Supporting Information

Assaf et al. 10.1073/pnas.1424949112

SI Text

1. Balancing Selection for Arbitrary Dominance and Selection Coefficients

For arbitrary dominance the fitnesses of the diploids in the population are

$$w_{OO/OO} = 1 \quad \text{[S1]}$$

$$w_{BD/OO} = (1+h_b s_b)(1-h_d s_d) = 1 + h_b s_b - h_d s_d - h_b s_b h_d s_d \quad \text{[S2]}$$

$$w_{BD/BD} = (1+s_b)(1-s_d) = 1 + s_b - s_d - s_b s_d. \quad \text{[S3]}$$

The change in frequency of a BD haplotype at frequency p_t at time t due to selection can be derived,

$$p_{t+1} = p_t^2 \log\left(\frac{w_{BD/BD}}{\overline{w}_{pop}}\right) + p_t q_t \log\left(\frac{w_{BD/OO}}{\overline{w}_{pop}}\right). \quad \text{[S4]}$$

Using a series expansion in p and taking the continuous time limit, this becomes a differential equation describing the dynamics of the frequency of the *BD* haplotype, p, due to selection

$$S(p) = (h_b s_b - h_d s_d)p + (s_b - 3h_b s_b - s_d + 3h_d s_d)p^2 + (-s_b + 2h_b s_b + s_d - 2h_d s_d)p^3, \quad \text{[S5]}$$

which is valid for small p, although also qualitatively correct for $p \sim 1$.

We can approximate Eq. **S5** with the assumption $s_b \ll s_d \ll 1$, giving

$$S(p) \approx s_b p - s_d p^2 + s_d p^3. \quad \text{[S6]}$$

The first term reflects selection on heterozygotes with fitness $\sim s_b$ that occur with probability $\sim p$, the second term reflects selection on homozygotes with fitness $\sim -s_b$ that occur with probability $\sim p^2$, and the third term comes from a decrease in population mean fitness at higher frequencies of *BD*. This cubic equation is the origin of the shape of the selection curve in Fig. 2. Because $p^* \ll 1$ when $s_b \ll s_d$, the cubic term can be ignored as it is smaller than the other terms by a factor of $\sim p^*$. The cubic term, however, is important if one enters a regime in which $Ns_d \lesssim 1$; however, this regime is not considered in the paper as it is biologically implausible for most populations.

If we wish to use Eq. **S5** to derive an equilibrium frequency for arbitrary dominance and selection coefficients, we find

$$p^* = \frac{h_d s_d - h_b s_b + h_b s_b h_d s_d}{s_b - 2h_b s_b - s_d + 2h_d s_d - s_b s_d + 2h_b s_b h_d s_d}. \quad \text{[S7]}$$

2. Drift to Extinction from Equilibrium No Recombination

2.1. Strong Selection, Weak Drift Regime ($\alpha > 1$). In the absence of recombination, the rate of extinction depends on two things: the strength of selection [$S(p)$] pushing the *BD* haplotype toward equilibrium p^*, and the variance in frequency due to drift [$D(p)$] that enables the *BD* haplotype to fluctuate to extinction against selection. Under these two effects, the frequency of the *BD* haplotype, p, is governed by the stochastic equation

$$\delta p \approx S(p)\delta t + \eta\sqrt{D\delta t} \quad \text{[S8]}$$

with $S(p) \approx s_b p - s_d p^2 + s_d p^3$ and where $D(p) = p(1-p)/N$ is the variance from drift and η is Gaussian distributed noise with zero mean and unit variance. In the long time limit, the probability density $\rho(p|t, p_0)$ generated by Eq. **S8**, which describes the likelihood of observing a *BD* haplotype at frequency p after time t, given that it started at p_0, has weights at $p = 0$ and $p = 1$ only (which must sum to unity). What this means for the *BD* haplotype is that after a long enough time, it must eventually fluctuate to extinction or fixation. In our particular case, because $s_b \ll s_d$, the probability of fluctuating to extinction is far larger than that of fluctuating to fixation.

Before this time, however, the probability density must move from being concentrated at $p = p_0$ (where it all started) to being concentrated at $p = 0$ (extinction). We want to estimate how quickly this process happens. Consider starting all of the probability density at $p = p^*$. Initially, this probability density spreads out a small amount around p^*, due to drift. However, because of strong selection, it remains sharply peaked around $p = p^*$ and reaches a selection–drift steady state: Its shape does not change, but it begins to decay (i.e., its amplitude decays) at some characteristic rate λ such that $\rho(p|t, p_0) \approx \exp(-\lambda t)$. The rate, λ, is small because we are in the limit of strong selection, and fluctuating to extinction is improbable. Our goal here is to estimate λ, which determines the rate at which *BD* haplotypes go extinct.

To do this, we consider the stochastic equation for δp from Eq. **S8**. For changes in frequency δp such that $\delta p/p \ll 1$, both $S(p)$ (selection) and $D(p)$ (drift) can be assumed to be constant, and Eq. **S8** reduces exactly to standard diffusion. The solution for probability density in this case is therefore a Gaussian over δp,

$$\rho(\delta p|t)d(\delta p) = \frac{d(\delta p)}{\sqrt{2\pi D t}} \exp\left[-\frac{(\delta p - St)^2}{2Dt}\right]. \quad \text{[S9]}$$

This makes sense: The mean change in frequency is St, which is the change that would have been expected without drift, whereas the variance around this grows linearly in time Dt, with a coefficient that is exactly the variance introduced by the drift term. Multiplying out the quadratic and exponentiating the constant out front, this can be written as

$$\rho(\delta p|t) = \frac{1}{\sqrt{2\pi}} \exp\left[\beta - \frac{\beta}{2}\left(\frac{t}{\tau} + \frac{\tau}{t}\right) - \frac{\ln(t/\beta\tau)}{2}\right] \quad \text{[S10]}$$

where

$$\tau = \frac{\delta p}{S} \quad \text{and} \quad \beta = \frac{S\delta p}{D}. \quad \text{[S11]}$$

Writing the density in this way is particularly useful because one can straightforwardly see when selection dominates or when drift dominates. The parameter β determines which of the two is most important: when $|\beta| \gg 1$, selection dominates, whereas for $|\beta| \ll 1$, drift dominates. One way of interpreting β is as a ratio of two timescales,

$$\beta = \frac{\text{time for drift to change frequency by } \delta p}{\text{time or selection to change frequency by } \delta p} = \frac{(\delta p^2/D)}{(\delta p/S)} = \frac{S\delta p}{D}. \quad \text{[S12]}$$

If the timescale for selection to change the frequency by δp is much faster than that for drift (large β), then selection drives the dynamics. The reverse is true for small β in which drift drives the dynamics. Focusing on the case of strong selection and considering Eq. **S10**, one can see that when $\beta \gg 1$, the $\ln(t/\beta\tau)$ term is small compared with the other two; thus it can be ignored. The log density then becomes

$$\ln \rho(\delta p|t) \approx \beta - \frac{\beta}{2}\left(\frac{t}{\tau}+\frac{\tau}{t}\right). \quad \textbf{[S13]}$$

In our specific case, we are interested in the *BD* haplotype fluctuating from $p=p^*$ to $p=0$ (i.e., negative δp) in the presence of a positive selection (S) pushing the *BD* haplotype upward toward $p=p^*$. This means that we are interested in the case of $\delta p<0$ and $S>0$. Considering the expressions for τ and β, one can see that such a case corresponds to both $\beta<0$ and $\tau<0$. One can see from Eq. **S13** that the probability density is sharply peaked at $t=|\tau|$, regardless of the sign of β. This somewhat counterintuitive result means the following: When selection is strong ($|\beta| \gg 1$), if the *BD* haplotype moves against the flow of selection, driven purely by fluctuations, then the path it takes will be similar to the path it would have taken if it had moved with the flow of selection [a result similar to that found by Maruyama (1)]. The key thing to note, however, is that when selection is strong, then the *BD* haplotype is very unlikely to fluctuate against the flow of selection.

To quantify just how unlikely this is, we remember that in the rare cases where movement occurs against the flow, then the most probable paths are close to a time $t=|\tau|$. Therefore, the probability of changing in frequency δp in the direction opposite of v (i.e., β negative) is obtained by substituting this time into Eq. **S13**, giving

$$Pr(\text{moving by } \delta p \text{ against } v) \approx \exp(-2|\beta|). \quad \textbf{[S14]}$$

The above probability result assumed that S and D were constant over a small interval δp. To calculate the probability of fluctuating over the entire selective barrier $S(p)$ to extinction, one must therefore take the product of the probabilities of many such jumps, each with its own $\beta(p)$. This product gives an integral in the exponent

$$P(p^* \to 0) \approx \exp\left[2\int_{p^*}^{0} dp \frac{S(p)}{D(p)}\right]. \quad \textbf{[S15]}$$

In our specific case,

$$S(p) \approx s_b p - s_d p^2 \quad \textbf{[S16]}$$

$$D(p) \approx p/N \quad \textbf{[S17]}$$

$$p^* \approx s_b/s_d. \quad \textbf{[S18]}$$

The integral can be solved exactly giving

$$P(p^* \to 0) \approx \exp\left(-\frac{Ns_b^2}{s_d}\right). \quad \textbf{[S19]}$$

This result has the rather simple interpretation that what primarily determines the probability is the barrier height. Evaluating the strength of selection and noise at this highest point and multiplying by the distance it has to move, p^*, also gives exactly the above result. The parameter

$$\alpha \equiv Ns_b^2/s_d \quad \textbf{[S20]}$$

is therefore particularly important in the dynamics because it alone determines the probability of fluctuating to extinction

$$P(p^* \to 0) \approx A\exp[-B\alpha] \quad \textbf{[S21]}$$

with A and B constants $O(1)$.

At present, we have calculated the probability of fluctuating to extinction. To translate this into an estimate for the rate of decay, λ, we must remember that in deriving the above result for the probability, we used the fact that each of the small jumps in frequency δp that the *BD* haplotype takes to reach $p=0$ lasts a time $\tau=\delta p/S$. We can therefore calculate how long it takes for that probability to decay by adding up all of the τ on its journey from $p=p^*$ to $p=0$. However, our calculation must only include the region where $\beta>1$ (i.e., where selection dominates). Therefore, the frequency limits in the integral are not $p=p^*$ and $p=0$ but rather $p=p^*-1/Ns_b$ to $p=1/Ns_b$, because these define the region of the barrier where $\beta>1$ (see Fig. 2*A*). Therefore

$$\text{time to extinction} = \int_{p^*-1/Ns_b}^{1/Ns_b} \frac{dp}{S(p)} = \int_{p^*-1/Ns_b}^{1/Ns_b} \frac{dp}{s_b p - s_d p^2} \quad \textbf{[S22]}$$

This integral will be dominated by where the flow of selection is slowest, which is the start and the finish. Because of the symmetry of the quadratic around the point $p=p^*/2$, we can integrate from $1/Ns_b$ to $p^*/2$ and then double the result. We can also realize that because $p \ll 1$ the quadratic term can be ignored, giving

$$\text{time to extinction} \approx 2\int_{1/Ns_b}^{p^*/2} \frac{dp}{s_b p} = \frac{2}{s_b}\ln\left(\frac{Ns_b p^*}{2}\right) \approx \frac{2}{s_b}\ln(\alpha). \quad \textbf{[S23]}$$

This result again makes sense: This time is exactly the characteristic time it would have taken the *BD* haplotype to rise from $1/Ns_b$ to p^*, were it to have done so deterministically under the influence of selection alone. Approximating the rate is now straightforward: Given that an amount of probability $\exp(-Ns_b^2/s_d)$ is lost in an amount of time $\sim 1/s_b \ln(\alpha)$, it means the rate of decay, λ, must be approximately given by

$$\lambda_l \approx \frac{e^{-\alpha}}{(1/s_b)\ln(\alpha)}, \quad \textbf{[S24]}$$

which is the result quoted Eq. **10** of the main text, such that mean time to extinction $\tau_l \approx 1/\lambda_l$, and which agrees very well with the simulations described in *SI Text*, section 4. We verify the slightly counterintuitive dependence of λ on s_b by performing 10^5 simulations at various s_b but at fixed α in Fig. S3*J*, as well as for various N, s_b, and s_d without fixing α (Fig. S3 *A*–*I*).

2.2. A Note on α for Arbitrary Dominance Coefficients. For the case of arbitrary dominance coefficients ($h_b \neq 1$, $h_d \neq 0$), a relatively straightforward extension can be made to our model by changing α to reflect the fitness of heterozygotes (i.e., replace s_b with $h_b s_b - h_d s_d$). In this case, our approximations will hold as long as there is still heterozygote advantage, or, in other words, $h_b s_b > h_d s_d$ and $h_b s_b - h_d s_d < s_d$. Thus, α becomes

$$\alpha \approx N\frac{(h_b s_b - h_d s_d)^2}{s_d}, \quad \textbf{[S25]}$$

which will still predict the rate of loss. However, note that in the rate of escape quoted in the main paper,

$$\lambda_e \approx rlNs_b \frac{s_b}{s_d} \approx rl\alpha, \quad \text{[S26]}$$

the first s_b value comes from the probability of establishment for a new *BO* recombinant haplotype, and the second s_b value comes from the equilibrium frequency p^*. Thus, for arbitrary dominance, the probability of escape becomes

$$\lambda_e \approx rlNs_b \frac{h_b s_b - h_d s_d}{s_d}. \quad \text{[S27]}$$

2.3. Strong Drift, Weak Selection Regime ($\alpha < 1$). When $\alpha < 1$, there is no region below p^* in which selection dominates over drift. However, there is a region above p^* where selection will dominate over drift. This will occur when the strength of selection $[S(p) = s_b p - s_d p^2]$ approximately matches the strength of drift ($\sim \pm 1/N$). Because this occurs above p^*, the dominant term in $S(p)$ is the quadratic term, so the balance between selection and drift is a balance involving selection against the homozygotes,

$$-s_d p^2 \approx \pm 1/N \quad p \approx \sqrt{N/s_d}\,. \quad \text{[S28]}$$

This means that when the *BD* haplotype is below the frequency of $\sqrt{N/s_d}$, its dynamics are governed by drift, i.e., are largely neutral. Under neutrality, the cumulative probability that a lineage which started at one copy ($n_o = 1$) has gone extinct ($n = 0$) by time t is

$$\Pr[n = 0|t] \approx e^{-n_o/t} \approx e^{-1/t} \approx 1 - 1/t. \quad \text{[S29]}$$

Thus, the probability that a neutral lineage goes extinct in an interval dt near t is $\sim dt/t^2$, giving a distribution of extinction times that are power law distributed as $\sim 1/t^2$.

The distribution of extinction times of the *BD* haplotype in the $\alpha < 1$ regime is expected to follow this neutral power-law distribution, but with a cutoff at $t \approx \sqrt{N/s_d}$ (due to the *BD* haplotype being unable to drift to high frequencies where the recessive deleterious mutation is exposed). The fact that the maximum time a *BD* haplotype can remain in the population in this regime is $t \approx \sqrt{N/s_d}$, which is the same as the maximum frequency it can reach before selection against the homozygotes pushes it down in frequency, is no accident. They are the same because, under neutrality, the number of generations it takes to change by order n is $t \approx n$. We verify this in the simulations section (section 4.3).

Recalling that the total probability of an escape event by time t is

$$\text{probability of escape} \approx rls_b N \int_0^t p(t')dt' \quad \text{[S30]}$$

(which is valid for a probability of escape $\ll 1$). Also recall that, as we just showed and as quoted in *Predictions for the Regime of Weak Selection and Strong Drift ($\alpha \ll 1$)*, in an interval dt, there is a fraction $\sim dt/t^2$ of *BD* haplotypes that go extinct and a fraction $\sim rls_b t^2$ of *BO* haplotypes that escape via recombination, giving a probability of escape $\sim rls_b\, dt$ that is constant in time as long as the BD haplotypes persist. Thus, the probability of fixation in this $\alpha \ll 1$ regime is

$$\text{probability of fixation} \approx rls_b \int_0^{\sqrt{N/s_d}} dt' = rls_b \sqrt{N/s_d} \quad \text{[S31]}$$

where the upper limit in the integral comes from the fact that it is unlikely for the *BD* haplotype to drift for longer than $\sim \sqrt{N/s_d}$ generations.

3. Zones of Altered Adaptation

3.1. Strong Selection, Weak Drift ($\alpha > 1$).

Zone of suppressed probability of fixation. In this regime, the probability that a *BD* haplotype fixes can be significantly reduced relative to the equivalent beneficial mutation with no deleterious hitchhiker, but only if it is likely to fluctuate to extinction before escape can occur. This means that $\tau_l < \tau_e$. Using the expressions for these times from Eqs. **10** and **11** in the main text, the condition for the fixation probability to be significantly reduces becomes

$$\frac{\ln[\alpha]}{s_b} e^{\alpha} \ll 1/rl\alpha. \quad \text{[S32]}$$

Rearranging this equation defines a base pair distance, l_l, around a recessive deleterious mutation within which the probability of fixation of beneficial mutations is reduced relative to the case of no hitchhiker,

$$l_l = \frac{s_b}{r\alpha \ln[\alpha]} e^{-\alpha}. \quad \text{[S33]}$$

Zone of increased sweep time. Even if a beneficial mutation is not driven to extinction, the duration of a beneficial mutation's sweep to fixation can be substantially extended if it's genetically linked to a recessive deleterious hitchhiker. To understand when this occurs, consider that the total time of a successful sweep will be prolonged by the time it takes for an escape event to occur,

$$\text{average sweep time} \approx \frac{\ln[Ns_b]}{s_b} + \left(\frac{1}{\tau_e} + \frac{1}{\tau_l}\right)^{-1}. \quad \text{[S34]}$$

The first term corresponds to the sweep time for a single adaptive mutation with no hitchhiker, and the second term comes from conditioning on fixation such that the extension in sweep time is determined by the minimum of τ_e or τ_l. Typically, a sweep will be significantly extended provided the additional time $\min(\tau_e, \tau_l) \gg \ln[Ns_b]/s_b$. Using the expressions for τ_e and τ_l from Eqs. **10** and **11** in the main text and rearranging, we can again cast this in terms of a base pair distance,

$$l_e = \frac{s_b}{r\alpha \ln[Ns_b]}, \quad \text{[S35]}$$

which is the distance to a recessive deleterious mutation within which a new beneficial mutation must land to have its sweep time significantly extended.

3.2. Strong Drift, Weak Selection ($\alpha < 1$).

Zone of suppressed probability of fixation. The probability of fixation in this regime, from Eq. **6** and *Predictions for the Regime of Weak Selection and Strong Drift ($\alpha \ll 1$)*, is

$$\Pr(\text{fixation}) \approx rls_b \int_0^{\sqrt{N/s_d}} dt = rls_b \sqrt{N/s_d}, \quad \text{[S36]}$$

where the upper limit in the integral comes from the fact that it is highly unlikely for a *BD* haplotype to drift for longer than $\sim \sqrt{N/s_d}$ generations. Compared with a beneficial mutation with no hitchhikers (where probability of fixation is $\sim s_b$), a recessive deleterious hitchhiker will significantly suppress the probability of fixation whenever $rl\sqrt{N/s_d} < 1$. This can be used to define a base pair distance, l_l, around any recessive deleterious mutation within which a new beneficial mutation of effect size $\sim s_b$ (or smaller) has a reduced chance of fixation,

$$l_l = \frac{1}{r}\sqrt{\frac{s_d}{N}}. \quad \text{[S37]}$$

Zone of increased sweep time. In this regime, there can be no significant increase in sweep time, because the loss time is always small relative to the sweep time.

3.3. Quantifying Coding Gene Density for *Drosophila melanogaster* and Humans. Given that deleterious regions are likely to occur in functional regions, and that functional regions are likely to be clustered in the genome (particularly for humans), we framed our estimates of functional density in terms of the density of coding genes around every coding gene. For this analysis in humans, we downloaded the University of California, Santa Cruz (UCSC) knownCanonical table of genes and their corresponding positions, where each gene is only represented by one isoform, and the set of gene predictions are "based on data from RefSeq, GenBank, Rfam, and the tRNA Genes track. . .This is a moderately conservative set of predictions" (UCSC website, genome.ucsc.edu/cgi-bin/hgTables). We then excluded all genes that do not appear in the knownGenePep list (i.e., excluded noncoding) and excluded all genes that are not located on autosomes. Then, using the annotated locations of all genes in this set, we quantified the number of coding genes that fall within a window centered around the midpoint of every coding gene, where the window sizes were defined by our zone sizes. Note that this means, for example, we had 19,353 data points for coding gene densities because there are currently 19,353 coding genes annotated on the human autosomes. Results can be seen in Fig. S1 *A* and *B*.

We did a similar analysis for *Drosophila*, where we downloaded from flybase.org the dmel-all-gene- list, excluded all genes that do not appear on chromosomes 2 or 3, excluded all genes that do not appear in the dmel-all-translation- list (i.e., excluded noncoding), and only used the first isoform listed for every gene (i.e., if multiple isoforms were present, we used the -PA isoform). Then, using the annotated locations of all genes in this set, we quantified the number of coding genes that fall within a window centered around the midpoint of every coding gene. In this case, for *Drosophila*, there are currently 11,631 annotated coding genes on chromosomes 2 and 3. Results can be seen in Fig. S1 *C–F*.

3.4. Estimates for *Drosophila melanogaster* and Humans for the Proportion of the Genome in Which the Fixation Probability of Beneficial Mutations Is Reduced. Table 1 gives predictions for the proportion of coding genes in the *Drosophila* and human genomes that are subject to reduced fixation probabilities for new beneficial mutations. These predictions are based on coding gene densities (described in section 3.3) and the predicted zone size around the deleterious mutations (described in sections 3.1 and 3.2). Because this zone is maximized in the $\alpha = Ns_b^2/s_d \leq 1$ regime, we used a zone size of $l_l = \sqrt{s_d/N}/r$. Note that this is independent of the beneficial mutation effect size, and thus all beneficial mutations that satisfy $s_b \leq \sqrt{s_d/N}$ behave similarly, emphasizing that weakly adaptive mutations may be particularly susceptible to the effects of hidden recessive deleterious variation in the genome. Details of how we constructed Table 1 are found below. Method for Table 1:

Column 1 = Organism and assumed population size

Column 2 = Number of coding genes on autosomes, see section 3.3

Column 3 = Recessive deleterious effect size of interest

Column 4 = l_l for $\alpha < 1 = \sqrt{s_d/N}/r$

where brackets indicate number of genes which appear in this zone, given by section 3.3

Column 5 = the zone size is the same for all $s_b \leq \sqrt{s_d/N}$

Column 6 = n_d/g_c where

n_d = number of recessive deleterious mutations in a haploid set of autosomes (2–4)
g_c = number of coding genes in reference genome

Column 7 = column 4 brackets × column 6.

3.5. Estimates for *Drosophila melanogaster* and Humans for the Proportion of the Genome in Which the Sweep Time of Beneficial Mutations Is Extended. If we consider the beneficial mutations that do reach fixation, we find that a substantial proportion of the human and *Drosophila* genomes will cause there to be a staggered phase during the beneficial mutation's selective sweep, particularly due to more mildly deleterious recessive mutations (Table S1). Although the impact of recessive deleterious variation on the probability of fixation of beneficial mutations can be simplified in terms of which mutations fall within the $\alpha = Ns_b^2/s_d < 1$ regime, the potential for extended sweep times is not as easily simplified. Extensions in sweep times do not occur in the $\alpha < 1$ regime; instead, genomes will be most affected by staggered sweeps for intermediate values of α. In this regime, the adaptive mutation can reach a stable equilibrium frequency ($\alpha \not\ll 1$) where it has a slow rate of loss ($\tau_l > \ln[Ns_b]/s_b$), but the rate of new recombinants being generated in the population is not so large that the staggered sweep is resolved quickly ($\tau_e > \ln[Ns_b]/s_b$). One example where this regime likely applies is experimental evolutions (in obligately sexual diploids, like fruit flies), where the beneficial effect sizes may be strong but the population size is small. For example, in a population of $N = 1{,}000$ flies, a beneficial mutation with effect size $s_b = 1\%$ will be subject to staggered sweep phases within ~65% of coding genes.

To give a better sense of the parameter range affected, see Fig. S1 *G–I*, where *G* and *H* are for more mildly deleterious mutations (which can affect substantial portions of the genome), and *I* is for recessive lethals (which do not have a substantial effect). Experimental evolutions in *Drosophila* correspond to blue and purple lines in Fig. S1 *G–I*. Details of how these plots were constructed can be found below. Note that the plots are stepwise due to the measurements of the clustering of genes (zone size is rounded to the nearest order of magnitude and translated to the mean number of genes within that distance; see section 3.3). Method for Table S1:

Column 1 = Organism and assumed population size

Column 2 = Number of coding genes on autosomes, see section 3.3

Column 3 = Recessive deleterious effect size of interest

Column 4 = $\dfrac{s_b}{r\alpha \ln[Ns_b]}$

where brackets indicate number of genes which appear in this zone, given by section 3.3

Column 5 = Beneficial effect size of interest

Column 6 = n_d/g_c where

n_d = number of recessive deleterious mutations in a haploid set of autosomes (2–4)
g_c = number of coding genes in reference genome

Column 7 = column 4 brackets × column 6.

Our method for plots in Fig. S1 *G–I* was as follows: For Fig. S1 *G–I*, the *y* axis is count × density. Count denotes the number of genes that appear in region l_e (see section 3.3); $l_e = s_b/r\alpha\ln[Ns_b]$, where r = 10–8 and $\alpha = Ns_b^2/s_d$. Density is 1/30 for $s_d = 0.01$, 1/12,000 for dmel $s_d = 1$, and 1/20,000 for human $s_d = 1$. The *x* axis is s_b, where minimum s_b is set by $\alpha = Ns_b^2/s_d$, due to moving into the $\alpha < 1$ regime where no substantial extension in sweep time occurs, and maximum s_b is set by s_d, due to the beneficial mutation being stronger than the deleterious mutation.

4. Simulations

Two-locus Wright–Fisher forward simulations of an adaptive mutation genetically linked to a recessive deleterious mutation were performed to test our predictions. A wide range of parameters were varied, including selection coefficients ($s_b = 0.001 - 0.1$, $s_d = 0.01 - 1$), recombination rates ($c = r/l = 10^{-8} - 10^{-1}$), and diploid population sizes ($N = 10^2 - 10^6$), corresponding to ranges of $\alpha = Ns_b^2/s_d = 10^{-3} - 10^5$. All simulations used $s_b \leq s_d$ and heterozygous effects of $h_d = 0$ and $h_b = 0.5$ such that the equilibrium frequency $p^* \approx s_b/2s_d \leq 1/2$. We also included a one-locus control beneficial mutation (with no hitchhiker) for comparison. Diploid fitness was calculated as multiplicative across loci (as in Eqs. **S1**–**S3**). Simulations tracked frequencies of haplotypes (i.e., *BD*, *OO*, *BO*, *DO*), where each generation consisted of recombination and selection in diploids, and then drift of haplotypes (i.e., multinomial sampling). Simulations concluded when the beneficial mutation approached fixation or extinction (i.e., $p \gg 1 - 1/2N$ or $p \ll 1/2N$).

4.1. Extinction Times in the $\alpha > 1$ Regime. These simulations seeded the *BD* haplotype at frequency p^* and the *OO* haplotype at frequency $1 - p^*$, used zero recombination rate, and then recorded the generation of extinction of the beneficial mutation (i.e., the *BD* haplotype). A range of parameters were used, and each parameter set was performed for 1,000 simulations. Parameters were chosen such that $\alpha = Ns_b^2/s_d > 1$, $Ns_d > 1$ and mean time to extinction $\tau_l = 1/\lambda_l \gg \ln[Ns_b]/s_b$ (to ensure extinction times are not confounded with the time it typically takes to traverse from p^* to 0).

The extinction times are predicted to be exponentially distributed with rate

$$\lambda_l \approx \frac{e^{-\alpha}}{(1/s_b)\ln(\alpha)} \quad \textbf{[S38]}$$

called the "rate of loss," where

$$\alpha = Ns_b^2/s_d. \quad \textbf{[S39]}$$

Results of a few sample parameter sets can be seen in Fig. S2, in which both histograms of extinction times (Fig. S2 *A*–*C*) and QQ plots of exponential quantiles (Fig. S2 *D*–*F*) highlight that the extinction times indeed look to be exponentially distributed (with our predicted rate parameter indicated in red).

To test whether the rate of loss $\lambda_l \approx \exp(-\alpha) * (s_b/\ln(\alpha))$ from Eq. **S24** accurately captures the correct scaling, or more specifically is truly exponential in the parameter $\alpha = Ns_b^2/s_d$, we performed a series of simulations in which a rate of loss was inferred from the distribution of extinction times using the R function fitdistr (from the downloadable R statistics package "MASS"). If λ_l is indeed exponential in α, then we would expect that if we hold all parameters constant except one (thus changing α via a single parameter), this observed rate of loss would behave such that

$$\text{rate} = \exp[-\alpha] * (s_b/\ln[\alpha]) \quad \textbf{[S40]}$$

$$\text{rate} * (\ln[\alpha]/s_b) = \exp[-\alpha] \quad \textbf{[S41]}$$

$$\ln[\text{rate} * (\ln[\alpha]/s_b)] = -\alpha. \quad \textbf{[S42]}$$

This is shown in Fig. S3, where changing a single variable (either N, s_b, or s_d) affects the observed rate of loss only via the parameter α, indicating that λ_l accurately captures the scaling in the exponent. The term λ_l will predict the scaling for the rate of loss but not its exact form, so for comparison with simulations, we fit constants to a subset of simulations [finding $\lambda_l = 0.1\exp(-0.3\alpha) * (s_b/\ln(\alpha))$] and used these coefficients throughout the rest of our predictions for the rate of loss, the probability of fixation, and the sweep time in simulations.

4.2. Escape Times in the $\alpha > 1$ Regime. These simulations seeded the *BD* haplotype at frequency p^* and the *OO* haplotype at frequency $1 - p^*$, used nonzero recombination rates, and then recorded the generation of fixation or extinction of the beneficial mutation. A range of parameters were used, and each parameter set was performed for 1,000 simulations. We tried to choose parameters that satisfied $\alpha = Ns_b^2/s_d \gg 1$ such that drift to extinction should not confound measurements of escape time. Furthermore, we chose parameters to satisfy $Ns_d > 1$ (such that *BD* does not drift to fixation) and to satisfy $\tau_e = 1/\lambda_e \gg \ln[Ns_b]/s_b$ (such that escape times are not confounded with the time it typically takes to traverse from p^* to 1).

The escape times are predicted to be exponentially distributed with rate

$$\lambda_e \approx c\alpha \quad \textbf{[S43]}$$

called the "rate of escape," where

$$\alpha = Ns_b^2/s_d. \quad \textbf{[S44]}$$

Results of a few sample parameter sets can be seen in Fig. S4, in which both histograms of escape times (Fig. S4 *A*–*C*) and QQ plots of exponential quantiles (Fig. S4 *D*–*F*) highlight that the escape times indeed look to be exponentially distributed (with our predicted rate parameter indicated in red).

To test whether $c\alpha = cNs_b^2/s_d$ captures the correct scaling for the rate of escape, we performed a series of simulations in which a rate of escape was inferred from the distribution of fixation times using the R function fitdistr (from the library MASS). This observed rate is expected to scale linearly in response to c, N, s_b^2 and $1/s_d$. This is shown in Fig. S5, where each plot has the results from a series of simulations in which three parameters have fixed values and one parameter varies (either N, s_b, s_d, or c). The term λ_e will predict the scaling for the rate of escape but not its exact form, so, for comparison with simulations, we fit a constant using a subset of simulations (finding $\lambda_e = 0.8c\alpha$), and used this coefficient throughout the rest of our predictions for the rate of escape, the probability of fixation, and the sweep time in simulations.

4.3. Extinction Times in the $\alpha < 1$ Regime. To observe the extinction times in the $\alpha < 1$ regime, we performed 100,000 simulations in the $\alpha < 1$ regime (also requiring $Ns_d > 1$ so there is selection against the recessive deleterious mutation), seeding the *BD* haplotype at frequency $1/2N$ and the *OO* haplotype at frequency $1 - 1/2N$, and then recording the generation of extinction. Results are shown in Fig. S6. One way to view a power law is to plot the complementary cumulative distribution on a log–log plot, such that $\ln(\Pr[T > t]) = \ln(t)$. The neutral expectation is indicated by the black line, and the distribution of extinction times for neutral simulations are shown in the light blue histogram (note that it follows the neutral expectation, although finite sampling leads to poorer resolution at the tip of the tail). The distribution of extinction times from simulations in the $\alpha < 1$ regime are shown in the dark blue, pink, and yellow histograms, where they follow the neutral expectation up until a cutoff time that scales with $\sim\sqrt{N/s_d}$. The term $\sqrt{N/s_d}$ will predict the scaling for the extinction time but not its exact form, so, for comparison with simulations, we fit a constant to a subset of simulations (finding $4\sqrt{N/s_d}$) and used this coefficient throughout the rest of our predictions for the loss time when $\alpha < 1$, the probability of fixation, and the sweep time in simulations.

4.4. Probability of Fixation of the Beneficial Mutation (Both α Regimes). These simulations seeded the haplotype at frequency $p=1/2N$ and the *OO* haplotype at frequency $1-1/2N$, used nonzero recombination rates, and then recorded the generation of fixation or extinction of the beneficial mutation. A range of parameters were used, and each parameter set was performed for $1{,}000/s_b$ simulations such that a one-locus control beneficial mutation (with no deleterious hitchhiker) is expected to fix in ∼1,000 simulations.

For comparison with simulation results, we used the following analytic predictions for the probability of fixation of a beneficial mutation that enters a population on a *BD* haplotype:

$$\Pr(\text{fixation}) \approx \begin{cases} \left(\frac{1-e^{-s_b}}{1-e^{-2Ns_b}}\right)\left(\frac{\lambda_e}{\lambda_e+\lambda_l}\right)\left(1-e^{-(\lambda_e+\lambda_l)t_{max}}\right) & (\text{for } \alpha>1) \\ \left(\frac{1-e^{-s_b}}{1-e^{-2Ns_b}}\right)\left(1-e^{-c4\sqrt{N/s_d}}\right) & (\text{for } \alpha<1) \end{cases} \quad \textbf{[S45]}$$

where

$$\lambda_l \approx \frac{s_b}{\ln[\alpha]} e^{-\alpha} \quad (\text{for } \alpha>1) \quad \textbf{[S46]}$$

$$\lambda_e \approx c\alpha \quad (\text{for } \alpha>1) \quad \textbf{[S47]}$$

$$\alpha = \frac{Ns_b^2}{s_d} \text{ and } c = rl. \quad \textbf{[S48]}$$

A full panel of simulation results with analytic predictions for the probability of fixation can be seen in Fig. S7.

The first term in the probability of fixation in both regimes is the probability of establishment, where we have replaced the approximation s_b with the more exact form $1-e^{-s_b}/1-e^{-2Ns_b}$ for the sake of comparison with simulations (particularly important for small population sizes). Note that in the $\alpha>1$ regime, the third term is the probability that escape occurs before the simulations end at time t_{max} (due to not being able to run simulations for infinite generations). Simulations ran for, at most, $t_{max}=10^9$ generations.

A beneficial mutation of effect s_b will have a decreased probability of fixation if it falls within a recombination distance c_l of the recessive deleterious mutation with effect s_d in a population of N. In the $\alpha>1$ regime, this is found by setting $\lambda_l=\lambda_e$, and in the $\alpha<1$ regime, this is found by setting $c=1/\sqrt{N/s_d}$. For comparison with simulations, we thus use the following analytic predictions:

$$c_l \approx \begin{cases} \max\left(\frac{s_b}{r\alpha\ln[\alpha]e^{\alpha}}, \frac{1}{\alpha\, t_{max}}\right) & (\text{for } \alpha>1) \\ \sqrt{s_d/N} & (\text{for } \alpha<1) \end{cases}. \quad \textbf{[S49]}$$

These predictions for c_l can be seen as the vertical dashed gray lines in Fig. S7.

4.5. Sweep Time of the Beneficial Mutation (Both α Regimes). These simulations seeded the *BD* haplotype at frequency $p=1/2N$ and the *OO* haplotype at frequency $1-1/2N$, used nonzero recombination rates, and then recorded the generation of fixation or extinction of the beneficial mutation. A range of parameters were used, and each parameter set was performed until 500 fixation events occurred.

For comparison with simulation results, we used the following analytic predictions for the sweep time of a beneficial mutation that enters a population on a *BD* haplotype:

$$\text{total sweep time} \approx \begin{cases} \frac{4\ln[2Ns_b]}{s_b} + \left(\frac{1}{\tau_e}+\frac{1}{\tau_l}\right)^{-1} & (\text{for } \alpha>1) \\ \frac{4\ln[2Ns_b]}{s_b} + \left(\frac{1}{\tau_e}+\frac{1}{4\sqrt{N/s_d}}\right)^{-1} & (\text{for } \alpha<1) \end{cases}. \quad \textbf{[S50]}$$

A full panel of simulation results can be seen in Fig. S8. Note that the first term (in both regimes) has added factors of 2 that come from the sweep time predictions for diploids. The second term (in both regimes) is derived from the faster of two processes—either the escape time (τ_e) or the loss time (τ_l for $\alpha>1$ and $\sqrt{N/s_d}$ for $\alpha<1$).

The sweep time will not be changed if $\ln[Ns_b]/s_b$ is larger than τ_e, τ_l, or $\sqrt{N/s_d}$. If $\ln[Ns_b]/s_b < \tau_e < (\tau_l \parallel \sqrt{N/s_d})$, then escape will generally occur before extinction of the *BD* haplotype, and the sweep will be extended by τ_e. If $\ln[Ns_b]/s_b < (\tau_l \parallel \sqrt{N/s_d}) < \tau_e$, then loss occurs before escape; however, we are conditioning on fixation, and thus, in the rare instances where the beneficial mutation does reach fixation, it must do so in a time τ_l or $\sqrt{N/s_d}$ (depending on the regime). For example, in Fig. 4 *C* and *D*, we have a case of $\ln[Ns_b]/s_b < \tau_l$, and thus there is still an extension in the sweep time (seen as the leveling off of sweep times at low recombination rates).

A beneficial mutation of effect s_b will have an increased sweep time if it falls within a recombination distance c_e of the recessive deleterious mutation with effect s_d in a population of N. This can be found by setting $1/\lambda_e = \ln[Ns_b]/s_b$; thus, for comparison with simulations, we use the following analytic predictions:

$$c_e = \frac{s_b}{\alpha\ln[Ns_b]} \quad (\text{for } \alpha>1 \text{ and } \alpha<1). \quad \textbf{[S51]}$$

These predictions for c_e can be seen as the vertical dashed red lines in Fig. S8. Note that in the $\alpha<1$ regime, the sweep time is not expected to be substantially altered, because, in this regime, where $\alpha=Ns_b^2/s_d<1$, the beneficial mutation must escape in a time $\sqrt{N/s_d}$, which is generally less than $\ln[Ns_b]/s_b$.

4.6. Testing Robustness of Model if Deleterious Allele Segregating at High Frequency. Our model predictions should be robust to cases where the recessive deleterious allele is segregating on an *OD* haplotype in many individuals in the population, because, in our case of $N\mu_d<1$, the deleterious mutation will typically reach, at most, $\sqrt{N/s_d}$ copies, or a frequency of $1/\sqrt{Ns_d}$. To understand this, consider that the deleterious variation will reduce the mean (log) fitness advantage of a new *BD* haplotype, due to causing the *BD* haplotype to now appear in *BD*/*OD* diploids in addition to appearing in *BD*/*OO* diploids,

$$\overline{w}_{BD} \approx s_b \times (\text{frequency of } OO) + (s_b - s_d) \times (\text{frequency of } OD) \quad \textbf{[S52]}$$

$$\overline{w}_{BD} \approx s_b\left(1-\frac{1}{\sqrt{Ns_d}}\right) + (s_b-s_d)\left(\frac{1}{\sqrt{Ns_d}}\right)$$

$$\overline{w}_{BD} \approx s_b - \frac{s_b}{\sqrt{Ns_d}} + \frac{s_b}{\sqrt{Ns_d}} - s_d\frac{1}{\sqrt{Ns_d}} \quad \textbf{[S53]}$$

$$\overline{w}_{BD} \approx s_b - \sqrt{\frac{s_d}{N}}. \quad \textbf{[S54]}$$

Thus, the mean selective advantage of the *BD* haplotype can be substantially reduced by segregating deleterious variation when $s_b \lesssim \sqrt{s_d/N}$, which, interestingly, is when $\alpha=Ns_b^2/s_d \lesssim 1$. However, as we have shown, whenever $\alpha<1$, the selective advantage

of the *BD* is irrelevant because the dynamics are determined by drift and selection against the recessive deleterious allele. Therefore, having multiple copies of an *OD* haplotype will not change our results. We confirmed that our analytics still hold in such a case, using simulations that seeded an *OD* haplotype at $\sqrt{N/s_d}$ copies when the *BD* haplotype appears.

5. Additional Signatures of Staggered Sweeps

Some additional statistics were calculated for the diversity around a beneficial mutation that had a recessive deleterious hitchhiker during its sweep to fixation (Fig. S9). Hard sweeps were simulated by seeding a beneficial mutation with effect s_b on a single haplotype at establishment frequency, staggered sweeps by doing the same but where the single haplotype also contained a recessive deleterious mutation with effect s_d a distance l away, and soft sweeps by seeding the beneficial mutation on a new haplotype at establishment frequency every t generations, where t was drawn from an exponential distribution with rate s_b (thus $\theta \approx NU_b = 1$ beneficial mutation entering the population every generation with an establishment probability of $\sim s_b$).

1. Maruyama T (1974) The age of an allele in finite population. *Gen Res* 23(2):137–143.
2. Kusakabe S, Yamaguchi Y, Baba H, Mukai T (2000) The genetic structure of the Raleigh natural population of *Drosophila melanogaster* revisited. *Genetics* 154(2): 679–685.
3. Gao Z, Waggoner D, Stephens M, Ober C, Przeworski M (2015) An estimate of the average number of recessive lethal mutations carried by humans. *Genetics* 199(4):1243–1254.
4. Latter BD (1998) Mutant alleles of small effect are primarily responsible for the loss of fitness with slow inbreeding in *Drosophila melanogaster*. *Genetics* 148(3):1143–1158.

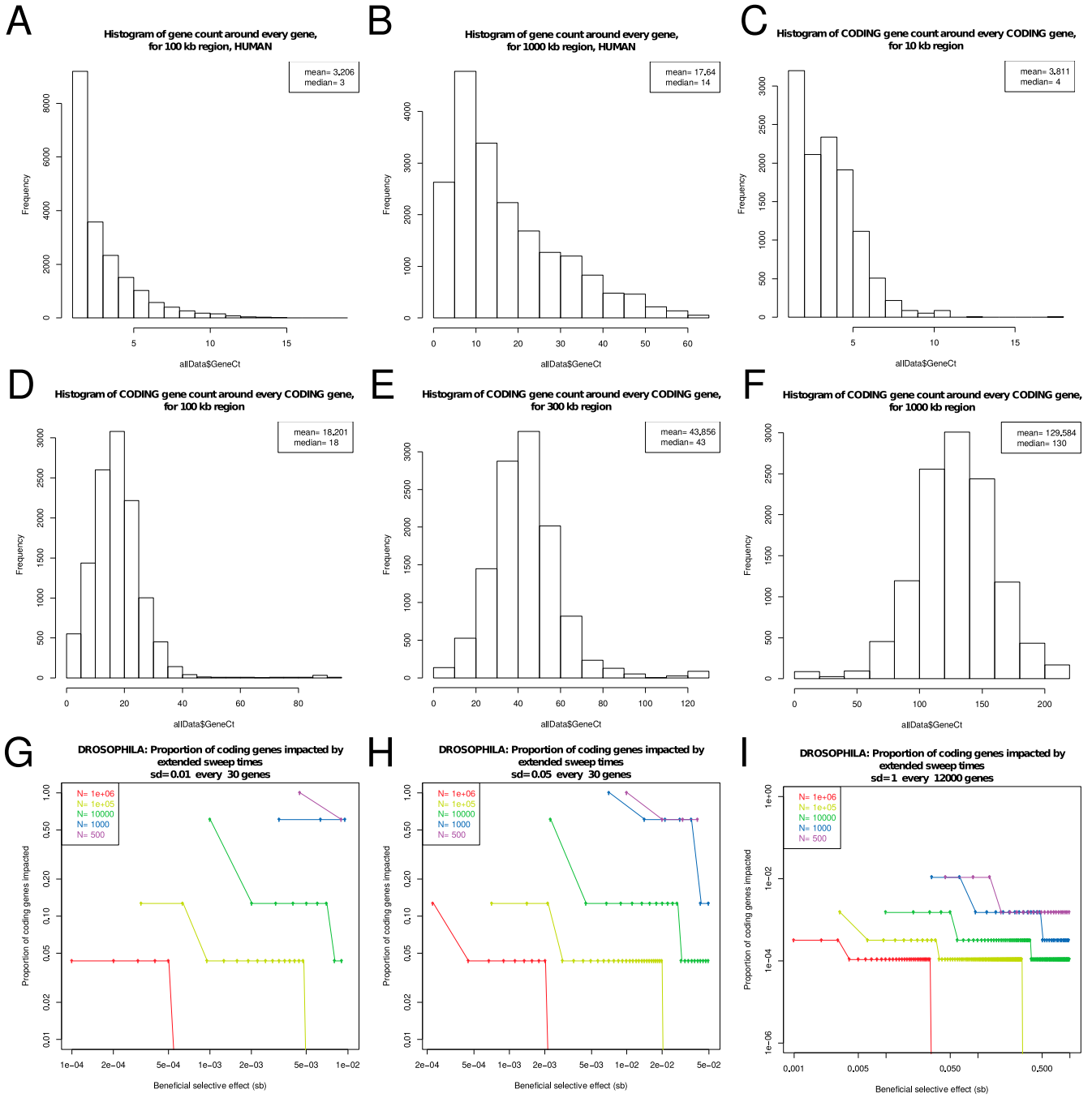

Fig. S1. Histograms of the density of coding genes around every coding gene in the reference genomes of human and *Drosophila* (*A*–*F*) and predictions for the proportion of coding genes affected in the *Drosophila* genome by staggered phases as a function of s_b (*G*–*I*). (*A* and *B*) For humans, the number of coding genes which appear in a (*A*) 100-kb and (*B*) 1-Mb window (respectively), where 19,353 windows, each centered on the midpoint of a coding gene, were used (per window size). (*C*–*F*) For *Drosophila*, the number of coding genes that appear in a (*C*) 10-kb, (*D*) 100-kb, (*E*) 300-kb, and (*F*) 1-Mb window (respectively), where 11,631 windows, each centered on the midpoint of a coding gene, were used (per window size). (*G*–*I*) Proportion of coding genes in the *Drosophila* genome impacted by extended staggered phases due to deleterious mutations with effect sizes (*G*) 1%, (*H*) 5%, and (*I*) 100%, as a function of s_b where the minimum is set by when $\alpha = 1$ and the maximum is set by $s_b = s_d$, and where each line indicates calculations for a different population size.

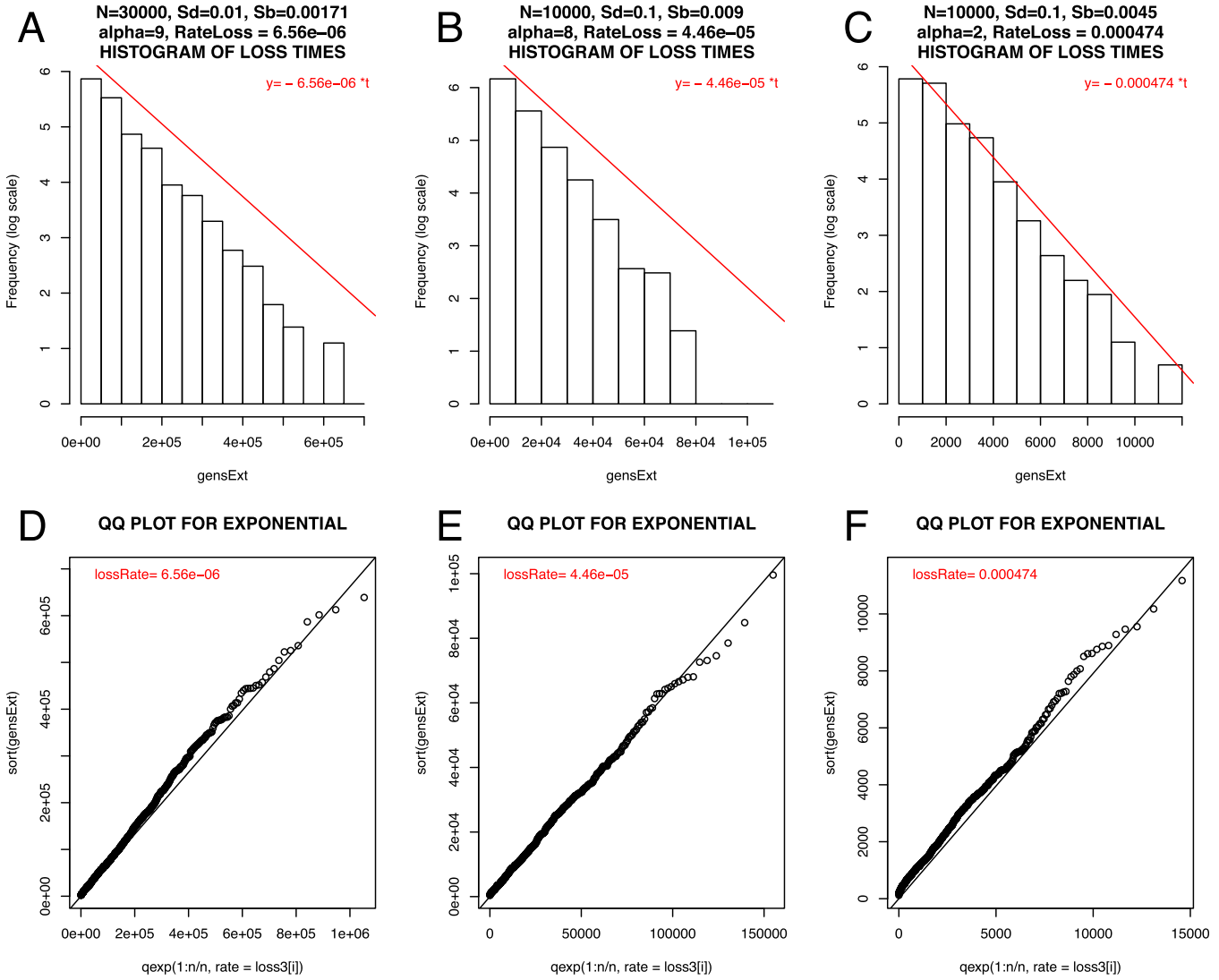

Fig. S2. Histograms and QQ plots demonstrating that the extinction times in the $\alpha > 1$ regime are exponentially distributed (a fuller set of simulation results across parameter regimes can be seen in Fig. S3). *A–C* consist of histograms of extinction times from 1,000 simulations (the parameter set used is indicated in the histogram title), such that the *x* axis is the extinction time *t* and the *y* axis is the frequency of events on a log scale. Note that if extinction times are indeed exponentially distributed with rate λ_l, then we expect these histograms to be described by the line $log(y) \approx -\lambda_l x$, which is indeed the case and can be seen by the red line, which is an exponential density curve with the analytically predicted rate parameter λ_l. Note that short extinction times ($1/\lambda_l = \tau_l \ll \ln[Ns_b]/s_b$) will not be exponentially distributed, sometimes causing an uptick in the distribution for extinction times near zero. *D–F* consist of the corresponding QQ plot for each histogram above it, where the *x* axis is the theoretical quantiles for an exponential distribution with rate λ_l and the *y* axis is the sample quantiles of extinction times obtained from simulations. The fact that the points lie along the straight line indicate that the extinction times are likely exponentially distributed.

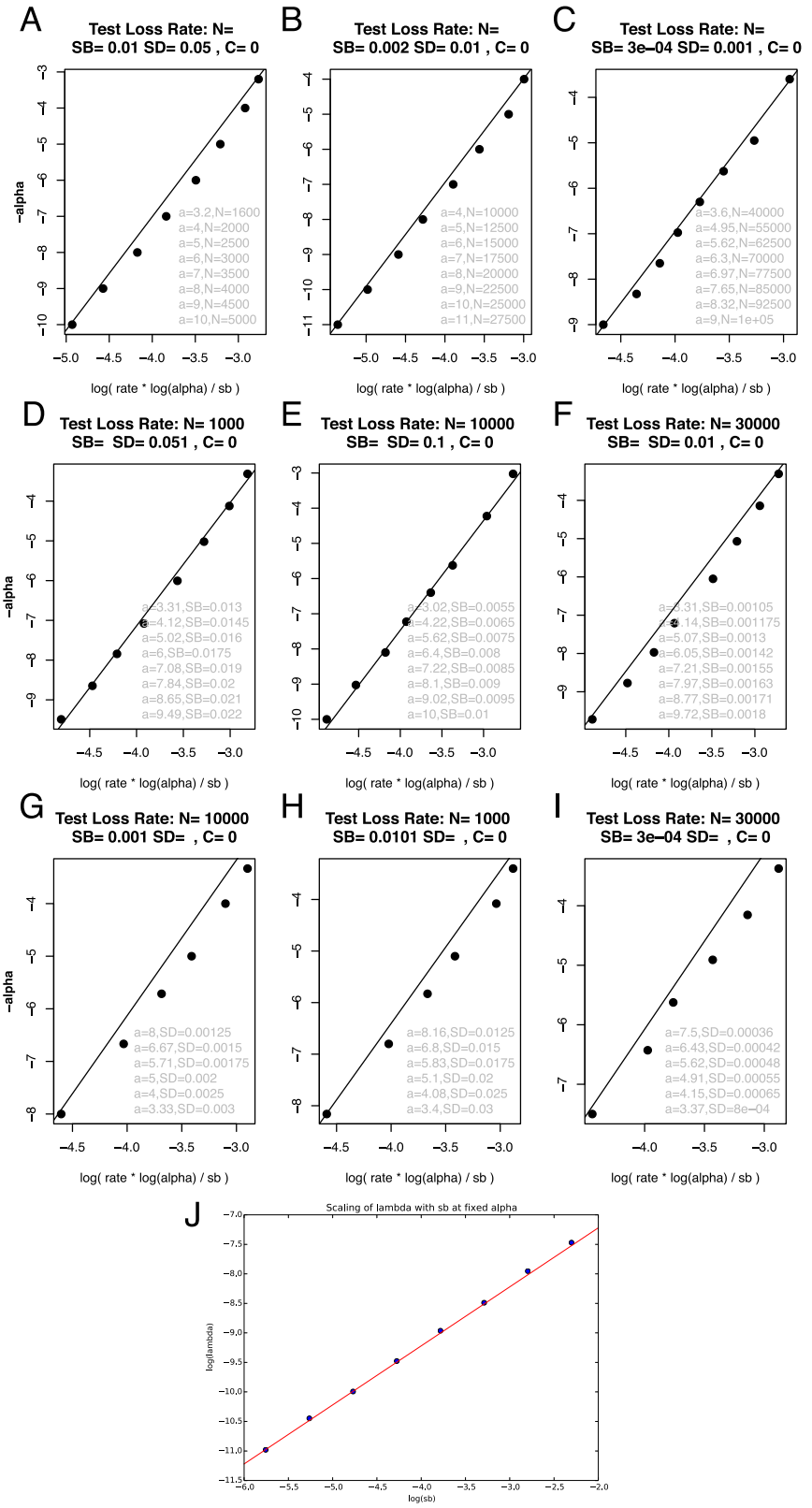

Fig. S3. Testing whether the rate of loss is indeed exponential in the parameter $\alpha = Ns_b^2/s_d$, as described by Eq. **S43**. Each subplot contains results from a set of simulations in which two parameters were held constant and one parameter varied; however, the observed rate of loss will be predicted through the parameter α (see Eq. **S43**). The rate of loss was inferred from 1,000 simulations, where an exponential distribution was fit to the extinction times using the R function fitdistr from the package MASS. Parameter values used (for N, s_b, s_d, and always $c = 0$) are indicated in the heading of each subplot (with blank values for the parameter varied, details for which are found in the gray text). For example, in *A*, each black data point indicates the results from 1,000 simulations in which $s_b = 0.01$ and $s_d = 0.05$ and N is varied, such that the value of N used is indicated by the gray text with the corresponding α (written "a") next to it. *A*–*C*

Legend continued on following page

use fixed values of s_b and s_d (values indicated in headings) and varies N (values indicated in gray text), *D–F* use fixed N and s_d and varied s_b, and *G–I* use fixed N and s_b and varied s_d. *J* verifies the s_b dependence of λ by plotting the rate of decay λ at fixed α over a range of s_b. These plots show that the rate of loss observed in simulations is indeed exponential in α.

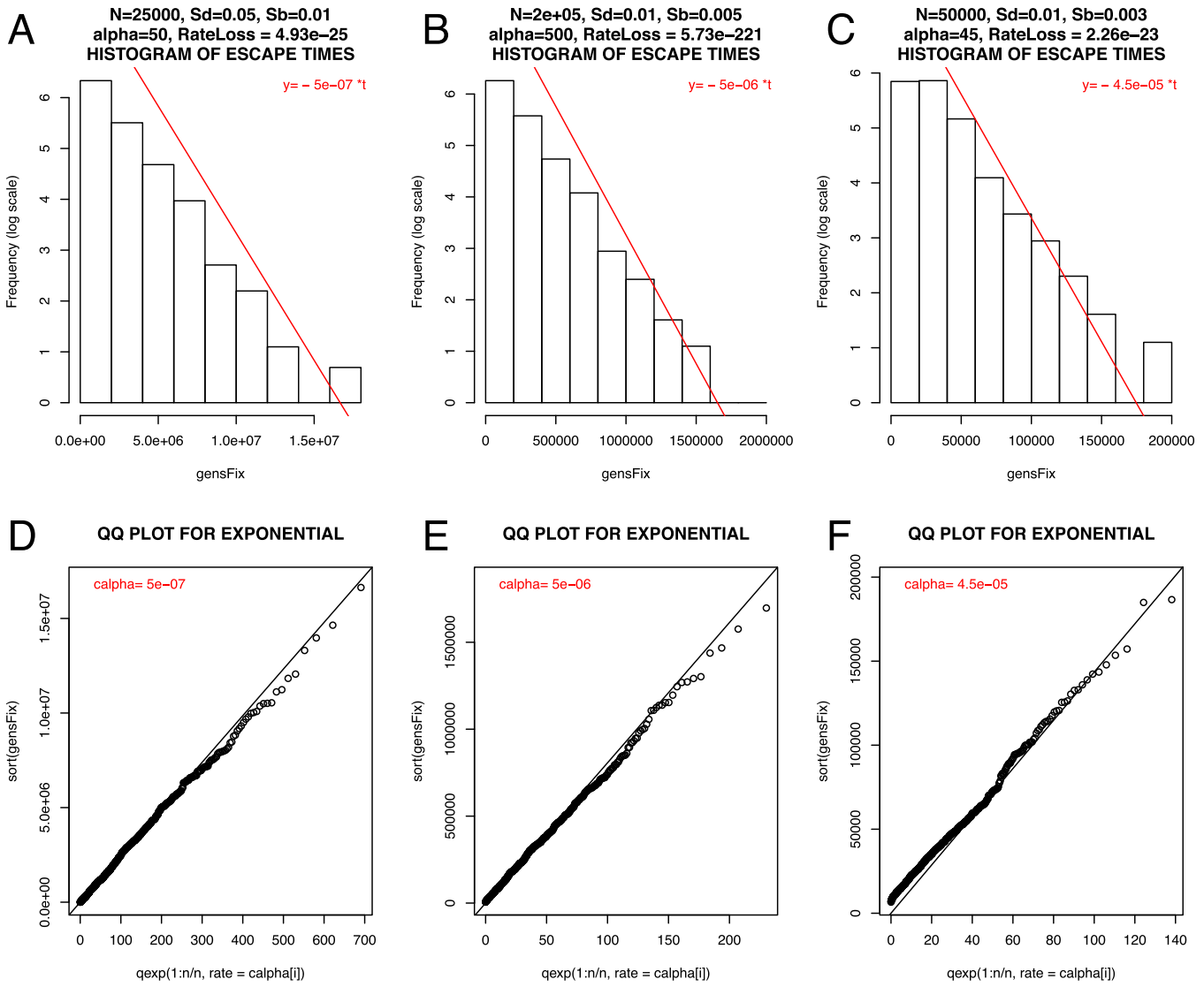

Fig. S4. Histograms and QQ plots demonstrating that the escape times in the $\alpha > 1$ regime are exponentially distributed (a fuller set of simulation results across parameter regimes can be seen in Fig. S5). *A–C* consist of histograms of escape times from 1,000 simulations, such that the *x* axis is the escape time *t* and the *y* axis is the frequency of events on a log scale, and where the parameter set used is indicated in the histogram title such that *A* uses $N = 25{,}000$, $s_b = 0.01$, $s_d = 0.05$; *B* uses $N = 200{,}000$, $s_b = 0.005$, $s_d = 0.01$; and *C* uses $N = 50{,}000$, $s_b = 0.003$, $s_d = 0.01$. Note that if escape times are indeed exponentially distributed with rate λ_e, then we expect these histograms to be described by the line $log(y) \approx -\lambda_e x$, which is indeed the case and can be seen by the red line, which is an exponential density curve with the analytically predicted rate parameter λ_e. Note that short escape times ($1/\lambda_e = \tau_e \ll \ln[Ns_b]/s_b$) will not be exponentially distributed, sometimes causing an uptick in the distribution for escape times near zero. *D–F* consist of the corresponding QQ plot for each histogram above it (such that *D* uses the same variables as *A*, etc.), where the *x* axis is the theoretical quantiles for an exponential distribution with rate λ_e and the *y* axis is the sample quantiles of escape times obtained from simulations. The fact that the points lie along the straight line indicate that the escape times are likely exponentially distributed.

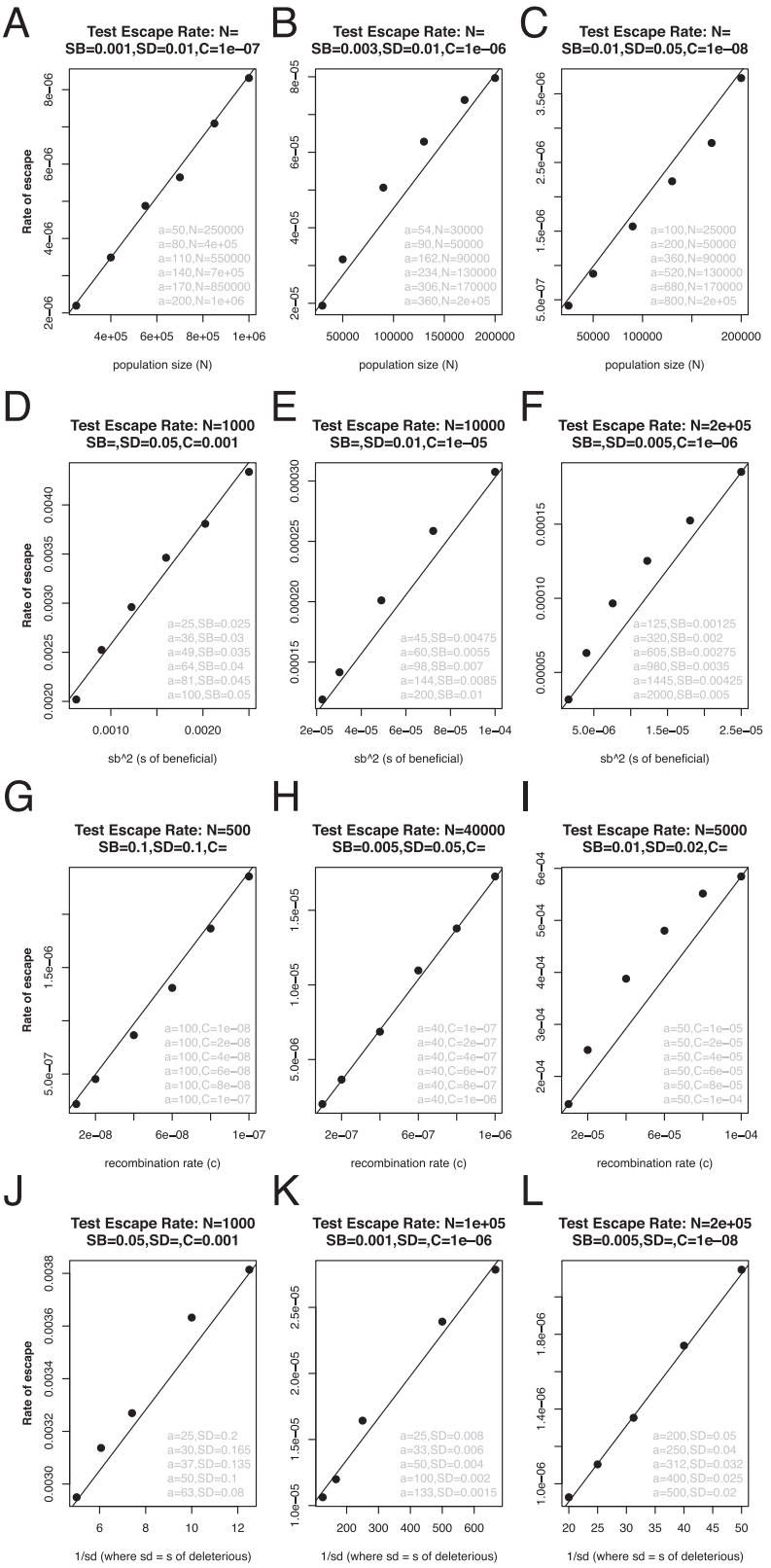

Fig. S5. Testing whether the rate of escape indeed scales with $c\alpha = Ns_b^2/s_d$. Each subplot contains results from a set of simulations in which three parameters were held constant and one parameter varied, such that the observed rate of escape should scale with the varied parameter (as written on the *x* axis). Each data point indicates the results from 1,000 simulations, where an exponential distribution was fit to the escape times using the R function fitdistr from the package MASS and this inferred rate of escape recorded (*y* axis). Parameter values used (for $N, s_b, s_d, c = 0$) are indicated in the heading of each subplot (with blank values for the parameter varied, details for which are found in the gray text). For example, in *A*, each black data point indicates the results from 1,000 simulations in which $s_b = 0.001$, $s_d = 0.01$, $c = 10^{-7}$, and N is varied across data points where the value of N used is indicated by the gray text with the

Legend continued on following page

corresponding α (written "a") next to it. *A*–*D* use fixed values of s_b, s_d, and *c* (values indicated in headings) and varied *N* (values indicated in gray text), *D*–*F* use fixed values of *N*, s_d, and *c* and varied s_b, *G*–*I* use fixed values of *N*, s_b, and s_d and varied *c*, and *J*–*L* use fixed values of *N*, s_b, and *c* and varied s_d. These plots show that the rate of escape scales with $c\alpha$.

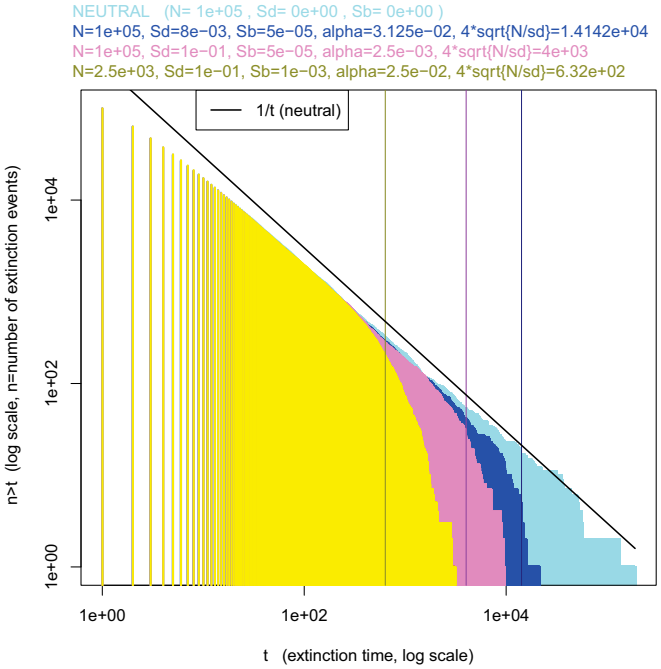

Fig. S6. Histogram of extinction times in the $\alpha < 1$ regime, where each color indicates the results from 100,000 simulations using the parameter set indicated at the top of the plot in the corresponding color. The light blue histogram indicates the results from a neutral simulation, where the black $y = 1/t$ line is the corresponding expectation for the distribution under neutrality. The dark blue, pink, and yellow histograms show that under the $\alpha < 1$ regime, the distribution of extinction times looks neutral up until a cutoff at $t \approx 4\sqrt{N/s_d}$.

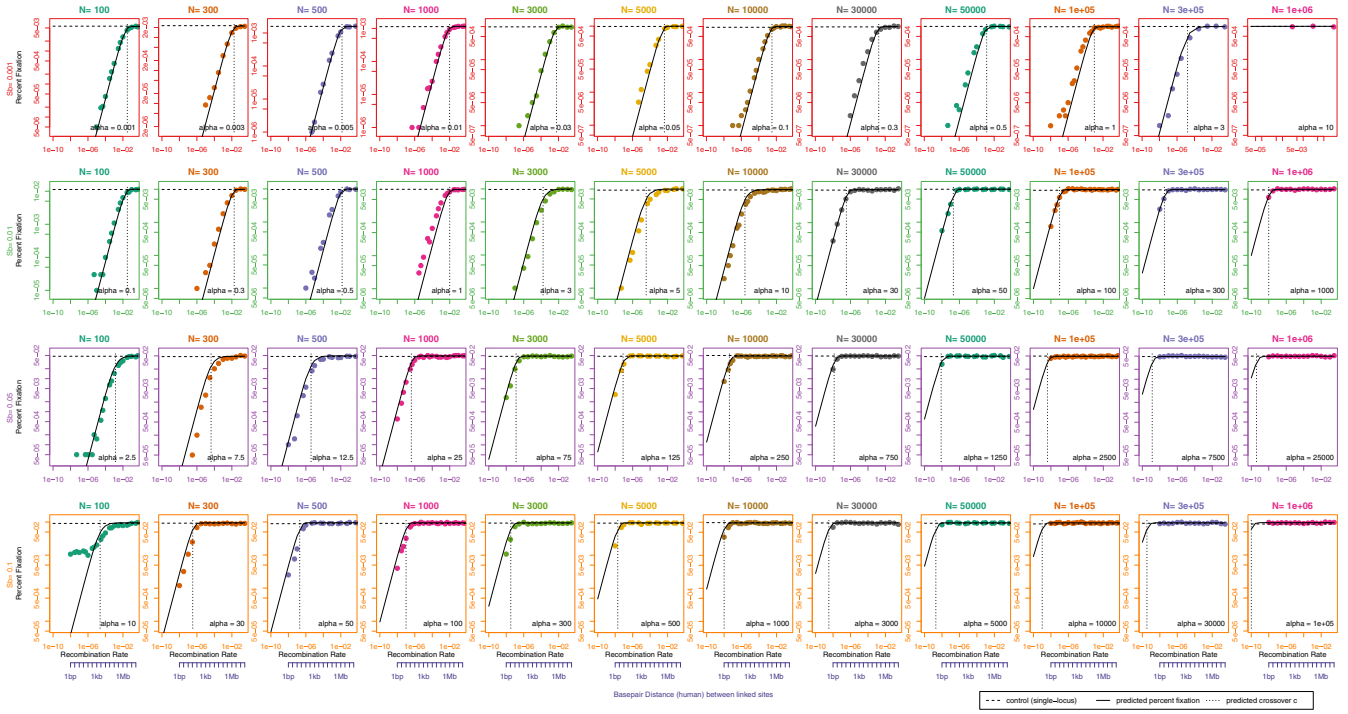

Fig. S7. Probability of fixation (*y* axis) vs. the recombination rate *c* (*x* axis). These plots show the comparison of simulations to analytics, where each data point is the result from 1,000/s_b simulations, and the horizontal dashed line is the expectation for a single beneficial mutation with no deleterious hitchhiker. Note that the population size *N* is indicated at the top of each column in a font color that corresponds to the color of the data points, the recessive deleterious effect is $s_d = 0.1$, and the beneficial mutation effect s_b is indicated at the left side of each row in a font color that corresponds to the color of the plot axes. Gray vertical dashed lines indicate the predicted crossover point c_l, such that recombination distances below this have a substantially decreased probability of fixation compared with a one-locus control beneficial mutation with no hitchhiker.

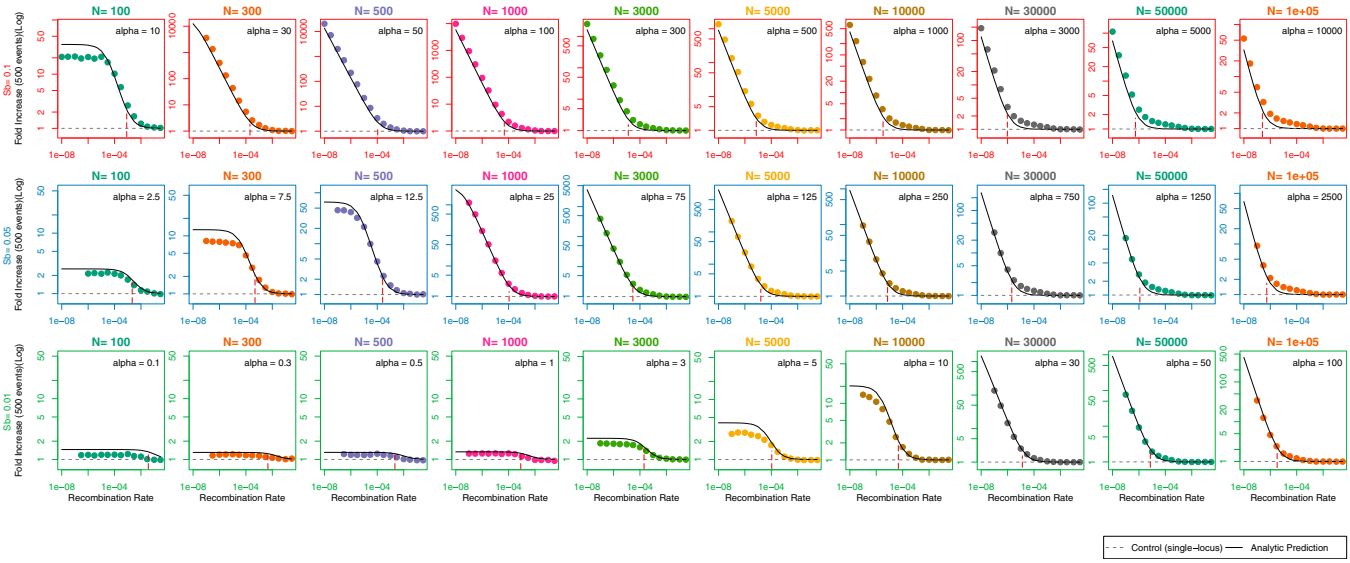

Fig. S8. Sweep time (*y* axis) vs. the recombination rate *c* (*x* axis). These plots show the comparison of simulations to analytics, where each data point is the result from 500 simulations in which fixation occurred, and the horizontal dashed line is the expectation for the sweep time of a single beneficial mutation with no deleterious hitchhiker. Note that the population size *N* is indicated at the top of each column in a font color that corresponds to the color of the data points, the recessive deleterious effect is $s_d = 0.1$, and the beneficial mutation effect s_b is indicated at the left side of each row in a font color that corresponds to the color of the plot axes. Red vertical dashed lines indicate the predicted crossover point c_e, such that recombination distances below this have an increased sweep time compared with a one-locus control beneficial mutation with no hitchhiker.

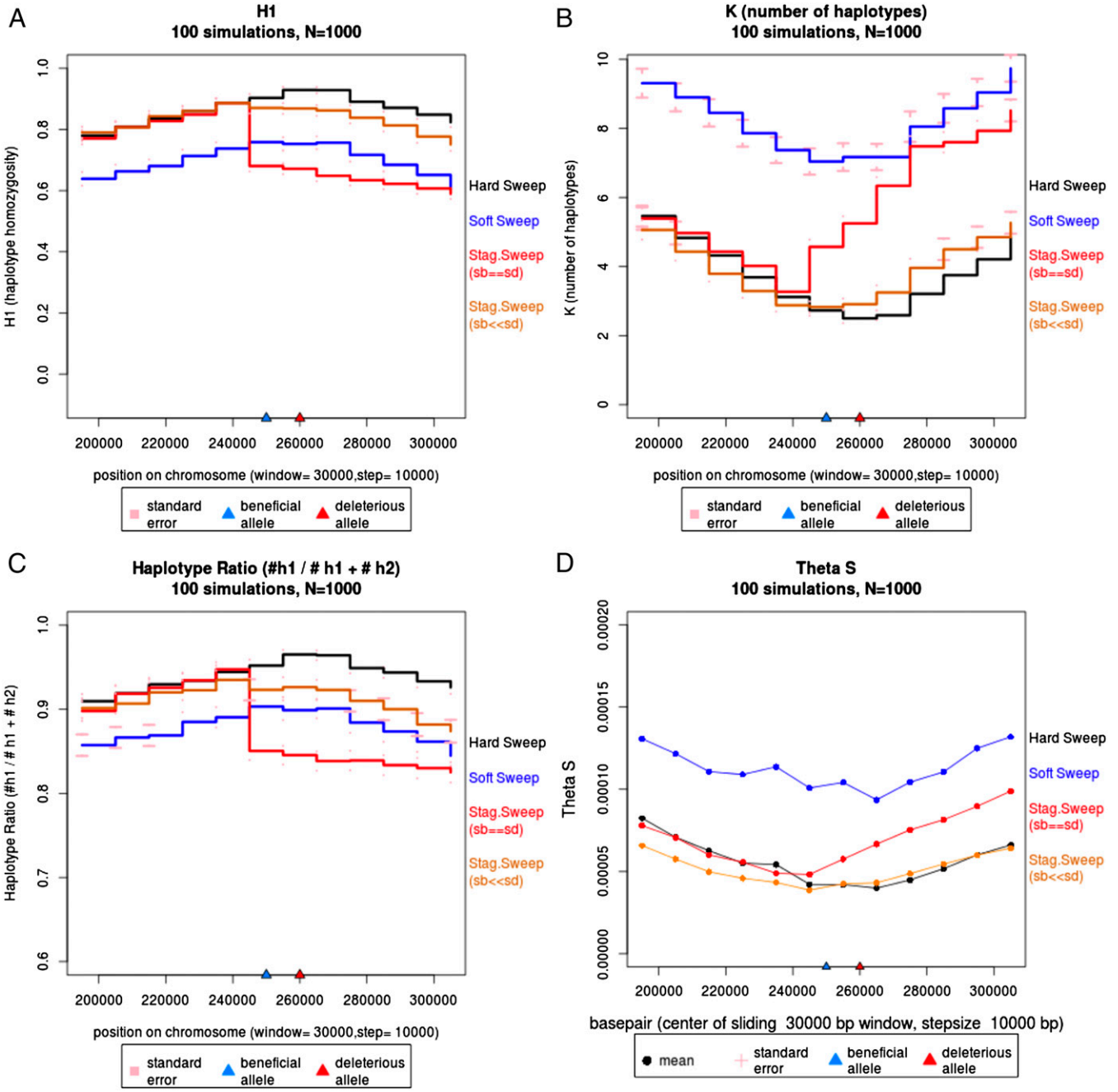

Fig. S9. Additional sweep signatures, averaged across 100 simulations using a diploid population size of N = 1,000 and a beneficial mutation effect size $s_b = 0.05$, where black lines indicate a hard sweep, blue lines indicate a soft sweep ($N\mu_b = 1$), red lines indicate a staggered sweep where $s_d = s_b = 0.05$, and orange lines indicate a staggered sweep where $s_d = 0.5$ such that $s_b \ll s_d$. For calculating statistics a window size of 30,000 base pairs with step size of 10,000 base pairs was used. Pink bars are SEMs. (*A*) The homozygosity of the most common haplotype ($H1 = p_1^2$), (*B*) the number of haplotypes in the population ("K"), (*C*) the ratio of the most common to the second-most common haplotype, and (*D*) θ_s diversity.

Table S1. Estimates for *Drosophila melanogaster* and humans for the proportion of the genome in which the sweep time of the beneficial mutation is extended

Organism and population size	Number of coding genes	Recessive deleterious effect (s_d), %	Zone of extended sweep time [genes in zone]	Beneficial effect (s_b) impacted, %	Density of recessive deleterious	Proportion of adaptive mutations impacted, %
Drosophila	~12,000					
$N = 10^6$		100	100 b [0.01 gene]	~5	1/(genome)	~0
		5	10 b [0.001 gene]	~5	1/(30 genes)	~0.003
		1	10 b [0.001 gene]	~1	1/(30 genes)	~0.003
$N = 10^3$		100	1 Mb [130 genes]	~5	1/(genome)	~1
		5	10 kb [3 genes]	~5	1/(30 genes)	~10
		1	100 kb [20 genes]	~1	1/(30 genes)	~65
$N = 10^2$		100	—	~5	1/(genome)	—
		5	1 Mb [130 genes]	~5	1/(30 genes)	~100
		1	10 Mb [10^3 genes]	~1	1/(30 genes)	~100
Human ($n = 10^4$)	~20,000	100	100 kb [3 genes]	~1	1/(genome)	~0.01
		1	100 kb [3 genes]	~0.10	1/(100 genes)	~3
					1/(30 genes)	~10
					1/(10 genes)	~30

Dashes indicate no effect on sweep time due to $\alpha \leq 1$, in which case the beneficial mutation effect size of interest is unlikely to sweep due to the recessive deleterious variation in the genome.