FILE S1: THE FULL CONDITIONAL CALCULATION

In the main text, we focused primarily on the non-conditional approximation to the coalescence probabilities, which led to our simple expression for the coalescence probabilities, Eq. (15). In this Supplementary File, we show how this approximation can be relaxed in our lineage-structure framework by carrying out the full conditional calculation for some of the simplest possible cases. We use this to understand the structure of the conditional results and discuss the validity of the non-conditional approximation. We note that the full conditional result can also be obtained from the sum of ancestral paths approach by keeping the higher order terms in Eq. (56) of Appendix A, as described in File S4, and the validity of the non-conditional approximation can be directly assessed with that approach.

We begin by considering the full conditional result for the probability that two individuals both sampled from class k coalesce in class $k - 2$. From Appendix A of the main text, we have

$$P_c^{k,k \to k-2} = I_x^{k-2} \int Q_{k,k}^{k-2}(t_1, t_2) \exp\left[-s(k-2)|t_1 - t_2|\right] dt_1 dt_2. \tag{S.1}$$

In order to evaluate this integral, we need to determine the probability distribution of mutant timings $Q_{k,k}^{k-2}(t_1, t_2)$. The time t_1 is the sum of the time for one individual to have mutated from class $k - 2$ to class $k - 1$ plus the time for it to have mutated from class $k - 1$ to class k, and analogously for t_2. However, in order for the two lineages to coalesce in class $k - 2$, they must *not* have coalesced in class $k - 1$. To illustrate the main point, we neglect the distortion in the mutant timings due to the fact that individuals did not coalesce in class k and focus only on the distortions due to the fact that coalescence did not occur in class $k - 1$; if desired, the former distortion can also be included using analogous methods. We refer to the probability distribution of the times when these individuals mutated from class $k - 1$ to class k conditional on them not having coalesced in class $k - 1$ as $Q_{k,k}^{k-1}(t_1, t_2|nc)$. The distribution of the times for these individuals to then have mutated from class $k - 2$ to class $k - 1$ is then given by

$$Q_{1step}^{k-2}(t_1, t_2) = [s(k-1)]^2 e^{-s(k-1)(t_1+t_2)}. \tag{S.2}$$

Thus the distribution of t_1 and t_2 is given by

$$Q_{k,k}^{k-2}(t_1, t_2) = Q_{k,k}^{k-1}(t_1, t_2|nc) \star Q_{1step}^{k-2}(t_1, t_2), \tag{S.3}$$

where $\star$ indicates a convolution. Note that much of the time when the individuals did coalesce in class $k - 1$, they did so because t_1 happened to be close to t_2 (since this increases the chance the two individuals mutated from the same lineage). Thus in $Q_{k,k}^{k-1}(t_1, t_2|nc)$, t_1 and t_2 are on average further apart than in $Q_{k,k}^{k-1}(t_1, t_2)$, and t_1 and t_2 are no longer independent random variables.

A. M. Walczak *et al.*

We now need to calculate $Q_{k,k}^{k-1}(t_1, t_2|nc)$. We have

$$Q_{k,k}^{k-1}(t_1, t_2|nc) = \frac{Q_{k,k}^{k-1}(t_1, t_2) - Q_{k,k}^{k-1}(t_1, t_2|c)P_c^{k,k\to k-1}}{1 - P_c^{k,k\to k-1}}, \tag{S.4}$$

where $Q_{k,k}^{k-1}(t_1, t_2|c)$ is the distribution of timings of mutations from class $k-1$ to k given that the lineages *do* coalesce in class $k-1$. Applying the general probability identity $P(t_1, t_2|c) = \frac{1}{P(c)}P(c|t_1, t_2)P(t_1, t_2)$, and reading off the coalescence probability given t_1 and t_2 from Eq. (13), we find that

$$Q_{k,k}^{k-1}(t_1, t_2|c) = \frac{I_x^{k-1}}{P_c^{k,k\to k-1}}Q_{k,k}^{k-1}(t_1, t_2)e^{-s(k-1)|t_1-t_2|}. \tag{S.5}$$

We therefore find

$$Q_{k,k}^{k-1}(t_1, t_2|nc) = \frac{1}{1 - P_c^{k,k\to k-1}}\left[(sk)^2 e^{-sk(t_1+t_2)} - I_x^{k-1}(sk)^2 e^{-2k(t_1+t_2)}e^{-s(k-1)|t_1-t_2|}\right]. \tag{S.6}$$

Plugging this into our convolution formula for $Q_{k,k}^{k-2}(t_1, t_2)$ and evaluating the integrals by separating out the possible time orderings, we find

$$Q_{k,k}^{k-2}(t_1, t_2) = \frac{k^2\left[s(k-1)\right]^2}{1 - P_c^{k,k\to k-1}}e^{-s(k-1)(t_1+t_2)}\left[\left(1 - e^{-st_1}\right)\left(1 - e^{-2t_2}\right) - \frac{I_x^{k-1}}{k-2}B\right], \tag{S.7}$$

where we have defined

$$\begin{aligned}
B &= \frac{1}{(k-2)}\left[1 - e^{-2s\min(t_1,t_2)} - \frac{2}{k}\left(1 - e^{-sk\min(t_1,t_2)}\right)\right.\\
&\quad\left. + \frac{1}{k}\left(1 - e^{-2k|t_1-t_2|}\right)\left(e^{-2s\min(t_1,t_2)} - e^{-sk\min(t_1,t_2)}\right)\right].
\end{aligned} \tag{S.8}$$

We can now use this expression in Eq. (S.1) to calculate the coalescence probability $P_c^{k,k\to k-2}$. Since the result is tedious and does not further illuminate the structure of the full conditional calculation, we do not do so explicitly here, but the integrals are straightforward to evaluate with the methods we have used above.

To motivate the validity of the non-conditional approximation, we need to consider the full calculation going back one additional step. Thus we consider the probability that two individuals both sampled from class k coalesce in class $k-3$, $P_c^{k,k\to k-3}$. This will be given by

$$P_c^{k,k\to k-3} = \int Q_{k,k}^{k-3}(t_1, t_2)\frac{x^2}{h_{k-3}^2}f_{k-3}(x)e^{-s(k-3)|t_1-t_2|}dt_1 dt_2 dx, \tag{S.9}$$

where here $Q_{k,k}^{k-3}(t_1, t_2)$ is the distribution of the time at which the ancestors of the two sampled individuals originally mutated from class $k-3$ to class $k-2$, conditional on them not coalescing in classes $k-2$ or $k-1$.

We can calculate $Q_{k,k}^{k-3}(t_1, t_2)$ in the same way we calculated $Q_{k,k}^{k-2}(t_1, t_2)$. Explicitly,

$$Q_{k,k}^{k-3}(t_1, t_2) = Q_{k,k}^{k-2}(t_1, t_2|nc) \star Q_{1step}^{k-3}(t_1, t_2), \tag{S.10}$$

where analogously to the expression in the previous step

$$Q_{k,k}^{k-2}(t_1, t_2|nc) = \frac{1}{1 - P_c^{k,k \to k-2}} \left[Q_{k,k}^{k-2}(t_1, t_2) - Q_{k,k}^{k-2}(t_1, t_2|c) P_c^{k,k \to k-2} \right]. \tag{S.11}$$

We note that $Q_{k,k}^{k-2}(t_1, t_2)$ is the expression in Eq. (S.7) we calculated above. As before, we have

$$Q_{k,k}^{k-2}(t_1, t_2|c) P_c^{k,k \to k-2} = I_x^{k-2} Q_{k,k}^{k-2}(t_1, t_2) e^{-s(k-2)|t_1-t_2|}, \tag{S.12}$$

hence we can write

$$Q_{k,k}^{k-2}(t_1, t_2|nc) = \frac{Q_{k,k}^{k-2}(t_1, t_2)}{1 - P_c^{k,k \to k-2}} \left[1 - I_x^{k-2} e^{-s(k-2)|t_1-t_2|} \right]. \tag{S.13}$$

Plugging the above expression back into Eq. (S.10), we obtain

$$
\begin{aligned}
Q_{k,k}^{k-3}(t_1, t_2) &= \frac{s^2(k-1)^2 k^2 s^2 (k-2)^2}{(1 - P_c^{k,k \to k-1})(1 - P_c^{k,k \to k-2})} e^{-s(k-2)(t_1+t_2)} \int_0^{t_2} \int_0^{t_1} e^{s(k-2)(y+z)} e^{s(k-1)(y+z)} \\
&\times \left[1 - I_x^{k-2} e^{-s(k-z)|y-z|} \right] \left[(1 - e^{-sy})(1 - e^{-sz}) - \frac{I_x^{k-1}}{k-2} B \right].
\end{aligned}
\tag{S.14}
$$

We could evaluate the integrals in the above expression for $Q_{k,k}^{k-3}(t_1, t_2)$ in the same way that we did in our calculation for $Q_{k,k}^{k-2}(t_1, t_2)$. We would then substitute this result for $Q_{k,k}^{k-3}(t_1, t_2)$ into an analogous calculation of $Q_{k,k}^{k-4}(t_1, t_2)$, and so on. In this way we can build up the full conditional results. The most useful way to go about this is to separate the results into powers of I_x, which is a small parameter related to the coalescent probability in each step. We see from the expression for $Q_{k,k}^{k-3}(t_1, t_2)$ that there is a term in $(I_x)^0$, which is exactly the non-conditional approximation. There are two terms involving $(I_x)^1$, and a single term involving $(I_x)^2$. In general, in the expression for $Q_{k,k}^{k-\ell}(t_1, t_2)$, we will have one $(I_x)^0$ term (which equals the result in the non-conditional approximation) plus ℓ terms proportional to I_x, $\binom{2}{\ell}$ terms proportional to $(I_x)^2$, and so on. Fortunately, the dependence on the population parameters is entirely contained within these powers of I_x. That is, the coefficients of these various powers of I_x depend *only* on k and ℓ, and not at all on the population parameters N, s, and U_d. Thus we could simply calculate a table of coefficients once, and then would be able to understand all the distributions of mutant timings (and from this all the coalescent probabilities).

In practice, it is easier to make these full conditional calculations within the sum of ancestral paths approach. As we show in File S4, that approach leads naturally to a power series in I_x of exactly the form described above, in which the leading order term is the non-conditional approximation and the additional terms represent the conditional corrections. This calculation shows that provided $I_x \ll 1$, which is true provided our usual condition that $N h_k s k \gg 1$ holds, these higher order terms are all small, and our non-conditional approximation is valid.

These full conditional results are, however, very complex and unilluminating. Therefore we focus here on

A. M. Walczak *et al.*

understanding the general structure of these results, and on showing why the non-conditional approximation is good description of the distribution of mutation timings. We can see that at each step back through the fitness distribution, the probability distribution of times shifts from the non-conditional results by a factor which is roughly proportional to the coalescence probability at that step. That is, in general we have

$$Q_{k,k}^{k-\ell}(t_1, t_2) = \frac{1}{1 - P_c^{k,k \to k-\ell}} \left[Q_{k,k}^{k-\ell}(t_1, t_2) - P_c^{k,k \to k-\ell} Q_{k,k}^{k-2}(t_1, t_2 | c) \right]. \tag{S.15}$$

The first term in square brackets reflects the fact that the probability distribution at a given step conditional on non-coalescence at that step is almost equal to the unconditional probability distribution at that step. The second term represents the correction: note that it is proportional to the coalescence probability in that step, $P_c^{k,k \to k-\ell}$. The nature of the correction can be seen by plugging in the distribution of times conditional on coalescence, giving

$$Q_{k,k}^{k-\ell}(t_1, t_2) = \frac{Q_{k,k}^{k-\ell}(t_1, t_2)}{1 - P_c^{k,k \to k-\ell}} \left[1 - I_x^{k-\ell} e^{-s(k-\ell)|t_1 - t_2|} \right]. \tag{S.16}$$

We see that the correction acts to reduce the probability that $|t_1 - t_2|$ is small — that is, it makes it more likely that t_1 and t_2 are further apart, because this is more likely to be the case given that coalescence did not occur.

Since at each step the shift in the distribution of mutant timings is proportional to the coalescence probability, and the coalescence probability at each step is small, it seems clear that the non-conditional approximation where we simply ignore this shift in mutant timings is reasonable. However there is one potential caveat we must consider: although the shift in the distribution of mutation timings due to conditioning on non-coalescence is small *in each step*, we typically take many steps before the lineages coalesce. In fact, since the shift in mutation timings is proportional to the coalescence probability, and we typically go back a number of steps of order one over the coalescence probability, in principle the shifts in mutation timings could add up to a substantial shift.

Fortunately, there are three factors which prevent this from happening. First, the shift in mutation timings at each step is always to reduce the probability of times t_1 and t_2 where $|t_1 - t_2| \lesssim \frac{1}{(k-\ell)s}$. Since at each step ℓ is increasing, and the range of separations between mutation timings at which coalescence can happen is also increasing, the shifts in mutation timings from many steps ago are not a huge factor in determining coalescence probabilities in a particular step. That is, though the shifts in mutation timings add up over many steps, the shifts most relevant to the coalescent probability in a given step do not. Second, the coalescence probabilities at each step are different. This reduces the chance that we take enough steps to shift the overall mutation timings substantially by the time we coalesce. Finally, and most importantly, we will see that the there is a substantial probability that the ancestors of the two individuals sampled do

not coalesce until they are in the most-fit class. This means that the total sum of coalescence probabilities (and hence the total possible weight in the shift of mutation timings) remains small even in the worst case where the two lineages do not coalesce for the maximum possible number of steps. The non-conditional approximation will always be good in the regime where this is true. All of these heuristic conclusions are reflected in the fact that the full conditional result we calculate in the sum of ancestral paths approach is equal to the non-conditional result plus corrections that are small provided $I_x \gg 1$.

A. M. Walczak *et al.*

FILE S2: THE NON-CONDITIONAL DISTRIBUTIONS OF MUTANT TIMINGS

Within the non-conditional approximation we need to calculate the distribution of mutant timings, as used in Eq. (48). Specifically, we need to calculate

$$Q_k^{k-\ell}(t) = Q_k^{k-1}(t) \star Q_{k-1}^{k-2}(t) \star Q_{k-2}^{k-3}(t) \star \ldots \star Q_{k-\ell+1}^{k-\ell}(t), \tag{S.17}$$

where $\star$ refers to a convolution and

$$Q_{k-\ell+1}^{k-\ell}(t) = s(k-\ell+1)e^{-s(k-\ell+1)t}, \tag{S.18}$$

as given by Eq. (6). In general, the convolution of n exponential distributions with parameters $\lambda_1 \ldots \lambda_n$ is given by

$$\sum_{i=0}^{n-1} \lambda_i e^{-\lambda_i t} \prod_{j=0,\neq i}^{n-1} \frac{\lambda_j}{\lambda_j - \lambda_i}. \tag{S.19}$$

Applying this identity with $\lambda_i = s(k-i)$, we find

$$Q_k^{k-\ell}(t) = \sum_{i=0}^{\ell-1} s e^{-s(k-i)t} \left(\frac{\prod\limits_{j=0}^{\ell-1} k-j}{\prod\limits_{j=0,\neq i}^{\ell-1} i-j} \right) \tag{S.20}$$

We can simplify this expression by noting that

$$\prod_{j=0}^{\ell-1}(k-j) = \frac{k!}{(k-\ell)!}, \tag{S.21}$$

and similarly that

$$\prod_{j=0,\neq i}^{\ell-1}(i-j) = i!(\ell-1-i)!(-1)^{\ell-1-i}. \tag{S.22}$$

This means we have

$$Q_k^{k-\ell}(t) = \sum_{i=0}^{\ell-1} s\ell e^{-s(k-i)t}(-1)^{\ell-i-1}\binom{\ell-1}{i}\binom{k}{k-\ell}. \tag{S.23}$$

We can evaluate this sum by recognizing the binomial expansion formula

$$(1+x)^n = \sum_{i=0}^{n} x^i \binom{n}{i}, \tag{S.24}$$

where we identify $x = -e^{st}$. We find

$$Q_k^{k-\ell}(t) = s\ell \binom{k}{\ell} e^{-skt} \left(e^{st}-1\right)^{\ell-1}. \tag{S.25}$$

More generally, we have

$$Q_a^b(t) = s(a - b)\binom{a}{b}e^{-sat}\left(e^{st} - 1\right)^{a-b-1}.$$

(S.26)

A. M. Walczak *et al.*

FILE S3: GENERAL COALESCENCE PROBABILITIES IN THE NON-CONDITIONAL APPROXIMATION

The probability of coalescence for two individuals originally in two different classes k and k', as defined in Eq. (48) can be rewritten as

$$P_c^{k,k' \to k'-\ell} = \frac{1}{1 + 2Nh_{k-\ell}s(k-\ell)} \left[I_1 + I_2 \right], \tag{S.27}$$

where we have defined

$$I_1 = \int_0^\infty Q_{k'}^{k-\ell}(t_1)e^{-s(k-\ell)t_1} \int_0^{t_1} Q_k^{k-\ell}(t_2)e^{s(k-\ell)t_2} dt_2 dt_1 \tag{S.28}$$

$$I_2 = \int_0^\infty Q_k^{k-\ell}(t_2)e^{-s(k-\ell)t_2} \int_0^{t_2} Q_{k'}^{k-\ell}(t_1)e^{s(k-\ell)t_1} dt_1 dt_2. \tag{S.29}$$

Note that both I_1 and I_2 involve integrals of the form

$$I_a = \int_0^t Q_a^b(t')e^{sbt'} dt'. \tag{S.30}$$

Plugging in the results for the non-conditional distributions of mutant timings, Eq. (S.26), and making use of the binomial expansion formula for $(1+x)^n$ noted in File S2, we find this integral becomes

$$I_a = s(a-b)\binom{a}{b} \int_0^t e^{s(b-a)t'} \left(e^{st'} - 1\right)^{a-b-1} dt' \tag{S.31}$$

$$= s(a-b)\binom{a}{b} \sum_{i=0}^{a-b-1} (-1)^{a-b-1+i}\binom{a-b-1}{i} \int_0^t e^{s(b-a+i)t'} dt' \tag{S.32}$$

$$= (a-b)\binom{a}{b}(-1)^{a-b} \sum_{i=0}^{a-b-1} \frac{(-1)^i}{a-b}\binom{a-b}{i} \left(e^{s(b-a+i)t} - 1\right) \tag{S.33}$$

$$= \binom{a}{b}(-1)^{a-b} \sum_{i=0}^{a-b} (-1)^i\binom{a-b}{i} \left(e^{s(b-a+i)t} - 1\right) \tag{S.34}$$

$$= \binom{a}{b}(-1)^{a-b}e^{s(b-a)t} \sum_{i=0}^{a-b} (-e^{st})^i \binom{a-b}{i} \tag{S.35}$$

$$= \binom{a}{b}e^{s(b-a)t} \left(e^{st} - 1\right)^{a-b}. \tag{S.36}$$

We now substitute this result for I_a into our expressions for I_1 and I_2. We note that both have terms of the form

$$I_b = \int_0^\infty Q_a^b(t)\binom{c}{b}e^{-sct} \left(e^{st} - 1\right)^{c-b} dt. \tag{S.37}$$

Using similar manipulations to those above, we find

$$I_b = (a-b)\binom{a}{b}\binom{c}{b}\int_0^\infty e^{-s(a+c)t}\left(e^{st}-1\right)^{a+c-2b-1}dt \tag{S.38}$$

$$= s(a-b)\binom{a}{b}\binom{c}{b}(-1)^{a+c-1}\sum_{i=0}^{a+c-2b-1}\binom{a+c-2b-1}{i}(-1)^i\int_0^\infty e^{-s(a+c-i)t}dt \tag{S.39}$$

$$= (a-b)\binom{a}{b}\binom{c}{b}(-1)^{a+c-1}\sum_{i=0}^{a+c-2b-1}(-1)^i\binom{a+c-2b-1}{i}\frac{1}{a+c-i}. \tag{S.40}$$

Using the partial fraction decomposition

$$\frac{1}{\binom{n+x}{n}} = \sum_{i=1}^n (-1)^{i-1}\binom{n}{i}\frac{i}{x+i}, \tag{S.41}$$

we find

$$I_b = \frac{\frac{a-b}{a+c-2b}\binom{a}{b}\binom{c}{b}(-1)^{a+c}}{\binom{-2b-1}{a+c-2b}} = \frac{\frac{a-b}{a+c-2b}\binom{a}{b}\binom{c}{b}(-1)^{2b}}{\binom{a+c}{a+c-2b}}. \tag{S.42}$$

We can now use this result for I_b to determine I_1 and I_2, and hence compute $P_c^{k,k'\to k'-\ell}$. We find

$$P_c^{k,k'\to k'-\ell} = \frac{1}{1+2Nh_{k-\ell}s(k-\ell)}\frac{\binom{k'}{k-\ell}\binom{k}{k-\ell}}{\binom{k+k'}{2\ell+k'-k}}. \tag{S.43}$$

As we noted in the main text, this is just

$$P_c^{k,k'\to k-\ell} = \frac{1}{1+2Nh_{k-\ell}s(k-\ell)}A_\ell^{k,k'}, \tag{S.44}$$

with $A_\ell^{k,k'}$ as defined in Eq. (16). Note that when $k = k'$, this result simplifies to $P_c^{k,k\to k-\ell}$ as defined in the main text, as expected.

FILE S4: COMPUTING SUMS OF ANCESTRAL PATHS

In this Supplementary File, we describe the calculation of $\phi_k^{k'}(\ell)$ using the sum of ancestral paths approach.

Calculation of $\phi_k^k(3)$: We begin by considering a simpler specific case, where $k = k'$ and $\ell = 3$. There are a total of $\binom{6}{3} = 20$ possible ancestral paths by which two individuals sampled from class k can coalesce in class $k - 3$. These can be separated into four types, according to whether the two ancestral lineages were ever together in classes $k - 1$ or $k - 2$. We can list all paths of each type, using the notation that A is a mutation event in the first lineage, and B is a mutation event in the second lineage. We have

$$
\begin{pmatrix}
ABABAB \\
ABABBA \\
ABBAAB \\
ABBABA \\
BAABAB \\
BAABBA \\
BABAAB \\
BABABA
\end{pmatrix}
\underbrace{\qquad}_{\binom{2}{1}\binom{2}{1}\binom{2}{1}=8 \text{ ways}}
\begin{pmatrix}
ABAABB \\
ABBBAA \\
BAAABB \\
BABBAA
\end{pmatrix}
\underbrace{\qquad}_{\binom{2}{1}\left(\binom{4}{2}-\binom{2}{1}\binom{2}{1}\right)=4 \text{ ways}}
\begin{pmatrix}
AABBAB \\
AABBBA \\
BBAAAB \\
BBAABA
\end{pmatrix}
\underbrace{\qquad}_{\binom{2}{1}\left(\binom{4}{2}-\binom{2}{1}\binom{2}{1}\right)=4 \text{ ways}}
\begin{pmatrix}
AAABBB \\
AABABB \\
BBBAAA \\
BBABAA
\end{pmatrix}
\underbrace{\qquad}_{\binom{6}{3}-\text{others}=4\text{ways}}
.
$$

The probabilities of all paths of a particular type are identical. We can calculate the probability of each of the four types of paths using the same logic as outlined in the main text. We find

$$
P(AAABBBc) \;=\; I_x^{k-3}\frac{k(k-1)(k-2)}{8(2k-1)(2k-3)(2k-5)}\left(1 - I_x^k\right), \tag{S.45}
$$

$$
P(AABBABc) \;=\; I_x^{k-3}\frac{k(k-1)(k-2)}{8(2k-1)(2k-3)(2k-5)}\left(1 - I_x^k\right)\left(1 - I_x^{k-1}\right), \tag{S.46}
$$

$$
P(ABAABBc) \;=\; I_x^{k-3}\frac{k(k-1)(k-2)}{8(2k-1)(2k-3)(2k-5)}\left(1 - I_x^k\right)\left(1 - I_x^{k-2}\right), \tag{S.47}
$$

$$
P(ABABABc) \;=\; I_x^{k-3}\frac{k(k-1)(k-2)}{8(2k-1)(2k-3)(2k-5)}\left(1 - I_x^k\right)\left(1 - I_x^{k-1}\right)\left(1 - I_x^{k-2}\right). \tag{S.48}
$$

Summing over all the possible paths, we find

$$
\phi_k^k(3) = I_{k-3}\frac{\binom{k}{k-3}\binom{k}{k-3}}{\binom{2k}{6}}\left[1 - \frac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}}I_{k-1} - \frac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}}I_{k-2} + \frac{\binom{2}{1}\binom{2}{1}\binom{2}{1}}{\binom{6}{3}}I_{k-1}I_{k-2}\right]. \tag{S.49}
$$

We now pause to consider the form of the probabilities of each type of ancestral path. These probabilities differ only by factors of $(1 - I_x^{k-i})$. One such factor arises each time the two ancestral lineages are together in class $k - i$. In other words, we can rewrite the probability of each path as the probability of an undistorted path (defined to be a path in which the contributions due to the possibility of coalescence in previous classes

are neglected), times a correction for each class in which the two lineages are together:

$$P(AAABBBc) \;=\; P(\text{Undistorted Path})\left(1 - I_x^k\right) \tag{S.50}$$

$$P(AABBABc) \;=\; P(\text{Undistorted Path})\left(1 - I_x^k\right)\left(1 - I_x^{k-1}\right) \tag{S.51}$$

$$P(ABAABBc) \;=\; P(\text{Undistorted Path})\left(1 - I_x^k\right)\left(1 - I_x^{k-2}\right) \tag{S.52}$$

$$P(ABABABc) \;=\; P(\text{Undistorted Path})\left(1 - I_x^k\right)\left(1 - I_x^{k-1}\right)\left(1 - I_x^{k-2}\right). \tag{S.53}$$

By definition, the "undistorted path" probability is the probability neglecting the contributions due to the possibility of coalescence in previous steps, and is therefore the same for all paths. We have

$$P(\text{Undistorted Path}) \;=\; \frac{k(k-1)(k-2)k(k-1)(k-2)}{2k(2k-1)(2k-2)(2k-3)(2k-4)(2k-5)}I_x^{k-\ell} \tag{S.54}$$

$$= \; \frac{\frac{k!}{(k-3)!}\frac{k!}{(k-3)!}}{\frac{2k!}{(2k-6)!}}I_x^{k-\ell}. \tag{S.55}$$

Using these results, we can write $\phi_k^k(3)$ as

$$\phi_k^k(3) \;=\; [\# \text{ of Paths}]\,P(\text{Undistorted Path})\left[F_k(1 - I_x^k) + F_{k,k-1}(1 - I_x^k)(1 - I_x^{k-1})\right.$$

$$\left. + F_{k,k-2}(1 - I_x^k)(1 - I_x^{k-2}) + F_{k,k-1,k-2}(1 - I_x^k)(1 - I_x^{k-1})(1 - I_x^{k-2})\right], \tag{S.56}$$

where we have defined $F_{\{a\}}$ to be the fraction of paths that are together in the set of classes $\{a\}$ (and are not together in any other class).

Calculation of $\phi_{k'}^k(\ell)$: We now use this approach to calculate the coalescence probability in the general case. The probability of any particular ancestral path from k and k' to $k - \ell$ is the product of the individual probabilities of each mutational step that makes up this path. Each such individual probability consists of three parts: a numerator, which depends only on the current class of the lineage that mutates, divided by a denominator, which depends only on the sum of the current set of classes for both lineages, times a correction factor of $(1 - I_x^{k-i})$ if the two lineages are in the same class at that step.

Although in each ancestral path the mutations will occur in a different order, all paths will ultimately consist of the same set of mutations ($k' \to k' - 1 \to \ldots \to k - \ell$ and $k \to k - 1 \to \ldots \to k - \ell$). Therefore, regardless of the path taken, the product of the numerators from each step will be identical. Similarly, the sum of the current set of classes will begin at $k' + k$, and decrement by one each time a deleterious mutation occurs, until both lineages are in the final class ($k' + k \to k' + k - 1 \to \ldots \to 2k - 2\ell$). Therefore, regardless of the path taken, the product of the denominators from each step will also be identical. Therefore, the paths will differ only by the correction factor $(1 - I_x^{k-i})$ for each class in which the two ancestral lineages are together. This means that, analogous to the case of $\phi_k^k(3)$ we described above, the probability of each

A. M. Walczak *et al.*

path is the probability of an "undistorted path" times the appropriate correction factor. The probability of the undistorted path is

$$P(\text{Undistorted Path}) = \frac{k'(k'-1)\dots(k-\ell+1)k(k-1)\dots(k-\ell+1)}{(k'+k)(k'+k-1)\dots(2k-2\ell+1)}I_x^{k-\ell}. \tag{S.57}$$

We can now sum up all possible paths to obtain

$$
\begin{aligned}
\phi_{k'}^k(\ell) &= [\# \text{ of Paths}]\, P(\text{Undistorted Path}) \left[F_\emptyset + \sum_{i=0}^{\ell} F_{k-i}(1 - I_x^{k-i}) \right. \\
&\quad + \sum_{i=0}^{\ell-1}\sum_{j>i}^{\ell} F_{k-i,k-j}(1 - I_x^{k-i})(1 - I_x^{k-j}) \\
&\quad \left. + \sum_{i=0}^{\ell-2}\sum_{j>i}^{\ell-1}\sum_{m>j}^{\ell} F_{k-i,k-j,k-m}(1 - I_x^{k-i})(1 - I_x^{k-j})(1 - I_x^{k-m}) + \dots \right],
\end{aligned}
\tag{S.58}
$$

where as before $F_{\{a\}}$ is the fraction of paths that are together in the set of classes $\{a\}$ (and are not together in any other class). Note that there are a total of $\ell + 1$ terms in this equation, representing the possibility that the two lineages can be together in anywhere from 0 to ℓ of the classes. We can rearrange these terms to write

$$
\begin{aligned}
\phi_{k'}^k(\ell) &= [\# \text{ of Paths}]\, P(\text{Undistorted Path}) \left[1 - \sum_{i=0}^{\ell} G_{k-i} I_x^{k-i} \right. \\
&\quad + \sum_{i=0}^{\ell-1}\sum_{j>i}^{\ell} G_{k-i,k-j} I_x^{k-i} I_x^{k-j} \\
&\quad \left. - \sum_{i=0}^{\ell-2}\sum_{j>i}^{\ell-1}\sum_{m>j}^{\ell} G_{k-i,k-j,k-m} I_x^{k-i} I_x^{k-j} I_x^{k-m} + \dots \right],
\end{aligned}
\tag{S.59}
$$

where we have defined $G_{\{a\}}$ to be the fraction of paths that are together in *at least* the set of classes $\{a\}$.

We can evaluate each of these factors of G. For example, the fraction of paths that are together in class $k - i$ equals the number of ways for the two lineages to descend from classes k' and k to be together in class $k - i$, $\binom{k'-k+2i}{i}$, times the number of ways for the two lineages to descend from class $k - i$ to be together in class $k - \ell$, $\binom{2i-2\ell}{i-\ell}$, divided by the total number of ways for the two lineages to descend from classes k' and k to be together in $k - \ell$, $\binom{k'-k+2\ell}{\ell}$. Using this logic, we find

$$
\begin{aligned}
\phi_{k'}^k(\ell) &= [\# \text{ of Paths}]\, P(\text{Undistorted Path}) \\
&\quad \times \left[1 - \sum_{i=0}^{\ell-1} \frac{\binom{k'-k+2i}{i}\binom{2\ell-2i}{\ell-i}}{\binom{k'-k+2\ell}{\ell}} I_x^{k-i} + \sum_{i=0}^{\ell-2}\sum_{j>i}^{\ell-1} \frac{\binom{k'-k+2i}{i}\binom{2j-2i}{j-i}\binom{2\ell-2j}{\ell-j}}{\binom{k'-k+2\ell}{\ell}} I_x^{k-i} I_x^{k-j} \dots \right].
\end{aligned}
\tag{S.60}
$$

The total number of paths is $\binom{k'-k+2\ell}{\ell}$, so we finally find that the full probability of coalescence in class

$k - \ell$ is

$$
\begin{aligned}
\phi_k^{k'}(\ell) \;=\;& I_x^{k-\ell} \frac{\binom{k'}{k-\ell}\binom{k}{k-\ell}}{\binom{k'+k}{k'-k+2\ell}} \left[1 - \sum_{i=0}^{\ell-1} \frac{\binom{k'-k+2i}{i}\binom{2\ell-2i}{\ell-i}}{\binom{k'-k+2\ell}{\ell}} I_x^{k-i} + \right.\\
& \left. \sum_{i=0}^{\ell-2}\sum_{j>i}^{\ell-1} \frac{\binom{k'-k+2i}{i}\binom{2j-2i}{j-i}\binom{2\ell-2j}{\ell-j}}{\binom{k'-k+2\ell}{\ell}} I_x^{k-i} I_x^{k-j} - \cdots \right] .
\end{aligned}
\tag{S.61}
$$

This is Eq. (56) from the main text. Note that it equals our non-conditional result for $P_c^{k,k'\to\ell}$ times a correction factor. There are a total of $\ell+1$ terms in this correction factor. This full correction factor can be arbitrarily complex for large ℓ, so we do not write out a general form here. However, it is straightforward to calculate for any values of k, k', and ℓ; a Mathematica script to do so is available on request.

A. M. Walczak *et al.*

FILE S5: THE CORRESPONDENCE BETWEEN STEPTIMES AND REAL TIMES

In this Supplementary File, we calculate the correspondence between steptimes and the actual times measured in generations. Our goal is to calculate the probability distribution of real coalescence times, $\psi(t|k,k',\ell)$, given that individuals were initially in classes k and k' and coalesced in class $k-\ell$.

To begin, we neglect the coalescence time within class $k-\ell$, and consider the distribution of the time at which an ancestor of one of the two sampled individuals first mutated from class $k-\ell$ to class $k-\ell+1$. We refer to this as $\psi_1(t|k,k',\ell)$. We first calculate the joint distribution of the times at which both ancestors mutated out of the class, $R_{k,k'}^{k-\ell}(t_1,t_2)$. Conditional on coalescence in class $k-\ell$, $R_{k,k'}^{k-\ell}(t_1,t_2)$, is given by the probability of t_1 and t_2 and coalescence divided by the total probability of coalescence. That is,

$$R(t_1,t_2) = \frac{P(coal|t_1,t_2)P(t_1,t_2)}{P(coal)}. \tag{S.62}$$

Substituting in the relevant expressions from the main text, this gives

$$R_{k,k'}^{k-\ell}(t_1,t_2) = \frac{1}{A_\ell^{k,k'}} Q_{k,k'}^{k-\ell}(t_1,t_2)e^{-s(k-\ell)|t_1-t_2|}. \tag{S.63}$$

The time at which the first ancestor mutated out of class $k-\ell$ is the longer of the two times t_1 and t_2,

$$\psi(t|k,k',\ell) = \left[\int_0^t R_{k,k'}^{k-\ell}(t_1,t)dt_1 + \int_0^t R_{k,k'}^{k-\ell}(t,t_2)dt_2 \right]. \tag{S.64}$$

Substituting in our expression for $R_{k,k'}^{k-\ell}(t_1,t_2)$ and carrying out the integrals as in File S3, we find

$$\psi_1(t|k,k',\ell) = s\pi_d e^{-s(k'+k)t}(e^{st}-1)^{\pi_d-1}\binom{k'+k}{\pi_d}, \tag{S.65}$$

where we have used $\pi_d = k'-k+2\ell$.

We can alternatively calculate $\psi_1(t|k,k',\ell)$ using our sum of ancestral paths approach. As before, we imagine two individuals sampled from classes k and k' and condition on them coalescing in class $k-\ell$. Consider a case where $k \neq k'$. Then the first event in the history of these two individuals must be a deleterious mutation. Since these mutations happen at rate sk and sk' in each lineage, the distribution of times since this mutation occurred in one of the two ancestral lineages is

$$P(t) = s(k+k')e^{-s(k+k')t}. \tag{S.66}$$

With probability $\frac{k'}{k+k'}$, this mutation is in the lineage sampled from class k', in which case the two lineages are now in classes k and $k'-1$. Alternatively, the mutaion occurred in the lineage sampled from k and the lineages are in classes $k-1$ and k'.

We can now consider the time to the next event backwards in time. If the two lineages are in the same

class (but not yet in class $k-\ell$), the distribution of times to the next deleterious mutation event is somewhat shorter, because we are conditioning on coalescence not occuring. However, provided that $2sk_1 \gg \frac{1}{Nh_k}$ (the condition we are already making elsewhere), this shortening of the time will be a small correction and neglecting it is a good approximation.

Making this approximation, the rate at which the next deleterious mutation event occurs when the two lineages are in classes k_1 and k_2 is just $s(k_1 + k_2)$. Regardless of the order in which these mutations happen between the two lineages, this sum is simply decreased by s at each step. This will continue until the both ancestral lineages are in class $k - \ell$. Therefore, the distribution of times until the original mutation out of class $k - \ell$ is given by:

$$\psi_1(t|k',k,\ell) = s(k'+k)e^{-s(k'+k)t} \star s(k'+k-1)e^{-s(k'+k-1)t} \star \ldots \star s(2k-2\ell+1)e^{-s(2k-2\ell+1)t}. \quad \text{(S.67)}$$

This can be written as

$$\psi_1(t|k',k,\ell) = \lambda_0 e^{-\lambda_0 t} \star \lambda_1 e^{-\lambda_1 t} \star \ldots \star \lambda_{k'-k+2\ell-1} e^{-\lambda_{k'-k+2\ell-1}t}, \quad \text{(S.68)}$$

where we have defined:

$$\lambda_i = s(k'+k-i). \quad \text{(S.69)}$$

We can compute this convolution as in File S2 (compare to Eq. (S.17) for $Q_{k+k'}^{2k-2\ell}(t)$). We find

$$\psi_1(t|k,k',\ell) = s\pi_d e^{-s(k'+k)t}(e^{st}-1)^{\pi_d-1}\binom{k'+k}{\pi_d}, \quad \text{(S.70)}$$

identical to the result of our lineage structure calculation above.

Distribution of Coalescence Times: To calculate the correspondence between steptimes and real times, we now need to add the time it takes two individuals two coalesce in class $k - \ell$, which we refer to as $\psi_2(t|k,k',\ell)$, to the time it took them both to get to that class, $\psi_1(t|k,k',k-\ell)$. The rate of coalescence once in class $k - \ell$ is $\frac{1}{Nh_{k-\ell}}$, so we have

$$\psi_2(t|k',k,\ell) = \left(2s(k-\ell) + 1/Nh_{k-\ell}\right)e^{-[2s(k-\ell)+1/Nh_{k-l}]t}. \quad \text{(S.71)}$$

Putting this together, the full distribution of times since coalescence is

$$\psi(t|k',k,\ell) = \psi_1(t|k',k,\ell) \star \psi_2(t|k',k,\ell). \quad \text{(S.72)}$$

Carrying out this convolution (and expanding the binomial factor $(e^{st}-1)^{\pi_d-1}$ in ψ_1), we find

$$\psi(t|k',k,\ell) = \sum_{i=0}^{\pi_d-1} s\pi_d(-1)^{\pi_d-i-1}\binom{\pi_d-1}{i}\binom{k'+k}{\pi_d}\frac{B}{A-B}\left(e^{-sBt}-e^{-sAt}\right), \quad \text{(S.73)}$$

A. M. Walczak *et al.*

where we have defined $A \equiv k' + k - i$ and $B \equiv 2\left(k - \ell\right) + \frac{1}{Nsh_{k-\ell}}$.

FILE S6: AN ALTERNATIVE APPROACH TO NEUTRAL DIVERSITY

Instead of calculating the distribution of neutral heterozygosity by first computing the distribution of real times, we could alternatively incorporate neutral mutations directly into the sum of ancestral paths framework. This completely bypasses the correspondence with real coalescence times. To do this, we characterize ancestral paths not only by the ordering of deleterious mutation and coalescence events, but also by the ordering of neutral mutations. This means that if we sample two individuals A and B, there are five types of events that can happen in their ancestral paths: a deleterious mutation (DM) in A or in B, a neutral mutation (NM) in either A or in B, and or a coalescence (C) event (if A and B are currently in the same class).

We now imagine that we sample two individuals from classes k and k', and that they coalesce in class $k - \ell$. Our goal is to calculate the probability distribution of π_n given k, k', and ℓ, $\rho(\pi_n | k, k', \ell)$. We will find it helpful to divide the five types of events that can occur into two classes: neutral mutations on the one hand, and deleterious mutations or coalescence (which we call "steps") on the other. We begin by computing the probability that a given number of NMs occur before the next DM or C events (i.e. the number of neutral mutations that occur at this "step"). We have

$$P\big(a \text{ NMs, then DM in } k' \text{ or } k' | k', k\big) = \left(\frac{\frac{2U_n}{s}}{k' + k + \frac{2U_n}{s}} \right)^a \frac{k + k'}{k' + k + \frac{2U_n}{s}}, \tag{S.74}$$

where we have made our usual assumption that $Nh_k sk \gg 1$, allowing us to neglect the rates of coalescence events (when $k = k'$) in writing this expression.

This probability only depends on the sum of the current classes the individulas are in. At each subsequent step, regardless of the path taken, this sum of the classes will decrease by one. Therefore, the probability that a_i neutral mutations occur at step i is independent of the path taken. This observation allows us to calculate the probability that a given total number of neutral mutations have occurred since coalescence. We first calculate the probability that a given number of neutral mutations have occurred since the first deleterious mutation out of the $k - \ell$ class. We will add in the additional neutral mutations once in the $k - \ell$ class at the end.

In order for π_n neutral mutations to have occurred since the first deleterious mutation out of class $k - \ell$, we require that a_0 mutations occurred at the first step, a_1 mutations occurred at the second step, and so on, such that $a_0 + a_1 + \ldots + a_{k'-k+2\ell-1} = \pi_n$. This gives

$$\rho(\pi_n = X | k', k, \ell) = \frac{\frac{(k'+k)!}{(2k-2\ell)!}}{\frac{(\frac{2U_n}{s}+k'+k)!}{(\frac{2U_n}{s}+2k-2\ell)!}} \sum_{|\vec{a}|=X} \left(\frac{2U_n/s}{2U_n/s + k + k'} \right)^{a_0} \cdots \left(\frac{2U_n/s}{2U_n/s + 2k - 2l + 1} \right)^{a_{k'-k+2l-1}}. \tag{S.75}$$

A. M. Walczak *et al.*

We can define $x \equiv 2U_n/s + k + k'$, recognize $\pi_d = k' - k + 2\ell$, and relabel the a_i as

$$a_0 \to X - b_0, \quad a_1 \to b_0 - b_1, \quad \ldots \quad a_{\pi_d-2} \to b_{\pi_d-3} - b_{\pi_d-2}, \quad a_{\pi_d-1} \to b_{\pi_d-2}. \tag{S.76}$$

This gives

$$
\rho(\pi_n = X | k', k, \ell) = \frac{\binom{k'+k}{\pi_d}}{\binom{\frac{2U_n}{s}+k'+k}{\pi_d}} \left(\frac{2U_n}{s}\right)^X \left(\frac{1}{x}\right)^X \sum_{b_0=0}^{X} \left(\frac{x}{x-1}\right)^{b_0}
$$
$$
\sum_{b_1=0}^{b_0} \left(\frac{x-1}{x-2}\right)^{b_1} \cdots \sum_{b_{\pi_d-2}=0}^{b_{\pi_d-3}} \left(\frac{x-\pi_d+2}{x-\pi_d+1}\right)^{b_{\pi_d-2}}. \tag{S.77}
$$

To simnplify this expression, it is helpful to define a function f such that:

$$
\mathbf{f}(A,B) \equiv \left(\frac{1}{x}\right)^X \sum_{b_0=0}^{X} \left(\frac{x}{x-1}\right)^{b_0} \tag{S.78}
$$
$$
\sum_{b_1=0}^{b_0} \left(\frac{x-1}{x-2}\right)^{b_1} \cdots \sum_{b_{A-1}=0}^{X} \left(\frac{x-A+1}{x-A}\right)^{b_0} \sum_{b_A=0}^{b_{A-1}} \left(\frac{x-A}{x-B}\right)^{b_A}
$$

In other words, $\mathbf{f}(A,B)$ is a set of A nested sums, each of the same form, except for the final sum, which can have a different denominator. Using this definition, we have

$$
P(\pi_n = X | k', k, \ell) = \frac{\binom{k'+k}{\pi_d}}{\binom{\frac{2U_n}{s}+k'+k}{\pi_d}} \left(\frac{2U_n}{s}\right)^X \mathbf{f}(\pi_d - 2, \pi_d - 1). \tag{S.79}
$$

The virtue of this definition is that this sum can be solved recursively. We have

$$
\sum_{b_A=0}^{b_{A-1}} \left(\frac{x-A}{x-B}\right)^{b_A} = \frac{x-B}{A-B} - \frac{x-A}{A-B} \left(\frac{x-A}{x-B}\right)^{b_{A-1}}. \tag{S.80}
$$

Therefore we have

$$
\mathbf{f}(A,B) = \frac{x-A}{B-A}\mathbf{f}(A-1,B) - \frac{x-B}{B-A}\mathbf{f}(A-1,A). \tag{S.81}
$$

Repeatedly inserting this result yields:

$$
\mathbf{f}(A,A+1) \to \frac{(x-A)(x-A-1)}{1} \left(\frac{\mathbf{f}(A-1,A+1)}{x-A-1} - \frac{\mathbf{f}(A-1,A)}{x-A}\right)
$$
$$
\mathbf{f}(A,A+1) \to \frac{(x-A+1)(x-A)(x-A-1)}{2} \left[\frac{\mathbf{f}(A-2,A+1)}{x-A-1} - \frac{2\mathbf{f}(A-2,A)}{x-A} + \frac{\mathbf{f}(A-2,A-1)}{x-A+1}\right]
$$
$$
\vdots
$$
$$
\mathbf{f}(A,A+1) \to (m+1)\binom{x-A-1+m}{m+1} \sum_{i=0}^{m} \frac{(-1)^{i+m}}{x-A-1+i} \binom{m}{i}\mathbf{f}(A-m,A+1-i). \tag{S.82}
$$

A. M. Walczak *et al.*

Note that $\mathbf{f}(-1, B) = 1/B^X$, since there are no more sums to compute. Thus, for $m = A + 1$ we have

$$\mathbf{f}(A, A+1) = (A+2)\binom{x}{A+2}\sum_{i=0}^{A+1}\frac{(-1)^{i+A+1}}{(x-A-1+i)^{X+1}}\binom{A+1}{i}. \tag{S.83}$$

Relabeling the sum and taking $A = \pi_d - 2$, we have

$$\mathbf{f}(\pi_d - 2, \pi_d - 1) = \pi_d\binom{x}{\pi_d}\sum_{i=0}^{\pi_d-1}\frac{(-1)^i}{(x-i)^{X+1}}\binom{\pi_d - 1}{i}. \tag{S.84}$$

We can now substitute these results into our expression for π_n, to find

$$\rho_1(\pi_n = X|k', k, \ell) = \pi_d\binom{k'+k}{\pi_d}\left(\frac{2U_n}{s}\right)^X\sum_{i=0}^{\pi_d-1}\frac{(-1)^i}{(2U_n/s + k + k' - i)^{X+1}}\binom{\pi_d - 1}{i} \tag{S.85}$$

Note, however, that this is only the distribution of neutral mutations since the first deleterious mutation out of class $k - l$. It is also possible for neutral mutations to occur prior to the coalescence event. Adding in this factor, we find

$$\begin{aligned}\rho(\pi_n = X|k', k, \ell) &= \pi_d\binom{k'+k}{\pi_d}\sum_{i=0}^{\pi_d-1}(-1)^i\binom{\pi_d - 1}{i} \\ &\quad\times \sum_{X=0}^{\pi_n}\frac{(2U_n/s)^X}{(2U_n/s + k + k' - i)^{X+1}}\left(\frac{2N_{k-l}U_n}{1 + 2N_{k-l}U_n + 2N_{k-l}s(k-l)}\right)^{\pi_n - X}.\end{aligned} \tag{S.86}$$

Rearranging this expression gives

$$\rho(\pi_n|k', k, \ell) = \sum_{i=0}^{\pi_d-1}\pi_d(-1)^{\pi_d-i-1}\binom{\pi_d - 1}{i}\binom{k'+k}{\pi_d}\frac{B}{A-B}\left(\frac{(\frac{2U_n}{s})^{\pi_n}}{(\frac{2U_n}{s} + B)^{\pi_n+1}} - \frac{(\frac{2U_n}{s})^{\pi_n}}{(\frac{2U_n}{s} + A)^{\pi_n+1}}\right), \tag{S.87}$$

where we have defind

$$A = k' + k - i, \qquad B = 2(k - \ell) + \frac{1}{Nsh_{k-l}}, \tag{S.88}$$

identical to our earlier result.

A. M. Walczak *et al.*