FILE S1

SUPPLEMENTARY INFORMATION

Contribution of doomed lineages to a fitness class:

In deriving the transition probabilities, we argue that the frequency of competing mutant lineages is frozen by the time a fitness class begins feeding mutants that are destined to establish into the nose. We arrive at this conclusion by calculating the number of establishing mutant lineages by the (typical) time this occurs. Since this number is typically large, and the contribution of each subsequent establishing lineage vanishes (despite the fact that these lineages arrive more and more quickly), the conclusion follows that the frequencies of lineages by this timepoint are frozen. This comes from the fact that the deterministic lineage size fraction, $n_L/n_1 \sim 1/L^{q/(q-1)}$ is convergent.

What could be the contribution of mutants that will not establish in a class, but may still contribute establishing mutants to the *next* fitness class? There are many more mutants introduced that will not establish relative to those that will, and these mutants grow roughly proportionally to the size of the class below them (i.e., at rate $(q-1)s$). By the time the population starts supplying establishing mutants, if these doomed lineages still comprise a significant fraction of the population, the result will be a deterministic drifting of the frequency, as these doomed lineages (which, because they are plentiful, will typically be split as x_0) become less and less of a contributing factor. We would like to determine if these combined lineages are negligible by the time a class begins feeding establishing mutants or still constituting some non-trivial fraction of the class.

The total number of individuals at time t derived from lineages that are destined to go extinct is given by

$$N_{doomed}(t) = \int_0^t \underbrace{dTU_b(1-qs)\frac{1}{qs}e^{(q-1)sT}}_{(a)} \underbrace{\langle n(t-T)|n(t\to\infty)=0\rangle}_{(b)}, \qquad (1)$$

where (a) is the expected number of non-establishing mutants occurring in a small interval $d\tau$, and (b) is the expected number of these mutants that still persist at a time $t-\tau$ later, given that these mutants are going to eventually go extinct. Here, $\langle n(t-\tau)|n(t\to\infty)=0\rangle$ is derived from standard branching process analyses.

The probability distribution of doomed (but not yet extinct) lineages at time t is given by

$$P(n>0,t|\text{doomed}) = P_{ext}(n)P(n,t)/P_{ext}(t), \tag{2}$$

where $P_{ext}(n)$ is the probability of extinction of a lineage composed of n individuals, $P_{ext}(t)$ is the probability that a lineage destined to go extinct is not yet extinct by time t, and $P(n,t)$ is the probability that a single mutant created at $t=0$ has n descendants at time t. Note that by a "doomed" lineage we mean a lineage that is destined to vanish.

Now, following standard branching process analysis, a new mutant at fitness qs will have the following distribution of ancestors a time t later (Desai and Fisher, 2007; Eq. 11):

$$P(n=0,t) = \frac{e^{qst}-1}{(1+qs)e^{qst}-1}, \tag{3}$$

so that

$$P(n=0,t|\text{doomed}) = \frac{(1+qs)(e^{qst}-1)}{(1+qs)e^{qst}-1}, \tag{4}$$

and furthermore (Desai and Fisher, 2007; Eq. 10)

$$P(n>0,t) = \frac{(qs)^2e^{qst}}{((1+qs)e^{qst}-1)((1+qs)e^{qst}-1-qs)}\left(\frac{((1+qs)e^{qst}-1-qs)}{(1+qs)e^{qst}-1}\right)^n. \tag{5}$$

In the branching process analysis, all lineages are independent, so

$$P_{ext}(n) = P_{ext}(1)^n = \left(\frac{1}{1+qs}\right)^n. \tag{6}$$

Furthermore,

$$P_{ext}(t) = 1 - P(n=0,t|\text{doomed}) = \frac{qs}{(1+qs)e^{qst}-1}. \tag{7}$$

Thus, using Eq. (2) gives

$$P(n>0,t|\text{doomed}) = \frac{qse^{qst}}{(1+qs)e^{qst}-1-qs}\left[\frac{e^{qst}-1}{(1+qs)e^{qst}-1}\right]^n. \tag{8}$$

Finally, this gives us

$$\langle n(t)|n(t)\neq 0, n(t\to\infty)=0\rangle = \sum_{n=1}^{\infty} nP(n,t|\text{doomed}) = \frac{qse^{qst}}{(1+qs)e^{qst}-1-qs}\sum_{n=1}^{\infty} n\xi^n \tag{9}$$

with

$$\xi = \frac{e^{qst}-1}{(1+qs)e^{qst}-1}. \tag{10}$$

This reduces to

$$\langle n(t)|n(t)\neq 0, n(t\to\infty) = \frac{((1+qs)-e^{-qst})}{qs} \tag{11}$$

Now, we include the possibility that the lineage has gone extinct by time t:

$$\langle n(t)|n(t\to\infty)=0\rangle = \frac{((1+qs)-e^{-qst})}{qs}P[n\neq 0|\text{doomed}] = e^{-qst}. \tag{12}$$

A fitness class, naively, has $\mathcal{O}(1/U_b)$ individuals when it typically starts creating establishing mutants. Hence,

$$\begin{aligned}\frac{N_{doomed}(t_1+\tau_k)}{n_k((t_1+\tau_k))} &= U_b^2\int_0^{(t_1+\tau_k)}(1-qs)\frac{1}{qs}e^{(q-1)sT}e^{-qs((t_1+\tau_k)-T)}dT\\ &\approx \frac{U_b^2 e^{qs(t_1+\tau_k)}}{2qs}\\ &\approx \frac{s}{2}.\end{aligned} \tag{13}$$

Thus, given our assumption that $s \ll 1$, the contribution of doomed lineages is indeed small and can be ignored.

Moments of the jump probability: The moments of the jump probability ρ are calculated as follows, analogous to the method used in Desai et al. (2013). Using the fact that

$$\left(\frac{1}{\sigma_k}\right)^n = \int_0^\infty \frac{z^{n-1}}{(n-1)!}e^{-z\sigma_k}dz, \tag{14}$$

the n-th moment (for the step $k\to k+1$) is derived as

$$\begin{aligned}\langle x^n\rangle = \left\langle\left(\frac{\nu}{\sigma_k}\right)^n\right\rangle &= \left\langle\int_0^\infty \frac{z^{n-1}}{(n-1)!}e^{-z\sigma_k}\nu^n dz\right\rangle\\ &= \int_0^\infty \frac{z^{n-1}}{(n-1)!}\left\langle e^{-z\sigma_k}\nu^n\right\rangle dz\\ &= \int_0^\infty \frac{z^{n-1}}{(n-1)!}\left\langle e^{-z(\nu+\tilde{\nu})}\nu^n\right\rangle dz\end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty \frac{z^{n-1}}{(n-1)!} \left\langle e^{-z\tilde{\nu}} \right\rangle \left\langle e^{-z\nu} \nu^n \right\rangle dz \\
&= \int_0^\infty \frac{z^{n-1}}{(n-1)!} e^{-(1-x_0)z^{(1-1/q)}} \left\langle e^{-z\nu} \nu^n \right\rangle dz. \qquad (15)
\end{aligned}$$

Now,

$$\langle \nu^n e^{-z\nu} \rangle = (-1)^n \frac{d^n}{dz^n} \langle e^{-z\nu} \rangle \tag{16}$$

and

$$\left\langle e^{-z\nu} \right\rangle = e^{-x_0 z^{(1-1/q)}}. \tag{17}$$

We can thus compute a general formula for the n-th derivative:

$$\frac{d^n}{dz^n} e^{-x_0 z^{(1-1/q)}} = e^{-x_0 z^{(1-1/q)}} \sum_{i=0}^{n-1} (-x_0)^{n-i} (1-1/q)^{n-i} z^{-(n-i)/q-i} (-1)^i \frac{\Gamma(1/q+i)}{\Gamma(1/q)}. \tag{18}$$

Thus,

$$\langle \nu^n e^{-z\nu} \rangle = e^{-x_0 z^{1-1/q}} \sum_{i=0}^{n-1} x_0^{n-i} (1-1/q)^{n-i} z^{-(n-i)/q-i} \frac{\Gamma(1/q+i)}{\Gamma(1/q)}. \tag{19}$$

This gives for the moments,

$$\begin{aligned}
\langle x^n \rangle = \left\langle \left(\frac{\nu}{\sigma_k} \right)^n \right\rangle &= \int_0^\infty \frac{z^{(n-1)}}{(n-1)!} e^{-(1-x_0)z^{1-1/q}} e^{-x_0 z^{1-1/q}} \sum_{i=0}^{n-1} x_0^{n-i} (1-1/q)^{n-i} z^{-(n-i)/q-i} \frac{\Gamma(1/q+i)}{\Gamma(1/q)} dz \\
&= \sum_{i=0}^{n-1} x_0^{n-i} (1-1/q)^{n-i} \frac{\Gamma(1/q+i)}{(n-1)!\Gamma(1/q)} \int_0^\infty z^{(n-1)} e^{-z^{1-1/q}} z^{-(n-i)/q-i} dz \\
&= \sum_{i=0}^{n-1} x_0^{n-i} \frac{\Gamma(1/q+i)\Gamma(n-i)}{(n-1)!\Gamma(1/q)} \left(\frac{q-1}{q} \right)^{n-1-i}. \qquad (20)
\end{aligned}$$

The first two moments are

$$\langle \Delta x \rangle = 0, \tag{21}$$

$$\langle (\Delta x)^2 \rangle = \frac{x_0(1-x_0)}{q}. \tag{22}$$

Equation (20) is readily generalized to k steps forward through the usual substitution,

$1 - 1/q = \alpha \Rightarrow \alpha^k$, which gives

$$\langle x_k^n \rangle = \sum_{i=0}^{n-1} x_0^{n-i} \frac{\Gamma(1-\alpha^k+i)\Gamma(n-i)}{(n-1)!\Gamma(1-\alpha^k)} \left(\alpha^k\right)^{n-1-i} . \tag{23}$$

Site Frequency Spectrum of Nearly Private Variants:

In this supplement we flesh out the derivation of the site frequency spectrum of nearly private variants. As noted in the text, several of the difficulties that arise in the derivation of the site frequency spectrum of common mutations may be neglected when dealing with rare variants. Specifically, the problem is vastly simplified with the inclusion of two approximations. First, since the lineage sizes of extremely rare variants are small and destined to go extinct, the lineages can be assumed to experience no further (establishing) beneficial mutations (specifically, this approximation holds if $nyU_b \ll 1$, where n is the lineage size and y the fitness of the mutant created by the lineage). Second, we assume that mutations are fed into a given fitness class k deterministically at rate $U_b n_{k-1}(t)$. This holds if $U_b n_{k-1}(t) \gg 1$ (certainly true in the bulk of the distribution), in which case fluctuations of incoming mutants around the expected number are small.

Because these lineages never comprise a significant fraction of the population, they can be studied through a standard branching process analysis with a constant death rate $d = 1$ and a time-varying birth rate $b(t) = 1 + y_0 - vt$, where t is time in generations, y_0 is the initial fitness of the mutant, and v the mean rate of adaptation of the population. First, we are interested in deriving the expected (time-averaged) number of mutations carried by n individuals with relative fitnesses y, $F_{rare}(n/N, y)$. This is obtained by considering the expected number of mutants introduced when a given fitness class was at relative fitness $y_0 > y$, and multiplying by the probability that, in the time it took for the relative fitness of the class to decrease to y, the lineage size of any of these mutants has increased to n. $F_{rare}(n/N, y)$ is then the integral over all possible landing fitnesses y_0.

Clearly, the number of mutants introduced at an initial fitness y_0 is simply $U_b N(y_0 - s)$ where $N(y)$ is the expected number of individuals at relative fitness y. So long as the individual does not reside in the distribution's high-fitness nose, $N(y)$ is well approximated by a Gaussian with variance v (Desai and Fisher, 2007):

$$N(y) = \frac{N}{\sqrt{2\pi v}} e^{-y^2/2v}. \tag{24}$$

Furthermore, a mutant that is introduced at initial fitness y_0 will be at fitness y in a time $t = (y_0 - y)/v$. Finally, the distribution in lineage sizes of a mutant with an initial birth rate $1 + y_0$ that is decreasing at a rate v per generation is a classic branching process problem that was solved by Kendall (1948). The distribution in lineage sizes is

$$P_1(n > 0, t) = \frac{e^{-y_0 t + \frac{vt^2}{2}}}{(\int_0^t e^{-y_0\tau + v\tau^2/2} d\tau + 1)^2} \left(1 - \frac{e^{-y_0 t + \frac{vt^2}{2}}}{\int_0^t e^{-y_0\tau + v\tau^2/2} d\tau + 1}\right)^{n-1}. \tag{25}$$

As a result, $F_{rare}(n/N, y)$ is given by

$$\begin{aligned} F_{rare}(n/N, y) &= \frac{1}{v} \int_y^{(q-2)s} P_1(n, (y_0 - y)/v) U_b N(y_0 - s) dx_0 \\ &= \frac{1}{v} \frac{U_b N}{\sqrt{2\pi v}} \int_y^{\infty} \frac{e^{\frac{y^2 + 2y_0 s - 2y_0^2 - s^2}{2v}}}{(\int_0^{((y_0-y)/v)} e^{-y_0\tau + v\tau^2/2} d\tau + 1)^2} \left(1 - \frac{e^{\frac{y^2 - y_0^2}{2v}}}{\int_0^{((y_0-y)/v)} e^{-y_0\tau + v\tau^2/2} d\tau + 1}\right)^{n-1} dy_0. \end{aligned} \tag{26}$$

where an arbitrary upper limit of $(q-2)s$ is imposed to restrict to regions where the deterministic supply rate of mutants is guaranteed to hold, with contributions from mutants founded when the class was at fitness greater than $(q-2)s$ already negligible for very rare mutations. The total SFS of semi-private variants is then obtained by integrating over all final fitnesses y:

$$\begin{aligned} F_{rare}(n/N) &= \int_{-qs}^{(q-2)s} F_{rare}(n/N, y) dy \\ &= \frac{1}{v} \frac{U_b N}{\sqrt{2\pi v}} \int_{-\infty}^{\infty} \int_y^{\infty} \frac{e^{\frac{y^2 + 2y_0 s - 2y_0^2 - s^2}{2v}}}{(\int_0^{((y_0-y)/v)} e^{-y_0\tau + v\tau^2/2} d\tau + 1)^2} \\ &\times \left(1 - \frac{e^{\frac{y^2 - y_0^2}{2v}}}{\int_0^{((y_0-y)/v)} e^{-y_0\tau + v\tau^2/2} d\tau + 1}\right)^{n-1} dy_0 dy. \end{aligned} \tag{27}$$

To extract the leading order behavior, we consider frequencies just at the mean, $y = 0$:

$$F_{rare}(n/N, 0) = \frac{1}{v}\frac{U_b N}{\sqrt{2\pi v}} \int_0^\infty \frac{e^{\frac{2y_0 s - 2y_0^2 - s^2}{2v}}}{(\int_0^{y_0/v} e^{-y_0\tau + v\tau^2/2} d\tau + 1)^2} \left(1 - \frac{e^{\frac{-y_0^2}{2v}}}{\int_0^{y_0/v} e^{-y_0\tau + v\tau^2/2} d\tau + 1}\right)^{n-1} dy_0. \tag{28}$$

Examining the integral inside the integrand, we observe that

$$\int_0^{y_0/v} e^{-y_0\tau + v\tau^2/2} d\tau = \sqrt{\frac{2}{v}} e^{-y_1^2} \int_0^{y_1} e^{y^2} dy = \sqrt{\frac{2}{v}} D(y_1) \tag{29}$$

for $y_1 = y_0/\sqrt{2v}$, and $D(y_1)$ is Dawson's integral, a well-studied special function. If y_1 is small, then

$$D(y_1) \approx y_1 - \frac{2}{3} y_1^3 + \frac{4}{15} y_1^5 ... \tag{30}$$

Since we are considering only small lineage sizes n, the integral will obtain its main contribution for y_0 small (meaning that most nearly-private variants were founded recently in the past). Thus we can use the first order expansion:

$$\begin{aligned}
F_{rare}(n/N, 0) &\approx \frac{U_b N}{v\sqrt{2\pi v}} \int_0^\infty \frac{e^{\frac{2y_0 s - 2y_0^2 - s^2}{2v}}}{(y_0/v + 1)^2} \left(1 - \frac{1}{y_0/v + 1}\right)^{n-1} dy_0 \\
&= \frac{U_b N v e^{-s^2/(2v)}}{\sqrt{2\pi v}} \int_0^\infty \exp\left(\frac{y_0 s - y_0^2}{v}\right) \frac{y_0^{n-1}}{(y_0 + v)^{n+1}} dy_0 \\
&\approx \frac{U_b N v e^{-s^2/(2v)}}{\sqrt{2\pi v}} \left(\int_0^\infty \exp\left(\frac{-y_0^2}{v}\right) \frac{y_0^{n-1}}{(y_0 + v)^{n+1}} dy_0 \right. \\
&\left. + \frac{s}{v} \int_0^\infty \exp\left(\frac{-y_0^2}{v}\right) \frac{y_0^n}{(y_0 + v)^{n+1}} dy_0 \right) \\
&= \frac{U_b N e^{-s^2/(2v)}}{\sqrt{2\pi}} \left(\int_0^\infty \frac{e^{-\xi^2} \xi^{n-1}}{(\xi + \sqrt{v})^{n+1}} d\xi + \frac{s}{\sqrt{v}} \int_0^\infty \frac{e^{-\xi^2} \xi^n}{(\xi + \sqrt{v})^{n+1}} d\xi \right),
\end{aligned} \tag{31}$$

where, in the last step, the substitution $\xi = \frac{y_0}{\sqrt{v}}$ was performed. Performing each integration separately:

$$\begin{aligned}
\int_0^\infty \frac{e^{-\xi^2} \xi^{n-1}}{(\xi + \sqrt{v})^{n+1}} d\xi &\approx \frac{1}{n\sqrt{v}}, \\
\int_0^\infty \frac{e^{-\xi^2} \xi^n}{(\xi + \sqrt{v})^{n+1}} d\xi &\approx -\frac{\gamma}{2} + \log\left(\frac{1}{n\sqrt{v}}\right) + \frac{1}{n}
\end{aligned} \tag{32}$$

for $\sqrt{v} \ll 1$. Note that the second approximation breaks down for n large; however, in the realm of validity of our approximation (derived below), the second term will generally be much smaller than the first term because of the log dependence on n.

Thus, the leading order behavior is

$$F_{rare}(n/N, y=0) = \frac{U_b N e^{-s^2/(2v)}}{\sqrt{2\pi v n}}. \tag{33}$$

Small n approximation condition: We would like to derive a realm of validity for all of our approximations. The primary assumption made in simplifying the Dawson's integral and exponential integrals is that $y_0/\sqrt{2v} \ll 1$, which is justified if the dominant contribution to the integral occurs for y_0 satisfying this condition. Since we are integrating against a Gaussian, the integral is sharply peaked around the maximum of

$$e^{\frac{2y_0 s - 2y_0^2 - s^2}{2v}} \exp\left(- \frac{(n-1)e^{\frac{-y_0^2}{2v}}}{e^{-y_0^2/2v}\int_0^{y_0/v} e^{vy^2/2}dy + 1} \right). \tag{34}$$

Thus we would like to find the location of the maximum of the exponent's argument:

$$\frac{2y_0 s - 2y_0^2 - s^2}{2v} - \frac{(n-1)e^{\frac{-y_0^2}{2v}}}{e^{-y_0^2/2v}\int_0^{y_0/v} e^{vy^2/2}dy + 1}. \tag{35}$$

Define

$$I = e^{-y_0^2/2v}\int_0^{y_0/v} e^{vy^2/2}dy. \tag{36}$$

The equation to be solved for the peak y_{max} is

$$0 = s - 2y_{max} + \frac{(n-1)y_{max}}{I+1}e^{-y_{max}^2/(2v)} + \frac{(n-1)y_{max} I e^{-y_{max}^2/(2v)}}{(I+1)^2}. \tag{37}$$

Note that $y_{max} = \mathcal{O}(s/2)$, with some n dependent correction that necessarily increases the location of the peak. Thus, for any of these Taylor expansions to hold for any n, we require at minimum that $s \ll \sqrt{v}$.

In this case it is true that

$$y_{max} = \frac{s}{2} + \frac{(n-1)y_{max}}{2(I+1)}e^{-y_{max}^2/(2v)} + \frac{(n-1)y_{max} I e^{-y_{max}^2/(2v)}}{2(I+1)^2} < \frac{s}{2} + \frac{(n-1)y_{max}}{2(I+1)} + \frac{(n-1)y_{max} I}{2(I+1)^2}. \tag{38}$$

This may be solved numerically to find the exact location of the peak. However, if we suppose that the n terms are a small perturbation on the $s/2$ peak, and (as we have already assumed) $s \ll \sqrt{v}$, we are justified in a first order expansion of Dawson's integral:

$$\begin{aligned} y_{max} &< \frac{s}{2} + \frac{(n-1)y_{max}}{2(y_{max}/v+1)} + \frac{(n-1)y_{max}^2/v}{2(y_{max}/v+1)^2} \\ &< \frac{s}{2} + (n-1)v. \end{aligned} \tag{39}$$

So long as $(n-1)v \ll 1$, it is indeed true that the new maximum is a small perturbation around the $s/2$ peak, and our original first order expansion of I was justified.

Thus, for the small n approximation to hold, we require that $y_{max} < \frac{s}{2} + (n-1)v \ll \sqrt{2v}$. Thus, the small n approximation holds for

$$n \ll \sqrt{\frac{2}{v}} - \frac{s}{2v}. \tag{40}$$

Derivation of the Neutral Site Frequency Spectrum:

In this section we derive the asymptotic form of the neutral site frequency spectrum for common alleles. To derive this, we start from the assumption of a class that is growing exponentially (and deterministically), such that

$$n_{k-1}(t) = \frac{e^{(q-1)st}}{qs}. \tag{41}$$

This class supplies neutral mutants at a rate U_n. Thus the expected number of individuals with lineage sizes n at time t, $F(n,t)$, is simply the integral over the expected number of mutants introduced at time τ multiplied by the probability for a mutant to reach a lineage size n in time $t-\tau$:

$$F(n,t) = \int_{-\infty}^{t} U_n n_{k-1}(\tau) P(n, t-\tau) d\tau. \tag{42}$$

The probability for a lineage to reach size n in a time t is given in Desai and Fisher (2007) and in the first section of this SI. This gives

$$F(n,t) = ((q-1)s)^2 \int_{-\infty}^{t} \frac{U_n d\tau}{qs} e^{(q-1)s\tau} \left(\frac{(1+(q-1)s)e^{(q-1)s(t-\tau)} - 1 - (q-1)s)}{((1+(q-1)s)e^{(q-1)s(t-\tau)} - 1)} \right)^n$$

$$\times \frac{e^{(q-1)s(t-\tau)}}{((1+(q-1)s)e^{(q-1)s(t-\tau)}-1)((1+(q-1)s)e^{(q-1)s(t-\tau)}-1-(q-1)s)}. \quad (43)$$

Defining $\sigma = (q-1)s, y = (1+\sigma)e^{\sigma k} - 1 - \sigma$, we obtain

$$F(n,t) = \frac{\sigma U_n e^{\sigma t}}{qs} \int_0^\infty \frac{dy y^{n-1}}{(y+1+\sigma)(y+\sigma)^{n+1}}. \quad (44)$$

An explicit series expansion in powers of n may now explicitly be derived. First we note that

$$\frac{1}{y+1+\sigma} = \frac{1}{y+\sigma}\left(\frac{1}{1+\frac{1}{y+\sigma}}\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(y+\sigma)^{k+1}}, \quad (45)$$

which gives

$$F(n,t) = \frac{\sigma U_n e^{\sigma t}}{qs} \int_0^\infty \sum_{k=0}^{\infty} \frac{dy y^{n-1}(-1)^k}{(y+\sigma)^{n+k+1}} = \frac{U_n e^{\sigma t}}{qs} \sum_{k=2}^{\infty} \frac{(-1)^k (k-1)!\Gamma(n)}{\sigma^{k-1}\Gamma(k+n)} = \frac{U_n e^{\sigma t}}{qs} \sum_{k=2}^{\infty} \frac{(-1)^k \beta(k,n)}{\sigma^{k-1}}, \quad (46)$$

where the Beta function $\beta(k,n)$ is defined as

$$\beta(k,n) = \frac{\Gamma(k)\Gamma(n)}{\Gamma(n+k)}. \quad (47)$$

The above is well approximated by the following expansion:

$$F(n,t) = \frac{U_n e^{\sigma t}}{qs} \sum_{k=2}^{\infty} \frac{(-1)^k (k-1)!}{\sigma^{k-1} n^k} \approx \frac{U_n e^{\sigma t}}{qs} \left(\frac{1}{\sigma n^2} - \frac{2}{\sigma^2 n^3} + \frac{6}{\sigma^3 n^4} \ldots\right). \quad (48)$$

As expected, for $\sigma \to \infty$ or $n \to \infty$ we recover the characteristic n^{-2} decay of an exponentially expanding population. Keeping the leading order term, we obtain

$$F(n,t) = \frac{U_n e^{\sigma t}}{qs\sigma n^2} \Rightarrow f(x)dx = \frac{U_n dx}{\sigma x^2}. \quad (49)$$

Details concerning forward-time Wright Fisher simulations:

We validate some of our results in the text by comparing theorized predictions to simulations. Toward this end, we implemented forward-time simulations that closely resemble evolution in the Wright-Fisher model. The details of the implementation of these simulations is described in detail in our previous work (Good et al., 2012).

To measure the transition probabilities and sojourn times of mutations between fitness classes, an initially clonal population is allowed to evolve for $2(q+1)\langle \tau_k \rangle$ generations until it reaches its steady state distribution of fitnesses. At this point, a mutation is seeded in at a frequency $x = 0.5$ in each fitness class. A new class k is allowed to establish, and shortly afterwords the mutation reaches some steady frequency x_k in this class. The population is then allowed to evolve until a class containing $\lceil k+q \rceil$ beneficial mutations establishes. At this point, class k is (roughly) at the population's mean fitness, and the frequency of the mutant in this class is recorded. This prescription certifies that frequencies of mutations attain their long-time steady values long before they are measured. Upon the establishment of class $\lceil k+q+1 \rceil$, the frequency of the mutant in class $k+1$ is recorded. Generally, when class $\lceil k+q+i \rceil$ establishes, the frequency of the mutant in class $k+i$ is recorded. In this way, a vector of transitions $\{x_k, x_{k+1}, ..., \}$ is generated, until $x_i = 0$ or $x_i = 1$ in the mean class for some i. The transition and sojourn probabilities are then collected from 60,000 such runs for each parameter set.

Implementation of our code in Python (used to obtain transition and sojourn times) and C (used to obtain site frequency spectra) are freely available upon request.

LITERATURE CITED

Desai, M. and Fisher, D. 2007. Beneficial mutation- selection balance and the effect of linkage on positive selection. Genetics 176:1759.

Desai, M., Walczak, A., and Fisher, D. 2013. Genetic diversity and the structure of genealogies in rapidly adapting populations. Genetics 193:565–585.

Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O., and Desai, M. M. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc. Natl. Acad. Sci. USA 109:4950–4955.

Kendall, D. G. 1948. On the generalized birth-and-death process. Annals of Mathematical Statistics 19:1–15.