

January 20, 1997

A Reduction of "Species"  
Resolves the Species Problem

Jody Hey  
Department of Biological Sciences  
Rutgers University  
Nelson Laboratories  
P.O. Box 1059  
Piscataway, NJ 08855-1059  
phone 908-445-5272  
fax 908-445-5870  
internet hey@mbcl.rutgers.edu

*Abstract* - The species problem is the persistent biological and philosophical debate on the meaning of the word "species" and the methods of species identification. With a meaning of "species" that follows from a simple model of DNA replication, species are shown to be real and non-arbitrary groups of organisms, under some circumstances. However, it also follows that many organisms do not belong to species. The criteria by which a group of organisms can be considered a species is whether they share in a process of genetic drift. This simplification is a negative resolution to the problem cases of species identification; it permits a concise listing of the causes of diversity and of the reasons why species can be very difficult to identify, but it does not simplify the process of species identification. For population biologists, a reduced species concept reveals a research plan for the study of organismic diversity that focuses on the determinants of structure in patterns of genetic drift. The finding that species exist, but that some organisms do not occur in species, reveals the central difficulty of systematic theories that assume the existence of species.

The diversity of life seems to have a pattern whereby organisms fall into a limited number of types. Although the existence of these types, or species, has

long been recognized (Mayr, 1982), the definition of the word "species" and the identification of species remain problematic. One advance is the understanding that the word "species", as often used by biologists, signifies a distinct kind of biological individuality (akin to "organism" or "cell"), and does not simply denote a group of similar organisms (Ghiselin, 1966, 1974; Hull, 1976, 1978). However, some biologists reject species as a distinct kind of individual (Nelson, 1989), and among those who do not, there persists a lack of consensus on the defining properties of this kind of individual (Endler, 1989). Two questions remain much discussed: is it useful to consider species as individuals?; and if so, what is the defining attribute of this individuality - the meaning of the word "species" - that applies to the different species of the world?

The approach taken in this paper does not begin with an assumption that species either are or are not individuals. Rather the approach has been to assume the existence of some simple natural phenomena concerning the nature of DNA replication, and then to consider whether these things will cause species. It is shown that these relatively modest phenomena will create a kind of individual that has a close correspondence with other concepts of the meaning of species.

The concept that is developed (the genetic species) is similar to some elements of the cohesion species concept (Templeton, 1989, 1994). In particular, both species concepts rely extensively on the idea of shared genetic drift. However, the two concepts differ in their motivation and their purpose. Templeton begins his discussion with the question "What is a species?" and the implicit assumption that species as individuals exist in nature. The genetic species concept arises from the basic question of whether organisms actually occur in groups that are individuals (i.e. species).

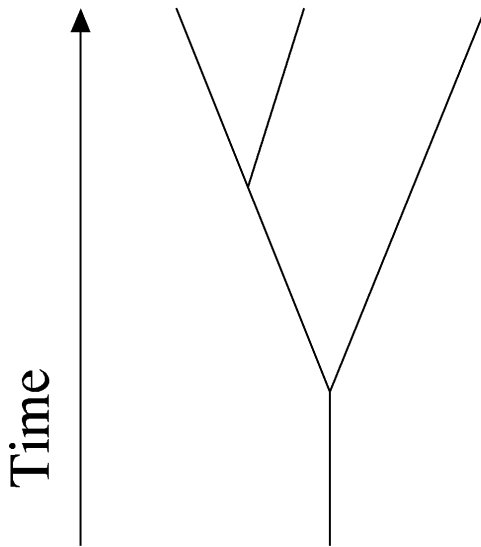
#### A SIMPLE SYSTEM

Consider the reproduction of a single-celled asexual organism. To simplify, focus solely on the replication of the DNA genome, and view the remainder of the organism as the machinery of DNA replication. This simplification follows from the genotype/phenotype relationship: the genotype is the information for the organism and is replicated; while the phenotype is the organism and is recreated each generation as a function of the genotype. The focus is on the transmission of information, and much of this discussion should apply in principle to any informational replicating system (Orgel, 1992).

Hereafter, "a DNA" will be used to refer explicitly to the molecule that contains the genotype information and that is replicated in this simple system. This choice of term is motivated by the same

reasoning that lead Dawkins to employ “replicator”, which he defined as an “entity that interacts with its world, including other replicators, in such a way that copies of itself are made” (Dawkins, 1978). The word “gene” is avoided

here, and by Dawkins (1978), because of its conventional meaning as a single unit of function in the expression of the phenotype. In this paper, a DNA is a contiguous double stranded molecule that is a replicator. A DNA may be physically connected to a longer contiguous stretch of DNA or it may correspond to a single chromosome, depending on the context. A DNA sequence is a particular order of the component nucleotides within a DNA and is not synonymous with a DNA. The terminology also includes “DNAs” to refer to multiple pieces of homologous DNA. In this



**Figure 1** The gene tree history of a sample of DNAs. The tree is a hypothetical depiction of a true history, not to be taken as an example of a tree estimated from data. Key features include: the directionality of time, from the past to the present; branches; branch tips; and nodes, the junctions of branches. Branch tips refer to different pieces of DNA that exist at the present moment. The remainder of the diagram below the tips is a description of history. The tip of the branch at the base of the tree is undefined because the true history is not known beyond this point. Branches refer precisely to the persistence of a DNA sequence through time. This persistence means at times the physical persistence, but also includes numerous cases of replication when it is the information in the sequence that persists. The nodes of the tree refer precisely to those cases of DNA replication when both daughter sequences that were produced are ancestors of sequences that are represented as tips of branches.

context, “homologous” means that the different DNAs are related by common ancestry and thus share a gene tree history (Fig. 1). A sample of DNAs may include one or multiple sequences.

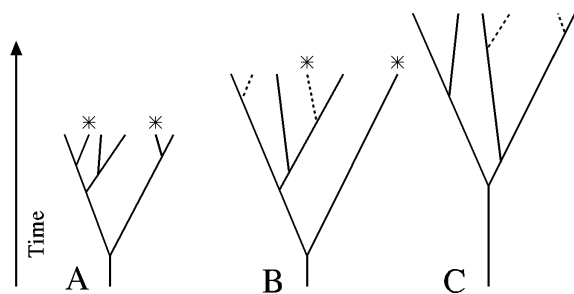
Consider a DNA that undergoes replication to form two daughter DNAs, and suppose that the replication depends upon both the DNA sequence and the local environmental resources. After replication, the fates of the daughter DNAs may be linked because they coexist under common circumstances and compete for the same pool of resources. If resources are limiting and competition occurs so that not all DNAs undergo replication, and if both daughter DNAs and all of their descendants are subject to the same circumstances (i.e. no mutational differences or geographical separations), then the long term persistence of both groups of descendants is mutually exclusive. After some time, perhaps after many rounds of replication, one group of descendants will have replaced the other, or both will have been replaced by the descendants of yet another DNA that also shares those circumstances.

Now consider that DNAs reside within organisms, and that the continuous random replacement of DNAs by the descendants of others is caused by a random birth and death process that happens within a group of organisms that share a finite set of resources. With an allele-based model of genetic variation, the effects of the random birth and death process within a population of organisms include random changes in allele frequencies that lead to the random loss and fixation of mutations. This process is called genetic drift. From a gene tree standpoint, genetic drift is manifested as a randomly shifting pattern of coancestry among a set of DNAs (Fig. 2). Consider the persistence through time of a group of organisms that experience a random process of birth and death, and then consider the gene tree, one DNA per organism, with the tips constantly moving forward with time. The random death of some organisms means that some gene tree tips do not persist, and the branches that lead to these tips disappear from the gene tree history that remains for those DNAs that do persist and replicate (Fig. 2). This shift forward in time of the pattern of ancestry proceeds continuously, and at intervals will include forward jumps for the most basal node representing the ancestor for an entire group of DNAs (Watterson, 1982).

Two kinds of events can cause the descendants of two daughter DNAs to not be mutually exclusive. First, the daughters may differ because of mutation, and this may cause differences in the circumstances of replication. Individuals carrying the mutation may utilize resources in a different way so that they do not compete directly with individuals not carrying the mutation. Second, one daughter DNA and respective

descendants may occur in a geographically distinct location from other DNAs. Under both mutation and geographic separation, the genetic drift experienced by the descendants of one DNA occurs partially independently of that experienced by the descendants of the other. To describe this in another way, the individual DNAs within a group compete more directly with one another, and are more likely to be replaced by the descendants of other DNAs within the same group than by the descendants of DNAs from the other group. In this way both mutation and geographic separation can lead to multiple groups of DNAs that are not mutually exclusive.

The model of replication that leads to multiple groups of DNAs that are not mutually exclusive has three components: a DNA with a sequence that causes replication; the possibility of mutations; and some kind of environmental structure such that the pool of resources used by one group of DNAs need not completely overlap those of another group of DNAs. This simple system probably existed early in the origin



**Figure 2** A gene tree at successive times under genetic drift. Asterisks at branch tips at times A and B indicate sequences that did not persist to the next time period. Solid lines at times B and C indicate branches that were present at the previous time and still remain (and are now longer) in the tree because of the persistence of DNAs at the branch tips. Dotted lines at times B and C indicate new branches leading to DNAs that arose by replication since the previous time period. All solid lines at times B and C correspond to a line (solid or dotted) at the previous time, and all lines at times A and B (solid or dotted) that do not lead to an asterisk, correspond to a solid line at the next time.

of life, though the actual nucleic acid may have been single stranded RNA (Gilbert, 1986). The model is also an approximation, for the multicellular case, of the transition from the reproductive cells of an organism in one generation to the reproductive cells in a descendant organism in the next generation. For a group of multicellular organisms, the appropriate gene tree history to consider is one in which a single DNA has been taken from each organism. In this case many but not all of the instances of DNA replication represented by branches on a gene tree will have occurred during germ line development and somatic growth (to the degree that somatic growth occurs prior to germ line development). The remaining replication events along gene tree branches, and all those replications represented by gene tree nodes, must have occurred within reproductive cells that gave rise to gametes or offspring.

### GENETIC SPECIES

The basic criteria by which something, a group of organisms for example, is considered to be an individual is whether or not the location of components are constrained in space and time (Ghiselin, 1966, 1974; Hull, 1976, 1978). In contrast the members of a class have no spatiotemporal constraints. If a species is a class, then the constituent organisms are members that simply happen to share the particular features by which the class is defined. If a species is an individual then the constituents are components that collectively form a distinct entity (Ghiselin, 1987), and thus has a real existence independent of an observer.

It is proposed that organisms whose DNAs share in a process of genetic drift constitute a kind of biological individual. A process of shared genetic drift among DNAs will occur within a group of organisms that share a birth and death process. A shared "birth and death process" means that at any point in time individual organisms can be physically replaced, and their function in the environment can be replaced, by the descendants of other organisms that are within the group. It also means that organisms compete, so that the survival and reproductive success of one organism has an affect on the survival and reproductive success of other organisms. This type of competition and potential for replacement has also been described as demographic exchangeability (Templeton, 1989) and occurs when organisms share the same niche, in an ecological sense (Hutchinson, 1958). The idea that organisms within a species share a niche is also a component of several other species concepts, including the evolutionary species concept (Simpson, 1961), as well as a recent formulation of the biological species concept (Mayr, 1982 p. 273), and it is contained within the ecological species concept (Van Valen, 1976). A similar view, emphasizing the competition

that occurs when organisms closely share resources, has been described by (Ghiselin, 1974) in an exposition on species as individuals (as opposed to classes). Ghiselin proposed that "species" be defined as "the most extensive units in the natural economy such that reproductive competition occurs among their parts".

"Genetic species" will be used to refer to a group of organisms that share genetic drift. This term has been chosen because "genetic" is often used in reference to the material of inheritance and because genetic species arise from the process of replication, which is the essential function of DNA. "Genetic" has also been used in the context of species concepts in reference to the process of gene exchange (Simpson, 1961; Masters and Spencer, 1989). The present meaning of "genetic species" may be appropriate, despite the limited prior use of "genetic", because replication is a more fundamental function of DNA than is recombination.

Genetic drift has long been recognized as a force that contributes both to uniformity within populations and to variation among separate populations (Gulick, 1872, 1888; Wright, 1931; Templeton, 1989). Genetic drift is proposed not as a mechanism that causes species, but as a description of the process of species existence. Genetic drift is the name of the instantaneous process (like a mathematical derivative taken with respect to time) of a genetic species. In this view, the ultimate cause of species is the reproduction and death of organisms that compete for shared resources. One could also say that the organisms in a genetic species share in a common birth and death process. However, organisms reproduce because of instructions in the genotype, and genetic species existed prior to the existence of complex cellular phenotypes (i.e. organisms).

#### THE ROLE OF SEX

In a gene tree view of the history of a sample of DNAs, sex is synonymous with recombination, and can be defined as any process that causes different portions along the sequence of a set of DNAs to have different gene tree histories. In the absence of sex, the gene tree history of a sample of DNAs is the same for all parts of the sequence. With sex, it is possible that the speed of genetic drift for one portion of a DNA is different from another portion. If there is a high recombination rate, then a large sample of DNAs will have a history of many different gene trees, perhaps as many as there are base pairs in the sequence.

For sexual organisms, a genetic species is the same as a Mendelian population, as defined by Dobzhansky:

*A Mendelian population is a reproductive community of sexual and cross-fertilizing individuals which share in a common gene pool.*

*... The smallest Mendelian populations are panmictic units (Wright, 1943), which are groups of individuals any two of which have equal probability of mating and producing offspring. (Dobzhansky, 1950)*

Thus by definition, organisms within a Mendelian population share in a probabilistic process of reproduction, and all pairs of organisms are equally subject to reproductive failure and equally likely to reproduce. Within a Mendelian population, each generation occurs with some distribution of reproductive success among the component organisms. The shape of this distribution may vary across generations, but at any point in time the particular pattern of reproduction is a major determinant of the gene tree for all portions of the genome. A sample of DNAs for a short region of the genome will have a particular history, while a different genomic region will have a different history; yet all of these histories must run through the same historical procession of organisms, with a different group of reproductives each generation. Thus a Mendelian population carries genomes with numerous gene trees that were all shaped by a common birth and death process.

#### GENETIC DRIFT AND NATURAL SELECTION

From a genetic perspective, natural selection can be defined as variation in reproductive success caused by genotypic variation (Lewontin, 1970), and it is often cast as a directed force of evolutionary change in contrast to the random force of genetic drift. However at the level of DNA where there is linkage, natural selection on functional DNA sequence variation contributes to the genetic drift that occurs among linked sequences. In a genetic species of asexual organisms, a mutation that changes a DNA sequence and causes natural selection, also causes a new pattern of genetic drift among organisms that carry that mutation. In effect, a new genetic species is created by the mutation; although one of the species will probably be replaced by the other. For the DNAs of organisms with recombination, the acceleration of genetic drift by natural selection depends on the degree of linkage, the number of sites of functional variation, and the strength of natural selection on the functional variation (Hill and Robertson, 1966; Felsenstein, 1974).

Natural selection on functional genotypic variation may play a major role in the formation of new genetic species. However, shared genetic drift, and not natural selection, is the appropriate description of the essence of genetic species. Genetic species will share in the process of natural selection on functional DNA sequence variation, and thus will share adaptations. However this process proceeds both concomitantly with, and as a contributor to, genetic drift. Furthermore, from a genealogical perspective

(Fig. 2), genetic drift proceeds even in the absence of DNA sequence variation and in the absence of natural selection caused by DNA sequence variation.

#### THE CAUSES OF MULTIPLE SPECIES

There are two kinds of events that can cause a single genetic species to become two genetic species. First is physical distance or the emergence of a physical barrier between organisms so that they do not draw from the same pool of resources. This geographic barrier to drift may be reversed if organisms are mobile or if geography changes. Second, is the appearance of a mutation that changes the sequence of a DNA so that an organism and its descendants undergo genetic drift separately from other organisms not carrying the mutation. This kind of speciation can be reversed only if all of the descendants of the organism fail to reproduce so that all copies of the new sequence ceased to exist. Back mutation could not reverse the speciation event unless all descendant copies of the mutant DNA underwent back mutation.

The effects of geographic barriers and environmental heterogeneity on genetic drift do not change as a function of sex. If geography constrains the replacement of some individuals by the descendants of others, then it also constrains the process of recombination between some pairs of individuals. This constraint on genetic drift will occur for all portions of the genome regardless of sex.

The kinds of mutations that can create genetic species are different for sexual and asexual organisms. In the absence of sex, a new advantageous mutation causes an organism and its descendants to undergo a different pattern of genetic drift from those organisms not carrying the advantageous mutation. Within a genetic species of sexual organisms, favorable mutations do not contribute directly to the multiplication of species. Regions of the genome under tight linkage to the site of the mutation will experience accelerated genetic drift and have a shortened gene tree history (Kaplan et al., 1989) while the gene trees of unlinked portions of the genome will not be affected. In short, sex prevents favorable mutations from causing one genetic species to split into two - organisms that lack the beneficial mutation are not excluded from the birth and death process that occurs among organisms that carry the mutation. However, there is a class of mutation that, in obligately sexual organisms, can contribute to the formation of new genetic species. These are mutations that cause recombination either not to occur between some individuals or cause the results of recombination to fail to reproduce. Included within this class of events are genomic changes that shift the mode of reproduction or the ploidy level of the genome. For example, interspecific hybrids of sexual organisms may be

polyploid or parthenogenetic. In either case, the progeny of the hybrids can no longer exchange genes with the original species and do not share in a common process of genetic drift with either original species.

The description of the causes of multiple species is intended as a simplification. By posing the discussion in terms of gene tree histories and genetic drift, mutations and barriers to the movement of DNAs appear to be the only possible causes of species diversity. Similarly the irreversible aspect of mutation, as a cause of diversity, contrasts with the reversible effect of barriers, and this distinction follows directly from the genetic drift perspective. However, these mechanisms are not novel ideas, but rather a simple version of the causes of speciation that have been discussed in other contexts. In particular, geographic isolation and the evolution of barriers to sex form the central model of speciation under the biological species concept (Mayr, 1942; Dobzhansky, 1951). Also, a similar depiction of the impact of different kinds of mutations, and their shifting effect as a function of sex, is contained within the cohesion species concept (Templeton, 1989).

#### GENETIC DRIFT AND POPULATION STRUCTURE

For a group of organisms to be a genetic species, genetic drift must be creating a kind of individual, meaning an entity with boundaries in space and time (Hull, 1976). Perhaps the clearest example of a genetic species would be a group of organisms that are completely panmictic (i.e. random mating) amongst themselves without the occurrence of any mating with organisms outside the group. For asexual organisms, the criteria of completely random mating is replaced by one of complete demographic exchangeability (Templeton, 1989). In an asexual group of organisms with complete demographic exchangeability, any one organism's physical place and environmental role could be taken by any other organism. However, complete panmixia or demographic exchangeability need not be present in order for there to be a sharp boundary to the pattern of genetic drift. In general, any time that a group of organisms occurs with gene flow or demographic exchangeability among the organisms, and where the level of gene flow or demographic exchangeability is high relative to the level with organisms outside of the group, a boundary will exist (Templeton, 1989).

It is also possible for there to be variation in the degree to which organisms share genetic drift, and this variation may not occur with sharp boundaries. Consider the case of isolation by distance (Wright, 1943) in which the times of possible coancestry for a given pair of DNAs are proportional to the physical distance between the members of that pair. Under this scenario the pattern of genetic drift, as well as the

pattern of genetic variation, among organisms may not be structured but may follow a continuous pattern over some environmental landscape. A sample of DNAs will still have a gene tree history, but the spacing of the nodes may vary widely and is expected to have a variance larger than expected under a simple demographic model of shared genetic drift (Slatkin, 1987; Hey, 1991).

Another kind of population structure may lead to nested levels of demographic exchangeability or gene flow, with multiple nested boundaries to the pattern of genetic drift. An example is the population structure of *Escherichia coli*. Genetic drift may occur over a short time scale among the cells in a single colony on a petri dish, and over a longer time scale among the population of cells within the intestine of a single mammal. On a larger scale, exchange of *E. coli* cells occurs between different individual intestinal populations, and this leads to turnover of individual intestinal populations

(Hartl and Dykhuizen, 1984). Thus the structure of *E. coli* populations seems to include a hierarchy of levels of genetic drift. Individuality, by the criterion of a boundary in the pattern of genetic drift, occurs at multiple nested levels.

Another pattern not clearly consistent with shared genetic drift can arise when two sexual populations share genetic drift for just a portion of the genome. They may share drift over the entire genome with the exception of a single region that is under natural selection, with different functional forms of the region maintained in different populations. Similarly, genetic drift may be shared for organelle genomes but not nuclear genomes. At the other extreme are two populations that share drift over very little of the genome because of natural selection. It is possible that these populations may still generate hybrids with some reproductive success and share drift at parts of the genome that are not linked to those that are under differential natural selection.

These scenarios of population structure and hybridization illuminate an area of uncertainty for the biological species concept (Mayr, 1963) (ch. 2). This can be seen by considering the very close parallel between Dobzhansky's (1950) concept of a Mendelian population, and the genetic species concept. Dobzhansky defined the smallest Mendelian population as a panmictic unit, and envisioned larger Mendelian populations to be groups of panmictic units that engaged in gene flow. Finally, the largest Mendelian populations are biological species. Dobzhansky envisioned the existence of a boundary, a partition in the magnitude of gene flow such that there was a point when species could be defined. However, the concept of a Mendelian population does not by itself imply the existence of such a boundary.

Dobzhansky's portrayal of nested levels of gene flow, beginning with high gene flow and panmixia at the lower limit, does not necessarily imply the existence of a sharp boundary at the upper limit where gene flow approaches zero. Under the biological species concept, the existence of a sharp boundary to gene flow is attributed to the presence of isolating mechanisms that prevent gene flow. The genetic species concept differs from the biological species concept in not having a necessary role for any process other than genetic drift.

A focus on genetic drift as the essence of species provides a form of negative answer to some aspects of the "species problem". All three of the situations described (complex population structure, isolation by distance, and sexual populations with some genomic partitioning of gene flow) are "problem cases" for which biologists are often at a loss for clear ways to delineate species. A positive resolution of these uncertainties would be, for example, some description of the meaning of "species" that permitted objective resolution of many problem cases. However, when these problem cases are considered with the reduced concept of genetic species, the uncertainty of these situations does not go away. In short, it appears that species are not a necessary consequence of those processes that do often give rise to species. A similar negative resolution was also proposed by Leven (1979) for many of the problem cases that occur among plants, especially those that rarely outcross or have limited gamete dispersal and experience isolation by distance. In these cases, and others where organisms do not occur in groups that share genetic drift, organisms do not occur as parts of species.

In the long term, it is expected that all organisms have histories that include periods when ancestors were part of a genetic species. This is because the causes of genetic species, both mutational and environmental, will sometimes create groups of organisms with periods of uniform genetic drift in which the probability of recent coancestry for any pair of DNAs has little variation. A low variance among pairs of DNAs for possible coancestry times is more likely for a group of DNAs among which genetic drift is proceeding rapidly. Rapid genetic drift may result either from ecological circumstances that sharply curtail reproduction or from the appearance of a strongly favored mutation. Also, with sex, rapid genetic drift can occur for a tightly linked portion of the genome as a result of advantageous mutations or an abundance of deleterious mutations (Maynard Smith and Haigh, 1974; Kaplan et al., 1989; Charlesworth et al., 1993). Thus environmental changes or mutations may create genetic species from groups of organisms that are not in genetic species.

## CONTEMPORANEOUS SPECIES AND HISTORY

One of the major dichotomies that arises in the discourse on species concepts is between contemporaneous, or “snapshot” viewpoints and historical viewpoints (Endler, 1989). A contemporaneous view is most likely to be useful to a population geneticist or ecologist focusing on ongoing population processes. The biological species concept, and the genetic species concept are contemporaneous concepts. However, systematists take a historical view and generally refer to species within the context of ancestor-descendant relationships. For example, (Simpson, 1961) defined a species as an “ancestral-descendant sequence of populations”. Cracraft defines a phylogenetic species as “an irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent” (Cracraft, 1989). One kind of resolution of these two different viewpoints of species (contemporaneous and historical) has been to argue that both are real but distinct kinds of entities (Donoghue, 1985; de Queiroz and Donoghue, 1988). In particular, de Queiroz and Donoghue (1988) argue that there are two processes, interbreeding and common descent, that are both valid species criteria. However, the two criteria differ fundamentally in their incorporation of time. Interbreeding is a process that can be viewed in a short time interval (e.g. a generation). Common descent is an explicitly historical criterion - whether or not a group of organisms have descent in common depends on what ancestors they had. The processes of interbreeding and common descent are distinct merely because one (i.e. common descent) includes the passage of time. If there exists a kind of species that can be defined by historical relationships, then these historical relationships must have arisen because of processes that occurred over time. For example, Simpson’s definition of an evolutionary species explicitly refers to the existence of a population that exists at a particular point in time. Cracraft’s definition of a phylogenetic species refers explicitly to a “parental pattern of ancestry and descent”. Though not defined precisely, this definition clearly implies some contemporaneous process by which groups of parents leave offspring. In both of these cases, Simpson’s and Cracraft’s, the definition of a historical species concept supposes the existence of some kind of cohesive group of organisms that exists for each slice of time in the history of the historical species.

The gap between historical and “snapshot” concepts can be bridged by considering the process that creates a contemporaneous species, and then extending that process through time to generate a picture of a historical species. Indeed much of the literature that forms the debate on the phylogenetic

species concepts includes this very exercise for the case of contemporaneous populations of interbreeding organisms (Hennig, 1979 p.73; de Queiroz and Donoghue, 1988; Nixon and Wheeler, 1990; Davis and Nixon, 1992; Baum and Shaw, 1995). This can also be done for the genetic species concept. The historical extensions of contemporaneous situations, with and without sex, can be considered.

1) For organisms that do not have sex, the common ancestors of a genetic species become fewer, the further one goes back into the past, and this reaches a limit of a single individual. For two genetic species that have recently diverged, there may be many shared ancestral organisms, and these may or may not have existed as a single genetic species. Thus whether or not a historical viewpoint includes ancestral and descendant species depends on how much time is being considered and the history of genetic drift. For an ancestor and descendant that are far apart in time, the ancestor is literally a single organism and could not be a species. Furthermore, this is true whether or not the descendant organisms are a contemporaneous genetic species.

2) For a group of sexual organisms, different portions of the genome will have different gene tree histories. The nodes of these gene trees are ancestral portions of the genome, and the spatial and temporal distribution of the organisms that carried these ancestral DNAs can be considered as a function of the historical pattern of genetic drift. From models of the variance of the coalescent process in the presence of population structure, it is clear that the variance of the time and the geographic location of gene tree nodes can become large, arbitrarily large, depending on the degree to which populations of ancestral organisms depart from panmixia (Slatkin, 1987; Hey, 1991). It is possible for a group of organisms to exist as a contemporaneous genetic species, and yet have a history of ancestors that did not occur in genetic species. This type of history would cause different gene tree estimates, for different portions of the genome, to have a very large variance for the pattern of node spacing (see EMPIRICAL CONSIDERATIONS). It is also possible for a group of organisms that do not exist as a contemporaneous genetic species to have a history of ancestors that did occur as genetic species.

In summary, when the genetic species concept is extended to a historical context, the exercise does not reveal the emergence of a distinct historical entity.

There is no epiphenomenon that could be called a historical species when the history has not included the persistence of contemporaneous genetic species. In the literature on phylogenetic systematics, one graphical tool is to depict historical species as a tube, with time as an axis that runs the length of the tube. In these diagrams, a view of a species or a population at a point in time can be represented as a cross section of a tube (e.g. (Hennig, 1979 p. 59). The point made here is that the histories of organisms need not have a shape that can be represented in this way. In the absence of sex, the more recent ancestors of a group of organisms may not have shared genetic drift. For times longer in the past, the ancestor of a group of asexual organisms, regardless of whether they occurred as part of a genetic species, is a single individual. In the presence of sex, the ancestors of a contemporaneous genetic species may not have been a genetic species. The ancestors may have been spread across a wide expanse of geography, in an isolation by distance relationship, or with a complex structure of multiple populations.

#### EMPIRICAL CONSIDERATIONS

Biologists often face the question of whether a sample of organisms comes from one or more than one species. Under the genetic species concept it is also possible that some or all of the organisms are not part of any species. In practice, assessments of genetic drift are expected to fall into two different categories, instantaneous and recent. An instantaneous determination is an assessment based on ongoing patterns of reproduction in contemporaneous organisms. For obligately sexual organisms, an assessment of genetic species status would be the same as an assessment of Mendelian population status (Dobzhansky, 1950). Those organisms that are exchanging genes are necessarily also sharing in a process of genetic drift. In this case, an assessment of genetic species status is reduced to determining if the pattern of genetic drift has sharp partitions such that there are distinct groups having gene flow within and little gene flow between. For organisms that are not exchanging genes, an instantaneous assessment of drift could be made on ecological and demographic grounds by an assessment of demographic exchangeability. However, this is tantamount to measuring the fundamental niche for each of the organisms in the sample (Templeton, 1989) which is generally impractical. An alternative to instantaneous measures are assessments of recent patterns of genetic drift based on patterns of genetic variation. With electrophoretic data on protein variation or with comparative DNA sequence data, the patterns of variation can be interpreted in terms of the genetic drift that has occurred in the time since the variation arose. However, the genetic species concept is a contemporaneous one, and so an assessment of genetic

drift based on patterns of genetic variation must admit two kinds of uncertainty. First, it is possible that a group of organisms that seem to have a recent history of genetic drift, may not currently share genetic drift. Secondly, a group of organisms may currently share genetic drift, but this may be due to a recent mixture of historically separated organisms that did not occur in a single species. The current situation may be one genetic species, but the recent history which is reflected in the pattern of genetic variation, may include zero or multiple genetic species.

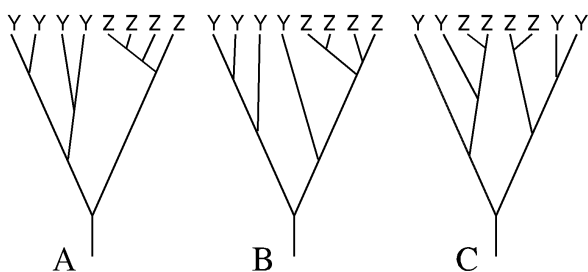
It is not the purpose of this report to develop practical criteria for the identification and delineation of genetic species. Rather the purpose is to show that an assessment of genetic drift is required if species are to be identified and distinguished. In this view, the task of identifying species and understanding the details of the causes of speciation falls squarely within the domain of population genetics. Since the time of (Wright, 1931) much of the field of population genetics has consisted of research on ways to assess genetic drift, and on the effects of genetic drift. The point that a population genetic approach must be used to identify species and understand the causes of species has been repeatedly emphasized by Templeton (1981, 1989, 1994).

It is useful to provide a qualitative description of some criteria that can be used for the case of a set of DNAs, one from each member in a sample of organisms. It is possible to describe the kinds of gene tree histories that can occur for a set of DNAs under different evolutionary models (Hudson, 1990) and it is possible to estimate the true tree for a sample of DNAs if there is some variation in their sequences. Although DNA sequences and gene tree estimates are not the only way to study genetic drift, they are increasingly referred to in the context of the population genetic causes of speciation (Hey, 1994; Templeton, 1994; Baum and Shaw, 1995).

Consider a sample of two sets of homologous DNAs, one from each of several organisms from each of two candidate species and consider the null model that the sample comes from a single species (Templeton, 1994). The gene tree history of the entire sample will, in the absence of recombination, be representable as a bifurcating diagram (Fig. 1). If the ancestral DNAs collectively underwent genetic drift, then the times between successive nodes of the tree are also a function of a genetic drift process. For several quite simple demographic and linkage/selection models the distribution of times between nodes has been solved (Tavaré, 1984; Hudson and Kaplan, 1988; Takahata, 1988). These theoretical compositions are called "coalescent" models (Kingman, 1982), reflecting the pattern of a collapsing sample as one proceeds from the present into the past. The general

prediction is that the most recent nodes of a tree will be more closely spaced in time than the more distant nodes, and that the expected time between successive nodes is proportional to the population size (Hudson, 1990). However the details of this prediction vary considerably depending on the demographic model (Tajima, 1989).

If a sample includes DNAs from two groups of organisms with a collective history of a single genetic drift process followed by a shift (speciation) to two drift processes, then the structure of the gene tree will differ in two general ways from the tree for a sample from a single genetic species. First, the distribution of times between successive nodes after the speciation will be the result of two independent systems of genetic drift. In the terms of a coalescent model, the waiting time between successive nodes will be a function of two random variables instead of a single one. Second, the tree will be partitioned into sections that are exclusive of DNAs from one of the sampled groups (Baum and Shaw, 1995). Figure 3 shows three examples of multiple DNAs from each of two independently drifting populations, one with a faster rate of drift (i.e. short time intervals between nodes) and one with a slower rate of drift. In Fig. 3A, the two samples form separate subtrees; while in Fig. 3B, one sample forms a gene tree that is nested within the gene tree for the other sample. Groups Y and Z of Fig. 3A and group Z in Fig. 3B are monophyletic, meaning that the group includes all of the descendants of a



**Figure 3** Three hypothetical gene trees. In each case, four DNAs have been randomly sampled from each of two genetic species, Y and Z. Species Y is larger with a slower rate of drift and larger intervals between nodes than species Z. See text for further explanation.

particular common ancestor (Hennig, 1979 p. 73). Figure 3C shows relatively little partitioning of the gene tree with respect to sample. This pattern could occur if the ancestors of one group recently became geographically separated from others of the same genetic species, so that relatively little drift has occurred within each group.

If genetic drift proceeds among the descendants of groups Y and Z, then the gene trees depicted for group Y in Fig. 3B and groups Y and Z in Fig. 3C will eventually be replaced by monophyletic gene trees (see e.g. (Avice and Ball, 1990)). The emergence of monophyletic gene trees is caused by the forward shift of the pattern of ancestry that occurs within a group of DNAs that share genetic drift (Fig. 2).

The condition of having previously identified candidate species (as in Fig. 3) is useful for articulating various patterns of genetic drift. However, in practice, researchers must consider cases where multiple species may exist, but where the DNAs have not been labeled beforehand as belonging to candidate species. This uncertainty adds considerably to the difficulty of identifying species simply because *a posteriori* hypotheses have much stricter criteria of statistical acceptance than *a priori* hypotheses. For sexual species, this burden can be reduced by generating hypotheses regarding genetic species using data from one locus. A second, unlinked, locus can then be studied to test these hypotheses, which are now *a priori* (Hey and Kliman, 1993; Hey, 1994).

For sexual organisms in a genetic species, different parts of the genome will have different gene tree histories, though nearly all portions are expected to share genetic drift. The actual rate of drift will vary among genomic regions by chance and because natural selection and variation in recombination rates will cause the rates of genetic drift to vary across the genome. One kind of natural selection that can frustrate a gene tree assessment of species status for sexual organisms, and thus require the study of multiple portions of the genome, is balancing selection. This type of natural selection occurs when there exists a stable pattern of multiple sequences, or alleles, for some region of the genome. The persistence of multiple functional forms will create gene trees like those in Fig. 3A, in this case with designations Y and Z referring to different alleles. Genetic drift will occur within each allele class, but natural selection prevents the replacement of one allele class by descendants of the other.

#### PROPOSALS FOR POPULATION BIOLOGISTS

Biologists often apply "species" without a clear meaning of the word or, if a specific meaning is articulated, with uncertainty over whether a group of organisms actually fit the meaning. For cases when "species" is used to convey some degree of

individuality on the part of a group of organisms, meaning some degree of spatio/temporal integrity, a two step convention is proposed. First, the specific meaning that should be used is that of genetic species. However, for most groups of organisms, including many that might be called species under other concepts, the population structure will not resemble panmixia. Thus the second component of species identification is the inclusion of a description of the pattern of genetic drift.

Having a meaning of "species" makes it possible for population biologists to avoid the word and associated uncertainties. If genetic drift underlies mechanistic species concepts, it follows that an assessment of the history of genetic drift for some group of organisms will obviate the need for species identification of those organisms. One or more species names may be applied, but will convey no additional information.

#### GENETIC SPECIES AND SYSTEMATICS

It has been proposed that some organisms do exist in groups that are real, and that the spatiotemporal distinctness of these groups arises when organisms share in a process of genetic drift. The genetic species is also a distinct kind of individual, different from the individual status that might be considered for larger groups of organisms. For example, it has been argued that species are not different from genera or other taxa, and that a taxon at any level can be thought of as an evolutionary unit (Nelson, 1989). However, the criteria of genetic species is shared genetic drift among organisms. It is possible that some organisms occur in a pattern of population structure with nested levels of genetic drift among organisms (see GENETIC SPECIES AND POPULATION STRUCTURE) so that genetic species may occur within larger genetic species. However this pattern is still caused by genetic drift among *organisms*, which is in turn caused by gene flow and demographic exchangeability. These causes of individuality at the species level need not be the same as processes that occur among species. While it may be argued that there are processes of species turnover that are analogous to genetic drift, these processes will not be identical to gene flow and demographic exchangeability. In short, the genetic species is a distinct kind of individual that includes multiple organisms.

The finding that an organism may not be part of a species, either in a contemporaneous sense or a historical sense, has implications for the study of the historical relationships among organisms and for classifying organisms. These implications may be roughly categorized as theoretical and practical. The primary theoretical implication is that a systematic theory that assumes that *all* organisms occur in nature

as parts of species, may be incorrect. For example, investigators in the field of phylogenetic systematics strive to use historical relationships among organisms as a guide for the classification of organisms (Hennig, 1979 p.73). It is implicit, and sometimes explicit, within this perspective that organisms truly do occur in phyla (e.g. species or higher taxa). For instance Hennig stated that species and higher taxa "are all segments of the temporal stream of successive 'interbreeding populations'" (1979 p. 81). If interbreeding populations are not, in fact, a necessary occurrence for organisms, then the theory of phylogenetic systematics has an error. However, as severe as the theoretical implications may be, the practical implications are largely unknown. It is not known how many organisms do not occur in genetic species, and it is not known to what degree the ancestors of present day organisms occurred in genetic species. Also, the estimation of gene trees as a way to estimate phylogenetic relationships will appear to be insensitive to whether or not ancestral organisms were in species. For instance, in the absence of sex there will still exist a bifurcating gene tree history for extant DNAs, regardless of the species status of ancestors (see CONTEMPORANEOUS SPECIES AND HISTORY). Such a tree may be misidentified as a phylogeny, when it actually does not represent phyla.

If the theoretical implications of the genetic species concept were to be included in a theory of systematics, then at least two possible courses can be considered. First, a theory of systematics could focus solely on the historical relationships of individual organisms (Vrana and Wheeler, 1992). This method would have the advantage that the identification of individual organism is vastly easier than for individual species. However, such a system would also face the difficulty that for any point in time, a sexual organism will likely have many ancestral organisms. This means that a bifurcating hierarchical tree will often not be a good historical model. An alternative proposal is that a systematic investigation could begin with a process of identifying candidate species and estimates of historical relationships among candidate species. For example, candidate species could be identified by relatively practical criteria (e.g. morphological similarity). This general approach is commonly used with other species concepts. In addition systematics would include a process of examining the genetic species status of candidate species. This kind of systematic protocol, with a theoretical component that includes the genetic species concept, would also need to address the issue of the classification of organisms that are not in species.

Both of these proposals, using organisms as the individuals of systematics and including the genetic species concept within systematic theory, address the

theoretical basis of systematics. A third alternative would be to forego biological theory, and develop a taxonomic system based on criteria identified by observers. For instance, organisms could be clustered on the basis of similarity (Sneath and Sokal, 1973) or on the basis of shared characters (Nelson, 1989) without assumptions of underlying biological processes. A taxonomic species identified in this way would be a class of organisms, and would have no necessary relationship to genetic species that occur as individuals in nature.

### CONCLUSIONS

#### *The genetic species concept within the context of species concept debates*

In this paper, the species problem has been identified with two ongoing scientific debates: first, over the meaning of the word "species"; and second, over methods of species identification. In practice, many disputes take specific forms that do not fall into just one of these categories of debate. Endler (1989) has outlined four dichotomies, each a broad category of species concept debate, and it may be informative to discuss the genetic species concept within the context of each of these.

*Taxonomic versus evolutionary* - Endler (1989) used "taxonomic" to refer to those species concepts motivated purely by the need for classification. The merit of a taxonomic concept of species is determined by how well it works for identification and grouping, without regard to the evolutionary relatedness of organisms, or the processes that gave rise to species. Endler's category of "evolutionary" applies to concepts that are defined in terms of, or make assumptions about, evolutionary processes. The genetic species concept is motivated solely by process, and thus would fall squarely in the "evolutionary" category.

*Theoretical vs operational* - A theoretical concept of species is concerned with the meaning of the word "species" in terms of the basic processes that give rise to species. Theoretical concepts, like the biological and the genetic species concepts, may not be very practical for purposes of species identification. Operational species concepts may include taxonomic concepts (see 1, above), because of the focus on applicability, but they also include species concepts that have a dual focus on evolutionary processes and practicality. Perhaps the best examples of these are the various attempts to develop a phylogenetic species concept (Rosen, 1979; Donoghue, 1985; Cracraft, 1989; Baum, 1992; Davis and Nixon, 1992).

*Contemporaneous vs clade* - Many evolutionary biologists, including population biologists, focus on extant organisms, and rely on meanings of "species" that hinge on ongoing biological processes. The biological and the ecological

species concepts are examples. In contrast, systematists are more likely to view species within the context of clades (i.e. monophyletic groups of species) and ancestor-descendant relationships. In this historical perspective, a species can be an explicitly historical entity (e.g. Simpson's evolutionary species), or a contemporaneous entity that is defined by historical relationships to others (e.g. (Cracraft, 1989) version of a phylogenetic species). Under the genetic species concept, a species is a group of organisms that share genetic drift. This is a contemporaneous definition.

*Reproductive vs Cohesive* - Endler (1989) classifies as "reproductive", those concepts that focus on the reproductive processes that maintain either isolation between species or recognition within species. He contrasts these with concepts that focus on species as units with genetic and phenotypic cohesion. The distinction between these, and thus the failure of some species concepts, must be recognized because of numerous examples of groups of organisms that appear to be one reproductive species because of the exchange of genes, but yet persist as multiple distinct cohesive units. The genetic species concept does not simply fit into either of Endler's categories. In a narrow sense, the genetic species concept is not a 'reproductive' concept because it does not apply only to sexual organisms. However, the reproductive successes of different organisms are linked within a genetic species, and this occurs regardless of sex. The genetic species concept does fit the 'cohesive' category in some respects. Genetic drift has long been recognized as a force that causes homogeneity among groups of organisms that share drift, and genetic drift is a significant component of Templeton's cohesion species concept (Templeton, 1989). However, under the genetic species concept, it is possible that a genetic species can form by the admixture of organisms that are not from genetic species or are from multiple genetic species. Such an admixture may exist as a group of organisms that share genetic drift, but at the time of initial mixing the group may be less cohesive than any of the groups from which the mixture was formed.

#### *The meaning of "reduction"*

The genetic species concept seems to be a reduction of those other species concepts that focus on biological processes. The terms of the genetic species concept (e.g. "DNA replication" and "genetic drift") refer to things that are components and properties of the things in other species concepts. The genetic species concept is not at odds with other concepts that are described in terms of biological processes. Other concepts are extensions of the genetic species concept in different contexts (e.g. ecological species, biological species, and evolutionary species). Components of

these concepts are also included within the cohesion species concept (Templeton, 1989). A cohesion species is "an evolutionary lineage that serves as the arena of action of basic microevolutionary forces, such as gene flow (when applicable), genetic drift and natural selection" (Templeton, 1994). Thus the cohesion concept resembles the evolutionary species concept with a population genetic emphasis on the causes of phenotypic similarity within species. However all of the basic microevolutionary forces can only occur when organisms share, to some degree, in a birth and death process and thus in the process of genetic drift.

The reduction of "species" has many parallels with the reduction of "gene". Beginning with Johanssen's strictly operational definition and extending to the time of the elucidation of the function of DNA sequences, the nature of genes and the meaning of "gene" were at the heart of great debate and energetic research (Dunn, 1965). Today it is clear that the function of DNA sequences takes many forms. For example, an incomplete list includes protein coding sequences, ribosomal RNA coding sequences, transfer RNA coding sequences, introns, promoters, enhancers, and telomeric sequences. Furthermore, the boundaries for some of these functional sequences, in terms of their location in a DNA chain, may not be distinct; and there are many cases where DNA sequences associated with one function are part of a larger sequence that is associated with a different function (e.g. enhancers lying within protein coding regions). Johanssen's single word is still useful as an approximation in many cases, but it does not begin to represent the breadth of functional diversity manifest in DNA sequences. In the day to day practice of molecular genetic research, the word "gene" has largely been replaced by a richer lexicon that has come with the growth of understanding of DNA function.

Like 'gene', the word 'species' may not be necessary or always useful in a research context, particularly when situations occur that do not fit a definition. However, understanding the meaning of "species" does simplify a plan for inquiry on the causes of species diversity. Speciation research should be the science of genetic history and of the biotic and abiotic phenomena that determine whether or not organisms share genetic drift and that shape the gene tree histories of closely related organisms. It is not necessary that the species status of organisms be established prior to research. If an assessment of the history of genetic drift for a sample of organisms reveals a relatively simple picture of one or more groups that share genetic drift, then species status may be delineated. However application of species status and species distinctions will not add information or meaning.

The process of species identification is not simplified by having a meaning of "species". Instead the task is shown to be as difficult as measuring genetic drift and then deciding if the data contain evidence of partitions in the structure of genetic drift. It may not be possible to identify species in many cases, but it is clearer why species can be so difficult to identify.

#### *Explaining the existence of species*

In practice, an investigation into the nature of species often begins with an implicit or explicit viewpoint that species exist. The approach taken in this paper has been to assume the existence of some relatively simple natural phenomena, and then to consider whether these things will cause something that corresponds to other ideas on species. In particular, the starting point includes the existence of replicating DNAs, an assumption that mutations can occur, and that the real world has limited resources and imposes geographic constraints on the location of DNAs. With this small set of observations it has been possible to predict several things about the diversity of DNAs that seem to correspond to common observations about the diversity of organisms and that are often attributed to species.

- 1) Mutations and geography will have the effect of partitioning DNAs into different groups, within each of which there is a shared process of genetic drift caused by the random replacement of some DNAs by the descendants of others.
- 2) Mutation will in the long run have a greater effect than geography on the formation of groups. Mutation is, for the most part, not reversible (see , THE CAUSES OF GENETIC SPECIES) while geography can change and can be overcome by the movement of DNAs.
- 3) There will often occur situations where the pattern of genetic drift does not occur with distinct boundaries among groups of DNAs, but rather occurs at varying rates across a geographic range. In these situations, the group status of DNAs will appear ambiguous.
- 4) DNAs that share a process of genetic drift will have a gene tree history shaped by that process. If they and their descendants persist in sharing genetic drift, then this gene tree will become monophyletic.
- 5) Groups of DNAs that share genetic drift can occur whether or not the DNAs exchange portions of their sequence (i.e. engage in sex).

In short, a very small set of basic observations lead to predictions that correspond very well to many of our most basic observations about groups of organisms, and that are often attributed to species. In this sense, the existence of species has been at least partly explained. Also, since the basic observations

(i.e. replicating DNA, mutations, geography and limiting resources) are not in doubt and are relatively simplistic (i.e. being chemical and physical, but not very biological), these observations seem to be necessary for any complete explanation of species. It is then appropriate to ask whether these observations are a sufficient explanation of species. There remain two distinct contexts in which the genetic species concept may be judged incorrect by reason of insufficiency. First, the genetic species concept may imply some unforeseen consequence that is not a true phenomenon. As described, genetic drift is an epiphenomenon of the basic components of replicating DNA, mutations, geography and limiting resources. If it were shown that the simple model that has been constructed also leads to predictions that do not occur in nature, then the model is incomplete and something must be added. Second, something may have been overlooked in the true nature of species. It may be that all real species share a property that cannot be predicted by the simple model developed here.

#### ACKNOWLEDGMENTS

I thank David Baum, Michael Bulmer, Rob DeSalle, Adam Eyre-walker, Douglas Futuyma, Michael Ghiselin, Anna Graybeal, Brandon Gaut, Holly Hilton, David Houle, Richard Kliman, Nicholas Leahy, Richard Lewontin, Margaret Saks, Elliot Sober, Alan Templeton, Robert Trivers, David Wake, John Wakeley, and Rong Lin Wang for comments on the manuscript. This research was supported by a grant from NSF (DEB-9306625).

#### REFERENCES

- Avise, J. C., and R. M. J. Ball. 1990. Gene genealogies and the coalescent process, pp. 45-67. *In* P. H. Harvey and L. Partridge (eds.), *Oxford surveys in evolutionary biology*, vol. 7, Oxford University Press, New York.
- Baum, D. 1992. Phylogenetic species concepts. *Tr. Ecol. Evol.* 7:1-2.
- Baum, D. A., and K. L. Shaw. 1995. Genealogical perspectives on the species problem, pp. 289-303. *In* P. C. Hock and A. G. Stevenson (eds.), *Experimental and molecular approaches to plant biosystematics*, Missouri Botanical Garden, St. Louis.
- Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular evolution. *Genetics* 134:1289-1303.
- Cracraft, J. 1989. Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation, p. mybook. *In* D. Otte and J. A. Endler (eds.), *Speciation and its consequences*, Sinauer Associates, Sunderland, Mass.
- Davis, J. I., and K. C. Nixon. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biol.* 41:421-435.
- Dawkins, R. 1978. Replicator selection and the extended phenotype. *Zeitschrift fur Tierpsychol* 47:61-76.
- de Queiroz, K., and M. J. Donoghue. 1988. Phylogenetic systematics and the species problem. *Cladistics* 4:317-338.
- Dobzhansky, T. 1950. Mendelian populations and their evolution. *Amer. Natur.* 84:401-418.
- . 1951. *Genetics and the origin of species*, 3rd ed. Columbia University Press, New York.
- Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist* 88:172-181.
- Dunn, L. C. 1965. *A short history of genetics*. McGraw-Hill Book Company, New York.
- Endler, J. A. 1989. Conceptual and other problems in speciation, pp. 625-648. *In* D. Otte and J. A. Endler (eds.), *Speciation and its consequences*, Sinauer Associates, Sunderland, Mass.
- Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.
- Ghiselin, M. T. 1966. On psychologism in the logic of taxonomic controversies. *Syst. Zool.* 15:207-215.
- . 1974. A radical solution to the species problem. *Syst. Zool.* 23:536-544.
- . 1987. Species concept, individuality, and objectivity. *Biology and Philosophy* 2:127-143.
- Gilbert, W. 1986. The RNA world. *Nature* 319:618.
- Gulick, J. T. 1872. On diversity of evolution under one set of external conditions. *J. Linn. Soc. Lond. Zool.* 11:496-505.
- . 1888. Divergent evolution through cumulative segregation. *J. Linn. Soc. Lond. Zool.* 20:189-274.
- Hartl, D. L., and D. E. Dykhuizen. 1984. The population genetics of *Escherichia coli*. *Annual Review of Genetics* 18:31-68.
- Hennig, W. 1979a. *Phylogenetic systematics*. University of Illinois Press, Chicago, 73.
- . 1979b. *Phylogenetic systematics*. University of Illinois Press, Chicago, 59.
- . 1979c. *Phylogenetic systematics*. University of Illinois Press, Chicago, 73.
- . 1979d. *Phylogenetic systematics*. University of Illinois Press, Chicago, 81.
- Hey, J. 1991. A multi-dimensional coalescent process applied to multi-allelic selection models and migration models. *Theor. Pop. Biol.* 39:30-48.
- . 1994. Bridging phylogenetics and population genetics with gene tree models. *In* B.

- Schierwater, B. Streit, G. Wagner and R. DeSalle (eds.), *Molecular Approaches to Ecology and Evolution*, Birkhäuser-Verlag, Basel.
- Hey, J., and R. M. Kliman. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. *Mol. Biol. Evol.* 10:804-822.
- Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. *Genet. Res. Camb.* 8:269-294.
- Hudson, R. R. 1990. Gene genealogies and the coalescent process, pp. 1-44. *In* P. H. Harvey and L. Partridge (eds.), *Oxford Surveys in Evolutionary Biology*, vol. 7, Oxford University Press, New York.
- Hudson, R. R., and N. L. Kaplan. 1988. The coalescent process in models with selection and recombination. *Genetics* 120:819-829.
- Hull, D. L. 1976. Are species really individuals. *Syst. Zool.* 15:174-191.
- . 1978. A matter of individuality. *Philosophy of Science* 45:335-360.
- Hutchinson, G. E. 1958. Concluding remarks. *Cold Spring Harbor Symposium in Quantitative Biology* 22:415-427.
- Kaplan, N., R. R. Hudson, and C. H. Langley. 1989. The "hitchhiking effect" revisited. *Genetics* 123:887-899.
- Kingman, J. F. C. 1982. The coalescent. *Stochastic Process. Appl.* 13:235-248.
- Levin, D. A. 1979. The nature of plant species. *Science* 204:381-384.
- Lewontin, R. C. 1970. The units of selection. *Annual Review of Ecology and Systematics* 1:1-18.
- Masters, J. C., and H. G. Spencer. 1989. Why we need a new genetic species concept. *Syst. Zool.* 38:270-279.
- Maynard Smith, J., and J. Haigh. 1974. The hitchhiking effect of a favourable gene. *Genet. Res. Camb.* 23:23-35.
- Mayr, E. 1942. *Systematics and the origin of species*. Columbia University Press, New York.
- . 1963. *Populations, species and evolution*. Harvard University Press, Cambridge, Mass.
- . 1982. *The growth of biological thought*. Harvard University Press, Cambridge, MA.
- Nelson, G. 1989. Species and taxa: systematics and evolution, pp. 60-81. *In* D. Otte and J. A. Endler (eds.), *Speciation and its consequences*, Sinauer Associates, Sunderland, Mass.
- Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. *Cladistics* 6:211-223.
- Orgel, L. E. 1992. Molecular replication. *Nature* 358:203-209.
- Rosen, D. E. 1979. Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative biogeography. *Bull. Am. Mus. Nat. Hist.* 162:267-376.
- Simpson, G. G. 1961. *Principles of animal taxonomy*. Columbia University Press, New York, 150-153.
- Slatkin, M. 1987. The average number of sites separating DNA sequences drawn from a subdivided population. *Theor. Pop. Biol.* 32:42-49.
- Sneath, P. H. A., and R. R. Sokal. 1973. *Numerical Taxonomy*. Freeman, San Francisco.
- Tajima, F. 1989. The effect of change in population size on DNA polymorphism. *Genetics* 123:597-601.
- Takahata, N. 1988. The coalescent in two partially isolated diffusion populations. *Genet. Res. Camb.* 52:213-222.
- Tavaré, S. 1984. Line-of-Descent and genealogical processes, and their applications in population genetics models. *Theor. Pop. Biol.* 26:119-164.
- Templeton, A. R. 1981. Mechanisms of speciation- a population genetic approach. *Annual Review of Ecology and Systematics* 12:23-48.
- . 1989. The meaning of species and speciation: a genetic perspective", b"Speciation and its consequences, p. xx and mybook. *In* D. Otte and J. A. Endler (eds.), *Speciation and its consequences*, Sinauer Associates, Sunderland, Mass.
- . 1994. The role of molecular genetics in speciation studies. *In* B. Schierwater, B. Streit, G. Wagner and R. DeSalle (eds.), *Molecular Approaches to Ecology and Evolution*, Birkhäuser-Verlag, Basel.
- Van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon* 25:233-239.
- Vrana, P., and W. Wheeler. 1992. Individual organisms as terminal entities: laying the species problem to rest. *Cladistics* 8:67-72.
- Watterson, G. A. 1982. Mutant substitutions at linked nucleotide sites. *Adv. Appl. Prob.* 14:206-224.
- Wright, S. 1931. Evolution in mendelian populations. *Genetics* 16:97-159.
- . 1943. Isolation by distance. *Genetics* 28:114-138.