A Comprehensive Study of Genic Variation in Natural Populations of *Drosophila melanogaster*. VI. Patterns and Processes of Genic Divergence Between *D. melanogaster* and Its Sibling Species, *Drosophila simulans*

M. Choudhary,[1] M. B. Coulthart[2] and R. S. Singh

Department of Biology, McMaster University, Hamilton, Ontario, Canada L8S 4K1

Manuscript received January 28, 1991
Accepted for publication December 26, 1991

ABSTRACT

We present here an extensive set of data on allelic differences between homologous proteins of *Drosophila melanogaster* and its sibling species, *Drosophila simulans*, obtained by nondenaturing one-dimensional, and denaturing two-dimensional gel electrophoresis. The data suggest that, for these two species, (1) approximately 10% of protein-coding loci have no alleles in common in our sample, (2) the extent of genic variation at a locus (mean heterozygosity) within a species is not correlated with the extent of divergence (Nei's genetic distance) at that locus between species, and (3) significant heterogeneity of divergence rates exists for different structural/functional classes of loci. These results are discussed in the context of the dynamics of genetic variation within and between species.

THE analysis of genetic differences between closely related species has historically been motivated by two basic types of questions. First, one would like to know *how many* and *what kinds* of genes are involved in the initial establishment of species as independent lineages, and in their subsequent phyletic divergence. Second, it is essential to try to understand the population-genetic processes underlying the genetic differences that are found between species.

The classical approach to the first type of question was pioneered in Drosophila, through the segregational analysis of genetic determinants of particular trait differences in interspecies crosses, by DOBZHANSKY (1936), and has been continued by contemporary workers (ZOUROS 1981; COYNE 1983, 1984, 1985; ORR and COYNE 1989). This approach is yielding important insights into the genetics of species differences (COYNE and ORR 1989). However, it suffers from two main practical weaknesses—its limited genetic resolution (since phenotypic traits and chromosome segments are analyzed, rather than individual genes or their products), and the requirement that the species be at least partially interfertile. A more serious limitation, though, is the fact that only specific classes of traits can be studied. This means that considerable detailed biological knowledge is required before genetic analysis can begin. It also means that the relevance of the specific genetic analysis to the general question of how many and what kinds of genes are involved in species differentiation depends on how astutely the choice of traits has been made.

These limitations are effectively complemented by a more recent approach to speciation genetics, one that employs molecular comparisons in an effort to broadly characterize amounts and patterns of interspecific differentiation in the genome. The original applications of this approach in Drosophila relied on data from electrophoretic mobility difference in homologous proteins (reviewed by LEWONTIN 1974; AYALA 1975; THROCKMORTON 1977). DNA restriction site and sequence data have also been used in this context (LEIGH BROWN and ISH-HOROWICZ 1981; LANGLEY, MONTGOMERY and QUATTLEBAUM 1982; COHN, THOMPSON and MOORE 1984; BODMER and ASHBURNER 1984; AQUADRO, LADO and NOON 1988), but so far studies of this kind are limited, for practical reasons, to single loci. Clearly, molecular methodology has its weaknesses and limitations as well: most importantly, it is usually very difficult to connect any general or specific features of molecular divergence with phenotypic traits that have been judged to contribute to species formation or divergence. What seems obvious, therefore, is that neither molecular nor nonmolecular approaches are sufficient in themselves, and that both will be necessary to fully answer the first type of question.

The second type of question, concerning population genetic mechanisms, is approachable only with molecular data (LEWONTIN 1974). In point of fact, however, most work in molecular population genetics to date has been directed toward understanding the causes of genetic polymorphism within species (*e.g.*, see SINGH and RHOMBERG 1987b; KREITMAN 1983, 1987; LANGLEY and AQUADRO 1987; RILEY, HALLAS and LEWON-

[1] Present address: Department of Ecology and Evolutionary Biology, Rice University, Houston, Texas 77251.
[2] Present address: Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7.

TABLE 1

Number of shared (common) and unshared (unique) electromorphs and genetic identity at allozyme loci in natural populations of *D. melanogaster* and *D. simulans*

Enzymes or proteins	No. of electromorphs				Genetic identity (mean ± SE × 10^2)[a]
	Unique to *D. melanogaster*	Unique to *D. simulans*	Common	Total	
Chromosome *I(X)*					
Carbonic anhydrase-3 (CA-3)	0	2	3	5	0.872 ± 1.094
Diaphorase-7 (DIA-7)	3	0	1	4	0.999 ± 0.005
Fumarase (FUM)	0	0	1	1	1
Glucose-6-phosphate dehydrogenase (G-6PD)	2	0	1	3	0.82 ± 1.014
Glutamate pyruvate transaminase (GPT)	0	1	3	4	0.404 ± 1.465
Hexokinase-3 (HEX-3)	0	0	1	1	1
Hydroxyacid dehydrogenase (HAD)	1	1	0	2	0
6-Phosphogluconate dehydrogenase (6PGD)	0	0	3	3	0.260 ± 1.367
Larval protein-11 (Pt-11)	3	0	5	8	0.977 ± 0.276
Alcohol dehydrogenase (ADH)	2	1	0	3	0
α-Amylase (α-AMY)	5	2	0	7	0
Dipeptidase-3 (DIP-3)	2	0	1	3	0.990 ± 0.018
β-Galactosidase (β-GAL)	0	2	1	3	0.990 ± 0.020
Glutamate oxaloacetate transaminase-1 (GOT-1)	1	1	0	2	0
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	0	0	1	1	1
α-Glycerophosphate dehydrogenase (α-GPD)	0	0	2	2	0.968 ± 0.375
Hexokinase-1 (HEX-1)	1	0	3	4	0.970 ± 0.278
α-Glycerophosphate oxidase (α-GPO)	0	0	1	1	1
Malate dehydrogenase-1 (MDH-1)	1	0	1	2	0.999 ± 0.004
Phenol oxidase (PHOX)	2	0	2	4	0.999 ± 0.005
Phosphoglucose isomerase (PGI)	0	0	1	1	1
3-Phosphoglycerate kinase (PGK)	1	1	0	2	0
Larval protein-10 (Pt-10)	0	0	6	6	0.196 ± 1.027
Succinate dehydrogenase (SDH)	0	2	3	5	0.984 ± 0.199
Sucrase (SUCR)	1	1	0	2	0
Trehalase (TREH)	0	0	4	4	0.616 ± 1.609
Chromosome *III*					
Acetylcholine esterase (ACE)	0	0	1	1	1
Acid phosphatase (ACPH)	0	4	3	7	0.911 ± 0.851
Aldehyde oxidase (AO)	5	0	4	9	0.246 ± 1.333
Aldolase (ALD)	0	0	1	1	1
Alkaline phosphatase (APH)	0	0	3	3	0.069 ± 0.649
Arginine kinase-1 (AK-1)	0	0	1	1	1
Catalase (CAT)	0	0	1	1	1
Dipeptidase-1 (DIP-1)	0	3	1	4	0.854 ± 1.194
Dipeptidase-2 (DIP-2)	1	1	0	2	0
Esterase-C (EST-C)	0	1	4	5	0.707 ± 1.919
Esterase-6 (EST-6)	0	2	4	6	0.859 ± 1.557
α-Fucosidase (α-FUC)	1	2	2	5	0.975 ± 0.587
Glucose oxidase (GO)	2	0	2	4	0.990 ± 0.132
β-Glucoronidase (β-GUS)	0	1	1	2	0.999 ± 0.008
Isocitrate dehydrogenase-1 (IDH-1)	1	3	1	5	0.088 ± 0.080
Isocitrate dehydrogenase-2 (IDH-2)	0	0	1	1	1
Leucine aminopeptidase-D (LAP-D)	2	0	1	3	0.327 ± 1.576
Leucine aminopeptidase-A (LAP-A)	2	0	1	3	0.917 ± 0.733
Malate dehydrogenase-2 (MDH-2)	1	0	1	2	0.999 ± 0.018
Malic enzyme (ME)	0	0	1	1	1
Octanol dehydrogenase (ODH)	0	0	2	2	0.999 ± 0.024
Phosphoglucomutase (PGM)	1	1	3	5	0.997 ± 0.069
Larval protein-9 (Pt-9)	2	1	2	5	0.981 ± 0.023
Larval protein-15 (Pt-15)	0	1	4	5	0.952 ± 0.056
Sorbitol dehydrogenase-1 (SODH-1)	0	1	3	4	0.999 ± 0.008
Superoxide dismutase-1 (SOD-1)	1	0	1	2	0.997 ± 0.056
Xanthine dehydrogenase (XDH)	3	0	5	8	0.948 ± 0.075
Autosomes					
Aconitase-1 (ACON-1)	3	0	1	4	0.921 ± 0.060
Adenylate kinase-1 (ADK-1)	2	0	1	3	0.999 ± 0.002

Enzymes or proteins	No. of electromorphs: Unique to *D. melanogaster*	Unique to *D. simulans*	Common	Total	Genetic identity (mean ± SE × 10^2)[a]
Adenylate kinase-2 (ADK-2)	0	0	1	1	1
Carbonic anhydrase-1 (CA-1)	0	1	3	4	0.926 ± 1.100
Carbonic anhydrase-2 (CA-2)	0	3	3	6	0.901 ± 0.093
Diaphorase-1 (DIA-1)	2	0	1	3	0.995 ± 0.062
Diaphorase-2 (DIA-2)	1	1	4	6	0.987 ± 0.283
Glucose dehydrogenase (GDH)	2	0	3	5	0.768 ± 1.394
α-Glucosidase-4 (α-GLU-4)	2	0	3	5	0.389 ± 1.442
Glutamate dehydrogenase (GLU-D)	2	0	3	5	0.992 ± 0.192
α-Mannosidase-4 (α-MAN-4)	2	0	1	3	0.046 ± 0.044
Mannose phosphate isomerase (MPI)	1	0	3	4	0.027 ± 0.397
Unmapped					
β-Acetyl glucosaminidase (HEXOS-1)	0	0	1	1	1
Arylsulfatase (ARS)	1	1	0	2	0
Diaphorase-6 (DIA-6)	0	0	1	1	1
Esterase-9 (EST-9)	1	0	4	5	0.759 ± 1.675
Esterase-10 (EST-10)	1	1	4	6	0.996 ± 0.071
α-Glucosidase-2 (α-GLU-2)	0	0	1	1	1
α-Glucosidase-3 (α-GLU-3)	0	2	1	3	0.999 ± 0.003
Hexokinase-2 (HEX-2)	0	0	1	1	1
Lactate dehydrogenase (LDH)	0	3	1	4	0.987 ± 0.163
Leucine aminopeptidase-6 (LAP-6)	3	0	1	4	0.898 ± 0.871
α-Mannosidase-1 (MAN-1)	0	0	1	1	1
α-Mannosidase-2 (MAN-2)	0	0	1	1	1
α-Mannosidase-3 (MAN-3)	0	0	1	1	1
Sorbitol dehydrogenase-2 (SODH-2)	0	0	1	1	1
Superoxide dismutase-2 (SOD-2)	1	0	1	2	0.999 ± 0.001
Larval protein-1 (Pt-1)	0	2	4	6	0.855 ± 1.184
Pt-2	3	0	1	4	0.999 ± 0.009
Pt-3	0	0	1	1	1
Pt-4	0	0	1	1	1
Pt-5	0	0	1	1	1
Pt-6	1	0	1	2	0.999 ± 0.023
Pt-7	0	0	1	1	1
Pt-8	0	0	1	1	1
Pt-12	2	0	1	3	0.999 ± 0.001
Pt-13	3	0	1	4	0.999 ± 0.018
Pt-16	0	1	1	2	0.999 ± 0.003
Adult protein-7A (Pt-7A)	1	0	1	2	0.998 ± 0.029
Pt-9A	2	0	2	4	0.999 ± 0.005
Pt-10A	1	0	1	1	1
Pt-15A	0	0	2	2	0.999 ± 0.006
Pt-26	1	0	1	2	0.962 ± 0.403
Pt-37	1	0	1	2	0.816 ± 1.284
14 other adult protein loci	0	0	1	1	1

[a] NEI's genetic identity.

TIN, 1989). Only in relatively few cases have mechanistic questions been asked about species divergence; for instance, whether divergence patterns are consistent with expectations of the neutral mutation-random drift model of molecular evolution (SKIBINSKI and WARD 1982; CHAKRABORTY and HEDRICK 1983; WARD and SKIBINSKI 1985; HUDSON, KREITMAN, and AGUADÉ 1987; AQUADRO, LADO and NOON 1988). The general conclusion from this limited number of studies has been that no strong inconsistency exists between the data and the neutral theory. (Except at the *Drosophila Adh* locus—see HUDSON, KREITMAN and AGUADÉ 1987).

Here we report the largest set of molecular data assembled to date on genetic divergence between *Drosophila melanogaster* and its sibling species, *Drosophila simulans* (for patterns of genetic variation among populations in each of these species, see CHOUDHARY and SINGH 1987a,b; and SINGH and RHOMBERG 1987a,b). The data were obtained by comparative electrophoretic analysis of the products of 112 structural loci encoding homologous soluble en-

zymes and other proteins. After presenting a general description of interlocus patterns of divergence for this sample of loci, we evaluate these patterns in light of the two types of questions addressed at the beginning of the Introduction. Four main conclusions emerge. First, the rate of divergence between *D. melanogaster* and *D. simulans*, as estimated by protein electrophoresis, appears to be lower than predicted on the basis of independent estimates of rates of base substitution in coding-DNA sequences. This indicates a strong constraining influence of purifying natural selection on the rate of electrophoretic divergence in this species pair. Second, a small but distinct minority of protein-coding loci apparently have diverged to the point that no alleles are shared by these two species. This type of divergence at a locus has been observed only rarely in previous studies, which have compared relatively small numbers of protein-coding loci between closely related species. Third, there appears to be significant heterogeneity of divergence rates for different structural/functional classes of loci, between *D. melanogaster* and *D. simulans*. And fourth, there is no correlation between the amount of genetic variation at a locus within either of the two species and the degree of divergence between the species at that locus. This result is not predicted by the neutral theory (KIMURA 1983).

MATERIALS AND METHODS

Fly stocks: The *D. melanogaster* stocks used in this study originated in 15 globally dispersed populations, as reported in detail elsewhere (SINGH and RHOMBERG 1987a,b). The populations studied are: Ottawa, Canada (20 lines); Hamilton, Ontario, Canada (40); Amherst, Massachusetts (30); Brownsville, Texas (30); La Plata, Argentina (9); Sweden (10); Ukraine (11); Central Asia, U.S.S.R. (10); Villeurbanne, France (20); Benin, West Africa (28); Central Africa (17); Seoul, Korea (33); Taiwan (20); Ho-Chi-Minh City, Vietnam (20); and Fairfield, Australia (25). *D. simulans* stocks came from five populations in Europe and Africa (SINGH, CHOUDHARY and DAVID 1987). The populations are South France (55), Tunisia (52), Congo (45), Cape Town, South Africa (32) and Seychelle Island (26). The stocks were maintained as isofemale lines at 20° on standard cornmeal medium.

Gel electrophoresis and scoring of genetic variants: Nondenaturing, one-dimensional polyacrylamide gel electrophoresis was carried out on vertical slab gels following standard methods (PRAKASH, LEWONTIN and HUBBY 1969). The enzyme staining procedures were adapted from HARRIS and HOPKINSON (1976) or SICILIANO and SHAW (1976), or were as referred to by SINGH and RHOMBERG (1987a). Abundant soluble proteins were stained with Coomassie brilliant blue R, or with silver as described by COULTHART and SINGH (1988a). Electrophoretic variants were designated according to the nomenclature of SINGH, HICKEY and DAVID (1982).

For most loci the identity of allozyme variants between species was checked under one set of electrophoretic conditions only (usually 5% gel and pH 8.9). This single criterion for comparison was adopted partly for practical reasons and partly because an earlier exhaustive study of 11 loci in these two species had revealed little electrophoretically cryptic variation (SINGH, CHOUDHARY and DAVID 1987). Also for practical reasons, the identity of alleles at highly polymorphic (*i.e.*, multiallelic loci) was checked for the major alleles; the less frequent alleles, for purposes of the present comparison, were assumed to be identical between the two species. For these reasons, the estimates of genetic distance and genetic identity, discussed below, must be treated as minimum and maximum values, respectively.

Methods for two-dimensional electrophoresis, used to obtain the data on interspecies divergence of male reproductive tract proteins listed in Table 2, were as described by COULTHART and SINGH (1988a). The criteria used to determine interspecies homologies of protein spots on these gels is discussed in detail by COULTHART and SINGH (1988b).

RESULTS

Joint distribution of genetic diversity: We compared *D. melanogaster* and *D. simulans* genetically by analyzing the electrophoretic mobilities of homologous proteins via one-dimensional gel electrophoresis under nondenaturing conditions. Of the 112 structural loci estimated (by inspection of gel patterns) to encode these proteins, 77 encode soluble enzymes and 35 encode abundant soluble proteins of unknown function (Table 1). The sets of electromorphs present in the two species are listed for the individual homologous loci in Table 1. As can be seen by inspection of the table, all possible degrees of interspecies overlap occur between the electromorph sets present at a locus. Both *D. melanogaster* and *D. simulans* carry many species-specific electromorphs; these electromorphs are distributed over a total of 65 loci.

The distribution of average within-population frequencies for all electromorphs identified is plotted in Figure 1, with the distribution of frequencies for species-specific electromorphs superimposed. In both species, most electromorphs are rare, with average within-population frequencies less than 0.01. The species-specific electromorph sets are also composed mostly of rare electromorphs. At the other end of the electromorph-frequency scale, in each species, is a distinct class of very common electromorphs (frequencies >0.95). Relatively few electromorphs occur at intermediate frequencies. *D. melanogaster*, with a total of 224 electromorphs encoded by the 112 loci, harbors a somewhat greater genic diversity than does *D. simulans*, with a total of 189 electromorphs. The differences in genic diversity between the two species are not simply the result of their disproportional sampling in the present study (15 populations with a mean of 21 lines per population in *D. melanogaster* and five populations with a mean of 42 lines per populations in *D. simulans*); these differences would persist even if we had done a more widely distributed geographic sampling of *D. simulans* (see SINGH, CHOUDHARY and DAVID 1987), the reason being that *D. simulans* has less geographic differentiation than *D. melanogaster*.

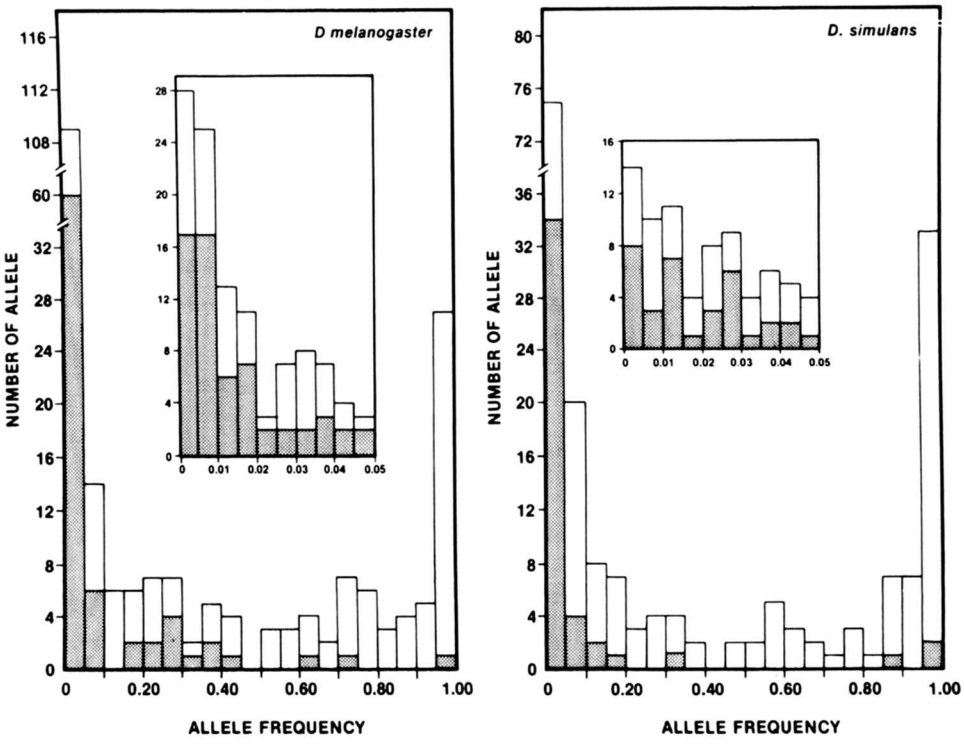

FIGURE 1.—Frequency distribution (in entire sample) of unique (shaded) and total (open) alleles at polymorphic loci in *D. melanogaster* and *D. simulans*. Insets: frequency distribution of alleles with frequency <5%.

To quantify the pattern of interspecies overlap for individual loci, we first considered the joint interspecies distribution of monomorphism and polymorphism for homologous loci. With the criterion for a polymorphic locus being a frequency of 99% or less for the most common allele, 47 loci are monomorphic in both species, and 37 polymorphic in both species; 21 are polymorphic only in *D. melanogaster* and 7 polymorphic only in *D. simulans*. On a more stringent criterion of polymorphism (most common allele at a frequency of 95% or less), these numbers become 62, 23, 16 and 11. The increase, from 7 to 11, in the number of loci polymorphic only in *D. simulans* under the 95% criterion is due to the fact that many loci that are polymorphic in both species under the 99% criterion remain polymorphic in *D. simulans* but are classed as monomorphic in *D. melanogaster* under the more stringent 95% criterion. On either criterion, there is a high degree of association between species with respect to the monomorphic or polymorphic state of a locus, as measured by 2 × 2 contingency tests (species × monomorphic/polymorphic state $- X_3^2 =$ 30.29 and 49.20, respectively, for 99% and 95% criteria; $p < 0.005$). At most of the loci for which one species is monomorphic and the other polymorphic, one of the species carries an allele belonging to the polymorphic set found in the other.

NEI (1987) genetic identity (I) and its standard error (NEI 1978; NEI and ROYCHOUDHURY 1974) were calculated for each locus (see Table 1). When the frequency distribution of these I values is plotted (Figure 2), there are clearly two major clusters of loci: one with I between 0.95 and 1.0, and another with I between 0 and 0.05. Of the 77 loci in the former class, 40 are identically monomorphic between *D. melanogaster* and *D. simulans*. Of the remaining 37 loci, 25 have I values of 0.99 or greater, and another 10 have values between 0.95 and 0.99, despite the presence at each of these loci of species-specific alleles. Thus, at a majority of loci interspecific divergence is undetectable or weak, involving only rare alleles. Spread more or less evenly along the scale of genetic identity is another, relatively ill defined group of loci with intermediate I values. In principle, values of I can vary because of differences in either allele frequency or allele composition at a locus. However, no significant correlation ($r^2 = -0.186$, $P > 0.05$) was found between I and the number of alleles at a locus. This result contrasts with what one might expect if the qualitative allele compositions of individual loci were major contributing factors to divergence; differences in allele frequency are thus emphasized as the primary factor underlying low-to-moderate I values in this sample of loci.

Finally, a relatively small, but distinct, class of loci occurs at the low end of the genetic identity scale. Eight loci (*Adh*, *α-Amy*, *Ars*, *Dip-2*, *Got-1*, *Had*, *Pgk*, and *Sucr*,) share no alleles ($I = 0$) between *D. melanogaster* and *D. simulans*. Interestingly, 6 of these 8 loci (all but *Adh* and *α-Amy*) are monomorphic within each species, and are thus "alternately fixed" between species. The difference between this ratio (6:2) and the ratio, in the sample of the number of loci monomorphic in both species to the number of loci that are

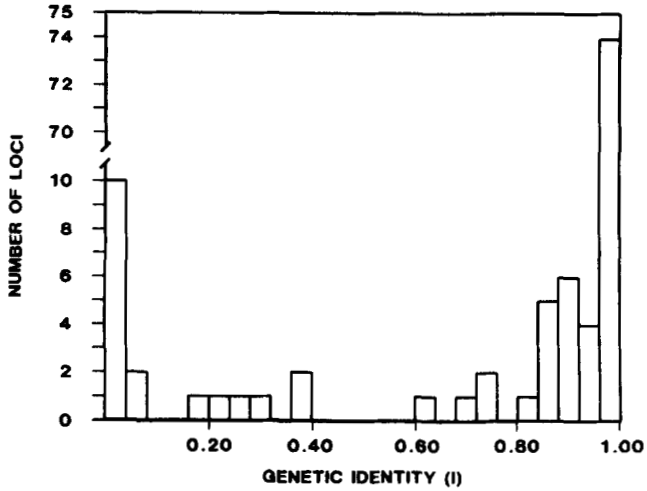

FIGURE 2.—Frequency distribution of single-locus Nei genetic identity (I) for 112 loci between *D. melanogaster* and *D. simulans*. The genetic identity (I) is defined as $I = Jxy/(Jx \cdot Jy)^{1/2}$, where Jx, Jy and Jxy are, respectively, the probability that two randomly sampled alleles from species x, y, and x and y are identical.

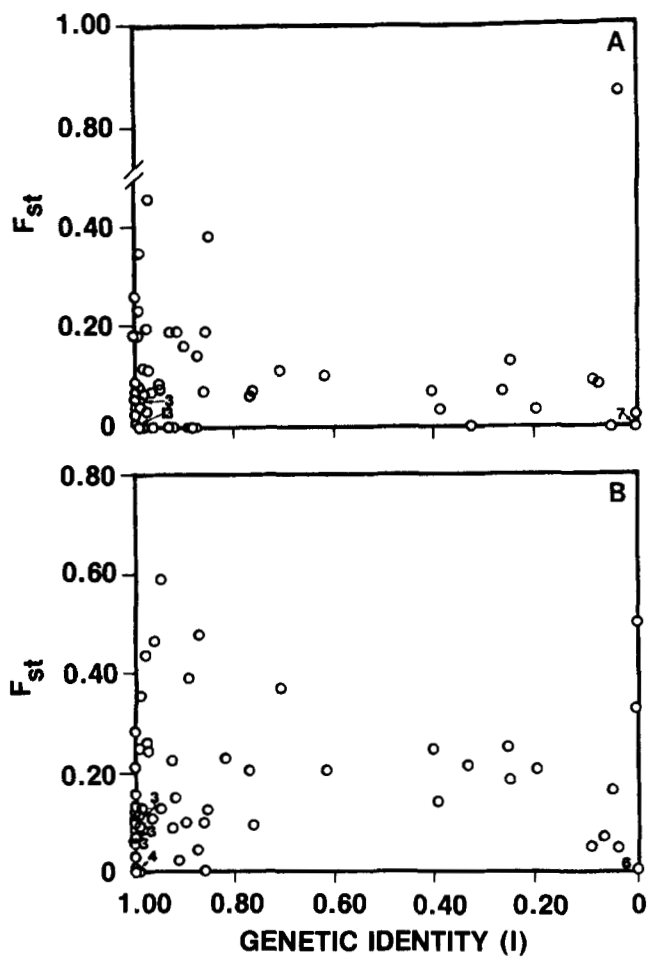

FIGURE 3.—Graphs showing relationship between genetic identity (I) and fixation index (F_{ST}) in *D. simulans* (A) and *D. melanogaster* (B). F_{ST} is a measure of geographic structure within species, and is defined as $F_{ST} = (H_T - H_S)/H_T$, where H_T and H_S are heterozygosity in the total and subpopulations, respectively. The numbers within graphs indicate number of loci with similar values.

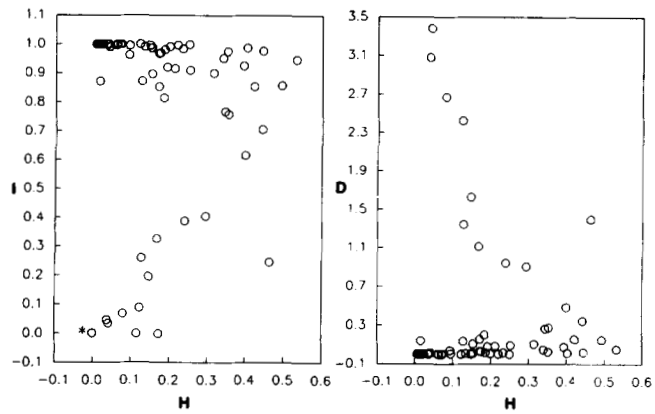

FIGURE 4.—A graph of NEI genetic identity (I) and genetic distance (D) against mean heterozygosity for the 64 polymorphic loci in *D. melanogaster* and *D. simulans*. $I = 0.823 - 0.116H$, and $D = 0.194 + 0.403H$. Note that 8 loci are completely diverged ($I = 0$) and 6 of these are alternatively fixed (indicated by asterisk), and 40 loci identically monomorphic between the two species.

not, is not statistically significant ("total sample" ratio = 47:65; $X_1^2 = 3.30$; $P = 0.07$).

Population structure, heterozygosity and divergence: One may ask whether the extent of interpopulation differentiation at a locus within a species is positively correlated with the degree of divergence between species at that locus. Such a positive correlation would be consistent with the notion that the same evolutionary process (or set of processes) acts as the major cause of genetic variation and change both within and between species. In Figure 3, A and B, we have plotted Wright's (1951) fixation index (f_{ST}), a measure of interpopulation differentiation within species, against NEI's genetic identity (I), for individual homologous loci between *D. melanogaster* and *D. simulans*. The data clearly show the lack of an overall positive correlation between F_{ST} and I for either species (*D. melanogaster*: $r^2 = 0.025$, *D. simulans* $r^2 = 0.032$). In fact, the loci with the highest F_{ST} values tend to cluster at the low end of the scale of divergence (high I); there is also a near-absence of loci with high F_{ST} and low I. In *D. melanogaster* for loci with major polymorphism (*i.e.*, where the second most common allele is, decided arbitrarily, ≥20% in the entire sample) there is a suggestion of negative but nonsignificant relationship ($r^2 = 0.33$, $P > 0.05$) between within- (F_{ST}) and between-species ($1 - I$) measures of divergence.

An analogous pattern is seen in Figure 4, wherein NEI's genetic identity (I) and genetic distance (D) are plotted against mean heterozygosity for each locus (H). Here, also, there is a lack of positive correlation between interspecies (D) and intraspecies (H) diversity. The correlation coefficients between I and H ($r^2 = -0.049$, $P > 0.05$) and D and H ($r^2 = 0.089$, $P > 0.05$) are nonsignificant. The data shown in Figure 4

are clearly inhomogeneous and there is an absence of loci showing both high *D* and high *H*.

Divergence in different classes of loci: The observation of a lack of allele sharing between *D. melanogaster* and *D. simulans* at a small number of loci encoding soluble enzymes and other proteins, as described above, raises the question whether such loci occur widely in the genome. We have recently carried out a preliminary comparative study of proteins found in male reproductive tract tissues of *D. melanogaster* and *D. simulans*, using high resolution two-dimensional gel electrophoresis (2DE) (COULTHART and SINGH 1988b). In Table 2 we summarize features of polymorphism and divergence within and between *D. melanogaster* and *D. simulans*, for the various sets of proteins which have been surveyed in the present study by one-dimensional gel electrophoresis (1DE) and in the 2DE study just mentioned. Several interesting points arise from inspection of this table.

First, the overall proportions of loci with no shared alleles are similar for the soluble enzymes and other proteins surveyed by 1DE (7.0%) and the abundant solubilized proteins surveyed by 2DE (13.2%). The difference is reduced if the 1DE set is restricted to soluble enzymes (10.1% completely diverged loci). This suggests by extrapolation that a fraction in the neighborhood of 10% of protein-coding loci in the genomes of *D. melanogaster* and *D. simulans* may share no alleles between the two. We note also that most loci that share no alleles are in fact fixed for alternative alleles in the two species.

The second point concerns the difference in proportions of loci with no shared alleles seen between, on one hand, male reproductive tract proteins that are found in both testis and gland tissue of the tract ("common" and "testis-elevated" proteins) and, on the other hand, those proteins found in testis but not gland tissue, or vice versa ("testis-specific" and "gland-specific" proteins, respectively). In the former two classes 9.8% and 3.5% of the loci, respectively, share no alleles in common, and in the latter two classes 17.5% and 26.1% of the loci, respectively, share no alleles in common. The figure of 9.8% loci sharing to alleles in common for "common" proteins is strikingly close to the 10.1% found with soluble enzymes.

Last, we draw attention to interspecific divergence of a different type, readily detectable with 2DE (where proteins are visualized by direct staining) but not as readily with 1DE (where proteins are mostly visualized by methods that depend on enzyme activity). Overall, 20–30% of *D. melanogaster* or *D. simulans* male reproductive tract protein loci (studied by 2DE) show "presence/absence" differences between *D. melanogaster* and *D. simulans* (COULTHART and SINGH 1988b). In these cases, a protein present in one of the two species lacks a detectable homologous spot in the other species; we have proposed (COULTHART and SINGH 1988b) that these seemingly species-specific proteins actually represent large interspecific differences in levels of gene expression. The proportions of loci showing such differences are given for the various tissue-distribution classes of male reproductive tract proteins in Table 2. Although presence/absence differences were seen frequently (18.3–39.1%) in all tissue-distribution classes of proteins, the class of proteins with a "common" expression pattern did carry the smallest proportion of loci (18.3%) with differences of this type.

DISCUSSION

Variation, divergence, selection and the neutral theory: Between the gene pools of *D. melanogaster* and *D. simulans* there is a large degree of nonoverlap on a qualitative basis, with *D. melanogaster* and *D. simulans* carrying totals of 81 and 45 identifiably species-specific alleles, respectively, in a sample of 112 loci. However, because of their low frequencies, the quantitative contribution of the great majority of alleles in the nonoverlapping subsets is slight. This overall genetic similarity is again reflected in the pattern of divergence at individual loci. Most gene products sampled are electrophoretically indistinguishable or nearly so, with a scattering of genes showing primarily quantitative divergence in allele frequency. It is only with respect to a small minority (*ca.* 10%) of structural loci that *D. melanogaster* and *D. simulans* share no alleles; however, these loci have a tendency to be monomorphic within each species.

The high degree of genetic similarity between *D. melanogaster* and *D. simulans* might be expected if the two species arose relatively recently from the same ancestral gene pool. Another explanation, and one that need not completely exclude this historical one, invokes similar kinds of natural selection on homologous loci in the two species.

Under the neutral mutation-random drift theory of molecular evolution (KIMURA 1983), the level of genic divergence between species is a function of neutral mutation rate, and time since divergence. NEI and TATENO (1975) modeled the divergence over time of two populations under neutral mutation and random drift, using computer simulation. The mean (0.832) of the distribution of genic identity between *D. melanogaster* and *D. simulans* in our study is similar to the mean (0.811) of NEI and TATENO's simulated distribution for two populations that have been isolated for $4N$ generations, where N is the effective population size. However, in the absence of independent estimates of effective population size or of chronological divergence time for *D. melanogaster* and *D. simulans*, it is impossible to test the goodness of fit of our data to the predictions of a model such as NEI and TA-

TABLE 2

Proportion of loci polymorphic and mean heterozygosity per population within species, and proportion of loci that lack any alleles in common between species of *D. melanogaster* and *D. simulans*

Enzymes and proteins	Species	No. of loci studied	Proportion of loci uniformly monomorphic	Proportion of loci polymorphic (mean ± SE)	Heterozygosity (mean ± SE)	Proportion of loci diverged	
						Homologous ($I = 0$)	Nonhomologous (showing +/−)
Enzymes and abundant soluble proteins (1DE)							
Enzymes	*D. melanogaster*	79[a]	0.418	0.412 ± 0.049	0.127 ± 0.018	0.101	– –
	D. stimulans	79[a]	0.443	0.329 ± 0.031	0.112 ± 0.016		
Larval (hemolymph) proteins	*D. melanogaster*	8[b]	0.375	0.550 ± 0.092	0.179 ± 0.069		
	D. stimulans	8[c]	0.250	0.650 ± 0.056	0.255 ± 0.081	0	– –
Larval (carcass) and adult proteins	*D. melanogaster*	30[a]	0.667	0.193 ± 0.057	0.037 ± 0.015		
	D. stimulans	27[c]	0.926	0.059 ± 0.020	0.004 ± 0.003	0	– –
Total (1DE)	*D. melanogaster*	117	0.470	0.420 ± 0.070	0.102 ± 0.014	0.070	– –
	D. stimulans	114	0.614	0.288 ± 0.022	0.094 ± 0.012		
Male reproductive tract proteins (2DE)[d]							
Common	*D. melanogaster*	110	0.927	0.041 ± 0.045	0.010 ± 0.008	0.098	0.183
	D. stimulans	131	0.954	0.038 ± 0.000	0.012 ± 0.005		
Testis-specific	*D. melanogaster*	74	0.878	0.101 ± 0.029	0.030 ± 0.007	0.175	0.234
	D. stimulans	87	0.896	0.092 ± 0.016	0.032 ± 0.010		
Gland-specific	*D. melanogaster*	26	0.769	0.173 ± 0.082	0.065 ± 0.033	0.261	0.391
	D. stimulans	25	0.680	0.280 ± 0.000	0.119 ± 0.019		
Testis-elevated	*D. melanogaster*	34	0.911	0.058 ± 0.042	0.023 ± 0.019	0.035	0.291
	D. stimulans	28	0.857	0.107 ± 0.051	0.028 ± 0.030		
Total (2DE)	*D. melanogaster*	244	0.893	0.080 ± 0.044	0.024 ± 0.010	0.132	0.231
	D. stimulans	271	0.900	0.085 ± 0.000	0.029 ± 0.001		

[a] Based on 9 populations.
[b] Based on 15 populations.
[c] Based on 5 populations.
[d] Based on 2 populations.

TENO's. A more empirically oriented approach to assess the role of selection can be taken as follows, with the aid of some heuristic calculations. The genetic distance between *D. melanogaster* and *D. simulans*, as measured by electrophoresis ($D = 0.179$), can be converted to chronological time (t) using the formula: $D/2\alpha$, where α is the rate of electrophoretically detectable amino acid substitution per year. Assuming $\alpha = 10^{-7}$ (NEI 1975, 1987), the divergence time obtained for *D. melanogaster* and *D. simulans* is 0.9 myr. We note that the value 0.8 myr is considerably lower than the range of values (2.0–3.5 myr) recently obtained for these two species by analysis of a compilation of DNA sequence data from protein-coding and noncoding regions (BODMER and ASHBURNER 1984; COHN, THOMPSON and MOORE 1984; EASTEAL and OAKESHOTT 1985; STEPHENS and NEI 1985).

To state this result differently: If one applies two independently derived scales of calibration (for the rates of nucleotide and amino acid substitution against absolute time) to the *D. melanogaster* and *D. simulans* data to estimate divergence time between the species, one obtains two discrepant results. Our intention here is not to attempt to resolve this discrepancy in order to obtain a more accurate estimate of divergence time, but rather to draw attention to its possible biological significance. Obviously, one or more parameters in the formulas may be incorrectly estimated; for instance, genetic divergence that was electrophoretically invisible in the current study may be revealed in future work. However, unless this effect is unexpectedly large (in fact the evidence is to the contrary, see CHOUDHARY and SINGH 1987a) it seems unlikely to nullify the discrepancy completely. The more likely prospect is that, for the loci we examined by electrophoresis in *D. melanogaster* and *D. simulans*, the actual value of α for *Drosophila*, is lower than expected. This suggests that natural selection may have been an important constraining factor lowering the rate of electrophoretic divergence of proteins between *D. melanogaster* and *D. simulans*. This putative constraint is also reflected in the prevalence of monomorphism among loci examined by electrophoresis (70% with 1DE, possibly 90% with 2DE; see COYNE 1982; LEWONTIN 1985; COULTHART and SINGH 1988a). It should also be noted that the present estimate of

genetic distance may *under*estimate the importance of natural selection in slowing down divergence, since other factors (*e.g.*, population bottlenecks) actually may have contributed in a positive way to the observed level of divergence, thus offsetting the effects of selective constraint. A complete resolution of the varying rates of divergence in proteins and DNA would require a comparison of silent and replacement changes in sequenced genes, which is outside the scope of this study.

A second, more surprising, kind of evidence for the influence of natural selection in interspecific divergence is constituted by our observation of lack of positive correlations between within-species polymorphism and between-species divergence. The neutral mutation-random drift theory predicts a positive, approximately linear relationship between heterozygosity and genetic distance (CHAKRABORTY, FUERST and NEI 1978; SKIBINSKI and WARD 1982; WARD and SKIBINSKI 1985). Population bottlenecks and small within-species samples are expected to give a nonzero intercept to a positively sloped plot of *D vs. H* [at loci with *major* polymorphisms] (CHAKRABORTY and HEDRICK 1983). On the basis of the worldwide population sample of *D. melanogaster* analyzed in this study and those of *D. simulans* in this study and elsewhere (SINGH, CHOUDHARY and DAVID 1987), gross underestimation of heterozygosity through sampling error seems unlikely. The absence of loci showing both high divergence and high F_{ST} or H, negates the idea that a positive correlation was originally present when the species became isolated, with subsequent decay to the point of undetectability. It seems much more likely in the case of *D. melanogaster* and *D. simulans* that such a relationship never existed.

We cannot exclude the possibility that the amounts of within-species variation we observe are not actually "equilibrium" values. In fact, recent evidence is consistent with the notion that most populations of *D. melanogaster* and *D. simulans* are not at evolutionary equilibrium, at least partly because of relatively recent worldwide colonization from a tropical African center of diversity (LEMEUNIER et al. 1985; CHOUDHARY and SINGH 1987a,b; HALE and SINGH 1987; SINGH, CHOUDHARY and DAVID 1987; LACHAISE *et al.* 1987). Again, however, it is not easy to understand how this would explain the lack of positive correlations between intraspecific variation and interspecific divergence. It will be interesting to see whether the same types of relationships between variation and divergence are found when species, *e.g.*, *Drosophila pseudoobscura*, that appear to be much closer to evolutionary equilibrium than are *D. simulans* and *D. melanogaster* (CHOUDHARY and SINGH 1987a), are compared with their close relatives. The comparison between different species pairs or groups, with respect to what relationships exist between parameters of variation and divergence, is something that has not yet been attempted with protein data. This approach should provide an interesting complement to that taken, for example, by SKIBINSKI and WARD (1982) and WARD and SKIBINSKI (1985). In their approach, data for a locus from different species are pooled to filter out species-to-species variation and to increase sample sizes so that overall relationships can be discerned. But there is no reason to expect that specific loci will exhibit the same degree and pattern of variation and divergence within and between all species examined. Thus, the weakness of this approach is that species-to-species variation in the loci contributing to correlations, or even in the shape of the relationship itself, may obscure a more widespread occurrence of the type of correlation we have observed. Studies of relationships between heterozygosity and divergence rates in closely related taxa can tell us more about the relative role of selective and non-selective processes in evolution than can divergence studies involving long-separated taxa (see GILLESPIE 1986a,b; 1988 for studies involving the latter approach).

One possible scenario is that balancing selection has lowered the average rate of divergence of polymorphic loci between species. In contrast, loci that show no alleles between the two species may be monomorphic because of tight constraint by purifying selection within species, while remaining capable of occasional allele substitution through directional selection. If the conditions for such allele substitutions coincided with the ecological conditions favoring interpopulation isolation, essentially monomorphic loci could play an important role in species formation (WRIGHT 1982).

Divergence in different classes of loci: With reference to the functional aspects of loci that might be involved in species formation and divergence, we were intrigued by the seeming tendency for hybrid male sterility to arise in the early stages of species differentiation in *Drosophila* (BOCK 1984). This prompted us to carry out a comparison of male reproductive tract proteins by 2DE between *D. melanogaster* and *D. simulans* (COULTHART and SINGH 1988b). The results bear on the present discussion in two ways. First, male reproductive tract proteins appear to be more divergent between *D. melanogaster* and *D. simulans* than are proteins of certain "non-reproductive" tissue such as imaginal wing discs. Whether this indicates preferential genetic involvement of male reproductive tract functions in *Drosophila* species formation and divergence, or something else—perhaps a difference between terminally differentiated tissues and embryonic ones—remains to be examined. The difference we have noted above between interspecific divergence in proteins with broad distributions within the male re-

productive tract and those with narrow distributions, and the similarity of divergence between the broadly distributed proteins of the reproductive tract and soluble enzymes, together suggest that multiple functional factors may be involved.

Second, an important aspect of 2DE, with its detection of proteins by direct staining, is that interspecies differences in protein amount can be readily detected. Such differences are very frequent among proteins of the Drosophila male reproductive tract, and most frequent among proteins with organ-specific distributions within the reproductive tract (COULTHART and SINGH 1988b). Thus, we raise the question whether the changes accompanying species formation are perhaps biased with respect to the *kinds* of genetic alterations that can occur. In comparison with electrophoretic data on protein structure, relatively few data are available on difference in expression of homologous proteins in closely related species, although important exceptions exist (DICKINSON 1980; DICKINSON, ROWNAN and BRENNAN 1984; PARKER, PHILLIP and WHITT 1985) (see review by DICKINSON 1988). Exploration of the role of such factors in genetic divergence could prove to be quite fruitful.

We would like to thank H. L. CARSON, R. C. LEWONTIN, S. K. JAIN and several anonymous reviewers for providing critical comments on various versions of this manuscript, and RANJIT CHAKRABORTY for his help with statistical computations. The research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to R.S.S.

LITERATURE CITED

AQUADRO, C. F., K. M. LADO and W. A. NOON, 1988 The *rosy* region of *Drosophila melanogaster* and *Drosophila simulans*. I. Contrasting levels of naturally occurring DNA restriction map variation and divergence. Genetics **119:** 875–888.

AYALA, F. J., 1975 Genetic differentiation during the speciation process. Evol. Biol. **8:** 1–78.

BOCK, I., 1984 Interspecific hybridization in the genus *Drosophila*. Evol. Biol. **18:** 41–70.

BODMER, M., and M. ASHBURNER, 1984 Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila*. Nature **309:** 425–430.

CHAKRABORTY, R., P. A. FUERST and M. NEI, 1978 Statistical studies on protein polymorphism in natural populations. II. Gene differentiation between populations. Genetics **88:** 367–390.

CHAKRABORTY, R., and P. W. HEDRICK, 1983 Heterozygosity and genetic distance of proteins. Nature **304:** 755.

CHOUDHARY, M., and R. S. SINGH, 1987a Historical effective size and the level of genetic diversity in *Drosophila melanogaster* and *Drosophila pseudoobscura*. Biochem. Genet. **25:** 41–51.

CHOUDHARY, M., and R. S. SINGH, 1987b A comparative study of genic variation in natural populations of *Drosophila melanogaster*. III. Variations in genetic structure and their causes between *Drosophila melanogaster* and its sibling species *Drosophila simulans*. Genetics **117:** 696–710.

COHN, V. H., M. A. THOMPSON and G. P. MOORE, 1984 Nucleotide sequence comparisons of the *Adh* gene in three Drosophilids. J. Mol. Evol. **20:** 31–37.

COULTHART, M. B., and R. S. SINGH, 1988a Low genic variation in male-reproductive-tract proteins of *Drosophila melanogaster* and *D. simulans*. Mol. Biol. Evol. **5:** 167–181.

COULTHART, M. B., and R. S. SINGH, 1988b High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*. Mol. Biol. Evol. **5:** 182–191.

COYNE, J. A., 1982 Gel electrophoresis and cryptic protein variation, pp. 1–32 in *Isozymes—Current Topics in Biological and Medical Research*, Vol. 6, edited by M. C. RATAZZI, J. G. SCANDALIOS and G. S. WHITT. Alan R. Liss, New York.

COYNE, J. A., 1983 Genetic basis of differences in genital morphology among three siblings species of *Drosophila*. Evolution **37:** 1101–1118.

COYNE, J. A., 1984 Genetic basis of male sterility in hybrids between two closely related species of *Drosophila*. Proc. Natl. Acad. Sci. USA **81:** 4444–4447.

COYNE, J. A., 1985 The genetic basis of Haldane's rule. Nature **314:** 736–738.

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

DICKINSON, W. J., 1980 Evolution of patterns of gene expression in Hawaiian picture-winged *Drosophila*. J Mol. Evol. **16:** 73–94.

DICKINSON, W. J., 1988 On the architecture of regulatory systems: evolutionary insights and implications. BioEssays **8:** 204–208.

DICKINSON, W. J., R. G. ROWAN and M. D. BRENNAN, 1984 Regulatory gene evolution: adaptative differences in expression of alcohol dehydrogenase in *Drosophila melanogaster* and *Drosophila simulans*. Heredity **52:** 215–225.

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

EASTEAL, S., and J. G. OAKESHOTT, 1985 Estimating divergence times of *Drosophila* species from DNA sequence comparisons. Mol. Biol. Evol. **2:** 87–91.

GILLESPIE, J. H., 1986a Natural selection and the molecular clock. Mol. Biol. Evol. **3:** 138–155.

GILLESPIE, J. H., 1986b Variability of evolutionary rates of DNA. Genetics **113:** 1077–1091.

GILLESPIE, J. H., 1988 More on the overdispersed molecular clock. Genetics **118:** 385–386.

HALE, L. R., and R. S. SINGH, 1987 Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster*. Mol. Biol. Evol. **4:** 622–637.

HARRIS, H., and D. A. HOPKINSON, 1976 *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland, Amsterdam.

HUDSON, R. R., M. KREITMAN and M. AGUADÉ, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

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

KREITMAN, M., 1983 Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature **304:** 412–417.

KREITMAN, M. V. 1987 Molecular population genetics. *Oxf. Sur. Evol. Biol.* **4:** 38–60.

LACHAISE, D., M. L. CARIOU, J. R. DAVID, F. LEMEUNIER, L. TSACAS and M. ASHBURNER, 1987 Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22:** 159–255.

LANGLEY, C. H., and C. F. AQUADRO, 1987 Restriction map variation in populations of *Drosophila melanogaster*: white-locus region. Mol. Biol. Evol. **4:** 651–663.

LANGLEY, C. H., E. MONTGOMERY and W. F. QUATTLEBAUM, 1982 Restriction map variation in the *Adh* region of *Drosophila*. Proc. Natl. Acad. Sci. USA **79:** 5631–5635.

LEIGH BROWN, A. J., and D. ISH-HOROWICZ, 1981 Evolution of

the 87A and 87C heat-shock loci in *Drosophila*. Nature **290:** 677–682.

LEMEUNIER, F., L. TSACAS, J. DAVID and M. ASHBURNER, 1985 The *Drosophila melanogaster* species group, pp. 147–256 in *The Genetics and Biology of Drosophila*, Vol. 3e, edited by M. ASHBURNER, H. L. CARSON and J. N. THOMPSON, JR. Academic Press, New York.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

LEWONTIN, R. C., 1985 Population genetics, pp. 3–18 in *Evolution: Essays in Honour of John Maynard Smith*. edited by P. J. Greenwood, P.H. Harvey and M. SLATKIN. Cambridge University Press, Cambridge.

NEI, M., 1975 *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam.

NEI, M., 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89:** 583–590.

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

NEI, M., and A. K. ROYCHOUDHARY, 1974 Sampling variance of heterozygosity and genetic distance. Genetics **76:** 379–390.

NEI, M., and Y. TATENO, 1975 Interlocus variation of genetic distance and the neutral mutation theory. Proc. Natl. Acad. Sci. USA **72:** 2758–2760.

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

PARKER, H. R., D. P. PHILLIP and G. S. WHITT, 1985 Gene regulatory divergence among species estimated by altered developmental patterns in interspecific hybrids. Mol. Biol. Evol. **2:** 217–250.

PRAKASH, S., R. C. LEWONTIN and J. L. HUBBY, 1969 A molecular approach to the study of genic heterozygosity in natural populations. IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. Genetics **61:** 841–858.

RILEY, M. A., M. E. HALLAS and R. C. LEWONTIN, 1989 Distinguishing the forces controlling genetic variation at the *Xdh* locus in *Drosophila pseudoobscura*. Genetics **123:** 359–369.

SICILIANO, M. J., and C. R. SHAW, 1976 Separation and visualization of enzymes on gels, pp. 185–209 in *Chromatographic and Electrophoretic Techniques*, Vol. 2, Ed. 4, edited by I. SMITH. Heinemann, London.

SINGH, R. S., M. CHOUDHARY and J. R. DAVID, 1987 Contrasting patterns of geographic variation in the cosmopolitan sibling species *Drosophila melanogaster* and *Drosophila simulans*. Biochem. Genet. **25:** 27–40.

SINGH, R. S., D. A. HICKEY and J. R. DAVID, 1982 Genetic differentiation between geographically distant populations of *Drosophila melanogaster*. Genetics **104:** 301–315.

SINGH, R. S., and L. R. RHOMBERG, 1987a A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. I. Estimates of gene flow from rare alleles. Genetics **115:** 313–322.

SINGH, R. S., and L. R. RHOMBERG, 1987b A comprehensive study of genic variation in *Drosophila melanogaster*. II. Estimates of heterozygosity and patterns of geographic differentiation. Genetics **117:** 255–271.

SKIBINSKI, D. O. F., and R. D. WARD, 1981 Relationship between allozyme heterozygosity and rates of divergence. Genet. Res. **38:** 71–92.

SKIBINSKI, D. O. F., and R. D. WARD, 1982 Correlations between heterozygosity and evolutionary rate of proteins. Nature **298:** 490–492.

STEPHENS, J. C., and M. NEI, 1985 Phylogenetic analysis of polymorphic DNA sequences at the Adh locus in *Drosophila melanogaster* and its sibling species. J. Mol. Evol. **22:** 289–300.

THROCKMORTON, L. H., 1977 Drosophila systematics and biochemical evolution. Annu. Rev. Ecol. Syst. **8:** 235–254.

WARD, R. D., and D. O. F. SKIBINSKI, 1985 Observed relationships between protein heterozygosity and protein genetic distance and comparisons with neutral expectations. Genet. Res. **45:** 315–340.

WRIGHT, S., 1951 The genetical structure of populations. Ann. Eugen. **15:** 323–354.

WRIGHT, S., 1982 The shifting balance theory and macroevolution. Annu. Rev. Genet. **16:** 1–20.

ZOUROS, E., 1981 The chromosomal basis of sexual isolation in two sibling species of *Drosophila*: *D. arizonensis* and *D. mojavensis*. Genetics **97:** 703–718.

Communicating editor: A. G. CLARK