Discord Between the Phylogenies Inferred From Molecular *Versus* Functional Data: Uneven Rates of Functional Evolution or Low Levels of Gene Flow?

Michael F. Palopoli, Andrew W. Davis[1] and Chung-I Wu[2]

Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637

Manuscript received August 14, 1995
Accepted for publication August 9, 1996

ABSTRACT

According to measures of molecular divergence, the three species of the *Drosophila simulans* clade are closely related to and essentially equidistant from each other. We introgressed 10% of the *D. sechellia X* chromosome into a pure *D. simulans* genetic background and found that males carrying this introgressed region were consistently fertile; in contrast, males carrying the same segment from *D. mauritiana* are sterile and suffer from incompatibilities at a minimum of four loci. Together with other recent results, these data suggest that *D. simulans* and *D. sechellia* are much more closely related to each other than either is to *D. mauritiana*. How can we reconcile the phylogeny inferred from the density of hybrid sterility genes with that inferred from molecular divergence? If the molecular phylogeny is correct, the discrepancy might be explained by uneven rates of functional evolution, resulting in the uneven accumulation of substitutions with corresponding negative effects in hybrids. If the functional phylogeny is correct, then low levels of gene flow across nascent species boundaries, particularly for loci not tightly linked to a hybrid sterility gene, may have erased the original pattern of lineage splitting. We propose tests that will allow us to discriminate between these hypotheses.

IN a landmark study, WILSON *et al.* (1974a) argued that there are two major types of molecular evolution. One is the steady accumulation of fixed differences in protein sequences, which occurs at roughly equal rates for a given protein in a wide range of species; the other is divergence in the regulatory systems responsible for anatomical/morphological evolution, which occurs at extremely variable rates in different lineages. To test this hypothesis, WILSON *et al.* contrasted measures of divergence between proteins of species pairs capable of producing viable hybrids in two broad taxonomic groups: frogs and mammals (50 and 31 species pairs, respectively). As a general rule, protein sequence differences between mammalian species capable of producing viable hybrids are an order of magnitude smaller than the corresponding values for pairs of frog species. To explain this discrepancy, the authors proposed that mammals, unlike frogs, have experienced rapid morphological evolution and correspondingly rapid evolutionary change in the systems regulating coordinated gene expression (see also WILSON *et al.* 1974b; KING and WILSON 1975). By contrasting the molecular evolutionary results for these taxa with an independent data set based on a functional assay, the authors presented an interesting hypothesis that was not apparent based on either type of data alone and might not have been considered otherwise.

Recently, several studies have addressed questions of molecular population genetics in the three sibling species of the *Drosophila simulans* clade (HEY and KLIMAN 1993; KLIMAN and HEY 1993; HILTON *et al.* 1994; BALLARD and GALWAY 1995; CACCONE *et al.* 1996). Perhaps the most consistent result reported in all of these analyses is that there is relatively little sequence divergence and many shared polymorphisms among these species. Together with earlier reports (COYNE and KREITMAN 1986; SOLIGNAC and MONNEROT 1986; CARIOU 1987; CACCONE *et al.* 1988; SATTA and TAKAHATA 1990), these molecular results suggest that these species shared a common ancestor in the very recent past, and that all three descendant species arose within an extremely short period of time. This conclusion is supported by extensive nucleotide sequence data for 11 nuclear and five mitochondrial genes, as well as comparisons of allozymes, DNA:DNA hybridization profiles, and mitochondrial DNA restriction fragment length polymorphisms. In accord with common practice for closely related, allopatric taxa, the numerous shared polymorphisms have generally been interpreted as the remnants of ancient polymorphisms (*i.e.*, polymorphisms that were segregating in the ancestral population and have persisted in more than one descendant population until the present day; see NEI 1987; TAKAHATA 1989; WU 1991); however, gene flow across species boundaries has also been invoked in some cases (see SOLIGNAC and MONNEROT 1986; HILTON *et al.* 1994; BALLARD and

Corresponding author: Michael F. Palopoli, Howard Hughes Medical Institute, MC1028, AMB N-101, 5841 South Maryland Ave., University of Chicago, Chicago, IL 60637. E-mail: palo@midway.uchicago.edu

[1] *Present address:* Department of Genetics, University of Cambridge, Cambridge CB2 3EH, United Kingdom.
E-mail: awd@mole.bio.cam.ac.uk.

[2] *E-mail address:* ciwu@pondside.uchicago.edu

GALWAY 1995). The species phylogeny is considered to be an unresolved trichotomy.

In parallel to these studies of molecular evolution, the genetic basis of reproductive isolation between these three sibling species has been analyzed extensively (COYNE and CHARLESWORTH 1986, 1989; JOHNSON *et al.* 1992, 1993; NAVEIRA 1992; PEREZ *et al.* 1993; WU *et al.* 1993; PALOPOLI and WU 1994; PEREZ and WU 1995; DAVIS and WU 1996; HOLLOCHER and WU 1996). As is typical between recently diverged species of Drosophila, male hybrids are sterile whereas female hybrids are fertile (reviewed in WU and DAVIS 1993). The genetical analyses have shown that these species have diverged at a large number of loci that have various deleterious effects on male fertility. For example, based on fine-scale analyses for the *D. mauritiana-D. simulans* species pair, it has been estimated that there are *at least* 40 such loci on the *X* chromosome alone (PALOPOLI and WU 1994; DAVIS and WU 1996). The question addressed in the present analysis is whether the phylogeny suggested by the number of genetic incompatibilities in hybrids agrees with the phylogeny that has been established using measures of divergence at molecular markers.

MATERIALS AND METHODS

Stocks and crosses: The crossing scheme used to generate introgressions was similar to the one outlined in Figure 1 of DAVIS and WU (1996). For 15 generations, *y v* [m^+g^+]*f*/*y v m g f* virgin females (brackets denote introgressed region) from each of four independent introgression lines were backcrossed to *D. simulans* males from the *y v m g f* stock. At this point, males harboring the [m^+g^+] introgressions were collected from each line and their fertility assessed (see below). All fly cultures were maintained at 22–23° and reared on standard cornmeal medium. The *y v m g f* and wild-type *D. sechellia* stocks were provided by J. COYNE.

DNA markers: To determine the lengths of introgressions more precisely, as well as to verify the species origin of the internal segments, the introgressions were mapped using four species-specific, single-strand conformation polymorphism (SSCP) markers as described in DAVIS and WU (1996). These four SSCP markers map to the 10C2-5 (*RpII215*), 11A7-8 (*gd*), 11B3-9 (P1 clone), and 12B-C (*Yp3*) band regions of the *D. melanogaster* polytene chromosomes (see Figure 1).

Determination of fertility: For each introgression, a population of males was examined to assess the effects of the introgressed segment on male fertility. Three criteria were employed, as follows: (1) Sperm motility: Males were separated by genotype and aged 4–7 days without access to females. Testes of individual males were dissected in a drop of Drosophila Ringer's solution, gently squashed under a coverslip, and examined using phase-contrast microscopy. Individual males were placed in one of four categories according to the approximate number of motile sperm that were visible under the coverslip, as follows: none (sterile), between one and 25, between 25 and 100, >100 (apparently quite fertile). (2) Spermatogenic phenotype: For most of these males, a qualitative examination of spermatogenesis was also conducted. Individual males were placed in one of three categories, as follows: normal onion cell stages, abnormalities in onion cell stage (*i.e.*, irregularities in the numbers and/or volumes of nuclei *vs.* mitochondrial derivatives within a given cyst), or testes completely atrophied (*i.e.*, visibly reduced in size and nothing but cellular debris found inside). (3) Progeny production: Multiple virgin males from each line were mated individually to *D. simulans C(1) y w* females in an attempt to establish a culture. Vials that failed to produce progeny after 7 days were classified as sterile. If a culture so established could not be maintained, it was considered to be subfertile.

TABLE 1

Analysis of motile sperm production in males carrying *D. sechellia* [m^+ g^+] introgressions

Line no.	Motile sperm estimate[a] 0	1–25	25–100	>100	Total[b]	Fertile by mating[c]
sec-A (I)	2	4	4	11	21 (0.905)	Yes
sec-B (II)[d]	2	1	5	10	18 (0.889)	Yes
sec-C (II)	2	3	1	4	10 (0.800)	Yes
sec-D (I)	3	10	3	1	17 (0.824)	Yes
Total	9	18	13	26	66 (0.864)	

[a] Virgin males were collected and held for 4–7 days before dissection.

[b] Numbers in parentheses are overall proportions of males with motile sperm.

[c] Five virgin males from each line were collected and allowed to mate with *D. simulans C(1) y w virgin* females for 7 days.

[d] Type I lines have introgression breakpoints distal of 10C2-5 (*RpII215* SSCP marker) while type II lines have distal breakpoints proximal of 10C2-5. See Figure 1.

RESULTS

In parallel with a similar study of introgressions covering the same region from *D. mauritiana* (DAVIS and WU 1996), we created four independent introgressions of the middle 10% of the *X* chromosome from *D. sechellia* into a pure *D. simulans* genetic background. These were then analyzed for their effects on motile sperm production (Table 1). In general, males harboring the introgressed region were able to produce sizable numbers of motile sperm. The only exception to this was the line designated sec-D: about half of these males produced <25 motile sperm under the assay conditions. If sec-D is excluded from analysis, the introgressions tested did not differ from one another in motile sperm production ($G_{adj} = 4.66$, $P > 0.25$, d.f. = 6, sec-D excluded). If sec-D is included in this comparison, there is a significant difference in motile sperm production ($G_{adj} = 21.13$, $0.01 < P < 0.005$). This result suggests the presence of a "partial" sterility factor within this single introgression; nevertheless, all lines were able to establish and maintain a culture with *D. simulans C(1) y w* females (Table 1).

The spermatogenic phenotypes exhibited by males harboring introgressions are summarized in Table 2. None of the lines differed significantly in the average degree of spermatogenic disruption. In each line, over 80% of the males exhibited normal onion cell stages, indicating little premeiotic disruption; however, some

TABLE 2

Spermatogenic phenotypes of males carrying *D. sechellia* [m^+ g^+] introgressions

Line no.	Spermatogenic phenotype[a] No. with normal onions	No. with onion cell defects	No. with atrophied testes	Total[b]
sec-A (I)	15	3	0	18 (0.833)
sec-B (II)[c]	15	3	0	18 (0.833)
sec-C (II)	8	0	1	9 (0.889)
sec-D (I)	13	2	0	15 (0.867)
Total	51	8	1	60 (0.850)

[a] Virgin males were collected and held for 4–7 days before dissection.

[b] Number in parentheses is overall proportion of males with normal onion cell morphology.

[c] Type I lines have introgression breakpoints distal of 10C2-5 (*RpII215* SSCP marker) while type II lines have distal breakpoints proximal of 10C2-5. See Figure 1.

males did exhibit noticeable amounts of subcellular debris in their seminal vesicles (*e.g.*, sperm tails, vacuoles). Overall, these introgressions caused only minor disruptions in spermatogenesis.

The breakpoints of the *D. sechellia* introgressions (Figure 1) were mapped using the species-specific SSCP markers introduced in DAVIS and WU (1996). All four introgressions extended from distal of *m* to proximal of *Yp3* at 12B-C; however, two sizes of introgressions were identified that differed in the location of their distal breakpoints with respect to the 10C2-5 marker, as follows (Figure 1): lines designated sec-B and sec-C had distal breakpoints that were proximal of 10C2-5 (making these the two smaller introgressions, designated type II), whereas sec-A and sec-D had distal breakpoints that were distal of 10C2-5 (making these the two larger introgressions, designated type I).

DISCUSSION

Although >10% of the euchromatic portion of the *D. sechellia X* chromosome was introgressed into a *D. simulans* genetic background, males harboring these introgressions were consistently fertile. In contrast, if the same region is introgressed from *D. mauritiana* into *D. simulans*, the resulting males are always sterile and suffer from hybrid incompatibilities at a minimum of four loci (DAVIS and WU 1996). Together, these genetic results suggest that *D. simulans* and *D. sechellia* are more closely related to each other than either is to *D. mauritiana*. This phylogenetic hypothesis contrasts sharply with the unresolved trichotomy that has been inferred from DNA sequence data.

The hypothesis that *D. simulans* and *D. sechellia* are more closely related to each other than either is to *D. mauritiana* is supported by a series of analyses of the male sterility associated with introgressions of various chromosomal regions between these species, as follows: the entire *Y* chromosome (JOHNSON *et al.* 1992, 1993; ZENG and SINGH 1993); a proximal region of the *X* chromosome, comprising ~5% of its length (PEREZ *et al.* 1993; PEREZ and WU 1995); the most distal 5% of the *X* chromosome (CABOT *et al.* 1994); and five regions on the second chromosome associated with visible genetic markers (HOLLOCHER and WU 1996). In every genomic region that has been examined in detail, there are more male sterility factors associated with a segment introgressed into *D. simulans* from *D. mauritiana* than are found when the corresponding segment has been introgressed from *D. sechellia*; this discrepancy is apparent in a compilation of the results for those regions of the *X* chromosome that have been analyzed in the most detail for both species (Table 3). In each of nine genomic regions analyzed in sufficient detail to make the comparison possible, at least one more hybrid sterility gene needs to be invoked for *D. simulans* × *D. mauritiana* than for *D. simulans* × *D. sechellia*. Overall, we must

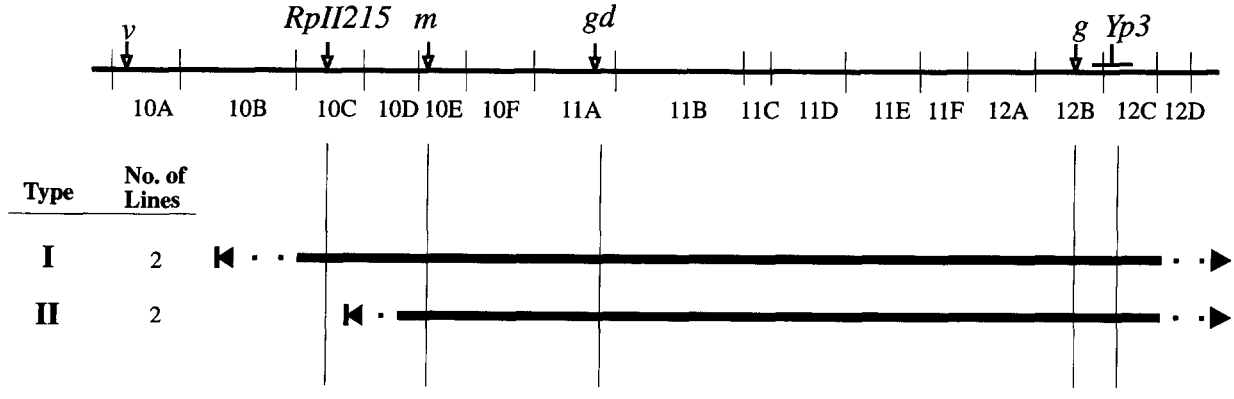

FIGURE 1.—Physical mapping of the *D. sechellia* introgressions (■) listed in Tables 1 and 2. Four SSCP markers were developed from *D. melanogaster* clones in this region as follows: *RpII215* (10C2-5), *gd* (11A7-8), P1 clone DS08832 from the Drosophila Genome Center (11B3-9), and *Yp3* (12B/C). Dotted lines indicate breakpoint uncertainty between the two markers shown. Proximal breakpoints were not determined beyond *Yp3* (12B/C). All introgressions shown were male fertile (Table 1). The polytene chromosome map depicted here is scaled to the number of bands in each letter division.

TABLE 3

Comparison of the number of hybrid male sterility factors identified in high-resolution mapping studies of X-linked introgressions from *D. mauritiana* or *D. sechellia* into *D. simulans*

Polytene region analyzed	No. of hybrid male sterility factors mapped[a]	
	D. mauritiana introgression	*D. sechellia* introgression
4A-5C[b]	3	1
10A-12B[c]	4	0
13F-18C[d]	2	0
Total	9	1

[a] To be included in the table, the region must have been analyzed with sufficient rigor for introgressions from both species. Using the criterion, we excluded three mapped *D. mauritiana* factors and a single mapped *D. sechellia* factor because the comparison for the region in question could not be made fairly. Note that not all factors in a given region were necessarily detected; for example, epistasis could conceal factors.

[b] CABOT *et al.* 1993. The region from polytene bands 1A-4A is not included because it was not analyzed adequately in *D. mauritiana;* as a result, one mapped *D. sechellia* factor was excluded.

[c] DAVIS and WU 1996. Two mapped *D. mauritiana* factors were excluded.

[d] PEREZ *et al.* 1993; PEREZ and WU 1995. The region proximal of polytene band 18C is not included because physical breakpoints were not determined beyond this location. Careful phenotypic analyses suggest, however, that there is a factor (or factors) proximal to 18C that are present in *D. mauritiana* but not *D. sechellia.*

invoke about nine times as many genes for the former cross as for the latter (*i.e.,* the density of functionally divergent genes contributing to hybrid male sterility appears to be about ninefold as great).

The most widely accepted, neo-Darwinian view is that the alleles contributing to hybrid male sterility evolved for reasons having nothing to do with their effect in hybrids (*i.e.,* they evolved via natural selection and/or random drift operating within each lineage separately); thus, the deleterious effects of these alleles when placed in a hybrid genetic background are considered incidental, negative, pleiotropic effects of the various substitutions within each lineage (*e.g.,* DOBZHANSKY 1937). Based on this view, with all other things being equal, the much greater capacity of the *D. mauritiana* genome than that of *D. sechellia* for causing sterility in a *D. simulans* background should indeed be interpreted as indicative of phylogenetic relatedness: *D. simulans* and *D. sechellia* would seem to be much more closely related to each other than either is to *D. mauritiana.* Thus there is a need to reconcile the phylogeny inferred from the density of hybrid sterility genes with the phylogeny inferred from sequence comparisons at arbitrarily chosen loci (*i.e.,* chosen primarily because molecular divergence could be measured at those loci). There are at least two different ways of explaining this disagreement between data sets.

Uneven rates of functional divergence: First, it is possible that the phylogeny originally suggested by the DNA sequence data is closer to correct, and the species actually diverged within an extremely short period of time (*i.e.,* the phylogeny is essentially a trichotomy). To explain the hybrid sterility data, then, we must assume that the rate of accumulation of genes capable of causing hybrid sterility varies between lineages. Two verbal models explaining this type of evolutionary scenario are summarized in Figure 2 and outlined below. First, if the process of accumulating functional divergence for the genes involved in male fertility occurred much more quickly in the *D. mauritiana* lineage than in either the *D. simulans* or *D. mauritiana* lineages, then this would help to explain the discrepancy between data sets (Figure 2, middle). There is no reason, however, to expect the *D. mauritiana* lineage to evolve at such a different rate; none of the three species exhibits significant morphological divergence from the other two and it is *D. sechellia* that is the only one of the three that has undergone clear adaptive evolution; it has evolved to specialize on the fruit of *Morinda citrifolia* (LOUIS and DAVID 1986; R'KHA *et al.* 1991). Nevertheless, it is possible that the *D. mauritiana* lineage has evolved more quickly, and this scenario makes a testable prediction: the density of hybrid sterility genes between *D. mauritiana* and *D. sechellia* should prove to be about the same as that already observed between *D. mauritiana* and *D. simulans* (*i.e.,* relatively high; refer to Table 3).

Second, if functional divergence occurred much more slowly in the *D. sechellia* lineage than in the other two lineages, then this would help to explain the discrepancy between data sets (Figure 2, bottom). This type of scenario makes a different, testable prediction: The density of hybrid sterility genes between *D. mauritiana* and *D. sechellia* should prove to be about the same as that already observed between *D. sechellia* and *D. simulans* (*i.e.,* relatively low; refer to Table 3). Interestingly, there is a plausible reason to expect the *D. sechellia* lineage to accumulate factors at a slower rate. Based on studies of molecular polymorphism, it appears that *D. sechellia* has experienced a much lower effective population size than the other two species (HEY and KLIMAN 1993; HILTON *et al.* 1994).

Presumably, a certain fraction of the substitutions that occurred in each lineage and now contribute to hybrid sterility were driven to fixation by positive selection (hybrid sterility being an incidental, pleiotropic effect of these substitutions). All other things being equal, the much smaller effective population size of *D. sechellia* would be expected to have slowed the rates at which such favorable alleles arose via mutation; in addition, reduced population size would be expected to have narrowed the window of selective coefficients

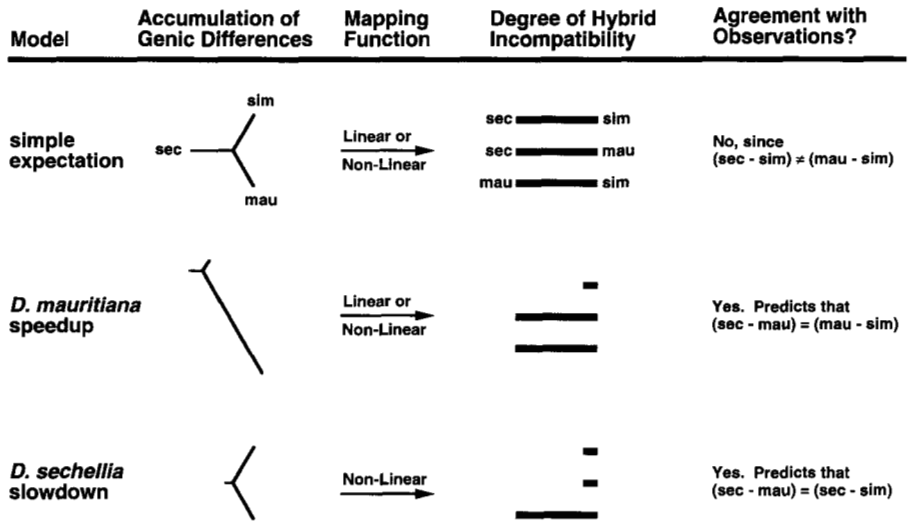

FIGURE 2.—The impact that differences in the rate of functional evolution for different lineages are expected to have on the degree of divergence measured as the relative density of mapped hybrid incompatibilities between species. These models all assume that the species phylogeny is essentially a trichotomy, which is suggested by extensive measures of divergence at arbitrarily chosen loci. The top row indicates the simplest model: that, in accord with the trichotomy suggested by the molecular results, the degrees of functional divergence, based upon estimates of the density of hybrid sterility genes, would suggest a trichotomy. This model is not correct (see Table 3). The middle row indicates a possible model that is consistent with all of the data collected so far: that functional evolution was speeded up dramatically in the *D. mauritiana* lineage, resulting in the faster accumulation of genes contributing to hybrid sterility. This model predicts that the density of hybrid sterility genes between *D. sechellia* and *D. mauritiana* will prove to be high, as has been observed for the *D. simulans* × *D. mauritiana* species pair. Finally, the bottom row indicates another possible model that is consistent with all of the data collected so far: that functional evolution was slowed down significantly in the *D. sechellia* lineage, perhaps because of the much smaller historical effective population size that has been inferred for this species based on the molecular results. This model predicts that the density of hybrid sterility genes between *D. sechellia* and *D. mauritiana* will prove to be low, as has been observed for the *D. simulans* × *D. sechellia* species pair. Note that a "Nonlinear" function mapping the relationship between genotype and phenotype is required for this third model in order to translate the "sec slowdown" into the observed ninefold difference in the density of hybrid incompatibilities (Table 3); there are excellent theoretical reasons to expect such a nonlinearity (ORR 1995). The species names *D. sechellia*, *D. simulans*, and *D. mauritiana* are represented in the top row as sec, sim, and mau, respectively. In the bottom two rows, these names are omitted to reduce visual clutter but the diagrams are in the same orientations as they are in the top row (*i.e.*, for both the hypothetical distance trees depicted in the second column and horizontal bars representing degree of hybrid incompatibility in the fourth column).

that could overcome the random effects of genetic drift (*i.e.*, only alleles with relatively *large* beneficial effects would have made it to fixation; *e.g.*, KIMURA 1983; NEI 1987). Hence, to the extent that the substitutions responsible for hybrid sterility were driven to fixation by positive selection in each lineage, the *D. sechellia* lineage would be expected to have accumulated substitutions more slowly than the other two lineages. As a result, the degree of divergence between *D. simulans* and *D. mauritiana* would be expected to exceed that between *D. sechellia* and either of these species. These basic conclusions concerning the effects of reduced population size are supported by a recent theoretical analysis of this problem (N. A. JOHNSON, unpublished results).

One potential problem with this explanation is that it is difficult to imagine any such rate discrepancy accounting for the several-fold difference observed in the density of hybrid sterility genes (see Table 3). In particular, even if there was *no evolution at all* in the *D. sechellia* lineage, and assuming roughly equal rates of evolution in the other two lineages, then the number of fixed differences between *D. simulans* × *D. mauritiana* would only approach twice that between either species and *D. sechellia*. This problem is alleviated, however, when we take into account the expected nonlinearity between the number of fixed differences between two species and the number of hybrid incompatibilities observed when these same two species are crossed (ORR 1995). There are convincing theoretical reasons to expect that the number of genic incompatibilities between any two taxa will increase *much faster than linearly* with the number of substitutions. Under certain circumstances, this "snowballing" effect of genetic interactions would be expected to make the relative density of hybrid incompatibilities a much more sensitive measure of the extent of divergence among taxa than would any measure that increased only linearly with time, such as DNA-sequence divergence at neutral sites. Indeed, something akin to the observed disagreement between the DNA-sequence and hybrid sterility data sets might have been predicted using ORR's model.

Permeable species boundaries: Alternatively, the phylogeny inferred from the degree of reproductive incompatibility could be correct in the sense that it

accurately reflects the relative order and the timing of lineage separations; this would place *D. simulans* and *D. sechellia* as the most closely related pair, with *D. mauritiana* as a distant out-group. In this case, we could explain the apparent trichotomy suggested by the DNA-sequence studies as due to interspecific introgression, whereby the genes chosen for molecular phylogeny studies have occasionally made it across species boundaries in the past, whereas genes that contribute to reproductive isolation were quickly eliminated upon introgression because of their severe fitness consequences. This scenario is illustrated in Figure 3. Under this model, the observed discrepancy between the phylogeny inferred from genes that cause hybrid sterility *vs.* randomly selected genes could be accounted for by the cumulative introgression of only the latter across permeable species boundaries. If correct, we expect that those neutral polymorphisms that are tightly linked to the genes responsible for hybrid male sterility between *D. mauritiana* and *D. simulans* [such as *Ods* (PEREZ *et al.* 1993) or *fixA* and *fixC* (CABOT *et al.* 1994)] will tend to depict a phylogeny that agrees with the patterns deduced from the overall density of hybrid sterility genes (*i.e.*, Figure 3C). Specifically, in these genomic regions, we might expect to observe greater neutral divergence between *D. mauritiana* and *D. simulans*, a higher frequency of species-diagnostic variants and perhaps reduced levels of intraspecific polymorphism in one of these species (due to a recent selective sweep). These are testable predictions, although to address them will ultimately require the molecular cloning of genes responsible for hybrid sterility to examine molecular divergence and polymorphism levels at nearby sites (*e.g.*, within 10 kb on either side of the primary genetic lesion responsible for a mapped sterility factor). Indeed, *falsifying* these predictions would require examining a large number of such regions, as only those polymorphisms that are tightly linked to sterility factors that became fixed early in divergence would tend to exhibit the original pattern.

Note that the permeable-species-boundaries hypothesis as presented here does not contradict the conclusion of KLIMAN and HEY (1993) that "The sequence data do imply an absence of gene flow between the species." They observed a pattern of within-species clustering of sequences in the estimated genealogies, which can be interpreted as lack of significant gene flow *in the recent past* (as diagrammed in the lower portion of Figure 3A). We are suggesting that low levels of gene exchange at earlier stages of divergence (see the upper portion of Figure 3A) could blur the phylogenetic pattern observed (Figure 3B), if that pattern is based on loci that are not tightly linked to genes that reduce hybrid fitness. In contrast, the genes that reduce hybrid fitness, and the regions tightly linked to those genes, would retain the underlying phylogenetic relationship more accurately (Figure 3C). It is important to realize that

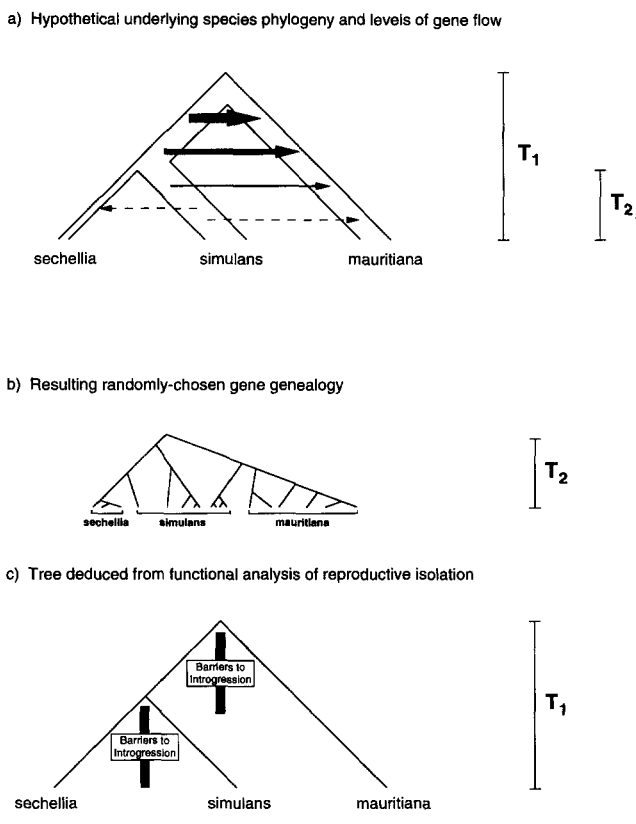

FIGURE 3.—The impact that introgressive hybridization (a) is expected to have on the species tree inferred from two types of data: the gene genealogy for a randomly chosen locus (b) *vs.* the tree deduced from a functional analysis of reproductive isolation (c). The arrows in a represent levels of gene flow, with wider arrows depicting less restricted gene flow; the opportunity for introgression is expected to decrease with time, as the number of genes contributing to reproductive isolation increases. *T1* and *T2* represent the times back to the appearance of substantial but incomplete reproductive isolation between *D. mauritiana* × [*D. simulans*/*D. sechellia*] ancestors and *D. simulans* × *D. sechellia* ancestors, respectively. Because genes that are not tightly linked to a gene that reduces hybrid fitness are able to move across incipient species' boundaries, the time back to a common ancestor inferred from a randomly chosen gene genealogy (b) is much less than the time inferred from the apparent density of hybrid male sterility factors (c).

to generate such a discrepancy would not necessarily require large levels of unrestricted gene flow, which would result in comparable levels of polymorphism in all three species. Rather, the introgressed materials depicted in Figure 3 might represent only a small subset of the total array of *D. simulans* variations. The amount of such early gene flow necessary to effectively erase (unlinked) gene genealogies can probably be inferred accurately from classical studies of population differentiation (*i.e.*, $4Nm \geq 1$; N. TAKAHATA, personal communication); it is clear that extremely low levels of gene flow are sufficient. Interestingly, low levels of gene flow among these sibling species has been suggested recently by DNA sequencing studies: there are two nonrecombining pieces of DNA (the mitochondrial genome and

the nonrecombining fourth chromosome) that appear to have moved across species boundaries within this clade in the recent past (HILTON *et al.* 1994; BALLARD and GALWAY 1995).

Speciation is often portrayed as a point event. This may be justified if the time to evolve ecological, behavioral, or physiological incompatibility is relatively short. During the process of speciation, however, there may be a protracted period of time when a limited amount of genetic exchange is possible, even in the absence of extended hybrid zones between taxa. This may apply to *D. simulans, D. mauritiana,* and *D. sechellia.* The geographical distributions of the latter two species are nested within that of the cosmopolitan *D. simulans.* Certainly, the three species can still exchange genes in a laboratory setting. Most importantly, the restriction of gene flow is not likely to be the same across the whole genome. Genes that are tightly linked to hybrid sterility factors (or other loci involved in hybrid fitness reduction) are less likely to be introgressed; however, even such restriction in gene flow is often unidirectional for a particular genomic region (WU and BECKENBACH 1983; JOHNSON *et al.* 1993; ZENG and SINGH 1993; PALOPOLI and WU 1994).

A general implication of this hypothesis is that a low level of gene flow during the time when incipient species remain incompletely isolated could blur the resulting phylogeny and confound sequence comparisons. It is conceivable that the time it takes to restrict gene flow completely could set a lower limit for the resolution of molecular phylogenies. In contrast, the genes responsible for reproductive isolation should be less affected by any such genetic exchange and therefore reflect the original pattern of lineage splitting more faithfully. Interestingly, this type of scenario could account for the apparent differences in observed "regional time depth" between molecular and morphological traits in modern human populations (FISCHMAN 1996; TISHKOFF *et al.* 1996).

We thank NORMAN JOHNSON and YUJUN XU for stimulating discussions and criticisms of earlier versions of this manuscript. We also thank NORMAN JOHNSON for suggesting the population-size explanation. This work was supported by a National Science Foundation (NSF) Graduate Fellowship and a National Institutes of Health (NIH) Genetics Training Grant to M.F.P., a William P. Grahams Fellowship from the Baxter Foundation to A.W.D., and NIH and NSF grants to C.-I.W. The data presented in this report were first published as part of A.W.D.'s dissertation.

LITERATURE CITED

BALLARD, J. W. O., and N. J. GALWAY, 1995 Assessing evolutionary hypotheses generated from mitochondrial DNA: inferences from Drosophila, in the Nicholson Centenary Meeting Proceedings, edited by R. FLOYD and A. SHEPHARD. Commonwealth Science and Industrial Research Organization, Canberra, Australia (in press).

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

CACCONE, A., G. D. AMATO and J. R. POWELL, 1988 Rates and patterns of scnDNA and mtDNA divergence within the *Drosophila melanogaster* subgroup. Genetics **118:** 671–683.

CACCONE, A., E. N. MORIYAMA, J. M. GLEASON, L. NIGRO and J. R. POWELL, 1996 A molecular phylogeny for the *Drosophila melanogaster* subgroup. Mol. Biol. Evol. (in press).

CARIOU, M.-L., 1987 Biochemical phylogeny of the eight species in the *Drosophila melanogaster* subgroup, including *D. sechellia* and *D. orena.* Genet. Res. **50:** 181–185.

COYNE, J. A., and B. CHARLESWORTH, 1986 Location of an X-linked factor causing sterility in male hybrids of *Drosophila simulans* and *Drosophila mauritiana.* Heredity **57:** 243–246.

COYNE, J. A., and B. CHARLESWORTH, 1989 Genetic analysis of X-linked sterility in hybrids between three sibling species of Drosophila. Heredity **62:** 97–106.

COYNE, J. A., and M. KREITMAN, 1986 Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia.* Evolution **40:** 673–691.

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

DAVIS, A. W., and C.-I. WU, 1996 The broom of the sorcerer's apprentice: the fine structure of a chromosomal region causing reproductive isolation between two sibling species of Drosophila. Genetics **143:** 1287–1298.

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

FISCHMAN, J., 1996 Evidence mounts for our African origins—and alternatives. Science **271:** 1364.

HEY, J., and R. M. KLIMAN, 1993 Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol. Biol. Evol. **10:** 804–822.

HILTON, H., R. M. KLIMAN and J. HEY, 1994 Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* species complex. Evolution **48:** 1900–1913.

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.

JOHNSON, N. A., D. E. PEREZ, E. L. CABOT, H. HOLLOCHER and C.-I. WU, 1992 A test of reciprocal *X-Y* interactions as a cause of hybrid sterility in Drosophila. Nature **358:** 751–753.

JOHNSON, N. A., H. HOLLOCHER, E. NOONBURG and C.-I. WU, 1993 The effects of interspecific *Y* chromosome replacements on hybrid sterility within the *Drosophila simulans* clade. Genetics **135:** 443–453.

KING, M. C., and A. C. WILSON, 1975 Evolution at two levels in humans and chimpanzees. Science **188:** 107–116.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

KLIMAN, R. M., and J. HEY, 1993 DNA sequence variation at the *period* locus within and among species of the *Drosophila melanogaster* complex. Genetics **133:** 375–387.

LACHAISE, D., M.-L. CARIOU, J. R. DAVID, F. LEMEUNIER, L. TSACAS *et al.,* 1988 Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22:** 159–225.

LOUIS, J., and J. R. DAVID, 1986 Ecological specialization in the *Drosophila melanogaster* species subgroup: a case study of *D. sechellia.* Acta OEcologica **7:** 215–229.

NAVEIRA, H. F., 1992 Location of *X*-linked polygenic effects causing hybrid sterility in male hybrids of *Drosophila simulans* and *D. mauritiana.* Heredity **68:** 211–217.

NEI, M., 1987 *Molecular Evolutionary Genetics.* Columbia University Press, New York.

ORR, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139:** 1805–1813.

PALOPOLI, M. F., and C.-I. WU, 1994 Genetics of hybrid male sterility between *Drosophila* sibling species: A complex web of epistasis is revealed in interspecific studies. Genetics **138:** 329–341.

PEREZ, D. E., C.-I. WU, N. A. JOHNSON and M.-L. WU, 1993 Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA marker-assisted mapping and characterization of a hybrid-male sterility gene, *Odysseus (Ods).* Genetics **134:** 309–318.

R'KHA, S., P. CAPY and J. R. DAVID, 1991 Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological,

behavioral, and genetical analysis. Proc. Natl. Acad. Sci. USA **88:** 1835–1839.

SATTA, Y., and N. TAKAHATA, 1990 Evolution of *Drosophila* mitochondrial DNA and the history of the *melanogaster* subgroup. Proc. Natl. Acad. Sci. USA **87:** 9558–9562.

SOLIGNAC, M., and M. MONNEROT, 1986 Race formation, speciation, and introgression within the three homosequential species *Drosophila simulans, D. mauritiana,* and *D. sechellia* inferred from mitochondrial DNA analysis. Evolution **40:** 532–539.

TAKAHATA, N., 1989 Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics **122:** 957–966.

TISHKOFF, S. A., E. DIETZSCH, W. SPEED, A. J. PAKSTIS, J. R. KIDD *et al.,* 1996 Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science **271:** 1380–1387.

WILSON, A. C., L. R. MAXSON, and V. M. SARICH, 1974a Two types of molecular evolution. Evidence from studies of interspecific hybridization. Proc. Natl. Acad. Sci. USA **71:** 2843–2847.

WILSON, A. C., V. M. SARICH, and L. R. MAXSON, 1974b The importance of gene arrangement in evolution: evidence from the rates of chromosomal, protein, and anatomical evolution. Proc. Natl. Acad. Sci. USA **71:** 3028–3030.

WU, C.-I., and A. W. DAVIS, 1993 Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. Am. Nat. **142:** 187–212.

WU, C.-I., D. E. PEREZ, A. W. DAVIS, N. A. JOHNSON, E. L. CABOT *et al.,* 1993 Molecular genetic studies of postmating reproductive isolation in Drosophila, pp. 191–203 in *Mechanisms of Molecular Evolution,* edited by N. TAKAHATA and A. G. CLARK. Sinauer Associates, Sunderland, MA.

Communicating editor: W. F. EANES