OXFORD GENETICS

GENETICS, 2026, 233(4), iyag127
https://doi.org/10.1093/genetics/iyag127
Advance Access Publication Date: 15 May 2026
Investigation

A classification of structured-coalescent processes with migration, conditional on the population pedigree

Sabin Lessard,[1] Terry Easlick ,[1] John Wakeley[2,*]

[1]Département de mathématiques et de statistique, Université de Montréal, Montreal, QC H3C 3J7, Canada
[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, United States

*Corresponding author: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, United States. Email: wakeley@fas.harvard.edu

Organismal genealogies or pedigrees of populations may have strong effects on patterns of variation. We extend recent analyses of the effects that pedigrees have on times to common ancestry to cases of population subdivision and migration. These analyses condition on the population pedigree, while traditional coalescent models marginalize over it. A finding of a "pedigree effect" implies that data analysis and interpretation should not be based on the corresponding traditional coalescent model but rather on a coalescent model obtained by conditioning on the pedigree. We apply a straightforward test based on the distribution of pairwise coalescence times to four previously described scenarios of subdivision and migration. These scenarios are defined by the relative magnitudes of four parameters: the number of local populations or demes, the deme size, the migration fraction, and the probability that migration can occur at all. We find pedigree effects in three scenarios. In two of the three, the effect is weak if the deme size is large. The one scenario without any pedigree effect corresponds to the well-known structured-coalescent model. The one scenario with a persistent pedigree effect even in the limit as the deme size tends to infinity involves long periods without gene flow interrupted by pulses of migration. We illustrate our results using simulations and numerical analysis. The implications of our results for statistical genetic inference in structured populations include that the structured-coalescent model may be applied to multiple unlinked loci, and that pedigrees will need to be accounted for when deme sizes are not large.

Keywords: conditional coalescent; conditional survival function; coalescence times; structured populations

Introduction

Coalescent models describe genetic ancestry in order to predict genetic variation in a sample from a population. These continuous-time models measure time backwards from the present to the past and make minimal reference to the population beyond the sample. Generally, they are obtained by taking the limit as the population size tends to infinity with time rescaled so the coalescence rate for a pair of ancestral lineages is equal to one. Their development can be traced through key papers such as Ewens (1972) and Watterson (1975), and considered mature in Kingman (1982a, 1982b, 1982c), Hudson (1983a, 1983b), and Tajima (1983). After these initial descriptions for well-mixed populations of constant size without selection, coalescent models were extended to various kinds of population structure, changes in population size, and selection; see the detailed overview by Nordborg (2001) or fuller treatments by Hein et al. (2005), Berestycki (2009), and Wakeley (2009).

When applied to diploid biparental organisms, these traditional coalescent models describe gene genealogies or other aspects of genetic ancestry without reference to the organismal genealogy or pedigree of the population. They instead describe ancestral genetic processes marginal to the pedigree. Their predictions are averages over all possible pedigrees. See Diamantidis et al. (2024) for a review of this practice and its historical development. Ball et al. (1990) were the first to ask whether distributions of gene genealogies conditional on the population pedigree might differ substantially from the marginal distributions. The motivation behind this question is that there is an actual pedigree of any population, and it is the structure of this pedigree that determines possible genetic ancestries of a sample. If the conditional and marginal distributions of gene genealogies differ, the utility of the corresponding traditional coalescent model may be limited.

Distributions of coalescence times or other aspects of genetic ancestry conditional on the population pedigree are predictions about what should be observed in data from a very large number of unlinked loci in a sample of individuals. This is the sampling structure implicit in many applications of traditional coalescent models, for example, when likelihoods are multiplied across loci in a sample of individuals using programs such as *∂a∂i* (Gutenkunst et al. 2009), *momi2* (Kamm et al. 2020), and *fastsimcoal2* (Excoffier et al. 2013, 2021), or when sample site-frequency spectra are compared to model predictions (Tajima 1989; Braverman et al. 1995; Fu 1995; Achaz 2009; Gao and Keinan 2016; Ragsdale et al. 2018; Patel et al. 2025). However, it is only conditional on the pedigree that patterns of variation at unlinked loci are independent.

We note that sometimes the predictions used in these applications come from forward-time Wright-Fisher diffusion theory rather than backward-time coalescent models (Ewens 2004). But predictions about site frequencies or other measures of genetic

Received on 18 February 2026; accepted on 10 May 2026
© The Author(s) 2026. Published by Oxford University Press on behalf of The Genetics Society of America. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

variation from diffusion theory are also marginal to the population pedigree. They are derived by averaging over the outcomes of reproduction. Traditional coalescent models are backward-time versions, or dual processes, of the corresponding forward-time diffusion models (Donnelly 1986; Ewens 1990; Möhle 1999).

The groundbreaking work of Ball et al. (1990) used simulations of a well-mixed population of 100 individuals and 50 independent loci to suggest that the effect of the pedigree on times to genetic common ancestry was limited. Subsequent simulations of much larger populations and numbers of loci reinforced this conclusion (Wakeley et al. 2012), in particular that distributions of coalescence times among unlinked loci in a sample of individuals conditional on their shared pedigree conform well to predictions from the standard neutral coalescent model of Kingman (1982a, 1982b, 1982c), Hudson (1983a, 1983b), and Tajima (1983). Tyukin (2015) proved this for a monoecious diploid, well-mixed population of constant size and limited variance of offspring numbers (Kingman 1982a): as the population size tends to infinity, conditioning on the pedigree or marginalizing over it leads to the same result which is the standard neutral coalescent model.

This coincidence of limiting results cannot be assumed to hold for populations outside the domain of the standard neutral coalescent model. For both modeling and inference, it is important to know for all kinds of populations: Does the limiting coalescent process conditional on the population pedigree differ from the corresponding traditional coalescent model? A negative answer, that there is no pedigree effect, means we can keep using the statistical models and methods of inference developed for the corresponding traditional coalescent model. A positive answer means we need to develop new models and methods. Here, we provide answers for different types of subdivided populations comprised of discrete local subpopulations, or demes, connected by migration.

We do this for an extended version of Wright's finite island model (Wright 1931, 1943; Maruyama 1970; Latter 1973) in which the times and sizes for migration events can be decoupled, as detailed in section "Structured population model with gametic migration." Specifically, in section "Four kinds of conditional coalescent models," we consider four limits of the model under different assumptions about the relative magnitudes of its parameters. We identify cases in which the conditional coalescent process and the corresponding pedigree-averaged process differ and cases in which they are the same. Briefly and before getting into the technical details, these limits and our findings are as follows.

The first limit is the one which gives the well-known structured-coalescent model in the pedigree-averaged case (Takahata 1988; Notohara 1990; Herbots 1997; Wilkinson-Herbots 1998). This is a model for large demes and correspondingly small migration probabilities. Of the four models we consider, this is by far the most frequently applied (Nordborg 1997; Beerli and Felsenstein 2001; Pinho and Hey 2010; Lohse et al. 2011, 2016; Costa and Wilkinson-Herbots 2017; Müller et al. 2017; Beerli et al. 2019). In this case, our results suggest that coalescence conditional on the pedigree can be described using the available structured-coalescent model.

The second limit is the many-demes limit (Wakeley 1998, 2003; Wakeley and Takahashi 2004) of the island migration model (Wright 1931, 1943; Maruyama 1970; Latter 1973). This is a model for finely subdivided populations or metapopulations (Pannell and Charlesworth 2000; Pannell 2003; Hanski and Gaggiotti 2004). In this case, we show that the conditional and pedigree-averaged pairwise coalescent processes differ, but only for ancestral lines starting in the same deme. An essentially similar result has recently been shown to hold for populations with appreciable levels of self-fertilization (Fan et al. 2025; Newman et al. 2025). Roughly speaking, and as elaborated in Lessard and Wakeley (2004) in the context of a traditional coalescent model, demes play the role of individuals and migration plays the role of outcrossing.

The third limit is the low-migration limit of Lande (1979, 1985) and Slatkin (1981). This is a separation of time-scales approximation in which events within finite-sized demes resolve quickly against a backdrop of long waiting times between migration events, due to a constant low probability of migration. It has been used to understand forward-time dynamics of frequency-dependent selection (Lande 1979, 1985; Slatkin 1981; Cherry 2003; Nishino and Tajima 2004, 2005; Ladret and Lessard 2007) as well as to describe patterns of coalescence and neutral genetic variation (Takahata 1991; Notohara 2001; Baronchelli and Pastor-Satorras 2009; Burden and Griffiths 2018). In this case, we show that the conditional and pedigree-averaged coalescent processes differ for ancestral lines starting in different demes.

The fourth limit is a model of stochastic migration (Sved and Latter 1977; Nagylaki 1979; Latter and Sved 1981; Whitlock 1992) in which a potentially large fraction of a deme is replaced by migrants but these events occur rarely, at random times and with no gene flow in between (Peniston et al. 2024). Our model is similar to the ones which Gaggiotti and Smouse (1996), Rice and Papadopoulos (2009), Yamaguchi and Iwasa (2013, 2017), and Aubree et al. (2022) studied in other contexts. We refer to this as the rare-migration limit. In this final case as well, we show that the conditional and pedigree-averaged coalescent processes differ as in the low-migration limit. We note that Newman (2026) recently also found such an effect for samples of arbitrary size in a general model of subdivision and migration, and that it is analogous to the pedigree effects of large reproduction events in well-mixed populations (Diamantidis et al. 2024; Alberti et al. 2025).

The three models for which we do find a pedigree effect have in common that the deme size is fixed and finite in the respective limit, though we do note that the total population size becomes infinite in the many-demes limit. Given that single finite populations show pedigree effects (Wakeley et al. 2012; Tyukin 2015), we further ask whether our results for these three models depend on the assumption of finite deme size. In the many-demes limit and the low-migration limit, the effect of conditioning on the pedigree becomes negligible as the deme size increases. In the rare-migration limit, the pedigree effect persists even as the deme size tends to infinity.

These will be our main findings. They primarily support the continued use of the widely used structured-coalescent model of population subdivision and migration with demes of large enough size, which was never justified with respect to the population pedigree. They also point toward the development of coalescent models conditional on the population pedigree for application to metapopulations, populations with very limited gene flow, and populations with large migration events or demes of small size. In addition, in the next section, we provide a new explanation of a method for assessing whether conditioning on the pedigree leads to a different coalescent model than averaging over the pedigree or whether both approaches give the same results in the limit.

The conditional survival function and its variance

In this work, we characterize distributions of single-locus pairwise coalescence times, conditional on the population pedigree and on

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

how the two genes are sampled in the current population. In general, we will be concerned with two types of samples: two genes at a locus in a single individual chosen at random from the population, and two genes in two individuals sampled from two different demes. Our methods will cover other sampling configurations as well. For illustration, in this section, we focus on the two genes within an individual.

In addition, we consider the most familiar way that coalescence times are defined for structured diploid populations undergoing discrete, nonoverlapping generations, namely in the limit as the deme size N tends to infinity with time measured in units of $2N$ generations. We denote the coalescence time by $\tau^{(N)}$, keeping the dependence on N explicit. This is the number of generations back it takes for the ancestral lines to coalesce. We use $\mathcal{A}$ to represent the population pedigree, leaving the dependence on N implicit to simplify the notation. We assume that $\mathcal{A}$ is the outcome of a random process of reproduction (as in section "Structured population model with gametic migration") and that it includes the information of which individual was sampled, so that the conditional coalescence time $\tau^{(N)}$ given $\mathcal{A}$ is well defined.

Following Diamantidis et al. (2024), we characterize the distribution of the coalescence time using the conditional survival function

$$F_N(t) = \mathbb{P}\left(\tau^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right) \tag{1}$$

for $t \geq 0$, where $\lfloor 2Nt \rfloor$ designates the largest integer less than or equal to $2Nt$. Thus, t is time back measured in units of $2N$ generations, and (1) is the probability that the two ancestral lines starting from the sampled individual coalesce more than $\lfloor 2Nt \rfloor$ generations ago, given the population pedigree. Note that $1 - F_N(t)$ is the cumulative distribution function (CDF) of the coalescence time.

Our goal is to describe the properties of $F_N(t)$ as $N \to \infty$. In particular, we wish to know whether it retains its dependence on the random population pedigree in the limit or it becomes a deterministic function only of t as in traditional coalescent models. Equation (1) can also be written as the conditional expectation of an indicator random variable given the population pedigree, namely,

$$F_N(t) = \mathbb{E}\left(\mathbf{1}_{\{\tau^{(N)} > \lfloor 2Nt \rfloor\}} \mid \mathcal{A}\right).$$

Then, by the tower property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}(F_N(t)) &= \mathbb{E}\left[\mathbb{E}\left(\mathbf{1}_{\{\tau^{(N)} > \lfloor 2Nt \rfloor\}} \mid \mathcal{A}\right)\right] \\ &= \mathbb{P}\left(\tau^{(N)} > \lfloor 2Nt \rfloor\right) \end{aligned} \tag{2}$$

for $t \geq 0$ as the mean of the survival function over all possible population pedigrees. The limit of (2) as $N \to \infty$ is how the coalescence time would be characterized under the traditional approach of coalescence theory. Now, if $\mathrm{Var}(F_N(t))$ tends to zero as $N \to \infty$, then $F_N(t)$ converges in mean square, and therefore in distribution, to its mean (2) for every fixed $t \geq 0$. This implies that the limiting conditional coalescent model must coincide with the limiting unconditional one. But if $\mathrm{Var}(F_N(t))$ does not tend to zero, then $F_N(t)$ remains a stochastic function, depending on the population pedigree, even as $N \to \infty$.

The variance of $F_N(t)$ can be computed by considering the coalescence times $\tau_1^{(N)}$ and $\tau_2^{(N)}$ of the ancestral lines that start from two pairs of genes at two unlinked loci in the same individual chosen at random, which are conditionally independent and identically distributed given the population pedigree. Again following Diamantidis et al. (2024), we have the identities

$$F_N(t) = \mathbb{P}\left(\tau_1^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right) = \mathbb{P}\left(\tau_2^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right)$$

and

$$\begin{aligned} &\mathbb{P}\left(\tau_1^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right)\mathbb{P}\left(\tau_2^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right) \\ &\quad = \mathbb{P}\left(\tau_1^{(N)} > \lfloor 2Nt \rfloor, \tau_2^{(N)} > \lfloor 2Nt \rfloor \mid \mathcal{A}\right), \end{aligned}$$

from which

$$\mathbb{E}\left(F_N(t)^2\right) = \mathbb{P}\left(\tau_1^{(N)} > \lfloor 2Nt \rfloor, \tau_2^{(N)} > \lfloor 2Nt \rfloor\right) \tag{3}$$

for $t \geq 0$. In (2) and (3), we have what we need to compute $\mathrm{Var}(F_N(t))$.

In section "Four kinds of conditional coalescent models," we apply this logic to four different limits of the population model presented in section "Structured population model with gametic migration," computing the mean and the variance of the survival function in each case. For each model, we ask whether the effect of the pedigree persists in the relevant limit. In fact, only one of these four models has $N \to \infty$. The other limits concern aspects of the migration process or the number of demes.

Note that (2) and (3) are averages over the population pedigree. Conveniently, there is no need to specify the whole pedigree or to compute the conditional survival function (1). For any population model which can be described as a Markov chain, (2) and (3) can be computed from a transition matrix for ancestral lines over one generation back in time. It is not even necessary to specify all aspects of the pedigree process (i.e. all details of reproductive outcomes among individuals) in the relevant Markov chain, insofar as these do not affect the averages (2) and (3).

To take advantage of these simplifications, we assume that reproduction occurs among monoecious individuals via random union of gametes from an infinite gamete pool which may include some proportion of migrant gametes. We do not keep track of individuals, only gametes because this is what is minimally necessary to account for joint coalescence and joint migration of ancestral lines at unlinked loci. Describing the model in this way minimizes the number of states in the Markov chains required to compute (2) and (3). Some key extensions (where we do need individuals) are treated in section "Robustness to diploid migration and dioecy" and Supplementary Sections 4 and 5. Extensions to samples larger than two are taken up in section "Discussion."

The method of assessing convergence using (2) and (3) was developed to analyze random walks in random environments (Bolthausen and Sznitman 2002; Birkner et al. 2013). The two copies of the process which are assumed to be conditionally independent in the general method here correspond exactly to the Mendelian transmission of two unlinked loci (having coalescence times $\tau_1^{(N)}$ and $\tau_2^{(N)}$). So, the computations involved are directly relevant to the genetic ancestries of unlinked loci. The statement that $\mathrm{Var}(F_N(t)) \to 0$ is the same as the statement that $\tau_1^{(N)}$ and $\tau_2^{(N)}$ retain their independence in the limit even when averaging over pedigrees as in (3). On the other hand, if $F_N(t)$ remains a stochastic function of the population pedigree, it means that $\tau_1^{(N)}$ and $\tau_2^{(N)}$ become nonindependent in the limit due to the random pedigree they share.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Structured population model with gametic migration

Our goal is to identify types of structured populations for which the limiting coalescent process depends on the population pedigree. For simplicity, we adopt a symmetric island model in which D demes each containing N diploid monoecious individuals are all equally connected by migration, and reproduction occurs via an effectively infinite gamete pool. Such a model would ordinarily include a single migration parameter, but here we will have two. The usual single parameter, which we also include, is denoted m. This is the proportion of each deme's gamete pool that comes from the other demes from one generation to the next (when migration occurs). The second parameter controls the timing of migration events across the entire population. Specifically, we define α as the probability that any migration can occur at all in a generation. With probability $1-\alpha$, no migration is possible, so each deme's gametes come from within that deme. With probability α, all demes' gamete pools include proportions $1-m$ and m of nonmigrant and migrant gametes, respectively. Migration being symmetric means that fractions $m/(D-1)$ come from each of the other $D-1$ demes.

We calculate the mean and the variance of the conditional distribution function of pairwise coalescence times. This entails describing backward-time Markov processes allowing (2) and (3) to be computed. Under our simple model, the average (2) can be computed by tracking just three gametic states in the ancestral process. Our starting sample includes the pair of genes at a single locus in an individual chosen at random in the current generation. These are the two genes in the two gametes which united to form the individual. In the model, they are no different than two genes in two gametes which went to different individuals in the same deme. Then it suffices to consider just three possible states of the ancestral lines of the two genes in the sample: [•] for one in one deme, i.e. the coalescent state, [••] for two in the same deme, and [•][•] for two in two different demes. Throughout, we use square brackets in this way to delineate demes. When necessary, we use parentheses to delineate individuals, so, for example, either [(••)] or [(•)(•)] for two genes in the same deme but in the same individual or in different individuals, respectively.

Under the assumption of discrete, nonoverlapping generation, the transition probability matrix for the ancestral process of a sample of two genes at a single locus is

$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2N}(1-\alpha\beta) & \left(1-\frac{1}{2N}\right)(1-\alpha\beta) & \alpha\beta \\ \frac{1}{2N}\frac{\alpha\beta}{D-1} & \left(1-\frac{1}{2N}\right)\frac{\alpha\beta}{D-1} & 1-\frac{\alpha\beta}{D-1} \end{pmatrix}, \tag{4}$$

in which

$$\beta = 2m(1-m) + m^2\frac{D-2}{D-1} = m(2-m) - \frac{m^2}{D-1}$$

represents the conditional probability for two gametes chosen at random in the same deme to come from two different demes given that migration occurs. Note that the conditional probability for two gametes chosen at random in different demes to come from the same deme is then given by $\beta/(D-1)$. In the next section, "Four kinds of conditional coalescent models," we use (4) to compute the mean of the conditional survival function (2) and describe how two-locus gametic states can be tracked in order to compute (3).

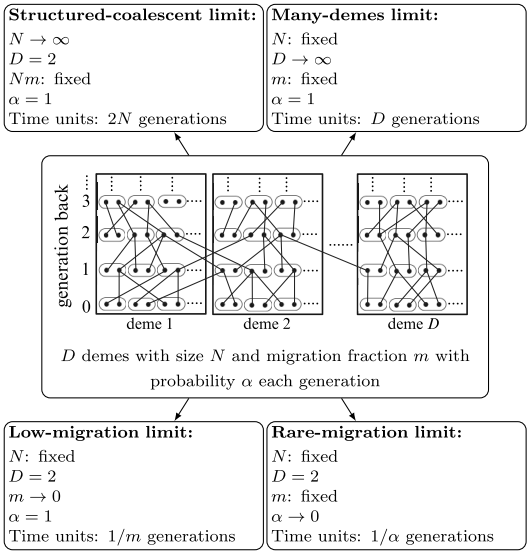


Fig. 1. Schematic of genealogies in a deme-structured system under four limiting regimes defined by the scaling of N, D, m, and α, corresponding, respectively, to the structured-coalescent, many-demes, low-migration, and rare-migration limits.

A reference guide to the notation for the main variables in this work is given in Supplementary Table S1.

Four kinds of conditional coalescent models

We consider a monoecious diploid population subdivided into D demes of size N and undergoing discrete, nonoverlapping generations. From one generation to the next, gametes are produced in equal large numbers by the individuals in all the demes. With probability α, a fraction $1-m$ of gametes stay in the deme where they were produced while a fraction m migrate uniformly to the other demes. Then, m is the fraction of gametes in each deme that come from the other demes and $m/(D-1)$ the fraction that come from each one. Otherwise, with probability $1-\alpha$, all gametes stay in the deme where they were produced. Random sampling of N pairs of gametes within each deme starts the next generation. We also consider dioecious populations and migration of (diploid) offspring, but we defer the details to Supplementary Material.

We are interested in the coalescence time of ancestral lines that start in the current generation from a pair of genes at a single locus in an individual chosen at random in the population or in two different individuals chosen at random in different demes. In Fig. 1, the central box shows a schematic of the monoecious diploid population with D demes each of size N and migration fraction m with per-generation probability of migration α where lines indicate within deme and across deme genealogies. We consider these coalescence times under four assumptions: (1) $N \to \infty$ with Nm fixed for $D = 2$ and $\alpha = 1$ (the structured-coalescent limit); (2) $D \to \infty$ with N and m fixed for $\alpha = 1$ (the many-demes limit); (3) $m \to 0$ with N fixed for $D = 2$ and $\alpha = 1$ (the low-migration limit); and (4) $\alpha \to 0$ with N and m fixed for $D = 2$ (the rare-migration limit). Setting $D = 2$ and $\alpha = 1$ except when necessary is for ease of

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

illustration. Arrows from the central box to the outer boxes in Fig. 1 indicate how the four assumptions above translate into the four different models in which we consider these coalescence times.

In case (1), the coalescence time, rescaled so that it is measured in units of $2N$ generations, is represented by $\tau^{(N)}$. We use $2N$ rather than N here, and similarly define $M = 4Nm$, for ease of comparison with the known results of this well-studied model (Wilkinson-Herbots 1998, Section 4). In cases (2), (3), and (4), the parameter N will be replaced by D, m, and α, respectively, and the unit of time $2N$ generations will be replaced by D, $1/m$, and $1/\alpha$ generations, respectively.

In each case, we will obtain the first two moments of the conditional survival function at rescaled time back $t \geq 0$, for example $\mathbb{E}(F_N(t))$ and $\mathbb{E}(F_N(t)^2)$ in case (1), by considering the transition probabilities from one generation to the previous one for possible states of the ancestral lines (in the same deme or in different demes, in the same individual or in different individuals). As described in section "The conditional survival function and its variance," the first moment corresponds to the probability that there are still two ancestral lines at time t starting from a single pair of genes at time 0. The second moment corresponds to the probability that there are still four ancestral lines at time t starting from two pairs of genes at two unlinked loci at time 0.

Structured-coalescent limit

In this subsection, we assume $D = 2$ demes of size N and migration fraction $m = M/(4N)$ with probability $\alpha = 1$ in each generation. Starting from a pair of genes at the same locus in an individual chosen at random in a deme chosen at random, the ancestral lines backward in time can be in three possible states ordered as follows: [•] for only one in one deme, [••] for two in the same deme, and [•][•] for two in two demes. Our assumptions of random union of gametes and gametic migration allow for this reduced state space, in that we do not need to specify whether two ancestral lines in the same deme, [••], are in the same individual or in different individuals.

The transition probabilities from one generation to the previous one are given by a transition matrix that can be written in the form

$$\mathbf{\Pi}_N = \mathbf{I} + \frac{\mathbf{G}_N}{2N},$$

where $\mathbf{I}$ is an identity matrix and

$$\mathbf{G}_N \to \mathbf{G} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -(1+M) & M \\ 0 & M & -M \end{pmatrix} \tag{5}$$

as $N \to \infty$. By the definition of the matrix exponential (see Möhle 1998a for a proof), we have

$$\mathbf{\Pi}_N^{\lfloor 2Nt \rfloor} \to \exp\{t\mathbf{G}\} \tag{6}$$

as $N \to \infty$, where now time is measured in units of $2N$ generations and rescaled time t corresponds to $\lfloor 2Nt \rfloor$ generations in the past. Diagonalizing the matrix $\mathbf{G}$ leads to

$$\exp\{t\mathbf{G}\} = \mathbf{V} \exp\{t\mathbf{D}\} \mathbf{V}^{-1},$$

where

$$\exp\{t\mathbf{D}\} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp\{t\lambda_1\} & 0 \\ 0 & 0 & \exp\{t\lambda_2\} \end{pmatrix}$$

and

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -(1+1/\lambda_1) & -(1+1/\lambda_2) \end{pmatrix}$$

with

$$\lambda_1 = \frac{-(1+2M) + \sqrt{1+4M^2}}{2},$$
$$\lambda_2 = \frac{-(1+2M) - \sqrt{1+4M^2}}{2}.$$

Using the convergence result in (6) above, the limit of

$$\mathbb{E}(F_N(t)) = (0,\ 1,\ 0)\mathbf{\Pi}_N^{\lfloor 2Nt \rfloor}(0,\ 1,\ 1)^T \tag{7}$$

as $N \to \infty$ is given by

$$\mathbb{P}(T > t) = (0,\ 1,\ 0)\exp\{t\mathbf{G}\}(0,\ 1,\ 1)^T. \tag{8}$$

Here the random variable T represents the time for a continuous-time Markov chain whose rate matrix is $\mathbf{G}$ to reach the coalescence state [•] starting from state [••] represented by (0, 1, 0) in (7) and (8). The vector $(0,\ 1,\ 1)^T = (0,\ 1,\ 0)^T + (0,\ 0,\ 1)^T$ in (7) and (8) specifies that the lineages have not yet coalesced at time back t because they are in state [••] or [•][•], as required for $T > t$. Note that the rate matrix $\mathbf{G}$ corresponds to the situation where two ancestral lines in the same deme coalesce at rate 1 while each one migrates from one deme to the other at rate $M/2$, all these events occurring independently of one another.

Next we consider the possible states for the ancestral lines of two pairs of genes at two unlinked loci, 1 and 2, in the same individual. They are ordered as follows, in which •, ○, and • represent gametes ancestral at locus 1 only, at locus 2 only, and at both loci, respectively, and brackets delineate demes:

1.	only one •, ○ or •	7.	[•• ○]
2.	[•• ○ ○]	8.	[••][○]
3.	[•• ○][○]	9.	[•][• ○]
4.	[• ○ ○][•]	10.	[• ○][•]
5.	[• ○][• ○]	11.	[••]
6.	[••][○ ○]	12.	[•][•]

Specifying that ancestral lines at two unlinked loci are in the same gamete (•) means that they travel together for a generation, either migrating or not, then separating into different gametes and going either to the same or to different individuals. In the terminology of Vitalis and Couvet (2001, p. 69), they are in the same haplotype. Because our focus is on the survival function, we lump all coalescent events into state 1. State 11 and state 12 are the initial states for the two conditionally independent copies of the process corresponding to samples [••] and [•][•], respectively, as needed in the computation of $\mathbb{E}(F_N(t)^2)$, because they specify identical initial states for the pairs at both loci, down to the gamete level.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

The transition matrix from one generation to the previous one takes the form

$$\tilde{\mathbf{\Pi}}_N = \tilde{\mathbf{A}} + \frac{\tilde{\mathbf{B}}_N}{2N},$$

where $\tilde{\mathbf{A}}$ is given by

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 1/4 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/4 & 0 \\
0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/4
\end{pmatrix}.$$

This is the transition probability matrix in the limiting case of demes of infinite size, which does not include any migration nor coalescence event. Transitions involving migration or coalescence are contained in the rate matrix $\tilde{\mathbf{B}}_N \to \tilde{\mathbf{B}}$ as the deme size $N \to \infty$ where $\tilde{\mathbf{B}}$ with respect to the subsets of states $S_1 = \{1, \ldots, 6\}$ and $S_2 = \{7, \ldots, 12\}$ is in the block form

$$\begin{pmatrix} \tilde{\mathbf{B}}_{11} & \tilde{\mathbf{B}}_{12} \\ \tilde{\mathbf{B}}_{21} & \tilde{\mathbf{B}}_{22} \end{pmatrix}$$

with $\tilde{\mathbf{B}}_{11}$ given by

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
2 & -4-2M & M & M & 0 & 0 \\
1 & M/2 & -3-2M & 0 & M & M/2 \\
1 & M/2 & 0 & -3-2M & M & M/2 \\
0 & 0 & M & M & -2-2M & 0 \\
2 & 0 & M & M & 0 & -2-2M
\end{pmatrix}$$

and

$$\tilde{\mathbf{B}}_{12} = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

Using this decomposition of $\tilde{\mathbf{\Pi}}_N$, we may apply the separation-of-time-scales result in Lemma 1 of Möhle (1998a). This guarantees that

$$\lim_{N\to\infty} \tilde{\mathbf{\Pi}}_N^{\lfloor 2Nt \rfloor} = \tilde{\mathbf{P}} \exp\{t\tilde{\mathbf{P}}\tilde{\mathbf{B}}\tilde{\mathbf{P}}\},$$

where $\tilde{\mathbf{P}} = \lim_{n\to\infty} \tilde{\mathbf{A}}^n$ is given by

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

The matrix $\tilde{\mathbf{P}}$ specifies the end result of the fast process in which genes at unlinked loci in the same gamete separate into different gametes, and the rate matrix $\tilde{\mathbf{P}}\tilde{\mathbf{B}}\tilde{\mathbf{P}}$ specifies the long-term processes of migration and coalescence.

This implies that, in the limit as $N \to \infty$, there are instantaneous transitions to states in the set $S_1 = \{1, 2, \ldots, 6\}$ if the initial state belongs to the set $S_2 = \{7, 8, \ldots, 12\}$ of all states with four ancestral lines, but where at least one pair of genes at different loci is in the same gamete. Moreover, once in the state space S_1, we have a continuous-time Markov chain whose rate matrix, denoted by $\tilde{\mathbf{G}}$, is given by the entries of $\tilde{\mathbf{P}}\tilde{\mathbf{B}}\tilde{\mathbf{P}}$ corresponding to the elements of S_1, that is,

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
2 & -2-2M & M & M & 0 & 0 \\
1 & M/2 & -1-2M & 0 & M & M/2 \\
1 & M/2 & 0 & -1-2M & M & M/2 \\
0 & 0 & M & M & -2M & 0 \\
2 & 0 & M & M & 0 & -2-2M
\end{pmatrix}. \quad (9)$$

This is the rate matrix for the ancestral lines of two pairs of genes at two loci, 1 and 2, if the ancestral lines at each locus coalesce or migrate as previously and independently of the ancestral lines at the other locus. Therefore, we have

$$\begin{aligned}&(0,\ 1,\ 0,\ 0,\ 0,\ 0)\exp\{t\tilde{\mathbf{G}}\}(0,\ 1,\ 1,\ 1,\ 1,\ 1)^{\mathsf{T}} \\ &\quad = \mathbb{P}(T_1 > t)\mathbb{P}(T_2 > t),\end{aligned}$$

where T_1 and T_2 represent two independent random variables having the same probability distribution as T previously introduced with $(0,\ 1,\ 0,\ 0,\ 0,\ 0)$ corresponding to state 2 in S_1, namely [•• ○ ○], and $(0,\ 1,\ 1,\ 1,\ 1,\ 1)^{\mathsf{T}}$ to the set of all states in S_1 with four ancestral lines still distinct (so both $T_1 > t$ and $T_2 > t$). But this is the limit of

$$\begin{aligned}&(0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0)\tilde{\mathbf{\Pi}}_N^{\lfloor 2Nt \rfloor} \\ &\cdot(0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1)^{\mathsf{T}}\end{aligned} \quad (10)$$

as $N \to \infty$, which is equal to $\mathbb{E}(F_N(t)^2)$, since in the limit of a large deme size, there is an instantaneous transition to state 2 in S_1 from state 11 in S_2, namely [••] corresponding to the vector on left in (10), while the vector on the right represents the set of all states in S_1 and S_2 with four ancestral lines still distinct. We conclude that

$$\lim_{N\to\infty} \mathbb{E}\left(F_N(t)^2\right) = (\mathbb{P}(T > t))^2 = \left(\lim_{N\to\infty} \mathbb{E}(F_N(t))\right)^2,$$

which implies that

$$\lim_{N\to\infty} \mathrm{Var}(F_N(t)) = 0 \quad (11)$$

for $t \geq 0$. The same conclusion can be drawn if the initial state for a pair of ancestral lines at a single locus is [•][•] instead of [••]. Conditioning on the pedigree and marginalizing over the pedigree give the same predictions about the distribution of pairwise coalescence times in the structured-coalescent limit.

Many-demes limit

Here we consider fixed N and m, with $D \to \infty$. We set $\alpha = 1$ for simplicity and for comparison with previous results from traditional

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

many-demes coalescent models (Lessard and Wakeley 2004; Wakeley 2004a). We denote the possible states for the pair of ancestral lines at a single locus as previously: [•] for only one in one deme, [••] for two in the same deme, and [•][•] for one in each of two different demes. Then, denoting (4) with $\alpha = 1$ as $\mathbf{\Pi}_D$, we have $\lim_{D\to\infty} \mathbf{\Pi}_D = \mathbf{A}$ where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{(1-m)^2}{2N} & (1-\frac{1}{2N})(1-m)^2 & m(2-m) \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\mathbf{P} = \lim_{k\to\infty} \mathbf{A}^k = \begin{pmatrix} 1 & 0 & 0 \\ \mathcal{F} & 0 & 1-\mathcal{F} \\ 0 & 0 & 1 \end{pmatrix},$$

in which

$$\mathcal{F} = \frac{(1-m)^2}{2Nm(2-m)+(1-m)^2}. \tag{12}$$

This coefficient is equivalent to one definition of Wright's F_{ST}, namely the probability that the first event in the ancestry of two genetic lineages in the same deme is a coalescent event rather than a migration event (Wright 1951; Slatkin 1991; Rousset 2002). Furthermore, the matrix $\mathbf{B} = \lim_{D\to\infty} (\mathbf{\Pi}_D - \mathbf{A})D$ is given by

$$\begin{pmatrix} 0 & 0 & 0 \\ \frac{m^2}{2N} & m^2\left(1-\frac{1}{2N}\right) & -m^2 \\ \frac{m(2-m)}{2N} & m(2-m)\left(1-\frac{1}{2N}\right) & -m(2-m) \end{pmatrix}$$

and

$$\mathbf{G} = \mathbf{PBP} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{(1-\mathcal{F})^2}{2N} & 0 & -\frac{(1-\mathcal{F})^2}{2N} \\ \frac{1-\mathcal{F}}{2N} & 0 & -\frac{1-\mathcal{F}}{2N} \end{pmatrix},$$

using the fact that

$$\frac{m(2-m)}{2N} + m(2-m)\left(1-\frac{1}{2N}\right)\mathcal{F} = \frac{1-\mathcal{F}}{2N}. \tag{13}$$

By Lemma 1 of Möhle (1998a), the limit $\lim_{D\to\infty} \mathbf{\Pi}_D^{\lfloor Dt \rfloor}$ is given by $\mathbf{P}\exp\{t\mathbf{G}\}$ which can be expressed as

$$\begin{pmatrix} 1 & 0 & 0 \\ 1-(1-\mathcal{F})\exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} & 0 & (1-\mathcal{F})\exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} \\ 1-\exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} & 0 & \exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} \end{pmatrix}.$$

Therefore, the conditional survival function defined here as $F_D(t) = \mathbb{P}(\tau^{(D)} > \lfloor Dt \rfloor \mid \mathcal{A})$ satisfies

$$\lim_{D\to\infty} \mathbb{E}(F_D(t)) = \exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} \tag{14}$$

if the initial state is [•][•], but

$$\lim_{D\to\infty} \mathbb{E}(F_D(t)) = (1-\mathcal{F})\exp\left\{-\frac{1-\mathcal{F}}{2N}t\right\} \tag{15}$$

if the initial state is [••], for $t \geq 0$. Thus, comparing (15) to (14), a sample in initial state [••] has probability $\mathcal{F}$ of instantaneous coalescence, whereas with probability $1-\mathcal{F}$ it follows the same coalescent process as a sample in initial state [•][•].

For two pairs of ancestral lines at two unlinked loci, we order the possible states as follows:

1. only one •, ○ or •
2. [•][○][•][○]
3. [••][○][○]
4. [•][•][○○]
5. [•][•○][○]
6. [•○][•○]
7. [•○][•]
8. [•][•][○]
9. [•][•]
10. [••○][○]
11. [•○○][•]
12. [••][○○]
13. [••○○]
14. [••][○]
15. [•][•○]
16. [••○]
17. [••]

State 17 and state 9 are the initial states for the two conditionally independent copies of the process for the samples [••] and [•][•], respectively, required for the computation of $\mathbb{E}(F_D(t)^2)$.

The transition matrix from one generation to the previous one can be expressed as

$$\tilde{\mathbf{\Pi}}_D = \tilde{\mathbf{A}} + \frac{\tilde{\mathbf{B}}_D}{D}, \tag{16}$$

where $\tilde{\mathbf{A}}$ is the transition matrix in the limit of an infinite number of demes. In this limit, all states are transient but states 1 and 2, which are absorbing. Moreover, with respect to the subsets of states $\tilde{S}_1 = \{1, \ldots, 9\}$ and $\tilde{S}_2 = \{10, \ldots, 17\}$, we have

$$\tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \mathbf{0} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{pmatrix}, \tag{17}$$

where $\tilde{\mathbf{A}}_{11}$ is given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_1 & p_2 & p_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_1 & p_2 & 0 & p_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 & p_3 & 0 & 0 & p_1 & 0 \\ 0 & p_2^2 & 0 & 0 & 2p_2p_3 & p_3^2 & 2p_1p_3 & 2p_1p_2 & p_1^2 \\ 0 & 0 & 0 & 0 & p_2/2 & p_3/2 & (p_1+p_3)/2 & p_2/2 & p_1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/4 & 1/2 & 0 & 1/4 \end{pmatrix}$$

with

$$p_1 = \frac{(1-m)^2}{2N}, \quad p_2 = m(2-m), \quad p_3 = 1-p_1-p_2. \tag{18}$$

By the ergodic theorem for discrete-time Markov chains on a finite state space (see, e.g. Karlin and Taylor 1975), we have

$$\tilde{\mathbf{P}} = \lim_{k\to\infty} \tilde{\mathbf{A}}^k = \begin{pmatrix} \tilde{\mathbf{P}}_{11} & \mathbf{0} \\ \tilde{\mathbf{P}}_{21} & \mathbf{0} \end{pmatrix},$$

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

where $\tilde{\mathbf{P}}_{11} = \lim_{k\to\infty} (\tilde{\mathbf{A}}_{11})^k$ takes the form

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{F} & 1-\mathcal{F} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{F} & 1-\mathcal{F} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{19}$$

and

$$\tilde{\mathbf{P}}_{21} = \begin{pmatrix} u_{10} & 1-u_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{11} & 1-u_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{12} & 1-u_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{13} & 1-u_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{14} & 1-u_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{15} & 1-u_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{16} & 1-u_{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ u_{17} & 1-u_{17} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{20}$$

for some absorption probabilities $u_{10}, \ldots, u_{17}$. From the set of equations (S2) through (S9) in Supplementary Subsection 1.1, we find that $u_{10} = u_{11} = u_{14} = u_{15} = \mathcal{F}$ and $u_{12} = 1-(1-\mathcal{F})^2$. These may be understood intuitively by noting that in the process specified by $\tilde{\mathbf{A}}$: the outcomes for lineages in different demes are independent; two • lineages or two ∘ lineages in the same deme will eventually coalesce with probability $\mathcal{F}$ or disperse to different demes with probability $1-\mathcal{F}$; but two lineages of different types in the same deme must eventually disperse to different demes. The lengthy expressions for u_{13}, u_{16}, and u_{17}, which satisfy equations (S7), (S8), and (S9), are shown in Supplementary Subsection 1.1.

On the other hand, the first two rows of

$$\tilde{\mathbf{B}} = \lim_{D\to\infty} \tilde{\mathbf{B}}_D = \begin{pmatrix} \tilde{\mathbf{B}}_{11} & \tilde{\mathbf{B}}_{12} \\ \tilde{\mathbf{B}}_{21} & \tilde{\mathbf{B}}_{22} \end{pmatrix}$$

correspond to state 1 with only one ancestral line at any one of the loci and state 2 with four ancestral lines, two at each locus, in all different demes. These are given by the first two rows of $\tilde{\mathbf{B}}_{11}$, which are

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{p_2}{N} & -6p_2 & p_2-\frac{p_2}{2N} & p_2-\frac{p_2}{2N} & 4p_2-\frac{2p_2}{N} & 0 & 0 & \frac{2p_2}{N} & 0 \end{pmatrix},$$

and the first two rows of $\tilde{\mathbf{B}}_{12}$, which exhibits 0 everywhere. Lemma 1 of Möhle (1998a) ensures that $\lim_{D\to\infty} \tilde{\mathbf{\Pi}}_D^{\lfloor Dt \rfloor}$ is given by $\tilde{\mathbf{P}} \exp\{t\tilde{\mathbf{G}}\}$, where

$$\tilde{\mathbf{G}} = \tilde{\mathbf{P}}\tilde{\mathbf{B}}\tilde{\mathbf{P}} = \begin{pmatrix} \tilde{\mathbf{P}}_{11}(\tilde{\mathbf{B}}_{11}\tilde{\mathbf{P}}_{11} + \tilde{\mathbf{B}}_{12}\tilde{\mathbf{P}}_{21}) & \mathbf{0} \\ \tilde{\mathbf{P}}_{21}(\tilde{\mathbf{B}}_{11}\tilde{\mathbf{P}}_{11} + \tilde{\mathbf{B}}_{22}\tilde{\mathbf{P}}_{21}) & \mathbf{0} \end{pmatrix}.$$

The first two rows of $\tilde{\mathbf{P}}_{11}(\tilde{\mathbf{B}}_{11}\tilde{\mathbf{P}}_{11} + \tilde{\mathbf{B}}_{12}\tilde{\mathbf{P}}_{21})$, which correspond to the first two rows of $\tilde{\mathbf{B}}_{11}\tilde{\mathbf{P}}_{11}$, are

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda & -\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{21}$$

where

$$\lambda = \frac{p_2}{N} + 2p_2\left(1-\frac{1}{2N}\right)\mathcal{F} = \frac{1-\mathcal{F}}{N} \tag{22}$$

owing to (13) and (18). The last row of $\tilde{\mathbf{P}}_{11}$ in (19) means that, in the limiting process with initial state 9 at two unlinked loci, or [•][•], so that both pairs of lines start from the same gametes in two different demes, there is an instantaneous transition to the dispersed state 2, or [•][∘][•][∘], with all four lines in different demes. Besides, the second row of $\tilde{\mathbf{G}}$ given in (21) means that this is followed by a transition to the fixation state 1 at rate $\lambda = (1-\mathcal{F})/N$. This implies that

$$\lim_{D\to\infty} \mathbb{E}\left(F_D(t)^2\right) = \exp\left\{-\frac{1-\mathcal{F}}{N}t\right\},$$

and then, owing to (14), that $\lim_{D\to\infty} \mathrm{Var}(F_D(t)) = 0$, for $t \geq 0$, if the initial state for a pair of lines at a single locus is [•][•].

With initial state 17 at two unlinked loci, or [••], such that both pairs of lines start from the same gametes in the same deme, there will be an instantaneous transition to either the coalescent state 1 with probability u_{17} or to the dispersed state 2 with the complementary probability $1-u_{17}$ owing to the last row of $\tilde{\mathbf{P}}_{21}$ in (20), from which the transition to the absorbing state 1 will occur at rate $(1-\mathcal{F})/N$ as previously. With the expression of u_{17} given in the Supplementary Subsection 1.1 denoted $\mathcal{U}$ for convenience, we conclude that

$$\lim_{D\to\infty} \mathbb{E}\left(F_D(t)^2\right) = (1-\mathcal{U})\exp\left\{-\frac{1-\mathcal{F}}{N}t\right\},$$

and then, owing to (15), that

$$\begin{aligned} &\lim_{D\to\infty} \mathrm{Var}\left(F_D(t)\right) \\ &\quad = \left(1-\mathcal{U}-(1-\mathcal{F})^2\right)\exp\left\{-\frac{1-\mathcal{F}}{N}t\right\} \end{aligned} \tag{23}$$

for $t \geq 0$, if the initial state for a pair of lines at a single locus is [••]. The quantity $1-\mathcal{U}-(1-\mathcal{F})^2$ depends on N and m and is complicated, but it is clearly not equal to zero (see Supplementary Subsections 1.1 and 1.3). This is the case because events of migration before coalescence for pairs of lines at unlinked loci within the same deme in the limit of a large number of small demes do not occur independently from one another. In this way, the many-demes coalescent process conditional on the population pedigree differs from the many-demes coalescent process marginal to the population pedigree.

So far, in this section, we have assumed that N and m are fixed as $D \to \infty$. Our result (23) contrasts with the result (11) in the previous section "Structured-coalescent limit" where the effect of the population pedigree became negligible. However, that was in a different limit: $N \to \infty$ with $m = M/(4N)$. Previously, traditional many-demes coalescent models which implicitly average over the pedigree have often included $M = 4Nm$ as a parameter, with comparatively little attention paid to finite N and fixed m (Lessard and Wakeley 2004; Wakeley 2004b).

Thus here we also ask whether the effect of the population pedigree in the many-demes limit becomes negligible if we further take $m = M/(4N)$ and N large enough. As shown in Supplementary Subsection 1.2 where we express $\tilde{\mathbf{A}}$ in (16) as $\tilde{\mathbf{A}} = \tilde{\mathbf{A}} + \tilde{\mathbf{B}}_N/(2N)$ and then apply Lemma 1 of Möhle (1998a) as $N \to \infty$, we have

$$\lim_{N\to\infty} \mathcal{U} = 1-\left(\frac{M}{M+1}\right)^2 \quad \text{and} \quad \lim_{N\to\infty} \mathcal{F} = \frac{1}{M+1}.$$

In other words, with $M = 4Nm$ kept constant as $N \to \infty$, the term $1-\mathcal{U}-(1-\mathcal{F})^2$ in (23) tends to zero and the pedigree effect in the many-demes limit becomes negligible.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Low-migration limit

Here we consider $D = 2$ demes of fixed size N in the limit $m \to 0$, with time rescaled accordingly and still with $\alpha = 1$. The analysis follows the pattern of the previous two sections, "Structured-coalescent limit" and "Many-demes limit," but we defer the details to Supplementary Section 2. Briefly, we define two state spaces and associated processes: for a single pair of ancestral lines (at one locus) and for two conditionally independent pairs of ancestral lines (at two unlinked loci). We write the transition matrix for the single-pair process as $\mathbf{\Pi}_m = \mathbf{A} + m\mathbf{B}_m$, and the transition matrix for the two-pair process as $\tilde{\mathbf{\Pi}}_m = \tilde{\mathbf{A}} + m\tilde{\mathbf{B}}_m$, then apply Lemma 1 of Möhle (1998a) with time measured in units of $1/m$ generations.

Two ancestral lines in initial state [••] will coalesce immediately regardless of the pedigree, because opportunities for migration are vanishingly rare in this limit. However, computing the first two moments of the conditional survival function given an initial state [•][•], represented here by $F_m(t) = \mathbb{P}(\tau^{(m)} > \lfloor t/m \rfloor \mid \mathcal{A})$ and in the limit as $m \to 0$, gives

$$\lim_{m\to 0} \mathrm{Var}(F_m(t)) = \exp\{-4t\}\left(\exp\left\{\frac{2}{N+1}t\right\} - 1\right) \tag{24}$$

for $t \geq 0$. Thus in the low-migration limit, the coalescent process conditional on the population pedigree differs from the coalescent process marginal to the population pedigree.

Note that this limiting variance (24) vanishes as $N \to \infty$. In "Structured-coalescent limit," which had $N \to \infty$ with $m \to 0$ proportionally, the conditional and marginal coalescent processes were the same. In addition, in "Many-demes limit," enforcing this same structured-coalescent limit on the end result (23) made the effect of the pedigree disappear. Similarly, if we apply the additional assumption $N \to \infty$ to (24) which already has $m \to 0$, the pedigree effect becomes negligible.

Rare-migration limit

Here, we consider $D = 2$ for fixed N and $m \in (0, 1)$ as $\alpha \to 0$. We defer the detailed analysis to Supplementary Section 3. For this model, we write the transition matrix for the single-pair process as $\mathbf{\Pi}_\alpha = \mathbf{A} + \alpha\mathbf{B}_\alpha$, and the transition matrix for the two-pair process as $\tilde{\mathbf{\Pi}}_\alpha = \tilde{\mathbf{A}} + \alpha\tilde{\mathbf{B}}_\alpha$, then apply Lemma 1 of Möhle (1998a) with time measured in units of $1/\alpha$ generations.

Similar to the case of low migration, here two ancestral lines in initial state [••] will coalesce immediately regardless of the pedigree, because opportunities for migration are so infrequent in this limit. Under this rare-migration model, the conditional survival function of the pairwise coalescence time given an initial state [•][•] is $F_\alpha(t) = \mathbb{P}(\tau^{(\alpha)} > \lfloor t/\alpha \rfloor \mid \mathcal{A})$ and our analysis gives

$$\left(\exp\left\{\left[\frac{2m(1-m)}{N+1} + \frac{4N^2m^2(1-m)^2}{(N+1)^2}\right]t\right\} - 1\right) \cdot \exp\{-4m(1-m)t\}$$

for $\lim_{\alpha\to 0} \mathrm{Var}(F_\alpha(t))$ for $t \geq 0$. Thus in the rare-migration limit as well, the coalescent process conditional on the population pedigree differs from the coalescent process marginal to the population pedigree.

In contrast to the low-migration regime of the previous subsection, if we consider the large population limit we obtain

$$\lim_{N\to\infty}\left[\lim_{\alpha\to 0}\mathrm{Var}(F_\alpha(t))\right] = \left(\exp\left\{4m^2(1-m)^2t\right\} - 1\right)\exp\{-4m(1-m)t\} \neq 0.$$

Table 1. Limiting variances of the conditional survival functions for the four limit models and two initial states of the sample.

Limit model	Initial state	Monoecy gametic	Monoecy diploid	Dioecy gametic	Dioecy diploid
Struct.	[••]	0	0	0	0
	[•][•]	0	0	0	0
Many	[••]	$V_1 \to 0$	$V_2 \to 0$	$V_3 \to 0$	$V_4 \to 0$
	[•][•]	0	0	0	0
Low	[••]	0	0	0	0
	[•][•]	$V_5 \to 0$	$V_6 \to 0$	$V_7 \to 0$	$V_8 \to 0$
Rare	[••]	0	0	0	0
	[•][•]	V_9	V_{10}	V_{11}	V_{12}

Of the 12 nonzero values, V_1 to V_{12}, demonstrating a pedigree effect for a particular case, eight tend to zero as $N \to \infty$ (indicated in parentheses). For the cases involving diploid migration rather than gametic migration, [••] and [•][•] above are shorthand for [(••)] and [(•)][(•)]. The full expressions for V_1 to V_{12} are given in Supplementary Table S2.

In this fourth case, the rare-migration limit, there is a persistent pedigree effect for initial state [•][•] even when we apply the additional assumption $N \to \infty$.

Robustness to diploid migration and dioecy

The four limit results in the previous subsections, "Structured-coalescent limit," "Many-demes limit," "Low-migration limit," and "Rare-migration limit," were obtained under the simplifying assumptions of monoecy and gametic migration. In Supplementary Section 4, we apply the same 4-fold analysis to a model of diploid migration. In Supplementary Section 5, we do the same for a model of dioecy with either gametic migration or diploid migration.

These extensions introduce additional states and transitions. Under diploid migration, for example, the two ancestral lines at a single locus can be in any of four possible states, not just the three considered so far: [•], [••], and [•][•]. Using parentheses to delineate individuals, we instead have [(•)] for one ancestral line in one individual in one deme, [(•)(•)] for two ancestral lines in two individuals in the same deme, [(•)][(•)] for two ancestral lines in two individuals in different demes, and [(••)] for two ancestral lines in the same individual in one deme. Furthermore, the two ancestral lines in [(••)] must travel together. They cannot be split apart by migration.

Despite such important additional details, we show that our characterizations of the effects of the pedigree on pairwise coalescence times for each of the four possible limits also hold for diploid migration and dioecy. There are differences in how the various constants and rates which arise in the each of four limits depend on N and m. Interested readers are referred to Supplementary Sections 4 and 5. Overall, if the pedigree does (respectively, does not) retain its effect on the pairwise coalescence time in the limit under monoecy and gametic migration, then it also does (respectively, does not) under dioecy and diploid migration. This is summarized in Table 1, which also notes the cases in which the pedigree effect becomes negligible for larger deme sizes. Supplementary Table S2 displays the full expressions for the limiting variances of the conditional survival functions, denoted simply V_1 through V_{12} in Table 1.

We anticipate that this robustness to different modes of reproduction and migration may hold more broadly, but we emphasize that all of the models we have considered are exceedingly simple. For example, in our model of monoecy and gametic migration, all gametes are of the same type. In contrast, monoecious plants

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

have "male" and "female" gametes with different migration patterns due to different dispersal mechanisms of pollen versus seeds (Levin and Kerster 1974; Heuertz et al. 2003; Sork et al. 2015). In our model of dioecy, there are two mating types but what might make them different is not specified. In fact, no differences between them are posited in the model. McLaughlin et al. (2023) review recent work on sex differences and mating strategies. Miyagi et al. (2025) point out that the interpretation of patterns of genetic variation under sex-biased migration or admixture depends on assumptions about intrasexual variation.

Simulation results and additional analysis

In this section, we illustrate variation among pedigrees using simulations of the prelimiting model (in section "Structured population model with gametic migration") for the case of two demes, and present additional analysis of the many-demes model with large deme size. Simulations were implemented in the Julia programming language. Additional analysis was done with the aid of Mathematica (Wolfram Research, Inc 2024). The associated code is available at https://doi.org/10.6084/m9.figshare.32144278.

Figures 2 and 3 display CDFs of coalescence times for the two initial states [••] and [•][•] on each of 1000 independently generated pedigrees under the prelimiting model with $D = 2$ demes and allowing migration in every generation ($\alpha = 1$). The CDF for a given sample on a given pedigree was estimated by simulating 10,000 independent coalescence times, which are conditionally independent given the pedigree as would be true for 10,000 unlinked loci in the sampled individual(s). The values of N and m were chosen so that $M = 4Nm = 0.5$ in order to illustrate convergence to the structured-coalescent limit. For the within-deme sample [••], Fig. 2 shows that as N increases from 24 to 124 to 640, the variation in CDFs among pedigrees decreases, converging on the marginal limiting prediction from (8). Figure 3 concerns the case $N = 124$ by contrasting the within-deme sample [••] results from Fig. 2 with the between-deme sample [•][•] results, showing that both vary around their respective limiting prediction from (8).

Figures 4 and 5 show the variability of CDFs of coalescence times for initial state [•][•] for two models with $D = 2$ and $N = 24$, but migration regimes representing the low-migration limit (Fig. 4) and the rare-migration limit (Fig. 5). In our prelimiting model, αm is the average per-generation probability that a single genetic line migrates. This is held constant at $\alpha m = 0.0003$ in both panels, but it is realized in two different ways: $\alpha = 1$ and $m = 0.0003$ in the low-migration regime versus $\alpha = 0.001$ and $m = 0.3$ in the rare-migration regime. As in Figs. 2 and 3, CDFs were estimated by simulating 10,000 conditionally independent coalescence times given each pedigree. Consistent with the variance among pedigrees tending to zero as $N \to \infty$ in the low-migration limit but not the rare-migration limit, the CDFs in Fig. 5 are more variable than those in Fig. 4. The CDFs in Fig. 5 show dramatic jumps when migration events (of size $m = 0.3$, not shown) appear in the pedigree in specific past generations. For both regimes, the empirical averages of the 10,000 CDFs are close to their limiting marginal predictions given in Supplementary Sections 2, 3, equations (S15) and (S19).

To illustrate the behavior of the many-demes limit, we do not display simulated CDFs of coalescence times but rather investigate the quantity $1 - \mathcal{U} - (1 - \mathcal{F})^2$ in the expression (23) for the limiting variance under this model. As previously noted, this quantity and consequently the limiting variance tends to zero when we add the structured-coalescent assumption, $N \to \infty$ for fixed $M = 4Nm$, to the many-demes limit. Here we consider the rate of approach of

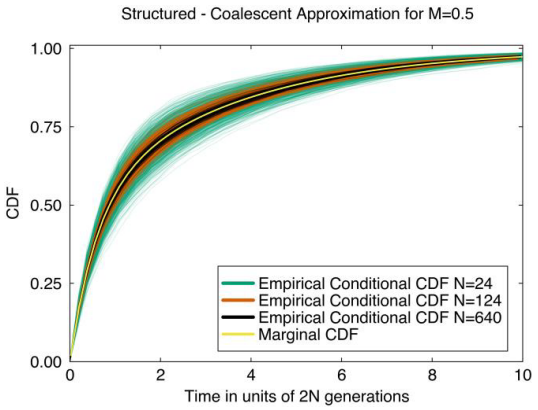


Fig. 2. Simulated CDFs of coalescence times for initial state [••] conditional on the pedigree, representing the structured-coalescent regime with N = 24, 124, and 640, together with the marginal CDF from (8).

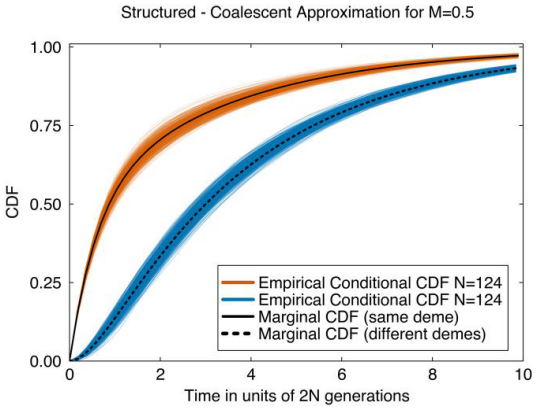


Fig. 3. Simulated CDFs of coalescence times for initial state [••] (same deme) and initial state [•][•] (different demes), for intermediate $N = 124$, together with their respective marginal CDFs.

$1 - \mathcal{U} - (1 - \mathcal{F})^2$ to zero when N is large. To leading order, this rate is equal to $g(M)/N$ with

$$g(M) = \frac{M(M^2 + 4M + 6)}{3(1 + M)^3}$$

from the limit given in Supplementary Subsection 1.4. The function $g(M)$ characterizes how quickly pedigree-induced fluctuations vanish as deme size increases, when $M = 4Nm$ is held constant.

Figure 6 plots $g(M)$ over a broad range of M. Fluctuations vanish fastest where $g(M)$ has its maximum value, near $M = 1$. This is in line with intuition that the structured-coalescent approximation will be most appropriate when m and $1/N$ are not too different in magnitude. For M close to zero ($m \ll 1/N$), much larger values of N are required to make the pedigree effect disappear. In this case, the low-migration approximation may be a better choice than the structured-coalescent approximation. At the other extreme, for large M, $g(M)$ approaches 1/3. Here, all effects of

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

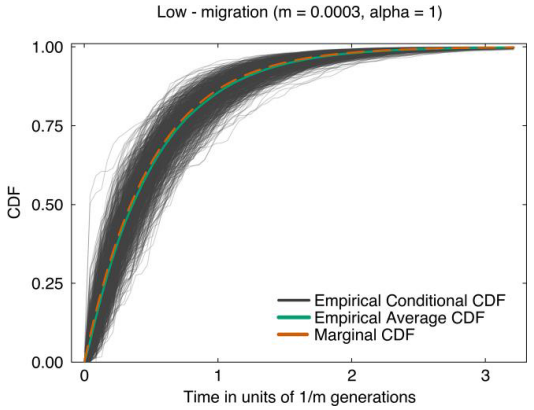


Fig. 4. Simulated CDFs of coalescence times for 1000 random pedigrees and sample state [•][•] under the low-migration regime with $N = 24$, and comparison to the marginal CDF $1 - \exp(-2t)$.

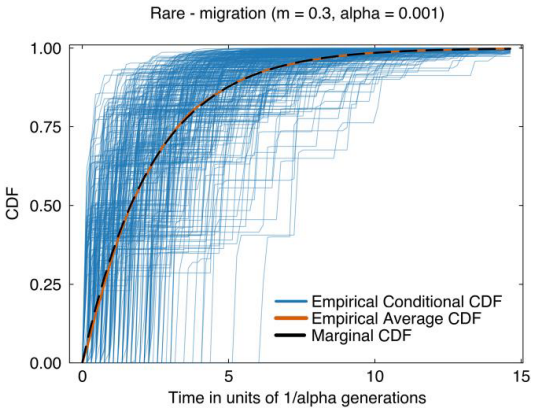


Fig. 5. Simulated CDFs of coalescence times for 1000 random pedigrees and sample state [•][•] under the rare-migration regime with $N = 24$, and comparison to the marginal CDF $1 - \exp(-2m(1-m)t)$.

migration-mediated population structure become negligible, and we may anticipate connections to related results for well-mixed populations. The quantity $1 - \mathcal{U} - (1 - \mathcal{F})^2$ is the excess probability that neither of the two pairs of sampled gene copies at two unlinked loci coalesce in the recent ancestry described by $\tilde{\mathbf{A}}$ in (16), the excess being due to the shared pedigree. This same value of 1/3 was obtained as the leading-order coefficient of decay for the correlation of coalescence times for unlinked loci in a monoecious randomly mating population (Kogan et al. 2026, p. 50) as well as for the squared coefficient of variation of single-locus heterozygosity, also in a monoecious randomly mating population, which Weir et al. (1980) attributed to the pedigree.

Discussion

We have classified structured-coalescent processes with migration in terms of how the shared population pedigree affects the distribution of pairwise coalescence times. We began with an idealized monoecious model of population subdivision and gametic migration, then extended our findings to diploid migration and dioecy. We quantified the effect of the pedigree in four different limits. Specifically, we asked whether the predicted distribution of pairwise coalescence times conditional on the pedigree differs from the one obtained by marginalizing or averaging over the pedigree.

Our idealized model has just four parameters: the number of demes D, the deme size N, the migration fraction m, and the migration probability α. The last of these, α, is the chance that any migration can occur at all, which, if it does, has migration fraction m. We did not find a pedigree effect in the structured-coalescent limit ($N \to \infty$ with m proportional to $1/N$ for $D = 2$ and $\alpha = 1$). But we did find a pedigree effect in the many-demes limit ($D \to \infty$ for fixed N, m and $\alpha = 1$), the low-migration limit ($m \to 0$ for fixed N, $D = 2$ and $\alpha = 1$), and rare-migration limit ($\alpha \to 0$ for fixed N, m and $D = 2$). These pedigree effects appear either for the within-deme sample [••] or the between-deme sample [•][•] depending on the particular limit as shown in Table 1, and they are similar in all the models. We also found that the pedigree effects in the many-demes limit and the low-migration limit are weak, in the sense that they become negligible as the deme size N increases. In the limits with a fixed number of demes, we have considered $D = 2$ for simplicity but the same conclusions are expected to hold for any $D \geq 2$.

Regarding samples larger than two, we argue as follows. In cases where samples of size two show a pedigree effect, we may be confident larger samples will as well, because the properties of samples of size two are marginal properties of larger samples. It would be odd for them to disagree in this way. Here we may appeal, at least heuristically, to the fundamental consistency property of subsamples under the Kingman coalescent (Berestycki 2009; Crane 2016). Indeed, Newman (2026) found a pedigree effect for larger samples in a generalized version of the rare-migration limit. In cases where samples of size two do not show a pedigree effect, additional work is needed to confirm that the full limiting conditional coalescent process for larger samples is identical to the traditional pedigree-averaged one. For a rigorous treatment of this in the structured-coalescent limit with an arbitrary number of demes, see Newman (2026). For the many-demes limit and the low-migration limit with the additional assumption $N \to \infty$, we leave this as a conjecture, with some confidence

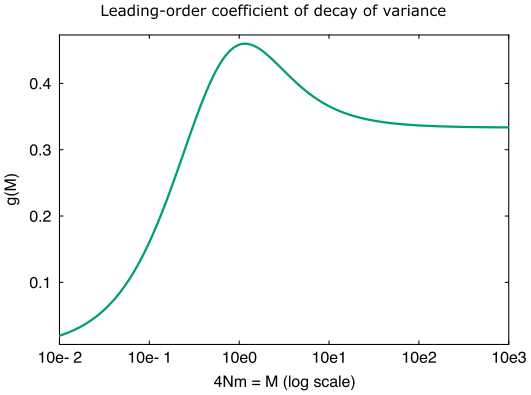


Fig. 6. The rate of decay of the pedigree effect for large N and fixed $M = 4Nm$ in the many-demes limit.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

from the results of Newman (2026) and analogous results for well-mixed populations (Tyukin 2015; Alberti et al. 2025).

A finding of little or no pedigree effect on coalescence implies that inference methods developed for traditional coalescent models, which have all been derived by implicitly averaging over the pedigree, may be reinterpreted as properly accounting for the pedigree. This is useful because all loci in the genome share the same pedigree and traditional coalescent models are typically applied to data from multiple loci in a sample of individuals. A finding of an appreciable pedigree effect implies that new coalescent models and associated inference methods must be developed. Diamantidis et al. (2024) and Alberti et al. (2025) discuss some of the challenges involved.

Our results can be understood with reference to previous findings for populations without subdivision and migration, as well as to the general phenomenon of separation of time scales. In well-mixed outbred populations with biparental reproduction, the ancestral genetic lineages of any sample disperse to different individuals in $O(1)$ generations (Möhle 1998b; Birkner et al. 2013, 2018) and the pedigree ancestries of all individuals overlap completely in $O(\log_2 N)$ generations (Chang 1999). In contrast, the coalescent process takes $O(N)$ generations.

Möhle (1998a, 1998b) introduced the formal separation-of-time-scales approach to coalescent theory. The complementary formalism for the forward-time diffusion models of population genetics was introduced by Ethier and Nagylaki (1980). Möhle (1998a) treated self-fertilization, in which there is a non-negligible probability that two genetic lineages in a single individual coalesce due to self-fertilization in the recent ancestry of the individual. This is a property of sampled individuals, true even in the coalescent limit $N \to \infty$, so self-fertilization is a fast-time-scale phenomenon. Möhle (1998b) treated dioecy, in which case the probability that two genetic lineages in a single individual coalesce in one generation is equal to zero, because the individual necessarily has two parents. This is also evident in the transition matrices (S28) and (S29) in Supplementary Section 5 for dioecious populations under either gametic migration or diploid migration. Biparental reproduction with two mating types is another fast-time-scale phenomenon.

In the case of self-fertilization, in a single population of size N, the coalescent process marginal to the population pedigree differs from the Kingman coalescent process (Nordborg and Donnelly 1997; Möhle 1998a) and the coalescent process conditional on the population pedigree differs from the marginal one (Newman et al. 2025). Self-fertilization is an extreme form of inbreeding, which is known more generally to produce non-Kingman marginal coalescent processes (Severson et al. 2019, 2021) and may similarly always show pedigree effects in conditional coalescent models (Newman et al. 2025). None of the models we have considered here have any form of inbreeding, beyond the background chance $1/N$ of self-fertilization within demes in the monoecious case. So we do not consider inbreeding as an explanation of the pedigree effects we observe.

At the same time, our finding of a pedigree effect in the many-demes limit for a sample initially in state [••] but none for a sample initially in state [•][•] appears quite similar to the corresponding result for partial self-fertilization (Newman et al. 2025), only with demes of size N playing the role of diploid individuals ($N = 2$, in effect). Our $\mathcal{F}$ in (12) is identical to F_{ST} in the island model of population structure (Wright 1951). It quantifies the increase in relatedness due to subdivision, relative to that expected in a random-mating population of the same total size, and is generally considered a type of inbreeding coefficient. However, it is Wright's F_{IS} that measures inbreeding owing to nonrandom-mating strategies, such as partial self-fertilization, within local populations or demes. Our $\mathcal{F}$ or F_{ST} can be made negligible simply by taking $N \to \infty$, whereas F_{IS} is insensitive to population size. In all of our models, F_{IS} is equal to zero.

We consider finite population size or, more particularly, finite deme size as the explanation of most of the pedigree effects on coalescence we observe. Such effects are expected in small populations even if the conditional and marginal coalescent processes are identical in the limit $N \to \infty$ (Wakeley et al. 2012; Tyukin 2015). An explanation in terms of finite N accords with our finding that the effects of the pedigree become negligible when we add the structured-coalescent assumption $N \to \infty$ for fixed $M = 4Nm$ to the many-demes model and when we add the assumption $N \to \infty$ to the low-migration model which already has $m \to 0$.

With four parameters, D, N, m, and α, there are a large number of possible limits we might have examined. The four we did examine were chosen because they are represented in the literature and they predict nontrivial levels of population structure, indicated by $F_{ST} > 0$. A fifth limit, which also appears in the literature but predicts $F_{ST} = 0$, is the strong-migration limit (Nagylaki 1980; Notohara 1993) here expressed most simply as $N \to \infty$ with $m \in (0, 1)$, $\alpha = 1$, and arbitrary D. In the strong-migration limit, the movement of genetic lineages among demes is so frequent relative to their coalescence that the distribution of lineages among demes reaches stationarity, and the coalescent process reduces to a Kingman coalescent process with a rescaled rate or effective population size (Notohara 1993).

With this in mind, we can check whether taking $N \to \infty$ for $m \in (0, 1)$ in the many-demes limit also makes the pedigree effect disappear for a sample initially in state [••]. From (23), this pedigree effect is revealed by the nonzero coefficient $1 - \mathcal{U} - (1 - \mathcal{F})^2$ in the limiting variance of the conditional survival function, with $\mathcal{F}$ as in (12) and $\mathcal{U}$ obtained by solving the equations for u_{13}, u_{16}, u_{17} and setting $\mathcal{U} = u_{17}$. Since the limiting conditional survival function becomes deterministic once in state [•][•], the variability comes only from the conditional probability it takes to reach state [•][•] from the initial state [••], whose expected value is $1 - \mathcal{F}$ and variance $1 - \mathcal{U} - (1 - \mathcal{F})^2$. Letting $N \to \infty$ makes this variance go to 0 as shown in Supplementary Subsection 1.2, and the effect of the pedigree does indeed become negligible. This holds not only with gametic migration in a monoecious population, but also with diploid migration and in a dioecious population (see Supplementary Sections 4, 5, and Supplementary Table S2).

We conclude that the pedigree effect we observe in the many-demes limit is due to finite deme size, especially of the deme that initially contains the pair of ancestral lines. As such, we expect a persistent pedigree effect for other types of finely subdivided populations, for example metapopulations in which genetic exchange is mediated by extinction and recolonization rather than migration (Pannell and Charlesworth 2000; Pannell 2003). In fact, taking the limit $N \to \infty$ may be seen as counter to the very idea of metapopulations, which are generally thought to be composed of small local subpopulations (Hanski and Gaggiotti 2004).

Nordborg and Krone (2002) extended the strong-migration limit to allow migration fractions of order $O(N^{-\theta})$ for some $\theta \in [0, 1)$, not just of order $O(1)$, and called this fast migration. Hössjer (2011) adopted this terminology. Sagitov et al. (2010) studied coalescent processes in a related strong-migration limit for a haploid population model conditional on the migration fractions in each generation being determined by a random process, and found that it too reduced to a Kingman coalescent process. In light of these results, we might conjecture for example that a limit with $N \to \infty$ for fixed D, $m \in (0, 1)$, and $\alpha \in (0, 1)$ would yield the same coalescent

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

process whether conditioning on or averaging over the pedigree, and that it would be a Kingman coalescent process.

The rare-migration limit is the only case in which we observe a pedigree effect on coalescence that we cannot attribute to finite deme size. In this limit, each migration event replaces a non-negligible fraction of a deme but the generations in which these events occur are separated by many generations without gene flow, with complete isolation between demes. Similarly, the times and sizes of large reproduction events have a strong effect on coalescence times in a single well-mixed population even in the limit $N \to \infty$ (Diamantidis et al. 2024; Alberti et al. 2025). Taking these two results together suggests that pedigree effects on coalescence will be observed whenever "large" events happened with some frequency on the coalescent time scale. Here, a "large" event is one which affects multiple ancestral genetic lineages or multiple pairs of ancestral genetic lineages in the same deme or in different demes. This is precisely what is shown in Newman (2026).

Such large, rare-migration events can also be referred to as pulse admixture or introgression (Iasi et al. 2021; Thawornwattana et al. 2023). Our findings, together with those of Newman (2026), of a pedigree effect in this case means that coalescence depends on the particular times and sizes of these events. In fact, pulse admixture and introgression are typically treated as having happened at fixed times in the past, even in situations where they are referred to as recurrent (Buzbas and Verdu 2018; Ferreira et al. 2020; Di Santo et al. 2023). Our results and those of Newman (2026) also support the use of prior models for the times and sizes of such events in Bayesian approaches to inference, as discussed in Diamantidis et al. (2024) for the case of large reproduction events.

In contrast to the rare-migration limit, we note that the reason the pedigree effect becomes negligible as $N \to \infty$ in the low-migration limit is that migration events will almost surely involve single individuals. Although pedigrees will contain many generations without any migration events, when N is large it will be unlikely for genetic lineages to trace back to those single migrant individuals in generations where migration has occurred. Conditionally independent genetic lineages will randomly use different migration events, making the distribution of pairwise coalescence times among unlinked loci insensitive to the details of the pedigree.

We emphasize that conditioning on the pedigree is the natural way to accommodate multiple loci in coalescent models. Averaging over the pedigree yields what may be thought of as a prior model for a single genetic locus without intragenic recombination (Diamantidis et al. 2024). In contrast, the distributions of single-locus coalescence times in conditional models are predictions relevant to what can actually be observed at multiple unlinked loci in a sample of individuals.

Conditional and marginal coalescent models thus have different interpretations. The absence of a pedigree effect means that the same mathematical formulas are found in the limit. Then the same limiting mathematical model can be applied under either interpretation. The presence of a pedigree effect means that the two differ both in their mathematical representations and their interpretations. Which model to use will then depend on the application: single versus multiple loci. Similar distinctions arise in coalescent models of temporal variation in population size (Donnelly and Tavaré 1995; Schaper et al. 2012; Diamantidis et al. 2024).

The classification of structured-coalescent processes with migration that we have proposed is based on the distribution of single-locus pairwise coalescence times. It is possible that different conclusions would be reached if some other properties of the ancestral genetic process were considered. We have no particular reason to think that this would be the case, but we note that there have not yet been any studies of pedigree effects for multiple loci with recombination.

The implications of our results for empirical work include especially that the structured-coalescent model (Takahata 1988; Notohara 1990; Herbots 1997; Wilkinson-Herbots 1998) may be applied to multilocus data. When deme sizes are large, as they are assumed to be in this model, it will be unlikely for patterns of genetic variation to be affected by the pedigree. That is, if one were able to know the pedigree, it would be unlikely to include very recent shared ancestors. In highly subdivided metapopulations, or generally when deme sizes are not large, we expect to see pedigree effects on patterns of genetic variation. This is clearly the case for example in the pedigrees of Florida Scrub-Jays (Aguillon et al. 2017; Chen et al. 2019) and of the settlers of Québec (Anderson-Trocmé et al. 2023). Then, the recent pedigree ancestry of samples of genetic data will need to be taken into account, as for example was done in Anderson-Trocmé et al. (2023) to assess how much current population structure among French Canadians is due to the pedigree and how much to older population structure within France. In addition, our results and those of Newman (2026) for the rare-migration limit are part of a broader theoretical background for analyses of large-scale movement of individuals based on genetic variation, for example, in Ralph and Coop (2013) and Shastry et al. (2025).

Data availability

No empirical data were used in this work. Many details of the analyses are presented in Supplementary Material, specifically: additional analyses of the many-demes model, analyses of the low-migration and rare-migration limits under gametic migration and monoecy, as well as analyses of all four limits under diploid migration and dioecy. Code in the Julia programming language used to simulate coalescence times conditional on the population pedigree and to make the graphs displayed in Figs. 2 to 6, together with the Mathematica notebook `q17manydemes.nb` displayed in Supplementary Section 6 and used for some analyses of the many-demes model in Supplementary Section 1, are available at https://doi.org/10.6084/m9.figshare.32144278.

Supplemental material available at GENETICS online.

Acknowledgments

We thank Yichen Si for helpful calculations and discussions in the initial stages of this project, and three anonymous reviewers for their careful consideration and comments.

Funding

TE and SL were supported in part by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant 8833.

Conflicts of interest

None declared.

Literature cited

Achaz G. 2009. Frequency spectrum neutrality tests: one for all and all for one. Genetics. 183:249–258. https://doi.org/10.1534/genetics.109.104042.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Aguillon SM et al. 2017. Deconstructing isolation-by-distance: the genomic consequences of limited dispersal. PLoS Genet. 13: e1006911.https://doi.org/10.1371/journal.pgen.1006911.

Alberti F, Birkner M, Fan WTL, Wakeley J. 2025. A conditional coalescent for diploid exchangeable population models given the pedigree [preprint], arXiv, arXiv:2505.15481. https://doi.org/10.48550/arXiv.2505.15481.

Anderson-Trocmé L et al. 2023. On the genes, genealogies, and geographies of Quebec. Science. 380:849–855. https://doi.org/10.1126/science.add5300.

Aubree F, Lac B, Mailleret L, Calcagno V. 2022. Migration pulsedness alters patterns of allele fixation and local adaptation in a mainland-island model. Evolution. 77:718–730. https://doi.org/10.1093/evolut/qpac067.

Ball RM, Neigel JE, Avise JC. 1990. Gene genealogies within the organismal pedigrees of random-mating populations. Evolution. 44: 360–370. https://doi.org/10.1111/j.1558-5646.1990.tb05205.x.

Baronchelli A, Pastor-Satorras R. 2009. Effects of mobility on ordering dynamics. J Stat Mech. 2009:L11001. https://doi.org/10.1088/1742-5468/2009/11/L11001.

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. Proc Natl Acad Sci U S A. 98: 4563–4568. https://doi.org/10.1073/pnas.081068098.

Beerli P, Mashayekhi S, Sadeghi M, Khodaei M, Shaw K. 2019. Population genetic inference with MIGRATE. Curr Protoc Bioinform. 68:e87. https://doi.org/10.1002/cpbi.87.

Berestycki N. 2009. Recent progress in coalescent theory. Ensaios Math. 16:1–193. https://doi.org/10.21711/217504322009/em161.

Birkner M, Černý J, Depperschmidt A, Gantert N. 2013. Directed random walk on the backbone of an oriented percolation cluster. Electron J Probab. 18:1–35. https://doi.org/10.1214/EJP.v18-2302.

Birkner M, Liu H, Sturm A. 2018. Coalescent results for diploid exchangeable population models. Electron J Probab. 23:1–44. https://doi.org/10.1214/18-EJP175.

Bolthausen E, Sznitman A. 2002. On the static and dynamic points of view for certain random walks in random environment. Methods Appl Anal. 9:345–376. https://projecteuclid.org/journals/methods-and-applications-of-analysis/volume-9/issue-3/On-the-Satic-and-Dynamic-Points-of-View-for-Certain/maa/1119027729.full. https://doi.org/10.4310/MAA.2002.v9.n3.a4.

Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics. 140:783–796. https://doi.org/10.1093/genetics/140.2.783.

Burden CJ, Griffiths RC. 2018. Stationary distribution of a 2-island 2--allele Wright–Fisher diffusion model with slow mutation and migration rates. Theor Popul Biol. 124:70–80. https://doi.org/10.1016/j.tpb.2018.09.004.

Buzbas EO, Verdu P. 2018. Inference on admixture fractions in a mechanistic model of recurrent admixture. Theor Popul Biol. 122:149–157. https://doi.org/10.1016/j.tpb.2018.03.006.

Chang JT. 1999. Recent common ancestors of all present-day individuals. Adv Appl Probab. 31:1002–1026. https://doi.org/10.1239/aap/1029955256.

Chen N et al. 2019. Allele frequency dynamics in a pedigreed natural population. Proc Natl Acad Sci U S A. 116:2158–2164. https://doi.org/10.1073/pnas.1813852116.

Cherry JL. 2003. Selection in a subdivided population with dominance or local frequency dependence. Genetics. 163:1511–1518. https://doi.org/10.1093/genetics/163.4.1511.

Costa RJ, Wilkinson-Herbots H. 2017. Inference of gene flow in the process of speciation: an efficient maximum-likelihood method for the isolation-with-initial-migration model. Genetics. 205: 1597–1618. https://doi.org/10.1534/genetics.116.188060.

Crane H. 2016. The ubiquitous Ewens sampling formula. Stat Sci. 31: 1–19. https://doi.org/10.1214/15-STS529.

Di Santo LN, Quilodrán CS, Currat M. 2023. Temporal variation in introgressed segments' length statistics computed from a limited number of ancient genomes sheds light on past admixture pulses. Mol Biol Evol. 40:msad252. https://doi.org/10.1093/molbev/msad252.

Diamantidis D, Fan WTL, Birkner M, Wakeley J. 2024. Bursts of coalescence within population pedigrees whenever big families occur. Genetics. 227:iyae030. https://doi.org/10.1093/genetics/iyae030.

Donnelly P. 1986. Dual processes in population genetics. In: Tautu P, editor. Stochastic spatial processes. Springer. p. 94–105. https://doi.org/10.1007/BFb0076240.

Donnelly P, Tavaré S. 1995. Coalescents and genealogical structure under neutrality. Annu Rev Genet. 29:401–421. https://doi.org/10.1146/annurev.ge.29.120195.002153.

Ethier SN, Nagylaki T. 1980. Diffusion approximations of Markov chains with two time scales and applications to population genetics. Adv Appl Probab. 12:14–49. https://doi.org/10.2307/1426492.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. Theor Popul Biol. 3:87–112. https://doi.org/10.1016/0040-5809(72)90035-4.

Ewens WJ. 1990. Population genetics theory—the past and the future. In: Lessard S, editor. Mathematical and statistical developments of evolutionary theory. Kluwer Academic Publishers. p. 177–227.

Ewens WJ. 2004. Mathematical population genetics, volume I: theoretical foundations. Springer-Verlag.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. PLoS Genet. 9:e1003905.https://doi.org/10.1371/journal.pgen.1003905.

Excoffier L et al. 2021. fastsimcoal2: demographic inference under complex evolutionary scenarios. Bioinformatics. 37:4882–4885. https://doi.org/10.1093/bioinformatics/btab468.

Fan LWT, Newman M, Wakeley J. 2025. Quenched coalescent for diploid population models with selfing and overlapping generations [preprint]. arXiv, arXiv:2510.26115. https://doi.org/10.48550/arXiv.2510.26115.

Ferreira MS et al. 2020. The legacy of recurrent introgression during the radiation of hares. Syst Biol. 70:593–607. https://doi.org/10.1093/sysbio/syaa088.

Fu Y. 1995. Statistical properties of segregating sites. Theor Popul Biol. 48:172–197. https://doi.org/10.1006/tpbi.1995.1025.

Gaggiotti OE, Smouse PE. 1996. Stochastic migration and maintenance of genetic variation in sink populations. Am Nat. 147: 919–945 (Also available at https://www.jstor.org/stable/2463185). https://doi.org/10.1086/285886.

Gao F, Keinan A. 2016. Inference of super-exponential human population growth via efficient computation of the site frequency spectrum for generalized models. Genetics. 202:235–245. https://doi.org/10.1534/genetics.115.180570.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 5: e1000695.https://doi.org/10.1371/journal.pgen.1000695.

Hanski I, Gaggiotti OE. 2004. Ecology, genetics, and evolution of metapopulations. Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-323448-3.X5000-4

Hein J, Schierup MH, Wiuf C. 2005. Gene genealogies, variation and evolution: a primer in coalescent theory. Oxford University Press.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Herbots HM. 1997. The structured coalescent. In: Donnelly P, Tavaré S, editors. Progress in population genetics and human evolution (IMA volumes in mathematics and its applications, vol. 87). Springer-Verlag. p. 231–255.
Heuertz M, Vekemans X, Hausman J, Palada M, Hardy OJ. 2003. Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. Mol Ecol. 12:2483–2495. https://doi.org/10.1046/j.1365-294x.2003.01923.x.
Hössjer O. 2011. Coalescence theory for a general class of structured populations with fast migration. Adv Appl Probab. 43:1027–1047. https://doi.org/10.1239/aap/1324045697.
Hudson RR. 1983a. Properties of a neutral allele model with intragenic recombination. Theor Popul Biol. 23:183–201. https://doi.org/10.1016/0040-5809(83)90013-8.
Hudson RR. 1983b. Testing the constant-rate neutral allele model with protein sequence data. Evolution. 37:203–217. https://doi.org/10.1111/j.1558-5646.1983.tb05528.x.
Iasi LNM, Ringbauer H, Peter BM. 2021. An extended admixture pulse model reveals the limitations to human-Neandertal introgression dating. Mol Biol Evol. 38:5156–5174. https://doi.org/10.1093/molbev/msab210.
Kamm J, Terhorst J, Durbin R, Song YS. 2020. Efficiently inferring the demographic history of many populations with allele count data. J Am Stat Assoc. 115:1472–1487. https://doi.org/10.1080/01621459.2019.1635482.
Karlin S, Taylor HM. 1975. A first course in stochastic Processes. 2nd ed. Academic Press.
Kingman JFC. 1982a. On the genealogy of large populations. J Appl Probab. 19:27–43. https://doi.org/10.2307/3213548.
Kingman JFC. 1982b. The coalescent. Stoch Process Their Appl. 13:235–248. https://doi.org/10.1016/0304-4149(82)90011-4.
Kingman JFC. 1982c. Exchangeability and the evolution of large populations. In: Koch G, Spizzichino F, editors. Exchangeability in probability and statistics. North-Holland. p. 97–112.
Kogan D, Diamantidis D, Wakeley J, Fan LWT. 2026. Correlation of coalescence times in a diploid Wright–Fisher model with recombination and selfing. Theor Popul Biol. 167:40–68. https://doi.org/10.1016/j.tpb.2025.11.003.
Ladret V, Lessard S. 2007. Fixation probability for a beneficial allele and a mutant strategy in a linear game under weak selection in a finite island model. Theor Popul Biol. 72:409–425. https://doi.org/10.1016/j.tpb.2007.04.001.
Lande R. 1979. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. Evolution. 33:234–251. https://doi.org/10.1111/j.1558-5646.1979.tb04678.x.
Lande R. 1985. The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. Heredity (Edinb). 54:323–332. https://doi.org/10.1038/hdy.1985.43.
Latter BDH. 1973. The island model of population differentiation: a general solution. Genetics. 73:147–157. https://doi.org/10.1093/genetics/73.1.147.
Latter BDH, Sved JA. 1981. Migration and mutation in stochastic models of gene frequency change. II. Stochastic migration with a finite number of islands. J Math Biol. 13:95–104. https://doi.org/10.1007/BF00276868.
Lessard S, Wakeley J. 2004. The two-locus ancestral graph in a subdivided population: convergence as the number of demes grows in the island model. J Math Biol. 48:275–292. https://doi.org/10.1007/s00285-003-0230-x.
Levin DA, Kerster HW. 1974. Gene flow in seed plants. In: Dobzhansky T, Hecht MK, Steere WC, editors. Evolutionary biology: volume 7. Springer US. p. 139–220. https://doi.org/10.1007/978-1-4615-6944-2_5
Lohse K, Chmelik M, Martin SH, Barton NH. 2016. Efficient strategies for calculating blockwise likelihoods under the coalescent. Genetics. 202:775–786. https://doi.org/10.1534/genetics.115.183814.
Lohse K, Harrison RJ, Barton NH. 2011. A general method for calculating likelihoods under the coalescent process. Genetics. 189:977–987. https://doi.org/10.1534/genetics.111.129569.
Maruyama T. 1970. Effective number of alleles in a subdivided population. Theor Popul Biol. 1:273–306. https://doi.org/10.1016/0040-5809(70)90047-x.
McLaughlin JF *et al.* 2023. Multivariate models of animal sex: breaking binaries leads to a better understanding of ecology and evolution. Integr Comp Biol. 63:891–906. https://doi.org/10.1093/icb/icad027.
Miyagi M, Huerta-Sánchez E, Richardson SS. 2025. Contextualizing population genetic models of sex-biased migration and admixture. Am J Hum Biol. 37:e70183. https://doi.org/10.1002/ajhb.70183.
Möhle M. 1998a. A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. Adv Appl Probab. 30:493–512. https://doi.org/10.1239/aap/1035228080.
Möhle M. 1998b. Coalescent results for two-sex population models. Adv Appl Probab. 30:513–520. https://doi.org/10.1239/aap/1035228081.
Möhle M. 1999. The concept of duality and applications to Markov processes arising in neutral population genetics models. Bernoulli (Andover). 5:761–777. https://doi.org/10.2307/3318443.
Müller NF, Rasmussen DA, Stadler T. 2017. The structured coalescent and its approximations. Mol Biol Evol. 34:2970–2981. https://doi.org/10.1093/molbev/msx186.
Nagylaki T. 1979. The island model with stochastic migration. Genetics. 91:163–176. https://doi.org/10.1093/genetics/91.1.163.
Nagylaki T. 1980. The strong-migration limit in geographically structured populations. J Math Biol. 9:101–114. https://doi.org/10.1007/BF00275916.
Newman M. 2026. The quenched structured coalescent for diploid population models on finite graphs with large migrations and uneven offspring distributions [preprint]. arXiv, arXiv:2601.21079. https://doi.org/10.48550/arXiv.2601.21079.
Newman M, Wakeley J, Fan WTL. 2025. Conditional gene genealogies given the population pedigree for a diploid moran model with selfing. Theor Popul Biol. 165:29–44. https://doi.org/10.1016/j.tpb.2025.05.003.
Nishino J, Tajima F. 2004. Effect of population structure on the amount of polymorphism and the fixation probability under overdominant selection. Genes Genet Syst. 79:41–48. https://doi.org/10.1266/ggs.79.41.
Nishino J, Tajima F. 2005. Effect of population structure on the amount of polymorphism and the fixation probability under overdominant selection. Genes Genet Syst. 80:287–295. https://doi.org/10.1266/ggs.80.287.
Nordborg M. 1997. Structured coalescent processes on different time scales. Genetics. 146:1501–1514. https://doi.org/10.1093/genetics/146.4.1501.
Nordborg M. 2001. Coalescent theory. In: Balding DJ, Bishop M, Cannings C, editors. Handbook of statistical genetics. John Wiley & Sons, Chichester. p. 179–212.
Nordborg M, Donnelly P. 1997. The coalescent process with selfing. Genetics. 146:1185–1195. https://doi.org/10.1093/genetics/146.3.1185.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Nordborg M, Krone SM. 2002. Separation of time scales and convergence to the coalescent in structured populations. In: Slatkin M, Veuille M, editors. Modern developments in theoretical population genetics: the legacy of Gustave Malécot. Oxford University Press. p. 194–232. https://doi.org/10.1093/oso/9780198599623.001.0001

Notohara M. 1990. The coalescent and the genealogical process in geographically structured population. J Math Biol. 29:59–75. https://doi.org/10.1007/BF00173909.

Notohara M. 1993. The strong-migration limit for the genealogical process in geographically structured populations. J Math Biol. 31:115–122. https://doi.org/10.1007/BF00171220.

Notohara M. 2001. The structured coalescent process with weak migration. J Appl Probab. 38:1–17. https://doi.org/10.1239/jap/996986639.

Pannell JR. 2003. Coalescence in a metapopulation with recurrent local extinction and recolonization. Evolution. 57:949–961. https://doi.org/10.1111/j.0014-3820.2003.tb00307.x.

Pannell JR, Charlesworth B. 2000. Effects of metapopulation processes on measures of genetic diversity. Philos Trans R Soc Lond B Biol Sci. 355:1851–1864. https://doi.org/10.1098/rstb.2000.0740.

Patel RA *et al.* 2025. Characterizing selection on complex traits through conditional frequency spectra. Genetics. 229:iyae210. https://doi.org/10.1093/genetics/iyae210.

Peniston JH, Backus GA, Baskett ML, Fletcher RJ, Holt RD. 2024. Ecological and evolutionary consequences of temporal variation in dispersal. Ecography. 2024:e06699. https://doi.org/10.1111/ecog.06699.

Pinho C, Hey J. 2010. Divergence with gene flow: models and data. Annu Rev Ecol Evol Syst. 41:215–230. https://doi.org/10.1146/annurev-ecolsys-102209-144644.

Ragsdale AP, Moreau C, Gravel S. 2018. Genomic inference using diffusion models and the allele frequency spectrum. Curr Opin Genet Dev. 53:140–147. https://doi.org/10.1016/j.gde.2018.10.001.

Ralph P, Coop G. 2013. The geography of recent genetic ancestry across Europe. PLoS Biol. 11:e1001555.https://doi.org/10.1371/journal.pbio.1001555.

Rice SH, Papadopoulos A. 2009. Evolution with stochastic fitness and stochastic migration. PLoS One. 4:e7130. https://doi.org/10.1371/journal.pone.0007130.

Rousset F. 2002. Inbreeding and relatedness coefficients: what do they measure? Heredity (Edinb). 88:371–380. https://doi.org/10.1038/sj.hdy.6800065.

Sagitov S, Jagers P, Vatutin V. 2010. Coalescent approximation for structured populations in a stationary random environment. Theor Popul Biol. 78:192–199. https://doi.org/10.1016/j.tpb.2010.06.008.

Schaper E, Eriksson A, Rafajlovic M, Sagitov S, Mehlig B. 2012. Linkage disequilibrium under recurrent bottlenecks. Genetics. 190:217–229. https://doi.org/10.1534/genetics.111.134437.

Severson AL, Carmi S, Rosenberg NA. 2019. The effect of consanguinity on between-individual identity-by-descent sharing. Genetics. 212:305–316. https://doi.org/10.1534/genetics.119.302136.

Severson AL, Carmi S, Rosenberg NA. 2021. Variance and limiting distribution of coalescence times in a diploid model of a consanguineous population. Theor Popul Biol. 139:50–65. https://doi.org/10.1016/j.tpb.2021.02.002.

Shastry V, Musiani M, Novembre J. 2025. Jointly representing long-range genetic similarity and spatially heterogeneous isolation-by-distance. PLoS Genet. 21:e1011612.https://doi.org/10.1371/journal.pgen.1011612.

Slatkin M. 1981. Fixation probabilities and fixation times in a subdivided population. Evolution. 35:477–488. https://doi.org/10.1111/j.1558-5646.1981.tb04911.x.

Slatkin M. 1991. Inbreeding coefficients and coalescence times. Genet Res (Camb). 58:167–175. https://doi.org/10.1017/S0016672300029827.

Sork VL, Smouse PE, Grivet D, Scofield DG. 2015. Impact of asymmetric male and female gamete dispersal on allelic diversity and spatial genetic structure in valley oak (*Quercus lobata* née). Evol Ecol. 29:927–945. https://doi.org/10.1007/s10682-015-9769-4.

Sved JA, Latter BDH. 1977. Migration and mutation in stochastic models of gene frequency change. J Math Biol. 5:61–73. https://doi.org/10.1007/BF00275807.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics. 105:437–460. https://www.genetics.org/content/105/2/437. https://doi.org/10.1093/genetics/105.2.437.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–595. https://doi.org/10.1093/genetics/123.3.585.

Takahata N. 1988. The coalescent in two partially isolated diffusion populations. Genet Res (Camb). 52:213–222. https://doi.org/10.1017/S0016672300027683.

Takahata N. 1991. Genealogy of neutral genes and spreading of selected mutations in a geographically structured population. Genetics. 129:585–595. https://doi.org/10.1093/genetics/129.2.585.

Thawornwattana Y, Huang J, Flouri T, Mallet J, Yang Z. 2023. Inferring the direction of introgression using genomic sequence data. Mol Biol Evol. 40:msad178. https://doi.org/10.1093/molbev/msad178.

Tyukin A. 2015. Quenched limits of coalescents in fixed pedigrees [master's thesis]. Johannes-Gutenberg-Universität Mainz.

Vitalis R, Couvet D. 2001. Two-locus identity probabilities and identity disequilibrium in a partially selfing subdivided population. Genet Res (Camb). 77:67–81. https://doi.org/10.1017/S0016672300004833.

Wakeley J. 1998. Segregating sites in Wright's island model. Theor Popul Biol. 53:166–174. https://doi.org/10.1006/tpbi.1997.1355.

Wakeley J. 2003. Polymorphism and divergence for island-model species. Genetics. 163:411–420. https://doi.org/10.1093/genetics/163.1.411.

Wakeley J. 2004a. Recent trends in population genetics: more data! More math! Simple models? J Hered. 95:397–405. https://doi.org/10.1093/jhered/esh062.

Wakeley J. 2004b. Metapopulation models for historical inference. Mol Ecol. 13:865–875. https://doi.org/10.1111/j.1365-294X.2004.02086.x.

Wakeley J. 2009. Coalescent theory: an introduction. Roberts & Company Publishers (Current publisher: Macmillan Learning).

Wakeley J, King L, Low BS, Ramachandran S. 2012. Gene genealogies within a fixed pedigree, and the robustness of Kingman's coalescent. Genetics. 190:1433–1445. https://doi.org/10.1534/genetics.111.135574.

Wakeley J, Takahashi T. 2004. The many-demes limit for selection and drift in a subdivided population. Theor Popul Biol. 66:83–91. https://doi.org/10.1016/j.tpb.2004.04.005.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 7:256–276. https://doi.org/10.1016/0040-5809(75)90020-9.

Weir BS, Avery PJ, Hill WG. 1980. Effect of mating structure on variation in inbreeding. Theor Popul Biol. 18:396–429. https://doi.org/10.1016/0040-5809(80)90061-1.

Whitlock MC. 1992. Temporal fluctuations in demographic parameters and the genetic variance among populations. Evolution. 46:608–615. https://doi.org/10.1111/j.1558-5646.1992.tb02069.x.

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026

Wilkinson-Herbots HM. 1998. Genealogy and subpopulation differentiation under various models of population structure. J Math Biol. 37:535–585. https://doi.org/10.1007/s002850050140.
Wolfram Research, Inc. 2024. Mathematica, version 14.2. Version 14.2.
Wright S. 1931. Evolution in Mendelian populations. Genetics. 16: 97–159. https://doi.org/10.1093/genetics/16.2.97.
Wright S. 1943. Isolation by distance. Genetics. 28:114–138. https://doi.org/10.1093/genetics/28.2.114.
Wright S. 1951. The genetical structure of populations. Ann Eugen. 15:323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x.
Yamaguchi R, Iwasa Y. 2013. First passage time to allopatric speciation. Interface Focus. 3:20130026. https://doi.org/10.1098/rsfs.2013.0026.
Yamaguchi R, Iwasa Y. 2017. A tipping point in parapatric speciation. J Theor Biol. 421:81–92. https://doi.org/10.1016/j.jtbi.2017.03.018.

Editor: P. Ralph

Downloaded from academic.oup.com/genetics/article/233/4/iyag127/8679135 by Universite de Montreal user on 12 November 2026