ORIGINAL ARTICLE

Dobzhansky–Muller incompatibilities and adaptation to a shared environment

RL Unckless and HA Orr

Department of Biology, University of Rochester, Rochester, NY, USA

Natural selection might drive the evolution of postzygotic reproductive isolation even when allopatric populations adapt to identical environments, an idea first suggested by Muller (1942). Here, we analyze this scenario mathematically, focusing on the evolution of a Dobzhansky–Muller incompatibility (DMI) between populations. Our results identify a potential problem with Muller's scenario: adaptation to identical environments can often involve substitution of the same alleles, precluding formation of a hybrid incompatibility.

We show that the probability of evolving a DMI falls as selection coefficients among beneficial alleles become less similar. The reason is that if one locus is under much stronger selection than the other, that locus is much more likely to experience a substitution first in both populations. This precludes the development of a DMI, which requires different substitutions in the two populations.
Heredity (2009) **102,** 214–217; doi:10.1038/hdy.2008.129; published online 14 January 2009

There can now be little doubt that natural selection plays an important role in the evolution of postzygotic reproductive isolation. Over the last decade or so, considerable work has shown that positive natural selection plays a part in the evolution of both extrinsic (ecological inviability or behavioral sterility) and intrinsic postzygotic (hybrid inviability or sterility) isolation (reviewed by Coyne and Orr, 2004). Although the former role is perhaps not surprising—it seems intuitively clear that selection shapes divergence in characters of ecological significance—the latter role was less expected. It was not obvious that substitution at the genes causing intrinsic hybrid sterility or inviability is driven by positive natural selection. Nonetheless, molecular population genetic analysis of several genes known to cause hybrid sterility or inviability, for example, *OdsH, Nup96, Lhr* and *Hmr*, shows strong evidence of positive Darwinian selection (reviewed by Coyne and Orr, 2004; Orr *et al.*, 2007). It seems clear, then, that many of the substitutions underlying genic (Dobzhansky–Muller) incompatibilities are driven by selection. (This selection was not, of course, selection for postzygotic isolation but for some other character; this evolution pleiotropically gives rise to lowered fitness in hybrids.)

Natural selection might drive the evolution of Dobzhansky–Muller incompatibilities (DMIs) in two ways. First, allopatric populations might adapt to their respective—and different—environments and hybrid sterility or inviability might arise as a pleiotropic side effect of this divergent selection. One population, for example, might adapt to a dry environment and another to a wet. The genes underlying these ecological adaptations might later cause incompatibilities in hybrids formed between the populations. Second, allopatric populations might adapt to the same environment but in different ways genetically. As populations can respond to identical selection pressures in different ways, two populations might still arrive at different genotypes. If so, their hybrids might still suffer incompatible gene combinations. This scenario was first described by Muller (1942).

Here, we point out a potential problem with this identical environment scenario. The problem is that it is harder than it might first seem for two populations to adapt to an identical environment in different ways, thereby allowing the evolution of postzygotic isolation. The reason is that natural selection, although not deterministic, can often cause the substitution of the same mutations in two populations. This can lower dramatically the chance of forming a hybrid incompatibility. This problem does not arise when incompatibilities result from adaptation to different environments.

In this note, we calculate the probability of evolving a DMI when natural selection drives adaptation to an identical environment. We consider the simplest scenario in which allopatric populations evolve by substitution of recurrent new mutations.

The model

We begin by considering two allopatric populations that experience no gene flow. These populations begin with identical genotypes (both are fixed for wild-type alleles at all relevant loci) and experience identical environments. At some point in time this shared environment changes and the populations begin to adapt to the new, but still shared, environment. (Environments need not be strictly identical; formally, our calculations assume only that those aspects of the environment that affect selection coefficients at the loci we follow are the same.) Mutation

is recurrent and evolution involves the substitution of new mutations. We assume that all mutations have independent effects on fitness. Following Gillespie (1984, 1991), we consider the strong selection ($Ns \gg 1$) and weak mutation ($N\mu \ll 1$) domain. Results involving weak selection and/or strong mutation (for example, $N\mu \sim 1$) might well differ.

Following Nei (1976), we can simplify our analysis by considering haploids. This captures the essence of our problem without requiring us to deal with unnecessary details about the dominance of incompatibilities. In the case of a two-locus DMI, both populations initially have genotype A_0B_0. In the new environment, mutations A_1 and B_1 are beneficial. But, when brought together, A_1 and B_1 cause an in compatibility. Relative fitnesses among genotypes are:

A_0B_0	1
A_1B_0	$1+s_A$
A_0B_1	$1+s_B$
A_1B_1	$1-t$,

where s_A, s_B, and t are all positive.

Note that, once A_1 is fixed in a population, B_1 cannot be fixed as this would yield the unfit A_1B_1 genotype; similarly, once B_1 is fixed, A_1 cannot be fixed. We assume that rates of mutation to A_1 and B_1 are equal; back-mutation to A_0 and B_0 is not allowed.

A DMI arises if one population evolves the A_1B_0 genotype and the other evolves the A_0B_1 genotype, allowing formation of unfit A_1B_1 hybrids if the populations were to come into contact. A DMI cannot occur if both populations evolve the same genotype, either A_1B_0 or A_0B_1. In this case, hybrids between the populations would retain a pure-population genotype and would be perfectly fit.

We calculate the ultimate probability that two loci form a DMI, P_{DM}, that is, the probability that populations arrive at incompatible genotypes after an extended (essentially infinite) period of time; in effect, then, we assume that taxa remain allopatric for very long periods of time. Note also that the loci we follow are capable of causing a hybrid incompatibility if they evolve in the 'right' way in our populations. P_{DM} represents, therefore, a conditional probability: given that the loci we follow are capable of causing a hybrid incompatibility, what is the probability that they do so?

The neutral case

Though mathematically trivial, it is worth first considering the case in which the evolution is neutral: $s_A = s_B = 0$. We will see that this scenario represents a limit in the case of evolving a hybrid incompatibility by natural selection. We still assume that $t > 0$, that is, A_1B_1 is unfit. Populations will ultimately either fix A_1 or B_1. Under neutrality, there is a one-fourth chance that both populations fix A_1 (arriving at A_1B_0) and a one-fourth chance that both fix B_1 (arriving at A_0B_1). The probability of an incompatibility is one minus the sum of these probabilities, $P_{DM} = \frac{1}{2}$.

This logic is extended trivially to more than two populations, each of which evolves independently. The probability of incompatibility, that is, that at least one pair of a allopatric populations arrives at incompatible

genotypes, is one minus the probability that all a populations arrive at the same genotype: $P_{DM} = 1 - 1/2^{a-1}$.

Selection: two loci

Now consider the evolution by natural selection. In the special case in which $s_A = s_B$, matters are again simple. Because A_1 and A_2 are equally likely to fix first in a population, there is again a one-fourth chance both populations evolve to be A_1B_0 and a one-fourth chance both evolve to be A_0B_1; the probability that two populations arrive at different genotypes, causing postzygotic isolation, is thus again $P_{DM} = \frac{1}{2}$, as noted by Gavrilets (2004). Finding the probability of incompatibility with arbitrary s_A and s_B is not so straightforward, however.

To find this probability, we take advantage of several results obtained by Gillespie (1984, 1991). Gillespie considered a large population in which beneficial mutations are sufficiently rare ($Nu \ll 1$) to have independent fates (that is, clonal interference does not occur) and in which beneficial mutations have definite fitness advantages ($Ns \gg 1$). Standard population genetic theory shows that a new beneficial mutation has a probability of fixation of $\Pi \approx 2s$, where the approximation assumes that s is fairly small (Haldane, 1927). Because most new mutations are lost accidentally, it takes time before recurrent mutation produces a copy of a beneficial allele that is destined to sweep to fixation. Gillespie showed that the waiting time to the appearance of such a mutation is geometrically distributed with a mean of $1/(2Nus)$ generations (in diploids, $1/(4Nus)$). For a A_0B_0 population, the mean waiting time to fixation of A_1 is $1/(2Nus_A)$ generations and the mean waiting time to fixation of B_1 is $1/(2Nus_B)$ generations. As each of these times is approximately exponentially distributed, the probability that A_1 fixes first in a population equals the probability that an exponential random variable with a mean of $1/(2Nus_A)$ is smaller than an exponential random variable with a mean of $1/(2Nus_B)$. Gillespie shows that this probability is $\Pi_A = s_A/(s_A + s_B)$. Conversely, the probability that B_1 fixes before A_1 is $\Pi_B = s_B/(s_A + s_B)$. Note that these values are independent of population size (so long as it is large).

As Π_A and Π_B are identical in our $a = 2$ populations, the probability that one population becomes A_1B_0 and the other A_0B_1, yielding a DMI, is

$$P_{DM} = \frac{2s_As_B}{(s_A + s_B)^2} \qquad (1)$$

This probability is maximized when $s_A = s_B$ and $P_{DM} = 1/2$. This maximum probability of incompatibility is identical to the probability of incompatibility under neutrality.

Equation (1) makes good sense. When s_A and s_B are equal, there is a good chance that one population will substitute A_1 and the other B_1, allowing the formation of a hybrid incompatibility. But when s_A and s_B are very different, both populations are likely to substitute the same allele—the one with the greatest advantage—arriving at the same genotype. This precludes formation of a hybrid incompatibility. When $s_A = 0.05$ and $s_B = 0.001$, for instance, the probability of incompatibility decreases to only 3.8%. The above logic is extended trivially to the case of a allopatric populations,

yielding

$$P_{DM} = 1 - [s_A/(s_A + s_B)]^a - [s_B/(s_A + s_B)]^a$$

Although we do not know the actual values of s for beneficial mutations in nature, a reasonable argument can be made that s is approximately exponentially distributed. Gillespie (1984, 1991) and Orr (2003) pointed out that, given a variety of distributions of fitnesses (for example, normal, gamma, exponential, logistic and so on), extreme value theory shows that the right-hand tail of these distributions declines exponentially. Because wild-type alleles are typically highly fit, this tail corresponds approximately to the distribution of selection coefficients among beneficial mutations (see Gillespie and Orr for details). Although this argument is far from decisive (see Joyce *et al.*, 2008), especially when considering mutations at different loci, an exponential distribution of s at least represents a reasonable guess about biological reality.

We can thus ask: What is the average probability of incompatibility between a pair of populations when s_A and s_B are drawn from a common exponential distribution? The answer is

$$E[P_{DM}] = \int_0^\infty \int_0^\infty \frac{2 s_A s_B}{(s_A + s_B)^2} \lambda^2 e^{-\lambda(s_A + s_B)} ds_A \, ds_B = \frac{1}{3}, \quad (2)$$

where $1/\lambda$ is the mean s among beneficial mutations. In words, Equation (2) shows that pairs of populations evolve in different directions—yielding a DMI in hybrids—one-third of the time. Surprisingly, this result is independent of mean selection coefficient, population size, and all other biological parameters. The finding is closely connected to that in Orr (2005) for parallel adaptation, who did not consider reproductive isolation and who took a different approach mathematically.

A similar calculation shows that, given a independently evolving populations, the expected probability of incompatibility is

$$E[P_{DM}] = 1 - 2 \int_0^\infty \int_0^\infty \frac{s_A}{(s_A + s_B)^a} \lambda^2 e^{-\lambda(s_A + s_B)} ds_A ds_B$$
$$= \frac{a-1}{a+1} \quad (3)$$

When $a = 2$, we recover Equation (2). As expected, postzygotic isolation is more likely with more allopatric populations.

Selection: complex incompatibilities

A DMI need not involve only two loci. Instead, genetic analyses of hybrid sterility and inviability have shown that complex incompatibilities, involving three or more loci, are common (Cabot *et al.*, 1994; Orr, 1995). We can extend our calculations to these cases.

It is again worth considering the neutral case. Consider two allopatric populations that experience recurrent mutation at n loci with no back-mutation. Though each mutation is neutral, any individual who carries all n alleles suffers a genic incompatibility. Given enough time, each population will ultimately fix $n-1$ of the alleles; the last allele cannot fix as this would yield the

unfit hybrid genotype. A hybrid incompatibility will thus form unless two populations fix the same genotypes, which occurs with probability $1/n$. The probability of incompatibility is therefore $P_{DM} = 1 - 1/n$. If $n = 3$ loci, $P_{DM} = 2/3$. Given a allopatric populations, we get $P_{DM} = 1 - 1/n^{a-1}$.

When substitutions are driven by natural selection, the n-locus case is messy mathematically. We thus present results for the three-locus ($n = 3$) case only, which captures the biological point we wish to make. Alleles A_1, B_1 and C_1 are beneficial when alone or in pairs; but any individual that carries all three alleles suffers a DMI and has fitness $1 - t$. We assume that fitnesses are otherwise independent, that is, multiplicative: $A_1 B_0 C_0$ has fitness $1 + s_A$, whereas $A_1 B_1 C_0$ has fitness $(1 + s_A)(1 + s_B)$, and so on.

Populations will ultimately fix two of the three mutations. Which two depends on the magnitudes of s_A, s_B and s_C. For example, the probability, $P(A_1, B_1)$, that a population ultimately fixes alleles A_1 and B_1 is the sum of two events: that A_1 is fixed first and then B_1 or that B_1 is fixed first and then A_1. Using Gillespie's arguments to calculate the probabilities of these events, straightforward calculations show that

$$P(A_1, B_1) = \frac{s_A s_B + \Pi s / \sum s}{s_A s_B + s_C \sum s} \quad (4)$$

where Πs is the product of the three selection coefficients and Σs is their sum. $P(A_1, C_1)$ and $P(B_1, C_1)$ have the same form, with the obvious change of subscripts.

As postzygotic isolation evolves unless both populations evolve the same genotype, the probability of incompatibility is

$$P_{DM} = 1 - P(A_1, B_1)^2 - P(A_1, C_1)^2 - P(B_1, C_1)^2 \quad (5)$$

Table 1 shows that, when selection coefficients among the three alleles are similar, evolution of a hybrid incompatibility is likely. In particular, when $s_A = s_B = s_C$, the chance of isolation is maximized at $P_{DM} = 2/3$. Again, then, the maximum probability of incompatibility under natural selection equals that under neutrality. But when the selection coefficients are very different, postzygotic isolation becomes much less likely. If, for instance, $s_A = 0.001$, $s_B = 0.03$ and $s_C = 0.05$, P_{DM} decreases to 7.7%.

Finally, we can ask about the mean probability of incompatibility when selection coefficients involved in three-locus incompatibilities are drawn from an exponential distribution. Although the integration required appears intractable, numerical work shows that

Table 1 Probabilities of postzygotic isolation as a function of selection coefficients among beneficial mutations

Number of loci	Selection coefficients (s_x)	P_{DM}
2	Neutral	0.500
	0.01, 0.01	0.500
	0.005, 0.01	0.444
	0.001, 0.01	0.165
3	Neutral	0.667
	0.01, 0.01, 0.01	0.667
	0.005, 0.075, 0.01	0.458
	0.001, 0.075, 0.01	0.168

$E[P_{DM}] \approx 3/7$. The mean probability of incompatibility thus increases somewhat when DMIs are complex.

Conclusions

Although the mathematics involved in our analysis is simple, the biological implications that emerge seem non-trivial. We have obtained two main results. First, when allopatric populations adapt to identical environments and the loci we follow can potentially form a DMI, the probability that they do so is constrained in a way that does not occur when populations adapt to different environments. The reason is that the evolution of a DMI requires that populations evolve different genotypes. But adaptation to an identical environment is somewhat repeatable: populations will often substitute the same alleles, arriving at the same genotype. (See Wood et al. (2005), who reviewed the evidence for parallel genetic evolution in the quantitative genetic and experimental evolution literature). When populations independently arrive at the same genotype, a DMI cannot arise. Of course, if selection is spatially heterogeneous, that is, populations adapt to different environments, incompatibilities can arise more easily (see also Navarro and Barton, 2003; Kondrashov, 2003; Gavrilets, 2004). Our analysis further shows that, as the number of loci involved in the incompatibility increases, the probability of evolving an incompatibility also increases.

Our second finding is that the probability of postzygotic isolation decreases as selection coefficients among beneficial mutations become less similar. In fact, the maximum probability of isolation occurs when selection coefficients at each locus are equal; in this case, the probability of incompatibility equals that under neutrality. Under both the equal-s and neutral scenarios, evolution randomly chooses alleles at the relevant loci for substitution. Populations thus often arrive at different genotypes, allowing postzygotic isolation. It should be noted, however, that the time-scales for the evolution of postzygotic isolation under the equal-s and neutral scenarios are very different (Gavrilets, 2004): neutral evolution is slow.

Most of our mathematical results have not been noted before and the reason seems clear. Many previous analyses of the Dobzhansky–Muller model (including our own, for example, Orr, 1995) typically focused on the long-term accumulation of alleles causing hybrid incompatibilities and were agnostic about whether substitutions were driven by natural selection or genetic drift. Selection coefficients among beneficial alleles did not, therefore, generally appear in previous theory (for important exceptions, see Nei (1976); Gavrilets (2003, 2004); Gavrilets, in particular, produced theory that considered the role of selection in the buildup of incompatibilities in spatially-structured populations).

In closing, it is important to guard against a possible misinterpretation of our results. We do not suggest that the evolution of postzygotic isolation in identical ecological environments is unlikely. Indeed, a common cause of reproductive isolation by DMIs could be genetic conflict, in which substitutions are driven by arm races among different sites in the genome. Such genetic-conflict-based reproductive isolation can obviously evolve in identical ecological environments. The reason our results do not militate against this form of speciation is subtle. Our model assumes that selection coefficients among beneficial mutations are independent. Thus, if one mutation fixes, selection coefficients at all other mutations remain unchanged. This need not, of course, be true and it not generally true under genetic conflict. With conflict, a mutation might become beneficial only after another has fixed. (For example, an allele that suppresses meiotic drive is favored only after a driver mutation has appeared.). Put differently, genetic conflict does, in fact, involve a kind of adaptation to different environments, but the relevant 'environments' are genetic, that is, involve genotypes at other loci.

Acknowledgements

References

Cabot EL, Davis AW, Johnson NA, Wu CI (1994). Genetics of reproductive isolation in the Drosophila simulans clade: complex epistasis underlying hybrid male sterility. Genetics 137: 175–189.

Coyne JA, Orr HA (2004). Speciation. Sinauer Associates Inc.: Sunderland, MA.

Gavrilets S (2003). Models of speciation: What have we learned in 40 years? Evolution 57: 2197–2215.

Gavrilets S (2004). Fitness Landscapes and the Origin of Species. Princeton University Press: Princeton, New Jersey.

Gillespie JH (1984). Molecular evolution over the mutational landscape. Evolution 38: 1116–1129.

Gillespie JH (1991). The Causes of Molecular Evolution. Oxford University Press: New York.

Haldane JBS (1927). A mathematical theory of natural and artificial selection, part V: selection and mutation. Proc Camb Phil Soc 28: 838–844.

Joyce P, Rokyta DR, Beisel CJ, Orr HA (2008). A general extreme value theory model for the adaptation of DNA sequences under strong selection and weak mutation. Genetics 180: 1627–1643.

Kondrashov AS (2003). Accumulation of Dobzhansky–Muller incompatibilities within a spatially structured population. Evolution 57: 151–153.

Muller HJ (1942). Isolation mechanisms, evolution and temperature. Biol Symp 6: 71–125.

Navarro A, Barton NH (2003). Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution 57: 447–459.

Nei M (1976). Mathematical models of speciation and genetic distance. In: Karlin S, Nevo E (eds). Population genetics and ecology. Academic Press Inc.: New York.

Orr HA (1995). The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139: 1805–1813.

Orr HA (2003). The distribution of fitness effects among beneficial mutations. Genetics 163: 1519–1526.

Orr HA (2005). The probability of parallel evolution. Evolution 59: 216–220.

Orr HA, Masly JP, Phadnis N (2007). Speciation in Drosophila: from phenotypes to molecules. J Hered 98: 103–110.

Wood TE, Burke JM, Rieseberg LH (2005). Parallel genotypic evolution: when evolution repeats itself. Genetica 123: 157–170.