

File S1
Supporting Methods

Derivation of the backward Master equation for $p(n, T)$

The fundamental quantity of the branching process is the probability distribution, $p(n, T)$, of observing n copies of the allele T generations after it originated. A backward equation for $p(n, T)$ can be derived by considering the probability $p(n, T|k, t, x)$ of having n copies at time T , given there were k copies on a background with fitness x at time t . $p(n, T|k, t - \Delta t, x)$ now can be expressed as a sum over possible intermediate states at time t :

$$\begin{aligned} p(n, T|k, t - \Delta t, x) = & p(n, T|k, t, x)(1 - \Delta t k(2 + x - \bar{x}(t) + s + r)) \\ & + \Delta t [k(1 + x - \bar{x}(t) + s)p(n, T|k + 1, t, x) + kp(n, T|k - 1, t, x)] \\ & + \Delta t r k \sum_{n'} \int_{x'} K_{xx'} p(n - n', T|k - 1, t, x) p(n', T|1, t, x'), \end{aligned} \quad (1)$$

where $\bar{x}(t)$ is the mean fitness of the population at time t . The different terms have straightforward interpretations: The first term is the probability that nothing happens in the small time interval Δt , the second accounts for a division of one of k individuals, which happens with rate $k(1 + x - \bar{x}(t) + s)$, the third term accounts for the death of one of the k , while the last term accounts for outcrossing of one of the k , producing a new individual with fitness x' and removing one with fitness x . The outcrossing term is then summed over all possible ways the $k - 1$ individuals with fitness x and the one with fitness x' can give rise to n individuals at time T . Sending Δt to zero and rearranging terms results in an ODE for $p(n, T|k, t, x)$

$$\begin{aligned} -\partial_t p(n, T|k, t, x) = & -p(n, T|k, t, x)k(2 + x - \bar{x}(t) + s + r) \\ & + k(1 + x - \bar{x}(t) + s)p(n, T|k + 1, t, x) + kp(n, T|k - 1, t, x) \\ & + rk \sum_{n'} \int_{x'} K_{xx'} p(n - n', T|k - 1, t, x) p(n', T|1, t, x') \end{aligned} \quad (2)$$

This is equation 1 from the main text.

Derivation of the equation for the generating function

To remove the convolution over n it is convenient to consider the generating function $\hat{p}(\lambda, T|k, t, x) = \sum_n \lambda^n p(n, T|k, t, x)$. Multiplying the above equation by λ^n and summing over n yields

$$\begin{aligned} -\partial_t \hat{p}(\lambda, T|k, t, x) = & -k(2 + x - \bar{x}(t) + s + r)\hat{p}(\lambda, T|k, t, x) \\ & + k(1 + x - \bar{x}(t) + s)\hat{p}(\lambda, T|k + 1, t, x) + k\hat{p}(\lambda, T|k - 1, t, x) \\ & + rk \int_{x'} K_{xx'} \hat{p}(\lambda, T|k - 1, t, x) \hat{p}(\lambda, T|1, t, x) \end{aligned} \quad (3)$$

All k initial individuals are independent, hence $\hat{p}(\lambda, T|k, t, x) = \hat{p}^k(\lambda, T|t, x)$ and the right hand side is $-\partial_t \hat{p}^k(\lambda, T|t, x) = -k\hat{p}^{k-1}(\lambda, T|t, x)\partial_t \hat{p}(\lambda, T|t, x)$. We can therefore divide the equation by $\hat{p}^{k-1}(\lambda, T|t, x)$ to obtain

$$-\partial_t \hat{p}(\lambda, T, t, x) = -(2 + x - \bar{x}(t) + s + r)\hat{p}(\lambda, T, t, x) + (1 + x - \bar{x}(t) + s)\hat{p}^2(\lambda, T, t, x) + 1 + r \int_{x'} K_{xx'} \hat{p}(\lambda, T, t, x) \quad (4)$$

The generating function has the boundary condition $\hat{p}(\lambda, T|T, x) = \lambda$, which follows from $p(n, T|k, T, x) = \delta_{nk}$. Substituting $\hat{p}(\lambda, T|t, x) = 1 - \phi(\lambda, T, t, x)$ removes the constant term.

$$\partial_t \phi(\lambda, T, t, x) = -r \int_{x'} K_{xx'} \phi(\lambda, T, t, x') - (x - \bar{x}(t) + s - r)\phi(\lambda, T, t, x) + (1 + x - \bar{x}(t) + s)\phi^2(\lambda, T, t, x) \quad (5)$$

which now has boundary condition $\phi(\lambda, T, T, x) = 1 - \lambda$. Assuming that selection is weak on the timescale of one generation ($\sigma \ll 1$), we can approximate $1 + x - \bar{x}(t) + s$ by 1 and arrive at Eq. 18 of the main text:

$$-\partial_t \phi(\lambda, T, t, x) = r \int_{x'} K_{xx'} \phi(\lambda, T, t, x') + (x - \bar{x}(t) + s - r) \phi(\lambda, T, t, x) - \phi^2(\lambda, T, t, x) \quad (6)$$

Solution for $\Phi(\tau)$ at low \tilde{r}

In the main text, we presented a solution to the two equations (Eqs. (21) and (23))

$$\psi(\tau, \chi) = \begin{cases} r e^{\theta^2/2} \int_0^\tau d\tau' \Phi(\tau') e^{-\theta'^2/2} & \theta < \Theta_c \\ \theta/\epsilon & \theta > \Theta_c \end{cases} \quad (7)$$

and

$$\partial_\tau \Phi(\tau) = \tilde{s} \Phi(\tau) - \epsilon \int d\chi P(\tau, \chi) \psi^2(\tau, \chi) \quad (8)$$

in a regime of intermediate \tilde{r} ($1/\sqrt{\log N} < \tilde{r} < 1$). An additional solution with qualitatively different properties exists at low \tilde{r} . While the assumption of a Gaussian fitness distribution is questionable in this range of \tilde{r} , we nevertheless present the solution here for completeness. Proceeding as before, we evaluate the integral in Eq. (7) by expanding exponent around its maximum. The maximum is located at $\tau' = \tau - \theta - \alpha(\tau')$ where $\alpha(\tau) = -\Phi(\tau)^{-1} \partial_\tau \Phi(\tau)$. Assuming $\alpha(\tau)$ changes slowly with time, we have

$$\psi(\tau) \approx \sqrt{2\pi\tilde{r}} \Phi(\tau) e^{\frac{(\theta+\alpha)^2}{2}} \quad (9)$$

This solution is valid below Θ_c , where $\psi(\tau)$ crosses over the linear saturated form θ/ϵ . In addition to the matching condition at Θ_c , we use Eq. 8 for $s = 0$ to determine α :

$$\partial_\tau \Phi(\tau) = -\alpha \Phi(\tau) = -\tilde{r} \Phi(\tau) e^{\Theta_c(\alpha - \tilde{r}) + \alpha^2/2 - \tilde{r}^2/2} \quad (10)$$

For $\tilde{r}\Theta_c \gg 1$ we recover the solution with small α given in the main text. For smaller \tilde{r} , however, we find $\alpha(\tau) \approx \frac{\log \tilde{r}}{\Theta_c}$. Solving the matching condition for Θ_c and differentiating with respect to τ yields $\partial_\tau \Theta_c = -\frac{1}{\Theta_c} \alpha(\tau)$. This is readily solved for this case of $\Theta_c \tilde{r} < 1$

$$\Theta_c(\tau) = (-3(\tau - \tau_0) \log \tilde{r} + \Theta_0^3)^{1/3} . \quad (11)$$

Substituting this solution for $\Theta_c(\tau)$ into the expression for the rescaled generating function yields

$$\Phi(\tau) = \frac{\Theta_c}{\epsilon \tilde{r}} e^{-\frac{(\Theta_c + \alpha)^2}{2}} \sim e^{-\frac{[-3(\tau - \tau_0) \log \tilde{r} + \Theta_0^3]^2/3}{2}} \quad (12)$$

Hence in this low \tilde{r} regime, the decay of the survival probability is qualitatively different from the regime of intermediate \tilde{r} .

Effective clone-based model

To rationalize the behavior of the continuous time branching process, we considered the following simplified model discussed in the main text: Genotypes expand clonally and produce recombinant offspring. The offspring start growing simultaneously in the next “effective” generation after all clones from the previous generation have disappeared. The relevant quantity now is the number of clones or distinct genotypes, rather than the number of individuals. To understand the dynamics of the number of clones, we need to know how many clones a single clone can produce.

Consider a single genotype with background fitness χ which is carrying a mutation of effect size \tilde{s} . The expected number of recombinant offspring from this genotype is $\xi = \tilde{r} \int_0^\infty dt n_\chi(t)$, where $n_\chi(t)$ is the copy number trajectory. The Laplace transform $\hat{p}(\xi)$ of $p(\xi)$ obeys the equation ($\phi(z) = 1 - \hat{p}(z)$)

$$\partial_\chi \phi(\chi, z) = \tilde{r} z + (\chi + \tilde{s} - \tilde{r}(1 - z)) \phi(\chi, z) - \phi(\chi, z)^2 , \quad (13)$$

which is a simpler version of the Eq. (42) (main text) since only one single clone is considered and recombination to daughter clone is ignored. This equation can be solved asymptotically in the regimes of large and small $\chi - \tilde{r}$.

$$\phi(\chi, z) = \begin{cases} \tilde{r}ze^{(\chi+\tilde{s}-\tilde{r}(1-z))^2/2} \int_{-\infty}^{\chi+\tilde{s}-\tilde{r}(1-z)} dx e^{-x^2/2} & \chi - \tilde{r} \ll \Theta_c \\ (\chi - \tilde{r}(1-z)) & \chi - \tilde{r} \gg \Theta_c \end{cases} \quad (14)$$

where $\Theta_c \approx \sqrt{-2 \log \tilde{r}z}$. The initial fitness of the genotype, χ , is Gaussian distributed and the Laplace transform has to be averaged over χ .

$$\phi(z) = \int_{-\infty}^{\infty} \frac{d\chi}{\sqrt{2\pi}} e^{-\chi^2/2} \phi(\chi, z) \quad (15)$$

Let's look at the mean number $\langle \xi \rangle$ of recombinant offspring first, which is given by differentiating with respect to z and setting $z = 0$, which in turn sends Θ_c to infinity. Integration by parts yields

$$\langle \xi \rangle = \partial_z \int_{-\infty}^{\infty} \frac{d\chi}{\sqrt{2\pi}} e^{-\chi^2/2} \phi(\chi) = \tilde{r} \int_{-\infty}^{\infty} \frac{d\chi}{\sqrt{2\pi}} e^{-\chi(\tilde{r}-\tilde{s})+\tilde{r}^2/2} \int_{-\infty}^{\chi-\tilde{r}+\tilde{s}} dx e^{-x^2/2} = \frac{\tilde{r}}{\tilde{r}-\tilde{s}} \quad (16)$$

Hence the mean number of recombinant offspring is 1 for a neutral mutation and approximately $1 + \tilde{s}/\tilde{r}$ for mutations with small effect. To evaluate the integral in Eq. (15) at finite z , we have to account for the cross-over of $\phi(\chi, z)$ at $\chi - \tilde{r} = \Theta_c \approx \sqrt{-2 \log \tilde{r}z}$. For small z , the cross-over translates into a cut-off of the integral. Again, the integral can be evaluated by parts:

$$\begin{aligned} \phi(z) &= \int_{-\infty}^{\Theta_c+\tilde{r}-\tilde{s}} \frac{d\chi}{\sqrt{2\pi}} \tilde{r}ze^{-\chi(\tilde{r}(1-z)-\tilde{s})+\tilde{r}^2(1-z)^2/2} \int_{-\infty}^{\chi-\tilde{r}(1-z)+\tilde{s}} dx e^{-x^2/2} + \int_{\Theta_c+\tilde{r}-\tilde{s}}^{\infty} \frac{d\chi}{\sqrt{2\pi}} e^{-\chi^2/2} (\chi - \tilde{r}(1-z)) \\ &\approx \frac{\tilde{r}z}{\tilde{r}-\tilde{s}} \left[1 - e^{-\Theta_c(\tilde{r}-\tilde{s})-\tilde{r}^2/2} \right] \end{aligned} \quad (17)$$

Hence, the generating function of the average number of recombinant offspring generated by a random genotype is given by

$$\hat{p}(z) = 1 - \phi(z) \approx 1 - \frac{\tilde{r}z}{\tilde{r}-\tilde{s}} \left[1 - e^{-\Theta_c(\tilde{r}-\tilde{s})} \right], \quad (18)$$

where the last expression is valid for $z \ll 1$ and $\tilde{r} \ll 1$.

The actual number m of recombinant offspring (novel clones) generated by one clone is Poisson distributed with mean ξ . The generating function of m is therefore

$$\sum_m \lambda^m P(m) = \sum_m \lambda^m \int_0^\infty d\xi p(\xi) e^{-\xi} \frac{\xi^m}{m!} = \int_0^\infty d\xi p(\xi) e^{-\xi(1-\lambda)} = \hat{p}(1-\lambda) = 1 - \phi(1-\lambda) \quad (19)$$

We will now use this result to calculate how the number of clones that descend from a particular genotype evolves over time.

The stochastic dynamics of the number genotypes

As long as the clones we are tracking constitute a small fraction of the population, different clones are independent. The probability to go from k to m clones in one effective generation has therefore the generating function $\hat{P}(\lambda, k) = \hat{P}^k(\lambda, 1) = (1 - \phi(1-\lambda))^k$. To study the dynamics of the number of clones over many generations, we need to know how this propagator behaves when iterated.

$$\begin{aligned} \sum_m \lambda^m \sum_{m_1, \dots, m_n} P(m, m_n) P(m_n, m_{n-1}) \dots P(m_1, k) &= \sum_{m_n} \hat{P}^{m_n}(1-\lambda) \sum_{m_1, \dots, m_{n-1}} P(m_n, m_{n-1}) \dots P(m_1, k) \\ &= \left[\hat{P}(1 - \hat{P}(1 - \hat{P}(1 - \dots \hat{P}(1 - \lambda)))) \right]^k = [1 - \phi \circ \phi \dots \circ \phi(1-\lambda)]^k = [1 - \Phi_n(1-\lambda)]^k \end{aligned} \quad (20)$$

Using the result for the Laplace transform in Eq. (18), we arrive at the difference equation

$$\Phi_{n+1} - \Phi_n \approx \left[\tilde{s}/\tilde{r} - e^{-\tilde{r}\sqrt{-2 \log \tilde{r}\Phi_n}} \right] \Phi_n \quad (21)$$

This is exactly the differential equation we have derived using the continuous time mode when the effective generation time is set to \tilde{r}^{-1} . The latter is reasonable since the \tilde{r}^{-1} is the turnover time by recombination.