J Mol Evol (1988) 27:250–260

Journal of Molecular Evolution

Molecular Evolution among Some *Drosophila* Species Groups as Indicated by Two-Dimensional Electrophoresis

Greg S. Spicer*

Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Illinois 60439, USA, and Committee on Evolutionary Biology, University of Chicago, Chicago, Illinois 60637, USA

Summary. The evolutionary and phylogenetic relationships of seven *Drosophila* species groups (represented by *D. melanogaster, D. mulleri, D. mercatorum, D. robusta, D. virilis, D. immigrans, D. funebris,* and *D. melanica*) were investigated by the use of two-dimensional electrophoresis. The resulting phylogeny is congruent with the current views of evolution among these groups based on morphological characters and immunological distances. Previous studies indicated that the ability of one-dimensional electrophoresis to resolve relationships between distantly related taxa extended to about the Miocene [25 million years (Myr) ago], but the present study demonstrates that two-dimensional electrophoresis is a useful indicator of phylogeny even back to the Paleocene (65 Myr ago). In addition, two-dimensional electrophoresis is shown to be a useful technique for detecting slowly evolving structural proteins such as actins and tropomyosins.

Key words: Molecular evolution — Two-dimensional electrophoresis — *Drosophila* systematics — Phylogenetic analysis — Molecular clocks — Actin — Tropomyosin

Introduction

Electrophoresis has long been used as a tool for determining the evolutionary relationships of organisms (Avise 1975; Buth 1984). However, its ability to resolve the phylogenetic relationships of distantly related taxa has been questioned (Avise 1975; Bush and Kitto 1978; Maxson and Maxson 1979; Matson 1984). The major reason for this uncertainty is that after the first mobility change in a protein, all subsequent substitutions reveal no further information regarding phylogenetic relationships (Maxson and Maxson 1979). However, this problem of "saturation" is greatly dependent on the rates of evolution of the proteins under study: proteins that evolve rapidly will only be useful in resolving recent events, but proteins with much slower rates of change can elucidate more distant relationships (Sarich 1977). In this context, it is known that different classes of proteins do evolve at different rates (Powell 1975; Wilson et al. 1977). Hence, an electrophoretic technique that can examine a more slowly evolving set of proteins will be more useful for inferring phylogenies among distantly related taxa than is the traditional allozyme electrophoresis.

Unlike traditional allozyme electrophoresis, the technique of two-dimensional electrophoresis surveys many classes of proteins in addition to those that have enzymatic activity (Klose and Feller 1981). The difference between one-dimensional and two-dimensional electrophoresis can best be shown in terms of the known correlation between the rate of change in a protein and its levels of polymorphism and heterozygosity (Skibinski and Ward 1981, 1982). Studies based on allozymic variation indicate that *Drosophila* have some of the highest polymorphism and heterozygosity levels in the animal kingdom (Nevo 1978). However, comparable studies using two-dimensional electrophoresis have given considerably lower estimates of polymorphism and heterozygosity in *Drosophila* (Leigh Brown and Langley 1979; Ohnishi et al. 1982; Coulthart 1986). Con-

Offprint requests to: G.S. Spicer
* *Current address:* Linus Pauling Institute, 440 Page Mill Road, Palo Alto, California 94306, USA

sequently, two-dimensional electrophoresis of these less variable and more slowly evolving proteins should result in an increased ability to resolve relationships among distantly related taxa (MacIntyre and Collier 1986).

As Beverley and Wilson (1982) mentioned, very few studies have used biochemical techniques to resolve the higher relationships within the genus *Drosophila.* This is unfortunate because the genus is one of the most widely studied in evolutionary biology and has been examined extensively at the lower taxonomic levels. However, to evaluate the evolution of the entire genus *Drosophila,* it is necessary to have an accurate phylogeny (Throckmorton 1962; Felsenstein 1985). For the most part, the current phylogeny of the Drosophilidae is based on the extensive morphological studies of Throckmorton (1962, 1969, 1975). The major biochemical studies that have proven most useful in determining the higher level systematic relationships of this group are those of Beverley and Wilson (1982, 1984, 1985), who used microcomplement fixation of the larval hemolymph protein to produce a phylogeny and molecular clock based on immunological distances. They found that both the phylogenetic relationships and evolutionary dates of divergence were congruent with those inferred from the morphological, fossil, and biogeographic evidence. Several other molecular studies, in which both microcomplement fixation (Duke and Glassman 1968; Collier and MacIntyre 1977; MacIntyre et al. 1978) and DNA–DNA hybridization (Entingh 1970) were used, have also tested these relationships and found them to be largely concordant with the currently accepted phylogeny of the genus.

This study was undertaken to elucidate the molecular evolution and phylogenetic relationships among several distantly related species groups of *Drosophila* [separated by about 60 million years (Myr)] by using two-dimensional electrophoresis. Although several other studies have also applied the technique of two-dimensional electrophoresis to infer relationships within the genus *Drosophila,* these other studies have been concerned primarily with the lower taxonomic levels (Ohnishi et al. 1983a,b; Spicer 1985; Lee and Pak 1986). Consequently, it seemed advisable to test the limits of resolution of this technique for phylogenetic reconstruction and to evaluate the accuracy of the data to function as a molecular clock.

Materials and Methods

Drosophila Strains. Most of the stocks used in this study came from the laboratory of Dr. Lynn H. Throckmorton, University of Chicago. The flies were raised and maintained on either a banana (B) or cornmeal (C) medium. The stocks that were used and their corresponding National Drosophila Species Resource Center stock numbers are as follows: *D. melanogaster,* Mount Carmel, Illinois, 1970 (C); *D. immigrans,* East Lansing, Michigan, November 1982 (B); *D. mulleri* (15081-1371.8), Roy Farm, Austin, Texas (B); *D. funebris* (15120-1911.2), Minneapolis, Minnesota (B); *D. mercatorum* (S-sl v pm vl-Br16), received from Dr. Alan R. Templeton (B); *D. virilis* (15010-1051), Pasadena, California (B); *D. virilis,* Whiteshell Provincial Park, Manitoba, Canada, August 2, 1974 (B); *D. virilis,* 16—Sapporo, Hokkaido, Japan, from Wheeler, 1971 (B); *D. robusta* (15020-1111.4), Jamestown, South Carolina (B); *D. melanica* (15030-1141.1), Cliff, New Mexico (B).

Sample Preparation and Electrophoresis. The two-dimensional electrophoresis was performed as outlined by O'Farrell (1975), with the modifications of Anderson and Anderson (1978a,b). All solutions and procedures for use of the ISO-DALT system can be found in Tollaksen et al. (1984).

Samples were prepared by homogenizing 12–60 (depending on the species) etherized adult male flies in a mixture containing 9 M urea, 2% Nonidet P-40 detergent, 2% mercaptoethanol, and 2% LKB ampholytes, pH 9–11. Each sample was prepared with a concentration of 0.2 mg wet weight of fly/μl of urea mix. These samples were centrifuged for approximately 1.5 min at 10,000 $\times$ *g*. The supernatants from these samples were then centrifuged at 435,000 $\times$ *g* (maximum) for 5 min in a Beckman TL-100 ultracentrifuge.

The gels used in the study measured 20 $\times$ 25 cm. Isoelectric focusing was performed in the first dimension with a 1:1 mixture of ampholytes, pH 3–10 Biolytes and pH 5–7 Biolytes. The amount of sample loaded onto each gel was 15 μl, which resulted in a concentration of 3 mg wet fly weight/gel. The first dimension was run at 30,000 volt-hours for an overall run time of 11 hours. The second dimension was a 9–18% computer-poured gradient sodium dodecyl sulfate (SDS) gel run at about 100–150 volts (0.6 amperes) overnight. The gels were stained overnight in a solution of 0.125% Coomassie Brilliant Blue (Serva Blue R) and 2.5% phosphoric acid and destained several times in 20% ethanol and water.

Data Analysis. The outgroup examined in this study, to determine the polarity of the protein spots, was *Drosophila melanogaster.* The choice of this species was based on the studies of morphological characters by Throckmorton (1962, 1969, 1975) and immunological distances by Beverley and Wilson (1982). Both of these data sets indicate that all the members of the ingroup are phylogenetically more closely related to each other than any one is to *D. melanogaster.*

The presence–absence data were converted into a binary data matrix, so that a phylogenetic analysis could be performed by using the PAUP program of Swofford (1984). Two separate cladistic analyses were performed on the data. The first is analogous to the locus-by-locus approach used by many workers (Throckmorton 1978; Wake et al. 1978; Baverstock et al. 1979; Avise et al. 1980; Honeycutt and Williams 1982; Arnold et al. 1983; Hillis et al. 1983; Patton and Avise 1983); however, none of the alleles were assigned polarity on the basis of ingroup commonality, as is done in some studies (Sites et al. 1981, 1984; Lanyon 1985). Any spot that was present in both the outgroup (*D. melanogaster*) and the ingroup was considered ancestral (plesiomorphic) and was eliminated from the analysis. These characters were not used in the phylogenetic analysis because plesiomorphic characters provide no information about branching sequences within the ingroup (Hennig 1966). Protein spots that were unique to *D. melanogaster* were retained in the analysis only to give an estimate of the branch length. A hypothetical ancestor containing only primitive characters was created, so that the absence of the unique characters possessed by *D. melanogaster* could not be considered as evidence of shared derived characters in the other

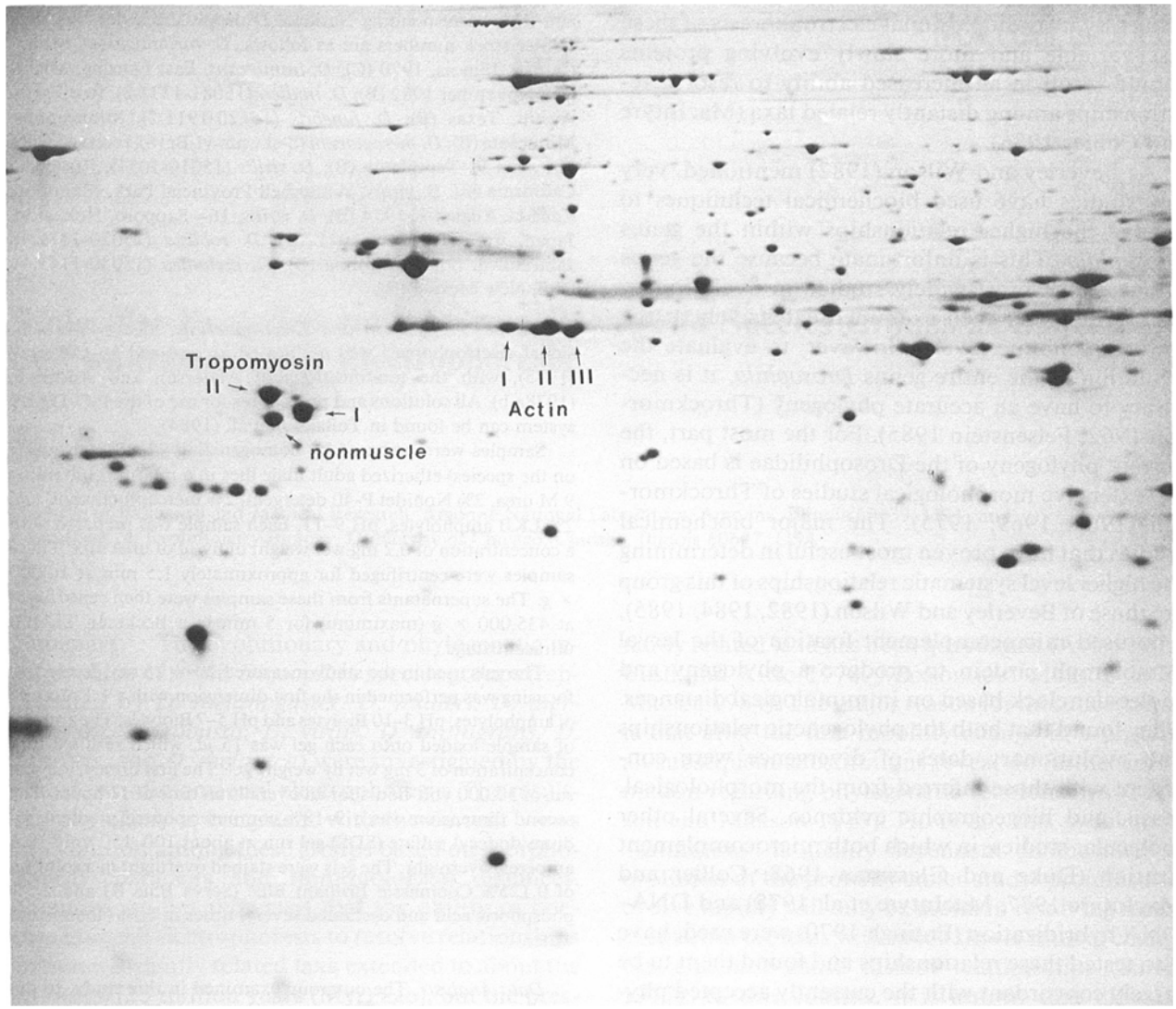

Fig. 1. Two-dimensional electrophoretic gel of *Drosophila mulleri* stained with Coomassie Brilliant Blue. The high-molecular-weight proteins are at the top of the gel, and the basic proteins are at the right.

species. This procedure is comparable to Lundberg rooting (Lundberg 1972).

The second analysis used the independent allele model, which considers the loss of an allele to be as important as the acquisition of a new one (Mickevich and Johnson 1976; Mickevich and Mitter 1981; Miyamoto 1981; Hillis et al. 1983; Sites et al. 1984; Hillis 1985). With this approach, all the characters are used and a most parsimonious tree is produced, in this case rooted with *D. melanogaster*. The data set was small enough so that an exhaustive search of all possible trees was performed by using the alltrees command. This procedure examines all possible trees and therefore guarantees that the most parsimonious tree(s) will be found. The CONTREE program of Swofford (1982) was used to produce both strict and Adams-2 consensus trees from the equally parsimonious trees that were found (Adams 1972; Rohlf 1982).

The binary data matrix was also analyzed phenetically under the assumption of a molecular clock (Wilson et al. 1977; Ayala 1982; Thorpe 1982). Unfortunately, it is difficult to homologize loci between different species with two-dimensional electrophoresis. Therefore, the generally used genetic distance methods are inappropriate for these data. As recommended by Sokal and Rohlf (1981), the most appropriate way to analyze binary data sets is to use the simple matching coefficient (S_{sm}) (Sokal and Sneath 1963; Sneath and Sokal 1973). However, I have modified this metric so that it will be appropriate under the assumption of a molecular clock by applying a logarithmic transformation to linearize the distance measure. This procedure results in a measure analogous to Nei's D (Nei 1972) with the same interval. The new distance measure is defined as the $-\log_e S_{sm}$. This distance measure is then clustered using UPGMA (Sokal and Sneath 1963). The phenetic analyses were performed using the NT-SYS programs of Rohlf et al. (1981).

Results

About 70 two-dimensional electrophoretic gels were run as part of this study. At least two gels were run per species, and many samples were coelectrophoresed to ensure the proper matching of the protein spots. An example of a Coomassie-stained gel, in this case for *Drosophila mulleri*, is presented in Fig. 1. A diagrammatic representation of this gel,

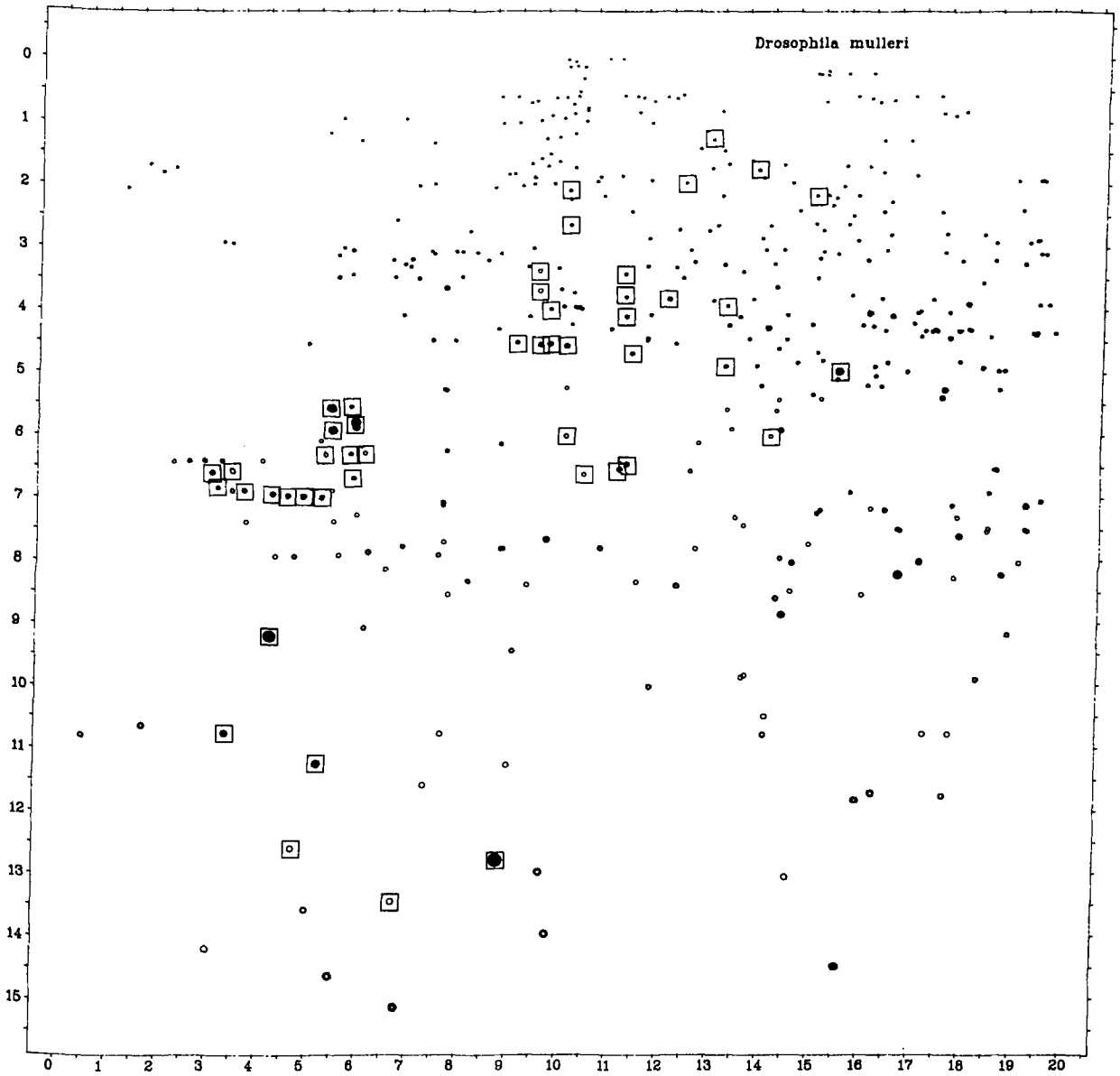

Fig. 2. A diagrammatic representation of the *Drosophila mulleri* two-dimensional electrophoretic gel in Fig. 1. The protein spots that are ensquared were scored for this species.

showing the protein spots that were scored for this species, is presented in Fig. 2.

Unfortunately, as Avise (1983) mentions, it can be difficult to homologize loci between different species when two-dimensional electrophoresis is used, because unlike the allozyme systems that stain for a specific enzyme, two-dimensional electrophoresis uses only general protein stains to detect the genetic products. Consequently, the homology of allelic products is assigned by position and general appearance of the protein spots. This method is very often inadequate for different species because the spots have changed position and shape to such an extent that the assignments become almost arbitrary. To minimize this problem the spots were simply scored as present or absent. Although this procedure does not take into account the genetics involved, it should not change the basic results unless many loci are differentially expressed and therefore excluded from the analysis.

A total of 135 protein spots were scored for this study. The presence–absence data set generated from the spot list is presented in Table 1. Not all spots on the gels were scored, because many were inconsistent from gel to gel. Hence, only the most reliably matched spots were used for the analysis. Of the 135 spots, 17 were constant among all the species and another 63 protein spots were unique to only one species. Another 23 proteins were shared between *D. melanogaster* and at least one other species. No intraspecific variation was found for the proteins scored in the three strains of *D. virilis*.

Six equally parsimonious cladograms with total lengths of 117 and a consistency index of 0.812 were

Table 1. Matrix of the 135 protein spots (characters) scored for the eight *Drosophila* species as resolved by two-dimensional electrophoresis

011111111000000000000100010000000101100100010111001001000110000111000110010001000111111111111111	*D. melanogaster*
010001110110000000000010010000111110001000101111000110000011000100001110001001000111111111111111	*D. immigrans*
011110010010000010000100000000000100001011000111010010001011000000110000011000000100000111111111	*D. funebris*
010000010110001001000010000011100111010000000111010000100111111000000000000100101100100000000010000000000000001000000000011000111111111111	*D. mercatorum*
100000010110010000100001000111100111010000001111100000101010111000000000001001011010000000000001000000000000011000000000011001111111111111	*D. mulleri*
100010010110000010001000001001111111010000011111010000100111110000000011001000000010000001000000000000001100000000000011111111111111	*D. virilis*
100000000110100110010000100011000111010000001111010000110011110000000011111010001000100010001001000000000000110000000000111111111111	*D. melanica*
010000010111000010001000000001000111010000001111010000100111111000000000001101000001010000000001010000000000000000000011000001111111111111	*D. robusta*

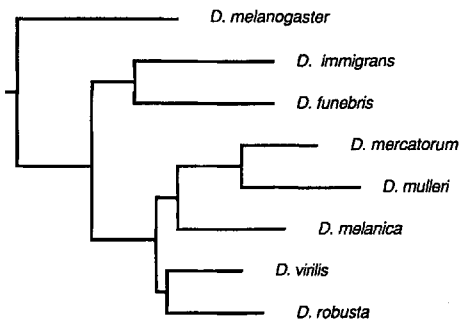

Fig. 3. The tree with the lowest *f*-value (100), analyzed by using the locus-by-locus approach. Total length = 117; consistency index = 0.812.

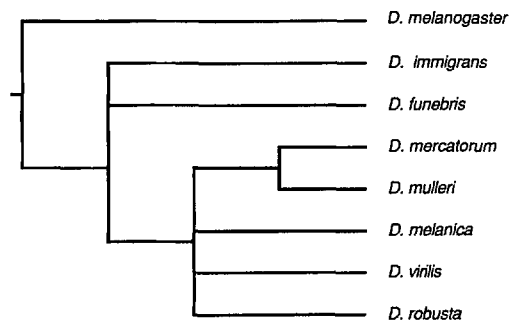

Fig. 4. The strict consensus tree produced from the 10 most parsimonious trees. This is considered the phylogeny as indicated by two-dimensional electrophoresis.

found after analysis with an exhaustive search algorithm that guarantees the shortest tree(s) when the data are analyzed by the locus-by-locus approach. The best tree, according to its low *f*-value of 100 (Farris 1972), is presented in Fig. 3. The other trees had *f*-values of 112, 112, 134, 176, and 180. When all the characters were used to infer a tree with the independent allele model, four equally most parsimonious trees with total lengths of 162 were resolved. These trees had a consistency index of 0.728 and *f*-values of 194, 212, 218, and 234. The strict and Adams-2 consensus tree produced from both the locus-by-locus approach and the independent allele model is presented in Fig. 4. There is no reason to suppose that the tree with the lowest *f*-value is more likely than the others, because they all invoke the same number of evolutionary steps. Therefore, the strict consensus tree presented in Fig. 4 is considered to represent the inferred phylogeny based on two-dimensional electrophoresis. Although consensus trees do not take into account partial agreement between trees (Rohlf 1982), they do summarize the parts of the trees that are in agreement and therefore can be considered to be the most conservative estimate of a phylogeny.

The equally parsimonious trees had some parts in common that were not expressed in the consensus tree presented as the phylogeny (Fig. 4). When analyzed with the locus-by-locus approach, *D. funebris* was never considered as part of the *virilis–repleta* clade. This was not the case, however, in the independent allele analysis, because one of the four equal trees relates *D. funebris* more closely to the *virilis–repleta* group. Similarly, within the *virilis–repleta* radiation, *D. robusta* was never part of the *repleta* species group clade (*D. mulleri* and *D. mercatorum*) in the locus-by-locus analysis, but was part of this clade in the analysis with the independent allele approach.

The tree produced when the data were analyzed with the distance measure values (Table 2) under the molecular clock hypothesis is presented in Fig. 5. This tree is unlike those produced with either the locus-by-locus or the independent allele approaches, although all of its parts are present in some of the equally parsimonious trees. For example, *D. im-*

Table 2. The values above the diagonal are the distances derived from the simple matching coefficient, and values below the diagonal are the $-\log_e$ of the simple matching coefficient. See Materials and Methods for explanation of this distance measure.

	mel	*imm*	*fun*	*mer*	*mul*	*vir*	*rob*	*mlc*
melanogaster	—	0.630	0.637	0.593	0.526	0.630	0.563	0.615
immigrans	0.462	—	0.696	0.681	0.644	0.704	0.622	0.689
funebris	0.451	0.362	—	0.674	0.637	0.667	0.615	0.711
mercatorum	0.523	0.384	0.395	—	0.830	0.800	0.763	0.815
mulleri	0.642	0.440	0.451	0.186	—	0.763	0.726	0.763
virilis	0.462	0.351	0.405	0.223	0.271	—	0.800	0.837
robusta	0.574	0.475	0.486	0.271	0.320	0.223	—	0.800
melanica	0.486	0.376	0.341	0.205	0.271	0.178	0.223	—

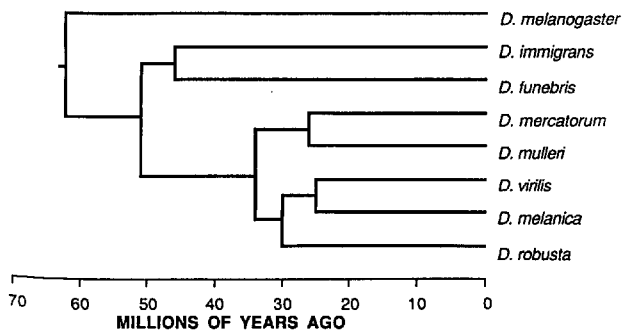

Fig. 5. The tree produced by analysis under the assumption of a molecular clock. The distance measure used is the $-\log_e S_{sm}$ (see text for details). Distances were clustered using UPGMA from the values in Table 2. Cophenetic correlation = 0.922.

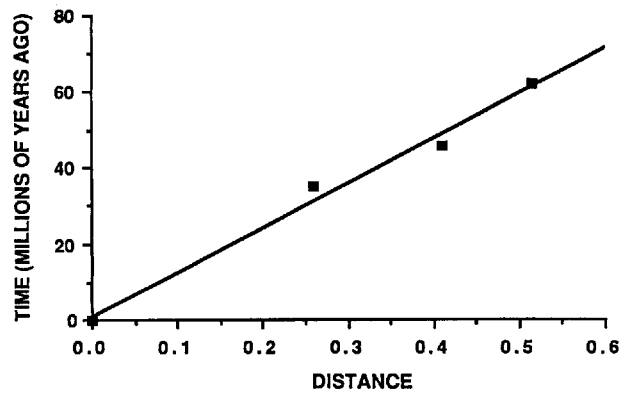

Fig. 6. The regression of divergence times from Beverley and Wilson (1984) on the distances obtained by two-dimensional electrophoresis. The slope is significantly different from zero (t = 13.23, df = 2, P = 0.0057).

migrans and *D. funebris* are clustered together in the molecular clock tree, and this relationship also occurs in several of the equally parsimonious trees. In addition, the branching sequence seen in the *virilis–repleta* radiation also occurs in some of the parsimonious trees, and the two *repleta* group species (*D. mercatorum* and *D. mulleri*) are clustered together as they should be. Also, as has been indicated by other studies, *D. melanogaster* is shown to be more distantly related to the other species in this study, which confirms its use as an outgroup.

The molecular clock tree (Fig. 5) was calibrated with the data of Beverley and Wilson (1984). This was accomplished by plotting their divergence estimates, based on immunological distance, against the distances generated in the present study. Four divergence times that were comparable between studies were used for the calibration. The first point is based on the assumption that at zero time of divergence there will be zero genetic distance (Fitch 1976). In this study, the assumption is substantiated in that no intraspecific variation was found in *D. virilis*; consequently, this can be considered an estimated point. The other three times correspond to speciation events common to both trees. The separation of the *robusta* species group from the *repleta* group is considered to have occurred 35 Myr ago by Beverley and Wilson (1984), who also set the divergence of the *immigrans–Hirtodrosophila* radiation from the *virilis–repleta* radiation at about 46 Myr ago. Finally, Beverley and Wilson (1984) indicate that the subgenus *Sophophora* diverged from the subgenus *Drosophila* around 62 Myr ago. The regression equation resulting from this comparison (time = 1.08 + 117·distance) has a significant slope (t = 13.23, df = 2, P = 0.0057, 95% confidence interval 78, 155) with an intercept not significantly different from zero (Fig. 6). This estimation sets the rate of divergence at one unit of distance ($-\log_e S_{sm}$) = 118 Myr. The slope is not significant however, when the zero–zero point is excluded from the analysis (t = 4.66, df = 1, P = 0.1345). The divergence times presented here are entirely dependent on those given by Beverley and Wilson (1984). However, even if the absolute times are in error, there is a remarkable correlation (r = 0.99) between the rates of molecular evolution in this study and theirs.

Discussion

The phylogeny of these *Drosophila* species as indicated by two-dimensional electrophoresis is entirely congruent with that produced by other data sets. The phylogeny proposed by Throckmorton (1975, 1982) on the basis of morphology is identical to that presented in Fig. 4. Throckmorton (1975)

identified four main groups within the flies examined in this study. The outgroup in this study, *D. melanogaster,* is a member of the sophophoran radiation that is considered the sister group to the members of the subgenus *Drosophila,* which includes all the ingroup species studied here. The subgenus *Drosophila* is represented by three separate lineages in this study: the *virilis–repleta* radiation, the *immigrans–Hirtodrosophila* radiation, and the *funebris* group. The relationships among the three species groups (*robusta, virilis, repleta*) within the *virilis–repleta* radiation are currently unresolved (Throckmorton 1982), but the two *repleta* group species are clustered together as they should be. All of these taxonomic levels are clearly resolved in Fig. 4, which further substantiates Throckmorton's (1975, 1982) proposed phylogeny for the genus *Drosophila.*

The major biochemical phylogeny of *Drosophila* (Beverley and Wilson 1982), based on immunological distances of larval hemolymph protein, is difficult to compare and evaluate with respect to this study. Of the five species in common to both studies (*D. melanogaster, D. mulleri, D. virilis, D. immigrans,* and *D. sordidula*), only two species were examined by using reciprocal tests (*D. melanogaster* and *D. mulleri*), and one of these was considered the outgroup in the present study. (Although *D. sordidula* was not examined in the present study, it is a member of the *robusta* species group and therefore is closely related to *D. robusta,* which was examined in the present study.) The other three species were placed on the tree using a unidirectional method, which is much less reliable than are reciprocal tests. To make comparisons more difficult, two of these latter three species (*D. immigrans* and *D. virilis*) were assigned positions on the tree on the basis of the morphological phylogeny of Throckmorton (1975) and not on the basis of immunological distances. Given these qualifications, the biochemical phylogeny of Beverley and Wilson (1982) is concordant with that based on morphology by Throckmorton (1975) and the phylogeny inferred by two-dimensional electrophoresis.

Another important molecular data set is that of MacIntyre et al. (1978), who examined the evolution of *Drosophila* by looking at the immunological distances deduced from microcomplement fixation tests of acid phosphatase-1 by conducting reciprocal comparisons among 10 species. Their study further confirmed the phylogeny proposed by Throckmorton (1975), with two exceptions. In the phylogeny of MacIntyre et al. (1978), *D. funebris* is considered a member of the subgenus *Sophophora* instead of the subgenus *Drosophila,* and *D. nebulosa* is clustered in the *saltans* species group instead of the *willistoni* species group. MacIntyre et al. (1978) explain the placement of *D. funebris* by the fact that the enzyme is extremely electronegative in this species, giving it some unusual immunological properties. They also state that the *D. nebulosa* placement seems to have resulted from a bias in the way immunological comparisons were conducted. Nevertheless, their overall phylogeny is otherwise congruent with both the morphological and two-dimensional electrophoretic data sets.

Several other biochemical studies have examined the higher relationships within the genus *Drosophila.* These are worthy of mention but do not really constitute tests of the phylogeny of *Drosophila,* because of various shortcomings in technique. The microcomplement fixation study of Collier and MacIntyre (1977) examined the evolution of α-glycerophosphate dehydrogenase in 34 species of *Drosophila.* The purpose of the study was not to produce a phylogeny, but simply to examine the evolutionary change in the enzyme. However, the resulting distances are in line with the general phylogenetic framework of the genus. Another immunological study is that of Duke and Glassman (1968) on xanthine dehydrogenase. These workers performed microcomplement fixation tests on 11 species of *Drosophila,* but performed no reciprocal tests and related all comparisons only to *D. melanogaster.* Still, the general relationships seem to substantiate the phylogeny presented here. The same is true for the DNA–DNA hybridization study of Entingh (1970). Although some of the comparisons were reciprocal, most were not, and consequently no phylogeny was reported. However, the distances presented are in accord with the currently accepted phylogeny of *Drosophila.*

Two other studies have presented phylogenies based on biochemical techniques, but are not congruent with any previous work. The study of MacIntyre and Dean (1978) measured the quantitative subunit hybridization of acid phosphatase-1 among 11 species of *Drosophila.* Although a dendrogram was produced from these data, the authors acknowledged that this should not be considered a useful technique for phylogeny reconstruction (MacIntyre and Dean 1978; MacIntyre and Collier 1986). The other study, by Vilageliu and Gonzalez-Duarte (1984), proposed phylogenetic relationships among seven divergent species of *Drosophila* on the basis of relative frequencies of the amino acids in alcohol dehydrogenase. However, this phylogeny is not congruent with any proposed phylogenetic hypothesis of the genus; therefore, the technique should also be considered suspect.

One criticism of two-dimensional electrophoresis is that it may not adequately discriminate electromorphs that differ in their primary structure (McLellan et al. 1983; McLellan and Inouye 1986).

Some workers have suggested that an inability to separate different proteins may cause the production of incorrect phylogenies (Coyne et al. 1979; Berlocher 1984) and may be responsible for much of the observed convergence on electrophoretic gels (Mickevich and Mitter 1981). However, the trees reported here both are congruent with other data sets (as mentioned above) and have a high consistency index for electrophoretic data sets (Mickevich 1978; Mickevich and Mitter 1981; Sites et al. 1984). These factors indicate that, at least for this study, two-dimensional electrophoresis can adequately separate proteins for systematic studies.

Another problem concerning electrophoresis is its ability to resolve phylogenetic relationships among distantly related organisms. Maxson and Maxson (1979) have suggested that traditional one-dimensional electrophoresis can be used only to about the Miocene (25 Myr ago), and Bush and Kitto (1978) suggested that its usefulness is only to the late Pliocene (10–15 Myr ago). The results presented here indicate that two-dimensional electrophoresis can be applied to taxa separated by much more time than can standard allozyme electrophoresis. Considering the fossil and biogeographic evidence presented for the separation of the subgenera of *Drosophila* (Throckmorton 1975, 1982; Beverley and Wilson 1984, 1985) and the corresponding dates of divergence based on the molecular clocks (corrected estimates of 61–65 Myr ago; Collier and MacIntyre 1977; Beverley and Wilson 1984, 1985; MacIntyre and Collier 1986), two-dimensional electrophoresis seems to be a useful indicator of phylogeny even back to the Paleocene (65 Myr ago).

The existence of molecular clocks has been much disputed (Goodman 1981; Ayala 1986), although the presence of some kind of clock seems undeniable (Fitch 1976; Thorpe 1982). However, virtually all the empirical data substantiating molecular clocks are based on either amino acid or nucleotide sequences. Electrophoretic clocks have been considered much more questionable because a 20-fold range of differences between calibrations has been observed (Avise and Aquadro 1982). Hence, a single electrophoretic clock seems unlikely, but it is well known that different proteins evolve at different rates (Dickerson 1971; Wilson et al. 1977) and even that the same protein can change at different rates in different lineages (Goodman 1976, 1981). Consequently, a large variance in electrophoretic clocks is not unexpected, particularly for studies that examine different proteins and survey comparatively few loci.

In this latter problem, two-dimensional electrophoresis potentially has a distinct advantage over traditional one-dimensional electrophoresis. It is known that the more independent molecular data available, the better the estimate of divergence time will be (Fitch 1976; Takahata and Nei 1985; Watterson 1985). Although this study examined only about 50 loci, it is possible to survey several hundred loci with two-dimensional electrophoresis (Klose and Feller 1981; Klose 1982; Jungblut and Klose 1985; Neel et al. 1985; Spicer 1985; Coulthart 1986). Even with so few loci examined, the correlation (Fig. 6) between this study and the results of Beverley and Wilson (1984) is remarkable.

However, the absolute times of divergence are of the greatest interest, and here a great deal of ambiguity still exists. With such a poor fossil record (only two fossils of *Drosophila* have been reported), the only alternative is to use biogeographical information combined with the phylogeny and present-day distributions of the species to reconstruct past events (Wilson et al. 1977; Throckmorton 1982; Cracraft 1983). Unfortunately, this approach gives only very rough estimates of divergence times, and therefore no firm dates for calibrating a molecular evolutionary clock. Beverley and Wilson (1984, 1985) have addressed this problem by examining the molecular evolution of larval hemolymph protein over a long expanse of geological time, enabling them to standardize a clock on the basis of several bound and unbound dates. They concluded that the subgenera *Sophophora* and *Drosophila* diverged approximately 62 Myr ago. Comparing the Beverley and Wilson (1984) divergence times with the two-dimensional electrophoretic data (Fig. 6) gives the time of divergence as 61 Myr ago.

Two other divergence times based on molecular studies have been reported. The study of Collier and MacIntyre (1977) based on α-glycerophosphate dehydrogenase gives the divergence date as 52 Myr ago, and another based on arginine kinase gives the date as 59 Myr ago (MacIntyre and Collier 1986; G. Collier, personal communication). However, these two dates were calibrated by taking the divergence of the *virilis–repleta* radiation from the *immigrans–Hirtodrosophila* radiation as 36 Myr ago, a date chosen because Throckmorton (1975) had inferred that the *virilis–repleta* radiation developed during the Oligocene and Miocene. Subsequently, Throckmorton (1982) indicated that this radiation had occurred about 30 Myr ago, and that "there is no ground to even speculate on how much before then they appeared." Consequently, these divergence dates should be considered minimum unbounded estimates.

To bring these estimates in line with the others, I used the regression procedure described above. To compare the studies of Collier and MacIntyre (1977) and MacIntyre and Collier (1986) with that of Beverley and Wilson (1984), I used the regression procedure described in the Methods section. The regres-

sion for the α-glycerophosphate dehydrogenase data is significant (t= 16.87, df = 1, P = 0.0377), but that for the arginine kinase data is not (t = 6.79, df = 1, P = 0.0931). These recalibrations correspondingly change the dates of 52 and 59 Myr ago to 63 and 65 Myr ago, respectively.

These suggested dates are still in accord with the biogeographic data as currently interpreted. Throckmorton (1975) considered the subgenus *Drosophila* to have been in existence by the Oligocene (36 Myr ago), but nothing can be conjectured about how much earlier it existed. Beverley and Wilson (1984) concluded on the basis of the biogeographic consideration of continental drift, that the Drosophilidae did not exist earlier than about 80 Myr ago. New Zealand, which is thought to have split from Australia at about that time, has no ancient native *Drosophila* fauna, while Australia has an extensive one. Therefore, the Drosophilidae must have originated sometime after the split of these two land masses (80 Myr ago). If the initial calibration of Beverley and Wilson (1984) is reliable, the subgenera must have diverged sometime between 61 and 65 Myr ago.

Previous studies investigating protein structure and function have revealed much about soluble proteins such as enzymes, but very little has been learned about the insoluble structural proteins (O'Brien and MacIntyre 1978). This discrepancy is partly due to the difficulty of effectively screening for structural protein mutants (Fyrberg 1984). Two-dimensional electrophoresis provides a simple method for screening for the insoluble proteins and examining the relative rates of protein evolution. By using this method it is now possible to identify the slowly evolving proteins.

Interestingly, even after roughly 60 Myr of divergence, the electrophoretic mobilities of some proteins have remained unchanged. Of the 12 proteins whose electrophoretic mobilities have not changed, 5 can be identified from previous studies on *Drosophila*. Four of these are the actins I–III (Storti et al. 1978), and the other is muscle tropomyosin II (Bautch et al. 1982; Mogami et al. 1982; Bautch and Storti 1983). Although six actin genes in *Drosophila* (Fyrberg et al. 1980; Tobin et al. 1980) are known to produce at least five different proteins (Fyrberg 1984), only three forms have previously been separated by electrophoresis (Storti et al. 1978; Fyrberg et al. 1983). The fourth form resolved here seems to be related to actin II. In addition, it seems that actin III is probably also resolved into two forms as well. However, more study is needed to determine these relationships with certainty.

The proteins identified here as evolving slowly are structural proteins. Both actin and tropomyosin are known to be highly conserved over evolutionary time (Fine and Blitz 1975; Firtel 1981; Hightower and Meagher 1986). However, all the tropomyosins are not equally conserved. The *Drosophila* tropomyosins are encoded by three tightly linked genes and consist of two muscle forms that are differentially regulated and a nonmuscle cytoplasmic form (Bautch et al. 1982). Only one of the muscle forms is apparently highly conserved, while the other two genes are much more variable. Given the functional-constraint hypothesis (Wilson et al. 1977; Kimura 1983), this difference in conservation probably indicates that tropomyosin II is selectively much more constrained than is either tropomyosin I or the nonmuscle cytoplasmic form.

This study seems to indicate that two-dimensional electrophoresis is useful both as a tool for phylogenetic reconstruction and as a molecular evolutionary clock. Even though only a few loci were surveyed (about 50) and a long expanse of geological time separated some of the taxa (roughly 60 Myr), this technique accurately reconstructed the higher-level relationships within the genus *Drosophila*. No other biochemical study, except for this one, is fully congruent with the phylogeny of *Drosophila* as it is now understood. Furthermore, the molecular evolutionary clock derived from this study seems to work as well as those based on immunological data sets. In addition, two-dimensional electrophoresis may well be a useful tool for examining the molecular evolution of insoluble proteins. Consequently, although it is technically a difficult procedure, two-dimensional electrophoresis should prove to be an effective technique for addressing problems in evolutionary biology.

Acknowledgments. I am most grateful to Lynn Throckmorton and Alan Templeton for supplying the flies used in this study. I also thank Bruce Carnes, Brian Charlesworth, Jerry Coyne, Carol Giometti, Karen Haugen, Scott Lanyon, Steve Orzack, Dave Townsend, and Jeff Vitale for their useful comments. I especially thank Dave Swofford for the use of his PAUP program and for the helpful discussions on data analysis. I also thank Eric Zurcher for the VAX 11/780 adapted version of NT-SYS. I am very appreciative to all the members of the Protein Mapping Group for assistance with the two-dimensional electrophoresis and to the Division of Educational Programs, Argonne National Laboratory, for their support. This work was supported by the U.S. Department of Energy, Office of Health and Environmental Research, under Contract no. W-31-109-ENG-38.

References

Adams EN (1972) Consensus techniques and the comparison of taxonomic trees. Syst Zool 21:390–397

Anderson NG, Anderson NL (1978a) Analytical techniques for cell fractions. XXI. Two-dimensional analysis of serum and tissue proteins: multiple isoelectric focusing. Anal Biochem 85:331–340

Anderson NL, Anderson NG (1978b) Analytical techniques for cell fractions. XXII. Two-dimensional analysis of serum and

tissue proteins: multiple gradient-slab electrophoresis. Anal Biochem 85:341–354

Arnold ML, Baker RJ, Honeycutt RL (1983) Genic differentiation and phylogenetic relationships within two New World bat genera. Biochem Syst Ecol 11:295–303

Avise JC (1975) Systematic value of electrophoretic data. Syst Zool 23:465–481

Avise JC (1983) Protein variation and phylogenetic reconstruction. In: Oxford GS, Rollinson D (eds) Protein polymorphism: adaptive and taxonomic significance. Academic Press, London, pp 103–130

Avise JC, Aquadro CF (1982) A comparative summary of genetic distances in the vertebrates. Evol Biol 15:151–185

Avise JC, Patton JC, Aquadro CF (1980) Evolutionary genetics of birds. I. Relationships among north American thrushes and allies. Auk 97:135–147

Ayala FJ (1982) Of clocks and clades, or a story of old told by genes of now. In: Nitecki MH (eds) Biochemical aspects of evolutionary biology. University of Chicago Press, Chicago, pp 257–301

Ayala FJ (1986) On the virtues and pitfalls of the molecular evolutionary clock. J Hered 77:226–235

Bautch VL, Storti RV (1983) Identification of a cytoplasmic tropomyosin gene linked to two muscle tropomyosin genes in *Drosophila*. Proc Natl Acad Sci USA 80:7123–7127

Bautch VL, Storti RV, Mischke D, Pardue ML (1982) Organization and expression of *Drosophila* tropomyosin genes. J Mol Biol 162:231–250

Baverstock PR, Cole SR, Richardson BJ, Watts CHS (1979) Electrophoresis and cladistics. Syst Zool 28:214–219

Berlocher SH (1984) Insect molecular systematics. Annu Rev Entomol 29:403–433

Beverley SM, Wilson AC (1982) Molecular evolution in *Drosophila* and the higher Diptera. I. Micro-complement fixation studies of a larval hemolymph protein. J Mol Evol 18:251–264

Beverley SM, Wilson AC (1984) Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J Mol Evol 21:1–13

Beverley SM, Wilson AC (1985) Ancient origin for Hawaiian Drosophilinae inferred from protein comparisons. Proc Natl Acad Sci USA 82:4753–4757

Bush GL, Kitto GB (1978) Application of genetics to insect systematics and analysis of species differences. In: Romberger JA, Foote RH, Knutson L, Lentz PD (eds) Beltsville symposia in agricultural research, vol 2. Wiley, New York, pp 89–118

Buth DG (1984) The application of electrophoretic data in systematic studies. Annu Rev Ecol Syst 15:501–522

Collier GE, MacIntyre RJ (1977) Microcomplement fixation studies on the evolution of α-glycerophosphate dehydrogenase within the genus *Drosophila*. Proc Natl Acad Sci USA 74:684–688

Coulthart MB (1986) Variation and evolution in proteins of the *Drosophila* male reproductive tract. Thesis, McMaster University, Hamilton, Ontario

Coyne JA, Eanes WF, Ramshaw JAM, Koehn RK (1979) Electrophoretic heterogeneity of α-glycerophosphate dehydrogenase among many species of *Drosophila*. Syst Zool 28:164–175

Cracraft J (1983) Cladistic analysis and vicariance biogeography. Am Sci 71:273–281

Dickerson RE (1971) The structure of cytochrome c and the rates of molecular evolution. J Mol Evol 1:26–45

Duke EJ, Glassman E (1968) Evolution of xanthine dehydrogenase in *Drosophila*. Genetics 58:101–112

Entingh TD (1970) DNA hybridization in the genus *Drosophila*. Genetics 66:55–68

Farris JS (1972) Estimating phylogenetic trees from distance matrices. Am Nat 106:645–668

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Fine RE, Blitz AL (1975) A chemical comparison of tropomyosins from muscle and non-muscle tissues. J Mol Biol 95: 447–454

Firtel RA (1981) Multigene families encoding actin and tubulin. Cell 24:6–7

Fitch WM (1976) Molecular evolutionary clocks. In: Ayala FJ (ed) Molecular evolution. Sinauer, Sunderland MA, pp 160–178

Fyrberg EA (1984) Structural and functional analyses of *Drosophila melanogaster* actin genes. In: Maclean N (ed) Oxford surveys of eukaryotic genes. Oxford University Press, Oxford, pp 61–86

Fyrberg EA, Kindle KL, Davidson N (1980) The actin genes of *Drosophila*: a dispersed multigene family. Cell 19:365–378

Fyrberg EA, Mahaffey JW, Bond BJ, Davidson N (1983) Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell 33:115–123

Goodman M (1976) Protein sequences in phylogeny. In: Ayala FJ (ed) Molecular evolution. Sinauer, Sunderland MA, pp 141–159

Goodman M (1981) Decoding the pattern of protein evolution. Prog Biophys Mol Biol 37:105–164

Hennig W (1966) Phylogenetic systematics. University of Illinois Press, Urbana, p 263

Hightower RC, Meagher RB (1986) The molecular evolution of actin. Genetics 114:315–332

Hillis DM (1985) Evolutionary genetics of the Andean lizard genus *Pholidobolus* (Sauria: Gymnophthalmidae): phylogeny, biogeography, and a comparison of tree construction techniques. Syst Zool 34:109–126

Hillis DM, Frost JS, Wright DA (1983) Phylogeny and biogeography of the *Rana pipiens* complex: a biochemical evaluation. Syst Zool 32:132–143

Honeycutt RL, Williams SL (1982) Genic differentiation in pocket gophers of the genus *Pappogeomys*, with comments on intergeneric relationships in the subfamily Geomyinae. J Mammal 63:208–217

Jungblut P, Klose J (1985) Genetic variability of proteins from mitochondria and mitochondrial fractions of mouse organs. Biochem Genet 23:227–245

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, p 367

Klose J (1982) Genetic variability of soluble proteins studied by two-dimensional eletrophoresis on different inbred mouse strains and on different mouse organs. J Mol Evol 18:315–328

Klose J, Feller M (1981) Genetic variability of proteins from plasma membranes and cytosols of mouse organs. Biochem Genet 19:859–870

Lanyon SM (1985) Molecular perspective on higher-level relationships in the Tyrannoidea (Aves). Syst Zool 34:404–418

Lee TJ, Pak JH (1986) Biochemical phylogeny of the *Drosophila auraria* complex. Drosophila Inf Serv 63:81

Leigh Brown AJ, Langley CH (1979) Reevaluation of level of genic heterozygosity in natural population of *Drosophila melanogaster* by two-dimensional electrophoresis. Proc Natl Acad Sci USA 76:2381–2384

Lundberg JG (1972) Wagner networks and ancestors. Syst Zool 18:1–32

MacIntyre RJ, Collier GE (1986) Protein evolution in the genus *Drosophila*. In: Ashburner M, Carson HL, Thompson JN (eds) The genetics and biology of *Drosophila*, vol 3e. Academic Press, London, pp 39–146

MacIntyre RJ, Dean MR (1978) Evolution of acid phosphatase-1 in the genus *Drosophila* as estimated by subunit hybridization. Interspecific tests. J Mol Evol 12:143–171

MacIntyre RJ, Dean MR, Batt G (1978) Evolution of acid

phosphatase-1 in the genus *Drosophila*. Immunological studies. J Mol Evol 12:121–142
Matson RH (1984) Applications of electrophoretic data in avian systematics. Auk 101:717–729
Maxson LR, Maxson RD (1979) Comparative albumin and biochemical evolution in plethodontid salamanders. Evolution 33:1057–1062
McLellan T, Inouye LS (1986) The sensitivity of isoelectric focusing and electrophoresis in the detection of sequence differences in proteins. Biochem Genet 24:571–577
McLellan T, Ames GF, Kikaido K (1983) Genetic variation in proteins: comparison of one-dimensional and two-dimensional gel electrophoresis. Genetics 104:381–390
Mickevich MF (1978) Taxonomic congruence. Syst Zool 27: 143–158
Mickevich MF, Johnson MS (1976) Congruence between morphological and allozyme data in evolutionary inference and character evolution. Syst Zool 25:260–270
Mickevich MF, Mitter C (1981) Treating polymorphic characters in systematics: a phylogenetic treatment of electrophoretic data. In: Funk VA, Brooks DR (eds) Advances in cladistics. New York Botanical Garden, New York, pp 45–58
Miyamoto MM (1981) Congruence among character sets in phylogenetic studies of the frog genus *Leptodactylus*. Syst Zool 30:281–290
Mogami K, Fujita SC, Hotta Y (1982) Identification of *Drosophila* indirect flight muscle myofibrillar proteins by means of two-dimensional electrophoresis. J Biochem 91:643–650
Neel JV, Baier L, Hanash S, Erickson RP (1985) Frequency of polymorphisms for alleles encoding for liver proteins of domesticated mice. J Hered 76:314–320
Nei M (1972) Genetic distances between populations. Am Nat 106:283–292
Nevo E (1978) Genetic variation in natural populations: patterns and theory. Theor Pop Biol 13:121–177
O'Brien SJ, MacIntyre RJ (1978) Genetics and biochemistry of enzymes and specific proteins of *Drosophila*. In: Ashburner M, Wright TRF (eds) The genetics and biology of *Drosophila*, vol 2a. Academic Press, London, pp 395–551
O'Farrell PH (1975) High resolution two-dimensional electrophoresis. J Biol Chem 250:4007–4021
Ohnishi S, Leigh Brown AJ, Voelker RA, Langley CH (1982) Estimation of genetic variability in natural populations of *Drosophila simulans* by two-dimensional and starch gel electrophoresis. Genetics 100:127–136
Ohnishi S, Kawanishi M, Watanabe TK (1983a) Biochemical phylogenies of *Drosophila*: protein differences detected by two-dimensional electrophoresis. Genetica 61:55–63
Ohnishi S, Kim K, Watanabe TK (1983b) Biochemical phylogeny of the *Drosophila montium* species subgroup. J Jpn Genet 58:141–151
Patton JC, Avise JC (1983) An empirical evaluation of qualitative Hennigian analyses of protein electrophoretic data. J Mol Evol 19:244–254
Powell JR (1975) Protein variation in natural populations of animals. Evol Biol 8:79–119
Rohlf FJ (1982) Consensus indices for comparing classifications. Math Biosci 59:131–144
Rohlf FJ, Kishpaugh J, Kirk D (1981) Numerical taxonomy system of multivariate statistical programs (NT-SYS). State University of New York, Stony Brook
Sarich VM (1977) Electrophoresis in evolutionary studies: rates, sample sizes, and the neutrality hypothesis. Nature 265:24–28
Sites JW, Greenbaum IF, Bickham JW (1981) Biochemical systematics of neotropical turtles of the genus *Rhinoclemmys* (Emydidae: Batagurinae). Herpetologica 37:256–264
Sites JW, Bickham JW, Pytel BA, Greenbaum IF, Bates BA (1984) Biochemical characters and the reconstruction of turtle phylogenies: relationships among batagurine genera. Syst Zool 33:137–158
Skibinski DOF, Ward RD (1981) Relationship between allozyme heterozygosity and rates of divergence. Genet Res 38: 71–92
Skibinski DOF, Ward RD (1982) Correlations between heterozygosity and evolutionary rate of proteins. Nature 298:490–492
Sneath PHA, Sokal RR (1973) Numerical taxonomy. WH Freeman, San Francisco, p 573
Sokal RR, Rohlf FJ (1981) Taxonomic congruence in the Leptopodomorpha re-examined. Syst Zool 30:309–325
Sokal RR, Sneath PHA (1963) Principles of numerical taxonomy. WH Freeman, San Francisco, p 359
Spicer GS (1985) Systematics of the *Drosophila virilis* species group as assessed by two-dimensional electrophoresis. Thesis, Texas Tech University, Lubbock
Storti RV, Horovitch SJ, Scott MP, Rich A, Pardue ML (1978) Myogenesis in primary cell cultures from *Drosophila melanogaster*: protein synthesis and actin heterogeneity during development. Cell 13:589–598
Swofford DL (1982) Consensus tree program (CONTREE) for ANSII standard Fortran 77. Illinois Natural History Survey, Champaign
Swofford DL (1984) Phylogenetic analysis using parsimony (ver 2.3.1). Illinois Natural History Survey, Champaign
Takahata N, Nei M (1985) Gene genealogy and variance of interpopulational nucleotide differences. Genetics 110:325–344
Thorpe JP (1982) The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. Annu Rev Syst Ecol 13:139–168
Throckmorton LH (1962) The problem of phylogeny in the genus *Drosophila*. Univ Tex Publ 6205:207–343
Throckmorton LH (1969) Concordance and discordance of taxonomic characters in *Drosophila* classification. Syst Zool 17:355–387
Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed) Handbook of genetics, vol 3. Plenum, New York, pp 421–469
Throckmorton LH (1978) Molecular phylogenetics. In: Romberger JA, Foote RH, Knutson L, Lentz PD (eds) Beltsville symposia in agricultural research, vol 2. Wiley, New York, pp 221–239
Throckmorton LH (1982) Pathways of evolution in the genus *Drosophila* and the founding of the *repleta* group. In: Barker JFS, Starmer WT (eds) Ecological genetics and evolution. Academic Press, Australia, pp 33–47
Tobin SL, Zulauf E, Sanchez F, Craig EA, McCarthy BJ (1980) Multiple actin-related sequences in the *Drosophila melanogaster* genome. Cell 19:121–131
Tollaksen SL, Anderson NL, Anderson NG (1984) Operation of the ISO-DALT system, ed 7. Argonne National Laboratory Report Publ ANL-BIM-84-1, Argonne National Laboratory, Argonne IL
Vilageliu L, Gonzalez-Duarte R (1984) Alcohol dehydrogenase from *Drosophila funebris* and *Drosophila immigrans*: molecular and evolutionary aspects. Biochem Genet 22:797–815
Wake DB, Maxson LR, Wurst GZ (1978) Genetic differentiation, albumin evolution, and their biogeographic implications in plethodontid salamanders of California and southern Europe. Evolution 32:529–539
Watterson GA (1985) Estimating species divergence times using multi-locus data. In: Ohta T, Aoki K (eds) Population genetics and molecular evolution. Springer-Verlag, Berlin, pp 163–183
Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639

Received July 28, 1987/Revised and accepted October 25, 1987