A., G. R. Abecasis, D. M. Altshuler, R. M. Durbin, *et al.*, 2015 A global reference for human genetic variation. Nature **526**: 68–74.

Begun, D. J. and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Nature **356**: 519–520.

Beichman, A. C., E. Huerta-Sanchez, and K. E. Lohmueller, 2018 Using genomic data to infer historic population dynamics of nonmodel organisms. Annual Review of Ecology, Evolution, and Systematics **49**: 433–456.

Beissinger, T. M., L. Wang, K. Crosby, A. Durvasula, M. B. Hufford, *et al.*, 2016 Recent demography drives changes in linked selection across the maize genome. Nature Plants **2**: 16084.

Bellon, M. R., A. Mastretta-Yanes, A. Ponce-Mendoza, D. Ortiz-Santamaría, O. Oliveros-Galindo, *et al.*, 2018 Evolutionary and food supply implications of ongoing maize domestication by mexican campesinos. Proceedings of the Royal Society B: Biological Sciences **285**: 20181049.

Bhaskar, A. and Y. S. Song, 2014 Descartes' rule of signs and the identifiability of population demographic models from genomic variation data. The Annals of Statistics **42**: 2469–2493.

Boyko, A. R., S. H. Williamson, A. R. Indap, J. D. Degenhardt, R. D. Hernandez, *et al.*, 2008 Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genetics **4**: e1000083.

Brandvain, Y. and S. I. Wright, 2016 The limits of natural selection in a nonequilibrium world. Trends in Genetics **32**: 201–210.

Burri, R., 2017 Interpreting differentiation landscapes in the light of long-term linked selection. Evolution Letters **1**: 118–131.

Cai, J. J., J. M. Macpherson, G. Sella, and D. A. Petrov, 2009 Pervasive hitchhiking at coding and regulatory sites in humans. PLoS Genetics **5**: e1000336.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza, 1994 *The history and geography of human genes*. Princeton university press.

Charlesworth, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genetics Research **68**: 131–149.

Charlesworth, B., 2009 Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics **10**: 195–205.

Charlesworth, B., 2012 The role of background selection in shaping patterns of molecular evolution and variation: evidence from variability on the *Drosophila* X chromosome. Genetics **191**: 233–246.

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

Comeron, J. M., 2014 Background selection as baseline for nucleotide variation across the *Drosophila* genome. PLoS Genetics **10**: e1004434.

Comeron, J. M., 2017 Background selection as null hypothesis in population genomics: insights and challenges from *Drosophila* studies. Phil. Trans. R. Soc. B **372**: 20160471.

Coop, G., 2016 Does linked selection explain the narrow range of genetic diversity across species? bioRxiv .

Corbett-Detig, R. B., D. L. Hartl, and T. B. Sackton, 2015 Natural selection constrains neutral diversity across a wide range of species. PLoS Biology **13**: e1002112.

Crow, J. F. and M. Kimura, 1970 *An introduction to population genetics theory*. Harper & Row, New York.

Cutter, A. D. and B. A. Payseur, 2013 Genomic signatures of selection at linked sites: unifying the disparity among species. Nature Reviews Genetics **14**: 262–274.

Cvijović, I., B. H. Good, and M. M. Desai, 2018 The effect of strong purifying selection on genetic diversity. Genetics **209**: 1235–1278.

David, J. R. and P. Capy, 1988 Genetic variation of *Drosophila melanogaster* natural populations. Trends in Genetics **4**: 106–111.

Dlugosch, K. and I. Parker, 2008 Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology **17**: 431–449.

Doebley, J. F., B. S. Gaut, and B. D. Smith, 2006 The molecular genetics of crop domestication. Cell **127**: 1309–1321.

Elton, C. S., 1924 Periodic fluctuations in the numbers of animals: their causes and effects. Journal of Experimental Biology **2**: 119–163.

Elyashiv, E., S. Sattath, T. T. Hu, A. Strutsovsky, G. McVicker, *et al.*, 2016 A genomic map of the effects of linked selection in *Drosophila*. PLoS Genetics **12**: e1006130.

Eyre-Walker, A., R. L. Gaut, H. Hilton, D. L. Feldman, and B. S. Gaut, 1998 Investigation of the bottleneck leading to the domestication of maize. Proceedings of the National Academy of Sciences **95**: 4441–4446.

Flowers, J. M., J. Molina, S. Rubinstein, P. Huang, B. A. Schaal, *et al.*, 2011 Natural selection in gene-dense regions shapes the genomic pattern of polymorphism in wild and domesticated rice. Molecular Biology and Evolution **29**: 675–687.

Gaut, B. S., D. K. Seymour, Q. Liu, and Y. Zhou, 2018 Demography and its effects on genomic variation in crop domestication. Nature Plants **4**: 512–520.

Gillespie, J. H., 2001 Is the population size of a species relevant to its evolution? Evolution **55**: 2161–2169.

Griffiths, R. C. and S. Tavare, 1994 Sampling theory for neutral alleles in a varying environment. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **344**: 403–410.

Henn, B. M., L. L. Cavalli-Sforza, and M. W. Feldman, 2012 The great human expansion. Proceedings of the National Academy of Sciences **109**: 17758–17764.

Hernandez, R. D., J. L. Kelley, E. Elyashiv, S. C. Melton, A. Auton, *et al.*, 2011 Classic selective sweeps were rare in recent human evolution. Science **331**: 920–924.

Hudson, R. R. and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics **141**: 1605–1617.

Itoh, M., N. Nanba, M. Hasegawa, N. Inomata, R. Kondo, *et al.*, 2009 Seasonal changes in the long-distance linkage disequilibrium in *Drosophila melanogaster*. Journal of heredity **101**: 26–32.

Ives, P. T., 1970 Further genetic studies of the South Amherst population of *Drosophila melanogaster*. Evolution **24**: 507–518.

Kaplan, N. L., R. Hudson, and C. Langley, 1989 The "hitchhiking effect" revisited. Genetics **123**: 887–899.

Keightley, P. D. and T. R. Booker, 2018 Understanding the factors that shape patterns of nucleotide diversity in the house mouse genome. Molecular Biology and Evolution **35**: 2971–2988.

Kimura, M., 1983 *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.

Leffler, E. M., K. Bullaughey, D. R. Matute, W. K. Meyer, L. Segurel, *et al.*, 2012 Revisiting an old riddle: What determines genetic diversity levels within species? PLoS Biology **10**: e1001388.

Lewontin, R. C., 1974 *The genetic basis of evolutionary change*. Columbia University Press, New York.

Lohmueller, K. E., A. Albrechtsen, Y. Li, S. Y. Kim, T. Korneliussen, *et al.*, 2011 Natural selection affects multiple aspects of genetic variation at putatively neutral sites across the human genome. PLoS Genetics **7**: e1002326.

Maruyama, T. and P. A. Fuerst, 1984 Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. Genetics **108**: 745–763.

Maruyama, T. and P. A. Fuerst, 1985 Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. Genetics **111**: 675–689.

McVicker, G., D. Gordon, C. Davis, and P. Green, 2009 Widespread genomic signatures of natural selection in hominid evolution. PLoS Genetics **5**: e1000471.

Myers, S., C. Fefferman, and N. Patterson, 2008 Can one learn history from the allelic spectrum? Theoretical Population Biology **73**: 342–348.

Nei, M., T. Maruyama, and R. Chakraborty, 1975 The bottleneck effect and genetic variability in populations. Evolution **29**: 1–10.

Nicolaisen, L. E. and M. M. Desai, 2013 Distortions in genealogies due to purifying selection and recombination. Genetics **195**: 221–230.

Nordborg, M., B. Charlesworth, and D. Charlesworth, 1996 The effect of recombination on background selection. Genetics Research **67**: 159–174.

Norén, K. and A. Angerbjörn, 2014 Genetic perspectives on northern population cycles: bridging the gap between theory and empirical studies. Biological Reviews **89**: 493–510.

Pool, J. E., I. Hellmann, J. D. Jensen, and R. Nielsen, 2010 Population genetic inference from genomic sequence variation. Genome Research **20**: 291–300.

Reed, F. A., J. M. Akey, and C. F. Aquadro, 2005 Fitting background-selection predictions to levels of nucleotide variation and divergence along the human autosomes. Genome Research **15**: 1211–1221.

Sabeti, P. C., D. E. Reich, J. M. Higgins, H. Z. Levine, D. J. Richter, *et al.*, 2002 Detecting recent positive selection in the human genome from haplotype structure. Nature **419**: 832–837.

Santiago, E. and A. Caballero, 2016 Joint prediction of the ef-

fective population size and the rate of fixation of deleterious mutations. Genetics **204**: 1267–1279.

Schraiber, J. G. and J. M. Akey, 2015 Methods and models for unravelling human evolutionary history. Nature Reviews Genetics **16**: 727–740.

Sella, G., D. A. Petrov, M. Przeworski, and P. Andolfatto, 2009 Pervasive natural selection in the *Drosophila* genome? PLoS Genetics **5**: e1000495.

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

Smith, J. M. and J. Haigh, 1974 The hitch-hiking effect of a favourable gene. Genetics Research **23**: 23–35.

Stankowski, S., M. A. Chase, A. M. Fuiten, P. L. Ralph, and M. A. Streisfeld, 2018 The tempo of linked selection: Rapid emergence of a heterogeneous genomic landscape during a radiation of monkeyflowers. bioRxiv p. 342352.

Tajima, F., 1989 The effect of change in population size on DNA polymorphism. Genetics **123**: 597–601.

Tang, H., U. Sezen, and A. H. Paterson, 2010 Domestication and plant genomes. Current Opinion in Plant Biology **13**: 160–166.

Tenaillon, M. I., J. U'ren, O. Tenaillon, and B. S. Gaut, 2004 Selection versus demography: a multilocus investigation of the domestication process in maize. Molecular Biology and Evolution **21**: 1214–1225.

Tennessen, J. A., A. W. Bigham, T. D. O'connor, W. Fu, E. E. Kenny, *et al.*, 2012 Evolution and functional impact of rare coding variation from deep sequencing of human exomes. science **337**: 64–69.

Terhorst, J. and Y. S. Song, 2015 Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. Proceedings of the National Academy of Sciences **112**: 7677–7682.

Thornton, K. R., 2014 A C++ template library for efficient forward-time population genetic simulation of large populations. Genetics **198**: 157–166.

Torgerson, D. G., A. R. Boyko, R. D. Hernandez, A. Indap, X. Hu, *et al.*, 2009 Evolutionary processes acting on candidate cis-regulatory regions in humans inferred from patterns of polymorphism and divergence. PLoS Genetics **5**: e1000592.

Torres, R., Z. A. Szpiech, and R. D. Hernandez, 2018 Human demographic history has amplified the effects of background selection across the genome. PLoS Genetics **14**: e1007387.

Voight, B. F., S. Kudaravalli, X. Wen, and J. K. Pritchard, 2006 A map of recent positive selection in the human genome. PLoS Biology **4**: e72.

Wiener, P. and S. Wilkinson, 2011 Deciphering the genetic basis of animal domestication. Proceedings of the Royal Society B: Biological Sciences **278**: 3161–3170.

Williamson, R. J., E. B. Josephs, A. E. Platts, K. M. Hazzouri, A. Haudry, *et al.*, 2014 Evidence for widespread positive and negative selection in coding and conserved noncoding regions of *Capsella grandiflora*. PLoS Genetics **10**: e1004622.

Wright, S. I., I. V. Bi, S. G. Schroeder, M. Yamasaki, J. F. Doebley, *et al.*, 2005 The effects of artificial selection on the maize genome. Science **308**: 1310–1314.

Xu, X., X. Liu, S. Ge, J. D. Jensen, F. Hu, *et al.*, 2012 Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nature Biotechnology **30**: 105–111.

Zeng, K., 2013 A coalescent model of background selection with recombination, demography and variation in selection coefficients. Heredity **110**: 363–371.

model	ancestral population size (N_e [N_{anc}])	bottleneck/expansion population size (N_e)	bottleneck time length (N_{anc} generations)	expansion time length (N_{anc} generations)*	final population size (N_e)
model 1	20000	NA	NA	NA	20000
model 2	20000	2000	1	NA	2000
model 3	20000	400	1	NA	400
model 4	20000	40000	NA	1	40000
model 5	20000	2000	0	1	200000
model 6	20000	400	0	1	200000
model 7	20000	2000	0.05	0.95	200000
model 8	20000	400	0.05	0.95	200000
model 9	20000	2000	0	0.1	200000
model 10	20000	400	0	0.1	200000
model 11	20000	2000	0.05	0.05	200000
model 12	20000	400	0.05	0.05	200000

*expansion in models 5-12 is exponential but in model 4 is instantaneous

Table S1 Demographic parameters for models 1-12

Supplement

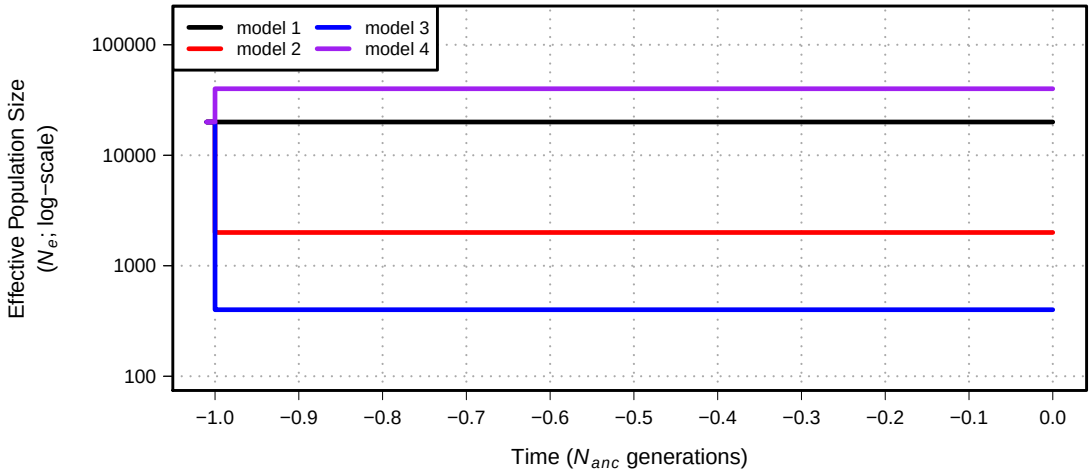


Figure S1 Demographic models 1-4 simulated in our study. Time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Demographic model 2 experiences a population contraction to 2000 individuals while demographic model 3 experiences a population contraction to 400 individuals. Demographic model 4 experiences a population expansion to 40,000 individuals. All population size changes are instantaneous for models 2-4. See Table S1 for additional model parameters.

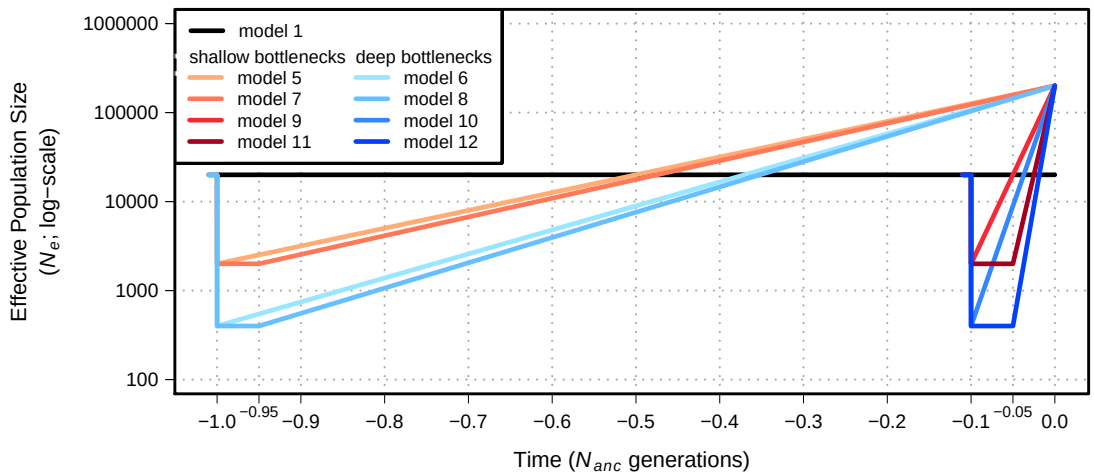


Figure S2 Demographic models 1 and 5-12 simulated in our study. Time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Demographic models with a shallow bottleneck (models 5, 7, 9, and 11) experience a population contraction to 2000 individuals while demographic models with a deep bottleneck (models 6, 8, 10, and 12) experience a population contraction to 400 individuals. After contraction, demographic models 5-12 undergo exponential growth to a final population size of 200,000 individuals. See Table S1 for additional model parameters.

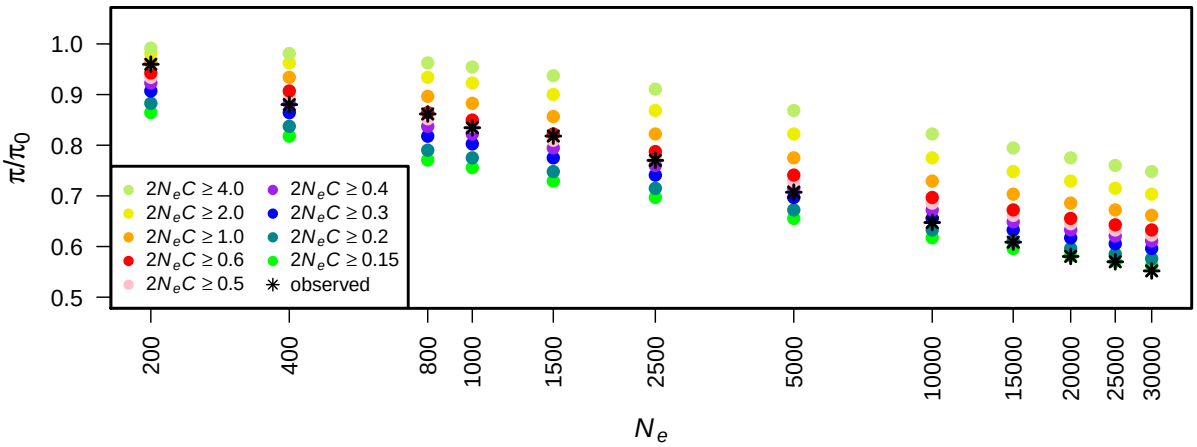


Figure S3 Estimate of π/π_0 from the classic model (Nordborg *et al.* 1996) across different population sizes and different truncation thresholds on selection. Different γ values used to truncate selection (C) for the classic model are shown in the legend ($2N_eC \geq \gamma$). Black stars represent the observed π/π_0 from running simulations of BGS.

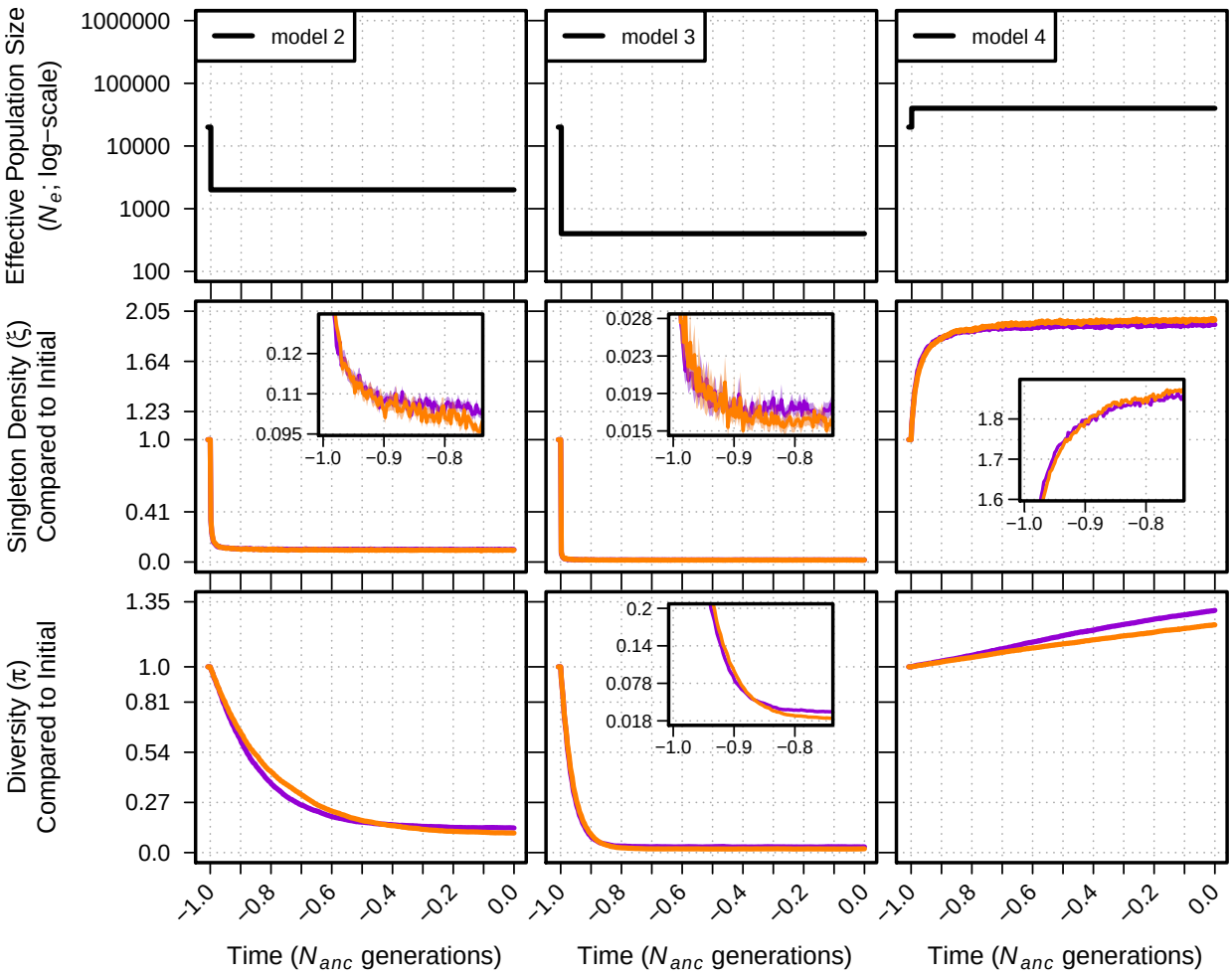


Figure S4 Singleton density (ξ) and diversity (π) for demographic models 2-4 under neutrality (orange lines) and BGS (violet lines) relative to their values in the initial generation prior to demographic change. The top panel shows each demographic model as in Figure 1. For greater detail, insets show data for generations over a smaller time scale and smaller y-axis (note: y-axes for insets are scaled linearly). Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data. The data used for this figure is identical to that of Figure 1.

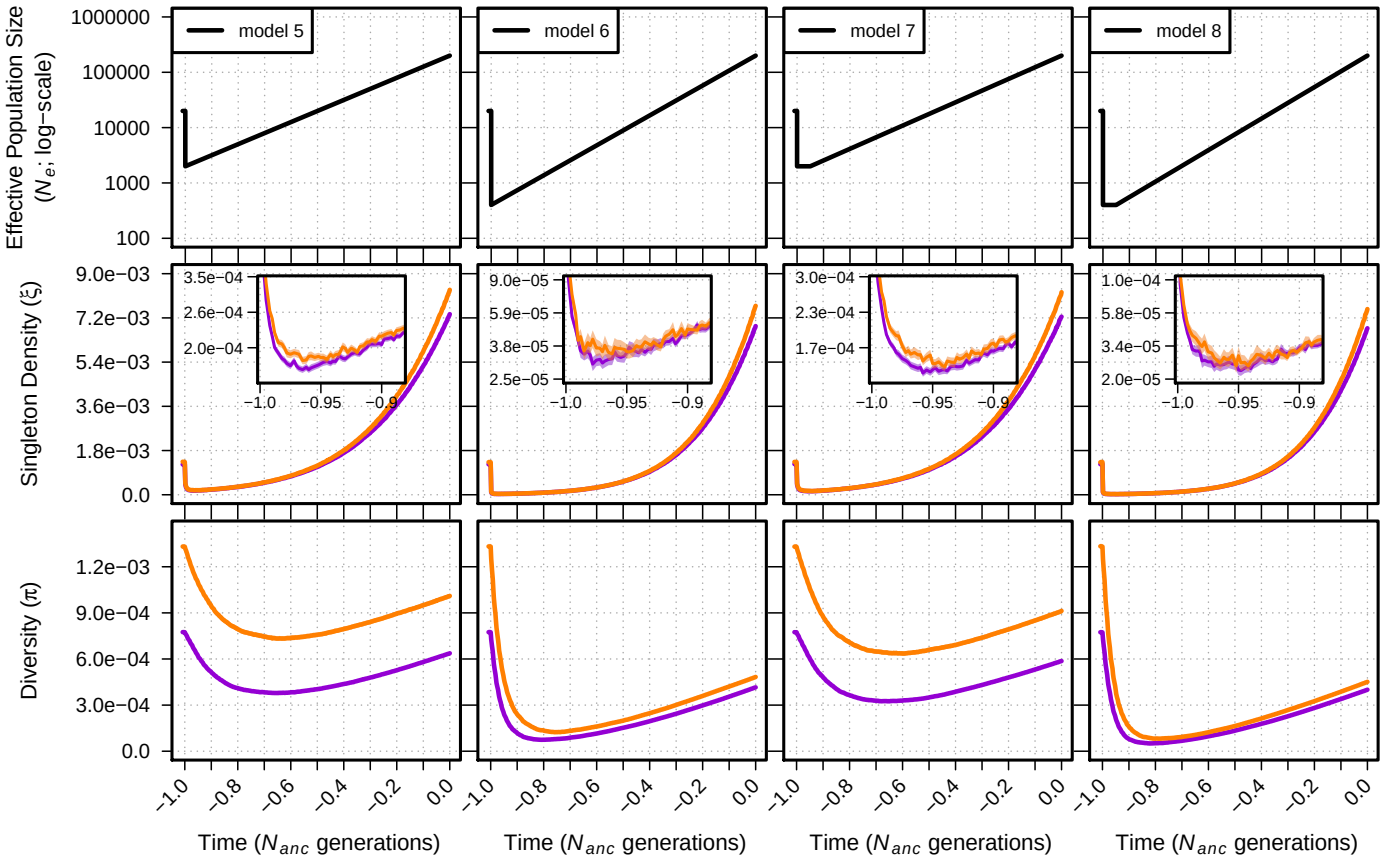


Figure S5 Singleton density (ξ per site) and diversity (π per site) for models 5-8. The top panel shows each demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Diversity statistics are shown for neutral simulations (orange lines) and simulations with BGS (violet lines). Insets show diversity using a log scale for improved detail. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

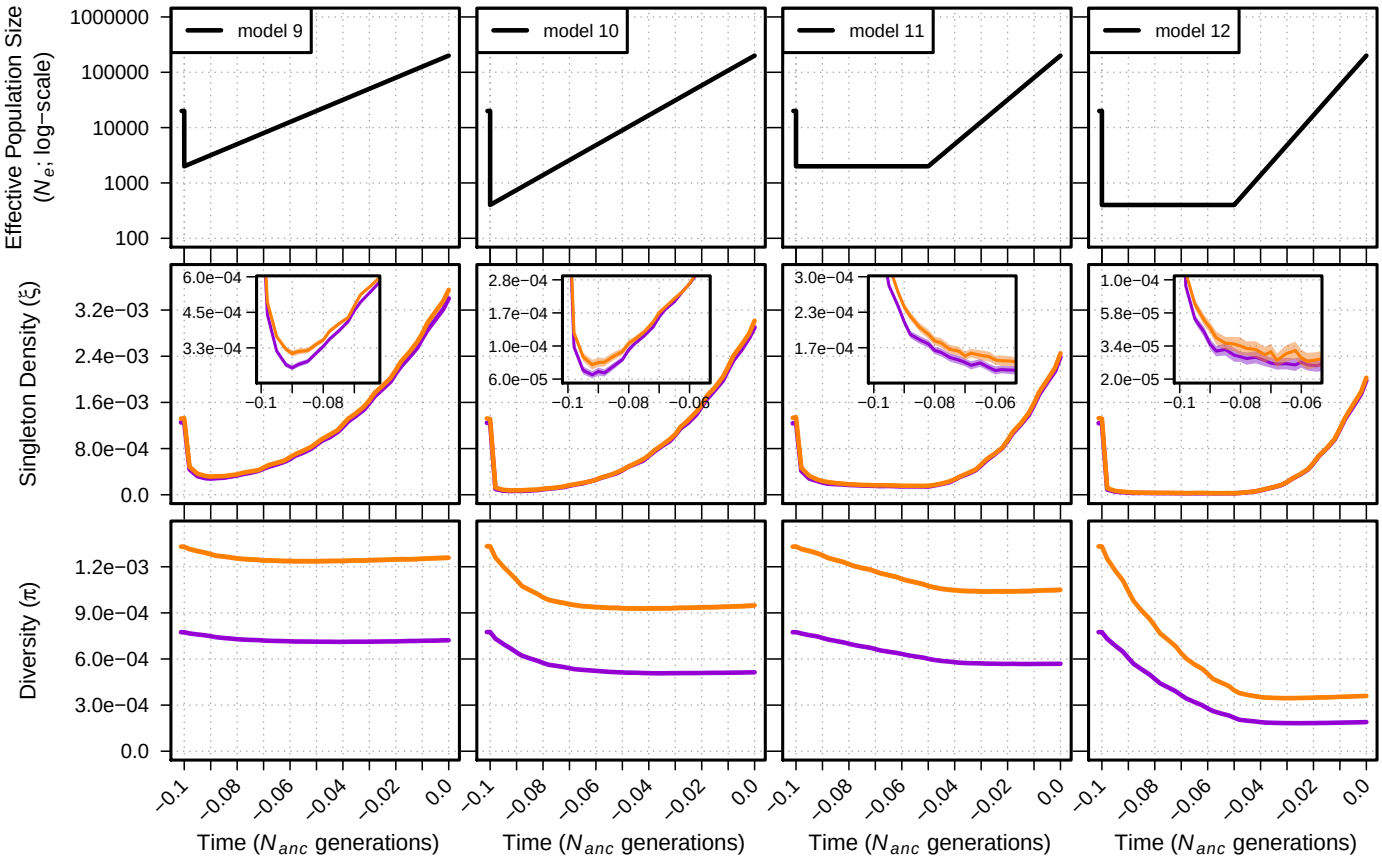


Figure S6 Singleton density (ξ per site) and diversity (π per site) for models 9-12. The top panel shows each demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Diversity statistics are shown for neutral simulations (orange lines) and simulations with BGS (violet lines). Insets show diversity using a log scale for improved detail. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

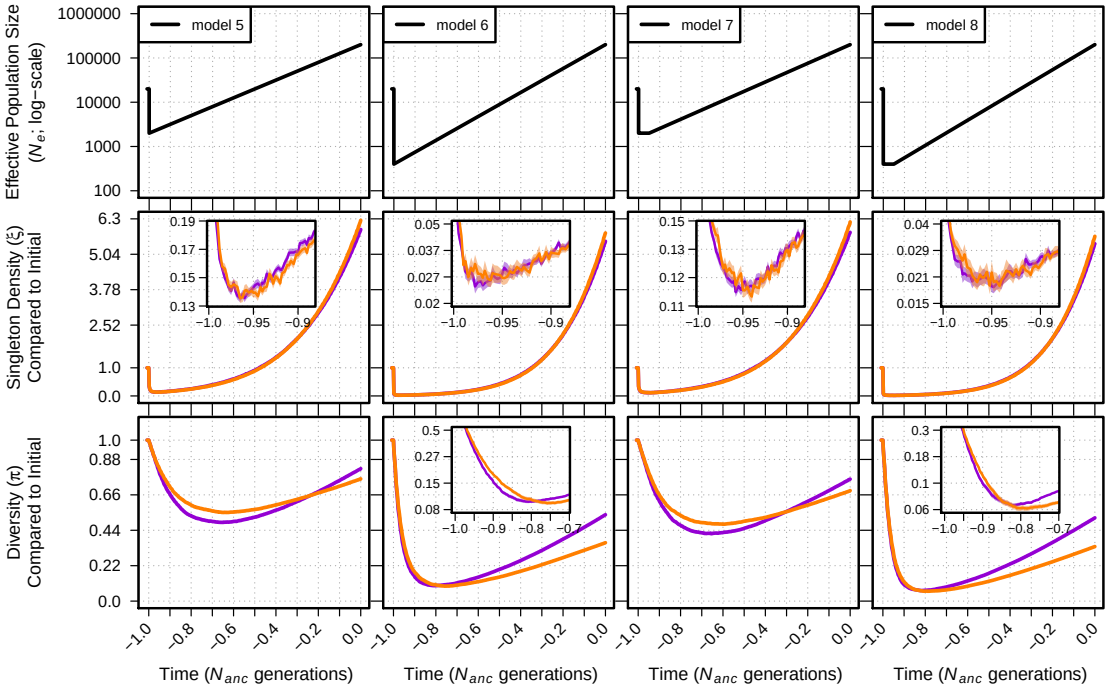


Figure S7 Singleton density (ξ) and diversity (π) relative to the initial generation for neutral (orange) and BGS (violet) simulations of demographic models 5-8. The top panel shows each demographic model as in Supplemental Figure S5. Insets show diversity over a shorter timescale and use a log scale for diversity for improved detail. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data. The data used for this figure is identical to that of Supplemental Figure S5.

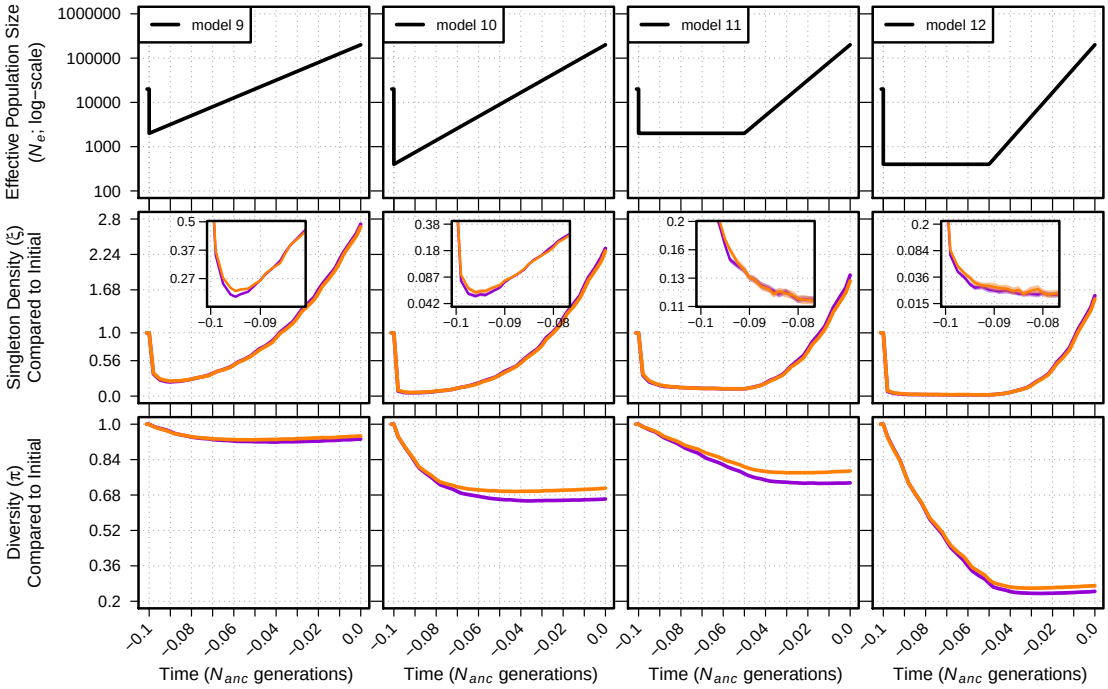


Figure S8 Singleton density (ξ) and diversity (π) relative to the initial generation for neutral (orange) and BGS (violet) simulations of demographic models 9-12. The top panel shows each demographic model as in Supplemental Figure S6. Insets show diversity over a shorter timescale and use a log scale for diversity for improved detail. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data. The data used for this figure is identical to that of Supplemental Figure S6.

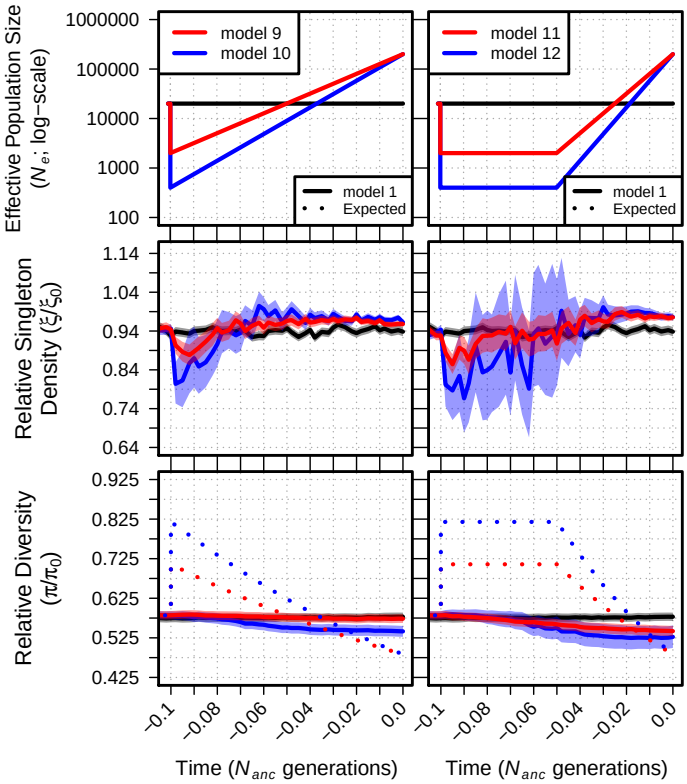


Figure S9 Relative singleton density (ξ/ξ_0) and relative diversity (π/π_0) across time for demographic models 1 and 9-12. The top panel shows each demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Black lines show ξ/ξ_0 and π/π_0 from simulations of a constant sized population (model 1). Dotted lines in the bottom panel show the equilibrium expectation of π/π_0 from Nordborg *et al.* (1996) given the specific selection parameters and the instantaneous N_e at each time point. Envelopes are 95% CIs calculated from 10,000 bootstraps of the original simulation data.

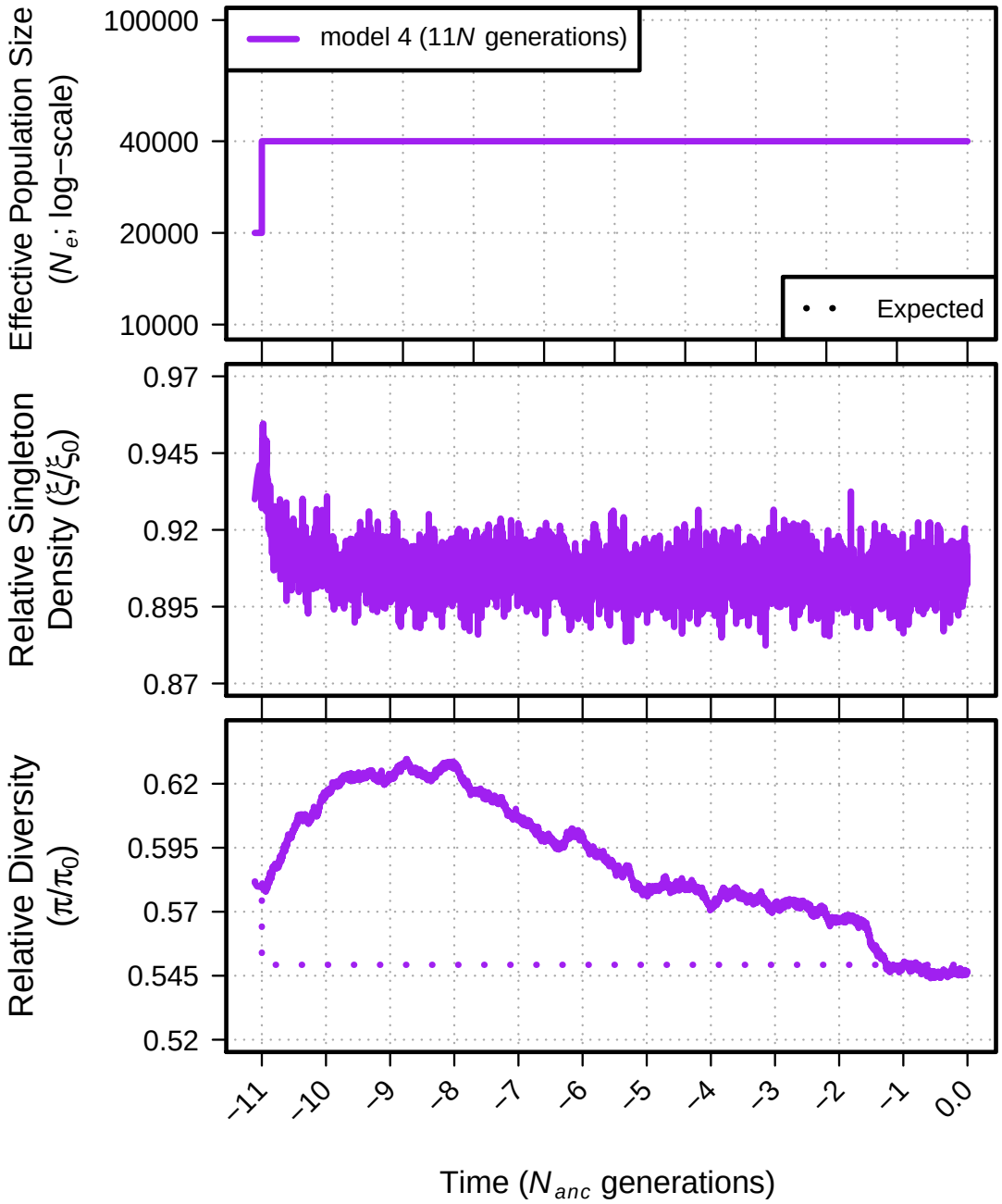


Figure S10 Relative singleton density (ξ/ξ_0) and relative diversity (π/π_0) across time for demographic model 4 over 11 N_{anc} generations. The top panel shows the demographic model; time proceeds forward from left to right and is scaled by the N_e of the population at the initial generation (N_{anc}; 20,000 individuals). Dotted lines in the bottom panel show the equilibrium expectation of π/π_0 from Nordborg *et al.* (1996) given the specific selection parameters and the instantaneous N_e at each time point.

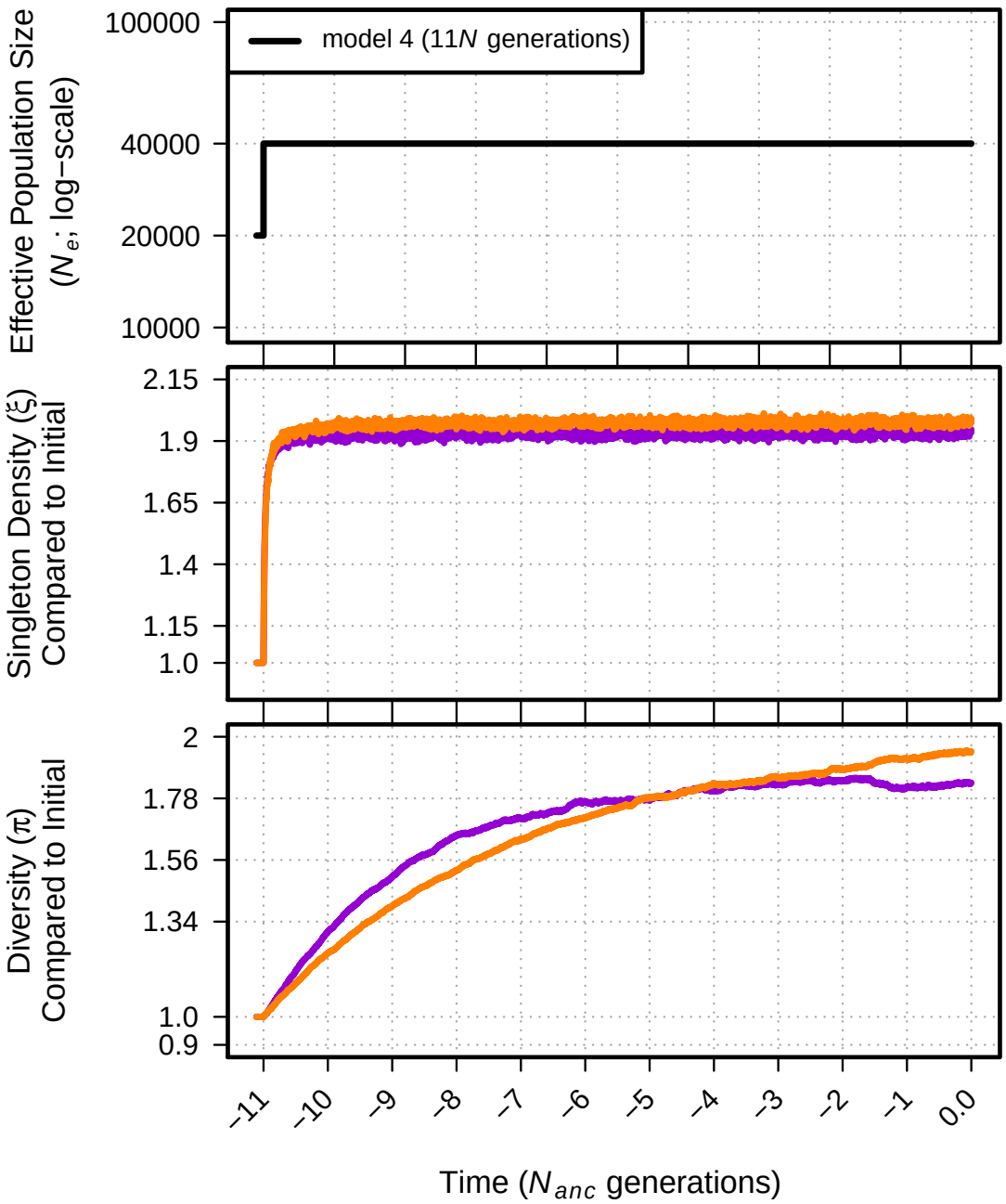
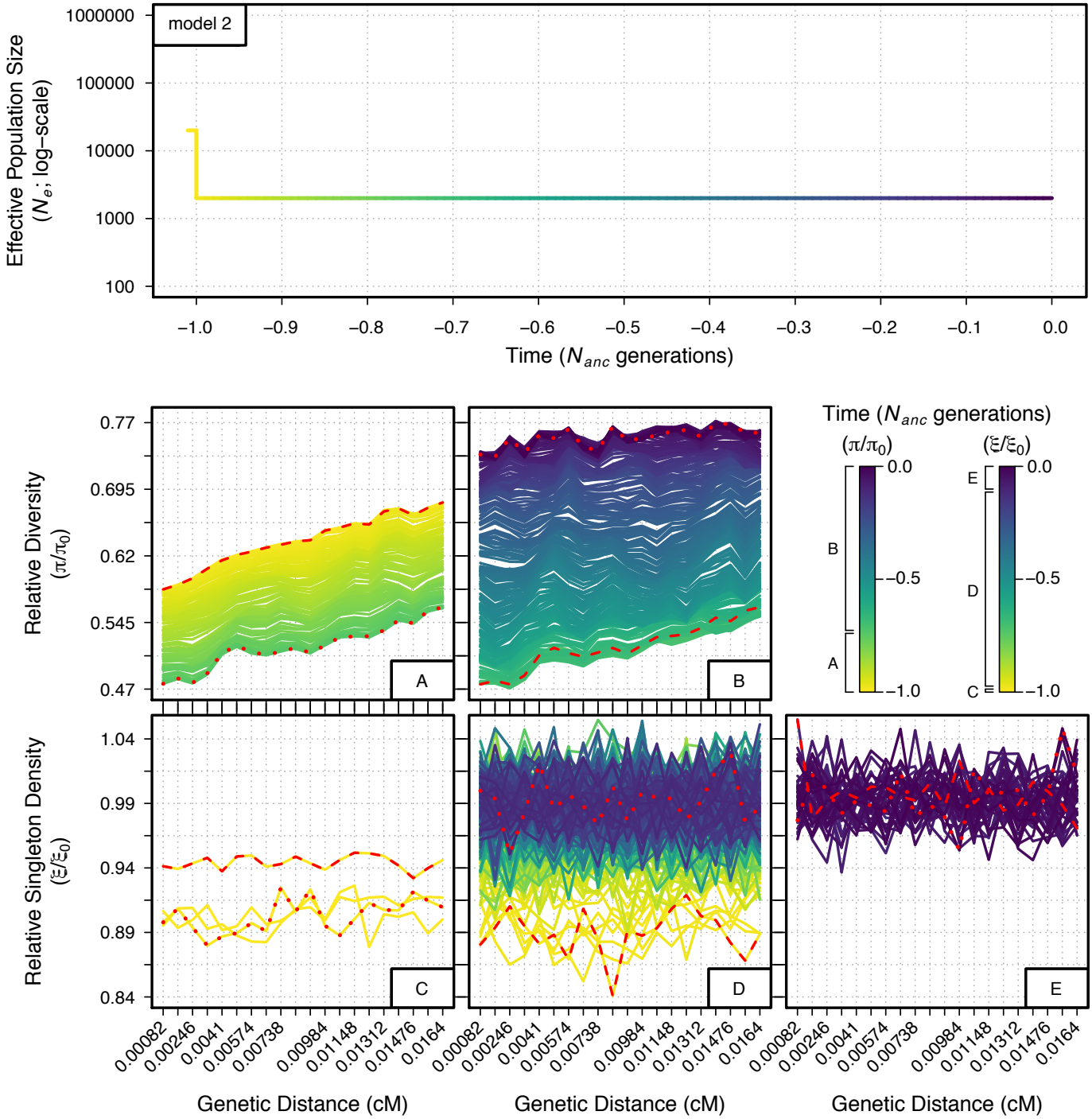
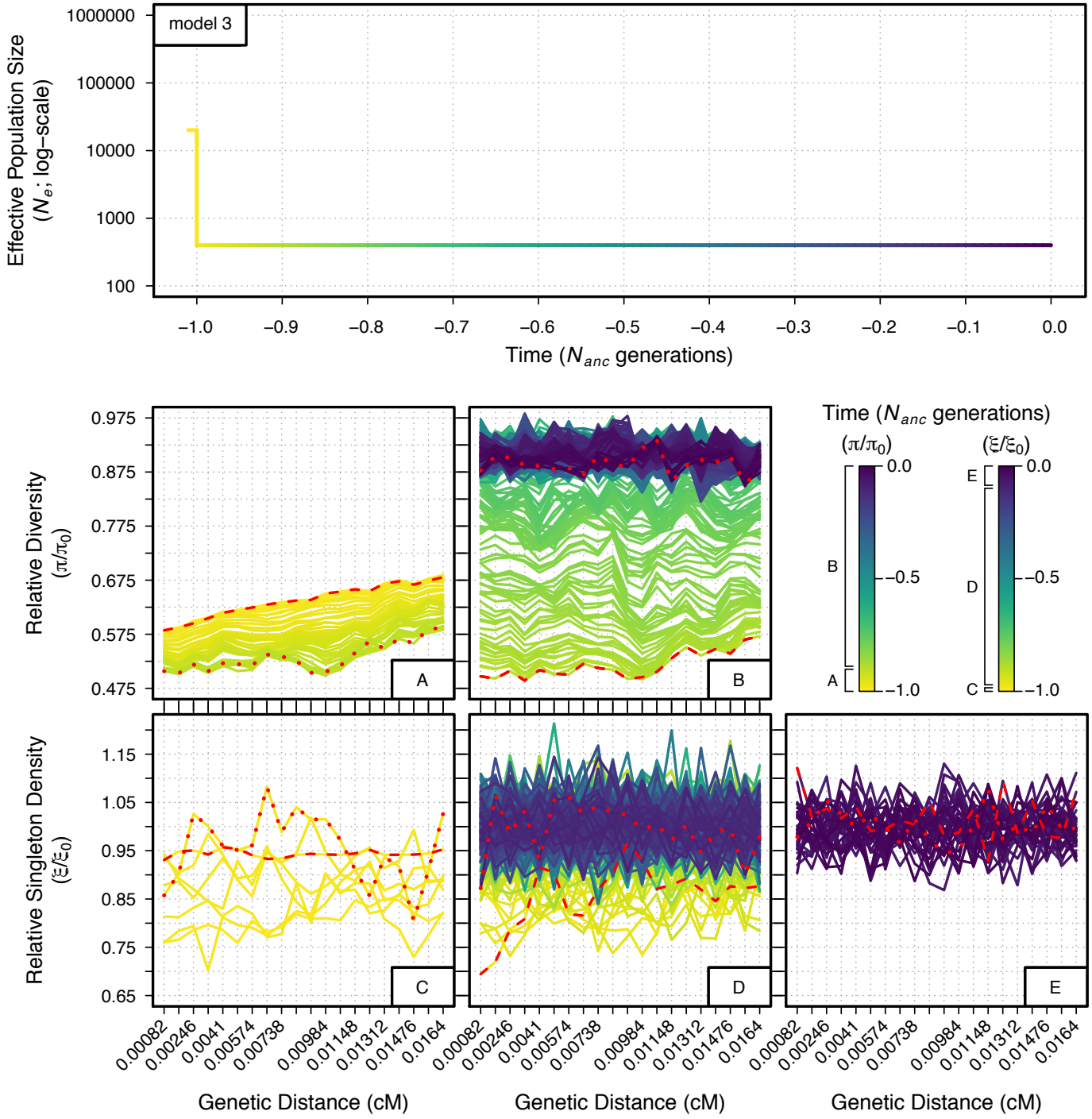
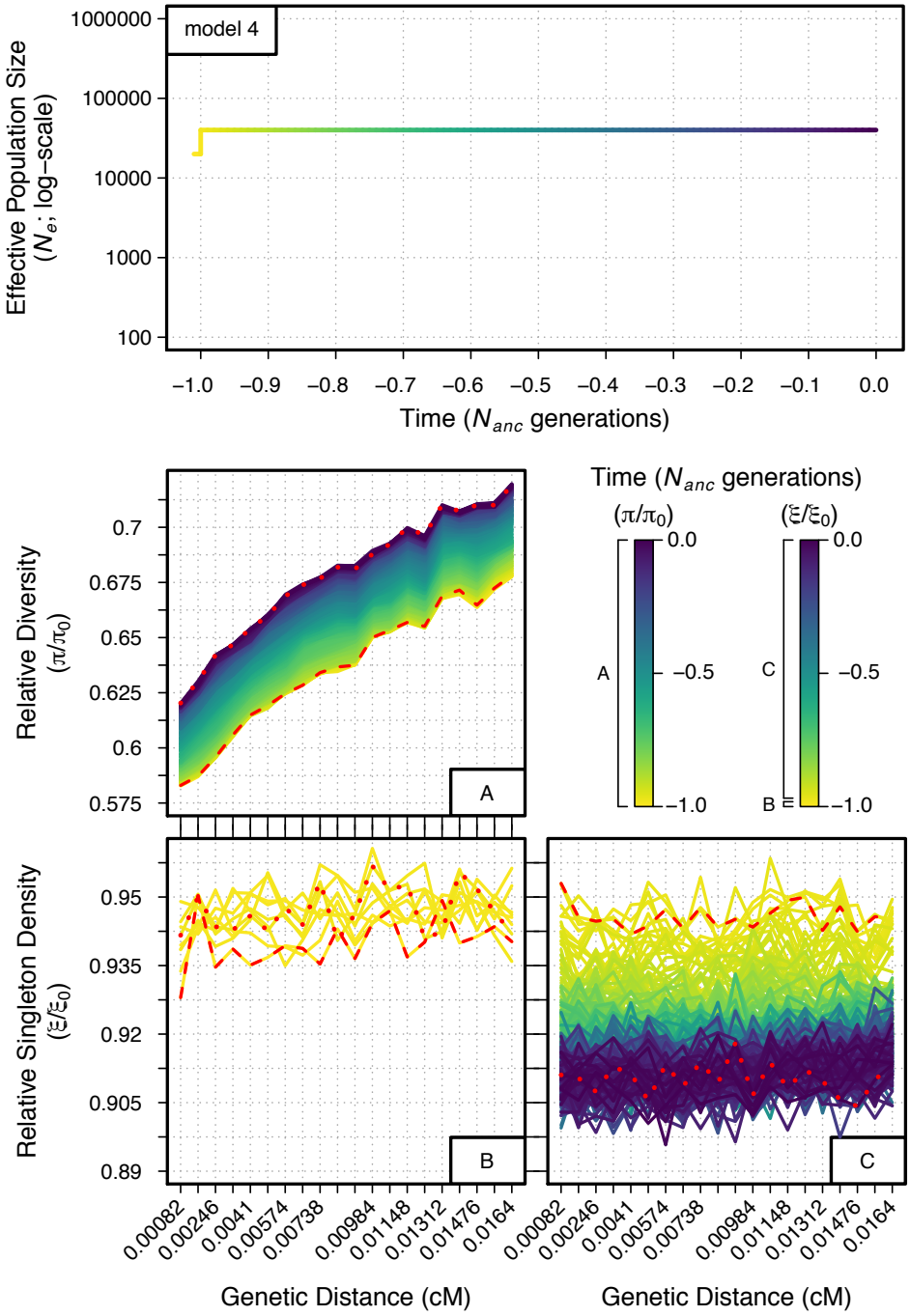
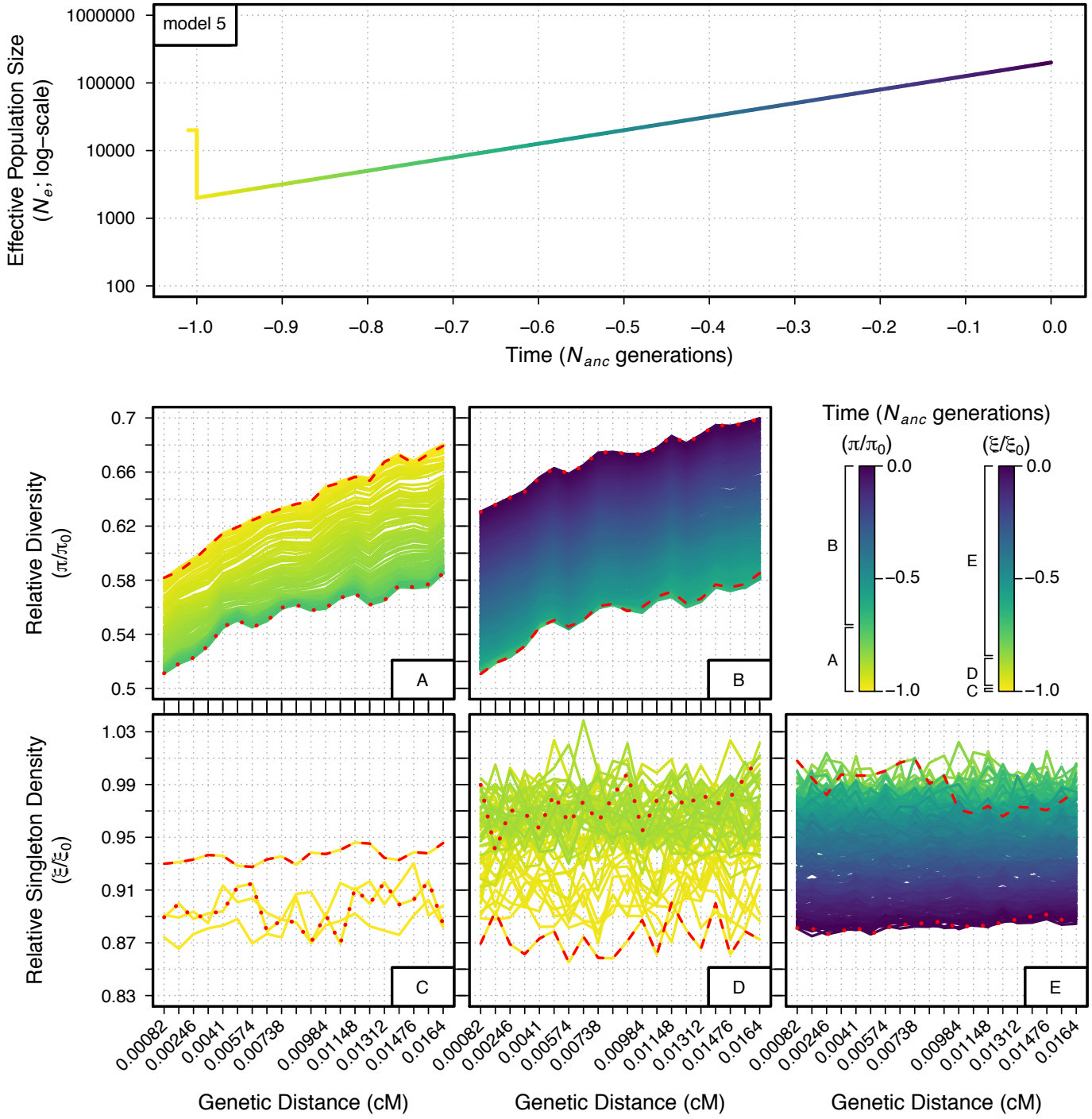


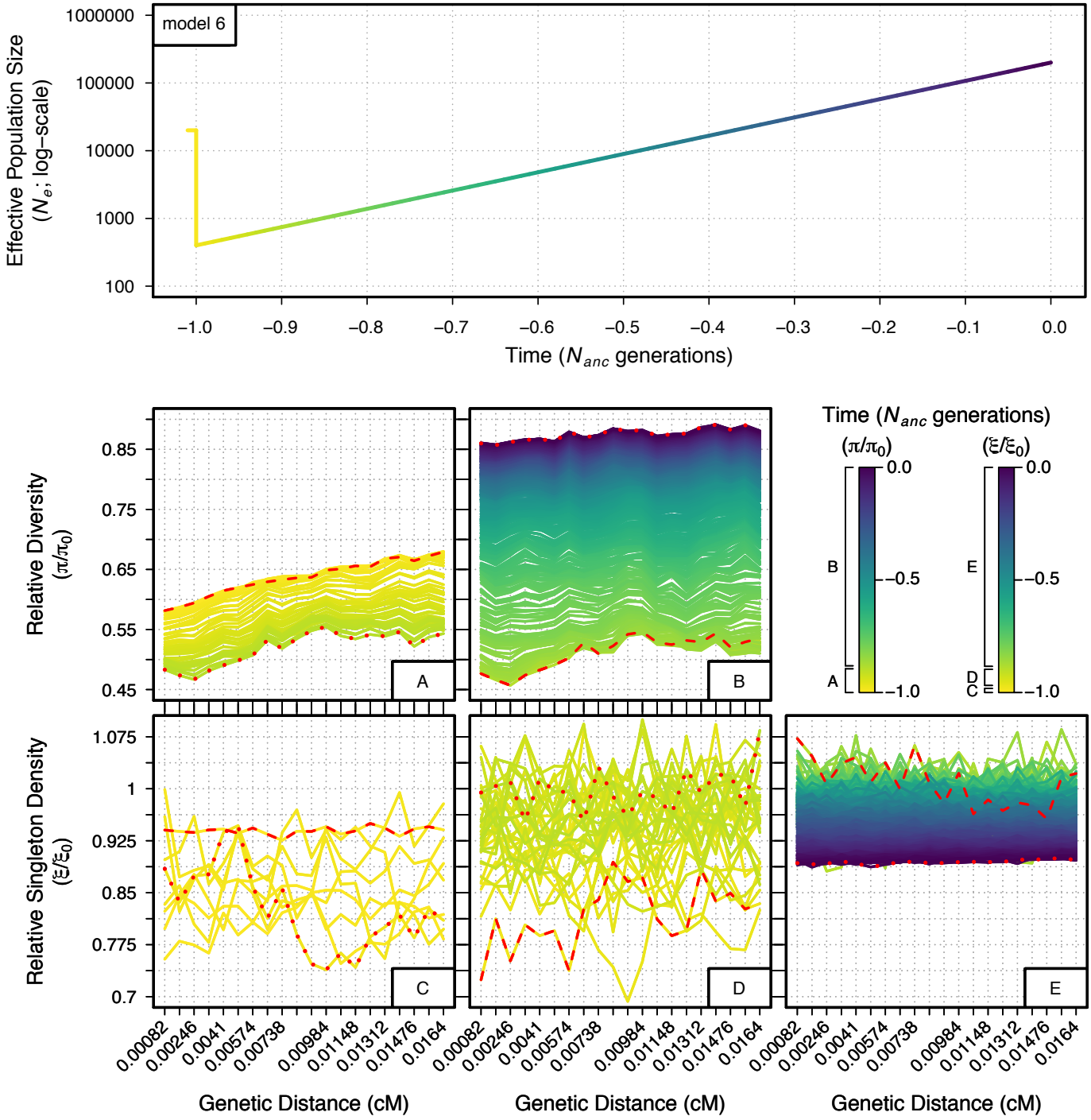
Figure S11 Singleton density (ξ) and diversity (π) relative to the initial generation for neutral (orange) and BGS (violet) simulations of demographic model 4 over 11 N_{anc} generations. The top panel shows the demographic model.

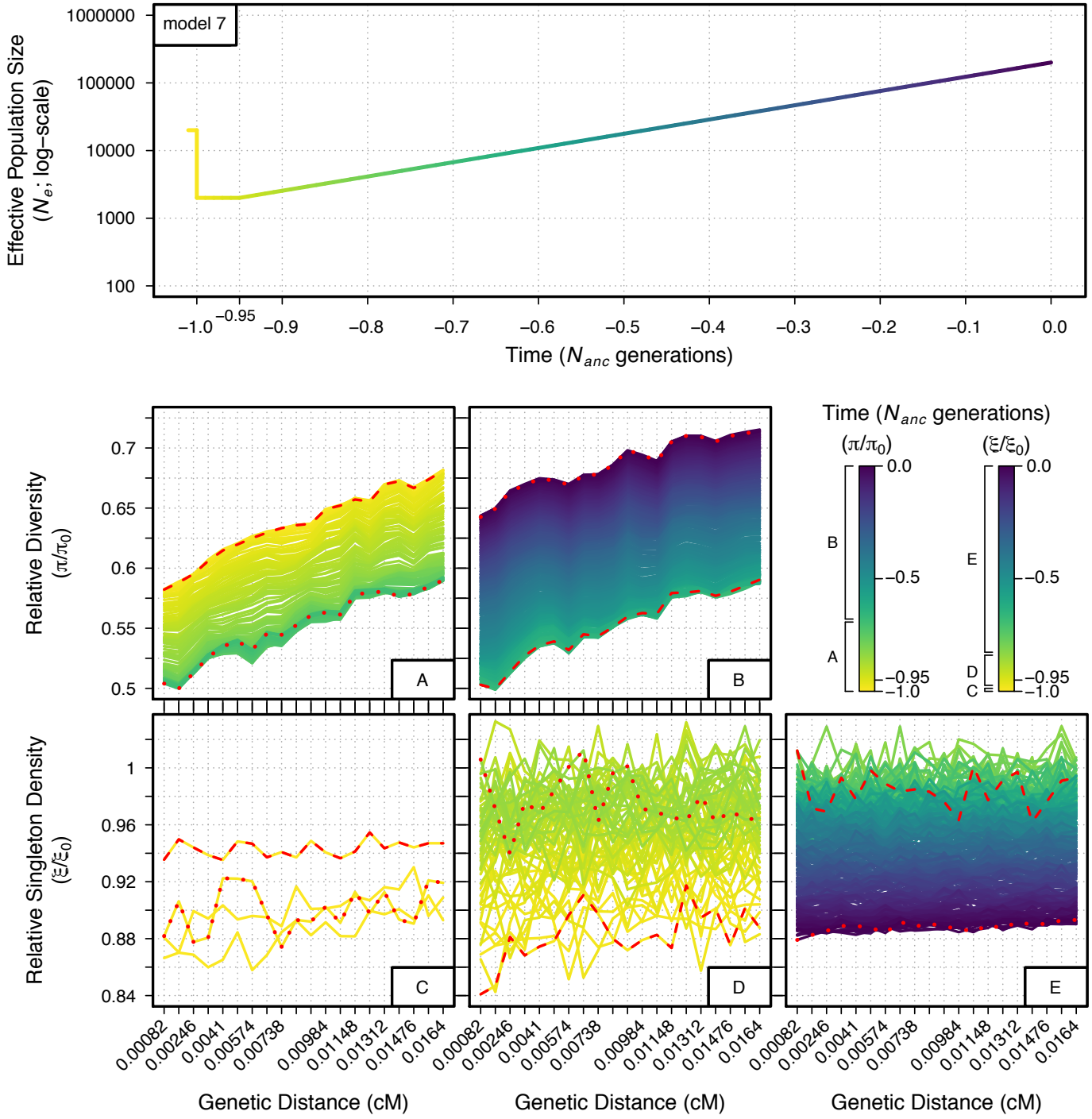

model 2
Effective Population Size
(N_e; log−scale)
1000000
100000
10000
1000
100
−1.0
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.77
0.695
0.62
0.545
0.47
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−1.0
A
B
C
D
E
Relative Singleton Density
(ξ/ξ_0)
1.04
0.99
0.94
0.89
0.84
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)

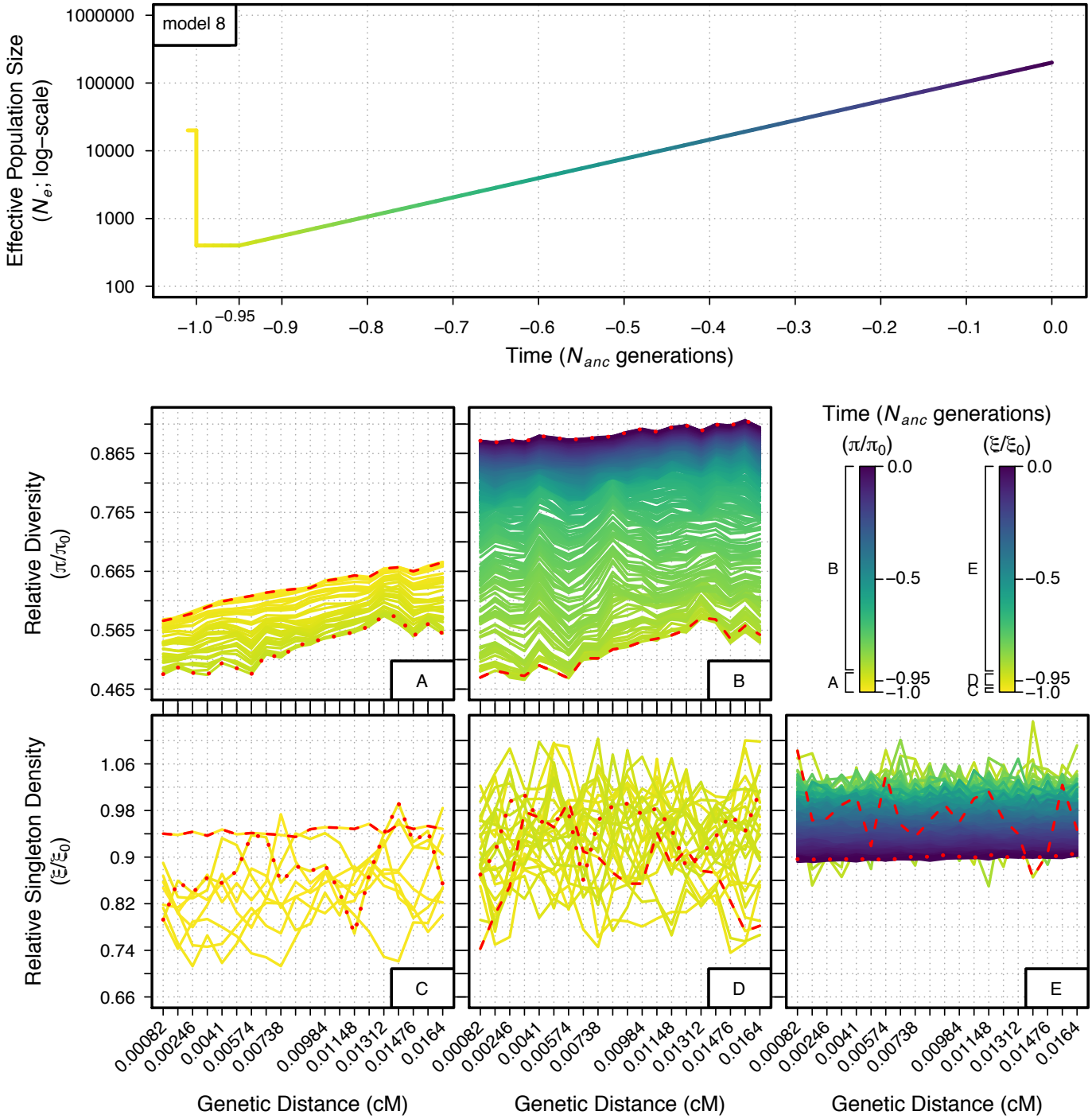

model 3
Effective Population Size
(N_e; log–scale)
1000000
100000
10000
1000
100
–1.0
–0.9
–0.8
–0.7
–0.6
–0.5
–0.4
–0.3
–0.2
–0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.975
0.875
0.775
0.675
0.575
0.475
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
–0.5
–1.0
A
B
C
D
E
Relative Singleton Density
(ξ/ξ_0)
1.15
1.05
0.95
0.85
0.75
0.65
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)

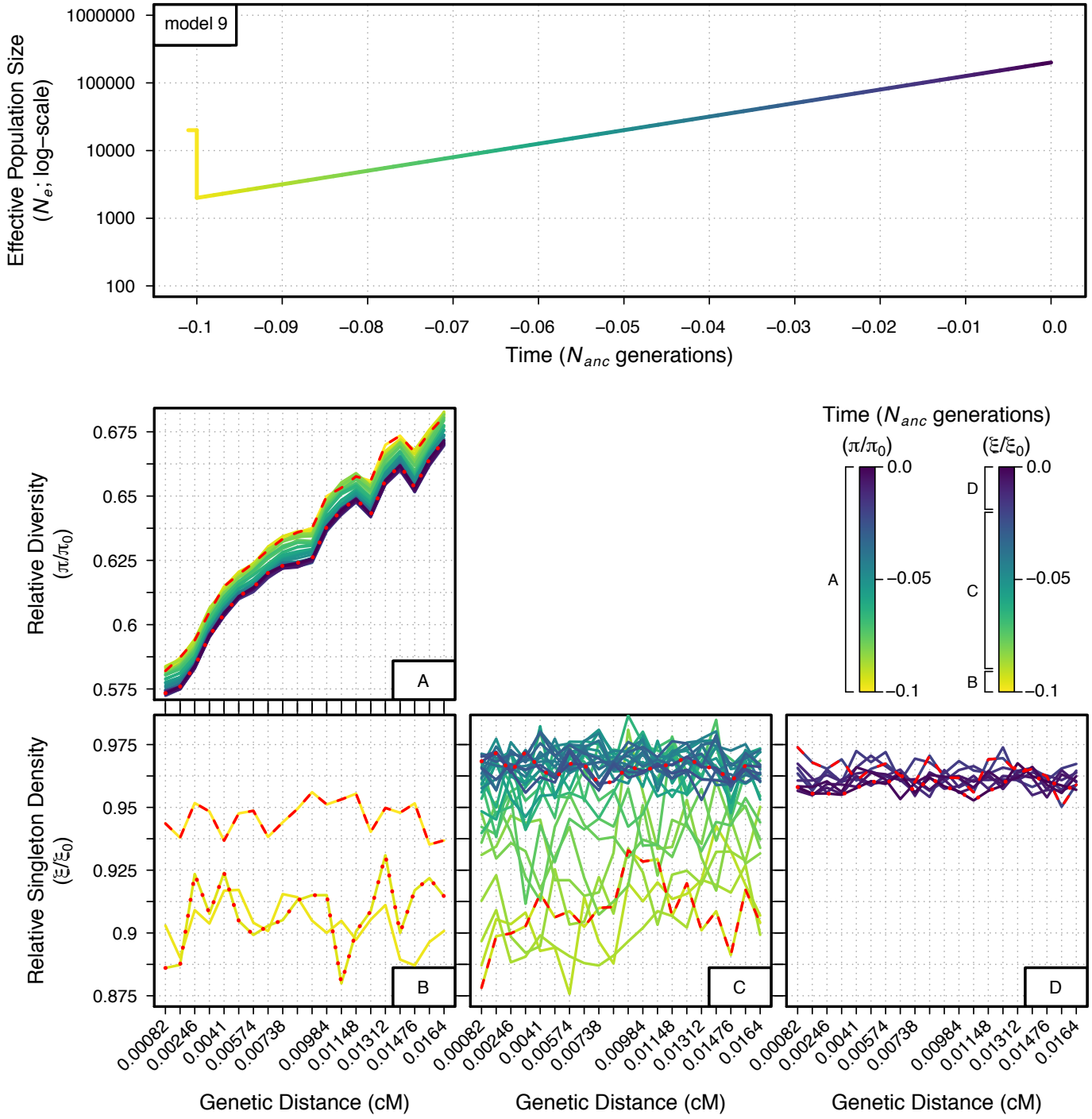

model 4
Effective Population Size
(N_e; log−scale)
1000000
100000
10000
1000
100
−1.0
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.7
0.675
0.65
0.625
0.6
0.575
A
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−1.0
A
C
B
Relative Singleton Density
(ξ/ξ_0)
0.95
0.935
0.92
0.905
0.89
B
C
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)

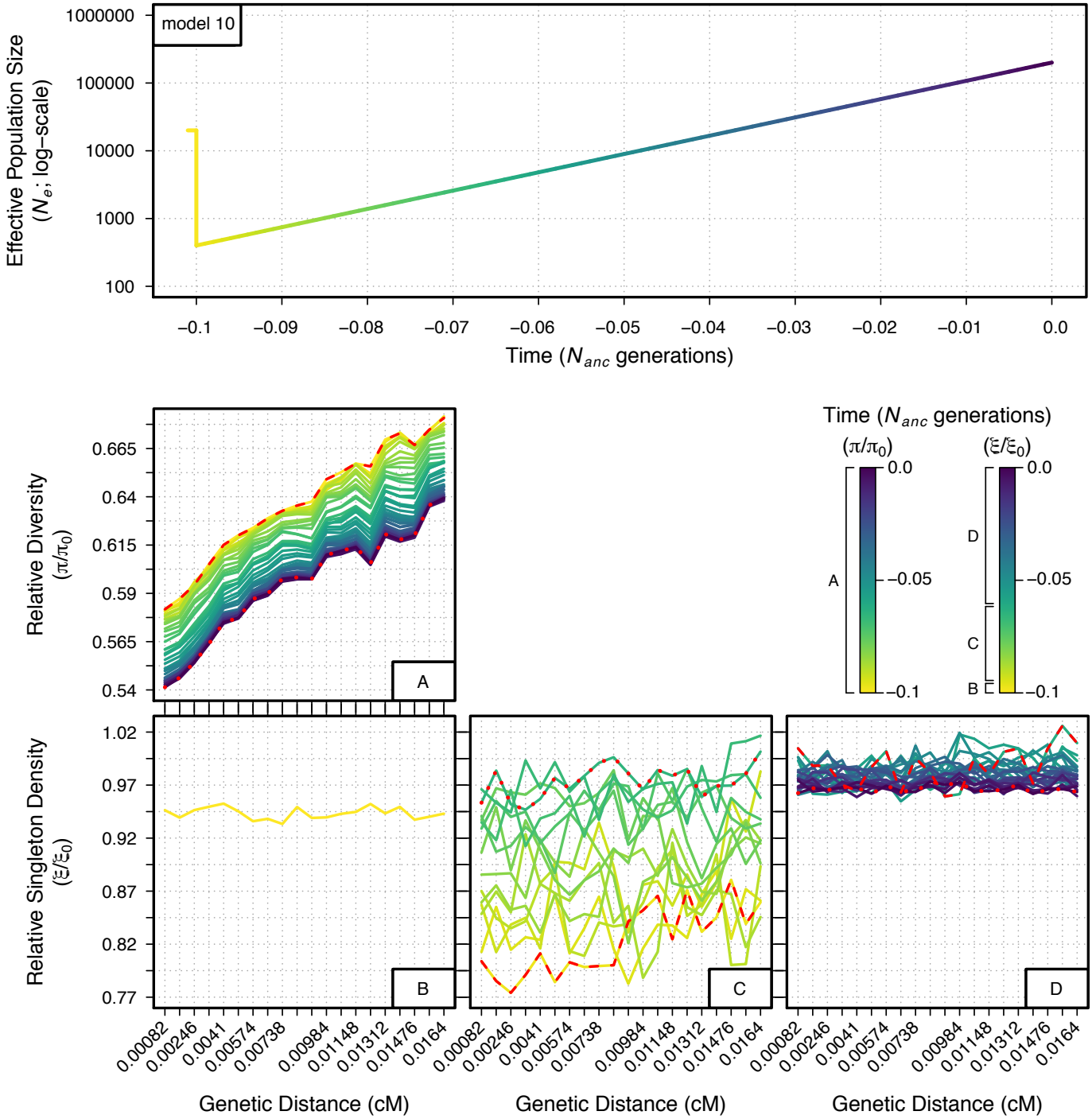

model 5
Effective Population Size
(N_e; log–scale)
1000000
100000
10000
1000
100
−1.0
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.7
0.66
0.62
0.58
0.54
0.5
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−1.0
B
A
E
D
C
Relative Singleton Density
(ξ/ξ_0)
1.03
0.99
0.95
0.91
0.87
0.83
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)

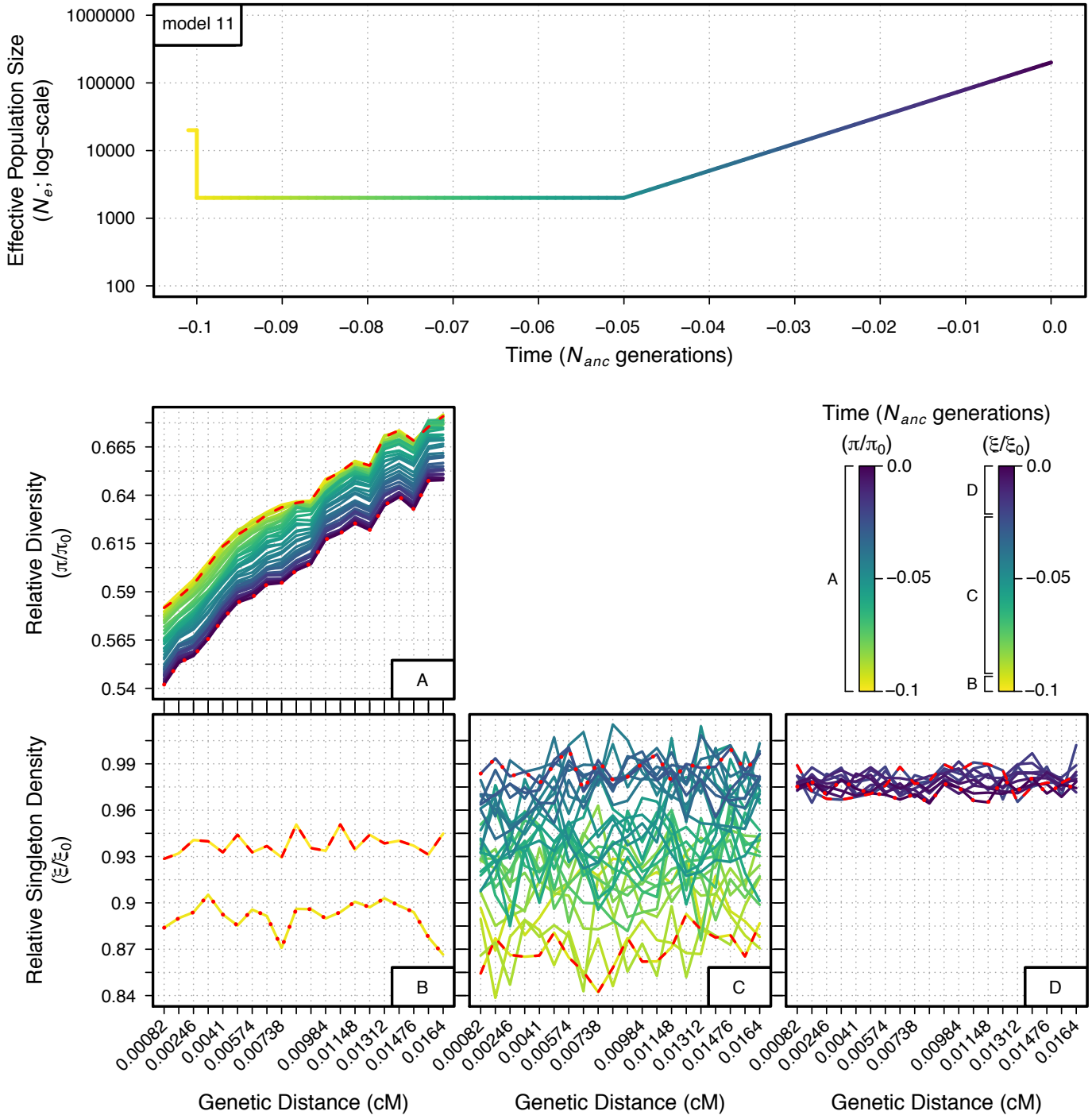

model 6
Effective Population Size
(N_e; log−scale)
1000000
100000
10000
1000
100
−1.0
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.85
0.75
0.65
0.55
0.45
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−1.0
B
A
E
D
C
Relative Singleton Density
(ξ/ξ_0)
1.075
1
0.925
0.85
0.775
0.7
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)


model 7
Effective Population Size
(N_e; log–scale)
1000000
100000
10000
1000
100
−1.0
−0.95
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.7
0.65
0.6
0.55
0.5
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−0.95
−1.0
B
A
E
D
C
Relative Singleton Density
(ξ/ξ_0)
1
0.96
0.92
0.88
0.84
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)


model 8
Effective Population Size
(N_e; log−scale)
1000000
100000
10000
1000
100
−1.0
−0.95
−0.9
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.865
0.765
0.665
0.565
0.465
A
B
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.5
−0.95
−1.0
Relative Singleton Density
(ξ/ξ_0)
1.06
0.98
0.9
0.82
0.74
0.66
C
D
E
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)


model 9
Effective Population Size
(N_e; log–scale)
1000000
100000
10000
1000
100
–0.1
–0.09
–0.08
–0.07
–0.06
–0.05
–0.04
–0.03
–0.02
–0.01
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.675
0.65
0.625
0.6
0.575
A
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
–0.05
–0.1
A
B
C
D
Relative Singleton Density
(ξ/ξ_0)
0.975
0.95
0.925
0.9
0.875
B
C
D
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)


model 10
Effective Population Size
(N_e; log−scale)
1000000
100000
10000
1000
100
−0.1
−0.09
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
0.0
Time (N_{anc} generations)
Relative Diversity
(π/π_0)
0.665
0.64
0.615
0.59
0.565
0.54
A
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
−0.05
−0.1
A
B
C
D
Relative Singleton Density
(ξ/ξ_0)
1.02
0.97
0.92
0.87
0.82
0.77
B
C
D
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)
Genetic Distance (cM)
Genetic Distance (cM)


model 11
Effective Population Size
(N_e; log–scale)
1000000
100000
10000
1000
100
–0.1
–0.09
–0.08
–0.07
–0.06
–0.05
–0.04
–0.03
–0.02
–0.01
0.0
Time (N_{anc} generations)
Time (N_{anc} generations)
(π/π_0)
(ξ/ξ_0)
0.0
–0.05
–0.1
A
B
C
D
Relative Diversity
(π/π_0)
0.665
0.64
0.615
0.59
0.565
0.54
Relative Singleton Density
(ξ/ξ_0)
0.99
0.96
0.93
0.9
0.87
0.84
0.00082
0.00246
0.0041
0.00574
0.00738
0.00984
0.01148
0.01312
0.01476
0.0164
Genetic Distance (cM)

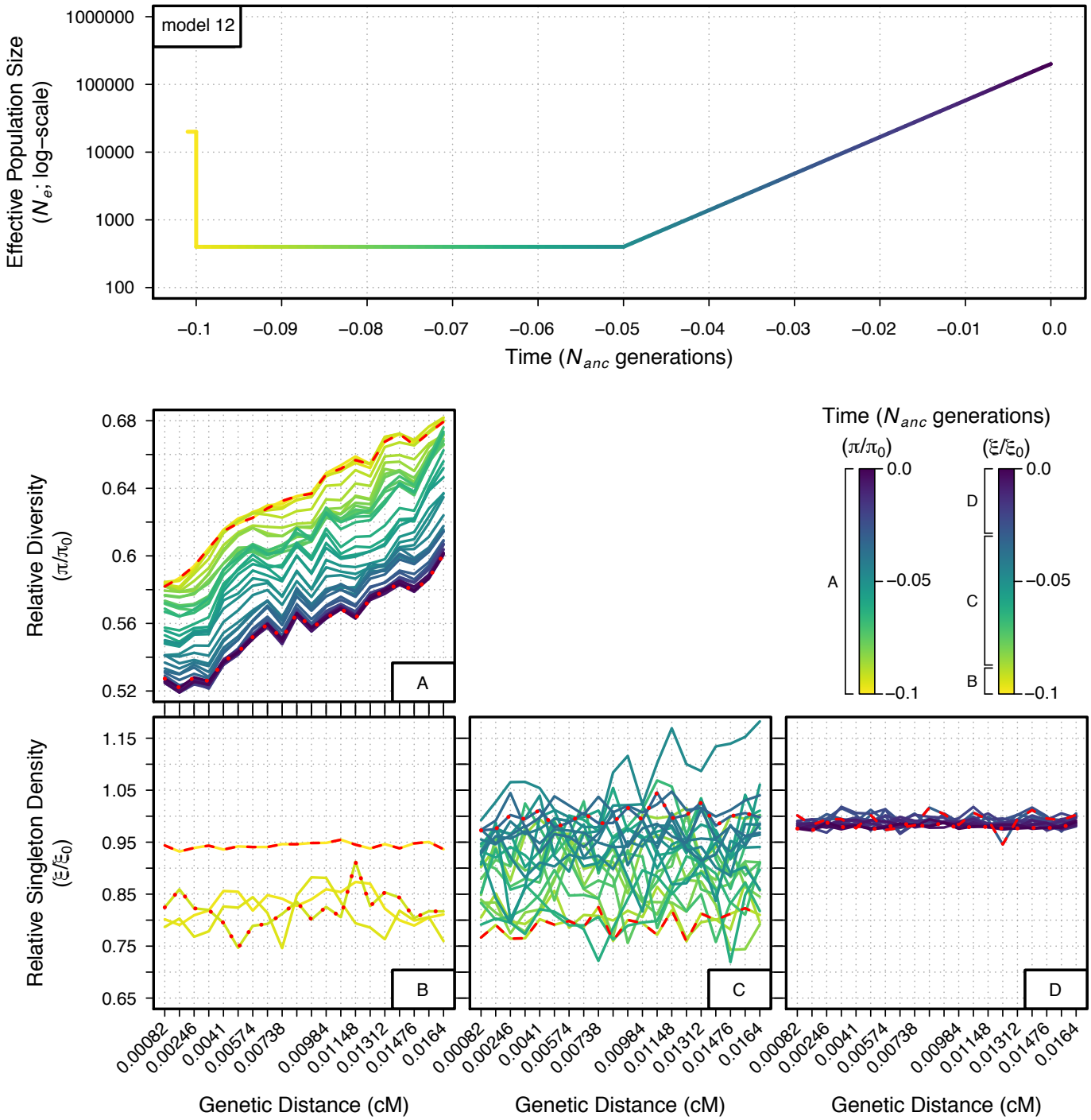


Figure S12 Relative diversity (π/π_0) and singleton density (ξ/ξ_0) through time for demographic models 2-12 measured across a neutral 200 kb region under the effects of BGS. The genetic distance of each 10 kb bin from the selected locus is indicated on the x-axes of the bottom two panels, with genetic distance increasing from left to right. Each line measuring π/π_0 and ξ/ξ_0 across the 200 kb neutral region represents a specific generation of the demographic model (401 discrete generations for demographic models 2-8, 41 discrete generations for demographic models 9-12). Specific generations are indicated by the color of the demographic model at the top of each figure (time is scaled in units of N_{anc} generations [20,000 individuals]) and in the figure legend. When necessary, multiple plots are given for π/π_0 and ξ/ξ_0 in order to prevent overlap of the measurements between generations (see legend for specific generations covered in each plot). Red dashed lines and red dotted lines indicate the first generation and last generation measured, respectively, for each specific plot.