Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

Edited by James F. Crow and William F. Dove

The Limits of Theoretical Population Genetics

John Wakeley[1]

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138

THE purpose here is to discuss the limits of theoretical population genetics. This 100-year-old field now sits close to the heart of modern biology. Theoretical population genetics is the framework for studies of human history (Reich *et al.* 2002) and the foundation for association studies, which aim to map the genes that cause human disease (Jorde 1995). Arguably of more importance, theoretical population genetics underlies our knowledge of within-species variation across the globe and for all kinds of life. In light of its many incarnations and befitting its ties to evolutionary biology, the limits of theoretical population genetics are recognized to be changing over time, with a number of new paths to follow. Stepping into this future, it will be important to develop new approximations that reflect new data and not to let well-accepted models diminish the possibilities.

It is valuable to define this field narrowly. Theoretical population genetics is the mathematical study of the dynamics of genetic variation within species. Its main purpose is to understand the ways in which the forces of mutation, natural selection, random genetic drift, and population structure interact to produce and maintain the complex patterns of genetic variation that are readily observed among individuals within a species. A tremendous amount is known about the workings of organisms in their environments and about interactions among species. Ideally, with constant reference to these facts—the bulk of which are undoubtedly yet to be discovered—theoretical population genetics begins by distilling everything into a workable mathematical model of genetic transmission within a species.

Taking this narrow view precludes the application of theoretical population genetics to studies of long-term evolutionary phenomena. This, instead, is the purview of evolutionary theory. For theoretical population genetics, processes over longer time scales are of interest only insofar as they directly affect observable patterns of variation within species. The focus on current genetic variation came to the fore during the 1970s and 1980s with the development of coalescent theory (Kingman 1982, 2000), or the mathematics of gene genealogies. Ewens (1990) reviews this transition from the forward-time approach of classical population genetics to the new, backward-time approach. It can be seen both in classical work (Fisher 1922; Wright 1931) and in coalescent theory (Kingman 1982; Hudson 1983; Tajima 1983), both of which are considered below, that the time frame over which the models of theoretical population genetics apply within a given species is a small multiple of N_{total} generations, where N_{total} is the total population size, or the count of all the individuals of the species. Looking at gene genealogies in humans, for example, it seems that this means roughly from 10^4 to 10^6 years (Harris and Hey 1999).

This allows us to suppose that the parameters affecting the species that we wish to model have remained relatively constant over time, compared to the situation in evolutionary theory. For purposes of discussion, consider the following simple model which, with embellishments, might serve to describe any species from *Homo sapiens* to *Bacillus subtilis*. The species is divided into D subunits, each of size N, so that the total population size is $N_{\text{total}} = ND$. Corresponding to the phenomena listed above, the other parameters of the model are the per-locus, per-generation probability of mutation u, the selective advantage or disadvantage, s, of some type relative to some other type in the population, and a parameter, m, which determines the extent of population structure.

The subunits in the model are used below to represent D diploid individuals, so that $N = 2$ is the number of copies of each chromosome within each individual. Note that this departs from the usual notation, in which N is the number of diploid individuals. The reason for

[1]*Address for correspondence:* Department of Organismic and Evolutionary Biology, 2102 Biological Laboratories, 16 Divinity Ave., Cambridge, MA 02138. E-mail: wakeley@fas.harvard.edu

this departure is to emphasize the similarities between the diploid model and other models of population structure. Thus, the same model is used to represent a population subdivided into D local populations, or *demes* (Gilmour and Gregor 1939), each containing N individual organisms.

Many details have been ignored in this model for the sake of simplicity. For example, mutation is a complex process, which includes various kinds of recombination, and natural selection is similarly not likely to be so simple that a single parameter captures all of its intricacies. In addition, the general term "population structure" encompasses dioecy, ploidy level, age structure, reproductive patterns such as partial selfing, as well as the various forms of geographical structure and dispersal. Finally, as noted above, all parameters are assumed to not change over time. However, with some flexibility in the interpretations of parameters, this model can be used to illustrate the limits of theoretical population genetics.

The ranges of the parameters are restricted by nature. Specifically, D and N are whole numbers, both of which it is natural to assume are ≥ 1. The other parameters can vary continuously, but also have natural ranges: $0 \leq u \leq 1$, $s \geq -1$, and $0 \leq m \leq 1$. The last two require some context. Let m be the fraction of each subunit (of which there are D) that is replaced by offspring randomly sampled from the entire population each generation. This is the island model of population subdivision and migration introduced by Wright (1931), but it can be used to represent other forms of structure as well. Subdivision is at its least when $m = 1$ and is at its most when $m = 0$. Selection is imagined between two types, one with fitness 1 and the other with fitness $1 + s$, and $s \geq -1$ precludes negative fitness values. With selection among more than two types, the fitness of one of them is taken to be equal to one and this establishes the relative selection coefficients (values of s) of the others.

The current and historical boundaries of theoretical population genetics can be understood with reference to the object of study, which is genetic variation within species, but also in terms of methodology. The ridiculously oversimplified model just described already has five parameters. Even with the restrictions above, there is an enormous five-dimensional space that defines all possible kinds of species under the model: $\{(D, N, u, s, m);\ D \geq 1,\ N \geq 1,\ 0 \leq u \leq 1,\ s \geq -1,\ 0 \leq m \leq 1\}$. Theoretical population geneticists obtain predictive equations by simplifying such complicated models, again ideally with close attention to the biological relevance of any assumptions made. Formally, this is done by taking mathematical limits. The hope is that by doing so, *i.e.*, by further restricting the ranges of parameters, tractable analytical results or simple approximations to the model can be obtained, which will be both useful and illuminating.

The first limiting result was established independently by Hardy (1908) and Weinberg (1908) for the case of two alleles, A and a, with frequencies p and $q = 1 - p$, respectively, in a population of diploid, monoecious organisms; see Crow (1988) for a perspective on this important result. In this case, the subunits in the model represent the organisms ($N = 2$), the population is supposed to be infinite ($D = \infty$), without mutation ($u = 0$) or selection ($s = 0$), and offspring are formed by either random mating or random union of gametes ($m = 1$). Then, the Hardy-Weinberg law states that the frequencies of the genotypes AA, Aa, and aa will be equal to p^2, $2pq$, and q^2 after a single generation, regardless of the initial genotype frequencies, and that they will remain in these frequencies forever. Provine (1971) discusses the important historical role of the Hardy-Weinberg law in evolutionary biology, which was to show that the mechanism of inheritance would not itself cause the variation upon which selection acts to be depleted in a population.

The simplicity of the Hardy-Weinberg law is a consequence of its very stringent assumptions. It exists only in the special case in which the values of all parameters are fixed and given by ($D = \infty$, $N = 2$, $u = 0$, $s = 0$, $m = 1$). Fisher (*e.g.*, 1930) and Haldane (*e.g.*, 1932), and a great number of workers who followed their lead were content with the assumption of infinite population size. They sought to establish the dynamics of allele frequencies in an expanded Hardy-Weinberg population that included mutation and selection. As a result, much of classical population genetics takes place in the restricted parameter space where $\{(D, N, u, s, m);\ D = \infty,\ N = 2,\ 0 \leq u \leq 1,\ s \geq -1,\ m = 1\}$. However, the overwhelming majority of results have been derived under the additional assumption that u and s are small.

Although every population is finite, so that $D = \infty$ can never be true, these classical predictions are valuable because they establish tendencies at work in populations of any size (*e.g.*, the frequency of a favored allele will increase over time). Further, these classical predictions should be close to true if the population is "large enough." Of course, it is only by considering a finite population that these two statements can be investigated and verified. In addition, some vital phenomena simply cannot be studied using an infinite population model. Questions concerning the fixation or loss of alleles from the population or, more generally, questions about the behavior of alleles in low copy number are outside the boundaries of classical, infinite-population-size theory.

No population is so large that finite size can be ignored as a factor contributing to patterns of genetic variation within a species. For example, in an infinite population with mutation but no selection, every possible allelic type will be present at the frequency determined by the pattern and rate of mutation. However, even a stretch of 100 nucleotides has $4^{100} \approx 10^{60}$ possible alleles, and no population comes even remotely close

to being this large. Considered further, the consequences of reproduction in finite populations are rather amazing. First of all, without mutation (and assuming at least some mixing: $m > 0$), all variation will eventually be lost from any population. More subtly, reproduction with any reasonable fidelity, which is assured by universally small rates of mutation (Drake *et al.* 1998), causes identical or related alleles to accumulate in the population even as they are all ultimately ephemeral (Watterson 1976).

Random genetic drift is the term used to describe the stochastic effects of reproduction in a finite population. Historically, the need to incorporate random genetic drift into population genetic models was motivated by observations of J. T. Gulick and others concerning geographic variation within species without apparent selective causes—see Provine (1986) for a thorough compilation of the history—and by the trenchant argument of Hagedoorn and Hagedoorn (1921), which demonstrated the need to understand the random effects of reproduction in finite populations. The result was the Wright-Fisher model of random genetic drift.

To be concrete, consider the population model as it was used above to illustrate the Hardy-Weinberg law in an infinite population of diploid organisms, but eliminate the assumption of infinite population size. This leaves $\{(D, N, u, s, m); D \geq 1, N = 2, 0 \leq u \leq 1, s \geq 1, m = 1\}$ for the parameter space. The total population size is $N_{\text{total}} = ND = 2D$ and is finite. The Wright-Fisher model of random genetic drift states that the D diploid individuals that form generation $t + 1$ are obtained by randomly sampling pairs of gametes, with replacement, from the adults of generation t. Generations are nonoverlapping, so all adults die and are replaced by offspring. If there are currently i copies of allele A among gametes, then the frequency of allele A now is $p = i/(2D)$, and the probability P_{ij} that there are $j = 0, 1, \ldots, 2D$ copies of allele A at the beginning of the next generation is given by the familiar binomial distribution with parameters $2D$ and $p = i/(2D)$.

Fisher used the above model of genetic drift implicitly, in many cases assuming a Poisson distribution of offspring number with the mean equal to one per individual, which is the large-D approximation to the above binomial distribution with $i = 1$. Wright used the model explicitly, as a null model for the dynamics of a randomly mating population of finite size. Fisher and Wright showed, among other things, that the rate of loss of heterozygosity in a population is equal to $1/N_{\text{total}} = 1/(2D)$. This illustrates the statement above that the time scale over which theoretical population genetics considers things is a small multiple of N_{total} generations.

The Wright-Fisher model of random genetic drift is a discrete time, discrete allele-frequency model. Time is measured in numbers of generations and P_{ij} describes changes in the numbers of alleles. This model is surprisingly difficult to analyze, and few exact results are available. Early on, Fisher (1922) and Wright (1931) considered a continuous time, continuous allele-frequency approximation to the model, which allowed many results of biological interest to be derived. Their results relied on a diffusion approximation to the discrete model (Komolgorov 1931). Malécot (1944, 1946) used the same ideas and rigorous methods to greatly extend the application of diffusion results in population genetics. Feller (1951) provided the general mathematical framework for these models, and Kimura (1955a,b) obtained the full solution of the time-dependent distribution of allele frequencies in a population.

The transition from the discrete model to the continuous one happens in the limit as the population size tends to infinity, but it relies on very different assumptions about the other parameters than are made in classical deterministic work. This model, which is often called *the* diffusion limit of population genetics, exists in the limit as D tends to infinity and assumes that $\lim_{D\to\infty} 4Du = \theta$ and $\lim_{D\to\infty} 4Ds = \sigma$ are finite. Time is rescaled so that it is measured in units of $N_{\text{total}} = ND = 2D$ generations. The continuous model holds in the limit because single generations and single copies of alleles represent, respectively, infinitesimal amounts of time on the new time scale and infinitesimal differences in allele frequency. This is the appropriate diffusion approximation when $1/D$, u, and s are all small and do not differ too greatly in magnitude. Finally, the limit is taken with the allele frequency $i/(2D)$ assumed to be fixed (*i.e.*, constant) in the limit as D tends to infinity, which means that for most purposes—but see Bürger and Ewens (1995)—this model is not appropriate when the number of copies of an allele in the population is not large.

Note that the apparent dependence of the parameters u and s on D in the assumptions $\lim_{D\to\infty} 4Du = \theta$ and $\lim_{D\to\infty} 4Ds = \sigma$ is not a statement about biology. The model does not suppose, for example, that if the population doubled in size, the mutation rate and the selection coefficient would drop by one-half. The standard diffusion limit is simply a mathematical approximation to the behavior of a large population in which the probability of mutation and the selection coefficient(s) are small. Like the classical ($D = \infty$) results, it applies in a particular region of the parameter space, one in which $D \to \infty$ but where the effects of random genetic drift are not negligible. Another possible point of confusion is the extra factor of two in the parameters θ and σ relative to the way in which time is rescaled. This practice was inherited from Wright and Fisher, and it simply reflects biologists' great concern for heterozygosity, or polymorphism between a pair of chromosomes.

Nearly all of modern population genetics is based upon this standard diffusion model, although much of the time it is used implicitly. It is the source of the common practice of simplifying expressions obtained from a discrete model by keeping only terms involving u, s, and $1/D$ and throwing out "small" terms like u^2, s^2,

$1/D^2$, u/D, etc. In this case it is clear why $1/D$, u, and s should not differ too greatly in magnitude. For example, if $D = 10^4$ and $u = 10^{-8}$, it does not make sense to ignore terms involving $1/D^2$ but keep terms involving u. Technically, the standard diffusion holds for any σ and θ, as long as these remain finite as D tends to infinity. Thus, the standard diffusion model can be used to model weak selection and mutation by making σ and θ small and to model strong selection and mutation by making σ and θ large. The risk in doing so is that the error of using these results to approximate the results for a finite population may be large unless D is very large (Ethier and Norman 1977).

There are more appropriate approximations than the standard diffusion, even other diffusion approximations, if one needs to model populations that fall into other regions of the parameter space (Feller 1951; Karlin and McGregor 1964). One which is well known and has been fairly well exploited to address questions of fixation probabilities since Fisher (1922) and Haldane (1927) is the branching-process approximation for the number of copies of an allele. The counts of an allele can be approximated by a branching process (without reference to the rest of the population) in the limit as D tends to infinity for fixed values of u and s but where the number of copies of the allele is not large. This complements the classical deterministic model, which makes the same assumptions about D, u, and s, but applies only when the counts of alleles are very large. For a recent example, see Wahl and DeHaan (2004).

Another approximation, the Gaussian diffusion, sits between the standard diffusion model and the classical deterministic results. In a somewhat neglected article—but see Nagylaki (1990) and Gillespie (2001)—Norman (1975) proved that with $s \to 0$ and $u \to 0$, but $Ds \to \infty$ and $Du \to \infty$, the trajectories of allele frequencies would tend strongly to the deterministic predictions but with small deviations. Further, these stochastic deviations in allele frequencies tend to zero as D becomes much larger than $1/s$ and $1/u$. Thus, if D is very much greater than $1/s$ and $1/u$, and the latter are large, the deterministic equations are very nearly correct (as long as the number of copies of each allele is large). Such concerns underlie the use of a stochastic treatment of allele frequencies close to zero or one and a deterministic treatment in the interior, for example, by Kaplan *et al.* (1989) and Gillespie (1991).

Returning to the ubiquity of the standard diffusion approximation, the addition of a single assumption, that the sample size n is constant, so that $n/D \to 0$, as D tends to infinity (roughly: $n \ll D$), yields coalescent theory (Kingman 1982; Hudson 1983; Tajima 1983; Krone and Neuhauser 1997; Neuhauser and Krone 1997). Coalescent theory describes the genetic ancestry of a sample and provides the tools for the analysis of intraspecies molecular data. Nordborg (2001) gives a recent thorough review of this field. Kingman (2000) gives a historical perspective, which includes credit to Gustave Malécot for having the original idea of tracing lineages back to common ancestors; see also Nagylaki (1989). Applications of coalescent theory to the problems of modern biology abound, from the geographic origin of *Plasmodium falciparum* (Joy *et al.* 2003) and the dynamics of HIV within infected individuals (Drummond *et al.* 2002) to the extent of gene flow between recently separated cichlid species (Hey *et al.* 2004).

The standard diffusion approximation has permeated the field so thoroughly that it shapes the way in which workers think about the genetics of populations. There are positive aspects of this. For example, the parameters $\theta = 4Du$ and $\sigma = 4Ds$ capture the important and fascinating fact that even very weak mutation and selection can have a strong effect if the population size is large. This illustrates the potentially important role of the population size in setting the time scale of population genetic change. However, in terms of the general model with parameter space $\{(D, N, u, s, m); D \geq 1, N \geq 1, 0 \leq u \leq 1, s \geq -1, 0 \leq m \leq 1\}$, the standard diffusion model can be viewed only as a model of weak mutation and weak selection. For selection, the term "strong" would best be reserved for cases in which s is either close to -1 or much greater than zero (*e.g.*, $s = 10$), while the typical usage is to say, roughly, that $|\sigma| > 10$ constitutes strong selection.

It is problematic when conclusions drawn from a special case of a general model become normative statements carried over to other situations. Under the assumptions of the standard diffusion model, in which $D \to \infty$ while θ and σ remain fixed, everything depends only on the products Du and Ds. This limiting result is responsible for the notion that it is impossible to estimate D and u, for example, separately and that only θ can be estimated. However, this is simply a consequence of the assumptions of the model, which might be expected to break down in cases outside the region of parameter space in which the standard diffusion is appropriate. For example, it breaks down for very large samples in a coalescent model ($n/D \to x$ as $D \to \infty$), allowing both D and u to be estimated (Wakeley and Takahashi 2003). While it may be true that there is low power to estimate D and u separately, questions about this cannot even be posed within the framework of the standard coalescent.

A parallel set of issues arises in the study of structured populations. The simple model adopted here includes Wright's (1931) island model of population subdivision and migration, which he proposed to help explain nonadaptive differences among different subunits of a species—recall the observations of Gulick—and which became part of his shifting balance theory of evolution (Provine 1986). Wright introduced the diffusion approximation to obtain the equilibrium distribution of allele frequencies on a single island under the assumption of a constant allele frequency among migrants.

Thus, in addition to $\theta = \lim_{N\to\infty} 4Nu$ and $\sigma = \lim_{N\to\infty} 4Ns$, which Wright defined for the single island of N diploid organisms rather than for the total population, the island model has a scaled migration parameter, $M = \lim_{N\to\infty} 4Nm$. The parameter M captures the notion that small amounts of migration over the time scale of N generations can have a very large effect; see also Nagylaki (1980). As with θ and σ, the relevance of the parameter M in the limiting model should not be taken to mean it will be impossible to separately estimate N and m in other cases—see Vitalis and Couvet (2001)—or that the dynamics of every subdivided population depend only on the product Nm.

Wright offered two possible justifications for the assumption of constant allele frequency among migrants: (1) that migrants come from an infinitely large, unstructured population, like the one that gave the Hardy-Weinberg law above, or (2) that migrants come from an infinitely large collection of islands, of which the focal island is a single example. This second possibility is easily represented using the present model. It is obtained by assuming that $D = \infty$, so that allele frequencies in the total population remain constant, as they do under the Hardy-Weinberg law described above. The assumptions of diploidy ($N = 2$) and random mating ($m = 1$) need to be relaxed so the demes can be of any size ($N \geq 1$) and receive migrants at any biologically reasonable rate ($0 \leq m \leq 1$).

Described in this way, it is helpful to think of Wright's infinite-island model as a classical population genetic model for idealized N-ploid organisms (the demes), with complications such as double reduction ignored. Reproduction is a little more complicated than in the classical diploid model—newborn individuals receive a fraction, m, of their gametes from the total parental population's pool of gametes and a fraction, $1 - m$, from a single parent's gamate pool—but these models share many features. It is clear, for example, that the allele frequencies in the total population will remain constant only if there is no selection and no mutation; otherwise they should change according to something like the classical deterministic theory. In addition, the infinite-island model suffers the same restrictions as the classical model: questions about stochastic trajectories of allele frequencies in the total population (*e.g.*, the fixation or loss of alleles) cannot be addressed.

By assuming a fixed, finite number of demes, Maruyama (1970), Latter (1973), and others studied the finite-island model and obtained results for fixation probabilities and other properties of the population. Without making any assumptions about the parameters, the finite island model is represented by the general version of the present model, with parameter space $\{(D, N, u, s, m); D \geq 1, N \geq 1, 0 \leq u \leq 1, s \geq -1, 0 \leq m \leq 1\}$. There are difficulties in analyzing the finite island model, as there are in the case of the Wright-Fisher model of an unstructured finite population. In fact, the difficulties are greater because subdivision, *i.e.*, when $m < 1$, increases the complexity of the system substantially. Because there are more parameters, there are more choices as to how the parameters might be related or restricted in approximations to the model.

The best known of these limits is the one that underlies the structured coalescent process (Notohara 1990; Wilkinson-Herbots 1998). This is the finite-island model with $N \to \infty$ and with parameters scaled as Wright (1931) did originally. This model frames most work on populations structured by migration. It is a model of a relatively small number of very large populations connected by limited migration, with weak mutation and, in nearly all cases, no selection. Another limit, which is to the island model what the standard diffusion is to the unstructured model, is the many-demes limit with weak mutation and selection, and any $m > 0$ and $N \geq 1$ (Wakeley 2003). Allele frequencies in the total population change according to the standard diffusion, but on a time scale that depends on N and m. At the same time, relatively strong migration and drift within demes keeps the collection of demes close to the kind of equilibrium described by Wright (1931), which is the analog in this model of Hardy-Weinberg genotype frequencies in the diploid model. It hardly needs stating at this point that neither the finite-D, $N \to \infty$ diffusion nor this finite-N, $D \to \infty$ diffusion should be applied or accepted without attention to its restrictions.

Why all this attention to the arcane subject of diffusion theory, which may seem to have peaked with Kimura's work in the 1950s? Possibly the most exciting new direction in theoretical population genetics is the study of a coupled (backward and forward) process that promises to unite diffusion theory and coalescent theory, while fully incorporating natural selection into the latter. This relates population genetic models to bodies of more abstract mathematics, such as the theory of interacting particle systems (Liggett 1985). The approach was introduced into population genetics by Donnelly (1984), developed further by Krone and Neuhauser (1997), and can also be seen in Darden *et al.* (1989). Recent articles include Donnelly and Kurtz (1999) and Barton *et al.* (2004). The challenge is to develop from this work a set of tools for making inferences from genetic data that can be applied in the way that the standard coalescent is being applied now.

Due to recent developments in biotechnology, the theory and methodology of population genetics are lagging behind the collection of data. The abundance of data now available, and soon to be available, holds the promise that it will finally be possible to infer the current and historical characteristics of populations with a high degree of precision. There is already a huge store of results in the historical literature of theoretical population genetics, which can be mined for present-day aims. However, at least since the introduction of coalescent theory 20 years ago, theoretical population genetics has

developed closely in response to newly available data, and now is the time to push the boundaries of the field.

A number of new limits are just over the horizon. For example, high-throughput genotyping techniques have increased sample sizes in two directions: the number of individuals and the number of base pairs per individual. Simplifications of complex models may arise in the limit as the number of individuals sampled tends to infinity or as the length of sequence per sample tends to infinity. If history is any guide, then looking back in a few years it will be apparent how new directions such as these will have shaped the way in which we think about patterns of genetic variation and the processes that conspire to maintain them.

This work was supported by a Presidential Early Career Award for Scientists and Engineers (DEB-0133760) from the National Science Foundation.

LITERATURE CITED

Barton, N. H., A. M. Etheridge and A. K. Sturm, 2004 Coalescence in a random background. Ann. Appl. Prob. **14:** 754–785.

Bürger, R., and W. J. Ewens, 1995 Fixation probabilities of additive alleles in diploid populations. J. Math. Biol. **33:** 557–575.

Crow, J. F., 1988 Eighty years ago: the beginnings of population genetics. Genetics **119:** 473–476.

Darden, T., N. L. Kaplan and R. R. Hudson, 1989 A numerical method for calculating moments of coalescent times in finite populations with selection. J. Math. Biol. **27:** 355–368.

Donnelly, P., 1984 The transient behaviour of the Moran model in population genetics. Math. Proc. Camb. Phil. Soc. **95:** 349–358.

Donnelly, P., and T. G. Kurtz, 1999 Genealogical models for Fleming-Viot models with selection and recombination. Ann. Appl. Prob. **9:** 1091–1148.

Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. Genetics **148:** 1667–1686.

Drummond, A. J., G. K. Nichols, A. J. Rodrigo and W. Solomon, 2002 Estimating mutation parameters, population history, and genealogies simultaneously using temporally spaced sequence data. Genetics **161:** 1307–1322.

Ethier, S. N., and M. F. Norman, 1977 An error estimate of the diffusion approximation to the Wright-Fisher model. Proc. Natl. Acad. Sci. USA **74:** 5096–5098.

Ewens, W. J., 1990 Population genetics theory—the past and the future, pp. 177–227 in *Mathematical and Statistical Developments of Evolutionary Theory*, edited by S. Lessard. Kluwer Academic, Amsterdam.

Feller, W., 1951 Diffusion processes in genetics, pp. 227–246 in *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman. University of California Press, Berkeley, CA.

Fisher, R. A., 1922 On the dominance ratio. Proc. R. Soc. Edin. **42:** 321–341.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Gillespie, J. H., 1991 *The Causes of Molecular Evolution*. Oxford University Press, New York.

Gillespie, J. H., 2001 Is the population size of a species relevant to its evolution? Evolution **55:** 2161–2169.

Gilmour, J. S. L., and J. W. Gregor, 1939 Demes: a suggested new terminology. Nature **144:** 333.

Hagedoorn, A. L., and A. C. Hagedoorn, 1921 *The Relative Value of the Processes Causing Evolution*. Martinus Nijho, The Hague.

Haldane, J. B. S., 1927 The mathematical theory of natural and artificial selection. Proc. Camb. Phil. Soc. **23:** 838–844.

Haldane, J. B. S., 1932 *The Causes of Natural Selection*. Longmans Green, London.

Hardy, G. H., 1908 Mendelian proportions in a mixed population. Science **18:** 49–50.

Harris, E. E., and J. Hey, 1999 X chromosome evidence for ancient human histories. Proc. Natl. Acad. Sci. USA **96:** 3320–3324.

Hey, J., Y.-J. Won, A. Sivasundar, R. Nielsen and J. A. Markert, 2004 Using nuclear haplotypes with microsatellites to study gene flow between recently separated cichlid species. Mol. Ecol. **13:** 909–919.

Hudson, R. R., 1983 Testing the constant-rate neutral allele model with protein sequence data. Evolution **37:** 203–217.

Jorde, L. B., 1995 Linkage disequilibrium as a gene mapping tool. Am. J. Hum. Genet. **56:** 11–14.

Joy, D. A., X. R. Fung, J. B. Mu, K. Chotinavich, A. U. Krettli *et al.*, 2003 Early origin and recent expansion of *Plasmodium falciparum*. Science **300:** 318–321.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics **123:** 887–899.

Karlin, S., and J. McGregor, 1964 On some stochastic models in genetics, pp. 245–271 in *Stochastic Models in Medicine and Biology*, edited by J. Gurland. University of Wisconsin Press, Madison, WI.

Kimura, M., 1955a Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. USA **41:** 144–150.

Kimura, M., 1955b Stochastic processes and the distribution of gene frequencies under natural selection. Cold Spring Harbor Symp. Quant. Biol. **20:** 33–53.

Kingman, J. F. C., 1982 The coalescent. Stochastic Process. Appl. **13:** 235–248.

Kingman, J. F. C., 2000 Origins of the coalescent: 1974–1982. Genetics **156:** 1461–1463.

Komolgorov, A., 1931 Über die analttischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. **104:** 415–458.

Krone, S. M., and C. Neuhauser, 1997 Ancestral processes with selection. Theor. Popul. Biol. **51:** 210–237.

Latter, B. D. H., 1973 The island model of population differentiation: a general solution. Genetics **73:** 147–157.

Liggett, T. M., 1985 *Interacting Particle Systems*. Springer-Verlag, New York.

Malécot, G., 1944 Sur un problème de probabilités en chaine pue pose la génétique. C. R. Acad. Sci. Paris **219:** 379–381.

Malécot, G., 1946 La diusion des gènes dans une population Mendélienne. C. R. Acad. Sci. Paris **221:** 340–342.

Maruyama, T., 1970 On the fixation probability of mutant genes in a subdivided population. Genet. Res. Camb. **15:** 221–225.

Nagylaki, T., 1980 The strong-migration limit in geographically structured populations. J. Math. Biol. **9:** 101–114.

Nagylaki, T., 1989 Gustave malécot and the transition from classical to modern population genetics. Genetics **122:** 253–268.

Nagylaki, T., 1990 Models and approximations for random genetic drift. Theor. Popul. Biol. **37:** 192–212.

Neuhauser, C., and S. M. Krone, 1997 The genealogy of samples in models with selection. Genetics **145:** 519–534.

Nordborg, M., 2001 Coalescent theory, pp. 179–212 in *Handbook of Statistical Genetics*, edited by D. J. Balding, M. J. Bishop and C. Cannings. John Wiley & Sons, Chichester, UK.

Norman, M. F., 1975 Approximation of stochastic processes by Gaussian diffusions, and applications to Wright-Fisher genetic models. SIAM J. Appl. Math. **29:** 225–242.

Notohara, M., 1990 The coalescent and the genealogical process in geographically structured population. J. Math. Biol. **29:** 59–75.

Provine, W. B., 1971 *The Origins of Theoretical Population Genetics*. University of Chicago Press, Chicago.

Provine, W. B., 1986 *Sewall Wright and Evolutionary Biology*. University of Chicago Press, Chicago.

Reich, D. E., S. F. Schaffner, M. J. Daly, G. McVean, J. C. Mullikin *et al.*, 2002 Human genome sequence variation and the influence of gene history, mutation and recombination. Nat. Genet. **32:** 135–142.

Tajima, F., 1983 Evolutionary relationship of DNA sequences in finite populations. Genetics **105:** 437–460.

Vitalis, R., and D. Couvet, 2001 Estimation of effective population size and migration rate from one- and two-locus identity measures. Genetics **157:** 911–925.

Wahl, L. M., and C. S. DeHaan, 2004 Fixation probability favors increased fecundity over reduced generation time. Genetics **168:** 1009–1018.

Wakeley, J., 2003 Polymorphism and divergence for island model species. Genetics **163:** 411–420.

Wakeley, J., and T. Takahashi, 2003 Gene genealogies when the

sample size exceeds the effective size of the population. Mol. Biol. Evol. **20:** 208–213.

Watterson, G. A., 1976 The stationary distribution of the infinitely many neutral alleles diffusion model. J. Appl. Prob. **13:** 639–651.

Weinberg, W., 1908 Über Vererbungsgesetze beim Menschen. Jaresh. Verein. f. vaterl. Naturk. Würtem. **64:** 368–382 (translations: 1963, *Papers on Human Genetics*, pp. 4–15, Prentice-Hall, Englewood Clis, NJ; 1977, *Evolutionary Genetics*, pp. 115–125, Dowden, Hutchinson, and Ross, Stroudsburg, PA).

Wilkinson-Herbots, H. M., 1998 Genealogy and subpopulation differentiation under various models of population structure. J. Math. Biol. **37:** 535–585.

Wright, S., 1931 Evolution in Mendelian populations. Genetics **16:** 97–159.