

7

Specific Models

7.1 Coalescent Processes: Reversed Branching

P. Jagers and S. Sagitov

7.1.1 Coalescence

Branching viewed backward is *coalescence*, the process of merging or clumping. It arises naturally in the structuring or formation of dispersed matter of various kinds and in various scales, from that of molecule aggregates in colloids (so-called *micelles*) to galaxies. It has been studied by physicists through computer simulation (e.g., Nilsson *et al.* 2000) and in a series of interesting mathematical articles by Aldous (1999).

Evolution can be viewed as a grand multi-type branching process, with new species that arise through mutation (see Jagers 1991; Jagers *et al.* 1992; Taib 1992). The study of the origin of species is then time-reversed branching (i.e., coalescence). In genetics, the latter is also used to trace the roots of the genetic composition of populations and its development. It was within this area that the first pure coalescence model, the Kingman coalescent (1982a), was formulated as a reverse counterpart to the diffusion approximation of the renowned Wright–Fisher model (Fisher 1930; Wright 1931; see also Ewens 1979).

The object of genetic models is thus population composition rather than size. Indeed, most population genetics even assumes that population size is completely constant over generations. As we show later, the Wright–Fisher model can be obtained as Galton–Watson branching with Poisson reproduction, conditioned at a constant population size. In the same vein, most population genetics simplifies the flow of time into generation counting. Instead, it is lineage that counts. What are the relations among n individuals sampled out of a total population of size N ?

For (diploid) genetics mating also cannot be disregarded. It is usually modeled as random, that is, the two genes of an individual are selected independently at random from any individual in the preceding generation. (The latter are thus, implicitly or explicitly, thought of as hermaphroditic.) Most of the time genes are passed unchanged, but sometimes new alleles appear through mutation. We consider only *neutral* mutations, so that the reproduction law remains one and the same throughout the process. In such cases, what is the distribution of alleles at a locus?

If mating is thus assumed to be random and population size is thought to be given, gene branching might be studied and simulated directly, with individual reproduction behavior bypassed by means of a general exchangeable population

model (to be further discussed in Section 7.1.4). A vector of random variables is said to be *exchangeable* if any permutation of its components has the same distribution. In symbols, if i_1, i_2, \dots, i_N is any reordering (permutation) of the numbers $1, 2, \dots, N$, then the vector $(v_{i_1}, v_{i_2}, \dots, v_{i_N})$ has the same joint distribution function as has (v_1, v_2, \dots, v_N) . In an exchangeable population model N is the size of the reproducing generation and $v_1 + v_2 + \dots + v_N$, which is usually also required to equal N , is the size of the following generation.

7.1.2 The Wright–Fisher population model

The best-known exchangeable reproduction model is the *Wright–Fisher* population. It depicts a population in (an imagined) equilibrium in which the population size N (the number of genes) is constant over generations, time thus being discrete. It is assumed that gene number i in the current generation is passed on to the next generation in v_i copies so that the joint distribution of gene offspring sizes (v_1, \dots, v_N) is symmetrically multinomial. In line with earlier notation, for binomial, geometric, and Poisson random variables this can be denoted by the vector being multinomial $(N, \frac{1}{N}, \dots, \frac{1}{N})$ (see the Appendix).

A remarkable feature of the Wright–Fisher model, which emphasizes its concentration upon lineage at the expense of size fluctuations, is the mentioned interpretation of it as a Galton–Watson process with Poisson reproduction law, conditioned to have constant population sizes. Indeed, if the offspring distribution is Poisson(m) and ξ_1, ξ_2, \dots are the numbers of offspring of the individuals in the n th generation (somehow numbered), then provided $i_1 + i_2 + \dots + i_N = N$,

$$\begin{aligned} & \mathbb{P}(\xi_1 = i_1, \xi_2 = i_2, \dots, \xi_N = i_N) \\ &= \mathbb{P}(\xi_1 = i_1, \xi_2 = i_2, \dots, \xi_N = i_N | \xi_1 + \dots + \xi_N = N) \\ &= \frac{\frac{m^{i_1}}{i_1!} e^{-mi_1} \frac{m^{i_2}}{i_2!} e^{-mi_2} \dots \frac{m^{i_N}}{i_N!} e^{-mi_N}}{\frac{(mN)^N}{N!} e^{-mN}} = \frac{N!}{i_1! i_2! \dots i_N!} \left(\frac{1}{N}\right)^N. \end{aligned} \quad (7.1)$$

Otherwise the probability is obviously equal to zero.

The reverse description of how Wright–Fisher populations evolved is particularly simple: the N children are partitioned into sibships with multinomially distributed sizes (v_1, \dots, v_N) by letting them select their parents independently at random from among the N individuals that constitute the preceding generation. This leads to a simple algorithm to simulate n alleles sampled from a Wright–Fisher population:

1. Build an ancestral tree for n sampled genes that result from n coalescing random walks over the set of N points;
2. Impose mutations on the ancestral tree.

The reverse approach renders simulation faster because it deals only with the fate of direct ancestors of the sampled genes. Figure 7.1 depicts a possible outcome of the reverse simulation algorithm with $n = 6$. Mutations are numbered in time sequence and the resultant six alleles are labeled according to the mutations that

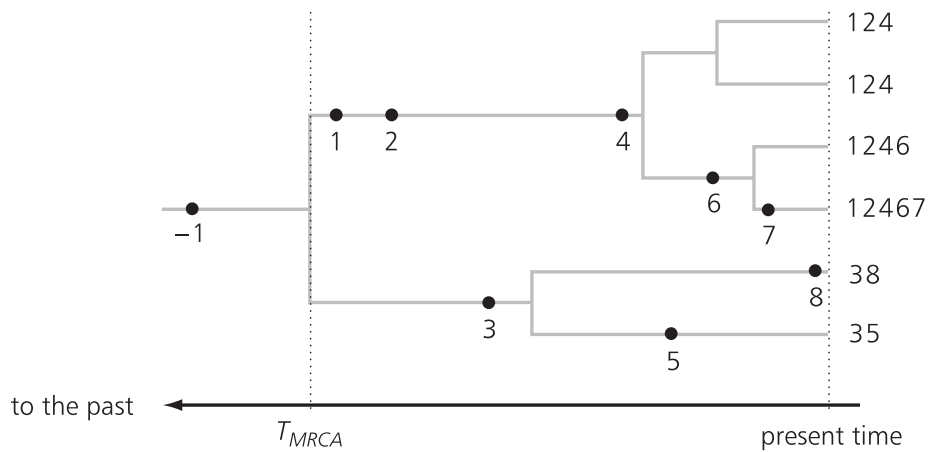


Figure 7.1 An example of simulating six sequences using the Kingman coalescent. Dots stand for mutations numbered in their order of occurrence.

created them. Mutations prior to the *most recent common ancestor* (MRCA) do not contribute to the sample polymorphism.

As $N \rightarrow \infty$, the forward picture of the Wright–Fisher model can be simplified by a diffusion approximation. However, there is a reverse counterpart to this diffusion approximation, called the Kingman coalescent (Kingman 1982a). It is an asymptotic ancestral process and the subject of Section 7.1.3. It is a basic coalescent process with several extensions that reflect such evolutionary and demographic features as recombination, population subdivision, variable population size, and selection. A reader interested in the theory of coalescent processes and its applications in population genetics is referred to Chapters 6–8 of Balding *et al.* (2001). A recent review by Rosenberg and Nordborg (2002) contains further references. Important earlier surveys are Tavaré (1984), Hudson (1991), and Donnelly and Tavaré (1995).

7.1.3 The Kingman coalescent

Consider the genealogical tree of n genes sampled at random from a Wright–Fisher population with constant population size N . Let R_k^N be the number of branches in the tree k generations before the generation sampled. For a fixed N the sequence $(R_k^N)_{k=0,1,\dots}$, the *ancestral process*, has monotone trajectories that decrease from $R_0^N = n$ down to $R_T^N = 1$ at the random time $T = T_{MRCA}$, when the tree reaches the MRCA of the sampled genes. It is a Markov chain and it has been shown (Watterson 1975) that its transition probabilities have the form

$$\mathbb{P}(R_{k+1}^N = j \mid R_k^N = i) = c_{i,j} N(N-1) \cdots (N-j+1) / N^i, \quad (7.2)$$

where $c_{i,j}$ are combinatorial constants independent of N (actually, so-called Stirling numbers of the second kind). In particular, $c_{i,i} = 1$ and $c_{i,i-1} = \binom{i}{2} = i(i-1)/2$, so

$$\begin{aligned} \mathbb{P}(R_{k+1}^N = i \mid R_k^N = i) &= 1(1 - 1/N) \cdots (1 - (i - 1)/N) \\ &= 1 - (1 + 2 + \cdots + i - 1)/N + o(1/N) \\ &= 1 - i(i - 1)/(2N) + o(1/N), \end{aligned} \tag{7.3}$$

as $N \rightarrow \infty$. Similarly,

$$\mathbb{P}(R_{k+1}^N = i - 1 \mid R_k^N = i) = i(i - 1)/(2N) + o(1/N), \tag{7.4}$$

and the remaining probabilities

$$\mathbb{P}(R_{k+1}^N = j \mid R_k^N = i) = o(1/N), \quad j \leq i - 2. \tag{7.5}$$

If N is large (as compared with n^2), then Equation (7.3) means that the chain R_k^N stays in the same state for many generations, until it jumps to a new state. The number of generations between two jumps should be comparable with the population size N . If such a jump takes place, in view of Equation (7.5), the probability is high that the number of branches in the gene tree reduces by just one. This corresponds to a pairwise coalescent. Equations (7.3)–(7.5) imply the following convergence result (cf. Kingman 1982a), which suggests a remarkably convenient way to simulate the gene tree of a large Wright–Fisher population.

Theorem 7.1 *If R_k^N is the number of branches in the gene tree of a Wright–Fisher population with population size N , then as $N \rightarrow \infty$ the process*

$$R_{[Nt]}^N \rightarrow R_t, \quad t \geq 0, \quad N \rightarrow \infty. \tag{7.6}$$

Here $[Nt]$ means the integral part of Nt (i.e., the largest integer that does not exceed Nt), and the limit process R_t , to be described below, is the *Kingman coalescent*. The arrow stands for so-called weak convergence, which is actually a fairly strong convergence concept in stochastic processes. For an introduction to this, see more advanced probability literature (Billingsley 1968). Suffice it here to say that it, usually, implies convergence of the distributions at a fixed t (problems possibly arise at values of t where the limit has a jump with positive probability). However, it also implies convergence of distributions of vectors,

$$(R_{[Nt_1]}^N, R_{[Nt_2]}^N, \dots, R_{[Nt_r]}^N) \rightarrow (R_{t_1}, R_{t_2}, \dots, R_{t_r}), \tag{7.7}$$

and even of functionals like the supremum of the process.

The Kingman coalescent is a Markov process that describes the asymptotic distribution of a gene tree for large Wright–Fisher populations. The process R_t yields the number of branches in the Kingman coalescent. The theorem shows that time t in the limit tree corresponds to $[Nt]$ generations in the Wright–Fisher population, counted backward in time. Equations (7.3)–(7.5) for the ancestral process are transferred into the defining properties of the Kingman coalescent,

$$\mathbb{P}(R_{t+\Delta} = j \mid R_t = i) = q_{ij} \cdot \Delta + o(\Delta), \quad j \neq i, \tag{7.8}$$

and

$$\mathbb{P}(R_{t+\Delta} = i \mid R_t = i) = 1 - \lambda_i \cdot \Delta + o(\Delta), \quad \Delta \rightarrow 0, \tag{7.9}$$

where

$$q_{ij} = \begin{cases} \binom{i}{2} & \text{if } j = i - 1, \\ 0 & \text{otherwise,} \end{cases} \tag{7.10}$$

and $\lambda_i = \sum_{j \neq i} q_{ij}$ is equal to

$$\lambda_i = \binom{i}{2} = i(i - 1)/2. \tag{7.11}$$

These relations lead to a simple wait-and-jump description of the limit process R_t . First, according to Equation (7.9) the waiting time T_i spent by the limit process at the state i is distributed exponentially with parameter λ_i , so that

$$\mathbb{E}[T_i] = \frac{2}{i(i - 1)}, \quad \text{Var}[T_i] = \frac{4}{i^2(i - 1)^2}, \quad i \geq 2. \tag{7.12}$$

Second, in view of Equations (7.10) and (7.11), the jump probabilities $p_{ij} = q_{ij}/\lambda_i$ satisfy

$$p_{ij} = \begin{cases} 1 & \text{if } j = i - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.13}$$

This means that the number of branches will be $i - 1$ immediately after it was i . In other words, the limit process of Theorem 7.1 is what is known as a pure death process, which consecutively visits all the states $\{n, n - 1, \dots, 2, 1\}$ with waiting times as described.

7.1.4 Exchangeable population models

The Kingman coalescent was derived as a limit of a very special, albeit classic, genetic model, the Wright–Fisher model. The convergence result of Theorem 7.1 can, however, be extended to a broad class of population models, which thus widens the applicability of the coalescent far beyond Wright and Fisher’s framework.

Exchangeable population models (cf. Cannings 1974) generalize the Wright–Fisher setup by allowing the offspring sizes (ν_1, \dots, ν_N) to have any symmetric (exchangeable) joint distribution. Exchangeability, as pointed out, means that the joint distribution does not change if the family indices are permuted. This implies that all offspring sizes ν_1, \dots, ν_N have the same distribution and therefore the same expectation and variance. Since

$$N = \mathbb{E}[\nu_1 + \dots + \nu_N] = \mathbb{E}[\nu_1] + \dots + \mathbb{E}[\nu_N], \tag{7.14}$$

the common expected offspring number must be $\mathbb{E}[\nu_i] = 1$. We denote the variance by σ_N^2 .

Exchangeability of gene inheritance is natural under the assumption of only neutral mutations. This class of population models also encompasses all Galton–

Watson processes conditioned to have a constant population size N in all generations.

Kingman (1982b) presents two simple conditions, namely

$$\sigma_N^2 \rightarrow \sigma^2, N \rightarrow \infty, 0 < \sigma^2 < \infty, \quad (7.15)$$

and

$$\sup_N \mathbb{E}[v_1^k] < \infty, k \geq 3, \quad (7.16)$$

which ensure a convergence to the coalescent,

$$R_{[Nt/\sigma^2]}^N \rightarrow R_t, N \rightarrow \infty, \quad (7.17)$$

for the ancestral process R_k^N of a population with exchangeable reproduction. The limit process is again the Kingman ancestral process, while the ancestral process of the exchangeable population is accelerated by a factor of σ^2 , which reflects that the ancestral lines merge faster in a population with greater variation in offspring size.

Minimal conditions for convergence to the Kingman coalescent were found by Möhle and Sagitov (1999). A simplified version of their result is given in Theorem 7.2.

Theorem 7.2 *Let $R_k^N, k = 0, 1, 2, \dots$ be the number of branches in the gene tree of an exchangeable population with constant population size N that satisfies Condition (7.15). The convergence (7.17) to a Kingman coalescent holds if and only if*

$$\mathbb{E}[v_1^3]/N \rightarrow 0, N \rightarrow \infty. \quad (7.18)$$

Biologically, Theorem 7.2 says that in a large population of N individuals, the ancestral process is of the Kingman type, which does not allow mergers of more than two lines simultaneously, if and only if a single sibship attains a very large size with a diminutive probability only (but see Box 7.1).

There are two types of natural follow-up questions:

1. How large does N need to be to ensure validity of the approximation (7.17) (a million, a billion, or only 100 or 1000)?
2. In general, when the restriction (7.15) is removed, what are the reproduction laws that lead to coalescent processes that allow triple mergers?

The answer to the first question is given in Möhle (2000), in which the distributions of two random vectors connected by

$$(R_{[Nt_1/\sigma^2]}^N, R_{[Nt_2/\sigma^2]}^N, \dots, R_{[Nt_r/\sigma^2]}^N) \rightarrow (R_{t_1}, R_{t_2}, \dots, R_{t_r}) \quad (7.19)$$

are compared. It was found that the corresponding probabilities differ at most by the value

$$c_1 \mathbb{E}[v_1^3]/N + c_2 \mathbb{E}[v_1^2 v_2^2]/N + c_3/N. \quad (7.20)$$

Here, the constants c_1 , c_2 , and c_3 depend, in a complicated way, on the sample size n , the dimension k , and the variance σ^2 . This implies, in particular, that under the Kingman moment conditions the convergence rate in Equation (7.17) is c/N . In general, the dominating term in the estimate is $c_1\mathbb{E}[v_1^3]/N$, in accordance with Theorem 7.2.

The second item was analyzed recently by Sagitov (1999, 2003). It turns out that the ancestral process allows mergers of more than two lines simultaneously if the largest sibship $v_{(1)}$ in the generation exceeds Nx with a probability that satisfies a smallness condition. We obtain convergence to the Kingman coalescent with a more flexible scaling of time,

$$(R_{\lfloor Nt/\sigma_N^2 \rfloor}^N)_{t \geq 0} \rightarrow (R_t)_{t \geq 0}, \quad N \rightarrow \infty, \tag{7.21}$$

if for any fixed number $x > 0$,

$$N\sigma_N^{-2}\mathbb{P}(v_{(1)} > Nx) \rightarrow 0. \tag{7.22}$$

If there is a non-zero limit, we could easily obtain multiple mergers with three or more branches that join at one node. If, furthermore, the second-largest sibship $v_{(2)}$ satisfies

$$N\sigma_N^{-2}\mathbb{P}(v_{(2)} > Nx) \rightarrow 0, \tag{7.23}$$

for all $x > 0$, it is impossible to encounter two large sibling groups in the same generation among $\lfloor N/\sigma_N^2 \rfloor$ consecutive generations, which effectively prohibits simultaneous mergers in a coalescence with multiple mergers.

Example 7.1 However, even though situations in which Equation (7.22) is not satisfied seem impossible in the realm of mammals, they could certainly be thought of among e.g., reptiles, fish, or insects. As an illustration, assume that $N = 100$ is large enough for Theory 7.2 to apply. In such a small population of, say, turtles, it could certainly happen once in a century that one of the turtles is the mother of, maybe, half the population. In other terms, the probability of begetting 50 offspring or more does not seem to be less than 0.0001.

As a consequence, the probability that the maximal sibship exceeds 50 individuals may well be larger than 0.01.



Schweinsberg (2003) considers an important family of the exchangeable reproduction models with fixed population size N that arise from supercritical Galton–Watson processes starting with N individuals. If X_1, \dots, X_N are the offspring numbers in the initial generation with $\mathbb{E}(X_1) = \mu$, then assuming $\mu > 1$ we expect the total offspring number $Y = X_1 + \dots + X_N$ to be much larger than N , with a high probability.

The offspring numbers v_1, \dots, v_N of the corresponding fixed-size model are defined in a two-step procedure. The first step yields the Galton–Watson offspring numbers X_1, \dots, X_N . The second step defines v_i as the number of survivors among X_i siblings after N individuals have been sampled randomly to survive from the Y offspring produced at the first step.

Box 7.1 Non-negligible sibships

The more mathematical reader might like to know what happens if the restrictions (7.15) and (7.18) are removed.

Theorem 7.3 Assume Equation (7.23). If for all $0 < x < 1$

$$N\sigma_N^{-2}\mathbf{P}(v_{(1)} > Nx) \rightarrow \int_x^1 y^{-2} dF(y), \quad N \rightarrow \infty, \quad (a)$$

F being an arbitrary probability distribution function on the unit interval not concentrated at zero, the weak convergence (7.21) holds toward a coalescent with multiple mergers.

Given (a), the limit process R_t has the transition rates

$$q_{ij} = \binom{i}{j-1} \int_0^1 x^{i-j-1}(1-x)^{j-1} dF(x) \quad (b)$$

and

$$\lambda_i = \int_0^1 \frac{1 - (1-x)^i - ix(1-x)^{i-1}}{x^2} dF(x). \quad (c)$$

In particular, if $F(x)$ is concentrated at zero, the limit process R_t corresponds to the Kingman coalescent with the transition rates given by Equations (7.10) and (7.11).

Furthermore, the expectation that the intensity of merger slows down as the number of branches in the gene tree decreases turns out to be correct, $\lambda_n > \lambda_{n-1} > \dots > \lambda_2 = 1$, and $\lambda_i \leq \binom{i}{2}$. Thus, the Kingman coalescent brings about the fastest merger for any given number of ancestral lines.

In the special case of the so-called *Beta-coalescent*, with $F(x)$ being a Beta(a, b) distribution function, the key quantities q_{ij} and λ_i are easy to compute using simple recursion relations that involve parameters $a > 0$ and $b > 0$. As an example of a coalescent pattern dramatically different from the Kingman coalescent, consider the Beta(1, 1)-coalescent with $F(x) = x$ that corresponds to the uniform distribution over the (0,1) interval (Bolthausen and Sznitman 1998). In this case, we have a linear relationship $\lambda_i = i - 1$ instead of the quadratic (7.11).

Theorem 7.3 suggests a simple way to model genealogies with rare multiple mergers based on the Beta($a, 1$)-coalescent with a small positive parameter $a \approx 0$: if $F(x)$ has density $f(x) = ax^{a-1}$, the transition rates become

$$q_{i(i-1)} \approx \binom{i}{2} \left(1 - \sum_{k=1}^{i-2} \frac{a}{k} \right), \quad q_{i(i-j)} \approx \binom{i}{2} \frac{2a}{(j-1)j(j+1)} \quad \text{for } j \geq 2, \quad (d)$$

so that $\lambda_i \approx \binom{i}{2}(1 - a\delta)$, where $\delta = \sum_{k=1}^{i-2} \frac{1}{k} - \frac{1}{2} + \frac{1}{i(i-1)}$.

For a thorough mathematical treatment of the coalescent with the transition rates as in Equation (b) consult Pitman (1999). A full classification of the coalescent processes that arise in exchangeable populations is discussed in Möhle and Sagitov (2001), which covers the possibility of simultaneous mergers of ancestral lines that correspond to a non-zero limit in Equation (7.23). The coalescent with simultaneous mergers is further analyzed by Schweinsberg (2000).

According to Theorem 4 in Schweinsberg (2003), if $\mu > 1$ and

$$P(X_1 > k) \sim Ck^{-a}, \quad k \rightarrow \infty \quad (7.24)$$

for some finite constants $C > 0$ and $a > 0$, then the weak convergence (7.21) holds with the coalescent limit R_t , depending on the parameter value a :

- If $a \geq 2$, the limit R_t is the Kingman coalescent.
- If $1 \leq a < 2$, the limit is the Beta($2 - a, a$)-coalescent and, in particular, if $a = 1$, it is the Bolthausen–Sznitman coalescent already mentioned. If $1 < a < 2$, the timescale N/σ_N^2 is proportional to N^{a-1} , and $N/\sigma_N^2 \sim \ln N$ in the case $a = 1$ (see Box 7.1).
- If $0 < a < 1$, the limit process belongs to a certain one-parameter class of coalescent processes with simultaneous multiple mergers.

7.2 Ancestral Inference in Branching Processes

S. Tavaré

7.2.1 Introduction

The topic of inference for branching processes is classic and many articles and books have been devoted to it. Common themes include estimation of the offspring mean, the offspring distribution, and the age of the process (cf. Stigler 1970; Guttorp 1991, 1995). In this subsection we illustrate some computational approaches to ancestral inference for branching processes when the effects of mutations among individuals in the population are taken into account. Our examples are from population genetics (in which the timescale is of the order of thousands of years) and from cancer biology (in which the timescale is of the order of years). The techniques illustrated here are but the tip of the inferential iceberg, but they serve to illustrate the crucial interplay between the simulation of a stochastic model and any inference about its parameters.

7.2.2 Inference in the coalescent

Coalescent trees. In Section 7.1 the *coalescent* was introduced as a model for ancestral relationships among a set of chromosomal segments sampled from an evolving population. In the case of a population that has a constant but large number N of chromosomal segments, we showed that when time is measured in units of N generations, the coalescent tree of a sample of n segments can be described as follows. We begin with n tips and wait for an amount of time T_n that has an exponential distribution with mean $2/n(n-1)$ time units before choosing at random two of the tips to coalesce. The coalescent tree now has $n-1$ nodes (which corresponds to $n-1$ ancestors of the sample), and we then wait a further time T_{n-1} that has an exponential distribution with mean $2/(n-1)(n-2)$ time units until, once again, choosing at random two of the nodes to coalesce. We can continue this description using mutually independent exponential random variables, the waiting time while there are j ancestors of the sample having a mean of $2/j(j-1)$ time units. Eventually, the segments in the sample can be traced back to a common

References

References in the book in which this section is published are integrated in a single list, which appears on pp. 295–305. For the purpose of this reprint, references cited in the section have been assembled below.

- Aldous D (1999). Deterministic and stochastic models for coalescence (aggregation, coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5**:3–48
- Balding D, Bishop M & Cannings C (2001) *Handbook of Statistical Genetics*. Chichester, UK: John Wiley & Sons
- Billingsley P (1968). *Convergence of Probability Measures*. New York, NY, USA: John Wiley & Sons
- Bolthausen E & Sznitman AS (1998). On Ruelle’s probability cascades and an abstract cavity method. *Communications in Mathematical Physics* **197**:247–276
- Cannings C (1974). The latent roots of certain Markov chains arising in genetics: A new approach. I. Haploid models. *Advances in Applied Probability* **6**:260–290
- Donnelly P & Tavaré S (1995). Coalescents and genealogical structure under neutrality. *Annual Review of Genetics* **29**:401–423
- Ewens WJ (1979). *Mathematical Population Genetics*. Berlin–Heidelberg–New York: Springer-Verlag
- Fisher R (1930). *The Genetical Theory of Natural Selection*. (Second Edition 1958). New York, NY, USA: Dover
- Hudson RR (1991). Gene genealogies and the coalescent process. In *Oxford Surveys in Evolutionary Biology*, eds. Futuyama D & Antonovics J, Vol. 7, pp. 1–44. Oxford, UK: Oxford University Press
- Jagers P (1991). The growth and stabilization of populations. *Statistical Science* **6**:268–283
- Jagers P, Nerman O & Taib Z (1992). When did Joe’s great-grandfather live? Or: On the time-scale of evolution. In *Selected Proceedings of the Sheffield Symposium on Applied Probability*, eds. Basawa I & Taylor RL, pp. 118–126. Monograph Series, IMS Lecture Notes, Volume 18. Beechwood, OH, USA: Institute of Mathematical Studies
- Kingman JFC (1982a). On the genealogy of large populations. *Journal of Applied Probability* **19A**:27–43
- Kingman JFC (1982b). Exchangeability and the evolution of large populations. In *Exchangeability in Probability and Statistics*, eds. Koch G & Spizzichino F, pp. 97–112. Dordrecht, Netherlands: North-Holland
- Möhle M (2000). Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Advances in Applied Probability* **32**:1547–1562
- Möhle M & Sagitov S (1999). Coalescent patterns in exchangeable diploid population models. *Journal of Mathematical Biology* **47**:337–352
- Möhle M & Sagitov S (2001). A classification of coalescent processes for haploid exchangeable population models. *Annals of Probability* **29**:983–993
- Nilsson M, Rasmussen S, Mayer B & Whitten D (2000). Constructive molecular dynamics lattice gases: 3-D molecular self-assembly. In *New Constructions in Cellular Automata*, eds. Griffeath D & Moore C, pp. 275–290. Oxford, UK: Oxford University Press
- Pitman J (1999). Coalescents with multiple collisions. *Annals of Probability* **27**:1870–1902
- Rosenberg NA & Nordborg M (2002). Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics* **3**:380–390
- Sagitov S (1999). The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability* **36**:1116–1125

- Sagitov S (2003). Convergence to the coalescent with simultaneous multiple mergers. *Journal of Applied Probability* **40**:839–854
- Schweinsberg J (2000). Coalescents with simultaneous multiple collisions. *Electronic Journal of Probability* **5**:1–50
- Schweinsberg J (2003). Coalescent processes obtained from supercritical Galton–Watson processes. *Stochastic Processes and their Applications* **106**:107–139
- Taib Z (1992). *Branching Processes and Neutral Evolution*. Lecture Notes in Biomathematics **93**. Berlin, Germany: Springer-Verlag
- Tavaré S (1984). Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoretical Population Biology* **26**:119–164
- Watterson GA (1975). On the number of segregating sites in general models without recombination. *Theoretical Population Biology* **7**:256–276
- Wright S (1931). Evolution in Mendelian populations. *Genetics* **16**:97–159