The Population Genetics of Speciation: The Evolution of Hybrid Incompatibilities

H. Allen Orr

Department of Biology, University of Rochester, Rochester, New York 14627

ABSTRACT

Speciation often results from the accumulation of "complementary genes," *i.e.*, from genes that, while having no deleterious effect within species, cause inviability or sterility when brought together with genes from another species. Here I model speciation as the accumulation of genic incompatibilities between diverging populations. Several results are obtained. First, and most important, the number of genic incompatibilities between taxa increases much faster than linearly with time. In particular, the probability of speciation increases at least as fast as the *square* of the time since separation between two taxa. Second, as Muller realized, all hybrid incompatibilities must initially be asymmetric. Third, at loci that have diverged between taxa, evolutionarily derived alleles cause hybrid problems far more often than ancestral alleles. Last, it is "easier" to evolve complex hybrid incompatibilities requiring the simultaneous action of three or more loci than to evolve simple incompatibilities between pairs of genes. These results have several important implications for genetic analyses of speciation.

I T is well known that little of *The Origin of Species* concerns the splitting of species. One of the reasons for this neglect is not generally appreciated, however. It was not simply that DARWIN was more interested in the forces shaping change within lineages. Instead, the origin of reproductive isolation posed a serious problem to DARWIN: as he admitted (1859, p. 264), it was unclear how something as patently maladaptive as the sterility or inviability of hybrids could evolve by natural selection. In modern parlance, it was unclear how two genotypes descended from a common ancestor could become separated by an adaptive valley unless one of the lineages passed through the valley. This would not, of course, be allowed by natural selection.

This fundamental problem was finally solved by DOB-ZHANSKY (1936) and MULLER (1939, 1940) early in the modern synthesis. Each produced genetic models showing that two populations could come to produce completely sterile or inviable hybrids even when no substitution caused any sterility or inviability within either population. Their models were very simple: two allopatric populations begin with identical genotypes at two loci (*aa,bb*). In one population, an *A* allele appears and is fixed; the *Aabb* and *AAbb* genotypes are perfectly viable and fertile. In the other population, a *B* mutation appears and is fixed; *aaBb* and *aaBB* are also viable and fertile. The critical point is that, although the *B* allele is compatible with *a*, it has not been "tested" on an *A* genetic background. It is thus possible that *B* has a deleterious effect that appears only when *A* is present. If the two populations meet and hybridize, the resulting *AaBb* hybrid may be inviable or sterile.

Author e-mail: haorr@darwin.biology.rochester.edu

As MULLER (1942) pointed out, it makes no difference whether the substitutions occur in both populations, as above, or in one only. If one population retains the ancestral *aabb* genotype and the other becomes *AAbb* and then *AABB*, the *B* allele may well be incompatible with the *a* allele among the *AaBb* hybrids. In either case, reproductive isolation results from "complementary" or "reinforcing" epistasis between loci *A* and *B* (CROW and KIMURA 1970, p. 81): the lethal or sterile effect of an allele at one locus depends on the background genotypes at other loci. Of course, complementary genes need not have a complete effect—a pair of complementary genes might cause only partial hybrid sterility or inviability.

DOBZHANSKY and MULLER's model of speciation is important for two reasons. First, it shows that the evolution of hybrid sterility or inviability need not involve any intermediate, maladaptive step. Perhaps more important, it shows that, while the problem of the origin of species can be reduced to the origin of reproductive isolation, this in turn—at least for postzygotic isolation—can be reduced to the building up of complementary genes.

It is now clear that postzygotic isolation usually results from complementary genes (ignoring sterility resulting from polyploidy). Among plants, complementary genes causing hybrid inviability have been found in *Mimulus* (CHRISTIE and MACNAIR 1984, 1987), *Crepis* (HOLLINGSHEAD 1930), and cotton (GERSTEL 1954). Similarly, complementary factors causing hybrid inviability or sterility have been found in many Drosophila hybridizations (MULLER and PONTECORVO 1940; MULLER 1942; PANTAZIDIS and ZOUROS 1988; H. A. ORR, unpublished data).

Given the success of this simple model in explaining the genetical facts of hybrid sterility and inviability, it is remarkable that almost no one has explored the mathematical consequences of viewing speciation as an accumulation of complementary genes [see NEI (1976) for a notable exception]. Although population geneticists have constructed more or less formal models of speciation (TEMPLETON 1981; SLATKIN 1982; WALSH 1982; NEI et al. 1983; BARTON and CHARLESWORTH 1984; CHARLESWORTH et al. 1987), these theories are largely concerned with the relative roles of genetic drift and natural selection in speciation and almost entirely ignore the simple mechanism by which postzygotic isolation evolves: the accumulation of complementary genes. The reasons for this neglect are not clear, but, as will become clear, they do not include the mathematical difficulty of the problem.

In this and in a later paper (M. TURELLI and H. A. ORR, unpublished data), we analyze models in which speciation is treated as the accumulation of complementary genes. In a sense, we try to see how far one can go in explaining the facts of postzygotic isolation by studying consequences of the accumulation of complementary genes. Several different problems are considered: the rate at which reproductive isolation arises, the roles of ancestral vs. derived alleles in reproductive isolation, the expected complexity of hybrid incompatibilities (will most incompatibilities involve pairs or triplets, etc. of genes?), the genetic basis of Haldane's rule and the large role of the X chromosome in reproductive isolation.

The first several issues can be addressed with simple models that ignore dominance (this paper), while Haldane's rule and the large X-effect require more complex models that take dominance into account (M. TURELLI and H. A. ORR, unpublished data). Both classes of model, however, are built around the same theme: the evolution of postzygotic isolation can be reduced to the mechanics of genic incompatibilities.

THE BASIC MODEL

The central assumption of the DOBZHANSKY–MULLER model of speciation is that alleles cause no sterility or inviability on their normal "pure species" genetic background. Instead, an allele can lower fitness only when brought together with genes from another species. Any particular hybrid incompatibility might cause partial or complete hybrid sterility or inviability. For most of this paper, I assume that hybrid incompatibilties involve interactions between pairs of genes, as in DOBZHANSKY and MULLER's verbal models. Later, I consider three-locus and higher interactions. I also assume that multiple substitutions do not occur at the same locus, an assumption that is reasonable during the early divergence of taxa. I assume nothing about the evolutionary causes of substitutions. The DOBZHANSKY–MULLER model of speciation requires only that substitutions occur and assumes nothing about whether they are brought about by natural selection or genetic drift.
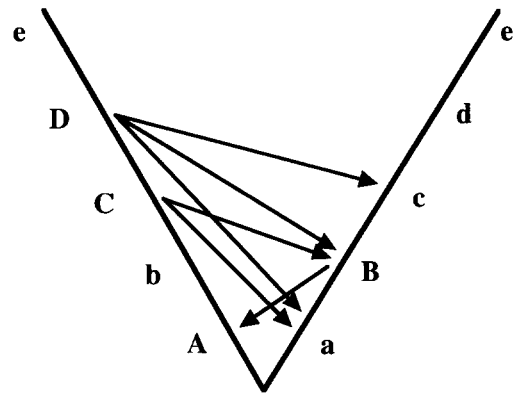

FIGURE 1.—History of substitutions fixed between two populations. Time runs upward. Both populations were initially fixed for lowercase alleles at all loci. The first substitution occurred at locus a, the second at b, and the third at c. Arrows show possible incompatibilities between loci from the two populations.

Because I consider the cumulative effects of interactions between many loci—which quickly gets complicated—it is useful to picture this process diagramatically. Figure 1 offers a simple way to picture the accumulation of complementary genes between two haploid populations. Each of the two heavy lines represents a lineage descended from a common ancestor. The two allopatric populations begin with identical "ancestral" lowercase genotypes at all loci (a b c . . .). Time runs upward, with the first substitution occurring at the a locus, the second at the b locus and so on.

The first substitution involves the replacement of the a allele by the A allele in population 1 (uppercase letters indicate only that an allele is "derived"; no dominance is implied). The A allele cannot cause any hybrid sterility or inviability: because A is obviously compatible with the genetic background of population 1, it must be compatible with the identical background of population 2. The second substitution, at the B locus (in population 2), could be incompatible with only one locus: A, as the B allele has not been "tested" for compatibility with A. The third substitution, at C, could be incompatible with the B or a alleles. As we continue this process, it is clear that we can identify all possible (i.e., evolutionarily allowed) incompatibilities by drawing an arrow from each derived allele to each "earlier" (lower) allele carried by the other species. Thus D can be incompatible with c, B, and a. This arrow-drawing device will repeatedly prove useful.

Several facts immediately emerge from Figure 1. First, hybrid incompatibilities only occur between two loci that have both experienced a substitution: in Figure 1, arrows never run up toward loci that have not diverged (e.g., locus e). This follows from the fact that the two populations carry identical alleles at all undiverged loci, so that any substitution must be compatible with these loci in both species.

Several other less trivial facts also emerge from Figure 1:

- All incompatibilities are asymmetric. For example, although B might be incompatible with A, b cannot be incompatible with a.
- Evolutionarily derived (uppercase) alleles are involved in more potential incompatibilities than ancestral (lowercase) alleles.
- Later substitutions cause more possible incompatibilities than earlier ones (*e.g.*, although the substitution of B produces one possible incompatibility, the later substitution of D produces three). This suggests that the strength of reproductive isolation might increase faster than linearly with time.

I consider each of these observations below. Because the last point has the most important evolutionary consequences, I consider it in the most detail.

THE ASYMMETRY OF INCOMPATIBILITIES

As first noted by MULLER (1942), all hybrid incompatibilities must be asymmetric. Figure 1 shows that, although B might be incompatible with A, a cannot be incompatible with b. The reason is simple: *aabb* represents an ancestral step in the divergence of these taxa (*i.e.*, *aabb* is either the genotype of the common ancestor or an intermediate step in the evolution of these taxa). Thus the required fertility/viability of all intermediate steps in the divergence of two taxa places constraints on which incompatibilities are possible and which are not, a point that will recur.

Notice, however, that the required asymmetry does *not* mean that both derived and ancestral alleles at a locus cannot cause hybrid problems. Instead, it means only that the ancestral and derived alleles at one locus (*e.g.*, A and a) cannot be incompatible with alleles at the *same* other locus (*e.g.*, the *A-B and a-b* incompatibilities are not both possible). Thus, even if the A allele of Figure 1 were incompatible with B, a could still be involved in another incompatibility, for example, with C. Indeed, this latter incompatibility is just as likely as any other.

The finding that both homologous alleles at a locus do not usually cause postzygotic isolation (*e.g.*, WU and BECKENBACH 1983) merely reflects the fact that p^2 is much smaller than p, where p is the probability that an allele causes detectable problems in hybrids.

THE ROLE OF DERIVED *VS.* ANCESTRAL ALLELES IN HYBRID INCOMPATIBILITIES

Figure 1 also shows that derived (uppercase) alleles tend to be involved in hybrid incompatibilities more often than ancestral (lowercase) alleles. Once again, this is a trivial consequence of the types of incompatibilities that are possible: when a new derived allele (say C) is substituted, it might be incompatible with either another derived allele (B) or with an ancestral allele (a). Thus, both derived–derived (DD) and derived–ancestral (DA) incompatibilities occur. The only type of incompatibility

that does not arise is ancestral–ancestral (AA). The reason is obvious: all ancestral alleles must be compatible as they represent the initial genotype.

Thus, restricting our attention to those loci that have experienced a substitution, the alleles causing postzygotic isolation will be derived more often than ancestral. We would like to know how much more often.

Imagine that a fraction f of all substitutions occur in population 1 and $1 - f$ in population 2. We consider only those alleles that are involved in a hybrid incompatibility. A new allele that arises in, say, population 1 can only be incompatible with those loci that have already experienced substitutions. By assumption, a fraction f of these loci in population 2 carry ancestral alleles and a fraction $1 - f$ carry derived alleles. Thus the probability that a new allele in population 1 will be incompatible with an ancestral allele in population 2 is proportional to f. Similarly, the probability that it will be incompatible with a derived allele in population 2 is proportional to $1 - f$. By making all such comparisons it is easy to see that the probabilities of the various possible incompatibilities are

$$P(D_1 A_2 \text{ incompatible}) = f^2,$$
$$P(A_1 D_2 \text{ incompatible}) = (1 - f)^2,$$
$$P(D_1 D_2 \text{ incompatible}) = 2f(1 - f), \qquad (1)$$

where D represents a derived allele, A an ancestral allele, and the subscripts identify the two populations.

From this, we can tabulate the expected frequency with which derived *vs.* ancestral alleles from each population will be involved in hybrid incompatibilities:

$$P(D_1) = \frac{f(2 - f)}{2},$$

$$P(A_1) = \frac{(1 - f)^2}{2},$$

$$P(D_2) = \frac{(1 - f)(1 + f)}{2},$$

$$P(A_2) = \frac{f^2}{2}. \qquad (2)$$

Obviously $P(D_1) + P(A_1) = P(D_2) + P(A_2)$, *i.e.*, the total frequency with which alleles from population 1 are involved equals the frequency with which alleles from population 2 are involved, as every incompatibility must involve an allele from each species.

Finally, the ratio $P(D):P(A)$ of derived-to-ancestral alleles causing hybrid incompatibilities is

$$1 + 2f(1 - f):1 - 2f(1 - f). \qquad (3)$$

In words, the number of derived *vs.* ancestral alleles causing reproductive isolation depends on the proportion of substitutions that occur in each population. When there are equal rates of evolution in the two lineages ($f = \frac{1}{2}$), *derived alleles are three times more likely*

to be involved in hybrid incompatibilities than ancestral alleles, as noted by ORR (1993). Indeed, derived alleles are always more likely to cause hybrid problems unless all evolution occurs in one lineage ($f = 1$). In that case, $P(D):P(A) = 1:1$ since the only possible type of incompatibility is derived–ancestral. The ratio $P(D):P(A)$ plays an important role in one possible explanation of Haldane's rule (see ORR 1993).

THE RATE OF SPECIATION

Later substitutions cause more potential incompatibilities than earlier ones (Figure 1). As already noted, the first substitution at the A locus cannot cause any hybrid incompatibility, while the second substitution could be incompatible with only one locus: the B allele has not been tested with the A allele. In general, the Kth substitution can be incompatible with $K - 1$ loci from the other population.

It is obvious, then, that the total number of incompatibilities separating two taxa increases faster than linearly with the number of substitutions that have occurred between them. This, in turn, implies that the strength of reproductive isolation—or the probability of speciation—between two taxa increases faster than linearly with time. This important effect is easily quantified. I consider two cases. First, I assume that complete reproductive isolation results from a single incompatibility between two complementary genes. Second, I assume that reproductive isolation results from the cumulative effects of many small incompatibilities. As we will see, both cases yield similar results.

Single-incompatibility speciation: Reproductive isolation here results from a single incompatibility between two complementary genes and the level of isolation suddenly leaps from zero to unity. Although it is unclear just how common this situation is, there is considerable evidence that reproductive isolation sometimes has a simple genetic basis (HOLLINGSHEAD 1930; GERSTEL 1954; WITTBRODT *et al.* 1989).

I calculate the probability that speciation has occurred as a function of the number of substitutions separating two diverging populations. For simplicity, assume that a new derived allele has a fixed probability, p, of being incompatible with each of the loci that has previously experienced a substitution. Because substitution K may be incompatible with $K - 1$ loci, the Kth substitution has a probability $1 - (1 - p)^{K-1}$ of causing speciation. The cumulative probability of speciation, S, is simply the probability that at least one incompatibility occurs, given by

$$S = 1 - \prod_{n=1}^{K} (1 - p)^{n-1} = 1 - (1 - p)^{K(K-1)/2}. \quad (4)$$

When p is small (which it surely is) and K is large,

$$S \approx 1 - e^{-K(K-1)p/2} \approx 1 - e^{-K^2 p/2}. \quad (5)$$

It is important to note that Equations 4 and 5 do *not* depend on the proportion of substitutions that occur in each population nor on the order in which substitutions arise in population 1 *vs.* 2.

Thus the probability of speciation increases much faster than linearly. Indeed, when $S \ll 1$, *the cumulative probability of speciation increases as the square of the number of substitutions.* How S varies with time depends, of course, on how the number of substitutions increases with time. If K increases approximately linearly with time, *i.e.*, if there is a rough molecular clock, the cumulative probability of speciation rises with the square of time since divergence. In any case, S increases faster with time than does K.

Although we have assumed that a single incompatibility can cause hybrid lethality or sterility, additional incompatibilities will obviously continue to accumulate after this initial speciation event. It is easy to show that because of the everincreasing probability of obtaining hybrid incompatibilities, the expected number of incompatibilities separating two taxa also increases very rapidly. Because Equation 5 describes a state-dependent Poisson process, the mean number of incompatibilities that have occurred up through substitution K is

$$I = \frac{K(K - 1)p}{2} \approx \frac{K^2 p}{2}, \quad (6)$$

i.e., the expected number of complementary lethals/steriles increases as K^2. This rapid increase in the number of "speciation genes" has important implications for genetic analyses of speciation (see DISCUSSION).

Interestingly, we can also find the expected "time" to speciation. If K_s is the number of substitutions required until the appearance of a hybrid incompatibility, then $P(K_s > K) = (1 - p)^{K(K-1)/2}$. Thus the expectation of K_s is $\bar{K}_s = \sum_{K=2}^{\infty} (1 - p)^{K(K-1)/2}$, which is approximately

$$\bar{K}_s \approx \int_0^{\infty} e^{-K^2 p/2} \, dK$$

$$= \sqrt{\frac{\pi}{2p}}. \quad (7)$$

Thus if the probability that any two genes are incompatible in hybrids is $p = 10^{-5}$, an average of 400 substitutions is required for speciation. Because the time to speciation is an inverse function of $\sqrt{p}$, it is not as sensitive to p as one might expect. Doubling the probability that an incompatibility occurs, for example, does not halve the time to speciation but reduces it by a factor of $1/\sqrt{2}$.

Multiple incompatibility speciation: I now consider the case where speciation results from the cumulative effects of several to many smaller incompatibilities. We will see that the above results do not depend on the assumption that speciation is caused by a single incompatibility of complete effect.

The reduction in hybrid fitness, r ($0 \le r \le 1$), resulting from an interaction between any two genes has some frequency distribution $f(r)$. Obviously, the mean

$\bar{r}$ must be very small or speciation would be nearly instantaneous; indeed, most interactions between genes in hybrids may have no effect on hybrid fitness ($r = 0$).

I assume that different incompatibilities act independently: if one incompatibility reduces fitness to $(1 - r_1)$ and another reduces fitness to $(1 - r_2)$, both together reduce fitness to $(1 - r_1)(1 - r_2)$. Thus the effect of substitution K on hybrid fitness is given by $L_K = 1 - w_K = 1 - \prod_{i=1}^{K-1} (1 - r_i)$, where w_K is fitness, considering only those incompatibilities that involve the Kth substitution. If the r_i are all fairly small, this is roughly $L_K \approx 1 - (1 - \bar{r})^{K-1}$. Considering the cumulative effects of all K substitutions,

$$L = 1 - w \approx 1 - \prod_{j=1}^{K} (1 - \bar{r})^{j-1} \approx 1 - e^{-K^2 \bar{r}/2}, \quad (8)$$

where L is the strength of reproductive isolation (or the fitness "load" among hybrids due to complementary gene interactions). Thus, early in the divergence of two taxa ($L \ll 1$), the strength of reproductive isolation increases as the square of the number of substitutions: $L \approx K^2 \bar{r}/2$.

Thus the chance of speciation increases much faster than linearly with K (or time) whether speciation typically results from a very small number of genes of large effect (as in the first model) or a large number of genes of smaller effect (as in the second).

The number of substitutions having a substantial effect on reproductive isolation also increases faster than linearly with time. Thus, if one were to double the time since divergence, one would more than double the number of genes having a large effect on hybrid fitness. This result has important consequences for genetic analysis of reproductive isolation (see DISCUSSION). Although it is very difficult to analytically determine the expected size of this effect (the problem involves the product of $K - 1$ integrals, each with different limits of integration), the effect is easily seen in Monte Carlo simulations (Table 1). If, for example, one could detect any gene that decreased hybrid fitness by $\geq 3\%$, the number of alleles with detectably large effects easily increases by five- to 19-fold when the time involved is doubled (the exact size of the effect depends on both $\bar{r}$ and the shape of $f(r)$; Table 1). The effect is even more dramatic if one uses larger, more realistic thresholds (*i.e.*, genetic analysis can only detect genes of fairly large effect).

The role of early *vs.* late substitutions: The above discussion might seem to imply that a gene of known large effect on hybrid fitness was more likely a later than earlier substitution. This is incorrect. Although the probability that a substitution causes hybrid sterility or inviability increases with time, any gene afflicting hybrids is just as likely to have been the first gene to diverge as the last. This is because a late diverging gene must be incompatible with *something*, in particular with some locus that diverged *earlier*.

TABLE 1

Number of substitutions having large effects on reproductive isolation increases faster than linearly

$f(r)$	Mean proportion ($\pm$ var)	Factor	Replicates
Gaussian	0.172 + 0.0008	5.8×	1000
Exponential	0.174 + 0.0009	5.8×	1000
U-shaped	0.052 + 0.0015	19.4×	1000

The second column shows the proportion of substitutions of large effect that occur during the first half of the time to speciation (from simulations). The third column shows the factor by which the number of substitutions of large effect increases when the time of divergence is doubled (reciprocal of column two). If genes of large effect accumulated at a constant rate, the total number of large factors would be 2 times the number fixed in the first half of speciation. For each substitution, $K - 1$ random numbers were drawn from the probability density, $f(r)$, where r is the strength of reproductive isolation between two loci. $\bar{r} = 10^{-3}$. Substitutions continued until hybrid fitness was $<5\%$ of that of the pure species. The number of substitutions of greater-than-threshold effect (threshold = 3%) was tabulated for the first *vs* second halves of the speciation process. Similar results were obtained when different periods in the time to speciation were studied (*e.g.*, the first and second quarters of the time to speciation). Larger thresholds yield even larger values in column three. "Gaussian" densities refer to right-hand tail of a Gaussian centered at zero with $\sigma = \bar{r}\sqrt{\pi/2}$. U-shaped $f(r)$ constructed by adding "spike" of alleles of large effect ($r = 0.5$) to an exponential density (0.05% of probability density fell in this spike).

Formally, if K loci have diverged, the mth locus to diverge ($m < K$) has a probability $P_b \approx 1 - e^{-(m-1)p}$ of being incompatible with a locus that diverged before it, and a probability $P_a \approx 1 - e^{-(K-m)p}$ of being incompatible with a locus that diverged after it. Thus the total probability that locus m affects hybrid fitness is approximately $1 - (e^{-(m-1)p})(e^{-(K-m)p}) = 1 - e^{-(K-1)p}$, which is independent of m. Thus, the fact that a gene has a large effect on hybrid fitness tells us nothing about its place in the chronology of substitutions between two species.

Complex incompatibilities: We have assumed that hybrid incompatibilities involve interactions between pairs of genes. This need not be the case (MULLER 1942). An incompatibility might, for instance, require an interaction among three loci: the hybrid genotype *Abc* might be completely or partially sterile while any genotype consisting of any other combination of alleles at these loci might be perfectly fertile.

Remarkably, such complex interactions have been repeatedly found among the few hybridizations that have been adequately analyzed (MULLER 1942; CABOT *et al.* 1994). In the *Drosophila pseudoobscura* Bogota-USA hybridization, for example, hybrid sterility appears only when males carry the "right" combination of alleles at at least three loci (H. A. ORR, unpublished data). [See GOTTSCHEWSKI (1940) [reviewed in

MULLER 1942], PONTECORVO (1943, p. 60), ORR and COYNE (1989) and CABOT et al. (1994) for similar results.] It thus appears that complex interactions may be quite common. Here I ask how these complex incompatibilities affect the rate of speciation. Later, I consider the problem of why complex incompatibilities are so common.

If speciation involves such complex interactions, the probability of speciation must rise even faster than K^2. In particular, if speciation can result suddenly from an incompatibility between n loci (where n ranges between 2 and K inclusive) with a probability of p_n for each type of incompatibility, then the cumulative probability of speciation is

$$S \approx 1 - \exp\left[-\sum_{n=2}^{K} p_n \binom{K}{n}\right], \qquad (9)$$

where the binomial coefficient $\binom{K}{n} = K!/(n![K-n]!)$, the number of combinations of K objects taken n at a time. Equation 9 shows that S must increase at least as fast as K^2 whether pair-wise incompatibilities are the commonest or the rarest type of incompatibility: $\binom{K}{3}$, for instance, is larger than $\binom{K}{2}$, so inclusion of "triplet" interactions can only increase the cumulative probability of speciation in Equation 9. A similar result holds if speciation results from the accumulation of many incompatibilities of smaller effect (not shown). Equation 9 also makes it clear that reproductive isolation increases faster than linearly as a trivial consequence of the fact that the number of pairs or triplets, and so on, of any object (including substitutions) increases faster than the number of objects itself.

THE FREQUENCY OF COMPLEX INCOMPATIBILITIES

It is not obvious why complex incompatibilities arise so often (MULLER 1942; CABOT et al. 1994). Why should the genetic architecture of reproductive isolation be constructed in such a way that hybrids suffer no decrease in fertility or viability until they carry the "right" alleles at three or more loci? Complex incompatibilities could arise for purely biochemical reasons. For instance, physical interactions between protein products from four loci may be more common than interactions between two gene products. This seems unlikely, however. On the other hand, complex incompatibilities might reflect the redundancy of biochemical pathways: it may be necessary to "knock out" both loci A and B and redundant loci C and D in hybrids to obtain any sterility or inviability (the redundant pathway might involve duplicate genes).

There are, however, other, more evolutionary, reasons why complex incompatibilities may be so common. First, there are simply more combinations of three (or more) than two substitutions. Second, it is "easier" to evolve complex interactions without incurring selection against any of the intermediate steps involved.
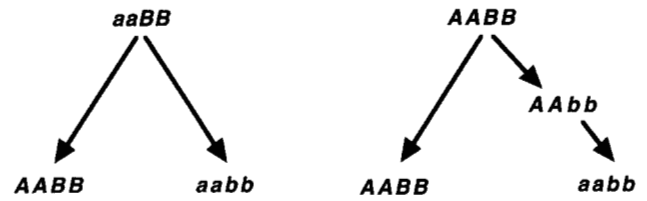

FIGURE 2.—Two imaginable paths for the evolution of two taxa from a common ancestor. One daughter species has genotype "AABB" and the other "aabb". The pathway on the left is allowed by natural selection. The pathway on the right is not allowed by selection as it requires passing through the sterile or inviable "AAbb" hybrid genotype.

As Equation 6 suggests, the number, I_n, of incompatibilities involving an interaction between n genes is

$$I_n = p_n \binom{K}{n}, \qquad (10)$$

i.e., the probability that any randomly chosen n substitutions will be incompatible with each other multiplied by the number of possible combinations of n substitutions, where K is the total number of substitutions that have occurred between two taxa. I assume that the probability that any set of loci is incompatible is independent of whether other such sets are incompatible.

One reason complex incompatibilities may be so common, then, is simply that $\binom{K}{n}$ increases with n (as long as $n < K/2$, which includes the range of biological interest). There are, for example, many more possible combinations of three than two loci.

There is, however, a second reason complex incompatibilities may be common that is not obvious from Equation 10 because it is buried in K: the evolution of simpler incompatibilities is often prevented because it would require passing through the unfit hybrid genotype, i.e., because the next substitution would cause sterility or inviability. To see this, consider the simplest case—an incompatibility involving a single locus ($n = 1$). One species has the genotype AA and the other aa; the sterile or inviable hybrid is Aa. As DOBZHANSKY (1937) and MULLER (1942) pointed out, it is virtually impossible to evolve such an incompatibility because, whether the ancestral genotype was AA or aa, the evolution of the other species' genotype would require passing through the unfit Aa genotype.

The same problem confronts the evolution of incompatibilities involving more than one gene: if an incompatibility involves two genes (e.g., species 1 = $AABB$, species 2 = $aabb$ and the unfit hybrid genotype = $AAbb$), then there are six possible ways to evolve these genotypes from some common ancestor, but only three of these paths do not either begin with or pass through the unfit hybrid genotype. (I assume that we do not know the genotype of the common ancestor and that evolution takes the most parsimonious path between two genotypes, i.e., multiple substitutions at a locus are not allowed. Different orders of substitutions count as different paths.) Two sample paths are shown in Figure
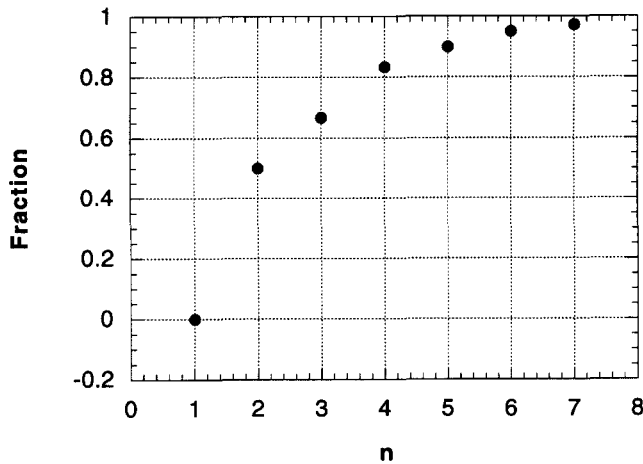
FIGURE 3.—The fraction of imaginable paths, F, connecting common ancestors to two incompatible genotypes that are "allowed" by natural selection (from Equation 11). n is the number of loci involved in the incompatibility. For convenience, I assume that the hybrid genotype is intermediate between the parental genotypes (*i.e.*, m is set equal to the integer closest to $n/2$).

2; the left one is allowable while the right one is not because it would require passing through the unfit *AAbb* genotype.

In general, there are $(n + 1)!$ paths leading from all imaginable common ancestors to the present two species. Some $(n - m)!m!$ of these paths, however, begin with a common ancestor having the genotype of the unfit hybrid, where m is the number of n loci at which the hybrid and one of the present species (it does not matter which) carry the same alleles. Another $n(n - m)!m!$ of all imaginable paths pass through the sterile or inviable hybrid genotype as an intermediate step. Thus, a total of $(n + 1)(n - m)!m!$ paths are proscribed by selection. Therefore, the fraction of *possible* paths is

$$F = 1 - 1 / \binom{n}{m}. \tag{11}$$

[Equation 11 also gives the fraction of possible paths if one begins with some fixed (known) ancestral genotype and considers all ways of evolving two species separated by an incompatibility involving n genes.]

The important point is simple: Equation 11 shows that the fraction of allowable paths is fairly small when n is small but quickly increases as n grows (see Figure 3). Thus the evolution of incompatibilities involving a small number of genes is often prevented by selection because it would require passing through an unfit genotype, for example, when $n = 2$ loci, 50% of the possible ways of evolving these incompatible genotypes from imaginable common ancestors are barred. Conversely, when incompatibilities involve more genes, more of the possible ways of evolving these incompatible genotypes are allowed by evolution. When $n = 4$, for instance, only 17% of all imaginable ways of obtaining these incompatible genotypes are prevented by selection.

In short, *for the same reason that it is "easier" to evolve*

incompatibilities involving two genes than one, so it is easier to evolve incompatibilities involving three (or more) genes than two. Complex incompatibilities may, therefore, be common simply because it is easier for evolution to arrive at such genotypes without passing through any maladaptive intermediate step.

CABOT *et al.* (1994) have also recently considered the evolution of complex incompatibilities and have independently come to essentially the same conclusion. It should be noted, however, that CABOT *et al.* derived the total *number* of possible pathways connecting two species; this number increases very fast with the number of loci involved, suggesting that incompatibilities involving 20 genes may be much easier to evolve than those involving 10. Equation 11 shows, however, that the *fraction* of allowable paths levels off quickly. By the time interactions involve six or seven loci, natural selection almost never bars any imaginable path (Figure 3). There is no reason, therefore, to expect the evolution of extraordinarily complex interactions involving a very large number of genes.

DISCUSSION

The alleles causing postzygotic isolation cannot have caused sterility or inviability when they first arose. This simple fact constrains the ways in which poatzygotic can evolve. Consequently, the genetics of speciation are expected to show several regularities.

First, as MULLER (1942) emphasized, hybrid incompatibilities should be asymmetric: if allele A from one species is incompatible with allele B from the other, alleles a and b must be compatible (as least early in speciation before multiple substitutions at loci are common). Second, derived alleles will be involved in hybrid incompatibilities more often than ancestral. When two lineages experience equal substitution rates, derived alleles are three times more likely to cause hybrid problems than ancestral. Third, it is easier to evolve hybrid incompatibilities between three (or more) than two loci; in the first place, there are simply more possible combinations of three (or more) than two substitutions. In addition, while the evolution of pair-wise incompatibilities is often prevented by natural selection (because the next required substitution would cause sterility or inviability), more complex incompatibilities are easily arrived at without incurring selection against any intermediate step. Last, and most important, hybrid sterility and inviability should evolve faster than linearly with time.

This fact has several interesting consequences. First, the evolution of reproductive isolation may differ from the evolution of "normal" character differences between species. It seems likely that morphological or physiological differences between species usually increase in proportion to the number of substitutions affecting such characters [these characters *can*, however, diverge faster than linearly with time if there is strong epistasis between the genes involved (LYNCH

1994)]. Reproductive isolation *must*, however, increase faster than linearly with time. Although we should not conclude that strong reproductive isolation will arise before morphological or physiological differences (this depends on the value of p or $\bar{r}$), this snowballing certainly improves the odds that substantial reproductive isolation will precede or accompany the evolution of other differences between taxa. Thus the genetics of hybrid sterility and inviability may fortuitously increase the chances that differences evolved in allopatry will be preserved by reproductive isolation upon secondary geographic contact. As FUTUYMA (1987) has argued, this "preserving" effect of reproductive isolation may cause an apparent association between speciation events and morphological evolution in the fossil record.

Second, the increasing rate of speciation requires that we interpret genetic analyses of reproductive isolation with caution. One purpose of such experiments is to estimate the number of genes causing speciation. Most experiments involve taxa (usually Drosophila) that produce sterile or inviable male hybrids only (reviewed by COYNE and ORR 1989a). A traditional concern about these experiments is that one might erroneously count genes that affect hybrid fertility or viability but that evolved *after* the appearance of complete male sterility/inviability. Indeed, in Drosophila, male sterility/inviability often arises quite early and there is a long lag before the evolution of female sterility/inviability (COYNE and ORR 1989b). The likely reasons for this stalling at male effects are complex and need not concern us here (see COYNE and ORR 1989a,b).

Nonetheless, this stalling poses a problem: genetic analysis cannot distinguish between the genes that actually caused the appearance of hybrid sterility (say between NEI's genetic distance $D = 0$ and $D = 0.25$, where we use D as a measure of time) and those that accumulated after the evolution of complete male sterility (say between $D = 0.25$ and $D = 0.50$). The important point is that this problem may be more serious than previously realized. In the single incompatibility model, for example, in going from $D = 0.25$ to $D = 0.50$, we *quadruple*, not double, the observed number of genes affecting hybrid fertility (Equation 6). The same result holds qualitatively if isolation is caused by the cumulative effect of genes of smaller effect. Indeed, the simulation results show that when doubling the time since divergence, it is possible to increase the number of genes of detectable effect by an order of magnitude. The genetics of reproductive isolation will, therefore, become very complicated very quickly. Consequently, we might easily overestimate the number of genes required to obtain hybrid male sterility. The same problem afflicts attempts to estimate the number of genes causing species-level reproductive isolation between taxa producing sterile or inviable hybrids of both sexes (MULLER and PONTECORVO 1940; ORR and COYNE 1989). The only solution, it seems, is to genetically analyze younger taxa.

It is certainly possible that the number of "speciation genes" detected here could be far smaller.

Finally, this work highlights several places where we remain remarkably ignorant about the genetics of speciation. First, we have surprisingly little information about the rate at which reproductive isolation accumulates. COYNE and ORR's (1989b) review of the Drosophila literature, in which the strength of reproductive isolation between taxa was compared with genetic distance, probably provides the best data. Although these data are rough (*e.g.*, partial sterility/inviability of a hybrid sex was ignored), they provide some support for a "snowballing" of postzygotic isolation. In particular, an approximate doubling of the number of incompatibilities separating two taxa does not require a doubling of time. COYNE and ORR's data show that the mean (+ SE) genetic distances at which hybrid males are sterile/inviable in only one *vs.* both directions of a cross do not significantly differ: $D = 0.369 + 0.183$ ($n = 5$) *vs.* $D = 0.226 + 0.021$ ($n = 28$), respectively (all female hybrids are viable and fertile). Similarly, the D at which nonreciprocal and reciprocal female sterility/inviability appear do not differ: $D = 0.950 + 0.123$ ($n = 6$) *vs.* $D = 1.009 + 0.087$ ($n = 20$), respectively (all male hybrids are sterile/inviable). Thus, for both males and females, reciprocal sterility/inviability (requiring at least two incompatibilities) seems to appear on the heels of nonreciprocal effects (requiring at least one incompatibility).

While the data are not voluminous, this result is consistent with the notion that the number of incompatibilities separating taxa accumulates at an accelerating rate (Equation 6). Unfortunately, a snowballing of isolation is not the only process that could yield this pattern: if hybrid fitness falls off suddenly (as with truncation selection), the interarrival times for one-way *vs.* reciprocal hybrid sterility/inviability become smaller even if contributions to the underlying additive "hybrid breakdown" scale occur at a steady, not accelerating, pace (TURELLI and ORR, unpublished data). Thus, the genetic distance data are consistent with, but do not prove, a snowballing of reproductive isolation.

Last, and most remarkably, we have very little information about the probability, p, that two genes will be incompatible with each other, causing hybrid inviability or sterility. Consequently, we have little feel for how "easy" speciation is; if many pairs (or triplets, etc.) of substitutions can form complementary lethals/steriles, speciation may not usually require a very large number of substitutions. If, on the other hand, complementary interactions are very rare, an enormous number of substitutions may be required before speciation is likely. We know only that complementary lethals and steriles can be found in nature (DOBZHANSKY 1946; DOBZHANSKY *et al.* 1959; KRIMBAS 1960) and that, at least in *Drosophila melanogaster*, they are very rare. TEMIN *et al.* (1969) found that ≈ 0.005 of all second and third chromosome combinations harbor complementary lethals. Unfortunately, we have little idea about the number of (coding) allelic

differences between such chromosomes and so cannot easily translate these figures into values of p.

Given this paucity of direct data, it may be worth noting that indirect estimates of p are possible. In Drosophila, COYNE and ORR (1989b, Table 2) showed that hybrid male sterility/inviability appears at an average genetic distance of $D \approx 0.25$. Male effects never appear before $D = 0.07$. Because Drosophila (at least D. *melanogaster*) appears to have ~5000 loci (most of which are essential) (ASHBURNER 1989), these taxa have diverged at roughly 1200 loci (we dangerously assume here that allozyme data are representative of most loci). If most cases of hybrid sterility/inviability have a simple genetic basis, Equation 6 suggests that $p \approx 10^{-6}$. This figure is, of course, the result of several wild assumptions and must be treated with skepticism. Nonetheless, it seems unlikely that this estimate is off by several orders of magnitude: for example, if $p \approx 10^{-4}$, only 400 loci would have to diverge and hybrid male sterility/inviability would typically appear long before $D \approx 0.25$.

Far more direct estimates of p and $\bar{r}$ will, of course, be possible once the number of "speciation genes" separating young taxa of known genetic distance is determined experimentally. Until then we are left with only the roughest and most indirect estimates of these critical parameters.

LITERATURE CITED

ASHBURNER, M., 1989 *Drosophila: A Laboratory Handbook.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BARTON, N. H., and B. CHARLESWORTH, 1984 Genetic revolution, founder effects, and speciation. Annu. Rev. Ecol. Syst. **15**: 133–164.

CABOT, E. L., A. W. DAVIS, N. A. JOHNSON and C.-I. WU, 1994 Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. Genetics **137**: 175–189.

CHARLESWORTH, B., J. A. COYNE and N. H. BARTON, 1987 The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. **130**: 113–146.

CHRISTIE, P., and M. R. MACNAIR, 1984 Complementary lethal factors in two North American populations of the yellow monkey flower. J. Hered. **75**: 510–511.

CHRISTIE, P., and M. R. MACNAIR, 1987 The distribution of postmating reproductive isolating genes in populations of the yellow monkey flower, *Mimulus guttatus.* Evolution **41**: 571–578.

COYNE, J. A., and H. A. ORR, 1989 Patterns of speciation in *Drosophila.* Evolution **43**: 362–381.

COYNE, J. A., and H. A. ORR, 1989 Two rules of speciation, pp. 180–207 in *Speciation and its Consequences*, edited by D. OTTE and J. ENDLER. Sinauer Associates, Sunderland, MA.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper and Row Publishers, New York.

DARWIN, C., 1859 *The Origin of Species.* Avenel Books Edition, New York.

DOBZHANSKY, T., 1936 Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics **21**: 113–135.

DOBZHANSKY, T., 1937 *Genetics and the Origin of Species.* Columbia University Press, New York.

DOBZHANSKY, T., 1946 Genetics of Natural Populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura.* Genetics **31**: 269–290.

DOBZHANSKY, T., H. LEVENE, B. SPASSKY and N. SPASSKY, 1959 Release of genetic variability through recombination. III. *Drosophila prosaltans.* Genetics **44**: 75–92.

FUTUYMA, D. J., 1987 On the role of species in anagenesis. Am. Nat. **130**: 465–473.

GERSTEL, D. U, 1954 A new lethal combination in interspecific cotton hybrids. Genetics **39**: 628–639.

HOLLINGSHEAD, L., 1930 A lethal factor in *Crepis* effective only in interspecific hybrids. Genetics **15**: 114–140.

KRIMBAS, C. B., 1960 Synthetic sterility in *Drosophila willistoni.* Proc. Natl. Acad. Sci. USA **46**: 832–833.

LYNCH, M., 1994 Neutral models of phenotypic evolution, pp. 86–108 in *Ecological Genetics*, edited by L. REAL. Princeton University Press, Princeton, NJ.

MULLER, H. J., 1939 Reversibility in evolution considered from the standpoint of genetics. Biol. Rev. Camb. Philos. Soc. **14**: 261–280.

MULLER, H. J., 1940 Bearing of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by J. HUXLEY. Clarendon Press, Oxford.

MULLER, H. J., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. **6**: 71–125.

MULLER, H. J., and G. PONTECORVO, 1940 Recombinants between Drosophila species, the F₁ hybrids of which are sterile. Nature **146**: 199.

MULLER, H. J., and G. PONTECORVO, 1942 Recessive genes causing interspecific sterility and other disharmonies between *Drosophila melanogaster* and *simulans.* Genetics **27**: 157.

NEI, M., 1976 Mathematical models of speciation and genetic distance, pp. 723–766 in *Population Genetics and Ecology*, edited by S. KARLIN and E. NEVO. Academic Press, New York.

NEI, M., T. MARUYAMA and C.-I. WU, 1983 Models of evolution of reproductive isolation. Genetics **103**: 557–579.

ORR, H. A., 1993 A mathematical model of Haldane's rule. Evolution **47**: 1606–1611.

ORR, H. A., and J. A. COYNE, 1989 The genetics of postzygotic isolation in the *Drosophila virilis* group. Genetics **121**: 527–537.

PANTAZIDIS, A. C., and E. ZOUROS, 1988 Location of an autosomal factor causing sterility in *Drosophila mojavensis* males carrying the *Drosophila arizonensis* Y chromosome. Heredity **60**: 299–304.

PONTECORVO, G., 1943 Viability interactions between chromosomes of *Drosophila melanogaster* and *Drosophila simulans.* J. Genet. **45**: 51–66.

SLATKIN, M., 1982 Pleiotropy and parapatric speciation. Evolution **36**: 263–270.

TEMIN, R. G., H. U. MEYER, P. S. DAWSON and J. F. CROW, 1969 The influence of epistasis on homozygous viability depression in *Drosophila melanogaster.* Genetics **61**: 497–519.

TEMPLETON, A. R., 1981 Mechanisms of speciation—a population genetics approach. Annu. Rev. Ecol. Syst. **12**: 23–48.

WALSH, J. B., 1982 Rate of accumulation of reproductive isolation by chromosome rearrangements. Am. Nat. **120**: 510–532.

WITTBRODT, J., D. ADAM, B. MALITSCHEK, W. MAUELER, F. RAULF *et al.*, 1989 Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus.* Nature **341**: 415–421.

WU, C.-I., and A. T. BECKENBACH, 1983 Evidence for extensive genetic differentiation between the sex-ratio and the standard arrangement of *Drosophila pseudoobscura* and *D. persimilis* and identification of hybrid sterility factors. Genetics **105**: 71–86.