CHAPTER 6

PRELIMINARY INVESTIGATION OF ELECTROPHORETICALLY CRYPTIC VARIATION AT THE *EST-2* LOCUS

6.1 INTRODUCTION

At the time zone electrophoresis was first used for the detection of allozyme variation it was acknowledged that the technique would detect only approximately one-third of amino-acid substitutions (Lewontin and Hubby, 1966). For many years electrophoretic surveys were made, and the data obtained were treated as true estimates of allelic frequencies. However, the application of additional techniques for the detection of primary sequence variation, such as heat denaturation, quickly revealed that electrophoretic "allele" classes were indeed heterogeneous groups of proteins (Bernstein *et al.*, 1973; Singh *et al.*, 1975). Simultaneously with the introduction of novel experimental techniques, the mathematical theory of selectively neutral and very slightly deleterious mutations was extended to include electrophoretically cryptic variation (King and Ohta, 1975; Ohta and Kimura, 1975). To stress the fact that electrophoretically detected mobility classes are actually phenotypes rather than genotypes, King and Ohta (1975) coined the term electromorph.

Attempts to measure directly the effects of selection on allozyme polymorphisms thus far have met with little success because of the many difficulties involved. Consequently, the majority of the debate between the neutralist and selectionist schools has been concerned with the nature of the statistical distributions of allele frequencies and heterozygosity. Clearly, if there is extensive electrophoretically cryptic variation, the observed distribution of these parameters, based on a single electrophoretic criterion, will need to be

drastically revised.

The urgent need for clarification of this problem has resulted in a variety of techniques being developed and applied to the search for cryptic variants. Methods which have been successfully employed in a number of species include heat denaturation (Bernstein *et al.*, 1973; Singh *et al.*, 1975), denaturation at differing concentrations of urea (Loukas *et al.*, 1981), molecular sieving imposed by varying concentrations of acrylamide in disc gels (Johnson, 1977) and sequential electrophoresis (Coyne, 1976; Singh *et al.*, 1976). The most systematic approach to the problem of cryptic variation has been that of Lewontin and co-workers with their programme of sequential electrophoresis. This involves the repeated testing of electromorph classes under a variety of buffers of differing pH and acrylamide gels of differing concentration.

6.1.1 Extra Allelic Variation

As a result of these studies several important observations have been made, and some potential generalisations have begun to emerge. The first observation is that all of the techniques mentioned above demonstrated a substantial increase in the number of alleles which could be detected for some loci. Possibly the most spectacular example is provided by the xanthine dehydrogenase (*Xdh*) locus of the sibling species pair *Drosophila pseudoobscura* and *D. persimilis*. Through the use of four different electrophoretic criteria and a heat stability test, Singh *et al.* (1976) uncovered 37 allelic classes in a sample of 146 chromosomes from 12 populations of *D. pseudoobscura*, where only six previously had been detected using their standard method of electrophoresis. An even more extreme result was obtained by Coyne (1976) for *D. persimilis*, where in a sample of only 60 chromosomes 23 alleles were found but only five previously were known.

From the studies which have been conducted to date a general pattern has been revealed (e.g. Singh, 1979; Loukas *et al.*, 1981), which may be summarised as follows:

(i) Highly polymorphic loci such as *Xdh* (Singh *et al.*, 1976), *Aldox* (Singh, 1979) and *Est-5* (Coyne *et al.*, 1978; Singh, 1979; Keith, 1983) in *D. pseudoobscura*, and *Est-5*, *Xdh* and *Aldox* of *D. subobscura* (Loukas *et al.*, 1981) all show large increases in the numbers of alleles found in natural populations.

(ii) Loci which are monomorphic, or which show only low levels of polymorphisms do not show large increases in the number of detectable alleles after the application of additional criteria. For example, in *D. pseudoobscura* the almost monomorphic locus *Odh* (Coyne and Felton, 1977) and several hexokinase loci (Beckenbach and Prakash, 1977) failed to reveal any significant increase in the number of alleles detectable. It has been suggested (e.g. Singh, 1979) that this apparent dichotomy of locus types reflects differential constraints on mutational alterations which are permissible in these two groups of loci, although the data are not yet extensive enough to generalise with any confidence. The *Adh* polymorphism in *D. melanogaster* provides an interesting and somewhat anomalous result when examined by sequential electrophoretic techniques (Kreitman, 1980). From a sample of 96 isochromosomal lines isolated from a single population, Kreitman was unable to detect any cryptic variation. This negative result did not appear to be attributable to a lack of sensitivity of the experimental technique, however, since Kreitman was able to distinguish two known thermostability variants by his sequential electrophoretic procedures. These lines had previously been indistinguishable from the standard electromorphs using conventional electrophoretic criteria. The *Adh* locus of *D. melanogaster* appears to be atypical, in that it is the only example examined to date in which a ubiquitous polymorphism has failed to harbour substantial electrophoretically cryptic primary sequence variation. The reason for this may be that strong negative, or purifying selection is acting on the ADH phenotype. Support for this idea has been provided recently with the nucleotide sequencing of a number of independently cloned *Adh* genes and their 5' and 3'

non-translated flanking sequences (Kreitman, 1983). In a sample of 11 cloned genes from five natural populations, Kreitman found 43 nucleotide polymorphisms, 13 of which were unique. However, although 14 of the 43 nucleotide differences were located in exon regions, only one of these was responsible for a difference in the amino-acid sequence of the protein product, and that was the alteration threonine *vs* lysine at codon 192 which results in the ubiquitous electrophoretic polymorphism.

6.1.2 Increased Differentiation Between Species and Local Populations

The application of additional criteria to the analysis of genetic variation has had a more profound effect than the simple increase of the number of alleles known at a locus. The extra variation discovered has greatly increased the genetic divergence between species (Bernstein *et al.*, 1973; Singh *et al.*, 1975; Coyne, 1976; Coyne *et al.*, 1979). These techniques also have been useful in revealing previously undetected local differentiation of populations within a species. The most striking example of this occurred in *D. pseudoobscura*, where the isolated population of Bogotá, Colombia appeared to be genetically very similar to the mainland populations of North America by conventional electrophoretic procedures (Prakash *et al.*, 1969). The application of sequential electrophoretic techniques quickly revealed that the Bogotá population was clearly differentiated from mainland populations at the *Xdh* (Singh *et al.*, 1976), *Adh-3* (Coyne and Felton, 1977) and *Est-5* (Coyne *et al.*, 1978) loci. There was also a suggestion of increased differentiation between some mainland populations of *D. pseudoobscura* for the *Xdh* and *Est-3* loci (Singh, 1979). These results clearly indicate that earlier observations regarding the genetic uniformity of populations within a species (e.g. Lewontin, 1974) can no longer be taken as correct. It is interesting, however, that application of extra criteria has failed to uncover any additional disequilibria between loci located on the same chromosome in *D. subobscura* (Loukas *et al.*, 1981).

6.1.3 Efficiency of Extra Techniques for the Recovery of Allelic Variation

The success of these more rigorous methods in uncovering previously cryptic variation automatically raises the question of what proportion of amino-acid substitutions the new techniques detect. Unfortunately the nucleotide sequence data for the *D. melanogaster Adh* genes cloned by Kreitman (1983) are not as useful in this context as might have been hoped, since the only base change leading to an amino-acid substitution in his sample was the one resulting in the routinely detected polymorphism. In an attempt to examine this problem more thoroughly at the protein level, Ramshaw *et al.* (1979) subjected a group of 20 haemoglobin variants, for which complete amino-acid sequence data were available, to their sequential electrophoretic programme. They found that 85% of the variants in this group were detected. Further experiments comparing groups of substitutions which were chemically identical, but located at different positions in the α- and β polypeptides, and comparing groups which were chemically different but charge equivalent at the same chain position suggested that their techniques were recovering 80-90% of the variants. While the generality of this result remains to be assessed by comparisons of other groups of proteins of known sequence, it does suggest that more sophisticated electrophoretic programmes may be capable of detecting quite a high proportion of those mutations which lead to an amino-acid substitution.

The apparent biochemical differences between allelic variants at the *Est-5* locus in *D. buzzatii*, reported in preceding chapters, make it important to determine whether there are substantial amounts of cryptic variation within electromorph classes at this locus. It was not possible within the limits of this thesis to undertake a comprehensive sequential electrophoretic analysis of this highly polymorphic locus. However, a preliminary investigation has been made of a sample of lines made isogenic for the two most common electromorph classes of this locus.

6.2 MATERIALS AND METHODS

6.2.1 Genetic Material

Drosophila bussarti were sampled from the Hemmant population (Locality 31 of Barker and Mulley, 1976), using fermented banana baits spread widely over the distribution of prickly pear plants in which the species breeds. This population has been sampled many times over different years and different seasons, and has consistently yielded substantial numbers of flies, suggesting that it harbours a relatively large and stable population of *D. bussarti*. Athough seven electromorph classes were known at the *Est-2* locus based on samples of Australian and South American populations (Barker, East and Sene, unpublished), the Hemmant population routinely exhibits five of these; $Est\text{-}2^{a}$, $Est\text{-}2^{b}$, $Est\text{-}2^{c}$, $Est\text{-}2^{d}$, and $Est\text{-}2^{e}$ at average frequencies of 0.490, 0.271, 0.173, 0.063, 0.002 respectively. As there was no balanced lethal marker stock available in *D. bussarti*, lines isogenic for the two *Est-2* electromorphs were established by the following crossing programme. A sample of wild-caught males were test-crossed individually to a stock homozygous for the $Est\text{-}2^{d}$ electromorph. As soon as larval activity was observed in the cultures, the male parents were analysed on the standard starch gel criterion of 11% w/v starch and Tris-borate-EDTA buffer, pH 8.0. Only those cultures in which the male parent was heterozygous $Est\text{-}2^{a/b}$ were kept for further development. The following protocol was employed:

(i) $Est\text{-}2^{a/b} \times Est\text{-}2^{d/d}$

(ii) Twenty single pair, full-sib matings were established between progeny from cross (i). Parents were assayed and only those matings which were $Est\text{-}2^{a/d} \times Est\text{-}2^{a/d}$ and $Est\text{-}2^{b/d} \times Est\text{-}2^{b/d}$ were kept. This established a group of lines containing the $Est\text{-}2^{a}$ electromorph, and a second group containing the $Est\text{-}2^{b}$ morph.

(iii) Within each set of 'a' and 'b' lines, twenty single pair full-sib matings were established between the progeny of cross (ii). When larvae were visible, parents were sacrificed and their genotypes determined.

(iv) Among the matings from cross (iii) a proportion were isogenic *Est-2*$^{a/a}$ x *Est-2*$^{a/a}$ or *Est-2*$^{b/b}$ x *Est-2*$^{b/b}$. For the majority of lines however, the crosses were between a homozygous and heterozygous individual. For these lines a further generation of twenty single pair full-sib matings were made to generate isogenic stocks.

(v) This four generation breeding programme yielded a sample of 19 lines isogenic for the *Est-2*^{a} electromorph and 20 lines isogenic for the *Est-2*^{b} morph.

This group of 49 isogenic lines were typed for their phenotype at the Est-1 locus, and karyotyped with respect to the polymorphic chromosome II inversion. Details of the genetic constitution of the lines are given in Table 6.1.

6.2.2 Electrophoretic Procedures

The EST-2 enzyme proved to be very refractory to the establishment of a broad range of electrophoretic criteria, in particular it was difficult to consistently maintain sharp bands. For the sequential electrophoretic analysis reported below the following set of five criteria was employed:

(i) Starch gels at a concentration of 11% w/v Connaught hydrolysed starch in Tris-borate-EDTA buffer pH 8.0 (Buffer III of Shaw and Prasad, 1970). Electrophoresis was for five hours at a constant voltage of 200V. This is the standard criterion used for population screening.

(ii) 2.5-16% linear gradient polyacrylamide gels in 0.1M Tris-borate-EDTA buffer pH 9.0. Electrophoresis was carried out for 2 hours

Table 6.1 Genetic constitution of the *Est-2* isogenic lines used in the cryptic variation survey

Line	*Est-2* genotype	*Est-1* genotype	Chromosome II Karyotype
1A	a/a	b/b	J
2A	a/a	b/b	J
3A	a/a	b/b	J
4A	a/a	b/b	J
5A	a/a	c/c	J
6A	a/a	b/b	J
7A	a/a	b/b	J
8A	a/a	x/x	S
9A	a/a	c/c	S
10A	a/a	b/b	S
11A	a/a	b/b	J
12A	a/a	b/b	J
13A	a/a	b/b	S
14A	a/a	b/b	J
15A	a/a	b/b	J
16A	a/a	b/b	J
17A	a/a	b/b	J
18A	a/a	b/b	J
19A	a/a	b/b	S
1B	b/b	n/n	J
2B	b/b	b/b	S
3B	b/b	b/b	S
4B	b/b	b/b	S
5B	b/b	n/n	J
6B	b/b	a/a	S
7B	b/b	b/b	S
8B	b/b	x/x	S
9B	b/b	b/b	S
10B	b/b	x/x	S
11B	b/b	c/c	S
12B	b/b	x/x	S
13B	b/b	c/c	S
14B	b/b	a/a	S
15B	b/b	b/b	J
16B	b/b	b/b	J
17B	b/b	b/b	J
18B	b/b	b/b	J
19B	b/b	b/b	J
20B	b/b	x/x	S

at 200 volts and approximately 20 milliamps per gel.

(iii) 2.5-16% linear gradient polyacrylamide gels in 0.1M Tris-borate-EDTA buffer pH 8.1. Electrophoresis was for 2 hours at 200 volts and approximately 20 milliamps per gel.

(iv) 2.5-16% linear gradient polyacrylamide gels in 0.1M Tris-borate-EDTA buffer ph 7.2. Electrophoresis was for 3 hours at 150 volts and approximately 15 milliamps per gel.

(v) Cellulose acetate plastic-backed plates (Titan III Zip Zone plates, Helena Laboratories, Beaumont, Texas, U.S.A.) in a complex amine-citric acid buffer, pH 6.2. Electrophoresis was for 4 hours at 250 volts and 6 milliamps per plate.

The fifth criterion is one of a set of several developed by Drs. J.G. Oakeshott and S. Easteal at the Australian National University, Canberra, A.C.T. for the sequential electrophoretic analysis of enzyme polymorphisms. Further details of the Helena cellulose acetate system are provided in Easteal and Boussy (1984).

6.2.3 Thermostability Studies

Several attempts were made to uncover electrophoretically cryptic variants by the application of a thermostability criterion. The methodology employed was that of Cochrane (1976) and Cochrane and Richmond (1979b). The following procedure was followed for the preparation of sample:

(i) 10 male and 10 female adults were homogenised in 200 µl of 0.1M phosphate, pH 7.0 containing 10% w/v sucrose.

(ii) Homogenate was centrifuged at 12,000 g for 15 minutes at 1^{o}C, and the resultant supernatant was divided into 4 x 20 µl aliquots.

(iii) Samples from (ii) were incubated at 50^{o}C ± 0.5^{o}C for 0, 3, 6 and 9 minutes.

(iv) After heat treatment samples were quenched immediately in an ice-water bath.

(v) 8 µl of treated sample were placed on 4 x 7 mm filter paper wicks such that each line x treatment group was represented in duplicate on any given origin of a standard starch gel. In this way, three lines could be compared on any single origin of a gel. Preliminary tests suggested that this procedure would provide a sensitive and efficient method for the analysis of thermostability variation. However, more detailed and carefully controlled studies reported in Section 6.3.2 indicated that the repeatability of this technique was very low.

6.3 RESULTS

The *Est-2*^{a} and *Est-2*^{b} electromorph classes were very clearly resolved on all five electrophoretic criteria, with the separation between the two being commonly 3-4 mm. An example of the amount of separation routinely observed between electromorph classes is shown in Fig. 6.1, where lines homozygous for the *Est-2*^{a}, *Est-2*^{b} and *Est-2*^{c} electromorphs are compared.

6.3.1 Cryptic Variants Revealed by Sequential Techniques

It quickly became apparent that some cryptic variation could be readily detected using the technique (Fig. 6.2). However, extensive testing of the 2.5-16% linear gradient polyacrylamide gels used in conditions (ii) to (iv) of the sequential programme outlined above has revealed that this is not likely to be an ideal electrophoretic support for the analysis of cryptic variation, even for a monomeric molecule the size of EST-2 with a molecular weight of only 54,000 daltons. It was found that, for any given pH criterion the same set of cryptic variants could be detected if the enzyme was allowed to proceed into the gel to an acrylamide concentration anywhere between 7 and 10 per cent. Beyond 10% however, the mobility differences were rapidly lost, presumably because of the greater retardation of molecules due to the reduced pore size at higher acrylamide concentrations.

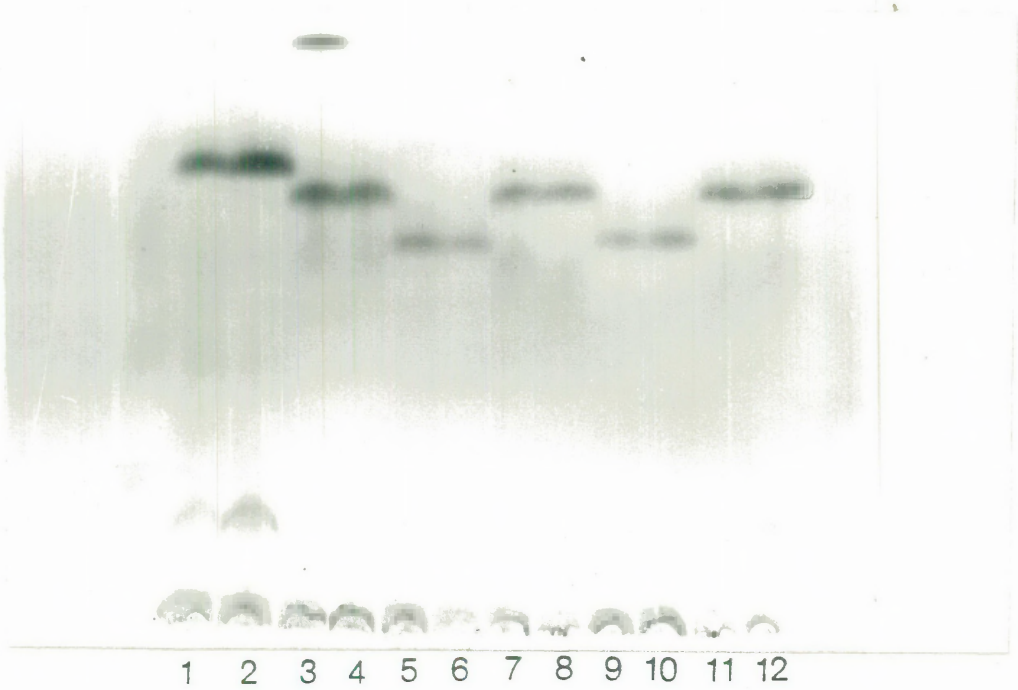

Figure 6.1 Sample gel showing mobility differences for electromorphs EST-2^a, EST-2^b and EST-2^c run under criterion (iii). Lanes 1,2 are EST-2^a, lanes 3,4,7,8,11,12 are EST-2^b and lanes 5,6,9,10 are EST-2^c. Electrophoretic matrix was 2.5-16% linear gradient polyacrylamide gel.

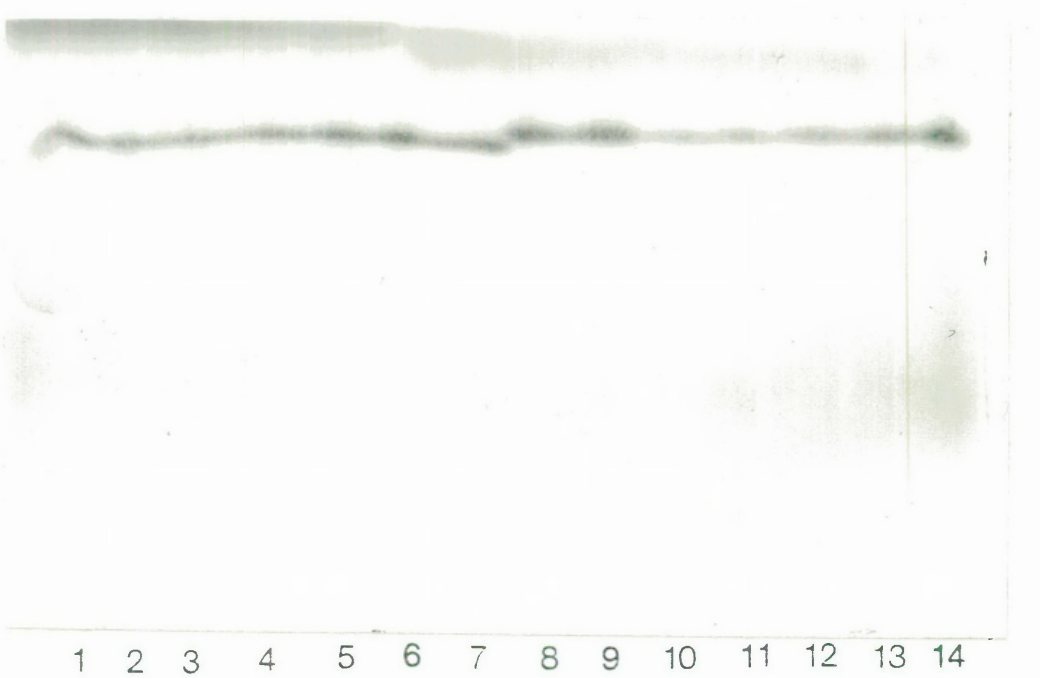

Figure 6.2 Sample gel showing recovery of cryptic variation within the EST-2^a electromorph under criterion (ii). Lanes 1,4,7 are line 1A, lanes 2,3 are line 2A, lanes 5,6 are line 3A, lanes 8,11,14 are line 8A and lanes 9,10,12,13 are line 9A. Electrophoretic matrix was 2.5-16% linear gradient polyacrylamide gel.

The acrylamide gels were run with 12 samples per gel and an arbitrarily chosen standard from each set of the 'a' and 'b' electromorph lines. Mobility was determined relative to this standard for each gel, and after each run the lines were re-grouped and compared against each other under the next condition of electrophoresis. The results of applying the three sequential acrylamide criteria (ii) to (iv) are presented in Tables 6.2 to 6.4, and the re-classification of the lines is presented in Table 6.5 using the nomenclature of Coyne *et al.* (1978). These results must be considered preliminary because it was not possible to exhaustively test and re-test each set of lines within a mobility class against themselves, or to test each class against all others. At this stage of the analysis the *Est-2*^{a} electromorphs had been resolved into four sub-classes, and the *Est-2*^{b} lines were divided into five sub-classes (Table 6.5).

Once this tentative classification had been achieved a genetic analysis of two cryptic variants was undertaken. Two lines, (16B and 2B) which had shown a relatively large mobility difference under criteria (ii) to (iv) were selected, and analysed by a back-cross experiment. The data are summarised in Table 6.6 and a typical gel showing the pattern of segregation in the first back-cross generation is shown in Figure 6.3. The results indicate that these two variants segregate as typical Mendelian alleles of a single locus. Although this preliminary sequential analysis revealed some apparent cryptic variation within both the *Est-2*^{a} and *Est-2*^{b} electromorph classes, and some of this variation has been shown to behave as simple Mendelian alleles of a single locus, a number of technical shortcomings of this method must be acknowledged. The greatest drawback to the use of the linear gradient acrylamide gels is the restricted distance of migration which can be utilised, beyond which the newly acquired resolution is lost again. Gradient acrylamide gels were chosen initially in preference to single concentration acrylamide because they consistently produced better resolution of the EST-2 isozyme. However, it was not realised at the time of starting

Table 6.2 Classification of *Esterase-2* electromorphs under sequential criterion (ii)

Line	Rm*	Line	Rm
1A	1.00	1B	1.00
2A	1.00	2B	1.00
3A	1.00	3B	1.00
4A	1.00	4B	1.00
5A	1.00	5B	1.00
6A	1.00	6B	1.00
7A	1.00	7B	1.00
8A	1.02	8B	1.00
9A	1.02	9B	1.00
10A	1.00	10B	1.00
12A	1.00	11B	1.00
13A	1.00	12B	1.00
14A	1.00	13B	1.00
15A	1.00	14B	1.00
16A	1.00	15B	1.00
17A	1.00	16B	1.04
18A	1.00	17B	1.02
19A	1.00	18B	1.02
		19B	1.02
		20B	1.00

*Rm is the mobility relative to arbitrarily selected standards; lines 2A and 1B for the *Est-2*^{a} and *Est-2*^{b} electromorphs respectively.

Table 6.3 Classification of *Esterase-2* electromorphs under sequential criterion (iii)

Line*	Rm±	Line	Rm
1A	1.00	1B	1.00
2A	1.00	2B	1.00
3A	1.00	3B	1.00
4A	1.00	4B	1.00
5A	1.00	5B	1.00
6A	1.00	6B	1.00
7A	0.98	7B	1.00
10A	1.02	8B	1.00
12A	1.00	9B	1.00
13A	1.00	10B	1.00
14A	1.00	11B	1.00
15A	1.00	12B	1.00
16A	1.00	13B	1.00
17A	1.00	14B	1.00
18A	1.00	15B	1.02
19A	1.02	20B	1.00
8A	1.02	17B	1.02
9A	1.02	18B	1.02
		19B	1.04
		16B	1.04

* Lines were regrouped according to their classification under criterion *ii).

± Rm is the mobility relative to lines 2A and 1B for the *Est-2*^{a} and *Est-2*^{b} electromorphs respectively.

Table 6.4 Classification of *Esterase-2* electromorphs under sequential criterion (iv)

Line*	Rm±	Line	Rm
7A	1.00	1B	1.00
1A	1.00	2B	1.00
2A	1.00	3B	1.00
3A	1.00	4B	1.00
4A	1.00	5B	1.00
5A	1.00	6B	1.00
6A	1.00	7B	1.00
12A	1.00	8B	1.00
13A	1.00	9B	1.00
14A	1.00	10B	1.00
15A	1.00	11B	1.00
16A	1.00	12B	1.00
17A	1.00	13B	1.00
18A	1.00	14B	1.00
		20B	1.00
10A	1.00		
19A	1.00	15B	1.00
8A	1.00	17B	1.00
9A	1.00	18B	1.00
		19B	1.02
		16B	1.02

*Lines were regrouped according to their classification under criteria (ii) and (iii).

±Rm is the mobility relative to lines 2A and 1B for the *Est-2*^{a} and *Est-2*^{b} electromorphs respectively.

Table 6.5 Allocation of *Esterase-2* isogenic lines to electromorph classes after the sequential application of electrophoretic criteria (i) to (iv)

Electromorph[*]	Number of lines
a/1.00/1.00/1.00	13
a/1.02/1.02/1.00	2
a/1.00/1.02/1.00	2
a/1.00/0.98/1.00	1
b/1.00/1.00/1.00	15
b/1.04/1.04/1.02	1
b/1.02/1.04/1.02	1
b/1.00/1.02/1.00	1
b/1.02/1.02/1.00	2

[*] Electromorphs are classified as w/x/y/z, where w is classification for criterion (i), x for criterion (ii) y for criterion (iii) and z for criterion (iv).

Table 6.6 Genetic analysis of cryptic variants of *Est-2*^{b} electromorph under the discriminatory criterion (ii)

a. Crossing programme

$Est\text{-}2^{1.04/1.04}$ x $Est\text{-}2^{1.00/1.00}$

↓

$Est\text{-}2^{1.04/1.00}$ x $Est\text{-}2^{1.04/1.04}$

↓

Progeny collected and analysed under electrophoretic criterion (ii)

b. Data analysis

	$Est\text{-}2^{1.04/1.04}$	$Est\text{-}2^{1.04/1.00}$
Observed	54	44
Expected	49	49

χ^2_1 = 1.02 : not significant at α = 0.05

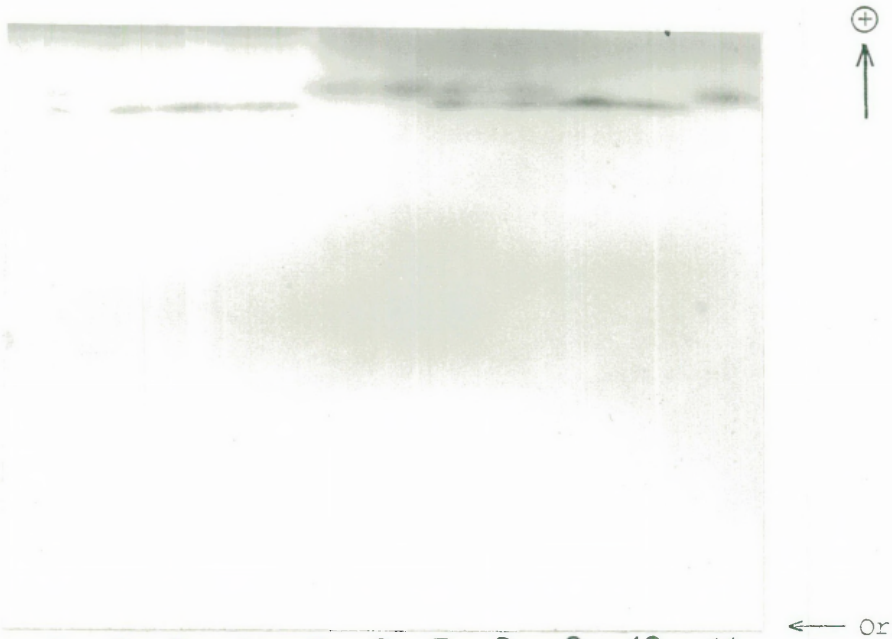

Figure 6.3 Gel showing progeny of backcross analysis of the cryptic variation in lines 2B and 16B under criterion (ii). Lanes 4,10 are line 2B controls, lanes 5,6,11 are line 16B controls, remainder are test-cross progeny. Electrophoretic matrix was 2.5-16% linear gradient polyacrylamide gel.

the study that the sensitivity of the sequential technique for the analysis of cryptic variants is, in large part, dependent on long running times to draw out the minute differences between lines. Another major difficulty in using these gels is that only 12 samples can be run per origin, and given the very large numbers of comparisons which need to be made to unambiguously determine that any group of lines is identical for a given electrophoretic condition, the time and expense would become prohibitive. At the end of the analyses reported above I was by no means convinced that sufficient cross-checking had been done to unambiguously define the new set of putative electromorphic classes.

At this stage, the fifth criterion employing cellulose acetate plates, was added. The relative speed and cheapness of this technique allowed exhaustive comparisons to be made between pairs and triplets of lines, and the resulting re-classification of lines was completely unambiguous. The re-grouping of lines under this criterion is given in Table 6.7, and a typical gel showing the recovery of a cryptic variant from the *Est-2*^{a} electromorph class is given in Fig. 6.4. In the main, the differences in electrophoretic mobility between cryptic variants were very small, and their reliable detection was dependent on repeated testing. However, when more extreme classes were compared at the end of the analysis it was clear that mobility differences could be quite large (Fig. 6.5). If the sequential criteria (ii) to (iv) are disregarded and only criterion (v) is used, the *Est-2*^{a} electromorph is broken down into 3 groups and the *Est-2*^{b} morph becomes 4 distinguishable groups.

6.3.2 Thermostability Tests for Electrophoretically Cryptic Variants

Several attempts were made to analyse electrophoretically cryptic thermostability variation using the methods of Cochrane (1976) and Cochrane and Richmond (1979b). These tests must be considered a total failure, since they were not repeatable from one experiment to the next. The reasons for

Table 6.7 Classification of *Esterase-2* electromorphs under sequential criterion (v)

Line[a]	Mobility group	Line	Mobility group
1A	3	1B	3
2A	2	2B	4
3A	2	3B	4
4A	2	4B	4
5A	2	5B	3
6A	2	6B	4
7A	3	7B	4
8A	1	8B	4
9A	1	9B	4
10A	1	10B	4
12A	3	11B	4
13A	1	12B	4
14A	1	13B	4
15A	2	14B	4
16A	2	15B	2
17A	2	16B	1
18A	2	17B	3
19A	1	18B	3
		19B	1
		20B	4

[a] Lines were assigned to mobility groups within the *Est-2*^{a} and *Est-2*^{b} electromorph classes on the basis of distance of migration towards the anode. *i.e.* group 1 is more anodal than group 2 *etc.*

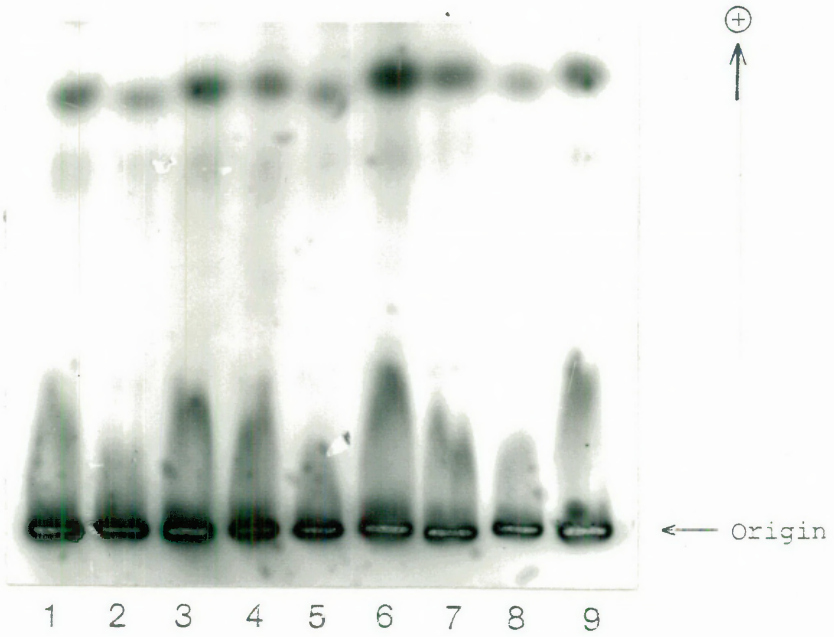

Figure 6.4 Cellulose acetate plate showing recovery of cryptic variation within the EST-2^a electromorph under criterion (v). Lanes 1,4,7 are line 13A, lanes 2, 5,8 are line 15A, lanes 3,6,9 are line 14A.

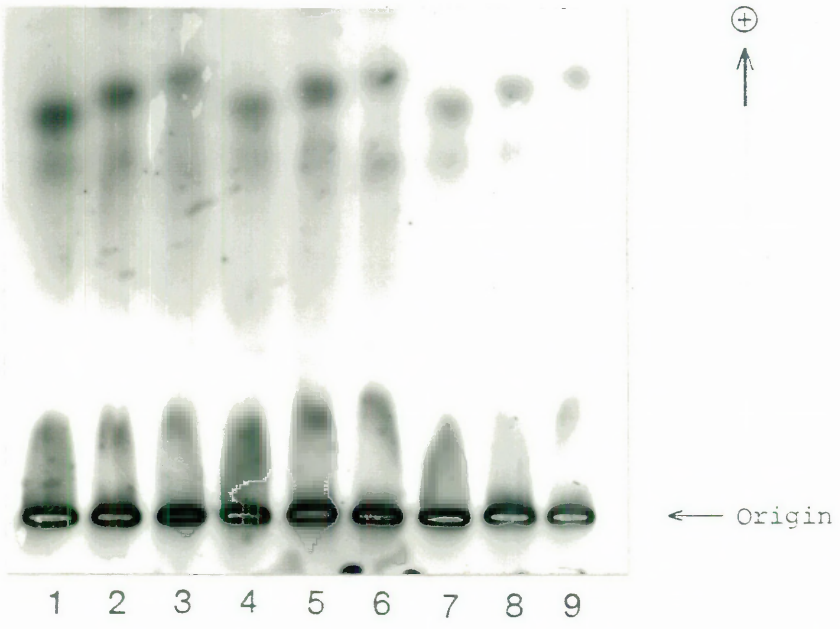

Figure 6.5 Cellulose acetate plate showing mobility differences between two cryptic variants of the EST-2^a electromorph, and an EST-2^b line. Lanes 1,4,7 are line 1B, lanes 2,5,8 are line 4A, lanes 3,6,9 are line 14A.

this lack of repeatability are manifold, at least four factors being shown to be relevant in the course of these investigations:

(i) The age of the flies used for the tests. It was found that EST-2 enzyme from flies 2-3 days of age was substantially more thermostable than enzyme extracted from flies 7-8 days of age. Therefore, flies must be carefully standardised for age before any thermostability tests are made.

(ii) The concentration of protein in the extract appears to exert a significant effect on EST-2 stability. If extracts were prepared at one fly per 10 µl of homogenising buffer the enzyme was much more stable than if flies were ground up at one fly per 20 µl of buffer. Furthermore, apparent thermostability variation could be detected at one fly per 10 µl, but not at one fly per 20 µl of buffer, where most activity had disappeared after 6 minutes of incubation.

(iii) The nutritional status of the flies appeared to be a contributing factor. ESTERASE-2 enzyme extracted from flies reared from uncrowded cultures appeared to be appreciably more thermostable than enzyme from flies raised under crowded conditions. This apparent effect of culture density may be wholly or partially confounded with observation (ii) above, since flies from crowded cultures are much smaller than those reared under uncrowded conditions.

(iv) Finally, the buffer employed for enzyme extraction was shown to be a critical factor. If flies were homogenised in 0.1M phosphate buffer, pH 7.0, no differences in thermostability could be detected between any of the isogenic lines, (Fig. 6.6a). However, if flies from the same cultures of the same lines were homogenised in buffer containing 1.5M urea, apparent thermostability differences were found (Fig. 6.6b).

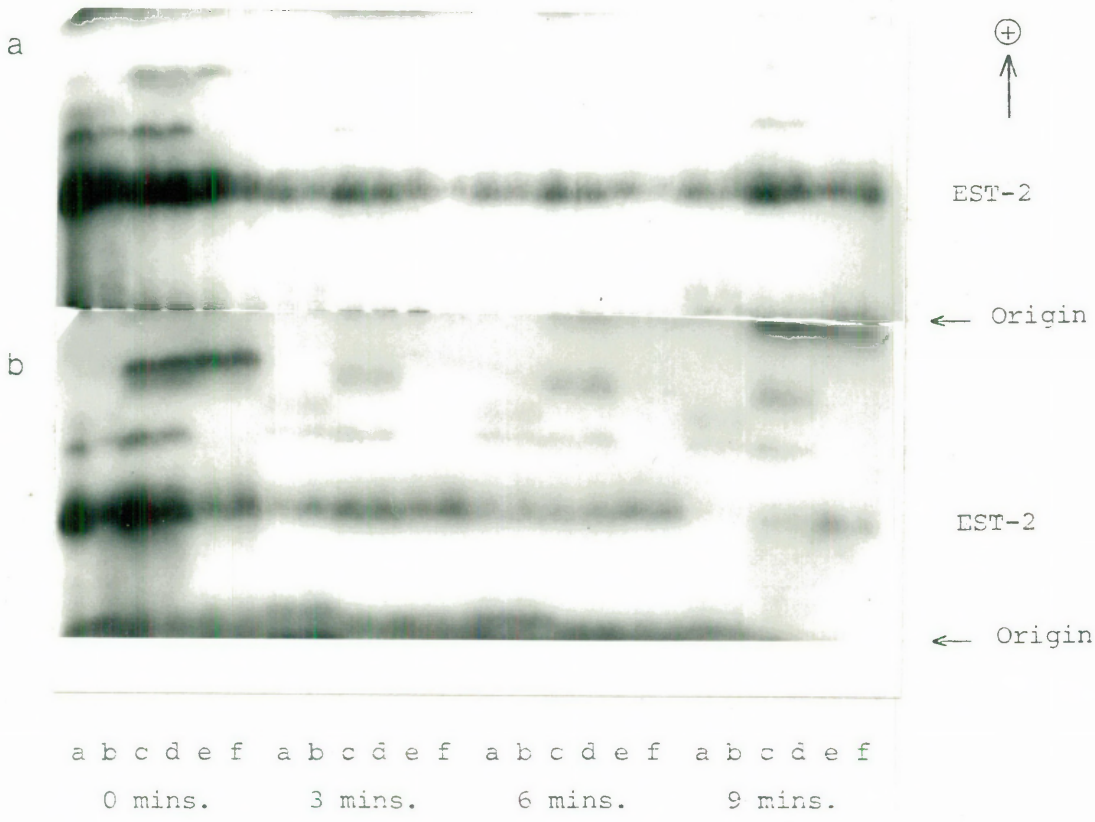

Figure 6.6 Analysis of thermostability variation among EST-2^{b} lines. (a) Flies homogenised in 0.1M phosphate buffer pH 7.0. (b) Flies homogenised in 0.1M phosphate buffer pH 7.0 containing 1.5M urea. Origins shown in Figs. 6.6a and 6.6b were run on the same gel under electrophoretic criterion (i). Line 15B, a and b; line 16B c and d; line 17B, e and f; heated at 50^{o}C for the times indicated.

The combined effects of so many factors apparently affecting the thermostability of EST-2 extracts renders this method of analysis practically useless for the study of cryptic variation for this enzyme.

6.4 DISCUSSION

The results presented above must be treated as only a preliminary investigation of the presence or absence of electrophoretically cryptic variation at the *Est-2* locus in *D. buzzatii*. I am not completely satisfied that the classification of electromorph classes under sequential criteria (i) to (iv) has been subjected to a sufficiently rigorous examination either for lines within apparently homogeneous classes, or for cross comparison of apparently different groups. I have no similar reservations concerning the data obtained under criterion (v), however, since it was possible to make all the necessary comparisons. Accordingly, the data may tentatively be treated as upper and lower estimates of the amount of extant cryptic variation which can be detected by sequential procedures. As a thoroughly tested lower estimate there are three sub-classes within the $Est\text{-}2^{a}$ electromorph and four classes within the $Est\text{-}2^{b}$ electromorph (Table 6.7). As an incompletely tested upper estimate, based on criteria (i) to (v) there are six classes for the $Est\text{-}2^{a}$ electromorph and a further six for the $Est\text{-}2^{b}$ electromorph (Table 6.8). Thus, the two electromorph classes $Est\text{-}2^{a}$ and $Est\text{-}2^{b}$ comprise at least seven and possibly up to 12 distinct electromorph groups, an increase of three to six-fold in the number of alleles. An increase in the number of allelic classes of this magnitude is entirely consistent with previously published data for other highly polymorphic loci. For example, in a sample of 237 lines from two populations of *D. pseudoobscura*, Keith (1983) detected 41 electromorph classes where only 12 previously were detected using a single electrophoretic criterion for the *Est-5* locus. Similar situations apply for the highly polymorphic *Adh* (Singh *et al.*, 1976) and *Ao* (Singh, 1979) loci of *D. pseudoobscura*.

Table 6.8 An upper estimate of the extent of electrophoretically cryptic variation at the Est-2 locus under five sequentially applied criteria

Electromorph class*	Number of lines
a/1.00/1.00/1.00/1	2
a/1.00/1.00/1.00/2	9
a/1.00/1.00/1.00/3	2
a/1.02/1.02/1.00/1	2
a/1.00/1.02/1.00/1	2
a/1.00/0.98/1.00/3	1
b/1.00/1.00/1.00/3	2
b/1.00/1.00/1.00/4	13
b/1.04/1.04/1.02/1	1
b/1.02/1.04/1.02/1	1
b/1.02/1.02/1.00/3	3
b/1.00/1.02/1.00/2	2

*Electromorph classification is the same as for Table 6.5 with the addition of a fifth mobility ranking for criterion (v).

Another feature which these data for *Est-2* of *D. buzzatii* share with the results for highly polymorphic loci in other species of *Drosophila* is the apparent existence of a modal class within an original electromorph group. In this instance, 9 of 18 lines of the $Est\text{-}2^{a}$ electromorph, and 13 of 20 lines of the $Est\text{-}2^{b}$ electromorph remain indistinguishable after sequential analysis (Table 6.8) Keith (1983) observed a similar phenomenon in her study of the *Est-5* polymorphism of *D. pseudoobscura* in which two common electromorphs detected under the standard electrophoretic criterion were each found to harbour a prevalent class after the addition of four extra criteria.

A further aspect to emerge from these data, which has not yet been widely investigated in other species, is the effect which the recovery of cryptic variants has on estimates of gametic disequilibria, both with other polymorphic loci, and also with polymorphic inversions. The parental population from which this set of isogenic lines were derived showed substantial disequilibrium between the *Est-2* locus and the polymorphic J inversion, both at the time these lines were sampled (East, unpublished) and in a subsequent sample (Knibb, East and Barker, unpublished). The *Est-1* locus also showed disequilibrium with this inversion, but no disequilibrium could be detected between the two *Esterase* loci. Although there was a highly significant disequilibrium between electromorph classes of the *Est-2* locus and the S and J arrangements of chromosome II, there were no fixed differences between electromorph classes, as can be seen in Table 6.1 where $Est\text{-}2^{a}$ and $Est\text{-}2^{b}$ lines both are represented in S and J chromosomes. However, if these data are re-examined using only the conservative estimate of cryptic variation provided under criterion (v), it is clear that gametic disequilibrium is now close to fixed differences between electromorph classes (Table 6.9). Obviously if we hope to assess the contribution of polymorphic inversions to allozyme frequencies in natural populations, it will be critical to obtain more

Table 6.9 The chromosome II constitution of the *Est-2* isogenic lines after reclassification according to sequential criterion (v)

Line	Mobility class	Chromosome II karyotype
8A	1	S
9A	1	S
10A	1	S
13A	1	S
14A	1	J
19A	1	S
2A	2	J
3A	2	J
4A	2	J
5A	2	J
6A	2	J
15A	2	J
16A	2	J
17A	2	J
18A	2	J
1A	3	J
7A	3	J
12A	3	J
16B	1	J
19B	1	J
15B	2	J
1B	3	J
5B	3	J
17B	3	J
18B	3	J
2B	4	S
3B	4	S
4B	4	S
6B	4	S
7B	4	S
8B	4	S
9B	4	S
10B	4	S
11B	4	S
12B	4	S
13B	4	S
14B	4	S
20B	4	S

precise information concerning the extent of cryptic variation at the enzyme loci being studied. There is, as yet, little empirical evidence to indicate the effect which the presence of cryptic variants will have on studies of disequilibria. Norman and Prakash (1980) found that third chromosome inversions in *D. pseudoobscura* and *D. persimilis* were actually associated with different alleles of the amylase locus, where previously they were believed to share the same allele. In contrast to this, Loukas *et al.* (1981) failed to uncover any previously undetected disequilibria between allozyme loci in *D. subobscura* when additional techniques were used to assess cryptic variation. Clearly this question requires a much closer analysis, and the *Est-1* and *Est-2* loci of *D. bussatii* which appear to be in disequilibrium with a polymorphic inversion, but not with each other, provide an excellent system for further investigation.

The results presented in Section 6.3 indicate that, when a thorough study is made, extensive extra variation will be found at the *Est-2* locus. Such a study must await the availability of a balanced lethal second chromosome in *D. bussatii*, so that large numbers of isogenic lines can be produced rapidly. In the interim, the significance of this enormous amount of variation must remain a matter for conjecture.

Singh (1979) reviewed the structural and functional components of genic variation and the contribution of studies of cryptic variants. One frequent observation is that there are more rare variants than expected under models of neutral mutation and random genetic drift. It would be exceedingly difficult to conceive a model of balancing selection which could account for the very large numbers of variants observed, and a possible explanation which must be considered for these highly polymorphic loci is that some proportion of the variants are being generated by intragenic recombination. The rapidly spreading application of recombinant DNA technology may provide some relevant information on this question in the near future.

Although selectionist models could not hope to account for the newly discovered amounts of genic variation, the data obtained to date do not fit well with neutralist models either. The theories of King and Ohta (1975) and Ohta and Kimura (1975) predict J-shaped distributions of alleles, with a common 'type' allele and many other rarer alleles. The fitness of alleles is predicted to decrease with increasing mutational divergence from the 'type' allele, and rarer electromorph classes are predicted to harbour less cryptic variants than more prevalent electromorphs. Pertinent data are scarce at this stage, but Singh (1979) found that common electromorphs did not necessarily have more cryptic variants than rarer classes. The *Est-2* polymorphism of *D. bussatii* will provide an excellent opportunity to investigate this further because of the ubiquitous occurrence of four relatively common electromorph classes at this locus.

CHAPTER 7

SUMMARY AND CONCLUDING COMMENTS

The studies which have been reported in this thesis were undertaken in the specific context of a multidisciplinary analysis of allozyme variation in the cactophilic Drosophilid *D. buzzatii*. The two polymorphic esterase isozymes were chosen for a variety of reasons; the principal factor being the abundant circumstantial evidence which suggested that the genetic variation at the loci coding for these enzymes was not neutral with respect to selection (Section 1.2.5). My other major reason for undertaking a biochemical study of the esterases was simply the fact that they are genetically interesting proteins. The observation that esterases tend to be ubiquitously polymorphic in *Drosophila* species (Powell, 1975) coupled with the facility with which ester hydrolase enzymes have acquired new functions in the course of evolution (Hartley, 1979) suggested that they might represent an excellent opportunity to assess the relative roles of isozymic and allozymic variation in adaptation and microevolution (Section 1.1.2).

The principal problem confronting the study of esterases is the fact that their *in vivo* function is unknown (Section 1.2.3). The first objective of this work therefore was to undertake a preliminary study of some physiological features of the EST-1 and EST-2 isozymes (Chapter 2). The most significant observations to come from this work are that the two predominant adult isozymes EST-1 and EST-2 are characterised by completely different ontogenic and tissue distributions, and that there is another major β-esterase (EST-J) which has a very limited period of ontogenic expression. These data have been useful for establishing potential homologies with esterases of other *Drosophila* species, in particular the two

β-esterases EST-1 and EST-J would appear to be homologues of the duplicated β-esterase loci in other members of the *mulleri* subgroup described by Zouros *et al.* (1982). These β-esterases are likely to provide an excellent model system for the molecular analysis of duplicated isozyme loci, since they have diverged sufficiently to exhibit very different patterns of developmental and tissue-specific expression, but are not so different that the protein products can no longer form an intermolecular hybrid. These loci are particularly attractive for molecular studies of gene regulatory alterations in evolution, not only because of the duplication, but also because the duplicated loci appear to be differently expressed in members of the *mulleri* subgroup (Zouros *et al.*, 1982).

The comparison of the EST-1 and EST-2 isozymes was continued in Chapter 3, where some general biochemical properties were considered. One of the first observations in this study was the apparent failure of the *D. buzzatii* esterases to conform closely to the classification system laid down for mammalian esterases. This has been found for esterases of other insect species, as noted in Chapter 1 (Section 1.2.4.2). The main conclusion to be drawn from the experiments of Chapter 3 is that, if the *Est-1* and *Est-2* loci evolved by gene duplication, then they have now diverged in function to the extent that every property examined was different for the two enzyme products. Substantial similarities in properties were found between the EST-1 enzyme of *D. buzzatii* and the β-esterase of *D. virilis* (Table 3.9), and EST-2 and the α-esterase of *D. virilis* (Table 3.10) suggesting that the functions of the α-esterases and of the β-esterases may have been conserved for a lengthy period of evolution. It was further noted that major α- and β-esterases are found in species of the sub-genus *Sophophora*, with molecular weights similar to those reported for α- and β-esterases of the *Drosophila* sub-genus, suggesting that if the α- and β-esterases arose originally as a duplication, it occurred prior to the division of these two branches of the *Drosophilidae*.

Once the differences between the isozymes had been compared, attention was turned to some of the allozymic variants produced by the *Est-1* and *Est-2* loci (Chapter 4). Although some differences between two allozymes of the *Est-1* locus could be detected, these were never large, and one isozyme almost invariably was superior in terms of catalytic activity. In contrast to this, substantial differences were detected between four variants produced by the *Est-2* locus, and the relative activity of the enzymes changed according to the substrate examined. These differences between EST-2 allozymes were examined further by a kinetic analysis involving the environmental variables temperature, pH and substrate type (Chapter 5), and again significant differences between allozyme variants were found. It is interesting that the allozyme encoded by the $Est\text{-}2^{C}$ allele proved to be consistently different in its biochemical properties from the other three EST-2 allozymes, since it is the $Est\text{-}2^{C}$ allele which is most conspicuous in other studies of allozyme polymorphism in *D. bussatii* (Mulley et al., 1979).

Perhaps the potentially most important factor to emerge from these experiments is the apparent difference in substrate specificity of EST-2 allozymes. This observation must seriously raise the possibility that, for the *Est-2* locus at least, environmental heterogeneity manifested as substrate variability, may be an important factor in the maintenance of the *Esterase* polymorphisms in *D. bussatii*. Clearly, without more detailed information concerning the *in vivo* function of the EST-2 enzyme, this will be a difficult hypothesis to test. Nevertheless, a number of observations following from this study allow the construction of a useful model to examine this suggestion. The experiments reported in Chapters 4 and 5 were suggestive of differences in substrate specificity between allozymes, especially with respect to acyl carbon chain length. Cactophilic yeasts appear to differ in their capacity to produce volatiles (East, 1982 and unpublished data), and cactophilic yeast species are known to vary in their temporal and spatial distribution in

Australia (Barker *et al.*, 1983). Other work with yeasts used in the brewing industry has shown that yeast esterases can synthesise esters (Parkkinen and Suomalainen, 1982), and that higher molecular weight esters such as ethyl caprylate are sequestered in the yeast cell (Nykanen *et al.*, 1977). Some of these esters may well be toxic to *Drosophila* at relatively low concentrations. For example, ethyl caprylate incorporated into standard food medium at a concentration of 1% v/v resulted in greater than 90% mortality to larvae placed in this food (East, unpublished). Clearly this system contains all the components required for an analysis of genotype-environment interaction in a laboratory context. The efficacy of this approach will be dependant on the demonstration that EST-2 is capable of hydrolysing esters such as ethyl caproate and ethyl caprylate *in vivo*, and that these esters, or others similar to them are produced by the cactophilic yeasts, and are present in natural rots. Whilst this is clearly rather speculative, it is nonetheless the logical synthesis of the results reported in this thesis with other published information. It represents a preliminary attempt to fulfil the four steps outlined by Clarke (1975) as an experimental strategy for testing the adaptive significance of allozymic variation. A single study such as this, beginning with no prior knowledge of the biochemistry or physiology of the enzymes could not possibly hope to exhaustively follow Clarke's protocol, but I believe that we are now in a position to ask more critical questions.

The results presented in Chapter 6 of this thesis, although representing only a preliminary analysis of electrophoretically cryptic variation at the *Est-2* locus, nonetheless have significant implications for future research. Firstly, the fact that there is extensive electrophoretically cryptic variation at this locus casts some doubt over the generality of the biochemical properties of allozymic variants reported in Chapter 4 and 5, since those studies were conducted on enzyme material isolated from single isofemale lines. Whether these data will be of specific predictive value in helping to explain the distribution of electromorph frequencies in natural populations therefore

seems dubious. However, I believe that the data are of some value in identifying possible components of the biotic and abiotic environments of *D. buzzatii* which may be influencing allele frequencies at these highly polymorphic loci. Chemical analysis of the metabolic by-products of cactophilic yeasts, and the chemical perturbation of synthetic laboratory culture media should provide significant insight into the nature of genetic polymorphism at these loci.

The second implication of the results presented in Chapter 6 is for the study of the population dynamics of electromorph variation in natural populations of *D. buzzatii*. It is clear from the data presented in Table 6.9 that the recovery of electrophoretically cryptic variants at the *Est-2* locus has drastically altered the nature of the gametic disequilibrium between the *Est-2* electromorphs and polymorphic inversions on the second chromosome of this species. It now becomes critical to establish a simple and practical protocol for the analysis of cryptic variation at the *Est-1* and *Est-2* loci, and then to examine in detail the nature of the spatial distribution of cryptic variants on large samples of second chromosomes from at least two populations, and the effect which this has on disequilibria between these loci and polymorphic inversions. This should be a significant step in helping to determine whether any putative selection operating on these loci is simply the result of disequilibrium with inversion complexes.