Allele frequency divergence reveals ubiquitous influence of positive selection in *Drosophila*

Jason Bertram[a,1]

[a]Environmental Resilience Institute & Department of Biology, Indiana University, 1001 E 3rd St., Bloomington, IN 47405.

This manuscript was compiled on March 15, 2021

Resolving the role of natural selection is a basic objective of evolutionary biology. It is generally difficult to detect the influence of selection because ubiquitous non-selective stochastic change in allele frequencies (genetic drift) degrades evidence of selection. As a result, selection scans typically only identify genomic regions that have undergone episodes of intense selection. Yet it seems likely such episodes are the exception; the norm is more likely to involve subtle, concurrent selective changes at a large number of loci. We develop a new theoretical approach that uncovers a previously undocumented genome-wide signature of selection in the collective divergence of allele frequencies over time. Applying our approach to temporally-resolved allele frequency measurements from laboratory and wild *Drosophila* populations, we quantify the selective contribution to allele frequency divergence and find that selection has substantial effects on much of the genome. We further quantify the magnitude of the total selection coefficient (a measure of the combined effects of direct and linked selection) at a typical polymorphic locus, and find this to be large (of order 1%) even though most mutations are not directly under selection. We find that selective allele frequency divergence is substantial at intermediate allele frequencies, which we argue is most parsimoniously explained by positive — not purifying — selection. Thus, in these populations most mutations are far from evolving neutrally in the short term (tens of generations), including mutations with neutral fitness effects, and the result cannot be explained simply as a purging of deleterious mutations.

Adaptation | Linked selection | Neutral theory | Effective population size

One of the central problems of evolutionary biology is to delineate the role of natural selection in shaping genetic variation. Most genetic variation consists of neutral mutations which, though having no appreciable effects on fitness, are not free from the influence of selection. When selection acts on non-neutral mutations, neutral mutations that share similar genetic backgrounds can be dragged along for the ride, a process called linked selection (1). The extent to which linked selection influences neutral variation is a major point of contention (2, 3) — one with practical implications because putatively neutral mutations are widely used to infer population demographic history (4) and as a baseline for detecting selection (2, 5). There is also ongoing debate about the particular modes of selection responsible for shaping genetic variation. Negative selection purging the influx of deleterious mutations is probably prevalent (6, 7), but positive selection on rarer advantageous mutations is crucial for adaptive evolution and likely also has a hand in shaping neutral variation (8).

Until recently, the bulk of the evidence entering the above debates rested on patterns of genetic variation measured at single snapshots in time. The interpretation of such evidence is complicated because the prospective signatures of selection are accumulated over an uncertain history during which other confounding processes (e.g. population demography) also shape genetic diversity (5, 9, 10). Crucially, single snapshot data is unable to reveal what the process of selection is doing at any point in time i.e. selectively changing allele frequencies.

A more direct approach is to analyze allele frequency data gathered from the same population at multiple points in time (10). Evolve and resequence (E&R) experiments (11–14) and studies on wild (15) and field populations (16) have identified allele frequency changes associated with rapid phenotypic adaptation. However, determining the full nature and extent of selective allele frequency change has been difficult. Numerous methods exist for inferring selection coefficients from allele frequency time series (10, 17–24), but are only reliable for selection that is strong relative to the intensity of random, non-selective allele frequency change (random genetic drift). This is a major limitation that likely precludes detection of most of the influence of selection. Fitness-relevant traits are often complex (influenced by a large number of genes) and harbor ample genetic variation. Selection on such traits will thus often cause modest allele frequency shifts distributed across many loci rather than be concentrated at a small number of strongly selected loci (25–27). Moreover, even if some genomic regions harbor strongly selected alleles, much of the associated linked selection could be undetectably weak. Thus, resolving the short-term ($\sim$ tens of generations) influence of selection across the genome remains an important challenge (28).

Here we present a new approach to analyze the genome-wide influence of selection using time-resolved allele frequency data. Our approach capitalizes on a distinctive pattern of among-locus temporal allele frequency divergence that to our knowledge has not previously been described. In contrast with

Significance Statement

Natural selection is the process fundamentally driving evolutionary adaptation; yet the specifics of how natural selection molds the genome are contentious. A prevailing neutralist view holds that the evolution of most mutations is essentially random. Here, we develop new theory that looks past the stochasticity of individual mutations and instead analyzes the behavior of mutations across the genome as a collective. We find that selection has a strong non-random influence on most of the *Drosophila* genome over short timescales (tens of generations), including the bulk of mutations that are not themselves directly targeted by selection. We show that this likely involves ongoing positive selection.

JB conceptualized the study, performed the formal analysis and wrote the manuscript.

No competing interests.

[2]To whom correspondence should be addressed. E-mail: jxb@iu.edu

DRAFT

single-locus approaches, this allele frequency divergence is a collective pattern incorporating alleles across the genome. We therefore lose the ability to identify particular loci under selection; in return are able to detect polygenic selective processes that are not detectable with single-locus approaches.

Traditionally the allele frequency variance in a cohort of neutral alleles with initial frequency p is assumed to have the binomial form

$$\mathrm{Var}(\Delta_t p|p) = C_t p(1-p) \qquad [1]$$

where $\Delta_t p$ denotes the change in allele frequency after t generations, and the variance coefficient C_t is independent of p (29, Chap. 3). The allele frequency divergence in Eq. (1) is largely a consequence of random genetic drift. However, selection can also cause allele frequency divergence. The influential effective population size literature argues that C_t is independent of p in a wide variety of circumstances (30). Crucially, a large body of work has attempted to subsume the effects of selection on neutral alleles into the frequency-independent value of C_t, including both the effects of unlinked fitness variation (31, 32), and some manifestations of linked selection (6, 33, 34). The effective population size literature thus views Eq. (1) as a broadly applicable model of neutral allele divergence, simply requiring a tuning of C_t to capture the effects of selection on neutral alleles, at least to a first approximation (3, 30).

Here we show, on the contrary, that linked selection causes among-locus neutral allele frequency variance to deviate from the binomial form Eq. (1), such that the variance coefficient C_t is frequency-dependent. We use this frequency-dependence to detect the presence of selection, analyze its influence on allele frequencies over time and estimate the typical magnitude of selection coefficients across the genome. Applying our approach to E&R and wild *Drosophila* single nucleotide polymorphism (SNP) data we find evidence of strong linked selection affecting most SNPs. We argue that the specific form of frequency-dependence we find implies a substantial role for positive selection.

Results

Neutral evolution implies binomial allele frequency variance. The binomial variance Eq. (1) classically arises in the neutral Wright-Fisher model, which assumes that genetic drift is caused by random sampling of gametes each generation. This yields the variance coefficient $C_t = 1-(1-1/2N)^t$, where N is the (diploid) population size. In its basic form the neutral Wright-Fisher model entails a number of biological simplifications including random mating, constant population size, non-overlapping generations, and the absence of fitness differences between individuals. Many of these assumptions can be relaxed without affecting the binomial form of Eq. (1) — a key finding of the effective population size literature — although typically some form of Wright-Fisher sampling is assumed (30).

To rule out the possibility that deviations from Eq. (1) are caused by neutral processes, we show that binomial variance applies quite generally for neutral alleles unlinked from selected loci (Supplement S1). Our argument is based on the Cannings exchangeable vector model (35), which encompasses a much broader class of genetic drift models allowing for arbitrary offspring number distributions (36). The Cannings model also has binomial variance, but is not strictly applicable to adapting populations because the exchangeability assumption precludes fitness variation between individuals. In Supplement S1 we extended the Cannings model to show that binomial variance still holds in the presence of fitness variation provided that the alleles under consideration are neutral and in linkage equilibrium with alleles under selection. Binomial variance is thus a generic property of neutral allele frequency evolution rather than of particular genetic drift models. Non-binomial variance (equivalently, frequency-dependent C_t) therefore signifies non-neutral evolution, meaning that some form of biased allele frequency change is occurring (migration, selection, biased mutation or gene drive).

Selection creates non-binomial allele frequency variance. We now show that selection will often cause deviations from binomial allele frequency variance. Selection causes allele frequency divergence when its strength or sign varies among alleles in a cohort. To quantify this effect, we consider the expected frequency change of one allele due to selection over one generation. This is given by

$$E[\Delta_1 p|p,s] = sp(1-p), \qquad [2]$$

where $s = (\overline{w}_F - \overline{w}_{NF})/\overline{w}$ is the selection coefficient, $\overline{w}_F$ is the mean fitness of the focal allele, $\overline{w}_{NF}$ is the mean fitness of all other alleles at the same locus, and $\overline{w}$ is population mean fitness. Here s is the "total" selection coefficient that captures the net effect of selection at linked loci and at the focal locus (37).

Next, we apply the law of total variance to $\Delta_1 p$ where s is allowed to vary between loci:

$$\begin{aligned}\mathrm{Var}(\Delta_1 p|p) = {} & E_s[\mathrm{Var}(\Delta_1 p|p,s)] \\ & + \mathrm{Var}_s(E[\Delta_1 p,|p,s]). \qquad [3]\end{aligned}$$

The second term in Eq. (3) represents the deterministic allele frequency divergence created by among-locus variability in s. Using Eq. (2), it can be written as $\sigma^2(s|p)[p(1-p)]^2$, where $\sigma^2(s|p)$ is the (possibly frequency-dependent) variance in total selection coefficients among loci. The presence of the $[p(1-p)]^2$ factor will tend to cause intermediate frequency alleles to have elevated variance relative to the binomial case. Thus, the allele frequency variance created by selection is in general not binomial. A similar $\sigma^2(s|p)[p(1-p)]^2$ frequency dependence has been noted before arising from temporal variability in s (38), but the selective among-locus divergence due to variation in genetic background has not been previously utilized to our knowledge.

To fully specify the contribution of selection to allele frequency divergence, we must also evaluate the first term in Eq. (3). This term represents the variance attributable to random genetic drift averaged over cohorts with different selection coefficients. Up to this point we have not assumed any particular population dynamic model. The first term in Eq. (3) involves an interplay between selection and drift, which does depend on the choice of population dynamic model. For simplicity we assume that the population obeys a Wright-Fisher model with neutral allele frequency variance given by $D_t p(1-p)$ (the relationship between C_t and D_t will become apparent below). Even with this Wright-Fisher simplification, the accumulation of variance over time $\mathrm{Var}(\Delta_t p|p)$ in the presence of selection has a complicated mathematical form

DRAFT

(e.g. (20, 21, 23)). We further simplify by assuming that selection coefficients are small $s < 0.05$, that the total time of divergence is short ($t \sim 10$ generations), and that initial cohort frequencies are not near the $p = 0, 1$ boundaries (major allele frequencies $p < 0.95$).

Under these assumptions, selective changes in allele frequency accumulate approximately linearly over time such that $E[\Delta_t p|p, s] \approx stp(1-p)$. Moreover, the first term in Eq. (3) can be approximated as $E_s[\mathrm{Var}(\Delta_t p|p, s)] = E_s[D_t p_t(1-p_t)] \approx D_t(1 + (1-2p)E_s[st])p(1-p)$, where $p_t = p + stp_t(1-p_t)$ is the expected frequency after t generations (dropping terms of order $D_t s^2$ and smaller). The $(1-2p)E_s[st]$ term represents the perturbation to the drift variance resulting from a mean nonzero bias in the allele cohort (accordingly, this term vanishes at $p = 0.5$ where the binomial variance factor $p(1-p)$ is flat). Thus, in addition to the effect of variation in s among loci discussed above, selection also causes non-binomial frequency variance by introducing a directional bias in allele frequency trajectories.

Combining variance contributions together we have

$$C_t(p) = D_t + D_t(1-2p)E_s[st|p] + p(1-p)\sigma^2(st|p). \qquad [4]$$

The variance coefficient $C_t(p)$ is thus partitioned into a frequency-independent genetic drift component, and two frequency-dependent selective components.

Intermediate frequency alleles have elevated variance in *Drosophila* E&R experiments. In light of the preceding results, we investigated whether binomial allele frequency variance is observed empirically. To check for binomial frequency dependence, we calculated $C_t = \mathrm{Var}(\Delta_t p|p)/p(1-p)$ using SNP allele frequencies measured in two sequencing snapshots, binning SNPs according to their initial frequency p. According to the above theory, if these populations are evolving neutrally then C_t should be independent of p.

In two fruit fly (*D. Simulans*) E&R experiments (11, 12), we observe systematically elevated C_t at intermediate frequencies (Fig. 1), incompatible with neutral evolution. We rule out measurement error as driving this pattern, because the major sources of pooled sequencing error (population sampling, read sampling, unequal individual contributions to pooled DNA) also create binomial variance rather than a systematic frequency-dependent bias (Supplement S2; (39, 40)). Since these E&R populations are closed to migration, we deduce that this pattern is most likely due to selection. Moreover, as will be discussed in the next section, systematically elevated variance cannot be explained by a few large effect loci, implying that a substantial fraction of SNPs across the genome are affected by selection at any one time. Similar results are found in a wild *D. Melanogaster* population (15) (not shown in Fig. 1; see Fig. 2), although in this case the population is not closed and non-neutral allele frequency divergence could also be attributed to migration.

Next we explored how the elevated variance shown in Fig. 1 behaves over time by following its accumulation within a frequency cohort for the two studies in which allele frequencies were measured at more than two timepoints (11, 15). At each measured timepoint we quantified the excess variance using the difference $C_t(p) - C_t(p^*)$, where p is the initial frequency of the cohort and $p^* > p$ is a reference frequency. In practice we choose $p = 0.5$ to maximize the contrast with the divergence

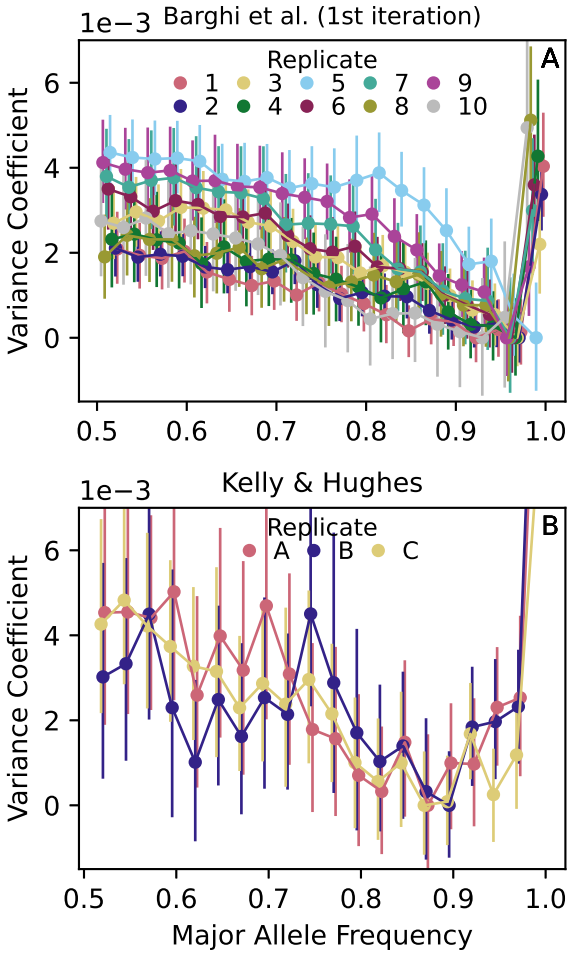

Fig. 1. Intermediate frequency SNPs in E&R *D. Simulans* populations (A (11); B (12)) have systematically elevated variance coefficients $C_t(p) = \mathrm{Var}(\Delta_t p|p)/p(1-p)$ relative to higher frequency SNPs after one round of evolution and resequencing ($t \approx 10$ in A; $t \approx 15$ in B), inconsistent with the binomial expectation for neutrally evolving alleles Eq. (1). $C_t(p)$ is calculated in 2.5% major allele frequency bins using all SNPs in the genome (circles). Vertical lines show 95% block bootstrap confidence intervals (1Mb blocks). We subtract the constant $\min_p C_t(p)$ from $C_t(p)$ in each replicate to prevent differences in the overall magnitude of $C_t(p)$ between replicates from obscuring p dependence within each replicate.

at the reference frequency, while $p^* \sim 0.8 - 0.9$ is chosen to be large enough that there is a meaningful contrast with $p = 0.5$ but safely displaced from the $p = 1$ boundary where allele frequency variances are not measured reliably (e.g. see sharp increases in Fig. 1 as $p \to 1$).

We find that excess variance accumulates over the course of the entire Barghi et al. (11) E&R experiment (Fig. 2A shows one replicate, other replicates are similar; Fig. S1), implying a sustained, polygenic selective divergence in allele frequencies. This pattern is consistent with the positive Δp temporal autocovariances documented by (28). Sustained divergence is intuitively what we might expect to occur in a novel but constant laboratory environment.

By contrast, excess variance in wild *D. Melanogaster* populations (15) does not exhibit continual accumulation of excess variance over time, with fluctuations evident in each cohort (Fig. 2B). Fluctuations imply a concurrent reversal in the direction of non-neutral allele frequency change across many loci

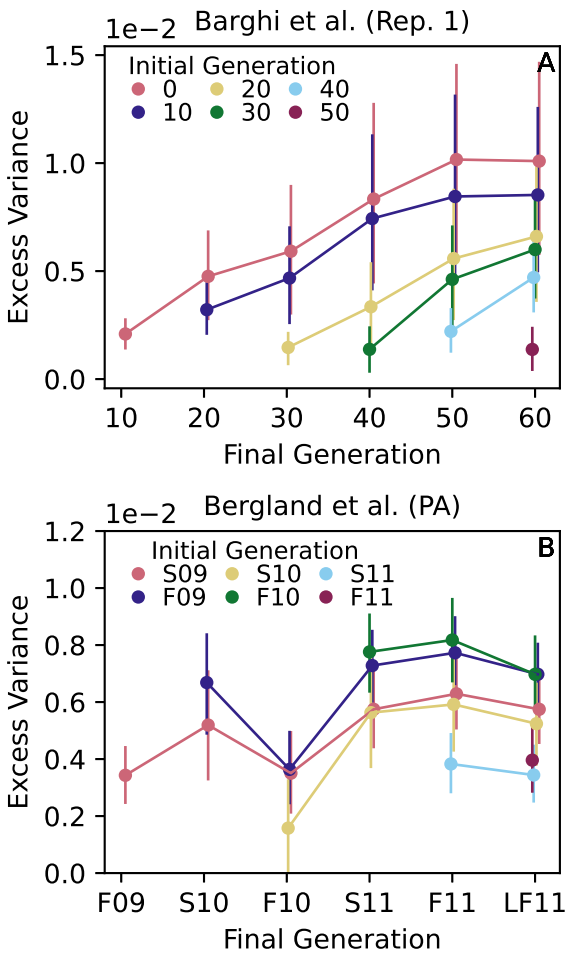

Fig. 2. Excess allele frequency variance (a measure of deviation from neutrality defined as $C_t(p) - C_t(p^*)$) accumulates over time in a *D. Simulans* E&R experiment (A; (11)), but remains relatively flat in a wild *D. Melanogaster* population (B; S=Spring, F=Fall, LF=Late Fall; 09=2009 etc.; (15)). The excess variance is calculated for alleles falling within a 2.5% major allele frequency bin at $p = 0.5$ (i.e. $0.475 < p < 0.525$), and similarly for the reference variance $C_t(p^*)$ with $p^* = 0.9$ (A) and $p^* = 0.8$ (B). Vertical lines show 95% block bootstrap confidence intervals (1Mb blocks).

such that non-neutral divergence is partly lost to a subsequent coordinated non-neutral convergence. Bearing in mind that migration may contribute to this pattern, the fluctuations shown in Fig. 2 are compatible with temporally fluctuating selection affecting a large proportion of the genome, as proposed by (15). However, while Ref. (15) attributed temporal fluctuations in selection to periodic seasonal change, we do not see a clear annual periodicity in the accumulation of variance. A similar lack of annual periodicity is found with a temporal autocovariance approach (28). These results suggest that a more complex selective (or migratory) regime may be involved of which seasonal fluctuations are only a part.

Linked selection strongly perturbs SNP frequencies in *Drosophila*. We next estimated the typical magnitude of the total selection coefficient s driving allele frequency divergence. Since s can be positive or negative, the mean $E[s]$ across loci is a poor measure of this magnitude. Indeed, we expect s to be symmetrically distributed about $s = 0$ at intermediate major allele frequencies $p \to 0.5$, implying $E[s] = 0$ (Figs. S2-S4 show empirical $E[\Delta p|p]$ curves from (11, 12, 15), which are consistent with this expectation). We therefore used the dispersion of s, as quantified by the standard deviation $\sigma(s|p)$, as our measure of the typical intensity of selection (note that because $E[s|p \to 0.5] = 0$ we have $\sigma(s|p \to 0.5) \approx \sqrt{E[s^2]}$).

We estimate $\sigma(s|p)$ from measured allele frequency divergence using Eq. (4). This requires removing the non-selective divergence contributions of genetic drift D_t and measurement error. In Materials and Methods we show that these latter contributions are cancelled out in the excess variance $C_t(p) - C_t(p^*)$, avoiding the complication of independently estimating them. However, some selective divergence is also cancelled out in the difference $C_t(p) - C_t(p^*)$, so that this approach only obtains a lower bound

$$\sigma^2(s|p) > \frac{C_t(p) - C_t(p^*)}{t^2 p(1-p)}. \qquad [5]$$

In all three *Drosophila* studies we evaluated, we find the above lower bound to be of order 10^{-4} (Fig. 3). This implies that total selection coefficients with magnitudes of order $\sigma(s|p) \sim 1\%$ are commonplace in both laboratory and wild *Drosophila* populations.

Intuitively, the standard deviation $\sigma(s|p)$ measures the intensity of a collective "polygenic" adaptive response shared across many loci. A substantial fraction of the loci in a cohort must have nonzero s in order for there to be a $\sigma(s|p)$ signal: if a fraction f of loci have $s = 0$, then $\sigma(s|p) = \sqrt{1-f}\,\sigma_{\rm nn}(s|p)$ where $\sigma_{\rm nn}(s|p)$ is the standard deviation in s among non-neutral loci. Thus, unrealistically strong selection would be needed for a few loci ($f \approx 1$) to drive a $\sigma(s|p)$ signal.

Positive selection explains consistently-positive excess variance. In the preceding sections we showed that selection strongly influences temporal allele frequency divergence throughout the *Drosophila* genome. In principle this could involve a mix of positive and negative selection, although *a priori* we might expect that the presence of such pervasive selection is more easily explained by negative selection acting continually to purge deleterious mutations. However, it is not clear how a negative selection scenario can be reconciled with the observed pattern of elevated variance coefficients at intermediate frequencies (Fig. 1). The effects of negative selection should be substantially stronger at high major allele frequencies, because deleterious mutations strong enough to cause detectable allele frequency divergence are unlikely to reach high frequencies.

To better understand the processes generating excess variance we performed forward-time population genetic simulations of negative and positive selection scenarios. We find that negative selection indeed does not consistently generate positive excess variance (Fig. 4A), because $\sigma^2(s|p)$ rapidly increases with p such that the overall selective divergence term $\sigma^2(s|p)p(1-p)$ in Eq. (4) is independent of frequency — just like a neutral allele frequency divergence. The effects of positive selection on the other hand are elevated at intermediate frequencies leading to consistently positive excess variance (Fig. 4A,C). Some amount of positive selection therefore seems to be essential to explain the elevated divergence in Figs. 1-3.

Discussion

Several lines of evidence support the view that selection strongly influences genetic variation in *Drosophila* (8, 12, 28,

DRAFT

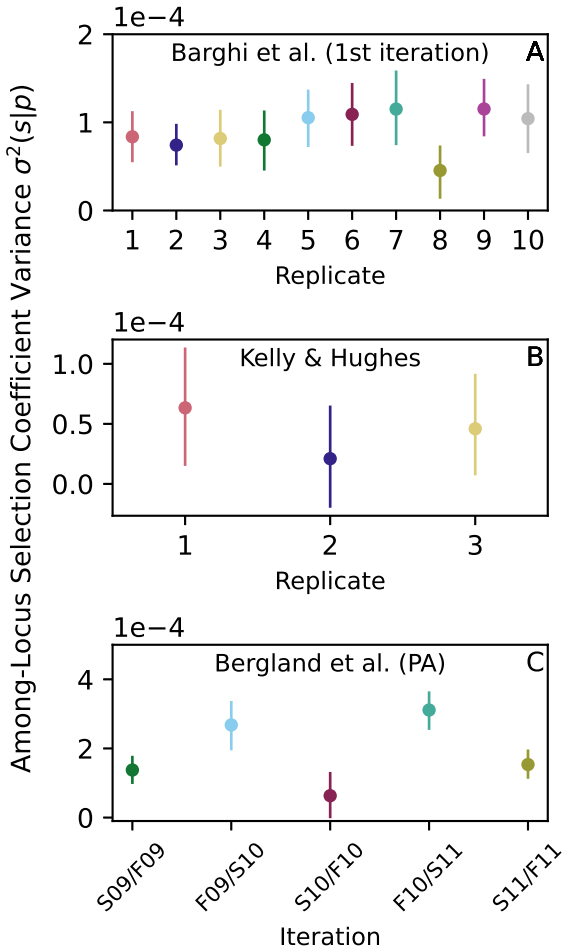

Fig. 3. Total selection coefficients show substantial among-locus variance in *Drosophila*. (A-C) Lower bound estimates of $\sigma^2(s|p)$ calculated from Eq. (5) (circles; vertical lines show 95% block bootstrap confidence intervals) are of order 10^{-4}, which implies typical s values of $\sim 1\%$. Following the original studies (11, 12, 15), we assume $t = 10$ (A); $t = 15$ (B) and $t = 10$ (C; for both summer and winter).

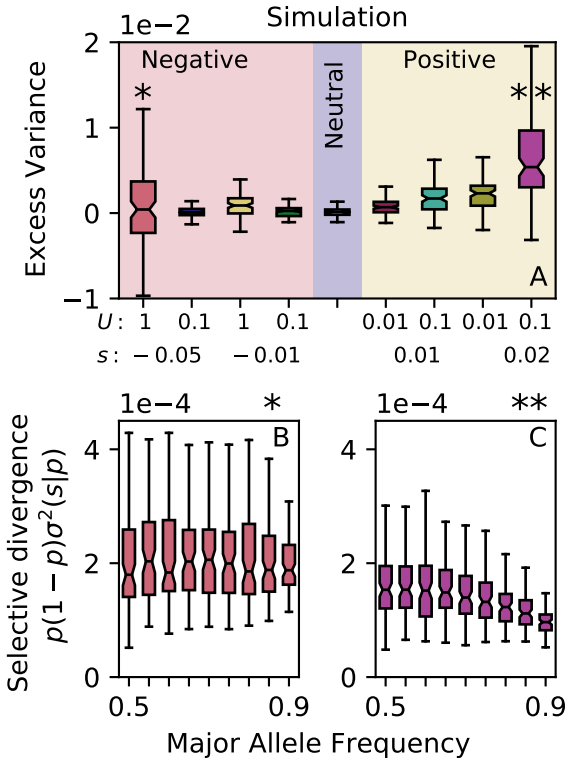

Fig. 4. (A) Forward-time population genetic simulations only consistently show elevated excess variance under positive selection. Excess variance defined as $C_t(p) - C_t(p^*)$ with major allele frequencies $0.5 < p < 0.55$ and $0.9 < p^* < 0.95$ and $t = 10$ generations. (B) Under strong negative selection (deleterious mutation rate $U = 1$/genome/generation, mutation selection coefficient $s = -0.05$), total selection coefficients are substantial at all frequencies but much stronger for high major allele frequencies resulting in a frequency-independent overall selective divergence $p(1-p)\sigma^2(s|p)$ like the neutral case. (C) In contrast, the selective divergence $p(1-p)\sigma^2(s|p)$ shows clear frequency dependence under positive selection, thus producing excess variance at intermediate frequencies. Population size $N = 1000$; 100 replicates per parameter combination.

41). Our results independently show that even over a short time interval (tens of generations), most intermediate frequency SNPs are influenced by selection mostly as a result of linked selection — total selection coefficients of $|s| \sim 1\%$ are the norm among intermediate frequency SNPs, including neutral alleles.

Yet, perhaps surprisingly given the extent of selection involved, it is only recently that genome-wide signals of selection of this sort have started to be documented (28, 42). It has even been argued that the usefulness of temporal data is severely curtailed by the power of random genetic drift to swamp biological plausible selection coefficients (10). The power of our approach stems from aggregating allele frequency behavior from many loci, thereby leveraging the sheer number of variants measured with whole-genome sequencing to discern a selective signal. Heuristically, the sampling error in the lower bound estimate Eq. (5) is proportional to $1/\sqrt{L}$ where L is the number of independent loci used to estimate $C_t(p)$ (Materials and Methods). With enough sequenced variants ($L \sim 10^5$), selection coefficients of order $|s| \sim 1\%$ should be detectable over a single generation even when allele frequency noise is of comparable magnitude (i.e. read depth or population size $= 100$). The intuition here is that when selection simultaneously influences many loci, even if its effects at individual loci are too subtle to be discerned, variants across the genome experience a detectable non-neutral shift as a collective.

The approach of (28, 42) is also an aggregate approach in the above sense, but based instead on the temporal autocovariance in Δp (i.e. the covariance between Δp measured at different times). Ref. (28) similarly found strong genome-wide signals of selection in the same *Drosophila* SNPs (11, 12, 15). Our findings expand upon these in two major respects. First, we go beyond variance/covariance metrics of selective allele frequency change to estimate the genome-wide magnitude of total selection coefficients. Second, our approach explicitly considers frequency dependence. Ref. (28) group all frequencies together, which in practice means temporal autocovariance predominantly reflects the behavior of high major frequency alleles as these are overwhelmingly more numerous. Background selection thus readily generates non-zero auto-covariances in simulations (28). Here in contrast, the frequency dependence of our selective signature allows us to discriminate between modes of selection — the most parsimonious explanation of elevated variance at intermediate frequencies is positive selection

driving allele frequency divergence for these alleles.

The theory underpinning our approach runs counter to the spirit of the influential effective population size literature (30). The variance coefficient C_t can be expressed in terms of the "variance effective population size" N_e as $C_t = 1 - (1 - 1/2N_e)^t$. Therefore, the entire premise of our approach is that selection makes N_e frequency-dependent, removing the intuitive appeal of a population size analogy. It has been shown that in the presence of positive selection the coalescent-effective population size (a close relative of the variance-effective N_e (30)) depends on timescale, with genetic drift dominating in the short term and the "drafting" effect of positive selective sweeps (33) becoming increasingly important over longer times (43). We have shown that even just focusing on the short term, genetic drift and selection contribute fundamentally differently, with no one N_e value fully accounting for the observed divergence of allele frequencies. Deviation in $N_e(p)$ as a function of p is however small compared to the average magnitude of N_e, so the main consequence of assuming a single N_e in the short term is overlooking a potentially valuable signature of polygenic selection.

Quantifying the bounds on how much selection is possible, and how much selection actually occurs in natural popoulations, is a long running controversy (44, 45). The strong total selection coefficients ($|s| \sim 1\%$) we find must predominantly reflect linked selection on neutral SNPs, because most SNPs are neutral. This implies a substantial risk of greatly overestimating the amount of direct selection when, as is commonly done, selection coefficients are inferred at individual loci and then attributed to direct selection. This "excess significance" is a well known difficulty in E&R experiments (12, 46), and similar challenges have arisen in wild populations (15). It is clear from our results that simply improving the accuracy of single-locus selection coefficient inferences, or better controlling for multiple comparisons, will not help to resolve this issue. Our simulation results (Fig. 4) suggest that the frequency structure of allele-frequency divergence is informative about the underlying structure of direct selection. Fully exploiting this information would ideally involve moving beyond variances and studying the shape of the entire distribution of allele frequency change $\Delta_t p$.

Materials and Methods

Data sources and availability. All processing and plotting code can be accessed at https://github.com/jasonbertram/polygenic_variance/. SNP frequency data were obtained from the open access resources published in (15) (wild *D. Melanogaster*, 1 replicate, $\sim 5 \times 10^5$ SNPs, 7 timepoints), (11) (*D. Simulans* E&R, 10 replicates, $\sim 5 \times 10^6$ SNPs, 7 timepoints) and (12) (*D. Simulans* E&R, 3 replicates, $\sim 3 \times 10^5$ SNPs, 2 timepoints). For the (15) data, only SNPs tagged as "used" were included. We performed no additional SNP filtering.

Block bootstrap confidence intervals. We use bootstrapping to estimate the variability of the quantities plotted in Figs. 1-3. These quantities are calculated as an average over loci, where nearby loci are unlikely to be statistically independent due to linkage. To account for the non-independence of individual loci when bootstrap sampling, 95% confidence intervals are calculated using a block bootstrap procedure (28). Each chromosome is partitioned into 1 megabase windows (~ 120 total windows). Bootstrap sampling is then applied to these windows. The plotted vertical lines span the 2.5% and 97.5% block bootstrap percentiles.

Estimation of the selection coefficient variance. To derive Eq. (5) we use the reference value $C_t(p^*)$ as an upper bound on the drift and measurement components of $C_t(p)$ for all p. That is, we argue that

$$\begin{aligned} C_t(p^*) &= M + D_t + (1-2p^*)E_s[st|p^*] \\ &\quad + p^*(1-p^*)\sigma^2(st|p^*) \\ &> M + D_t \end{aligned} \qquad [6]$$

The first line above is Eq. (4) evaluated at the reference frequency p^* with an additional measurement error term M included. M is constant because measurement error is binomial (Supplement S2; (39, 40)).

To justify the inequality above, first note that $p^*(1-p^*)\sigma^2(st|p^*)$ must always be positive. Second, we find empirically in all three datasets we study (11, 12, 15) that $E[\Delta p|p]$ is negligibly small at $p \approx 0.5$, while at higher frequencies as $p \to 1$, $E[\Delta_t p|p]$ consistently becomes negative (i.e. allele cohorts at high frequencies move on average towards intermediate frequencies; Figs. S2-S4). Thus $(1-2p^*)E_s[st|p^*]$ is also positive. Removing these two positive terms thus gives the inequality Eq. (6).

Taking the difference $C_t(p) - C(p^*)$ and using Eq. (6), we have

$$C_t(p) - C_t(p^*) < p(1-p)\sigma^2(st|p) + (1-2p)E_s[st|p] \qquad [7]$$

Taking the difference $C_t(p) - C(p^*)$ thus eliminates the need to estimate D_t and M. To obtain Eq. (5) we simply drop the $(1-2p)E_s[st|p]$ term, which is negligibly small for the $p \approx 0.5$ case that we focus on.

Simulations. We used SLiM (47) to simulate a closed population with $N = 10^3$ individuals, a 100Mb diploid genome, a recombination rate of 10^{-8}/base pair/generation, and a neutral mutation rate of 10^{-8}/base pair/generation. Non-neutral mutations were introduced at rate U/chromosome/generation, where in each simulation non-neutral mutations were assumed to have the same fixed selection coefficient. Four background selection regimes ($U = 1, 0.1 \times s = -0.05, -0.01$), one neutral regime ($U = 0$), and four positive selection regimes ($U = 0.1, 0.01 \times s = 0.01, 0.02$) were evaluated (Fig. 4). In each regime, 100 replicates were simulated with complete genotypes recorded at generations 10^4 and $10^4 + 10$, mimicking the $t = 10$ generation interval in the empirical studies after a burn in period of $10N = 10^4$ generations. Total selection coefficients in Fig. 4 computed using Eq. (2) from genotype data at generation 10^4.

Estimation limits. Our analysis relies on detecting differences in $C_t(p)$ between cohorts with different values of p. The ability to detect such differences is determined by the sampling error in $C_t(p)$ arising due to the calculation of $\mathrm{Var}(\Delta_t p|p)$ from a finite number of loci. To estimate this sampling error, we assume that $\Delta_t p$ is approximately normally distributed, in which case the sample variance in $\mathrm{Var}(\Delta_t p|p)$ is $2\mathrm{Var}(\Delta_t p|p)^2/(L-1) \approx 2\mathrm{Var}(\Delta_t p|p)^2/L$ where $L \gg 1$ is the number of independent loci used to estimate $\mathrm{Var}(\Delta_t p|p)$. The standard error in $C_t(p) = \mathrm{Var}(\Delta_t p|p)/p(1-p)$ is thus given by $\sqrt{2/L}C_t(p)$. This defines the scale of statistically detectable differences in $C_t(p) - C_t(p^*)$, which in turn determines the statistically detectable lower bound estimate on $\sigma^2(s|p)$ Eq. (5). For example, to detect $\sigma^2(s|p) \sim 10^{-4}$ at $p = 0.5$ (i.e. a typical selection coefficient of $\sigma(s|p) \sim 1\%$) after one generation of evolution with $C_1 \sim 10^{-2}$ (i.e. a population sample of ~ 100 individuals, an average read depth of ~ 100 and fairly strong genetic drift $D_1 \sim 10^{-2}$), we need at least $L \sim 10^5$ independent SNPs.

ACKNOWLEDGMENTS. We thank Matthew Hahn for insightful discussions.

1. NH Barton, Genetic hitchhiking. *Philos. Transactions Royal Soc. London. Ser. B: Biol. Sci.* **355**, 1553–1562 (2000).
2. AD Kern, MW Hahn, The neutral theory in light of natural selection. *Mol. biology evolution* **35**, 1366–1371 (2018).
3. JD Jensen, et al., The importance of the neutral theory in 1968 and 50 years on: a response to kern and hahn 2018. *Evolution* **73**, 111–114 (2019).
4. RN Gutenkunst, RD Hernandez, SH Williamson, CD Bustamante, Inferring the joint demographic history of multiple populations from multidimensional snp frequency data. *PLoS genet* **5**, e1000695 (2009).
5. RJ Haasl, BA Payseur, Fifteen years of genomewide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Mol. ecology* **25**, 5–23 (2016).

DRAFT

6. B Charlesworth, M Morgan, D Charlesworth, The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).
7. F Pouyet, S Aeschbacher, A Thiéry, L Excoffier, Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inferences. *Elife* **7**, e36317 (2018).
8. E Elyashiv, et al., A genomic map of the effects of linked selection in drosophila. *PLoS genetics* **12**, e1006130 (2016).
9. J Li, et al., Joint analysis of demography and selection in population genetics: where do we stand and where could we go? *Mol. ecology* **21**, 28–44 (2012).
10. M Lynch, WC Ho, The limits to estimating population-genetic parameters with temporal data. *Genome biology evolution* **12**, 443–455 (2020).
11. N Barghi, et al., Genetic redundancy fuels polygenic adaptation in drosophila. *PLoS biology* **17**, e3000128 (2019).
12. JK Kelly, KA Hughes, Pervasive linked selection and intermediate-frequency alleles are implicated in an evolve-and-resequencing experiment of drosophila simulans. *Genetics* **211**, 943–961 (2019).
13. NO Therkildsen, et al., Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. *Science* **365**, 487–490 (2019).
14. JP Castro, et al., An integrative genomic analysis of the longshanks selection experiment for longer limbs in mice. *elife* **8**, e42014 (2019).
15. AO Bergland, EL Behrman, KR O'Brien, PS Schmidt, DA Petrov, Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in drosophila. *PLoS Genet.* **10**, e1004775 (2014).
16. PJ Monnahan, J Colicchio, L Fishman, SJ Macdonald, JK Kelly, Predicting evolutionary change at the dna level in a natural mimulus population. *PLOS Genet.* **17**, 1–25 (2021).
17. JP Bollback, TL York, R Nielsen, Estimation of 2nes from temporal allele frequency data. *Genetics* **179**, 497–502 (2008).
18. CJ Illingworth, V Mustonen, Distinguishing driver and passenger mutations in an evolutionary history categorized by interference. *Genetics* **189**, 989–1000 (2011).
19. AS Malaspinas, O Malaspinas, SN Evans, M Slatkin, Estimating allele age and selection coefficient from time-serial data. *Genetics* **192**, 599–607 (2012).
20. AF Feder, S Kryazhimskiy, JB Plotkin, Identifying signatures of selection in genetic time series. *Genetics* **196**, 509–522 (2014).
21. BS Khatri, Quantifying evolutionary dynamics from variant-frequency time series. *Sci. reports* **6**, 32497 (2016).
22. Z He, X Dai, M Beaumont, F Yu, Estimation of natural selection and allele age from time series allele frequency data using a novel likelihood-based approach. *Genetics* **216**, 463–480 (2020).
23. JG Schraiber, SN Evans, M Slatkin, Bayesian inference of natural selection from allele frequency time series. *Genetics* **203**, 493–511 (2016).
24. T Taus, A Futschik, C Schlötterer, Quantifying selection with pool-seq time series data. *Mol. biology evolution* **34**, 3023–3034 (2017).
25. JK Pritchard, JK Pickrell, G Coop, The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. biology* **20**, R208–R215 (2010).
26. I Höllinger, PS Pennings, J Hermisson, Polygenic adaptation: From sweeps to subtle frequency shifts. *PLoS genetics* **15**, e1008035 (2019).
27. N Barghi, J Hermisson, C Schlötterer, Polygenic adaptation: a unifying framework to understand positive selection. *Nat. Rev. Genet.* **21**, 769–781 (2020).
28. V Buffalo, G Coop, Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proc. Natl. Acad. Sci.* **117**, 20672–20680 (2020).
29. B Walsh, M Lynch, *Evolution and selection of quantitative traits*. (Oxford University Press), (2018).
30. B Charlesworth, Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* **10**, 195–205 (2009).
31. A Robertson, Inbreeding in artificial selection programmes. *Genet. Res.* **2**, 189–194 (1961).
32. E Santiago, A Caballero, Effective size of populations under selection. *Genetics* **139**, 1013–1030 (1995).
33. JH Gillespie, Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* **155**, 909–919 (2000).
34. E Santiago, A Caballero, Effective size and polymorphism of linked neutral loci in populations under directional selection. *Genetics* **149**, 2105–2117 (1998).
35. C Cannings, The latent roots of certain markov chains arising in genetics: a new approach, i. haploid models. *Adv. Appl. Probab.* **6**, 260–290 (1974).
36. R Der, CL Epstein, JB Plotkin, Generalized population models and the nature of genetic drift. *Theor. population biology* **80**, 80–99 (2011).
37. Z Gompert, SP Egan, RD Barrett, JL Feder, P Nosil, Multilocus approaches for the measurement of selection on correlated genetic loci. *Mol. Ecol.* **26**, 365–382 (2017).
38. J Felsenstein, Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution*, 1229–1242 (1981).
39. Y Zhu, AO Bergland, J González, DA Petrov, Empirical validation of pooled whole genome population re-sequencing in drosophila melanogaster. *PloS one* **7**, e41901 (2012).
40. M Gautier, et al., Estimation of population allele frequencies from next-generation sequencing data: pool-versus individual-based genotyping. *Mol. Ecol.* **22**, 3766–3779 (2013).
41. G Sella, DA Petrov, M Przeworski, P Andolfatto, Pervasive natural selection in the drosophila genome? *PLoS Genet.* **5**, e1000495 (2009).
42. V Buffalo, G Coop, The linked selection signature of rapid adaptation in temporal genomic data. *Genetics* **213**, 1007–1045 (2019).
43. DB Weissman, NH Barton, Limits to the rate of adaptive substitution in sexual populations. *PLoS Genet.* **8**, e1002740 (2012).
44. JBS Haldane, The cost of natural selection. *J. Genet.* **55**, 511 (1957).
45. NH Barton, Linkage and the limits to natural selection. *Genetics* **140**, 821–841 (1995).
46. SV Nuzhdin, TL Turner, Promises and limitations of hitchhiking mapping. *Curr. opinion genetics & development* **23**, 694–699 (2013).
47. BC Haller, PW Messer, Slim 3: forward genetic simulations beyond the wright–fisher model. *Mol. biology evolution* **36**, 632–637 (2019).

DRAFT

Supporting Information Text

S1 Frequency dependence of C for neutrally evolving alleles

Here we show that the allele frequency variance of neutrally evolving alleles has the form $\mathrm{Var}(\Delta_t p|p) = C_t p(1-p)$ where $\Delta_t p$ is the change in allele frequency over t generations, the variance is evaluated across loci with the same initial allele frequency p, and the variance coefficient C_t is independent of p. Neutrally evolving alleles are defined as having no effect on fitness and being in linkage equilibrium with alleles that do affect fitness.

Our argument is based on a generalization of the Cannings model (1). Let y_i be the number of copies of an allele descended from the i'th copy of that allele t generations ago. There are a total of $Y = \sum_{i=1}^{X} y_i$ copies of the allele in the population, where X is the number of allele copies present t generations ago. The allele frequency change that occurs is $\Delta_t p = \frac{Y}{N(t)} - \frac{X}{N(0)}$, where $N(0)$ and $N(t)$ denote the census population size t generations ago and at the present time, respectively. To simplify our equations we assume a haploid population; for a diploid population we replace $N \to 2N$ without otherwise affecting our argument.

Since all loci share the same demography, there is no among-locus variance in $N(t)$. It follows that the among-locus variance in Y for alleles with the same initial frequency p (and hence same initial abundance X) can be written as $\mathrm{Var}(Y|p) = \mathrm{Cov}(Y, Y|p) = -\mathrm{Cov}(Y, N(t) - Y|p) = -\mathrm{Cov}(\sum_{i=1}^{X} y_i, \sum_{j=X+1}^{N} y_j|p) = -X(N-X)\overline{\mathrm{Cov}(y_i, y_j|p)}$, where $\overline{\mathrm{Cov}(y_i, y_j|p)} = \frac{1}{X(N-X)} \sum_{i=1}^{X} \sum_{X+1}^{N} \mathrm{Cov}(y_i, y_j|p)$ is the descendant number covariance averaged over allele copies. Consequently, we have

$$\begin{aligned}
\mathrm{Var}(\Delta p|p) &= \mathrm{Var}\left(Y/N(t) - X/N(0)|p\right) \\
&= \mathrm{Var}\left(Y/N(t)|p\right) \\
&= \frac{N(0)^2}{N(t)^2} \frac{\mathrm{Var}(Y|p)}{N(0)^2} \\
&= -\frac{N(0)^2}{N(t)^2} \overline{\mathrm{Cov}(y_i, y_j|p)} p(1-p). \qquad [1]
\end{aligned}$$

To prove our claim that C_t is independent of p we therefore need to show that the averaged covariance $\overline{\mathrm{Cov}(y_i, y_j|p)}$ is independent of p.

Frequency-independent C_t is straightforward to show in the Cannings model (2). Neutrality in the Cannings model takes the form of a symmetry assumption: the random variables y_i are assumed to be exchangeable (i.e. the joint probability distribution of the y_i is invariant with respect to rearrangement of their labels i; (3)). Exchangeability allows for a wide variety of marginal distributions for y_i (the allelic offspring distribution), giving the Cannings model enormous biological flexibility (e.g. (2)). When the y_i are exchangeable, $\overline{\mathrm{Cov}(y_i, y_j|p)} = \mathrm{Cov}(y_i, y_j|p)$ is independent of i, j for $i = 1, \ldots, X$ and $j = X+1, \ldots, N$ because $\mathrm{Var}(N(t)) = 0$ implies $\mathrm{Cov}(y_i, y_j|p) = -\sigma^2/(N(0)-1)$ where $\sigma^2 = \mathrm{Var}(y_1) = \ldots = \mathrm{Var}(y_{N(0)})$ is the variance in the number of allele descendants. This value of $\mathrm{Cov}(y_i, y_j|p)$ is independent of p because, from the definition of exhcangeability, which allele identities are attached to the y_i has no bearing on their joint probability distribution.

However, the Cannings exchangeability assumption has an important limitation as a model of genetic neutrality, namely that no non-neutral loci can be present in the genome. Non-neutral loci introduce fitness variation among genetic backgrounds which implies differences in the marginal distributions of the y_i corresponding to those genetic backgrounds. Nevertheless, provided that the loci over which $\mathrm{Var}(\Delta_t p|p)$ is being calculated are themselves neutral and in linkage equilibrium with loci under selection, $\overline{\mathrm{Cov}(y_i, y_j|p)}$ will still be frequency independent. To see this, note first that $\overline{\mathrm{Cov}(y_i, y_j|p)}$ is invariant with respect to reshuffling $i, j \to i', j'$ of the allele copy labels at each locus:

$$\begin{aligned}
\overline{\mathrm{Cov}(y_i, y_j|p)} &= \frac{1}{X(N(0)-X)} \mathrm{Cov}\Big(\sum_{i=1}^{X} y_i, \sum_{j=X+1}^{N} y_j\Big) \\
&= \frac{1}{X(N(0)-X)} \mathrm{Cov}\Big(\sum_{i'=1}^{X} y_{i'}, \sum_{j'=X+1}^{N} y_{j'}\Big) \\
&= \overline{\mathrm{Cov}(y_{i'}, y_{j'}|p)} \qquad [2]
\end{aligned}$$

In particular, at each locus we can assign i, j labels at random such that y_i and y_j are sampled at random from one focal and one non-focal background respectively. For neutral alleles in linkage equilibrium with non-neutral alleles, the focal and non-focal distributions of genetic backgrounds thus sampled are identical. This leads to a situation analogous to strict exchangeability: the allele identities of the y_i have no bearing on the distribution of pairs y_i, y_j that enter the covariance implying that $\overline{\mathrm{Cov}(y_i, y_j|p)}$ is independent of p. This proves our claim.

S2 Measurement error

In this section we show that the major sources of allele frequency measurement error do not systematically alter the frequency independence of C_t in main text Eq. (1). Population allele frequencies are denoted p while estimated allele frequencies are denoted $\hat{p}$.

Suppose that the measurement variance satisfies $\mathrm{Var}(\hat{p}|p) = Mp(1-p)$ where M is frequency independent and can also differ between generations. Applying the law of total variance (for any two random variables X and Y, $\mathrm{Var}(Y) = E_X[\mathrm{Var}_Y(Y|X)] + \mathrm{Var}_Y(E_X[Y|X])$), alleles with the same initial frequency p_0 will have variance,

$$\begin{aligned}
\mathrm{Var}(\Delta_t\hat{p}|p_0) &= \mathrm{Var}(\hat{p}_t - \hat{p}_0|p_0) \\
&= \mathrm{Var}(\hat{p}_t|p_0) + \mathrm{Var}(\hat{p}_0|p_0) \\
&= E_{p_t}[\mathrm{Var}(\hat{p}_t|p_t, p_0)] + \mathrm{Var}_{p_t}(E[\hat{p}_t|p_t, p_0]) + M_0p_0(1-p_0) \\
&= E_{p_t}[M_tp_t(1-p_t)|p_0] + \mathrm{Var}_{p_t}(p_t|p_0) + M_0p_0(1-p_0) \\
&= M_t(E[p_t|p_0] - E[p_t^2|p_0]) + C_tp_0(1-p_0) + M_0p_0(1-p_0) \\
&= M_t(E[p_t|p_0] - (\mathrm{Var}(p_t|p_0) + E[p_t|p_0]^2)) + (C_t + M_0)p_0(1-p_0) \\
&= (C_t + M_0 + M_t - M_tC_t)\, p_0(1-p_0)
\end{aligned} \qquad [3]$$

where we have used the fact that $E[p_t|p_0] = p_0$ for neutral alleles. Therefore, measurement error does not change the frequency dependence of neutral cohort variances, provided that our assumption of frequency independent M_t is true. We now show the latter accounting for different sources of error. Following (4), we divide measurement error into three main sources that will be considered in turn: population sampling, pooling and amplification of the sampled DNA, and finite read depth.

Population sampling. Random sampling of individuals generates a binomial sampling error, such that the sample allele frequency p_S has variance $\mathrm{Var}(p_S|p) = p(1-p)/n_S$ where n_S is the sample size. Using the law of total variance, the contribution of this sampling error is

$$\begin{aligned}
\mathrm{Var}(\hat{p}|p) &= E_{p_S}[\mathrm{Var}(\hat{p}|p_S)] + \mathrm{Var}_{p_S}[E(\hat{p}|p_S)] \\
&= E_{p_S}[\mathrm{Var}(\hat{p}|p_S)] + p(1-p)/n_s
\end{aligned} \qquad [4]$$

The term $\mathrm{Var}(\hat{p}|p_S)$ represents the error in $\hat{p}$ introduced after population sampling, which we now evaluate.

Read sampling. Let r be the number of reads of the reference allele A recorded by the sequencer at a locus with a total of r_T reads. The allele frequency estimate of A is $\hat{p} = r/r_T$. We initially restrict our attention to loci that have the same read depth r_T; below we incorporate the effects of among-locus variation in r_T. Reads are assumed to be generated by random sampling with replacement from the DNA pool. Reference allele reads are thus binomially distributed $r \sim \mathrm{Bin}(r_T, p_P)$ where p_P is the allele frequency of A in the sequenced DNA pool.

If all allele copies sampled from the population contribute equally to the sequenced DNA pool, then $p_P = p_S$. Errors introduced during pooling and amplification cause unequal contributions and thus variance in p_P among sequenced loci. We assume that pooling+amplification error do not systematically bias allele frequencies such that $E_{p_P}[p_P] = p_S$. Applying the law total variance to $\hat{p}$ for fixed r_T gives

$$\begin{aligned}
\mathrm{Var}(\hat{p}|p_S, r_T) &= \frac{\mathrm{Var}(r|p_S, r_T)}{r_T^2} \\
&= \frac{E_{p_P}[\mathrm{Var}_r(r|p_S, p_P, r_T)] + \mathrm{Var}_{p_P}(E_r[r|p_S, p_P, r_T])}{r_T^2} \\
&= \frac{E_{p_P}[r_Tp_P(1-p_P)] + \mathrm{Var}_{p_P}(r_Tp_P|p_S)}{r_T^2} \\
&= \frac{p_S(1-p_S) + (r_T-1)\mathrm{Var}(p_P|p_S)}{r_T}
\end{aligned} \qquad [5]$$

where we have used the identity $E[X^2] = \mathrm{Var}[X] + E[X]^2$ in the last line. Note that the conditionality on r_T has been dropped in $\mathrm{Var}_{p_P}(p_P|p_S)$ because the DNA pool frequency p_P is independent of r_T. We now incorporate the effects of pooling and amplification to show that $\mathrm{Var}_{p_P}(p_P|p_S) \propto p_S(1-p_S)$.

Pooling+amplification. Let F_i denote the proportion of the DNA pool at the focal locus originating from the i'th sampled allele copy, where $i = 1, \ldots, 2n_S$ for diploids. In the case of equal contributions from each sampled allele copy we simply have $F_i = 1/2n_S$, but in general the F_i are random variables with unknown probability distributions denoted $P(F_i)$ (ref. (4) assumes a Dirichlet distribution for the F_i). We have $p_P = \sum_{i=1}^{2n_Sp_S} F_i$ where $2n_Sp_S$ is the number of sampled allele copies containing the focal allele (i.e. we order the haplotype labels i such that the reference allele copies appear first). Since the F_i sum to one, we have analogous to Eq. (1),

$$\mathrm{Var}_{p_P}(p_P|p_S) = -\overline{\mathrm{Cov}(F_i, F_j)}p_S(1-p_S) \qquad [6]$$

where $\overline{\mathrm{Cov}(F_i, F_j)} = \frac{1}{4n_S^2p_S(1-p_S)}\sum_i^{2n_Sp_S}\sum_j^{2n_S(1-p_S)}\mathrm{Cov}(F_i, F_j)$ is the mean covariance of the DNA pool contributions between the focal and non-focal alleles.

Similar to exchangeability in the Cannings model, pooling and amplification are blind to allele copy identity and should therefore affect allele copy probabilities symmetrically (see S1). This implies $\overline{\mathrm{Cov}(x_i, x_j)} = \mathrm{Cov}(x_i, x_j) = \mathrm{Var}(F_i)/(2n_s - 1)$ independent of i and p_S. Combining the this result with Eq. (5) we therefore obtain

$$\mathrm{Var}(\hat{p}|p_S, r_T) = f(r_T)p_S(1 - p_S) \qquad [7]$$

where $f(r_T)$ is independent of p_S.

Variable read depth. Read depth varies across loci according to some probability distribution $P(r_T)$. This variability in incorporated by integrating out the dependence on r_T

$$\mathrm{Var}(\hat{p}|p_S) = \int \mathrm{Var}(\hat{p}|p_S, r_T)P(r_T)dr_T = fp_S(1 - p_S) \qquad [8]$$

where $f = \int f(r_T)dr_T$. Substituting this result into Eq. (4) we finally obtain

$$\mathrm{Var}(\hat{p}|p) = \left(\frac{1}{n_S} + f + \frac{f}{n_S}\right) p(1 - p), \qquad [9]$$

where the term multiplying $p(1 - p)$ is frequency-independent as claimed at the start of S2.

References

1. C Cannings, The latent roots of certain markov chains arising in genetics: a new approach, i. haploid models. *Adv. Appl. Probab.* **6**, 260–290 (1974).
2. R Der, CL Epstein, JB Plotkin, Generalized population models and the nature of genetic drift. *Theor. population biology* **80**, 80–99 (2011).
3. YS Chow, H Teicher, *Probability theory: independence, interchangeability, martingales.* (Springer Science & Business Media), (2012).
4. M Gautier, et al., Estimation of population allele frequencies from next-generation sequencing data: pool-versus individual-based genotyping. *Mol. Ecol.* **22**, 3766–3779 (2013).

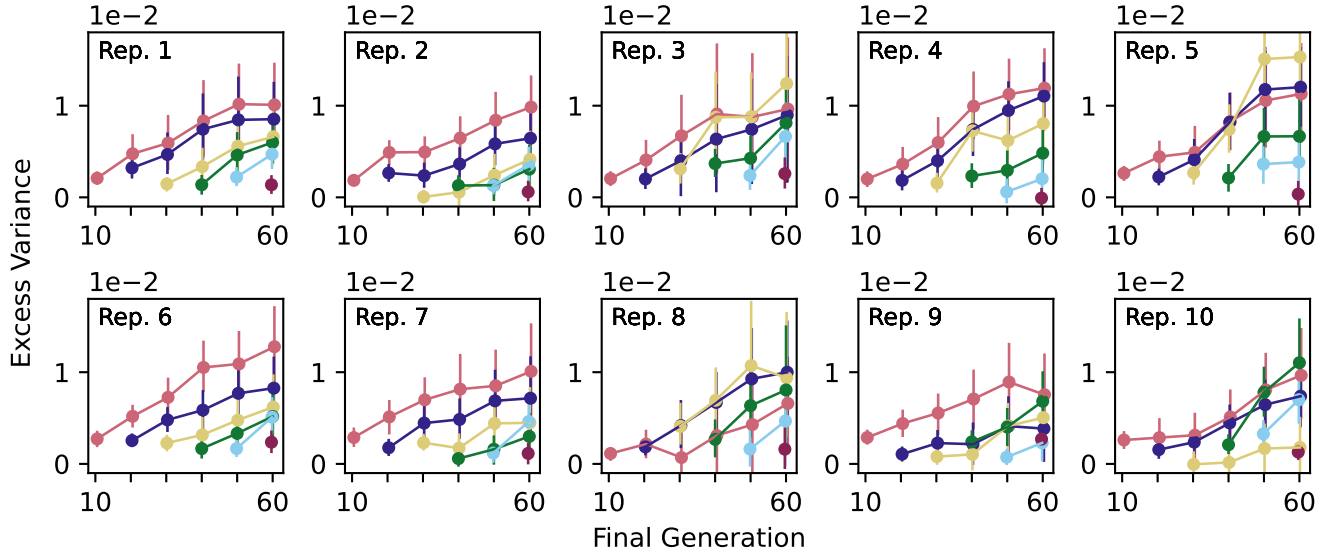

Fig. S1. Identical to Fig. 2A in the main text but including all 10 replicates from Barghi et al.

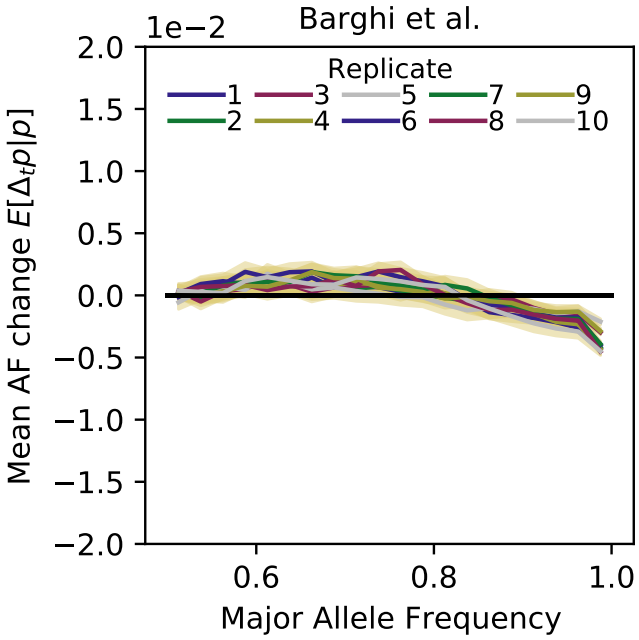

Fig. S2. Barghi et al. expected $\Delta_t p$, first iteration, all replicates. 95% 1MB block bootstrap confidence intervals are too narrow to be easily visible.

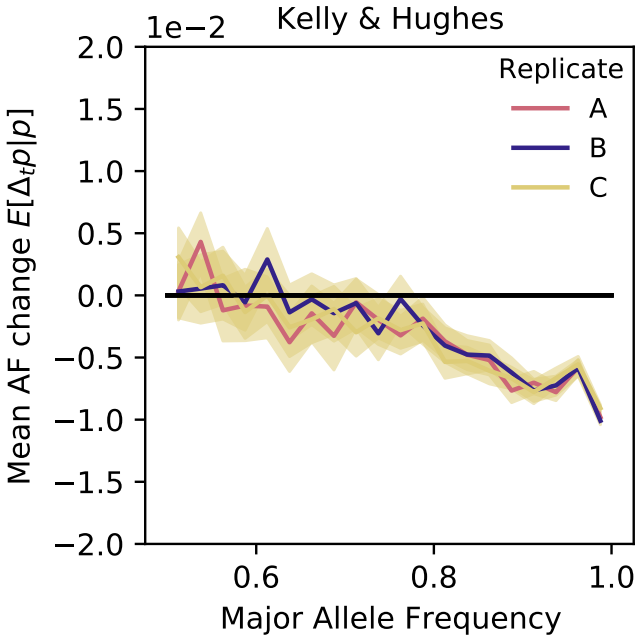

Fig. S3. Kelly et al. expected $\Delta_t p$, all replicates, with 1MB Block bootstrap 95% confidence intervals.

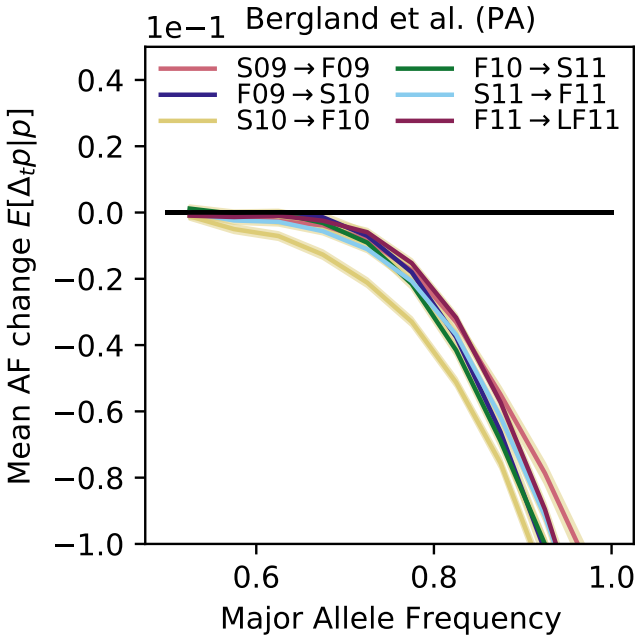

Fig. S4. Bergland et al. expected $\Delta_t p$ over all seasonal iterations. 95% 1MB block bootstrap confidence intervals are too narrow to be easily visible.