A SURVEY OF HEAT SHOCK PROTEIN LENGTH VARIATION WITHIN AND BETWEEN SPECIES OF DROSOPHILA

A SURVEY OF HEAT SHOCK PROTEIN LENGTH VARIATION WITHIN AND BETWEEN SPECIES OF DROSPHILA

BY

WILLIAM G. McMILLAN B.A., B.Sc.

A Thesis
Submitted to the School of
Graduate Studies in Partial
Fulfillment of the Requirements
for the Degree
Master of Science

McMaster University
September, 1987

Master of Science (1987) McMaster University
Hamilton, Ontario

Title: A Survey of Heat Shock Protein Length Variation Within and Between Species of Drosophila.

Author: William G. McMillan, B.A. B.Sc. (McMaster University)

Supervisor: Professor Rama S. Singh

Number of Pages: V, 96.

ABSTRACT

Forty iso-females lines of *Drosophila melanogaster* were surveyed by SDS-PAGE for length polymorphisms in their heat shock proteins. None was observed. The 70K and 27K HSP's were further analyzed by peptide mapping using *Staphylococcus aureas* V8 protease, acid hydrolysis and chymotrypsin. Again, no variation was observed.

Thirty-seven species of *Drosophila* were surveyed for length variation in their heat shock proteins by SDS-PAGE. No variation was observed in HSP84 or HSP70 among any of the species. HSP36 was shown to vary in mobility only, especially among species from different species groups. The small HSP's showed the greatest interspecies variation, especially in the *Repleta* and *Virilis* groups. This variation included mobility and band number differences.

The results from the 36K HSP's were used to construct a *Drosophila* species group phylogeny. These groups, shown in descending order of age, are *Saltans*, *Victoria*, *Melanogaster*, *Willistoni*, *Repleta*, and *Virilis*.

ACKNOWLEDGEMENTS

After a very long time, I am pleased to at last be able to acknowledge my friends in the McMaster Biology Department. Particular thanks go to Dr. Rama S. Singh, my supervisor, the most patient man I know. I am also grateful to Dr. S.F.H. Threlkeld and Dr. R.A. Morton for reading this thesis on such short notice , and to Dr. Morton for his many attempts to impart to me some of the finer points of baseball.

To my friends, Dr. M.C. Coulthart, Madhu Choudhary, Larry Hale, John Hudson and Shanta Thomas, I enjoyed our time together and will miss you. Thanks, Mike, for all your help with my project.

Finally, and most important, to my wife Judy who has put up with being married to a student for ten years with no end in sight, I dedicate this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Before the advent of the techniques of molecular biology, population geneticists were limited in their ability to study genetic variation. Studies of visible mutants and lethal genes in natural populations and inbreeding and selection experiments in laboratory populations were the only available methods and provided little information concerning the overall levels of genetic polymorphism and heterozygosity in natural populations (for review, see Lewontin, 1974).

In 1966, Hubby and Lewontin introduced the technique of gel electrophoresis of proteins to population genetics, allowing protein polymorphisms to be studied in some detail (Hubby and Lewontin, 1966). A wide range of studies on many species from the plant, animal and bacterial kingdoms have shown that an average species contains polymorphisms at about one third of its structural gene loci, with individuals being heterozygous at from five to twenty per cent of their loci (Ayala, 1982).

One problem with the early, one dimensional gel electrophoresis was that at many of the loci which appeared polymorphic, a lot of cryptic variation was being missed due to insensitivity of the technique.

The "charge-ladder" model of Ohta and Kimura (1974) suggested that electrophoresis distinguished only those amino acid substitutions in which the substituted amino acids differed in electrical charge, accounting for less than thirty per cent of all substitutions. This situation was partially rectified by the development of techniques using heat denaturation (Singh et al., 1975), biochemical kinetics (Modiano et al., 1979) and sequential gel electrophoresis (Singh et al., 1976; Coyne, 1976; Coyne and Felton, 1977), all of which detected variation which was previously missed by early electrophoresis. Of the three methods, sequential gel electrophoresis was the most useful because it was less time consuming and avoided the cumbersome measurements and biochemical procedures of the other two methods. In the 1976 study of Singh et al., the xanthine dehydrogenase gene product of Drosophila pseudoobscura was subjected to gel electrophoresis during which two different gel concentrations and two different buffer pH's were combined to produce four different sets of electrophoretic conditions. This resulted in a greater than 300% increase in the detection of electromorphs, from eight to twenty-seven. Subsequent work has shown that polymorphic loci are of three types. The xanthine

dehydrogenase study showed that at some loci, there exist one common allele and several rare alleles, while an alcohol dehydrogenase study showed that it is possible to have a polymorphism consisting of two major alleles only (Kreitman, 1980). The third type, as seen with esterase-5 in D. pseudoobscura, seems to be a combination of types one and two, with a major, two allele polymorphism, and many rare alleles (Keith,1983).

Although sequential gel electrophoresis drastically increased the number of alleles which could be detected, there was still no way of knowing whether most of the hidden variation had been discovered. In 1979, Ramshaw et al., using thirty-two variants of human haemoglobin, all of known amino acid sequence, were able to detect eighty-five per cent of the variants using sequential gel electrophoresis, illustrating the power of the technique. Since that time, the refinement of peptide mapping (Fey, 1983) and two dimensional gel electrophoresis (O'Farrell, 1975, O'Farrell et al., 1977) have further increased the scope of the study of genetic polymorphisms.

Most electrophoretic studies of protein variability have employed nondenaturing conditions in which the net charge of the protein determines its mobility. Under such conditions, abundant Drosophila

proteins such as those of the salivary glands or hemolymph have shown lots of variation (Korge, 1977; Roberts and Evans-Roberts, 1979), and it has been shown that, especially among invertebrates, enzyme polymorphism is also abundant (Ayala,1984).

Considerably fewer electrophoretic studies have employed denaturing conditions. The dogma of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) says that this method separates proteins on the basis of size, rather than charge differences (Weber and Osborn, 1969). Early SDS studies showed low variability among proteins (Postlethwait and Jowett, 1980), and this fact, along with the fact that most variation has been determined under non-denaturing conditions, which can not distinguish between size differences and amino acid substitutions, has led to the assumption by most population geneticists that protein variability is the result of amino acid substitutions, rather than size differences. This leads to the conclusion that the majority of genic variation arises by changes in sequence, rather than by deletions or insertions.

Studies at the level of DNA have also considered this question. In 1983, Kreitman compared the DNA sequences of eleven alcohol dehydrogenase (Adh) genes from five populations of D. melanogaster.

He found forty-three polymorphisms in introns and exons, all but one of which were silent. He also found that the adult intron and 3' flanking regions contained six length polymorphisms. In another study, Aquadro et al. (1986) examined forty-eight lines of D. melanogaster and found that eighty per cent of them contained at least one insertion or deletion when compared to the consensus restriction map, but that none of these ocurred in the Adh coding region. However, one insertion was correlated with an unusual pattern and level of Adh expression, suggesting that length polymorphisms could be important in the evolution of the locus.

As was previously mentioned , studies of the size variation in proteins have been very few. However, Whalen and Wilson (1986) have shown that considerable size variation exists among the male accessory gland proteins in different lines of D. melanogaster. It is also known that a ubiquitous group of proteins known as the heat shock proteins show varying degrees of length polymorphism among Drosophila species (Sinibaldi and Storti, 1982). The heat shock proteins, because they are universal and because a wealth of heat shock protein and DNA sequence information is available, are excellent candidates for the study of protein variation and the

evolution of protein size.

THE HEAT SHOCK RESPONSE: A SHORT REVIEW

The heat shock proteins (HSP's) are a universal family of polypeptides of unknown function. Since the discovery, more than twenty-five years ago, of chromosome puffs after the heat stressing of Drosophila salivary glands (Ritossa, 1962), and the linking of these heat shock puffs, twelve years later, to the induction of a new set of proteins coupled with the repression of normal protein synthesis (Tissieres, 1974), the HSP's have been extensively studied. Many of the heat shock genes in Drosophila and other organisms have been cloned and sequenced, and they have proven to be highly conserved over great phylogenetic distances. Organisms as divergent as Escherichia coli and man have shown close to fifty per cent homology among some of their heat shock proteins (Bardwell and Craig, 1984). Their widespread presence and their sequence conservation would suggest that the function(s) of the HSP's, although presently unknown, are of major importance in the survival of the organisms.

Although the majority of heat shock studies have been carried out on Drosophila, the induction of

this group of proteins has been noted in all organisms so far surveyed in the plant, animal, protist and fungal kingdoms (Baszczynski et al., 1982; Ashburner and Bonner, 1979; Hauser and Levy-Wilson, 1981; Loomis and Wheeler, 1980). While heat induction of these proteins is the experimental method of choice, such a wide variety of agents are capable of causing the same response (Ashburner and Bonner, 1979; Tanguay, 1983; Lanks, 1986) that a more appropriate term for the HSP's would be "stress proteins". The list of HSP inducers presently includes developmental hormones, uncouplers of oxidative phosphorylation, inhibitors of electron transport, hydrogen acceptors and inhibitors of many enzymes (Tanguay, 1983). It is not known whether these agents act to induce the production of HSP's by a common pathway, or whether induction is accomplished by independent or convergent pathways. However, induction experiments using heat shock, cadmium or adenovirus E1a protein exposure, or the addition of serum to serum-starved cells have shown that the sequences responsible for cadmium and heat shock induction map in different regions from the sequences responsible for serum and E1a induction, (Wu, 1985) suggesting that a common pathway is probably not employed.

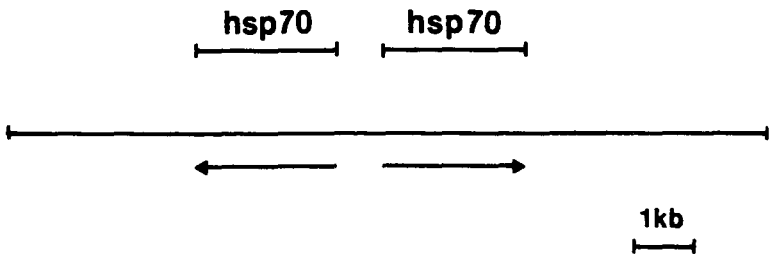
The numbers and molecular weights of the

HSP's vary among different organisms, as shown by SDS-PAGE. A number of other proteins, synthesized at normal temperatures and migrating with HSP's are often mistaken for HSP's. Many of these proteins are related to the HSP's and are called heat shock cognates (Kurtz et al., 1986; Lowe and Moran, 1984, Craig et al., 1983; Ingolia and Craig, 1982). While these proteins are not induced by heat shock, their genes must be considered to belong to heat shock multi-gene families by virtue of DNA sequence homologies of up to seventy-two per cent and amino acid sequence homologies of up to seventy-four per cent (Ingolia and Craig, 1982).

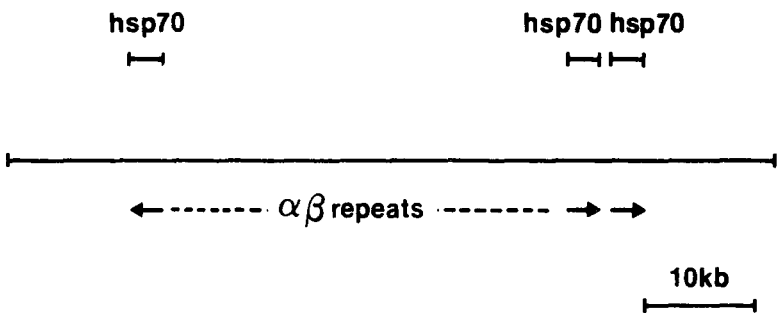
In Drosophila, the most prominent group of HSP's belong to the 68-70 kilodalton (K) family. HSP70 is coded for by cytological loci 87A (two copies per haploid genome) and 87C (three copies per haploid genome) on chromosome 3 (Figure 1) in most Drosophila species, (Craig and Ingolia, 1982) while HSP68 is coded for by locus 95D on chromosome 3 (Holmgren et al., 1979). The remaining members of this multi-gene family are heat shock cognates (HSC's)

In D. melanogaster and D. pseudoobscura, the hsp70 loci are located about five hundred kilobases apart and in corresponding chromosomal locations (Pierce and Lucchesi, 1980). Since these two species

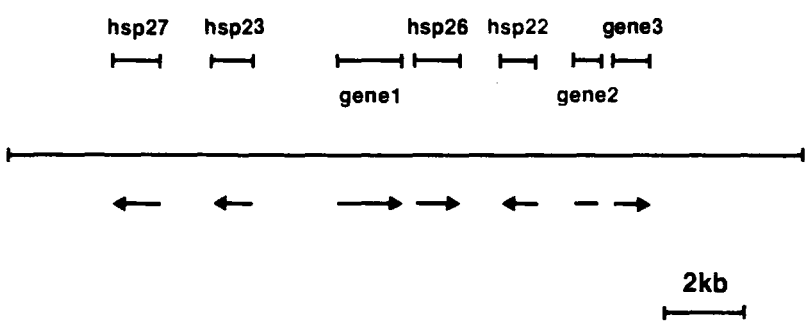
FIGURE 1. The organization of heat shock genes at cytological loci 87A, 87C and 67B. Arrows indicate the orientation of the transcription unit. Where an arrow is omitted, orientation is unknown. (Artavansis-Tsakonas and Schedl, 1979; Ayme and Tissieres, 1985; Sirotkin and Davidson 1982)



Locus 87A



Locus 87C



Locus 67B

are thought to have diverged more than thirty-five million years ago (Throckmorton, 1975), this suggests that the original duplication occurred at least that long ago. The DNA sequences in the five gene copies have been found to be ninety-seven per cent conserved (Craig and Ingolia, 1982). Since it is probable that these genes have a vital biological function, this could explain their interspecies sequence conservation. However, it does not expain why five copies of the gene show such a high degree of homology, especially since mutant flies which contain only one gene copy have proven to be viable and to have a normal response to heat shock (Udvardy et al., 1982). This point will be considered further in the Discussion.

The expression of the various members of the hsp70 gene family varies with the species and the type of induction. While hsp70 and hsp68 in Drosophila are induced by heat shock, other genes are expressed at ambient temperatures (Artavanis-Tsakonas and Schedl, 1979). Transcription of heat shock cognates hsc1, hsc2 and hsc4, located at loci 70C, 87D and 88E of chromosome 3, while not induced by heat shock, is induced at different stages of development (Craig et al., 1983). Also, induction of proteins related to HSP70 in mouse and rat cells has been observed at

normal temperatures (Hughes and August, 1982).

While the functions of the hsp70 gene family have not been proven, a number of studies have provided evidence to support roles in the negative regulation of the heat shock response as well as the aggregation of cellular proteins. In Drosophila, transcription regulation depends upon the accumulation of HSP70. When HSP70 synthesis is stopped, the stability of the other hsp messages is increased and their transcription continues without repression (DiDomenico et al., 1982). Similar results were obtained with Saccharomyces cerevisiae, in which double HSP70 mutant strains showed constitutive synthesis of the other HSP's (Craig and Jacobsen, 1984), and with Escherichia coli, in which mutants of the dna K gene, which codes for the only 70K protein induced by that bacteria, were unable to suppress the transcription of the rest of the HSP'S (Tilly et al., 1983).

A number of experiments have been performed in order to test the effect of the HSP70 family upon protein aggregation. First, one of the two prominent HSP70 family proteins of rat embryo fibroblasts was shown to concentrate heavily in nucleoli, both during and after heat shock (Welch and Feramisco, 1984). Next, a Drosophila hsp70 gene was placed into a

mammalian genome, and its HSP70 protein showed a similar distribution (Lindquist, 1986). Also, a Drosophila hsp70 gene was separated from its usual promoter and placed next to a normal promoter. This hybrid gene was introduced into mouse L cell and monkey COS cell genomes, followed by inhibition of protein synthesis and then heat shock. Nucleolar structure, transiently damaged by such a heat shock, was shown to return to normal much more quickly in these cells than in cells undergoing the same treatment which did not possess the Drosophila gene (Pelham, 1984). Finally, it has been shown that while transcription is hardly affected by heat, mRNA splicing is blocked at the upper temperature range of the heat shock response (Yost and Lindquist, 1986). These results support the suggestion that the HSP70 proteins may catalyze the repair of heat damaged ribonucleoproteins. There have also been suggestions that members of the hsp70 gene family can act as ATPases and proteases (Chappell et al.,1986; Phillips et al., 1984). Together, these results have led to the hypothesis that the 70K HSP's may be able to use energy from the breakdown of ATP to build up and break down cellular protein aggregates.

Another prominent Drosophila HSP is HSP84. It is coded for by locus 63BC on chromosome 3

(Holmgren et al., 1979). Unlike the members of the HSP70 family, this protein, which is soluble, is found only in the cytoplasm (Levinger and Varshavsky, 1981). To date, all eukaryotes studied have been shown to produce a heat shock protein in the 83-90K range, and sequencing studies have suggested that this is the second most highly conserved family of HSP's. The amino acid sequences of Drosophila and yeast proteins have shown sixty per cent homology (Farrelly and Finkelstein, 1984), and antibodies raised against the chicken HSP89 cross react with human, mouse, frog and Drosophila proteins (Schlesinger et al.,1982).

While the function of the hsp84 family is unknown, these proteins have been shown to form a number of complexes. They are part of the avian 8S progesterone receptor (Schuh et al., 1985), are found with a chicken oviduct glucocorticoid receptor (Sanchez et al., 1985), with a membrane ATPase (Burdon and Certmore, 1982) and with a golgi protein in mammals (Welch and Feramisco, 1982). It has been suggested that this group of proteins may act to move other molecules around the cell (Lindquist, 1986). However, since only a small fraction of the total cellular HSP90 protein is found in these complexes (Lindquist, 1986), it is unclear why it is found in such excess.

The small HSP's, the least conserved of all the HSP's, comprising a heterogeneous family of proteins between 20K and 30K in most species, are coded for by locus 67B on chromosome 3 in D. melanogaster (Voellmy et al., 1981). This locus has recently been shown to contain seven small heat shock genes in a fifteen kilobase region (Ayme and Tissieres, 1985). The originally noted four small heat shock genes code for HSP22, HSP23, HSP26 and HSP27, while the three newly discovered genes have been shown to code for heat inducible transcripts, one of which, from gene 1, would be translated into a 26.5K protein (Ayme and Tissieres,1985). The organization of this locus is shown in Figure 1 (Ayme and Tissieres, 1985; Sirotkin and Davidson, 1982).

The pattern of induction of the Drosophila small HSP's shows a striking difference from that of the 70K family. Substantial amounts of mRNA are produced from the third instar larval stage to the middle pupal stage, with the hsp23 transcript showing an amount two to five fold greater than the hsp26 or hsp27 transcripts. The hsp22 transcript is barely detected. Genes 1 and 3 have been shown to be expressed in the same development stages (Ayme and Tissieres, 1985). This induction is due to the presence of the molting hormone, ecdysterone, which

does not induce HSP70 (Ireland et al., 1982). Induction by heat shock during the same developmental stages increased transcription from five to twenty-five fold (Ayme and Tissieres,1985).

As with the other HSP families, the function(s) of the small HSP's have not been proven. However, the small HSP's from most species are homologous with an aggregating protein of the eye, the alpha-crystallin lens protein (Craig et al., 1982), and have been shown to form aggregates of their own (Arrigo et al., 1985). Upon heat shock, the small HSP's migrate to the nucleus, showing up to 80% localization, the highest such percentage of any of the HSP's. They associate with nuclear RNA, forming particles with the same morphology as the prosome ribonucleoprotein particles seen in mouse and duck cells (Arrigo et al., 1985).

Although the specific functions of most of the HSP's are at present speculative, there is little doubt that together, they are very important in conferring thermotolerance. Since the effects of heat vary in different circumstances, it may be that the different HSP's are necessary in order to deal with these circumstances. (Linquist, 1986 review) Whatever their roles may turn out to be, their universality and high degree of conservation make it

certain that the HSP's are necessary for life.

EVOLUTION OF GENES AND PROTEINS

A comprehensive study of the evolution of genes and proteins would require the consideration of a number of parameters. These include:

1. gene organization within the genome. Is the gene duplicated? Are the duplications organized as tandem arrays, located at different loci, or both?
2. species differences in genome organization and nucleotide sequence.
3. the evolution of regulatory sequences, protein coding sequences and introns within the genome and among species.
4. differences between rates of amino acid substitution in different areas of the coding sequences (i.e. functional domains) within the genome and among species.
5. the type of amino acid substitution (i.e. conservative or non-conservative) and its relation to the functional domains.
6. amino acid length polymorphisms within the genome and among the species. Are differences due to deletions or additions? Where do these occur?

In the past ten years, the development of the

techniques of molecular biology has allowed population geneticists to progress beyond the search for an answer to the question of how much protein variation exists in natural populations. DNA sequencing studies have provided the information necessary for making accurate phylogenetic predictions. Mitochondrial sequences, because they are inherited maternally and without segregation, have proven to be particularly useful (Lewontin, 1985). The original mtDNA study showed that three of nineteen restriction sites in D. melanogaster mtDNA were polymorphic (Shah and Langley, 1979), while a subsequent study showed how the mtDNA from different gopher populations varied with location (Avise et al., 1979), allowing the phylogenies to be worked out. The sequencing of protein coding genes has also provided useful phylogenetic data which will be dealt with in the discussion.

DNA sequences have also been used by the participants in the "selectionist-neutralist" controversy. Because it is possible to compare the rates of nucleotide substitution in the various coding and non-coding regions of a gene and its flanking DNA, the roles of selection and drift at a particular locus can be tested. For example, Kreitman (1983) showed that selection has acted at non-silent sites of the Adh locus in D. melonogaster to prevent amino acid

polymorphisms, but that silent sites in the same exons have not been so constrained, and show six per cent polymorphism. Also, much less polymorphism (one per cent) in different non-coding and flanking regions suggests that selection also works upon these sequences.

In the future, population genetics may be expected to change along the lines predicted by Lewontin in 1985.

> "The next task in the study of variation is to put aside polymorphic genes and to study monomorphic ones... Thus, the question of what causes monomorphism is dual to the question of polymorphism. The ability to follow gene phylogenies will answer questions about migration, about effective population sizes and about the repeatability of mutational and cytogenetic events. In the end, population genetics is conceptually the study of gene lineages. Until now, the data to study such lineages have not really been available."

In this thesis, one dimensional SDS-PAGE was used in order to study size differences among three different families of heat shock proteins, the major (70K) HSP's, the 84K HSP's and the small HSP's, in a number of iso-female lines of D. melanogaster and in one strain each of thirty-seven other species of Drosophila. Also, peptide mapping was used in order

to probe for heat shock protein differences in various iso-female lines within a species. Finally, published amino acid length polymorphisms from various protein families among a variety of species were considered and compared to similar data from the heat shock proteins.

MATERIALS & METHODS

1. DROSOPHILA STOCKS

Iso-female lines of *D. melanogaster* from North American, European and African populations were used in the population survey. These lines have been kept in the laboratory for many generations. *Drosophila* species were obtained from the *Drosophila* Species Resource Centre, Bowling Green, Ohio (see Table 1).

BIOCHEMICALS USED

Acrylamide, Tris, SDS, and all electrophoresis purity reagents were obtained from BIO-RAD.N,N'-methylenebisacrylamide (bis) was obtained from Bethesda Research Laboratories, Inc. Ammonium persulphate was obtained from BDH Chemicals. Glycerol was obtained from Caledon Laboratories Ltd. Hydrochloric acid and EDTA were obtained from J.T. Baker Chemical Company. Imadazole, 2-mercaptoethanol, phenylmethylsulfonylfluoride (PMSF) and N,N,N',N'-tetramethylethylenediamine (TEMED) were obtained from Sigma Chemical Company.

Table 1 *Drosophila* species surveyed for HSP's

Group	Species	Stock Number	Collection Site
-	*D. busckii*	13000-0081.23	Blue Mnt. Lake, New York
Obscura	*D. pseudoobscura*	14011-0121.0	Tucson Arizona
Saltans	*D. saltans*	14045-0911.0	San Jose, Costa Rica
	D. prosltans	14045-09010	Turrialba, Costa.Rica
Melanogaster	*D. melanogaster*	-	France
	D. simulans	14021-025515	Lima, Peru
	D. yakuba	14021-0261.0	Ivory Coast, Africa
	D. teissieri	128.6	Mt. Selinda, Rhodesia
	D. takahashii	14022-0311.5	Yun Shui, Taiwan
	D. ananassee	2507.18	Angra Dos Reis, Brazil
Willistoni	*D. willistoni*	14030-0811.2	Royal Palm Park, Florida
	D. nebulosa	2309.3	Key West Florida
	D. paulistorum	-	unknown
	D. equinoxialis	14030-0741.1	Teffe, Brazil
Virilis	*D. texana*	15010-1041.22	New Orleans, Louisiana
	D. novamexicana	15010-1031.12	Antlers, Colorado
	D. americana	15010-0951.1	Anderson, Indiana
	D. lummei	15010-1001.0	Finland
	D. virilis	15010-1051.48	Texmelucan, Puebla, Mexico
	D. littoralis	15010-1001.2	Kuhn Aargau, Switzerland
	D. lacicola	15010-0991.13	Beaver Creek, Manitoba
Repleta	*D. mercatorum*	15082-1521.15	Cochabamba, Bolivia
	D. paranaensis	15082-1541.4	Campo Grande, Brazil
	D. mojavensis	15081-1351.0	Esperanza, Sonora, Mexico
	D. arizonensis	-	unknown
	D. mulleri	15081-1371.0	Lake Travis, Texas
	D. peninsularis	15081-1401.0	Rio Piedras, Puerto Rico
	D. fulvimaculoides	15084-1581.0	Turriabla, Costa Rica
	D. hydei	15085-1641.4	Sao Paulo, Brazil
	D. neohydei	15085-1651.0	Carpentaro, Venezuela
	D. eohydei	15085-1651.0	Santa Marta Mts., Columbia
Melanica	*D. melanica*	15030-1141.1	Cliff, New Mexico
Polychaeta	*D. polychaeta*	15100-1711.0	Leticia, Columbia
Victoria	*D. easteeli*	11010-0011.0	Veyo, Utha
	D. lebanonensis	11010-0021.0	Beirut, Lebanon
	D. pattersoni	11010-0031.0	Beirut, Lebanon
	D. victoria	-	unknown

3. RADIOCHEMICALS USED

^{35}S methionine (850-1410 Ci/mmole) was obtained from Amersham/ Searle.

4. ENZYMES USED

Staphylococcus aureas V8 protease and chymotrypsin were obtained from the Sigma Chemical Company.

5. SOLUTIONS

5.1 SAMPLE PREPARATION

5.1.1 GRACE'S INSECT MEDIUM

This insect cell culture medium was made by Dr. D. Morgan as described previously. (Grace, T.D.C., 1962)

5.1.2 EXTRACTION BUFFER

5.1.2.1 SOLUTION B

Tris	5.5g
1N Hcl	48 ml
2M Tris	to pH 6.8
dd H2O	to 100 ml

5.5g Tris was dissolved in 48 ml 1N Hcl. 2M Tris was added to pH 6.8. ddH_2O was added to 100 ml. This was known as solution B.

5.1.2.2. BUFFER STOCK WITHOUT SDS

Solution B	12.5 ml
ddH_2O	50 ml
PMSF	0.0174g
2-mercaptoethanol	1 ml
glycerol	10 ml
1M Tris	to pH 6.8
ddH_2O	to 100 ml

0.017g PMSF was dissolved in a mixture of 12.5 ml Solution B, 50 ml ddH2O, 10 ml glycerol and 1 ml 2-mercaptoethanol. 1M Tris was added to pH 6.8 and ddH_2O added to 100 ml. This solution was known as Extraction Buffer Stock and was kept in the refridgerator.

5.1.2.3. EXTRACTION BUFFER ALIQUOTS

SDS	0.01g
Extraction Buffer Stock	1 ml

0.01g SDS was dissolved in 1 ml aliquots of Extraction Buffer Stock as needed. Aliquots were stored at -20°C for up to two months.

5.2 GEL ELECTROPHORESIS (modified from Laemmli, 1970)

5.2.1. SDS STOCK (10%)

SDS 10g

ddH_2O to 100 ml

10g SDS WAS dissolved in ddH_2O and the solution brought to 100 ml. This solution was known as SDS Stock.

5.2.2. RESERVOIR BUFFER (0.1% SDS)

Tris 6g

glycine 28.8g

ddH_2O to 990 ml

SDS Stock (10%) 10 ml

6g Tris and 28.8g glycine were dissolved in 990 ml ddH_2O. 10 ml SDS stock was added.

5.2.3. RESOLUTION GEL BUFFER (X4)

Tris	18.3g
IN Hcl	20 ml
ddH_2O	to 80 ml
IN Hcl	to pH 8.8
ddH_2O	to 100 ml

18.3g Tris was dissolved in 20 ml 1N Hcl and ddH_2O was added to 80 ml. 1N Hcl was added dropwise to pH 8.8, and ddH_2O was added to 100 ml.

5.2.4. STACKING GEL BUFFER (X4)

Tris	5.5g
IN Hcl	48 ml
ddH_2O	to 80 ml
2M Tris	to pH 6.8
ddH_2O	to 100 ml

5.5g Tris was dissolved in 48 ml IN Hcl and ddH_2O was added to 80 ml. 2M Tris was added dropwise to pH 6.8 and ddH_2O was added to 100 ml.

5.2.5. ACRYLAMIDE/BIS STOCK (30%)

acrylamide	15g
bis	0.4g
ddH_2O	to 50 ml

15g acrylamide and 0.4g bis were dissolved in ddH_2O and the solution brought to 50 ml.

5.3 PEPTIDE MAPPING (modified from Cleveland et al., 1976)

5.3.1 RESOLUTION GEL BUFFER (X4)

This is the same as previously described in section 5.2.3.

5.3.2 RESERVOIR BUFFER

This is the same as previously described in section 5.2.2.

5.3.3 ACRYLAMIDE/BIS STOCK (30%)

This is the same as previously described in section 5.2.5.

5.3.4 PEPTIDE MAPPING STACKING GEL BUFFER (X4)

Tris	3.04g
Imidazole	1.72g
IN Hcl	to pH 6.8
ddH_2O	to 100 ml

Tris and imidazole were dissolved in 1N Hcl to pH 6.8., and ddH_2O was added to 100 ml.

5.3.5. SAMPLE BUFFER

Tris	0.81g
Imidazole	0.46g
1N Hcl	to pH 6.8
glycerol	10 ml
EDTA	0.123g
2-mercaptoethanol	3.3ml
SDS	0.5g
ddH_2O	to 100 ml

Tris and imidazole were dissolved in 1N Hcl to pH 6.8. Glycerol, EDTA, SDS and 2-mercaptoethanol were added with stirring. ddH_2O was added to 100 ml.

5.3.6. SAMPLE BUFFER (20% GLYCEROL)

sample buffer	0.9 ml
glycerol	0.11 ml

0.11 ml glycerol was added to 0.9 ml sample buffer.

5.3.7. ACID HYDROLYSIS SOLUTION

ddH_2O	91 ml
methanol	91 ml
glacial acetic acid	18 ml

ddH_2O, methanol and acetic acid were mixed immediately before use.

5.3.8. ENZYME SOLUTION

Staphylococcus aureas V8 protease and chymotrypsin were used at a concentration of 30 ug/ml sample buffer.

6. SAMPLE PREPARATION

Late third instar larvae from each isofemale line were heat shocked at 37°C for forty-five minutes. For each

sample, seven of the best larvae were selected and the salivary glands dissected out in Grace's medium.The dissected glands (with attached fat bodies and some mouth parts) were heat shocked in Grace's medium in an incubator for forty-five minutes at 37°C. They were then transferred to a 4 ul drop of ^{35}S-methionine (20 uCi) and Grace's medium and heat shocked for another thirty minutes at 37°C. It should be noted that the temperature of the incubator varied by ± 1°C, and this may have caused differences in HSP induction. (see RESULTS) Heat shocking in a water bath could avoid this problem. The glands were then rinsed three times with Grace's medium using a 5 ul micropipette and placed into 30 ul of extraction buffer (1% SDS). Each sample was immediately placed into a boiling water bath for two minutes and then frozen in liquid nitrogen and stored at -72°C for up to two months.

7. SDS POLYACRYLAMIDE GEL ELECTROPHRESIS

Electrophoresis in SDS gels was carried out in a slab gel apparatus. The procedure is a modification of the system described by Laemmli (1970). The gel box was assembled and sealed with 1% agar. A 30 ml (9%,13% or 15% acrylamide) resolution

gel solution was made as follows: 9 ml (or 13 ml or 15 ml) acrylamide/bis stock, 7.5 ml resolution gel buffer and 13.1 ml.(or 9.1 ml or 7.1 ml) ddH_2O were mixed and degassed for thirty minutes. After degassing, 0.3 ml 10% SDS was added. To polymerize the gel, 16 ul TEMED and 100 ul 10% ammonium persulphate were used. After pouring, the solution was overlaid with 0.1% SDS and allowed to polymerize for sixty minutes. It should be noted that 9% gels had a thickness of 1.5 mm while 13% gels were 0.75 mm thick.

A 10 ml (6% acrylamide) stacking gel solution was made as follows:

2 ml acrylamide/bis stock, 2.5 ml stacking gel buffer and 5.4 ml ddH_2O were mixed and degassed until polymerization of the resolution gel was complete. 0.1 ml 10% SDS was added, and to polymerize the gel, 7 ul TEMED and 40 ul 10% ammonium persulphate were used. The spacer comb was then inserted and the gel was allowed to polymerize for sixty minutes.

Following polymerization, reservoir buffer was added to the upper and lower chambers. To produce a dye front in the gel, 1 drop of phenol red indicator was added to the upper reservoir. Samples were thawed, placed in a boiling water bath for two minutes and spun in the microcentrifuge for three minutes before being placed in the gel slots. Electrophoresis

was carried out at 10 mA and from 20-30 V overnight

8. PREPARATION OF BANDS FOR PEPTIDE MAPPING

The procedure used for band preparation was a modification of that described by Cleveland et al. (1976). The method of band preparation was dependent upon the protein cleavage treatment to be used. For enzyme cleavage, the gel was placed upon filter paper (Whatman #1, medium fast) and dried immediately. For acid hydrolysis, the gel was placed in 200 ml of an acid hydrolysis solution and 0.1% Coomassie Blue stain was added until the solution was dark blue. Staining was continued for six hours, followed by destaining overnight. The gel was then rinsed with ddH_2O, placed upon filter paper and dried.

Autoradiography was carried out using Kodak XAR5 film exposed for twenty-four hours. After developing the film, an acetate sheet was placed over the film and the positions of the 70K, 27K and 22K HSP bands were marked on the sheet. This sheet was then placed over the gel and stapled to it in order to identify the positions of the desired bands. Bands were then excised and placed in 2 ml of sample buffer for thirty minutes. This procedure was repeated and then the sample buffer was discarded. The filter

paper was then carefully removed from each band and the band was stored at -20°C for up to one month.

The preparation of the resolution gel solution was the same as previously described for the first dimension 9% gel. However, the solution was poured to a depth which would allow a stacking gel of about 3-4 cm.

A 10 ml (3% acrylamide) stacking gel solution was made as follows:

1 ml acrylamide/bis stock, 2.5 ml Peptide Mapping Stacking Gel Buffer and 6.4 ml ddH_2O were mixed and degassed until the resolution gel had polymerized. Polymerization was carried out by the additon of 7 ul TEMED and 40 ul 10% ammonium persulphate. The solution was then poured and the spacer comb inserted. Polymerization was complete within one hour.

The previously prepared protein bands were thawed and placed lengthwise into the slots. The best results were achieved with a tight fit.

For enzyme cleavage, the following solutions were added to each slot:

10 ul sample buffer containing 20% glycerol

10 ul sample buffer

Reservoir buffer was then added to the upper and lower reservoirs. 1 drop of phenol red indicator was added to the upper reservoir and electrophoresis was carried out at 8 mA and 20 V. When the dye front had moved to within 0.5 cm of the end of the stacking gel, the current was turned off for 30 minutes to enhance the stacking of the proteins. Electrophoresis was then continued overnight.

After completion of electrophoresis, the gels were placed on filter paper and dried. Autoradiography was carried out using Kodak XAR5 film. Films were exposed for up to fourteen days.

RESULTS

Variation in HSP's Within D. Melanogaster

While a few studies have considered the variation in HSP's between different species of Drosophila (Sinibaldi and Storti, 1982), and others have considered the variation among the different HSP's within a species (Buzin and Petersen, 1982), a systematic study of each of the HSP's for variation among different lines within a species has not been made. Here, I present a comparison of the 84K, 70K and small (including 36K) heat shock proteins from iso-female lines of D. melanogaster.

SDS-PAGE ANALYSIS

In order to look for intraspecies HSP variation in D. melanogaster, HSP's from individual strains were examined by SDS-PAGE electrophoresis. Forty iso-female lines from four populations of D. melanogaster were surveyed. The laboratory strain, Canton S, was used as a standard on all gels. The four populations were from Villeurbane, France;

Benin, West Africa; Brownsville, Texas and Hamilton, Ontario. In all cases, the mobility of each of the heat shock proteins (84K, 70K, 36K and 22-27K) was unchanged among all of the lines. Therefore, if variation existed, it was not uncovered by SDS-PAGE. However, since in SDS-PAGE the separation of proteins is based on the size, and since this method is not sensitive to charge differences among different allelic forms of a protein, a modified gel electrophoresis, substituting the zwitterionic detergent CHAPS for SDS, was attempted. Theoretically, CHAPS should be able to solubilize proteins while maintaining their electrical charge over a range of pH values. It should also be nondenaturing, preserving the secondary and tertiary structure of the protein, as well as its enzymatic activity, if applicable (Hjelmeland et al., 1983). Unfortunately, for reasons which were not understood, this system was unable to solubilize the HSP's, and the CHAPS gels were unsuccessful.

PEPTIDE MAPPING

A further comparison of the SDS-PAGE separated HSP's was made using the method of peptide

mapping. Peptide mapping was introduced by Ingram (1956, 1959) and provides a dimension beyond SDS-PAGE for the separation of different polypeptides. It has previously been shown that detection of hidden variation in D. melanogaster alcohol dehydrogenase may be accomplished by peptide mapping (Ayala, 1982). The method of Cleveland et al. (1976), modified by Fey (1984), which allows for the mapping of polypeptides that have previously been subjected to SDS-PAGE, was used. By this method, proteins which have been cut from SDS gels are analyzed by partial digestion using chemical means or various proteases in an SDS buffer, followed by SDS-PAGE. In this study, the enzymes chymotrypsin and Staphylococcus aureas V8 protease were used, as was the method of acid hydrolysis. V8 protease has been shown to cleave peptides on the carboxy terminal side of glutamic acid and aspartic acid residues (Houmard and Drapeau, 1972), while chymotrypsin cleaves on the carboxy terminal side of tyrosine, phenylalanine and tryptophan residues (Fey et al. 1983). Acid hydrolysis cleaves between aspartic acid and proline residues (Sonderegger et al.,1982). However, the original method used formic acid, while in this study, hydrolysis using nine per cent glacial acetic acid in the SDS gel fixative was

used. For this reason, the residues which were cleaved may have differed from those which would be expected in formic acid cleavage.

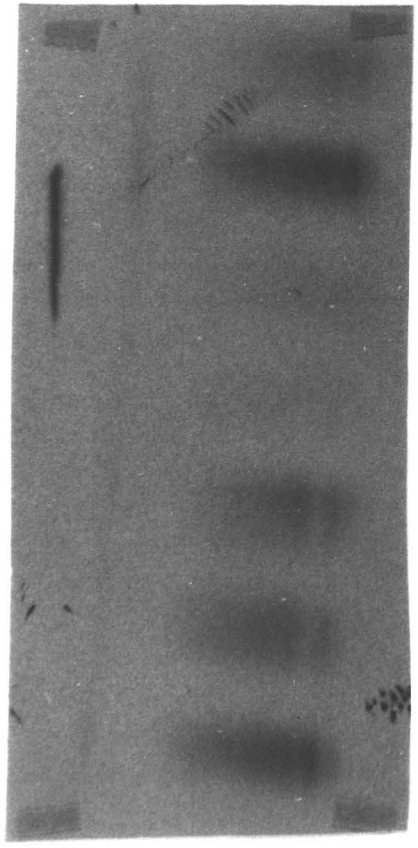
Cleveland et al. (1976) have shown that peptide mapping is quite insensitive to changes in substrate and enzyme concentrations, as well as incubation time. This helps to ensure that any differences observed between proteins within a single gel are real, not procedural artifacts. However, because this procedure involves partial cleavage, and because conditions may differ from gel to gel, some artifactual differences may occur from one gel to another. By comparing the peptide banding patterns of different strains on the same gel, the question of whether or not differences are real can be resolved.

Peptide mapping of the same iso-female lines was attempted using Staphylococcus aureas V8 protease, acid hydrolysis and chymotrypsin. Cleavage of 70K and 27K HSP's was successful. Cleavage of the 22K HSP's did not occur.

Sixteen lines (seven from France, nine from Benin) were surveyed for variation in their 70K HSP's after treatment with V8 protease (Figure 2a). Five bands were resolved in all cases. No variation was observed. Twenty-one lines (ten from France, ten from

FIGURE 2 A) An example of peptide mapping using Staphylococcus aureas V8 protease to cleave HSP 70 Iso-female lines are indicated.

B) An example of peptide mapping using Staphylococcus aureas V8 prtease to cleave HSP 27. Iso-female lines are indicated.



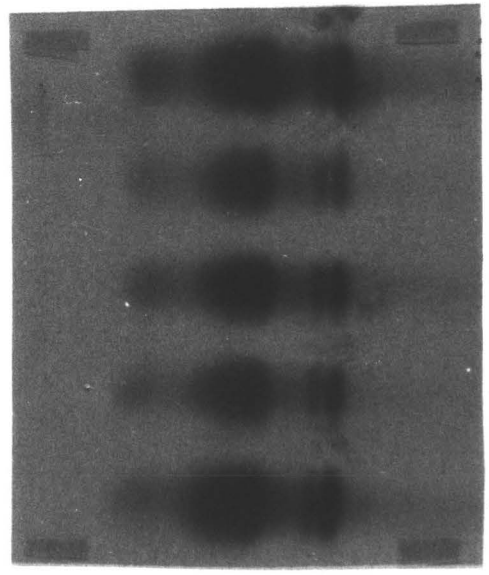
27K B23C

27K B19CA

27K B27

27K V#19-2

27K V#3-1



70K B27A

70K V#12-1

70K V#7-2

70K V#14-2

70K V#19-2

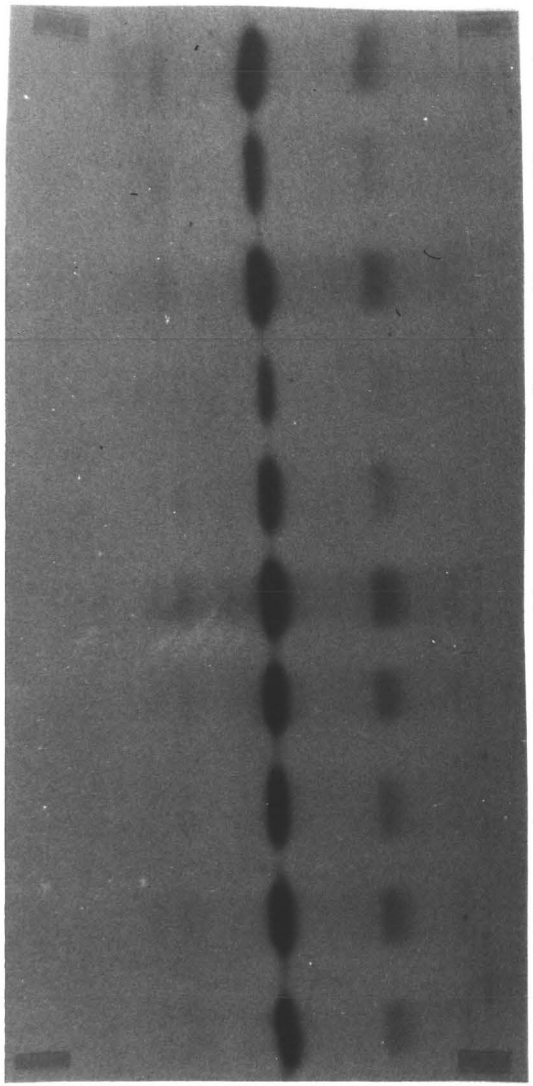
Brownsville, one from Benin) were surveyed for variation in their 70K HSP's after acid hydrolysis (Figure 3a). Ten major bands were resolved. While some variation was observed between gels, there was no variation seen within a gel. Therefore, it was concluded that the variation seen between gels was artifactual. Ten lines (six from France, two from Benin, two from Hamilton) were surveyed for variation in their 70K HSP's after treatment with chymotrypsin (Figure 4). Eleven major bands were resolved. No variation was observed.

Five lines (two from France, three from Benin) were surveyed for variation in their 27K HSP's after treatment with V8 protease (Figure 2b). Three bands were resolved, but no variation was observed. Twenty lines (ten from France, ten from Brownsville) were surveyed for variation in their 27K HSP's after acid hydrolysis (Figure 3b) Five bands were resolved, and no variation was observed. Six lines (two from France, two from Benin, two from Hamilton) were surveyed for variation in their 27K HSP's after treatment with chymotrypsin (Figure 4) Six bands were resolved, but no variation was observed.

Thus, from our SDS-PAGE analysis and peptide mapping of separated HSP's from about forty iso-female

FIGURE 3 A) An example of peptide mapping using acid hydrolysis to cleave HSP 70. Iso-female lines are indicated.

B) An example of peptide mapping using acid hydrolysis to cleave HSP 27. Iso-female lines are indicated



27K Br14

27K Br13

27K Br3

27K Br8

27K Br18

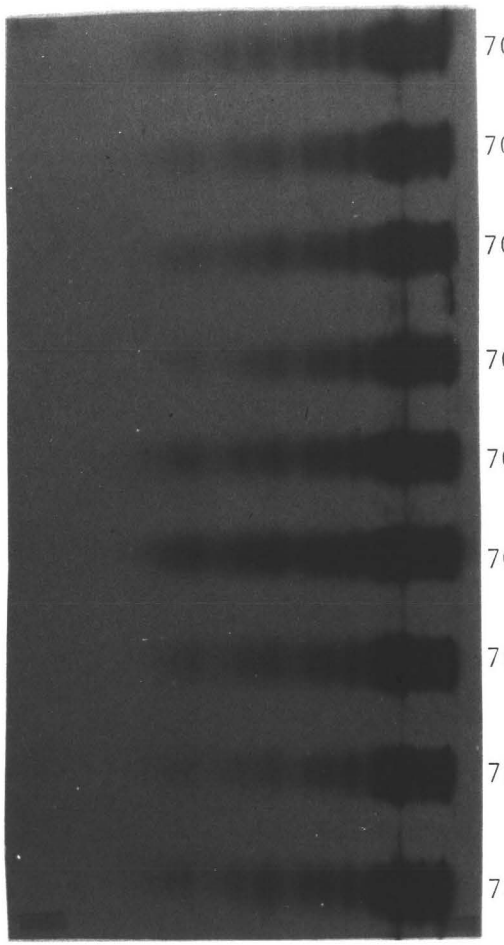
27K Br9

27K Br11

27K Br15

27K Br12

27K Br19



70K V#2-1

70K V#18-2

70K V#13-1

70K V#16-2

70K V#9-1

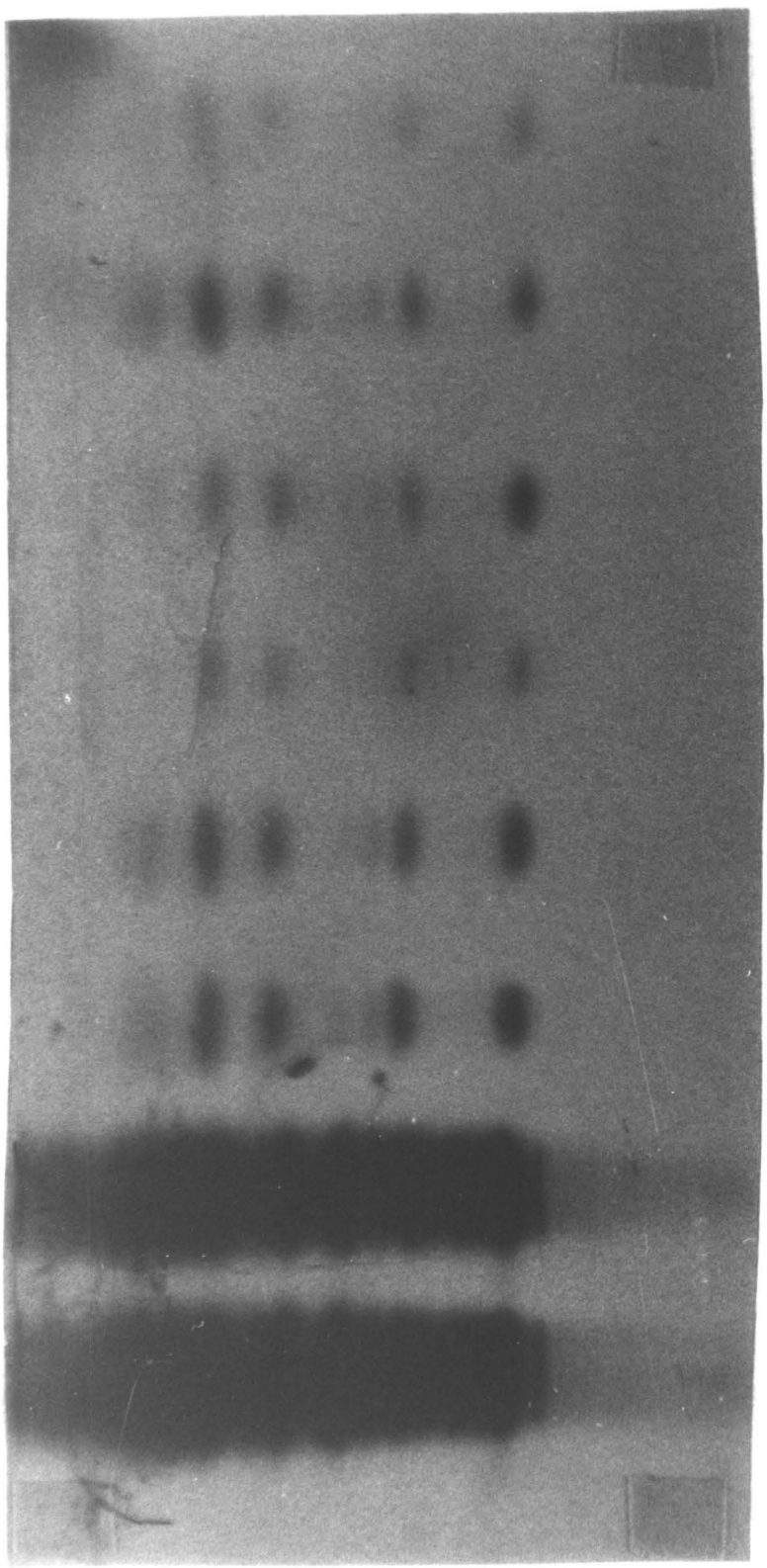
70K V#11-2

70K V#1-2

70K V#8-1

70K V#10-1

FIGURE 4 An example of peptide mapping using chymotrypsin to cleave HSP 27 and HSP 70. Iso-female lines are indicated


27K Da11
27K Db17
27K B23
27K B20CA
27K V#11-2
27K V#13-1
70K V#8-1
70K V#19-2

lines, we conclude that natural populations of D. melanogaster are monomorphic for their heat shock proteins.

VARIATION IN HSP'S AMONG DROSOPHILA SPECIES

Figure 6 represents a composite of SDS gels, such as the one shown in Figure 5, showing the relative mobilities of the 84K, 70K, 36K and small HSP bands from thirty-seven species of Drosophila. It illustrates two important results concerning HSP variation. First, the amount and type of variation depends upon the particular HSP being considered. Neither HSP84 nor HSP70 show any variation across the thirty-seven species. These results agree with those of a previous studies (Sinibaldi and Storti, 1982) and extend the survey beyond the Virilis and Melanogaster groups. The variation in HSP36 manifests itself as band mobility differences only, while the more extensive variation in the small HSP's includes differences in both mobility and number of bands. Second, the amount of HSP variation differs depending upon the phylogenetic relationships of the species being considered. For HSP36, the greatest variation in band mobility occurs among species from

FIGURE 5 An example of a 13% SDS-PAGE gel. *Drosophila* species names and the mobilities of various HSP's are shown.

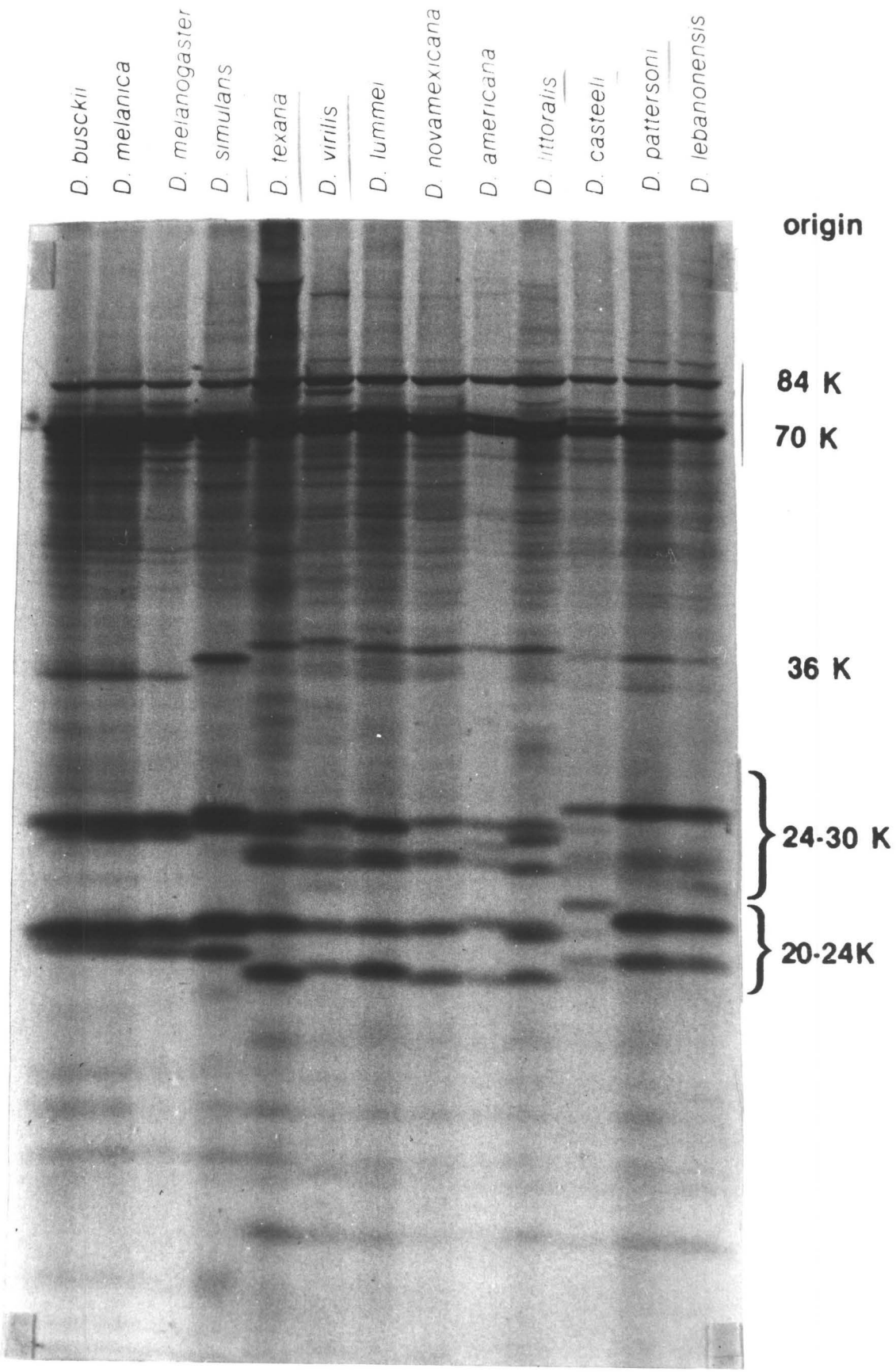
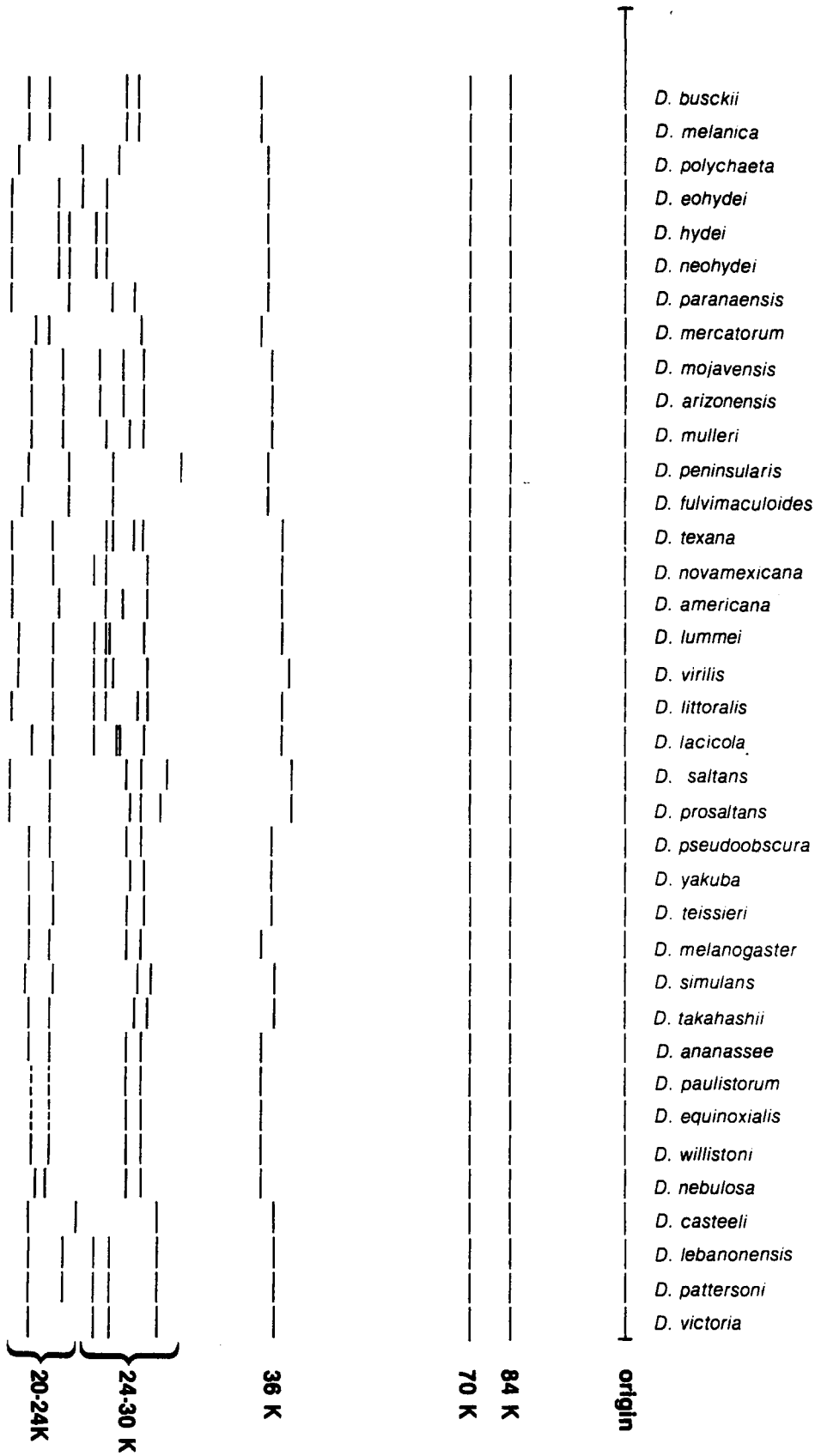

D. buscki
D. melanica
D. melanogaster
D. simulans
D. texana
D. virilis
D. lummei
D. novamexicana
D. americana
D. littoralis
D. casteeli
D. pattersoni
D. lebanonensis
origin
84 K
70 K
36 K
24-30 K
20-24K

FIGURE 6 A composite derived from 13% SDS-Page gels. Drosophila species names and the mobilities of various HSP's are shown.


D. busckii
D. melanica
D. polychaeta
D. eohydei
D. hydei
D. neohydei
D. paranaensis
D. mercatorum
D. mojavensis
D. arizonensis
D. mulleri
D. peninsularis
D. fulvimaculoides
D. texana
D. novamexicana
D. americana
D. lummei
D. virilis
D. littoralis
D. lacicola
D. saltans
D. prosaltans
D. pseudoobscura
D. yakuba
D. teissieri
D. melanogaster
D. simulans
D. takahashii
D. ananassee
D. paulistorum
D. equinoxialis
D. willistoni
D. nebulosa
D. casteeli
D. lebanonensis
D. pattersoni
D. victoria
origin
84 K
70 K
36 K
24-30 K
20-24K

different species groups, with less variation being observed among species of the same group. It must be noted that within the Virilis group, the mobility of HSP36 in D. virilis differs from the mobilities of the proteins in the other six species surveyed. This observation differs from the results of Sinibaldi and Storti (1982), who found no variation among the 36K proteins of these species. For the small HSP's, most of the differences in band number occur between different species groups, but some differences in band number occur between species belonging to the same group. (i.e., Virilis, Repleta and Victoria groups)

The differences in the number of bands among species sometimes made analysis difficult, especially in the Virilis and Repleta groups, which show additional small HSP's, different from the usual four (i.e. 22K, 23K, 26K, 27K). In order to treat the data objectively, all bands in each species were scored. This would not have been a problem, except that induction of these proteins has been shown to be very sensitive to even slight temperature changes (Sinibaldi and Storti, 1982). When determining the per cent similarity between any pair of species, the total number of bands shared was divided by the total number of bands of different mobilities. A lack of

induction caused by slight experimental temperature changes could distort these values in one of two ways. If the noninduced protein would have been shared, then the variation value would be inflated by decreasing the numerator and increasing the denominator. If the noninduced protein would not have been shared, then the variation value would be reduced by reducing the denominator. This was not a problem in the other groups because although mobilities were often different, mobility patterns were the same, making induction deficiencies obvious.

HSP VARIATION IN SIBLING VERSUS NON-SIBLING SPECIES

In the present work, analysis of HSP36 has been included with the analysis of the small HSP's. Tables 2-6 show the number of shared small HSP's among individual pairs of species from the Melanogaster, Willistoni, Repleta, Victoria and Virilis groups respectively. Tables 2-6 also show the total number of small heat shock variants between members of each species pair. Table 7 provides a species group comparison of the average number of shared small HSP's among pairs of species with different intragroup relationships. The pairs may be composed of

Table 2. Number of shared small HSP's (above diagonal) and the total number of small heat shock protein variants (below diagonal) among species pairs within the Melanogaster group.

	Mel.	Sim.	Yak.	Teis.	Tak.	Anan.
Mel.	-	0	1	2	2	5
Sim.	10	-	1	1	1	0
Yak.	9	9	-	4	1	1
Teis.	8	9	6	-	1	2
Tak.	8	9	9	9	-	2
Anan.	5	10	9	8	8	-

Table 3 Number of shared small HSP's (above diagonal) and the total number of small heat shock protein variants (below diagonal) among species pairs within the Willistoni group.

	Will.	Paul.	Equin.	Neb.
Will.	-	5	5	3
Paul.	5	-	5	3
Equin.	5	5	-	3
Neb.	7	7	7	-

Table 4 Number of shared small HSP's (above diagonal) and the total number of small heat shock protein variants (below diagonal) among species pairs in the Repleta group.

	Mer.	Par.	Moj.	Ari.	Mul.	Pen.	Ful.	Hyd.	Neo.	Eoh.
Mer.	-	0	0	0	0	0	0	0	0	0
Par.	9	-	0	0	0	3	3	3	3	2
Moj.	10	11	-	6	4	1	0	0	0	0
Ari.	10	11	6	-	4	0	0	0	0	0
Mul.	10	11	8	8	-	0	0	1	1	1
Pen.	9	7	11	11	11	-	3	2	2	1
Ful.	8	6	10	10	10	6	-	2	2	1
Hyd.	9	8	12	12	11	9	8	-	6	4
Neo.	9	8	12	12	11	9	8	6	-	4
Eoh.	9	8	11	11	10	9	8	7	7	-

Table 5. Number of shared small HSP's (above diagonal) and the total number of small heat shock protein variants (below diagonal) among species pairs within the Victoria group.

	Cast.	Leb.	Patt.	Vict.
Cast.	-	3	3	3
Leb.	7	-	6	5
Patt.	7	6	-	5
Vict.	7	6	6	-

Table 6. Number of shared small HSP's (above diagonal) and the total number of small heat shock protein variants (below diagonal) among species pairs within the Virilis group

	Tex.	Am.	Nov.	Lum	Vir.	Litt.	Lac.
Tex.	-	3	4	4	3	4	3
Am.	10	-	4	2	2	4	1
Nov.	9	8	-	4	4	6	3
Lum.	10	11	9	-	4	4	4
Vir.	11	11	9	10	-	4	2
Litt.	10	9	7	10	10	-	3
Lac.	11	12	11	10	12	11	-

subspecies, sibling species, non-sibling species from the same subgroup or non-sibling species from different subgroups.

A general result, shown by the data in Table 7, seems to be that the level of small HSP variation decreases as the phylogenetic relationships of the species pairs become closer. Sibling species show an average of sixty-eight per cent shared small HSP's. This decreases to thirty-nine per cent among non-sibling species from the same subgroup, while non-sibling species from different subgroups share only about eighteen per cent of their small HSP's. It should be noted that although subspecies share only thirty-six per cent of their small HSP's, this value is probably affected by a lack of data, since only two subspecies comparisons were made. It should also be noted that as the number of species comparisons increased, the shared fraction among species within a group decreased. Also, within particular groups (Victoria, Melanogaster and Virilis) the expected hierarchical arrangement of values does not occur. In the Victoria and Melanogaster groups, this result is probably due to the small sample size, but in the Virilis group, in which the two categories of non-siblings show nonsignificant variation in per cent

Table 7. Proportion of small HSP variants shared among sub-, sibling- and non-sibling species pairs within various groups of Drosophila.

Group*	Species relationship	No. of species pairs	No. of Shared small HSP variants (mean ± S.E.)	Per cent shared small HSP variants (mean ± S.E.)
Victoria	Subspecies	1	3	42.9
	Sibling	2	4.00 ± 1.41	63.1 ± 28.6
	Non-sibling	3	4.67 ± 1.53	75.4 ± 29.3
Saltans	Sibling	1	4	50.0
Willistoni	Sibling	3	5.00 ± 0.00	100 ± 0.0
	Non-sibling	3	3.00 ± 0.00	42.9 ± 0.0
Melanogaster	Sibling	1	0	0
	Non-sibling$_a$	5	1.80 ± 1.30	25.0 ± 24.1
	Non-sibling$_b$	9	1.56 ± 1.51	24.4 ± 29.6
Virilis	Subspecies	1	3	30.0
	Non-sibling$_a$	10	3.40 ± 0.84	35.4 ± 11.7
	Non-sibling$_b$	10	3.50 ± 1.35	37.0 ± 20.8
Repleta	Sibling	3	4.00 ± 3.46	66.7 ± 57.7
	Non-sibling$_a$	7	2.43 ± 1.99	31.9 ± 27.3
	Non-sibling$_b$	35	0.86 ± 1.14	11.1 ± 16.0

continued

Table 7 continued...

TOTALS	Species relationship	No. of species pairs	No. of shared small HSP variants (mean ± S.E.)	Per cent shared small HSP variants (mean ± S.E.)
	Subspecies	2	3.00 ± 0.00	36.4 ± 9.1
	Sibling	10	3.90 ± 2.23	67.6 ± 41.6
	Non-sibling$_a$	28	2.96 ± 1.50	37.8 ± 23.7
	Non-sibling$_b$	54	1.48 ± 1.58	18.3 ± 21.7

*-groups shown in descending order of age

a-species from same subgroup

b-species from different subgroups

shared small HSP's, sample size should not affect the result. Table 7 also shows that within the Willistoni and Repleta groups, the per cent shared small HSP's follows the expected hierarchical arrangement.

VARIATION IN HSP'S AMONG SPECIES GROUPS

Table 8 provides an intergroup comparison of the overall averages of the number of shared small HSP's among species within each group, regardless of the phylogenetic relationships of the species pairs. For example, the value for the Melanogaster group was derived by averaging values obtained from Table 2. By dividing these values by the average number of small heat shock protein variants in each group, the percentage of shared small HSP variants was determined. Based upon these results, the Repleta group shows the greatest variation among the small HSP's (82%), followed by the Melanogaster (77%), Virilis (64%), Saltans (50%), Victoria (34%) and Willistoni (29%) groups. While there appears to be an inverse relationship between the number of species examined and the average number of shared bands, the Virilis group appears to show a larger fraction of small HSP's shared among species.

Table 8. Comparison of shared HSP variants among species groups

Group	No. of Species	No. of shared small HSP variants (mean ± S.E.)
Victoria	4	4.17 ± 1.33
Saltans	2	4
Willistoni	4	4.00 ± 1.10
Melanogaster	6	1.60 ± 1.35
Virilis	7	3.43 ± 1.08
Repleta	10	1.31 ± 1.72

HSP SIZE VARIANTS AS AN INDICATOR OF SPECIES PHYLOGENY

The results of this study indicate that the 70K and 84K HSP's do not possess enough variation in length to be useful in the determination of the phylogenetic relationships among the various *Drosophila* species groups. The small HSP's, excluding HSP36, show too many variants to be of much use in a study of phylogeny. However, since each species appears monomorphic for HSP36, and making the parsimonious assumption that each 36K variant has arisen only once, this data may be used to construct a species group phylogeny.

Table 9 considers the 36K HSP variants separately. The proteins are numbered in ascending order of increasing mobility and the variants for each group are shown.

The *Victoria* and *Saltans* groups have been established as the oldest of those shown in Table 9 (Throckmorton, 1975). There is no variation within either group, but since *Saltans* is the only group with variant #1, while *Victoria* shares #4 with *Melanogaster*, and since *Melanogaster* is known to be more recent than either, it is reasonable to suggest that *Victoria* is more recent than *Saltans*.

Table 9 HSP 36 variation among species groups. Variants are numbered in increasing order of mobility. Group subscripts show assumed order of origin.

	No. of species	No. of 36k bands 1	2	3	4	5	6	7
Saltans$_1$	2	2	-	-	-	-	-	-
Victoria$_2$	4	-	-	-	4	-	-	-
Melanogaster$_3$	6	-	-	-	2	2	-	2
Willistoni$_4$	4	-	-	-	-	-	-	4
Repleta$_5$	10	-	-	-	-	3	6	1
Virilis$_6$	7	-	1	6	-	-	-	-
Buscckii$_3$	1	-	-	-	-	-	-	1
Melanica$_3$	1	-	-	-	-	-	-	1
Obscura$_3$	1	-	-	-	-	1	-	-
Polychaeta$_5$	1	-	-	-	-	-	1	-

Melanogaster also shares variant #7 with Willistoni. Assuming that #5 arose from #4, and that #7 arose from either #4 or #5, then the Willistoni group should be newer than the Melanogaster group. Melanogaster also shares #5 and #7 with Repleta, but a majority of Repleta species possess the unique #6, which is assumed to have arisen from #5 or #7. Therefore, Repleta seems to be newer than Melanogaster. The Virilis group contains variants #2 and #3, both of which are unique to this group. Since Virilis and Repleta are known to be recent groups, the data suggests that Repleta is older than Virilis.

Table 9 also shows four other groups represented by single species. These results are not as useful because it is not possible to determine whether or not the variants are unique within each group. However, since Melanica and Busckii possess variant #7, as does Melanogaster, it is possible that these three groups arose in similar time periods. Also, since the total heat shock protein pattern of Busckii and Melanica matches completely the pattern of Melanogaster (Figure 6), this also supports the suggested relationship. Polychaeta may have arisen at the same time as Repleta since it possesses one of the newer variants, #6. Finally, D. pseudoobscura, a member of the Obscura

group which shares variant #5 with Melanogaster and Repleta, may have arisen in a time period similar to that of either of the latter two groups.

In summary, this data supports the idea that the descending order of age of the HSP36 variants is #1, #4, #5, #7, #6, #2, #3 (or #3, #2). The descending order of age of the species groups is suggested to be Saltans, Victoria, Melanogaster (Obscura, Busckii, Melanica), Repleta (Polychaeta) and Virilis. The groups in which only one species was surveyed are shown in parentheses to emphasize the fact that their positions in this hierarchy are particularly tenuous.

DISCUSSION

The study of sequence conservation among heat shock proteins has proved interesting because the level of conservation seems to vary from one protein to another among different organisms. This work has shown that among the 84K and 70K proteins of different Drosophila species, no variation was detectable by SDS-PAGE, while the smaller proteins showed abundant variation. In order to fully understand the evolutionary significance of these differences in levels of conservation, the specific function(s) of each of the proteins will have to be elucidated and many more DNA sequences from closely related and highly diverged species will have to be obtained. The present level of understanding concerning HSP function(s) has already been considered (see Introduction) and this discussion will present and analyze heat shock gene and protein sequence data with particular emphasis on its relevance to questions concerning the variation and evolution in the size of proteins.

The evolution of multigene families is a topic which has been studied for the past thirty

years. Its relevance to the study of heat shock gene evolution is obvious, since the hsp genes are organized into two types of family. The first type is illustrated by the hsp70 genes. In Drosophila, HSP70 is coded for by at least four gene copies located at two widely spaced loci. As already mentioned, it is probable that these genes provide for the same function, since restriction analysis and sequencing studies have shown almost complete homology, and since deletion mutants with only one copy of the gene are viable and show a normal response to heat shock. The second type is illustrated by the small hsp genes. While these genes obviously code for different proteins, as shown by SDS-PAGE studies, their homology and common regulation suggest a common origin, and there seems to be only one copy of each.

The formation of such families of genes has been hypothesized to be the result of a number of events, the first, and most important of which is the duplication of an ancestral gene. (Maeda and Smithies, 1986). This event is also the most difficult to study, since later mutations are likely to have changed the sequence so much that the original event is unrecognizable. Maeda and Smithies (1986) have suggested four scenarios which could produce this

initial duplication. The first is a nonhomologous chromosomal breakage and reunion. Maeda et al. (1984) have shown that an allele (Hp2) at the haptoglobin-gene locus which is almost twice as long as the other two common alleles was formed by an unequal crossover between introns found in each of the other alleles. Since the 5' and 3' crossover points show no sequence homology, this was a nonhomologous event. The second is an unequal crossing over between two homologous sequences which are found on either side of a gene. Supporting evidence was provided by Shen et al. (1981), who showed that the human G and A fetal globin genes contain homologous one hundred and twenty base pair regions in the 5' and 3' areas of each gene and between the genes. The third is an RNA mediated DNA duplication, by which a DNA copy of an RNA transcript is randomly inserted in the genome. Such genes are recognizable because they do not possess introns, they often have 3' poly (A) regions, and they usually turn out to be pseudogenes, because gene copies made from RNA transcripts do not carry regulatory sequences. Some examples are the immunoglobulin, tubulin and actin gene families, which must have acquired the necessary regulatory sequences following the duplication (Vanin, 1985). The fourth

is gene amplification, caused by many replication initiation events at a single origin (Maeda and Smithies, 1986). These genes are recognized because their orientations differ and their endpoints are variable. However, they are so far unknown among higher eukaryotes (Maeda and Smithies, 1986).

Following the original duplication event, genes may diverge by the accumulation of subsequent events. These may include point mutations, deletions, insertions and more duplications. If the original function is satisfactorily provided by one copy, a duplicate may be free to change and either acquire a new function, as well as the selection constraints which go along with that function, or become a functionless pseudogene. However, as is the case with the hsp70 gene family in Drosophila, it has been shown that independent evolution of the duplicated genes does not always occur (Leigh Brown and Ish-Horowicz, 1981).

In evolutionary studies, two types of protein and DNA sequence data are important. First, a comparison of different gene and protein sequences within the same species can provide useful information concerning the evolution of multigene families. Second, a comparison of sequences for the same gene or

protein from different species can provide useful information concerning phylogeny, as well as the selective constraints acting upon that particular gene.

Of the heat shock genes, hsp 70 has been the most extensively characterized in the largest number of species. It has previously been noted that the degree of conservation among the five hsp70 gene copies of D. melanogaster is greater than would be expected if they were evolving separately. Restriction analysis by Leigh Brown and Ish-Horowicz (1981) showed that in three pairs of Drosophila hsp70 coding elements, restriction sites changed in parallel. Since the probability of six sites out of thirty-six surveyed changing randomly in parallel was very small, it was concluded that these elements were not evolutionarily independent. Similar situations are known to occur among gene families in higher eukaryotes. Zimmer et al., (1980), noticed that the previously mentioned human alpha-globin genes were more similar than expected, and suggested that this could be the result of many unequal crossovers. This idea was confirmed by Oliviers et al., (1985), who showed that multiple crossovers had occurred at this locus, and by Metzenberg and Wurzer (1986 unpublished

data), who demonstrated by sequence data that the Lepore fusion genes were the result of unequal crossovers between human delta and beta globin genes. Because the hsp70 loci include inverted genes between which unequal crossovers can not occur, a different mechanism, involving chromosome bending and crossing over, was suggested by Leigh Brown and Ish-Horowicz (1981).

One of the most highly conserved regions of the heat shock gene is the regulatory sequence. A consensus heat shock regulatory element (HSE) has been identified in Drosophila by Pelham (1982). He used deletion mapping to show that heat induction is caused by a fourteen base sequence which forms an inverted repeat (dyad), a property commonly seen among protein recognition sites. This sequence, which is located upstream from the TATA box, was found in all of the Drosophila heat shock genes, with hsp23 showing the least homology with the consensus sequence (fifty per cent), and hsp83, hsp22 and hsp26 showing the greatest homology with the consensus sequence (ninety-three per cent). It has also been found in organisms as divergent as the soybean (Schoffl et al., 1984), in which it showed seventy per cent homology with the Drosophila consensus

sequence. This remarkable conservation over great phylogenetic distances suggests that all of these proteins are induced by a common mechanism, supporting the idea of related, if not common, functions.

Other sequencing studies have shown the high degree of sequence conservation among the members of the hsp70 gene family (Craig et al., 1983; Craig and Ingolia, 1982). Table 11 shows that two genes from the 87A and 87C loci of D. melanogaster are ninety-seven per cent homologous in their amino acid sequences and ninety-four per cent homologous in their nucleotide sequences. It also shows that the three heat shock cognates, while clearly related to the hsp70 genes, have diverged from them to a much greater extent than the hsp70 genes have diverged from each other. This data suggests that between the 87A and 87C loci, there is a constraint upon independent evolution which can not be entirely explained by the functional constraints placed upon the proteins. Because the DNA sequence is almost as highly conserved as the protein sequence, it is evident that many of the silent codon sites are being prevented from diverging. This supports the previously mentioned idea that these loci are being

prevented from evolving separately by some mechanism related to DNA structure, rather than by strictly functional constraints upon the proteins. Hunt and Morimoto (1985) have also shown that the same phenomenon has occurred between species, since the observed divergence is only forty-five per cent between Drosophila and human hsp70 genes at the third codon position. Table 10 also supports the idea that the heat shock cognates, which are induced at normal temperatures, provide for functions different from those of the hsp70 genes.

Much more hsp70 sequence data is available for the comparison of widely diverged species than for closely related ones. Hunt and Morimoto (1985) and Lindquist (1986) have compared the amino acid sequences among five species; Table 11 shows comparisons of these sequences in different parts of the proteins. One particularly interesting result is the fact that in the central domain between amino acid residues 250-375, much less homology is observed in comparisons between E. coli and each of the four eukaryotes than among the four eukaryotes themselves. Therefore, it is reasonable to suggest that this central domain may define functional differences between the HSP70 proteins of prokaryotes and

Table 10 Sequence comparisons among selected members of Drosophila heat shock genes and proteins. Nucleotide sequence homologies (per cent) are shown below the diagonal. Amino acid sequence homologies (per cent) are shown above the diagonal. (data from Craig and Ingolia, 1982; Craig et al., 1983)

	HSC1	$HSC1_{sim}$	HSC2	HSC4	$HSP70_{87A}$	$HSP70_{87C}$
HSC1	---	100	-	-	-	75
$HSC1_{sim}$	98.5	---	-	-	82	75
HSC2	74-79	-	---	-	-	77
HSC4	74-79	74-79	-	---	-	82
$HSP70_{87A}$	-	-	-	-	---	97
$HSP70_{87C}$	-	-	77	76	94	---

Table 11 Amino acid homologies (per cent) among human, yeast, *E. coli*, *Xenopus* and *D. melanogaster* HSP70 proteins

	Human	Dros.	Yeast	Xen.	amino acid position
E. coli	42	42	52	44	
Human	-	75	74	86	1-125
Dros.	-	-	66	74	
Yeast	-	-	-	73	
E. coli	70	70	66	69	
Human	-	87	84	92	125-150
Dros.	-	-	79	83	
yeast	-	-	-	82	

continued...

Table 11 continued...

	human	Dros.	yeast	Xen.	amino acid position
E. coli	35	35	38	36	
Human	-	73	81	90	250-375
Dros.	-	-	72	77	
Yeast	-	-	-	82	
E. coli	62	61	63	66	
Human	-	80	82	93	375-500
Dros.	-	-	80	82	
Yeast	-	-	-	85	

continued..

Table 11 continued...

	Human	Dros.	Yeast	Xen.	amino acid position
E. coli	24	26	30	29	
Human	-	56	50	67	500-625
Dros.	-	-	48	46	
Yeast	-	-	-	52	
E. coli	20	12	20	08	
Human	-	40	60	60	625-650
Dros.	-	-	48	20	
Yeast	-	-	-	48	

eukaryotes. Another interesting result is that the carboxy termini of these proteins have diverged to a greater extent than the other domains, and most of the differences have been the result of deletions, rather than amino acid mismatches. The implications of this will be considered in the discussion of the evolution of protein size.

While the only sequence data available for HSP84 comes from Drosophila, yeast and E. coli, it appears that this is the second most highly conserved of the heat shock proteins. Farrelly and Finkelstein (1984) have shown that the Saccharomyces cerevisiae amino acid sequence is sixty-five per cent homologous with the Drosophila sequence, with one stretch of about fifty amino acids at the amino terminus showing ninety per cent homology, while the Drosophila and E. coli amino acid sequences show about thirty-six per cent homology and also contain regions of up to ninety per cent homology. (Bardwell and Craig, unpublished data). The lack of variation in HSP84 among the Drosophila species is not inconsistant with the above data, since the Drosophila species are so closely related.

As mentioned in the Introduction, antibodies for the chicken HSP84 protein cross react with the

proteins from Drosophila, frog, mouse and human. This fact, combined with the above mentioned sequence homologies, lack of protein variation on SDS gels and the observation that these proteins are all slightly acidic, with pH values ranging from 5.1 to 5.4 (Lindquist, 1986), suggests that the HSP84 proteins might provide similar functions in the different organisms. Whether or not there is a constraint upon third position codon variation, as was the case among the hsp70 genes, remains to be determined.

In Drosophila, all four of the small hsp genes have been sequenced, as well as one of the related, non-heat induced genes (gene 1). The levels of amino acid homology have been shown to vary across the different domains in the proteins (Ayme and Tissieres,1985). The first homologous region is in the fifteen amino-terminal amino acids. The second region is the most interesting; It corresponds to an eighty-three amino acid stretch, running from positions 122-204 in each of the small HSP's, which shows substantial homology to the bovine alpha-crystalline B2 chain. This region shows a twenty-two per cent homology among all six proteins, and a fifty-one per cent homology among five of the six. Also, the homology among the heat shock proteins

continues for twenty-five amino acids beyond the end of the alpha-crystalline homology.

It has been suggested that the conserved domains are responsible for the function(s) of the small HSP,s (Ingolia and Craig, 1982; Ayme and Tissieres, 1985). Because the alpha-crystalline protein is known to aggregate, and because the small HSP's have been shown to produce aggregates in the nuclei of heat shocked cells, it is possible that the conserved region causes aggregation of these proteins. It is also possible that this is the only functionally constrained domain in the small HSP's.

Assuming that no polymorphisms exist within species, there are two possible explanations for the variation between species. First, the variants could be neutral or have a high selective advantage. In this case , the lack of intraspecies polymorphism could be explained by a very low mutation rate. The second possibility is the exact opposite. Heterozygote disadvantage would cause fixation of one or the other variants by random genetic drift, in which case, the only variation would be between

species.

Why do the small HSP's show more variation than the 70K and 84K HSP's? There are two possible answers. The variants may be selected for, or they may be neutral, in which case the lack of polymorphism within species could be explained by a mutation rate which is lower than the rate of speciation.

Another question which must be addressed is "How does the level of variation among the heat shock proteins within a species compare to the variation seen at other loci?" Singh and Rhomberg (1987) have shown in a study of variation which included the same four populations of D. melanogaster as were surveyed in this heat shock study, that approximately two thirds of the one hundred seventeen loci studied were monomorphic. Of the remaining polymorphic loci, the mean number of alleles was about 1.55 per locus. The heat shock protein loci, all of which seem to be monomorphic as shown by SDS-PAGE and peptide mapping, seem to belong to the larger group.

Lewontin has suggested that proteins may be divided into two groups; those whose divergence causes speciation events to occur and those whose divergence follows speciation events. In Drosophila, the first category would include the male reproductive tract

proteins, which were studied by Coulthart (1985, Ph.D. thesis). Proteins of this category are restricted to specific tissues, and in this case, it can be postulated that the speciation events which caused the separation of the sibling species D. melanogaster and D. simulans occurred because of differences in the male reproductive tract proteins. The outstanding difference between this type of protein and the heat shock proteins is that the HSP's are not localized to a particular type of tissue. Also, as shown here, they show no variation among populations within a species. This suggests that the HSP,s belong to the group of proteins whose divergence follows speciation events.

The detailed study of the evolution of protein length has not been included in evolutionary studies because of the previously mentioned belief of most population geneticists that protein evolution is the result of substitutions, rather than additions or deletions in the coding regions of a gene. However, a few introductory studies of this question have been carried out. Whalen et al. (1986), found many length polymorphisms in the accessory gland proteins of Drosophila males. He postulated that these had been allowed to occur because of a lack of selection, which

was based upon the idea that these are mainly nutritive proteins.

Other studies have also provided data on protein length without addressing the question of its significance. In her 1986 review, Lindquist shows that the 70K HSP's of Drosophila, yeast, Xenopus, E. coli and humans have slightly different lengths. It is interesting to note that the eukaryotes all possess an internal deletion with respect to the prokaryotic sequence, which could have evolutionary significance. In the present work, length polymorphisms were shown to be absent from the large heat shock proteins of D. melanogaster and the other Drosophila species. However, while the result was the same for the small HSP's within D. melanogaster, there was ample variation in length observed among species.

While sequences for intraspecies or closely related interspecies comparisons are not available, the Genebank DNA and protein data bank provides data for a wide variety of protein lengths and sequences from a large number of species. A sample of this data is provided in Table 12. Although the number of species compared for each protein varies widely, it can be seen that length polymorphisms have occurred in most of the proteins examined, with standard errors

ranging from less than one per cent to greater than thirteen per cent. This data, by itself, does not say very much about the role of length variation in the evolution of proteins. However, a comparison of length variation with known rates of amino acid substitution would be useful. Also a comparison of the sequences from the available species, in order to provide information about the sites of the polymorphisms, as well as the sequencing and analysis of proteins from within a species will be necessary in order to properly address this question.

In conclusion, the present study has shown that the heat shock proteins of D. melanogaster do not vary in length. However, among different species of Drosophila, the small heat shock proteins show plenty of length variation. It is suggested that the heat shock proteins belong to the group of proteins which is monomorphic within a species and that their divergence has resulted from, speciation rather than causing it to occur.

In the future, a number of studies should be undertaken in order to pursue further the question of heat shock protein variation. Phylogenetic studies involving HSP36 should include more species, especially from the four groups which were represented by only one species each. The question of whether HSP polymorphism does, in fact, exist within species

Table 12 Interspecies length variation among selected proteins. Data was derived from the Genebank data bank

Protein	No. of Species Surveyed	Protein Length (mean ± S.E.)
Actin	12	374.3 ± 5.53
Adh	5	335 ± 45.6
alpha-crystallin A chain	43	172.2 ± 6.9
alpha-crystallin B chain	3	175.7 ± 1.15
aspartate aminotransferase (cytoplasmic)	2	411.5 ± 0.7
aspartate aminotransferase (mitochondrial)	2	403 ± 2.8
ATP-ase protein 6	5	231.6 ± 13.7
calcitonin	6	32 ± 0.0
calmodulin	12	149.6 ± 2.7
cytochrome b (mitochondrial)	8	380 ± 7.5

continued...

Table 12 continued...

Protein	No. of Species Surveyed	Protein Length (mean ± S.E.)
cytochrome c	80	110.4 ± 8.0
G-3 PD	9	332.8 ± 4.2
small HSP's	4	195.5 ± 18.6
HSP 70	5	641.8 ± 3.4
Histone H2A	9	130.4 ± 12.9
Histone H3	9	136.5 ±. 5.1
Osteocalcin	6	48.8 ± 1.0
Superoxide dismutase	6	152.8 ± 2.4

should be addressed by the survey of more species. Peptide analysis of size variation could be used in order to see whether the variants really are all size variants, or whether some are site variants. Amino acid sequence analysis of the proteins could be used to see whether size variants result from internal deletion/addition events or those occurring closer to the end of the protein. Finally, the question of why all of the heat shock genes are found on chromosome 3 should be addressed.

LITERATURE CITED

Aquadro, C.F., Desse, S.F., Bland, M.M. Langley C.H. and Laurie-Ahlberg, C.C., (1986), Molecular population genetics of the Adh gene region of Drosophila melanogaster. Genetics 114:1165-1190

Arrigo, A.P., Darlix, J.L., Kandjion, E.W., Simon, M. and Spahr, P.F., (1985), Characterization of the prosome from Drosophila and its similarity to the cytoplasmic structures formed by the low molecular weight heat-shock proteins. E.M.B.O. J. 4:399-406

Artavanis-Tsakonas, S., and Schedl, P., (1979), Genes for the 70,000 Dalton heat shock protein in two cloned D melanogaster DNA Segments. Cell 17:9-18

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

Avise, J.C., Gibllin-Davidson, C., Laerm, J., Patton, J.C. and Lansman, R.A., (1979), Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher. P.N.A.S. USA 76:6694-6698

Ayala, F.J., (1982) Genetic variation in natural populations: problems of electrophoretically cryptic alleles. Proc. Nat. Acad. Sci. U.S.A.79:550-554

Ayala, F.J., (1984), Molecular polymorphism: how much is there and why is there so much? Dev. Genet. 4: 379-391

Ayme, A. and Tissieres, A., (1985), Locus 67B of Drosophila melanogaster contains seven, not four, closely related heat shock genes. E.M.B.O. J. 4:2949-2954

Bardwell J. and Craig, E.A., (1984), Major heat shock gene of Drosophila and the Escherichia coli heat-inducible dnaK gene are homologous. P.N.A.S. USA 81:848-852

Basyczynski, C.L., Walden, D.B. and Atkinson, B.G., (1982), Regulation of gene expression in corn (Zea mays L.) by heat shock. Can. J. Biochem. 60:569-579

Basyczynski, C.L., Walden, D.B. and Atkinson, B.G., (1983), Regulation of gene expression in corn (Zea mays L.)

by heat shock.II. in vitro analysis of RNA's from heat-shocked seedlings. Can. J. Biochem. Cell Biol. 61:395-403

Burdon, R.H. and Certmore, C.M.M., (1982), Human heat shock gene expression and the modulation of plasma membrane Na+, R+ -ATPase activity. FEBS Letters 140:45-48

Buzin, C.H. and N.S. Petersen, (1982) A comparison of the multiple Drosophila heat shock proteins in cell lines and larval salivary glands by two-dimensional electrophoresis. J. Mol. Biol. 158:181-202.

Chappell, T.G., Welch, W.J., Schlessman, D.M., Palter, K.B., Schlesinger, M.J. and Rothman, J.E., (1986), Uncoating ATPase is a member of the 70 Kiladalton family of stress proteins. Cell 45:3-13

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K., (1976), Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. 252:1102-1106

Coulthart, M.B. (1986) Evolution in Drosophila male reproductive tract proteins. Ph.D. Thesis: MacMaster University, Hamilton, Ontario.

Coyne ,J.A., (1976), Lack of genic similarity between two sibling species of Drosophila as revealed by varied techniques. Genetics 84:593-507

Coyne, J.A., and Felton, A.A., (1977), Genic heterogeneity at two Adh loci in Drosophila pseudoobscura and Drosophila persimilis. Genetics 87:285-304

Craig, E.A., Ingolia, T.D. and Manseau, L.J., (1983), Expression of Drosophila heat-shock cognate genes during heat shock and development. Dev. Biol. 99:418-426

Craig, E., Ingolia, T., Slater, M., Manseau, L. and Bardwell, J., (1982), Heat shock- from bacteria to man. pages 11-18

Craig, E.A.and Jacobsen, K., (1984), Mutations of the heat inducible 70 Kilodalton genes of yeast confer temperature sensitive growth. Cell 38:841-849

DiDomenico, B.J., Bugarisky, G.E. and Lindquist, S., (1982), The heat shock response is self-regulated at both the transcriptional and posttranscriptional levels. Cell 31:593-603

Farrelly, F.W. and Finkelstein, D.B., (1984), Complete sequence of the heat shock inducible HSP90 gene of saccharomyces cerevisial. J. Biol. Chem. 259:5745-5751

Fey, S.J., Bravo, R., Mose Larsen, P. and Celis, J.E., (1983), Correlation between mouse and human two-dimensional gel patterns: Peptide mapping of proteins extracted from two-dimensional gels. pp. 170-193 in Celis, J.E., and Bravo, R., (eds.), Two-dimensional gel electrophoresis of proteins:methods and applications., New York: Academic Press

Hauser, L. and Levy-Wilson, B., (1981), Kinetic changes in protein synthesis to a sublethal heat shock in starved T. thormophila. J. Biol. Chem. 256:3612-3614

Hjelmelland, L.M., Nebert, D.W. and J.C. Osborne. (1983) Sulfobetaine derivatives of bile acids: non-deaturing surfactants for membrane biochemistry. Anal. Biochem. 130:72-82.

Holmgren, R., Livak, K., Morimato, R. and Meselson, M., (1979), Studies of cloned sequences from four drosopila heat shock loci. Cell 18:1359-1370

Houmard, J., and Drapeau, G.R., (1972), Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds. P.N.A.S. USA 69: 3506-3509

Hubby, J.L. and Lewontin, R.C., (1966), A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in Drosophila pseudoobscura. Genetics 54:577-594

Hughes, E.N. and August, J.T., (1982), Co-precipitation of heat shock proteins with a cell surface glycoprotein. P.N.A.S. USA 79:2305-2309

Hunt, C. and R.I. Morimoto. (1985) Conserved features of eukaryotic HSP70 genes revealed by comparison with the nucleotide sequence of human HSP70. Proc. Nat. Acad. Sci. U.S.A. 82:6455-6459.

Ingolia, T.D. and Craig, E.A., (1982), Drosophila gene related to the major heat shock induced gene is transcribed at normal temperatures and not induced by heat shock. P.N.A.S. USA 79:525-529

Ingram, V.M., (1956), A specific chemical difference between the globins of normal human and sickle cell anaemia haemoglobin. Nature 178: 792-794

Ingram, V.M. (1959) Abnormal human haemoglobin. 111. The chemical difference between normal and sickle-cell anaemia haemoglobin. Nature 178:792-794.

Ireland, R.C., Berger, E., Sirothin, K., Yund, M.A., Osterbur, D. and Fristrom, J., (1982), Ecolysterone induces the transcription of four heat shock genes in Drosophila S3 cells and imaginal discs. Dev.Biol. 93:498-507

Ish-Horowicz, D. and Pinchin, S.M., (1980), Genomic organization of the 87A7 and 87C1 heat induced loci of Drosophila melanogaster. J. Mol. Bio. 142:231-235

Keith, T.P., (1983) Frequency distribution of esterase-5 alleles in two populations of D. pseudoobscura. Genetics 105: 135-155.

Kimura, M., and Ohta, T., (1974) On some principles governing molecular evolution. P.N.A.S. USA 71:2848-2852

Korge, G., (1977), Larval saliva in Drosophila melanogaster: production, composition and relationship to chromosome puffs. Dev. Biol. 58: 339-355

Kreitman, M.,(1980) Assessment of variability within electromorphs of alcohol dehydrogenase in Drosophila melanogaster. Genetics 95:467-475.

Kreitman, M., (1983), Nucleotide polymorphism at the alcohol dehydrogenase locus of Drosophila melanogaster, Nature 304:412-417

Kurtz, S., Rossi, J., Petco, L.and Lindquist, S., (1986), An ancient developmental induction: heat-shock proteins induced in spowlation and oogenesis. Science 231:1154-1157

Laemmli, U.K., (1970), Cleavage of the structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

Lanks,R.,(1986), Modulators of the eukeryotic heat shock response. Experimental Cell Research 165:1-10

Leigh Brown, A.J. and Ish-Horowicz, (1981), Evolution of the 87A and 87C heat-shock loci in Drosophila. Nature 290:677-682

Levinger, L. and Varshavsky, A., (1981), Heat-shock proteins of Drosophila are associated with nuclease-resistant, high-salt-resistant nuclear structures. J. Cell Biol. 90:793-796

Lewontin, R.J., (1974),. The genetic basis of evolutionary change. New York: Columbia University Press.

Lewontin, R.C., (1985), Population Genetics. Ann. Rev. Genet. 19:81-102

Lindquist, S., (1986), The heat shock reponse. Ann. Rev. Biochem. (1986) 55:1151-1191

Loomis, W.F. and Wheeler, S., (1980), Heat shock response of dictyostelium. Dev. Biol. 79:399-408

Lowe, D.G. and Moran, L.A., (1984), Proteins related to the mouse L-cell major heat shock protein are synthesized in the absence of heat shock gene exrpession. P.N.A.S. USA 81:2317-2321

Maeda, N. and O. Smithies. (1986) The evolution of multigene families: Human haptoglobin genes. Ann. Rev. Genet. 20:81-108.

Maeda, N., Yang, F., Barnett, D.R., Bowman, B.H. and Smithies. (1984) Duplication within the haptoglobin Hp gene. Nature 309:131-135.

Modiano,G., Battistuzzi, G., Esan, G.J.F., Testa, U. and Luzzatto, L., (1979), Genetic heterogeneity of normal human erythrocyte G6PD: An isoelectophoretic polymorphism. P.N.A.S. USA 76:852-856

Neidhardt, F.C., Van Bogelen, R.A. and Vaughn, V., (1984), The genetics and regulation of heat-shock proteins. Ann. Rev. Genet. 18:295-329

O'Farrell, P.H., (1975), High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250: 4007-4021

O'Farrell, P.Z., Goodman, H.M. and P.H. O'Farrell. (1977) High-resolution two-dimension gel electrophoresis of basic as well as acidic proteins. Cell 12:1133-1142.

Oliviers, S., DeMarchi, M., Bermini, L.F. Bensi, G., Raugi, G. and A.O. Carbonara. (1985) Molecular evidence of triplication in the haptoglobin Johnson variant gene. Hum. Genet 71:49-52.

Pelham, H.R.B. (1982) A regulatory upstream promoter element in the Drosophila HSP70 heat-shock gene. Cell 30:517-528.

Pelham, H.R.B., (1984), Hsp 70 accelerates the recovery of nuclealar morphology after heat shock. E.M.B.O. J. 3:3095-3100

Phillips, T.A., Van Bogelen, R.A. and Neidhardt, F.C., (1984), Con gene product of escherichia coli is a heat-shock protein. J. Bacterial. 159:283-287

Pierce, D.A. and Lucehesi, J.C., (1980), Chromosoma 76:245-254

Postlethwait, J.H. and Jowett, T., (1980), Genetic analysis of the hormonally regulated yolk polypeptide genes in D. melanogaster. Cell 20:671-678

Ramshaw, J.A.M., Coyne, J.A. and Lewontin, R.C., (1979), The sensitivity of gel electrophoresis as a detector of genetic variation. Genetics 93: 1019-1037

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

Roberts, D.B. and Evans-Roberts, S., (1979), The genetic and cytogenetic localization of the three structural genes coding for the major protein of Drosophila serum. Genetics 93:663-679

Sanchez, E.R., Taft, D.O., Schlesinger, M.J. and Pratt, W.B., (1985), Evidence th t the 90-kDa phosphoprotein associated with the untransformed L-cell glueacorticaid receptor is a menine heat shock protein. J. Biol. Chem. 260:12398-12401

Schlesinger, M.J., Kelly, P.M., Aliperti, G. and C. Malfer. (1982) Properties of three major chicken heat-shock proteins and their antibodies. pp. 243-250 in Schlesinger, M.J., Ashburner, M. and A. Tissieres (eds.) Heat-shock:from bacteria to man. Cold Spring Harbor, N.Y.

Schoffl, F., Raschke, E. and T.T. Nagao. (1984) The DNA sequence analysis of soybean heat-shock genes and identification of possible regulatory promoter elements. EMBO J. 3:2491-2497.

Schuh, S., Yanemato, W., Brugge, J., Bauer, V.J., Riehl, R.M., Sullivan, W.P. and Taft, D.O., (1985), A 90,000 -Dalton binding protein common to both steroid receptors and the Rous Sarcoma Virus transforming protein, pp60^{v-src}. J. Biol. Chem. 260:14292-14296

Shah, D.M. and Langley, C.H., (1979), Inter and intraspecific variation in restriction maps of Drosophila mitochondrial DNA's. Nature 281:1-3

Shen, S., Slightom, J.L. and O.Smithies. (1981) A history of the human fetal globin gene duplication. Cell 26:191-203.

Singh, R.S., Hubby, J.L. and Throckmorton, L.H., (1975), The study of genic variation by electrophoretic and heat denaturation techniques at the octanol dehydrogenase locus in members of the Drosophila virilis group. Genetics 80:637-650

Singh, R.S., Lewontin, R.C. and Felton, A.A., (1976), Genetic heterogeneity within electrophoretic alleles of xanthine dehydrogenase in Drosophila pseudoobscura. Genetics 84:609-29

Sinibaldi R. and Storti, R.V., (1982), One- and Two-Dimenional polyacrylamide gel analysis of heat shock proteins of the virilis group of Drosophila. Biochem. Genet. 20:791-

Sirotkin, K. and Davidson, N., (1982), Developmentally regulated transcription from Drosophila melanogaster chromosomal site 67B. Dev. Biol. 89:196-210

Sondregger, P., Jaussi, R., Gehring, H., Bruschweiler, K. and P. Christen. (1982) Peptide mapping of protein bands from polyacrylamide gel electrophoresis by chemical clevage in gel pieces and re-electrophoresis. Anal. Biochem. 122:298-301

Tanguay, R.M., (1983), Genetic regulation during heat shock and function of heat shock proteins: a review. Can. J. Biochem. Cell Biol. 61:387-394

Throchmorton, L.H., (1975), The phyloginy, ecology and geography of Drosophila. In Handbook of Genetics, (ed. R.C. King) Vol. 3, p.421-470, Plenum Press, N.Y.

Tilly, K., McKittrick, N., Zylicz, M. and Georgopoulos, C., (1983), The dna K protein modulates the heat shock response of Escherichia coli. Cell 34:641-646

Tissieres, A., Mitchell, H.K. and Tracy, U.M., (1974), Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J. Mol. Biol. 84:389-398

Udvardy, A., Sumegi, J., Csordas Toth, E., Gausz, J., Gyurkovics, H., Schedl, P. and Ish-Horowicz, D., (1982), Genomic organization and functional analysis of a deletion variant of the 87A7 heat shock locus of Drosophila melanogaster. J. Mol. Biol. 155:267-280

Vanin, E.F. (1985) Processed pseudogenes: characteristics and evolution. Ann. Rev. Genet. 19:253-272

Voellmy, R., Goldschmidt-Clermont, M., Southgate, R. and Tissieres, A., (1981), A DNA segment isolated from chromosomal site 67B in D. melanogaster contains four closely linked heat-shock genes. Cell 23:261-270

Weber, K. and Osborn, M., (1969), The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem. 244:4406-4412

Welch, W.J. and Feramisco, J.R., (1984), Nuclear and nuclealar localization of the 72,000 Dalton heat shock protein in heat-shocked mammalion cells. J. Biol. Chem. 259:4501-4513

Whalen, M. and Wilson, T.G., (1986), Variation and genomic localization of genes encoding Drosophila melanogaster male accessory gland proteins separated with SDS-PAGE. Genetics 114: 77-92

Wu, B., Hunt, C. and Morimoto, R., (1985), Structure and Expression of the human gene encoding major heat shock protein HSP70. Mol. Cell. Biol. 5:330-341

Yost, H.J. and Lindquist, S., (1986), RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis.

Zimmer, E.A., Martin, S.L., Beverly, S.M., Kan, Y.W. and A.C. Wilson. (1980) Rapid duplication and the loss of genes coding for the chains of haemoglobin. Proc. Nat. Acad. Sci. U.S.A. 77:2158-2162