

The biological reality of species: gene flow, selection, and collective evolution

Loren H. Rieseberg¹ & John M. Burke¹

Summary

Rieseberg, L. H. & Burke, J. M.: The biological reality of species: gene flow, selection, and collective evolution. – *Taxon* 50: 47–67. – ISSN 0040-0262.

Some authors have claimed that species are passive end products of evolution and thus not substantially different than higher taxa. This claim is based on reports that (1) levels of intraspecific gene flow may be too low to account for species' integration, and (2) populations are likely to diverge rather than evolve in parallel when exposed to uniform selection pressures. These conclusions are premature. A review of the plant literature reveals that there is sufficient gene flow to enable the efficient spread of strongly favourable alleles ($s > 0.05$), the most likely agents of collective evolution. Moreover, estimates of s for major quantitative trait loci (QTLs) are sufficiently large to enable their spread across the range of a species, although minor QTLs seem more likely to evolve locally. In addition, evidence that intraspecific variation in genetic background affects the response of alleles to selection is rare, but examples of parallel genotypic evolution are becoming increasingly common. We conclude that, as traditionally believed, species are the most inclusive entities that directly participate in evolutionary processes. However, we also note that the traditional role of gene flow as a force that constrains differentiation due to genetic drift or local adaptation has been over-emphasised relative to its creative role as a mechanism for the spread of advantageous mutations.

Keywords: advantageous mutations, gene flow, genetic background, QTLs, selection, species.

Introduction

The discussion of species concepts has become a cottage industry. Not only does the number of papers and full-length books devoted to the topic continue to grow (e.g., Claridge & al., 1997; Wilson, 1999; Wheeler & Meier, 2000), but new concepts of species proliferate at an extraordinary rate. For example, Mayden (1997) lists 22 different concepts, and several additional concepts have arisen since this list was compiled (e.g., de Pinna, 1999; Levin, 2000). Unfortunately, the singular attention given to species concepts by scientists and philosophers has not had the desired effect of moving the field toward consensus. Rather, opinions on species appear to be diversifying as a function of increased discussion. It might even be argued that the increased attention given to this topic has done more to bolster Curricula Vitae than to clarify our understanding of what species are, how they arise, and how to discover them in nature. This sentiment is perhaps best articulated by Coyne (1992), who writes that "it is clear that arguments [about species concepts] will persist for years to come, but equally clear that, like barnacles on a whale, their main effect is to retard slightly the progress of the field". In a similar vein, Schemske (2000) asks, tongue in cheek, that proposals of new species concepts be limited to "one per century".

¹Department of Biology, Indiana University, Bloomington, Indiana 47405, U.S.A. E-mail: lriesebe@indiana.edu (author for correspondence); jburke@bio.indiana.edu.

Given the discordant nature of these discussions and the poor prospects for consensus, we must admit to misgivings about authoring a paper on species concepts. We neither wish to contribute more fuel to the debate, nor add another barnacle to Coyne's whale. Nonetheless, ideas have been put forward about species that offer challenges to students of speciation. In particular, there is growing skepticism about the biological reality of species or at least about the distinction between species and higher taxa (Levin, 1979; Mishler & Donoghue, 1982; Luckow, 1995; Bachmann, 1998; Ereshefsky, 1999; Mishler, 1999).

The traditional view is that species are fundamental units of natural diversity because they represent the most inclusive entities that directly participate in evolutionary processes. The process most commonly invoked is gene flow, which can act conservatively to prevent diversification or creatively through the diffusion of favourable mutations. However, the role of gene flow as a cohesive evolutionary force has been challenged because levels of gene flow are said to be too low in some taxa, particularly certain plant species, to account for their apparent integration (Ehrlich & Raven, 1969; Grant, 1980). Parallel responses to selection have been posited as another means by which populations might evolve collectively (Lande, 1980; Templeton, 1989), but others have argued that populations are most likely to diverge in response to uniform selection pressures (Wade & Goodnight, 1998; Levin, 2000). If populations within a species do not evolve collectively, it becomes difficult to argue that they are different from higher taxa. Like higher taxa, they would simply be passive end products of evolution or aggregates of evolving units. From a philosophical standpoint, they would fit best in the category *class* or *kind* rather than in the category *individual* (Hull, 1999).

This conclusion, if correct, also has interesting ramifications for studies of speciation. In particular, analyses of reproductive isolating mechanisms—the chief occupation of students of speciation—would have to be considered as an unfortunate scientific detour, having sidetracked the careers of such promising intellectuals as Ernst Mayr and Theodosius Dobzhansky. Local populations or metapopulations would presumably replace species as the unit of evolution (Luckow, 1995; Levin, 2000), increasing the intellectual divide between micro- and macroevolutionary biology. As with higher taxa, there would be no criteria for ranking clusters of populations in a comparable way, providing impetus for a rank-free taxonomy, as urged by Mishler (1999) among others.

However, as we will attempt to show in this paper, the conclusion that species are passive end products of evolution and thus no different from higher taxa is premature. We make this argument by first reviewing the literature on gene flow levels among plant populations. As will quickly become apparent, most claims about too little gene flow are based on early studies that, for technical reasons, grossly underestimated actual gene flow levels. In addition, many authors have also failed to consider that (1) only very low rates of gene flow are required for the diffusion of strongly advantageous alleles, the most likely agents of collective evolution, and (2) there have been large and rapid changes in the distributions of many plant species, which also suggest that gene dispersal rates are likely underestimated. Following our review of the gene flow literature, we will discuss what is known about the likely response of populations to uniform selection pressures. Although there is little

relevant information available on this topic, we will describe some striking examples of parallel evolution, as well as the paucity of evidence for strong genetic background effects on adaptive evolution within species. Our focus throughout will be on plant species because, perhaps more than any other group, they offer the greatest diversity of challenges to a monistic view of species. However, our conclusions should be generalisable to other organismal groups.

Gene flow

Theory. — Gene flow may unify populations in two ways (Slatkin, 1987). First, it may be a constraining force that acts to prevent differentiation due to genetic drift or adaptation to local conditions. Second, it may be a creative mechanism that promotes the collective evolution of populations; advantageous mutations that arise in a local population may spread to other populations through the movement of gametes, individuals, or groups of individuals.

Most theoretical work has focused on the conservative role of gene flow (migration), and two simple rules have been discovered that summarise its relationship to drift (rule 1) and selection (rule 2):

Rule 1. If, on average, one or more individuals are exchanged between two populations per generation ($Nm \geq 1$), differentiation through drift will be prevented (Wright, 1931). In contrast, if migration rates fall below this level ($Nm < 1$), then differences are likely to accumulate. This rule is unaffected by population size because the efficacy of drift declines inversely with population size and the two forces remain in balance. Thus, if population sizes are large, the fraction of immigrants required to prevent differentiation through drift is very small.

Rule 2. Gene flow will prevent local selective differentiation at a given locus unless the strength of selection (s , the fitness difference between alternative alleles) exceeds the fraction of immigrants (m). Note that unlike drift, which affects all loci equally, the effects of selection will vary across loci. Thus, it is possible for gene flow to be high enough to prevent divergence at some loci, but to be overcome by selection at other loci. This leads to the paradoxical situation in which populations are likely to be evolving collectively at some loci, while simultaneously diverging at others.

Unfortunately, there are no simple rules for predicting the rate of spread of advantageous alleles. The primary problem is that most population genetic models do not consider discontinuous population structure and sporadic long-distance migrants. This is a serious concern. Discontinuous population structure greatly reduces the rate of spread because a new allele arriving in a population has only a very small chance of becoming fixed. Long-distance dispersal may be even more critical, as ecological models of biological invasions indicate that even when long-distance dispersal is rare, it still governs the invasion speed (Neubert & Caswell, 2000).

As far as we are aware, the only model that adequately accounts for discontinuous population structure and long-distance dispersal is that of Slatkin (1976). He models the amount of time it would take for a new mutant allele to spread across populations in a stepwise fashion (Fig. 1) and then asks whether it would occur more quickly if long-distance dispersal occurred at a very low frequency. The primary

shortcomings associated with Slatkin's model are that gene flow occurs in only one dimension and the likely synergistic effects of both short- and long-distance dispersals cannot be assessed. Both shortcomings are likely to lead to an underestimate of rate of spread. Nonetheless, a number of conclusions can be drawn from the model (Fig. 2):

- (1) Migration rate has a small effect on the number of generations required for the spread of an advantageous allele, except when migration rates are extremely low.
- (2) The selective advantage of the mutant allele is the primary factor governing rate of spread.
- (3) Long-distance migration greatly enhances rate of spread.

The take-home message is that even very low levels of gene flow should be sufficient for the rapid advance of advantageous alleles, as long as selection coefficients are fairly large ($s > 0.05$; Fig. 2). However, any conclusions regarding whether the spread of favourable alleles is likely to promote the collective evolution of plant populations require empirical estimates of migration rates and selection coefficients.

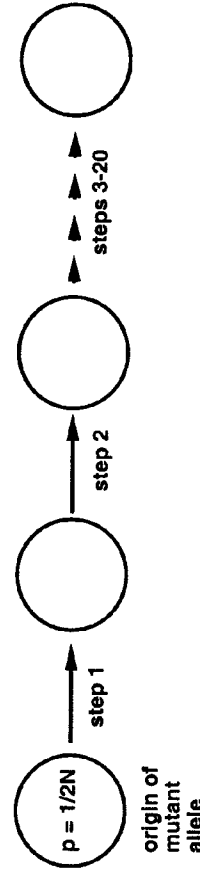


Fig. 1. Illustration of stepping-stone model of gene flow employed by Slatkin (1976) to study the spread of mutant alleles across a species' range.

Empirical estimates of gene flow. — Early studies of seed and pollen dispersal appear to have systematically underestimated actual levels of gene flow (Levin, 1981). The most important reasons for this are reviewed below, but more detailed discussions can be found elsewhere (e.g., Schaal, 1980; Levin, 1981; Ellstrand, 1992; Hamrick & Nason, 2000). With respect to gene flow via pollen, most early experiments either studied pollinator foraging behaviour or the movement of marked pollen from a central source. Gene flow estimates based on pollinator foraging behaviour are problematic because they do not take into account pollen carry-over, biological factors involved in mating success, and occasional long-distance dispersals. Studies of pollen dispersal based on tracking pollen from a central source also underestimate gene flow, but for different reasons. Simply put, source pollen from a single plant follows a leptokurtic function (Handel, 1982), but a given recipient plant accumulates pollen from the tails of many leptokurtic distributions. Thus, more accurate estimates of pollen dispersal can be obtained by analysing sink rather than source populations.

As with pollen flow, most early estimates of seed dispersal were obtained by tracking the movement of seed dispersal agents or by monitoring the movement of seed from a central source (Howe & Smallwood, 1982). The problems associated with these approaches are similar to those described above for pollen dispersal, and

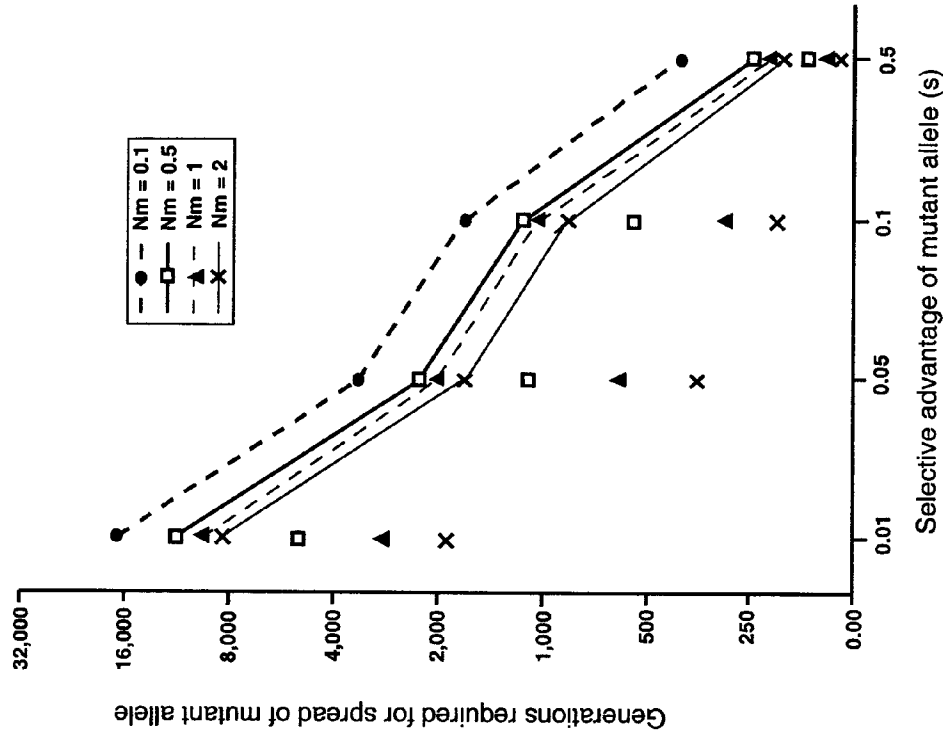


Fig. 2. Number of generations required for a mutant allele to reach a population 20 steps away from the one in which it originated (cf. Fig. 1). Symbols connected by lines show the numbers of generations required for allelic spread in the absence of long-distance dispersal. Symbols without lines indicate the number of generations required for allelic spread with long-distance dispersal. The effects of long-distance dispersal were calculated by assuming that the frequency of migrant individuals reaching a population 20 steps away is 1% of that for reaching an adjacent population (Slatkin, 1976). Note that for $N_m = 0.1$, long-distance dispersal had no impact on rates of spread because long-distance dispersals were so rare as to be non-existent. All values shown in this figure are extrapolated from Table 1 of Slatkin (1976).

Note that the number of steps required for a mutant allele to spread across the range of a species will vary depending on the distribution of the species, the location of the population in which the mutant allele originates, migration rates, and the frequency of long-distance dispersal. Our choice of 20 steps follows Slatkin (1976) and should be viewed as a very rough guess of stepwise distances across a typical species' range.

seed dispersal appears to have been consistently underestimated in these early studies.

The development of molecular markers, combined with improved experimental designs, has led to much more accurate estimates of gene flow (Sork & al., 1999; Hamrick & Nason, 2000). One approach, which may be best suited for estimating the role of gene flow as a cohesive evolutionary force over long durations, uses the amount of differentiation among populations to provide an estimate of historical levels of gene flow or N_m (the effective number of migrants coming into a population per generation). Differentiation may be measured in many different ways, including F_{st} , G_{st} , ϕ_{st} , AMOVA, and private alleles, but all make the same fundamental assumptions in deriving N_m (Whitlock & McCauley, 1999). These include no selection, no mutation, all populations are equivalent in size and make equal contributions to the migrant pool, migration is random, and populations are at equilibrium. These assumptions have been challenged by many authors (e.g., Bossart & Prowell, 1998; Whitlock & McCauley, 1999), but as noted by Whitlock & McCauley (1999), comparisons of large groups of species may be informative because differences between estimated and actual gene flow levels are likely to average out.

N_m values calculated using this approach make sense when interpreted with respect to what is known about the life histories of the species being compared (Hamrick & Godt, 1996). As one might predict, historical gene flow levels are very low for the average selfer ($n = 123$; $N_m = 0.24$), moderate for mixed-maters ($n = 142$; $N_m = 0.90$), and quite high for outcrossers ($n = 481$; $N_m = 1.43$). These values suggest that although some divergence is likely through drift in selfers and mixed-maters, in all cases, gene flow levels are high enough to allow the rapid spread of advantageous alleles, as long as the selective advantage of mutant alleles is sufficiently high. Even in a worst-case scenario, such as in a selfing annual with an N_m value of 0.1, a mutant allele with a selective advantage of 0.05 should take less than 4,000 years to spread throughout the range of the species (Fig. 2). This is a short interlude in the life of a plant species, which has been estimated for herbs as 10 million years (Levin & Wilson, 1976). Thus, limited gene flow should not prevent the collective evolution of populations as mitigated by the spread of advantageous alleles.

One drawback with the historical or "indirect" estimates of gene flow described above is that they fail to measure contemporary gene flow. For example, it is possible that populations that have had historically high levels of gene flow (and thus high levels of N_m) may no longer be exchanging genes. Another concern relates to the many assumptions that are violated when calculating N_m from estimates of population differentiation (Whitlock & McCauley, 1999). Thus, direct estimates of contemporary gene flow are critical to ensure that the indirect, historical values are meaningful. For reasons discussed earlier, the least biased approach for measuring contemporary gene flow levels is to analyse the number of seeds found in a sink population that have mothers or fathers from outside the population. This evaluation is typically conducted using some kind of paternity or parentage analysis.

Direct estimates of gene flow through pollen (Table 1) generally seem higher than necessary to explain N_m values for species with similar life histories (Hamrick &

Godt, 1996). However, direct estimates typically are for populations separated by fairly short distances, whereas N_m provides an average value for the entire species. Moreover, pollen is haploid, but N_m measures the immigration of diploid individuals. If these factors are considered, direct estimates of pollen flow appear to be compatible with N_m values.

Table 1. Percentage of seeds or seedlings within natural populations that were sired by pollen from outside the population as assessed by molecular markers.

Species	(%)	Distance (m)	Immigrant pollen	Reference
herbs:				
<i>Asclepias exaltata</i> L.	29–50	>50 to >1000		Broyles & al., 1994
<i>Cucurbita texana</i> A. Gray	5	1300		Kirkpatrick & Wilson, 1988
<i>Datisca glomerata</i> Baill.	27	>500		Nakazato, Wolf & Rieseberg, unpubl.
<i>Helianthus annuus</i> L.	4–42	3–400		Whitton & al., 1997
<i>Lathyrus latifolius</i> Vis.	40	>70		Godt & Hamrick, 1993
	23	>21		Godt & Hamrick, 1993
<i>Persoonia mollis</i> R. Br.	21	>300		Krauss, 2000
<i>Raphanus sativus</i> L.	3–18	100–1000		Ellstrand & al., 1989
	20	>150		Devlin & Ellstrand, 1990
<i>Silene alba</i> Burnat	6–47	>10 to >80		Richards & al., 1999
trees¹:				
<i>Calophyllum longifolium</i> Wall.	62	>210		Stacy & al., 1996
<i>Cecropia obtusifolia</i> Bertol.	37	>1000		Kaufman & al., 1998
<i>Cordia alliodora</i> Cham.	3	>280		Boshier & al., 1995
<i>Ficus</i> (three spp.)	>90	>1000		Nason & Hamrick, 1997
<i>Gleditsia triacanthos</i> L.	40–70	>200		Schnabel & Hamrick, 1995
<i>Magnolia obovata</i> Thunb.	>36	?		Isagi & al., 2000
<i>Pinus flexilis</i> James	7	>4000		Schuster & Mitton, 2000
<i>Pinus sylvestris</i> L.	26	>200		Harju & Muona, 1989
	38	>100		Nagasaka & Szmidt, 1984
	21–36	>100		El-Kassaby & al., 1989
<i>Pinus taeda</i> L.	28–48	>200		Friedman & Adams, 1985
<i>Pinus taeda</i> L.	15	>350		Chase & al., 1996
<i>Platanus occidentalis</i> L.	36–77	>100		Hamrick & Murawski, 1990
<i>Pseudotsuga menziesii</i> (Milib.) Franco	40	>100		Neale, 1984
	29–52	>100		Smith & Adams, 1983
	4–25	>500		Wheeler & Jech, 1986
<i>Quercus macrocarpa</i> Endl.	57	>200		Dow & Ashley, 1998
<i>Spondias mombin</i> L.	44	>200		Nason & Hamrick, 1997
<i>Tachigala versicolor</i> Standl. & L.O. Williams	25	>500		Loveless & al., 1998
<i>Turpinia occidentalis</i> G. Don	1–27	>130		Stacy & al., 1996
<i>Yucca filamentosa</i> L.	≥10	>200		Massey & Hamrick, 1999

¹Much of the information for the tree examples is from Table 6.2 of Hamrick & Nason (2000).

The question of seed dispersal is more difficult. In the few studies that have estimated the parentage of seeds or seedlings in populations (e.g., Schnabel & Hamrick, 1995; Dow & Ashley, 1996; Isagi & al., 2000), seed immigration ranges

from about 2.1% (honey locust) to 40% (*Magnolia obovata* Thunb.). Another way to estimate the contribution of seed dispersal to gene migration is by comparing levels of interpopulation differentiation for maternal, paternal, and biparentally inherited genes. Studies that have employed this method have reported ratios of pollen to seed flow between 4 (for the selfing annual, *Hordeum spontaneum* C. Koch) and 500 (for the wind-pollinated, gravity-dispersed *Quercus petraea* Liebl.) (Ennos, 1994; El-Mousadik & Petit, 1996; Latta & Mitton, 1997; McCauley, 1997; Latta & al., 1998). Because of occasional very long-distance dispersal events by seeds, the overall contribution of seed dispersal appears to increase at greater geographic distances. For example, in the perennial herb *Silene alba* Burnat, the ratio of pollen to seed movement was 124 at the scale of meters compared to 3.4 at the scale of kilometers. In addition, long-distance dispersal of seeds is required to explain the rapid expansion of many forest trees following the last glacial epoch (Hewitt, 1996; Cain & al., 2000).

Ehrlich & Raven (1969) also use biogeographical evidence to question the role of gene flow in species cohesion. They point out that populations of the same species separated by major barriers (e.g., oceans, mountain ranges, deserts, etc.) often appear to be similar morphologically in the apparent absence of gene flow. However, there are two problems with this argument. First, it may be that the disjunct distribution is very recent and there has been inadequate time for species divergence. Because of climatic changes, species distributions are dynamic, and populations that are apparently isolated today may not have been a few thousand years ago. Likewise, it may be that currently isolated populations will be reunited in the future. A second possible problem is that these barriers may not be as impermeable as previously believed. The efficacy of long-distance dispersal for many plant taxa has been demonstrated repeatedly by the rapid colonisation of islands—sometimes by species with no obvious means of long-distance dispersal (Cain & al., 2000). In addition, rapid Holocene migrations of both herbs and trees cannot be explained in the absence of repeated long-distance dispersals (Hewitt, 1996; Cain & al., 1998; Clark & al., 1999). There is no reason to believe that the long-distance dispersal events that create the widespread distribution of many species come to a sudden halt once this distribution is achieved. Rather, rates of gene flow may actually increase because potential source populations are now more numerous and widespread.

In conclusion, there appears to be sufficient gene flow among populations of most or all plant species to allow the efficient spread of favourable alleles as long as the fitness advantage conferred is fairly large ($s > 0.05$). This is particularly true if long-distance dispersal is included in these calculations (Fig. 1). However, in many plant species (selfers and some mixed-maters) gene flow levels are not high enough to prevent local differentiation due to drift, and in most if not all plant species, gene flow levels are not high enough to prevent adaptation to local conditions. Thus, we should view plant (and animal) species as groups of populations that are evolving collectively at some loci, but diverging at others. Note that species' cohesion need not require identity at all loci, but may be maintained by collective evolution at a small number of major loci (see below).

Selection

Empirical estimates of selection coefficients. — The above conclusion regarding the adequacy of gene flow rates is contingent on the distribution of s for alleles that contribute to fixed differences between species. Unfortunately, as far as we are aware, this distribution is unavailable for any pair of plant species. However, there are several lines of evidence that suggest that s may be sufficiently large to have enabled allelic spread across the range of a species.

One line of evidence comes from the analysis of morphological traits in segregating progenies produced from crosses between plant species. The early literature in this area is comprehensively reviewed by Gottlieb (1984), who demonstrates that most of the taxonomic characters used to distinguish plant species (i.e., differences in the structure, shape, orientation, and presence versus absence of such characters) are governed by one or two genes. In contrast to these "discrete" characters, Gottlieb (1984) noted that differences in length, width, weight, and number tend to show quantitative patterns of inheritance and are controlled by a larger number of genes.

Unfortunately, the genetic methods available at that time (e.g., Lande, 1981) provided imprecise estimates of the number and magnitude of infinitesimal factors. It was impossible, for example, to distinguish between Fisher's infinitesimal model, in which quantitative traits are controlled by a very large number of factors of small effect individually (Fisher, 1930), and the oligogenic model, in which continuous patterns of variation are explained by a few major factors (Tanksley, 1993). Recently, methods for analysing quantitative trait loci (QTLs) using mapped molecular markers provided a solution to this problem (Paterson & al., 1988). In the QTL approach, phenotypic effects are assigned to a specific chromosomal block, making it feasible to estimate QTL numbers and magnitudes.

What generalisations can be made from QTL studies? First, Gottlieb's conclusions regarding the simple inheritance of discrete, taxonomic characters have generally been confirmed (e.g., Vlot & al., 1992; Doebley & Stec, 1993; Battjes & al., 1994; Homburg & Bachmann, 1994; Van Houten & al., 1994; Bradshaw & al., 1998). An apparent exception to this general rule was the discovery of more complex inheritance of phyllary and seed shape in interspecific sunflower crosses (Kim & Rieseberg, 1999), but shape was measured as a ratio of length to width in this study. Second, many length, width, weight, and number traits that display classic quantitative variation patterns are characterised by oligogenic inheritance (e.g., Paterson & al., 1988, 1991; Doebley & Stec, 1993; Bradshaw & al., 1995, 1998; Lin & al., 1995; Lin & Ritland, 1997; Kim & Rieseberg, 1999). Probably the best example is from *Mimulus*, where analysis of a large interspecific mapping population detected major QTL (> 25% phenotypic variance explained) for 9 of 12 floral characters measured (Bradshaw & al., 1998). Many of these were length, width, and volume traits. Of course, there are exceptions to the general rule of oligogenic inheritance for interspecific quantitative differences in plants. For example, Lin & Ritland (1997) failed to detect QTL for flower width in a 247-individual mapping population of *Mimulus guttatus* DC. $\times$ *M. platycalyx* Greene. This failure would be expected under the assumptions of the infinitesimal model.

The discovery that many of the traits that differentiate plant species are governed by one or a few factors of moderate to large effect is consistent with a model of species evolution in which widely distributed populations evolve collectively through the spread of advantageous alleles. However, this conclusion is preliminary for several reasons. First, it is not clear whether an allele with a large effect on phenotype will have an equally large effect on fitness. Second, the effects of any single locus may result from multiple mutations, some of which may have arisen and/or gone to fixation in the time since speciation. Third, many of the QTL studies cited above employed small mapping populations, likely leading to an overestimate of QTL magnitudes (Beavis, 1998). Nonetheless, current evidence does favour moderate to large values of s .

More direct evidence comes from studies of segregating hybrid populations grown under natural conditions (Table 2). These studies provide selection differentials for segregating traits, although in most cases the genetic basis of the traits is unknown. It is straightforward, however, to calculate crude estimates of the likely selective advantage (s) of individual QTL alleles based on assumptions about genetic architecture gleaned from studies of interspecific mapping populations (Table 2). For studies that report selective differentials for the QTLs themselves (e.g., Schemske & Bradshaw, 1999), s can be calculated without making additional genetic assumptions.

Although the number of relevant studies is small, they do indicate that the selective differential for traits/QTLs that vary between subspecies or species are often quite large, with the magnitude of the maximum differential for a given trait averaging 0.40 (Table 2). However, the magnitude of s for individual alleles appears to be considerably smaller, averaging 0.11 for a hypothetical leading QTL (Table 2). One interpretation of these data is that s for major QTLs is likely to be large enough to facilitate their rapid spread, but that minor QTLs will spread slowly and are less likely to contribute to fixed differences between species. Alternatively, s may be underestimated in segregating hybrid populations because the variance in fitness is high relative to a more realistic situation in which a mutant allele is introduced into a less polymorphic population. It also is noteworthy that estimates of s were highest for the one experiment (Schemske & Bradshaw, 1999) in which selection on QTLs rather than traits was assessed (Table 2).

We also calculated s from selection differentials for traits that were polymorphic within species, but similar in kind to those that differentiate species (Table 2). The average magnitude of the maximum selection differential for each trait (0.68) actually exceeded those for interspecific differences, suggesting there may be some validity to the concerns (above) about excessive variance in fitness in hybrid populations. These selection differentials translate into an average value of s of 0.12 for a hypothetical leading QTL (Table 2). Again, these data suggest that at least for major QTLs, s is sufficiently high to allow populations of most plant species to evolve collectively.

Of course not all differences between species need to have spread across a geographic range. In some instances, it seems likely that speciation has occurred locally followed by range expansion (Levin, 2000). In these cases, some or all of a species' unique features may result from common descent rather than persistent gene

flow. It is also possible that only a few major alleles spread across a species range and that modifier loci with smaller individual effects evolve locally. One final criticism of the hypothesis that favourable alleles spread across a species range is that mutation may not limit evolution in a large population (Barton, in press). If this is true, alleles of equivalent value may be more likely to arise independently throughout the range of a species than to spread through gene flow.

The response of populations to selection. — Another means by which populations might evolve collectively is through parallel responses to selection (Ehrlich & Raven, 1969; Lande, 1980). However, this hypothesis is not uniformly accepted. Some authors, for example, argue that species populations are more likely to diverge in response to uniform selection pressures because of variation in genetic background (Wade & Goodnight, 1998; Goodnight, 2000; Levin, 2000). Unfortunately, we know little about the effects of intraspecific variation in genetic background on the responses of alleles to selection. Likewise, few studies have assessed the possibility of parallel evolution at the genotypic level. Nonetheless, accumulating evidence seems most consistent with a deterministic view of adaptive evolution, in which the effects of genetic background are minor and parallel genotypic changes are common. (Note that there is general agreement that genetic background or epistasis plays a significant role in the generation of hybrid incompatibilities in crosses between genetically divergent populations or species. It is the contribution of epistasis to adaptive evolution that is controversial.)

Until recently, evidence for genetic background effects or epistasis was primarily inferred from studies that partition phenotypic variance in populations of known pedigree (Wolf & al., 2000). In these studies, the component of phenotypic variance not accounted for by additive, dominance, or environmental effects was assumed to be epistatic. However, recent advances in molecular techniques have made it possible to measure interactions among individual loci, providing more confident discrimination between epistatic and other components of phenotypic variation. One method tests for interactions among loci that are already known or suspected to contribute to a particular phenotype. This "candidate locus" approach has several advantages (reviewed in Templeton, 2000), but its use has largely been restricted to humans and fruit flies—organisms for which numerous genes have been cloned and sequenced. An alternative method involves the use of mapped molecular markers as surrogates for linked QTLs. Interactions among marker regions are used to detect epistasis. We focus on evidence that has accumulated by using this QTL approach.

Early searches for epistasis using QTL methods were disappointing due to statistical difficulties (Tanksley, 1993; Lynch & Walsh, 1998). In particular, the F_2 mapping strategy that was promoted (e.g., Lander & Botstein, 1989) and widely employed (Tanksley, 1993) was poorly designed for the detection of epistasis because of unequal sizes of two-locus genotypic classes. A second problem was the numerous, semi-independent comparisons required to detect multilocus interactions in a genome-wide scan. Multiple comparisons increase the likelihood of false positives. This problem was minimised in the early QTL literature by restricting analyses to previously identified QTLs (thus reducing the number of comparisons) or by employing very stringent statistical criteria. The former approach created a bias toward finding QTLs with large individual effects and against QTLs whose

Table 2. Selection differentials for phenotypic traits and QTLs. The upper panel refers to interspecific trait differences, whereas the lower panel refers to selection on intraspecific variation. Plausible values for the selective advantage (s) of individual alleles are also provided. For individual QTLs, the selection differential was halved to estimate s for individual alleles. To provide a crude estimate of s for quantitative traits, we first calculated the average percentage variance explained (PVE) by the leading QTL for 50 traits (average PVE = 31.1%) in seven empirical studies that involved interspecific or inter-subspecific crosses between plant taxa and for which mapping populations were fairly large (> 225 plants; Paterson & al., 1998, 1991; Doebley & Stec, 1993; Lin & al., 1995; Lin & Ritland, 1997; Bradshaw & al., 1998; Kim & Rieseberg, 1999). If we assume a similar genetic architecture for the phenotypic traits reported on below, s can be calculated by multiplying the selection differential for the trait by 0.31 and then halved to account for diploidy.

Taxa	Trait/QTL	Selection Differential (s)	s	Reference
<i>Ipomopsis aggregata</i> (Pursh) V.E. Grant x <i>I. tenuituba</i> Rydb.	corolla width	0.21/-0.02	0.03/-	Campbell & al., in press
	corolla length	0.29/-0.27	0.05/-0.04	
	In optical density	0.56/0.05	0.09/0.008	
	flower colour	0.64/0.40	0.10/0.06	
<i>Mimulus cardinalis</i> Dougl. ex Benth. x <i>M. lewisii</i> Pursh	carotenoid concentration QTL	0.80	0.40	Meléndez-Ackerman & Campbell, 1998
	nectar production QTL	0.50	0.25	Schemske & Bradshaw, 1999
<i>Mimulus guttatus</i> DC. x <i>M. nasutus</i> Greene	corolla width	-0.31	-0.05	Fenster & Ritland, 1994
	corolla tube length	-0.02	-0.003	
	stigma/anther separation	-0.26	-0.04	
<i>Agrostis tenuis</i> Vasey	metal tolerance	0.53/-1.00	0.08/-0.16	McNeilly, 1968
	metal tolerance	0.62/0.78	0.10/0.12	Hickey & McNeilly, 1975
<i>Anthoxanthum odoratum</i> L.	metal tolerance	0.78/0.99	0.12/0.15	Hickey & McNeilly, 1975
<i>Lotus corniculatus</i> L. ²	cyanogenesis	0.55/0.88	0.14/0.22	Foulds, 1977
<i>Mimulus guttatus</i> DC.	corolla width	-0.05/0.83	-0.01/0.13	
	corolla tube length	-0.72/-0.23	-0.11/-0.04	Fenster & Ritland, 1994
	stigma/anther separation	0.01/0.03	0.002/0.005	
<i>Phlox drummondii</i> Hook. ³	corolla colour	0.16	0.03	Levin, 1969
<i>Phlox pilosa</i> L.	corolla colour	0.30	0.05	Levin & Kerster, 1967
<i>Plantago lanceolata</i> L.	metal tolerance	0.93/0.11	0.14/0.02	Hickey & McNeilly, 1975
<i>Rumex acetosa</i> L.	metal tolerance	0.61/0.74	0.09/0.11	Hickey & McNeilly, 1975
<i>Salvia columbariae</i> Benth. ⁴	seed colour	0.87	0.13	Brayton & Capon, 1980
<i>Trifolium repens</i> L. ²	cyanogenesis	0.59/0.48	0.15/0.12	Foulds, 1977

¹Multiple selection differentials are given when selection was measured in different habitats, multiple experiments, or with different pollinators.

²Cyanogenesis is a two-locus trait in *Lotus* and *Trifolium*. Therefore, assuming that each locus makes an equivalent contribution to fitness, we halved the total selection differential to estimate the selection differential for a single locus. We then halved the single locus selection differential to estimate s for a single allele.

³The selection differential for *Phlox drummondii* was calculated as the mean selection differential between all possible pairs of seven different flower colours within this species.

⁴The genetic basis of seed color in *Salvia columbariae* is unknown. Therefore, to provide a conservative estimate of s , we treated it as a quantitative trait.

effects were primarily epistatic. The latter approach reduced the likelihood of finding epistatic QTLs, regardless of their individual effects. A third difficulty was recombination between the molecular markers used as surrogates for QTLs in statistical analyses and the QTLs themselves. The further the markers were from the QTLs, the weaker the test. Finally, there was a focus on detecting significant epistatic interactions rather than on asking whether interaction effects actually exceeded the individual effects of the interacting QTLs. Background dependent responses to selection will only occur when epistatic effects exceed those of individual QTLs.

Recent studies of epistasis have employed a variety of strategies to minimise these difficulties. Modified mapping designs such as backcross populations, recombinant inbred lines, and near isogenic lines reduce problems associated with unequal sample sizes, control for genetic background effects outside of regions of interest, and/or minimise recombination between markers and QTLs. Also, statistical approaches have been developed that provide biologically realistic statistical thresholds for genome-wide scans of epistatic interactions (Cheverud, 2000). However, even with these improved methods, the number of significant epistatic interactions detected for most traits in most studies has been close to that expected by chance (reviewed in Lynch & Walsh, 1998). Strong epistasis has sometimes been reported (Doebley & al., 1995; Lark & al., 1995; Cockerham & Zeng, 1996; Eshed & Zamir, 1996; Routman & Cheverud, 1997; Yu & al., 1997; Lukens & Doebley, 1999; Yamamoto & al., 2000), but it appears to be infrequent.

A typical situation is exemplified by a study of 15 phenotypic traits that differentiate two wild sunflower species, *Helianthus annuus* L. and *H. debilis* Nutt. (Kim & Rieseberg, 1999). Significant epistasis was observed for six traits, but only two of the detected interactions (one each for stem pigmentation and phyllary pubescence) had magnitudes that exceeded the individual effects of the interacting loci and thus are likely to affect adaptive evolution (Kim & Rieseberg, in press).

A more direct test of how genetic background impacts the response of alleles to selection has recently been conducted by Ungerer (2000). He analysed the response of the *Arabidopsis* genome to three generations of selection for increased viability and fertility. QTL alleles governing life history traits, as well as other genomic regions, almost always responded uniformly to selection even when in populations with different genetic backgrounds. This study is important because it represents the first genome-wide survey of genetic background effects within a plant species and supports a model of adaptive evolution in which allelic selection values are not strongly influenced by variation in genetic background.

Animal studies also support the view that parallel selection rarely leads to significant diversification. For example, Rice & Hostert (1993) cite 13 studies in which divergent selection in allopatry led to the evolution of prezygotic and/or postzygotic isolation in experimental populations of fruit flies or house flies (isolation failed to evolve in four studies). In contrast, when divergent selection was replaced by parallel selection, isolation consistently failed to evolve (Kilias & al., 1980; Dodd, 1989).

Experimental evidence for minimal genetic background effects is complemented by studies of historical evolutionary changes in isolated populations. Under Fisher's

infinitesimal model, parallel genotypic changes were assumed to be rare and of little evolutionary importance (Fisher, 1930). This view is no longer tenable in the light of (1) empirical evidence for the oligogenic inheritance of species' differences as discussed above, and (2) several excellent examples of independent trait evolution involving the same genetic loci and even the same nucleotide substitutions (see below).

One of the best examples of parallel genotypic changes within a plant species comes from a study of heavy metal tolerance in *Silene vulgaris* Garcke, in which heavy metal tolerance evolved independently at the same major loci in Germany and Ireland (Schat & al., 1996). Parallel evolution has been documented for more distantly related taxa as well. For example, Paterson & al. (1995) showed that the same major QTLs have contributed to the domestication of three major cereals (sorghum, rice, and maize). That is, lineages separated by as many as 50 million years appeared to have responded in essentially the same way to selection.

The evidence for parallel evolution is even more striking when measured at the nucleotide level. For example, half of the amino acid substitutions that accumulated in two bacteriophage lines under strong uniform selection were identical (Whichman & al., 1999). Likewise, evidence of parallel nucleotide substitutions has been documented for HIV isolates from patients undergoing drug therapy (Crandall & al., 1999), as well as for stomach lysozyme sequences from ruminant species (Zhang & Kumar, 1997). The bottom line is that parallel genotypic changes are likely to be a common outcome of uniform selection, although differences are expected as well. Presumably, these parallel changes will occur most frequently in closely related individuals or populations because of greater similarity in ecological and genetic environment, but as of yet there is no empirical evidence relevant to this prediction. With the accumulation of QTL and sequence data from an ever-widening pool of plant and animal populations, it should become feasible to estimate the degree to which parallel phenotypic changes result from parallel genotypic changes, as well as the sensitivity of this relationship to variation in the ecological and genetic environment.

Discussion

The ongoing discussion of the species problem has largely shifted from a scientific debate based on facts about the natural world to a philosophical debate, which relies largely on clever arguments that are often disengaged from biological reality. New information about species is rarely incorporated into these discussions, and the biology that is included tends to be dated and qualitative. This is unfortunate since it is unlikely that progress toward understanding species will derive from further philosophical debate. Rather, we need to carefully consider data that have been gathered over the past 30 years and incorporate it into ongoing discussions. In addition, we need to determine what new information would move us closer toward a solution to this problem.

This paper compiles information about several key parameters that are critical to the debate about whether plant species are actively evolving units and thus fundamentally different than higher taxa, or whether they are simply aggregates of actively evolving units. The parameters considered include (1) levels of gene flow

within plant species, (2) the magnitude of selection coefficients for advantageous alleles that contribute to species' differences, (3) the role of genetic background variation in mediating responses to selection, and (4) parallel genotypic evolution. We found that there is sufficient gene flow to allow the efficient spread of strongly favourable alleles ($s > 0.05$) among populations. Moreover, estimates of s for major QTL alleles are sufficiently large to enable their spread across the range of a species, although minor QTL alleles seem more likely to evolve locally. In addition, our review of the literature suggests that intraspecific variation in genetic background infrequently affects allelic selective responses, thereby increasing the opportunity for the collective evolution of populations through parallel responses to selection. The probable prevalence of this latter means of collective evolution was illustrated by several striking examples of parallel evolution at the genotypic level.

Although our arguments favouring the biological reality of species may seem similar to those of other authors that have emphasised evolutionary genetic processes (e.g., Dobzhansky, 1937; Mayr, 1963; Grant, 1971; Templeton, 1989), there are significant differences. Most of these authors have focused on the conservative role of gene flow in preventing diversification. However, gene flow in selfing plant species is not sufficient to prevent diversification due to drift, and even very weak local selection should be sufficient to overcome the homogenising effects of gene flow in species with mixed mating or outcrossing breeding systems. Thus, the primary contribution of this article may be its emphasis on how gene flow acts creatively by enabling the diffusion of favourable mutations. In this regard, it is of interest that Ehrlich & Raven (1969) recognised the ease with which selection could overcome gene flow, but were pessimistic with regard to the diffusion of advantageous alleles, noting that the fate of a new allele will "presumably be governed by the kind of gloomy odds facing mutant genes". However, as we have shown in this paper, the odds are not nearly as gloomy as was earlier presumed, primarily because of empirical evidence for an oligogenic basis of species' differences. Finally, our work differs from earlier papers in that we have presented quantitative estimates of gene flow and selection coefficients rather than relying entirely on verbal arguments that derive from a few selected examples.

Our discussion of selection also differs from that of prior authors in its emphasis on parallel responses to selection. In the past, most students of speciation have tended to focus on either the role of selection as a conservative force, preventing phenotypic or genetic evolution (e.g., Ehrlich & Raven, 1969; Lande, 1980; Templeton, 1989), or as a diversifying force, leading to the divergence of population systems (Levin, 2000). Our emphasis on parallel evolution is not intended to minimise the importance of either the stabilising or diversifying role of selection. Rather, we feel that the frequency of parallel responses to selection and their contribution to collective evolution have been underestimated by most students of speciation. One objective of this paper was to address this imbalance.

The recognition that species are biologically real entities does not necessarily make them easier to recognise. What we need is a better understanding of the processes responsible for the collective evolution of species populations. Specific needs include (1) data on the frequency with which favourable mutations arise and the likely distribution of mutational effects (see Orr, 1998 for a theoretical treatment

of the latter); (2) better models for the spread of these mutations across species' ranges; (3) greater information on the ages of species and their historical distributions and demographies; (4) better estimates of the contributions of different forms of selection to the collective evolution of species; and (5) a more sophisticated understanding of how variation in the ecological and genetic environment affect responses to selection. Our optimistic view is that once we understand the means by which species' populations evolve collectively, it should be straightforward to find genetic or phenotypic signatures of these processes that can then be employed to recognise species. These signatures may already be employed by some of the more widely accepted species concepts. For example, we suspect that prezygotic reproductive barriers serve as a good indicator of when isolated groups of populations are no longer likely to evolve collectively as a result of gene flow. Postzygotic barriers probably mark the point at which isolated populations are more likely to diverge than evolve in parallel when responding to similar selective pressures. We also expect the more historically oriented definitions of species to function well when based on the genealogies of loci that effect the collective evolution of population systems (e.g., loci that contribute to reproductive isolation). Alleles for isolation loci should coalesce early in the speciation process, and reproductively isolated groups will likely be reciprocally monophyletic and diagnosable at these loci (Ting & al., 2000). Thus, essentially all versions of the phylogenetic species concept would recognise the same species, which would be recovered under the biological species concept as well.

In conclusion, considerable progress is being made on the species problem, but mostly through empirical study rather than philosophical debate. As a greater number of genes that contribute to speciation are genetically mapped and cloned, we expect to recover a wealth of knowledge regarding the effects of individual genes on assortative mating and fitness in nature. We also expect that sequencing the timing and demography of speciation. What we request from biologists is the wisdom to recognise the need for more information, the drive to gather the needed data, and the patience to hold out for these data before concluding that the species' problem is intractable.

Acknowledgements

We thank R. Rieseberg for editorial suggestions and the NIH, NSF, and USDA for supporting our research on species and speciation.

Literature cited

- Bachmann, K. 1998. Species as units of diversity: an outdated concept. *Theor. Biosci.* 117: 213–230.
 Barton, N. H. In press. The role of hybridisation in evolution. *Molec. Ecol.*
 Batjes J., Chambers K. L. & Bachmann, K. 1994. Evolution of microsporangium numbers in *Microseris* (Asteraceae: Lactuceae). *Amer. J. Bot.* 81: 641–647.
 Beavis, W. D. 1998. QTL analyses: power, precision, and accuracy. Pp. 145–162 in: Paterson, A. (ed.), *Molecular dissection of complex traits*. Boca Raton, Florida.
 Boshier, D. H., Chase, M. R. & Bawa, K. S. 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 2. Mating system. *Amer. J. Bot.* 82: 476–483.

- Bossart, J. L. & Prowell, D. P. 1998. Genetic estimates of population structure and gene flow: limitations, lessons, and new directions. *Trends Ecol. Evol.* 13: 202–206.
- Bradshaw, H. D., Jr., Otto, K. G., Frewen, B. E., McKay, J. K. & Schemske, D. W. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* 149: 367–382.
- , Wilbert, S. M., Otto, K. G. & Schemske, D. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* 376: 762–765.
- Brayton, R. D. & Capon, B. 1980. Productivity, depletion, and natural selection of *Salvia columbariae* seeds. *Aliso* 9: 581–587.
- Broyles, S. B., Schnabel, A. & Wyatt, R. 1994. Evidence for long-distance dispersal in milkweeds (*Asclepias exaltata*). *Evolution* 48: 1032–1040.
- Cain, M. L., Damman, H. & Muir, A. 1998. Seed dispersal and the Holocene migration of woodland herbs. *Ecol. Monogr.* 68: 325–347.
- , Milligan, B. L. & Strand, A. E. 2000. Long-distance seed dispersal in plant populations. *Amer. J. Bot.* 87: 1217–1227.
- Campbell, D. R., Waser, N. M. & Meléndez-Ackerman, E. J. In press. Analyzing pollinator-mediated selection in a plant hybrid zone: hummingbird visitation patterns on three spatial scales. *Amer. Naturalist*.
- Chase, M. R., Moller, C., Kesseli, R. & Bawa, K. S. 1996. Distant gene flow in tropical trees. *Nature* 383: 398–399.
- Cheverud, J. M. 2000. Detecting epistasis among quantitative trait loci. Pp. 58–81 in: Wolf, J., Brodie, B. & Wade, M. J. (eds.), *Epistasis and the evolutionary process*. New York.
- Claridge, M. F., Dawah, H. A. & Wilson, M. R. (eds.) 1997. *Species: the units of biodiversity*. New York.
- Clark, J. S., Silman, M., Kern, R., Macklin, E. & HilleRisLambers, J. 1999. Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology* 80: 1475–1494.
- Cockerham, C. C. & Zeng, Z.-B. 1996. Design III with marker loci. *Genetics* 143: 1437–1456.
- Coyne, J. A. 1992. Much ado about species. *Nature* 357: 289–290.
- Crandall, K. A., Kelsey, C. R., Imamichi, H., Lane, H. C. & Salzman, N. P. 1999. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate to detect selection. *Molec. Biol. Evol.* 16: 372–382.
- de Pinna, M. C. C. 1999. Species concepts and phylogenetics. *Rev. Fish Biol. Fisher.* 9: 353–373.
- Devlin, B. & Ellstrand, N. C. 1990. The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* 44: 248–259.
- Dobzhansky, T. 1937. *Genetics and the origin of species*. New York.
- Dodd, D. M. B. 1989. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution* 39: 1388–1392.
- Doobley, J. & Stec A. 1993. Inheritance of the morphological differences between maize and teosinte: comparison of results for two F₂ populations. *Genetics* 134: 559–570.
- , – & Gustus, C. 1995. *Teosinte branched* and the origin of maize—evidence for epistasis and the evolution of dominance. *Genetics* 141: 333–346.
- Dow, B. D. & Ashley, M. V. 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molec. Ecol.* 5: 617–627.
- & – 1998. High levels of gene flow in oak revealed by paternity analysis using microsatellites. *J. Heredity* 89: 62–70.
- Ehrlich, P. R. & Raven, P. H. 1969. Differentiation of populations. *Science* 165: 1228–1232.
- El-Kassaby, Y. A., Rudin, D. & Yazdani, R. 1989. Levels of outcrossing and contamination in two *Pinus sylvestris* L. seed orchards in northern Sweden. *Scand. J. Forest Res.* 4: 41–49.
- Ellstrand, N. C. 1992. Gene flow among seed plant populations. *New Forest* 6: 241–256.
- , Devlin, B. & Marshall, D. L. 1989. Gene flow by pollen into small populations—data from experimental and natural stands of wild radish. *Proc. Natl. Acad. U.S.A.* 86: 9044–9047.
- El-Mousadik, A. & Petit, R. J. 1996. Chloroplast DNA phylogeography of the argan tree of Morocco. *Molec. Ecol.* 5: 547–555.
- Ennos, R. A. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* 72: 250–259.
- Ereshfsky, M. 1999. Species and the Linnaean hierarchy. Pp. 285–306 in: Wilson, R. A. (ed.), *Species: new interdisciplinary essays*. Cambridge, Massachusetts.
- Eshed, Y. & Zamir, D. 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143: 1807–1817.
- Fenster, C. B. & Ritland, K. 1994. Evidence for natural selection on mating system in *Mimulus* (*Scrophulariaceae*). *Int. J. Plant Sci.* 155: 588–596.
- Fisher, R. 1930. *The genetical theory of natural selection*. Oxford.
- Foulds, W. 1977. The physiological response to moisture supply of cyanogenic and acyanogenic phenotypes of *Trifolium repens* L. and *Lotus corniculatus* L. *Heredity* 39: 219–234.
- Friedman, S. T. & Adams, W. T. 1985. Estimation of gene flow into two seed orchards of loblolly pine (*Pinus taeda* L.). *Theor. Appl. Genet.* 69: 609–615.
- Godt, M. J. W. & Hamrick, J. L. 1993. Patterns and levels of pollen-mediated gene flow in *Lathyrus latifolius*. *Evolution* 47: 98–100.
- Goodnight, C. J. 2000. Quantitative trait loci and gene interaction: the quantitative genetics of metapopulations. *Heredity* 84: 587–598.
- Gottlieb, L. D. 1984. Genetics and morphological evolution in plants. *Amer. Naturalist* 123: 681–709.
- Grant, V. 1971. *Plant speciation*. New York.
- 1980. Gene flow and the homogeneity of species populations. *Biol. Zentralbl.* 99: 157–169.
- Hamrick, J. L. & Godt, M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 351: 1291–1298.
- & Murawski, D. A. 1990. The breeding structure of tropical tree populations. *Pl. Spec. Biol.* 5: 157–165.
- & Nason, J. D. 2000. Gene flow in forest trees. Pp. 81–111 in: Young, A. & Boshier, D. (eds.), *Forest conservation genetics*. Collingwood, Australia.
- Handel, S. N. 1982. Dynamics of gene flow in an experimental population of *Cucumis melo* (*Cucurbitaceae*). *Amer. J. Bot.* 69: 1538–1546.
- Harju, A. & Muona, O. 1989. Background pollination in *Pinus sylvestris* seed orchards. *Scand. J. Forest Res.* 4: 513–520.
- Hewitt, G. M. 1996. Some consequences of ice ages and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58: 247–276.
- Hickey, D. A. & McNeilly, T. 1975. Competition between metal tolerant and normal plant populations: a field experiment on normal soil. *Evolution* 29: 458–464.
- Homborgen, E.-J. & Bachmann, K. 1994. RAPD mapping of three QTLs determining trichome formation in *Microris hybrid H27* (*Asteraceae: Lactuceae*). *Theor. Appl. Genet.* 90: 853–858.
- Howe, H. F. & Smallwood, J. 1982. Ecology of seed dispersal. *Annual Rev. Ecol. Syst.* 13: 201–228.
- Hull, D. L. 1999. On the plurality of species: questioning the party line. Pp. 23–48 in: Wilson, R. A. (ed.), *Species: new interdisciplinary essays*. Cambridge, Massachusetts.
- Isagi, Y., Kanazashi, T., Suzuki, W., Tanaka, H. & Abe, T. 2000. Microsatellite analysis of the regeneration process of *Magnolia obovata* Thunb. *Heredity* 84: 143–151.
- Kaufman, S. R., Smouse, P. E. & Alvarez-Buylla, E. R. 1998. Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (*Moraceae*): a paternity analysis. *Heredity* 81: 164–173.
- Killias, G., Alahiotis, S. N. & Pelecanos, M. 1980. A multifactorial genetic investigation of speciation theory using *Drosophila melanogaster*. *Evolution* 34: 730–737.
- Kim, S.-C. & Rieseberg, L. H. 1999. Genetic architecture of species differences in annual sunflowers: implications for adaptive trait introgression. *Genetics* 153: 965–977.
- & – In press. The contribution of epistasis to species differences in annual sunflowers. *Molec. Ecol.* 10.
- Kirkpatrick, K. J. & Wilson, H. D. 1988. Interspecific gene flow in *Cucurbita*: *C. texana* and *C. pepo*. *Amer. J. Bot.* 75: 519–527.

- Krauss, S. L. 2000. Patterns of mating in *Persoonia mollis* (Proteaceae) revealed by an analysis of paternity using AFLP: implications for conservation. *Austral. J. Bot.* 48: 349–356.
- Lande, R. 1980. The response to selection on major and minor mutations affecting a metrical trait. *Amer. Naturalist* 116: 463–479.
- 1981. The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* 99: 541–553.
- Lander, E. S. & Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199.
- Lark, K. G., Chase, K., Alder, F., Mansure, L. M. & Orf, J. F. 1995. Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. *Proc. Natl. Acad. U.S.A.* 92: 4656–4660.
- Latta, R. G., Linhart, Y. B., Fleck, D. & Elliot, M. 1998. Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine. *Evolution* 52: 61–67.
- & Mitton, J. B. 1997. A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). *Genetics* 146: 1153–1163.
- Levin, D. A. 1969. The effect of corolla color and outline on interspecific pollen flow in *Phlox*. *Evolution* 23: 444–455.
- 1979. The nature of plant species. *Science* 204: 381–384.
- 1981. Dispersal versus gene flow in plants. *Ann. Missouri Bot. Gard.* 68: 233–253.
- 2000. *The origin, expansion, and demise of plant species*. New York.
- & Kerster, H. W. 1967. Natural selection for reproductive isolation in *Phlox*. *Evolution* 21: 679–687.
- & Wilson, A. C. 1976. Rates of chromosomal evolution in seed plants: net increase in diversity of chromosome numbers and species numbers through time. *Proc. Natl. Acad. U.S.A.* 73: 2086–2090.
- Lin, J. Z. & Ritland, K. 1997. Quantitative trait loci differentiating the outbreeding *Minulus guttatus* from the inbreeding *M. platycalyx*. *Genetics* 146: 1115–1121.
- Lin, R.-R., Schertz, K. R. & Paterson, A. H. 1995. A comparative analysis of QTLs affecting plant height and maturity across the *Poaceae*, in reference to an interspecific sorghum mapping population. *Genetics* 141: 391–411.
- Loveless, M. D., Hamrick, J. L. & Foster, R. B. 1998. Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. *Heredity* 81: 134–143.
- Luckow, M. 1995. Species concepts—assumptions, methods, and applications. *Syst. Bot.* 20: 589–605.
- Lukens, L. H. & Doebley, J. 1999. Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. *Genetical Research* 74: 291–302.
- Lynch, M. & Walsh, B. 1998. *Genetics and analysis of quantitative traits*. Sunderland, Massachusetts.
- Massey, L. K. & Hamrick, J. L. 1999. Breeding structure of a *Yucca filamentosa* (Agavaceae) population. *Evolution* 53: 1293–1298.
- Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. Pp. 381–424 in: Claridge, M. F., Dawah, H. A. & Wilson, M. R. (eds.), *Species: the units of biodiversity*. New York.
- Mayr, E. 1963. *Animal species and evolution*. Cambridge, Massachusetts.
- McCauley, D. E. 1997. The relative contributions of seed and pollen movement to the local genetic structure of *Silene alba*. *J. Heredity* 88: 257–263.
- McNeilly, T. 1968. Evolution in closely adjacent populations. III. *Agrostis tenuis* on a small copper mine. *Heredity* 23: 99–108.
- Meléndez-Ackerman, E. & Campbell, D. R. 1998. Adaptive significance of flower color and inter-trait correlations in an *Ipomopsis* hybrid zone. *Evolution* 52: 1293–1303.
- Mishler, B. D. 1999. Getting rid of species. Pp. 307–316 in: Wilson, R. A. (ed.), *Species: new interdisciplinary essays*. Cambridge, Massachusetts.
- & Donoghue, M. J. 1982. Species concepts: a case for pluralism. *Syst. Zool.* 31: 503–511.

- Nagasaka, K. & Szmidt, A. E. 1984. Multilocus analysis of external pollen contamination of a Scots pine (*Pinus sylvestris* L.) seed orchard. No. VII. Pp. 1–10 in: Szmidt, A. E. (ed.), *Genetic studies of Scots Pine* (*Pinus sylvestris* L.) domestication by means of isozyme analysis. Umeå.
- Nason, J. D. & Hamrick, J. L. 1997. Reproductive and genetic consequences of forest fragmentation: two case studies of neotropical canopy trees. *J. Heredity* 88: 264–276.
- Neale, D. B. 1984. Population genetic structure of Douglas fir shelterwood regeneration system in southwest Oregon. Ph.D. dissert. Oregon State Univ., Corvallis, Oregon.
- Neubert, M. G. & Caswell, H. 2000. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology* 81: 1613–1628.
- Orr, A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* 52: 935–949.
- Paterson, A. H., Danon, S., Hewitt, J. D., Zamir, D., Rabinovitch, H. D., Lincoln, S. E., Lander, E. S. & Tanksley, S. D. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics* 127: 181–197.
- , Lander, E. S., Hewitt, J. D., Peterson, S., Lincoln, S. E. & Tanksley, S. D. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335: 721–726.
- , Lin, Y.-R., Li, Z., Schertz, K. F., Doebley, J. F., Pinson, R. M., Lu, S.-C., Stansel, J. W. & Irvine, J. E. 1995. Convergent domestication by independent mutations at corresponding genetic loci. *Science* 269: 1714–1718.
- Rice, W. R. & Hostert, E. E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* 47: 1637–1653.
- Richards, C. M., Church, S. & McCauley, D. E. 1999. The influence of population size and isolation on gene flow by pollen in *Silene alba*. *Evolution* 53: 63–73.
- Routman, E. J. & Cheverud, J. M. 1997. Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. *Evolution* 51: 1654–1662.
- Schaal, B. A. 1980. Measurement of gene flow in *Lupinus texensis*. *Nature* 248: 450–451.
- Schat, H., Vooijs, R. & Kuiper, E. 1996. Identical major gene loci for heavy metal tolerances have independently evolved in different local populations and subspecies of *Silene vulgaris*. *Evolution* 50: 1888–1895.
- Schemske, D. W. 2000. Understanding the origin of species. *Evolution* 54: 1069–1073.
- & Bradshaw, H. D., Jr. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Minulus*). *Proc. Natl. Acad. U.S.A.* 96: 11910–11915.
- Schnabel, A. & Hamrick, J. L. 1995. Understanding the population genetic structure of *Gleditsia triacanthos* L.: the scale and pattern of pollen gene flow. *Evolution* 49: 921–931.
- Schuster, W. S. F. & Mitton, J. B. 2000. Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity* 84: 348–361.
- Slackin, M. 1976. The rate of spread of an advantageous allele in a subdivided population. Pp. 767–780 in: Karlin, S. & Nevo, E. (eds.), *Population genetics and ecology*. New York.
- 1987. Gene flow and the geographic structure of natural populations. *Science* 236: 787–792.
- Smith, D. B. & Adams, W. T. 1983. Measuring pollen contamination in clonal seed orchards with the aid of genetic markers. Pp. 69–77 in: Guries, R. P. (ed.), *Proceedings of the 17th Southern Forest Tree Improvement Conference*. Athens.
- Sork, V. L., Nason, J., Campbell, D. R. & Fernandez, J. F. 1999. Landscape approaches to historical and contemporary gene flow in plants. *Trends Ecol. Evol.* 14: 219–224.
- Stacy, E. A., Hamrick, J. L., Nason, J. D., Hubbell, S. P., Foster, R. B. & Condit, R. 1996. Pollen dispersal in low density populations of three neotropical tree species. *Amer. Naturalist* 148: 275–298.
- Tanksley, S. D. 1993. Mapping polygenes. *Annual Rev. Genet.* 27: 205–233.
- Templeton, A. R. 1989. The meaning of species and speciation: a genetic perspective. Pp. 3–27 in: Otte, D. & Endler, J. A. (eds.), *Speciation and its consequences*. Sunderland, Massachusetts.
- 2000. Epistasis and complex traits. Pp. 41–57 in: Wolf, J., Brodie, B. & Wade, M. J. (eds.), *Epistasis and the evolutionary process*. New York.

- Ting, C. T., Tsaui, S. C. & Wu, C. I. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc. Natl. Acad. U.S.A.* 97: 5313–5316.
- Ungerer, M. 2000. Selection, genetic environment, and adaptive evolution: an analysis of microevolutionary dynamics in experimental plant populations. Ph.D. dissert. Indiana Univ., Bloomington.
- Van Houten, W., Van Raamsdonk, L. & Bachmann, K. 1994. Interspecific evolution of *Microseris pygmaea* (Asteraceae, Lactuceae) analyzed by cosegregation of phenotypic characters (QTLs) and molecular markers (RAPDs). *Pl. Syst. Evol.* 190: 49–67.
- Vlot, E. C., Van Houten, W. H., Mauthe, S. & Bachmann, K. 1992. Genetic and nongenetic factors influencing deviations from five pappus parts in a hybrid between *Microseris douglasii* and *M. bigelovii* (Asteraceae, Lactuceae). *Int. J. Pl. Sci.* 153: 89–97.
- Wade, M. J. & Goodnight, C. J. 1998. Perspective: The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. *Evolution* 52: 1537–1553.
- Wheeler, N. & Jech, K. 1986. Pollen contamination in a mature Douglas-fir seed orchard. Pp. 160–171 in: Hatcher, A. V. & Weir, R. J. (eds.) *Proceedings. IUFRO conference on breeding theory, progeny testing and seed orchards*. Williamsburg.
- Wheeler, Q. D. & Meier, R. (eds.). 2000. *Species concepts and phylogenetic theory*. New York.
- Whitchman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. 1999. Different trajectories of parallel evolution during viral adaptation. *Science* 285: 422–424.
- Whitlock, M. C. & McCauley, D. E. 1999. Indirect measures of gene flow and migration. *Heredity* 92: 117–125.
- Whitton, J., Wolf, D. E., Arias, D. M., Snow, A. A. & Rieseberg, L. H. 1997. The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. *Theor. Appl. Genet.* 95: 33–40.
- Wilson, R. A. (ed.). 1999. *Species: new interdisciplinary essays*. Cambridge, Massachusetts.
- Wolf, J., Brodie, B. & Wade, M. J. (eds.). 2000. *Epistasis and the evolutionary process*. New York.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.
- Yamamoto, T., Lin, H. X., Sasaki, T. & Yano, M. 2000. Identification of heading date quantitative trait locus *Hd6* and characterization of its epistatic interactions with *Hd2* in rice using advanced backcross progeny. *Genetics* 154: 885–891.
- Yu, S. B., Li, J. X., Xu, C. G., Gao, Y. J., Li, X. H., Zhang, Q. F. & Maroof, M. A. S. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. U.S.A.* 94: 9226–9231.
- Zhang, J. Z. & Kumar, S. 1997. Detection of convergent and parallel evolution at the amino acid sequence level. *Molec. Biol. Evol.* 14: 527–536.