PROTEIN DIFFERENCES IN DROSOPHILA. II. COMPARATIVE SPECIES GENETICS AND EVOLUTIONARY PROBLEMS[1]

JOHN L. HUBBY AND LYNN H. THROCKMORTON

Department of Zoology, The University of Chicago, Chicago, Illinois

Received April 14, 1965

FOR many years biologists have been challenged by problems of comparative species genetics (STURTEVANT and NOVITSKI 1941; SPENCER 1949). Questions concerning the amount of genetic identity between species and the amount of genetic reorganization accompanying speciation have been subjects of much controversy. Most attempts to provide the estimates necessary for resolving these issues have been frustrated because fertile crosses between organisms are required to demonstrate alleles. However, technical advances in protein chemistry (see CANFIELD and ANFINSON 1963) coupled with refinements in theories of genes and gene action (YANOFSKY, CARLTON, GUEST, HELINSKI, and HENNING 1964) now make it possible to demonstrate the existence of alleles in related species even when fertile crosses cannot be made. If current theories of gene action are taken as reasonable operational assumptions, it may be possible to obtain a direct estimate of gene change or stability through analyzing the primary structure of proteins, which reflects the DNA sequence of structural genes.

Most existing studies of protein homology have been made at the broadest levels. Such investigations (insulin—SANGER 1956; ribonuclease—ANFINSON, ÅQUIST, COOKE, and JÖNSSON 1959; SMYTH, STEIN, and MOORE 1963; hemoglobin —INGRAM 1963; cytochrome c—MARGOLIASH 1963) have shown that the basic integrity of the gene has been maintained over long periods of time. The organisms studied in these instances (cow, pig, whale, man, etc.) are only distantly related and the intimate details of their ancestry have never been fully demonstrated. These studies have been, for the most part, studies of single proteins. As such, they can provide information regarding the differences between loci among related species. If we would achieve an understanding of the evolutionary history of gene pools, however, we must study a number of proteins simultaneously. A first approach to the study of large numbers of proteins has been made possible through the use of techniques developed to study proteins from Drosophila (HUBBY 1963).

The purpose of this paper is to present evidence derived from electrophoretic analysis of over 360 proteins from species of the virilis group of the genus Drosophila. The species in the virilis group represent forms whose morphological divergence has been minimal (they are a group of nine sibling and near-sibling species) and whose ancestry has been reconstructed from analysis of chromo-

[1] This research was supported in part by grants from the National Science Foundation (G 19085 and GB 630) and the Public Health Service (GM 11216).

somal rearrangements (STONE, GUEST, and WILSON 1960). Thus, when information concerning the evolutionary history of the group is combined with information derived from protein analyses, some interesting, though tentative, conclusions can be drawn regarding the amount of genetic change involved during the evolution of these species.

MATERIALS AND METHODS

Laboratory stocks of species were obtained from the Genetics Foundation of the University of Texas. The reference designations of the stocks are University of Texas accession names and numbers. The forms studied are: *virilis* (Pasadena), *novamexicana* (1714.4), *americana americana* (Anderson, Indiana), *americana texana* (1128.1), *littoralis* (2096), *ezoana* (2531.1), *montana* (1942.6), *lacicola* (1756.26), *borealis* (2077.5a), and *flavomontana* (1951.3). Of these *a. americana* and *a. texana* are subspecies, the remainder are species. Thus, with the exception of *americana*, only one strain has been investigated for each species.

Methods for culture and collection of flies, fractionation procedures, and zone electrophoresis of protein samples have all been described (HUBBY 1963). For the present study the fractionation procedure was changed to the following: all material precipitable in 45% saturated ammonium sulfate solution was discarded; the remaining material was precipitated in four fractions. These were: 45 to 60%, 60 to 75%, 75 to 85%, and 85 to 100% saturation. Bovine serum albumin was applied to each gel as a reference standard. All mobilities were calculated relative to the mobility of this protein. Protein bands were visualized by staining with Analine Blue Black. As many as ten independent fractionations were made of some species. All analyses are based on no fewer than six fractionations. At least four gels were analyzed from each fractionation.

Several methods for specific enzyme detection on acrylamide gels have been published (FINE and COSTELLO 1963) or have been developed in this laboratory (SIMS 1965). The assays used for the enzyme data reported herein are those of SIMS, and they detect alpha-glycerolphosphate dehydrogenase and malic dehydrogenase.

RESULTS

Each of the four protein fractions had distinct characteristics. There was a great deal of similarity between similar fractions from the nine different species. Figure 1 illustrates the four fractions. The diagrams are a composite from all fractionations and from the several gels made from each fractionation. Proteins that stained faintly and appeared inconsistently have been omitted from the figure. In each comparison of fractions between two different species, proteins can be classed either as shared or not shared. A species may have a protein with the same electrophoretic mobility as that found in another species or it may not. Using this criterion the data from the nine species has been reorganized for presentation in Figure 2. In this figure the proteins are regrouped without regard for fraction or relative mobility. When two or more species share proteins with identical mobilities, these proteins are shown at the same horizontal level in the appropriate species columns. The sequence of proteins in the columns is determined in the following way. Those proteins that had mobilities different from the mobilities of all other proteins from the same fraction from all other species are grouped at the top and are designated as *unique*. Proteins that shared identical mobilities with proteins from one or more species belonging to the same cytological phylad are grouped next and are designated *phylad*. Proteins that

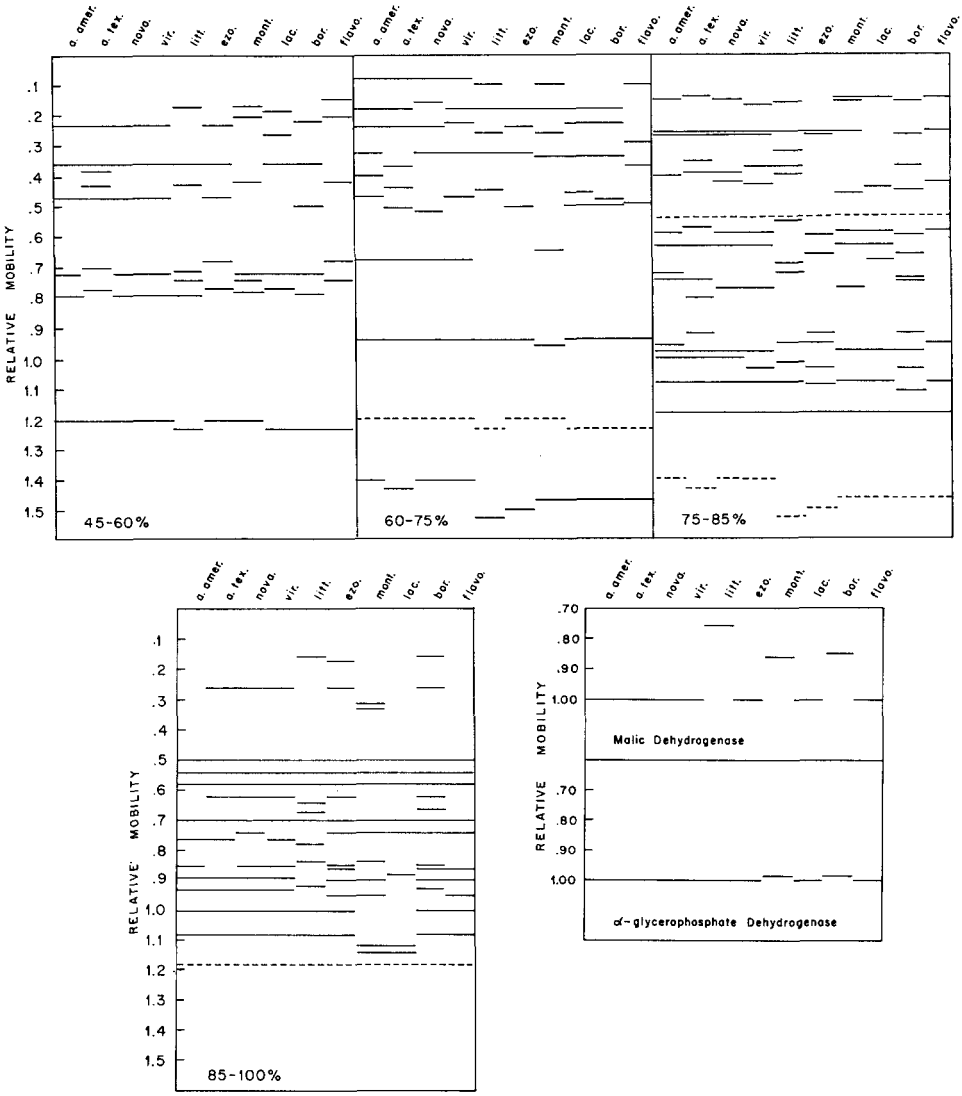

FIGURE 1.—(Top and left) Relative mobilities of proteins from members of the virilis group derived from electrophoresis in acrylamide gels. Dashed lines indicate proteins which appear in two adjacent fractions. (Lower right) The relative mobilities of α-glycerolphosphate dehydrogenase and malic dehydrogenase from members of the virilis group derived from electrophoresis in acrylamide gels. Mobilities are calculated using *virilis* as standard. Only substrate specific bands are shown.

shared identical mobilities with proteins from species in *both* cytological phylads are designated as *ancestral*. The grouping of species into two cytological phylads is based on STONE, GUEST, and WILSON (1960). Their cytological phylogeny for the group is shown in Figure 4a. We have interpreted the virilis phylad as including *virilis*, *novamexicana*, and *americana*, with its two subspecies, *a. ameri-*

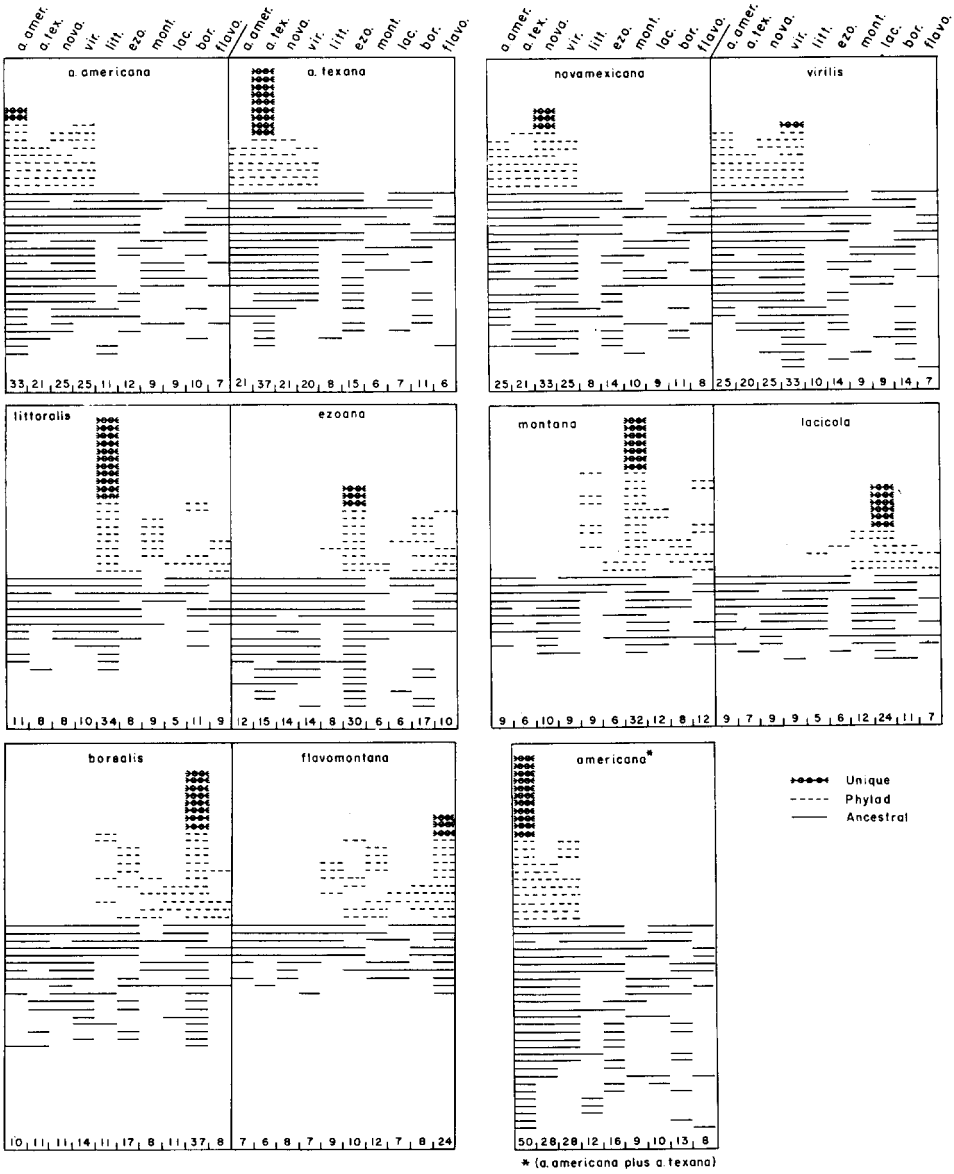

FIGURE 2.—The distribution of proteins from species of the virilis group. Proteins that are unique to a species appear at the top, those restricted to members of a phylad are in the middle, and those occurring in members of both phylads are at the bottom. Proteins in adjacent diagrams are not necessarily in the same order.

cana and *a. texana*. The montana phylad may be broken down into two subphylads on cytological grounds. One subphylad includes *ezoana* and *littoralis*, the other the remaining four species. This latter subdivision is not strongly reflected in our protein data so we have not attempted to indicate it in Figure 2.

There are five proteins that are shared by all species. For compactness these have been omitted from Figure 2, although they could have been included there among ancestrals. In Figure 2 the sequence of proteins in adjacent diagrams is not necessarily the same. The numbers at the bottom of each column are the number of proteins that each species has in common with the species named in the diagram. The first diagram (Figure 2, upper left) shows the way in which the total complement of proteins from *a. americana* is shared by other species in the group, identifying shared proteins on the basis of identical mobilities. Thus, *a. americana* has two proteins unique to it, nine in common with one or more species from its cytological phylad, and 22 in common with one or more species from both cytological phylads. For convenience the subspecies, *a. americana* and *a texana*, are treated separately. However, it was of interest to note the consequences of pooling the information from these two subspecies in order to obtain a composite for the species, *americana*. This composite is also included in Figure 2 (toward the lower right).

The results from the enzyme studies are shown to the lower right in Figure 1. Malic dehydrogenase exists in four electrophoretic variants among these species, with one variant common to *americana*, *novamexicana*, *virilis*, *ezoana*, *lacicola*, and *flavomontana*. Each of the remaining species, *littoralis*, *montana*, and *borealis*, has its own unique electrophoretic variant of this enzyme. Alpha-glycerolphosphate dehydrogenase exists in only two electrophoretic variants among these species. One form is common to *montana* and *borealis*, the other to the remaining species.

DISCUSSION

The basic premise for the discussion that follows is that the amino acid sequences in proteins are colinear with the nucleotide sequence in the DNAs of the structural genes. Hence, an analysis of a protein is, for all practical purposes, an analysis of a gene. The clearest direct proof of colinearity has been obtained by YANOFSKY *et al.* (1964) in their researches on tryptophan synthetase in *E. coli*. Except in the special case of degeneracy of the amino acid code, colinearity is expected to be exact. With degeneracy, analysis of protein structure will underestimate the amount of genetic change. Protein studies allow information to be obtained only from those elements of the genome that specify protein structure, and these may not represent all genetic elements. They comprise, however, a sufficiently large category so that their analysis may provide useful information concerning evolutionary changes in genes and gene pools.

Ideally, protein studies should include determination of amino acid sequence if direct inferences are to be made regarding the DNAs that are reflected by these proteins. Such studies are expensive and time-consuming, and, as a first approximation, informative results can be obtained by simpler procedures. Zone electrophoresis in a supporting medium such as polyacrylamide is a sensitive method for detecting minor differences in overtly identical molecules. The most readily detectable differences revealed by this method are amino acid substitutions resulting in a charge difference in a protein molecule, either because the substituted

amino acid itself carries the different charge or because its substitution results in a configurational change in the molecule and a consequent over-all net charge change due to covering or uncovering ionizable groups. HENNING and YANOFSKY (1963) found that seven out of nine mutationally altered forms of tryptophan synthetase show differences in migration rate in polyacrylamide, and it seems probable that a similar large proportion of all single amino acid substitutions can be detected by this means. When evolutionary comparisons are made, even among closely related forms, double or higher order substitutions may be expected in addition to single substitutions. A second substitution may occur at the same site as the first or at a second site. At present we are unable to predict the respective probabilities of such events as they would affect a sample such as ours. Second site substitutions or substitutions of greater rank may often produce configurationally altered proteins. Such proteins may not often have the same mobility when compared with an unaltered molecule, but no systematic investigation is available from which to draw conclusions.

In the present study, two kinds of discriminating procedures have been employed. The first, ammonium sulfate fractionation, separates the proteins only into major solubility classes. The second, zone electrophoresis, subdivides these classes according to the mobility of the proteins in polyacrylamide. The latter makes a much more critical distinction between proteins, but it still affords the opportunity for different molecules to migrate at the same rate. Thus, proteins in the same ammonium sulfate fraction and having the same electrophoretic mobility may be of three general kinds. (1) At one extreme, will be those whose identical migration rates are entirely fortuitous. Their DNAs will have little in common and will not reflect homology, much less gene identity. (2) A second category would include those proteins reflecting phylogenetically related and thus homologous DNAs (alleles) but not having identical amino acid sequences. The proteins in the first and second categories would lead to spuriously high estimates of genetic identity between species if they were included in the sample. (3) The last category would contain the identical protein molecules. It is these proteins that will give us a direct indication of genetic identity between species.

One may undertake to estimate the proportion of spurious identities in several ways, none completely satisfactory. Table 1 presents the average number of proteins of identical mobility shared by each pair of species. It was necessary to use an average figure for each species pair since these are calculated as percent total proteins for each species and few species pairs had the same total number of proteins. The species are arranged in the table according to their cytological relationships. In the lower left part of the table, are shown the average number of identities by phylads and subphylads. Thus, species within the virilis phylad share, on the average, 71.6% of their proteins with each other, 43.8% of their proteins with members of the littoralis subphylad, and 38.3% of their proteins with members of the montana subphylad. There is, then, a direct correlation between relationship and the number of proteins having identical mobilities that are shared between species. If, now, we make the rather extreme assumption that it is unlikely that species in the montana subphylad share *any* identical proteins

TABLE 1

Average numbers of identities between pairs of species, as percent total proteins for the species

	a.amer	*a.tex*	*nova*	*vir*	*litt*	*ezo*	*mont*	*lac*	*bor*	*flavo*
a.amer		65.2	78.9	78.9	41.6	46.8	37.3	42.6	37.6	36.5
a.tex			65.2	62.7	32.1	52.4	28.0	35.0	38.1	32.1
		71.6								
nova				78.9	33.8	52.2	40.0	42.6	40.0	39.5
vir					39.0	52.2	37.3	42.6	47.6	36.5
litt						35.2	36.9	30.1	39.6	42.1
		43.8			**35.2**					
ezo							30.6	34.7	57.7	47.3
mont								52.3	33.1	52.3
lac									46.7	41.4
		38.3			**39.9**		**44.0**			
bor										37.9
flavo										

Average number of identities by phylads and subphylads are shown in bold face type in the lower left portion of the table.

with species in the virilis phylad, we could conclude that on the average 38.3% of the identical mobilities were fortuitous. This figure would estimate the sum of the first two categories mentioned earlier. It could represent the nonhomologous proteins having identical mobilities, plus the homologous but nonidentical proteins having identical mobilities. The latter category might be estimated independently from the observation of HENNING and YANOFSKY (1963) that 7/9 (~78%) of all single substitutions result in proteins with changed electrophoretic mobilities. Thus, about 22% of the homologous proteins having identical mobilities might have different amino acid sequences. If 38% equals category one plus category two, and 22% equals category two, then 16% (38–22) equals category one, and 62% equals category three. It is probable that 62% is an underestimate of the proportion of proteins with identical mobilities that are indeed identical, but this figure can serve as a general indication of the possible error in our estimates.

From the standpoint of characterizing a species, analysis of proteins from a single strain will not give complete information regarding that species. Protein polymorphisms from other species are known. Over 100 laboratory stocks of *D. melanogaster* have been analyzed and six protein differences have been recorded (HUBBY 1963). If the species of the virilis group are subject to this much protein polymorphism, or to more, the analysis of single strains will give somewhat biased results. The extent of the bias will depend in part on the extent of polymorphism. If the species has little genetic variability in its gene pool, a single sample will represent it quite well. If there is much variability the sample may be quite in-

complete since, as a rule, only one of several possible alleles would be included in the sample. The more alleles included in the sample, the greater the possibility for increasing both the fortuitous and the true identities. Thus, while our method of estimating genetic identity tends toward an overestimate, the limitation in sample numbers tends to bias toward an underestimate. The two biases may not completely cancel each other, but we cannot be certain which more strongly influences our results. Under the circumstances, it seems reasonable to accept the data provisionally as giving a good approximation of the amount of genetic divergence that has occurred during the evolution of this species group. In the discussion that follows, identical protein mobilities will be treated as indicating genetic identity. However, the reader may convert these data to *minimal* estimates of genetic identity by introducing a factor somewhere between 60 and 80%.

Table 2 is a summary of the distribution of proteins among species in the virilis group. The numbers in the column headed Total Proteins include the five bands common to all species but not shown in Figure 2. The column headed Unique shows the percent total proteins for a species that are restricted to that species. The column headed Phylad shows the percent total proteins that are shared with other members of the same cytological phylad but not with members

TABLE 2

Summary of the distribution of the proteins among species of the virilis group

		Percent total proteins			
				Ancestral*	
	Total proteins	Unique	Restricted to phylad	minimum	maximum
Virilis phylad					
americana†	55	20.4	20.4	40.8	59.3
a. americana	38	5.3	23.7	55.3	71.1
a. texana	42	21.4	16.7	42.9	61.9
novamexicana	38	7.9	21.1	57.9	71.1
virilis	38	2.6	21.1	57.9	76.3
Average, virilis phylad	39.0	9.3	20.7	53.5	70.1
Montana phylad					
littoralis	39	28.2	25.6	38.5	46.2
ezoana	35	8.6	25.7	44.4	65.7
montana	37	18.9	37.8	37.8	43.2
lacicola	29	20.7	20.7	48.3	58.6
borealis	42	19.0	28.6	26.2	40.5
flavomontana	29	10.3	37.9	41.4	51.7
Average, montana phylad	35.0	17.6	29.4	39.4	51.0
Average, virilis group	36.6	14.3	25.9	45.1	58.6

* The minimum estimate of ancestral proteins is based on proteins that have identical mobilities in at least two species from each phylad. The maximum estimate is based on all proteins that occur at least once in each phylad.

† The composite of *a. americana* and *a. texana*. Values for the composite are not included in the phylad and group averages.

of the other phylad. The next two columns are estimates of the amount of ancestral genetic material retained by each species. The term "ancestral" is used here in a somewhat restricted sense. It refers to proteins (and by extension to alleles) from the common ancestor for the species group, although alleles from the ancestor of a phylad are also "ancestral" insofar as a species is concerned. The maximum estimate of ancestral alleles is based on all proteins that occur at least once in members of both phylads. Since single, distant recurrences (as the protein common only to *a. texana* and *flavomontana*, Figure 2) might represent recurrence through mutation rather than through retention from a common ancestor, a second, minimum estimate was made. This is based only on proteins that appear in at least two species from each phylad. These columns are averaged by phylad and for the group as a whole.

From Table 2 it can be seen that, as a group average, a species retains from 45 to 60% of the genetic material from the common ancestor for the species group. An additional 25% is retained from the common ancestor for a phylad, and only 15% is unique to a species. Both the individual species and the two phylads may differ rather markedly from each other in each of the major categories. It may be noted in passing that species differences cannot be interpreted solely as the genetic divergence that occurred during speciation. Instead, these differences represent changes during speciation plus those changes added since the form achieved full species status. In this study the minimum estimate (2.6%) for species divergence is that for *virilis*, if the measure for divergence is taken as the percent Uniques. However, the uniqueness of a species is not really an appropriate measure of the degree of divergence accompanying speciation. The percentage of the ancestral gene pool that is altered during and following speciation should be a better measure, and such estimates can be made from Table 1. The minimum amount of divergence (maximum percent identity) is between *virilis*, *a. americana*, and *novamexicana*, where 78.9% of their sampled proteins are in common. The other species have considerably less in common, a feature that possibly reflects, in part, some of the time intervals involved. These data suggest that the divergence of species within the virilis phylad may have been more recent than divergence within the montana phylad, although this is not necessarily the case. However that may be, the estimate of about 80% in common following speciation is of interest, and a comparable figure can be arrived at in a different manner, using data from the montana phylad as well as from the virilis phylad. This approach involves the partial reconstruction of the gene pools through which the alleles of the present species passed in order to achieve their present distribution. The results of this analysis are shown in Figure 3.

Figure 3 was constructed from the data in Figure 2 by a sequential pooling of the data from different evolutionary levels, as these are indicated by the cytological phylogeny of the group. Thus, when information from *montana*, *lacicola*, *borealis*, and *flavomontana* is pooled a composite is formed that may be treated as the hypothetical gene pool (AIV of Figure 3) for the ancestor for these four species. By convention, Uniques may be excluded from this hypothetical gene pool since they are present in only one species and they may have had a muta-

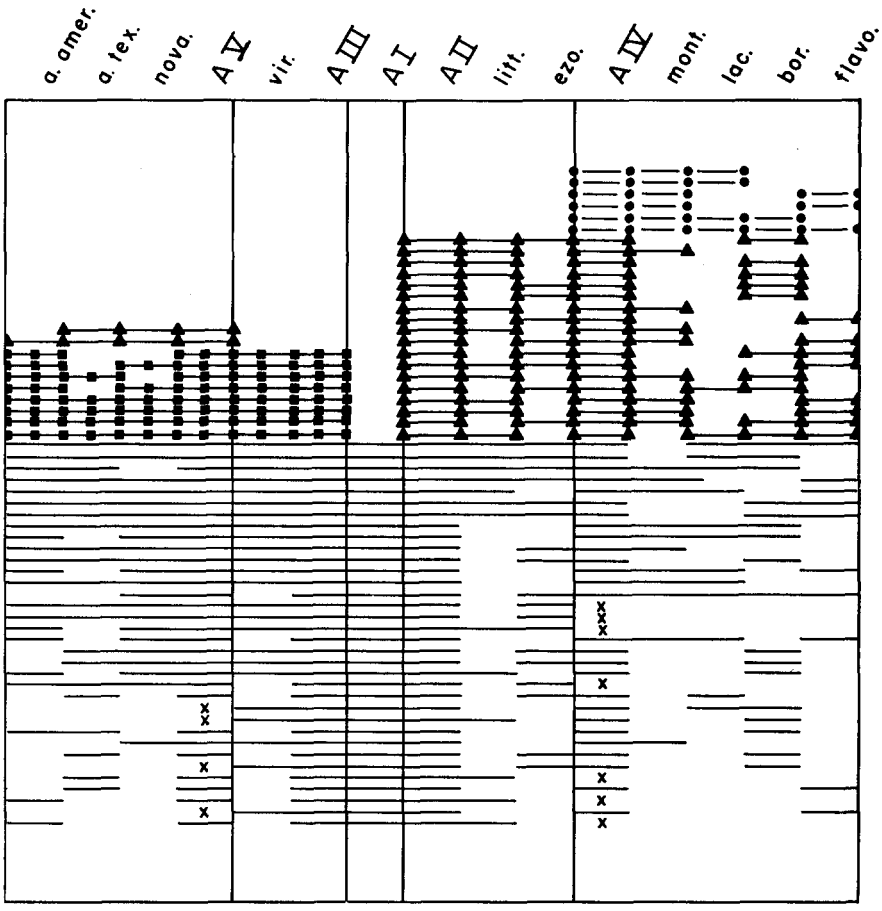

FIGURE 3.—Reconstruction of ancestral gene pools for species in the virilis group. Ancestral columns are headed by roman numerals preceded by A. Thus, AI indicates the ancestor for the group as a whole, AIII the ancestor for the virilis phylad, and so on. Unadorned lines toward the bottom of the figure indicate proteins derived from the ancestor for the species group. Lines adorned with solid circles indicate proteins present only in *montana*, *flavomontana*, *borealis*, and *lacicola*. Lines adorned with squares indicate proteins restricted to the virilis phylad. Lines adorned with triangles *to the left* of AI indicate proteins restricted to *americana* and *novamexicana*. Lines adorned with triangles *to the right* of AI indicate proteins restricted to the montana subphylad. With the exception of Ancestrals, proteins to the left and right of AI are not identical. Uniques are not included in the figure. The phylogenetic positions of the Ancestral populations are indicated in Figure 4.

tional origin within the species where they are detected. Next, the hypothetical aggregation (AIV) is pooled with data from *littoralis* and *ezoana* to derive a second hypothetical pool (AII), which would be that of the ancestor for the montana phylad. This would be the biochemical equivalent of primitive gene arrangement III (Figure 4a) of the cytological phylogeny. In this last instance pooling was not complete, since six of the proteins of the montana subphylad (AIV) are restricted to this subphylad and so, as the same convention used in

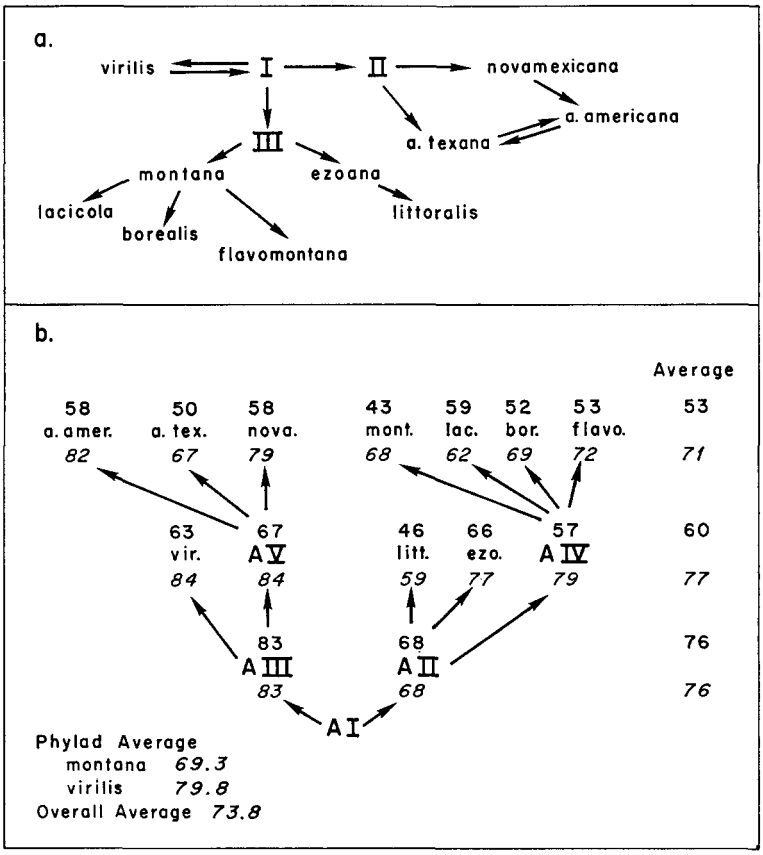

FIGURE 4.—(a) The cytological phylogeny of the virilis group (after STONE, GUEST, and WILSON 1960). The roman numerals indicate the primitive gene arrangements of these authors. (b) Species divergence during the evolution of the virilis group. The numbers above each name indicate the percent of the proteins that a species has in common with the ancestor of the species group (AI). The numbers under the names are the percent of the proteins that a species has in common with its immediate ancestor. To the right are shown the averages for the three evolutionary levels. To the lower left are shown these same figures averaged by phylad and for the group as a whole.

handling Uniques, are excluded from the next lower ancestral gene pool. The remainder of the figure is constructed in a similar manner.

This figure has several features of interest, and from it we can estimate the divergence that occurred at different evolutionary levels in time. This can be done by calculating the percent identity between the species and the hypothetical population from which they were most probably derived. Hypothetical populations themselves are also treated as species. The results of these calculations are shown in Figure 4b. The major skeleton of this figure indicates the most probable speciation events that occurred during the evolution of the group, except that *a. americana* and *a. texana* are, for simplicity, treated separately. (If the pooled data from these two subspecies are used, *americana* then has 97.8% of its proteins

in common with AV, suggesting that *americana is* AV.) In Figure 4 the italicized numbers *below* each name are the percent of proteins that species has in common with its immediate ancestor (*novamexicana* with AV, AV with AIII, AIII with AI, etc.). Thus, *novamexicana* has 79% of its proteins in common with Ancestral V, *virilis* has 84% of its proteins in common with Ancestral III, and Ancestral III has 83% of its proteins in common with the ancestor for the group as a whole (AI). The averages for each level are shown at the right and are roughly comparable (71, 77, 76) at each of the three levels. The averages of these values by phylad are shown to the lower left in the figure, as is the total over-all average. In general, these values are higher in the virilis phylad than in the montana phylad. Within the virilis group as a whole, there has been an average decrement of about 26% during the evolution of a species from its ancestor. This value is not too different from the estimate ($100-78.9 = 21.2\%$) that was derived from Table 1. Indeed, if the estimate for the virilis *phylad* from Table 1 (21.2%) is compared with the estimate for the virilis *phylad* from Figure 4 ($100-78.9 = 20.2\%$) these are found to be virtually identical.

The same general feature of the data can be shown in a second way by computing the percent of proteins that each species or hypothetical gene pool had in common with the ancestor (AI) for the species group. These numbers are shown *over* the names in Figure 4b, and their averages for each level are shown at the right of the figure. From these averages it can be seen that the percent loss of proteins common to the ancestor falls off in a rough progression (24, 16, 7) as the evolutionary levels are passed. This would be expected if the erosion of ancestral characters involved a relatively constant decrement during species divergence. However, there is a wide range (58 to 84) among the estimates for percent identity, and it seems unlikely that divergence has proceeded at a constant rate. As an average, 20 to 30% difference between ancestor and descendent may be a useful figure, but it must be accepted with reservation for the present.

In addition to providing a model for estimating species divergence, Figure 3 gives a simple picture of gene change that is quite compatible with what common sense suggests as the consequence of evolution. The loss of ancestral alleles is gradual and, from the standpoint of our being able to predict which will be lost, haphazard. At each succeeding level some new alleles are gained and some are lost (indicated by an X in the columns headed AV and AIV in Figure 3), but neither the gain nor the loss appears to be precipitous, even granting some corrections for fortuitous identical mobilities. Certainly the gene pool as a whole was not replaced during the descent of these forms. Changes in some lineages (compare AIII and its descendents with AII) seem to be more extreme than in others. This might be anticipated since we would not expect uniform selection pressures to impinge on all forms. Alternately, such differences may reflect different periods of time involved, but more probably they reflect the effects of adaptive stress and lapse of time, both intermingled in a way that cannot be unraveled by our experimental approach.

SUMMARY

Over 360 proteins from nine species of the virilis group of the genus Drosophila

have been compared using acrylamide gel electrophoresis. The study indicated apparent molecular identity extending through many members of the species group. The estimates were based on methods that might overestimate the amount of identity, but the number of samples involved would lead to an underestimate. The data indicate that approximately 60% of the sampled genetic material of these species was derived from the common ancestor of the species group, 25% was retained from the common ancestor of the subgroup (cytological phylad), and the remaining 15% was unique to the species. An analytical method is described by which it is possible to reconstruct, in part, some of the gene pools that existed early in the history of the species group. From this reconstruction it could be estimated that, on the average, species divergence within this group involved a change in about 20 to 30% of that part of the gene pool that was sampled.

LITERATURE CITED

ANFINSEN, C. B., S. E. G. ÅQUIST, J. P. COOKE, and B. JÖNSSON, 1959 A comparative study of the structure of Bovine and Ovine pancreatic ribonuclease. J. Biol. Chem. **234:** 1118–1123.

CANFIELD, R. E., and C. B. ANFINSEN, 1963 Concepts and experimental approaches in the determination of the primary structure of proteins. pp. 311–378. *The Proteins-Composition, Structure and Function*, Vol. 1. Edited by H. NEURATH. Academic Press, New York and London.

FINE, I. H., and L. A. COSTELLO, 1963 The use of starch electrophoresis in dehydrogenase studies. pp. 958–972. *Methods in Enzymology*, Vol. 6. Edited by S. P. COLOWICK and N. O. KAPLAN. Academic Press, New York and London.

HENNING, U., and C. YANOFSKY, 1963 An electrophoretic study of mutationally altered A proteins of the tryptophan synthetase of *Escherichia coli*. J. Mol. Biol. **6:** 16–21.

HUBBY, J. L., 1963 Protein differences in Drosophila. I. *Drosophila melanogaster*. Genetics **48:** 871–879.

INGRAM, V. M., 1963 *The Hemoglobins in Genetics and Evolution*. Columbia University Press, New York.

MARGOLIASH, E., 1963 Primary structure and evolution of cytochrome c. Proc. Natl. Acad. Sci. U.S. **50:** 672–679.

SANGER, F., 1956 The structure of insulin. pp. 434–459. *Currents in Biochemical Research*. Edited by D. E. GREEN. Interscience, New York.

SIMS, M., 1965 Methods for detection of enzymatic activity after electrophoresis on polyacrylamide gel in Drosophila species. Nature (In press.)

SMYTH, D. G., W. H. STEIN, and S. MOORE, 1963 The sequence of amino acid residues in Bovine pancreatic ribonuclease: revisions and confirmations. J. Biol. Chem. **238:** 227–234.

SPENCER, W. P., 1949 Gene homologies and the mutants of *Drosophila hydei*. pp. 23–44. *Genetics, Paleontology, and Evolution*. Edited by G. L. JEPSEN, G. G. SIMPSON, and E. MAYR. Princeton University Press, Princeton, N.J.

STONE, W. S., W. C. GUEST, and F. D. WILSON, 1960 The evolutionary implications of the cytological polymorphism and phylogeny of the virilis group of Drosophila. Proc. Natl. Acad. Sci. U.S. **46:** 350–361.

STURTEVANT, A. H., and E. NOVITSKI, 1941 The homologies of the chromosome elements in the genus Drosophila. Genetics **26:** 517–541.

YANOFSKY, C., B. C. CARLTON, J. R. GUSET, D. R. HELINSKI, and U. HENNING, 1964 On the colinearity of gene structure and protein structure. Proc. Natl. Acad. Sci. U.S. **51:** 266–272.