GENETIC AND MORPHOLOGICAL VARIATION IN THE MUSSEL MYTILUS IN NEWFOUNDLAND

JENNIFER ANN BATES, B.Sc.

GENETIC AND MORPHOLOGICAL VARIATION IN THE MUSSEL MYTILUS IN NEWFOUNDLAND

BY

© JENNIFER ANN BATES B.Sc.

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science

Department of Biology
Memorial University of Newfoundland

April 1992

St. John's Newfoundland

ISBN 0-315-78108-4

Abstract

The genus Mytilus in Newfoundland is thought to be differentiated into two taxa, presently called M. trossulus and M. edulis. In the present study a survey of the distribution of these two taxa was undertaken. Ten sites along the east coast of Newfoundland were sampled; two of these sites were sampled on a finer scale. Four partially diagnostic enzyme loci, Pgm, Est, Lap, and Ap differentiated two groups. Both groups were found at most sites. Heterozygote deficiencies (relative to Hardy Weinberg Equilibrium) were found in one or more loci at all but one site. However the presence of some heterozygotes suggests that the two taxa interbreed. Pgm, Est, and Lap genotypes were significantly associated and together were useful in separating the two taxa.

Inner shell colour was correlated with enzyme genotype, M. trossulus being relatively darker than M. edulis, but this failed to be a robust taxonomic character across all sites. In a fine scale analysis at Traytown, genotype frequency was found to be correlated with increasing shell length. This suggests variation in recruitment, growth and/or mortality between the two taxa.

Acknowledgements

I would like to thank my supervisor, Dr. David J. Innes for his support and guidance in all matters throughout this thesis. In addition thanks to Dr. Ray Thompson and Dr. Pat Dabinett for input and suggestions. A special thank-you to friends and family for their support throughout this degree.

This research was funded in part by a Fellowship from Memorial University and through grants to Dr. D. J. Innes.

Contents

List of Tables

List of Figures

List of Appendices

Introduction

1.1 Mytilus taxonomy

The smooth-shelled edible blue mussels of the genus Mytilus have been the centre of recent taxonomic debate. They are bivalves found commonly in the intertidal and shallow nearshore waters in many parts of the world. Early classifications based on external shell morphology tended to be complex and confusing. For example, Soot-Ryen (1955) lists 17 species and varieties of mussel from the west coast of North America which had been defined on the basis of external shell morphology and proposed that these all be considered as Mytilus edulis. Many of the proposed varieties or species were based on the collection of a few unusual individuals. Mytilids are known to be highly plastic phenotypically (Seed 1978, 1969a). The shell shape can be influenced profoundly by variations in growth rates due to biological factors (i.e. densities of mussels) and environmental factors (i.e. temperature, substrate).

The view that all edulis-like mussels form one species persists. This is largely based on overlap in morphological characters and on the fact that a long planktonic larval stage of 4-6 weeks results in high dispersal. This should result in extensive genetic exchange between geographically separated populations which would serve to counteract genetic differentiation through random genetic drift or selection (Scheltema 1971).

Genetic evidence from protein electrophoresis suggests that Mytilus edulis comprises three genetically distinct taxa, M. edulis L., M. galloprovincialis L. and M. trossulus Gould (Koehn 1991, Varvio et al. 1988). The status of these taxa as species or subspecies is still a matter of controversy (Gardner 1992, Gosling 1984). Several partially diagnostic enzyme loci have been used to map a rough world-wide distribution (McDonald et al. 1991, Koehn 1991, Varvio et al. 1988). A locus is considered diagnostic if an individual can be assigned to the correct species, (as defined by morphological or biochemical characters or on the basis of other diagnostic loci) with a probability of at least 99% (Avise 1974).

The species M. edulis is proposed for the form of Mytilus found in the North and South Atlantic Ocean, Arctic Ocean, White Sea and North Sea (Koehn 1991, McDonald et al. 1991).The species M. galloprovincialis is proposed for the form in the Mediterranean and southward in the eastern Atlantic region, northward in the Atlantic as far as northern France, southwest England and southern Ireland, on the west coast of North America south of Oregon, in the Pacific Ocean in Japan, China and Australia, and in the Indian Ocean (Coustau et al. 1991, Koehn 1991, McDonald et al. 1991, McDonald et al. 1990, McDonald and Koehn 1988). The proposed range of M. trossulus is the North Atlantic in

Canada, Scandinavia and Russia, the North Pacific in Russia and Canada and the Arctic Ocean (Varvio et al. 1988, McDonald and Koehn 1988, Koehn at al. 1984, McDonald et al. 1989). There is a debate as to whether the trossulus - type is also present in the Baltic Sea or whether Baltic mussels represent a M. edulis population adapted to the low salinity conditions of the area (Varvio et al. 1988, Johannesson et al. 1990, Bulnheim and Gosling 1988).

M. trossulus is the most recently named taxon (McDonald and Koehn 1988) and therefore has not been extensively studied. It is considered to be a cold water species in the northern hemisphere (Koehn 1991), though very little is known of its fine scale distribution or ecology. The Atlantic provinces of Canada are a potential "hybrid zone" between M. edulis and M. trossulus (McDonald et al. 1991) likely to be similar in nature to the well documented "hybrid zone" between M. edulis and M. galloprovincialis in England and France (Skibinski et al 1983, Coustau et al. 1991). The present study uses protein electrophoresis to provide further information on the presence of M. trossulus and M. edulis in Newfoundland.

1.2. Protein electrophoresis as a population genetics tool

Protein electrophoresis is a technique whereby proteins are separated by differential migration in an electrical field according to the net charge and/or molecular size of the molecule. This technique has become a useful tool in population genetics. Some mutations or alternative forms of the protein may not be detectable by electrophoresis because they do not alter the net charge (Hartl 1988). The detectable differences in mobility can be used as genetic markers because these differences are due to alternative alleles at the structural loci which code for the production of these proteins (Murdock et. al. 1975).

Enzyme polymorphisms have been shown to be heritable in Mendelian fashion in Mytilus (Hvilsom and Theisen 1984), so they can be used as genetic markers in the analysis of the breeding structure of populations (Bryant 1974). The allele frequencies can be tested for conformance to Hardy-Weinberg predictions. Deviation from predicted allele frequencies indicates that one of the assumptions of the model is being violated (Hartl 1988). The use of hierarchial F-Statistics can yield information on the population structure of mussels over wide geographic ranges within and among populations.

There are two general theories used to explain how protein polymorphisms can be maintained in populations of

mussels: 1) The selectionist school - The variation is selectively maintained through, for example, physiological variation (for example the cline in Ap-1 polymorphism in M. edulis on the United States east coast is thought to be maintained by natural selection acting on biochemical differences of the Ap-1 electromorphs in different salinities and temperatures (Hilbish and Koehn 1985) or 2) The neutralist school considers most variation physiologically irrelevant and that variation in enzyme structure is maintained as a result of random processes (i.e. mutation and random drift) (Levinton and Lassen 1978, Murdock et al. 1975). It can be difficult to distinguish between genetic drift and selection in a descriptive study - primarily because the selective factors and their strengths will not be known and most survey studies represent a 'snapshot' of the results of complex interactions (Johanesson et al. 1990). This can result in widely different interpretations of the same data (e.g. Varvio et al. 1988 vs Johanesson et al. 1990)

1.3 Mytilus edulis - Mytilus galloprovincialis - a model

Protein electrophoresis has been used to address a complex taxonomic issue in the differentiation of the two taxa Mytilus edulis and Mytilus galloprovincialis in the British Isles, France and Spain - a similar situation to

what is now thought to exist in Newfoundland between M. edulis and M. trossulus. M. edulis and M. galloprovincialis were first established as distinct species based on morphology (Seed 1978) but the great phenotypic variation within each taxon and the presence of intermediate forms, especially in areas where both taxa occur, has left the question of the status of the taxa as "distinct species" unanswered (Gosling 1984). Genetic studies indicate that there are two distinct groups which do hybridise where they co-occur, but to a different extent in different areas (Coustau et al. 1991, Skibinski et al. 1983). The mechanisms which keep the taxa separate in other areas are not well understood though different spawning times have been demonstrated (Gardner and Skibinski 1988, Seed 1971). The fact that no 100% diagnostic character has ever been found, despite the wide variety of techniques applied to the problem (i.e. morphology, gel electrophoresis, cytology, immunology and artificial hybridization), led Gosling (1984) to conclude that M. galloprovincialis and M. edulis are too similar to be considered separate species. The status of a subspecies of M. edulis has been suggested as the most appropriate classification for galloprovincialis (Gardner 1992, Gosling 1984). It is also argued that because the taxa maintain distinct genetic identities despite hybridisation and introgression, it is more appropriate to regard them as

separate species (McDonald et al. 1991).

1.4 Mytilus edulis - Mytilus trossulus

On the northeast coast of North America, M. trossulus was not originally described on the basis of morphological characters but rather by electrophoretic discrimination. Based on allozyme data from five enzyme loci (Ap-1, Ap, Gpi, Odh, Pgm) Koehn et al. (1984) distinguished three genetically distinct groups; Group I - Populations south of Cape Cod; Group II - Populations throughout the Gulf of Maine, Gulf of St. Lawrence, areas of both southern and northern Newfoundland and southern Hudson Bay; Group III - Populations in southeastern Nova Scotia, northern Newfoundland and Hudson Strait, Quebec.

The differentiation Koehn et al. (1984) found between Group I and Group II was due almost entirely to a single locus, Ap-1. This was interpreted as being a cline within a single species maintained by natural selection acting on biochemical differences of the Ap-1 electromorphs in different salinities and temperatures (Hilbish and Koehn 1985). This finding supported earlier work by Koehn et al. (1976).

Group I/ II and Group III are partially differentiated at several loci which was interpreted as evidence that the two groups represent distinct species (Koehn et al. 1984).

Later studies identified the same genetic groups in various parts of the world. The allele frequencies characteristic of Group III were also found in mussels from the North American northwest coast (McDonald and Koehn 1988), the Baltic (Varvio et al. 1988, Bulnheim and Gosling 1988), and the Sea of Japan (McDonald et al. 1989). This widespread distribution of the same allele frequencies was taken by Varvio et al. (1988) as evidence that M. trossulus and M. edulis represent two distinct species. This view is supported by recent work which shows that the mtDNA profiles of the two taxa are significantly different (Zouros et al., unpubl.).

The name M. trossulus has historical precedence from a variety named by Gould (1850) on the northwest coast of North America (McDonald and Koehn 1988). This mussel was distinguished from other mytilids morphologically, primarily on shell characters, it being described as more cylindrical than M. edulis (Gould 1850). Although the type specimen has been lost, the name was revived by McDonald and Koehn (1988) because mussels sampled today from the type locality exhibit allele frequencies characteristic of 'Group III' from previous studies (Koehn et al. 1984, Varvio et al. 1988). The use of this name is recent and its status as a species or subspecies is still debated (Bulnheim and Gosling 1988, Johannesson et al. 1990). McDonald and Koehn (1988) suggest

that M. trossulus is a northern mussel with a distinct hybrid zone with M. galloprovincialis to the south on the North American west coast. (They do not consider M. edulis to be present at all on this coast.) There is a general implication in the literature that trossulus is a 'northern' form (Koehn et al. 1984, Varvio et al. 1988, McDonald and Koehn 1988, McDonald et al. 1989).

The extent of reproductive isolation, (if there is any), of M. trossulus from either M. edulis or M. galloprovincialis has not been extensively studied. Koehn et al. (1984) report an "intermediate population sample" from Bellevue, Newfoundland. This sample was intermediate because it contained a simple mixture of individuals from Group II and Group III populations rather than hybrids. The apparent lack of intermediate forms was taken as evidence that interbreeding between groups II and III does not occur. However McDonald and Koehn (1988) give evidence for interbreeding between M. galloprovincialis and M. trossulus on the west coast of North America. Interbreeding with M. edulis in the Baltic is implied or stated by most researchers (Varvio et al. 1988, Bulnheim and Gosling 1988, Johannesson et al. 1990). This is based on the presence of the 'wrong' combinations of partially diagnostic allozymes (i.e. an individual with an allozyme characteristic of M. edulis at one locus, but an allozyme characteristic of M.

trossulus at another locus), as well as loci having differing patterns of variation (Johannesson et al. 1990).

M. trossulus is always found in a complex spatial pattern with either Mytilus edulis or M. galloprovincialis, (with the possible exception of the Baltic Sea), i.e. there are no clear cut boundaries between the distributions of the two taxa; instead there are pockets of the two types interspersed with each other along a coastline. Two types have also been found in the same bed (Koehn et al. 1984, McDonald and Koehn 1988). This is similar to the situation with M. edulis and M. galloprovincialis in the British Isles (Skibinski et al. 1983). There are no reported cases of all three taxa occurring at the same site. The study of Koehn et. al (1984) had too few sample sites to draw any conclusions about the spatial boundaries between populations of Groups II and III in the Atlantic Provinces and Northern Canada though both trossulus and edulis forms were found in Newfoundland and Nova Scotia.

Genetic variation within each taxon has also been reported. Differentiation of allele frequencies has been found between mussel samples taken within the same bays and even within tens of meters of each other (Levinton and Suchanek 1978, Gartner-Kepkay et al. 1980, 1983, Gosling and McGrath 1990, Murdock et al. 1975). Various environmental factors have been suggested as important including

temperature (Hall 1985 Levinton and Lassen 1978), salinity (Koehn and Mitton 1972, Johannesson et al. 1990), wave exposure (Gosling and Mcgrath 1990, Skibinski et al. 1983), tidal exposure (Gosling and Wilkins 1981), and extremes of conditions over the year (Gartner-Kepkay et al. 1980, 1983). Isolation due to hydrographic features over what would otherwise be considered homogenous areas may also lead to differentiation through genetic drift.

1.5 Morphological - genetic correlations

External shell morphology has been extensively studied in Mytilus in an attempt to find diagnostic characters. Some of the characters listed by Seed (1978) as the most useful in distinguishing M. galloprovincialis and M. edulis are as follows:

1) Length of adductor muscle scar. M. galloprovincialis has smaller anterior and posterior adductor scars. Despite some overlap this difference is quite pronounced in some populations.

2) Length and shape of hinge plate. The hinge plate of M. edulis tends to be shaped as a gentle curve, while M. galloprovincialis tends to have a tight arc. The hinge plate is also longer in M. edulis.

These characters tend to show maximum differentiation between the two types when the mussels are growing in pure

(only one type) populations in uncrowded conditions, with the majority of mussels of a relatively young age/small size. In areas where both types occur they tend to interbreed, (Skibinski et al. 1983), producing intermediate forms (Seed 1978). Mussels which are crowded tend to grow more elongated. Also as mussels increase in size (age) their growth rate decreases causing them to become wider and often incurved giving the shall a hooked appearance (Seed 1978, 1969a). All of these factors result in a convergence of shapes which can make it impossible to accurately score mussels on the basis of a single morphological character (Gosling 1984).

Consequently there have been attempts to use multivariate techniques to distinguish species based on morphology. Ferson et al. (1985) used an elliptic Fourier analysis on a two dimensional image of the whole valve in an attempt to distinguish M. edulis and M. trossulus. This yielded 40 normalised coefficients which described the image. These coefficients can then be treated as individual morphological characters in a multivariate analysis such as principal components and discriminant analysis. Samples from Sunnyside and Bellevue, Newfoundland were collected and genotyped using Pgm and Gpi allozymes. Two groups were then constructed; Group 1 containing all individuals which had either a Pgm^{111} or Pgi^{98} allele in a heterozygous or a

homozygous combination, presumably M. trossulus and Group 2 containing all other individuals, presumably M. edulis. The mean shell shape indicated an overall significant difference in the shape of the valve outline of the two groups. A discriminant analysis incorrectly assigned a mussel to genotype based on shell shape at a rate of 15 - 18%. While this indicates a fairly good ability to separate the electrophoretic groups by shape alone, it is not sensitive enough to replace electrophoresis, which is estimated to have an approximate error rate of 7% (Ferson et al. 1985).

A worldwide study of all three taxa, M. edulis, M. galloprovincialis and M. trossulus, used a multivariate analysis of 18 morphometric variables (McDonald et al. 1991). Sites were specifically chosen as single species sites based on allozyme variation in an attempt to find the characters or combinations which would be most useful in a more in depth study. It was found that no single character provided good discrimination, but the canonical variates from an analysis of all 18 characters separated each species pair very well. M. trossulus and M. galloprovincialis showed some morphological overlap and southern hemisphere M. edulis were somewhat intermediate between northern hemisphere M. edulis and M. trossulus but closer to the northern M. edulis. It is not known however, if these results would apply in an area of hybridisation and there are no data on

whether the differences are genetic, environmental or the result of genetic-environment interaction.

Sperm ultrastructure is another morphological character which may be useful in discriminating among the three species. Hodgson and Bernard (1986) found differences in the length and width of the acrosome and in the nucleus shape in sperm of M. edulis and M. galloprovincialis. An allozyme study by McDonald et al. (1990) confirmed an earlier report of the existence of two separate species in the Sea of Japan and the Avancha Inlet based on differences in sperm ultrastructure (Drosdov and Reunov 1986 cited in McDonald et al. 1990), concluding that mussels from the White Sea are M. edulis, while those from Peter the Great Bay, the Sea of Japan and the Avancha Inlet are M. trossulus. Further studies on these microstructural differences between the mussel types are required to determine if they are consistent from region to region.

The colour of the mantle edge has been used by many different authorities, including Seed (1971, 1972), as a diagnostic character for distinguishing between M. galloprovincialis and M. edulis. M. galloprovincialis typically has a dark purple-violet mantle edge while M. edulis is typically a light yellow-brown. Mytilus edulis found growing in direct or strong sunlight had a greater abundance of darker thicker shells which may be a protective

mechanism against the effect of sunlight on the tissue (Seed 1969b). Gosling (1984) discounts the usefulness of mantle colour as a distinguishing character between these two taxa stating that in some populations over 60% of individuals would have been misclassified if colour alone had been used as a diagnostic character.

The division between body colour and shell colour is artificial, the colour of the shell being correlated with that of the shell secreting organs (Comfort 1951). External shell colours are considered to be of great interest as the ostracum is secreted by the mantle edge and as such the colours of the ostracum represent a linear record of pigment deposition. This contrasts with the formation of the nacreous layer which is secreted continually by the whole outer layer of the mantle (Meglitsh, 1972). As variation in mantle colour has been demonstrated in Mytilus, inner shell colour may be a useful character for distinguishing species provided that it is not under strong environmental influence. It should therefore be possible to keep the shells of individuals used in an allozyme study, to score the shells for colour and determine whether there is a consistent difference in shell colour among allozyme genotypes.

1.7 Size dependent variation in allele frequencies

Changes in allele frequencies with increasing size have been used as natural markers to follow cohorts of mussels (Milkman and Koehn 1977) and have been interpreted as support for the hypothesis of natural selection (Gardner and Skibinski 1988, Boyer 1974). Gardner and Skibinski (1988) give several explanations for this phenomenon in M. galloprovincialis and M. edulis. If the two types of mussels experience differential mortality after settlement, then one would expect allele frequencies to change with time. Depending on the reproductive parameters of each type this could also result in a "historical change" of the population over time with one type replacing another. Furthermore if one type has a greater growth rate or attains a greater maximum size causing a change in the relative proportions of the two types with increasing size, then one would expect a change in the allele frequencies with increasing size.

There are thus two levels on which size dependent variation can be examined. Change within a size class over time (indicating different cohorts growing through and possible replacement of one type with another) and constant(consistent) change of allele frequencies with increasing size (indicating a faster growth rate or differential mortality). In this study evidence is presented of changing allele frequencies with size. However, change

within size classes over time was not examined.

1.8 Objectives

This thesis provides further information on the population structure of Mytilus and on the presence/absence of the mytilid taxon M. trossulus from several sites in eastern Newfoundland. There is an increasing amount of literature suggesting that M. trossulus is a separate northern taxon which may be ecologically different from the 'traditional' M. edulis. (McDonald et al 1991, Varvio et al. 1988, Koehn et al. 1984). Atlantic Canada is thought to be a potential hybrid zone between M. edulis and M. trossulus making it an area of interest for further study (Koehn et al. 1984, McDonald et al. 1991). The possibility of an inferior aquaculture product if M. trossulus is collected (due to decreased meat weights and thin shells (Thompson, pers. comm.)) also suggests that further information on the distribution of the two species should be collected.

Population structure was examined on three levels using data collected by protein electrophoresis. First, it was examined on a macrogeographic scale (hundreds of kilometers) from Notre Dame Bay to Trinity Bay. Secondly microgeographic differentiation on a scale of tens to hundreds of meters was examined at Bellevue and Traytown. Thirdly differentiation within a bed at Traytown was examined with emphasis on

differentiation within and among size classes. This provides information on the abundance of the trossulus type and gives some of the baseline data needed to examine finer scale ecological processes affecting distributions and differences in growth or mortality rates.

In addition, data are presented on correlations between morphology (inner shell colour) and genotype in an atttempt to find a diagnostic morphological character distinguishing the two taxa.

Materials and Methods

2.1 Sampling

Ten sites were sampled along the coast of eastern Newfoundland during an 18 month period in 1989-1990 (Figure 1). Samples consisted of approximately 50-100 individuals taken from the intertidal or very shallow subtidal (i.e. approximately 0.5m of water at low tide) with the exception of one of the Bellevue sites (BV0), where mussels were retrieved by SCUBA at a depth of approximately 3m. Nine of the sites were only sampled once but two sites, Bellevue and Traytown, were sampled more extensively (Figure 2). The sixteen samples taken from the above sites will be collectively referred to as the survey samples.

Individual mussels were picked by hand from throughout a mussel bed rather than removing all individuals from within a quadrat. This allowed for the collection of a representative sample from a mussel bed which can contain thousands of mussels. All mussels used in the allozyme study were measured for shell length (greatest anterior-posterior axis). Mussels used ranged from 16mm - 89mm with a grand mean of 40.51mm (Table 1).

From the preliminary results of the survey, Traytown-3 was found to have a relatively high proportion of *trossulus*, though it was not a pure *trossulus* site as reported by Koehn et al. (1984). On August 22 1990, this site was resampled to

take random samples and determine the relative frequencies of the trossulus type and the edulis type as a function of shell length.

The Traytown-3 site consists of a relatively long and narrow (17m by 1m) bed of mussels on a mud/gravel substrate. Mussels were found on the surface or partially buried in the sediment. Three 1/16m^2 samples were taken from three areas of the bed (Figure 2b) to ensure a random collection of all sizes. All mussels within the quadrat were removed and taken to the Ocean Sciences Centre for holding. Shell length was measured to the nearest millimeter and individuals were placed in 5mm size classes. A subsample of individuals from each size class were electrophoretically typed.

2.2 Electrophoretic Methods

Mussels were held in ambient sea water tanks at the Ocean Sciences Centre until they could be analyzed. The number typed from each locality varied from 19 - 86 with a mean of 29.

Mussels > 30mm in length were dissected and a small piece of digestive gland and adductor muscle tissue taken and pooled for homogenisation. In smaller mussels, 10-30mm in length, the whole body was dissected out of the shell and used. In very small individuals, <10mm in length, the whole animal, shell and soft tissue, was used. There was

difficulty in resolution in the electrophoretic analysis of the two smallest size classes (0-5mm and 6-10mm). This is consistent with previous studies where resolution was found not to be possible in animals with a shell length <10 mm (Gosling and McGrath 1990, Skibinski et al. 1983).

An approximately equal volume of homogenisation buffer (Tris-HCl, pH 8.0 with 20% glycerol) was added to each sample. The sample was homogenised with a glass rod for approximately 30 seconds and centrifuged at 0°C and 10 000 rpm for approximately 5 minutes. The supernatant was then applied to horizontal starch gels using filter paper wicks (McDonald 1985).

Four enzyme loci were investigated - Phosphoglucomutase (PGM E.C. 5.4.2.2), Esterase (EST, E.C. 3.1.1.1), Leucineaminopeptidase (=aminopeptidase-1) (PEP, 3.4.-.-), called Lap in this study, and aminopeptidase (PEP, 3.4.-.-), called Ap in this study. These loci were chosen because they are known to be partially diagnostic between M. edulis and M. trossulus (Koehn et al. 1984). Mannose phosphate isomerase (MPI, E.C. 5.3.1.8), the most diagnostic locus for separating M. edulis and M. trossulus, (McDonald and Koehn 1988), was not used due to difficulty in obtaining clear resolution.

Two of the enzymes, Pgm and Est were run immediately after dissection to avoid the problems associated with the

degradation of proteins during freezing (e.g. ghost bands or extra bands making the interpretation of gels difficult). Pgm was occasionally found to develop ghost bands after samples were frozen and Est could usually not be resolved after sample freezing. The remainder of the sample was stored at -80°C for later comparisons of Pgm and for analysis of the other two enzyme systems, Ap and Lap. Samples still gave good results for Ap and Lap after one year of storage at -80°C.

2.2.1 Buffers and Stains Used in the Electrophoretic Analysis

All buffers and stains used are after McDonald (1985) and Selander et al. (1971). Gels were made of 11% (18.7g in 160 ml gel buffer) Sigma starch (hydrolysed for electrophoresis), and were 14cm x 14cm x 6mm.

Pgm was separated on a tris maleic buffer pH 7.4 (12.12g Tris, 11.61g maleic acid, 3.72g EDTA, disodium salt, 2.03g $MgCl_2$ in 1l water). The gel buffer was a 1:9 dilution of the electrode buffer. Electrophoresis was carried out at 50 - 60 mA for approximately 5 hours at 5°C. The stain was an agar overlay of stain mixture as given in McDonald 1985. All agar overlay stains were saved by placing a sheet of filter paper over the stain and lifting the stain onto the paper. The paper was then dried in an oven overnight. This

stained paper overlay provided an accurate record which could be stored indefinitely.

Est was separated on an acetate buffer pH 5.6, (27.2g sodium acetate in 1l water, acetic acid to pH 5.6). The gel buffer was a 1:19 dilution of the electrode buffer. Electrophoresis was carried out at 50-60 mA for approximately 4 hours at 5°C. This enzyme was resolved with 4-methyl-umbelliferyl acetate dissolved in acetate buffer which was poured over a piece of filter paper placed on the gel (McDonald 1985). After several minutes the paper was removed and the gel viewed under U.V. light. This stain could not be preserved so gels were sketched to provide a permanent record.

Both Lap and Ap were separated on a LiOH buffer, (LiOH-A 1.26g LiOH, 11.75g boric acid in 1l water pH 8.1; LiOH-B 6.10g Tris, 1.54 g citric acid in 1l water, pH to 8.4 with NaOH). The gel buffer was a 1:9 dilution of LiOH-A and LiOH-B (McDonald 1985). Electrophoresis was carried out at 50-60 mA for approximately 3-4 hours at 5°C. Lap was stained using an agar overlay consisting of 15ml acetate buffer, 10mg L-leucine-B-napthylamide, and 2mg fast garnet GBC salt. Ap was stained using an agar overlay of 15ml Tris-HCl, pH 8.0, 20mg DL-leucylglycine, 20mg peroxidase, 10 mg L-amino acid oxidase and 10 mg O-dianisidine (McDonald 1985). The agar overlays were saved as described above.

2.3 Shell Colour

The inner shell colour of the individuals used in the electrophoretic survey varied in a continuum from white to dark grey; many shells had spots or patches of pinks and purples. The outer border of the inner shell was always very dark and was easily distinguished in white and lighter coloured shells but was virtually indistinguishable by colour in the darkest of shells. From the preliminary results of the survey, this internal shell colour seemed to be correlated with genotype - *edulis* being relatively lighter than *trossulus*.

All shells used in this analysis were laid out in a continuum from white to dark on a white background in front of a well lit window. One of the most difficult tasks in dividing the shells into groups was to make evenly spaced intervals between the groups. The Munsell Neutral Value Scale (1971 edition) was used in an effort to score shells as objectively as possible. Four groups were created; white, grey and dark, which were continuous, and pink which contained shells which were entirely purple, pink, red or brown and as such could not be placed into one of the other three groups (see below). The white, grey and dark groups were defined on the Munsell Neutral Value Scale (1971) as follows:

White	9.5 - 8.0
Grey	7.5 - 6.0
Dark	5.5 - 3.0

Many shells had a varying degree of red to purple to brown coloration ranging from small stripes or blotches to covering the entire inner surface of the shell. These colours were easily seen in shells with a light background colour, and although still present, were difficult to distinguish in darker shells. Shells with an obviously white/light background and any amount of pigment stripe/coverage were classified as white.

The inside surface of most shells was glossy, although some were characterised by a 'chalky' (easily scratched) finish. This difference was not investigated but it should be noted that the classification of some of the chalky shells was troublesome as a darker colour could be seen through the lighter chalky layer. Such shells were usually classified as grey.

There were a few shells which showed a discrete change in colour from a light/white colour near the umbo to a darker colour near the margin. As these were rare and difficult to classify by colour they were excluded from the analysis. There were no shells which showed the opposite change in coloration.

2.4 Statistical Analysis

2.4.1 Survey Samples

2.4.1.1 Genetic Variation Within Sites

For each of the sites studied, deviations from the Hardy-Weinberg equilibrium (HWE) were calculated using Levene's correction for small sample size (Biosys-1; Swofford and Selander 1989). Heterozygote deficiencies (D) were calculated as $D = \frac{H_o}{H_e} - 1$

Significance values for D were calculated using X^2 tests of the observed number of genotypes versus the number expected for HWE. These measures give an indication of the level of deviations within a sample.

2.4.1.2 Genetic Differentiation Among Sites

Randomised X^2 tests (Monte Carlo program, REAP, McElroy et al. 1991) were carried out on allele frequencies, (expressed as the number of each allele found in each sample - see Appendix 1), to determine whether there was significant genetic differentiation among samples. The three samples at Traytown, TT1, TT2, and TT3 and the four Bellevue sites BV0, BV1, BV2, and BV3 were also tested for allele frequency differences on smaller scales - tens and hundreds of metres respectively.

Wright's Hierarchical F-Statistics were calculated for

each sample using the Biosys program (Swofford and Selander 1989). There are three F-Statistics which are related by the following equation:

$$F_{IS} = (F_{IT} - F_{ST}) / (1-F_{ST})$$

F_{ST} is a measure of variation among subdivisions, independent of allele frequency. F_{IT} is a measure of the deviation from HWE in the compound population obtained by pooling all subdivisions. It is calculated for each allele by pooling all other alleles of that locus into a synthetic allele. F_{IS} is a measure of deviation from HWE for an individual subdivision and is therefore related to D, the measure of the deficiency of heterozygotes relative to HWE (Skibinski et al. 1983). Defining a statistical test of significance of F_{ST} is difficult so the numbers are generally used as an indication of the magnitude of differentiation while the X^2 test outlined above provides a test for significant differentiation among samples.

Non-metric Multidimensional Scaling (MDS) was used to summarise the pattern of genetic differentiation among localities (Koehn et al. 1984). The analysis was carried out on a matrix of Nei's genetic distance (Nei 1972), based on allele frequencies, in which only individuals for which all four loci had been scored were used. The resulting figures were shown with a minimum spanning tree (MST) superimposed to represent the amount of distortion involved in reducing

the data to two dimensions. A MST connects sites using the original distance matrix such that the total length of the tree is minimised (Sneath and Sokal 1973). NTSYS was used to produce Nei's genetic distance matrix, the MDScale plots and the MST (Rohlf 1988).

2.4.1.3 Association Among Loci

Association among each of the four loci was tested using randomised X^2 tests (REAP, McElroy et al. 1991). A contingency table was constructed by recording the number of individuals of a two locus genotype and a X^2 value calculated (see Appendix 2). The Monte Carlo procedure then randomised the table 1000 times, keeping the row and column totals constant, and calculated a X^2 value each time. A probability value was calculated as the number of times the program randomly generated a X^2 value greater than or equal to the observed X^2 value. This procedure avoids the necessity of making assumptions about the distribution of the data and allows for a test when some cell sizes are < 5. All possible 2-way comparisons of loci were made. (i.e. Pgm - Est, Pgm - Lap, etc.). Interlocus associations were also displayed graphically by plotting allele frequency vs. site because correlated loci should show a similar pattern of allele frequency variation.

Pattern in the allozyme data was explored further in a

principal component analysis carried out on the same individuals (all four loci scored). Individuals were coded for each allele at each locus as 2, homozygous for that allele, 1, heterozygous for that allele and 0, having no copy of that allele (McDonald and Koehn 1988). These data were used to generate the Spearman's rank correlation coefficient matrix which was the basis for the principal components analysis. The data were plotted on the first two principal axes. Histograms of the first principal component values were then plotted for the whole data set as well as for samples from each site.

2.4.1.4 Shell Colour - Genotype Association

A randomised X^2 test (REAP program, McElroy et al. 1991) was used to test independence of genotype and colour at each locus. The contingency table was set up as locus genotype vs. colour. Shell colour was then examined in relation to the observed differentiation of individuals in the principal component analysis based on the allozyme data. The two extreme groups of the colour analysis, dark and white, were qualitatively assessed for correlation with the groupings of individuals for the first principal component values. Colour was also assessed separately at Traytown because this was a site with relatively even proportions of *edulis* and *trossulus*.

2.4.2 Random Traytown Samples

2.4.2.1 Genetic Variation Among Size Classes

Tests of genetic differentiation among size classes were performed using the data from the three random Traytown samples. A randomised Monte Carlo method (REAP, McElroy et al. 1991) was used to test for differences in allele frequencies among classes and for association among loci. The contingency table was set up the number of each allele in each size class (See Appendix 4). D-values were calculated and X^2 probabilities determined as described for the survey sample analysis to test for deviation from HWE within size classes. The principal components analysis described above was repeated using size classes rather than locality as the grouping. This analyses was carried out to determine the pattern of change in genotype frequency with an increase in length.

Results

3.1 Enzyme Variation

3.1.1 Pgm

A total of 6 alleles was recorded for this locus. The slowest allele, Pgm1, probably represented the pooling of several slow alleles with mobilities that were difficult to distinguish. The alleles probably correspond to those reported in Koehn et al. (1984) as follows; Pgm1(89+93+96), Pgm2(100), Pgm3(106), Pgm4(111), Pgm5(114), Pgm6(118) (Table 2). Koehn's alleles 104 and 108 were either not found or were pooled with another allele, probably Pgm3. This is not of concern as both of these alleles were reported in very low frequency, (0.002 - 0.008) in the Koehn et al. (1984) study. Allele Pgm2(100) was the most common. The fastest allele, Pgm6, was not present in most populations and was rare when it was found. Alleles Pgm1 and Pgm5 were more common in some samples but rarely exceeded a frequency of 0.100. Allele Pgm3 was usually present in a frequency greater than 0.100 but did not show a wide range of variation (Table 2).

Koehn et al. (1984) used the frequency of alleles Pgm2(100) and Pgm4(111) as a partially diagnostic character to define their Group II and Group III mussels. These groups have subsequently become known as M. edulis and M. trossulus respectively. McDonald and Koehn (1988) report that M.

trossulus has a greater proportion of faster alleles than M. edulis. Considerable variation was found in the frequencies of the two alleles Pgm2 and Pgm4 in the present study (Table 2).

3.1.2 Est

Est is a dimer forming three banded heterozygotes. Only two common alleles were found, Est1 and Est2 (Table 3). Allele Est8 represents a rare slow allele, slower than Est1, which was not found in the initial samples. Alleles Est3 and Est4 are faster than Est2 and were rare. The alleles probably correspond to those of McDonald and Koehn (1988) as follows; Est8(80), Est1(90), Est2(100), Est3(110). The rare 95 allele reported by McDonald and Koehn (1988) was not found in this study and the rare allele Est4 found in this study was not reported by McDonald and Koehn (1988).

McDonald and Koehn (1988) interpret allele Est1(90) as a trossulus marker and allele Est2(110) as an edulis marker. These are the two common alleles in this study and show a large amount of variation among the samples (Table 3).

3.1.3 Ap

A total of six alleles were scored for this locus. This locus corresponds to that of aminopeptidase (Ap) of Koehn et al. (1984) and the alleles correspond as follows; Ap1(85),

Ap2(90), Ap3(95), Ap4(100), Ap5(105), Ap6(108). The scoring system of Koehn et al. (1984) is somewhat modified in order to maintain consistency with the literature. McDonald and Koehn (1988) report very little differentiation at this locus which is confirmed in the present study, although there was considerable variation among samples (Table 4).

3.1.4 Lap

This locus corresponds to the Ap-1 locus of Koehn et al. (1984) and alleles correspond as follows; Lap1(92), Lap2(94), Lap3(96), Lap4(98), Lap5(100). The allele designations of Koehn et al. (1984) do not represent relative electrophoretic mobilities in the usual sense (i.e. 100 is not the most common allele), but probably represent an effort to maintain consistency with the previously published literature. In the present study alleles Lap2, Lap3 and Lap4 were the most common (Table 5).

3.2 Statistical Analysis

3.2.1 Survey samples

3.2.1.1 Genetic variation within sites

Table 6 gives D-values which are used as a measure of deviation from Hardy - Weinberg expectations (HWE). A negative D-value indicates a deficit of heterozygotes; a positive value indicates an excess; a 0 value indicates no

deviation. Most samples (55/64 tests) show a deficit of heterozygotes. Eight out of sixteen samples showed significant heterozygote deficiency at the Est and Pgm loci. Only one sample showed significant deviation at both Pgm and Est. Only 2 samples showed significant deviation at the Lap locus and 1 at the Ap locus. No sample showed significant deviation at all four loci.

3.2.1.2 Genetic differentiation among sites

Allele frequency differentiation among all samples was significant ($p<0.001$) for the Pgm, Est and Lap loci. Ap did not show significant differentiation among all samples (Table 7). The four Bellevue samples, which were taken within hundreds of meters of each other, also showed significant heterogeneity at the Pgm, Est ($p<0.001$) and Lap loci ($p<0.01$) but not at the Ap locus (Table 7). Finally the three Traytown samples, taken within tens of meters of each other, showed significant genetic differentiation for the Pgm, Est, and Lap ($p<0.001$) loci but again differentiation at the Ap locus was not significant (Table 7).

The hierarchical F Statistics are presented for each allele using pooled data from all sites in Table 8. Mean F_{ST} values are relatively high for Pgm and Est, 0.142 and 0.280, respectively. In contrast Lap and Ap F_{ST} values are low;

0.079 and 0.034 respectively. This implies a greater degree of differentiation among subpopulations at the Pgm and Est loci than at the Ap locus. The same trend is seen with the F_{IS} (the equivalent of the inbreeding coefficient) which is high for Est (0.461), intermediate for Pgm (0.244) and Lap (0.186), and low for Ap (0.026).

Geographic patterns of genetic differentiation are summarised using nonlinear multidimensional scaling. The populations do not fall into discrete clusters in the two dimensional plot (Figure 3). Instead, the populations are stretched out in a line with a small cluster of samples at the bottom right. The minimum spanning tree has no back or criss-crossing of lines connecting samples which indicates that the two-dimensional plot gives a reasonable representation of the genetic distances among samples. The final stress value was less than 0.1 after 40 iterations also indicating that the two dimensional plot is a reasonable representation of the full data set.

The same analysis was performed on each individual locus to determine how much of a contribution each locus made to the overall pattern (Figure 4). Patterns generated by Est and Pgm correspond most closely to the overall pattern in that the samples are positioned in almost the same order relative to each other (final stress values were 0.002 and 0.044 respectively). Lap has some deviations from

this pattern, most noticeably in the relative positions of BV1, BV3, and EM (final stress value = 0.015). The clustering at either end is also not as strong. Ap is the most different with a high degree of scatter and a high level of distortion (final stress value = 0.156).

3.2.1.3 Association among loci

Significant association was found between Pgm and Est ($p<0.001$), and Pgm and Lap ($p<0.05$), (Table 9). All other pairwise combinations were not significantly associated. When the most common alleles of these loci were plotted by sample, the similarity in the patterns of variation of their frequencies illustrated the correlation among loci (Figure 5). The Pgm2, Est2 and Lap4 alleles produced one similar pattern of variation and the Pgm4, Est1 and Lap2 alleles formed another. The two most common Ap alleles were not plotted because they were not similar to either group. The Spearman Rank Correlation coefficients calculated for the principal components analysis (see below) also show these 6 alleles to be significantly correlated ($p<0.05$); Pgm2 - Est2 0.65, Pgm2 - Lap4 0.40, Est2 - Lap4 0.46; Pgm4 - Est1 0.69, Pgm4 - Lap2 0.46, Est1 - Lap2 0.41 (Table 10). This association of loci is expected based on the description of allozyme characters of M. edulis and M. trossulus given by McDonald and Koehn (1988).

The pattern of genetic variation in the survey data was further summarised using principal components analysis based on a Spearman Rank correlation coefficient matrix. Only the 300 individuals from the survey samples for which all 4 loci were scored were included. The principal components analysis explained 19% of the variation on the first principal axis and an additional 10% on the second. The scatter plot of these individuals projected onto the first two principal axes shows a tendency towards a separation of two groups on the first axes, although some individuals fall between the two groups (Figure 6). The highest component loadings are for the alleles Pgm4, Est1, Pgm2 and Est2 (Table 11). The component loadings are interpreted in exactly the same manner as product moment correlation coefficients. The square of their values indicates the proportion of the variance in the individual variable which can be associated with the component (Johnston 1980). Thus 61% of the variance in Pgm2 was explained in the first principal component. Est1 had the highest correlation with 81% of the variation explained by the first principal component (Table 11). The highest loading on the second principal component, which explained 10% of the total variation, was on Ap4 (52% of the variance explained).

The first group was therefore defined by the presence/absence of the Est2 and Pgm2 alleles. These alleles

correspond to the M. edulis alleles of McDonald and Koehn (1988). The second group was defined by the presence/absence of the Est2 and Pgm4 alleles which correspond to the M. trossulus alleles of McDonald and Koehn (1988).

A histogram of the first principal component values for each individual was constructed, (following the procedure of McDonald and Koehn 1988). A bimodal distribution was found with very few individuals with intermediate scores (Figure 7). When the data were plotted according to sample, most samples contained individuals from each of the two peaks (Figure 8). FH and EM seem to be pure peak 1 populations. CC1 and BV2 are dominated by peak 2 mussels though there does not appear to be a pure peak 2 site. Traytown represents a mixture with fairly high proportions of both peaks.

Based on the interpretation of the component loadings, the first peak can be explained by the contribution of the Pgm2 and Est2 alleles which have large negative loadings on the first axis, while the second peak can be explained by the contribution of the Pgm4 and Est1 alleles which have large positive loadings on the first axis. These are the alleles referred to by McDonald and Koehn (1988) as M. edulis and M. trossulus alleles respectively.

3.2.1.4 Shell colour - genotype association

A significant association was found between genotype and shell colour for Lap ($p<0.001$), Pgm ($p<0.001$) and Est ($p<.01$) (Table 12). In addition, the proportion of mussels with dark shells increased with an increase in the first principal component score (i.e. more trossulus like - see above), for both the survey samples and the random Traytown samples (described in following section) (Figure 9). However, the correlation was not perfect i.e. there were dark negatives and white positives. The correlation was much stronger in the Traytown analysis with colour being strongly correlated with first principal component score.

3.2.2 Genotypic and size variation at Traytown

The three random samples taken from Traytown were dominated by individuals in the 0-5 and 6-10 mm size classes (Figure 10 and Table 13). Unfortunately, it was not possible to obtain consistent electrophoretic results with individuals of this size. In particular, the Est and Lap loci could not be resolved at all, and Pgm showed poor resolution. Only Ap gave clear resolution but unfortunately it is the least informative locus (see previous results). All calculations involving all four loci could therefore be based only on individuals > 10mm in length.

Alleles found for each locus in the Traytown random samples were identical to alleles described above for the survey samples (Tables 14 - 17), although the rare extremely fast and extremely slow alleles for Est were not present. The variation in sample size among loci was primarily due to problems obtaining clear resolution.

3.2.2.1 Variation among size classes

Allele frequency variation among size classes was significant for all loci except Ap (Table 18). Deviations from HWE, as measured by the D-value, were calculated for each locus in each size class. There were deviations at some loci in most size classes : 5 size classes showed significant heterozygote deficiencies at both the Pgm and Est loci, 2 showed significant heterozygote deficiencies at neither, and 2 showed significant heterozygote deficiencies at the Est locus only (Table 19). Only one size class showed significant heterozygote deficiency at the Lap locus and 5 out of the 9 showed significant heterozygote deficiencies at the Ap locus. There was, however, no consistent change in D-values with size (Figure 11). Est had D-values close to zero in two size classes, 26-30mm and 50+mm. These two classes were found to be predominantly trossulus (26-30mm) and edulis (50+mm) in the PCA described below. All other classes had large negative D-values and were found to contain a

mixture of genotypes. This pattern in Est was not reflected in any other loci.

3.2.2.2 Size-genotype association in random Traytown samples

The frequencies of the Pgm2, Est2 and Lap4 alleles increased with increased shell length and conversely those of the Est1, Pgm4 and Lap2 alleles decreased (Figure 12). There was a reversal of the Pgm and Est genotype correlation in smaller mussels (<10mm). However the sample sizes are small due primarily to difficulty in resolving Est in small mussels. It is therefore uncertain how significant this trend is.

The principal component analysis described earlier was repeated using the size class data and the data collected from the original three samples taken from Traytown (TT1, TT2, and TT3). 19% of the variation was explained on the first principal axis and 10% on the second axis. Pgm2 and Est2 had the largest postive loadings on the first axis. Pgm4 and Est1 and were again the alleles which had the largest negative loadings on the first axis (Table 20). The highest loading on the second axes was Ap5.

A histogram of the first principal component values of all individuals does not show the same bimodal distribution seen in the first PCA due to an intermediate peak (Figure

13). However the individual histograms for the three survey samples, TT1, TT2 and TT3 are very similar to those generated in the first PCA using the whole survey sample data set and lend support to the two PCAs being comparable (Figure 13). Although the second PCA indicated a large group of intermediate values the D-values still show a defeciency of heterozygotes in most size classes.

When the data were plotted by sample using individual first component values, it was seen that most samples contained individuals from throughout the range of first principal component values (Figure 14). Size class 6 (26-30 mm) and size class 11 (50+mm) appeared to be unimodal for peak 2 (M. trossulus) and peak 1 (M. edulis) respectively. This interpretation is difficult to support as the overall histogram does not clearly delineate two groups.

Table 1 : Range, mean and standard deviation of lengths of mussels sampled at each survey site.

Pop	N	Minimum	Maximum	Mean	Std Dev
BV0	86	26.00	89.00	64.62	13.63
BV1	20	32.00	57.00	41.85	7.49
BV2	17	16.00	29.00	19.24	2.99
BV3	20	30.00	59.00	46.90	6.75
CC1	20	25.00	47.00	30.70	5.29
CC2	20	26.00	53.00	40.35	7.17
LHE	30	21.00	49.00	34.73	7.76
THL	20	27.00	52.00	37.60	7.35
TT1	34	28.00	60.00	46.47	8.35
TT2	30	37.00	63.00	47.20	6.51
TT3	74	37.00	61.00	46.16	5.11
BB	21	28.00	57.00	39.81	7.28
EM	21	41.00	55.00	46.76	4.05
FH	20	31.00	55.00	42.50	6.39
IA	19	26.00	47.00	35.42	6.16
TW	20	21.00	35.00	28.00	3.51

Table 2 : Pgm allele frequencies for all survey sites. Alleles are arranged in order of increasing mobility.

		Allele					
Site	N	1	2	3	4	5	6
BV0	86	0.076	0.791	0.116	0.017	0.000	0.000
BV1	20	0.100	0.350	0.225	0.325	0.000	0.000
BV2	17	0.029	0.059	0.147	0.647	0.118	0.000
BV3	20	0.025	0.325	0.225	0.350	0.075	0.000
CC1	20	0.025	0.150	0.175	0.500	0.150	0.000
CC2	20	0.125	0.625	0.150	0.075	0.000	0.025
LHE	30	0.083	0.300	0.200	0.367	0.033	0.017
TH	20	0.050	0.550	0.200	0.175	0.025	0.000
TT1	34	0.059	0.588	0.206	0.132	0.015	0.000
TT2	30	0.033	0.317	0.283	0.300	0.050	0.017
TT3	74	0.020	0.257	0.216	0.453	0.041	0.014
BB	21	0.214	0.548	0.167	0.071	0.000	0.000
EM	21	0.071	0.738	0.190	0.000	0.000	0.000
FH	20	0.125	0.800	0.075	0.000	0.000	0.000
IA	19	0.079	0.526	0.079	0.289	0.026	0.000
TW	19	0.105	0.395	0.132	0.368	0.000	0.000

Table 3 : Est allele frequencies for all survey sites. Alleles are arranged in order of increasing mobility.

		Allele				
Site	N	1	2	3	4	8*
BV0	85	0.059	0.912	0.018	0.006	0.006
BV1	18	0.472	0.528	0.000	0.000	0.000
BV2	16	0.844	0.094	0.031	0.000	0.031
BV3	20	0.625	0.375	0.000	0.000	0.000
CC1	20	0.900	0.100	0.000	0.000	0.000
CC2	20	0.225	0.750	0.000	0.000	0.025
LHE	30	0.500	0.500	0.000	0.000	0.000
THL	20	0.350	0.650	0.000	0.000	0.000
TT1	34	0.297	0.688	0.016	0.000	0.000
TT2	30	0.633	0.350	0.000	0.000	0.017
TT3	74	0.669	0.310	0.007	0.000	0.014
BB	21	0.143	0.810	0.048	0.000	0.000
EM	21	0.071	0.929	0.000	0.000	0.000
FH	20	0.075	0.925	0.000	0.000	0.000
IA	19	0.444	0.556	0.000	0.000	0.000
TW	19	0.500	0.500	0.000	0.000	0.000

*8 is slower than 1

Table 4 : Ap allele frequencies for all survey sites. Alleles are arranged in order of increasing mobility.

		Allele					
Site	N	1	2	3	4	5	6
BV0	26	0.019	0.269	0.019	0.269	0.385	0.038
BV1	20	0.000	0.100	0.050	0.450	0.375	0.025
BV2	17	0.000	0.088	0.059	0.471	0.353	0.029
BV3	17	0.000	0.118	0.059	0.441	0.382	0.000
CC1	20	0.000	0.075	0.025	0.600	0.275	0.025
CC2	19	0.000	0.211	0.053	0.263	0.474	0.000
LHE	17	0.029	0.176	0.059	0.353	0.324	0.059
TH	20	0.000	0.075	0.075	0.500	0.325	0.025
TT1	19	0.000	0.237	0.053	0.316	0.395	0.000
TT2	20	0.000	0.275	0.075	0.300	0.300	0.050
TT3	21	0.000	0.238	0.048	0.381	0.310	0.024
BB	19	0.000	0.105	0.079	0.211	0.553	0.053
LM	21	0.000	0.167	0.024	0.381	0.405	0.024
FH	20	0.000	0.250	0.000	0.400	0.350	0.000
IA	19	0.000	0.079	0.000	0.368	0.553	0.000
TW	20	0.000	0.150	0.175	0.275	0.400	0.000

Table 5 : Lap allele frequencies for all survey sites. Alleles are arranged in order of increasing mobility.

		Allele				
Site	N	1	2	3	4	5
BV0	26	0.000	0.038	0.346	0.596	0.019
BV1	19	0.026	0.105	0.342	0.526	0.000
BV2	17	0.059	0.324	0.471	0.147	0.000
BV3	17	0.000	0.088	0.265	0.647	0.000
CC1	19	0.026	0.342	0.526	0.105	0.000
CC2	19	0.026	0.105	0.368	0.474	0.026
LHE	15	0.067	0.300	0.400	0.233	0.000
TH	20	0.025	0.175	0.350	0.400	0.050
TT1	19	0.026	0.211	0.316	0.421	0.026
TT2	20	0.025	0.275	0.475	0.225	0.000
TT3	21	0.119	0.429	0.357	0.071	0.024
BB	19	0.000	0.105	0.421	0.447	0.026
EM	21	0.071	0.167	0.429	0.286	0.048
FH	20	0.050	0.075	0.200	0.675	0.000
IA	19	0.053	0.184	0.421	0.342	0.000
TW	20	0.000	0.275	0.425	0.275	0.025

Table 6 : D values - Deficiency (-) or Excess (+) of heterozygotes; per locus, per sample. Probability values are Chi-Square tests of significant deviations from Hardy-Weinberg expectations. Numbers in brackets indicate samples which are reduced from total number analysed for that site.

Sample	N	Pgm	Est	Ap	Lap
BV0	86	-.219***	-.078	-.037(26)	-.351(26)
BV1	20	-.178	-.675(18)***	-.167	-.396(19)
BV2	17	-.157*	-.345(16)	0.067	0.056
BV3	20	-.386	-.896***	0.067(17)	-.320(17)
BCC	20	-.276	-.458*	-.126***	-.051(19)
LCC	20	-.137**	-.495	0.011(19)	-.428(19)
LHE	30	-.279	-.390*	-.064(17)	0.119(15)
THL	20	-.218	-.571**	-.152	-.002
TT1	34	-.215	-.720***	-.103(19)	-.470(19)
TT2	30	-.322	-.725***	-.007(20)	-.322(20)
TT3	74	-.408***	-.632***	-.136(21)	-.166(21)*
BB	21	-.327**	-.134	-.024(19)	-.077
EM	21	0.123	0.051	0.123	-.070
FH	20	-.424***	0.054	-.032	-.116
IA	19	-.345*	-.344(18)	-.073	-.387***
TW	20	-.322(19)*	-.283(19)	-.109	-.343

* p<.05 ** p<.01 *** p<.005

Table 7 : Heterogeneity among samples. Chi-square values calculated using a Monte Carlo program (REAP, McElroy 1991). Tests are based on 1000 random iterations. See Appendices 1A-C for contingency tables.

Group	N[1]	Pgm	Est	Lap	Ap
All samples	16	345.73***	308.94***	152.83***	90.32 N.S.
Traytown	3	34.69***	29.23***	20.26*	3.49 N.S.
Bellevue	4	142.53***	127.81***	35.40***	13.96 N.S.

[1] number of samples in group tested.
* p<0.05 ** p<.01 *** p<.005 N.S. not significant

Table 8 : F_{IS}, F_{IT}, F_{ST} for each allele at each locus for pooled data from all survey sites, mean F_{IS}, F_{IT}, F_{ST} for each locus and grand means over all loci (generated by BioSys). Only individuals for which all four loci were scored are included. N=300.

Locus/Allele	F_{IS}	F_{IT}	F_{ST}	
Pgm1	0.090	0.121	0.034	
Pgm2	0.347	0.484	0.210	
Pgm3	0.064	0.097	0.035	
Pgm4	0.389	0.508	0.195	
Pgm5	0.083	0.143	0.065	
Pgm6	-0.032	-0.004	0.028	
mean Pgm	0.244	0.351	0.142	(0.118)[1]
Est8	-0.028	-0.007	0.021	
Est1	0.492	0.638	0.289	
Est2	0.472	0.624	0.289	
Est3	-0.039	-0.008	0.030	
mean Est	0.461	0.612	0.280	(0.289)[1]
Ap1	-0.031	-0.003	0.027	
Ap2	-0.101	-0.058	0.039	
Ap3	-0.032	0.005	0.035	
Ap4	0.112	0.148	0.041	
Ap5	0.032	0.058	0.027	
Ap6	-0.038	-0.022	0.016	
mean Ap	0.026	0.059	0.034	(0.035)[1]
Lap1	0.028	0.058	0.030	
Lap2	0.073	0.141	0.074	
Lap3	0.169	0.192	0.028	
Lap4	0.332	0.429	0.145	
Lap5	-0.037	-0.014	0.022	
mean Lap	0.186	0.250	0.079	(0.082)[1]
grand means	0.199	0.298	0.124	

[1] mean F_{ST} value calculated using alleles with $p>0.05$

Table 9 : Associations between loci. Chi-square values calculated using a Monte Carlo program (REAP, McElroy 1991). Tests are based on 1000 random iterations. See Appendices 2A-F for contingency tables.

Loci	Chi-Square Value	P	
Pgm-Est	384.57	0.00 ± 0.00	***
Pgm-Lap	269.12	0.03 ± 0.02	*
Pgm-Ap	190.88	0.43 ± 0.05	N.S.
Est-Lap	127.10	0.10 ± 0.03	N.S.
Est-Ap	105.96	0.21 ± 0.04	N.S.
Lap-Ap	189.35	0.31 ± 0.05	N.S.

Table 10 : Spearman Rank Correlation Coefficients used in the Principal Components Analysis. Boldface indicates r values significantly different from 0. (Table tested at $p<0.05$ using a sequential Bonferroni test (Rice 1988)).

	Pgm 1	2	3	4	5	6	Est 0	1	2	3
Pgm										
1	1.00									
2	-0.14	1.00								
3	-0.12	**-0.47**	1.00							
4	**-0.28**	**-0.63**	-0.09	1.00						
5	-0.07	-0.19	-0.15	-0.11	1.00					
6	-0.03	-0.10	0.03	0.02	-0.01	1.00				
Est										
0	0.08	0.00	0.03	-0.05	-0.01	-0.01	1.00			
1	**-0.28**	**-0.65**	**0.22**	**0.69**	**0.17**	0.10	0.03	1.00		
2	**0.29**	**0.65**	**-0.22**	**-0.69**	**-0.17**	-0.10	-0.10	**-0.98**	1.00	
3	-0.05	0.07	0.04	-0.08	-0.02	-0.01	-0.01	-0.11	-0.03	1.00
Lap										
1	-0.04	-0.18	0.14	0.13	0.04	-0.02	-0.02	0.18	-0.18	-0.03
2	-0.15	**-0.42**	0.05	**0.46**	0.17	-0.06	-0.06	**0.41**	**-0.40**	-0.08
3	-0.03	-0.04	0.05	0.07	-0.09	0.08	-0.02	0.16	-0.15	0.04
4	0.16	**0.40**	-0.16	**-0.42**	-0.08	-0.02	0.08	**-0.48**	**0.46**	0.04
5	0.09	0.00	0.01	-0.04	-0.03	-0.01	-0.01	-0.12	0.13	-0.02
Ap										
1	-0.03	0.10	-0.05	-0.05	-0.01	-0.01	-0.01	-0.08	0.08	-0.01
2	0.07	0.06	-0.10	-0.07	0.11	0.03	-0.05	-0.16	0.16	-0.02
3	-0.02	**-0.25**	0.18	0.12	0.14	-0.03	0.10	0.17	-0.17	-0.04
4	-0.07	-0.12	-0.02	0.20	-0.02	0.08	0.04	0.19	0.18	0.01
5	0.07	0.18	0.00	**-0.21**	-0.09	-0.09	-0.03	-0.16	0.15	0.04
6	-0.04	-0.02	0.01	0.00	-0.04	-0.02	-0.02	0.00	0.00	-0.02

Table 10 (cont.) : Spearman Rank Correlation Coefficients used in the Principal Components Analysis. Boldface indicates r values significantly different from 0 (Table tested at $p<0.05$ using a sequential Bonferroni test (Rice 1988)).

	Lap 1	2	3	4	5	Ap 1	2	3	4	5	6
Lap											
1	1.00										
2	0.01	1.00									
3	-0.11	**-0.22**	1.00								
4	-0.21	**-0.46**	**-0.64**	1.00							
5	-0.05	-0.12	-0.06	-0.03	1.00						
Ap											
1	-0.02	-0.06	0.12	-0.08	-0.01	1.00					
2	-0.06	-0.01	-0.13	0.16	-0.02	0.03	1.00				
3	-0.01	0.15	0.06	-0.19	0.06	-0.03	-0.03	1.00			
4	0.03	0.08	0.13	-0.19	0.06	-0.09	**-0.33**	**-0.25**	1.00		
5	0.02	-0.16	-0.05	0.18	-0.10	-0.03	-0.31	**-0.16**	-0.63	1.00	
6	-0.06	0.06	-0.01	0.01	-0.04	-0.02	-0.11	-0.02	-0.10	-0.09	1.00

Table 11 : Component loadings on the first two axes of the principal components analysis on the survey samples. n=300, 21 variables.

	PRIN1	PRIN2
Pgm1	-0.28862	0.09197
Pgm2	-0.78193	-0.26540
Pgm3	0.30304	0.31020
Pgm4	0.77934	-0.06314
Pgm5	0.18992	0.35963
Pgm6	0.10106	-0.17484
Est0	0.00649	0.04018
Est1	0.90243	0.07380
Est2	-0.86959	-0.07332
Est3	-0.07558	-0.08056
Lap1	0.23291	0.15248
Lap2	0.57911	0.23705
Lap3	0.23196	-0.53734
Lap4	-0.68698	0.26992
Lap5	-0.08814	-0.11932
Ap1	-0.07745	-0.15784
Ap2	-0.18015	0.22497
Ap3	0.27065	0.35569
Ap4	0.30752	-0.72096
Ap5	-0.31439	0.42859
Ap6	0.01574	0.02522

Table 12 : Association of genotype with inner shell color. Chi-square values calculated using a Monte Carlo program (REAP, McElroy 1991). Tests are based on 1000 random iterations. (See Appendices 3A-D for contingency tables.)

Locus	X^2 Value	N	P
Pgm	374.44	708	0.001 ± 0.001 ***
Est	292.69	627	0.010 ± 0.003 **
Ap	52.29	546	0.231 ± 0.013 N.S.
Lap	178.45	536	0.000 ± 0.000 ***

Table 13 : Number per size class in each of three replicate 1/16m2 samples taken at TT3.

Size class	Sample 1	Sample 2	Sample 3	Total
0-5 mm	73	378	88	539
6-10 mm	214	408	149	771
11-15 mm	55	38	28	121
16-20 mm	10	12	26	48
21-25 mm	6	24	40	70
26-30 mm	22	32	56	110
31-35 mm	18	22	23	63
36-40 mm	31	24	13	68
41-45 mm	21	38	6	65
46-50 mm	25	31	18	74
51+ mm	15	12	3	30
Total	490	1019	450	1959

Table 14 : Pgm allele frequencies calculated by size class for combined random Traytown samples. Alleles are arranged in order of increasing mobility.

		Allele					
Size Class	N	1	2	3	4	5	6
1-5 mm	7	0.000	0.000	0.286	0.571	0.143	0.000
6-10	38	0.013	0.039	0.224	0.618	0.066	0.039
11-15	7	0.000	0.071	0.357	0.500	0.071	0.000
16-20	19	0.105	0.237	0.368	0.263	0.026	0.000
21-25	22	0.023	0.159	0.182	0.568	0.068	0.000
26-30	41	0.049	0.159	0.220	0.537	0.037	0.000
31-35	27	0.037	0.167	0.278	0.463	0.056	0.000
36-40	22	0.045	0.250	0.295	0.318	0.068	0.023
41-45	19	0.026	0.553	0.053	0.289	0.053	0.026
46-50	24	0.042	0.375	0.208	0.313	0.063	0.000
50 +	21	0.095	0.762	0.095	0.048	0.000	0.000

Table 15 : Est allele frequencies calculated by size class for combined random Traytown samples. Alleles are arranged in order of increasing mobility.

		Allele		
Size class (mm)	N	1	2	3
1-5	1	1.000	0.000	0.000
6-10	1	1.000	0.000	0.000
11-15	7	0.143	0.714	0.143
16-20	19	0.632	0.368	0.000
21-25	19	0.868	0.132	0.000
26-30	20	0.850	0.150	0.000
31-35	22	0.795	0.205	0.000
36-40	21	0.524	0.429	0.048
41-45	17	0.500	0.441	0.059
46-50	24	0.500	0.500	0.000
50 +	17	0.353	0.647	0.000

Table 16 : Ap allele frequencies calculated by size class for combined random Traytown samples. Alleles are arranged in order of increasing mobility.

		Allele					
Size Class	N	1	2	3	4	5	6
1-5 mm	3	0.000	0.000	0.000	0.500	0.500	0.000
6-10	32	0.000	0.219	0.078	0.375	0.297	0.031
11-15	7	0.000	0.286	0.071	0.429	0.214	0.000
16-20	17	0.000	0.235	0.088	0.471	0.176	0.029
21-25	22	0.000	0.182	0.045	0.386	0.364	0.023
26-30	41	0.000	0.232	0.037	0.390	0.329	0.012
31-35	27	0.000	0.111	0.074	0.537	0.278	0.000
36-40	22	0.000	0.159	0.023	0.477	0.318	0.023
41-45	19	0.000	0.132	0.053	0.447	0.316	0.053
46-50	23	0.000	0.152	0.065	0.435	0.326	0.022
50 +	21	0.000	0.190	0.000	0.167	0.571	0.071

Table 17 : Lap allele frequencies calculated by size classes for combined random Traytown samples. Alleles are arranged in order of increasing mobility.

		Allele				
Size Class	N	1	2	3	4	5
1-5mm	1	0.000	1.000	0.000	0.000	0.000
6-10	5	0.200	0.500	0.300	0.000	0.000
11-15	7	0.214	0.071	0.643	0.071	0.000
16-20	14	0.107	0.286	0.393	0.214	0.000
21-25	21	0.048	0.286	0.500	0.167	0.000
26-30	40	0.125	0.400	0.387	0.075	0.013
31-35	26	0.288	0.308	0.288	0.115	0.000
36-40	22	0.068	0.273	0.500	0.159	0.000
41-45	19	0.132	0.263	0.395	0.211	0.000
46-50	24	0.063	0.292	0.333	0.292	0.021
50 +	19	0.000	0.211	0.368	0.421	0.000

Table 18 : Heterogeneity in allele frequency among size classes. Chi-square values calculated using a Monte Carlo program (REAP, McElroy 1991). Tests are based on 100 random iterations. See Appendices 4A-D for contingency tables.

Locus	X^2 Value	P-Value
Pgm	157.00	0.000 ± 0.000 ***
Est	88.21	0.0070 ± 0.0026 **
Lap	70.47	0.0120 ± 0.0034 *
Ap	40.14	0.4210 ± 0.0156 N.S.

Table 19 : D values, (deficiency (-) or excess (+) of heterozygotes), per size class for combined random Traytown samples. Probability values are Chi-Square tests of significant deviations from HWE.

Size Class	N	Pgm	Est	Ap	Lap
11-15 mm	7	-.133	-1.00***	-.418	-.250
16-20 mm	14	-.229*	-.856***	-.417***	-.123
21-25 mm	19	-.152***	-.776***	-.161*	0.252*
26-30 mm	20	-.119	-.235	-.154	0.177
31-35 mm	22	-.524***	-.863***	-.437*	0.063
36-40 mm	21	-.065*	-.828***	0.016	-.365***
41-45 mm	18	-.375	-.897***	-.156***	-.159
46-50 mm	24	-.431***	-.671***	-.028	-.288[2]
50 + mm	16	0.079[3]	-.155	-.445***	-.235

* p<0.05 ** p<0.01 *** p<0.005
[1] n=17 [2] n=23 [3] n=15

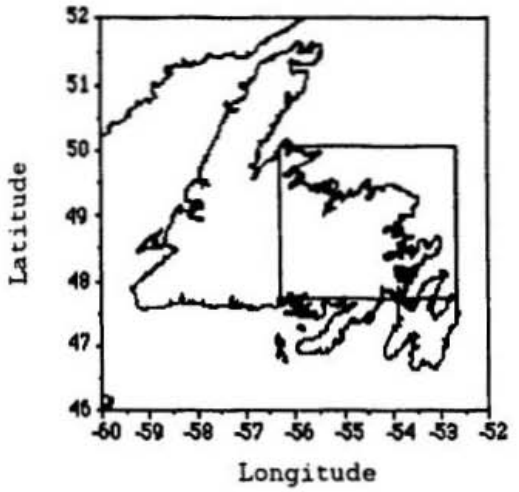
Table 20 : Component loadings of the first two axes of the principal component analysis on the Traytown samples. n=161, 19 variables.

	PRIN1	PRIN2
Pgm1	-0.23153	0.02208
Pgm2	-0.77369	0.06740
Pgm3	0.29618	0.51931
Pgm4	0.62413	-0.41999
Pgm5	0.10749	-0.00475
Pgm6	-0.03880	-0.17388
Est1	0.86259	0.01785
Est2	-0.84536	-0.09256
Est3	-0.10485	0.31663
Lap1	0.45159	0.07540
Lap2	0.43033	-0.32225
Lap3	-0.06268	-0.03090
Lap4	-0.68679	0.29500
Lap5	-0.08407	0.12113
Ap2	-0.13497	-0.09498
Ap3	0.25315	-0.12142
Ap4	0.32188	0.75601
Ap5	-0.26672	-0.68512
Ap6	-0.06583	-0.09412

Figure 1: Location of samples. Abbreviations used for each site through out text as follows:

Reference Number	Sample Site	Abbreviation
1	Fortune Harbour	FH
2	Embree	EM
3	Indian Arm	IA
4	Birchy Bay	BB
5	Twillingate	TW
6	Traytown*	
7	Little Hearts' Ease	LHE
8	Big Chance Cove	CC1
	Little Chance Cove	CC2
9	Bellevue*	
10	Thornlea	TH

* see Figure 2 for sample abbreviations


Latitude
52
51
50
49
48
47
46
-60 -59 -58 -57 -56 -55 -54 -53 -52
Longitude

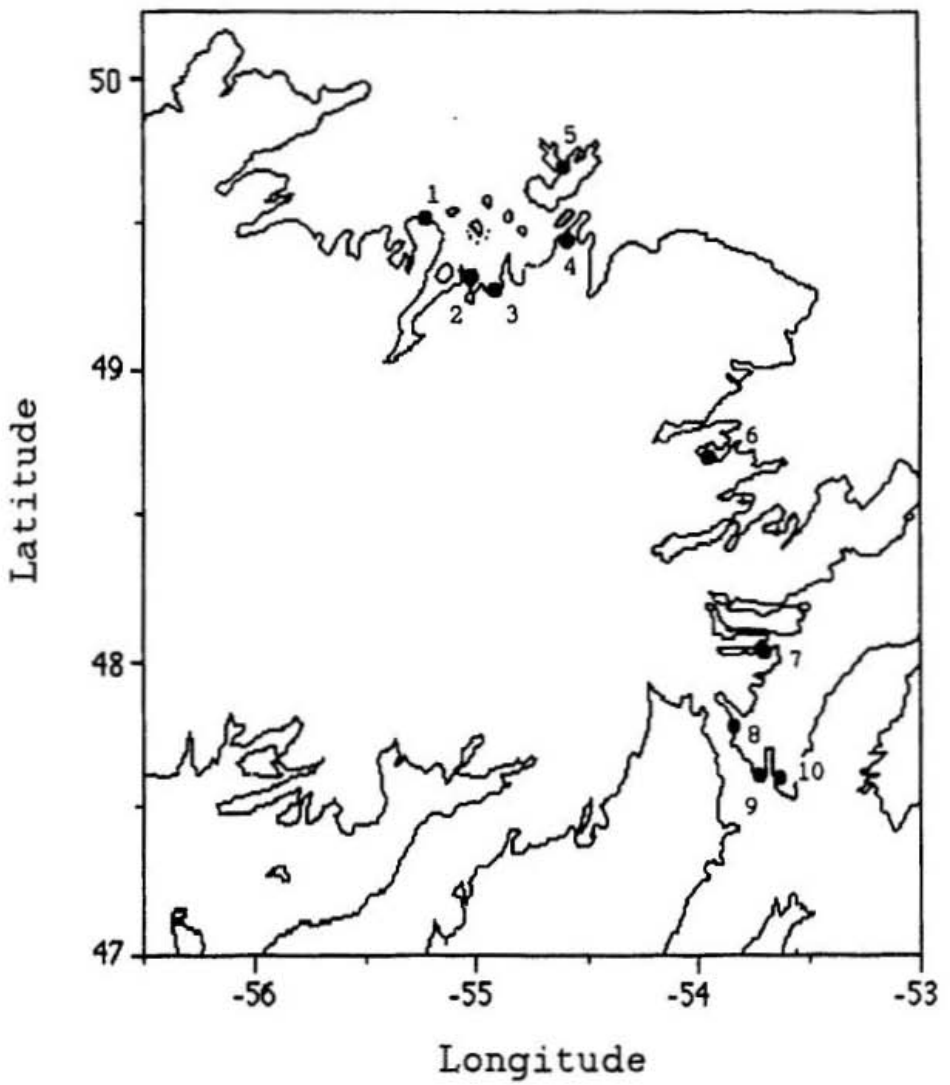

Latitude
50
49
48
47
-56
-55
-54
-53
Longitude
1
2
3
4
5
6
7
8
9
10

Figure 2A : Location of samples at Bellevue site. Abbreviations used for each sample throughout text as follows:

Abbreviation	Sample Site
BV0	Bellvue subtidal
BV1	Bellevue 1
BV2	Bellevue 2
BV3	Bellevue 3

Figure 2B : Location of samples at Traytown site. Locations of the random samples taken from the bed from which the original TT3 sample was taken are also shown. Dashed lines indicate roads. Abbreviations used for each sample throughout text as follows:

Abbreviation	Sample Site
TT1	Traytown 1
TT2	Traytown 2
TT3	Traytown 3
1	1st random sample (TT3-1)
2	2nd random sample (TT3-2)
3	3rd random sample (TT3-3)

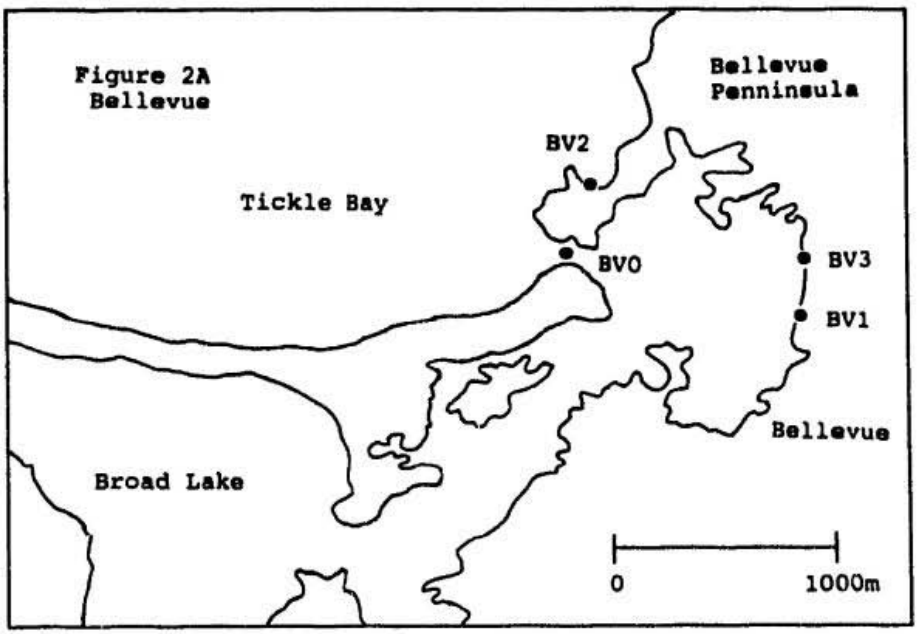


Figure 2A
Bellevue
Bellevue
Penninsula
BV2
Tickle Bay
BV0
BV3
BV1
Bellevue
Broad Lake
0
1000m

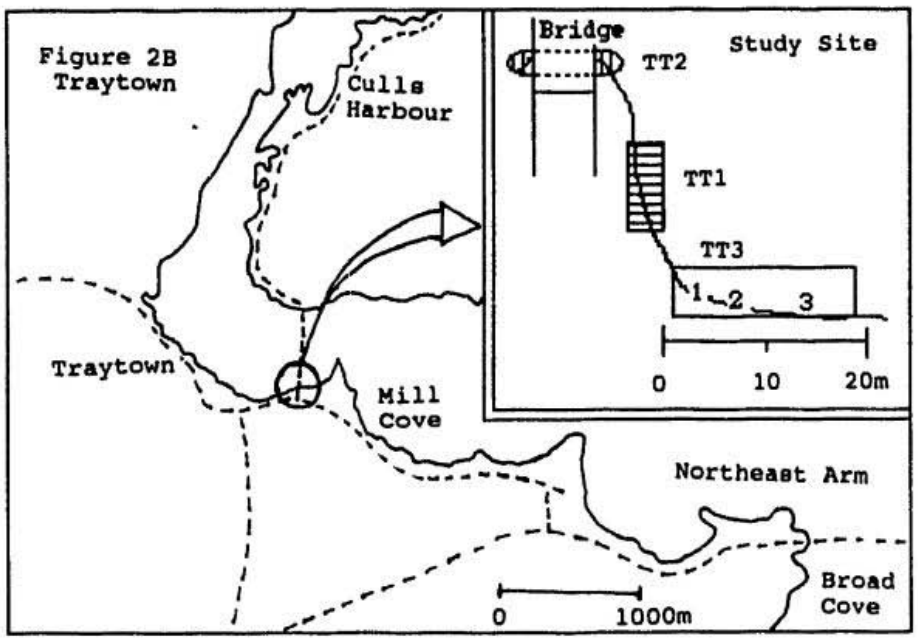


Figure 2B
Traytown
Culls
Harbour
Bridge
Study Site
TT2
TT1
TT3
1
2
3
0
10
20m
Traytown
Mill
Cove
Northeast Arm
Broad
Cove
0
1000m

Figure 3 : Multidimensional scaling diagram summarising genetic differentiation among sites with Minimum Spanning Tree superimposed.

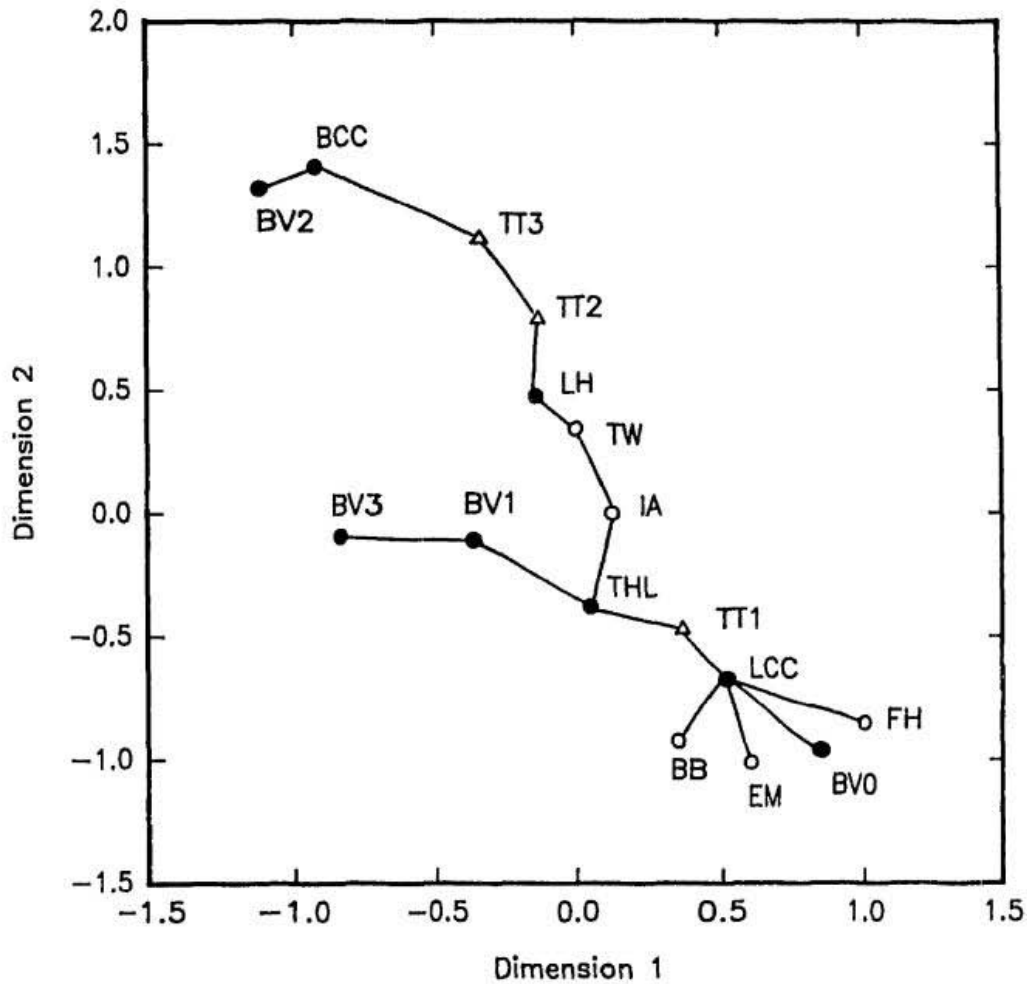

2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Dimension 2
Dimension 1
BCC
BV2
TT3
TT2
LH
TW
IA
BV3
BV1
THL
TT1
LCC
FH
BB
EM
BV0

Figure 4 : Multidimensional scaling diagrams showing genetic differentiation among sites for each locus with minimum spanning trees superimposed.

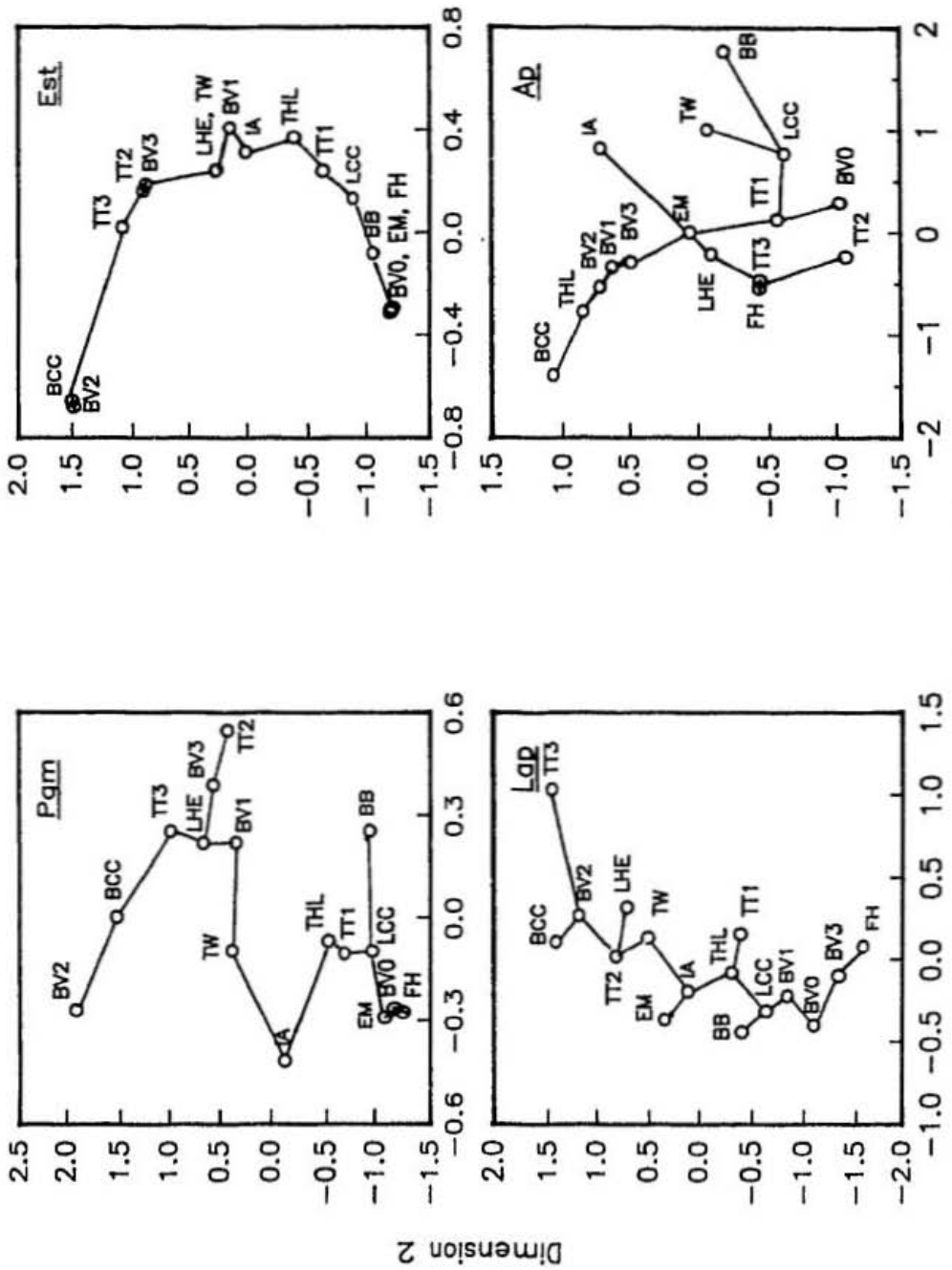

Est
Ad
Pgm
Lgp
BCC
BV2
TT3
TT2
BV3
LHE, TW
BV1
IA
THL
TT1
LCC
BB
BV0, EM, FH
TW
EM
FH
LHE
Dimension 1
Dimension 2

Figure 5 : Covariation of alleles in survey samples.
● = first pattern
o = second pattern

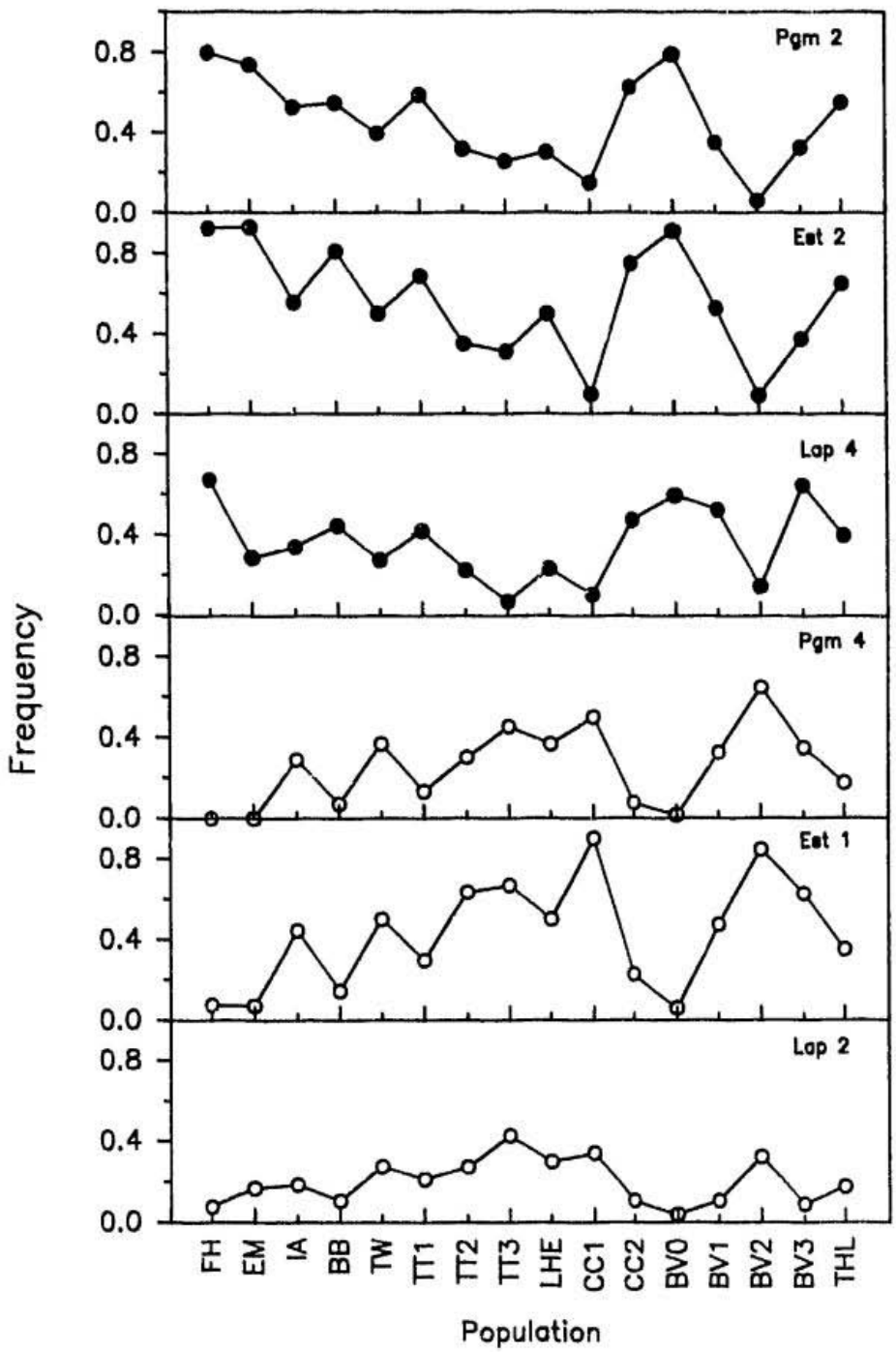

Pgm 2
Est 2
Lap 4
Pgm 4
Est 1
Lap 2
0.0
0.4
0.8
Frequency
FH
EM
IA
BB
TW
TT1
TT2
TT3
LHE
CC1
CC2
BV0
BV1
BV2
BV3
THL
Population

Figure 6 : Scatter plot of first two principal components for all survey sample individuals scored for all four loci.

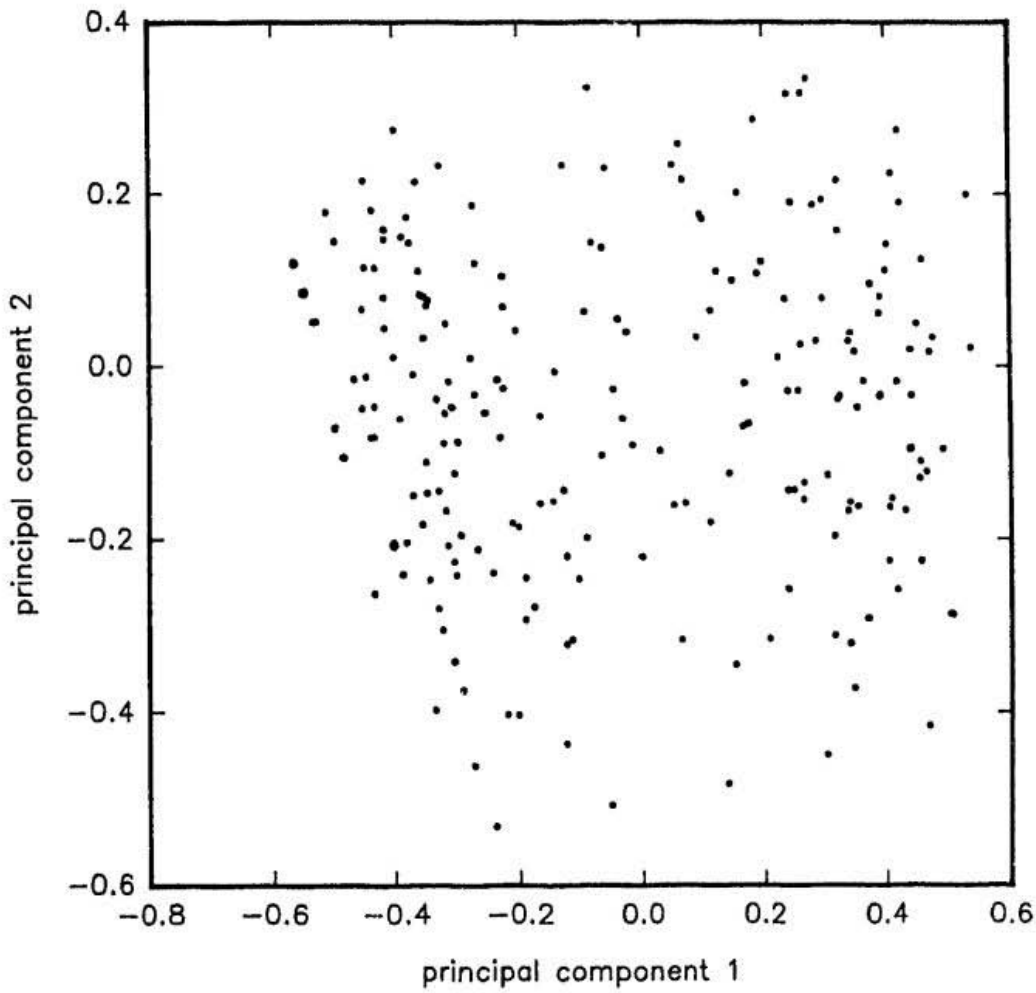

0.4
0.2
0.0
−0.2
−0.4
−0.6
principal component 2
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
principal component 1

Figure 7 : Histogram of first principal component scores for all survey individuals with all four loci scored.

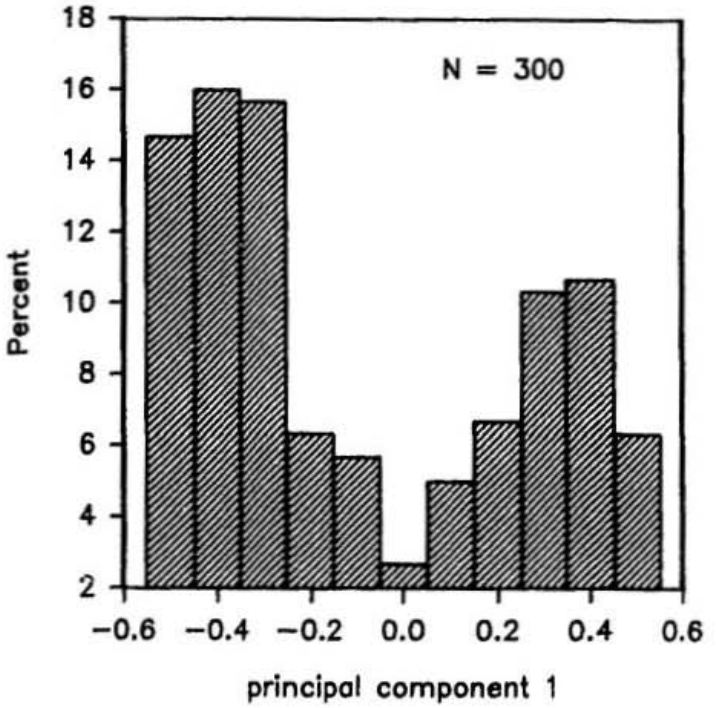

N = 300
Percent
18
16
14
12
10
8
6
4
2
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
principal component 1

Figure 8 : Histograms of first principal component scores broken down by sample locality overlayed on map of sample sites.

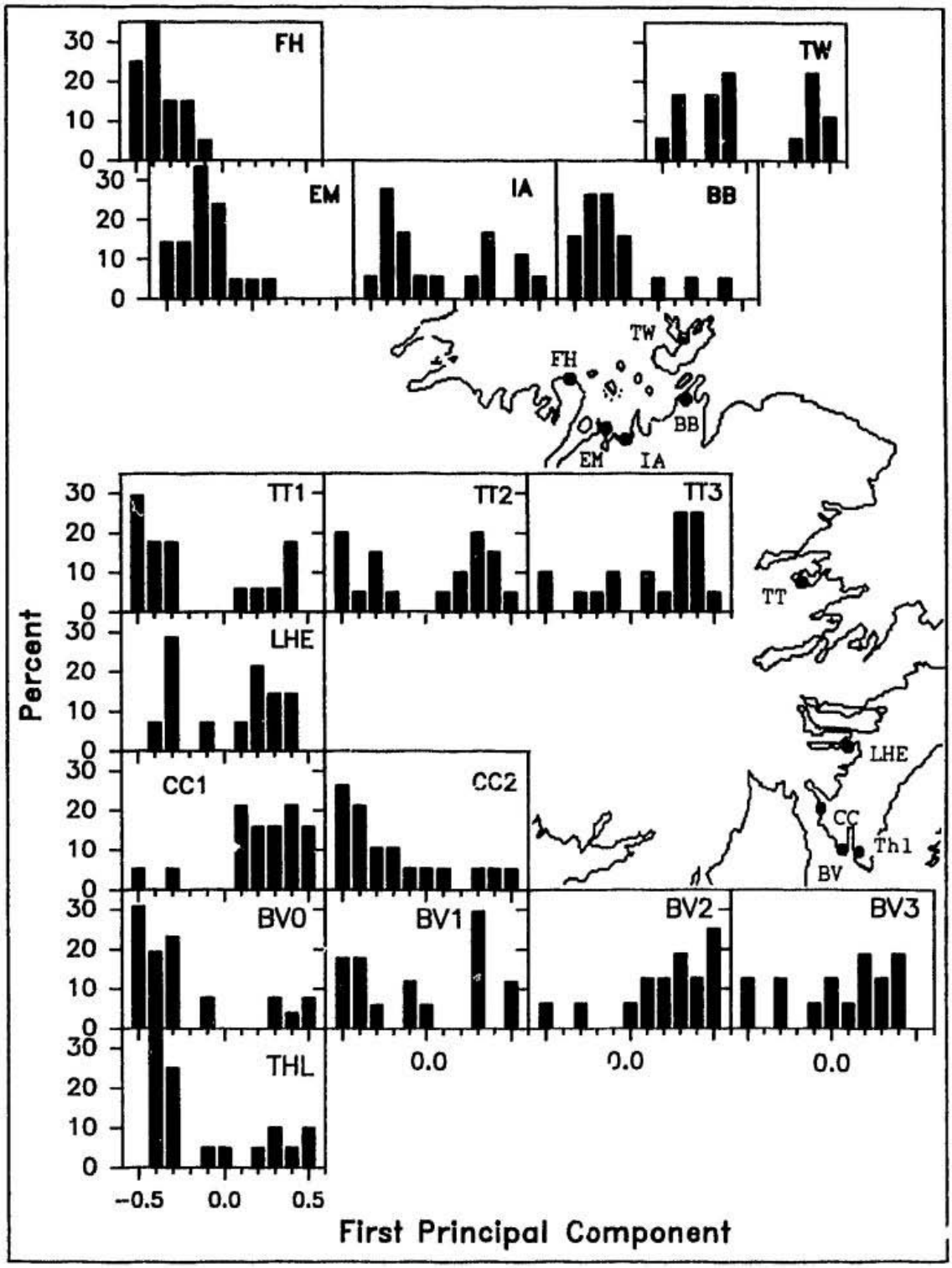

FH
TW
EM
IA
BB
TT1
TT2
TT3
LHE
CC1
CC2
BV0
BV1
BV2
BV3
THL
30
20
10
0
-0.5
0.0
0.5
Percent
First Principal Component
TT
LHE
CC
Th1
BV

Figure 9 : Association of colour with first principal component score; Survey and Traytown analyses.

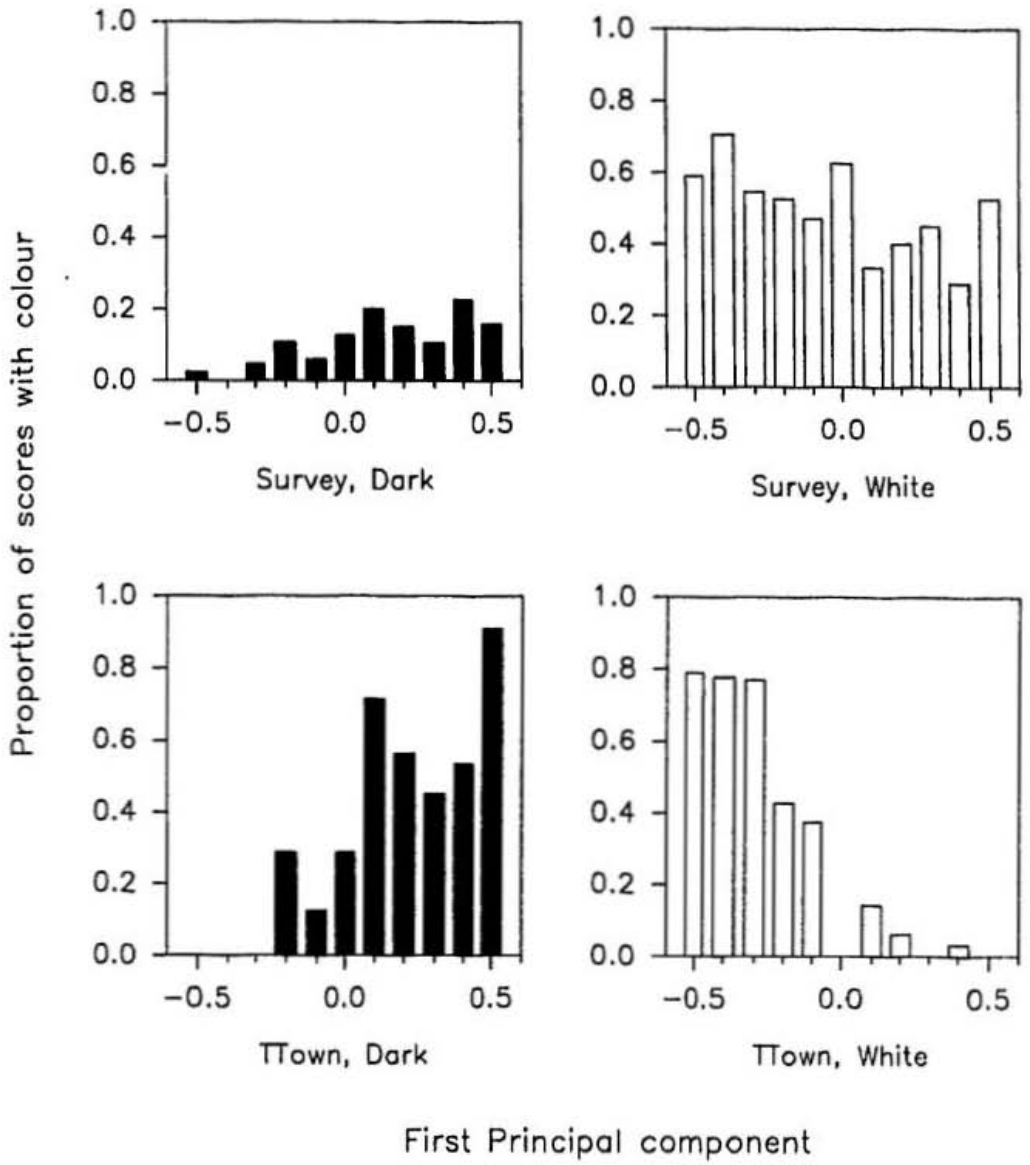

Proportion of scores with colour
1.0
0.8
0.6
0.4
0.2
0.0
−0.5
0.0
0.5
Survey, Dark
Survey, White
TTown, Dark
TTown, White
First Principal component

Figure 10 : Number of individuals used in Traytown size class analysis.

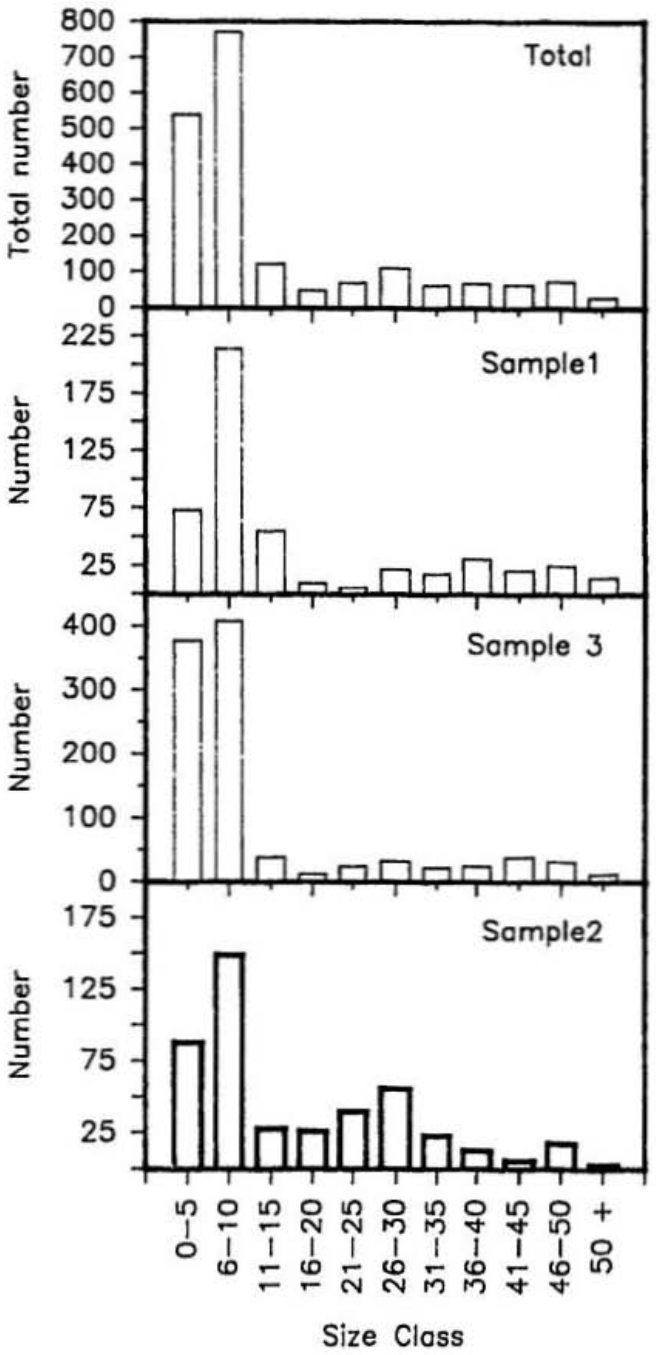

Total number
800
700
600
500
400
300
200
100
0
Total
Number
225
175
125
75
25
Sample1
Number
400
300
200
100
0
Sample 3
Number
175
125
75
25
Sample2
0–5
6–10
11–15
16–20
21–25
26–30
31–35
36–40
41–45
46–50
50 +
Size Class

Figure 11 : D-values plotted versus size class for all four loci in combined random Traytown samples.

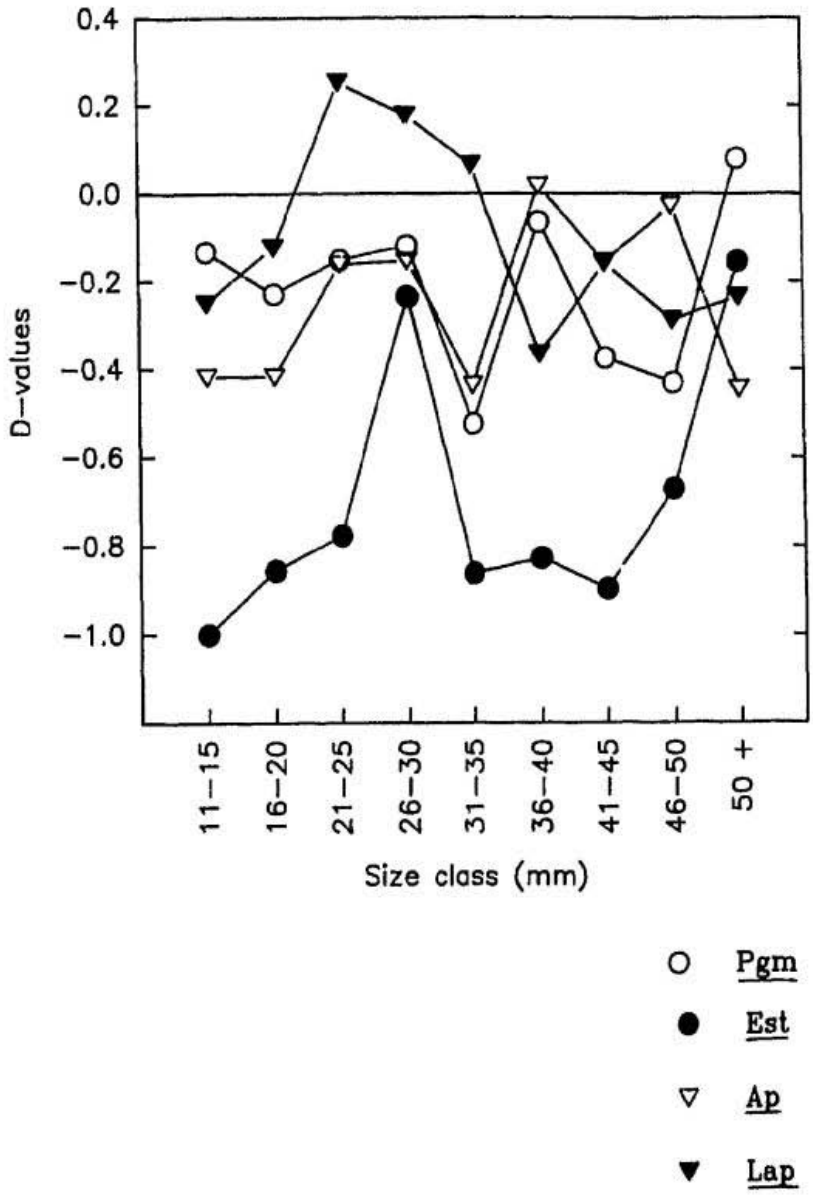

0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
D−values
11−15
16−20
21−25
26−30
31−35
36−40
41−45
46−50
50 +
Size class (mm)
Pgm
Est
Ap
Lap

Figure 12 : Covariation of loci in random Traytown samples.
● = first pattern
o = second pattern

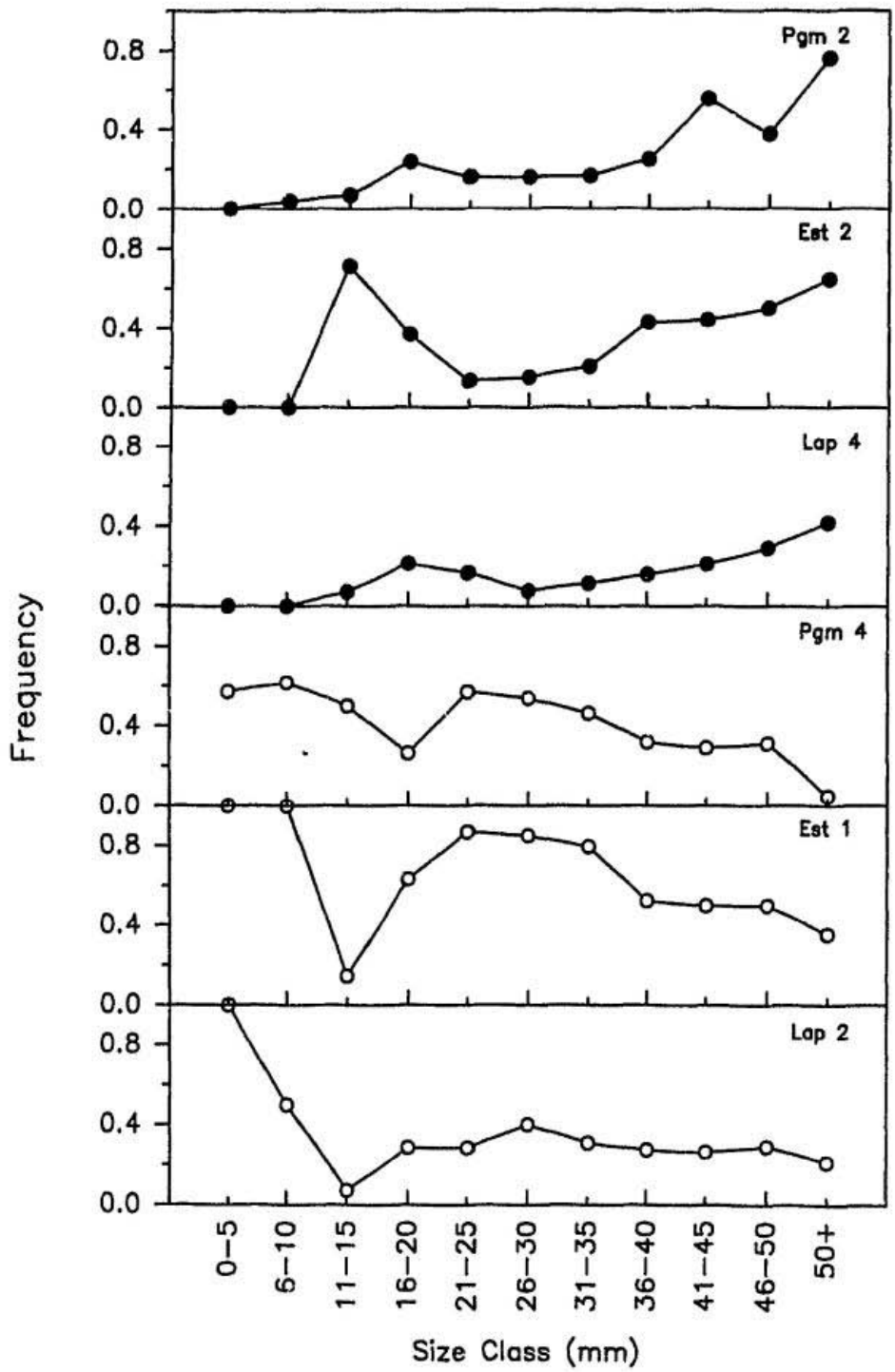

Pgm 2
Est 2
Lap 4
Pgm 4
Est 1
Lap 2
0.0
0.4
0.8
Frequency
0–5
6–10
11–15
16–20
21–25
26–30
31–35
36–40
41–45
46–50
50+
Size Class (mm)

Figure 13 : First principal component scores for all individuals sampled from Traytown and comparison of Traytown 1, 2, 3 results from each pca analysis.

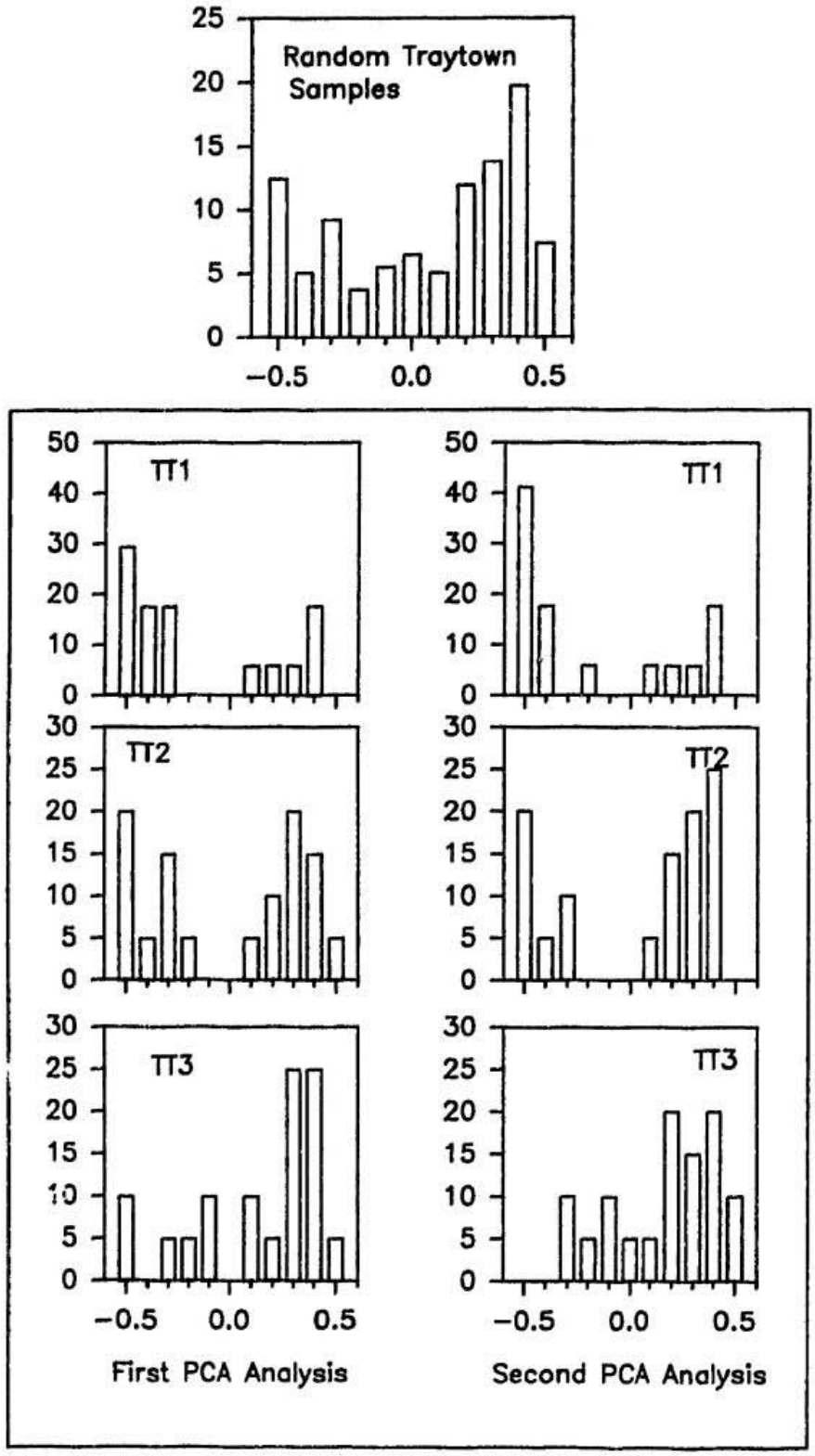
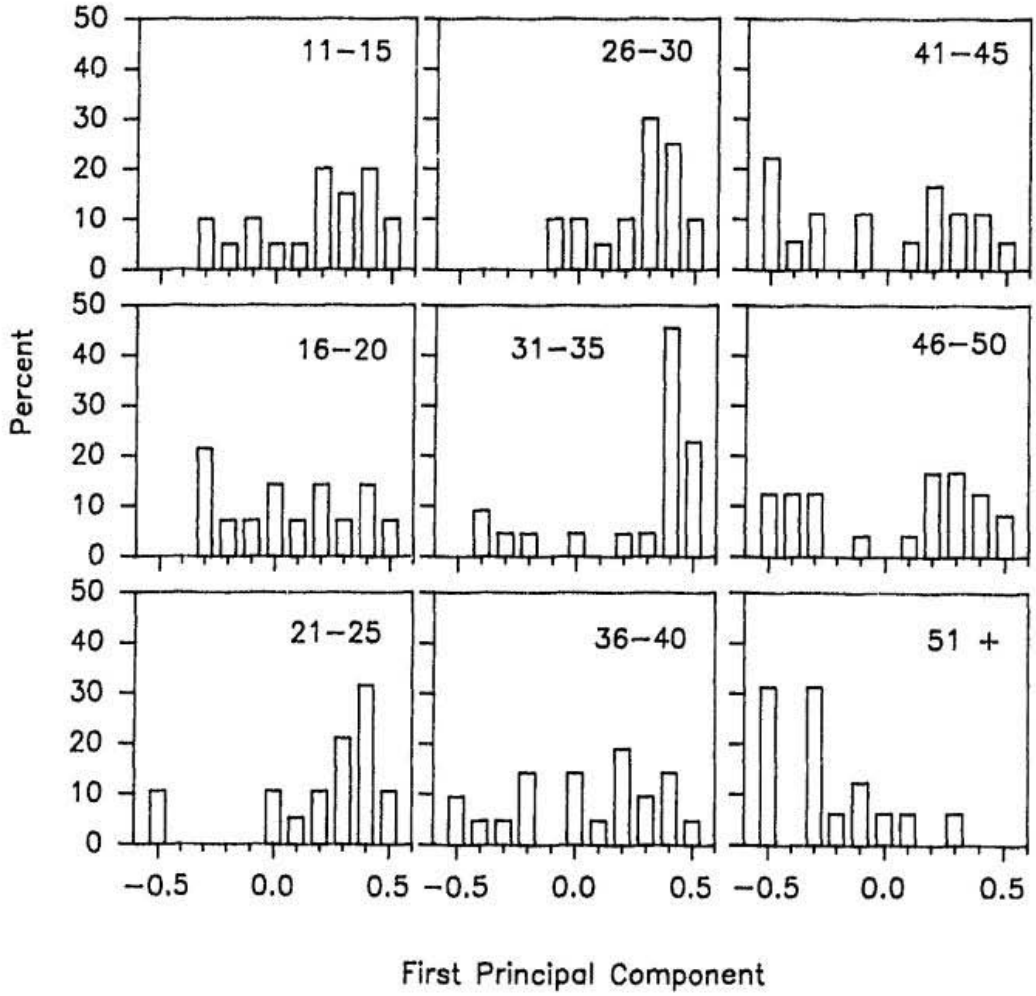

Random Traytown Samples
TT1
TT2
TT3
TT1
TT2
TT3
First PCA Analysis
Second PCA Analysis

Figure 14 : Values of first principal component for combined random Traytown samples presented by size class.


11–15
26–30
41–45
16–20
31–35
46–50
21–25
36–40
51 +
Percent
0
10
20
30
40
50
−0.5
0.0
0.5
First Principal Component

Discussion

Allozyme variation can be used as taxonomic tool to discriminate between morphologically cryptic species or to determine the population structure of a widespread species. Overall patterns of allozyme variation in Mytilus were compared to predictions under the hypothesis of random mating in order to determine if the pattern of genetic variation is best described by the presence of partially reproductively isolated groups. In this study allozyme variation has confirmed the existence two genetically distinct groups of Mytilus which show a complex population structure in Newfoundland. One of the complications in the analysis is the lack of an alternate independent character which can discriminate between the two groups defined by the allozymes. This combined with varying patterns of differentiation among the loci make it impossible to define any one character as diagnostic (as defined by Avise 1974).

Macrogeographic variation

Significant differentiation among populations at three of four loci was found in the present study. This supports previous findings in the region (Koehn et al. 1984, Gartner-Kepkay et al. 1980). Differentiation was also found on three different spatial scales - hundreds of kilometers, hundreds of meters and tens of meters. Differentiation on such small

scales in Mytilus, an organism characterised by a life cycle of large dispersal abilities, is often viewed as an indicator of the presence of two or more taxonomic groups which are reproductively isolated to some degree (Skibinski et al. 1983, Koehn 1991) or as evidence of selection (i.e. Lap cline in northeastern United States, Hilbish and Koehn 1985).

Genetic differentiation is further seen in the MDS analysis, in which the samples are arranged in a two dimensional space according to the multidimensional allele frequency space. The samples did not cluster into discrete groups. Discrete groups would not be expected in a situation where the intragroup variation overlaps or exceeds the intergroup variation. Furthermore if there is hybridisation and introgression in areas where the groups co-occur, this will blur the distinction among the groups. The pattern that does emerge does not correspond to the relative geographic positioning of the samples. In some cases samples hundreds of kilometers apart fall closer together (i.e. are more genetically similar) in the MDS analysis than samples within tens of meters. This indicates that the genetic differentiation is not a simple function of geographic isolation of populations. Similar lack of concurrence of genetic and geographic distance was reported by Koehn et al. (1984) who found Group III samples in south-west N.S. and

Eastern Newfoundland with populations of Group II inbetween.

Population Structure

Several samples were not in HWE at the Pgm and Est loci due primarily to deficiencies of heterozygotes although only one population showed significant heterozygote deficiency at both loci. No sample showed significant heterozygote deficiency at all four loci. A deficiency of heterozygotes is a phenomenon frequently reported in Mytilus and other bivalves (Mallet et al. 1985, Singh and Green 1984). The Wahlund effect (i.e. the pooling of samples with different allele frequencies perhaps due to the subdivision of the breeding population into small partially isolated populations), has been advanced as a mechanism resulting in a deficiency of heterozygotes relative to predicted numbers under HWE (Koehn et al. 1976, Lassen and Turano 1978). Inbreeding or another non-random breeding structure could also result. In addition the presence of null alleles (allozymes devoid of any enzyme activity) and the possibility of mis-scoring gels have been discussed as the reason for the observed deficiency (Skibinski et al. 1983). A model of selection in which there is a heterozygote disadvantage in the larval stage and a heterozygote advantage post-settlement has been proposed based on the observation that in many bivalves the heterozygote

deficiency is greatest in younger age classes (Mallet et. al. 1985, Zouros and Foltz 1984, Singh and Green 1984). None of these theories is mutually exclusive and some combination of factors is likely acting to produce heterozygote deficiencies in Mytilus.

The results from the F-statistics analysis further test the assumption of panmixia. F_{ST} values were high for both Pgm and Est indicating significant genetic subdivision. In addition, F_{IS} values were high which indicates an excess of homozygotes within subpopulations. This was expected in view of the negative D-values which indicate a deficiency of heterozygotes for each locus. F_{IT} was also greater than F_{IS} in all cases due to the significant allele frequency differentiation among the samples (Skibinski et al. 1983). A similar finding in M. edulis / M. galloprovincialis populations in England was interpreted as evidence for the Wahlund effect, a consequence of two reproductively isolated groups that differ in allele frequency being pooled in the samples (Skibinski et al. 1983).

Multilocus associations

Variation in allele frequency at the Pgm and Est loci was not independent and these two loci were the most differentiated among samples. The Lap locus was also not independent of Pgm but was independent of Est. Random mating

within taxa will eventually result in a breakdown of locus associations even among linked loci (Hartl 1988). Therefore associations among loci may be maintained by an absence of interbreeding. Covariation among loci has been used as evidence for the presence of non-interbreeding groups (Skibinski et al. 1983). If the pattern was due to a cline in allele frequency involving many loci then the genetic distance among samples would be expected to correspond to the geographic distances among samples (Koehn et al. 1984).

Ap was not associated with any of the other three loci. In a world wide survey of Mytilus, Ap exhibited a pattern of geographical variation not consistent with proposed species boundaries of the three species (Varvio et al. 1988). It was suggested that this discordant pattern could result from genetic structures (i.e. chromosomes or segments of chromosomes) possessing different fitness values (Varvio et al. 1988). Genes located on these units could therefore cross taxonomic boundaries more easily than those which are more tightly linked to loci responsible for the systematic identity of lineages. According to this interpretation, discordant patterns such as that exhibited by Ap do not affect the interpretation of the main multilocus covariation pattern.

The fact that all loci are not equally differentiated may also be a result of sampling variance, linked loci

responding in groups to neutral processes or selection at the Pgm and Est loci, or a closely linked gene or gene cluster. Variation in the amount of differentiation among loci has been interpreted as the effects of natural selection operating on those loci or on loci to which they are strongly linked. Lewontin and Krakauer (1974) proposed the use of the variance of F_{ST} as a measure of population structure based on the theory that all loci should behave in the same manner if only neutral processes were involved. The difficulty in assigning meaningful probability values to the variance estimates has caused the statistic to fall out of use. If natural selection is operating, F_{ST} should differ for different enzyme functions, so that alleles involved in one ecologically adaptive function may vary independently of those involved in others (Bryant 1974). In addition, if some loci are selectively neutral in their polymorphisms and are not tightly linked to loci that are under the influence of natural selection, the alleles at these loci would show patterns of differentiation more closely correlated with the relative geographic positioning of the samples (Varvio et al. 1988).

A principal component analysis divided individuals into two groups based on their four-locus genotype. The alleles which contribute most strongly to the groups found were Pgm2, Est2, Pgm4 and Est1 as indicated by the loadings on

the eigen vectors. This corresponds to the previous taxonomic descriptions of M. edulis and M. trossulus (McDonald and Koehn 1988), with M. edulis having the faster Est2 allele and slower Pgm2 allele.

Consideration of data from individual sites reveals that each site contained a mixture of the two taxa, since there are always a few individuals from each PCA group present. This mixture of taxa is reflected in the MDS analysis, which creates two dimensional coordinates from the allele frequencies at all loci to describe each site. If the alleles which most strongly affect the PCA analysis are delimiting two discrete taxa, and if both taxa are present at each site, then the allele frequencies for each site are affected by the varying proportions of the two taxa and may result in the absence of discrete clusters.

These results present a more complicated picture than that presented by Koehn et al. (1984) who generally found pure edulis - type and pure trossulus - type populations in Atlantic Canada and only one population with both types. In the present study, most populations appeared to be a mixture of the two taxa. It was rare for a population to consist only of one taxon.

Evidence of interbreeding

The amount of interbreeding between the two taxa is unknown. Koehn et al. (1984) concluded there was no interbreeding based on the one sample in which they found both taxa. Using the technique of MD-Scale analysis it was proposed that this sample consisted of a simple mixture of the two taxa of mussel i.e. M. edulis (Group II) and M. trossulus (Group III) rather than hybrids. In the present study there are deficiencies of heterozygotes from HWE at most sites indicating that mating is not random, however the data from the Traytown analysis suggest that some interbreeding does take place at a microgeographic level. It is quite possible, if not probable, that the two taxa interbreed to a different extent in different localities as is the case for Mytilus edulis and M. galloprovincialis in England (Skibinski et al. 1983), for M. trossulus and M. galloprovincialis on the west coast of North America (McDonald and Koehn 1988), and for M. edulis and M. trossulus in the Baltic Sea region (Vainola and Hvilsom 1991).

The mussels in eastern Newfoundland do not seem to represent one panmictic population and the population structure appears to be complex. This complex breeding structure may be explained as a consequence of Newfoundland being an area where both species M. edulis and M. trossulus

co-exist with some hybridisation. The extended larval dispersal period of several weeks would allow for a 'hybrid zone' on this scale (Coustau et al. 1991). If there are differing levels of interbreeding, as reported between M. edulis and M. galloprovincialis (Skibinski et al. 1983), then the pooling of samples in the present study could yield a data set in which variation from several different sources is combined resulting in many intermediates between the proposed taxa.

Inner Shell Colour Variation

The colour of the inner shell is associated with genotype at Traytown (M. trossulus being darker than M. edulis), but this is not consistent on a macrogeographic scale, perhaps as a result of environmental factors affecting shell coloration. However, shell colour is difficult to quantify. It is probably no more useful than mantle edge colour as a taxonomic tool except that the shells can be stored for extended periods of time. (Mantle edge colour has been reported to have a 60% misclassification between M. edulis and M. gallprovincialis in some populations (Gosling 1984).) However the strength of the correlation at Traytown does suggest that some difference between the two taxa, physiological or biochemical, is resulting in different inner shell colour

but the extent to which this is under genetic control is not known. Previous studies on shell coloration, though primarily on external shell colour, have suggested many possible physiological or biochemical links. The external colour of Mytilus ranges from solid blue-black to solid brown to blue-black with lighter stripes and is under genetic control (Innes and Haley 1977, Newkirk 1980). Mitton (1977) proposed a theory of selection on external shell colour based on thermoregulation. Waldron et al. (1976) showed that variation in foot and mantle pigmentation is due to variation in the pigmentation of the epithelial cells. Seed (1971) also reprted that mantle colour was due to dense granules of pigment, of four colours, in the outer ciliated epithelium and noted that the overall colour of the mantle was due to the proportion and distribution of these granules. This suggests that a physiological/developmental effect causes the variation rather than cell damage or loss which would be more related to the condition of the animal. A study by Shurova and Zolotarev (1988) correlated shell colour phenotype with survival of epithelial gill cells in different salinities. Newkirk (1980) reported that brown shelled mussels had a slower growth rate and suggested that there was a link between the gene or genes determining shell colour and those controlling growth; perhaps these genes have primarily a metabolic role and shell colour varies as

an ancillary expression of the gene products. Many of the external colours of the clam Macoma balthica are due to the presence of different 'structural melanins or similar pigments'(Cain 1988). The presence of these pigments, or modifications made to allow their incorporation into the shell, may confer an additional strength or 'hardness' to the shell.

Unlike the outer shell, the inner shell is laid down continuously along the whole length of the shell (Meglitsh 1972). Inner shell colour may be a reflection of waste materials and may indicate physiological differences between the two types (Cain 1988, Comfort 1951). It is therefore likely to be sensitive to environmental fluctuations. The presence of many irregular spots of various colours of red and pink in the shells of both taxa supports this but more work on the direct causes of such variation is needed before any conclusions can be drawn.

Genotypic variation with length

There were significant differences in allele frequencies among size classes at all loci except Ap at the Traytown site. In particular, the frequency of edulis alleles (Pgm2, Est2), increased with increasing size. The fact that the trend is only slightly significant at the Lap locus and not significant at the Ap locus is not surprising

as the size-dependent effect is expected to be greatest for those loci which show the strongest differentiation between taxa.

Length-dependent genetic variation can be a result of different cohorts moving (i.e. growing) through a population. Different cohorts of larvae will almost certainly differ in allele frequencies at the time of settlement. This can be due to many factors including fluctuating allele frequencies of the parental population(s), interannual differences in local hydrodynamic patterns resulting in the recruitment of larvae from different populations of parents, and varying environmental factors (i.e. food availability) which could affect larval viability (Singh and Green 1984). Milkman and Koehn (1977) followed different cohorts of mussel larvae using allozyme markers indicating that differences in allele frequencyy among age classes may be simply due to local factors in a given year. Johnson and Black (1984) advanced a theory of patches in the environment becoming available for settlement in different years and being colonised by cohorts of differing allele frequencies resulting in "chaotic" microgeographic patterning of allele frequencies.

The effects of differential mortality and differential growth rates on length dependent genetic variation are difficult to separate. The increase in frequency of an

allele with increased size could be due to that individual living longer or growing faster. In previous studies (Skibinski and Roderick 1991, Milkman and Koehn 1977), selective mortality has been favoured as the cause of changing allele frequencies with increased size. Skibinski and Roderick (1991) showed this by demonstrating that while the relative frequency of galloprovincialis alleles increased with increasing size, the increase would be expected if the galloprovincialis individuals from the smaller size class survived in relatively greater numbers than edulis. (i.e. the absolute number of galloprovincialis alleles did not increase but the relative frequency increased as edulis alleles were removed). Unfortunately this approach cannot be repeated here because individuals cannot be easily separated into 'edulis' and 'trossulus' as there are many alleles at the loci involved, especially at the Pgm locus, which do not fit this simple definition. Boyer (1974) found similar increases in the frequency of one allele with increasing size and related his findings to environmental factors because mussels from different ecological areas had opposite trends. This was taken as support of selective mortality.

Gardner and Skibinski (1988) state that if differential mortality is acting to result in an increase in one mussel type, this must be balanced by immigration or by greater

reproductive capacity of the other, otherwise there will be a change in the most common taxa or allele over time. This potentially has happened at Traytown as Koehn et al. (1984) reported Traytown to be a pure trossulus site, however it is likely that all size classes were not represented in their sample and that only one patch with a high frequency of trossulus was sampled. Work done by Gardner and Skibinski (1988) on M. edulis and M. galloprovincialis in England suggests that allele frequencies do remain constant over time, at least in the short term (approximately 5 years).

Conclusions

The taxonomy of Mytilus species is highly complex due to the lack of diagnostic characters (Varvio et al. 1988). This study confirms that there are two genetically distinct groups which correspond to the descriptions of the species or taxa M. edulis and M. trossulus based on allozyme data (McDonald and Koehn 1988). These two taxa were found to co-exist over a broad area of eastern Newfoundland. However no single 100% diagnostic character was found so it was necessary to use the multivariate technique of PCA to describe the genetic variation. Recent studies on a world-wide basis have used allozyme data to conclude that M. trossulus is a valid species (Varvio et al. 1988, McDonald and Koehn 1988). Some evidence for interbreeding of the two

taxa was found in the present study which may imply that the status of species is not appropriate for M. trossulus. However the work on M. trossulus is not extensive and many basic questions including distribution, reproductive biology, physiology, growth rates, and mortality rates have not been adequately addressed.

This study has used allozyme data to address the question of the distribution of the edulis and trossulus mussel types in eastern Newfoundland. The distribution is patchy and is not "organised" by geographic proximity. Work is needed to determine if there is ecological and/or physiological differentiation. In addition, differentiation among size classes was found. This brings to light the problems of representative sampling. Fine scale distributions must be addressed in future surveys.

Literature Cited

Avise, J.C. (1974) Systematic value of electrophoretic data. Syst. Zool. 23, 465-481.

Boyer, J.F. (1974) Clinal and size dependent variation at the LAP locus in Mytilus edulis. Biol. Bull. 147, 535-549.

Bryant, E.H. (1974) On the adaptive significance of enzyme polymorphism in relation to environmental variability. Amer. Natur. 108, 1-19.

Bulnheim, H.P. and E. Gosling (1988) Population genetic structure of mussels from the Baltic Sea. Helgolander Meeresuntersuchungen 42, 113-129.

Cain, A.J. (1988) The scoring of polymorphic colour and pattern variation and its genetic basis in molluscan shells. Malacologia 28, 1-15.

Comfort, A. (1951) The pigmentation of molluscan shells. Biol. Rev. 26, 285-301.

Coustau, C., F. Renaud and B. Delay (1991) Genetic characterisation of the hybridisation between Mytilus edulis and M. galloprovincialis on the Atlantic coast of France. Mar. Biol. 111, 87-93.

Ferson, S., F.J. Rohlf and R.K. Koehn (1985) Measuring shape variation of two-dimensional outlines. Syst. Zool. 34,59-68.

Gardner, J.P.A. and D.O.F. Skibinski (1988) Historical and size dependant genetic variation in hybrid mussel populations. Heredity 61, 93-105.

Gardner, J.P.A. (1992) Mytilus galloprovincialis (Lmk) (Bivalvia, Mollusca) : The taxonomic status of the Mediterranean mussel. Ophelia 35, 219-243.

Gartner-Kepkay, K.E., L.M. Dickie and K.R. Freeman (1983) Genetic differentiation in the face of gene flow: a study of mussel populations from a single Nova Scotian embayment. Can. J. Fish. Aquat. Sci. 40, 443-451.

Gartner-Kepkay, K.E., L.M. Dickie, K.R. Freeman and E. Zouros (1980) Genetic differences and environments of mussel populations in the maritime provinces. Can. J. Fish. Aquat. Sci. 37, 775-782.

Gosling, E. (1984) The systematic status of Mytilus galloprovincialis in western Europe: a review. Malacologia. 25, 551-568.

Gosling, E.M. and D. McGrath (1990) Genetic variability in exposed shore mussels, Mytilus spp., along an environmental gradient. Mar. Biol. 104, 413-418.

Gosling, E.M. and N.P. Wilkins (1981) Ecological genetics of the mussels Mytilus edulis and M. galloprovincialis on Irish coasts. Mar. Ecol. Prog. Ser. 4, 221-227.

Gould, A.A. (1850) Shells from the United States exploring expedition. Proc. Boston Soc. Nat. Hist. 3, 343-348.

Hall, J.G. (1985) Temperature-related kinetic differentiation of Glucosephosphate isomerase alleleoenzymes isolated from the blue mussel Mytilus edulis. Biocem. Genet. 23, 705-728.

Hartl, D.L. (1988) A Primer of Population Genetics, Second Edition. Sinauer Associates, Massachussettes, U.S.A.

Hilbish, T.J. and R.K. Koehn (1985) The physiological basis of natural selection at the LAP locus. Evolution 39, 1302-1317.

Hodgson, A.N. and T.F. Bernard (1986) Observations on the ultrastructure of the spermatozoan of two mytilids from the south-west coast of England. J. Mar. Biol. Ass. U.K. 66, 385-390.

Hvilson, M.M. and B.F. Theisen (1984) Inheritance of allozyme variations through crossing experiments with the blue mussel, Mytilus edulis L. Herditas 101, 1-7.

Innes, D. and Haley (1977) Inheritance of a shell colour polymorphism in the mussel. J. Heredity 68, 203-204.

Johannesson, K., N. Kautsky and M. Tedengren (1990) Genotypic and phenotypic differences between Baltic and North Sea populations of Mytilus edulis evaluated through reciprocal transplantations.11. Genetic variation. Mar. Ecol. Prog. Ser. 59, 211-219.

Johnson, M.S. and R. Black (1984) Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. Evolution 38, 1371-1383.

Johnston, R.J. (1980) Multivariate Statistical Analysis in Geography; A primer on the General Linear Model. Longman Group Limited, New York.

Koehn, R.K. (1991) The genetics and taxonomy of species in the genus Mytilus. Aquaculture 94, 125-145.

Koehn, R.K., J.G. Hall, D.J. Innes and A.J. Zera (1984) Genetic differentiation of Mytilus edulis in eastern North America. Mar. Biol. 79, 117-126.

Koehn, R.K., R. Milkman and J.B. Mitton (1976) Population genetics of marine pelecypods. lV. Selection, migration and genetic differentiation in the blue mussel Mytilus edulis. Evolution 30, 2-32.

Koehn, R.K. and J.B. Mitton (1972) Population genetics of marine pelecypods. 1. Ecological heterogeneity and adaptive strategy at an enzyme locus. Am. Nat. 106, 47-56.

Lassen, H.H., and F.J. Turano (1978) Clinal variation and heterozygote deficit at the Lap-locus in Mytilus edulis. Mar. Biol. 49, 245-254.

Levinton, J.S. and H.H. Lassen (1978) Selection,ecology and evolutionary adjustment within bivalve mollusc populations. Phil. Tran. R. Soc. Lond. 284, 403-415.

Levinton, J.S. and T.H. Suchanek (1978) Geographic variation,niche breadth, and genetic differentiation at different geographic scales in the mussels Mytilus californianus and M.edulis. Mar. Biol. 49, 363-375.

Lewontin, R.C. and J. Krakauer (1973) in Genes in Populations 2nd Edition (1989). ed. E.B. Speiss. pp 600-601.

Mallet, A.L., E. Zouros, K.E. Gartner-Kepkay, K.R. Freeman and L.M. Dickie (1985) Larval viability and heterozygote deficiency in populations of marine bivalves: evidence from pair matings of mussels. Mar. Biol. 87, 165-172.

McDonald, J.H. (1985) No Bad Gels unpubl. manual.

McDonald, J.H. and R.K. Koehn (1988) The mussels Mytilus trossulus and M. galloprovincialis on the Pacific coast of North America. Mar. Biol. 99, 111-118.

McDonald, J.H., R.K. Koehn, E.S. Balakirev, G.P. Mancheko, A.I. Pudovkin, S.O. Sergievski, and K.V. Krutovskii (1990) Species identity of the "common mussel" inhabiting the Asiatic coasts of the Pacific Ocean. translated from Biologiya Morya 1, 13-22, in the Soviet Science translation series.

McDonald, J.H., R. Seed and R.K. Koehn (1991) Allozymes and morphometric characters of three species of Mytilus in the northern and southern hemispheres. Mar. Biol. 111, 323-333.

McElroy, D., P. Moran, E. Bermingham, and I. Kornfield (1991) REAP: The Restriction Enzyme Analysis Package, Version 4.0. University of Maine.

Meglitsh, P.A. (1972) Invertebrate Biology, Second Edition. Oxford University Press. Toronto.

Milkman, R. and R.K. Koehn (1977) Temporal variation in the relationship between size,numbers and an allele frequency in a population of Mytilus edulis. Evolution 31, 103-115.

Mitton, J.B. (1977) Shell color and pattern variation in Mytilus edulis and its adaptive significance. Chesapeake Sci. 18,

Murdock, E.A., A. Ferguson and R. Seed (1975) Geographical variation in leucine aminopeptidase in Mytilus edulis L from the Irish coasts. J. Exp. Mar. Biol. Ecol. 19, 33-41.

Munsell Neutral Value Scale (1971 Edition)

Nei, M. (1972) Genetic distance between populations. Amer. Nat. 106, 283-292.

Newkirk, G.F. (1980) Genetics of shell colour in Mytilus edulis L. and the association of growth rate with shell colour. J. Exp. Mar. Biol. Ecol. 47, 89-94.

Rice, W.R. (1989) Analyzing tables of statistical tests. Evol. 43, 223-225.

Rohlf, F.J. (1988) NTSYS-PC : Numerical Taxonomy and Multivariate Analysis System, Version 1.40. Exeter Publishing Ltd.

Scheltema, R.(1971) Larval dispersal as means of genetic exchange between geographically separated populations of shallow-water benthic marine gastropods. Biol. Bull. 140, 284-322.

Seed, R. (1969a) Factors influencing shell shape in the mussel Mytilus edulis. J. Mar. Biol. Assoc. U.K. 48, 561-584.

Seed, R. (1969b) The ecology of Mytilus edulis L. (Lamellibranchiata) on exposed rocky shores. I. Breeding and settlement. Oecologia 3, 277-316.

Seed, R. (1969c) The ecolgy of Mytilus edulis L. (Lamellibranchiata) on exposed rocky shores. II. Growth and mortality. Oecologia 3, 317-350.

Seed, R. (1972) Morphological variations in Mytilus from the french coasts in relation to the occurrence and distribution of M. galloprovincialis (Link). Cah. de Biol. Mar. 13, 357-384.

Seed, R. (1978) The systematics and evolution of Mytilus galloprovincialis (Link). In B. Battaglia and J.A. Beardmore (eds.), Marine Organisms: genetics, ecology and evolution. Plenum Press, New York, pp. 447-468.

Selander R.K., M.H. Smith, S.Y. Yang, W.E. Johnson, and J.B. Gentry. (1971) Biochemical polymorphisms and systematics in the genus Peromyscus. I. Variation in the old-field mouse. Studies in Genetics VI. Tex. Univ. Publ. 7103, 49-90.

Shurova, N.M. and V.N. Zolotarev (1990) Phenotypic differences in the survival of epithelial gill cells of Black Sea mussels in water with different salinities. Translated in the Soviet Translation Series from Biol. Morya 4, 31-35. (1988)

Singh, S.M. and R.H. Green (1984) Excess of allozyme homozygosity in marine molluscs and its possible biological significance. Malacologia 25(2), 569-581.

Skibinski, D.O.F., J.A. Beardmore and T.F. Cross (1983) Aspects of the population genetics of Mytilus (Mytillidae:Mollusca) in the British Isles. Biol. J. Linn. Soc. 19, 137-183.

Skibinski, D.O.F and E.E. Roderick (1991) Evidence of selective mortality in favour of the Mytilus galloprovincialis Lmk. phenotype in British mussel populations. Biol J. Linn. Soc. 42, 351-366.

Soot-Ryen, T. (1955) A report on the family Mytilidae (Pelecypoda). Allan Hancock Pacif. Exped. 20, 1-175.

Sneath, P. and R. Sokal (1973) Numerical Taxonomy. W.H. Freeman and Company. San Francisco, U.S.A.

Swofford, D.L. and R.B. Selander (1981) Biosys-1 : A Fortran program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. 72, 281-283.

Varvio, S.L., R.K. Koehn and R. Vainola (1988) Evolutionary genetics of the Mytilus edulis complex in the North Atlantic region. Mar. Biol.

Vainola, R. and M.M. Hvilsom (1991) Genetic divergence and a hybrid zone between Baltic and North Sea Mytilus populations (Mytilidae: Mollusca). Biol. J. Linn. Soc. 43, 127-148.

Waldron, M., R.M. Packie, and F.L. Roberts (1976) Pigment polymorphism in the blue mussel Mytilus edulis. Veliger 19, 82-83.

Zouros E. and D.W. Foltz (1984) Possible explanations of heterozygote deficiency in bivalve molluscs. Malacologia 25, 583-591.

Appendix 1
A : Tables used in Monte Carlo tests of heterogeneity for all survey samples using all individuals.

Locus/Sample	1	2	3	4	5	6
Pgm						
BV0	13	136	20	3	0	0
BV1	4	14	9	13	0	0
BV2	1	2	5	22	4	0
BV3	1	13	9	14	3	0
TT1	4	40	14	9	1	0
TT2	2	19	17	18	3	1
TT3	3	38	32	67	6	2
LHE	5	18	12	22	2	1
THL	2	22	8	7	1	0
CC1	1	6	7	20	6	0
CC2	5	25	6	3	0	1
FH	5	32	3	0	0	0
IA	3	20	3	11	1	0
EM	3	31	8	0	0	0
BB	9	24	6	3	0	0
TW	4	15	5	14	0	0
Est[1]						
BV0	10	155	3	1	1	
BV1	17	19	0	0	0	
BV2	27	3	1	0	1	
BV3	25	15	0	0	0	
TT1	19	44	1	0	0	
TT2	38	21	0	0	1	
TT3	95	44	1	0	2	
LHE	29	29	0	0	0	
THL	14	26	0	0	0	
CC1	36	4	0	0	0	
CC2	9	30	0	0	1	
FH	3	37	0	0	0	
IA	16	20	0	0	0	
EM	3	39	0	0	0	
BB	6	34	2	0	0	
TW	19	19	0	0	0	

[1] Est are alleles numbered as 1 2 3 4 8; 8 is slower than 1

A (cont.) : Tables used in Monte Carlo tests of heterogeneity for all survey samples using all individuals.

Locus/Sample	1	2	3	4	5	6
Lap						
BV0	0	0	2	18	31	1
BV1	0	1	4	13	20	0
BV2	1	1	11	16	5	0
BV3	0	0	3	9	22	0
TT1	0	1	8	12	16	1
TT2	0	1	11	19	9	0
TT3	0	5	18	15	3	1
LHE	0	2	9	12	7	0
THL	0	1	7	14	16	2
CC1	0	1	13	20	4	0
CC2	0	1	4	14	18	1
FH	0	2	3	8	27	0
IA	0	2	7	16	13	0
EM	0	3	7	18	12	2
BB	0	0	4	16	17	1
TW	0	0	11	17	11	1
Ap						
BV0	1	14	1	14	20	2
BV1	0	4	2	18	15	1
BV2	0	3	2	16	12	1
BV3	0	4	2	15	13	0
TT1	0	9	2	12	15	0
TT2	0	11	3	12	12	2
TT3	0	10	2	16	13	1
LHE	1	6	2	12	11	2
THL	0	3	3	20	13	1
CC1	0	3	1	24	11	1
CC2	0	8	2	10	18	0
FH	0	10	0	16	14	0
IA	0	3	0	14	21	0
EM	0	7	1	16	17	1
BB	0	4	3	8	21	2
TW	0	6	7	11	16	0

B : Tables used in Monte Carlo tests of heterogeneity for Bellevue sites BV0, BV1, BV2, and BV3.

Locus/sample	1	2	3	4	5	6
Pgm						
BV0	13	136	20	3	0	0
BV1	4	14	9	13	0	0
BV2	1	2	5	22	4	0
BV3	1	13	9	14	3	0
Est[1]						
BV0	10	155	3	1	1	-
BV1	17	19	0	0	0	-
BV2	27	3	1	0	1	-
BV3	25	15	0	0	0	-
Lap						
BV0	0	0	2	18	31	1
BV1	0	1	4	13	20	0
BV2	1	1	11	16	5	0
BV3	0	0	3	9	22	0
Ap						
BV0	1	14	1	14	20	2
BV1	0	4	2	18	15	1
BV2	0	3	2	16	12	1
BV3	0	4	2	15	13	0

[1] Est alleles are numbered 1 2 3 4 0; 0 is slower than 1.

C : Tables used in Monte Carlo tests of heterogeneity among Traytown sites TT1, TT2 and TT3.

Locus/Sample	1	2	3	4	5	6
Pgm						
TT1	4	40	14	9	1	0
TT2	2	19	17	18	3	1
TT3	3	38	32	67	6	2
Est[1]						
TT1	19	44	1	0	0	
TT2	38	21	0	0	1	
TT3	95	44	1	0	2	
Lap						
TT1	0	1	8	12	16	1
TT2	0	1	11	19	9	0
TT3	0	5	18	15	3	1
Ap						
TT1	0	9	2	12	15	0
TT2	0	11	3	12	12	2
TT3	0	10	2	16	13	1

[1] Est are numbered alleles 1 2 3 4 0; 0 is the slower than 1

Appendix 2

A : Contingency table of Est - Lap genotypes used in Monte Carlo X^2 test. (McElroy) N=302.

	Lap											
	11	12	13	14	22	23	24	33	34	35	44	45
Est												
11	1	6	7	0	13	31	4	21	7	1	5	0
12	0	2	0	0	4	10	6	8	8	0	11	0
13	0	0	0	0	0	0	1	0	0	0	0	0
22	0	0	2	3	2	9	11	26	32	3	55	5
23	0	0	0	0	0	0	0	1	2	0	1	0
81	0	0	0	0	0	1	0	1	0	0	0	0
82	0	0	0	0	0	0	0	0	1	0	1	0

B : Contingency table of Ap - Est genotypes used in Monte Carlo X^2 test. (McElroy) N=306.

	Ap															
	12	15	22	23	24	25	26	33	34	35	36	44	45	46	55	56
Est																
11	0	0	2	5	7	7	1	1	3	6	1	27	28	2	9	1
12	0	0	1	2	7	6	0	0	0	5	0	8	10	1	10	0
13	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
22	1	1	4	2	20	28	0	0	3	3	0	17	32	2	29	5
23	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
82	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0

C : Contingency table of Pgm - Est genotypes used in Monte Carlo X^2 test. (McElroy) N=468.

	Est							
	11	12	13	22	23	24	81	82
Pgm								
11	0	0	0	5	0	0	0	0
12	1	3	0	37	0	0	0	1
13	0	4	0	10	0	0	0	0
22	1	24	0	136	5	1	0	2
23	4	10	0	36	3	0	0	2
24	13	7	0	2	0	0	0	0
33	15	3	0	5	0	0	1	0
34	32	5	0	2	0	0	0	0
35	5	2	0	0	0	0	0	0
36	1	0	0	0	0	0	0	0
44	55	10	1	0	0	0	1	0
45	16	1	0	0	0	0	0	0
46	4	0	0	0	0	0	0	0
55	2	0	0	0	0	0	0	0

D : Contingency table of Lap - Ap genotypes used in Monte Carlo X^2 test. (McElroy) N=310.

	Ap															
	12	15	22	23	24	25	26	33	34	35	36	44	45	46	55	56
Lap																
3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
12	0	0	1	0	0	1	0	0	0	0	0	3	2	0	1	0
13	0	0	0	0	0	0	0	0	1	1	0	1	4	0	1	0
14	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0
22	0	0	0	1	3	1	1	1	1	3	0	4	2	0	2	0
23	0	0	1	4	4	4	0	0	0	3	1	16	13	2	4	1
24	0	0	0	1	4	5	0	0	1	0	0	1	3	2	6	0
33	1	1	0	2	8	4	0	0	1	4	1	11	17	1	8	1
34	0	0	1	1	6	8	0	0	0	2	0	3	20	1	8	1
35	0	0	0	0	1	0	0	0	1	0	0	2	0	0	0	0
44	0	0	2	0	10	17	0	0	1	1	0	10	14	0	16	3
45	0	0	1	0	0	0	0	0	1	0	0	1	0	0	3	0

E : Contingency table of Pgm - Ap genotypes used in Monte Carlo X^2 test. (McElroy) N=314.

	Ap															
	12	15	22	23	24	25	26	33	34	35	36	44	45	46	55	56
Pgm																
11	0	0	0	0	0	2	0	0	0	0	0	0	1	1	1	0
12	0	0	0	0	5	6	0	0	2	1	0	2	8	0	7	0
13	0	0	1	0	3	0	0	0	0	1	0	0	2	0	1	0
22	1	1	3	2	11	20	0	0	0	1	0	13	27	1	20	4
23	0	0	1	1	4	4	0	0	1	1	0	5	4	1	12	1
24	0	0	0	0	2	1	0	0	0	0	0	5	5	0	1	0
33	0	0	1	0	0	0	0	0	0	4	0	4	3	1	1	0
34	0	0	0	2	0	3	0	0	1	3	0	7	12	0	2	1
35	0	0	0	2	2	0	0	0	0	1	0	0	1	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
44	0	0	2	2	7	4	1	1	3	2	1	10	13	2	2	0
45	0	0	0	0	1	1	0	0	0	0	0	5	3	0	2	0
46	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0

F : Contingency table of Pgm - Lap genotypes used in Monte Carlo X^2 test. (McElroy) N=310.

	Lap											
	11	12	13	14	22	23	24	33	34	35	44	45
Pgm												
11	0	0	0	0	0	0	2	0	3	0	0	0
12	0	0	0	1	1	2	2	3	10	1	9	2
13	0	0	0	1	0	0	0	3	2	0	2	0
22	0	0	1	1	1	7	6	18	27	2	41	1
23	0	0	2	0	1	3	3	10	1	0	13	2
24	0	0	1	0	1	4	2	4	1	0	1	0
33	0	1	1	0	1	3	0	5	0	0	3	0
34	0	4	2	0	3	10	2	5	2	0	0	1
35	0	0	0	0	2	1	1	1	0	0	1	0
36	0	0	0	0	0	0	0	1	0	0	0	0
44	0	3	1	0	7	21	5	5	2	1	3	0
45	1	0	0	0	1	1	0	5	2	0	1	0
46	0	0	0	0	0	0	0	0	1	0	0	0
55	0	1	0	0	1	0	0	0	0	0	0	0

Appendix 3
A : Contingency table of genotype versus inner shell colour used in Chi-square tests of association using a Monte Carlo technique (Reap, McElroy 1991).

Genotype	Dark	Grey	Pink	White
Pgm				
11	0	0	0	6
12	5	2	0	49
13	3	1	0	10
14	6	0	0	0
22	19	8	1	198
23	14	3	1	54
24	18	2	0	9
25	2	0	0	0
33	26	5	0	13
34	64	9	0	6
35	12	1	0	2
36	1	0	0	0
44	103	6	0	18
45	16	0	0	7
46	3	0	0	0
55	3	0	0	0
56	2	0	0	0
Est				
11	179	11	0	37
12	41	7	1	59
13	0	0	0	1
22	23	7	1	247
23	0	0	0	8
24	0	0	0	1
33	0	0	0	4

A (cont.) : Contingency table of genotype versus inner shell colour used in Chi-square tests of association using a Monte Carlo technique (Reap, McElroy 1991).

Genotype	Dark	Grey	Pink	White
Ap				
12	0	0	0	1
15	0	0	0	1
22	8	1	0	12
23	3	0	0	5
24	26	7	0	31
25	23	2	1	42
26	3	0	0	0
33	0	0	0	1
34	14	1	0	3
35	12	0	0	7
36	1	0	0	2
44	47	10	0	39
45	70	7	0	55
46	3	0	0	4
55	36	7	1	52
56	1	0	0	5
66	1	1	0	0
Lap				
11	7	0	0	0
12	30	5	0	1
13	17	3	0	4
14	4	0	0	3
22	35	4	0	9
23	63	8	0	29
24	11	2	0	16
25	2	1	0	0
33	41	7	1	44
34	16	2	1	63
35	3	0	0	3
44	10	3	0	82
45	1	0	0	5

Appendix 4
A : Tables used in Monte Carlo tests of heterogeneity among size classes in the random Traytown samples.

Locus/size	Allele 1	2	3	4	5	6
Pgm						
1-5mm	0	0	4	8	2	0
6-10	1	3	17	47	5	3
11-15	0	1	5	7	1	0
16-20	4	9	14	10	1	0
21-25	1	7	8	25	3	0
26-30	4	13	18	44	3	0
31-35	2	9	15	27	3	0
36-40	2	11	13	14	3	1
41-45	1	21	2	13	2	1
46-50	2	18	11	15	4	0
51+	5	33	4	4	0	0
Est						
1-5mm	0	2	0	0	-	-
6-10	0	2	0	0	-	-
11-15	0	2	10	2	-	-
16-20	0	24	14	0	-	-
21-25	0	33	5	0	-	-
26-30	0	34	6	0	-	-
31-35	0	37	9	0	-	-
36-40	0	22	18	2	-	-
41-45	1	17	15	3	-	-
46-50	0	26	24	0	-	-
50+	0	12	24	0	-	-
Ap						
1-5mm	0	0	0	3	3	0
6-10	0	14	5	24	19	2
11-15	0	4	1	6	3	0
16-20	0	8	3	16	6	1
21-25	0	8	2	17	16	1
26-30	0	19	3	32	27	1
31-35	0	6	4	29	17	0
36-40	0	7	1	21	14	1
41-45	0	6	2	17	13	2
46-50	0	7	3	21	16	1
51+	0	9	0	8	26	3

Appendix 4
A (cont.) : Tables used in Monte Carlo tests of heterogeneity among size classes in the random Traytown samples.

Locus/size	Allele 1	2	3	4	5	6
Lap						
1	0	2	0	0	0	-
2	2	5	3	0	0	-
3	3	1	9	1	0	-
4	3	8	11	6	0	-
5	2	12	21	7	0	-
6	10	32	31	6	1	-
7	16	16	16	6	0	-
8	3	12	22	7	0	-
9	5	11	16	8	0	-
10	4	15	16	14	1	-
11	1	9	15	17	0	-