CHAPTER 4

INVESTIGATION OF PUTATIVE MAJOR GENE

The experimental results presented so far in this thesis do not provide any firm, direct evidence supporting the Shifting Balance Theory. However, as has been mentioned several times, results for Treatment D, block 1, where an extreme phenotype appeared in a single deme and was subsequently spread throughout the system of demes, may be relevant to the theory. For this reason, some effort was made to understand the genetic basis of these results.

This chapter includes descriptions of the occurrence and behaviour under selection of this phenotype, and of attempts made to ascertain its genetic control (if any). No separate review of literature concerning major genes, genetic analysis of selection lines etc. will be presented in this chapter; rather, appropriate references will be discussed within each section of the work.

4.1. Description of the Extreme Phenotype Under Selection

The treatment in which this effect appeared consisted a population subdivided into 10 small population units, or demes, each of 5 pairs of parents each generation. Selection was initially on a within-deme basis each generation, followed by migration between demes, at an overall rate of 10% (i.e. 10% of selected adults each generation were migrants) imposed such that migrants moved from phenotypically superior demes to inferior ones, every generation. It was hoped that if any significant favourable epistatic systems were "discovered" through the action of drift, they would be multiplied within their deme of origin, then spread to other demes through the selective migration process. This mode of diffusion of favourable epistatic systems could operate just as, or more, effectively, for major genes affecting the selected trait. Such major genes could have existed in the population at low frequencies, or arisen via mutation at any stage in the selection period (Hill,1982).

This major effect was first noticed in deme 4 of this treatment at G_{13} of selection. In the preceding generations, this deme had received one immigrant at G_{11}, and again at G_{12}; prior to that it had not received genetic imputs from other demes since G_5. Neither source deme at G_{11} and G_{12} contained any extreme individuals (phenotypically). The phenotypic distribution for deme 4 at G_{12} approximated a normal distribution reasonably well, with no outliers on the curve of either males or females, at either extreme of the distribution. In G_{13}, a number of extremely large individuals were noted. At this stage in selection, the overall female mean was c. 1.85 mg (the starting point having been 1.40 mg). Amongst the extreme individuals recorded, 5 had bodyweights in excess of 2.50 mg, with the heaviest being over 3.00 mg. These weights are 8-10 standard deviations from the original mean, and approximately 5 standard deviations heavier than the mean after 12 generations of selection.

Morphologically, these extreme individuals were characterised by generally increased dimensions, but most noticeable were very swollen abdomens, which were filled with clear fluid. An immediate response to their appearance was to check for viral or microsporidial infection; however, no trace of either agent was found (P. Christian, pers. comm.). A second possibility was that the extreme phenotype was produced by a single mutation such as Giant, or Lethal Translucida (Lindsley and Grell, 1967). However, none of those checked seemed to fit the observed morphology and development. The principal discrepancies were that there were no obvious effects on larval development, in terms of either morphology or time period, and the adults appeared reasonably viable.

All the female parents selected in deme 4 at G_{13} (i.e. to be parents of G_{14}) showed the extreme phenotype to some extent, but none of the males selected were extreme (Fig. 4.1, over page). In G_{14}, the number of extreme phenotype females was reduced, with only 3 individuals appearing outside the main distribution. The male distribution was however different from the previous generations, with 5 outliers appearing. As with the extreme females, these individuals were generally somewhat larger than "normal" selected flies, and had swollen, lengthened abdomens. In the following generations, the

stretched form of the distributions was maintained: an apparently normally distributed cluster, then a varying number of extreme individuals at the heavy end of the range.

The effect of these individuals on migration patterns was immediate. At G_{13}, deme 4 contributed 3 females to other demes (having received 1 at G_{12}), and in the following generations contributed successively 5,5,3,4,4,6,6, and 2 individuals. From G_{23} to G_{28}, deme 4 was a recipient of immigrants every generation, but from that point to the end of selection at G_{34}, was again a contributor of migrants at every generation. It should be noted that, even assuming all selected individuals contributed equally to progeny numbers in every generation, which seems quite unlikely, the level of inbreeding in deme 4 after 10 generations of not receiving any immigrants (from G_{13} to G_{22}) would have been approximately 0.40.

The migration model used in this treatment was designed to allow for the spread of favourable interaction systems from source demes to neighbours. It is of interest thereby to observe the patterns of migration from deme 4 (and later from other demes containing the effect) and its relationship to the pattern of appearance of the effect in other demes. Table 4.1 (page 114) shows the pattern of migration from demes that had previously exhibited the effect, and its pattern of occurrence throughout the deme system. There were marked differences between demes in the number of immigrants received from demes containing the effect needed to "introduce" it to the host deme. This number ranged from only 2 in the case of deme 3 to 13 in the case of deme 1. In deme 5, the effect initially appeared after receipt of 3 large migrants, but a further 4 were required for the effect to become established. As well as the number of migrants received from demes containing the effect, the proportion that these represent of all migrants received over the period of influx into the deme might affect the rate of establishment of the effect in new demes. To examine this possibility, the regression of number of generations after first receiving affected migrants before establishment of the effect, on the proportion of affected migrants/total migrants received

FIG. 4.1: Frequency Distributions, Deme 4,
Treatment D, Block 1 (Females only)

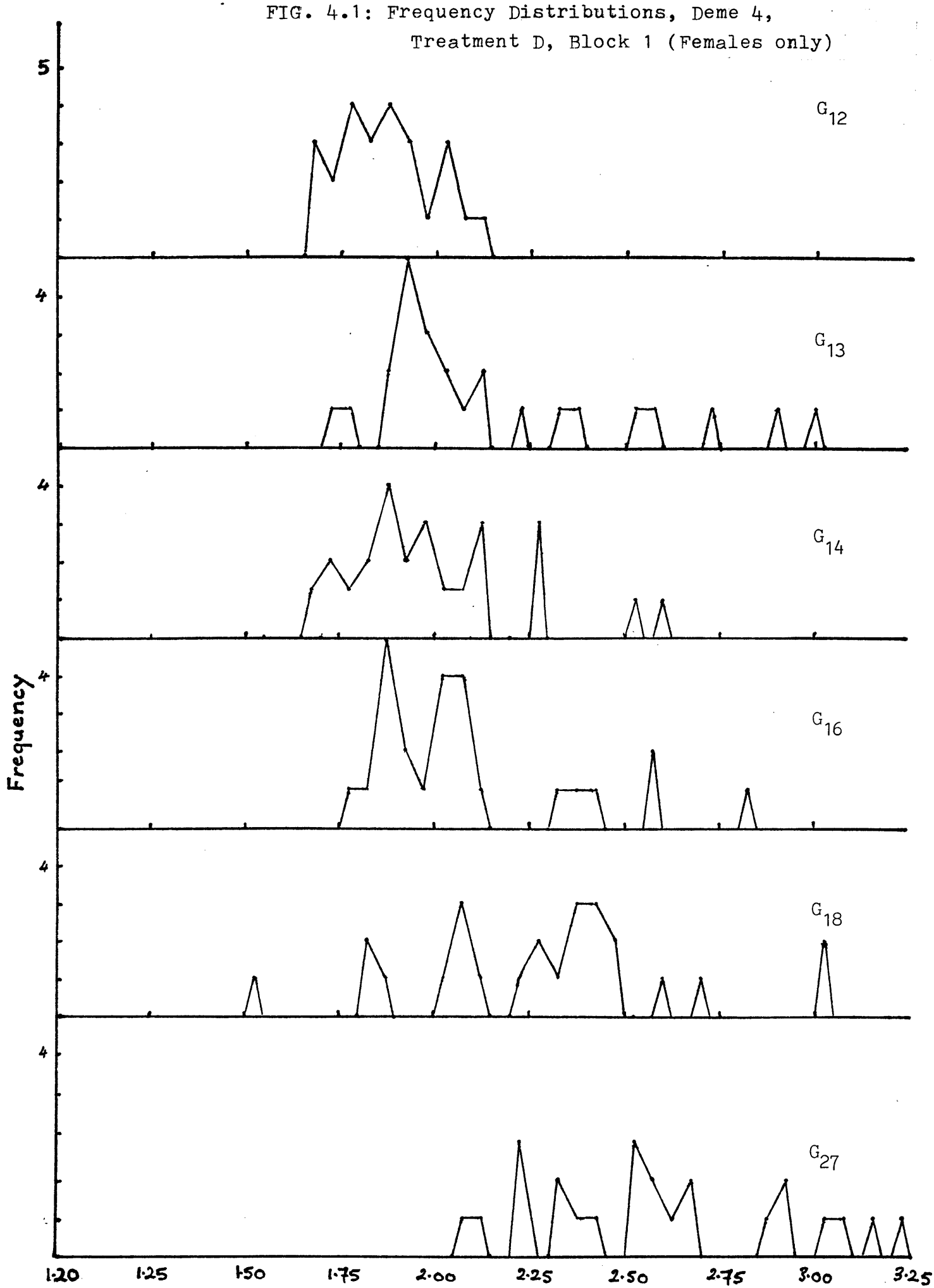

Table 4.1: Number of Migrants received by each deme from Demes previously showing Extreme Phenotype

Generation	Deme Number									
	1	2	3	4	5	6	7	8	9	10
13		1(5)		*	(2)		1(1)		2(2)	
14		2(5)			2(3)				2(2)	
15	2(5)	2(3)			(1)				1(1)	
16	3(5)	1(2)			1(1)				*	
17	3(5)	1(1)				(1)				
18		1(3)			*	1(3)	1(1)			1(2)
19	1(1)	2(3)				2(5)				1(1)
20		*				2(4)	*			3(3)
21	2(2)		2(3)			1(1)				3(3)
22	1(1)		*		1(1)	1(1)				1(2)
23	1(2)				3(3)	*				1(1)
24	*				*			2(2)		*
25								3(3)		
26								*		
Numbers in brackets: total number of immigrants received in each generation										
Total 1	13	11	2	-	7	9	2	5	5	10
Total 2	21	22	3		11	15	2	5	5	12

Total 1 : Total Number of Migrants received from demes showing extreme phenotype

Total 2 : Total Number of Migrants received from all demes over the same period as total 1.

* : Generation in which major effect became established in a deme (NB: effect appeared established in deme 5 at G_{18}, but was lost and was re-established at G_{24}).

over that period was calculated. Although the regression was not significant, the regression coefficient was negative (i.e. when more immigrants were received, fewer generations were required to establish the effect). A similar result was obtained when number of affected migrants received before establishment was plotted against the proportion of such migrants out of total received over that period, although here the regression coefficient was larger. While not conclusive, these results suggest that "establishment" in the new demes was most effective when a high proportion of the migrants received exhibited the effect. Variation around this pattern could result from fitness differences (i.e. in mating rate, egg-lay etc.) between affected and non-affected migrants, and also would depend on the underlying genetic basis of the effect (for example, spread of a dominant major gene with complete penetrance would be very rapid).

Table 4.1 also shows that by G_{26} the large effect had become established in all 10 demes of the population. Thirteen generations, while a large number in applied/livestock terms, is insignificant in evolutionary time; viewed in this context the genetic basis of the selected trait in the entire population system was transformed in a very short period of time. (This rapid change originating in a small population is reminiscent of a widely invoked mode of speciation: namely, rapid changes in a semi-isolated peripheral population which then acts as a source of new genetic material for the rest of the population. Since the pattern of change observed in this experiment occurred in a population under directional selection for a single trait, the analogy with speciation events cannot be closely drawn, but it does support the possibility of rapid change even in subdivided populations with only limited gene flow amongst demes).

The series of frequency distributions presented in Fig. 4.1 illustrates an important aspect of the behaviour of this effect: even in those subpopulations that had presumably been selected to moderate/high frequencies of the gene(s) controlling the effect, its incidence pattern did not change greatly. The main peak became less obvious, and in later generations there was no obvious normal distribution, but the occurrence of a few extreme individuals at intervals at the heavy end of the distribution remained. After many generations of selecting these extreme individuals, no obvious single peak emerged, as might be expected if the effect were caused by a single fixable effect. Further, there was no clear bimodality between affected individuals and the rest, rather, the pattern seems to have been increasing weight of the entire population, with a small increase in the incidence of the extreme individuals.

Several tests were conducted in order to examine the genetic basis of this effect. These included comparison of parent-offspring regressions in the treatment containing the effect with those in the unselected control line and the mass selected treatment A, reciprocal crosses with the unselected control line to produce F_1 and F_2 generations; and an attempt to locate the gene(s)

contributing to differences between the line containing the effect and the mass-selected treatment A, by means of chromosomal substitution methods.

4.2. Parent-Offspring Regression in Experimental Populations

4.2.1. Materials and Methods

Six experimental populations were tested; those of treatments A, D, and E from both blocks. This estimation was conducted at G_{24}, after 23 generations of selection, at which stage the selected populations were 3-5 standard deviations heavier than the unselected control. For each treatment/block combination, 10 pairs of flies were sampled at random from each of the 10 bottles comprising the experimental population. These 10 pairs were weighed in the usual way, and the two heaviest and two lightest pairs retained to be parents. These were then mated assortatively (on a within-bottle basis) and females allowed to lay eggs in 3-inch vials containing F_1 medium. After emergence, 10 pairs of progeny were collected and scored, and male and female progeny means recorded. The heritability of adult bodyweight in each of the treatment/block populations could then be estimated as regression of offspring on parent. Assortative mating increases the precision of such estimates (Falconer, 1981).

4.2.2. Results

The regression coefficients estimated are presented in Table 4.2 (over page). Regressions of progeny on mid-parent were calculated using both the raw data, and data transformed by expressing each observation as the difference from its respective population mean, standardised by the appropriate standard deviation. In general, this had the effect of increasing male progeny-parent regressions, and decreasing those between female progeny and parents. The regression coefficients presented for cross-sex pairings were estimated using transformed data, those within sexes are based on the raw data.

As with the parent-offspring estimates obtained in the base population study (Chapter 2), the regression coefficients for treatments A and E in both blocks were not significantly different from zero. Given the problems discussed concerning this method of estimating the heritability of bodyweight

in this population (see Chapter 2), no conclusions can be drawn concerning the levels of additive genetic variance in bodyweight in treatments A, D, and E at G_{24}. These populations do serve as a useful basis for comparison for the treatment D populations. The continued response to selection in treatment A in both blocks, after G_{24}, is evidence of available genetic variance not detected in the parent-offspring regression.

Table 4.2: Offspring-Parent Estimates of Heritability:Treatments A, D, and E.

Treatment	Block 1.			Block 2.		
Regression	b	S.E.b	r^2	b	S.E.b	r^2
A:						
O on Sire	0.082	0.074	-	0.0245	0.0780	0.4
O on Dam	0.106	0.090	-	0.0278	0.1420	0.2
O on Dam	-	-	-	-0.0079	0.0850	0.0
O on Sire	-	-	-	0.0899	0.1299	2.0
O on Mid-Par.	0.068	0.073	-	0.0102	0.0834	0.1
O on Mid-Par.	0.106	0.098	-	0.0647	0.1894	0.9
O+O on Mid-Par.	0.087	0.077	-	0.0874	0.0968	0.6
D:						
O on Sire	0.683	0.215	-	0.221	0.075	23.8
O on Dam	0.554	0.209	-	0.348	0.117	23.9
O on Dam	-	-	-	0.151	0.093	8.6
O on Sire	-	-	-	0.334	0.101	28.2
O on Mid-Par.	0.503	0.196	-	0.212	0.087	17.4
O on Mid-Par.	0.826	0.256	-	0.378	0.112	28.9
O+O on Mid-Par.	0.664	0.207	-	0.295	0.081	32.1
E:						
O on Sire	-0.001	0.084	-	0.020	0.083	0.3
O on Dam	0.165	0.099	-	0.246	0.268	4.5
O on Dam	-	-	-	0.005	0.111	0.0
O on Sire	-	-	-	0.260	0.197	1.9
O on Mid-Par.	0.020	0.100	-	0.016	0.099	0.1
O on Mid-Par.	0.170	0.097	-	0.284	0.236	7.4
O+O on Mid-Par.	0.095	0.080	-	0.150	0.147	5.4

In both blocks, all regression coefficients for the treatment D populations were significantly different from those observed in both the A or E populations. A contributing factor in this increased offspring-parent regression coefficient could be higher inbreeding levels expected within demes of the subdivided treatment D. This is the simplest explanation for the difference between treatment D and treatments A and E in block 2. The

important feature of these results from the point of view of the large flies present in treatment D, block 1, (D1) is that the offspring-parent regressions in that treatment were all much higher than those observed in treatment D, block 2. Given that the total population for treatment D1 may well have been even more inbred than D2, as a result of constant inbreeding to flies exhibiting the effect, which originally numbered only 5, the larger regression coefficients may again reflect nothing more than increased homozygosity increasing resemblance between offspring and parents. However, the difference in regression coefficients between D1 and D2 was very large: those in D1 were 2 to 3 times bigger than those of D2, and this strongly suggests that some fairly major effect was influencing the regressions in D1. Roberts and Smith (1982) discussed the effects of major genes on genetic parameters, and showed that for a single gene, the heritability in the presence of this gene becomes

$$h^2 = \frac{V_A + 2pq^2}{V_p + 2pqx^2 + (2pqd)^2}$$

where α, d, p, and q follow Falconer's (1981) terminology.

From this formula it follows that major genes segregating can significantly increase the heritability of the trait. Such increases have been investigated by Latter (1965), Frankham and Nurthen (1981) and Smith and Webb (1981). In the latter two papers, increases in heritability due to major genes were documented.

Close investigation of the data underlying these results showed marked differences between regressions within individual demes for treatment D1, whereas results for the other treatments were more homogeneous. Such heterogeneity could result from the combined effect of two factors: firstly, differences in frequency of the gene(s) of large effect between demes, and secondly, interactions between the large effect and fitness, resulting in matings where both parents were extremely large often failing to produce progeny. Thus the regression estimates, already based on a fairly small number of observations, include only a very small number of successful matings

involving the very large flies. Because of the problems in interpretation of the results for treatment D1, the estimation of parent-offspring regressions was repeated.

4.2.3. Parent-Offspring Regressions in Treatment D1 at G26

For this estimation, 10 pairs of flies were scored in each of 6 bottle populations (comprising a reserve population set up by sampling at random from treatment D1 at G_{24}, and allocating 30 pairs of adults at random to each of 6 F_1 bottles) and mated assortatively, each pair in a 3x1 inch F_1 mating vial. Parental and progeny mean bodyweight were scored, as well as total progeny per mating, and incidence of progeny showing the large effect.

4.2.3.2. Results

The pooled estimate of heritability from offspring-midparent regression was 0.289 ± 0.074. This estimate is lower than those obtained for D1 at G_{24}, reflecting correction for between-deme differences, and also possibly changes in gene frequency, although these would probably not be great over only two generations. Correcting the raw estimates for the treatments A and E obtained at G_{24} for between-bottle differences does not alter them significantly; therefore, these results support the conclusion that a major genetic effect was influencing bodyweight in treatment D1.

For increases in the heritability of this magnitude, assuming a single gene at intermediate frequency, the genotypic values of the homozygote would have to be approximately 3 standard deviations greater than the mean. This is of the same order as the phenotypic range observed in the presence of the effect, the most extreme individuals being approximately 5 standard deviations larger than the mean at the time they first appeared.

While the results from this second estimation are more reliable than at G_{24}, they still suffer from a problem that plagued the earlier set; namely, the fact that matings where both parents were extremely large only rarely succeeded. Thus conclusions regarding the high bodyweight section of the regression are dependent on a very small number of successful matings involving the large effect (only 4 out of 30 "heavy" matings). The regression

of progeny number on parental mean bodyweight in this second trial was significant and negative, supporting the suggestion of lowered fitness in the extremely large flies.

The other important observation in these results is that appearance of extreme individuals in the progeny generation was not limited to those matings where one or both parents exhibited the effect. Six matings between parents at the light end of the scale produced progeny showing the effect to some degree. When all matings except those where either parent exhibited the phenotype are included, extreme individuals were present in the progeny of apparently unaffected parents in 17 out of 43 cases. For this to occur with a single gene, it would have to show low penetrance and/or recessiveness. An alternative explanation, that the effect involves some sort of interaction, can explain this observation: if the alleles required for the interaction are at moderate frequencies in this population, many individuals will not have the complete interaction set but matings with other members of the same population could produce the required combination. The frequency of such an occurrence depends on the number of genes involved in the interaction, their frequencies in the population, degree of extreme phenotype produced by different components of the interacting system, linkage and linkage disequilibrium. The high proportion of matings where this occurred (17/43) suggests that if interaction was involved, it must have been fairly simple, or the component alleles were all present at very high frequencies.

4.3. Reciprocal Crossing with Unselected Flies

4.3.1. Materials and Methods

For this experiment, virgins were collected and scored from the D1 and E1 populations at the end of the selection experiment. Fifty single-pair matings were set up in both combinations (i.e. D1 ♂ x E1 ♀, and D1 ♀ x E1 ♂). Parents were chosen at random except that as many of the most extreme D1 flies as were available were used. Ten pairs of progeny were collected from each successful mating; these F_1 individuals were individually weighed, then mated within F_1 families to produce the F_2. F_2 progeny were weighed en masse to give family

means. Number of progeny produced from each F_0 cross was also recorded.

4.3.2. Results

Means and standard deviations of male and female bodyweights for the F_0, F_1, and F_2 generations are presented in Table 4.3 (over page). Family numbering reflects F_0 parentage as follows:

Families 1-50: D1 ♂ x E1 ♀

Families 51-100: E1 ♂ x D1 ♀

Differences in proportions of matings that were fertile between the two groups were large: 6 matings in the 1-50 group were completely infertile, while none of the 51-100 group failed. Of those in the 1-50 group that produced F_1 progeny, 5 failed to produce progeny in the F_2. The 51-100 group produced F_2 progeny in 44 cases. (Means and variances presented are restricted to those families represented in all three generations). The main differences in fertility were in the F_0, and this may reflect differences in the effect of extreme size on male and female mating ability: the swollen abdomen and resultant reduced flexibility did not appear to reduce mating success in the case of the females mated to wild-type males, since the males could position themselves normally. However, the large abdomen made it almost impossible for the largest males to position themselves successfully on "wild-type" females. In the F_1, fewer families failed to produce any progeny since there was a range of phenotypes from normal to large, so that within any F_1 family, some matings would involve unaffected individuals.

In both the F_1 and F_2, male and female means differed from each other and from the F_0 means, for both family groups. Differences between male means in the F_1, and between female means in the F_2, between 1-50 and 51-100 groups, were highly significant ($P<0.001$); those between F_1 female means and F_2 male means for the two groups were not significant. In both the F_1 and the F_2, means for both sexes were lower for families 1-50. Productivity was higher in the F_0 matings for the 51-100 group, so it seems unlikely that crowding effects could have produced this difference. The pattern of differences between male means in the F_1 and F_2 for the two groups is suggestive of a

Table 4.3: Means and Variances in Crosses with the Unselected Control

Group	1-50					
	Male			Female		
	n	x	s	n	x	s
All F_0	50	1.5284	0.2511	50	1.6278	0.1359
Fertile F_0	40	1.4885	0.2293	40	1.6295	0.1228
All F_1; Family Means	40	1.0799	0.1910	40	2.0989	0.2331
All F_1; Individuals	200	1.0803	0.2597	199	2.0976	0.3179
F_1; Families having F_2 Progeny	39	1.0707	0.1843	39	2.0991	0.2361
F_2; F_0 Family Means*	39	1.1614	0.1267	39	2.0237	0.1468
F_2; F_1 Family Means**	165	1.1514	0.1553	165	2.0047	0.1984
Group	51-100					
	Male			Female		
	n	x	s	n	x	s
All F_0	50	0.9076	0.0611	50	2.6330	0.3473
Fertile F_0	45	0.9047	0.0608	45	2.6273	0.3037
All F_1; Family Means	45	1.2381	0.2025	45	2.1952	0.2409
All F_1; Individuals	225	1.2368	0.3031	223	2.1937	0.3837
F_1; Families having F_2 Progeny	44	1.2321	0.2008	44	2.1916	0.2424
F_2; F_0 Family Means*	44	1.2070	0.1296	44	2.1499	0.1447
F_2; F_1 Family Means**	170	1.1832	0.1606	172	2.1262	0.1925

* based on the mean of all F2 progeny within each F0 family
** based on means of F1 matings within F0 families

gene(s) on the X chromosome affecting bodyweight. This would result in the difference between male means disappearing in the F_2 since the expected frequency of such an allele in the F_2 is the same in the two groups (see Appendix A). However, this hypothesis cannot completely explain the observed results, since in both crossing groups, individuals showing the effect appear in the F_1. A sex-linked allele would only appear in the F_1 of the 51-100 group, where the female parents included affected individuals.

The observation of the effect in the F_1 means that it must involve dominance, at least against a wild-type background. This apparent dominance was not however straighforward: particularly in the F_2, individuals clearly showing the effect were produced from parents not affected to the same degree. This again suggests either an interaction or the involvement in the extreme phenotype of recessive alleles.

Support for an epistatic basis for the effect is derived from incidence patterns in the F_1 and F_2. In no cases did all progeny in an F_1 family show the effect, the highest observed proportion being 8/10 affected individuals. Similarly, even where both parents of an F_2 family showed the effect, the F_2 progeny exhibited a range of phenotypes, from obviously extreme to apparently unaffected. As shown in Appendix A, this pattern is consistent with even a simple two-factor interaction, and an F_0 population segregating for the alleles involved.

Table 4.4 (over page) shows the correlations between relatives in this data set. It is apparent that presence of the effect increased the correlation significantly (e.g. compare the correlation of F_1 progeny with male and female parents; in families 1-50, correlation of F_1 male and female family mean was much higher with male parent than female, the latter being not significantly different from zero; this situation was reversed in families 51-100). This observation is similar to the offspring-parent regressions observed in the three treatments discussed earlier; where the effect was segregating, the regression coefficient was high and significant, while in lines not segregating for the effect, such regressions were not significantly different from zero.

Table 4.3: Means and Variances in Crosses with the Unselected Control

Group	1–50					
	Male			Female		
	n	$\bar{x}$	s	n	$\bar{x}$	s
All F_0	50	1.5284	0.2511	50	1.6278	0.1359
Fertile F_0	40	1.4885	0.2293	40	1.6295	0.1228
All F_1; Family Means	40	1.0799	0.1910	40	2.0989	0.2331
All F_1; Individuals	200	1.0803	0.2597	199	2.0976	0.3179
F_1; Families having F_2 Progeny	39	1.0707	0.1843	39	2.0991	0.2361
F_2; F_0 Family Means*	39	1.1614	0.1267	39	2.0237	0.1468
F_2; F_1 Family Means**	165	1.1514	0.1553	165	2.0047	0.1984
Group	51–100					
	Male			Female		
	n	$\bar{x}$	s	n	$\bar{x}$	s
All F_0	50	0.9076	0.0611	50	2.6330	0.3473
Fertile F_0	45	0.9047	0.0608	45	2.6273	0.3037
All F_1; Family Means	45	1.2381	0.2025	45	2.1952	0.2409
All F_1; Individuals	225	1.2368	0.3031	223	2.1937	0.3837
F_1; Families having F_2 Progeny	44	1.2321	0.2008	44	2.1916	0.2424
F_2; F_0 Family Means*	44	1.2070	0.1296	44	2.1499	0.1447
F_2; F_1 Family Means**	170	1.1832	0.1606	172	2.1262	0.1925

* based on the mean of all F2 progeny within each F0 family
** based on means of F1 matings within F0 families

Strong correlations were also observed between F_0 parent and F_2 progeny, both male and female, where the F_0 parents were segregating for the effect (i.e. correlations large with F_0 male for families 1-50, with F_0 female for families 51-100). Correlations were also calculated with F_1 and F_2 standard deviation (within family), and with numbers of male and female progeny in the F_2. Correlations including individuals or groups of individuals segregating for the effect were significant and positive for standard deviation, and negative for numbers of progeny. Correlations between the F_1 and F_2 variables were all significant suggesting either that the effect appeared at similar frequency in the F_1 and the F_2, or that its presence boosted covariances markedly.

These results are all suggestive of a non-additive basis for the effect. They show that it must involve dominance, be of large effect, and possibly include sex-linked genes. Some of the observations are very hard to explain without invoking some form of variable expression, which could reflect the involvement of some sort of regulator system, or of metabolic thresholds.

Three problems of interpretation should be noted, one of which was under experimental control. Firstly, the variable expression of the effect, with individuals of widely ranging bodyweights producing affected offspring, made determination of frequencies of occurrence, and thus determination of the genetic basis, very difficult. Secondly, the observed negative interaction with fitness, particularly among males, again reduced the usefulness of the data, as the potentially most informative matings failed to produce progeny. Thirdly, the data would have been more valuable had individual F_2 progeny weights been recorded, rather than weighing being done en masse. Frequency estimation within the F_2 would have then been possible. This also means that analysis of variance in the F_2 was limited to describing proportions between and within F_0 families. Had individual weights been recorded, variance components for F_1 families within the F_0 could have been estimated. The reason for not obtaining individual weight was the amount of work that would have been involved - weighing c. 4000 flies (approx. 35 hours). Had the entire experiment been spread over several consecutive days, this could have been

achieved.

4.4. Chromosomal Analysis of Lines A1 and D1

4.4.1. Introduction

The technique of chromosomal substitution for determining the location and minimum number of genes affecting quantitative traits has been used extensively (Robertson and Reeve, 1953; Robertson, 1954; King and Somme, 1958; Seiger, 1966; Kidwell, 1969; Frankham, 1969, 1970; Frankham and Nurthen, 1981). By means of marked inversions in one or more chromosomes, stocks whose genetic content (in terms of source of chromosomes) is known precisely can be produced, and their phenotypes observed.

The technique was used here in an attempt to determine the genetic basis/location of the gene(s) underlying the large effect observed in this study, following the procedure of Frankham (1969). Difficulties arose due to low fitness of some genotypes, in particular of the large flies of interest, resulting in a fairly unbalanced data set, and low representation of the most important phenotypes.

In this experiment, the two selection lines A1 and D1 were compared. Initially it was intended to compare these lines with each other and with the unselected control from block 1 in order that the effects of additive genes accumulated by the selection process could be distinguished from those of the gene(s) producing the large flies. However, problems encountered in the initial A1-D1 comparison resulted in the procedure having to be repeated, and the stock of large flies was lost before further tests could be done.

4.4.2. Materials and Methods

The procedure of Frankham (1969) was followed, using the same marker stock (Basc;SM1,Cy/Pm;TM2,Ubx/Sb;spa^{pol}). The initial sample from lines A1 and D1 was 72 males, but only 40 A1 line and 39 D1 line males were represented in the generation scored. In order that all flies could be weighed at as close to 3 days after emergence as possible, matings to produce the measured generation were spread over 6 consecutive days. Final emergences were spread over nearly

9 days, and flies were collected in both mornings and evenings. These factors; day of mating, day of emergence and time of collection (and thus weighing) were all included in the models for analysis. To account for possible differences among parents in the initial sample, particularly within the D1 line where there was considerable variation in the degree of expression of the extreme phenotype, the number of the initial samples contributing the A-line and D-line chromosomes was also included in the model. To account for maternal effects related to maternal genotype, this also was included in the model.

The procedure produced families of flies of all combinations (18 in males and 27 in females) of chromosomes from the two parent lines. From each family produced, 10 males and 10 females were weighed en masse.

4.4.3. Results

Analyses of variance for male and female bodyweight are presented in Table 4.5. The first feature to note is the low proportion of variation explained (33% for males and females). This may reflect low levels of genetic variation within the selection lines, and large amounts of environmental variation due to factors other than those identified.

Table 4.5: Analyses of Variance for Male and Female Bodyweight, Chromosomal Substitution Experiment

Source	Male			Female		
	df	Mean Square	P	df	Mean Square	P
Day of Mating	5	0.1693	0.000	5	0.2202	0.000
Day of Collection	7	0.0390	0.001	7	0.0839	0.001
Time of Weighing	1	0.0911	0.002	1	0.0915	0.084
Sample No: A Line	40	0.0306	0.000	40	0.1011	0.000
Sample No: D Line	39	0.0328	0.000	39	0.1182	0.000
Maternal Genotype	7	0.0240	0.019	7	0.0611	0.055
First Chromosomes	1	0.0018	0.669	2	0.0109	0.699
Second Chromosomes	2	0.0117	0.298	2	0.1576	0.006
Third Chromosomes	2	0.1201	0.000	2	0.7216	0.000
First x Second	2	0.0119	0.289	4	0.0599	0.098
First x Third	2	0.0402	0.017	4	0.0249	0.515
Second x Third	4	0.0160	0.160	4	0.0644	0.078
First x Second x Third	4	0.0169	0.138	8	0.0309	0.412
Residual	924	0.0096		931	0.0301	
	r^2 = 33.047			r^2 = 32.507		

Both Day of Mating and Day of Collection were significant at the 1% level in both sexes. Time of Weighing was significant at this level in males, and nearly significant at the 5% level in females. In both sexes, sample number within both selection lines was highly significant ($P<0.001$) indicating that there were differences between individuals in both lines in progeny mean weights. Maternal genotype also significantly influenced progeny weight (P-0.01 for males, P-0.05 for females).

Of the chromosomal effects, differences between A and D first chromosomes were not significant in either sex. Second chromosomes were significantly different in females ($P<0.01$) but not so in males, while third chromosomes were significantly different ($P<0.001$) in both sexes. Of the interaction terms, only the first x third interaction in males was significant, although not so in females ($P = 0.515$), and probability for second x third in females, approached significance level ($P = 0.077$).

Table 4.6: Least Squares Means for Single Chromosomes, and Interactions between Chromosomes

Chromosome(s)	Male			Female	
	L.S.M.	S.E		L.S.M.	S.E
First					
A	1.1505	0.0682	AA	2.2850	0.1220
D	1.1797	0.0680	AD	2.3180	0.1248
			DD	2.3327	0.1202
Second					
AA	1.1571	0.0667	AA	2.3203	0.1205
AD	1.1693	0.0708	AD	2.3384	0.1248
DD	1.1688	0.0674	DD	2.2573	0.1201
Third					
AA	1.1617	0.0667	AA	2.3276	0.1202
AD	1.1856	0.0707	AD	2.3481	0.1249
DD	1.1479	0.0674	DD	2.2402	0.1201
First x Third					
A.AA	1.1493	0.0645	AA.AA	2.3180	0.1180
.AD	1.1728	0.0711	.AD	2.3263	0.1267
.DD	1.1293	0.0716	.DD	2.2108	0.1293
D.AA	1.1742	0.0706	AD.AA	2.3425	0.1255
.AD	1.1984	0.0708	.AD	2.3454	0.1253
.DD	1.1644	0.0647	.DD	2.2661	0.1263
			DD.AA	2.3223	0.1235
			.AD	2.3726	0.1254
			.DD	2.2438	0.1122

Least squares means for the individual chromosomes, and for the first x third interaction are presented in Table 4.6 (previous page). The non-significant single chromosome effects are included for comparison. Means for D1 first chromosomes were greater than those for A1, and in females there was slight heterosis between first chromosomes suggesting frequency differences between the lines for sex-linked genes influencing bodyweight. For second and third chromosomes (with the exception of second chromosomes in males), A1-line homozygotes were heavier than D1-line homozygotes, and in all cases there was heterosis between chromosomes from the two lines.

4.4.4. Discussion

These results do not contribute greatly to understanding the putative major gene(s). The analysis suggests involvement of genes located on the first (X) chromosome in large D-line flies. However, the results of the backcrossing trial do not support the involvement of X-linked genes. The results do suggest that on all three chromosomes different genes were selected, or selected genes reached different frequencies in the two lines, providing a basis for heterosis (Falconer, 1981).

As a means of locating and defining the major gene(s), this experiment was not a success. Several factors contributed to this, the primary one being the reduced viability of those flies showing the effect, resulting in low representation of this phenotype in the measured generation.

Secondly, because of this low frequency in the measured generation, and the variable expression of the effect, the likelihood of discerning the effect against the background of additive genetic and environmental variation in bodyweight would be reduced.

Thirdly, it would have been advantageous to compare the D1-line with a more uniform line than the A1, which even though apparently plateaued or nearly so, obviously still contained genetic variation (as evidenced by the highly significant variation between A1-line samples). Similar variation also existed between D1-line samples, and this contributed a further difficulty, in that detection and definition of the effect would have been more likely had

all samples contained the effect and expressed it equally.

As with other attempts to define this effect, individual weighing would have been highly advantageous, but again, logistically more difficult. The most useful result of having individual bodyweights would be that the distribution of individuals showing the effect would be easily obtained, and thus the chromosome(s) required for its expression would be known. This procedure would still be subject to sampling effects, since only a small number of individuals in the scored generation showed the effect. Failure to record individual bodyweight, or at least presence/absence of flies showing the effect, was an oversight, and the data allow no accurate means of rectifying it.

Listings of the heaviest families for males and females (Table 4.7, over page) reveal that among D1-line samples, three comprised a high proportion of the most extreme (if the heaviest 15 families are taken, the same three, Nos. 12, 20, and 31, accounted for 10 of the 15 in males and 7 of the 15 in females). No such increased frequencies were seen in the A1-line listings where only two samples occured more than once in both males and females.

In the 17 families in which D1-line samples 12, 20, and 31 appeared, the genotypes include D1-line first, second, and third chromosomes in 12 cases; D1-line second and third chromosomes in two cases; and in the remaining three cases, only the third chromosomes derived from the D1 line. These observations provide circumstantial evidence that the extreme phenotype involved genes on the second and third chromosomes.

Estimation of the differences between A1 and D1, and of the genetic basis of the extreme phenotype, would have been much easier had both lines been isogenic. Considerable variation existed between the original samples in both lines, making interpretation of the analysis of variance more complex, and reducing the likelihood of detecting any chromosomal effects and interactions. Least squares means without sample number fitted in the model, for the first x third chromosome interaction in males, are given in Table 4.8. The effect of replacing A1-line first chromosomes with D1-line is much more pronounced than

with sample number fitted (see Table 4.6), suggesting that differences existed between samples for this interaction.

Table 4.7: Listing Of Heaviest Families in Chromosomal Analysis: a) Male

Rank	Chromosomal Content			Sample No. Source of A Chromosomes	D Chromosomes	Mean Bodyweight (mg)
	1st	2nd	3rd			
1	D	DD	AD	12	12	1.8231
2	D	DD	DD	-	12	1.7250
3	D	AD	DD	12	12	1.6850
4	D	AD	DD	3	3	1.6475
5	A	AD	AD	33	31	1.6457
6	D	AD	AD	15	12	1.5872
7	D	DD	AD	18	20	1.5640
8	D	AD	DD	29	25	1.5616
9	D	DD	AD	24	22	1.5400
10	A	AD	DD	14	12	1.4393
11	D	AD	AD	24	27	1.4382
12	A	DD	AD	12	11	1.4261
13	A	AA	AD	18	20	1.4250
14	D	AD	AD	29	31	1.4200
15	A	AA	DD	40	34	1.4177

Overall Male Mean = 1.1651 mg, Standard Deviation = 0.0984 mg
Dotted lines indicate weights 3 and 4 standard deviations larger than the overall mean.

Table 4.7: Listing of Heaviest Families in Chromosomal Analysis:(b)Female

Rank	Chromosomal Content			Sample No. Source of A Chromosomes	D Chromosomes	Mean Bodyweight (mg)
	1st	2nd	3rd			
1	DD	AD	DD	3	12	2.8614
2	AA	AD	AD	36	31	2.8400
3	AA	AA	AD	24	24	2.7200
4	AA	AA	AA	43	-	2.7100
5	AA	AA	AD	16	12	2.6775
6	AD	AD	AD	24	26	2.6300
7	AD	AD	AD	33	31	2.6243
8	AD	AA	AA	38	40	2.6200
9	AA	AA	AA	12	-	2.6000
10	AD	DD	AD	11	12	2.5969
11	AA	AA	AD	18	20	2.5900
12	AA	AA	AA	4	-	2.5900
13	DD	DD	DD	-	12	2.5881
14	AD	AA	DD	40	34	2.5850
15	DD	AD	AD	12	15	2.5843

Overall Female Mean = 2.1042, Female Standard Deviation = 0.2365
Dotted line indicates weight 3 standard deviations larger than the overall mean.

Table 4.8: Least Squares Means for First x Third Chromosomes in Males: Model not including Sample No. from A1- and D1-lines

Chromosomes		LSM	SE
1st	3rd		
A	AA	1.1907	0.0224
A	AD	1.1783	0.0142
A	DD	1.1251	0.0192
D	AA	1.1652	0.0180
D	AD	1.1946	0.0147
D	DD	1.1733	0.0209

4.5. General Discussion

The experimental work reported in this chapter was aimed at determining the genetic basis of the extreme phenotype seen in treatment D1 of the selection experiment. That treatment was designed to model the conditions of the Shifting Balance Theory. Since the theory is concerned with selection of favourable interaction systems, it was important to determine the nature of the extreme phenotype.

The diffusion pattern of the extreme phenotype in treatment D1 follows that suggested by Wright for the third phase of the Shifting Balance process. Over a relatively short period, the effect was introduced to, and established in, all demes of the subdivided population. This in itself provides no evidence as to the genetic basis of the effect, since single genes would be spread in a similar way, with their speed of establishment depending on degree of penetrance and dominance.

The observed diffusion pattern results from the interaction of a number of factors; the overall migration rate, the composition of the migrant pool, mating behaviour of migrants, the genetic basis of the effect, and accidents of sampling. Diffusion would presumably be enhanced by a higher overall migration rate, which might result from altered environmental conditions. Restriction of gene flow is however a feature of the necessary conditions for the Shifting Balance process, so that increases in overall migration rate after occurrence of novel favourable interaction systems would have to be seen

as fortuitous coincidences. Changes in overall migration rates must be distinguished form the increased dispersal form demes in which new interactions have arisen, which is an essential component of the Shifting Balance Theory.

The sexual composition of the migrant pool would also affect the diffusion pattern. In species where sex ratios at mating are unbalanced, migration of the rare sex would have far greater effect than of the more common. This has the potential both for negating any environmentally imposed population subdivision, and for the spread of interaction systems. Where both sexes migrate, the possibility exists for spread by competitive displacement rather than by simple diffusion (Shields, pers. comm.). This involves migration of pairs (or larger units) when then competitively displace original members of the deme.

The possibility of migration by both sexes raises the question of mating behaviour of migrants. Highly assortative mating behaviour where both sexes migrate, or after introduction of the interaction system, would also increase the rate of diffusion by increasing the rate of production of individuals carrying the system. Such assortative mating could result from intense selection within the new demes, and from behavioural patterns inherent in the organism or generated by the interaction itself.

As has been suggested previously, the genetic basis of the effect would affect the rate of diffusion from the source deme. More complex interactions would require more intense inbreeding in new demes, longer periods of time, or higher migration rates, for establishment, than two-locus interactions, or single genes. Dominance of the alleles comprising an interaction would reduce the number of generations required for successful introduction, since F_1 progeny in new demes would be more likely to be selected than where the alleles were neutral or recessive.

Lastly, chance effects, operating via mishaps, sampling of migrants and their offspring etc, would affect the rate and pattern of diffusion.

These predictions would suggest that rapid spread of novel interaction

systems will be most likely when they are relatively simple genetically, have a large selective advantage, and where migrants, if possible of both sexes, form a large proportion of recipient deme mating pools. In this experiment, the extreme phenotype was spread throughout the subdivided population in only 13 generations. It had a large selective advantage (i.e. high adult bodyweight), and migrants expressing the extreme phenotype comprised up to 30% of the mating pool in recipient demes. However, while results suggest dominance of the effect (or its component alleles), no precise conclusions as to its genetic basis are possible.

Crossing with unselected flies showed that F_1 progeny could express the effect, and that matings between individuals not showing the effect could produce extreme offspring. There was a suggestion of F_2 breakdown of the effect, which along with the previous point, suggests an epistatic basis for the extreme phenotype (although F_2 breakdown can be produced by dominance). Heritability estimates from parent-offspring regression suggested the presence of a major non-additive factor.

The chromosomal substitution analysis did not provide much information about the extreme phenotype. Significant interactions were found in both sexes between first and third chromosomes, and individuals homozygous for D1 third chromosomes were significantly lighter than A1 homozygotes, or heterozygotes. The significant first x third interaction does not fit with the observation of the extreme phenotype in both sexes in the F_1 progeny in matings with unselected flies.

The chromosomal substitution analysis would have been much more useful had individual phenotypes been noted in the scored generation. This would have shown which D1 chromosome(s) were necessary for appearance of the extreme phenotype, and would not have interfered with the group weighings. Difficulty in identifying lighter affected individuals would have caused some problems, but would equally affect the quantitative analysis.

In all the experiments aimed at understanding the extreme phenotype, the reduced fitness of its carriers, and the lack of clear separation of affected from non-affected individuals were problems. No simple methods of overcoming

either are obvious. In the selection experiment, large changes in bodyweight were achieved in both A1 and D1, but with quite different affects on fitness. In A1, where the changes presumably resulted solely from accumulation of additive genes, no reduction in fitness was observed, whereas marked reductions accompanied the spread of the extreme phenotype in D1. Single-pair matings involving the most extreme individuals were almost always infertile, and mating bottles (5 pairs) where all individuals showed the effect produced lower numbers of offspring. This resulted from both greater difficulty in mating, and from lower egg production from mated females.

This series of experiments investigating the extreme phenotype was brought to a halt, and rendered of somewhat academic interest, by the loss of the stock segregating for it in an air-conditioning breakdown. The temperature dropped to 12^{o}C for several hours, and no extreme individuals survived. As they had been set up for mating and a backup stock was stored in the same room, this meant that the stock was lost.

CHAPTER 5

ELECTROPHORETIC SURVEY OF EXPERIMENTAL POPULATIONS

5.1. Introduction

The aim of this study was to examine responses to directional selection imposed on differently structured populations. The populations differed in the degree of isolation between component sub-populations, and thus might be expected to differ in the partitioning of total genetic variance into between- and within-subpopulation components (see Chapter 3 for details of the population structures used in the selection experiment). Simple techniques for estimating this subdivision from phenotypic parameters were discussed in Chapter 3. A more sophisticated description of the genetic variation is available from electrophoretic surveys of the populations, using allozyme loci. Comparisons between phenotypic and genetic population differences and within-population differentiation are very difficult, and must be viewed with caution (Lewontin, 1984), but population comparisons using the electrophoretic data are valid. In this chapter, the results of a small-scale survey of allozyme variation in the experimental populations are presented, together with estimates of the genetic differences among populations, and descriptions of population structuring using three sets of statistics.

5.2. Literature Review

Since Hubby and Lewontin (1966) first showed that isozymic variation could be detected using electrophoretic techniques, they have been applied in a wide range of species and populations. These techniques provide a number of methods of describing the genetic variation present in a population. The basic parameter is allele frequency at a locus, which can be used to provide measures of diversity such as percentage of polymorphic loci, average number of alleles per locus, measures of the evenness of allele frequencies, and measures of the degree of heterozygosity and of deviations from Hardy-Weinberg conditions. Loci may be considered

individually, or information from two or more loci combined in order to study correlations among alleles at different loci. Statistical complexity and sampling errors increase with numbers of loci considered. Finally, estimates of allele frequencies at one or more loci can be used to describe patterns of genetic differentiation among sub-populations. In that these provide an indication of the amount of gene flow between subpopulations, they provide a means of observing the interaction of drift and migration in natural populations. The common procedures used are Wright's F-statistics (1951, 1965), analysis of within and between-population heterozygosity (Nei, 1973), and correlation based measures (Cockerham, 1969). Recently, Weir and Cockerham (1985) have presented F-statistics for finite and variable sample sizes.

Wright's F-statistics (1951) are as follows:

F_{IT} : the deviation of entire populations from Hardy-Weinberg equilibrium

$$F_{IT} = \frac{1}{n} \sum_{k=1}^{n} \left(1 - \frac{H_k}{2p_k q_k} \right)$$

where H_k = observed heterozygote frequency in entire population

p_k, q_k = allele frequencies for the total population at the k^{th} locus

n = number of loci

F_{IS} : average deviation of each subpopulation from Hardy-Weinberg frequencies

$$F_{IS} = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{m} \frac{N_i}{N} \left(1 - \frac{H_{ik}}{2p_{ik} q_{ik}} \right)$$

p_{ik}, q_{ik} = frequencies of the 2 alleles at the k^{th} locus in the i^{th} subpopulation

m = number of subpopulations

N_i = size of the i^{th} subpopulation

N = size of total population

H_{ik} = observed heterozygote frequency in the i^{th} subpopulation at k^{th} locus.

and F_{ST}, which measures the degree of differentiation among subpopulations

$$F_{ST} = \frac{1}{n} \sum_{k=1}^{n} \sigma^2_k \; / \; \bar{p}_k \, \bar{q}_k$$

σ^2_k = variance of gene frequencies between subpopulations

These statistics are inter-related as follows:

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

Wright (1969, 1978) has discussed the widespread use of these statistics in population and evolutionary genetics. Nei (1977) showed how genetic diversity could be ascribed to different population levels in a manner analogous to that of Wright, but based on observed and expected genotypic and allelic frequencies. Weir and Cockerham (1984) have discussed problems of interpretation of both sets of statistics, of sampling problems affecting both, and present a method whereby the correlations among genes in structured populations may be expressed using statistics outlined by Cockerham (1969, 1973). Full details of their methods are given in Weir and Cockerham (1984). For the present purpose it is sufficient to note that identical descriptions of population structuring are available using these statistics. The three statistics are:

$F = F_{IT}$

$\theta_W = F_{ST}$

and $f = F_{IS}$

where

F : the correlation of genes within individuals ("inbreeding")

θ_W : the correlation of genes of different individuals of the same population ("Co-ancestry")

and f : the correlation of genes within individuals within populations.

Levels of gene flow in subdivided populations may also be estimated using the conditional average frequency (the average frequency of an allele conditioned on the number of local populations it appears in), particularly for rare alleles (Slatkin, 1981; 1985). Spatial autocorrelation estimates, (Sokal and Oden, 1978a,b) which measure the dependence of variables, e.g. allozyme frequencies, on values of that variable in neighbouring locations, can provide an indication of the possible importance of gene flow and other evolutionary forces.

The literature on allozymic variation in natural populations has been extensively reviewed by Wright (1978). His review includes reinterpretations in some cases, and discussions of conclusions drawn in some studies. Selander and Whittam (1983) reviewed the study of population genetic structure through analysis of protein polymorphism. Their discussion included the phenomenon of hidden genetic variation, and the effects of major evolutionary forces acting both historically and at the time of survey.

Descriptions of patterns of allozymic variation in artificially structured populations in the literature are rare. Altukhov and Bernashevskaya (1978) describe gene frequencies at allozyme loci in a circular stepping-stone model population of D. melanogaster. This system was simulated by Altukhov et al (1984), the results showing that in this system of subpopulations significant local differentiation of gene frequencies could develop and that stability of population mean frequency was enhanced.

Supporting Lewontin's (1984) demonstration of the difficulties inherent in comparing levels of quantitative and electrophetic variation, Giles (1984) reported on quantitative and allozymic variation in Wild Barley, Hordeum murinum. Most populations surveyed were monomorphic at allozyme loci with only limited genetic differentiation between populations. In contrast, there was significant quantitative variation both between and within populations.

5.3. Materials and Methods

From each experimental treatment (both blocks) at generations 9 and 19 of the selection program, 25 pairs of flies were sampled at random from each deme and stored at -5.0°C. At the completion of the selection experiment, approximately half the stored flies were assayed electrophoretically for the following allozyme loci: Phosphoglucomutase (Pgm); Esterase-6 (Est-6); Alcohol dehydrogenase (Adh); α-Glycerophosphate dehydrogenase (α-Gpdh); Glucose-6-Phosphate dehydrogenase (G-6-Pdh) and 6-Phosphogluconate dehydrogenase (6-Pgd). These loci were chosen on the basis that each arm of the three major chromosomes was thereby sampled, clear results could usually be obtained for these loci on stored material, and all six exhibited variation in samples taken at generation 28 from treatments A, D, and E, with two alleles segregating at each locus.

From the data, gene frequencies for each treatment, sex and locus in both blocks and generations were estimated. The gene frequency estimates provide the basis of estimates of Genetic Identity and Distance (Nei, 1972).

Three sets of F-statistics were estimated from the data: Wright's (1965); Nei's (1977), and those presented by Weir and Cockerham (1984). Because of the small numbers sampled from each deme, no attempt was made to estimate linkage disequilibrium.

5.4. Results and Discussion

Mean gene frequencies at the 6 loci surveyed, across demes, and within sexes for each population sample are shown in Table 5.1 (over page). All frequencies relate to the fast (F) allele at each locus. At generation 9 (Tables 5.1a & c), there was already differentiation in levels of genetic variation between treatments. In block 1, treatments B and D were polymorphic at all six loci, treatments A and C at five of the six loci, while treatment E was effectively fixed at two of the six loci. In block 2, treatments A and B were polymorphic at five of the six loci, and treatments C, D, and E, at four of the six. At generation 19, (Tables 5.1b & d), block 1, treatment A was polymorphic at 5 out of 6 loci, as at generation 9. Treatment B had become

Table 5.1(a): Summary of Gene Frequency Data: Generation 9, Block 1

Locus	Sex	Mean Gene Frequency (no. in sample) Treatment A	B	C	D	E
Pgm	M	1.000(109)	0.899(69)	0.593(108)	0.844(112)	0.896(115)
	F	0.793(29)	0.951(82)	0.505(111)	0.822(118)	0.940(116)
Est-6	M	0.292(118)	0.226(93)	0.357(115)	0.161(112)	0.000(116)
	F	0.379(112)	0.184(68)	0.426(115)	0.158(120)	0.000(117)
Adh	M	0.350(137)	0.406(90)	0.500(112)	0.482(111)	0.356(111)
	F	0.570(114)	0.159(82)	0.563(112)	0.433(120)	0.417(115)
∝-Gpdh	M	0.345(97)	0.263(38)	0.640(111)	0.490(105)	0.325(103)
	F	0.434(99)	0.525(40)	0.838(114)	0.496(113)	0.273(110)
G-6-Pdh	M	0.367(98)	0.471(17)	0.869(107)	0.865(111)	0.400(105)
	F	0.514(105)	0.684(19)	0.842(111)	0.941(119)	0.451(92)
6-PgD	M	1.000(123)	0.559(93)	1.000(115)	0.920(113)	0.991(116)
	F	1.000(109)	0.306(67)	1.000(115)	0.920(119)	1.000(117)

Table 5.1(b): Summary of Gene Frequency Data: Generation 19, Block 1

Locus	Sex	Mean Gene Frequency (no. in sample) Treatment A	B	C	D	E
Pgm	M	0.960(164)	1.000(186)	0.657(137)	0.852(166)	1.000(121)
	F	0.957(163)	1.000(195)	0.678(135)	0.852(165)	1.000(120)
Est-6	M	0.038(169)	0.070(185)	0.376(141)	0.131(168)	0.000(121)
	F	0.034(161)	0.075(194)	0.358(169)	0.145(165)	0.000(120)
Adh	M	0.156(141)	0.109(184)	0.740(167)	0.866(157)	0.512(41)
	F	0.163(166)	0.115(195)	0.762(168)	0.909(149)	0.500(32)
∝-Gpdh	M	0.312(85)	0.435(153)	0.635(137)	0.117(94)	0.409(110)
	F	0.369(134)	0.459(160)	0.543(164)	0.132(87)	0.391(115)
G-6-Pdh	M	0.516(93)	- *	- *	- *	- *
	F	0.483(89)	- *	- *	- *	- *
		*:no scorable gels				
6-Pgd	M	1.000(145)	0.837(92)	1.000(144)	0.860(157)	1.000(120)
	F	1.000(141)	0.704(113)	0.929(169)	0.891(151)	1.000(120)

Table 5.1(c): Summary of Gene Frequency Data: Generation 9, Block 2

Locus	Sex	Mean Gene Frequency (no. in sample) Treatment A	B	C	D	E
Pgm	M	0.814(129)	1.000(116)	0.809(115)	0.882(110)	1.000(95)
	F	0.782(133)	1.000(115)	0.845(116)	0.902(112)	1.000(96)
Est-6	M	0.088(130)	0.056(116)	0.000(110)	0.000(110)	0.226(95)
	F	0.125(132)	0.030(115)	0.000(120)	0.000(117)	0.297(96)
Adh	M	0.572(97)	0.555(73)	0.800(80)	0.669(86)	0.506(77)
	F	0.543(94)	0.591(77)	0.791(86)	0.702(94)	0.431(80)
∝-Gpdh	M	0.409(127)	0.303(104)	0.513(114)	0.523(108)	0.759(93)
	F	0.386(127)	0.311(114)	0.566(113)	0.478(112)	0.632(72)
G-6-Pdh	M	0.400(115)	0.500(76)	0.740(104)	0.455(77)	0.161(62)
	F	0.505(103)	0.556(72)	0.472(108)	0.388(76)	0.617(77)
6-Pgd	M	1.000(181)	0.839(112)	1.000(118)	1.000(112)	1.000(95)
	F	1.000(158)	0.741(114)	0.915(118)	1.000(117)	1.000(96)

Table 5.1(d): Summary of Gene Frequency Data: Generation 19, Block 2

Locus	Sex	Mean Gene Frequency (no. in sample) Treatment A	B	C	D	E
Pgm	M	0.931(102)	1.000(167)	0.136(103)	1.000(118)	1.000(120)
	F	0.935(107)	1.000(168)	0.246(118)	1.000(120)	1.000(110)
Est-6	M	0.058(120)	0.000(167)	0.000(108)	0.000(99)	0.597(62)
	F	0.071(119)	0.000(168)	0.000(120)	0.000(96)	0.627(51)
Adh	M	0.479(119)	0.476(166)	0.523(108)	0.716(102)	0.702(104)
	F	0.525(118)	0.524(168)	0.438(120)	0.750(106)	0.691(97)
∝-Gpdh	M	0.339(109)	0.195(128)	0.391(101)	0.449(99)	0.691(76)
	F	0.221(111)	0.158(92)	0.549(113)	0.486(104)	0.750(72)
G-6-Pdh	M	0.426(108)	0.128(94)	0.762(844)	0.762(105)	1.000(69)
	F	0.554(102)	0.282(94)	0.700(95)	0.662(117)	0.775(91)
6-Pgd	M	1.000(120)	0.000(167)	0.991(106)	1.000(118)	1.000(119)
	F	0.895(119)	1.000(168)	0.954(120)	1.000(120)	1.000(110)

fixed at 1 locus (Pgm) in block 1, and two loci (Pgm and Est-6) in block 2. Treatment C was unchanged, in fact showing variation at 6-Pgd that had not been detected at generation 9. The only change in treatment D was that Pgm became fixed in block 2. Treatment E became fixed for Pgm in block 1.

It is perhaps surprising that treatment E, a random mating, single population, should show less variation than the subdivided populations, at least in terms of loci polymorphic. This could be due to the combination of several factors; sampling having unpredictable effects in different populations, the small number of loci sampled, and the possibility that selection for bodyweight may have assisted in the retention of genetic variation at these loci via the "hitch-hiking" effect.

Table 5.2 shows the changes in gene frequency at each locus for all five treatments from generation 9 to 19 in both blocks. Most of the changes are small ($<|0.10|$), and not dissimilar between blocks. No significance can be attached to these changes; there are no consistent patterns, and changes as large as any were seen in Treatment E, where no selection was taking place.

Table 5.2: Changes in Gene Frequency from Generation 9 to 19

Treatment	Block	Pgm	Est-6	Adh	α-Gpdh	G-6-Pdh	6-Pgd
A	1	+0.16	-0.30	-0.29	-0.04	+0.05	0.00
	2	+0.14	-0.04	-0.06	-0.12	+0.04	-0.05
B	1	+0.07	-0.14	-0.16	+0.05	-	+0.31
	2	0.00	-0.04	-0.07	-0.12	-0.32	+0.21
C	1	+0.12	-0.03	+0.22	-0.15	-	-0.04
	2	-0.63	0.00	-0.32	-0.07	+0.13	+0.01
D	1	+0.02	-0.02	+0.43	-0.37	-	-0.05
	2	+0.11	0.00	+0.05	-0.02	+0.29	0.00
E	1	+0.08	0.00	-0.12	+0.10	-	+0.01
	2	0.00	+0.35	+0.23	+0.01	+0.15	0.00

Of interest in the context of the search for a major gene(s) in Treatment D, block 1, is the simultaneous increase at Adh and decrease at α-Gpdh (located on the two arms of the second chromosome), for that treatment, to a greater extent than observed in any other treatments in block 1. This could be produced by sampling via selection of second chromosomes carrying a major gene(s) and being homozygous F for Adh and S for α-Gpdh. Unfortunately, the evidence discussed in Chapter 4 does not clearly suggest involvement of second chromosomes in that effect.

The experimental populations represent samples from one population subjected to various sampling regimes. The extent to which those regimes have caused genetic differentiation among samples can be measured by genetic identity, or distance measures. Nei's (1972) Genetic Identity and Distance measures for the five treatments within each block and generation are presented in Table 5.4, and for the four cases of each population treatment in Table 5.3 (following pages). Looking first within treatments (Tables 5.3), in general, genetic identities were high, and higher within blocks over time than between blocks. This is logical, since there is an autocorrelation between the set of frequencies within a population over time. The identities within treatment A were consistently high, those of the other treatments slightly lower.

In the Tables 5.3 and 5.4, samples are described as follows: G9.1 : Generation 9, Block 1 G9.2: Generation 9, Block 2 G19.1 : Generation 19, Block 1 G19.2 : Generation 19, Block 2

Table 5.3(a): Genetic Identities within Treatment A

(Standard Error)

	G9.1	G19.1	G9.2	G19.2
G9.1	-	0.958 (0.253)	0.986 (0.270)	0.977 (0.241)
G19.1	0.043	-	0.957 (0.285)	0.973 (0.271)
G9.2	0.014	0.043	-	0.990 (0.256)
G19.2	0.024	0.028	0.010	-

Table 5.3(b): Genetic Identities within Treatment B

(Standard Error)

	G9.1	G19.1	G9.2	G19.2
G9.1	-	0.894 (0.262)	0.941 (0.268)	0.872 (0.265)
G19.1	0.112	-	0.887 (0.293)	0.931 (0.231)
G9.2	0.061	0.119	-	0.965 (0.274)
G19.2	0.136	0.079	0.035	-

Table 5.3(c): Genetic Identities within Treatment C

(Standard Error)

	G9.1	G19.1	G9.2	G19.2
G9.1	-	0.797 (0.389)	0.902 (0.229)	0.910 (0.256)
G19.1	0.227	-	0.876 (0.242)	0.768 (0.320)
G9.2	0.103	0.132	-	0.879 (0.365)
G19.2	0.094	0.264	0.129	-

Table 5.3(d): Genetic Identities within Treatment D

(Standard Error)

	G9.1	G19.1	G9.2	G19.2
G9.1	-	0.886 (0.205)	0.920 (0.280)	0.962 (0.233)
G19.1	0.121	-	0.948 (0.215)	0.947 (0.210)
G9.2	0.084	0.053	-	0.976 (0.305)
G19.2	0.038	0.053	0.024	-

Table 5.3(e): Genetic Identities within Treatment E

(Standard Error)

	G9.1	G19.1	G9.2	G19.2
G9.1	-	0.958 (0.280)	0.960 (0.295)	0.808 (0.351)
G19.1	0.043	-	0.940 (0.289)	0.733 (0.434)
G9.2	0.041	0.061	-	0.911 (0.348)
G19.2	0.063	0.310	0.093	-

There was a wider range in genetic identities between treatments than within, and some consistent features emerged (Tables 5.4,a-d). Firstly, the genetic identities tended to be higher (i.e. distances smaller) at generation 9 than generation 19. Nei (1972) has pointed out that this is to be expected if the alleles involved are not subject to differential selection. Secondly, identities with Treatment C were generally the lowest (conversely, distances from C the highest).

The averages over all identities with other treatments and samples for the five treatments were:

A = 0.924 B = 0.896 C = 0.876 D = 0.917 E = 0.901

There is no obvious reason why treatment C should have been more differentiated from all other treatments than any other.

Table 5.4: Nei's Genetic Identity and Distance Measures Among the Experimental Populations

(a) Generation 9, Block 1

	A	B	C	D	E
A	-	0.90 (0.18)	0.90 (0.30)	0.94 (0.25)	0.97 (0.25)
B	0.11	-	0.81 (0.09)	0.91 (0.18)	0.91 (0.24)
C	0.10	0.21	-	0.95 (0.24)	0.84 (0.29)
D	0.07	0.10	0.05	-	0.93 (0.27)
E	0.03	0.09	0.18	0.08	-

(b) Generation 9, Block 2

	A	B	C	D	E
A	-	0.98 (0.22)	0.97 (0.27)	0.99 (0.28)	0.97 (0.28)
B	0.03	-	0.96 (0.25)	0.97 (0.27)	0.95 (0.28)
C	0.03	0.04	-	0.99 (0.28)	0.94 (0.26)
D	0.01	0.03	0.01	-	0.97 (0.29)
E	0.03	0.05	0.06	0.03	-

(c) Generation 19, Block 1

	A	B	C	D	E
A	-	0.93 (0.23)	0.81 (0.28)	0.86 (0.32)	0.92 (0.30)
B	0.08	-	0.85 (0.30)	0.78 (0.31)	0.95 (0.26)
C	0.21	0.16	-	0.86 (0.20)	0.93 (0.27)
D	0.15	0.24	0.14	-	0.89 (0.22)
E	0.08	0.05	0.07	0.12	-

(d) Generation 19, Block 2

	A	B	C	D	E
A	-	0.98 (0.26)	0.85 (0.38)	0.97 (0.29)	0.84 (0.33)
B	0.02	-	0.78 (0.42)	0.92 (0.34)	0.76 (0.40)
C	0.17	0.25	-	0.84 (0.40)	0.74 (0.36)
D	0.03	0.09	0.17	-	0.90 (0.32)
E	0.17	0.28	0.31	0.11	-

NB: Throughout Table 5.4, Genetic Identities are above the diagonal, Genetic Distances below.

Measures of genetic similarity (or divergence) can provide a guide for taxonomic and evolutionary comparisons. Identities will be highest for sets of local populations, lower for races, lower still for sub- or sibling species and so on. Using the values of Ayala and Valentine (1979), most of the comparisons suggest relationhip at the level of local populations, but at the later sampling, particularly for treatment C, some of the identities are close to the value for sub-species. This shows that even medium-sized populations under-going similar selection can diverge appreciably, as a result of drift.

In terms of the principal aim of this study, to compare the effects of population structure on responses to artificial selection, the distribution of genetic variation between units of population in the different treatments is of great significance. Descriptions of these distributions are provided by the various sets of statistics discussed in the literature review. Values obtained from this data summed over loci, are presented in Tables 5.5-5.7 (following pages), and complete lists of Wrights' and Weir and Cockerham's statistics for individual loci in Appendix B. The general view obtained from all three sets of statistics will be discussed first, with a brief comparison of the three to follow.

Examining the values of F_{ST} (and θ_W) (Tables 5.5, 5.6a and 5.7a), there was reasonably good agreement within treatments, across blocks and generations. The lowest values of F_{ST} were found in treatments A and E, which

would be expected since no subdivision was imposed in either of these treatments. Generally the agreement between sexes for these treatments was good. The subdivided treatments showed generally higher values of F_{ST}, with those for Treatment C being somewhat higher than either B or D. In terms of the migration patterns applied in treatments C and D, this suggests that the variable migration model (D) leads to less genetic differentiation between subpopulations/demes than the circular stepping-stone model (C).

The similarity in F_{ST} between treatments B and D suggests that selection among inbred sub-lines followed by further sublining (B) does not produce as great a degree of between-line variation as does permanent semi-isolation (C). Nei's F_{ST}, and 0_W values for treatment B were lower at generation 19 than 9, suggesting a reduction in diversity of sub-lines as a result of two episodes of between-line sampling between generations 9 and 19. No clear change was evident for treatment D, while all three sets of statistics showed a marked increase in F_{ST} in treatment C over this period.

This similarity between F_{ST} values at generations 9 and 19 for treatment C suggests that while migration (in the circular stepping-stone pattern) was included in this treatment, it did not necessarily result in effective gene flow. The relative values of F_{ST} for the five treatments were in agreement

Table 5.5: Nei's (1977) Modified F-Statistics:: F_{ST} Summed Over Loci

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.101	0.044	0.083	0.072
	F	0.255	0.034	0.074	0.131
B	M	0.176	0.198	0.136	0.062
	F	0.260	0.233	0.104	0.097
C	M	0.265	0.143	0.274	0.380
	F	0.161	0.152	0.320	0.458
D	M	0.355	0.062	0.159	0.094
	F	0.176	0.060	0.151	0.093
E	M	0.081	0.072	0.013	0.013
	F	0.040	0.072	0.010	0.026

Table 5.6(a): Wright's F-Statistics Summed over Loci: F_{ST}

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.152	0.089	0.139	0.155
	F	0.163	0.090	0.111	0.215
B	M	0.231	0.323	0.284	0.156
	F	0.596	0.348	0.187	0.256
C	M	0.243	0.453	0.507	0.736
	F	0.244	0.336	0.450	0.692
D	M	0.418	0.311	0.236	0.230
	F	0.305	0.326	0.238	0.229
E	M	0.282	0.151	0.047	0.034
	F	0.133	0.151	0.039	0.054

Table 5.6(b): Wright's F-Statistics Summed over Loci: F_{IS}

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.250	0.463	0.538	0.326
	F	0.078	0.268	0.358	0.326
B	M	0.413	0.432	0.569	0.607
	F	0.243	0.282	0.448	0.451
C	M	0.394	0.586	0.257	0.615
	F	0.151	0.330	0.265	0.504
D	M	0.474	0.513	0.376	0.603
	F	0.391	0.449	0.161	0.400
E	M	0.685	0.423	0.290	0.233
	F	0.481	0.274	0.286	0.041

Table 5.6(c): Wright's F-Statistics Summed over Loci: F_{IT}

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.283	0.470	0.359	0.479
	F	0.109	0.270	0.201	0.175
B	M	0.360	0.357	0.657	0.645
	F	0.105	0.220	0.513	0.459
C	M	0.469	0.631	0.571	0.757
	F	0.192	0.260	0.392	0.719
D	M	0.433	0.550	0.359	0.678
	F	0.177	0.493	-0.185	0.446
E	M	0.705	0.427	0.474	0.242
	F	0.493	0.349	0.466	0.056

Table 5.7(a): Weir and Cockerham's θ_W Summed Over Loci

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.126	0.070	0.101	0.088
	F	0.124	0.059	0.079	0.117
B	M	0.242	0.257	0.175	0.117
	F	0.435	0.343	0.132	0.244
C	M	0.314	0.355	0.456	0.906
	F	0.183	0.249	0.413	0.675
D	M	0.349	0.193	0.200	0.193
	F	0.227	0.194	0.132	0.194
E	M	0.195	0.125	0.039	0.016
	F	0.083	0.120	0.030	0.031

Table 5.7(b): Weir and Cockerham's F Averaged Over Loci

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.196	0.343	0.436	0.267
	F	-0.011	0.092	0.025	0.092
B	M	0.319	0.227	0.409	0.181
	F	0.349	0.128	0.266	-0.017
C	M	0.331	0.462	0.423	0.675
	F	0.011	0.119	0.387	0.720
D	M	0.355	0.269	0.261	0.329
	F	0.153	0.156	-0.293	-0.080
E	M	0.518	0.061	0.504	-0.513
	F	0.076	0.053	-0.690	-0.415

Table 5.7(c): Weir and Cockerham's f Averaged Over Loci

Treatment	Sex	Population/Sample			
		G9.1	G9.2	G19.1	G19.2
A	M	0.080	0.295	0.373	0.309
	F	-0.155	0.035	-0.058	-0.028
B	M	0.102	-0.040	0.284	0.072
	F	-0.152	-0.328	0.155	-0.345
C	M	0.023	0.166	-0.061	-2.464
	F	-0.210	-0.173	-0.045	0.140
D	M	0.010	0.094	0.077	0.169
	F	-0.095	-0.048	-0.490	-0.341
E	M	0.401	0.074	-0.632	-0.542
	F	-0.007	-0.076	-0.649	-0.460

with the distributions of phenotypic and genetic variance estimated in Chapter 3. The differences between treatments in F_{ST} were clearer than those from the estimated phenotypic partitioning; this presumably reflects the absence of any environmental component in the F_{ST} values.

Examination of the values of F_{IT} (only Wright's and Weir and Cockerham's are presented because of the distortionary effect on Nei's F_{IT} of fixed loci) reveals no obvious pattern over treatments. Both sets of estimates (Wright's and Weir and Cockerham's) suggest moderate to high levels of inbreeding in the populations as a whole. There was no clear pattern of increase in F_{IT} from generation 9 to 19, except in treatments B and C, and even in these the pattern differed between sexes. In all treatments at generation 9, and at least A and D at generation 19, responses to selection were still occurring. Thus while the level of inbreeding estimated from electrophoretic loci was high, this was either not the case at loci influencing bodyweight, or had not reached sufficient levels to restrict responses to selection. This observation agrees with the deductions of Lewontin (1984) and the experimental observations of Giles (1984). The high F_{IT} levels of treatment E, more marked in Wright's statistics, as compared with the other treatments, suggest that the selection applied in treatments A-D may, via the hitch-hiker effect, have maintained genetic variation throughout the genome.

Inbreeding within subpopulations, estimated by F_{IS}, was less affected by population structure than either F_{ST} or F_{IT}. No clear distinction between the subdivided and single population treatments was evident. This suggests that the levels of gene flow into demes/subpopulations, though quite different for the different treatments, were equally effective in maintaining genetic variation within the demes. Had the level of migration been reduced in the subdivided treatments from 10%, to 5% or lower, differences in F_{IS} between treatments may have become apparent. Wright (1978) and others have stated that levels of migration of one immigrant every other generation are sufficient to maintain populations effectively panmictic. While the subdivisions applied here did produce inter-deme genetic variation, in all treatments genetic variation remained within individual demes. Since in all cases this was the primary area of selection, with different methods and amounts of between-deme selection superimposed, there was, at least for allozyme loci, ample genetic variation available for within-deme selection. (Once again it must be stressed that the existence and estimated level of allozymic variation provides no real

indication of the situation at quantitative loci).

F_{IT} and F_{IS} values differed markedly between sexes. This is largely due to the inclusion of sex-linked loci in the sample and in the estimates summed over loci. Where genetic variation exists at a sex-linked locus, descriptions of that variation and its partitioning between population levels which are based on, or include, observed and/or expected amounts of heterozygosity, will underestimate genetic variation present in the hetero-chromosomal sex since no heterozygotes will be present. In this study, the most likely explanation for differences in F_{IT} or F_{IS} at the autosomal loci is sampling; larger numbers would be necessary before sex-differences in selection coefficients became statistically significant. The differences between sexes observed in F_{IS} and F_{IT} do not affect F_{ST} and θ_W, since neither is based upon observed heterozygosity (Weir and Cockerham's θ_W includes observed heterozygosity in both the numerator and denominator, in both cases as components of statistics including other quantities unaffected by sex-linkage).

Estimates of effective migration rates between demes can be obtained from the formula for expected F_{ST} (Wright, 1951):

$$F_{ST} = \frac{1}{4Nm + 1}$$

Values of effective migration rate, m_e, obtained using the average F_{ST} (or θ_W) values for each treatment and effective population sizes (for demes) of 9, are presented in Table 5.8 (over page). The three sets of estimates show reasonable agreement, particularly those based on Wright's F_{ST} and Weir and Cockerham's θ_W. The m_e estimates estimates based on Nei's F_{ST} are mostly larger than the others, but all agree fairly well with what might be expected from the population structures applied. The highest estimated migration rate was in Treatment E, where there was random allocation of parents to bottles each generation.

The same mating structure was applied in treatment A, with the additional factor of selection of those parents. From Weir and Cockerham's θ_W, this additional sampling factor seems to have produced only a very small increase

in genetic differentiation over that observed for treatment E, and consequently only a small decrease in effective migration rate, but on the basis of Nei's F_{ST}, there was a lower effective migration rate between bottles/demes in treatment A than in E. If this was in fact the case, it would

Table 5.8: Estimates of Effective Migration Rates from F_{ST}/θ_w Values

Treatment	F_{ST} Estimation Procedure					
	Wright's		Nei's		Weir and Cockerham's	
	F_{ST}	m_e	F_{ST}	m_e	θ_w	m_e
A	0.139	0.172	0.099	0.252	0.096	0.263
B	0.298	0.066	0.158	0.148	0.243	0.086
C	0.458	0.033	0.269	0.075	0.444	0.035
D	0.287	0.069	0.144	0.165	0.210	0.104
E	0.111	0.222	0.041	0.652	0.080	0.320

presumably have arisen from departures from random sampling in allocating parents, although it is hard to see how a consistent departure from randomness could arise. In fact, the subpopulations might be expected to be more similar in Treatment A, due to the effect of selection in reducing genetic variation among the selected sample (Falconer, 1981).

The estimated effective migration rates for the three subdivided treatments were all lower than for treatments A and E. The lowest rate was in treatment C, a circular stepping-stone model with migration of one female (selected) per generation. It should be noted here that since demes finally analysed in the electrophoretic survey were chosen mainly to give even spread over the population systems, rather than randomly, and that this was usually done by using demes 1,3,5,7 and 9, the degree of genetic differentiation between demes in treatment C may overestimate the average degree of differentiation that existed between any random pair of demes. This would not occur with any other treatment, since migration was random with respect to deme number in those. (Treatment C is the one treatment in this comparison

where spatial autocorrelation techniques could be applied usefully). Since in the absence of more extensive data it would be very difficult to accurately assess the effect of this sampling pattern on estimates of F_{ST}/θ_W for treatment C, the estimates will be treated as valid although perhaps too large. From both Nei's and Weir and Cockerham's estimates, the effective migration rate for treatment C was less than half that of the other subdivided treatments. In addition to the sampling method already mentioned, this suggests that drift was more important in this treatment, and that the migration imposed was less effective in "binding" gene frequencies.

The estimate of effective migration rate for treatment B suggests that by the end of the period of sublining, the degree of differentiation between demes was quite marked (estimates of 0.158 and 0.243 for F_{ST}/θ_W, analogous to Slatkin's (1981) populational heritability). Continuing the analogy with populational heritability, the accuracy of inter-deme selection on the basis of the F_{ST}/θ_W estimates would be in the range 40%-49%, which is not high enough to support responses to such selection. If estimates of the degree of genetic differentiation for this population structure are accurate, they suggest that failure to find advantage in sub-lining-with-crossing regimes in this and previous experiments may be due to insufficient inter-deme variation for selection at this level to be effective.

The estimated migration rate for treatment B was slightly lower than that for D, suggesting that the mixing of lines every five generations in treatment B was a less effective cohesive force than the migration pattern in treatment D. The estimates of 10% and 16% for effective migration in treatment D are the highest of any of the subdivided treatments, and at least for Weir and Cockerham's θ_W, very close to the imposed rate of 10%. On a simply additive basis, the resulting degree of genetic differentiation was too low to allow accurate inter-deme selection (accuracy 38%-46%). In this light, the results of the selection experiment could be interpreted as providing no evidence of any disadvantage in subdivision using the treatment D model, and in fact that provided sufficient inter-demal migration is allowed, panmixis effectively

prevails.

Mention of the different sets of statistics leads to brief discussion of differences between the descriptions provided by the three of the observed genetic variation. Agreement between estimators of F_{ST} was fairly good, with closer agreement over the entire set of treatments and population samples between Nei's and Weir and Cockerham's statistics, than between either of these and Wright's F_{ST}. The agreement between all three was best when deme sample sizes were more even, and the variance of gene frequencies among demes small. Departures from these conditions tended to produce lower values for Wright's F_{ST}, than for Nei's F_{ST} or Weir and Cockerham's θ_W. These conclusions based on estimates summed over loci hold also when individual loci are considered. Weir and Cockerham presented simulation results estimating F_{ST} (= θ_W) in a number of ways, and found the traditional (Wright's) F_{ST} to be consistently biased upwards in comparison with their own θ_W.

When Wright's and Weir and Cockerham's F_{IT} are compared, the most obvious difference is that the estimates of overall inbreeding by Wright's method were in almost all cases higher than Weir and Cockerham's. The best agreement was again seen where sample sizes were similar among demes, and the average gene frequencies intermediate, although this was not a universal condition. For individual loci, the agreement between estimates was quite good, although Wright's F_{IT} tended to be larger than Weir and Cockerham's. This points to a part of the difference when values are summed over loci being due to sex-linked loci, at least in males, where F_{IT} (and F_{IS}), according to Wright, will always be 1.0, but to Weir and Cockerham, 0.0.

A similar situation is found for the estimates of F_{IS}: both summed over loci, and at individual loci. F_{IS} via Wright was generally higher than via Weir and Cockerham. For individual loci, agreement was again best where sample sizes were even across demes, and gene frequencies were intermediate. As with F_{IT}, exceptions to this tendency can be found. Even if the values for the sex-linked loci are excluded, the overall F_{IS} via Wright was still higher than via Weir and Cockerham.

The use of the F_{ST}/θ_W to provide estimates of effective migration rates,

which can then be compared with expected rates, allows a very crude test of their relative performance. On this basis, Nei's and Weir and Cockerham's statistics are more useful than Wright's. When the estimates of F_{IS} and F_{IT} are included in the comparison, the statistics introduced by Weir and Cockerham appear to handle both the inclusion of sex-linked loci and the unbalanced data set, better than those of Nei and Wright. In their paper introducing these statistics, Weir and Cockerham (1984) pointed out that different assumptions were usually made regarding sample sizes (or they were ignored), and then compared the performance of a number of estimtors of F_{ST} using simulation. Over a range of conditions, their θ_W showed the least bias in a majority of cases, and compared favourably in terms of its mean square error, particularly at higher values of 0. The results obtained in this survey suggest that their estimates perform satisfactorily for real population data.

These observations regarding the behaviour of the three sets of estimates can be summarised as follows: for this data set, Weir and Cockerham's method suggests lower levels of inbreeding at both total population and individual deme level, and a greater degree of inter-deme genetic differentiation, than Wright's method, with Nei's producing estimates intermediate between the two. Differences in derivation result in different results for sex-linked loci, but these cannot explain all the observed differences, since the pattern was seen in both sexes. Especially in the case of F_{ST}, the estimates obtained via Nei's and Weir and Cockerham's procedures seem more realistic than via Wright's.

Finally, the estimates of genetic differentiation are important in the context of the aims and design of the selection experiment from which the samples for electrophesis were taken. The level of F_{ST}/θ_W determines the accuracy of inter-deme selection, and on the basis of the estimates presented, would range from c. 45% for treatments B and D, to c. 60-70% for treatment C (from θ_W). The latter level would support effective selection among demes, but the levels observed in treatments B and D are slightly too low. Paradoxically, inter-deme selection was not applied in treatment C, but was applied in both treatments B and D. This raises the question of whether the inbreeding applied

in treatments B and D should not have been more intense, producing more genetic differentiation and hence more accurate inter-demal selection. An alternative approach would be to modify the selection procedures to incorporate all the genetic information available, in the manner of James (1966b). On a long-term basis, this might require regular estimation of genetic structuring.