RESTRICTION ENZYME POLYMORPHISMS IN THE REGION OF THE SMALL HEAT STOCK GENES IN DROSOPHILA

RESTRICTION ENZYME POLYMORPHISMS IN THE REGION OF THE SMALL HEAT SHOCK GENES IN DROSOPHILA

By

DAVID R. TOWNSEND B.Sc.

A Thesis

Submitted to the School of

Graduate Studies in Partial

Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

December, 1982

Master of Science (1982) McMaster University

Hamilton, Ontario

Title: Restriction Enzyme Polymorphisms in the Region of the Small Heat Shock Genes in Drosophila

Author: David R. Townsend, B.Sc. (McMaster University)

Supervisor: Professor Rama S. Singh

Number of Pages: viv; 128

ABSTRACT

A survey of restriction enzyme polymorphisms was performed with fifteen iso-female lines of D. melanogaster amd eighteen single representatives of other Drosophila species. Three enzymes: Bam Hl, Eco Rl, and Pst 1 were used to probe the genetic structure of the region containing the genes Hsp 22, Hsp 23, Hsp 26, and Hsp 27. The results for within D. melanogaster show that all variation in the DNA sequence is limited to the non-coding region. The restriction patterns confirm the hypothesis that the hsp 22- 27 genes are a result of tandem duplications. The values for Nei's estimate of sequence diversity (δ) are 0.034 between populations of D. melanogaster, 0.113 between sibling species, and 0.123 between nonsibling species. These estimates were compared to values obtained for protein and enzyme variation. DNA sequence divergence between nonsibling species versus nonsibling species show less differentiation than protein and enzyme divergence. The restriction enzyme phenotype was used to generate phylogenies which is in approximate agreement with previously reported phylogenies. Molecular drive and selectionist hypotheses of differential rates of evolution during cladogenesis and anagenesis are discussed.

ACKNOWLEDGEMENTS

I wish to thank Dr. Rama S. Singh for encouragement, guidance, and patience throughout my time as his student. His help in all facets of my education have been invaluable. The encouragement of Dr. S. F. H. Threlkeld and Dr. R. A. Morton have been of great help during my time at McMaster University.

The technical assistance of Dr. Larry Moran and Dr. Thomas Chen and all the members of their laboratories is greatly appreciated. I wish to express my gratitude to Dr. Lorenz Rhomberg and Dr. J. McNeill for assistance in the use of the NT-SYS computer package and their help in the interpretation of the results.

The daily help, encouragement and friendship of Mr. M. Coulthart, Mr. B. Holwerda, and Mr. D. Morgan was greatly appreciated. Lastly the kindness and constant assistance of Mrs. S. Thomas was very much appreciated.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES CONTINUED

INTRODUCTION

The methods of studying genetic variation have changed greatly in this century and most extremely so in the last fifteen years. Starting in the nineteen thirties and continuing through to the nineteen fifties most studies of genetic variation were based on visible mutants segregating in the natural populations (Spencer 1947, Lewontin 1974 for review). The values obtained from these types of studies are approximately one half a visible mutant per genome. A second paradigm used for measuring variation at this time was to measure concealed variability of entire chromosomes for some fitness character (Dobzhansky and Spassky 1953, 1954). The concealed variability is usually expressed as percent lethal and sublethal chromosomes, whose effect is seen when the chromosomes are made homozygous. Some species, Drosophila prosaltans for example, have low levels of lethals and sublethals (10%), while others like D. melanogaster and D. persimilis have higher levels (approximately 30%). Both of these approaches have their drawbacks. The first depends on some large scale phenotypic change for its detection and there is no way to determine the number of genes involved in the variation in the second.

Lewontin and Hubby (1966) and Harris (1966) were the first to apply the technique of gel electrophoresis to population genetic studies. These authors and others in thousands of studies since then have used native gel electrophoresis of proteins to determine variation at strucural

gene loci. This technique seperates protein variants based on their conformational and charge characteristics. Extracts are prepared from individuals and placed on a porous gel (starch, polyacrylamide, agarose). An electrical potential is applied across the gel and proteins migrate at different rates, based on their amino acid composition and size. A dye system is used to stain proteins (nonspecifically) or catalytic enzymes (specifically) and therefore visualize the position of the electromorph. This technique is limited to structural genes with a stainable protein product. This will limit the number and type of loci that can be studied.

The study of genetic variability by electrophoresis is reviewed by Nevo (1978), Lewontin (1974) and Selander (1976). The heterozygosity per locus per individual ranges from 0.054 to 0.242 and the proportion of polymorphic loci, between 0.13 and 0.62 within the genus Drosophila (Nevo 1978). By the late nineteen seventies many studies have attempted to determine the amount of hidden electrophoretic variation which had not been resolved by standard gel electrophoresis technique. Utilizing either sequential gel electrophoresis, heat sensitivity, or other criteria of identity of proteins these electrophoretic alleles can be further subdivided (Singh 1979, 1982). This demonstrated one of the faults of standard gel electrophoresis, namely that it samples a subset of the variation in the protein product.

In the last five years a new experimental system for studying variation has become available. This approach is to determine the sequence variation of the polynucleotide that encodes the gene. Using restriction enzyme fragment length polymorphism and gene sequencing, researchers are now attempting to quantify the true genetic variation at the level of DNA.

DNA restriction enzymes recognize specific sequences in DNA and catalyse endonucleolytic cleavages. These cleavages result in fragments of defined length. These restriction fragments can be separated by size on agarose gels utlizing gel electrophoresis. Restriction enzyme cleavage sites are inherited as simple Mendelian codominant markers. Besides treating each site as a locus for analysis one can estimate sequence divergence between any two DNA sources by loss or gain of restriction sites.

As this technique does not depend on products of the gene it can be used on loci with a detectable protein or RNA product or loci with no known product. Previous to this technique only genes having a protein or abundant RNA product could be studied. The restriction enzyme recognizes only the primary DNA sequence and does not depend on coding or noncoding capabilities of the sequences. Restriction sites from coding and noncoding sequences can therefore be compared. Interesting evolutionary problems as codon usage, rates of transversions compared to transitions and rates of silent mutations can now be studied.

The object of this study is to characterize the sequence variation in the DNA encoding the small heat shock proteins in Drosophila. Heat shock protein genes are a group of genes which change in activity as a consequence of subjecting the organism to a wide variety of environmental stresses, including a brief heat shock (Ritossa 1962). A review of the heat shock response in Drosophila is described by Ashburner and Bonner (1979).

In Drosophila melanogaster seven heat shock proteins are known.

Two of these, hsp 83 and 68 (with molecular weights of 83,000 and 68,000) are coded for by unique sequences in the polytene chromosome regions 63BC and 95D respectively (Holmgreen et al. 1979). The most abundant heat shock protein, hsp 70 (70,000 daltons) has two heat shock puff sites: 87A and 87C. There are two copies of the gene at 87A and three copies at 87C (Ish-Horowicz et al. 1979).

There are four more smaller hsps having molecular weights of 27,000, 26,000, 23,000 and 22,000. Their cytological localization is 67B on the polytene map. Although clustered within an interval of 12 kb, the four genes are not a single transcription unit (Corces et al. 1980).

In this study we chose this gene group for the following reasons. The induction of gene function has been shown in many eukaryotes- slime molds (Frances and Lin 1980), soybeans (Key et al. 1981), chickens (Kelley and Schlesinger 1978), and mice (Hammond et al. 1981) to name a few. Apparently a similar set of genes are always induced. Most organisms have a major heat shock protein with a molecular weight of approximately 70,000 and a second group of smaller heat proteins in the range of 15,000 to 30,000 daltons. The number and molecular weight vary widely (Table 1).

Secondly, these genes are the subject of intensive study with respect to their regulation. Evidence for control at the transcriptional level has been shown (Mirault et al. 1979). Post-translational control by modificiations of the ribosomal proteins has been suggested (Scharf and Nover 1982). Linkage of the heat shock response to mitochondrial functions has been shown (Sin 1975). A model of autoregulation of the 70K protein has been proposed (Lindquist et al. 1982). Therefore any

Table 1. Heat Shock Proteins of Some Organisms Currently Being Studied.

Organism	Tissue	Molecular Weights (kilodaltons)	Reference
Achyla ambisexualis (mold)		78,70,46,44	Gwynne and Brandhorst 1982
Drosophila americana		83,72,70,68,38,28,27,26.5,25.5 23,20	Sinibaldi and Storti 1982
texana		83,72,70,68,38,28,27,26.5,25.5 23,20	" " "
novamericana		83,72,70,68,38,27,25.5,23,20	" " "
lummei		83,72,70,68,38,28,27,26.5,25.5 23,21	" " "
virilis		83,72,70,68,38,28,27,26.5,25.5 23,21	" " "
ezoana		83,72,70,68,38,28,27,25.5,23 21	" " "
littoralis		83,72,70,68,38,28,27,25.5,23 21	" " "
kanekoi		83,72,70,68,38,38,27.5,27,26 23,22	" " "
flavomontana		83,72,70,68,38,28,27.5,27,26 23.5,22	" " "

Table 1. Continued.

Organism	Tissue	Molecular Weights (kilodaltons)	Reference
Drosophila montana		83,72,70,68,38,28,27,26,23,22	Sinibaldi and Storti 1982
lacicola		83,72,70,68,38,28,27.5,26.5 26,23,22	" " "
melanogaster		82,70,68,36,27,26,23,22	Ashburner and Bonner 1979
hydei		70,67,38,26,25,20	Sondermuller and Lubsden 1978
Gallus gallus (chicken)	fibroblasts	95,76,22	Kelley and Schlesinger 1978
Glycine max (soybean)		94,80,75,72,19,16	Altshuler and Mascarenhas 1982
		103,99,85.5,82,80,76,63,37,25 14,11	Barnett et al. 1980
Homo sapiens (man)	HeLa cells	100,74,72,37	Slater et al. 1981
Mus musculus (mouse)	heart tissue	71	Hammond et al. 1982
Nicotiana tabacum (tobacco)		110,100,90,85.5,82,76,63,53,37	Barnett et al. 1980

Table 1. Continued.

Organism	Tissue	Molecular Weights (kilodaltons)	Reference
Polysphondyllium pallidium (slime mold)		105,87,74,33	Francis and Lin 1980
Saccharomyces cerevisiae (brewer's yeast)		90,78,36,30	Finkelstein et al. 1982
Tetrahymena pyriformis		91,75,45,30,29,24,23	Guttman et al. 1980
	thermophilia	97,95,87,81,80,38,37,35,34,33, 21	Hauser and Levy-Wilson 1980

evolutionary study of these genes can draw on a large literature of molecular information about the system.

Thirdly as the number of genes induced by heat shock is large, it is possible to do parallel studies on a group of coordinately controlled genes. The information that there is high homology within the small heat shock gene cluster in Drosophila makes it possible to study it as a duplicated gene family (Corces et al 1980, Ingola and Craig 1982).

Lastly, another advantage of studying this gene system is that parallel studies with both protein variation and DNA variation on the same genes are possible. This facet opens up the means to study regulatory variants of the same protein variants in a rigorous way.

This study will address the question of variation at three levels of genome divergence. These are within species, between closely related species, and between species in the genus Drosophila. The genus Drosophila has been a major source of experimental material for evolutionary studies. The organism is amenable to growth in the laboratory. It has one of the most closely studied genetic system available. The seminal work on heat shock has been done on Drosophila melanogaster, and it is the organism about which the heat shock response is best understood. It is for these reasons that the genus Drosophila was chosen as the experimental material.

The genus Drosophila is broken up into six subgenera. These are in order of their presumed divergence: Scaptodrosophila, Sophophora, Hirtodrosophila, Dorsilopha, Drosophila, and Phloridosa (Figure 1)

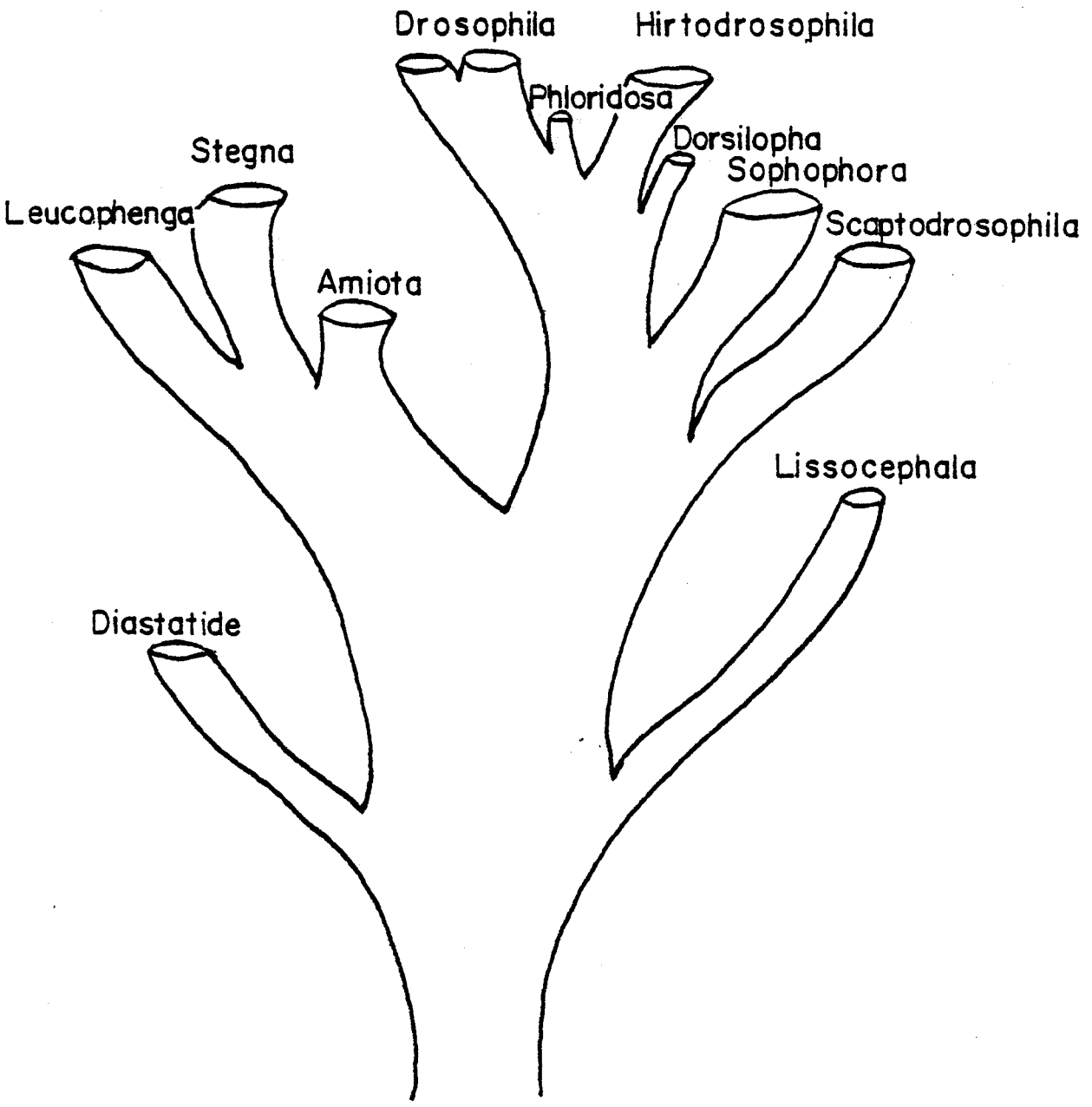

Drosophila
Hirtodrosophila
Phloridosa
Dorsilopha
Sophophora
Stegna
Leucophenga
Scaptodrosophila
Amiota
Lissocephala
Diastatide

FIGURE 1.

(Throckmorton 1975). While attempting to sample most of the major radiations of this genus care was taken to use different levels of divergence whenever possible. A sample of Drosophila melanogaster strains was used as a means to estimate within species divergence. For between species variation extensive use of species triads was made. A triad is a grouping of three species, two of which are sibling species, and the third while still within the same species group is not a sib of either of the other two species. A sibling species pair is defined as a pair of species which are morphologically indistinguishable but reproductively isolated in their natural environment. Also used were species belonging to the various groups of Drosophila which are quite divergent.

MATERIALS AND METHODS

MATERIALS

1. SOURCE OF DROSOPHILA CELL LINE

The Drosophila cell line used in this study was a subline, S-2 (Schneider-2) which was obtained from Dr. Larry Moran, University of Toronto. This cell line was derived from late embryonic stages of Drosophila melanogaster Oregon-R strain (Schneider 1972).

2. SOURCES OF DROSOPHILA SPECIES AND STRAINS

The Drosophila species and strains used have been maintained in our laboratory for sometime. The information concerning their original source and collection site is described in Table 2 .

3. SOURCE OF RECOMBINANT PLASMID

The DNA probe used in this study was dm67B. This is a cloned heat shock mRNA which is homologous to the coding sequence of one of the Drosophila small heat shock loci. It is contained in pBR322 interrupted at the PST 1 recognition site in the ampicillin gene. This plasmid was used to transform E. coli K-12 SF8 strain (Lis etal 1981). This recombinant strain was given to us by Dr. J. Lis, Cornell University.

Table 2 ; Information Concerning Sources of Drosophila Species and Strains

Species or Strain	Source	Collection Site
affinis	National Drosophila Species Resource Centre Austin, Texas (Stock Number 2069.2)	Crystal Lake Hastings, Nebraska U.S.A.
americana	NDSRC (1893.10)	Lake Travis, Texas U.S.A.
ananassae	NDSRC (2507.18)	Keller's Bay, Vermont U.S.A.
arizonensis	NDSRC (E2.2)	Angra dos Reis Brazil
emarginata	NDSRC (H158.2)	Turrialba Costa Rica
eohydei	NDSRC (H186.58)	Santa Mario Mountains Columbia
hydei	NDSRC (H338.7)	Sao Paulo Brazil
lacicola	NDSRC (1756.2B)	Ely, Minnesota U.S.A.
lebanonensis	NDSRC (1733.1)	Beirut Lebanon

Table 2 ; Continued

Species or Strain	Source	Collection Site
malerkotliana	NDSRC (3253.3)	Mysore India
mauritiana	Dr. J.R. David Laboratoire de Biologie de CNKS Gif-Sur-Yvette France (163.1)	The Fourice, Fin Juillet France
melanogaster		
Canton-S		Canton, Ohio U.S.A.
Dalewood (7 strains)	Dr. Richard Morton McMaster University Hamilton, Ontario Canada	Hamilton, Ontario Canada
Brownsville (9 strains)	Dr. D. Hickey Université de Ottawa Ottawa, Ontario Canada	Brownsville, Texas U.S.A.
miranda	Dr. S. Prakash University of Rochester Rochester, N.Y. U.S.A. (Mather 3)	Mather, California U.S.A.

Table 2 ; Continued

Species or Strain	Source	Collection Site
mojavensis	NDSRC (not known)	not known
montana	NDSRC (not known)	not known
mulleri	NDSRC (E45.1)	Panuco, Vera Cruz Mexico
nebulosa	NDSRC (2373.9)	Tingo Mario Peru
neohydei	NDSRC (H207.26)	Carpentaro Brazil
pattersoni	NDSRC (3359.1)	Beirut Lebanon
paulistorum	NDSRC (1975.21)	Belem Brazil
persimilis	Dr. S. Prakash (Mather 220)	Mather, California U.S.A.
prosaltans	NDSRC (H163.13)	Turrialbo Costa Rica
pseudoobscura pseudoobscura	Dr. Lewontin Harvard University Cambridge, Massachusettes U.S.A. (SC-17)	Strawberry Canyon, California, U.S.A.

Table 2 ; Continued

Species or Strain	Source	Collection Site
pseudoobscura bogotana	Dr. Lewontin (BO-16)	Bogota Columbia
robusta	NDSRC (2069.3)	Crystal Lake, Hastings, Nebraska, U.S.A.
saltans	NDSRC (H180.40)	Sao Paulo Brazil
simulans	NDSRC (2394.3)	Lima Peru
takahashii	NDSRC (3146.16)	Yan-Shui Taiwan
tesserie	Dr. J.R. David (128.6)	Mount Selinda Rhodesia
texana	NDSRC (1880.6A)	Swift Creek, South Richmond, Virginia, U.S.A.
tropocalis	Dr. F. Ayala	not known
virilis	NDSRC (1801.1)	Texmelucan Mexico
victoria	Dr. M. Napp Université Fed. R.G. SUL. Porto Alegre Brazil	not known

Table 2 ; Continued

Species or Strain	Source	Collection Site
willistoni	NDSRC (1156.1)	Royal Palm Park, Florida U.S.A.
yakuba	Dr. J.R. David (115)	Kounoleu Carn Croun

4. RADIOCHEMICALS

α^{32}P-dCTP and α^{32}P-dGTP, with a specific activity of greater than 600 Ci/mmol, were obtained from New England Nuclear.

5. BIOCHEMICALS

Ribonuclease, deoxyribonuclease 1, proteinase-K, DNA polymerase (Kornberg's polymerase), lysozyme, and restriction enzymes (ECO R1, PST 1, BAM H1) were supplied by Boehringer Mannheim. An alternate source of the restriction enzymes was Bethesda Research Laboratories. Streptomycin, penicillin-G, tetracycline, salmon testes DNA, bovine serum albumin, agarose, nucleotides, and all vitamins used in the cell culture media were obtained from Sigma. Fetal calf sera was supplied by Flow Laboratories. Lactalbumin hydrolysate was purchased from GIBCO. Bacto-Tryptone, Bacto-Yeast Extract and Agar were supplied by DIFCO. Yeast, malt syrup, and corn syrup were purchased at Health Services Centre. Sugar was obtained from Hickson-Langs Supply Co. Ltd..

6. REAGENTS

All reagents were reagent grade and purchased from Sigma, Fisher or BDH. The exceptions were: Urea-Schwartz Mann, and Tris*BIORAD. Nitrocellulose filter paper was purchased from Schleicher and Schuell.

7. SOLUTIONS

7.1 Drosophila Cell Culture Medium

D22 medium was prepared from a recipe received from Dr. Larry Moran. It was supplemented by ten percent fetal calf sera. The complete recipe is given in Appendix A.

7.2 Drosophila Culture Media

Stock cultures were maintained on two types of media. These

---

* tris-hydroxymethyl-aminomethane

were Carpenter's Medium and Banana Medium. The recipes for these media are stated in Appendix A.

7.3 Media For Growth of Bacteria

7.3.1 Luria Broth

10 g	Bacto-Tryptone
5 g	Bacto-Yeast Extract
10 g	NaCl

The dry ingredients are dissolved in distilled water and pH adjusted to 7.8 with NaOH. The final volume was one liter. If antibiotics were required they were added when media is below 65°C in temperature. Commonly used in this experiment was tetracycline at a concentration of 20 µg/ml. If plates were required 11 g of agar was added before being autoclaved.

7.4 Solutions For DNA Extractions

7.4.1 Drosophila Homogenization Buffer

140 mM	Tris-HCl pH 8.5
90 mM	NaCl
175 mM	Sucrose
275 mM	Na_2EDTA*
1 in 100 volume	Diethylpyrocarbonate

7.4.2 Proteinase-K Stock Solution

50 mM	Tris-HCl pH 8.0
10 mM	Na_2EDTA
10 mM	NaCl
1 mg/ml (w/v)	Proteinase-K

---

* ethylenediaminetetraacetic acid

7.4.3 RNase Stock Solution

50 mM	Tris-HCl pH 8.0
10 mM	Na EDTA
10 mM	Na_2Cl
1 mg/ml (w/v)	Ribonuclease

7.4.4 Phenol-Chloroform Mixture

50% (v/v)	Phenol
50% (v/v)	Chloroform

7.5 Restriction Enzyme Digestion Solutions

7.5.1 10X ECO R1 Buffer

1000 mM	Tris-HCl pH 8.0
50 mM	$MgCl_2$
500 mM	NaCl
20 mM	2-Mercaptoethanol

7.5.2 10X BAM H1 Buffer

500 mM	Tris-HCl pH 8.0
100 mM	$MgCl_2$
500 mM	NaCl
10 mM	Dithiothreitol

7.5.3 10X PST 1 Buffer

200 mM	Tris-HCl pH 8.0
100 mM	$MgCl_2$
500 mM	NaCl
10 mM	Dithiothreitol

7.6 DNA Agarose Gel Solutions

7.6.1 10X DNA Gel Buffer

400 mM	Tris
50 mM	Sodium Acetate
10 mM	Na_2EDTA

This solution is adjusted to pH 7.8 with glacial acetic acid and made up to final volume.

7.6.2 50% Ficoll Solution

50% (w/v)	Ficoll
0.02% (w/v)	Brom Phenol Blue

7.6.3 Ethidium Bromide Solution

0.5 µg/ml (w/v)	Ethidium Bromide

7.7 Southern Blot Solutions

7.7.1 0.25M HCl

23.25 ml	HCl (concentrated)

The concentrated HCl is made up to one liter with distilled water.

7.7.2 Denaturation Solution

0.5 M	NaOH
1.5 M	NaCl

7.7.3 Neutralization Solution

0.5 M	Tris-HCl pH 7.4
3.0 M	NaCl

7.7.4 20X SSC

3.0 M	NaCl
0.3 M	Sodium Citrate

7.8 Nick Translation Solutions

7.8.1 2X Nick Translation Buffer

100 mM	Tris-HCl pH 7.9
10 mM	$MgCl_2$
20 mM	2-Mercaptoethanol
200 µg/ml (w/v)	BSA*

7.8.2 DNase 1 Solution

0.01 M	HCl
5 mg/ml (w/v)	DNase 1

*Bovine Serum Albumin

7.8.3 DNase Activation Buffer

10 mM	Tris-HCl pH 7.5
5 mM	$MgCl_2$
1 mg/ml (w/v)	BSA

7.8.4 Stop Mix

50% (v/v)	0.1 M Na_2EDTA
50% (v/v)	Glycerol

7.8.5 dXTPs Solution

330 µM	dATP*
330 µM	TTP†
330 µM	dCTP∿

or

330 µM	dATP
330 µM	TTP
330 µM	dGTP↓

7.9 Hybridization Solutions

7.9.1 50X Denhardt's Solution

1.0% (w/v)	BSA
1.0% (w/v)	Ficoll
1.0% (w/v)	Polyvinylpyrrolidone

7.9.2 Phosphate Buffer

0.2 M	NaH_2PO_4
0.2 M	Na_2HPO_4

These two solutions are mixed until a pH of 6.5 is reached.

7.9.3 Salmon Testes DNA Solution

2 mg/ml (w/v)	Salmon Testes DNA

---

* 2'-deoxyadenosine 5'-triphosphate
† thymidine 5'-triphosphate
∿ 2'-deoxycytidine 5'-triphosphate
↓ 2'-deoxyguanosine 5'-triphosphate

7.9.4 Dextran Sulphate Solution

50% (w/v)	Dextran Sulphate

7.9.5 Acetate Buffer

0.2 M	Sodium Acetate

Adjust to pH 6.5 with glacial acetic acid and make up to final volume with distilled water.

7.9.6 Hybridization Solution

9.0 ml	Formamide
7.5 ml	20X SSC
0.6 ml	50X Denhardt's
3.0 ml	Acetate Buffer
1.0 ml	Salmon Testes DNA Solution
6.0 ml	Dextran Sulphate Solution
2.9 ml	Water (containing probe DNA)

7.9.7 Prehybridization Solution

9.0 ml	Formamide
7.5 ml	20X SSC
3.0 ml	50X Denhardt's
6.0 ml	Phosphate Buffer
1.5 ml	Salmon Testes DNA Solution
3.0 ml	Dextran Sulphate Solution

7.9.8 Filter Wash Solutions

Stock solutions of different concentrations of SSC were used. These concentrations are:

5X	SSC
0.1% (w/v)	SDS*
2X	SSC
0.1% (w/v)	SDS

* sodium dodecyl sulphate

0.2X	SSC
0.1% (w/v)	SDS

7.10 Plasmid Extraction Solutions

7.10.1 TE 7.0

10 mM	Tris-HCl pH 7.0
1 mM	Na_2EDTA

7.10.2 Triton Salt

65 mM	Na_2EDTA
50 mM	Tris
0.1% (w/v)	Triton-X100

7.11 Scintillation Fluids

7.11.1 Toluene Based Fluors

16.0 g	PPO*
2.0 g	POPOP†
3.8 l	Toluene

---

* 2,5-diphenoloxazole

† 1,4-bis[2-(5-phenyloxazolyl)]-benzene

METHODS

1. GROWTH OF DROSOPHILA CELL CULTURE

S-2 cells were maintained in suspension culture at a concentration of approximately 1 X 10^7 cells/ml. The cells were grown on D22 medium which was supplemented with ten percent fetal calf serum. The cultures were grown at 25°C.

2. GROWTH OF DROSOPHILA STRAINS

Stock cultures were grown at between 19 and 22°C. The species of the Melanogaster and Obscura group were grown on Carpenter's Medium. All other species were grown on Banana Medium. Density was maintained at approximately 200 flies per bottle.

Adult flies were harvested and frozen at -70°C. When adequate numbers had been collected the sample was used.

3. GROWTH OF E. COLI

The procedures for the maintainance and large scale growth of E. coli carrying recombinant plasmids is described in Maniatis et al. (1982). Those protocols were followed closely.

4. EXTRACTION OF DNA

4.1 Extractions from S-2 Cells

The extraction method used was a modification of a procedure by Blin and Stafford (1976). Confluent cells were pelleted (5000Xg for 5 minutes in a JA10 rotor at 4°C) and resuspended in drosophila homogenization buffer (K. Livak personal communication). Cells were lysed by

the addition of one-tenth volume 5% (w/v) SDS and gentle agitation. This was incubated at 65°C for 30 minutes. The following extraction series was performed: once Phenol-Chloroform, twice Chloroform, and twice Ether. The homogenate was made up to 50 µg/ml Proteinase-K and incubated at 37°C for 16 to 18 hours. The previously mentioned organic extractions were repeated and then the homogenate was made up to 100 µg/ml RNase. This was incubated at 37°C for 3 hours. The extraction series was repeated for a third time. The treated homogenate was made up to 0.2 M potassium acetate and two volumes of absolute ethanol was added. The DNA was allowed to precitate overnight. The precitate was pelleted (12,000Xg for 30 minutes in a SS34 rotor at 4°C) and washed twice with 70% (v/v) ethanol. A final pelleting was performed and the pellet was dried. The DNA was dissolved and optical density measurements were taken. A 260 nm to 280 nm ratio of 2 was the required criterion of purity.

4.2 Extraction from Flies

A method similar to the one used on cells was used on the flies. However it was necessary to modify the initial steps.

One hundred to 400 flies were ground in liquid nitrogen to a fine powder using a mortor and pestle. The dry powder was transferred to the drosophila homogenation buffer and vortexed for a few seconds. One one-tenth volume of 5% SDS was added and the mixture was gently agitated. All subsequent steps were the same as the technique for cells.

5. RESTRICTION DIGESTIONS

All restriction digestions were performed in a manner suggested by BRL. Incubations were for 16 to 18 hours at 37°C.

6. GEL ELECTROPHORESIS

DNA gel electrophoresis was performed as suggested by E. Southern (Southern 1980). A tris-sodium acetate-EDTA buffer pH 7.8 was used. Gel concentrations were maintained at 0.8% (w/v) agarose throughout the survey. Electrophoresis was performed at 30V (approximately 30 to 35 mA) for 16 to 18 hours. Gels were stained in electrophoresis buffer containing one one-hundredth volume ethidium bromide solution. Photographs were taken under UV light of all gels.

7. SOUTHERN BLOTS

Blotting of DNA onto nitrocellulose was performed by the method of Southern (Southern 1980). Two techniques were attempted. One included a 15 minute soak in 0.25 M HCl solution before denaturation of the DNA (Wahl et al. 1981). Gels were then soaked in denaturation solution for 15 minutes while being gently agitated. This soak was performed twice. Gels were then soaked in neutralization solution for 15 minutes while being gently agitated. This soak was also performed twice. Blotting by capillary action was allowed to proceed for at least 18 hours. The DNA was then fixed to the nitrocellulose filter by heating the filter to 80°C under vacuum for two hours.

8. NICK TRANSLATIONS

8.1 Activation of DNase

To activate the DNase the following protocol was used (Southern 1980). 50 µl of DNase 1 was added to 450 µl of DNase activation buffer and incubated for 2 hours at 0°C. This solution was further diluted one in one thousand in the DNase activation buffer.

8.2 Nick Translation

Routinely one microgram of plasmid DNA was radioactively labelled at a time. The procedure is a modification of the technique of Southern (Southern 1980) that was suggested by T. Chen (personal communication). The reaction was set as the following: 50 µl of 2X nick translation buffer, 8 µl of dXTPs (minus radiolabelled nucleotide), 10 µl containing 1 µg plasmid DNA, 4 µl activated DNase 1 (containing 1 ng), and sterile water to make the required volume. This solution was incubated at 12°C for 15 minutes. 100 µCi of α^{32}P-dCTP or α^{32}P-dGTP was added and then 2 µl of DNA polymerase (1 Unit) was added. This solution was incubated at 12°C for one hour and 15 minutes. To stop the reaction 60 µl of stop mix was added. Unincorporated labelled nucleotides were separated from the labelled DNA by passage through a Sephadex G-50 column using 0.1% SDS as elutant. The peak fractions were pooled and counted using a toluene based fluor.

9. HYBRIDIZATION PROTOCOL

9.1 Prehybridization

The nitrocellulose filter was placed in a plastic bag and 20 ml of prehybridization solution was added. Bubbles were removed and the bag was sealed. The prehybridization was performed at 42°C for at least 12 hours.

9.2 Hybridizations

The prehybridization solution was removed and replaced with the hybridization solution containing at least 0.5 µg of nick translated DNA with a specific activity of at least 1 X 10^8 cpm/µg. This nick translated DNA had been heat denatured for 5 minutes at 100°C before

addition to the hybridization solution. Hybridization was performed at 42°C for 3 days.

9.3 Filter Washes

A series of washes was usually used. The initial wash was at 5X SSC 0.1% SDS at room temperature. The filter was exposed to Kodak XAR film. If the background was not acceptable further washes were performed. The next level of stringency was 5X SSC, 0.1% SDS at 63°C. Again if the background was till too high when exposed to film a third wash was performed. This last wash was at 2X SSC, 0.1% SDS at 65°C. It was found that the third was not usually necessary.

10. PLASMID EXTRACTION

E. coli were pelleted (5000Xg fro 20 minutes in a JA 10 rotor at 4°C) and resuspended in 10 mls of TE pH 7.0. The method used is a modification of the technique of Godson and Vapnek (1973). The bacteria were made up to 20 mg/ml of lysozyme and incubated at room temperature for 10 minutes. 10 ml of 0.5 M EDTA pH 8.0 was added and mixed well. 40 ml of Triton salt was added and vortexed. The homogenate was centrifuged at 9000Xg for 30 minutes in a SS34. The supernatant was extracted with phenol twice. The homogenate was made up to 0.2 M sodium acetate and two volumes of 70% ethanol were added. The nucleic acids were allowed to precitate overnight at -20°C. The nucleic acids were pelleted (12,000Xg for 30 minutes in SS34 rotor at 4°C). The pellet was dried and resuspended in 3 ml of TE pH 7.0. The solution was made up to 1.0 g per ml CsCl. 0.3 ml of ethidium bromide solution was added. The mixture was centrifuged in a Sw 50.1 rotor at

45,000Xg for 24 hours at 20°C. The plasmid band was removed and made up to 1.0 ml with TE pH 7.0. Ethidium bromide was removed three times by extraction with butanol. The plasmid DNA was precipitated, washed twice with 70% ethanol and pelleted. The pellet was dried and then suspended in sterile water and its concentration determined.

RESULTS

A survey of restriction enzymes fragment polymorphism was performed using enzymes Bam Hl, Eco Rl, and Pst 1. The survey included sixteen strains of Drosophila melanogaster and one each of the following species from the Melanogaster group; D. simulans, D. tesserie, D. yakuba, D. mauritiana, D. takahashii, D. ananassae, and D. malerkotliana.

Single strains from five other species groups outside the Melanogaster group were also analysed using Bam Hl, Eco Rl and in some cases Pst 1.

1. STANDARDIZATION OF GELS

To determine molecular weight of the resulting fragments a means of extrapolating numbers of base pairs was required. For this reason molecular weight markers were run on all gels. These standards were a Hind III digest of lambda phage DNA and Pst 1 digest of D. melanogaster Oregon-R cell culture DNA. The lambda phage digest results in fragments ranging from 23,500 to 2100 base pairs on a 0.8% agarose gel. The Pst 1 digest of the cell culture DNA results in three fragments between 3700 and 1000 base pairs which hybridize to the dm67B probe. Figure 2 is an ethidium bromide stained gels showing DNA digestions. Figure 3 shows the autoradiograms of these standards. There is sufficient homology between the lambda phage DNA and the plasmid that hybridization occurs to the molecular weight standards.

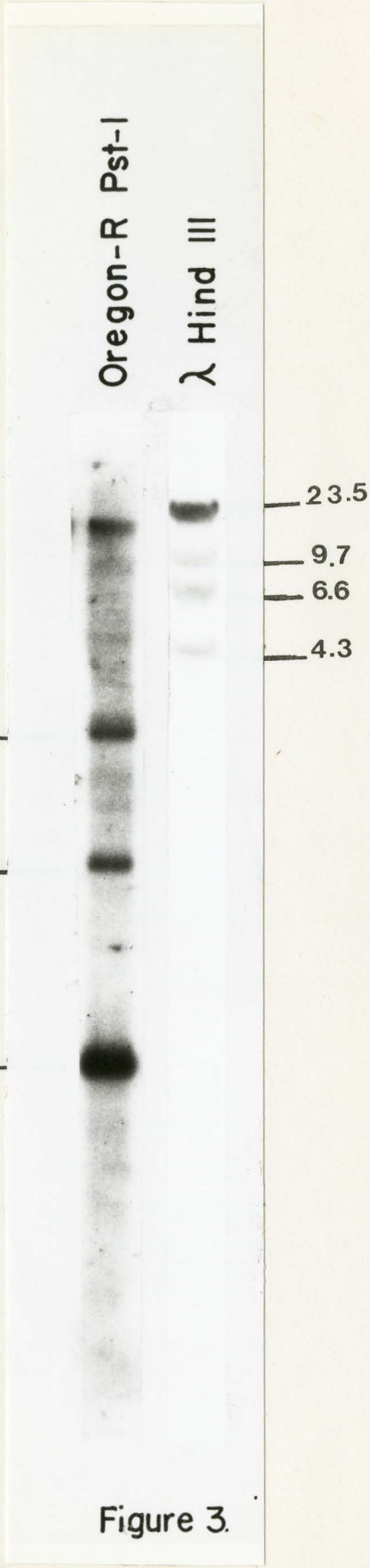


Figure 3.

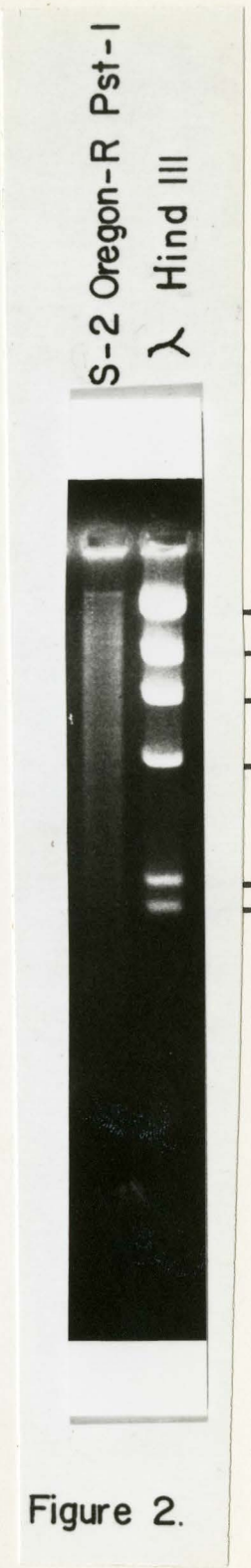


Figure 2.

Both linear and nonlinear regression models were tested to determine the best fit line. As shown in figure 4 the relationship is not linear over the range of mobilities studied. A fourth order polynomial regression of R_f on $\log_{10}$ BP was used to extrapolate the base pair number of unknown DNA fragments.

2. VARIATION WITHIN DROSOPHILA MELANOGASTER

A total of seven isofemale strains from Hamilton, Ontario (Dalewood) and eight isofemale strains from Brownsville, Texas (Brownsville) were tested for restriction enzyme polymorphisms. Canton-S, a standard laboratory strain, was used as a control.

2.1 Bam Hl Digestions

Seven restriction patterns were observed within D. melanogaster (figure 5). There are fragments in the range of 10.2 to 12 Kb in all cases. In five of the fifteen digestions one or two fragments were noted at 4.4 or 5.1 Kb. A diagrammatic representation of the patterns and the frequency of these patterns is shown in figure 6.

The restriction map for Oregon-R cell culture DNA for this region is shown in figure 7 (Craig and McCarthy 1980). The position of coding sequences have been determined and are shown on the figure for reference. Based on this described map it is possible to propose tentative maps for the restriction patterns observed in this study. The coding sequences of the four proteins have been shown to be very similar. The same nucleotide is found at 77% of the positions when one compares any three genes, and the same nucleotide is found in all four genes at 37% of the positions (Ingolia and Craig 1982). Based on these levels of homology then it is expected that hybridization of the cDNA probe to all

Figure 4. Graph of Base Pair Number versus Relative Mobility for the standards used on each gel. Both the best fit line using linear and nonlinear regression models are shown.

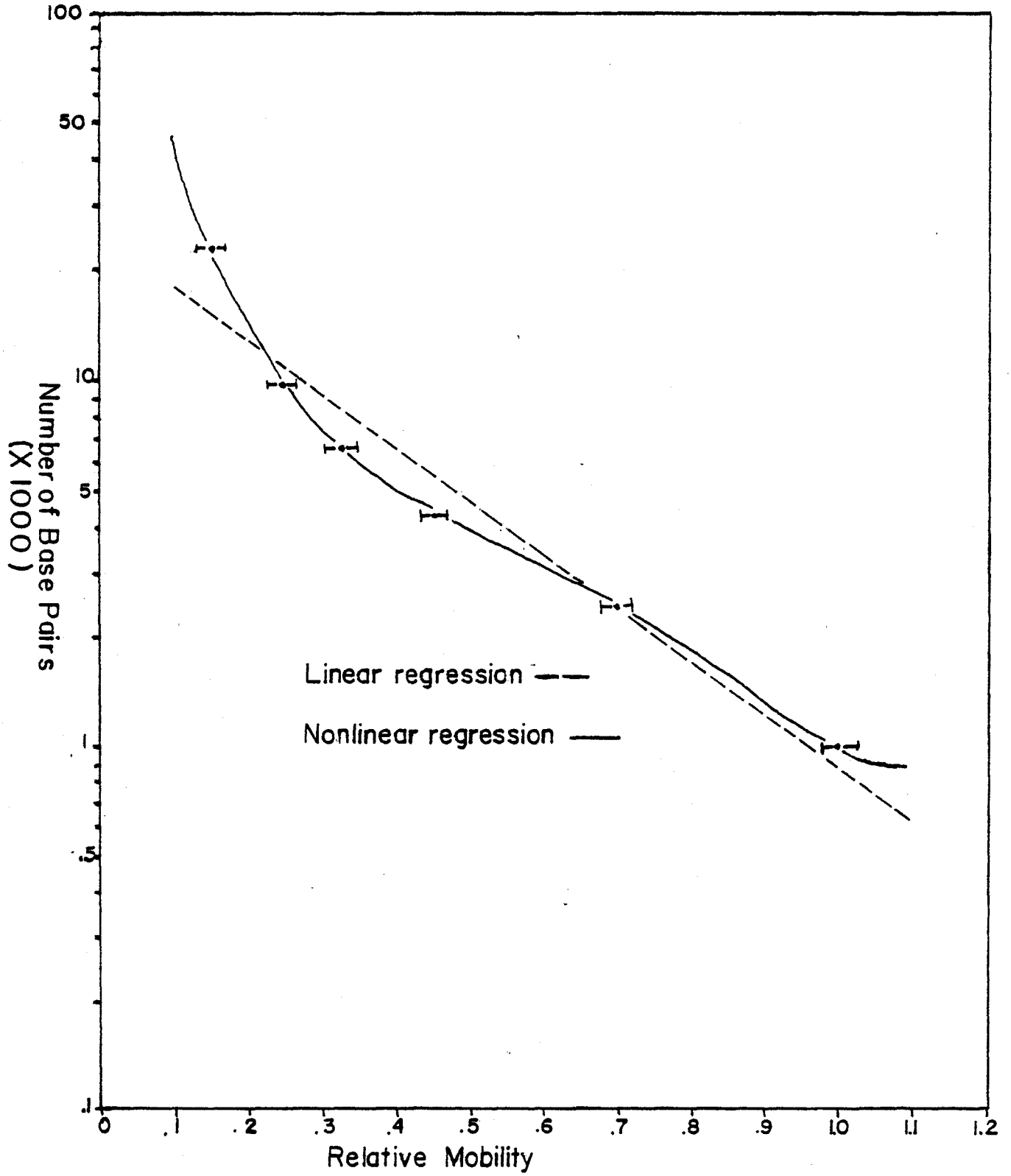

100
50
10
5
1
.5
.1
Number of Base Pairs (X 1000)
Linear regression
Nonlinear regression
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
1.1
1.2
Relative Mobility

Figure 5. Bam Hl digestion patterns of genomic digests of D. melanogaster strains.

10 µg of DNA from fly stocks were digested with Bam Hl, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, and hybridized to ^{32}P dm67B probe. 0.75 µgm of ^{32}P dm67B probe at a specific activity of 2.0×10^8 cpm/µgm was used in hybridization. Hybridization proceeded for 76 hours at $42^{\circ}C$. The filter was washed at 5X SSC 0.1% SDS at room temperature. The exposure shown was to XAR film for 7 days.

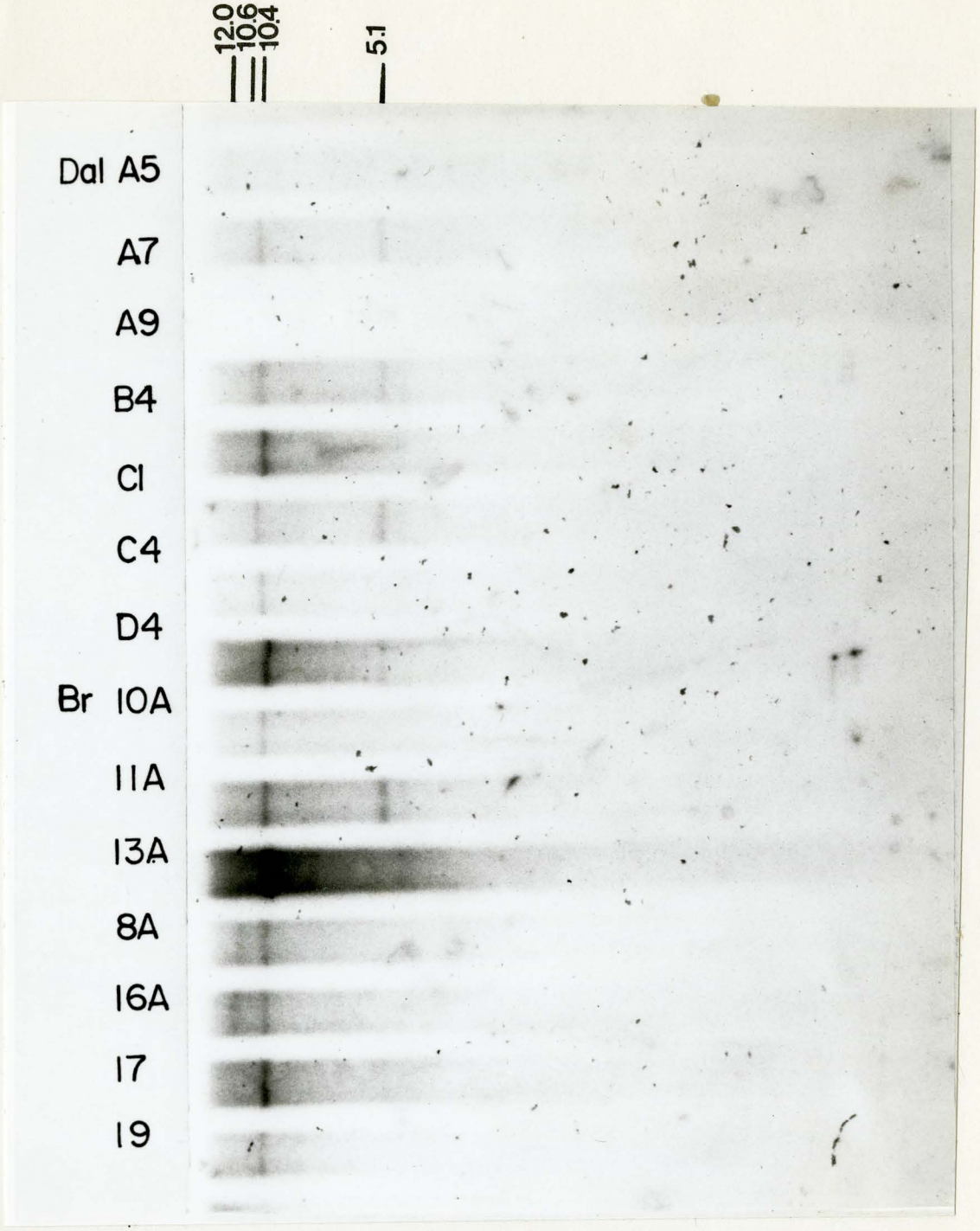
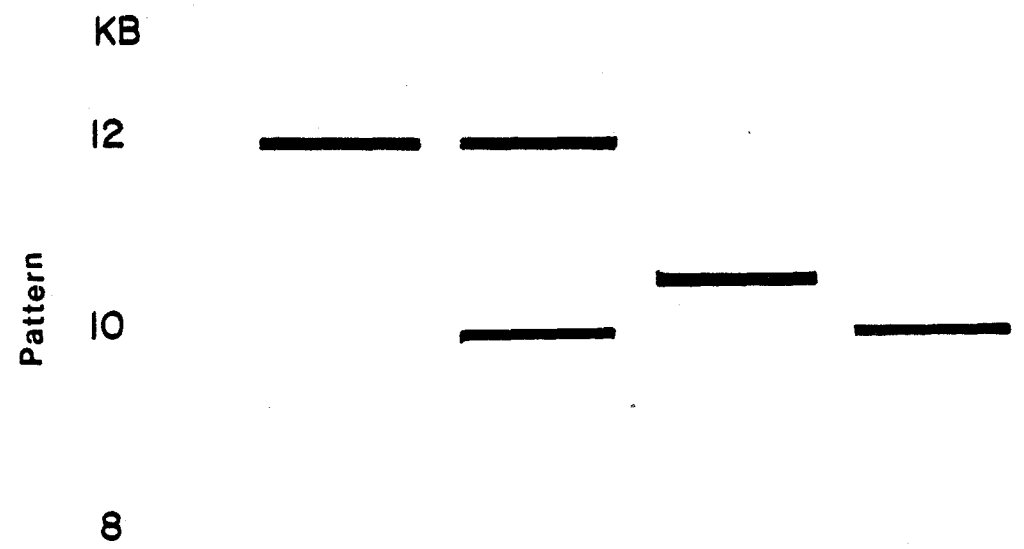

12.0
10.6
10.4
5.1
Dal A5
A7
A9
B4
C1
C4
D4
Br 10A
11A
13A
8A
16A
17
19

Figure 6. Diagrammatic representation of the pattern obtained with Bam Hl digestion of D. melanogaster strains. Also shown are the frequencies in the two populations and in the total sampling.



Frequency

Total	1	2	2	10	/15
Dalewood	1	0	2	4	/7
Brownsville	0	2	0	6	/8

Figure 7. A diagram representing the map positions of the restriction sites using Bam Hl, Eco Rl, and Pst 1 to digest D. melanogaster Oregon-R cell culture DNA. Redrawn from Craig and McCarthy 1980.

BAM HI
ECO RI
PST I
27
23
12
26
22
4.4
4.8
0.9
0.8
2.8
6.5
2.4
1.0
2.4
5.8

of the coding sequences is possible.

In view of this the 10.2, and 10.6 Kb fragments appear to be variants of the originally described 12 Kb fragment containing the coding sequences for hsps 27, 23, and 26. The weakly hybridizing bands at 4.4 and 5.1 Kb are proposed to be homologous to the 4.4 Kb fragment found by Craig and McCarthy (1982). The weakness of the hybridization makes the interpretation of this variation in 4.4. and 5.1 Kb fragments difficult and in the absence of clear data on this region, the two regions will be analysed separately.

Ascribing these fragments patterns to genotypes is possible based on these proposed maps. By designating the 12 Kb fragment as allele B1, 10.6 Kb as allele B2 and the 10.2 Kb fragment as allele B3, genotypic frequencies can be used to calculate fixation index, expected heterozygosity, and genetic distance between the two populations (table 3).

2.2 Eco R1 Digestions

Five restriction patterns were observed when fly DNA was digested with Eco R1. The patterns for within D. melanogaster are shown in figures 8 and 9. Referring to figure 7 Oregon-R cell culture DNA has three major fragments 6.5, 4.8 and 2.8 Kb. Based on the homology to the cDNA clone would expect little hybridization to the 2.8 Kb fragment as little if any of the coding regions for any of the genes are contained on it.

Using the Oregon-R map as a standard, the fragments 6.5 and 4.8 Kb are expected. As shown in figure 9, a 4.8 Kb fragment is found in all fifteen strains. However a second fragment is found only in six

Table 3. Allele frequencies of Bam Hl digestion products of *D. melanogaster* strains. Also shown are F_{ST}, and Genetic Distance.

Allele	Dalewood	Brownsville	Total
B1	0.143	0.125	0.133
B2	0.286	0.000	0.133
B3	0.571	0.875	0.733
Expected heterozygosity	0.572	0.219	0.427

$$\text{Fixation Index } (F_{ST}) = \frac{H_T - H_S}{H_T} \quad \text{where} \quad H_S = \frac{H_{Dal} + H_{Br}}{2}$$

$$= 0.074$$

$$\text{Genetic Distance } (D) = -\log_e I \quad \text{where} \quad I = \frac{\Sigma\, p_x p_y}{[(p_x^2)\,(p_y^2)]^{\frac{1}{2}}}$$

$$= 0.041$$

of these.

Ingolia and Craig (1981) have shown a Eco R1 cleavage site at position minus 100 bases from the hsp 23 5'end, and a second cleavage site at minus 170 bases in the flanking region of the hsp 26 gene (5' end). A possible interpretation of this, taking the homology of the coding sequence into consideration, is that these genes are duplicate loci (Corces et al. 1980). These genes have diverged in sequence as well as in size since the initial duplication event(s). Based on the placement of the genes and their position, the duplication events may have occurred in three steps. The original gene was duplicated to form two copies, these may or may not have been in the inverted repeat form. A second duplication event caused the repeat of these two loci to form four copies. If the genes were in the inverted repeat originally then the inversion of hsp 27 is implied, and if they were in tandem repeat in the original duplication then the inversion of hsp 22 is implied. This makes the axis of symmetry of the two regions to be about 2.25 Kb from the end of each of the hsp 23 and 26 loci.

Based on this argument it would be expected that it is possible that the 4.8 Kb fragment would be retained in both of the gene pairs. Therefore the following maps are ascribed to the patterns found. These are 6.5 and 4.8 Kb as in the cell culture DNA (allele E1), 4.8 and 5.1 (allele E2) as found in three Dalewood lines and one Brownsville line, and a map of 4.8 and 4.8 Kb as found in the other nine lines (allele E3). Ten micrograms of DNA was added to each pocket on the gel shown in figure 8. It appears that the lanes with only a 4.8 Kb fragment on the average have higher density of silver grains. This in turn implies that

Figure 8. Eco R1 digestion patterns of genomic digests of D. melanogaster strains.

10 µg of DNA from fly stocks were digested with Eco R1, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, and hybridized to ^{32}P dm67b probe. Lane 1 is a Pst 1 digest of Oregon-R cell culture DNA. 0.66 µgm of ^{3}P dm67B probe at a specific activity of 1.68×10^{8} cpm/µgm was used in the hybridization. Hybridization proceeded for 76 hours at $42^{o}C$. The filter was washed at 5X SSC 0.1% SDS at room temperature. The exposure shown was to XAR film for 14 days.

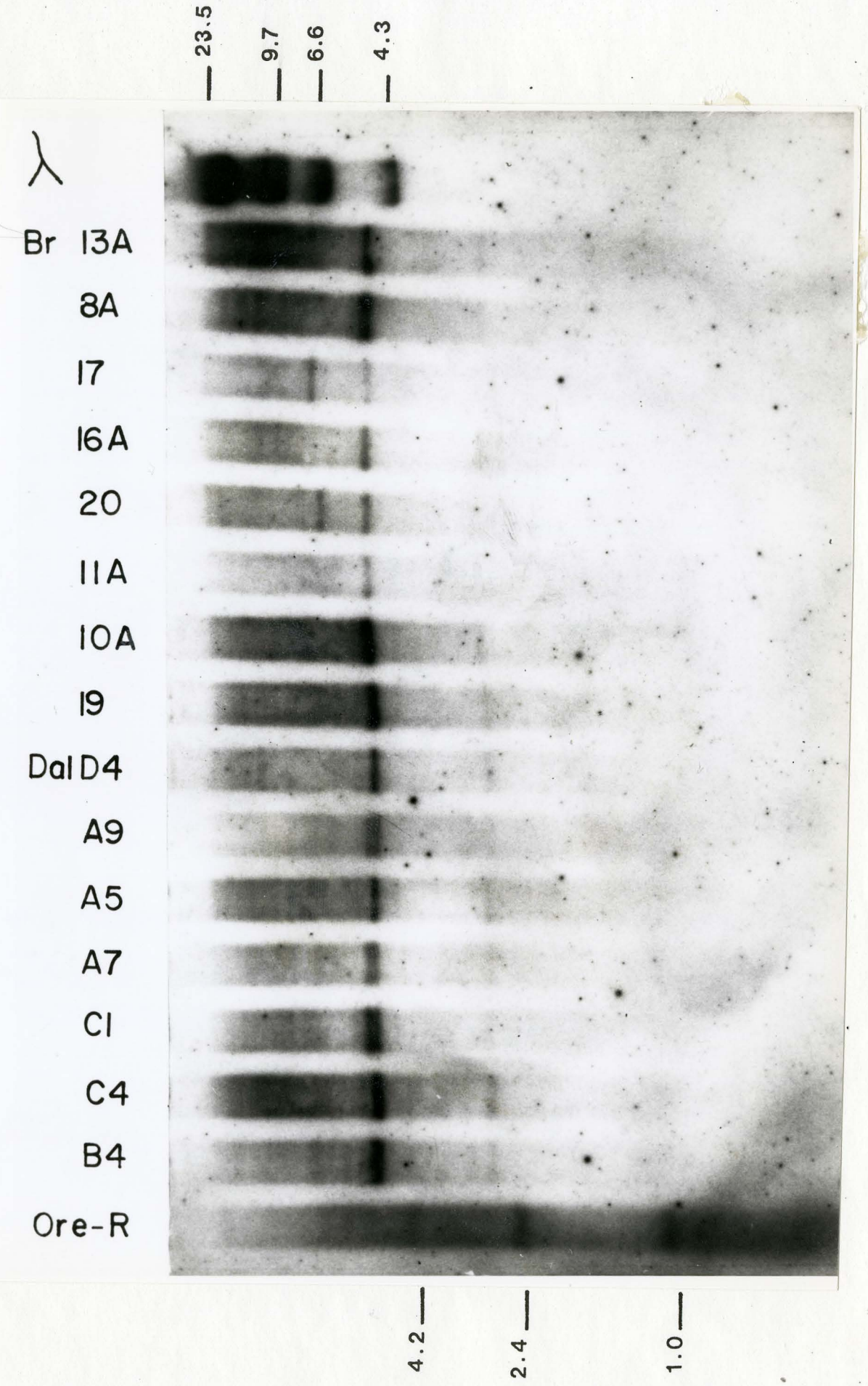
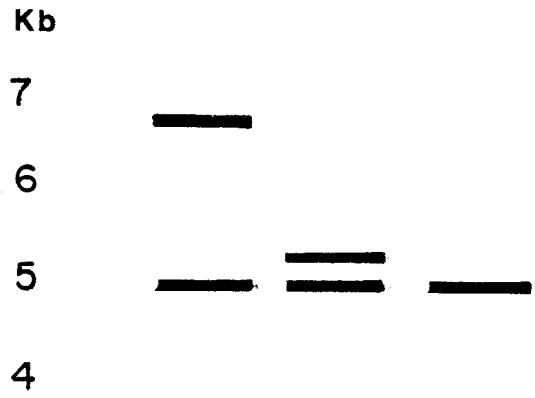

23.5
9.7
6.6
4.3
λ
Br 13A
8A
17
16A
20
11A
10A
19
Dal D4
A9
A5
A7
C1
C4
B4
Ore-R
4.2
2.4
1.0

Figure 9. Diagrammatic representation of the pattern obtained with Eco R1 digestion of D. melanogaster strains. Also shown are the frequencies in the two populations and the total sampling.

Pattern



Frequency

Dalewood	0	3	4	/7
Brownsville	2	1	5	/8
Total	2	4	9	/15

the amount of DNA in that band is greater. This observation though qualitative is consistent with the hypothesis of retention of 4.8 Kb fragments in both gene pairs.

A second point of interest is that the migration of the 6.5 Kb fragment varies slightly between the two strains showing it. In view of the short migration distances any inferences about true differences will not be made until a gel concentration that would maximize the difference has been attempted.

Frequencies of these alleles in the two populations are shown in Table 4, along with F_{ST} values, expected heterozygosity, and genetic distance based on this digestion pattern.

2.3 Pst 1 Digestion

Five restriction patterns were observed when DNA was digested with Pst 1. The patterns for within D. melanogaster are shown in figure 10. Refering to figure 7, Oregon-R cell culture DNA has four known fragments 2.4, 1.0, 2.4. and 5.8 Kb. The size of the fragments on the left end of the gene group is unknown. As a cDNA probe would not be expected to hybridize to the 2.8 Kb fragment in the untranscribed region between hsp 23 and 26 only one of the 2.4 Kb fragments would hybridize.

In contrast to the earlier report only 2.4, 1.0, and 3.7 Kb fragments were found in Oregon-R digests in this study. It is hypothesized that either the genomic fragment used to generate the map was from a line fixed or heterozygous for a sequence, which, while present in the original Oregon-R fly stocks from which this cell culture was initiated, is not present in the subline used in this study. A second hypothesis is that a somatic mutation has occurred in the cells during

the time since their original initiation that is different in the subline in our laboratory than in others.

There were six strains demonstrating a 5.8 Kb fragment under Pst 1 digestion. The 3.5, 2.4, and 1.0 Kb fragments were found in all digestions. However there was variation for the presence of 3.7, 3.9, and 5.8 fragments. In figure 11 is a diagrammatic representation of the patterns observed in the two populations.

There are many combinations of tentative maps possible using this data. I am assuming that 3.7, 3.5, 2.4, and 1.0 Kb fragments represent a homozygote, and 3.9 is an alternate fragment homologous to the 3.7 Kb as these are the simplest patterns. If this is assumed then no information as to whether 5.8 is homologous to the 3.5 or 3.7-3.9 fragments is suggested. If the 5.8 fragment is homologous to only 3.7 or 3.9 then in the case of all three fragments occurring together three alleles at the 'locus' are implied. However to explain all the patterns a 5.8 Kb fragment must be homologous to both the 3.5 and 3.7-3.9 Kb fragments. This implies again symmetry of the two gene pairs to some extent.

This hypothesized sets of alleles are the lowest number needed to explain the patterns found. Gene frequencies can be calculated from these patterns and are shown in table 5. Also shown are the F_{ST} value, the expected heterozygosity, and the genetic distance estimate from these data.

2.4 Phylogenetic Estimates Based on Restriction Data

The inferences to be made on the data on similarity within

Table 4. Allele frequencies of the Eco R1 digestion products of *D. melanogaster* strains. Also shown are expected heterozygosity and F_{ST} and D.

Allele	Dalewood	Brownsville	Total
E1	0.000	0.250	0.133
E2	0.429	0.125	0.267
E3	0.571	0.625	0.600
Expected heterozygosity	0.490	0.531	0.551

$F_{ST} = 0.0726$

$D = 0.176$

Figure 10. Pst 1 digestion patterns of genomic digests of D. melanogaster strains.

10 µg of fly DNA was digested with Pst 1, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, and hybridized to ^{32}P dm 67b probe. Lanes designated by letters refer to patterns shown in figure 11.

0.75 µgm of ^{32}P-dm 67B probe at a specific activity of 2.2×10^8 cpm/µgm was used in hybridization. Hybridization proceeded for 79 hours at 42°C. The filter was washed at 5X SSC, 0.1% SDS at room temperature. The exposure was to XAR film for 14 days.

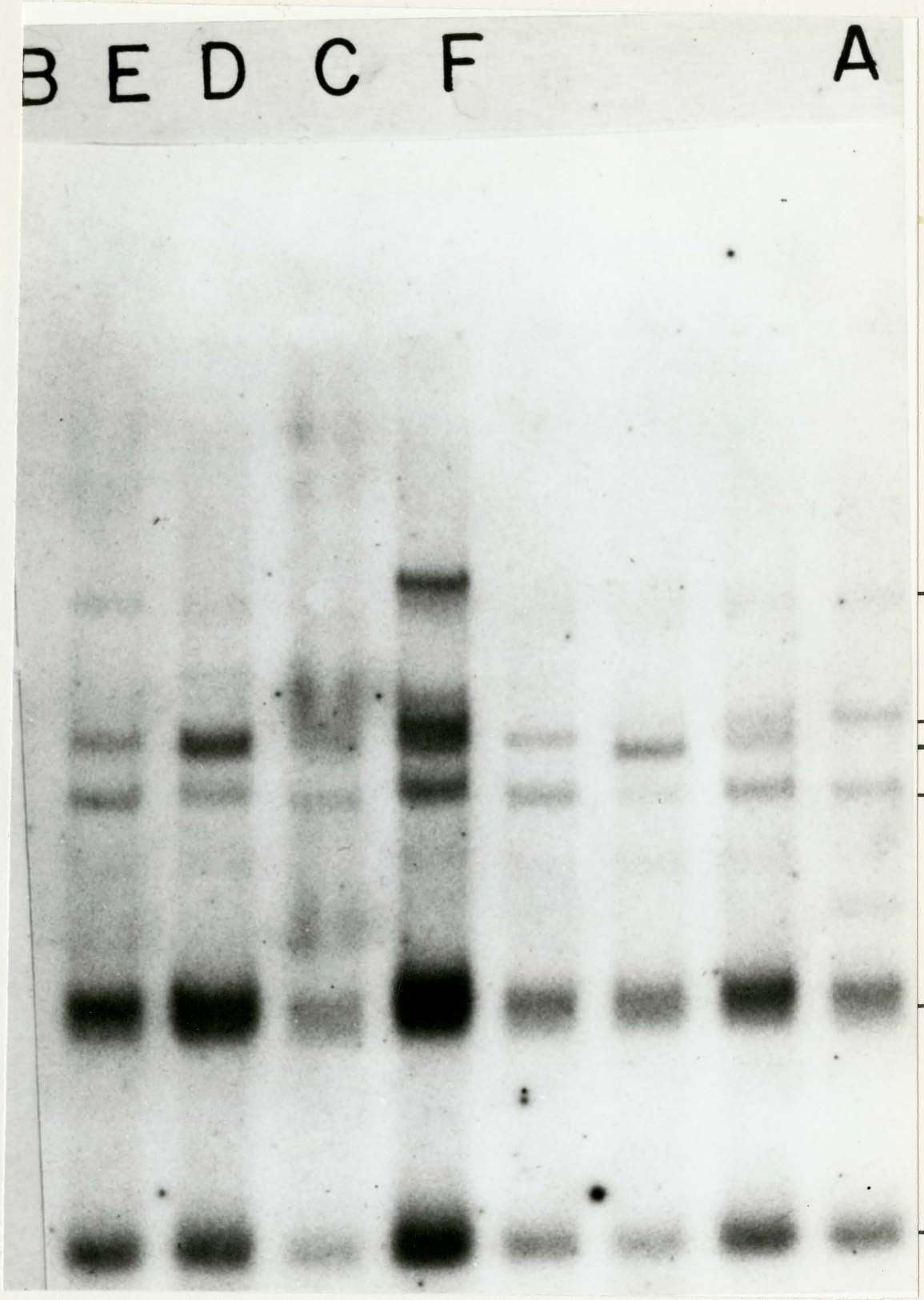
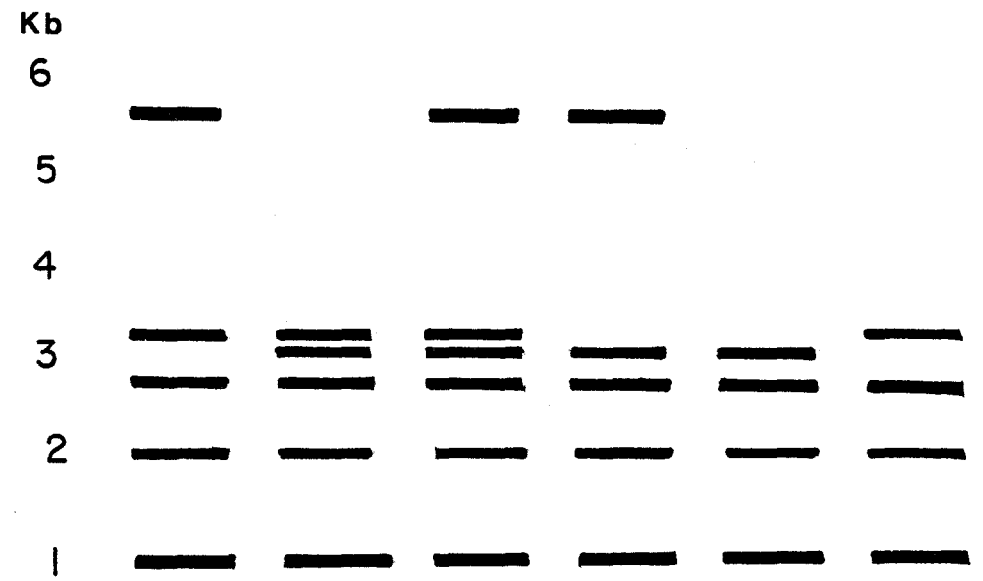

B
E
D
C
F
A
5.8
3.9
3.7
3.5
2.4
1.0

Figure 11. Diagrammatic representation of the pattern obtained with Pst 1 digestion of D. melanogaster strains. Also shown are the frequencies in the two populations studied and the total sampling. In order from left to right the patterns for reference to figure 10 are A, E, C, B, D, and F.



Pattern

Frequency

Dalewood	2	0	0	0	1	4	/7
Brownsville	1	2	1	2	2	0	/8
Total	3	2	1	2	3	4	/15

Table 5. Allele frequencies of Pst 1 digestions products of *D. melanogaster* strains. Also shown are expected heterozygosity, F_{ST}, and D.

Alleles	Dalewood	Brownsville	Total
P1	0.571	0.250	0.400
P2	0.286	0.500	0.400
P3	0.143	0.125	0.133
P4	0.000	0.125	0.067
Expected heterozygosity	0.572	0.656	0.652

$F_{ST} = 0.058$

$D = 0.234$

D. melanogaster can be approached in two ways. The first is to follow the assumption in the previous sections about tentative maps and treat the data in the manner of electrophoretic alleles and determine genetic distance treating the entire sequence as one locus (Nei 1974). The second is to use the estimates of similarity or diversity specific to restriction data. The analysis can be performed essentially free of assumptions of maps. If the existence of a specific fragment can be considered a character then the presence or absence of that fragment is the character state for that character in the sample studied. Upholt (1977), Nei and Li (1979) and Li (1981) have all suggested ways to measure sequence diversity using restriction enzyme fragment polymorphism data in this manner. Nei and Li (1979) have proposed a simple relationship between the sequence divergence and the restriction fragment polymorphisms observed on gels. Basically it is that the number of shared cleavage sites between two lineages can be used to estimate the sequence divergence between them. If a restriction map is known then the estimate of site divergence is S where

$$S = \frac{n_{xy}}{n_0}$$

n_{xy} is the number of shared sites between lineages x and y, and n_0 is the original number of sites at the time of separation of the two lineages. The estimate of n_0 is n_x and n_y or the sum of the sites in both lineages.

The amount of sequence divergence can be estimated in the absence of exact knowledge of shared restriction sites. Using F as the shared fragment similarity where;

$$\hat{F} = \frac{2n_{xy}}{n_x + n_y}$$

where n_{xy}, n_x, and n_y are the number of shared fragments, the number of fragments in lineage x and in lineage y respectively.

Knowing F then one can estimate the nucleotide diversity or base substitution rate in a lineage. Nei and Li (1979) have derived $\hat{F} \tilde{=} P^4/(3-2P)$ where $P = e^{-r\lambda t}$ where r is the number of bases in the recognition sequence. δ is defined as $2\lambda t$ and this relationship was used to determine the corresponding value of δ for F. When intra-population estimates are possible then a value for sequence diversity can be used. π(index of nucleotide diversity) is derived as

$$\pi = \Sigma x_i s_j \pi_{ij}$$

where x_i and x_j are the frequencies of the ith and jth sequences in the population and π_{ij} is the value for the base subsitution number between the two sequences i and j. Shown in appendix B is calculation of δ and π.

Table 6 lists the genetic distances based on the treatment of the data as gene frequencies. The normalized genetic distance for the three enzymes is 0.1503.

Shown in Table 7 are the values for F and the corresponding δ for the within D. melanogaster strains. The observed value attained for $\hat{\pi}$ was 0.034. This value for $\hat{\pi}$ may be inflated as the data may include heterozygous individuals. In this case the divergence will be overestimated as certain individuals will be segregating for different alleles while sharing one in common. If this is true then the F will be less than actual if every individual was homozygous. By taking this into

Table 6. Estimates of Genetic Distance Between the Hamilton and Brownsville populations of *D. melanogaster*.

Enzyme	Genetic Distance
Bam H1	0.041
Eco R1	0.176
Pst 1	0.234
Normalized Genetic Distance (all three enzymes)	0.150

Table 7. Values for F and π for within D. melanogaster strains. Above the diagonal are values for π ($X10^{-1}$) and below are values for F (X1). The diagonal equals zero and one respectively for π and F.

	O-R	C-S	A5	A7	A9	B4	C1	C4	D4	8A	10A	11A	13A	16A	17	19	20
Ore-R	----	.21	.46	.81	.91	.91	.81	.91	.91	.65	.55	.55	.65	.29	.39	.55	1.00
C-S	.83	----	.21	.55	.55	.46	.55	.46	.21	.39	.29	.29	.39	.09	.17	.29	.65
Dal A5	.67	.83	----	.29	.29	.21	.29	.39	.46	.65	.55	.29	.39	.46	.29	.29	.25
A7	.50	.62	.77	----	0	.09	.39	.25	.29	.25	.39	.29	.25	.39	.46	.17	.25
A9	.46	.62	.77	1.0	----	.09	.39	.25	.29	.25	.39	.17	.25	.39	.46	.17	.25
B4	.46	.67	.83	.92	.92	----	.29	.29	.29	.39	.29	.09	.17	.29	.39	.09	.17
C1	.50	.62	.77	.71	.71	.77	----	.08	.55	.46	.39	.21	.25	.65	.74	.39	.25
C4	.46	.67	.71	.80	.80	.77	.93	----	.65	.32	.46	.46	.32	.74	.81	.46	.32
D4	.50	.83	.67	.77	.77	.77	.62	.57	----	.17	.17	.09	.17	.09	.17	.09	.29
Br 8A	.57	.71	.57	.80	.80	.71	.67	.75	.86	----	.08	.25	.14	.17	.32	.25	.39
10A	.62	.77	.62	.71	.71	.77	.71	.67	.86	.93	----	.17	.08	.17	.25	.17	.25
11A	.62	.77	.77	.86	.86	.92	.93	.67	.92	.80	.86	----	.08	.17	.25	0	.25
13A	.57	.71	.71	.80	.80	.86	.80	.75	.86	.88	.93	.93	----	.25	.32	.08	.14
16A	.77	.92	.67	.71	.71	.77	.57	.53	.92	.86	.86	.86	.80	----	.08	.17	.46
17	.71	.86	.72	.67	.67	.71	.53	.50	.86	.75	.80	.80	.75	.93	----	.25	.32
19	.62	.77	.77	.86	.86	.92	.71	.67	.92	.80	.86	1.0	.93	.86	.80	----	.25
20	.43	.57	.80	.80	.80	.86	.80	.75	.77	.71	.80	.80	.88	.67	.75	.80	----

account the estimate of $\hat{\pi}$ is 0.030.

By partitioning the comparisons to within a population the following estimates were obtained. The population Dalewood had a $\hat{\pi}$ of 0.0291 with variance of 0.00237. The Brownsville population had a $\hat{\pi}$ of 0.0206 with a variance of 0.00108.

The last method of analysis is the phenetic and cladistic analysis of the similarity coefficients generated by using Nei and Li's method of comparison. The matrix of similarity was used in the SAHN technique of cluster analysis unweighted pair-group method (UPGM) (Sokal and Sneath 1973) to generate a phenogram of restriction enzyme phenotype. The Minimum Spanning Tree Method (Sokal and Sneath 1973) of tree-building was used. The method generates a tree of restriction enzyme fragment type. The diagrams showing these dendrograms are given in figures 12 and 13.

3. VARIATION WITHIN THE GROUP MELANOGASTER

Eight species were examined from the Melanogaster group. These can be divided based on conventional morphological characters into three groups; subgroups Melanogaster, Takahashii, and Ananassae. From the Melanogaster subgroup the species *D. melanogaster* and its sibling species *D. simulans*, and the nonsibling species *D. yakuba*, *D. tesserie*, and *D. mauritiana*. From the Takahashii subgroup represented by *D. takahashii*. Representatives of the Ananassae subgroup are *D. ananassae*, and *D. malerkotliana*.

3.1 Bam Hl Digestion

All species within the subgroup Melanogaster had an approximately

Figure 12. Phenogram of within D. melanogaster strains. Diagram is a reproduction of phenogram generated by UPGM using NT-SYS phylogenetic analysis computer package.

PHENOGRAM FROM MATRIX SIM 00

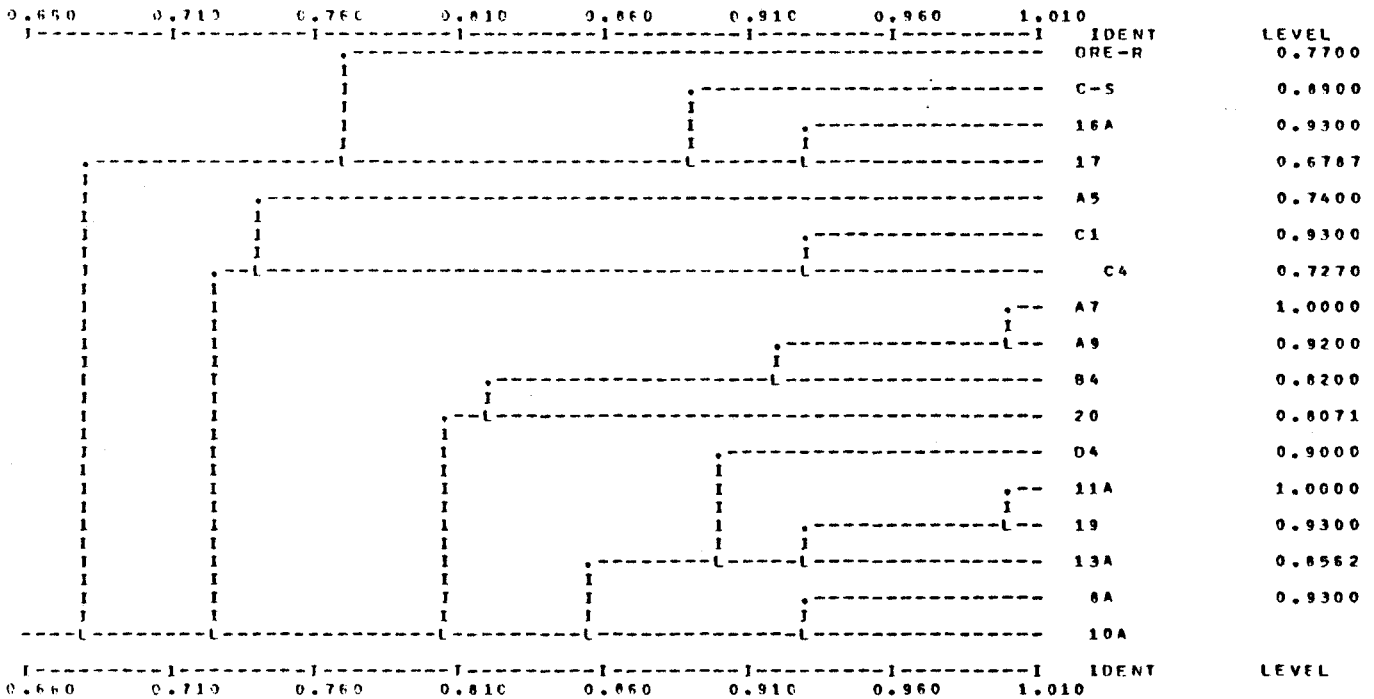
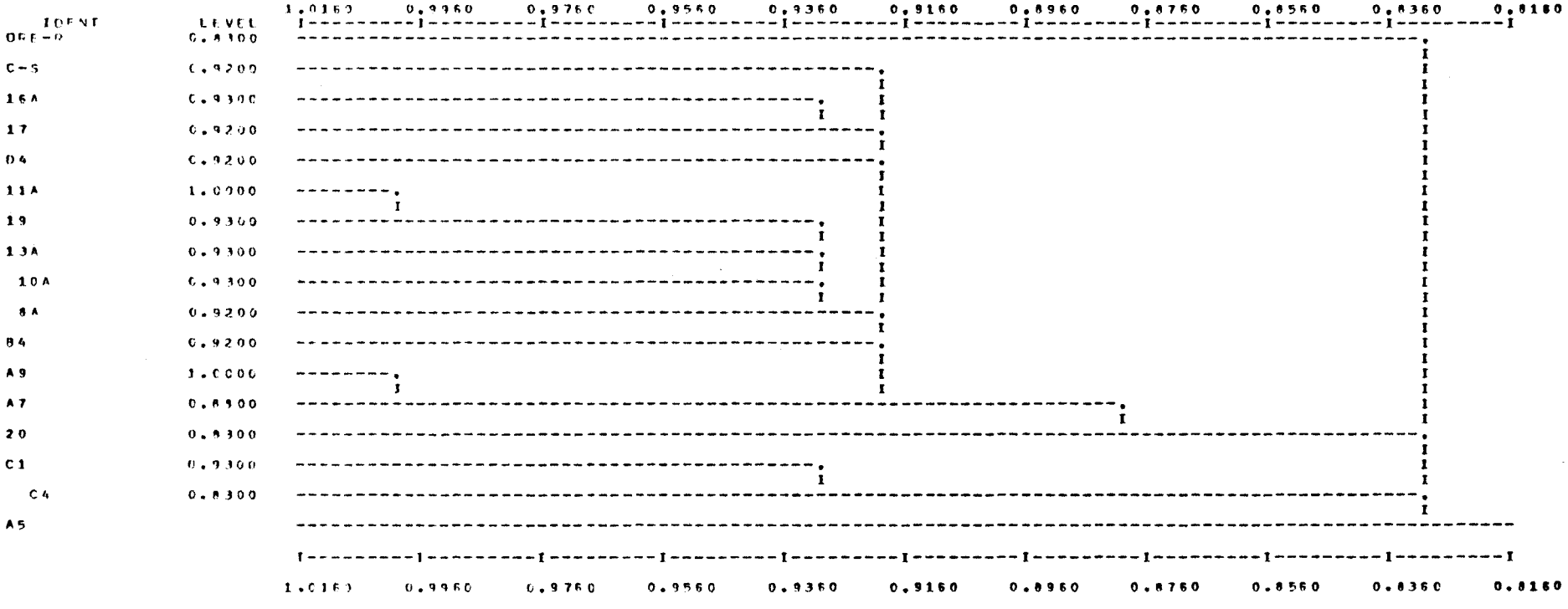


Figure 13. Minimum Spanning Tree of similarity values generated by restriction analysis of within D. melanogaster strains. (MST of NT-SYS generated output).


IDENT
LEVEL
1.0160
0.9960
0.9760
0.9560
0.9360
0.9160
0.8960
0.8760
0.8560
0.8360
0.8160
ORE-O 0.8400
C-S 0.9200
16A 0.9300
17 0.9200
D4 0.9200
11A 1.0000
19 0.9300
13A 0.9300
10A 0.9300
8A 0.9200
B4 0.9200
A9 1.0000
A7 0.8900
20 0.8300
C1 0.9300
C4 0.8300
A5

12 Kb fragment which hybridized to the dm67b probe (figure 14). Slight differences were noted in one gel (figure 15) but were not replicated in other attempts. Regarding the other three species, they did not have this Bam H1 fragment. Both D. takahashii and D. ananassae had a single fragment of 5.5. Kb (figure 15). D. malerkotliana had three fragments of size 5.5, 4.7 and 3.9 Kb.

3.2 Eco R1 Digestion

There was considerable restriction enzyme variation in the results of Eco R1 restriction enzyme digestion within the Melanogaster group. The Canton-S strain was the reference strain of comparisons of other species to D. melanogaster variation. D. melanogaster, D. simulans D. takahashii, D. ananassae, and D. malerkotliana all had strongly hybridizing bands at 4.8 Kb.

Both D. yakuba and D. tesserie show a strongly hybridizing band at 3.2 Kb and lastly D. mauritiana has a strongly hybridizing band at 5.0 Kb. Many of the species had other bands showing less strong hybridization; D. tesserie had three bands in the region of 7.5 to 7.0 Kb, while D. takahashii and D. mauritiana had a band at 1.5 Kb and D. ananassae at 1.9 Kb. D. takahashii also had bands hybridizing at 2.5 and less strongly at 4.5 and 4.0 Kb.

In the cases of D. tesserie and D. takahashii either incomplete digestion or nonspecific binding are possible explanations as these bands were variable between digestions. The patterns are shown in figures 16 (autoradiographs) and figure 17 (diagram of fragment patterns).

Figure 14. Bam Hl digestion patterns of within the Melanogaster Group. 10 µg of DNA from fly stocks were digested with Bam Hl, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 18 hrs, blotted, and hybridized to ^{32}P dm67B probe. Lane labelled SIM2 is a Pst 1 digest of D. simulans DNA, all others are Bam Hl digestions. 0.5 µgm of ^{32}P dm67B probe at specific acitivty of 1.6 X 10^8 cpm/µgm was used in the hybridization. Hybridization proceeded for 83 hours at 42°C. The filter was washed with 5X SSC and 0.1% SDS at room temperature. The exposure shown was to XAR film for 5 days.

Figure 15. Bam Hl digestion patterns of within Melanogaster Group. 10 µg of DNA from fly stocks were digested with Bam Hl, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, and hybridized to ^{32}P dm67B probe. 3.0 µgm of ^{3}P dm67B probe at a specific activity of 1.2 X 10^8 cpm/µgm was used in the hybridization. Hybridization proceeded for 71 hours at 42°C. The filter was washed with 5X SSC and 0.17 SDS at room temperature. The exposure shown was to XAR film for 19 days.

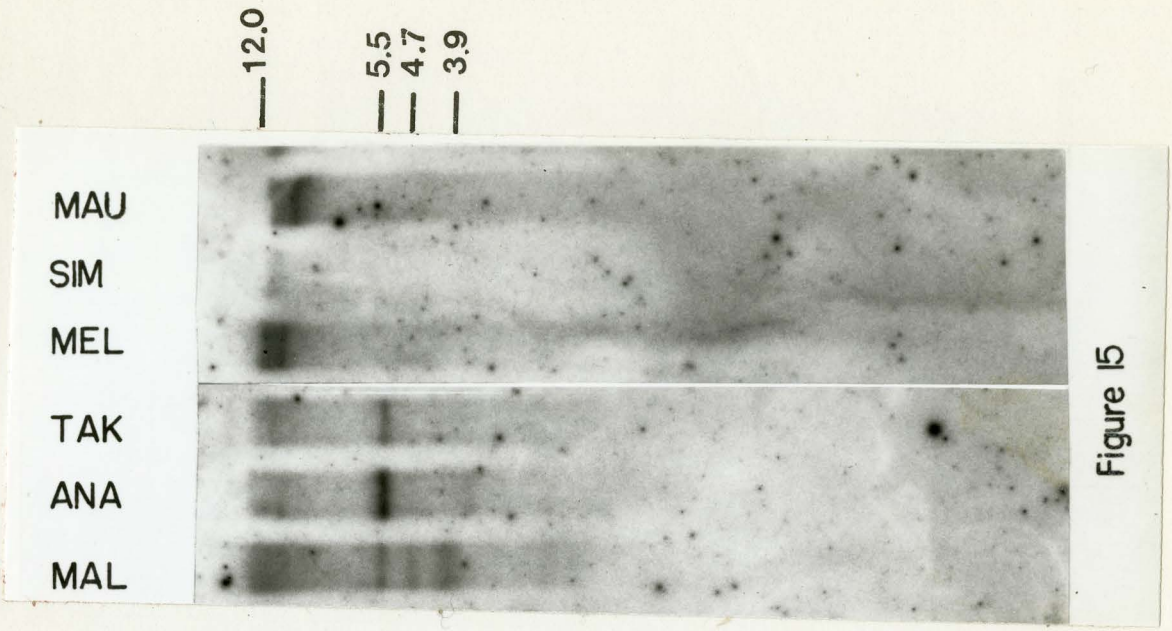


Figure 15

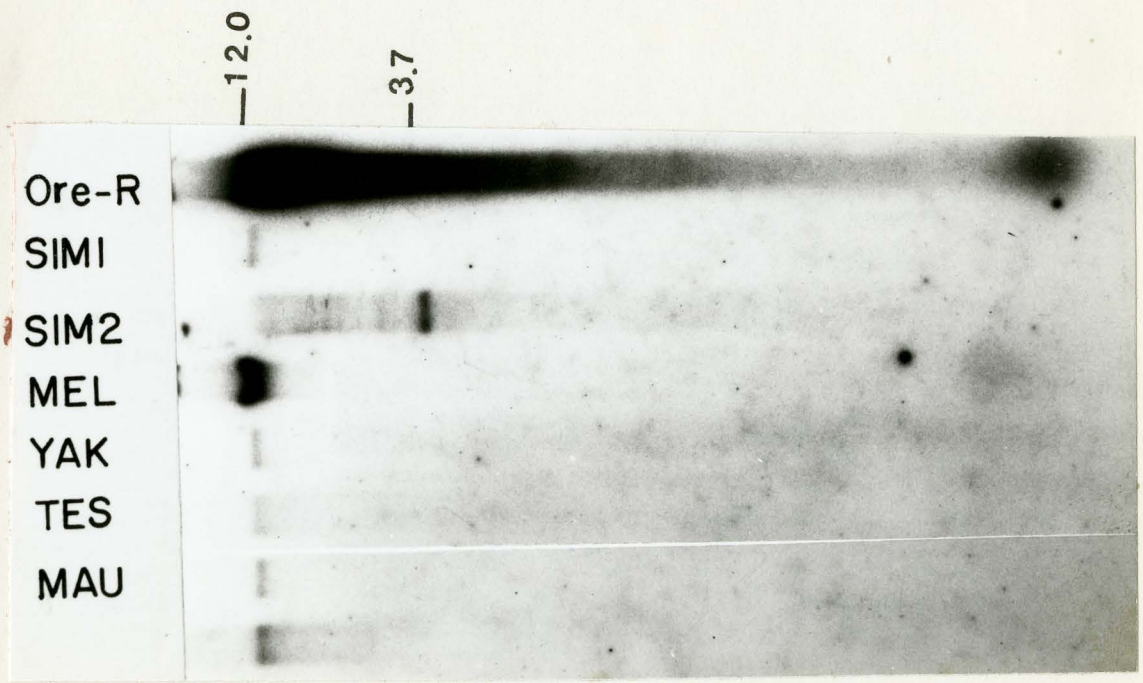


Fig. 14

Figure 16. Eco R1 digestion patterns of genomic digests of Melanogaster group species.

10 µg of fly DNA was digested with Eco R1, separated by electrophoresis in 0.8% agarose gels for 16 hours at 3V/cm, blotted, then hybridized with ^{32}P dm67b probe. 0.7 µgm was used in the hybridization. Hybridization proceeded for 72 hours. The filter was washed with 5X SSC and 0.1% SDS at room temperature. The figure shown is a composite of two exposures. Lanes one to three (Ore-R, Mau, Tak) are of a 21 day exposure to XAR film. Lanes four to six (C-S, Sim, Tes) are of a 7 day exposure to XAR film.

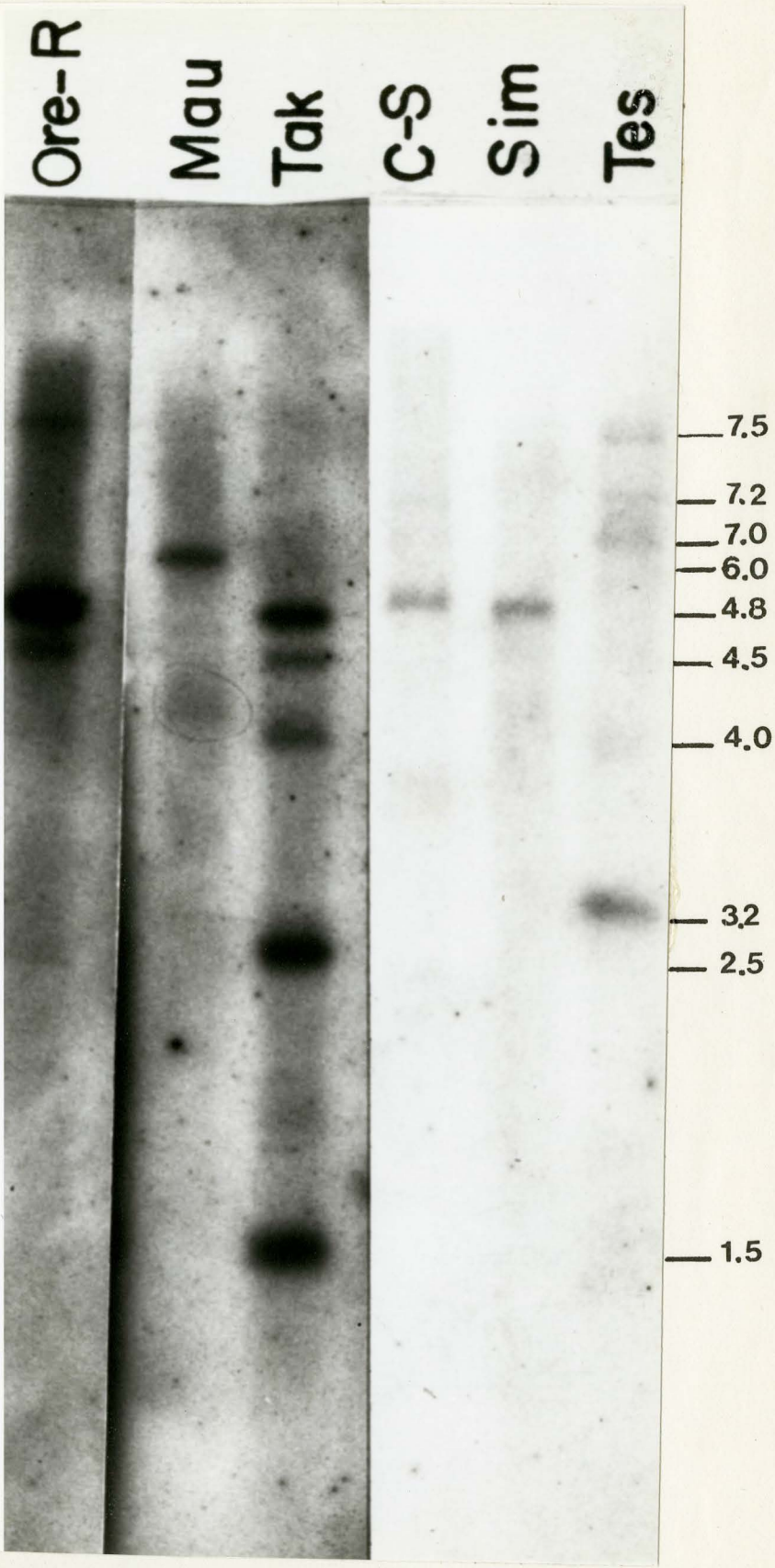
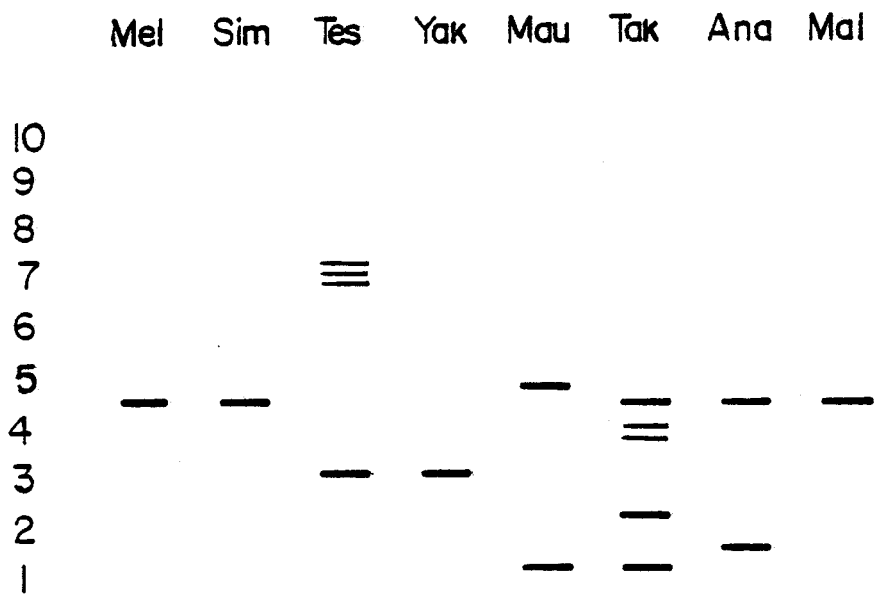

Ore-R
Mau
Tak
C-S
Sim
Tes
7.5
7.2
7.0
6.0
4.8
4.5
4.0
3.2
2.5
1.5

Figure 17. Diagrammatic representation of the pattern of Eco R1 digestion of within Melanogaster group species.


Mel
Sim
Tes
Yak
Mau
Tak
Ana
Mal
10
9
8
7
6
5
4
3
2
1

3.3 Pst 1 Digestion

As shown in figure 18 the patterns of fragments when Pst 1 is used was more clear cut. In fact the same pattern was shared by pairs of species. D. melanogaster Oregon-R and D. melanogaster Canton-S had three identical fragments 3.7, 2.4 and 1.0 Kb. D. tesserie and D. yakuba have four bands 5.5, 3.7, 1.0 Kb and a weakly hybridizing band at 0.5 Kb. D. simulans has one band at 3.7 Kb and D. mauritiana has two bands at 5.5. and 3.0 Kb. D. takahashii and D. ananassae have one band at 4.7 Kb, and D. malerkotliana has one band at 4.5 Kb.

3.4 Phylogenetic Estimates Based on Restriction Data

Using Nei and Li's Method the sequence diversity between species within Melanogaster Group has been calculated. Table 8 shows the estimate of F and δ for the species within this group. The data was analysed in two ways; within subgroup and within group. The value for mean δ was 0.153 with a variance of 0.015 for the Melanogaster subgroup, and 0.397 with a variance of 0.157 for within the Melanogaster group. The value for $\hat{F}$ was also used as a measure of similarity for phenetic and cladistic analysis. The resulting phenogram and phylogenetic tree are shown in figures 19 and 20.

4. VARIATION WITHIN THE GENUS DROSOPHILA

A total of eleven non-Melanogaster group species were examined for restriction enzyme fragment polymorphism using Bam H1 and Eco R1. These represent six species groups from two subgenera. The subgenus Scaptodrosophila is represented by a triad of species in the group Victoria. These are D. victoria and its sibling species D. lebanonensis,

Figure 18. Diagrammatic representation of the patterns obtained with Pst 1 digestion of within Melanogaster group species.

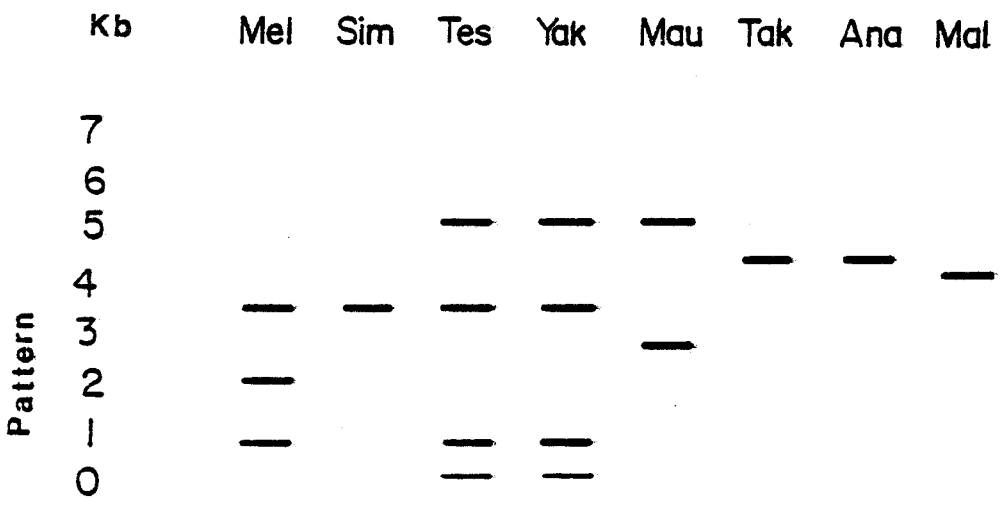

Kb
Mel
Sim
Tes
Yak
Mau
Tak
Ana
Mal
7
6
5
4
3
2
1
0
Pattern

Table 8. The values for F and δ for species within Melanogaster group. Below the diagonal are the values for F and above are the values for δ. The diagonal is equal to zero or one for F and δ respectively.

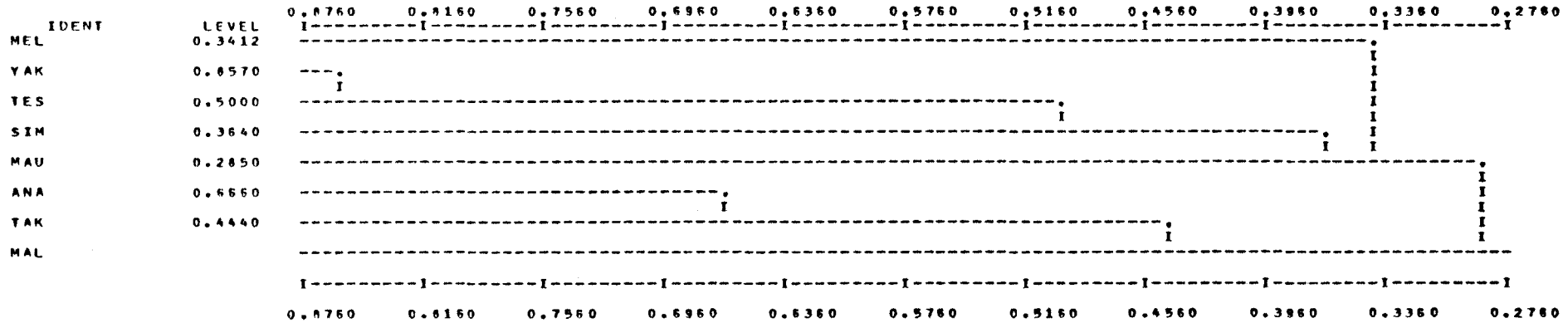
	Mel	Sim	Tes	Yak	Mau	Tak	Ann	Mal
Mel	-----	.137	.147	.129	.458	.221	.209	.221
Sim	.321	-----	.081	.095	.169	.169	.152	.185
Tes	.295	.500	-----	.017	.152	1.000	1.000	1.000
Yak	.341	.444	.857	-----	.121	1.000	1.000	1.000
Mau	.032	.250	.286	.364	-----	.198	1.000	1.000
Tak	.169	.250	0.00	0.00	.200	-----	.046	.109
Ann	.185	.285	0.00	0.00	0.00	.667	-----	.096
Mal	.169	.222	0.00	0.00	0.00	.400	.444	-----

Contractions:
Mel = D. melanogaster
Sim = D. simulans
Tes = D. tesserie
Yak = D. yakuba
Mau = D. mauritiana
Tak = D. takahashii
Ann = D. ananassae
Mal = D. malerkotliana

Figure 19. Phenogram of within Melanogaster group species. Diagram is a reproduction of the phenogram generated by UPGM using NT-SYS phylogenetic analysis computer package.

PHENOGRAM FROM MATRIX SIM 00

Figure 20. Minimum Spanning Tree of similarity values generated by restriction analysis of within the Melanogaster subgroup. (MST of NT-SYS generated output)


IDENT
LEVEL
MEL 0.3412
YAK 0.8570
TES 0.5000
SIM 0.3640
MAU 0.2850
ANA 0.6660
TAK 0.4440
MAL
0.8760 0.8160 0.7560 0.6960 0.6360 0.5760 0.5160 0.4560 0.3960 0.3360 0.2760

and the nonsibling species D. pattersoni. Within the subgenus Sophophora are the groups Melanogaster, Saltans, Obscura, and Willistoni. The species of the Melanogaster group have been described in the earlier section. The Saltans group is represented by D. saltans, and its sibling D. prosaltans of the subgroup Saltans. The third member of the triad is D. emarginata of the subgroup Elliptica. The group Obscura is represented by D. persimilis, and its sibling species D. pseudoobscura, the third member of the triad D. miranda. D. pseudoobscura is represented by two subspecies D. p. pseudoobscura and D. p. bogotana. The last group is the Willistoni group represented by only one species D. insularis.

4.1 Bam Hl Digestion

4.1.1 Victoria Group

As shown in figure 21 the species of this group demonstrates only one band in each species. The level of hybridization is low but bands are discernible. The size of the fragments are 4.7 Kb for D. victoria and D. lebanonensis and 5.1 Kb for D. pattersoni.

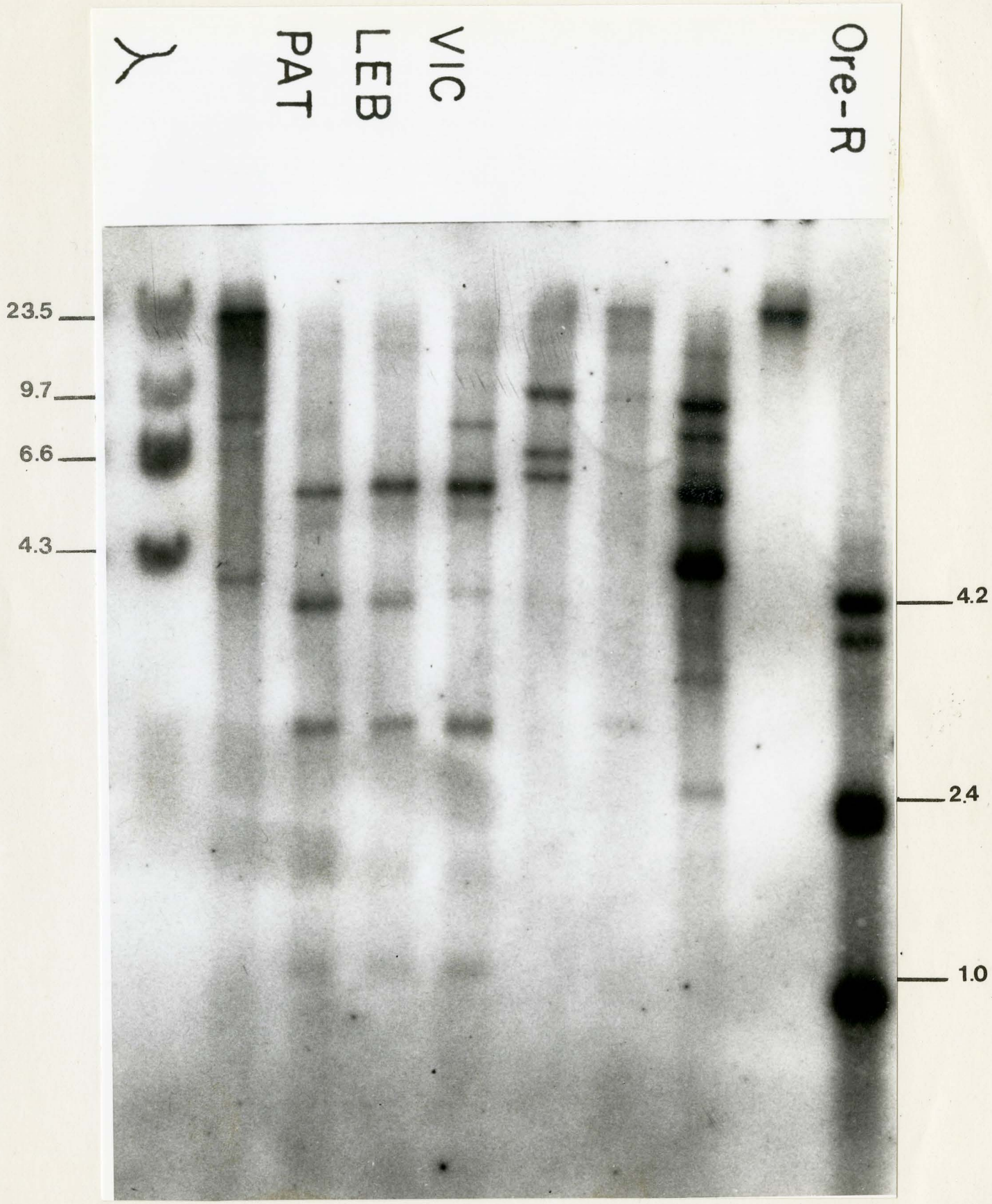
4.1.2 Saltans Group

The species D. saltans and D. prosaltans have one band of 5.1 Kb while D. emarginata has one band at 5.4 Kb.

4.1.3 Obscura Group

The level of hybridication was low in this group and the results equivocal. However it appeared that in all four strains a 0.9 Kb fragment was present. D. p. psuedoobscura and D. p. bogotana also share a 3.0 Kb fragment. D. p. pseudoobscura also had another band at 4.4. Kb and D. persimilis had a fragment at 6.7 Kb.

Figure 21. Victoria group species pattern when digested with Bam H1. 10 µg of fly DNA was digested with Bam H1, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, then hybridized to ^{32}P dm67b probe. 0.5 µgm of ^{32}P-dm67B probe at a specific acitivty of 1.6 X 10^8 cpm/µgm was used in the hybridization. Hybridization proceeded for 80 hours. The filter was washed with 5X SSC and 0.1% SDS at room temperature. The exposure shown was to XAR film for 14 days.


Y
PAT
LEB
VIC
Ore-R
23.5
9.7
6.6
4.3
4.2
2.4
1.0

4.1.4 Willistoni Group

The species D. insularis demonstrated one band at 4.7 Kb as shown in figure 21.

4.2 Eco R1 Digestion

4.2.1 Victoria Group

As shown in figure 22 the pattern for Eco R1 in the Victoria group is consistent for all three members. All three have fragments of 5.7, 4.0, 3.0 and 1.2 Kb. D. victoria also has a fragment at 7.8 Kb.

4.2.2 Saltans Group

The species D. saltans and D. prosaltans shared no fragments. The pattern for D. saltans was 5.8, 3.5 and 3.1 Kb. D. prosaltans had one band at 9.8 Kb. This 9.8 Kb fragment was shared with D. emarginata. D. emarginata also had a fragment at 3.5 Kb which it shared with D. saltans.

4.2.3 Obscura Group

The pattern of hybridization shows only one band for each strain. The sizes found were D. persimilis, 4.8 Kb, D. p. pseudoobscura, 4.4 Kb, D. p. bogotana, 4.3 Kb, and D. miranda, 4.3 Kb.

4.2.3 Willistoni Group

D. insularis demonstrated one fragment at 4.8 Kb.

4.3 Phylogenetic Estimates Based on Restriction Data

The estimates of sequence divergence was analysed in steps; between sibling species pairs, between non-sibling species in the same group, and between species groups in the genus Drosophila.

Shown in table 9 is the value for F and δ for the sibling species pairs studied. The mean of the δ for the five comparisons is 0.158

Figure 22. Victoria group pattern when digested with Eco R1. 10 µg of DNA was digested with Eco R1, separated by electrophoresis in 0.8% agarose gels at 3V/cm for 16 hours, blotted, then hybridized to ^{32}P dm67B probe. 0.66 µgm of ^{32}P dm67B probe at a specific activity of 1.6×10^8 cpm/µgm was used in the hybridization. Hybridization proceeded for 80 hours. The filter was washed with 5X SSC and 0.1% SDS at room temperature. The figure shown is a composite of two exposures. The lane labelled INS was a 21 day exposure to XAR film. The lanes Leb, Pat, and Vic were from an exposure of 7 days to XAR film.

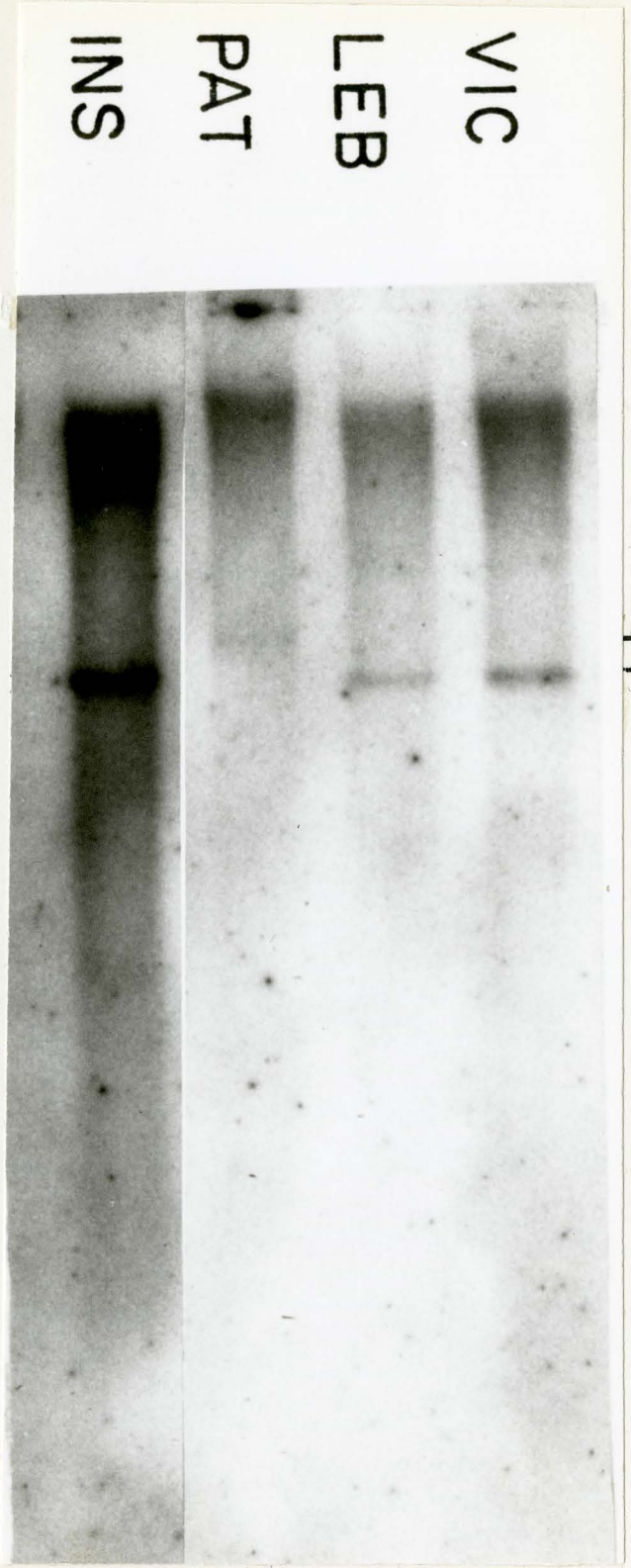

INS
PAT
LEB
VIC
5.1
4.7

Table 9. The values for F and δ for all sibling species pairs studied.

Species 1	Species 2	F	δ
Mel	Sim	0.321	.137
Vic	Leb	0.910	.010
Sal	Pro	0.333	.132
Per	Psp	0.333	.132
Per	Psb	0.285	.152
All			
Mean (Variance)			.1126 (.0034)
Without Vic-Leb comparison			.1385 (.0000896)
Mean (Variance)			

Contractions:
Vic = *D. victoria*
Leb = *D. lebanonensis*
Sal = *D. saltans*
Pro = *D. prosaltans*
Per = *D. persimilis*
Psp = *D. p. pseudoobscura*
Psb = *D. p. bogotana*

with a variance of 0.0034.

Shown in table 10 is the value for F and δ for all non-sibling species pairs within the same subgroup. The mean of the value for δ is 0.1233 with a variance of 0.00143. As shown in table 11 the values are frequently 0 and 1 for F and δ respectively, when comparisons outside a species group are made.

The values for F will be used as coefficients of similiarity for analysis. The phenogram of the species relationship is shown in figure 23. The phylogenetic tree using cladistic analysis is shown in figure 24.

Table 10. The values for F and δ for all nonsibling species from the same subgroup.

Species 1	Species 2	F	δ	Group Mean
Mel	Tes	0.295	.147	
Mel	Yak	0.341	.129	
Mel	Mau	0.032	.458	
Sim	Tes	0.500	.081	
Sim	Yak	0.444	.095	
Sim	Mau	0.250	.169	
Tes	Yak	0.857	.017	
Tes	Mau	0.286	.152	
Yak	Mau	0.364	.121	.1521
Ana	Mal	0.444	.096	.096
Pat	Vic	0.600	.059	
Pat	Leb	0.600	.059	.059
Mir	Per	0.400	.109	
Mir	Psb	0.800	.025	
Mir	Psp	0.333	.132	.08867
Grand Mean				0.1233
Variance				0.00143

Contractions: Pat = *D. pattersoni*
Mir = *D. miranda*

Table 11. Values of F and δ for all species studied in the genus Drosophila. Below the diagonal are the values for F and above are δ. The values of the diagonal are zero and one respectively.

	Mel	Sim	Tes	Yak	Mau	Tak	Ann	Mal	Sal	Pro	Ema	Per	Psp	Psb	Mir	Ins	Vic	Leb	Pat
Mel	-----	.137	.147	.129	.458	.221	.209	.221	1.00	1.00	1.00	.117	1.00	1.00	1.00	.060	1.00	1.00	1.00
Sim	.321	-----	.081	.095	.169	.169	.152	.185	1.00	1.00	1.00	.109	1.00	1.00	1.00	.081	1.00	1.00	1.00
Tes	.295	.500	-----	.017	.152	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yak	.341	.444	.857	-----	.121	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Mau	.032	.250	.286	.364	-----	.198	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Tak	.169	.250	0	0	.200	-----	.046	.109	1.00	1.00	1.00	.152	1.00	1.00	1.00	.133	1.00	1.00	1.00
Ann	.185	.285	0	0	0	.667	-----	.096	1.00	1.00	1.00	.132	1.00	1.00	1.00	.109	1.00	1.00	1.00
Mal	.169	.222	0	0	0	.400	.444	-----	.169	.132	.152	.132	1.00	1.00	1.00	.133	1.00	1.00	1.00
Sal	0	0	0	0	0	0	0	.250	-----	.132	.152	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.2564
Pro	0	0	0	0	0	0	0	.333	.333	-----	.109	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.152
Ema	0	0	0	0	0	0	0	.286	.286	.400	-----	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Per	.374	.400	0	0	0	.285	.333	.333	0	0	0	-----	.152	.132	.109	.081	1.00	1.00	1.00
Psp	0	0	0	0	0	0	0	0	0	0	0	.285	-----	.132	.025	1.00	1.00	1.00	1.00
Psb	0	0	0	0	0	0	0	0	0	0	0	.333	.333	-----	.064	1.00	1.00	1.00	1.00
Mir	0	0	0	0	0	0	0	0	0	0	0	.400	.800	.572	-----	1.00	1.00	1.00	1.00
Ins	.596	.500	0	0	0	.333	.400	.333	0	0	0	.500	0	0	0	-----	.185	.169	1.00
Vic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.222	-----	.010	.059
Leb	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.250	.910	-----	.059
Pat	0	0	0	0	0	0	0	0	.222	.285	0	0	0	0	0	0	.600	.600	----

Contractions: Ins = D. insularis

Figure 23. Phenogram of the species studied in the genus Drosophila. Diagram is a reproduction of the phenogram generated by UPGM using NT-SYS phylogenetic analysis computer package.

PHENOGRAM FROM MATRIX SIM 00

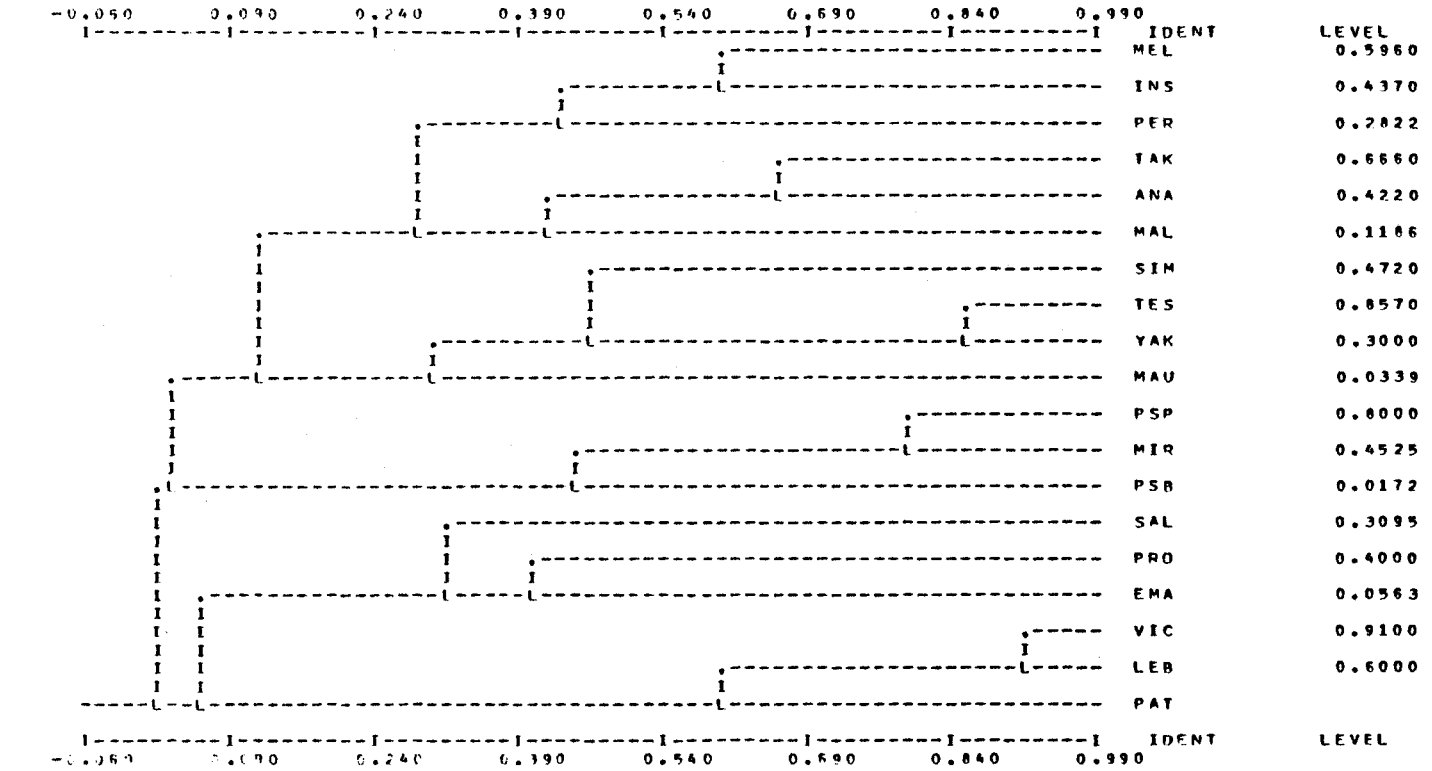
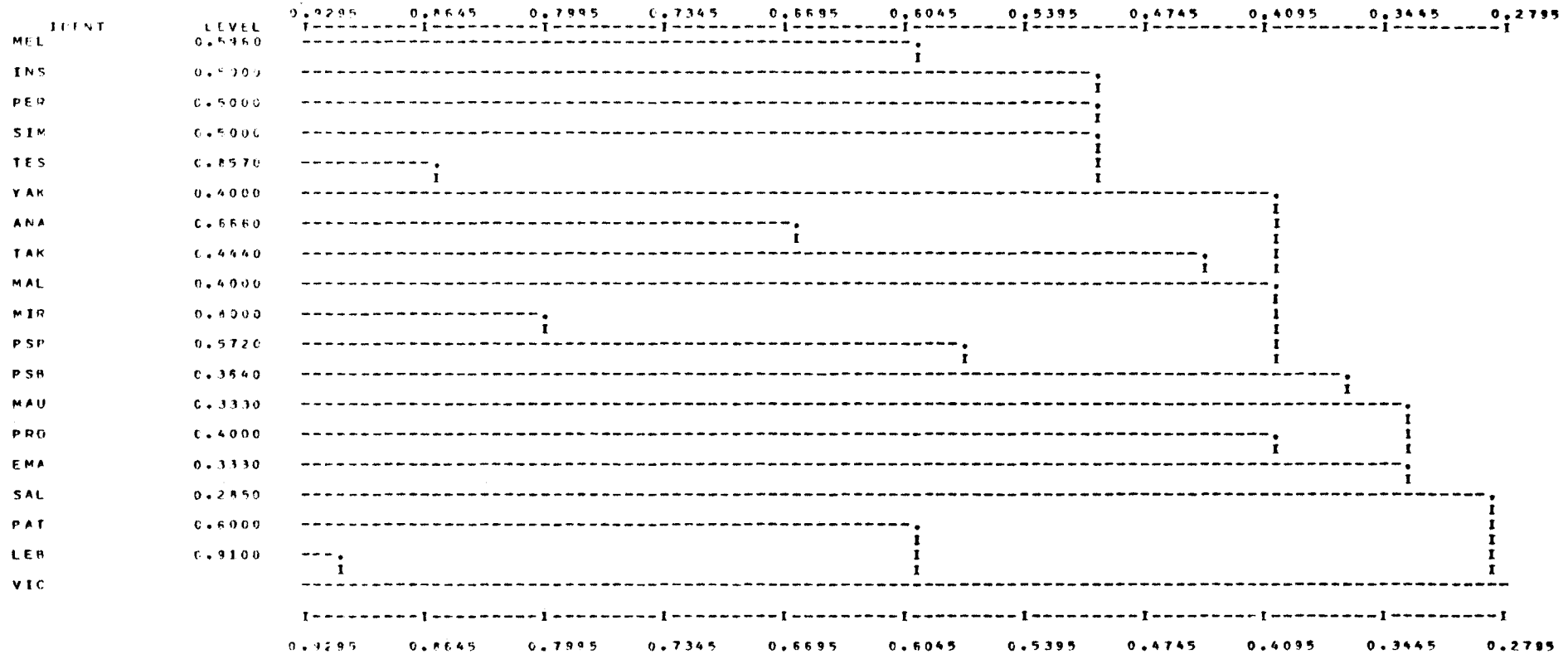


Figure 24. Minimum Spanning Tree of the similarity values generated by restriction analysis of within the genus Drosophila (MST of NT-SYS generated output).


IDENT
LEVEL
0.9295
0.8645
0.7995
0.7345
0.6695
0.6045
0.5395
0.4745
0.4095
0.3445
0.2795
MEL 0.5960
INS 0.5000
PER 0.5000
SIM 0.5000
TES 0.8570
YAK 0.4000
ANA 0.6660
TAK 0.4440
MAL 0.4000
MIR 0.4000
PSP 0.5720
PSB 0.3640
MAU 0.3330
PRO 0.4000
EMA 0.3330
SAL 0.2850
PAT 0.6000
LEB 0.9100
VIC

DISCUSSION

1. DROSOPHILA PHYLOGENY

The Drosophila phylogeny has been described in detail by Throckmorton (Throckmorton 1975, 1977, 1982). He proposes that there are five major radiations in the genus Drosophila. These are in order of divergence, the subgenera Scaptodrosophila, Sophophora, Hirtodrosophila, and Drosophila. The subgenus Drosophila has two lineages of interest- the Virilis-Repleta and the Tripunctata radiations (see figure 1). By use of fossil evidence and geological and climatic change estimates divergence times have been estimated for these radiations and speciation events within the subgenera. Proposed times of divergence are: Scaptodrosophila-Sophophoran splitting in mid-Oligocene (approximately 35 million years ago), the Melanogaster-Saltans splitting in mid-Miocene (approximately 15 MYA), and the D. miranda-D. persimilis splitting about 5 MYA (Throckmorton 1975).

The subgenus Scaptodrosophila is thought to be the most ancient of the subgenera, arising in tropical Asia in the Oligocene period (Throckmorton 1975, Bock and Parsons 1978). The Victoria group is a representative of the subgenus and the three species studied are from this group (Patterson and Stone 1952).

The subgenus Sophophora is broken into two major radiations. The Melanogaster group is proposed to have arisen as a protomelanogaster lineage in the mid-Oligocene (Throckmorton 1975). The major radiations are the Willistoni-Saltans radiation and the Melanogaster-Obscura

radiation. The Willistoni and Saltans groups are more similar to each other morphologically than they are to the Melanogaster group or the Obscura group, which are closely related to each other morphologically (Throckmorton 1975).

The Saltans group is made up of five recognized subgroups. These are in order of primitive to advanced: Cordata, Elliptica, Sturtevanti, Parasaltans, and Saltans. The species studied in this project were D. emarginata of the Elliptica subgroup and the species D. prosaltans and D. saltans of the Saltans subgroup (Throckmorton 1975, Patterson and Stone 1952).

The group Willistoni is made up of three subgroups: Willistoni, Fumipennis, and Alagitans-Bocainensis (Patterson and Stone 1952). The species D. insularis is a member of the Willistoni subgroup (Burla etal 1949).

The Obscura group has two subgroups: Obscura and Affinis. There are eleven recognized species of the subgroup Obscura (Patterson and Stone 1952). D. miranda, D. persimilis, and D. pseudoobscura are a triad of closely related species in this subgroup (Dobzhansky and Epling 1944). Lakovaara (1972) clusters the species D. persimilis, D. miranda, D. pseudoobscura, and D. lowei as a separate lineage that diverged from the rest of the subgroup early in the phylogeny of this subgroup.

The Melanogaster group is made up of over 150 species in thirteen subgroups (Bock and Wheeler 1972, Bock 1980). The subgroups of the Melanogaster group are thought to be a cohesive group showing strong morphological similarities (Bock 1980). The relationships within the group are not clearly defined. Represented in this study are species

of the subgroups Melanogaster, Takahashii, and Ananassae.

D. takahashii, D. ananassae and D. malerkotliana are not sibling species but are fully reproductively isolated from each other. The species of the Melanogaster subgroup on the other hand are very closely related and in some cases can mate and produce offspring of low fertility and viability (Lemeunier and Ashburner 1976). The species studied are D. melanogaster, D. simulans, D. tesserie, D. yakuba, and D. mauritiana. Lemeunier and Ashburner (1976) propose a phylogeny based on polytene chromosome bands to be the following. The ancestral pattern of bands is found in D. simulans and D. mauritiana. One inversion difference leads to D. melanogaster. Four changes lead to a hypothesized ancestor of D. erecta, D. yakuba, D. tesserie. D. yakuba and D. tesserie differ for four inversions but share one. D. erecta differs from the hypothetical ancestor by five chromosomal changes.

2. GENE ORGANIZATION

Before dicussing the results for sequence variation I would like to comment on the organization of the region in D. melanogaster. If one referes to figure 7 then one notes that four cleavage sites fall within the coding region or the S' noncoding region of these genes. While proposed maps are tentative, there are no changes at these sites based on the maps. Corces et al. (1980) have shown that the sequence of the coding regions of these genes are similar. From this work the noncoding sequences also appear to be similar between gene pairs. The maps proposed in the case of Eco R1 digestion pattern describe a symmetry about the middle of the cluster. In the Pst-1 digestion pattern a common

Table 12. The mean and variance of δ, and the number of comparisons used to generate this estimate for differing level of species divergence.

Level	δ	Var. (δ)	Number
Within a population (Between individual)	.0249	.000367	2
Within a species (Between populations)	.03398		1
Between sibling species	.1126	.0034	5
Between nonsibling species	.1233	.00143	15
Between species of different group	.8291	.1229	69
Between species of different subgenera	.9480	.0413	48

sized fragment may contain both gene pairs separately. This confirms the hypothesis that these four loci are results of tandem duplications.

3. SEQUENCE VARIATION AND GENETIC DIFFERENTIATION

I have attempted to quantify the amount of sequence variation at various levels of species divergence in the genus Drosophila. The results are summarized in Table 12.

I wish to compare the results of DNA divergence with the data on protein divergence at the same levels. However, because of the variety of data available, different authors have used various indices of similarity. They vary between genetic distance which takes into account the allele frequencies at the loci studied (Nei 1976) to percent shared protein bands when the genetic basis is not clearly known (Hubby and Throckmorton 1968).

The variation in proteins in the population of D. melanogaster have been studied including the two populations used in this study. The summary of the information is shown in Table 13.

In Table 14 are estimates of divergence for protein variation at differing levels of species divergence. The values attained in studies of the Mulleri and Willistoni subgroup are on enzyme loci whereas all other data is on protein loci stained nonspecifically. In Table 15 are the results of Hubby and Throckmorton's work on protein loci in triads of species in the genus Drosophila (Hubby and Throckmorton 1968).

The variation of heat shock proteins in the 20 to 30 Kilodalton range has alse been investigated by one dimensional SDS PAGE. The

Table 13. Summary of I, D, Fst and δ for populations of D. melanogaster studied in this survey. (Data from Singh, Hickey and David, 1982, Singh and Coulthart 1982)

No. Populations Studied	Enzyme or Protein	No. of Loci	I	D[1] or δ	Fst
9	Enzyme	22	0.866	0.149	.174
(Worldwide)	Protein	20	0.963	0.038	.069
2	Enzyme	22	0.898	0.097	0.160
(Dalewood-Brownsville)	Protein	20	0.989	0.011	0.107
	DNA	8[2]		0.034	0.068

[1] D for enzyme or protein data and δ for DNA data

[2] Number of restriction sites

Table 14. Some Similarity Estimates at Different Levels of Diveregence in Drosophila

Nei's Genetic Distance Estimates		
Level	Mulleri subgroup[1]	Willistoni subgroup[2]
Within a population	0.001	0.003
Within subspecies	0.025	0.228
Between sibling species	0.168	0.538
Between nonsibling species	0.292	1.214

Amino Acid Substitutions	
Level	Obscura subgroup[3]
Within European species	2.7
Within American species	2.0
Between American and European species	4.1

General similarity between monomorphic and polymorphic loci	Hawaiian Drosophila[4]
Both species monomorphic for same allele (1)	0.31
One monomorphic and one polymorphic with common alleles (2)	0.21
One monomorphic and one polymorphic with no common alleles (3)	0.162
Both polymorphic for different alleles (4)	0.155
Both polymorphic for same and different alleles (5)	0.121
Both polymorphic with no common alleles (6)	0.034
Both polymorphic for same alleles (7)	0.007
1+2+7	0.527
3+4+5+6	0.473

1 Zouras (1973)
2 Ayala etal (1973, 1974)
3 Lakovaara (1972)
4 Rockwood etal (1971)

Table 15. Shared proteins between species of species triads in Drosophila. I means that sibling species shared bands only, II means siblings and nonsibling species share bands only, and III means bands shared between all three members of the triad. (From Hubby and Throckmorton 1968)

	I	II	III	I+III	II+III
arizonensis mojavensis mulleri	42.1	6.3	6.3	48.4	12.6
mercatorum paranensis peninsularis	55.0	11.8	11.8	66.8	23.6
hydei neohdei eohydei	43.8	3.2	6.3	50.4	9.8
fulvimaculata fulvimaculoides lemensis	50.0	13.2	15.8	65.8	29.0
melanica paramelanica nigromelanica	26.3	10.0	5.3	31.6	15.3
melanogaster simulans takahashii	52.9	7.9	0.0	52.9	7.9
saltans prosaltans emarginata	36.8	7.7	10.5	47.3	18.2
willistoni paulistorum nebulosa	7.1	11.6	15.4	22.5	27.0
victoria lebanonensis pattersoni	42.0	0.0	21.4	85.7	21.4
Average	42.0	7.9	10.4	52.4	18.3

population differentiation for heat shock proteins was zero by definition as no variants were found. Morgan and Singh (personal communication) have found that 49.7% of the small hsps proteins are common between two sibling species. They also found that nonsibling species share 35.5% of the same heat shock proteins in this molecular weight range. In an examination of within species variation, they found no variants within D. melanogaster using the same strains studied here for DNA sequence variation. These values are the same estimates as I+III and II+III of Hubby and Throckmorton (1968).

These results show that within populations on the average sequences of DNA studied here are less variable than proteins and enzymes. It is difficult to equate the two measures of variation in a numerical sense but the use of F_{st} index calculated for the three types of 'loci' studied have different values. Enzyme loci show the greatest fixation index and DNA the least. A second approach to compare the sequence results obtained is to describe coding region changes. The level of coding region sequence variation within D. melanogaster for the four genes studied was zero. This is also shown by the lack of apparent molecular weight variants in the same populations and strains for these proteins (Morgan and Singh personal communication). This argues for apparent conservation of sequence in both the amino acid constitution and DNA sequence for the small heat shock proteins within a species. In comparison to the results for protein and enzyme loci these values of divergence are extremely low as in D. melanogaster the proportion of monomorphic protein or enzyme is .278 (Singh, Hickey and David 1982), and the probability of chosing randomly four monomorphic

loci is 0.006.

Looking at divergence above the species level it is better to compare the DNA sequence diversity data and the protein data in Drosophila in terms of magnitude of the change at differing levels of species divergence. Refering to Table 12 the levels for divergence between sibling species, and between nonsibling species are 0.1126 and 0.1233 respectively for DNA sequence. In studies of enzyme loci (Table 14) they are 0.168 to 0.538, and 0.292 to 1.214 for the same two levels. For protein loci (Table 15) the values are 50.0% similar for sibling species, and 18.3% similar for nonsibling species. In studies of the small heat shock proteins similarity of 50% and 35.5% are found for sibling species and nonsibling species.

One aim of these studies is to be able to compare the rate of evolution of various classes of sequence at various levels of taxonomic divergence. One way to study rates of evolution when no fossil evidence is available is to study the differentiation exhibitied by classes of sequences at differing levels of taxonomic divergence. Even if the rates of evolution differ between classes of sequences, their ratio at the two levels indicated above should be similar if they are being acted on by the same mechanisms to moderate variation. Following from this, if there is differences between the ratio for different types of sequences then different mechanisms would be indicated.

The amount of divergence observed between nonsibling species is 1.74 (Zouros 1973) to 2.26 times (Ayala et al 1973, Ayala et al 1974), or approximately twice as large as between sibling species using enzymatic loci as a

means of measuring variation. The same measure of difference in divergence is 2.9 times for protein loci, 1.4 times for heat shock proteins, and 1.08 times for DNA sequence of region of the four small heat shock proteins.

One expects greater differentiation of nonsibling species and this is clearly shown for protein and enzyme loci. However the heat shock proteins show that differentiation between nonsiblings is not much more than sibling species. This protein result agrees with the very little amount of divergence for the DNA coding sequences for nonsibling compared to sibling species.

Before discussing this point further I would like to present some results on sequence variation in Drosophila and other organisms.

4. COMPARISON OF DNA SEQUENCE VARIATION

In current studies of satelite DNA in Drosophila, sequence variation within a species has been studied. In the following species the divergence between two homologous repeat classes was determined to be between 10 and 12%, in *D. grimshawi*, *D. gymnobasis*, and *D. silvarentis* (Miklos and Gill 1980). In this study there were examples of divergence due to both base changes and deletions. Carlson and Brutlag (1979) by restriction enzyme analysis found that the 1.688 gm/cm^3 satellite of *D. melanogaster* is highly conserved. While three variants of this class were found each variant was in high frequency and homogeneous within a variant class. One of these variants differed by a base change in the repeat sequence that apparently swept through an array, and the second was a deletion variant which again had reached

high frequencies.

The satellite and main-band DNA of the species of the Melanogaster group were studied by Barnes et al. (1978). A review of these results are given in Dover (1982). The main-band DNA was common between all members of this subgroup. For the ten satellite classes found, some were shared by two or more species, some were unique to a single species. Mullens and Blumenfeld (1979) found a short 7 bp repeat present in D. virilis and in D. melanogaster. In an attempt to find similar sequences or variants of this sequence in the genome of either species these researchers failed to find any homologous repeat from which these arose independently in either lineage.

The ribosomal RNA genes of the Melanogaster subgroup have been studied for sequence variation and spacing. The Tm of heteroduplexes between D. melanogaster rRNA to the genes of these sequences are indistinguishable. The Tm of heteroduplexes between D. melanogaster rRNA and rDNA of the species D. erecta, D. tesserie, and D. yakuba are reduced by 2.4 to 3.1 ^{o}C implying approximately 4.7% mismatch. However when D. melanogaster rDNA spacer sequence is hybridized to the genomic fragments of all six species there is less mismatch (D. simulans 2.1%, D. mauritiana 2.9%, D. erecta 0.07%, D. tesserie 3.0% and D. yakuba 2.17%) (Tartoff 1979). Tartoff concludes that rDNA space region is more highly conserved than the coding region of these genes within this subgroup.

In the 28S ribosomal genes in both D. melanogaster and D. virilis there is a 9.6 Kb intervening sequence. In both species it occurs at approximately the same place in the gene. However they differ markedly for sequence. There are in the genome of D. melanogaster sequences similar to the intervening sequence of this species 28S gene. In D. virilis also there are sequences similar to the intervening sequence of the 28S gene. But there is no sequence in D. melanogaster homologous to the sequence found in D. virilis at any of these sites (Barnett and Rae 1979).

Very little single copy or small family sequence divergence information is available in Drosophila. Work in mammalian genomes is available however. Miyata (1982) has reviewed some of this data. In comparisons of β-Globin, α-Globin, preproinsulin and growth hormone between rat, mouse and man the mean sequence divergences (δ) are the following: 5' noncoding 0.247 ± 0.060 SD, 5' portion of the 3' noncoding 0.451 ± 0.037, 3' portion of 3' noncoding 0.195 ± 0.024, and synonymous site changes in the coding region was 0.487 ± 0.020, and 0.129 ± 0.039 amino acid substitution site changes in the coding region. The time of divergence for these three organisms is thought to be similar and on the order of 20 MYA.

In man work has been performed on variation in sequence for the globin genes. Jeffreys (1979) chose 60 unrelated individuals to test for sequence variation. He found three restriction enzyme variants all of which probably fell in the intervening sequences of these genes. He found no variants in the spacer DNA between the coding

blocks. In terms of sequence diversity the estimate of δ was 0.01 for the sixty individuals studied.

The last report I wish to present is on DNA variation at the 70 Kilodalton hsp loci in D. melanogaster, D. simulans, and D. mauritiana (Leigh - Brown and Ish-Horowicz 1981). In a study of one individual of each species these researchers determined the sequence variation at all copies of the 70K hsp. There are four copies in D. simulans and D. mauritiana, two at each of the cytological loci 87A and 87C. In D. melanogaster there are five copies, two at 87A and three 87C. All copies are in the inverted repeat form except for 87C in D. melanogaster where a 40 Kb insert containing one copy of the gene and a highly repeated sequence are inserted between the two copies. In a comparison of both coding and noncoding changes the following estimates were determined. D. melanogaster-D. simulans 0.0213 coding and 0.0295 noncoding, D. melanogaster -D. mauritiana 0.0213 coding and 0.0327 noncoding, and D. simulans-D. mauritiana 0.00 coding and 0.0257 noncoding.

Returning to explanations for the differences in rate of evolution between sequences and between taxonomic groups the following hypothesis are suggested. The first is that there are differences in the rate of evolution for different classes of sequence. Evolutionary rates are highly variable for proteins. The differences in amino acid substitution rate between fibrinopeptide and cytochrome c are very large and cytochrome c has a much lower rate of change (Miyata 1982).

A recent review by Gojobori (1982) examined the relationship between heterozygosity at a locus and its function. He was able to

class heterozygosity by function of the protein coded by the loci in question. Lower heterozygosities were associated to their involvement in main metabolic pathways, and higher heterozygosities to enzymes with less stringent substrate affinities and involvement in more than one pathway. Singh, Hickey, and David (1982), and Coulthart and Singh (1982) demonstrate similar relationships for groups of enzymes and proteins.

In the hypothesis the heterozygosity is a function of neutral mutation rate. If the increase in heterozygosity is not a function of balancing selection, then this equates variation to functional constraints of the molecule. I feel this is an adequate explanation for the lack of variation for the coding sequence of the heat shock proteins. There appears to be a continuum of the rate of evolution for coding sequence based on structural constraints on the molecule. The more variable proteins show more differentiation because differentiation is a function of time and nonsibling species have diverged for a longer period of time than sibling species.

A second factor that relates intensity of selection to rate is the extremely low rate of 5' flanking sequence diversity. This flanking region of the gene is the area shown to be involved in gene regulation. Both the results quoted from Miyata (1982) and in this study describe low substitution rates in this region. This supposes that the 5' flanking region is a noncoding sequence under intense selection and therefore shows little variation due to the constraints on its function.

However, this hypothesis does not explain the differences in the rates of evolution at different taxonomic levels as demonstrated

in the ratio of sibling species to nonsibling species diversity. This has lead G. Dover (1983, Dover et al. 1982) to propose a mode of evolution he terms molecular drive. Sequence homogenization between families of genes on chromosomes (ie. clustered repeat families, rDNA, hsp 70 loci). Homogenization between families of repeated sequences within and between chromosomes (ie. repetitive sequences, satellite DNA, dispersed gene families), and sequence homogenization of DNA sequence between individuals within a species is said to occur as a result of molecular drive. The mechanisms he envisions are unequal crossing over, gene conversion, and DNA transposition (Dover 1982). This phenomena would maximize variation between species and minimize it within a species. The splitting off of a gene pool may result in a new incipient species that rapidly become genomically dissimilar as a result of fixation of variable sequences. This fixation of variable sequence would be due to chance that the new incipient species does not represent the variation in the larger population exactly. This difference could be fixed rapidly because of molecular drive. After isolation molecular drive would reduce the rate of sequence change. This predicts that the rate of change during speciation (cladogenesis) is greater than the rate of change during phyletic evolution (anagenesis).

This hypothesis is essentially a stochastic process acting on the genome. A second hypothesis is selectionist. Two researchers currently have commented on rates of change during speciation.

Mayr (1982) has attributed to the genome of a species a genotypic cohesion. This causes an evolutionary stasis of the phenotype

and genotype of a species. He proposes that genotype cohesion is loosened up at the time of speciation. The genotype is broken down because of new population numbers and a different environment. Mayr in this model of peripatric speciation emphasizes that natural selection can result in rapid genetic change during speciation. The new species achieves a new balanced system which induces relative stasis (Mayr 1982). Sewall Wright (1982) on the other hand predicts stasis of the genotype as a result of random genetic drift in his shifting balance theory. Rapid change of the genome, predicted during speciation is because the local differentiated populations responds to ecological opportunity. The incipient species already is divergent from the other populations and becomes more so rapidly because peak-shifts are favoured in the new population size and environment.

Thus as Mayr (1982) and Wright (1982) suggest, if the rate of evolution is relatively rapid than that presumed during anagenesis, it would lead to a differential rate of evolution at the various taxonomic levels as seen in the protein and DNA data in this study.

5. DNA SEQUENCE VARIATION USES IN PHYLOGENETIC ANALYSIS

Use of DNA sequence data for phenetic and cladistic analysis has been proposed recently as a means of building phylogenies. Sokal and Sneath (1973) feel that this type of similarity estimate can be treated as any other character. It may or may not be a true idicator of phylogenetic distance as based on conventional morphological characteristics.

The analysis of similarity coefficients shows that DNA sequence variation does cluster species in a phylogenetically correct manner.

Refering to figure twenty four one sees that the dendrogram diverges the subgenera Sophophora and Scaptodrosophila in a manner that agrees with the phylogenies based on morophological characteristics (Throckmorton 1975). The clustering of the Saltans group, three members of the Obscura group and most of the Melanogaster group in a manner roughly consistent with the scheme based on non-molecular characters. However, in contrast to the established phylogenies the clustering placed D. insularis and D. persimilis in the Melanogaster subgroup.

The cluster analysis was attempted as a rough test of random sequence convergence. If the occurrance of the same fragment between two widely divergent source was random the resulting phenogram would also have been random. However the phenogram is, as stated earlier, roughly consistent with morphological phenograms. This implies that the similarities seen are for at least some part lineal random similarities.

The use of DNA sequence similarity is just one way of increasing the number of characters used in taxonomic analysis and at this time cannot be weighted above or below other characters.

6. FUTURE RESEARCH

The use of Southern Blot Hybridization analysis of whole genomic digests has drawbacks. The technique leads to possibilities of artifacts in migration distances due to salt conditions of the sample applied to the gel. As seen in the autoradiograms of the Victoria group (figures 21 and 22) the level of hybridization is variable and the level above background low. Due to these problems I have conservative in interpertation of

migration distance and identity of bands. Beyond commenting on these problems, I would hope that these problems are considered when the results specific to this thesis are considered. Furthermore I feel future work should include the isolation of the region of interest by means of subcloning the fragment from each strain and analysis of this clone in a more rigorous way.

LITERATURE CITED

Altschuler, M. and J.P. Mascarenhas, (1982), Heat shock proteins and effects of heat shock in plants, Plant Mol. Biol. 1: 103-115.

Ashburner, Michael and J. Jose Bonner, (1979), The induction of gene activity in Drosophila by heat shock, Cell 17: 241-254.

Ayala, F.A., M.T. Tracey, L.G. Barr, J.F. McDonald, S. Perez-Salas, (1974), Genetic Variation in Natural Populations of Five Drosophila Species and the Hypothesis of Selective Neutrality of Protein Polymorphisms, Genetic 77: 343-384.

Ayala, F.J., M.L. Tracey, D. Hedgecock, and R.C. Richmond, (1973), Genetic Differentiation During the Speciation Process in Drosophila, Evol. 28: 576-592.

Barnes, Stephen, David Webb, and Gabriel Dover, (1978), The Distribution of Satellite and Main-Band DNA Components in the Melanogaster Species Subgroup of Drosophila; I. Fractionation of DNA in Actinomycen D and Distamycin A Density Gradients, Chromosoma 67: 341-363.

Barnett, T. and P.M.M. Rae, (1979), A 9.6 Kb Intervening Sequence in D. virilis rDNA and Sequence Homology in rDNA Interruptions of Diverse Species of Drosophila and Other Diptera, Cell 16: 763-775.

Barnett, T., M. Altschuler, C.N. McDanile, and J.P. Mascarenhas, (1980), Heat Shock Induced Proteins in Plants, Devel. Genet. 1: 331-340.

Bock, Ian, (1980), Current status of the Drosophila melanogaster species - group (Diptera), Syst. Ento. 5: 341-356.

Bock, I., and P.A. Parsons, (1978), The subgenus Scaptodrosophila (Diptera: Drosophilidae), Syst. Ento. 3: 91-102.

Bock, Ian and Marshall R. Wheeler, (1972), I. The Drosophila melanogaster Species Group, Studies in Genetics VII., Univ. Texas Publ. 7213: 1-102.

LITERATURE CITED CONTINUED

Burla, H., A.B. DaCunha, A.R. Cordeiro, Th. Dohzhansky, C. Malogolowkin, and C. Pavan, (1949), The willistoni group of sibling species of Drosophila, Evol. 3: 300-314.

Carlson, Marion and D. Brutlag, (1979), Different Regions of a Complex Satellite DNA Vary in Size and Sequence of the Repeating Unit, J. Mol. Biol. 135: 483-500.

Carson, H.L., D.E. Hardy, H.T. Speith, and W.S. Stone, (1970), The Evolutionary Biology of the Hawaiin Drosophilidae, Essays in Evolution and Genetics, Appleton-Century Croffs: New York, pp. 437-543.

Charlesworth, Brian, Russell Lande, and Montgomery Slatkin, (1982), A Neo-Darwinian Commentary on Macroevolution, Evol. 36: 474-498.

Corces, V., R. Holmgreen, R. Freund, R. Morimoto, and M. Meselson, (1980), Four heat shock proteins of Drosophila melanogaster coded within a 12-kilobase region in chromosome subdivision 67B, Proc. Natl. Acad. Sci. U.S.A. 77: 5390-5393.

Coyne, Jerry A., (1976), Lack of Genic Similarity Between Two Sibling Species of Drosophila as revealed by varied techniques, Genetics 84: 593-607.

Dobzhansky, Th. and C. Epling, (1944), Contributions to the Genetics, Taxonomy and Ecology of Drosophila pseudoobscura and its Relatives, Carnegie Inst. Wash. Publ. 554: 1-183.

Dobzhansky, Th. and B. Spassky, (1953), Genetics of Natural Populations XXI. Concealed Variability in Two Sympatric Species of Drosophila, Genetics 38: 471-484.

Dobzhansky, Th. and B. Spassky (1954), Genetics of Natural Populations XXII. A Comparison of the Concealed Variability in Drosophila prosaltans With That in Other Species, Genetics 39: 472-487.

Dover, Gabriel, (1982), Molecular drive: a cohesive mode of species evoluation, Nature 299: 111-117.

Dover, Gabriel, Stephan Brown, Enrico Coen, John Dallas, Tom Strachan, and Martin Trick, (1982), The Dynamics of Genome Evolution and Species Differentiation, in, Genome Evolution, eds., G.A. Dover and R.B. Flavell, Academic Press: London, pp. 343-372.

LITERATURE CITED CONTINUED

Eldredge, N. and S.J. Gould (1972), Punctuated equilibria: and alternate to phyletic gradualism, in, Model of Paleobiology, ed., T.J.M. Schopf, Freeman: San Francisco, pp. 82-115.

Finkelstein, D.B., S. Strausberg, and L. McAllister, (1982), Alteration of Transcription during Heat Shock of Saccharomyces cerevisiae, J. Biol. Chem. 257: 8405-8411.

Francis, David and Lois Lin (1980), Heat Shock Response in Cellular Slime Mold, Polysphondyllium pallidum, Devel. Biol. 79: 238-242.

Godson, G.N. and D. Japnek, (1973), A simple method of preparing large amounts of $\phi\chi$ 174 RF1 supercoiled DNA, Biochim, Biophys. Acta 229: 516.

Gojobori, T., (1982), Means on Variances of Heterozygosity and Protein Function, in, Molecular Evolution, Protein Polymorphism and The Neutral Theory, ed., M. Kimura, Japan Scientific Press: Tokyo / Springer Verlag: Berlin, pp. 137-148.

Goldschmidt, Richard, (1940), The Material Basis of Evolution, Yale University Press: New Haven.

Guttman, Susan D., C.V. Glover, C.D. Allis, and M.A. Gorovsky, (1980), Heat shock, deciliation, and release from anoxia induce the synthesis of same set of polypeptides in staved T. pyriformis, Cell 22: 299-307.

Gynne, D., and B.P. Brandhorst, (1982), Alterations in gene expression during heat shock of Achyla ambisexualis, J. Bacter. 149: 488-493.

Hammond, G.L., Y-K. Lai, and C.L. Markert, (1982), Diverse forms of stress lead to new patterns of gene expression through a common and essential metabolic pathway, Proc. Natl. Acad. Sci. U.S.A. 79: 3485-3488.

Hauser, Loren, and B. Levy-Wilson, (1981), Kinetic changes in protein synthesis to a sublethal heat shock in starved T. thermophila, J. Biol. Chem. 256: 3612-2614.

Holmgreen, R., K. Kivak, R. Morimoto, R. Freund, and M. Meselson, (1979), Studies of Cloned Sequences from Four Drosophila Heat Shock Loci, Cell 18: 1359-1370.

LITERATURE CITED CONTINUED

Hubby, J.L. and Lynn H. Throckmorton, (1968), Protein Differences in Drosophila, IV. A study of sibling species, Am. Nat. 102: 193-205.

Ingolia, T.D. and E.A. Craig (1982), Four small Drosophila heat shock proteins are related to each other and to mammalian α-crystallin, Proc. Natl. Acad. Sci. U.S.A. 79: 2360-2364.

Ish-Horowicz, D., S. Pinchin, P. Schedl, S. Artavanis-Tsakonas, and M.-E. Mirault, (1979), Genetic and Molecular Analysis of the 87A7 and 87C1 Heat-Inductable Loci of D. melanogaster, Cell 18: 1351-1358.

Jeffreys, A.J., (1979), DNA sequence variants in the Gγ, Aγ, δ and β-Globin Genes of Man, Cell 18: 1-10.

Kelley, P.M. and M.J. Schlesinger, (1978), The effect of Amino Acid Analogues and Heat Shock on Gene Expression in Chicken Embryo Fibroblasts, Cell 15: 1277-1286.

Key, J.L., C.Y. Lin, and Y.M. Chen, (1981), Heat shock proteins of higher plants, Proc. Natl. Acad. Sci. U.S.A. 78: 3526-3530.

Lakovaara, S., A. Saura, and C. Falk, (1972), Genetic distance and evolutionary relationships in the Drosophila obscura group, Evol. 26: 177-184.

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. and M. Ashburner, (1976), Relationships within the melanogaster species group of the genus Drosophila (Sophophora) II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences, Proc. R. Soc. London B. 193: 275-294.

Lewontin, R.C., (1974), The Genetic Basis of Evolutionary Change, Columbia University Press: New York, N.Y.

Lewontin, R.C. and J.L. Hubby, (1966), A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura, Genetics 54: 595-606.

LITERATURE CITED CONTINUED

Li, Wen-Hsiung, (1981), A simulation study on Nei and Li's Model of Estimating DNA Divergence from Restriction Maps, J. Mol. Evol. 17: 251-255.

Lindquist, S., B. Di Domenico, G. Bugaisky, S. Kurtz, and L. Petko, (1982), Regulation of Heat Shock Response in Drosophila and Yeast, (Cold Spring Harb. Symp. Quant. Biol. in press).

Maniatis, T., E.T. Fritsch, and J. Sambrook, (1982), Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory: Cold Spring Harbour, N.Y.

Mayr, Ernst, (1966), Animal spcies and Evolution, Harvard University Press: Cambridge.

Mayr, Ernst, (1982), Speciation and Macroevolution, Evol. 36: 1119-1132.

Miklos, G.L.G. and A.C. Gills, (1981), The DNA sequence of Cloned Complex Satellite DNAs from Hawaiian Drosophila and Their Bearing on Satellite DNA Sequence Conservation, Chromosoma 82: 409-427.

Mirault, M.-E., M. Goldsmidt-Clermont, L. Moran, A.P. Arrigo and A. Tissieres, (1979), The effect of Heat Shock on Gene Expression in Drosophila melanogaster, Cold Spring Harb. Symp. Quant. Biol., 42: 819-827.

Miyata, T. (1982), Evolutionary Changes and Functional Constraints in DNA Sequences, in, Molecular Evolution, Protein Polymorphism and the Neutral Theory, ed., Motoo Kimura, Japan Scientific Press: Tokyo/Springer Verlag: Berlin, pp. 233-266.

Mullins, J.I., and M. Blumenfeld, (1979), Satellite 1c: a Possible Link between Satellite DNAs of D. virilis and D. melanogaster, Cell 17: 615-651.

Nei, M., (1976), Mathematic Models of Speciation and Genetic Distance, in, Proceedings of the International Conference on Population Genetics and Ecology, Israel 1975, Academic Press: New York, pp. 723-765.

Nei, M. and W.-H. Li, (1979), Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc. Natl. Acad. Sc. U.S.A., 76: 5269-5273.

Nevo, Evitar, (1978), Genetic Variation in Natural Populations: Pattern and Theory, Theor. Pop. Biol. 13: 121-177.

LITERATURE CITED CONTINUED

Patterson , J.T. and W.S. Stone, (1952), Evolution in the Genus Drosophila, MacMillan Company: New York.

Rensch, Bernhard, (1954), Evolution Above the Species Level, Columbia University Press: New York.

Ritossa, F., (1962), A new puffing pattern induced by heat shock and DNP in Drosophila, Experientia 18: 571-573.

Rockwood, E.S., C.G. Kanapi, M.R. Wheeler and W.S. Stone, (1971), X. Allozyme Changes During Evolution of Hawaiian Drosophila, in, Studies in Genetics VI, Univ. Texas Publ. 7103: 193-212.

Scharf, K.-D. and L. Nover, (1982), Heat Shock Induced Alterations of Ribosomal Protein Phosphorylation in Plant Cell Cultures, Cell 30: 427-437.

Schneider, Imogene, (1972), Cell lines derived from late embryonic stages of DROSOPHILA MELANOGASTER, J. Embryol. Exp. Morph. 27: 353-365.

Sin, Y.T., (1975), Induction of puffs in Drosophila salivary gland cells by mitochondrial factors, Nature 258: 159-160.

Sinibaldi, R.M. and R.V. Storti, (1982), One and Two-Dimensional Polyacrylamide Gel Analysis of the Heat Shock Proteins of the Heat Shock Proteins of the Virilis group of Drosophila, Biochem. Genet. 20: 791-807.

Singh, R.S., (1979), Genic Heterozygosity Within Electrophoretic "Alleles" and the Pattern of Variation Among Loci in Drosophila pseudoobscura, Genetics 93: 997-1018.

Singh, R.S., (1982), Genetic Differentiation for Allozymes and Fitness Characters Between Mainland and Bogota Populations of Drosophila pseudoobscura, Can. J. Genet. Cytol. (in press).

Singh, R.S., and M. Coulthart, (1982), Genic variation in abundant soluable proteins of Drosophila melanogaster and Drosophila pseudoobscura, Genetics 102: 437-453.

Singh, R.S., D. Hickey, and J. David, (1982), Genetic Differentiation Between Geographically Distant Populations of Drosophila melanogaster, Genetics 101: 235-256.

LITERATURE CITED CONTINUED

Slater, A., A.B. Cato, G.M. Shlar, J. Kioussis and R.H. Burdon, (1981), Pattern of protein synthesis induced by heat shock in HeLa cells, Biochem. 117: 341-346.

Sondermuller, P.J.A. and N.H. Lubsden, (1978), Heat Shock Peptides in Drosophila hydei and their synthesis in vitro, Eur. J. Biochem. 88: 331-339.

Southern, E.M., (1980), Agarose Gel Electrophoresis, in, Methods in Enzymology, Vo. 68; Recombinant DNA, ed., R. Wu, Academic Press: New York, N.Y.

Spencer, W.P., (1947), Mutations in wild populations of Drosophila, Advan. Genet. 1: 359-402.

Tartoff, K.D., (1979), Evolution of Transcibed and Spacer Sequences in Ribosomal RNA Genes in Drosophila, Cell 17: 607-614.

Throckmorton, L.H., (1975), The phylogeny, ecology, and geography of Drosophila, in, Handbook of Genetics, Vol. 3, ed., R.C. King, Plenum: New York, pp. 421-469.

Throckmorton, L.H., (1977), Drosophila systematics and biochemical evolution, Ann. Rev. Ecol. Syst. 8: 235-254.

Throckmorton, L.H., (1983), Evolution and Speciation in Selected Species Groups, in, The Genetics and Biology of Drosophila, eds., Ashburner, M., H.L. Carson, and J.N. Thompson, Jr., Academic Press: New York, (in press).

Wright, Sewall, (1982), Character Change, Speciation, and the Higher Taxa, Evol. 36: 427-443.

Zouros, E., (1973), Genic Differentiation Associated with the early stages of Speciation in the Mulleri Subgroup of Drosophila, Evol. 27: 601-621.

APPENDIX A

D22 MEDIUM

Solution 1

73.5 g	Glutamic acid
37.4 g	Glycine

Adjust to pH 7.0 with 10 N KOH and make up to 1 l.

Solution 2

73.5 g	Glutamic acid
37.4 g	Glycine

Adjust to pH 7.0 with 10 N NaOH and make up to 1 l.

Solution 3A

10 g	$MgCl_2 \cdot 6H_2O$
37 g	$MgSO_4 \cdot 7H_2O$
4.16 g	$NaH_2PO_4 \cdot 1H_2O$

Make up to 350 ml with distilled water.

Solution 3B

30 g	Yeast Extract

Make up to 100 mls with water.

Solutions 1, 2, 3A, 3B and 4 should be autoclaved.

Solution 5: Grace's Vitamin Solution

10 mg	Thiamine.HCl
10 mg	Riboflavin
10 mg	Ca.Pantothenate
100 mg	p-aminobenzoic acid
10 mg	Folic acid
100 mg	Nicotinic acid
100 mg	Inositol

Solution 5 Continued

50 mg	Biotin
1000 mg	Choline choride

Make up to 1 l with distilled water.

Mixing of D22

1) Mix 40 ml 3A, 355 ml distilled water and add 5 ml of 3B.
2) Add 54 ml of 1, 94 ml of 2, 5 ml of 4 and 2 ml of Grace's Vitamin Solution.
3) Add the following dry ingredients:

Malic acid	670 mg
Succinic acid	60 mg
Sodium Acetate	15.1 mg
Glucose	2 mg

4) In 200 mls hot water dissolve 15 g of lactalbumin hydrosylate and after it has cooled add to mixture.
5) Add 0.1 g streptomycin and 0.15 g penecillin.
6) Adjust to pH 6.7 with 1 N KOH and bring up to 1100 mls with distilled water.
7) Filter sterilize and distribute to sterile bottles.

Carpenter's Medium

Solution A

900 ml	water
50 g	yeast
100 g	sugar
15 g	agar
1 g	KH_2PO_4

Solution B

200 ml	water
0.5 g	$CaCl_2$
0.5 g	$FeCl_2$
0.5 g	$MnCl_2$
0.5 g	NaCl
8.0 g	Sodium Potassium Tartrate

Solution A is autoclaved for 30 minutes at 250°F and immediately upon removal from autoclave Solution B is added. When the mixture has cooled to between 42°C and 45°C add 5.5 ml of propionic acid and distribute to bottles.

Banana Medium

Solution A

1200 ml	water
13.3 g	agar

Solution B

1 tablespoon	malt syrup
1.5 tablespoons	corn syrup
1 large	banana
40 g	yeast
27 ml	water
27 ml	ethanol

Solution A is brought to boil and cooked for 10 minutes. Solution B is thoroughly mixed in a blender. The two are mixed and cooked for ten minutes. When the mixture has cooled to 40 to 45^{o}C then 24 ml of 10% (w/v) methyl-p-hydroxy benzoate in ethanol is added. The mixture is distributed to bottles.

APPENDIX B

To calculate the value for δ for the experimentally derived value of F a computer simulation was run. Nei and Li (1979) have shown that

$$\hat{F} \triangleq P^4/(3-2P)$$

where $P = e^{-r\lambda t}$. δ is defined as $2\lambda t$. From this then

$$P = e^{-3\delta}$$

as r equals 6 in all the restriction enzymes used in this study. Placing $P = e^{-3\delta}$ in the original equation results in

$$\hat{F} \triangleq (e^{-3\delta})^4/(3-2e^{-3\delta}).$$

This relationship was used to find F for all δ beween 0 and 1 by 0.01 increments.

The value of π was estimated by calculation of π where

$$\pi = \Sigma x_i x_j \pi_{ij}$$

As I weighted all lines equally then x_i and x_j were equal to 1/N for all i and j. The value of π_{ij} was the pairwise sequence divergence measure δ_{ij}.