Evolution, 35(1), 1981, pp. 124–138

SKEPTICISM TOWARDS SANTA ROSALIA, OR WHY ARE THERE SO FEW KINDS OF ANIMALS?[1]

JOSEPH FELSENSTEIN
Department of Genetics, University of Washington, Seattle, Washington 98195

Received January 3, 1980. Revised May 21, 1980

In a classic paper, Hutchinson (1959) set the tone for much of the ecological work done during the past 20 years by suggesting that ecologists try to explain the numbers of species of animals. Hutchinson's approach was ecological, explaining diversity by explaining how the species could coexist. There can be little question that competitive exclusion sets an upper limit on species diversity, but it is not obvious that this upper limit will be achieved. There may be additional constraints on the process of speciation, constraints set by genetics rather than ecology. It has been usual for evolutionists to reject the possibility of sympatric speciation, and this amounts to asserting the existence of such genetic constraints. Even if the ecological opportunity for coexistence is present, under the conventional view lack of geographic isolation can prevent speciation. The attempt by Rosenzweig (1975) to give a comprehensive explanation of continental species diversity takes as its starting point the assumption that geographic isolation is a necessary prerequisite to species formation.

A number of workers have made and analyzed detailed population genetic models of sympatric or parapatric speciation, in particular Maynard Smith (1966), Dickinson and Antonovics (1973), and Caisse and Antonovics (1978). Balkau and Feldman (1973) made a model of migration modification which can also be regarded as a model of parapatric speciation. The upshot of these models is that it is not difficult for sympatric speciation to occur. While these authors have largely been concerned with showing that sympatric speciation is possible, one might come away from some of these papers with the disturbing impression that it is all but inevitable.

If this were the case, one would expect to find nearly infinite numbers of species, a different species on every bush. More precisely, if there were no genetic constraints on speciation, an "island-biogeography" model of speciation could be constructed, resembling that of Rosenzweig but with no requirement for geographical isolation in order to get speciation. There would be a balance between speciation and the extinction of small species. Without having such a model in hand, it is my impression that the number of species in nature is far smaller, and their size far larger, than such a model would predict.

This raises the question of whether there is any genetic constraint on speciation. If there is, we would wish to know in population genetic terms what forces were opposing speciation. This paper presents a simple model of speciation, in an attempt to search for and characterize evolutionary factors acting against speciation. Such a force is found, and in fact can also be seen to have been acting in at least three of the earlier population genetic models. The picture of speciation which emerges here involves a new distinction between two kinds of speciation, one which cuts across the usual allopatric-sympatric spectrum. The distinction made is between speciation in which the reproductive isolating mechanisms come into existence by the substitution of different alleles in the two nascent species, and speciation in which the same alleles are substituted in both species. This distinction seems to be an important one.

Caisse and Antonovics (1978) have dis-

[1] Dedicated to Sewall Wright, in celebration of his 90th birthday.

cussed many of these same issues, and reached conclusions completely consistent with those of this paper. The present model differs from theirs in detail, and may help round out our picture of the genetics of speciation. It must be emphasized that the model presented here is not intended as a realistic genetic model of the speciation process. Rather, it is the simplest model I can find which exhibits many of the genetic effects which will be found in more complex, more realistic models of speciation. The purpose of this paper is to clarify the nature of the genetic forces involved in speciation, and to get some sense of the direction in which they act and their relative strengths.

The Model

We consider an infinite haploid population with discrete generations. Haploidy is considered primarily to avoid mathematical and computational complications: we shall see below that qualitatively identical results are obtained with a diploid model whenever the two cases are compared. We suppose that there are three loci: A, B, and C, each with two alleles. Loci B and C are under natural selection, and locus A controls a cue for assortative mating. We start by considering the case of sympatric speciation. Our model resembles that of Maynard Smith (1966). The population consists of two subpopulations, within each of which there is separate density-dependent regulation of population size, according to the model of Levene (1953).

Each generation the organisms are distributed randomly into the two subpopulations. Selection (which it is convenient to think of as involving differential viability) takes place within each subpopulation. Each subpopulation regulates its density separately, and the survivors of selection and of density regulation emerge into a single mating pool containing individuals from both subpopulations. We take the two subpopulations to have adult populations of equal size (after density regulation), so that they contribute equally to the mating pool.

Natural selection acts at loci B and C according to the following fitness scheme:

Genotype	Subpopulation I	Subpopulation II
BC	$(1 + s)^2$	1
Bc	$1 + s$	$1 + s$
bC	$1 + s$	$1 + s$
bc	1	$(1 + s)^2$

In population I alleles B and C are at an advantage, and in population II alleles b and c are at an advantage. The selection coefficients have all been taken to be s for simplicity. Fitnesses at different loci are assumed to combine multiplicatively. Initially we shall assume that the three loci are unlinked.

If we were only confronted with loci B and C, the outcome would be fairly simple. A globally stable polymorphism will be maintained at both loci, and it will involve gene frequencies of 0.5 for all alleles, as well as coupling linkage disequilibrium (gametic phase disequilibrium), with an excess of BC and bc genomes. The larger the selection coefficient s, the stronger this disequilibrium will be. When s is infinite, only genotypes BC and bc will exist after selection, and each will exist in only one of these two subpopulations.

In effect, the natural selection creates a partial postzygotic isolation between the two types, as expressed in the nonrandom association between B and C. We now add to the model the locus A, which embodies a prezygotic isolating mechanism, assortative mating. We assume that in the mating pool, which is formed from the survivors of selection and density regulation, the frequencies of different mating types are as follows:

	A	a
A	$p^2(1 - d) + pd$	$p(1 - p)(1 - d)$
a	$p(1 - p)(1 - d)$	$(1 - p)^2(1 - d) + (1 - p)d$

This is the simplest possible type of assortative mating. A fraction d of the individuals mate with their own type, and a fraction $1 - d$ mates at random. If the

A locus were the only locus segregating, this type of assortative mating would not result in any change of gene frequency at that locus. Some mechanisms leading to this type of assortative mating are briefly explored in Appendix 1.

We have not yet specified how the assortative mating affects loci B and C. The assumption will be that the genotypes at these loci are carried along passively in the assortative mating. For example, the probability that a mating is ABC × AbC will simply be the overall probability of an A × A mating, times the product of the fraction of all A genomes which are also BC, and the fraction of all A genomes which are also bC.

Possible Outcomes

We now have a model involving postzygotic partial isolation, plus a potential partial prezygotic isolating mechanism. There are two parameters, s and d. We know that a two-locus polymorphism will be maintained at the B and C loci. If there were initially linkage equilibrium (gametic phase equilibrium) between locus A and loci B and C, then it is relatively easy to see what will happen. The natural selection at the latter two loci will have no effect on locus A, which will maintain its initial gene frequency. There will of course be nonrandom mating at locus A, but this in turn will have no implications for loci B and C as long as this linkage equilibrium continues to hold.

It is only when there is initial linkage disequilibrium that we see interaction between the assortative mating and the natural selection. Linkage disequilibrium is particularly critical to this model. We interpret the disequilibrium between loci B and C to be partial postzygotic isolation between two entities, BC and bc, the intermediate forms Bc and bC being less frequent. We are interested in whether the prezygotic isolating mechanism A becomes nonrandomly associated with the genes B and C which demarcate the two entities. We interpret such an association as progress in the direction of speciation.

Of course, the disequilibria between A and loci B and C involve three parameters: the pairwise disequilibria between A and B and between A and C, as well as a third-order disequilibrium parameter which measures whether the disequilibrium between B and C is the same among A genomes as among a. In the numerical iterations, disequilibrium between A and the other loci has been monitored by following the pairwise linkage disequilibrium between loci A and C. In most of the cases run, the symmetries of the initial population composition were such that the disequilibrium between A and B would remain the same as between A and C. In many of these runs, the third-order disequilibrium would be forced to remain zero by these symmetries. The disequilibrium between B and C was followed separately as an indication of the effectiveness of postzygotic isolation. If BC individuals tend to be A as well, and bc to be a, then this association serves to prevent the mating of BC with bc. This in turn reduces the rate of production of the maladapted Bc and bC types. That these are in fact maladapted is seen by an average fitness calculation. The average fitness of BC or bc is

$$1 + s + s^2/2,$$

whereas the average fitness of Bc or bC is $1 + s$, which will always be smaller.

Thus the association of isolating mechanisms increases mean fitness. This in itself does not guarantee that the association will be brought about, and we are particularly interested in cases in which it will not. The reader who is skeptical of whether this association between isolating mechanisms really constitutes a step in the direction of speciation is invited to contemplate the case in which $d = 1$. Then if only ABC and abc were present, they would be completely isolated species. Note of course that the labelling of the alleles at the A locus is arbitrary: we would be equally interested in an association involving an excess of aBC and Abc types.

Numerical Results

The model just described has eight haploid genotypes, so that it is described by seven variables. A formal description of the model in algebraic terms will be found in Appendix 2. While some special cases can be treated analytically, this is quite tedious. A computer program has been written to iterate genotype frequencies, using the language PASCAL on an 8080 (SOL-20) microcomputer. This iteration is entirely deterministic. A variety of different cases have been run. They involve a large range of values of s and of d. Initial gene frequencies at the three loci and initial values of linkage disequilibrium have been less thoroughly explored.

In all cases it has been found that B and C reach equilibrium gene frequencies of one-half, and that there is an excess of BC and bc coupling genomes at equilibrium. When initial linkage disequilibria between A and the B-C complex are introduced, these disequilibria may or may not disappear. Which happens has never proven to be dependent on the sizes of the initial disequilibria. In the case of disequilibrium between A and C (or between A and B), if linkage disequilibrium is destined to become established, it seems to do so even if the initial disequilibrium is very slight. There is, of course, a dependence of the sign of the final disequilibrium on the sign of the initial disequilibrium, but no dependence of the magnitude of the ultimate disequilibrium.

Under the assumption that these observations constitute a general pattern, we can investigate which values of s and d allow the establishment of linkage disequilibrium. Figure 1 shows the division of the (s, d) plane into these two regions. Above the curve is the region in which disequilibrium between A and B-C can become established. As we approach the top of the region (d = 1) we find more and more complete disequilibrium. As we approach the curve from above, the disequilibrium between A and B (and also that between A and C) becomes less. Below the curve no disequilibrium will persist.

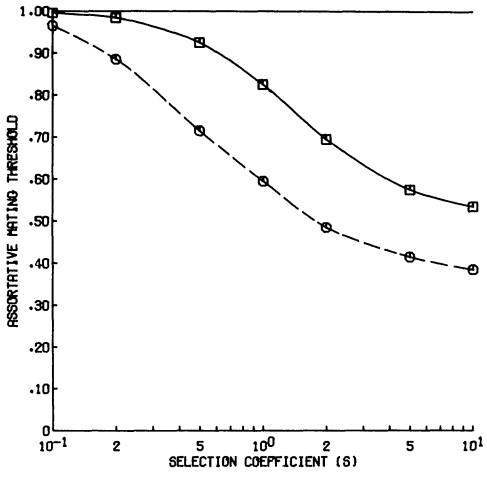

FIG. 1. The minimum amount of assortative mating (d) required to get stable association between the locus A and loci B and C, plotted for various selection coefficients in a haploid moel. The squares (interpolated by the solid curve) are for m = 0.5, the circles (interpolated by the dashed curve) are for m = 0.1.

The immediate conclusion which we can draw from these results is that it is possible to construct a sympatric speciation model which sometimes does not speciate. (As we shall see, this property was also possessed by some earlier models.) Speciation in this case requires sufficiently strong selection and an isolating mechanism of sufficient potential strength. There is a complex tradeoff between the two. Note that the selection required in order that progress be made toward speciation is very strong, the major dip in the curve in Figure 1 being in the vicinity of s = 1.

We are interested in this model, not as a particularly realistic model of speciation but as a means of investigating the direction in which various evolutionary forces are operating. In order to look into this, it is necessary to alter various aspects of the model. The first alteration will involve the amount of gene flow between the two subpopulations.

TABLE 1. *Threshold values of the amount of assortative mating in various cases involving different amounts of migration and linkage. The threshold value of d lies between the two values given in the appropriate column. SMR and SRM refer to the two orderings of events (Selection, Migration, and Recombination) in the life cycle.*

				d	
s	r_{12}	r_{23}	m	SMR	SRM
1.0	.5	.5	.5	.82–.83	.81–.82
1.0	.5	.5	.1	.59–.60	.57–.58
1.0	.5	.5	.01	.51–.55	.53–.54
1.0	.1	.5	.1	.28–.29	.27–.28
1.0	.5	.1	.1	.86–.87	.86–.87

Restricting Migration

The dashed curve on Figure 1 shows the result when we restrict migration between the two subpopulations at the point in the life cycle following selection and density regulation. Instead of having individuals from both subpopulations contributed to a common mating pool, we make a conventional migration model. Each generation, there is exchange of individuals between the two subpopulations. Following this migration, a fraction m of each subpopulation will consist of new arrivals from the other subpopulation. Mating and recombination follow this migration. The case of a single mating pool will be seen to be the same as that in which m = 0.5.

The dashed curve is for the case of m = 0.1, which is in effect a case of parapatric speciation rather than sympatric speciation. In every case which has been investigated, restriction of migration eases the conditions for speciation. Whereas with m = 0.5 one could not get progress toward speciation unless d exceeded 0.5, no matter how strong the selection, this restriction on d is relaxed when m is small, and when m = 0.1 very strong selection can lead to speciation even with a fairly small value of d.

This pattern is quite consistent with conventional theory, which asserts the difficulty of sympatric speciation, but considers allopatry as a situation favorable to speciation. In the present model the sympatry-allopatry spectrum is represented by the value of m. Sympatry corresponds to m = 0.5, which implies complete random mating and random distribution of offspring among habitats, and allopatry requires that m be zero.

When m < 0.5, BC is the most frequent genotype in subpopulation I, and bc the most frequent genotype in subpopulation II. If there is a linkage disequilibrium of these loci with A, this is reflected in a higher frequency of one allele (say A) in subpopulation I, and a higher frequency of the other in subpopulation II. In effect, there are clines in all three loci. Slatkin (1975) has noted that when two loci have clines in the same region, there will be linkage disequilibrium between the two loci, and it will tend to steepen both clines and increase the adaptation of individuals to their environment. The cline in the A locus is favored in the present context because it reinforces this disequilibrium.

Recombination Fractions

Table 1 shows the results of changing the recombination fractions between the loci. The implicit genetic map is A-B-C. The recombination fraction between A and B is given by r_{12}, and between B and C by r_{23}. The cases in the Table are for partially restricted migration (m = 0.1), although the same patterns are seen when m = 0.5. For different recombination fractions the Table shows the lower limit of d which gives speciation (the actual lower limit is between the two values given in the table). When r_{12} is reduced, speciation becomes easier to envisage. When r_{23} is reduced, conditions for speciation become more restrictive. This same pattern has been seen in all cases examined.

EVOLUTIONARY FORCES AND SPECIATION

We are now in position to make a preliminary interpretation of the direction in which different evolutionary forces are working. In this model, natural selection produces linkage disequilibrium between B and C. With this disequilibrium established, natural selection will act so as to

increase the magnitude of any disequilibrium between A and the B-C complex.

That natural selection will have such an effect in this case can be seen from a heuristic example. Suppose that linkage disequilibrium between A and the B-C complex were nearly complete, selection very strong, and assortative mating nearly complete. If almost all individuals were either ABC or abc, an Abc individual would tend to mate with an ABC. This union would give rise to some ABc and AbC gametes, but as a result of the near absence of Bc and bC individuals, there would be no corresponding production of bc from matings of Bc and bC individuals. This destruction of bc genotypes by recombination occurs among Abc individuals, which are nonrandomly likely to mate with ABC. By contrast, abc individuals will tend to mate with other abc individuals, and the bc genotype will be preserved in their offspring. There is a corresponding preservation of ABC and elimination of aBC.

The result is a reduction of the frequency of A among bc individuals, and of a among BC individuals. This reinforces the linkage disequilibrium. Note that the effect involves selection against bC and Bc as well as recombination between loci B and C.

Another way of intuiting this result is to realize that there are in effect two parts of the population, one consisting of the A individuals and the other of the a's. In the A pool the genotypes are mostly BC, with a few bc's, and in the a pool the reverse is true. When there is selection against Bc and bC and in favor of BC and bc, recombination within each of these nearly isolated mating pools will result in a reduced frequency of whichever of the two types (BC and bc) is in the minority. This will reinforce the disequilibrium by raising the frequencies of ABC and of abc and lowering the frequencies of Abc and of aBC.

With an understanding of the part played by recombination between loci B and C in this process, it becomes apparent why reduction of this recombination makes speciation more difficult. In both of the above intuitive arguments, recombination between B and C plays a crucial role in the selection against (say) A alleles in bc genomes, and against a alleles in BC genomes. If the recombination between B and C is eliminated, there is no selection creating disequilibrium between A and these loci.

An understanding of the role of recombination between A and the other loci is even more important. It tends to break down the association between the prezygotic and postzygotic isolating mechanisms, so that it is always eroding the degree of progress toward speciation. In this light it should be clear why restricting recombination between A and B makes speciation easier. There is a continual conflict between selection, which increases the association between isolating mechanisms, and recombination, which reduces it. Restricting this recombination can only improve chances for speciation.

We have now identified the evolutionary force responsible for favoring speciation—natural selection—and the force opposing it—recombination. However it should be kept in mind that the recombination between B and C is favorable for the recruitment of the prezygotic isolating mechanism A/a into the scheme of reproductive isolation. We can describe the pattern in teleological terms: the assortative mating is associated with the substantive adaptations at loci B and C "in order to" prevent the formation of maladapted bC and Bc genotypes. It should be noted that although the recombination between B and C is favorable for the maintenance of disequilibrium between A and the B-C complex, this is achieved at the cost of having less disequilibrium between B and C (and hence less effective postzygotic isolation) in the first place.

Variations on the Model

Now that we have a general picture of the forces at work, we can make some further changes in the model to see how they alter this pattern. In particular, we want to know how dependent the results are on the ordering of events in the life cycle, on

diploidy, and on the type of interaction between loci B and C.

Timing of Migration

In our model, the order of events in a single generation has been Selection-Migration-Recombination. Without violating the haploid model, it would be possible to have recombination (and hence mating) precede migration. In Table 1, the column labelled SMR shows the standard ordering of events which we have discussed, and the column labelled SRM shows this alternative ordering. The patterns are the same in both columns, although the value of d needed to prevent disappearance of the linkage disequilibrium is somewhat smaller with the SRM ordering than with the SMR ordering. This result is typical of the effect of changing from SMR to SRM ordering, in all haploid cases. In some diploid cases speciation was more difficult with SRM than with SMR, as will be mentioned below, but in no case have the other qualitative patterns been much altered by the order of events.

Interaction Between Loci

The heuristic rationale presented above for the patterns seen involved a lower fitness for Bc and bC, averaged over both subpopulations, than for BC and bc. This in turn depends on the multiplicative combination of fitnesses at the two loci. If the fitness of BC in subpopulation I is instead $1 + 2s$, there is no difference in average fitness of different genotypes, if one averages arithmetically over both subpopulations. Would locus A be recruited into the nascent speciation in this case?

To check this, a version of the computer program was constructed in which the fitness of BC in subpopulation I (and correspondingly of bc in subpopulation II) was taken to be

$$1 + 2s + ks^2$$

and runs were made for different values of k with $m = 0.5$. When $k = 1$, the fitnesses combine multiplicatively, and when $k = 0$ they combine additively. A striking pattern immediately emerged. If $k \leq 0$, it becomes impossible to find any value of d for which locus A is recruited into the nascent speciation. As k is raised above zero, association of A with B-C becomes possible, but only if d is very near 1. The larger k is taken to be, the more parameter combinations there are in which one gets association of A with B-C.

This would make it seem that association of the prezygotic and postzygotic isolating mechanisms depends on the rather subtle distinction between additive and multiplicative combination of fitnesses. However, when m is reduced this is not the case. With $m = 0.1$, progress toward speciation was possible even with negative values of k as extreme as -0.5.

Thus the lower mean fitness of types Bc and bC is necessary for speciation only in the case of the Levene Model, where $m = 0.5$, while in parapatric situations conditions for speciation seem to be less restrictive. Although only a few cases have been studied numerically, the condition for there to be any values of d enabling progress toward speciation seems likely to be that migration and selection lead to positive disequilibrium within each subpopulation between B and C. It is interesting in this context that when $k = 0$ in a Levene Model, there will be no disequilibrium within each subpopulation between B and C. This phenomenon was found by Gillespie and Langley (1976) in two-locus Levene Models with fitnesses varying randomly in time and combining additively between loci. It seems to result from an exact cancellation between the positive linkage disequilibrium which results from mixture of individuals from the different populations (Wahlund, 1928), and the negative disequilibrium which results from directional selection within populations with additive fitnesses (Felsenstein, 1965).

Interestingly, the value $k = 0$ is also the boundary of the set of k values which allow the two-locus polymorphism for loci B and C to be stable. We then have a pleasing correspondence between the conditions for polymorphism in the postzygotic loci and the conditions for a prezygotic mechanism to become associated

with them. To see whether this correspondence might be general, a few additional runs were done for m = 0.1. Based on this limited information, it seems that the correspondence may be general. Those values of k which allow progress towards speciation when there is a sufficiently high value of d also seem to be precisely those which allow the two-locus polymorphism at loci B and C to persist.

If loci B and C do not interact and control separate phenotypes, and if these phenotypes affect viability, then the natural assumption for the way their viabilities combine is to assume that they are multiplicative (k = 1). If the loci are physiologically related (as when they contribute to the same phenotype), or if they affect fertility, multiplicative combination is less plausible as a null hypothesis.

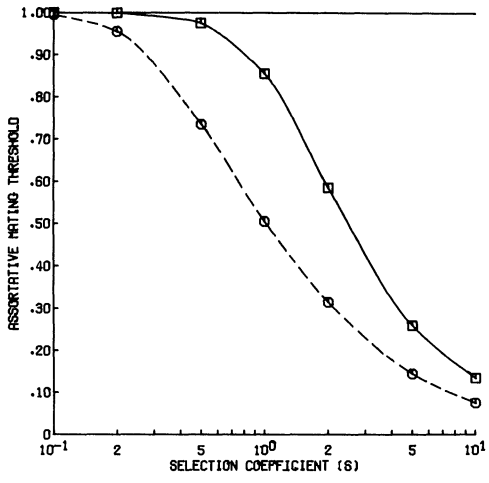

FIG. 2. The minimum amount of assortative mating (d) required to get stable association between the locus A and loci B and C, plotted for various selection coefficients in a diploid model. The squares (interpolated by the solid curve) are for m = 0.5, the circles (interpolated by the dashed curve) for m = 0.1.

Initial Gene Frequencies

The picture that has been presented here is not sensitive to different initial gene frequencies (or gamete frequencies) at loci B and C. They will rapidly move to their equilibrium gene and gamete frequencies, and since the association between A and B-C is insensitive to the initial amount of disequilibrium between them (except in regard to its sign), any increase or decrease of this disequilibrium which occurs as a byproduct of these initial changes of frequency at loci B and C will have only a transitory effect.

It is a bit less obvious whether the outcome is sensitive to the initial gene frequency at the A locus. In the absence of disequilibrium between A and the other loci, there is no selection changing gene frequencies at locus A. When disequilibrium between these loci does exist, it is possible in principle that it results in effective selection to change gene frequencies at the A locus.

Computer iterations to check this starting with one allele rare at the A locus showed a pleasing pattern: if the values of s and d were such as to result in progress toward speciation, and if there was initial disequilibrium between loci A and B-C, then selection at the latter loci resulted in changes in gene frequency at the A locus, so that A proceeded to a gene frequency of one-half. Thus if selection is able to be effective in the face of recombination between A and the other loci, it will not only result in association between the isolating mechanisms, but will also act to actively maintain polymorphism for a prezygotic mechanism such as the one we are studying.

Little time was spent investigating cases in which the associations between A and B-C would disappear. Since the effective selection at locus A results from the disequilibrium with the other loci, we would expect in those cases that as the disequilibrium disappeared, gene frequency change at the A locus would gradually cease, and it might end up not having changed much from the initial gene frequencies.

Modifiers of Assortative Mating

When the assortative mating locus is in association with the selected loci, speciation has not yet been fully achieved unless d = 1. We can only count this linkage disequilibrium as progress toward speciation

TABLE 2. *Threshold values of the amount of assortative mating necessary to allow progress toward speciation in some diploid cases. The threshold value of d lies between the two values given in the appropriate column. SMR and SRM refer to the two orderings of events in the life cycle.*

				d	
s	r_{12}	r_{23}	m	SMR	SRM
1.0	.5	.5	.5	.86–.87	.66–.67
1.0	.5	.5	.1	.50–.51	.46–.47
1.0	.5	.5	.01	.44–.45	.44–.45
1.0	.1	.5	.1	.28–.29	.25–.26
1.0	.5	.1	.1	.55–.56	.53–.54

if further events are capable of carrying the population to complete reproductive isolation. One way in which this might happen is by the substitution of modifiers which increase the strength of assortative mating.

To see whether this could happen, a program was written in which a fourth locus, D, was added whose alleles modified the strength of assortative mating. D × D matings were assumed to involve assortative mating with $d = d_1$, D′ × D′ to have $d = d_2$, and D × D′ to have $d = (d_1 + d_2)/2$. The object was to see whether an allele, D′, which increased the strength of assortative mating, would increase in a situation in which linkage disequilibria are established between A and B-C. Of course, there are many other ways modification of assortative mating could be modelled. This scheme was chosen because it places no direct selection on the modifier, and in hopes that the result will prove to be insensitive to the specific modification scheme.

Only a few runs of this program could be made, as the presence of the fourth locus greatly increased the number of computations. In those runs, the allele which increased d was observed to increase, albeit very slowly. In our model, the natural selection on locus A is a consequence of its linkage disequilibrium with B and C. The natural selection on the modifier locus D is a consequence of its disequilibrium with A, and hence indirectly its disequilibrium with B-C, and is therefore very weak. When there is no disequilibrium established between A and B-C, we expect no selection on the modifier alleles D and D′.

Diploidy

All of the above models have been haploid. To see whether this simplification has a qualitative effect on the results, a diploid model was investigated. In this diploid model, fitnesses were taken to be multiplicative both between and within loci. The fitnesses were thus:

	Fitness in subpopulation	
Genotypes	I	II
BBCC	$(1 + s)^4$	1
BbCC, BBCc	$(1 + s)^3$	$1 + s$
BBcc, BbCc, bbCC	$(1 + s)^2$	$(1 + s)^2$
Bbcc, bbCc	$1 + s$	$(1 + s)^3$
bbcc	1	$(1 + s)^4$

The assortative mating scheme used was also analogous to the haploid case. Locus A was assumed to code for the probability that the individual would enter the first of the two mating pools. The probability that the individual joined pool 1 was taken to be:

AA	Aa	aa
$(1 + d)/2$	$1/2$	$(1 - d)/2$

It was assumed that each individual entered one of the two mating pools. Within each mating pool mating was at random, and the number of offspring expected from each mating was taken to be equal, aside from effects of loci B and C. As in the haploid case, the loci B and C were assumed to play a passive role in the mating process, the probability of two genotypes mating depending only on their genotype at the A locus. Some further algebraic details concerning this simple system of assortative mating are given in Appendix 1.

In spite of the fact that the haploid case is not a special case of the diploid case, the two models give qualitatively similar results. Figure 2 and Table 2 show the minimum values of d which will enable

permanent association between locus A and loci B and C for some diploid cases. As can be seen, the qualitative patterns are the same as in the haploid case, giving us some confidence that we are not being seriously misled by concentrating our attention on haploid models. One exception to this is the effect of tightening linkage between B and C, which does not seem to cause as much restriction in the conditions for speciation as it does in the haploid case. When we discuss the work of Dickinson and Antonovics (1973) below, we will see that this effect is significant.

Two Kinds of Speciation

The picture of speciation which has emerged from considering the present models involves two opposing forces, selection and recombination, with movement towards speciation only being possible when selection is strong enough that its effect is not overwhelmed by recombination. We may contrast this picture with the situation in the model by Balkau and Feldman (1973). They imagined two populations and two loci. At one locus there were two alleles, each well-adapted to one of the populations, so that a cline was set up. At the other locus, a modifier locus, an allele was introduced which had the effect of reducing the migration rate of its bearers. They showed that this modifier allele would always increase when rare, irrespective of the recombination fraction between the two loci.

Their model was intended to model selection pressures for modification of migration, but we could just as well regard it as a model of speciation. If the process of modification were to continue, the outcome would be two allopatric reproductively isolated populations. Note that in their model, there is neither a threshold amount of selection necessary to allow speciation, nor any effect of linkage of the modifier locus to the selected locus.

The key to the different behavior of Balkau and Feldman's model is that speciation in their model proceeds by substituting the same allele in both populations. If the modifier locus has the same frequency in both populations, then recombination between individuals from the two populations has no tendency to destroy reproductive isolation. Selection is then unopposed by recombination, and can be effective even when weak. By contrast, in the model presented in this paper speciation proceeds by substituting different alleles in the two nascent species. This means that selection is at risk of being overwhelmed by recombination, and that speciation can proceed only when there is sufficiently strong selection at loci B and C, or sufficiently weak gene flow between the two nascent species.

We may tentatively call these two classes of models of speciation "one-allele" and "two-allele" models. The critical distinction between them is whether reproductive isolation is strengthened by substituting the same or different alleles in the two nascent species. Of course, there is nothing to prevent both kinds of processes from going on at the same time, but at different sets of loci. This distinction cuts across the traditional sympatric-parapatric-allopatric spectrum. Both the present model and the Balkau-Feldman model can be formulated with initial random mating between the two populations, or with little gene flow between them. Thus neither is intrinsically sympatric or intrinsically allopatric.

A distinction similar to this one has been made by Endler (1977), who distinguishes between "Type I" and "Type II" modifiers. In the parapatric cases which Endler considers, these are modifiers which increase fitnesses of locally-adapted genotypes in both regions of a cline (Type I) or which increase fitnesses of the genotypes adapted to one region but decrease fitnesses of genotypes adapted to the other (Type II). Endler does not discuss the effect of recombination between genotypes from the two regions on the fate of these two kinds of modifiers. In fact, they are incidental to his model of parapatric speciation, a one-allele model in which a modifier of the amount of assortative mating spreads in both nascent species.

The sympatric speciation models of

Maynard Smith (1966) include both one-allele and two-allele models. Maynard Smith's basic model framework has a Levene Model with two subpopulations, and one locus, A, whose two alleles are adapted to the two subpopulations with the following fitness scheme:

Subpopulation	Genotype AA	Aa	aa
I	1 + K	1 + K	1
II	1	1	1 + k

Maynard Smith notes that a modifier causing "habitat selection," which in his context means a tendency to remain in the subpopulation where the animal was born, will spread in both subpopulations, and cause speciation. This is in effect a one-allele speciation model very similar to that of Balkau and Feldman, although Maynard Smith did not attempt a full mathematical analysis to confirm his intuition. Another of Maynard Smith's models involves a modifier of the amount of assortative mating. This is also a one-allele model.

Yet another of his models involves a second locus, B, which codes for two phenotypes according to which there is assortative mating. Except for his assumption that there is complete dominance at that locus, it has exactly the same properties as the assortative mating locus A in the present model. His model is of the two-allele type, and we should expect to see the same sort of threshold effect as in the present model. Maynard Smith does not discuss this matter, but presents a numerical example involving complete assortative mating, strong selection, and some restriction of migration between the subpopulations. In that case speciation occurs, and it is implicit in Maynard Smith's discussion that there are parameter combinations for which speciation would not occur. He was more concerned with showing that there were sympatric models for which speciation would occur.

Dickinson and Antonovics (1973) made a more extensive study of a model which was quite similar to this model of Maynard Smith's. They found that strong selection, strong assortative mating, and restricted migration all favored the establishment of an association between the assortative mating gene and the selected locus. Their results are thus completely consonant with the picture presented above. They found that "any tendency toward isolation could be swamped if the level of gene flow increased," although they did not present information on the exact parameter combinations for which this swamping occurred.

Their work complements the present model in that the phenomenon which speciation prevents is not recombination between two loci in a situation of coadaptation, but formation of heterozygotes between two alleles adapted to different environments. It is therefore questionable whether a counterpart to their model could be made to work in haploids. In fact, if we eliminate segregation at one of the selected loci in our diploid model, it would come to resemble their model. A few runs have been made on the diploid model with segregation at loci A and B only. Interestingly enough, speciation turns out to be possible in our diploid model even when there is segregation only at one locus (or what is nearly equivalent, when there is complete linkage between loci B and C). This apparently results from the fact that the formation of ill-adapted heterozygotes is avoided by the occurrence of assortative mating.

Caisse and Antonovics (1978) presented a model involving parapatric speciation, with one locus (A) having a selection-migration cline, and another (B) for which assortative mating occurs. They find many of the same patterns, including parameter combinations for which the B locus did not show a cline of its own. They find that tightening the linkage between the A and B loci makes speciation happen more readily. They explicitly raise the question of why speciation does not always occur, and conclude that "the conditions leading to isolation are far more

stringent than those permitting genetic divergence." It is not known whether a haploid version of their model would show the same behaviors as their diploid model. Since the "purpose" of the assortative mating in their model is to prevent formation of ill-adapted heterozygotes, I suspect that it would not.

Predictions

From the picture presented here, a few predictions can be made as to how the patterns of speciation seen in nature should relate to the distinction between one- and two-allele speciation models. We have found that increased migration between the subpopulations is an unfavorable condition for speciation in a two-allele model. In a one-allele model we would expect it to make little difference. Thus there should be a correlation between the sympatry-allopatry spectrum and the distinction between one- and two-allele speciation. Allopatry is a situation favorable to either, but in sympatric situations, two-allele cases should be rarer. Thus we expect that after speciation there will be less genetic differentiation with respect to the genes involved in isolating mechanisms in cases of sympatric speciation than in cases of allopatric (or near-allopatric) speciation.

In a case of sympatry, speciation would be nearly impossible unless it were based on genetic variation which could lead to one-allele reproductive isolation. I find it easier to imagine genetic variation of the two-allele sort than of the one-allele sort. If there were a shortage of genetic variation of the one-allele sort, then recombination between the loci involved in pre- and postzygotic mechanisms could form an effective block to speciation. On the other hand, if genetic variation of both sorts is readily available, then there will be no block to speciation. If evolution is never limited by a shortage of any particular type of genetic variation, then the block to speciation presented here will never be relevant.

There may also be implications for the linkage relations between genes affecting the isolating mechanism and the genes which affect adaptation to the different environments. We have seen that this linkage has no influence in Balkau and Feldman's one-allele model. In the two-allele models it makes speciation easier. We may therefore find that there has been some tendency for two-allele reproductive isolating mechanisms to have arisen at loci linked to loci which affect the substantive adaptations. This need not be the result of linkage modification, but could simply result from the fact that unlinked loci coding for two-allele isolating mechanisms failed to become involved in a successful speciation. This may not be a large effect.

One test for the occurrence of two-allele isolation suggests itself. In a speciation carried out wholly by one-allele isolating mechanisms, laboratory crosses of the two species, if possible at all, should result in no breakdown of the prezygotic isolation, although the postzygotic mechanisms which involve differentiation between the two species will become randomized. To the extent that prezygotic isolation is based on two-allele mechanisms, it will be broken down by crossing. Of course, care must be taken to observe the behavior or phenotype which leads to isolation, not merely the isolation itself. Otherwise we might conclude that a hybrid had lost the isolation when it was in reality still engaging in a behavior such as stringent mate selection, but was selecting other hybrids as mates.

Limitations of the Model

There are a number of directions in which it would be desirable to expand this model. One would involve making the genetic determination of both sets of traits, the adaptations as well as the reproductive isolation, polygenic. It is not clear a priori whether the results found here are sensitive to the number of loci assumed involved in the traits, although some effects, such as the breakdown of isolation by crossing, must also occur in polygenic cases. Extension of the model to examine effects of unequal subpopulation sizes is also called for, in part to examine situa-

tions where a new species arises in a peripheral population.

A more important extension would involve making a model of postzygotic isolation. The B and C loci in our model may be thought of as constituting a postzygotic mechanism causing partial reproductive isolation—after all, Bc and bC gametes are being eliminated by the selection. But one could also imagine a modifier locus which had an allele which intensified postzygotic isolation by decreasing the fitness of Bc and bC, or by increasing the fitness of BC and bc. If the same allele at that locus increased the fitnesses of both BC and bc, this would be a one-allele isolating mechanism, while if one allele increased the fitness of BC and the other increased the fitness of bc, this would be a two-allele case.

The papers of Bazykin (1965, 1969) and Slatkin (1975) present phenomena relevant to the construction of models of postzygotic isolation in clines. It is less than obvious how to construct a simple canonical model of postzygotic isolation. One could imagine evolution bringing into play more loci like B and C, or modifiers which increase the fitness of BC and bc, or modifiers which decrease the fitness of Bc and bC. All of these would behave differently.

An additional direction for future work is suggested by the effects of gene interaction studied above. When the mean fitnesses of BC and bc, averaged across the two subpopulations, did not exceed the fitnesses of Bc and bC, then sympatric speciation became impossible no matter how strong the assortative mating. In this case the generalists Bc and bC do not have lower average fitness than the specialists. This suggests that we may be able to relate these fitness patterns to measurements of niche overlap between the forms adapted to the two subpopulations, and that when this overlap is too great, speciation will not occur.

Only when we can bring genetic and ecological constraints on speciation into a common framework will we begin to have a satisfactory overview of the speciation process. Only then will geneticists be able to join ecologists in paying homage to Santa Rosalia.

Summary

A model of speciation has been constructed involving two loci under natural selection in two subpopulations, with different alleles adapted to the different subpopulations. Progress toward speciation in this model consists of association of a third locus, at which there is assortative mating, with the original two loci. Cases can be found in which speciation cannot occur. The evolutionary force acting against speciation turns out to be recombination, which acts to randomize the association between the prezygotic isolating mechanism (assortative mating) and the adaptations to the two environments. This model suggests that there is an important distinction between two kinds of speciation. One involves speciation by substitution of the same alleles in the two nascent species, the other by substitution of different alleles. Only in the latter case does recombination act as a force retarding or blocking speciation.

Acknowledgments

I am grateful to Monty Slatkin, Takeo Maruyama, Alan Templeton, Janis Antonovics, Freddy Christiansen, and Marcy Uyenoyama for helpful suggestions and discussions. Apologies are due G. Evelyn Hutchinson for misuse of the title of his classic paper. This work was funded by Task Agreement No. DE-AT06-76EV71005 under contract No. DE-AM06-76RLO2225 between the United States Department of Energy and the University of Washington.

Literature Cited

Balkau, B., and M. W. Feldman. 1973. Selection for migration modification. Genetics 74:171–174.

Bazykin, A. D. 1965. On the possibility of sympatric species formation (Russian). Byulleten' Moskovskogo Obshchestva Ispytateley Prirody. Otdel Biologicheskiy 70 (1):161–165.

———. 1969. Hypothetical mechanism of speciation. Evolution 23:685–687.

Caisse, M., and J. Antonovics. 1978. Evolution

in closely adjacent plant populations. Heredity 40:371–384.

DICKINSON, H., AND J. ANTONOVICS. 1973. Theoretical considerations of sympatric divergence. Amer. Natur. 107:256–274.

ENDLER, J. A. 1977. Geographic Variation, Speciation, and Clines. Monographs in Population Biology No. 10, Princeton Univ. Press, Princeton.

FELSENSTEIN, J. 1965. The effect of linkage on directional selection. Genetics 52:349–363.

GILLESPIE, J., AND C. LANGLEY. 1976. Multilocus behavior in random environments. I. Random Levene models. Genetics 82:123–137.

HUTCHINSON, G. E. 1959. Homage to Santa Rosalia *or* why are there so many kinds of animals? Amer. Natur. 93:145–159.

LEVENE, H. 1953. Genetic equilibrium when more than one ecological niche is available. Amer. Natur. 87:331–333.

MAYNARD SMITH, J. 1966. Sympatric speciation. Amer. Natur. 100:637–650.

ROSENZWEIG, M. L. 1975. On continental steady states of species diversity, p. 121–140. *In* M. L. Cody and J. M. Diamond (eds.), Ecology and Evolution of Communities. Belknap Press, Cambridge.

SLATKIN, M. 1975. Gene flow and selection in a two-locus system. Genetics 81:787–802.

WAHLUND, S. 1928. The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity (German). Hereditas 11:65–106.

Corresponding Editor: D. B. Wake

APPENDIX 1. Further comments on the models of assortative mating.

The assortative mating scheme in the haploid model was specified by stating that a fraction d of the population mated with individuals of the same genotype at the A locus, and 1 − d mated at random. Another mechanism which would lead to the same mating type frequencies would be the following: The A locus controls the division of the mating pool into two mating groups (according to space, time, or phenotype). Type A has probability (1 + x)/2 of entering mating group 1, and probability (1 − x)/2 of entering mating group 2. For allele a the probabilities are the reverse, (1 − x)/2 and (1 + x)/2. Within each of the two mating groups mating is at random. In this model it is specified that the likelihood of an individual reproducing is not affected by its genotype at the A locus or by the mating group it is in. It is not hard to show that the resulting mating type frequencies are then as previously given, with

$$d = \frac{x^2[1 - (p - q)^2]}{1 - x^2(p - q)^2} \qquad \text{(A1-1)}$$

where p and q are respectively the frequencies of alleles A and a.

In the case of the diploid model, we use a very similar model of assortative mating. Every individual joins one of the two mating pools, and the probabilities of joining pool 1 are (1 + d)/2, 1/2, and (1 − d)/2 for the three genotypes AA, Aa, and aa. This form of assortative mating differs from the haploid case in that when d = 1 some heterozygotes will still be formed, while in the haploid case all matings would then be A × A or a × a. However, if d = 1 the frequency of heterozygotes at the A locus will gradually decline to zero through time. In a single population without effects of other loci, this scheme of assortative mating will reach equilibrium with gene frequency unchanged from its initial value. When the initial gene frequency is p, the final genotype frequencies of AA, Aa, and aa will be P, 2p − 2P, and 1 − 2p + P, where P is the smaller of the two solutions to the quadratic equation

$$-d^2P^2 + P[1 - 8d^2p^2 + 12d^2p - 4d^2] - p^2 = 0 \qquad \text{(A1-2)}$$

which for p = 0.5 becomes

$$-1/4 + P - d^2P^2 = 0. \qquad \text{(A1-3)}$$

Note that the frequency of AA individuals in the diploid case is not the same as the frequency of A × A matings in the haploid case, so that we cannot directly compare haploid and diploid cases with the same value of d.

It must be emphasized that the division of a population into mating groups has no necessary relation to the question of allopatry. The offspring of these matings are in the same ecological population, in that they compete perfectly with each other. The mating groups might simply be those individuals who mate in the morning and those who mate in the afternoon.

APPENDIX 2. A more formal statement of the haploid model.

Let $x_{ijk,u}^{(t)}$ be the frequency of genotype ijk in subpopulation u in generation t, where i, j, and k are indices which take on the values 0 for (respectively) alleles a, b, and c, and the values 1 for A, B, and C. The fitnesses of these genotypes are given by the values $w_{ijk,u}$. After selection, the genotype frequencies are given by

$$y_{ijk,u}^{(t)} = x_{ijk,u}^{(t)} w_{ijk,u} / w_u^{(t)} \qquad \text{(A2-1)}$$

where $w_u^{(t)}$ is the mean fitness of subpopulation u in generation t. Migration takes place according to a conventional island model:

$$z_{ijk,u}^{(t)} = \sum_{v=1}^{2} m_{uv} y_{ijk,v}^{(t)} \qquad \text{(A2-2)}$$

where m_{uv} is 1 − m if u = v and m otherwise, in the two-population case. Mating now takes place in each subpopulation. If $M_{ijk,fgh,u}^{(t)}$ is the frequency of matings between genotypes ijk and fgh, under the present scheme of assortative mating this is given by

$$M_{ijk,fgh,u}^{(t)} = (1 - d)z_{ijk,u}^{(t)}z_{fgh,u}^{(t)} + dz_{ijk,u}^{(t)}z_{fgh,u}^{(t)}\delta_{if}/z_{i..,u}^{(t)} \quad \text{(A2-3)}$$

where δ is the Kronecker delta function which is 1 if i = f and 0 otherwise. Here $z_{i..,u}^{(t)}$ is the overall frequency of allele i in subpopulation u in generation t, after selection.

Recombination then occurs in each subpopulation. To avoid tedium, it will not be described here, except to say that it follows the autosomal Mendelian rules with no interference, with recombination fractions r_{12} and r_{23}, and with the resulting gametes being denoted by $x_{ijk,u}^{(t+1)}$.

This Appendix has described the haploid model with the SMR ordering of life cycle events. The other ordering and the diploid model can be formalized in a similar fashion.

Evolution Society News

Society Election

The results of the 1980 election of officers are as follows:

President elect	Otto Solbrig
Vice President I	Wyatt W. Anderson
Vice President II	Linda Maxon
Vice President III	A. H. D. Brown
Councilor I	Michael Rosenzweig
Councilor II	Mary F. Willson

New Associate Editors

The associate editors for 1981–1983 are as follows:

Dr. Joseph Felsenstein, University of Washington
Dr. Subbodh Jain, University of California, Davis
Dr. Robert Ricklefs, University of Pennsylvania

Dr. Stevan Arnold, University of Chicago, will fill Dr. Futuyma's 1981–1982 term as associate editor.