The Origin of Interspecific Genomic Incompatibility via Gene Duplication

Michael Lynch; Allan G. Force

The American Naturalist, Vol. 156, No. 6. (Dec., 2000), pp. 590-605.

Stable URL:
http://links.jstor.org/sici?sici=0003-0147%28200012%29156%3A6%3C590%3ATOOIGI%3E2.0.CO%3B2-D

The American Naturalist is currently published by The University of Chicago Press.

VOL. 156, NO. 6 THE AMERICAN NATURALIST DECEMBER 2000

The Origin of Interspecific Genomic Incompatibility via Gene Duplication

Michael Lynch* and Allan G. Force

Department of Biology, University of Oregon, Eugene, Oregon 97403

Submitted April 2, 2000; Accepted July 12, 2000

ABSTRACT: One of the great unsolved mysteries of evolutionary biology concerns the genetic mechanisms underlying the origin of genomic incompatibilities between species. Two prevailing thoughts are that such incompatibilities often result from epistatically interacting genes that act as loss-of-function alleles in hybrid backgrounds or from chromosomal rearrangements that result in mis-segregation during meiosis in hybrids. However, it is unclear how genes that cause a radical breakdown in hybrids arise without reducing fitness within species, and numerous cases of speciation appear to be unassociated with obvious chromosomal rearrangements. Here we suggest that duplicate genes, and more generally any kind of genomic redundancies, provide a powerful substrate for the origin of genomic incompatibilities in isolated populations. The divergent resolution of genomic redundancies, such that one population loses function from one copy while the second population loses function from a second copy at a different chromosomal location, leads to chromosomal repatterning such that gametes produced by hybrid individuals can be completely lacking in functional genes for a duplicate pair. Under this model, incompatibility factors accumulate with essentially no loss of fitness within populations as postulated under the Bateson-Dobzhansky-Muller (BDM) model of speciation and despite the fact that they arise from degenerative mutations. However, unlike the situation often envisioned under the BDM model, no change in the mode of gene action in hybrid backgrounds need be invoked. The plausibility of this model derives from a number of recent observations, including the fact that most genomes harbor substantial numbers of gene duplicates whose turnover is common and ongoing process and the fact that many genes have complex regulatory regions that facilitate their divergent resolution in sister taxa.

Keywords: complementation, gene duplication, gene expression patterns, genomic redundancy, reproductive incompatibility, speciation.

* To whom correspondence should be addressed; e-mail: mlynch@oregon.uoregon.edu.

Am. Nat. 2000. Vol. 156, pp. 590–605.

A long-standing goal of evolutionary biology is the elucidation of the genetic mechanisms underlying the speciation process. Although the precise sets of genes that are responsible for reproductive isolation between species pairs are likely to be serendipitous outcomes of mutation, selection, and random genetic drift, a general mechanism for the origin of interspecific genomic incompatibility was suggested by Bateson, Dobzhansky, and Muller (BDM; see Orr 1996 for a review). Bateson, Dobzhansky, and Muller envisioned a situation in which one daughter taxon would become fixed for an allele at one locus (e.g., an **A** to **a** mutation), whereas the other daughter taxon would become fixed for a second allele at another locus (e.g., a **B** to **b** mutation). The key to this model is the assumption that although both of these mutations are neutral (or advantageous) within the populations in which they arise, they cause inviability and/or infertility when expressed together in hybrid offspring. Some mathematical features of this model have been explored (Nei et al. 1983; Wagner et al. 1994; Orr 1995; Gavrilets and Gravner 1997; Gavrilets 1999).

Although alternative frameworks relating epistasis and speciation have been discussed (Templeton 1981; Wade and Goodnight 1998), to our knowledge, no significant challenge to the classical BDM model of speciation has been mounted, and we do not suggest one here. However, despite the appeal of this model, an explicit genetic mechanism by which mutations with neutral (or beneficial) effects within recently isolated taxa commonly yield negative effects when expressed together in hybrids has not yet been elucidated. Here we suggest that the divergent evolution of gene duplicates provides a plausible and powerful isolating mechanism that is entirely compatible with the basic features of the BDM model—evolutionary changes can be essentially neutral within species, while having substantial deleterious effects in hybrid backgrounds.

Divergent Degeneration of Ancestral Gene Duplicates

It is now well established that the genomes of most eukaryotic organisms contain thousands of gene duplicates produced by an array of events, including tandem dupli-

cation of single genes, replicative translocation, chromosomal duplication, and polyploidization. For example, 40%–50% of the genes in the nematode *Caenorhabditis elegans* and in the fly *Drosophila melanogaster* are recognizable duplicates of various ages (Rubin et al. 2000). The genomes of virtually all vertebrates (Nadeau and Sankoff 1997; Postlethwait et al. 1998) contain two or more copies of gene family members for large numbers of functional proteins. Genomic analyses suggest that the members of the great ape lineage (humans, gorillas, and chimpanzees), which diverged only 5 Ma, differ substantially in terms of the numbers and locations of duplicate genes (Pennisi 1998); and similar inferences have been made for the congeneric nematodes *C. elegans* and *Caenorhabditis briggsae* (Robertson 1998). A complete analysis of human chromosome 22 revealed several duplicated regions separated by substantial physical distances (Dunham et al. 1999). Probably at least one-third of all flowering plants and even higher proportions of mosses and ferns are polyploid derivatives (Lewis 1980), and most diploid plants also carry many duplicate loci (Gottlieb 1982). The relatively streamlined genome of *Arabidopsis thaliana* contains large numbers of gene duplicates (Lin et al. 1999; Mayer et al. 1999); many of these duplicates are tandemly arrayed, but there are also several examples of replicative translocation of chromosomal regions containing dozens to hundreds of genes, and the species may have even experienced an ancient genome duplication event (Grant et al. 2000).

It has often been argued that duplicate-gene preservation by natural selection requires the acquisition of a mutation to a novel and beneficial function by one of the two copies, with the other copy retaining the original function (Ohno 1970; Walsh 1995; Sidow 1996; Cooke et al. 1997; Nadeau and Sankoff 1997). However, because the vast majority of mutations are deleterious and the null mutation rate per gene is on the order of 10^{-6} per year or higher (Lynch and Walsh 1998), the most common fate of a pair of gene duplicates is thought to be silencing of one member of the pair within a few million years (Nei and Roychoudhury 1973; Takahata and Maruyama 1979; Li 1980; Watterson 1983; Walsh 1995). With this idea in mind, Werth and Windham (1991) suggested a very simple genetic mechanism for the allopatric origin of new species. Although their focus was on polyploid plants, the mechanism that they envisioned applies to any duplicate gene in any organism.

Consider a fully functional (and fully redundant) pair of duplicate genes in an ancestral species, fixed respectively for alleles **A** and **B** (which are not necessarily different), and suppose that two subpopulations of this species become spatially isolated. Letting **a** and **b** denote null alleles for the two loci, then ignoring rare neofunctionalization events, each descendent taxon will randomly lose function at one of the two loci. Thus, divergent resolution of the two loci, with one sister species becoming fixed for **A** and **b** alleles and the other for **a** and **B** alleles, would occur with 50% probability. Hybridization between two such taxa would then lead to **AaBb** progeny. Supposing the haploid product of these loci was essential to gamete function, then if the two loci segregate independently, one-fourth of the gametes produced by the F_1 individuals would be of type **ab** and lacking in function. If the loci were essential to zygote function, then one-sixteenth of the F_2 progeny would be of genotype **aabb** and completely lacking in viability and/or fertility, and another one-fourth would contain three null alleles and possibly have reduced function. With divergent resolution of multiple loci, the magnitude of hybrid inviability/infertility would be magnified accordingly.

An attractive feature of this model is that it is totally driven by degenerative mutations, which we know are much more common than beneficial mutations. Moreover, the model is functionally compatible with the BDM model in all respects, except that there is no need to invoke a change in gene action in hybrid backgrounds. Finally, the premise on which the model is built—that silencing of one member of the pair is a common fate of duplicate genes—is well documented in plants (Wilson et al. 1983; Gastony 1991; Soltis and Soltis 1993; Gottlieb and Ford 1997) and animals (Ferris and Whitt 1979; Force et al. 1999).

The level of species incompatibility predicted by the model of Werth and Windham (1991) is likely to be an underestimate for the following reason: Under the classical model of gene duplication, a gene is viewed as having either a single function or several functions that are not subject to independent evolution. However, it is now known that many eukaryotic genes have multiple, independently mutable subfunctions that govern the timing and spatial location of gene expression (Jack and DeLotto 1995; Kirchhamer et al. 1996; Arnone and Davidson 1997; Gerhart and Kirschner 1997; Yuh et al. 1998). With this type of modular gene structure, degenerative mutations can lead to duplicate-gene preservation rather than gene loss by a process that we refer to as "duplication, degeneration, and complementation" (DDC; Force et al. 1999). If one copy becomes fixed for a mutation that eliminates one subfunction, the second copy will be permanently preserved, but if the second copy subsequently becomes fixed for a mutation that eliminates a different subfunction, then the first copy will be reciprocally preserved. As a consequence of this process, the two members of a duplicate pair of genes can eventually partition the original expression pattern of the ancestral gene. This phenomenon is consistent with a substantial amount of expression-pattern data on duplicate genes (Ferris and Whitt 1979; Wes-

tin and Lardelli 1997; Normes et al. 1998; Force et al. 1999; Lynch and Force 2000).

Under this more general model for the fate of duplicate genes, species incompatibilities can arise even in the absence of gene loss. Unless the subdivision of expression patterns by duplicate genes follows an identical pattern in two sister taxa, an incompatibility will exist in their F_2 descendants because some two-locus genotypes will be null for one or more subfunctions. The probability that gene duplicates become preserved by subfunctionalization is a function of the number of independently mutable subfunctions (z), the rate at which degenerative mutations arise for each subfunction (μ_r), and the rate at which mutations completely silence a locus (μ_c). As the total rate of subfunctionalizing mutations ($z\mu_r$) becomes large relative to μ_c, the probability of duplicate-gene preservation by subfunctionalization (P_s) approaches 1, while the probability of nonfunctionalization by gene silencing ($P_n = 1 - P_s$) converges on 0. We now provide a more formal description of the probability of a species incompatibility arising from the divergent resolution of a gene duplicate.

In the absence of gene flow, the fates of duplicate-gene pairs will be resolved independently in two isolated populations. Both populations will experience subfunctionalization with probability P_s^2, one will experience subfunctionalization; and the other, nonfunctionalization with probability $2P_sP_n$, and both will experience nonfunctionalization with probability P_n^2. As noted above, in the latter case, the probability of divergent resolution is one-half, and given its occurrence, the probability that a gamete produced by an F_1 hybrid will be lacking in function for a gamete-specific gene is one-fourth, while the probability that an F_2 offspring will be lacking in function for a zygote-specific gene is one-sixteenth. In the second case, one species experiencing subfunctionalization and the other nonfunctionalization, one-fourth of the F_1 gametes and one-sixteenth of the F_2 offspring will again be lacking in function. When both species experience subfunctionalization, assuming z unresolved subfunctions at the time of geographic isolation, there is a $[1 - (1/2)^z]$ probability that different patterns of subfunctionalization will have arisen, and this again leads to one-fourth of the F_1 gametes and one-sixteenth of the F_2 zygotes being nonfunctional (fig. 1). Summing up, the expected incidence of nonfunctional F_1 gametes resulting from divergent resolution of a pair of duplicate genes in two sister taxa is

$$I_g = \frac{1}{4}\left\{P_s^2\left[1 - \left(\frac{1}{2}\right)^z\right] + 2P_sP_n + \frac{P_n^2}{2}\right\}, \tag{1a}$$

whereas that for F_2 zygotes is

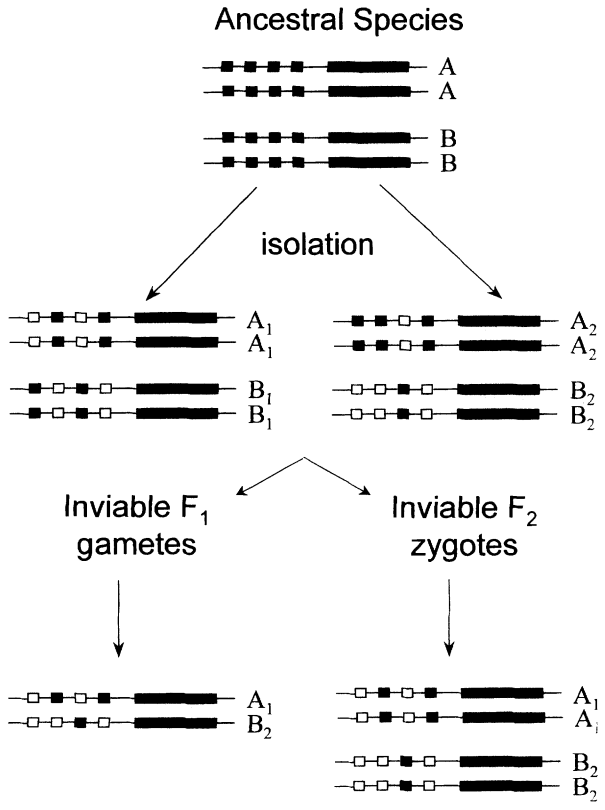

Figure 1: An example of divergent resolution of an ancestral pair of duplicate genes in two sister taxa (1 and 2). The four small boxes to the left denote independently mutable subfunctions (e.g., four regulatory regions), whereas the long boxes to the right denote coding regions. A black box denotes a functional gene component, whereas an open box denotes nonfunctionality. In this example, the last three subfunctions are resolved in the same way in the sister species. However, divergent resolution occurs for the first subfunction—species 1 loses function at the **A** locus and retains it at the **B** locus, while the opposite occurs in species 2. As a consequence of this divergence, one-fourth of the F_1 gametes and one-sixteenth of the F_2 zygotes are lacking in functional genes.

$$I_z = \frac{I_g}{4}. \tag{1b}$$

The remaining issue is the relative level of P_s and P_n. Under the assumption of z equally mutable subfunctions, each of which needs to be present in only a single copy to endow a zygote with perfect fitness, an analytical approximation for the probability of subfunctionalization is

$$P_s = \sum_{i=2}^{z}\left[\frac{z}{(\mu_c/\mu_r) + z}\right]\prod_{j=0}^{i-2}\left[\frac{(z-j-1)}{(\mu_c/\mu_r) + 2(z-j-1)}\right] \tag{2}$$

(Force et al. 1999). Derivation of this expression relies on

the assumption that fixations of mutations at the different loci occur as essentially nonoverlapping events, with double null recessive genotypes being sufficiently rare that selection can be ignored as a factor influencing fixation events. Computer simulations show equation (2) to be very accurate for $N\mu_c < 0.01$, where N is the effective population size (Lynch and Force 2000).

The solution to equations (1a), (1b), and (2) shows that the incidence of incompatibilities increases with the number of independently mutable subfunctions per gene (fig. 2). For F_1 gametes, I_g increases from one-eighth under the classical model to one-fourth under the DDC model with high z; and for F_2 zygotes, I_z increases from one-thirty-second to one-sixteenth. Assuming that complete absence of a subfunction is lethal, with n loci, the expected viability of F_1 gametes is $(1 - I_g)^n$, whereas that for F_2 zygotes is $(1 - I_z)^n$. Thus, since most eukaryotic organisms contain on the order of 10^4–10^5 genes, it is clear that only a very small fraction of the genome must be present as unresolved duplicates in an ancestral species to provide a powerful mechanism for the passive origin of reproductive isolation. For example, under the preceding (double-recessive) model, as few as eight unresolved pairs of duplicate genes in an ancestral species can be sufficient to reduce F_1 gamete function to <10%, whereas as few as three dozen are sufficient to drive F_2 zygote viability and/or sterility down to this level. As noted above, these conditions are likely to be satisfied in most species.

If more than a single operable copy is required for some subfunctions in zygotes (haploinsufficiency), then the expected degree of incompatibility revealed in the F_2 generation will be greater than that suggested above. Haploinsufficiency is a common phenomenon in mammals, as illustrated by many examples of dominant loss-of-function mutations involved in genetic disorders in mice and humans (e.g., Cheng et al. 1997; Oda et al. 1997; Fero et al. 1998; Lamande et al. 1998; Wilm et al. 1998). There is no reason to expect the situation to be different in other species, and it is conceivable that haploinsufficiency would be exacerbated in a hybrid background. In a survey of the behavior of 114 small chromosomal deficiencies from *D. melanogaster* on a *melanogaster* × *simulans* hybrid background, Coyne et al. (1998) found nine that caused either lethality or severe reductions in viability.

The simplest (double-recessive) model presented above is only able to explain F_1 sterility for cases in which haploid gene expression is essential to gamete function. The latter condition is fulfilled in plants, for which gametophytic performance is critical (Walsh and Charlesworth 1992; Xu et al. 1999), as well as in algae, fungi, and other organisms with a conspicuous and transcriptionally active haplophase. However, the situation is less clear in metazoans

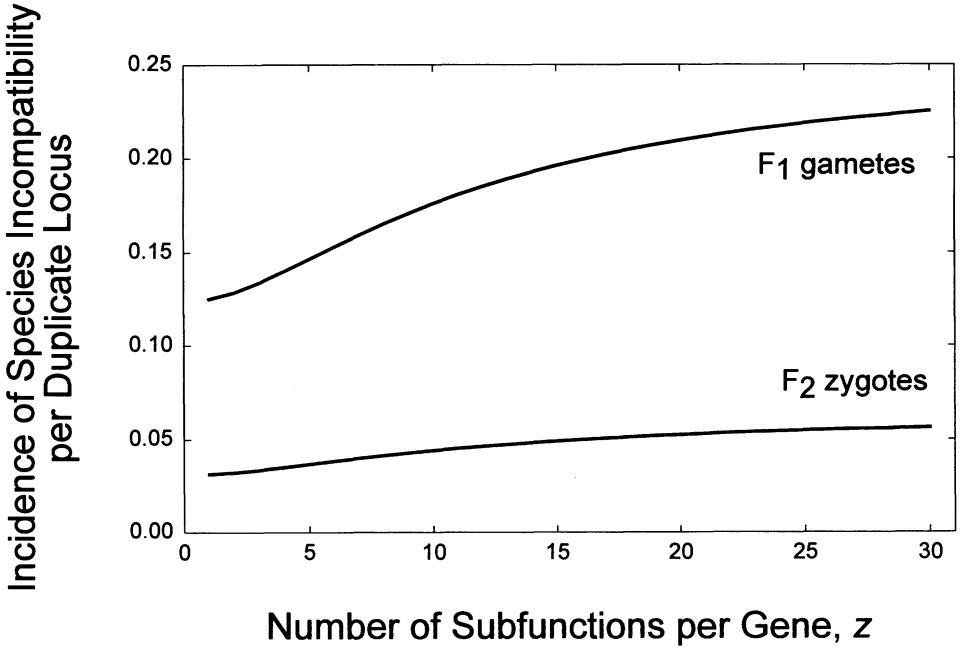

Figure 2: The relationship between the ultimate incidence of interspecific incompatibility resulting from divergent resolution of an ancestral gene duplication and the number of independently mutable subfunctions per gene. Results are given for the situation in which the degenerative mutation rate for complete gene function is 10 times that per subfunction ($\mu_c/\mu_r = 10$).

with a reduced haploid stage. For example, in *Drosophila*, there is no known postmeiotic transcription in sperm (Gould-Somero and Holland 1974), and sperm lacking in one or more chromosomes can be fully functional (Muller and Settles 1927; Lindsley and Lifschytz 1972). On the other hand, early postmeiotic mammalian sperm are transcriptionally active (Eddy et al. 1993; Yiu and Hecht 1997), and circumstantial evidence suggests some haploid gene expression in the sperm of *C. elegans* (Argon and Ward 1980), although the functional significance of such gene expression is poorly understood in both groups. Recent work has shown that the paternal genome is silent during early embryonic development in plants (Vielle-Calzada et al. 2000), while much of the maternally derived genome may be transcriptionally inactive during early embryogenesis in mammals (Mayer et al. 2000). Thus, the quantitative predictions for I_g may in many cases describe F_2 inviability more accurately than do those for I_z.

Although F_1 sterility or inviability is a common feature of wide interspecific crosses, small introgressions often have indiscernible heterozygous effects while being lethal or sterilizing in the homozygous state (reviewed in Turelli and Orr 2000). Such observations are qualitatively consistent with the predictions of the model proposed above, although complete compatibility with the theory will require the demonstration that such introgressions are a consequence of null homozygotes induced by microchromosomal map changes resulting from divergent resolution of duplicate genes.

Gene Duplication Subsequent to Prezygotic Isolation

The results in the previous section pertain to gene duplicates that are fixed in an ancestral population, such that all members of the incipient sister taxa initially contain two functional loci. However, gene duplication events that arise subsequent to the isolation of two sister species can also passively promote the origin of postzygotic isolating barriers. Such events may be quite common, as evidenced by the dozens to hundreds of young, unlinked gene duplicates that are found in the genomes of most eukaryotic species (Lynch and Conery 2000). Consider an ancestral locus fixed for the functional **A** allele, with a subsequent duplication event creating a functionally redundant locus (with allele **B**) at an unlinked (or weakly linked) site in one of two descendent daughter taxa. Should the taxon containing the two loci retain **B** while losing some or all of the function of **A**, then a genomic incompatibility will arise in hybrids with an **AaB.** genotype, since **aa.**. derivatives would be completely lacking in gene function.

In order for this condition to be met, the original gene duplication (present in a single individual) must first drift to fixation, and then either **A** must be nonfunctionalized or **A** and **B** must be jointly subfunctionalized. Assuming there is no intrinsic advantage to having two functionally redundant copies of the locus, the probability of the first event is simply equal to the initial frequency of the functional **B** allele, $1/2N$, where N is the effective population size. The subsequent probability that the **A** allele will be lost by nonfunctionalization is then $P_n/2$, where $P_n = 1 - P_s$ is the probability of nonfunctionalization and the one-half accounts for the fact that there is an equal probability that either of the loci will be silenced. The probability that both loci will be retained by subfunctionalization is again P_s. Assuming a rate of origin of gene duplicates per haploid genome in each lineage equal to U_D and an adult population size of N in each lineage, then the asymptotic rate of origin of loci involved in postzygotic isolation is

$$R = (4NU_D)\frac{1}{2N}\left(\frac{1 - P_s}{2} + P_s\right)$$
$$= U_D(1 + P_s). \quad (3)$$

Because the processes of fixation and silencing are dependent on random genetic drift and mutation, it is expected to take approximately $1/2\mu_c$ generations to reach this asymptotic rate when $N\mu_c < 0.1$ and on the order of $10N$ generations when $N\mu_c > 1$ (Watterson 1983; Lynch and Force 2000). Once this point has been reached, depending on the value of P_s, the rate of origin of loci involved in postzygotic isolation is then between U_D and $2U_D$. This rate is not entirely independent of population size because although P_s is independent of N and defined by equation (2) for $N\mu_c < 0.01$, P_s declines with larger N, until $P_s \simeq 0$ when $N\mu_c > 1$ (Lynch and Force 2000).

Letting T denote the number of generations of divergence between two lineages, then the expected viability of F_2 hybrids resulting from gene duplicates originating during this period is approximately $(15/16)^{U_D(1+P_s)T}$ under the double null recessive model. Based on estimates of the age distributions of gene duplicates in the completely characterized genomes of the nematode *Caenorhabditis elegans* and the yeast *Saccharomyces cerevisiae*, we have estimated the rate of duplication events to be on the order of 0.002–0.02/gene/million years (m.yr.; Lynch and Conery 2000). Thus, assuming a moderate genome size of 20,000 loci, U_D is on the order of 40–400/m.yr. When we are conservative and let $P_s = 0$ and $U_D = 20$/m.yr., the results indicate that 1 m.yr. of divergence would result in F_2 viability of approximately 28%. When we are liberal and let $P_s = 1$ and $U_D = 200$/m.yr., the expected F_2 viability after only 100,000 yr would be approximately 8% and, after 1 m.yr., essentially 0%.

The Consequences of Chromosomal Location

In the preceding discussion, we assumed a pair of gene duplicates present on independently segregating autosomes, with the fitness consequences of divergent resolution being the same for both sexes. However, given the common observation that interspecific hybrids of the heterogametic sex (assumed to be males below) in animals are much more inviable and/or infertile than those of the homogametic sex (Haldane 1922; Coyne and Orr 1989; Coyne and Orr 1997; Laurie 1997; Orr 1997), it is worth considering the extent to which divergent resolution of duplicate genes can generate sex-specific incompatibilities.

In principle, the model presented above can directly accommodate sex-specific differences in F_2 performance if autosomal gene duplicates commonly have sex-specific expression. To explain Haldane's rule, autosomal genes influencing hybrid male performance would have to be much more common or more mutable than female-specific genes. This condition appears to be met in *Drosophila*, where large numbers of autosomal genes are known to be specifically essential for male fertility (about twice as many as for female fertility; Lindsley and Tokuyasu 1980; Castrillon et al. 1993). Recent studies also support the idea that the total rate of accumulation for hybrid male sterility genes is substantially greater than that for hybrid female sterility in both *Drosophila* (Hollocher and Wu 1996; True et al. 1996; Sawamura et al. 2000) and mosquitoes (Presgraves and Orr 1998), and male-sterility genes appear to have unusually high mutation rates (Zhang and Stankiewicz 1998).

We now consider the consequences of one or both members of a pair of duplicate genes residing on a sex chromosome. Although the idea that the X chromosome is disproportionately involved in the origin of isolating barriers in *Drosophila* (Coyne and Orr 1989) has not held up to close scrutiny (Hollocher and Wu 1996; True et al. 1996), it remains clear that genes on the X play a prominent role. In lepidopterans, large numbers of genes involved in reproductive isolation have been located on the X and Y (Prowell 1998). In addition to their potential bearing on the issue of sex-specific incompatibilities, the following scenarios provide insight into how divergent resolution of gene duplicates might contribute to incompatibilities observed in the F_1 generation, as well as a possible explanation for asymmetric incompatibilities that are commonly seen in reciprocal F_1 progeny of very closely related species.

Genes on Both Sex Chromosomes

In the absence of recombination between the X and Y chromosomes, X- and Y-linked alleles can potentially evolve independently of each other, thereby leading to the same type of evolutionary partitioning of subfunctions that arises with duplicate autosomal genes—each copy may lose one or more subfunctions, with the other copy permanently retaining them. In this case, subfunctionalization (or complete gene loss) can only occur for genes with expression specific to the heterogametic zygote, since any X-linked genes that are non–sex specific or specific to the homogametic sex must retain complete functionality.

If male-specific subfunctions of an X/Y gene divergently degenerate in two sister taxa, then the consequences of a reciprocal cross are expected to be asymmetrical. The F_1 male offspring of fathers from the species lacking in a critical Y-linked subfunction are 100% inviable (or sterile) when the mother lacks the same subfunction on the X, but male progeny of the reciprocal cross have fully competent X- and Y-linked copies (fig. 3). Such locus-specific asymmetries are expected to accumulate randomly in two taxa over evolutionary time (with each species losing subfunctions on the X for some genes and on the Y for others), so that pronounced asymmetry in the performance of F_1 males is expected only in the early stages of reproductive isolation (when few genes are involved).

Recalling that divergent resolution occurs at levels of 50%–100%, depending on the extent to which duplicate genes can become subfunctionalized, these results suggest that X/Y genes with male-specific subfunctions provide a potentially powerful substrate for the origin of postreproductive isolation—each such gene in an ancestral species will ultimately yield an incidence of incompatibility in F_1 male progeny of 0.25–0.50, so just six such genes could reduce F_1 male viability (or fertility) to between 18% and 1.5% of normal level. Because it extends to taxa in which females are the heterogametic sex (birds and lepidopterans), divergent resolution of genes on sex chromosomes provides a potentially more general explanation for Haldane's rule than does an explanation based on more frequent or more mutable male-specific genes, which runs into problems with species in which females are heterogametic (Wu and Palopoli 1994).

What is the evidence for the divergence of male-specific genes on the X and Y chromosomes? In most *Drosophila* species, there is no compelling evidence that the current X and Y chromosomes share a simple common ancestry (Hackstein et al. 1996). Rather, duplication and recruitment of genes between the autosomes and sex chromosomes and between the X and Y appear to be ongoing processes. This implies that the creation of potential substrate for the divergent resolution of sex chromosomes does not end with the initial establishment of the X and Y. The Y chromosome is often not necessary for normal somatic development in *Drosophila*, but in 17 of 22 species tested, it is essential for male fertility (Laurie 1997). In addition, a role for the Y chromosome has been implicated in a number of cases of hybrid male sterility (Coyne and

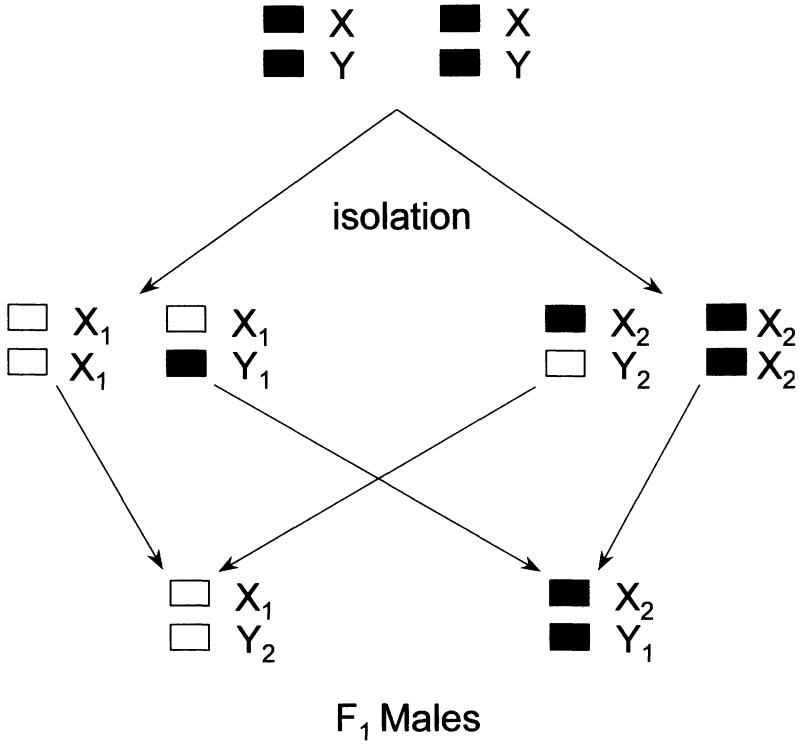

Figure 3: The consequences of a divergently resolved male-specific gene distributed on the X and Y chromosomes. Solid squares denote a functional gene or gene subfunction, open squares a nonfunctional gene. In species 1, loss of (sub)function has occurred on the X and is preserved on the Y, while the opposite has occurred in species 2. Divergent resolution results in an asymmetry in F_1 male inviability and/or sterility.

Orr 1989; Johnson et al. 1992; Lamnissou et al. 1996; review in Turelli and Orr 2000).

Two fairly well-studied genes, "suppressor-of-stellate" (Livak 1984; Kalmykova et al. 1997) and the rRNA array (Lohe and Roberts 1990), are known to have different distributions on the X and Y chromosomes of the closely related species *Drosophila melanogaster* and *Drosophila simulans.* In addition, *Drosophila miranda* carries a pair of secondary sex chromosomes, the X2 and the neo-Y, which have captured an entire linkage group of normally autosomal genes, some of which appear to have acquired substantial structural rearrangements relative to their ancestral states (Steinemann et al. 1996; Steinemann and Steinemann 1997). A potentially revealing experiment would be to compare the performance of a single X, a single Y, and both the X and the Y of one species on a pure autosomal background of another. If reproductive isolation is due to divergent resolution of X/Y genes, the latter construct should yield viable and fertile males because it contains a complete set of X/Y genes.

Gene Duplicates on an Autosome and a Sex Chromosome

Duplicate genes that are present on an autosome and on the Y chromosome may be common if male-specific autosomal genes are recruited to the Y via the selective mechanism envisioned by Fisher (1931). The evidence for such recruitment is most pronounced in mammals, where a small region at the tip of the Y is homologous to a portion of the X chromosome (the pseudoautosomal region) and undergoes recombination with it. In her addition-attrition hypothesis, Graves (1995) raised the idea that the Y chromosome of mammals represents an ongoing historical development of additions, expansions, and degenerations of autosomal sequences. Well-documented examples of human autosomal genes that have been duplicated on to the Y chromosome are *DAZ* (Saxena et al. 1996; Delbridge et al. 1997; Elliot et al. 1997; Ruggio et al. 1997), *RBM1* (Delbridge et al. 1997; Elliot et al. 1997), and *TSPY* (Manz et al. 1991; Jacubiczka et al. 1993; Delbridge et al. 1997).

If divergent resolution of a gene with male-specific subfunctions results in the autosomal copy retaining function in one daughter species and the Y copy retaining function in the other, then all F_1 male progeny will have one functional autosomal copy and half will have a nonfunctional Y-linked gene (fig. 4). In the F_2 generation, one-fourth of the male progeny of males lacking a functional Y-linked gene will be lacking in function, and if multiple genes are involved, a much higher fraction of hybrid males will lack

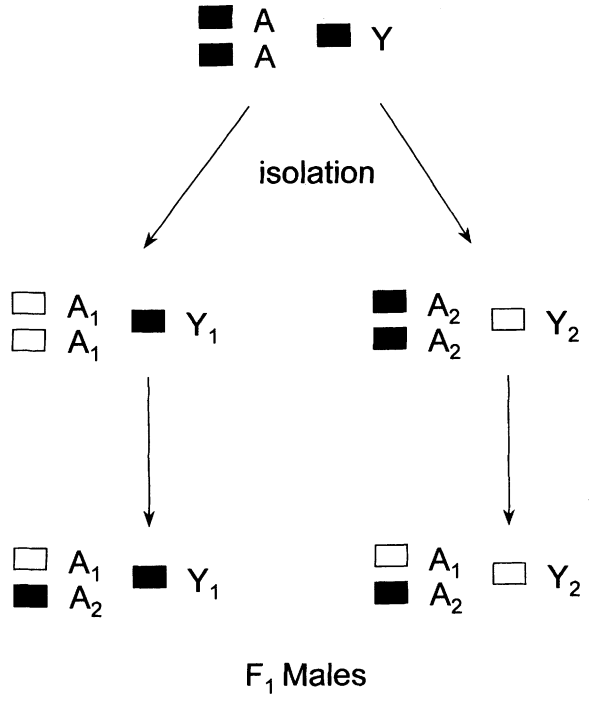

Figure 4: The consequences of a divergently resolved, male-specific gene function distributed on an autosome and on the Y chromosome. Solid squares denote a functional gene; open squares, a nonfunctional gene. Divergent resolution results in an asymmetry in the properties of F_1 males: if the gene influences zygote viability, F_2 male progeny of males on the right will be 25% inviable.

function at least at one locus. Some situations may be envisioned in this case in which F_1 males are nonfunctional. For example, if the autosomal locus is subject to gametic imprinting, a common phenomenon in mammals and plants (Hall 1990; Barlow 1995; Jegalian and Page 1998), the viability of F_1 males carrying the loss-of-subfunction allele on the Y chromosome will depend on which parental gamete transmits suppressed allelic expression. If suppression occurs in paternally derived genes, then F_1 males lacking in Y subfunction are expected to be inviable/infertile, whereas those resulting from the reciprocal cross should be viable/fertile.

The movement of genes between autosomes and sex chromosomes is by no means unidirectional. For example, cases of transfer from the X chromosome to the autosomes have been identified within the great ape lineage (Eichler et al. 1996, 1997) as well as within *Drosophila* (Yuan et al. 1996). Should gene duplicates distributed on an autosome and the X become divergently resolved in two daughter taxa, the consequences for hybrid function will again depend on the mode of regulation of gene expression. For example, for a divergently resolved autosomal/X gene duplicate with male-specific functions, F_1 males will have zero, one, or two expressed copies depending on the mode of gene expression. In the absence of gametic imprinting, the two reciprocal crosses will yield male progeny with one versus two active copies (fig. 5). However, if maternal autosomal gene silencing were to occur, then hybrid male progeny of mothers from the species lacking an active X-linked gene will be completely lacking in function.

Discussion

Although a universal mechanism for the origin of interspecific genomic incompatibility is highly unlikely, certain properties (unrelated to specific gene functions) may endow some genes with features that promote their involvement in the evolution of postmating reproductive isolation. Three general lines of evidence suggest that gene duplicates may be particularly common contributors to this process: first, the large numbers of duplicate genes and the redundant genetic pathways that exist within all genomes and their relatively rapid rate of origin; second, the direct observation of partitioning and/or alternative silencing of gene duplicates; and finally, the frequent trans-

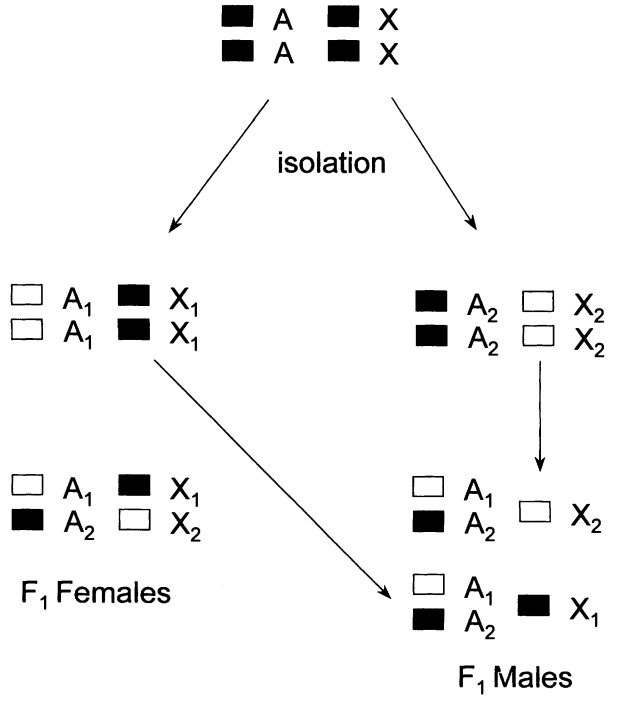

Figure 5: The consequences of a divergently resolved gene distributed on an autosome and on the X chromosome. Solid squares denote a functional gene; open squares denote a nonfunctional gene. Divergent resolution results in an asymmetry in the number of null alleles carried by F_1 males.

port of small genomic segments among nonhomologous chromosomes. Thus, the mechanism that we propose for the origin of interspecific genomic incompatibility is likely to be a common one, although we are by no means suggesting that it is the explanation for speciation in all organisms. As noted above, the divergent evolution of gene duplicates may be a more common mechanism for genomic incompatibility in plants and fungi than in animals, but there are many plausible situations in which it may play a role in animal speciation.

The model that we have presented for the origin of postzygotic isolating mechanisms is completely compatible with the classical BDM model, in that none of the postulated genetic changes within taxa have negative fitness consequences until they are exposed as negative epistatic effects on a hybrid background. The usual explanation for such behavior is that mutations that have gone to fixation within one species have never been tested on the genetic background of other species. This leaves to the imagination the nature of the genetic changes that are driven by neutral or adaptive mutations within populations but that suddenly give rise to a dramatic loss of fitness when combined in hybrid backgrounds. The gene-duplication model provides an explicit and biologically plausible mechanism for such behavior—complementing pairs of gene duplicates (or, more generally, genomic redundancies) are essential and fully functional in their own species backgrounds, but single members of the pair act as loss-of-function alleles when introgressed into a species with divergent resolution at the same loci.

There are three additional attractive features of a model for the origin of genomic incompatibilities based on the divergent resolution of gene duplicates. First, the genetic changes postulated to occur within populations are entirely a product of degenerative mutations, which we know to be much more frequent than beneficial mutations. As a consequence of genomic redundancy, such mutations can often accumulate in an effectively neutral manner (Li 1980; Force et al. 1999; Lynch and Force 2000). Second, there is no need to invoke a change in gene action in a hybrid genetic background. Loss of function arises as a simple by-product of Mendelian segregation and independent assortment, and coadaptive gene complexes within species are nothing more than mutually complementing sets of genes. Third, depending on the frequency of null alleles segregating within duplicate loci in an ancestral species, the time to complete silencing in a daughter taxon will be on the order of the reciprocal of the null mutation rate or less (Watterson 1983; Lynch and Force 2000). Thus, given that the null mutation rate per locus is on the order of 10^{-6} to 10^{-5} per generation, substantial genomic incompatibilities are expected to arise passively by the time sister species have been isolated for 10^5 to 10^6 generations. Within *Drosophila*, the only genus for which the data are substantial, complete postzygotic isolation is generally complete within 0.2–2.0 m.yr. (Coyne and Orr 1997), and qualitatively similar conclusions have been drawn for fishes (Parker et al. 1985; McCune and Lovejoy 1998).

Two aspects of genomic structure may be particularly conducive to the origin of species incompatibilities via gene duplication. First, as noted above, since the genes on nonrecombining portions of sex chromosomes are effectively isolated from each other, they are free to evolve independently in the same fashion as duplicate genes on different autosomes. Thus, by simultaneously creating a large pool of independently evolving genetic elements, the origin of sex chromosomes (or the replicative translocation of a large block of autosomal genes to the X and/or Y) can provide a substantial amount of substrate for the divergence of two isolated populations. Only those X/Y genes with effects specific to the heterogametic sex would be subject to divergent resolution, making them candidates for involvement in Haldane's rule. Second, the DDC process is expected to be a major genetic mechanism of speciation in taxa that have experienced a polyploidization event, as this will subject thousands of loci to the divergent resolution process once the level of chromosomal divergence has reached the point at which the genome behaves in meiosis as a diploid.

When large numbers of loci are subject to the DDC process, the stochastic divergence of locus-specific subfunctions over long periods of time could provide the fuel for many nested reproductive isolation events. In other words, after an ancestral polyploid species (or one that has recently evolved sex chromosomes) has diverged into two lineages by the DDC process, an enormous number of gene duplicates will generally remain to be resolved. Such residual redundancies would be expected to contribute to subsequent isolation events, leading to a clustering of speciation episodes around periods of polyploidization or sex chromosome evolution.

As a possible example of such an event, we note the genome modification that appears to have occurred in the common ancestor of the major vertebrate lineages. Relative to other deuterostome lineages, the vertebrates are quite species rich, and the ray-finned fishes are by far the most speciose of the vertebrate groups. It has often been suggested that two complete genome duplications directly fostered the diversification of the vertebrates by providing new substrate for adaptive divergence (Ohno 1970; Lundin 1993; Holland et al. 1994; Sidow 1996). Although the evidence for such polyploidization events in the basal vertebrate remains statistically equivocal (Skrabanek and Wolfe 1998; Hughes 1999; Martin 1999), there is no question that a massive amount of gene duplication occurred, and the evidence for subsequent genomic duplication

events in ray-finned fishes is substantial (Allendorf et al. 1975; Ferris and Whitt 1979; Amores et al. 1998; Postlethwait et al. 1998). Direct evidence for the divergence of duplicate-gene expression patterns among species exists for catastomids, pufferfish, and zebrafish (Ferris and Whitt 1979; Aparicio et al. 1997; Amores et al. 1998), and Ferris et al. (1979) found a statistical association between the silencing of duplicate-gene expression and the rate of speciation in polyploid fish. Thus, an alternative view of the radiation of the vertebrates, and ray-finned fishes in particular, is that large numbers of gene duplicates initially provided an enhanced opportunity for the passive origin of reproductive incompatibilities by the divergent resolution process, which was subsequently followed by morphological divergence (which may or may not have required duplicate genes).

Wilson and colleagues (Wilson et al. 1974, 1975; Levin and Wilson 1976; Bush et al. 1977) have argued that phyletic lineages with higher rates of chromosomal evolution also exhibit higher rates of morphological diversification and higher rates of origin of postzygotic reproductive isolation. Although a specific mechanism was not presented, they promoted the idea of a causal link between all three types of change—genomic rearrangements precipitate changes in patterns of gene regulation, which in turn foster the origin of reproductive isolation. The types of genomic reorganization emphasized by Wilson and colleagues were large-scale events, such as chromosomal fusions, translocations, and inversions (White 1978). The requirements for fixing such rearrangements are thought to be stringent because of the strong underdominance for fitness expected in chromosomal heterozygotes (Walsh 1982; Lande 1984), although some empirical evidence suggests otherwise (Coyne et al. 1993).

On the other hand, genomic reorganization via the process envisioned above promotes the origin of strong reproductive isolation by chromosomal repatterning without either sister taxon going through a bottleneck in fitness. The divergent resolution of genomic redundancies in effect induces small-scale changes in the genetic maps of sister taxa, thereby erecting a natural isolating barrier. Although the BDM–incompatibility factor model and the chromosomal model are often viewed as competing explanations for the origin of postzygotic isolation, the gene-duplication model that we present is consistent with both explanations—chromosomal repatterning resulting from divergent resolution of gene duplicates induces small-scale negative epistatic interactions between locations of ancestral and descendent sites of genes. Studies such as Fischer et al. (2000) that reject the chromosomal model of speciation on the basis of an observed absence of large chromosomal translocations ignore the more likely possibility that closely related species harbor many small-scale map changes.

In the field of animal speciation, the litmus test for a fully general model of genomic incompatibility is often viewed as the ability to explain the fact that the first step in postzygotic isolation generally involves the heterogametic sex (Haldane's rule) and that such problems quickly advance from the F_2 to the F_1 generation. The extent to which explanations for Haldane's rule provide general explanations for the genetic mechanisms of speciation is unclear, since the vast majority of eukaryotic species do not have sex chromosomes. Nevertheless, the fact that scenarios involving gene duplications with at least one resident on a sex chromosome yield predictions that are compatible with Haldane's rule suggests a possible mechanistic connection with gene duplication. Because the autosomes and/or the X chromosome must jointly contain the full set of genes essential for female function (assuming females to be the homogametic sex), F_1 females necessarily carry a complete set of functional genes. On the other hand, male-specific genes can be distributed over the autosomes, the X, and the Y chromosomes. Since male offspring acquire only two of these three sets of genes from each parent, there are numerous ways of obtaining noncomplementing sets of male-specific genes in F_1 hybrids, especially when issues of haploinsufficiency, gametic imprinting, and dosage compensation are taken into consideration. The divergent resolution of X/Y genes provides a particularly simple mechanism for generating viability and/or fertility problems in F_1 males, although it remains to be seen whether such events are common. Our model is entirely compatible with the dominance theory for Haldane's rule, which posits that reduced hybrid viability and/or fertility in the heterogametic sex is a consequence of X-linked mutations that act as deleterious recessives on a hybrid background (Muller 1942; Orr 1993; Turelli and Orr 1995).

The models outlined above are overly simplistic in many respects, but in ways that are likely to underestimate the role of gene duplication in the origin of species isolating barriers. First, significant loss of gene expression was assumed to occur only in double loss-of-function homozygotes, whereas haploinsufficiency may often result in inviability/sterility of hybrid progeny carrying a single active gene copy. In addition, loss of gene (sub)function was assumed to result entirely from degenerative mutations incurred by regulatory and/or coding regions, while the possible importance of trans-acting regulatory elements was ignored. A poorly understood but frequent consequence of interspecific hybridization is allelic suppression (the absence of expression of one or both parental alleles despite their full expression in the parental species; Whitt et al. 1973; Parker et al. 1985; Sampsell and Held 1985; Houchins et al. 1997). The occurrence of allelic suppression further broadens the generality of the gene-duplication model of

reproductive isolation because it provides still another mechanism by which complete absence of gene function may occur in F_1 hybrids.

Second, although we have couched the gene-duplication model for the origin of genomic incompatibility in terms of the divergent resolution of duplicate genes, the fundamental mechanisms that we envision are not strictly limited to orthologous gene duplicates. Divergent resolution can arise with any type of genetic redundancy, including that involving nonorthologous gene families that carry out similar functions. Although the mechanisms responsible for their origin and maintenance are poorly understood, it is well established that such redundant genetic pathways are very common in multicellular organisms (Brookfield 1997*a*, 1997*b*; Cooke et al. 1997), and the phenomenon of co-option (Gerhart and Kirschner 1997) suggests a potential role in the origin of postzygotic reproductive barriers.

Third, although our consideration of the divergent resolution of duplicate genes focused entirely on degenerative mutations, it should be noted that origins of new functions by one member of a duplicate pair can create the same level of incompatibility as loss-of-function mutations, provided the copy experiencing neofunctionalization does so at the expense of the original function. Although neofunctionalization is expected to be a very rare fate of gene duplicates in small populations (Walsh 1995), it becomes increasingly common with increasing *N*. Thus, the gene-duplication model for the origin of postzygotic isolation operates in large as well as small populations, although the mechanism of divergent resolution is expected to differ. We note with interest that the one gene that has so far been associated with reproductive incompatibility in interspecies crosses of *Drosophila*, the *Ods* locus, appears to be a duplicate gene that has experienced strong positive selection (Ting et al. 1998).

How can we test whether gene duplicates actually provide a common basis for the origin of interspecific incompatibilities? Although detailed mapping and molecular studies may eventually identify the specific sets of genes underlying an interspecies incompatibility (Ting et al. 1998; Sawamura et al. 2000), the necessary work is enormous. For situations in which pairs of gene duplicates in closely related taxa are known, an alternative approach is to directly study divergent resolution by comparing expression patterns of both gene copies in both sister taxa as well as in an informative out-group containing only single-copy genes (Force et al. 1999). Perhaps the simplest approach to testing the gene-duplication model is the mapping of incompatibility genes in the descendants of polyploid species that are functionally diploid. Assuming substantial genomic rearrangements have not occurred since the polyploidization event, epistatically interacting pairs of sterility/inviability genes should commonly map to the same locations on homeologous chromosomes if divergent resolution of gene duplicates is a common mechanism for the origin of genomic incompatibilities. Given the large amount of substrate for divergent resolution in polyploid species, failure to observe such complementary mapping would weaken the gene-duplication theory for speciation.

One of the best studied examples of hybrid sterility/inviability in plants, that involving the *indica* and *japonica* cultivars of rice (*Oryza sativa*), provides evidence that is at least qualitatively consistent with the origin of postzygotic reproductive isolation by the mechanisms suggested above. No chromosomal rearrangements have been detected between the two lineages (Li et al. 1997), although cryptic structural rearrangements cannot be ruled out. However, crossing experiments with isogenic lines suggest the presence of duplicate gametic lethals that have been divergently resolved in the two cultivars (Oka 1974). One-fourth of the gametes produced by the F_1 progeny of the two cultivars are nonfunctional, as expected with the independent assortment of a divergently resolved pair of loci. The exact nature of the genes underlying the genomic incompatibilities in rice is unknown, but Oka (1974, 1988) has speculated that a large number of complementary sterility genes have arisen in the rice genome as a consequence of ancestral chromosomal doublings.

Acknowledgments

This research evolved out of a collaborative effort fostered in part by a National Science Foundation Training Grant in Genetic Mechanisms of Evolution and a National Institutes of Health Training Grant in Developmental Biology, both of which supported A.F., and by a National Institutes of Health grant to M.L. (RO1-GM36827). We thank T. Birkhead, J. Graves, D. Howard, A. Orr, D. Presgraves, L. Rieseberg, S. Ward, C. Wedekind, and C. Werth for discussion that helped stimulate and refine our ideas.

Literature Cited

Allendorf, F. W., F. M. Utter, and B. P. May. 1975. Gene duplication within the family *Salmonidae*. II. Detection and determination of the genetic control of duplicate loci through inheritance studies and the examination of populations. Pages 415–432 *in* C. L. Markert, ed. Isozymes. Vol. 4. Genetics and Evolution. Academic Press, New York.

Amores, A., A. Force, Y. L. Yan, L. Joly, C. Amemiya, A. Fritz, R. K. Ho, et al. 1998. Zebrafish *hox* clusters and vertebrate genome evolution. Science (Washington, D.C.) 282:1711–1714.

Aparicio, S., K. Hawker, A. Cottage, Y. Mikawa, L. Zuo,

B. Venkatesh, E. Chen, R. Krumlauf, and S. Brenner. 1997. Organization of the *Fugu rubripes Hox* clusters: evidence for continuing evolution vertebrate *Hox* complexes. Nature Genetics 16:79–83.

Argon, Y., and S. Ward. 1980. *Caenorhabditis elegans* fertilization-defective mutants with abnormal sperm. Genetics 96:413–439.

Arnone, M. I., and E. H. Davidson. 1997. The hardwiring of development: organization and function of genomic regulatory systems. Development 124:1851–1864.

Barlow, D. P. 1995. Gametic imprinting in mammals. Science (Washington, D.C.) 270:1610–1613.

Brookfield, J. F. Y. 1997*a*. Genetic redundancy. Advances in Genetics 36:137–155.

———. 1997*b*. Genetic redundancy: screening for selection in yeast. Current Biology 7:R366–R368.

Bush, G. L., S. M. Case, A. C. Wilson, and J. L. Patton. 1977. Rapid speciation and chromosomal evolution in mammals. Proceedings of the National Academy of Sciences of the USA 74:3942–3946.

Castrillon, D. H., P. Gonczy, S. Alexander, R. Rawson, C. G. Eberhart, S. Viswanathan, S. DiNardo, and S. A. Wasserman. 1993. Toward a molecular genetic analysis of spermatogenesis in *Drosophila melanogaster*: characterization of male-sterile mutants generated by single P element mutagenesis. Genetics 135:489–505.

Cheng, T., N. S. Peachey, S. Li, Y. Goto, Y. Cao, and M. I. Naash. 1997. The effect of peripherin/rds haploinsufficiency on rod and cone photoreceptors. Journal of Neuroscience 17:8118–8128.

Cooke, J., M. A. Nowak, M. Boerlijst, and J. Maynard Smith. 1997. Evolutionary origins and maintenance of redundant gene expression during metazoan development. Trends in Genetics 13:360–364.

Coyne, J. A., and H. A. Orr. 1989. Two rules of speciation. Pages 180–207 *in* D. Otte and J. Endler, eds. Speciation and its consequences. Sinauer, Sunderland, Mass.

———. 1997. "Patterns of speciation in *Drosophila*" revisited. Evolution 51:295–303.

Coyne, J. A., A. W. Meyers, A. P. Crittenden, and P. Sniegowski. 1993. The fertility effects of pericentric inversions in *Drosophila melanogaster*. Genetics 134:487–496.

Coyne, J. A., S. Simeonidis, and P. Rooney. 1998. Relative paucity of genes causing inviability in hybrids between *Drosophila melanogaster* and *D. simulans*. Genetics 150: 1091–1103.

Delbridge, M. L., J. L. Harry, R. Toder, R. J. W. O'Neill, K. Ma, A. C. Chandley, and J. A. M. Graves. 1997. A human candidate spermatogenesis gene, *RBM1*, is conserved and amplified on the marsupial Y chromosome. Nature Genetics 15:131–136.

Dunham, I., A. R. Hunt, J. E. Collins, R. Bruskiewich, D. M. Beare, M. Clamp, L. J. Smink, and R. Ainscou. 1999. The DNA sequence of human chromosome 22. Nature (London) 402:489–495.

Eddy, E. M., J. E. Welch, and D. A. O'Brien. 1993. Gene expression during spermatogenesis. Pages 182–232 *in* D. de Kretser, ed. Molecular biology of the male reproductive system. Academic Press, New York.

Eichler, E. E., F. Lu, Y. Shen, R. Antonacci, V. Jurecic, N. A. Doggett, R. K. Moyzis, et al. 1996. Duplication of a gene-rich cluster between 16p11.1 and Xq28: a novel pericentromeric-directed mechanism for paralogous genome evolution. Human Molecular Genetics 5:899–912.

Eichler, E. E., M. L. Budarf, M. Rocchi, L. L. Deaven, N. A. Doggett, A. Baldini, D. L. Nelson, and H. W. Mohrenweiser. 1997. Interchromosomal duplications of the adrenoleukodystrophy locus: a phenomenon of pericentromeric plasticity. Human Molecular Genetics 6: 991–1002.

Elliot, D. J., M. R. Millar, K. Oghene, A. Ross, F. Kiesewetter, J. Pryor, M. McIntyre, et al. 1997. Expression of *RBM* in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm. Proceedings of the National Academy of Sciences of the USA 94:3848–3853.

Fero, M. L., E. Randel, K. E. Gurley, J. M. Roberts, and C. J. Kemp. 1998. The murine gene *p27-Kip1* is haploinsufficient for tumour suppression. Nature (London) 396:177–180.

Ferris, S. D., and G. S. Whitt. 1979. Evolution of the differential regulation of duplicate genes after polyploidization. Journal of Molecular Evolution 12:267–317.

Ferris, S. D., S. L. Portnoy, and G. S. Whitt. 1979. The roles of speciation and divergence time in the loss of duplicate gene expression. Theoretical Population Biology 15:114–139.

Fischer, G., S. A. James, I. N. Roberts, S. G. Oliver, and E. J. Louis. 2000. Chromosomal evolution in *Saccharomyces*. Nature (London) 405:451–454.

Fisher, R. A. 1931. The evolution of dominance. Biological Reviews 6:345–368.

Force, A., M. Lynch, F. B. Pickett, A. Amores, and J. Postlethwait. 1999. The preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531–1545.

Gastony, G. J. 1991. Gene silencing in a polyploid homosporous fern: paleopolyploidy revisted. Proceedings of the National Academy of Science of the USA 88: 1602–1605.

Gavrilets, S. 1999. A dynamical theory of speciation on holey adaptive landscapes. American Naturalist 154:1–22.

Gavrilets, S., and J. Gravner. 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. Journal of Theoretical Biology 184:51–64.

Gerhart, J., and M. Kirschner. 1997. Cells, embryos, and evolution. Blackwell Science, Malden, Mass.

Gottlieb, L. D. 1982. Conservation and duplication of isozymes in plants. Science (Washington, D.C.) 216: 373–380.

Gottlieb, L. D., and V. S. Ford. 1997. A recently silenced, duplicate *PgiC* locus in *Clarkia*. Molecular Biology and Evolution 14:125–132.

Gould-Somero, M., and L. Holland. 1974. The timing of RNA synthesis for spermiogenesis in organ cultures of *Drosophila melanogaster* testes. Wilhelm Roux' Archiv für Entwicklungsmechanik der Organismen 174: 133–148.

Grant, D., P. Cregan, and R. C. Shoemaker. 2000. Genome organization in dicots: genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. Proceedings of the National Academy of Sciences of the USA 97:4168–4173.

Graves, J. A. M. 1995. The origin and function of the mammalian Y chromosome and Y-borne genes—an evolving understanding. BioEssays 17:311–321.

Hackstein, J. H. P., R. Hochstenbach, E. Hauschteck-Jungen, and L. W. Beukeboom. 1996. Is the Y chromosome of *Drosophila* an evolved supernumerary chromosome? BioEssays 18:317–323.

Haldane, J. B. S. 1922. Sex-ratio and unisexual sterility in hybrid animals. Journal of Genetics 12:101–109.

Hall, J. G. 1990. Genomic imprinting: review and relevance to human diseases. American Journal of Human Genetics 46:857–873.

Holland, P. W. H., J. Garcia-Fernandez, N. A. Williams, and A. Sidow. 1994. Gene duplications and the origins of vertebrate development. Development 1994(suppl.): 125–133.

Hollocher, H., and C.-I. Wu. 1996. The genetics of reproductive isolation in the *Drosophila simulans* clade: *X* vs. autosomal effects and male vs. female effects. Genetics 143:1243–1255.

Houchins, K., M. O'Dell, R. B. Flavell, and J. P. Gustafson. 1997. Cytosine methylation and nucleolar dominance in cereal hybrids. Molecular and General Genetics 255: 294–301.

Hughes, A. L. 1999. Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. Journal of Molecular Evolution 48:565–576.

Jack, J., and Y. DeLotto. 1995. Structure and regulation of a complex locus: the *cut* gene of *Drosophila*. Genetics 139:1689–1700.

Jacubiczka, S., F. Schnieders, and J. Schmidtke. 1993. A bovine homolog of the human *TSPY* gene. Genomics 17:732–735.

Jegalian, K., and D. C. Page. 1998. A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature (London) 394:776–780.

Johnson, N. A., D. E. Perez, E. L. Cabot, H. Hollocher, and C.-I. Wu. 1992. A direct test of X-Y interactions as a cause of hybrid sterility in *Drosophila*. Nature (London) 358:751–753.

Kalmykova, A. I., Y. Y. Shevelyov, A. A. Dobrista, and V. A. Gvozdev. 1997. Acquisition and amplification of a testis-expressed autosomal gene, *SSL*, by the *Drosophila* Y chromosome. Proceedings of the National Academy of Sciences of the USA 94:6297–6302.

Kirchhamer, C. V., C.-H. Yuh, and E. H. Davidson. 1996. Modular cis-regulatory organization of developmentally expressed genes: two genes transcribed territorially in the sea urchin embryo, and additional examples. Proceedings of the National Academy of Sciences of the USA 93:9322–9328.

Lamande, S. R., J. F. Bateman, W. Hutchison, R. J. M. Gardner, S. P. Bower, E. Byrne, and H. H. M. Dahl. 1998. Reduced collagen VI causes Bethlem myopathy: a heterozygous *COL6A1* nonsense mutation results in mRNA decay and functional haploinsufficiency. Human Molecular Genetics 7:981–989.

Lamnissou, K., M. Loukas, and E. Zouros. 1996. Incompatibilities between Y chromosome and autosomes are responsible for male hybrid sterility in crosses between *Drosophila virilis* and *Drosophila texana*. Heredity 76: 603–609.

Lande, R. 1984. The expected fixation rate of chromosomal inversions. Evolution 38:743–752.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937–951.

Levin, D. A., and A. C. Wilson. 1976. Rates of evolution in seed plants: net increase in diversity of chromosome numbers and species numbers through time. Proceedings of the National Academy of Sciences of the USA 73:2086–2090.

Lewis, W. H., ed. 1980. Polyploidy: biological relevance. Plenum, New York.

Li, W.-H. 1980. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. Genetics 95:237–258.

Li, Z., S. R. M. Pinson, A. H. Paterson, W. D. Park, and J. W. Stansel. 1997. Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. Genetics 145:1139–1148.

Lin, X., S. Kaul, S. Rounsley, T. P. Shea, M.-I. Benito, and C. D. Town. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature (London) 402:761–768.

Lindsley, D. L., and E. Lifschytz. 1972. The genetic control of spermatogenesis in *Drosophila*. Pages 203–222 *in* R. A.

Beatty and S. Gluecksohn-Waelsch, eds. The genetics of the spermatozoon. Bogtrykkeriet Forum, Copenhagen.

Lindsley, D. L., and K. T. Tokuyasu. 1980. Spermatogenesis. Pages 225–294 *in* M. Ashburner and T. R. F. Wright, eds. The genetics and biology of *Drosophila.* Vol. 2d. Academic Press, New York.

Livak, K. J. 1984. Organization and mapping of a sequence on the *Drosophila melanogaster* X and Y chromosomes that is transcribed during spermatogenesis. Genetics 107:611–634.

Lohe, A. R., and P. A. Roberts. 1990. An unusual Y chromosome of *Drosophila simulans* carrying amplified rDNA spacer without rRNA genes. Genetics 125:399–406.

Lundin, L. 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16:1–9.

Lynch, M., and J. Conery. 2000. The evolutionary fate and consequences of gene duplication. Science (Washington, D.C.) (in press).

Lynch, M., and A. Force. 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics 154: 459–473.

Lynch, M., and J. B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, Mass.

Manz, E., F. Schnieders, A. M. Brechlin, and J. Schmidtke. 1991. Cloning and sequence analysis of a microheterogeneous gene family organized as constitutive elements in *DYZ5* tandem repeat units on the human Y chromosome. Genomics 17:726–731.

Martin, A. P. 1999. Increasing genomic complexity by gene duplication and the origin of vertebrates. American Naturalist 154:111–128.

Mayer, K., C. Schuller, R. Wambutt, G. Murphy, G. Volckaert, T. Pohl, A. Dusterhoft, and W. Stiekema. 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana.* Nature (London) 402:769–777.

Mayer, W., A. Niveleau, J. Walter, R. Fundele, and T. Haaf. 2000. Demethylation of the zygotic paternal genome. Nature (London) 403:501–502.

McCune, A. R., and N. R. Lovejoy. 1998. The relative rate of sympatric and allopatric speciation in fishes: tests using DNA sequence divergence between sister species and among clades. Pages 172–185 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford University Press, New York.

Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biological Symposium 6:71–125.

Muller, H. J., and F. Settles. 1927. The non-functioning of the genes in spermatozoa. Zeitshrift für Induktive Abstammungs und Verebungslehre 43:285–312.

Nadeau, J. H., and D. Sankoff. 1997. Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. Genetics 147: 1259–1266.

Nei, M., and A. K. Roychoudhury. 1973. Probability of fixation of nonfunctional genes at duplicate loci. American Naturalist 107:362–372.

Nei, M., T. Maruyama, and C.-I. Wu. 1983. Models of evolution of reproductive isolation. Genetics 103:557–579.

Normes, S., M. Clarkson, I. Mikkola, M. Pedersen, A. Bardsley, J. P. Martinez, S. Krauss, and T. Johansen. 1998. Zebrafish contains two *Pax6* genes involved in eye development. Mechanisms of Development 77:185–196.

Oda, T., A. G. Elkahloun, B. L. Pike, K. Okajima, I. D. Krantz, A. Genin, D. A. Piccoli, et al. 1997. Mutations in the human *Jagged1* gene are responsible for Alagille syndrome. Nature Genetics 16:235–242.

Ohno, S. 1970. Evolution by gene duplication. Springer, Heidelberg.

Oka, H.-I. 1974. Analysis of genes controlling F_1 sterility in rice by the use of isogenic lines. Genetics 77:521–534.

———. 1988. Functions and genetic bases of reproductive barriers. Pages 181–205 *in* H. Oka, ed. Origin of cultivated rice. Elsevier, New York.

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

———. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139: 1805–1813.

———. 1996. Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335.

———. 1997. Haldane's rule. Annual Review of Ecology and Systematics 28:195–218.

Parker, H. R., D. P. Philipp, and G. S. Whitt. 1985. Relative developmental success of interspecific *Lepomis* hybrids as an estimate of gene regulatory divergence between species. Journal of Experimental Zoology 233:451–466.

Pennisi, E. 1998. A mosaic of gene duplications. Science (Washington, D.C.) 280:1693.

Postlethwait, J. H., Y.-L. Tan, M. Gates, S. Horne, A. Amores, A. Brownlie, and A. Do. 1998. Vertebrate genome evolution and the zebrafish gene map. Nature Genetics 18:345–349.

Presgraves, D. C., and H. A. Orr. 1998. Haldane's rule in taxa lacking a hemizygous X. Science (Washington, D.C.) 282:952–954.

Prowell, D. P. 1998. Sex linkage and speciation in lepidoptera. Pages 309–319 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford University Press, New York.

Robertson, H. M. 1998. Two large families of chemoreceptor genes in the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae* reveal extensive gene duplication, diversification, movement, and intron loss. Genome Research 8:449–463.

Rubin, G. M., M. D. Yandell, J. R. Wortman. G. L. Gabor Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, et al. 2000. Comparative genomics of the eukaryotes. Science (Washington, D.C.) 287:2204–2215.

Ruggiu, M., R. Speed, M. Taggart, S. J. McKay, F. Kilanowski, P. Saunders, J. Dorin, and H. J. Cooke. 1997. The mouse *Dazla* gene encodes a cytoplasmic protein essential for gametogenesis. Nature (London) 389:73–77.

Sampsell, B., and W. A. Held. 1985. Variation in the major urinary protein multigene family in wild-derived mice. Genetics 109:549–568.

Sawamura, K., A. W. Davis, and C.-I. Wu. 2000. Genetic analysis of speciation by means of introgression into *Drosophila melanogaster.* Proceedings of the National Academy of Sciences of the USA 97:2652–2655.

Saxena, R., L. G. Brown, T. Hawkins, R. K. Alagappan, H. Skaletsky, M. P. Reeve, R. Reijo, et al. 1996. The *DAZ* gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. Nature Genetics 14:292–299.

Sidow, A. 1996. Gen(om)e duplications in the evolution of early vertebrates. Current Opinion in Genetics and Development 6:715–722.

Skrabanek, L., and K. H. Wolfe. 1998. Eukaryote genome duplication—where's the evidence? Current Opinion in Genetics and Development 8:694–700.

Soltis, D. E., and P. S. Soltis. 1993. Molecular data and the dynamic nature of polyploidy. Critical Reviews in Plant Sciences 12:243–273.

Steinemann, M., and S. Steinemann. 1997. The enigma of Y chromosome degeneration: *TRAM*, a novel retrotransposon is preferentially located on the Neo-Y chromosome of *Drosophila miranda.* Genetics 145:261–266.

Steinemann, M., S. Steinemann, and W. Pinsker. 1996. Evolution of the larval cuticle proteins coded by the secondary sex chromosome pair: X2 and Neo-Y of *Drosophila miranda.* I. Comparison at the DNA sequence level. Journal of Molecular Evolution 43:405–412.

Takahata, N., and T. Maruyama. 1979. Polymorphism and loss of duplicate gene expression: a theoretical study with application to tetraploid fish. Proceedings of the National Academy of Sciences of the USA 76:4521–4525.

Templeton, A. R. 1981. Mechanisms of speciation—a population genetic approach. Annual Review of Ecology and Systematics 12:23–48.

Ting, C.-T., S.-C. Tsaur, M.-L. Wu, and C.-I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science (Washington, D.C.) 282:1501–1504.

True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans.* Genetics 142:819–837.

Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane's rule. Genetics 140:389–402.

———. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663–1679.

Vielle-Calzada, J.-P., R. Baskar, and U. Grossniklaus. 2000. Delayed activation of the paternal genome during seed development. Nature (London) 404:91–94.

Wade, M. J., and C. J. Goodnight. 1998. Perspective: the theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution 52:1537–1553.

Wagner, A., G. P. Wagner, and P. Similion. 1994. Epistasis can facilitate the evolution of reproductive isolation by peak shifts: a two-locus two-allele model. Genetics 138: 533–545.

Walsh, J. B. 1982. Rate of accumulation of reproductive isolation by chromosome rearrangements. American Naturalist 120:510–532.

———. 1995. How often do duplicated genes evolve new functions? Genetics 110:345–364.

Walsh, N. E., and D. Charlesworth. 1992. Evolutionary interpretations of differences in pollen tube growth rates. Quarterly Review of Biology 67:19–37.

Watterson, G. A. 1983. On the time for gene silencing at duplicate loci. Genetics 105:745–766.

Werth, C. R., and M. D. Windham. 1991. A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. American Naturalist 137:515–526.

Westin, J., and M. Lardelli. 1997. Three novel *notch* genes in zebrafish: implications for vertebrate *Notch* gene evolution and function. Development, Genes, and Evolution 207:51–63.

White, M. J. D. 1978. Modes of speciation. W. H. Freeman, San Francisco.

Whitt, G. S., W. F. Childers, and P. L. Cho. 1973. Allelic expression at enzyme loci in an intertribal hybrid sunfish. Journal of Heredity 64:55–61.

Wilm, B., E. Dahl, H. Peters, R. Balling, and K. Imai. 1998. Targeted disruption of *Pax1* defines its null phenotype and proves haploinsufficiency. Proceedings of the National Academy of Sciences of the USA 95:8692–8697.

Wilson, A. C., L. R. Maxson, and V. M. Sarich. 1974. Two types of molecular evolution: evidence from studies of interspecific hybridization. Proceedings of the National Academy of Sciences of the USA 71:2843–2847.

Wilson, A. C., G. L. Bush, S. M. Case, and M.-C. King. 1975. Social structuring of mammalian populations and rate of chromosomal evolution. Proceedings of the National Academy of Sciences of the USA 72:5061–5065.

Wilson, H. D., S. C. Barber, and T. Walters. 1983. Loss of duplicate gene expression in tetraploid *Chenopodium.* Biochemical Systematics and Ecology 11:7–13.

Wu, C.-I. and M. F. Palopoli. 1994. Genetics of postmating reproductive isolation in animals. Annual Review of Genetics 27:283–308.

Xu, H., I. Swoboda, P. L. Bhalla, and M. B. Singh. 1999. Male gametic cell-specific gene expression in flowering plants. Proceedings of the National Academy of Sciences of the USA 96:2554–2558.

Yiu, G. K., and N. B. Hecht. 1997. Novel testis-specific protein-DNA interactions activate transcription of the mouse protamine 2 gene during spermatogenesis. Journal of Biological Chemistry 24:26926–26933.

Yuan, X., M. Miller, and J. M. Belote. 1996. Duplicated proteasome subunit genes in *Drosophila melanogaster* encoding testes-specific isoforms. Genetics 144:147–157.

Yuh, C.-H., H. Bolouri, and E. H. Davidson. 1998. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. Science (Washington, D.C.) 279:1896–1902.

Zhang, P., and R. L. Stankiewicz. 1998. Y-linked male sterile mutations induced by *P* element in *Drosophila melanogaster*. Genetics 150:735–744.

Editor: Joseph Travis