DNA Markers and Genetic Stock Structure in Commercial Species of Penaeid Prawns in the East Coast Fishery

Dr Shane Lavery

DPI
DEPARTMENT OF PRIMARY INDUSTRIES

FISHERIES RESEARCH & DEVELOPMENT CORPORATION

Project No.94/165

Table of Contents

List of Figures

List of Tables

Summary

- The DNA techniques of mitochondrial DNA sequencing, PCR-RFLP's and microsatellite DNA VNTR's are extremely useful in detecting large amounts of genetic variation in species of penaeid prawn.

- All these techniques have been refined to allow rapid screening of genetic variation in populations of *P. plebejus* and *P. esculentus* along the east coast of Queensland.

- *P. plebejus* appears to have a relatively panmictic population structure along the east coast of Qld. After examining genetic variation in 295 individuals from 13 locations throughout the species' distribution, considerable variation was found within populations, but little difference was detected among populations.

- *P. esculentus* appears to have considerably greater population structure throughout its Australian distribution, but very little along its east coast distribution. After examining genetic variation in 175 individuals from 10 locations throughout the species' distribution (all but two sites in Qld waters), considerable variation was found within populations, but, once again, little difference was detected among east coast populations, although it did appear to be greater than that found in *P. plebejus* . In contrast, the outlying Western Australian sample (from Shark Bay), which was used for comparison, is substantially genetically differentiated from all east coast populations.

- It therefore appears that genetically distinct stocks of prawn of either species do not exist along the Qld east coast. This does not necessarily suggest that the proportions of the stock of either species that successfully migrate and breed are a considerable proportion of the species, but may merely mean that there is a small proportion that does so. This proportion could well be less than 1%, but is difficult to quantify more accurately. Genetic techniques are best suited to estimating levels of migration that are lower than what appears to be occurring among east coast prawn populations.

- These results suggest that there are no genetic reasons why the east coast fisheries of both species should not be managed as unit stocks. In the case of *P. plebejus* , this means that the fisheries in both NSW and Qld are likely to be very interdependent, and largely sharing the same resource. In the case of *P esculentus*, this means that there appears, on this evidence at least, to be no genetic justification for segregating the stock into different management units. Although these are "negative" results, they are still of considerable importance to the fisheries.

- The high levels of genetic variation detected in these species, and the refinement of the techniques for doing so, provides a very important base for further genetic research in these and other species of penaeid prawn, be that in stock discrimination on a larger scale, or in the rapidly-expanding field of prawn aquaculture.

1. Background

Penaeid prawns constitute an extremely valuable fisheries resource in Australian coastal waters ($259 million in 1992/93 - 19% of Australia's total fisheries production). Over 40 species of penaeid prawn are found in waters around the continent, with at least ten of these forming the basis of significant commercial fisheries. Among the most valuable species are *Penaeus esculentus* (the brown tiger prawn) and *P. plebejus* (the eastern king prawn). These species comprise a considerable component (approx. 60%) of the catch of commercial prawns along the Queensland coast.

A knowledge of the stock structure of species taken in wild fisheries is essential for the successful management of the resource. Without information on the number and extent of discrete stocks of any one species, it is impossible to determine a suitable regional basis for management plans, or the likely impact that management strategies will make.

The existing understanding of the population dynamics of the eastern king prawn includes a northerly-migrating breeding population, recruitment to estuaries assisted by southerly drift of the east Australian current, growth, and subsequent recruitment into the estuarine and oceanic fishery. Principally based on tagging studies, the assumption has been that a single interbreeding population exists. However, this is based on evidence from only a small number of tag returns. Recent evidence suggests that only a small proportion of the stocks actually migrate, and that these individuals may contribute very little to subsequent recruitment.

Management of the eastern king fishery is shared between Queensland and New South Wales. However, quite different management issues have arisen in each state. The expansion of effort to the south and east in New South Wales has been coupled to declines in the oceanic catch of eastern king prawns in southern Queensland. Coincidentally, there has been an expansion of the Queensland fishery into deeper waters north-east of Fraser Island. Management is faced with uncertainty with regard to the efficacy of some of the basic assumptions about this fishery, and require more details of the stock structure of the species in order to make future management decisions. Similar uncertainties face management of the brown tiger fishery.

There are two primary methods of obtaining information on stock structure: physical tag-recapture studies and genetic studies. Tag-recapture studies can provide very useful information. However, they also have some major drawbacks, particularly for prawns. There are restrictions on the size of prawns that can be tagged, the rate of tag loss can be high due to moulting, and tagging can result in increased mortality. Furthermore, physical tags provide no information on the movement of larval or juvenile stages, and their data on adult movement is restricted to only one point in time. Finally, considerable cost and effort is required to undertake a satisfactory tag-recapture study.

Genetic studies have previously not proven particularly useful in prawns, due largely to the absence of suitable genetic markers (see below). However, recently-developed highly-variable DNA markers now allow the advantages of genetic stock structure studies to be exploited in prawns. Genetic studies detect migration of all growth stages, over many generations. They also utilise natural genetic tags, without interfering with normal behaviour, and are extremely cost- and time-efficient.

Genetic analysis provides a conservative estimator of stock structure. That is, if movement between stocks has been restricted only in recent years, or if movement between stocks is substantial (e.g. > 20%), then genetic differences between stocks are unlikely to be detected. On the other hand, this

conservatism also means that where genetic stock structure is detected, this is strong and compelling evidence that movement among locations has been substantially restricted for many years, and thus is a very significant aspect of the species' population dynamics. In other words, all stocks identified genetically are true stocks in the fisheries sense (i.e., "relatively homogeneous and self-contained populations whose losses by emmigration and accessions by immigration, if any, are negligible in relation to the rates of growth and mortality" - from R. Harden-Jones, 1994), but weakly-defined stocks may not be detected genetically. The recent rapid development of both laboratory and statistical techniques in this field now allows much greater precision in identifying distinct stocks than previously.

The importance of examining genetic stock structure across a large range of the species cannot be over-stated. Some knowledge of stock structure can certainly be obtained from sampling a restricted range, but the importance and true meaning of the results can be properly interpreted only in the context of wide-range sampling. Firstly, the suitability of the genetic markers used can be determined only by examining widely-separated locations that are known to have distinct stocks. Secondly, the significance of local-scale stock differences can be interpreted only by examining the type, pattern and magnitude of stock structure on a broad scale.

The initial genetic research on Australian penaeid prawns was undertaken by Mulley and Latter using allozyme electrophoresis[1,2]. The general interpretation of their work in the literature has been that tropical penaeid species are homogeneous throughout their range[3,4]. This is an overly-conservative interpretation, particularly as their information for many species was based on limited data (which they acknowledge). In addition, their conclusions were limited by the genetic techniques employed in their studies. Allozyme electrophoresis can detect only a small fraction of the genetic variation which actually exists in a species, and has been shown in many other studies to reveal very little variation in decapod crustaceans[5,6,7,8,9]. Various DNA techniques are now available which directly detect variation in the genetic code. Of particular use in population genetic studies is mitochondrial DNA (mtDNA), which has properties which make it relatively easy to analyse, and which appears to evolve at a faster rate than most nuclear DNA[10]. This means that DNA analysis, in particular mtDNA analysis, can detect far more of the existing population subdivision in a species than allozyme techniques, and with smaller sample sizes. Such DNA techniques have been used very successfully in a number of species, including decapod crustaceans[11,12,13], revealing much higher levels of genetic variation and population subdivision than is apparent from allozyme analyses.

2. Need

This project addresses the need for two distinct, but closely linked types of data: the need for highly variable genetic markers in penaeid prawns, and the need for details of stock structure in the major east coast prawn fisheries.

2.1 Highly variable genetic markers

Previously, the greatest need for such markers has been for the purpose of genetic stock discrimination. Now, however, there are many additional requirements for these markers. Particularly in prawns, physical tags have not proven very effective, and alternatives are now being sought for a variety of purposes, including: natural genetic tags for re-stocking, for tracing dispersal,

survival and growth, and for forensic identification (post-capture determination of species and/or location of capture).

Prior to the current research, there have not been sufficiently variable genetic markers available in penaeid prawns (see Background). Highly-variable DNA markers are now being developed rapidly for a great range of species. Up till now, very little of this work has been undertaken in prawns, but there is an enormous potential for this newly-developed technology in Australian prawn fisheries.

2.2 Stock structure in the east coast prawn fisheries

Both industry and management are keenly interested in actual stock structure of the eastern king prawn. Currently, this fishery is jointly managed by Queensland and New South Wales on the assumption of it being a single interbreeding east coast stock. However, recent evidence suggests that recruitment in this species may in fact be quite localised (see Background). Genetic information which could either support or refute the single-stock assumption would be extremely valuable. If genetic information refutes the assumption, then managers have indicated that separate management regimes may be considered. If there are distinct stocks of eastern king prawns, managers in each State would have quite different management strategies. In particular, the evolution of fishing effort and its direction would be viewed quite differently.

Similarly, industry and management are interested in whether regional genetic differences may exist in *P. esculentus*, given the interest in area management as a concept. The information from the proposed project will help managers decide on whether there is any biological justification to considering such management. If distinct prawn stocks are detected along the eastern coast, a major implication will be that management by zoning is required, with distinct assessment and management of the resources within each zone. Depending on the geographical extent of the unit stocks that will be defined, existing fishing closures may be seen to be unnecessary, and thus re-opened, or new closures may be recommended.

In addition, if current suggestions for re-stocking *P. esculentus* in Moreton Bay are carried through, there will be a critical need for background information on the current genetic diversity and local stock structure of this species.

Apart from the management requirement for the accurate delineation of unit stocks, such delineation is also of critical importance in determining the existence and type of stock-recruitment relationships (SRRs) in penaeid prawns. At the 1995 FRDC-funded SRR Workshop held at Bribie Island, a number of speakers, including Dr Ray Hilborn, Dr Derek Staples and Dr Nick Caputi highlighted the importance of being aware of the existence of distinct stocks in examining SRRs. It is vital that stock size and recruitment indices are actually from the one stock, and that they be subdivided if distinct stocks are present. It is only then that the correct relationship between stock size and recruitment can be determined. Thus knowledge of the stock structure is crucial to the setting of sustainable harvesting yields, and to the recognition of recruitment overfishing. The need for more detailed knowledge of stock structure was identified at the SRR Workshop as having high priority in the east coast fishery, and genetic studies were recognised as the best means of acquiring this knowledge.

While it is may be that socio-political factors are often paramount in fisheries management decisions, it is also obvious that the biological realities must be considered as well - at least by the scientists and managers entrusted with assessing the stocks and maintaining their sustainable

harvesting, if not by the politicians. The potential cost of not recognising the existence of distinct stocks could be enormous, with unsustainable fishing pressure being unwittingly permitted on specific stocks, with resulting stock declines. Few world fisheries have shown themselves to be immune to this.

Finally, a number of other Penaeid species fisheries are commercially important and are similarly lacking in any detailed knowledge of stock structure. The techniques developed and refined in this research will permit rapid and cost-efficient analyses to be undertaken on the population genetics of other Penaeid species.

3. Objectives

- Assess the utility of a range of DNA techniques for examining genetic variation within species of penaeid prawn.
- Refine those techniques for examining east coast prawn stocks.
- Describe the genetic stock structure of two commercial species (*Penaeus esculentus* and *P. plebejus*) throughout their East Coast distribution.
- Determine whether genetically distinct stocks of these species exist in the east coast fishery and, if so, locate the boundaries between such stocks.
- Quantify the potential rates of effective prawn migration between regions for each species.
- Provide specific advice on the management implications of the results.

In the initial phase of the work it was intended to examine a range of DNA techniques for their suitability in examining population structure in penaeid species. The results of this research could prove to be of far-reaching value for the future study of penaeid prawns in Australia. Very efficient and sensitive genetic techniques would permit highly cost-effective and detailed population studies to be easily undertaken on many of the other commercial species. In addition, these techniques will provide invaluable tools for employing natural genetic tags in future studies of re-stocking and dispersal.

4. Methods

4.1 Sampling

All samples were taken from research or commercial trawlers, and immediately frozen (at -20°C or below) until transport to the laboratory, where they were kept at -70°C or -20°C. Remaining tissue samples are still kept at -20°C.

4.1.1 *P. plebejus*

Samples of *P. plebejus* were collected from throughout the species' distribution along the east coast of Australia. The size and dates of the collections are shown in Table 4.1.

Table 4.1 - Sizes and dates of the *P. plebejus* sample collections

Location	Collection Date	Number of Individuals
Swain Reefs	25-5-92	150
Lady Musgrave I.	15-6-92	160
Mooloolaba	7-7-92	62
Moreton I.	5-7-93	20
Moreton Bay	9-4-93	20
Coffs Harbour	10-11-93	100
Wallis Lake	29-3-94	45
Lake Tuggerah	29-3-94	25
Lake Illawarra	29-3-94	25
Lake Brau	16-3-94	19
Bombo Lake	16-3-94	5
Tilba Lake	16-3-94	5
Mallacoota	16-3-94	25

4.1.2 *P. esculentus*

Samples of *P. esculentus* were collected from throughout the species' distribution along the east coast of Queensland, plus a collection from the extreme of the species' distribution in Western Australia. The size and dates of the collections are shown in Table 4.2.

Table 4.2 - Sizes and dates of the *P. esculentus* sample collections

Location	Collection Date	Number of Individuals
Moreton Bay	9-4-93	10
Mackay	11-11-94	15
Turtle Group	1-6-93	10
	6-8-94	39
Torres Strait	15-9-94	27
Nichol Bay (W.A.)	29-10-93	14

4.2 Development of DNA markers

4.2.1 DNA extraction

4.2.1.1 Sample storage

A variety of storage methods were examined to determine which methods could be used successfully to store tissue samples prior to DNA extraction. The method primarily used, and the one which generally proved easiest for collectors on board trawlers, was to freeze the samples as soon as possible at -20°C or below. Where trawlers did not have blast freezers, (e.g. Moreton Bay) samples were thrown onto dry ice until they could be placed in a freezer.

To test what alternative storage methods would prove feasible, two other methods were tried: storage in 95% alcohol, and drying the tissue. For storage in alcohol, rapid penetration of the tissue by the alcohol is necessary, therefore only the pleopods were placed in a 1.5ml microcentrifuge tube with ~1ml of 95% alcohol. These samples could be stored at room temperature for short periods (up to a week or so) before placing at 4°C. The same DNA extraction protocols could be used on these tissues, except that some prior treatment was required. The pleopods were removed from the alcohol, drained, and soaked in at least one bath of TE buffer for 5 minutes to completely remove the alcohol.

For drying the tissue, the objective is to obtain rapid drying of the tissue, before decomposition can begin. Small masses of tissue work best, so walking legs (periopods) were chosen. Two or more legs were removed from an animal, and crushed by fingers between a small piece of folded filter paper, which was then sealed in a small paper envelope, and kept in a dry, ventilated container. These envelopes were simply stored in a drawer in an air-conditioned laboratory (i.e., kept in dark, cool, low humidity conditions).

Samples stored in both these ways provided perfectly adequate DNA (using a phenol/chloroform extraction protocol) after more than one year's storage. An additional storage medium which has also proved adequate for similar samples is a saturated NaCl / 20% DMSO (dimethylsulfoxide) solution. Again, samples stored in this buffer can be kept at room temperature for short periods, and stored at 4°C in the longer term.

4.2.1.2 Tissue choice

Four prawn tissues (abdominal muscle, pleopod, hepatopancreas and ovary) were tested for their suitability for DNA extraction. Total DNA could be extracted from all tissues, but with varying degrees of success. Hepatopancreas requires little prior homogenisation of the tissue before DNA extraction can begin, and can provide abundant amounts of DNA if the tissue is fresh or preserved in good condition. However, the high concentrations of autolytic and nuclease enzymes in this tissue also means that the tissue (and DNA) can degrade very quickly. After only 5 minutes at room temperature, the DNA yield from a well-preserved frozen hepatopancreas dropped by approximately 90% compared to a sample processed immediately. This tissue is therefore not recommended for routine DNA extraction.

Abdominal muscle could also provide perfectly adequate DNA yields for most purposes. However, in comparison to pleopod muscle ("red" muscle), abdominal muscle has a much tougher texture, requiring considerable additional initial homogenisation and digestion. Furthermore, pleopod muscle provided DNA yields approximately 5 times that of abdominal muscle, as well as a beneficial higher ratio of mtDNA to nDNA (due to a higher concentration of mitochondria in this tissue).

Ovary provides the richest source of mtDNA, and was used successfully for at least one entire sample collection, however its very high lipid concentration can prove difficult in standard extractions.

4.2.1.3 Protocols

A number of DNA extraction protocols were compared in an attempt to determine the best rapid protocol for processing large numbers of samples. Before briefly describing the protocols tried and their performance, the overall simple conclusion that can be drawn is that, in general, the more

effort expended in extraction and purification of DNA from prawn tissue, the greater the yield of DNA, the better the quality of DNA produced, the longer it can be stored without degradation, the easier it is to amplify in a PCR, and the fewer the individuals that will not amplify. Therefore, if DNA extraction is required from a large number of well-preserved samples for immediate once-only use in PCR, then the most simple of extraction protocols will almost certainly suffice. However, if any of these parameters are not optimal (that is, only a small number of individuals in the sample, or samples are not well preserved, or DNA extractions are wanted for repeated or long-term use) then it is usually more efficient to use a more refined extraction protocol from the start.

4.2.1.3.1 Simplified phenol/chloroform extraction

DNA was extracted from frozen pleopod muscle which was removed from the prawn while still frozen. About 100 mg of tissue (that from about two large pleopods) from each individual was placed in a microcentrifuge tube. 500 ml lysis buffer (100 mM EDTA, 10 mM Tris, 1% SDS, pH7) and 10ml Proteinase K (10 mg/ml) were added and the tissue was homogenised immediately with a small pestle. The samples were digested in a water bath (60°C) for 2 hours, with regular inversion for mixing. To remove RNA (necessary only if spectrophotometric determination of DNA concentration was required), 5ml of DNA free RNAse was mixed into the solution and the tubes were placed in a 37°C water bath for 1 hour.

The following steps were carried out at room temperature. The tubes were spun for 10 minutes at 13,000 rpm to pellet cell debris. The supernatant was transferred to clean tubes and extracted three times, carefully pipetting the upper fraction each time into clean tubes, being careful not to take any of the interface. The first extraction was with an equal volume of (pH8) buffered phenol (500ml) which was mixed gently for one minute and centrifuged at 13,000 rpm for 10 mins. The second extraction was done with phenol/chloroform/isoamyl alcohol (49:49:2 - 500ml), which was mixed gently for one minute and centrifuged at 13,000 rpm for 5 mins. The final extraction used chloroform/isoamyl alcohol (24:1 - 500ml) which was mixed gently for one minute and centrifuged at 13,000 rpm for 3 mins. This final supernatant (approximately 250ml) was transferred to clean tubes and precipitated with ½ volume (125ml) of 7.5 M ammonium acetate and 3 volumes (750ml) of ethanol. This mixture was inverted slowly to precipitate the DNA (which could often be seen) and allowed to sit for at least 10 mins before centrifuging at 13,000 rpm for 10 mins. The supernatant was carefully discarded and the pellet (loosened by flicking) was washed with 1 ml of 70% ethanol, with gentle inversion. The tube was again centrifuged at 13,000 rpm for 10 mins, the supernatant very carefully discarded and the pellet dried under vacuum in a SpeedVac for 5 mins or air dried for 1 hr.

Finally, the pellet of DNA was resuspended in 50-100ml of TE in a water bath at 37°C overnight. The presence of RNA and the approximate concentration of large fragment DNA was checked the next day by electrophoresis of 3ml of sample on a 0.6% TBE agarose gel run at 80 volts (~50 mA) for 1-2 hours with a 1 sample control. Also, a 1:50 dilution was checked on a spectrophotometer to ascertain the concentration of nucleic acid and protein.

4.2.1.3.2 Guanidine hydrochloride extraction

In this protocol, about 100mg of tissue was ground in a microcentrifuge tube after freezing in liquid nitrogen. This ground material was then quickly transferred to a 10ml Falcon tube already prepared with 2.5ml of 20:5 TE and 250µl 10% sarcosyl. 35µl of 20mg/ml proteinase K was added, and the solution mixed and incubated at 60°C for 2 hr, with mixing every 15 mins. This solution was then

added to a 50ml Falcon tube containing 7.5ml 8M guanidine hydrochloride and 0.5ml 7.5M ammonium acetate. The tubes were placed on a shaker tray for 1hr before adding 27ml 100% ethanol and gently inverting for a minute. The precipitate was pelleted at 3000rpm for 5 min. and the supernatant decanted off before washing the pellet with 10ml 70% alcohol into a microcentrifuge tube. The DNA was again pelleted, allowed to air-dry, and resuspended in 50µl TE overnight.

4.2.1.3.3 CTAB extraction

Once again, 100mg of tissue was ground in liquid nitrogen, before adding to 500µl of CTAB buffer (0.1M Tris-HCl pH8.3, 0.02M EDTA pH8.0, 1.4M NaCl, 0.2% ß-mercaptoethanol, 0.05 cetyltrimethylammonium bromide) and incubated at 65°C for 2hr with occasional shaking. The tissue debris was pelleted by centrifugation for 10 min., followed by two chloroform/isoamyl alcohol extractions (500µl each, 2 min mixing, 2 min. centrifugation, transfer supernatant to new tube). The DNA was precipitated with 50µl of 7.5M ammonium acetate and 500µl of -20°C ethanol, and spun for 15 min. The pellets were dried and resuspended in 50µl TE overnight.

4.2.1.3.4 Rapid methods

These rapid commercial DNA extraction protocols strictly followed the recommendations of the suppliers (BioRad: Instagene and Chelex; Molecular Research Centre Inc.: DNAzol; Qiagen: QIAamp).

Instagene and Chelex are very rapid preparation types where a matrix is added to a crude preparation containing a minute piece of tissue, and subjected to a short period of heating or boiling followed by centrifugation. Most PCR contaminants should become bound to the matrix, leaving some free DNA to serve as template for PCR. Instagene-prepared samples did not appear to work well on prawn tissues. Chelex-prepared samples appeared to perform better in PCR reactions, however, the DNA extracts did not seem to last well, and successful PCRs could not be performed using these samples after a couple of months storage. Further experimentation using these techniques may prove useful in the longer term for very rapid preparation for large numbers of samples, but were not pursued for this project.

QIAamp DNA extraction used a standard proteinase K digestion followed by the binding of DNA to a matrix in a microcentrifuge column format, washing and elution. This technique produced good quality DNA of relatively high molecular weight (although somewhat sheared) in a shorter time than regular phenol/chloroform extractions (and without the use of harmful phenol), but was relatively expensive, and could not be realistically used for large numbers of samples.

A compromise between these two extremes in commercial protocols was DNAzol, which was used in this project more extensively. This method employs a guanadine-based buffer to rapidly digest the tissue and extract the DNA, before a crude ethanol precipitation. Although this extraction protocol produced only low yields and relatively poor quality DNA, it was very fast, and the DNA extracted was of sufficient purity to act successfully as PCR template. This protocol would not be recommended for samples in poor condition or of great value.

4.2.2 mtDNA amplification

4.2.2.1 PCR primers, amplification and optimisation

Each of the following PCR fragments amplified were checked that they were actually of mitochondrial DNA origin. This was done by comparing the amplification products from total DNA template against the products from highly purified (caesium-chloride ultracentrifugation prep) mitochondrial DNA template. These products were compared in both size and sequence. In no case were alternative fragments, of possible nuclear origin, found.

4.2.2.1.1 16s and COI

The PCR amplification used 1 µl of 1/10 dilution of template in a 50 µl reaction. This was added to 49 µl of reaction mixture which differed for each set of primers, as set out in Table 4.3, below. The initial universal primers used for both cytochrome oxidase I (COI) and 16s rRNA (16s) genes were from Simon et al. (1991). They were:

COIa (21mer) 5' - AGTATAAGCGTCTGGGTAGTC -3'
COIf (20mer) 5' - CCTGCAGGAGGAGGAGAYCC -3' (Y - C or T)
16sar (20mer) 5' - CGCCTGTTTAACAAAAACAT -3'
16sbr (22mer) 5' - CCGGTCTGAACTCAGATCACGT -3'

Table 4.3 - Concentration and volumes of solutions in reaction mixtures for PCR amplification of COI and 16s mtDNA

Solution	Initial concentration	COI volume(µl)	16s volume(µl)
dd H_20		31.3	40.8
10x PCR buffer	*see below	5	5
dNTP	5 mM	2.5	1
Primer 1	10 µM	2.5	1
Primer 2	10 µM	2.5	1
MgCl	25 mM	5	-
Taq	5 u/µl	0.2	0.2
Total volume		49.00	49.00

*PCR buffer: 100 mM Tris-HCl (pH8.3), 15 mM $MgCl_2$, 500 mM KCl.

The reaction mixtures and template were added to Perkin-Elmer thin walled 200 µl microtubes and placed into a Perkin-Elmer 9600 Thermocycler. Initial denaturation was at 94°C for 90 s. A

reaction cycle (94°C for 5 s, 45°C for 20 s, 72°C for 20 s) was then repeated 35 times with a final extension step of 72°C for 5 minutes. To examine the PCR products, 5 μl of each sample was run on a 1.4% TBE agarose gel for 1 hr at 80 volts, stained with 0.5 μg ml^{-1} of ethidium bromide.

4.2.2.1.2 12s

A small number of individuals were also tested for the variability of a portion of the mitochondrial 12s rDNA sequence. Procedures were very similar to the above, and once again, conserved primer pairs from Simon et al. (1991) were used. These were:

12sai (25mer) 5' - AAACTAGGATTAGATACCCTATTAT-3'

12sbi (20mer) 5' - AAGAGCGACGGGCGATGTGT-3'

4.2.2.1.3 Control Region

The non-coding mitochondrial control region has exhibited considerable useful genetic variation in vertebrate animals. Although not expected to provide the same degree of variation in invertebrates (largely due to its extreme A-T nucleotide bias in these animals), it was considered that its non-coding nature may provide a region of higher variation than those previously described. A large number of primer combinations from the literature, based on known sequences from a variety of invertebrates (primarily *Drosophila* and other insects, but also including *Artemia* and *Daphnia*) were tried in an attempt to amplify this region in prawns, but without success. It was assumed that either prawn sequences in these priming locations were sufficiently different to inhibit the binding of the universal primers used, or that in fact a rearrangement had occurred in the location of the tRNAs flanking this region. Without additional technical and financial assistance at the early stages, it was not possible to pursue the variation in the control region (CR). For the population analyses, it was decided to pursue instead the known (but relatively low levels of) nucleotide variation in the mitochondrial 16s, and (particularly) the COI regions.

At a later stage in the project, it was possible to further explore possible primers for amplifying this region, with the assistance of Dr Monique Monnerot (C.N.R.S., France) and Erik Garcia-Machado (University of Havana, Cuba) who kindly supplied an unpublished sequence of a portion of this region from the prawn *P. notialis*. Alignments of DNA sequences of the flanking regions of the control region were made. The CR is flanked at one end by the 12rRNA gene, and at the other by three tRNAs (Ile (I), Gln (Q) and Met (M)) and the ND2 gene. Sequences available for alignment were taken from *Drosophila yakuba* and *D. virescens*, *Artemia franciscana* and *Daphnia pulex*. This procedure confirmed not only that *P. notialis* had the same gene arrangement as *Drosophila* (and distinct from *Artemia*), but that there was sufficient homology among all these species in some flanking regions to suggest that good PCR primers for Australian species of *Penaeus* could be designed.

```
                          1 1111111112 2222222223 3333333334 4444444445
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       ---------- ---------- --------GC CAGGTGTGTT ATTAGGATCA
Daphnia pulex    .......... .......... ........-- ----AACTAA .CA.AAC.TT
Artemia fran.    .......... .......... .....AAA.. TCCTCTGT.C .CCTAA.AT.
Drosophila vir.  .......... .......... .......TAT A.TTG....A .CCGC...TG
Drosophila yak.  .......... .......... .........- --TTG..A.A .CCGC..CTG

                                                                       1
                 5555555556 6666666667 7777777778 8888888889 9999999990
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       TCAATTCCAA CTTACATTTA TTTAATTAGT ACCAG-GCAC TGAGACTTTA
Daphnia pulex    .ACTC.A.T. AAATTC..CG .C...CCGCG ..GGCT.GCA C..TTT.AGT
Artemia fran.    .TTGC..T.G GGCCTTG.AT GACCGCGGA. G.TG.CA..A A.T.GTCCCC
Drosophila vir.  CTGGCA.... T...ATC.AT AC.CT...A. .TTGCTATTT CT.A.T..CT
Drosophila yak.  CTGGCA.... T..GGTCAAT AC.TT...A. .TTGCTATTT CT.A.T..CT

                 1111111111 1111111111 1111111111 1111111111 1111111111
                 0000000001 1111111112 2222222223 3333333334 4444444445
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       CTATTCAACG GAGGCCCCTA CCTCTTTCTA ATTCATACTC CTC----TAG
Daphnia pulex    .A....T.T. -..TTTAACT AGG.CCAAAT ..--T.GT.T .CTAAAACTT
Artemia fran.    ....CTTTT- -C.TT.G... TAC..AAGA. G.AGGC.A.A A.G...TC.A
Drosophila vir.  T..ACT..TA A.ATTAAT.. .TG.GA.TAT .A.A.ATT.A A.A....-.T
Drosophila yak.  T..A.T..TA ATATTAAT.. .TG.GAATA. ..AAT.TA.A A.ATATT..T

                 1111111111 1111111111 1111111111 1111111111 1111111112
                 5555555556 6666666667 7777777778 8888888889 9999999990
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       CTTGTATA-A ATTACATTGG CTCATTAAAT TTTACATGTG CCATTAACCT
Daphnia pulex    AC.ACTACTT ...GTT..AT T.A.A.CCCC .G.TT..A.A TAGA.TTAAC
Artemia fran.    TCAC.GC..- C..CT.ACCT A.A..G.T.. C..G...A.. A-C.CGTAGA
Drosophila vir.  T..T..A.TT ....A..ATT .A..CA...A .......A.A A-..C..AT.
Drosophila yak.  T..T..A.T. .A..T.AATT .A..CA...A .......A.A A-..C..A..

                 2222222222 2222222222 2222222222 22222222
                 0000000001 1111111112 2222222223 33333333
                 1234567890 1234567890 1234567890 12345678

P.notialis       AATAATGAAA GAACCAGCTA GGATAAAACT TTTA----
Daphnia pulex    TCA..AT.G. .CCAA...C. .AT.C..... ..CCCTAA
Artemia fran.    ..G.G.A..T A...T..... .A..C..... ...-....
Drosophila vir.  .....CA..C TTTTA...C. AA........ ..AT....
Drosophila yak.  .....CA..T TTTTA...C. AA........ ..A.....
```

Figure 4.1 - Aligned sequences of last 200bp of 12s rDNA from 5 arthropods. Position of primer 12s-2 is underlined.

```
                          1 1111111112 2222222223 3333333334 4444444445
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       AATGAAGTGC CTGAAAAAG- -GATTATCTT GATAGGGTAG AACATGTAAG
Daphnia pulex    .G.TTGA... ....TT..A. G.GC..C... .......... .TG...C..C
Artemia fran.    G...GG.... ....T...AA G.....CG.. ....CT...- .TT.....GA
Drosophila vir. ......T... ....T...A. G.G....... .........A .TT.....G.
Drosophila yak. ......T... ....T...AA G.G...C... .........A .TT...C.GT

                                                                        1
                 5555555556 6666666667 7777777778 8888888889 9999999990
                 1234567890 1234567890 1234567890 1234567890 1234567890

P.notialis       CAAAACCTTA CCTTCATTA- ---------- ---------- ----------
Daphnia pulex    TTGTTG..C. AA..A----. .......... .......... ..........
Artemia fran.    T..T...CCC ATC.-----. .......... .......... ..........
Drosophila vir. ATTCTA.A.T .A..G----. .......... .......... ..........
Drosophila yak. TTTCTG.A.T .A..G.C.GA TTTATATATT ATTTAAAAAG AAGGTTTTAT

                 11111
                 00000
                 12345

P.notialis       -----
Daphnia pulex    .....
Artemia fran.    .....
Drosophila vir. .....
Drosophila yak. ATTTA
```

Figure 4.2 - Aligned sequences of Ile tRNA from 5 arthropods. Positions of primers Ile-2 and Ile-3 are underlined

Segments of DNA sequence of around 20 base pairs that were close to the CR and had highest homology among species were selected. Primer pairs were designed from these segments and compared using (in decreasing order of importance) the criteria of: maximum homology with the *P. notialis* sequence, the first five 3' bases anchored in the most conserved sites, weakest formation of primer dimers (both with self and with paired primer) and hairpin loops, and maximum and matching primer Tm's. Two sets of primers were selected in this way (with the assistance of the program OLIGO) (Figure 4.3) and tested on both target species. Both sets of primers worked successfully, with the 12s-2 / Ile-2 set ultimately used for sequencing a small number of individuals to test their utility.

```
12s-2          AAG AAC CAG CTA GGA TAA AAC TTT
Ile-2          ATC AAG ATA ATC CTT TTT CAG GCA
Ile-3          GCT TAC ATG TTC TAC CCT ATC AAG
```

Figure 4.3 - *Penaeid* control region primers (listed 5' - 3')

4.2.2.2 DNA sequencing

On examination of the above PCR reaction, if there were other background amplified fragments in addition to the expected one, the extraneous products were removed by gel separation. To accomplish this, firstly the DNA was precipitated with propanol and concentrated to 10 µl in TE.

The entire sample was then run on a 3 mm thick 1% TAE gel at 100 V. After ethidium bromide staining and photography each target band was excised with a clean scalpel blade and placed into a microcentrifuge tube with clean forceps. This was spun down to estimate the volume (~100 µl of gel in each tube).

Both the PCR product (if a single band) or the gel-excised band were purified from primers, dNTPs and buffer (and agarose) using either Gene-Clean (BIO-101) or QIAGEN quick-spin PCR purification columns, using their elution protocols. The purified DNA was resuspended in 20 µl of 0.1 x TE. 2 µl of sample was run on a 1.4% TBE agarose gel and the concentration of double-stranded PCR product was quantified by eye with reference to a known concentration sample.

Approximately 200 ng of PCR product was used as the template in a cycle-sequencing reaction with fluorescently labelled di-deoxynucleotides (using the ABI PRISM kit and protocols). Each cycle-sequencing reaction used one of the same primers as those in the initial amplification. After phenol/chloroform extraction to remove excess fluorescent nucleotides and ethanol precipitation, the single-stranded extension products were electrophoresed and analysed on an ABI 373A automated sequencer. Approximately 400-500 bases were routinely sequenced in each direction for mtDNA gene fragments in each individual.

4.2.2.3 RFLP markers

For routine screening of large numbers of individuals, a procedure of PCR amplification followed by digestion of the fragment by restriction enzymes and then gel electrophoresis (PCR-RFLP) was used. Once the location of polymorphic nucleotide sites was determined via DNA sequencing of a small number of randomly-chosen individuals, these sites were examined to see if they corresponded with any recognition sites for commercially-available restriction enzymes. In both species, only a proportion of the variable sites could be screened in this way, but the bulk of the species' haplotypes could be distinguished. In *P. plebejus*, one important divergent clade could not be distinguished using this technique, therefore a mismatch primer approach was used.

In the mismatch primer procedure, a new, long (at least 25bp) PCR primer is designed such that its 3' end anneals imperfectly to the template strand adjacent to the variable site. The primer sequence is altered close to the variable site so that a variable restriction enzyme recognition site is created. All the amplified fragments will now contain this site, and when cut by the restriction enzyme, a small fragment of about 20bp (from the 5' end of the primer) will be produced. Although this small fragment may not itself be visible on an agarose gel (it can be visualised on an acrylamide gel if necessary), the change in length of the adjoining fragment should be easily visible. To ensure this, high resolution agarose was used to allow discrimination of 20bp changes in length.

This procedure was used to screen variation at the *P. plebejus* site 158 using *Taq*I. Figure 4.4 shows the original template sequence, the mismatch primer used, and the resulting restriction site.

```
Template sequence A:
     5'..GGTATAATTTACGCAATGCTAGCTATTGGGGTTCT...3'

Mismatch primer:                    AATTTACGCAATGCTAGCTATTCG

Resulting PCR fragment:             AATTTACGCAATGCTAGCTATTCGGGTTCT...


Template sequence E:
     5'..GGTATAATTTACGCAATGCTAGCTATTGGAGTTCT...3'

Mismatch primer:                    AATTTACGCAATGCTAGCTATTCG

Resulting PCR fragment:             AATTTACGCAATGCTAGCTATTCGAGTTCT...
```

Figure 4.4 - Design of mismatch primer. Shows introduction of variable *Taq*I restriction site (TCGA - underlined in resulting PCR fragment for allele E only).

4.2.3 nDNA

4.2.3.1 microsatellites

4.2.3.1.1 Cloning and sequencing

All laboratory work undertaken in this sub-project was undertaken with the assistance and direction of Dr Steve Moore from the CSIRO Molecular Animal Genetics laboratory.

Cloning began with 15µg of high molecular weight DNA from a *P. esculentus* individual from Moreton Bay (Pe #16). The restriction enzyme *Sau*3AI was used to fragment the DNA (using 20 units of enzyme at 37°C O/N in 50µl), which was then electrophoresed through a 0.6% agarose gel alongside a size standard. All fragments of approximately 400bp in length (~350-450 bp) were cut from the gel and purified using Geneclean. These fragments were ligated to 25 ng of pUC18 plasmid DNA precut with *Bam*HI to produce compatible ends, using 2.5 units of T4 DNA ligase at 15°C overnight. Competent cells of *E. coli* were transformed with the ligated plasmids and grown overnight at 37°C.

Colonies were then transferred to nylon membranes and screened for the presence of CA dinucleotide repeats by hybridisation using a radioactively labelled CA oligonucleotide probe. Positive colonies from the initial screening were selected and again grown overnight for another round of radioactive hybridisation screening. Double positive colonies were selected and the recombinant plasmid DNA isolated and stored at -70°C.

The microsatellite-containing insert sequences were sequenced using standard forward and reverse sequencing primers in a flourescently labelled dye-terminator cycle-sequencing reaction (ABI PRISM kit). The extension products were run on an ABI 373 automated sequencer.

Sequences were aligned to the plasmid sequence, which was removed, and the microsatellites identified and characterised. Those clones containing suitable-length simple microsatellites

(between ~ 50 and 300 bp, with a simple repeat structure, and for which clear flanking sequences could be determined), were compared against known sequences in the Genbank database for identification of spurious DNA sequences. No strong matches were found.

4.2.3.1.2 PCR optimisation

PCR primer pairs were designed to flank the microsatellites, and to give a range of PCR product sizes to allow multiplexing of microsatellite loci during screening. Primers were designed (with the assistance of the program OLIGO) to maximise annealing specificity and minimise formation of primer dimers and hairpin loops. One primer of each pair was labelled with one of three flourescent dyes, the dye-primer combinations chosen so that amplification products of similar size had different dye labels. This meant that nearly all microsatellite loci could be easily discriminated by their size and label. Primer pairs were tested on a small number of individuals under a range of conditions in order to optimise the amplification reactions, while screening on simple ethidium-stained agarose gels.

4.2.3.1.3 Flourescent screening of variation

Once the PCR reactions were optimised, large-scale screenings were carried out on an ABI 373 DNA sequencer using GENESCAN software. All mutiplexed products were identified to their locus by size and colour, and the allele lengths determined automatically by calculation from an internal lane standard.

4.3 Analysis

4.3.1 Sequence alignment

Chromatogram traces and sequences from the ABI automated sequencer were firstly checked for obvious errors in base-calling and questionable base-calls due to low signal or high background noise. All sequences were trimmed to include only reliable sequence. Both forward and reverse sequences were obtained and aligned for all fragments so as to confirm the sequences and check that variable sites were indeed due to polymorphisms and not due to sequencing errors. All sequences of the same fragment from prawn species could be easily manually aligned in the ABI sequence editor SEQED. Where significant insertions and deletions (indels) were encountered in comparisons across many species, CLUSTALW was used for multiple alignment. The first sequence of each fragment of each species was aligned against known sequence from the most closely related species available to check for authenticity of the fragment. For coding sequence (COI), the amino acid sequence was also checked in order to authenticate the alignment and verify that base substitutions were feasible (i.e., stop codons were not introduced, substitutions were predominantly silent, etc.).

4.3.2 Sequence analyses

Sequences were exported into MEGA for analysis of basic sequence statistics (such as base content), calculation of intra- and inter-species sequence divergences and comparisons (including transition/transversion ratios, numbers of synonymous and non-synonymous changes, etc.), and creation of neighbour-joining trees. Parsimony analyses of relationships among haplotypes were

performed in PAUP. Measures of haplotype and sequence variability were calculated using ARLEQUIN.

4.3.3 Population analyses

4.3.3.1 mtDNA sequences

Sequence divergences within and between populations were calculated using REAP. ARLEQUIN was used to perform analysis of variance tests of non-random distribution of haplotypes both using and excluding sequence divergences (AMOVA). Standard chi-square permutation tests of non-random distributions of haplotypes were performed using CHIRXC and GENEPOP. Weir and Cockerham F-statistics were calculated and tested for significance using FSTAT. Pairwise estimates of migration rate (M-hat) were calculated from F_{ST} values in ARLEQUIN and FSTAT and tested for the existence of isolation by distance in GENEPOP. The patterns of population differentiation were displayed using both neighbour-joining trees constructed from net sequence divergences in PHYLIP, and multidimensional scaling analyses performed in NTSYS.

5. Results

5.1 Genetic variation detected in east coast prawns

5.1.1 mtDNA

5.1.1.1 12s rDNA

5.1.1.1.1 Sequences

Nucleotide sequences of a 410bp fragment of the mitochondrial 12s ribosomal RNA gene were determined from *P. esculentus*, and aligned with existing sequences from the most closely related species available, these being *P.vannamei, P. notialis, Cherax, Daphnia, Artemia* and *Drosophila*. These aligned sequences are shown in Figure 5.1.

```
                    1 1111111112 2222222223 3333333334 4444444445 5555555556
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus GTAGAT?CCC TATTAATACT TTATTTAAAT TATGAA--TA CCTGGGTACT AAACAGTTAT
P.vannamei   ---------- ---------- ---------- ------.--- ---------- ----------
P.notialis   T.....A... .G...-C... ....G..... ...A.CAT.. ........T. .......A..
Cherax       ---------- ---------- ---------- ------.--- ---------- ----------
Daphnia      AACC.GGATT AGA..CCCTG .....CT..A ..AT.GTG.. AA.A..GTAG T..A....G.
Artemia      A.G.CCAAA. A.C..GG.T. AG..ACCCTA .TA.TTCA.T T..AA-.TAA CCTGG..AG.
Drosophila   C...GATTAG ATACCC..T. A.T.AA..TG ..AAT.AA.T G..AAAGTAG T..T......


                                                        1 1111111111 1111111111
               6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus TATCTTGAAA CTCAAAGGAT TTGGCGGTAA CTTAGTCTAT TTAGAGGAAC CTGTCCTGTA
P.vannamei   ---------- ---------- ---------- ---------- ---------- ----------
P.notialis   .G........ .......... .......... T......... .......... ..........
Cherax       ---------- ----...A.. ........GT T......... .......... ..........
Daphnia      --....C... .C.....AT. .......C.T ....TC...A .-......G. T.....-A..
Artemia      AT.A..T... .CT....A.. ........TT ..A.ACT..C .-........ ...C..CA..
Drosophila   GT........ ..T...AA.. .........T T......... CC........ ....TT....


               1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
               2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus ATCGATAGTC CAC--GAAGA ATCTTACTCT ATCTTGTA-A ATACAGTTTG TATACCGTCA
P.vannamei   .......... .........A. ?.......T. ........G. .AT....... ..........
P.notialis   .......... ..ACGAG.A. ........T. G...G...G. .A........ ..........
Cherax       .......A.. ......TT.G .C......AA GGT....T.- -.T..A.... ....T..C.G
Daphnia      ..T....... .C...ATTTT .......C.. T.T....T.- ----.AGC.. .......C.G
Artemia      .A...C.A.. .......GAC .A.CG..C.G T.........- .A....C..A .........G
Drosophila   .......A.. .......T.G .C......TA .AT......- -AT......A .........G
```

```
                1111111111 1111111112 2222222222 2222222222 2222222222 2222222222
                8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus    TTATTAGGTA ACTTTTATAA AA-TTAGAGT TATTAAAATA GTTTAGCTAA T--ATATTAG
P.vannamei      .......A.. .....A.... ..T...A... .G..GG.... ..C....... ..........
P.notialis      .......A.. .....A.... ..TA..A... .G..GC..G. .......... ..........
Cherax          ....C..AAT .TC..A.A.G .GGAATA.T. .TC....T.T AA.A.TA... ATG..G.C..
Daphnia         ..G.C..-.T .....A.A.G ...---CGT. .TCAGG..A. A.A.TTT.CT ....CT.C..
Artemia         .--.....GG .T.AAATA.. T..---ACTG CTC.T..T.G C..AG..... ....A..C..
Drosophila      ....C..AAT .T...ATA.G ..TAATA.TA .TCA.T...T T.AATAAA.. .TT....C..


                2222222222 2222222222 2222222222 2222222222 2222222222 2222222223
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus    ATCAAGGTGC AGCTTATGGT AGAGAAAGAG ATGGGTTACA ATATATTAA- TATAGACGGA
P.vannamei      .......... ........A. .A........ .......... T....G...A ....A.....
P.notialis      .......... ........A. .A..G..... .......... ....GAC..A ....A.....
Cherax          G......... ..T....AT. TT..T.G-.. .......... ...A...T.. .T..T...TT
Daphnia         G......... ..T....AAA ..G.C.G-T. ...A.C.... ..T.TAAT.. G.A.T.....
Artemia         .........T ........TC C.G..G.-G. G......... ...AGAACT. --..A.....
Drosophila      .........T .......AT. TA..T..-TA .......... ...A...T.. .T..A.....


                3333333333 3333333333 3333333333 3333333333 3333333333 3333333333
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus    ACAATAAATG AATATTTGTT GATGAAGGTG GATTTAAGAG TAAAGTGATT TTAATAAGGT
P.vannamei      ....C..... ..GG...... .T........ .......A.. GT..TA---- ----------
P.notialis      GT.....G.. ....C..... .C........ .......A.. ...GAA.... ........A.
Cherax          .A.T.GTT.. C.ATAA.... --.T...... .......A.. ....A.A... AAT.A..-T.
Daphnia         TTTTAGCT.A ..AC.CA.G. ---........ ......TT.. ....T.T.C. .A.G..CA..
Artemia         ..TT.C.T.T T..CA.G.AG A-A....... ......GA.. ...G..CTA. CA--------
Drosophila      TA..ATT... ..A.AA.T.. --........ .....GGT.. ....A.T..A AAG..T.-A.


                3333333333 3333333333
                6666666667 7777777778
                1234567890 1234567890
P.esculentus    TACTTGATTT TAGCTCTAGG
P.vannamei      ---------- ----------
P.notialis      CTT....C.C ..........
Cherax          ..T......A ........AA
Daphnia         A.G....A.A AG..AA....
Artemia         -..C...AGG AG.....GA.
Drosophila      A.T....... ........AA
```

Figure 5.1 - Sequence alignment of mt 12s rDNA among *Penaeus* spp. and other arthropods.

There is considerable divergence between the three *Penaeus* sequences, with a (Jukes-Cantor) corrected nucleotide sequence divergence of 12.2% between *P. esculentus* and *P. vannamei*, and 14.3 % between *P. esculentus* and *P. notialis* (Table 5.1). Five small indels were observed between *Penaeus* species, and relatively few indels observed in comparison to the outgroup species. All *Penaeus* sequences are relatively AT-rich (*P. esculentus* 70.3%, *P. notialis* 71.5%), compared with 73.4% for *Cherax*, 66.6% for *Daphnia, 61.3%* for *Artemia* and 76.7% for *Drosophila.* The

transition / tranversion ratio was 1.3 - 2.2 between *Penaeus* spp. but only 0.5 - 0.7 between *Penaeus* and other species.

Table 5.1 - 12s rDNA nucleotide divergences (& SEs) between species

	OTUs	1	2	3	4	5	6	7
1	P.esculentus		0.1220	0.1429	0.3613	0.5349	0.5618	0.4343
2	P.vannamei	0.0250		0.1308	0.4775	0.5295	0.5246	0.3801
3	P.notialis	0.0212	0.0258		0.4108	0.5656	0.6195	0.4692
4	Cherax	0.0428	0.0612	0.0469		0.4778	0.5503	0.2503
5	Daphnia	0.0528	0.0682	0.0552	0.0537		0.6663	0.5899
6	Artemia	0.0554	0.0670	0.0602	0.0610	0.0653		0.5778
7	Drosophila	0.0440	0.0514	0.0465	0.0332	0.0573	0.0566	

(Distances in the upper-right matrix. Standard Errors in lower-left matrix)

5.1.1.1.1.1 P. esculentus

Sequences of the 12s region were obtained from 3 individuals from *P. esculentus* from 2 locations spanning the species' entire distribution. The variability within the species is shown in Figure 5.2, where all the haplotypes found are listed.

	2 3 3
	2 9 1 4
	7 6 1 7
A	A T G T
B	G . . .
C	. G A G

Figure 5.2 - *P. esculentus* 12s haplotypes and polymorphisms.

Only 4 sites (1.1 %) proved to be polymorphic. Diversity in this segment is relatively low, the average nucleotide diversity being 0.7%, with the maximum pairwise divergence being only 1.1%. 50% of nucleotide substitutions were transitions, giving a transition/transversion ratio of 1.0. No indels were observed in this segment within the species. The haplotype diversity was high among the three individuals examined (1.0).

A maximum parsimony network of the phylogenetic relationships among haplotypes was constructed. Figure 5.3 shows these relationships. Haplotypes A and B were found in individuals collected from Moreton Bay, while haplotype C was found in a Western Australian individual.

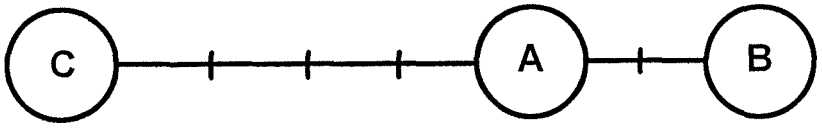


Figure 5.3. Network of *P. esculentus* 12s haplotypes.

5.1.1.2 16s rDNA

5.1.1.2.1 Sequences

Nucleotide sequences of a 560bp fragment of the mitochondrial 16s ribosomal RNA gene were determined from both *P. plebejus* and *P. esculentus*, and aligned with existing sequences from the most closely related species available, these being a partial sequence of *Daphnia pulex* and *Artemia franciscana*, as well as *Drosophila yakuba*.. These aligned sequences are shown in Figure 5.4.

```
                    1 1111111112 2222222223 3333333334 4444444445 5555555556
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus   TTATATAAAG TCTAGCCTGC CCACTGATCA AATTTTAAAG GGCCGCGGTA TACTGACCGT
P.plebejus     .C........ .......... ........TT G-........ .......... ..T.......
Daphnia pulex  ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.  ---.GGG.G. ..G....... ........G. .---...... .....T.... ........A.
Drosophila y.  A......... ....A..... .......--. .........T ......A... .TT....T..

                                                          1 1111111111 1111111111
               6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus   GCGAAGGTAG CATAATCATT AGTCTTTTAA TTGAAGGCTT GTATGAATGG TTGGACAAAA
P.plebejus     .......... .......... .......... ...G.....C .......... ..........
Daphnia pulex  ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.  .......... .......... ..C.....G. ..TG.....G .A........ ..T...G.G.
Drosophila y.  ..A....... .......... .......... .........G .A........ ......G...

               1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
               2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus   AGTAAGCTGT CT-CAGTTAT AAT---AATT GAATTTAACT TTTGAGTGAA AAGGCTTAAA
P.plebejus     .......... ....G.C... .......... .......... .......... ......C...
Daphnia pulex  ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.  GA.GGT.... ...--C..CG .T...A.... ...G....TC ...A...... ..A.......
Drosophila y.  TA.T.A.... T....T...A .T.TAA...A ......T.T. ...T...C.. ..A...A...

               1111111111 1111111112 2222222222 2222222222 2222222222 2222222222
               8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus   TAAATTGAGG GGACGATAAG ACCCTATAAA GCTTGACAGT AAGTTGATTA TATTATAAAT
P.plebejus     .GGC.CAG.. .......... .......... ........A. ..C..TG... A.........
Daphnia pulex  ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.  .GT.C.TG.A ..G....... ........G. T...T...T. T.A..CT..T GTC.TGCGG.
Drosophila y.  .T....T.AA A.....G... .......... T...T.T.T. TTA..T...T ..A.TAT..A
```

```
                2222222222 2222222222 2222222222 2222222222 2222222222 2222222223
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    TGTTAGTATA ACTTGATTTT AGTTAGCGTC TGTTACGTTG GGGCGACGAG AATATAATAA
P.plebejus      .......... .......... .AC...G..T ....T..... .......... ..........
Daphnia pulex   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.   A.G..A.-.. GACA..G.AA .ACA.----- ....CG.... ........GT ..G.AC.G..
Drosophila y.   GA...A.T.. .T..T.A.A- .A...AAA-- .A..TTA... ...T..TATT ..A..T-...

                3333333333 3333333333 3333333333 3333333333 3333333333 3333333333
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    GTAACTGTTC TTAAGTATCT AATAACAACT ATAGTTGGTA AATAAATGAT CCTCTATTAG
P.plebejus      A......... ..TTAA..A. ...T....AA ..GC.....G G..G.T.... ...T.....A
Daphnia pulex   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Artemia fran.   TA...AC..- ---.CA.CA. ..AC...T.A ...AA.---- G.CC.T.... ...TAGA.GA
Drosophila y.   AA....T..- ----AAT.T. ..--.A...A T..A..TA.G .....T.... ..AT..A..A

                3333333333 3333333333 3333333333 3333333334 4444444444 4444444444
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    GGATTAGAAG ATTAAGTTAC TTTAGGGATA ACAGCGTAAT CTTCTTTGAG AGTCCATATC
P.plebejus      A.....A... ..A....... .......... .......... .......... ..........
Daphnia pulex   ?????????? ?????????? .A.......G G.T.-A.... T.CT..---- .ACT.G..CT
Artemia fran.   AT.----... .CC....... C......... .......... TC.T...... ...T..A...
Drosophila y.   T.....A..A .......... .......... .......... T..T..G... ...T......

                4444444444 4444444444 4444444444 4444444444 4444444444 4444444444
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    GACAAGAAGG TTTGCGA-CC TCGATGTTGA ATTAAGGTAT CCTTATGA-T GCAGCAGTTA
P.plebejus      .......... .......... .......... .......... ......A... ..........
Daphnia pulex   T...CA.--- -..TG..CAA AGATG..CTG T.A..AA... TT.CCCC.AG .T.AA.A???
Artemia fran.   .....A.GA. .......G.. .........G -..C...G.C ...AC.CGG. .......CCT
Drosophila y.   ..T..A..A. A......... .........G ......A... AA..T..GG. .T...C...C

                4444444444 4444444445 555555555
                8888888889 9999999990 000000000
                1234567890 1234567890 123456789

P.esculentus    TAAAGGAAGG TCTGTTCGAC CTTTAAACC
P.plebejus      .......G.. .....C.... ..CC...T.
Daphnia pulex   ?????????? ?????????? ?????????
Artemia fran.   AG.GA.GCA. .......... TG.......
Drosophila y.   A..TTTT.A. .......... T......TT
```

Figure 5.4 - Sequence alignment of mt 16s rDNA among *Penaeus* spp. and other arthropods.

There is considerable divergence between the two *Penaeus* sequences, with a (Jukes-Cantor) corrected nucleotide sequence divergence of 11.2% (Table 5.2). Five indels were observed between *Penaeus* species, and relatively few indels observed in comparison to the outgroup species. Both sequences are relatively AT-rich (*P. esculentus* 65.4%, *P. plebejus* 64.4%), compared with 58.9%

for *Artemia*, and 76.8% for *Drosophila*. The transition/transversion ratio was 2.1 between *Penaeus* species, but only approximately 0.8 between *Penaeus* and *Drosophila*.

Table 5.2 - 16s nucleotide divergences (& SEs) between species.

(Distances in the upper-right matrix. Standard Errors in lower-left matrix)

	OTUs	1	2	3	4	5
1	P.esculentus		0.1115	0.8597	0.3496	0.2874
2	P.plebejus	0.0158		0.8597	0.3529	0.3157
3	Daphnia pulex	0.1696	0.1696		0.8776	0.7424
4	Artemia fran.	0.0328	0.0330	0.1726		0.3509
5	Drosophila y.	0.0283	0.0301	0.1440	0.0330	

5.1.1.2.1.1 P. plebejus

Sequences of the 16s region were obtained from 10 individuals from *P. plebejus* from 3 locations spanning the species' entire distribution. The variability within the species is shown if Figure 5.5, where all the haplotypes found are listed.

Haplotype	1 1 3 4 7 7 3 1 2 9 6 2
A	C T G T
B	. C . .
C	. . A .
D	T . . C

Figure 5.5 - *P. plebejus* 16s haplotypes and polymorphisms.

Only 4 sites (0.8 %) proved to be polymorphic. Diversity in this segment is relatively low, the average nucleotide diversity being 0.18% (+/- 0.00014%), with the maximum pairwise divergence being only 0.6%. All nucleotide substitutions were transitions. No indels were observed within this segment.

A maximum parsimony network of the phylogenetic relationships among haplotypes was constructed. Figure 5.6 shows these relationships, and the relative frequency of each of the haplotypes.

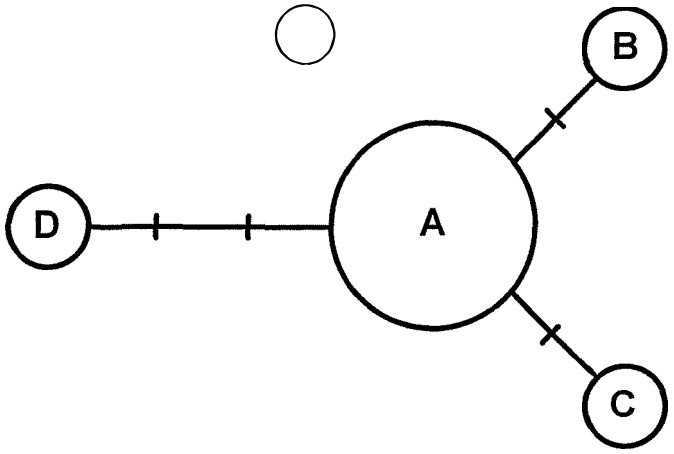


Figure 5.6 - Network of *P. plebejus* 16s haplotypes. (The size of each circle indicates the relative frequency of each haplotype.)

5.1.1.2.1.2 P esculentus

Sequences of the 16s region were obtained from 31 individuals from *P esculentus* from 6 locations spanning the species' entire distribution. The variability within the species is shown if Figure 5.7, where all the haplotypes found are listed.

Haplotype	1 1 1 2 2
	3 4 6 2 5
	3 0 8 0 3
A	G T A T C
B	A . . . .
C	. . . C .
D	. . . . T
E	. C . . .
F	. . G . .

Figure 5.7 - *P. esculentus* 16s haplotypes and polymorphisms.

Only five sites (1.0%) proved to be polymorphic. Diversity in this segment is again relatively low, the average nucleotide diversity being 0.18 %, with the maximum pairwise divergence being only 0.4%. All nucleotide substitutions were transitions. No indels were observed in this segment within *P. esculentus*. Average haplotype diversity was 0.63 (+/- 0.020).

A maximum parsimony network of the phylogenetic relationships among haplotypes was constructed. Figure 5.8 shows these relationships, and the relative frequency of each of the haplotypes.

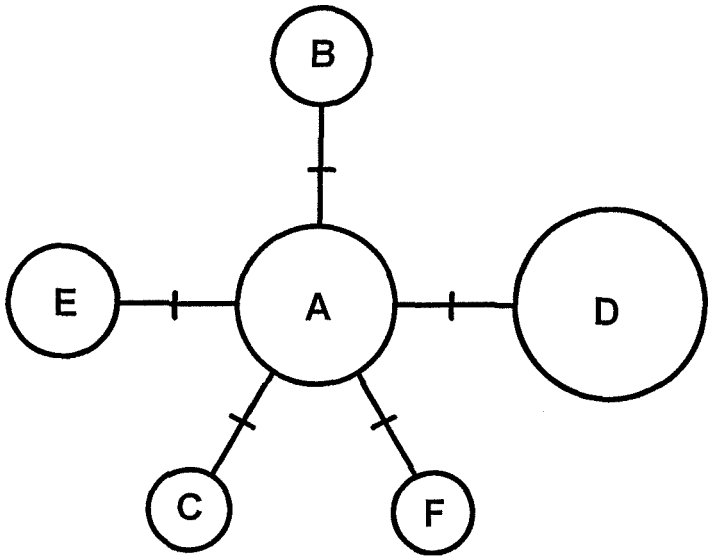


Figure 5.8 - Network of *P. esculentus* 16s haplotypes.

5.1.1.3 COI

5.1.1.3.1 Sequences

Nucleotide sequences of a 647bp fragment of the mitochondrial cytochrome oxidase subunit I (COI) gene were determined from both *P. plebejus* and *P. esculentus*, and aligned with existing sequences from the most closely related species available, these being *Daphnia pulex* and *Artemia franciscana*, as well as *Drosophila yakuba*. These aligned sequences are shown in Figure 5.9, and their translated amino acid sequences are shown in Figure 5.10.

```
                    111 111 111 122 222 222 223 333 333 333 444 444 444
              123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678
P.esculentus  TTG TAC CAG CAT TTA TTC TGA TTT TTT GGT CAC CCA GAA GTC TAT ATC
P.plebejus    GAT .GA ..A ..C ... ..T ..G ... ..C ..G ... ..G ... ..A ... ..T
Daphnia pulex ..A ... ... ... ..G ..T ..G ..C ... ... ..T ..T ... ..A ..C ..T
Artemia fran. C.T ..T ..A ... ... ..T ... ... ... ..C ..T ..T ... ..G ... ..T
Drosophila y. ... ... ..A ... ... ..T ... ... ... ... ... ..T ... ..T ... ..T

              455 555 555 556 666 666 666 777 777 777 788 888 888 889 999 999
              901 234 567 890 123 456 789 012 345 678 901 234 567 890 123 456
P.esculentus  TTG ATC TTA CCT GCC TTC GGA ATA ATT TCA CAT ATT ATT AGA CAA GAA
P.plebejus    ..A ..T A.. ..C ..T ..G ..T ... ..C ..C ..C ..._ ... ..T ... ..G
Daphnia pulex ..A ..T ... ..C .GG ..T ..T ... ... ..T ... ... ... ..C ..C ...
Artemia fran. ..A ..T ... ... .GA ... ..G ... G.G ..C ..C ... ... ..C ... ...
Drosophila y. ..A ..T ... ..G .GA ..T ... ... ... ..T ... ... ... ... ... ...

                  111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
              999 000 000 000 011 111 111 112 222 222 222 333 333 333 344 444
              789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234
P.esculentus  TCT GGT AAA AAA GAA GCA TTT GGA ACA CTA GGA ATA ATT TAT GCT ATA
P.plebejus    ..A ... ..._ ... ..G ..C ..C ..T ..T T.. ..T ... ... ..C ..A ..G
Daphnia pulex AGA ... ... ..G ... ... ..C ..C ..G ..T ..G ... ... ..C ... ...
Artemia fran. AGA ... ..G ..G ... ... ... ..T ... T.. ..T ..G ... ... ... ...
Drosophila y. ... ... ... ..G ... A.T ..C ..T T.T T.. ... ... ..C ... ... ...
```

```
                111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
                444 445 555 555 555 666 666 666 677 777 777 778 888 888 888 999
                567 890 123 456 789 012 345 678 901 234 567 890 123 456 789 012
P.esculentus    TTA GCA ATT GGA GTA TTA GGA TTT GTA GTA TGG GCT CAC CAC ATG TTT
P.plebejus      C.. ..T ... ... ..T C.. ... ... ... ..C ..A ... ..T ..T ..A ..C
Daphnia pulex   C.. ..T ..C ..G ..G ... ..C ..C ... ... ... ..C ... ..T ..A ...
Artemia fran.   C.T ... ... ..T A.T C.T ..T ... ..T ..C ... ... ..T ..T ..A ...
Drosophila y.   C.T ..T ... ... T.. ... ... ... A.T ..T ..A ... ..T ..T ..A ...

                111 111 122 222 222 222 222 222 222 222 222 222 222 222 222 222
                999 999 900 000 000 001 111 111 111 222 222 222 233 333 333 334
                345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
P.esculentus    ACT GTA GGT ATA GAC GTT GAT ACT CGT GCC TAC TTT ACA TCT GCT ACA
P.plebejus      ..A ... ..A ..G ... ... ..C ..G ..G ..T ... ... ..G ..A ..G ...
Daphnia pulex   ... ..G ..C ..G ..T ... ..C ..A ..A ..G ..T ... ..T G.A ... ...
Artemia fran.   ... ..G ..A ..G ... ..G ..C ... ..A ..T ..T ... ..T G.. ..A ..T
Drosophila y.   ..A ..T ..A ... ... ... ... ..A ..A ..T ..T ... ..T ... ... ..T

                222 222 222 222 222 222 222 222 222 222 222 222 222 222 222 222
                444 444 444 555 555 555 566 666 666 667 777 777 777 888 888 888
                123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678
P.esculentus    ATG ATT ATT GCT GTA CCT ACC GGT ATC AAG ATT TTC AGT TGA TTA GGG
P.plebejus      ... ... ... ... ..G ... ..A ..A ..T ..A ..C ..T ..C ... C.T ..A
Daphnia pulex   ... ... ... ..G ..T ... ..A ..A ..T ... ..C ... ..A ... C.T ..C
Artemia fran.   ... ... ... ..A A.T ... ..G ..G ..T ... ... ..T ..A ... A.C ..A
Drosophila y.   ..A ... ... ..G ..T ... ..A ..A ..T ..A ... ..T ..A ... ... .CT

                222 222 222 223 333 333 333 333 333 333 333 333 333 333 333 333
                899 999 999 990 000 000 000 111 111 111 122 222 222 223 333 333
                901 234 567 890 123 456 789 012 345 678 901 234 567 890 123 456
P.esculentus    ACT CTT CAT GGA ACT CAA TTA AAC TAC AGC CCT TCT CTC ATT TGA GCC
P.plebejus      ... ... ... ... ... ... ..G ..T ..T ..T ... ..A T.A ..C ..G ..T
Daphnia pulex   ..A ... ..C ... ..C ..G C.T GTA .TT .C. ... ... ... T.G ... ..G
Artemia fran.   ... T.A ..C ..G ..A .G. C.T .CT ATA .CT ..A ... A.G T.G ..G ...
Drosophila y.   ... T.A ... ... ... ... C.T TCT ..T TCT ..A G.. A.T T.A ... ..T

                333 333 333 333 333 333 333 333 333 333 333 333 333 333 333 333
                333 444 444 444 455 555 555 556 666 666 666 777 777 777 788 888
                789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234
P.esculentus    CTA GGT TTC GTA TTT TTA TTT ACA GTA GGG GGT CTA ACA GGG GTA GTG
P.plebejus      T.. ..A ... ..C ... C.. ... ..T ... ..A ..A T.. ... ... ..T ..C
Daphnia pulex   GC. ..G ..T ..G ... C.G ... ... ... ... ..G ..T ..T ... ... ..T
Artemia fran.   ... ..A ..T ... ... ..G ... ..T ..G ..A ..G ..T ... ..A ..T ..C
Drosophila y.   T.. ..A ..T ..T ... ... ..C ... ... ..A ..A T.. ... ..A ..T ..A

                333 333 333 333 333 444 444 444 444 444 444 444 444 444 444 444
                888 889 999 999 999 000 000 000 011 111 111 112 222 222 222 333
                567 890 123 456 789 012 345 678 901 234 567 890 123 456 789 012
P.esculentus    CTA GCC AAT TCA TCA ATC GAT ATT ATC TTA CAT GAT ACC TAT TAT GTT
P.plebejus      ... ..A ... ..T ... ..T ..C ..C ... C.C ..C ... ..T ... ... ..A
Daphnia pulex   T.. ... ..C ... AG. ..T ..C ... ..T C.T ... ..C ..G ..C ... ..C
Artemia fran.   ..G T.G ..C ..T AG. ..T ... ... G.T C.C ... ... ..T ... ... ..A
Drosophila y.   T.. ..T ... ... ... G.T ... ... ..T ... ... ... ..T ... ... ..A
```

```
                444 444 444 444 444 444 444 444 444 444 444 444 444 444 444 444 444
                333 333 344 444 444 445 555 555 555 666 666 666 677 777 777 778
                345 678 901 234 567 890 123 456 789 012 345 678 901 234 567 890
P.esculentus    GTA GCT CAT TTC CAC TAC GTC CTT TCA ATA GGA GCC GTA TTT GGT ATT
P.plebejus      ... ... ..C ..T ..T ..T ..T ..C ... ... ... ..T ... ... ..A ..C
Daphnia pulex   ... ... ... ..T ... ... ..T T.A ..T ... ..G ..A ... ..C .C. ..C
Artemia fran.   ..T ..C ... ... ... ..T ..T T.A ... ..G ..T ..T ..C ... .C. ..C
Drosophila y.   ... ... ... ... ... ... ..T T.A ... ... ... ..T ... ... .C. ...

                444 444 444 444 444 444 455 555 555 555 555 555 555 555 555 555 555
                888 888 888 999 999 999 900 000 000 001 111 111 111 222 222 222
                123 456 789 012 345 678 901 234 567 890 123 456 789 012 345 678
P.esculentus    TTT GCT GGT ATT GCT CAT TGA TTC CCC TTA TTT ACT GGC CTT ACT TTA
P.plebejus      ..C ..G ..A ... ... ..C ..G ... ..T C.T ... ... ..T ... ..A ...
Daphnia pulex   ... ..A ..A G.A ..C ... ... ... ..T C.T ... ..C ..G T.A ..C ..G
Artemia fran.   A.A ..A ..A T.. .TA ... ... ..T ..T C.. A.G ..C ..A T.A .GA A.G
Drosophila y.   A.A ..A ... T.. AT. ..C ... .A. ..A ... ... ... ..A T.G ..A ...

                555 555 555 555 555 555 555 555 555 555 555 555 555 555 555 555 555
                233 333 333 334 444 444 444 555 555 555 566 666 666 667 777 777
                901 234 567 890 123 456 789 012 345 678 901 234 567 890 123 456
P.esculentus    AAC CCG AAA TGA CTA AAA ATC CAC TTT TTA GTA ATG TTT ATT GGG GTA
P.plebejus      ..T ..T ... ... ..C ... ... ... ..C C.. ..T ..A ..C G.. ..A ...
Daphnia pulex   C.A G.C CGT ..G T.. ..G ..T ... ..C ..T AC. ..A ..C G.A ..A ...
Artemia fran.   ..T .A. TTT CTT ..T ..G G.A ..T ..C ..T A.T ..A ... T.A ..A ...
Drosophila y.   ..T AAT ... ..G T.. ... .GT ..A ... A.T A.T ... ... ... ..A ...

                555 555 555 555 555 555 555 556 666 666 666 666 666 666 666 666 666
                777 888 888 888 899 999 999 990 000 000 000 111 111 111 122 222
                789 012 345 678 901 234 567 890 123 456 789 012 345 678 901 234
P.esculentus    AAT ATT ACT TTC TTC CCT CAA CAT TTC CTA GGA CTT AAT GGT ATA CCA
P.plebejus      ..C ... ..A ..T ... ..C .T. ..C ..T T.. ..T ..G ... ..A ... ..T
Daphnia pulex   ... T.G ..C ..T ..T ..G ... ... ... T.. ... T.A GCA ... ... ..C
Artemia fran.   ..C T.A ... ..T ... ... ... ... ... T.. ..T ..C GC. ..A ... ...
Drosophila y.   ... T.A ..A ..T ... ..C ... ... ..T T.. ... T.A GCA ..A ... ..T

                666 666 666
                222 223 333
                567 890 123
P.esculentus    CGA CGT TAC
P.plebejus      ..G ..A ...
Daphnia pulex   ..G ... ...
Artemia fran.   ..T ... ...
Drosophila y.   ... ... ..T
```

Figure 5.9 - Aligned COI nucleotide sequences from *Penaeus* spp. and three other arthropods. (Primer locations are underlined)

```
                      1 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    LYQHLFWFFG HPEVYILILP AFGMISHIIS QESGKKEAFG TLGMIYAMLA IGVLGFVVWA
P.plebejus      DW........ ........M. .L........ .......... .......... ..........
Daphnia pulex   .......... .......... G......... H......... .......... ..........
Artemia fran.   .......... .......... G...V..... .......... .......... ..I.......
Drosophila y.   .......... .......... G......... .......T.. S......... ..L...I...
```

```
                                                          1 1111111111 1111111111
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    HHMFTVGMDV DTRAYFTSAT MIIAVPTGIK IFSWLGTLHG TQLNYSPSLI WALGFVFLFT
P.plebejus      .......... .......... .......... .......... .......... ..........
Daphnia pulex   .......... .......A.. .......... .......... ...VFT...L ..A.......
Artemia fran.   .......... .......A.. ....I..... ....I..... .R.TMT..ML ..........
Drosophila y.   .......... .......... .......... .....A.... ...S...AIL ..........

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

P.esculentus    VGGLTGVVLA NSSIDIILHD TYYVVAHFHY VLSMGAVFGI FAGIAHWFPL FTGLTLNPKW
P.plebejus      .......... .......... .......... .......... .......... ..........
Daphnia pulex   .......... .......... .......... ........A. ...V...... ......QAR.
Artemia fran.   .........S ......V... .......... ........A. M..FV..... M...SM.QFL
Drosophila y.   .......... ...V...... .......... ........A. M..FI..Y.. .......N..

                1111111111 1111111112 2222222222 2
                8888888889 9999999990 0000000001 1
                1234567890 1234567890 1234567890 1

P.esculentus    LKIHFLVMFI GVNITFFPQH FLGLNGMPRR Y
P.plebejus      .........V ........L. .......... .
Daphnia pulex   .....FT..V ...L...... ....A..... .
Artemia fran.   ..V..FI..L ...L...... ....A..... .
Drosophila y.   ..SQ.II... ...L...... ....A..... .
```

Figure 5.10 - Aligned COI amino acid sequences from Penaeus spp. and three other arthropods.

There is considerable divergence between the two *Penaeus* sequences, with a Jukes-Cantor corrected nucleotide sequence divergence of 26%, compared to a 34% divergence from *Artemia* (Table 5.3). 90% of the nucleotide changes between Penaeus species are synonymous, compared with only 62% for comparisons with *Artemia*. 87% of the substitutions between *Penaeus* species occur in the third codon position, 11% in the first, and 2% in the second. The transition / transversion ratio between *Penaeus* species is 1.5, but is only 0.9 between *Penaeus* and *Drosophila*. There is 2.8% amino acid sequence divergence between *Penaeus* species, compared to 14% between *Penaeus* and *Artemia* (Table 5.4). Both Penaeus sequences are relatively AT-rich (63%), compared with 58% for *Daphnia*, and 71% for *Drosophila*.

Table 5.3 - COI nucleotide divergences between Penaeus spp. and 3 other arthropods.

OTUs		1	2	3	4	5
1..	P.esculentus		0.2643	0.3059	0.3449	0.2509
2..	P.plebejus			0.3349	0.3251	0.2756
3..	Daphnia pulex				0.2988	0.2941
4..	Artemia fran.					0.2688
5..	Drosophila y.					

Table 5.4 - COI amino acid divergences between Penaeus spp. and 3 other arthropods.

OTUs		1	2	3	4	5
1..	P.esculentus		0.0284	0.0853	0.1422	0.1090
2..	P.plebejus			0.1043	0.1659	0.1374
3..	Daphnia pulex				0.1185	0.1232
4..	Artemia fran.					0.1422
5..	Drosophila y.					

Internal *Penaeus*-specific primers were designed to allow more efficient amplification of this gene fragment in these species. Primers PI and P2 were designed for *P. plebejus* and primers P3 and P4 for *P. esculentus*. See Figure 5.9 for the location of these primers.

5.1.1.3.1.1 P. plebejus

Sequences of the COI region were obtained from 34 individuals from *P. plebejus* from 8 locations spanning the species' entire distribution. The variability within the species is shown in Figure 5.11, where all the haplotypes found are listed. For most individuals, sequences were obtained from the 525 bp fragment amplified using *P. plebejus* -specific internal primers P1 and P2.

```
          122 33
   112 679 012 24
   282 953 538 12

A  TAC CCA GTA GC
B  ... ..G ... A.
C  ... ... ..G ..
D  ... ... ... .T
E  .G. ... ... ..
F  ... ... ... A.
G  ... ... .C. ..
H  .GT ... ... ..
K  CG. ... ... ..
I  .G. T.. ... ..
J  ... .T. A.. ..
L  ... ... A.. ..
```

Figure 5.11 - *P. plebejus* COI haplotypes and their polymorphic nucleotide sites.

11 sites (2.6%) proved to be polymorphic. Diversity in this segment is higher than that seen in the 16s fragment, the average nucleotide diversity being 0.28% (+/- 0.00003%), and the overall haplotype diversity 0.7411 (+/- 0.0038), with the maximum pairwise divergence being 1.0%. All nucleotide substitutions were transitions and were synonymous, with nearly all (93.8%) of all substitutions occurring in the third base position, only 1 (6.3%) in the first position, and none in the second.

A maximum parsimony network of the phylogenetic relationships among haplotypes was constructed. Figure 5.12 shows these relationships, and the relative frequency of each of the haplotypes.

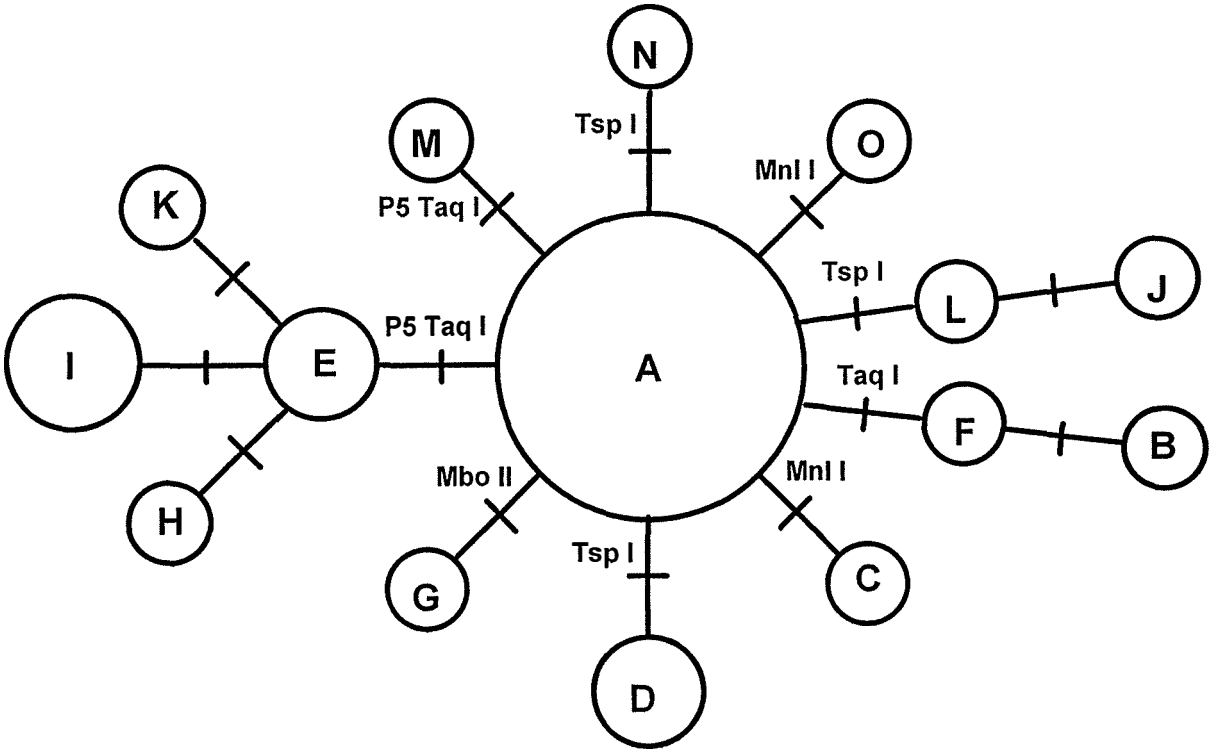


Figure 5.12 - Network of *P. plebejus* COI haplotypes.

Size of circles indicate relative frequency of haplotypes. Polymorphic sites screened for PCR-RFLP variation are labelled with the restriction enzyme used. "P5 Taq I" indicates this site was screened using mismatch RFLP analysis after amplification with the primer pair P5-P2, instead of the usual P1-P2.

5.1.1.3.1.2 P esculentus

Sequences of the COI region were obtained from 35 individuals from *P esculentus* from 6 locations spanning the species' entire distribution. The variability within the species is shown in Figure 5.13 and Figure 5.14, where all the haplotypes found are listed. For most individuals, sequences were obtained from the 525 bp fragment amplified using *P. esculentus*-specific internal primers P3 and P4.

```
     1113444
   583889346
   715093912

B  AGGGTAACT
K  ...A.G...
A  .....G...
E  ....CG...
F  GA...G...
C  .....GGA.
D  ..A..GGA.
I  ...A.....
G  .A...G...
H  .....G..A
J  .....G..C
```

Figure 5.13 - *P. esculentus* COI haplotypes and their polymorphic nucleotide sites.

```
   11
   45
   74

B  II
K  ..
A  ..
E  ..
F  ..
C  V.
D  V.
I  ..
G  ..
H  .M
J  ..
```

Figure 5.14 - *P. esculentus* COI haplotypes and their polymorphic amino acid sites.

9 sites (1.9%) proved to be polymorphic. Diversity in this segment is again higher than that seen in the 16s fragment, the average nucleotide diversity being 0.18% (+/- 0.00002%), and overall haplotype diversity 0.60 (+/- 0.008), with the maximum pairwise divergence being 1.3%. 80% of nucleotide substitutions were transitions, giving a transition/transversion ratio of 4.0. 82% of substitutions were synonymous, with 91% of all substitutions occurring in the third base position, 9% in the first position, and none in the second.

A maximum parsimony network of the phylogenetic relationships among haplotypes was constructed. Figure 5.15 shows these relationships, and the relative frequency of each of the haplotypes.

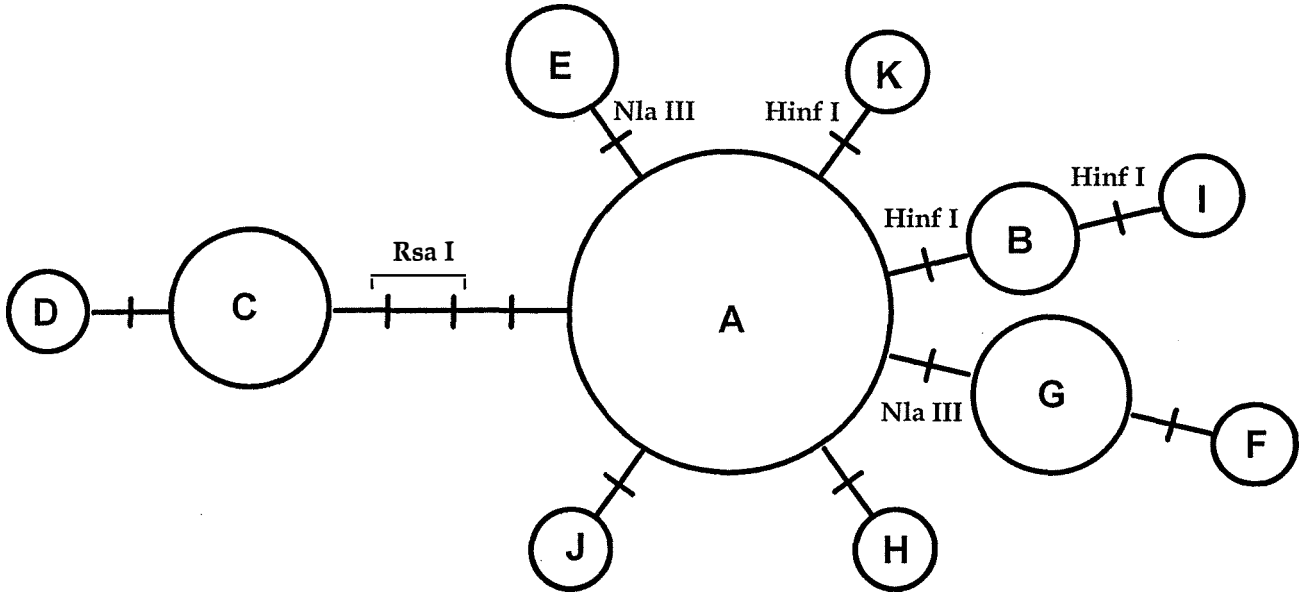


Figure 5.15 - Network of *P. esculentus* COI haplotypes.

Size of circles indicate relative frequency of haplotypes. Polymorphic sites screened for PCR-RFLP variation are labelled with the restriction enzyme used.

5.1.1.3.2 RFLPs

5.1.1.3.2.1 P. plebejus

Figure 5.12 indicates which variable base sites were screened for variation using restriction enzymes. The locations of restriction sites along the COI fragment are indicated in Table 5.5, as are details of the patterns of restriction fragments that are formed by the different haplotypes for each variable restriction enzyme. This table also lists the composite restriction enzyme patterns for each haplotype, and the haplotype groups which can be distinguished using PCR-RFLP only.

Table 5.5. *P. plebejus* restriction fragment patterns and overall haplotypes.

For each enzyme, for each restriction site, a '+' indicates a cut. . "P5 Taq I" indicates this site was screened using mismatch RFLP analysis after amplification with the primer pair P5-P2, instead of the usual P1-P2.

R. Enzyme	R. Site Posn.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Taq I	374	+		+	+	+		+	+	+	+	+	+	+	+	+
pattern		A	F	A	A	A	F	A	A	A	A	A	A	A	A	A
Mnl I	19	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
	275	+	+		+	+	+	+	+	+	+	+	+	+	+	+
	338	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	432	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
pattern		A	A	C	A	A	A	A	A	A	A	A	A	A	A	O
Tsp 509 I	48	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	159										+		+			
	180	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	221	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	226	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	307	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	315	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	328?														+	
	393				+											
	405	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
pattern		A	A	A	D	A	A	A	A	A	L	A	L	A	N	A
Mbo II	200	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	256							+								
	303	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	388	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
pattern		A	A	A	A	A	A	G	A	A	A	A	A	A	A	A
P5 Taq I	23	+	+	+	+		+	+			+		+		+	+
	324	+	+	+	+	+	+	+	+	+	+	+	+		+	+
pattern		A	A	A	A	E	A	A	E	E	A	E	A	M	A	A
Overall Haplotype		A	F+	C	D	E+	F+	G	E+	E+	L'	E+	L'	M	N	O

5.1.1.3.2.2 P esculentus

Figure 5.15 indicates which variable base sites were screened for variation using restriction enzymes. The locations of restriction sites along the COI fragment are indicated in Table, as are details of the patterns of restriction fragments that are formed by the different haplotypes for each variable restriction enzyme. This table also lists the composite restriction enzyme patterns for each haplotype, and the haplotype groups which can be distinguished using PCR-RFLP only.

Table 5.6. *P. esculentus* restriction fragment patterns and overall haplotypes.

For each enzyme, for each restriction site, a '+' indicates a cut.

R. Enzyme	R. Site Posn.	A	B	C	D	E	F	G	H	I	J	K
Rsa I	174	+	+	+	+	+	+	+	+	+	+	+
	468			+	+							
pattern		A	A	C	C	A	A	A	A	A	A	A
Hinf I	14	+	+	+	+	+	+	+	+	+	+	+
	208	+	+	+	+	+	+	+	+		+	
	420		+							+		
pattern		A	B	A	A	A	A	A	A	I	A	K
Nla III	109	+	+	+	+	+			+	+	+	+
	218	+	+	+	+		+	+	+	+	+	+
	338	+	+	+	+	+	+	+	+	+	+	+
pattern		A	A	A	A	E	G	G	A	A	A	A
Overall Haplotype		A+	B	C+	C+	E	G+	G	A+	I	A+	K

5.1.1.4 Control Region

5.1.1.4.1 Sequences

Nucleotide sequences of an approximately 1050bp fragment including the mitochondrial control region (of approx. 976bp) were determined from both *P. plebejus* and *P. esculentus*, and aligned with existing sequences from the most closely related species available, these being *P. notialis*, *Daphnia pulex* and *Artemia franciscana*, as well as *Drosophila yakuba*.. The lengths of the control region vary widely among taxa, with *P. plebejus* having a length of 972 bp, *P. esculentus* a length of 976bp, *P. notialis 983bp* and *P. semisulcatus* a length of 992bp. The length in *Daphnia* is very short at 689bp, and that of *Artemia* is quite large (1770bp), while *Drosophila* is very much intermediate at 1029 bp for *D.* virescens and 1077bp for *D. yakuba*.. As expected, there are considerable indels among the sequences. These aligned sequences are shown in Figure 5.16. The amplified fragment includes very conserved flanking regions of 26bp of 12 rDNA and 49 bp of the Ile tRNA.

```
                    1 1111111112 2222222223 3333333334 4444444445 5555555556
               1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus   ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus     .......... .......... .......... .......... .......... ..........
P.notialis     .......... .......... .......... .......... .......... ..........
Daphnia pulex  .......... .......... .......... .......... .......... ..........
Artemia fran.  ACTGCTGGAT GCTGGGATAA GAACGACTGG ATAAAATATG AAAATGATAT TGTGGCGGAA
Drosophila y.  .......... .......... .......... .......... .......... ..........
```

```
                                                                1 1111111111 1111111111
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus       .......... .......... .......... .......... .......... ..........
P.notialis       .......... .......... .......... .......... .......... ..........
Daphnia pulex    .......... .......... .......... .......... .......... ..........
Artemia fran.    AGTAATAGTA ACGGACAAAG AAGTAAGATG ATAGTTATTT ACTGATTTGA GTGTGTTTAA
Drosophila y.    .......... .......... .......... .......... .......... ..........

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus       .......... .......... .......... .......... .......... ..........
P.notialis       .......... .......... .......... .......... .......... ..........
Daphnia pulex    .......... .......... .......... .......... .......... ..........
Artemia fran.    GAATGAAGCT TATGACTGCG AATAATGGGA GATCAGATGT ATAAGGTAAT TCAAAGCAGT
Drosophila y.    .......... .......... .......... .......... .......... ..........

                1111111111 1111111112 2222222222 2222222222 2222222222 2222222222
                8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus       .......... .......... .......... .......... .......... ..........
P.notialis       .......... .......... .......... .......... .......... ..........
Daphnia pulex    .......... .......... .......... .......... .......... ..........
Artemia fran.    CTCCCCGGCG AGCGACCGAA ACCCATTATC TAGCCGCCAT GAAGAGTGAA ATATCTCCTC
Drosophila y.    .......... .......... .......... .......... .......... ..........

                2222222222 2222222222 2222222222 2222222222 2222222222 2222222223
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus       .......... .......... .......... .......... .......... ..........
P.notialis       .......... .......... .......... .......... .......... ..........
Daphnia pulex    .......... .......... .......... .......... .......... ..........
Artemia fran.    GATAACCTTA TTTAAATGTG GAAAAATACT CTAGACGTTG CTGCAGATAT CTGAAGTAAG
Drosophila y.    .......... .......... .......... .......... .......... ..........

                3333333333 3333333333 3333333333 3333333333 3333333333 3333333333
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     ---------? ???GGCC??? TTT?AA?GTA AC????GTAG GATAATAATT GTT--ATTAT
P.plebejus       .........- --------GC ....?C.ACC ...CT.A..A C.CCT...A. T.A...C.C.
P.notialis       .........- ---------- ..AA.GTAAC TTTCTTA..T AT.TC...AA T..GA..C.C
Daphnia pulex    .........- ---------- ---------- ---------- ---------- ---.-.-----
Artemia fran.    AGATAAATAG TGT..ATAAC .AAT.CATC. GGTATAT.GT ATG.T.TGAA ..GACTACT.
Drosophila y.    ..AAATATAC AATAAAAAAA ...ATTTA.. .TATAAA.TT A....A.T.. A.CAA.A..A

                3333333333 3333333333 3333333333 3333333334 4444444444 4444444444
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TTTATGTAAA TAA-----TT TTGAATATAA AGCTAATAAC TAGACCAACT ATAAATTTTA
P.plebejus       ...T.A.... GT.ACAAC.. ......TA.. TAA..C.TTA .T.TTT..TA .C..T.AA.T
P.notialis       A.A.AAC..T GT.GTAGT.. ...T...ATT .CT...A..A .TA.TTT.-A T...T.AAC.
Daphnia pulex    ---------- ---.....-- ---------- ---------- ---------- ----------
Artemia fran.    G..TG..TGT .G.ATCTTA. C..TC..G.T .AT..CATCA .TACGAT.AA ....G.AAA.
Drosophila y.    ...T.A.... ...TAAATAA AAT..A.A.T .ATAG....T .TA.AT..TA ..T...AAAT
```

```
                4444444444 4444444444 4444444444 4444444444 4444444444 4444444444
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     GATAATTTTA AATTTAGAAT CAATTCCTGC TG--TGGATT CCCCCTTTTT TT-ATATAA-
P.plebejus       T..T.A.C.T T.A.....  ...A...A.. --........ T.T.....  ..?G......
P.notialis       T.A...AAAG .TG.....  .T.AC..CCT --.....T.. TTT..CC...  ...G......
Daphnia pulex    ---------- ---------- ---------- --.------ ---------- --.------.
Artemia fran.    ATGG..A.G. .TGA.GA.G. T..CAGAATT G.AA.T.TA. TGAATACG.G .AA..G.G.G
Drosophila y.    A.A..C.... T.AAC.A..A T....TAATT .A...AT..A TATATA.A.A ..AG....TT

                4444444444 4444444445 5555555555 5555555555 5555555555 5555555555
                8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     --ACTCTTAA CTTGTGATAC A?CTAA?GGC --AGTAAAAA CCTTAATATG CTTTAATTTT
P.plebejus       ...T?.G... T.AA.T.ACT GCAC..CATT .....GTT.C T...TGCT.. ---...AA.A
P.notialis       ...ACT.GT. TC.CAT.ATA .A....CTA. TA..GTCT.. .....CAC.A --...G....
Daphnia pulex    ..-------- ---------- ---------- ..-------- ---------- ----------
Artemia fran.    TA.AAT.G.. A..A.TTA.A TA..TGCATT TACTC.TTT. TGA.TG..AC GA.G.TAAAA
Drosophila y.    ...AATAA.. T.CAAAT..T TTT.T.TCAT AA.TA.TT.. TAA.T.A.AT TAG...AA..

                5555555555 5555555555 5555555555 5555555555 5555555555 5555555556
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     AGATCAAATT TACTATTGAT TTTAGCTA-- GTACTTTT-- -TGCTAACGC ATATTTTTTT
P.plebejus       .TTCG..G.. ..T......A GA..AAAG.. AGGTCA..AA TAT.C?CT.T .C..A.....
P.notialis       ...-.CT.. A.T...A.G. .A..AG.... ACC..C..A. .AAT...AAA TCT.GA...A
Daphnia pulex    ---------- ---------- --------.. --------.. .--------- ----------
Artemia fran.    .A.AAT..AA AGG.TAAATA GGGGAT..AC CC..CCC.AA A.A.CT.AAA T..CCA.A.A
Drosophila y.    TTTAA...AA ..A..AAAT. .G..A..T.. --.TAA.AAA C.AT.T.AAT .AT.A.....

                6666666666 6666666666 6666666666 6666666666 6666666666 6666666666
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     --GAAAATTA AAAAATGC-- CTCGAC--TA TTTTTTA--G AAATTAACAG TTTCTACAAC
P.plebejus       .....C-... ....TA?... AGTC.TCT.. .......... G.T..TTT.C A.GG.TT.CA
P.notialis       TA.G.TT.CT T..G.CC... .......... .......... ..TG.C.TGA .A.GATATG.
Daphnia pulex    AATTT...CC .G..GGA... ..TA.AAG.. .GGA.A...A ..C.CGGA.C CACT.TAGCG
Artemia fran.    AATG.T.C.T .GTT..A.AG .CTC..GCA. C.AACA.CTA ..C..G..GA CCAT.TA..A
Drosophila y.    AATC.CTAA. TCTG..AA.. ..TATTCCCC .A.AA....A ..TA...T.A .....TA..T

                6666666666 6666666666 6666666666 6666666667 7777777777 7777777777
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TTTACTTTAA ATTTAAAATA TTTTAT-TAA TTTTATTAAT TAAATGAATT ATATAGTACA
P.plebejus       ...GAA.A.G G......TGT ..A..-.GGT ......ATT. .....A.... .......GG.
P.notialis       A..CTA..TG .AG..T.TA. C..A...... ..A...A... -...A..C.. ........G.
Daphnia pulex    CAAGTA..T. .CGACCGC.T .GAA..T.GC CCGCCAA... ---TCA.TC. T.CG..GGG.
Artemia fran.    GGAGAACGG. TGG.GG.... .AG..CA.G. .AACG.ATT. ..TGAA.TG. .G..T.GTG.
Drosophila y.    .AA.TACCTT .....C.TAT ..A.T.T.T. A.A.T.AT.C A..TAA.... ..T.TT.TT.

                7777777777 7777777777 7777777777 7777777777 7777777777 7777777777
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TATAT-TT?G GAGAGTTAG? ---TAATTTA TAATATATAT A--TATTA-- -----GCCCT
P.plebejus       ..C.A...TA T..CA.ACTT A....TA.AC .G.GT..AT. .GA.G.A.TT AA...ATA?C
P.notialis       .....A..GC T.TTA...AT G.....AGA. C.T.T....C ..A.T...AG AA....A.A.
Daphnia pulex    .CCCCCCCCC T...ACCCCC ...--CCCCC ..TC...A.. GCC.TACCTC AT...TTA..
Artemia fran.    G...ATA.AT TTAGA.AC.T AGGCT.CC.C ...G.C.AGG CCT.G...AA GAGTCAAG.C
Drosophila y.    ..C...-.AA A.T.A.A.AA A...T.AAAT .....A.AT. TAC...ATTT AT...TAAAC
```

```
                7777777777 7777777778 8888888888 8888888888 8888888888 8888888888
                8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TAAATAGATT T?ATA?AC?? ?????????? ???????GGG ?T-AACC?GA A-GC-ATCTA
P.plebejus       .T.TA.A... .AT.TA.TTA .ATATT.AAT AATA.GTA.T .-.....CTT ....C.....
P.notialis       ..T...TC.G .C..TA.TTA TATATTAATT TAGATGT..T TA....AG.T .TT.T.....
Daphnia pulex   .TTT..A.AA .CTCTTTTTT ATTTTTTTTT TAGATTTATA A..TTTAGC. .AT.CGCT.G
Artemia fran.   ..CG..C.CA CACATA.TTT AATATTAGTA TATATAAC.T C.GT.TAGAT .CATTCA.CG
Drosophila y.   A.....AT.A .TTATA.ATA AAAATT--AT AAAACAATAT T..GTTAATT .AATA.AT..

                8888888888 8888888888 8888888888 8888888888 8888888888 8888888889
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     AG?AG?TCTA ATCCCATATG G??ACCAAAG GCTTTCCAAT ?TCTT--ATA TCCAAAAAAA
P.plebejus       .TT.AT.-.. .AG?.-...A TGG.AT-.TA ?G-..GT... C.A.A..... .TT.GTC.?T
P.notialis       .AT.AAA-.. TAAATG.C.A TAT.GTGTT. ATA..AT..A A.A..T.... .TA...T...
Daphnia pulex   .AG.AT.--- ----...GCA ATG.AGC.T. TTCG.AT... T.TAGC..A. ...GCTTG..
Artemia fran.   CTCTTT.A.T .GAAT.ATAA AAG.G.GGT. AA.G.A.GT. A.T.ATATC. .TT.T.T..T
Drosophila y.   .AA.AT.T.. TAAAT---.A ATT.AT-..A TTA.AAT... T.T.AAA... .TA.TTTT..

                9999999999 9999999999 9999999999 9999999999 9999999999 9999999999
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     AGTATTAGGG GTAGTTAAAG GATGGTCTAT G?T?TATAA- CCTAAGT?G? GTAAATAC?T
P.plebejus       TAA-.A.AAC .......... C..AT...TA .A.CA...-. TA...A.AAT .G..T..TT.
P.notialis       ...-.A.A.T .......... A...T...T. .A.CAT..-. .A...A.A.T .C.....TT.
Daphnia pulex   GAA------T AAGA...T.A T.CTAAT.T. CA.TAG..-. --.G.A.ATC T...T.T.TC
Artemia fran.   .T.--A.ATT T..T.....C CC.CT.AAGA .GGT.TCT-. --A..A.AAC C..G.G.TA.
Drosophila y.   .A.---TATA T..A.....T A.ATT.T.C. AA.TA....T TA.TTA.AAA A...TA.AAA

                                                          1 1111111111 1111111111
                9999999999 9999999999 9999999999 9999999990 0000000000 0000000000
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     A?GTCGCTTC TCCC?TGTA? G??AC?TTAA TTATATATAG ACTAGTAAGT ACTTCTAG?T
P.plebejus       .TAAT.-AG. .TAAT....T TTG.AAC..G G.....TA.. .T.G....A. .AA.....A.
P.notialis       .T..GT--AT GT.GT....T TTATATA..T GC...A.A.. ........A. .AA.....T.
Daphnia pulex   CTTGAAGAGT GAAAG.C.TC AGTCTCA.TC ..TCT.T.G. GGCTTGG..A G.CC..TCT.
Artemia fran.   TGA.TT-A.A .TTAT..CCT .GA.GA..TG G.G.TCCAG. .TA.A...AA T.A.-..TC.
Drosophila y.   .CC.ATTC.T .TTTT.T.TT TAA.AAAA.. A.T.T.G..T ....AGTTC. .AA.-..ATA

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                0000000000 0000000000 0000000000 0000000000 0000000000 0000000000
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     CC?TTTG?AA TTC?CTAAC? TTATATCTTA TTAAATAAAA TAGG-GTCTT ?TTTCTGTG?
P.plebejus       TGA...AT.. ..AA....AA ?......... AA........ ...T...... C..ATCC?.A
P.notialis       ..T.C.AT.. ...AG...GA ....T....G .......... ...C...... C..AT.-..A
Daphnia pulex   TTTCAAAG.. AGAA..G.GA C.T.GA.C.T AGTC.A..GG G...GA.A.. T..AT.C.AA
Artemia fran.   .ATC.AAT.C AA.C..CTTA AG.GGGT... AA.TT...T. G.TCTA.AAA C..AT.TGTA
Drosophila y.   GATAA.CT.T A.ATA..TAA A.G.T.AA.. .ATT....T. ...TTAA.AA T..AAATAAA
```

```
                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                0000000000 0000000001 1111111111 1111111111 1111111111 1111111111
                8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     AAC?--TTAA TTATTCTCGG ?----TTAAA GTTTCACCC? TC?AGGAGAC ACTATTTTTT
P.plebejus       ...T..AA.. .-...TA... G......TT. ......GA.T ..A..A.... ..........
P.notialis       ...T..CA.. .-TC.AGAA. A.......G. ......TA.A AAA..A.... ..........
Daphnia pulex   .G.TAT...T CA.C.AA.AA ACATT....T ....GGT..A AAA.TT...A .AGGGG...A
Artemia fran.   GTATTT..TG .CC.AACACT AAGTG...GG ACAGGGAAAA AAT.CT.CGG .TA.G...A.
Drosophila y.   ..AT....C. CA..C.AAAA A.....GGT. ACA.A.TTTG .AA.AA.A.A T....A..CA

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                1111111111 1111111111 1111111111 1111111111 1111111111 1111111112
                4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     A?AC?TTTAA ATTAAAT--T AATG?TAATC TATTTAATTA AAAAG?TTAA TT--TTAAAA
P.plebejus       .T..A.A.CT ..AGGT.A.. ...ATAT..A ...A.G.... T.T..A...G ..........
P.notialis       .T..CAA.C. T.....A... ..AATC.... .......... .....ACA.. .....A...T
Daphnia pulex   .ACTGAA..G ..ACTTAT.. .GGAACC.CT .TCG.G.GG. .CTTTTCA.C ..TT..G..G
Artemia fran.   .TTTAA.... ..GT..CTA. ...AAG.TAT .....GTC.T .TTTTA.ATC A.TA.A.TG.
Drosophila y.   .ATAT..AT. TAAC.T.C.. TGGAT.T..A ..AA....AT ..T.AT.... ..AA...TT.

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                2222222222 2222222222 2222222222 2222222222 2222222222 2222222222
                0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TAGGATATGG -GG?TAC?TA AGGTTAATTT AATATATAT- --TAGCCG?A ATCTTAGCTT
P.plebejus       ..AATGT..T C..T..TAA. CTA.......  .......TG. ...T..TTTT .ATCATTA..
P.notialis       ..TTT.TCTC .--T..TA.. .AT.......  .......TAA ....A.T.A. ..TAGTTT..
Daphnia pulex   A.A.G.C.T. C..CCCAAA. .A...G.AAA GT.CCTCCGA A...ATATTG .AAAGT.G..
Artemia fran.   CGTTC.G.AT A..T..AA.T GTT.C..A.. G.A.AT.GA. ..-.AA.CCC G.AAA.T.G.
Drosophila y.   ..TAT.TATA TATT..TA.. TA...G.AGA TT.....TAC ATATATATCT ..A.AGAAAA

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                2222222222 2222222222 2222222222 2222222223 3333333333 3333333333
                6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TTTACTTTTG ?TTAATAAAT TG?CTATTTT AAGCTCGC?T TTTAATGCCG C?TTAAAACT
P.plebejus       .C.T...CA. TC..GACTT. .AT.A.CAAC T.CT..T.A. ....G.A.AT AA.AGT..T.
P.notialis       ...T.AC.CC CCG...CTT. ACCAACCGGC T.A.....AA ........A- AG.AGT.TG.
Daphnia pulex   CCG.G....C A..G...TT. .TATGCAAGC ...TGA.TAA ...TT.TAAA GT..TTT.AA
Artemia fran.   .AA.AA.CA. C....AGTT. GCGT.TAGGG ..TT.T.TGA CCCC..TTTA .ACGTTT.A.
Drosophila y.   A.A.AA..AT T.......T. .TTA..AAAA T.TT.TATGA A..CC.AAAA TG.GTT.T..

                1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
                3333333333 3333333333 3333333333 3333333333 3333333333 3333333333
                2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
                1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus     TATTTT--TT CG?TATTTAT TTTTTTTT-T -CTTTTTTAC ATAAAATG?A GTGCCTGAAA
P.plebejus       ..G..A...C GTG.TAC.T. ...AGG.... A..GA??.-. ?.?.....A? .?...?????
P.notialis       .TAA.G.... T.T...G.TG C......... A........T ....------ ----------
Daphnia pulex   A.CAA-..-- ---------- --------.- .--------- ---------- ----------
Artemia fran.   GC.AAGGT.C TCGATA..TA CC.CAAAAAC AGCC....T. .GG.TC.TGC C.TTAATCCT
Drosophila y.   A..A.AAA.C AAT..A..T. .AAAAA..G. AA.C.....A .A....AAT. ..AAA.A.T.
```

```
              1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
              3333333333 3333333334 4444444444 4444444444 4444444444 4444444444
              8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
              1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus  AA-------- ---------- ---------- ---------- ---------- ----------
P.plebejus    ..G..... .......... .......... .......... .......... ..........
P.notialis    --.......... .......... .......... .......... .......... ..........
Daphnia pulex --.......... .......... .......... .......... .......... ..........
Artemia fran. GTAAAAAAGT AAATTAGCCC ACCTCAAATT TTTCTTAAAA AAAAATGTAA AATAGATAAT
Drosophila y. ..AAAAAAAA AAAAAAAAGA TGAGTTTTTT ATTATT.... .......... ..........

              1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
              4444444444 4444444444 4444444444 4444444444 4444444444 4444444445
              4444444445 5555555556 6666666667 7777777778 8888888889 9999999990
              1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus  ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus    .......... .......... .......... .......... .......... ..........
P.notialis    .......... .......... .......... .......... .......... ..........
Daphnia pulex .......... .......... .......... .......... .......... ..........
Artemia fran. TTTATTCTAT TTCAGCAAGG GAGGGCGGCA GCAACCTGGA TTATCTCCTT ACAGGAGATA
Drosophila y. .......... .......... .......... .......... .......... ..........

              1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
              5555555555 5555555555 5555555555 5555555555 5555555555 5555555555
              0000000001 1111111112 2222222223 3333333334 4444444445 5555555556
              1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus  ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus    .......... .......... .......... .......... .......... ..........
P.notialis    .......... .......... .......... .......... .......... ..........
Daphnia pulex .......... .......... .......... .......... .......... ..........
Artemia fran. ACTCAATATA GGTTGATCTC CCCAATAAAT TGAAGGACCG ATAATTTTAT CTTTTTTGAG
Drosophila y. .......... .......... .......... .......... .......... ..........

              1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
              5555555555 5555555555 5555555555 5555555556 6666666666 6666666666
              6666666667 7777777778 8888888889 9999999990 0000000001 1111111112
              1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus  ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus    .......... .......... .......... .......... .......... ..........
P.notialis    .......... .......... .......... .......... .......... ..........
Daphnia pulex .......... .......... .......... .......... .......... ..........
Artemia fran. GTGAGTACCA AGAATCTGAT CCTGCTCCTA GTCCGGGGTA TAAACGTAAT AGTTAATTAG
Drosophila y. .......... .......... .......... .......... .......... ..........

              1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
              6666666666 6666666666 6666666666 6666666666 6666666666 6666666666
              2222222223 3333333334 4444444445 5555555556 6666666667 7777777778
              1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus  ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus    .......... .......... .......... .......... .......... ..........
P.notialis    .......... .......... .......... .......... .......... ..........
Daphnia pulex .......... .......... .......... .......... .......... ..........
Artemia fran. AAAAAGAATT ATATGTTTGA TAAGAGTATT ATATTGTCTA CTGTTAAAAA AGGATAGAGT
Drosophila y. .......... .......... .......... .......... .......... ..........
```

```
             1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
             6666666666 6666666667 7777777777 7777777777 7777777777 7777777777
             8888888889 9999999990 0000000001 1111111112 2222222223 3333333334
             1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
P.esculentus ---------- ---------- ---------- ---------- ---------- ----------
P.plebejus   .......... .......... .......... .......... .......... ..........
P.notialis   .......... .......... .......... .......... .......... ..........
Daphnia pulex .......... .......... .......... .......... .......... ..........
Artemia fran. TATTTTATTT AGATTAGCCT GGAAGAAGAA CTCCCTCCCT CGTTACGCTT CAGTTTGAAG
Drosophila y. .......... .......... .......... .......... .......... ..........

             1111111111 1111111111 1111111111 1111111111 11
             7777777777 7777777777 7777777777 7777777777 77
             4444444445 5555555556 6666666667 7777777778 88
             1234567890 1234567890 1234567890 1234567890 12
P.esculentus ---------- ---------- ---------- ---------- --
P.plebejus   .......... .......... .......... .......... ..
P.notialis   .......... .......... .......... .......... ..
Daphnia pulex .......... .......... .......... .......... ..
Artemia fran. CATAAGCTCG GGGGGGGGGG GGTTCTTCTC TCCCAGGGGT GT
Drosophila y. .......... .......... .......... .......... ..
```

Figure 5.16. Aligned Control Region nucleotide sequences from *Penaeus* spp. and three arthropods.

There is considerable divergence between the three *Penaeus* sequences, with a Jukes-Cantor corrected nucleotide sequence divergence of over 0.56, compared to a value of about 1.3 for divergence from *Artemia* (Table 5.7). There are considerable indels among these species. The transition / transversion ratio between all species is 0.4 - 0.5. All *Penaeus* sequences are very AT-rich (75 - 80%), compared with 67% for *Daphnia*, and 93% for *Drosophila*..

Table 5.7. Control region Jukes-Cantor nucleotide divergences between Penaeus spp. and 3 other arthropods.

Distances in the upper-right matrix. Standard Errors in lower-left matrix

OTUs		1	2	3	4	5	6
1..	P.esculentus		0.5510	0.5230	1.3167	1.3721	0.9504
2..	P.plebejus	0.0342		0.6036	1.1965	1.2786	0.8722
3..	P.notialis	0.0327	0.0359		1.2330	1.1102	0.9409
4..	Daphnia pulex	0.1148	0.0945	0.0980		1.4224	1.2343
5..	Artemia fran.	0.0993	0.0861	0.0694	0.1225		1.1759
6..	Drosophila y.	0.0585	0.0515	0.0561	0.0976	0.0721	

As described in section 4.2.2.1.3, *Penaeus* - specific primers were developed to amplify this region. and these are now being used in a number of laboratories in Australia and overseas.

5.1.1.4.1.1 P. plebejus

Sequences of the control region were obtained from 6 individuals of *P. plebejus* from 3 locations spanning the species' entire distribution. The variability within the species is shown in Figure 5.17, where all the haplotypes found are listed..

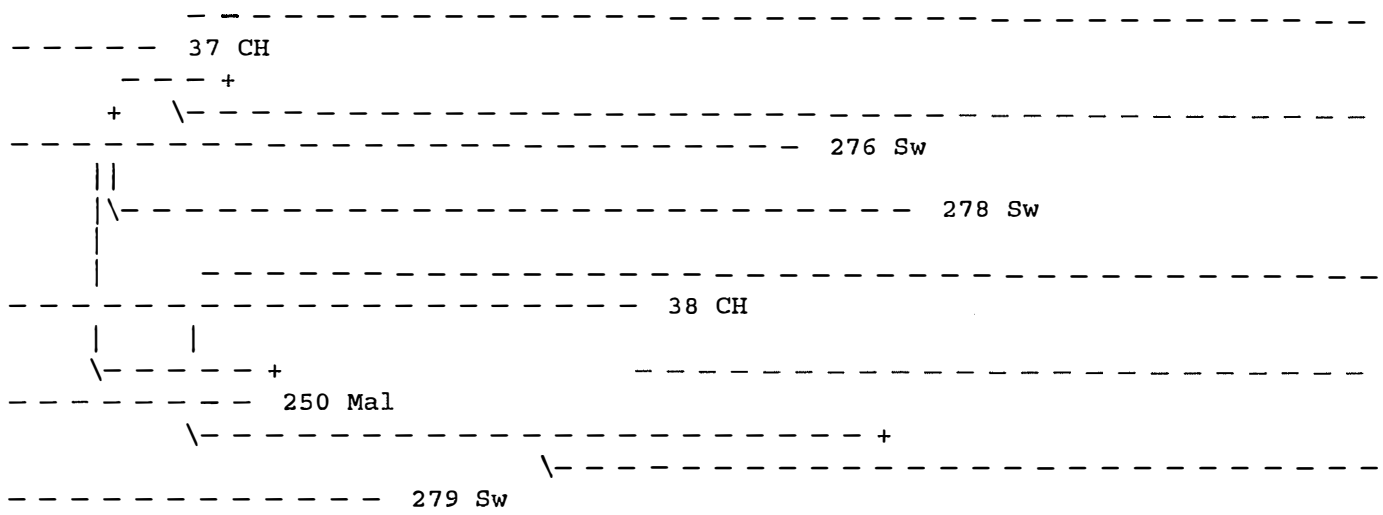
```
                112 2223333334 4444555556 6677788888 8888999999 99
          12578790 1463348991 3377556895 8901202247 7788011256 79
        6907069153 8244507269 4758681309 1502156822 3709401953 17

37 CH   TACCATAATG CGTGAGGCTG ATAAGCTATA TCAATTAGGT AATGTCTTAG GC
38 CH   C.....T.CA T..A..A... ........C. ...GCCG.A. ...AC.ACGA .T
250 Mal CG....T..A TTCA...T.A ..T.AA.G.G ..G.....A. G...C.A..A ..
276 Sw  C..TG.TG.. T..A.A.T.. GCTG...... ..G..CG.AC .GC....... C.
278 Sw  C..T.CT... T..A..A... ..T.AAC... ..G..C..A. ....C..... ..
279 Sw  C.A...T..A T.CAG..TCA ..T.AA.... CTG..CGAA. ....CTA..A ..
```

Figure 5.17 - *P. plebejus* control region haplotypes and their polymorphic nucleotide sites.

52 sites (5.3%) proved to be polymorphic. Diversity in this segment is far higher (up to an order of magnitude greater) than that seen in the COI fragment, the average nucleotide diversity being 2.2% (+/-0.54%), and the overall haplotype diversity 1.0, with the maximum pairwise divergence being 2.8%. No indels were observed among these individuals. The transition / transversion ratio between haplotypes was very high, ranging from 4.0 to 9.0 with an average of 5.8.

A neighbour-joining tree of the phylogenetic relationships among haplotypes was constructed, and is shown in Figure 5.18.



Scale: each — is approximately equal to the distance of 0.00021

Figure 5.18 - Neighbour-joining tree of *P. plebejus* control region haplotypes.

5.1.1.4.1.2 P esculentus

Sequences of the control region were obtained from 4 individuals of *P. esculentus* from 2 locations spanning the species' entire distribution. The variability within the species is shown in Figure 5.19, where all the haplotypes found are listed..

```
            11111 1112223339 999999
       2237800366 7890233480 235578
       7995567278 1592518950 221350

15 MB  GATGATAGAC GGCAATATCC TTCCAT
10 MB  .......... ......G?.. .C?...
12 SB  CCCA.CGC.T AA.GGC?CG? C.?GGC
13 SB  CCCAGCGCG. A.AGGC?.GT C.AGGC
```

Figure 5.19 - *P. esculentus* control region haplotypes and their polymorphic nucleotide sites.

26 sites (2.7%) proved to be polymorphic. Diversity in this segment is again higher than that seen in the COI fragment, the average nucleotide diversity being 3.5% (+/- 0.82%), and overall haplotype diversity 1.0, with the maximum pairwise divergence being 3.9%. No indels were observed among these individuals. The transition / transversion ratio between haplotypes was very high, averaging 2.6, and ranging from 2.0 upwards.

A neighbour-joining tree of the phylogenetic relationships among haplotypes was constructed, and Figure 5.20 shows these relationships.

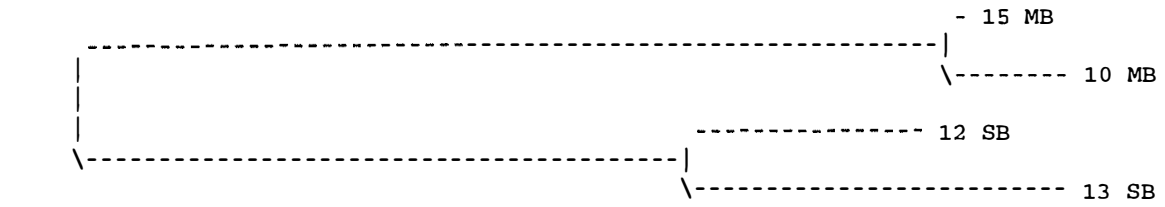


Figure 5.20 - Neighbour-joining tree of *P. esculentus* control region haplotypes.

5.1.2 nDNA

5.1.2.1 Microsatellites

5.1.2.1.1 *P esculentus*

A partial library (including only those fragments around 400bp in length after digestion with *Sau*3A) was constructed from DNA from the individual #16 from Moreton Bay. Twenty clones which positively hybridised to the CA dinucleotide probe were re-hybridised to the probe. Of these, eleven clones were again positive, and were sequenced. The general characteristics of each microsatellite (repeat type, number & length) were determined and are shown in Table 5.8 - *P. esculentus* microsatellites..

Table 5.8 - *P. esculentus* microsatellites.

MS #	Repeat Type	No. Repeats	PCR Product Length
1.1	CA	35	350
1.8	CA	38	211
2.3	CA	50	216
2.7	CA	48	253
1.4	CT	48	300
	CA	42	
2.11	CA	25	380
	TA	15	
1.2	TA	49	420
1.7	CA	64	
	TA	53	
2.6	TA	21	161
	CA	29	
2.9	CA	52	
	TA	15	
1.3	none		

Of these sequences, one (1.3) did not prove to contain microsatellite repeats. Of the others, one (2.9) contained a microsatellite that was very long and complex, that is, it was a complex mixture of repeat types (n(GA)n(CA)n(TA)n(CA)), which is not particularly suitable for VNTR (variable number of tandem repeats) analysis. One clone (1.4) contained two separate microsatellites. Flanking PCR primers were designed for the remaining ten microsatellites, and are listed in Table 5.9. All PCR primer pairs were tested and optimised for PCR conditions before analysing for length variation using the ABI 373 and GENESCAN.

Table 5.9 - Microsatellite primer pairs

Primer Name	Flourescent Label	Sequence
PE 2.3 A - T	TET	CACAAGAGTGTGAGTTCAATGCTC
PE 2.3 B		CTTGCTGATAACATCACTGGCATT
PE 1.7 A - T	TET	GTGATGACACTTACGACGACGATA
PE 1.7 B		GATCTCCCCGCACTGTCTGGATGC
PE 1.1 A - T	TET	AAATGCTCTCATGAATAACCTGTA
PE 1.1 B		TCTTGGACCACTTGTAAACGACTC
PE 1.2 A - F	6-FAM	AAATGAAGAGTGTTTTACGAATCA
PE 1.2 B		TTGTATGTGGTCACGGAGTAGACA
PE 1.8 A		GGGAACAGCAGAAGTAATACTACA
PE 1.8 B - F	6-FAM	AGTATTAGAGTGTGTTGACCATAC
PE 2.11 A - H	HEX	GAATAAAGCAGGAGATGCGAGTGA
PE 2.11 B		ATCGGTGATGACACTTACGACGAC
PE 1.4 A - H	HEX	TATTATGCCTGAAGAGATGAGG
PE 1.4 B		ACACATGCACATACATGGATAG
PE 1.4 C		TCCATGTATGTGCATGTGTATC
PE 1.4 D - H	HEX	CATTACGAGAATTTGCTGAGAG
PE 2.6 A		TTTTCTTTTCCTCTTACTGTGT
PE 2.6 B		GTGAACATATACATGCAC
PE 2.7 A - H	HEX	GAATAAAGCAGGAGATGCGAGTGA
PE 2.7 B		GTGTGGTACAAAACGGTAAATCAT

One of the sequences (2.6) included only a short flanking sequence at the 3' end, complicating the design of a flanking PCR primer. This fragment was not consistently amplified by PCR, but could perhaps be used in the future with a redesigned primer. Two loci (1.2 & 1.7) could not be satisfactorily or consistently amplified. One locus (2.7) gave only a weak product, and had unfortunately been labelled with the weakest dye (yellow). Redesigned and/or relabelled primers should allow this locus to also be used in the future. Three loci (2.11 plus the 2 loci from clone 1.4 - 1.4AB & 1.4CD) appeared to have few alleles and low heterozygosity. Of the remaining four loci, the three that amplified best (1.1, 1.8 & 2.3) were screened over a number of individuals, to ascertain if they would be likely to be useful in population analyses. The levels of variability are shown in Table 5.10 - Microsatellite loci variation.

Table 5.10 - Microsatellite loci variation.

Locus	2N	Heterozygosity
1.1	80	0.777
1.8	53	0.903
2.3	41	0.888
2.7	22	0.918
2.11	14	0.0

All polymorphic loci exhibited very high levels of variation, ranging from H=0.777 to 0.918, with an average level of heterozygosity of 0.872. This is in stark contrast to the average level of allozyme H in prawns of 0.019 [8].

5.1.2.1.2 P. plebejus

The microsatellite loci primer pairs developed for *P esculentus* were also tested on *P. plebejus*. It was expected that a large proportion of such loci would also be amplified in the congeneric species by the same primers, as had been found in many other species groups (e.g., turtles). Unfortunately, this did not prove to be the case. None of the primer pairs developed from the *P esculentus* clones consistently amplified any single locus in *P. plebejus*. This was a considerable set-back in the analysis of nuclear DNA variation in *P. plebejus*, as considerable time and effort had been invested in the original set of primers, and this could not be repeated for *P. plebejus* within the time and budget restrictions. This is, however, a very significant finding which, if confirmed, has considerable implications for the development of nuclear DNA markers in penaeid prawns.

5.2 Analysis of stock structure

5.2.1 P. plebejus

5.2.1.1 16s rDNA

5.2.1.1.1 Sequences

The geographic distribution of 16s haplotypes, as determined by complete sequencing , is shown in Table 5.11. The relatively low level of diversity did not suggest that this marker would prove particularly effective in analysing population structure within this species, and only a small number of individuals were analysed as a result. These initial results indicate that there is not an obvious difference in 16s genetic composition of this species from one region to another. Contingency table analysis of the distribution of haplotypes shows no significant differences among regions (P = 0.17), although a conventional F_{ST} calculated from haplotype frequencies (F_{ST}=0.14) was close to significance (P=0.053). Analyses including the information on sequence divergence among haplotypes supports this conclusion (F_{ST} = 0.06, P = 0.15). There is some indication that the diversity of haplotypes decreases from north to south, but sample sizes were not sufficient to reliably test this. The unpromising nature of this data for population analysis meant that further samples or analyses were not deemed particularly useful to pursue.

Table 5.11 - Geographic distribution of *P. plebejus* 16s haplotypes, as determined by complete sequencing.

Haplotypes	Swains	Moreton Bay	Mallacoota
A	1	2	4
B	0	1	0
C	1	0	0
D	1	0	0

Table 5.12 - Matrices of pairwise F_{ST}'s and P's for *P. plebejus* 16s haplotypes.

```
-------------------------
Population pairwise FSTs
-------------------------

Computing conventional F-Statistics from haplotype frequencies

                    Swains    Moreton Bay

Moreton Bay        -0.07143
Mallacoota          0.35135     0.11111


------------
FST P values
------------

Number of permutations : 992

                     Swains              Moreton Bay

Moreton Bay          0.99899+-0.0002
Mallacoota           0.15307+-0.0085     0.00000+-0.0000

------------------------
Population pairwise FSTs
------------------------


Distance method: Jukes & Cantor

                    Swains     Moreton Bay

Moreton Bay         0.00000
Mallacoota          0.11023     0.11111
```

```
------------
FST P values
------------

Number of permutations : 992

                        Swains          Moreton Bay

Moreton Bay           0.62135+-0.0131
Mallacoota            0.14099+-0.0113       0.00000+-0.0000
```

5.2.1.2 COI

5.2.1.2.1 Sequences

Diversity within this fragment is much greater than that within 16s. The geographic distribution of haplotypes from complete sequencing is shown in Table 5.13. A larger data set was analysed than for the 16s fragment, but there was no clear pattern of population subdivision within the species. Contingency table analyses of haplotype frequency showed no significant differences among samples (P = 0.08). F_{ST} analyses again supported this conclusion. F_{ST} calculated from haplotypes alone was approximately 0.00, while F_{ST} calculated using sequences was 0.04 (P = 0.2)

Table 5.13 - Geographic distribution of *P. plebejus* COI haplotypes, as determined by complete sequencing.

Haplo	Sw	MI	MB	CH	WL	LI	LT	MALL
A	3	2	1	3	2	3	1	2
B	0	0	1	0	0	0	0	0
C	0	0	0	0	0	0	0	1
D	2	0	0	0	0	0	0	0
E	2	0	1	0	0	1	0	0
F	0	0	1	0	0	0	0	0
G	0	0	0	1	0	0	0	0
H	0	0	0	0	1	0	0	0
I	0	1	0	0	0	0	2	0
J	0	0	0	0	0	0	1	0
K	0	0	0	0	1	0	0	0
L	0	1	0	0	0	0	0	0

Table 5.14 - Matrices of pairwise F_{ST}'s and P's for *P. plebejus* COI haplotypes from sequencing data.

```
------------------------
Population pairwise FSTs
------------------------

Computing conventional F-Statistics from haplotype frequencies

               Swains  Moreton I. Moreton Bay Coffs H. Wallis L.  L. Illaw.  L. Tuggerah

Moreton I.     -0.00984
Moreton Bay    -0.05698  -0.04762
Coffs Harbour  0.04533   -0.06667   0.07692
Wallis L.      -0.00984  -0.11111  -0.04762  -0.06667
L. Illawarra   -0.06203  -0.06667   0.00000  -0.14286  -0.06667
L. Tuggerah    0.11239   -0.11111   0.02222   0.17949   0.04762   0.17949
Mallacoota     -0.01402  -0.13580  -0.01961  -0.15254  -0.13580  -0.15254   0.08911


------------
FST P values
------------

Number of permutations : 100

               Swains   Moreton I. Moreton Bay  Coffs H.  Wallis L.  L. Illaw.  L. Tuggerah

Moreton I.     0.54455
Moreton Bay    0.68317    0.99010
Coffs Harbour  0.19802    0.99010    0.46535
Wallis L.      0.46535    0.99010    0.99010    0.99010
L. Illawarra   0.71287    0.99010    0.63366    0.99010    0.99010
L. Tuggerah    0.16832    0.99010    0.64356    0.31683    0.13861    0.32673
Mallacoota     0.32673    0.63366    0.40594    0.99010    0.49505    0.99010    0.18812

------------------------
Population pairwise FSTs
------------------------

Distance method: Jukes & Cantor

               Swains  Moreton I.  Moreton B.  Coffs H.   Wallis L.  L. Illaw.  L. Tuggerah

Moreton I.     -0.01343
Moreton Bay    0.11996    0.02604
Coffs Harbour  0.07861    0.00007    0.13338
Wallis L.      0.01590   -0.05536    0.04850    0.13393
L. Illawarra  -0.10265   -0.14334    0.03705    0.00040   -0.08337
L. Tuggerah    0.13711   -0.18002    0.11143    0.18985   -0.00027    0.05439
Mallacoota     0.06855   -0.03686    0.08450    0.01495    0.08504    0.01495    0.12992


------------
FST P values
------------

Number of permutations : 100

                Swains  Moreton I. Moreton Bay  Coffs H.  Wallis L.  L. Illaw.  L. Tuggerah

Moreton I.      0.42574
Moreton Bay     0.14851    0.35644
Coffs Harbour   0.22772    0.24752    0.21782
Wallis L.       0.27723    0.07921    0.15842    0.22772
L. Illawarra    0.69307    0.99010    0.29703    0.64356    0.99010
L. Tuggerah     0.18812    0.81188    0.20792    0.43564    0.47525    0.42574
Mallacoota      0.19802    0.84158    0.33663    0.75248    0.15842    0.73267    0.47525
```

Table 5.15 - Matrix of net COI sequence divergences among *P. plebejus* populations using sequence data.

SW	MI	MB	CH	WL	LI	LT	MALL
0.0000000							
0.0000000	0.0000000						
0.0002857	0.0001062	0.0000000					
0.0002285	0.0000000	0.0004000	0.0000000				
0.0000000	0.0000000	0.0001041	0.0000000	0.0000000			
0.0000000	0.0000000	0.0001000	0.0000000	0.0000000	0.0000000		
0.0008000	0.0001000	0.0011187	0.0016000	0.0007104	0.0007000	0.0000000	
0.0002285	0.0000000	0.0004000	0.0000000	0.0000000	0.0000000	0.0016000	0.0000000

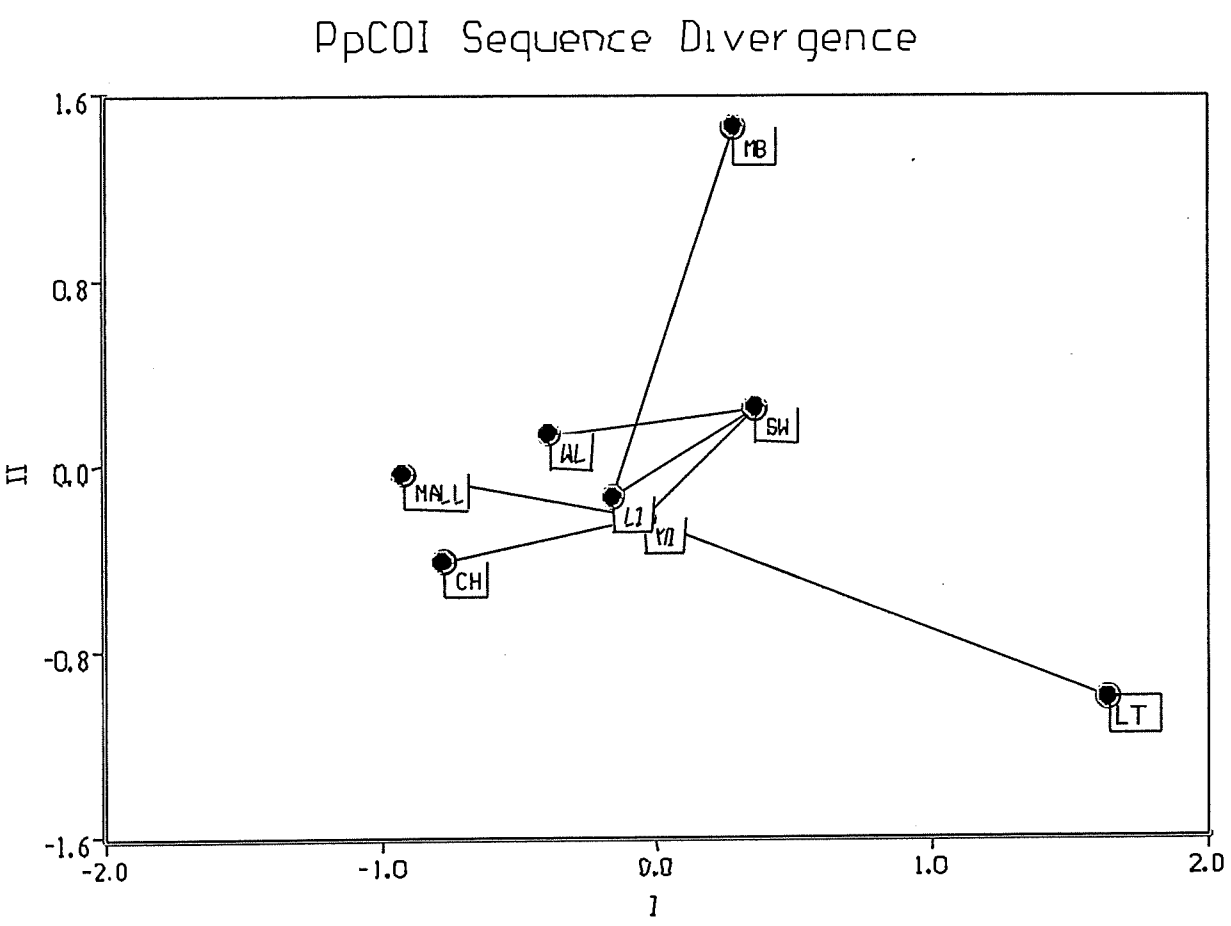


Figure 5.21 - Multidimensional Scaling (MDS) pattern of net COI sequence divergences among *P. plebejus* populations using sequence data.

(Population sample points connected using a minimum spanning tree.)

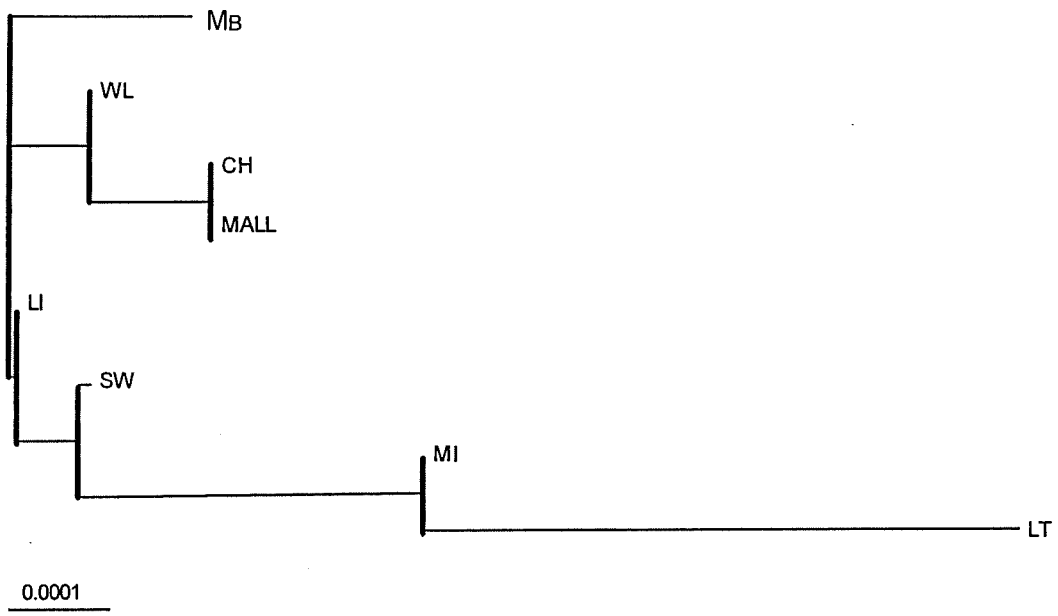


Figure 5.22 - Neighbour-joining tree of net COI sequence divergences among P. plebejus populations using sequencing data.

The high frequency of single haplotypes found in locations meant that population subdivision could not be tested properly without additional samples being analysed. This was undertaken using PCR-RFLPs.

5.2.1.2.2 RFLPs

A reduced set of haplotypes could be distinguished using this method, however all major clades could be distinguished. An additional 160 individuals were genotyped using this more rapid technique The total distribution of COI haplotypes (from both sequencing and RFLPs) among locations is shown in Table 5.16. Contingency table analyses of haplotype frequency showed marginally significant differences among samples (P = 0.04, 95% CI: 0.036-0.044). F_{ST} analyses did not support this significant subdivision. F_{ST} calculated from haplotypes alone was 0.015 (P = 0.15), while F_{ST} calculated using sequences was 0.014 (P = 0.15)

Table 5.16 - Geographic distribution of P. plebejus COI haplotypes, as determined by PCR-RFLP's.

Haplos	SW	MI	MB	CH	WL	LT	LI	MALL
A	10	11	4	19	21	11	15	11
C	1	1	0	1	1	0	1	1
D	3	0	0	0	0	0	0	1
E'	7	3	4	4	17	10	4	7
F'	0	0	2	1	0	0	1	0
G	0	0	0	1	0	0	0	0
L'	1	1	0	1	2	2	0	0
M	0	0	0	0	0	2	0	0
N	0	0	0	0	0	0	0	1
O	1	0	0	0	0	0	0	1

Table 5.17 - Matrices of pairwise F_{ST}'s and P's for P. plebejus COI haplotypes from PCR-RFLP data.

```
------------------------
Population pairwise FSTs
------------------------

Computing conventional F-Statistics from haplotype frequencies

                 Swains  Moreton I.  Moreton B. Coffs H.  Wallis L.  L. Tuggerah  L. Illaw.

Moreton I.       0.02218
Moreton Bay     -0.02165   0.06340
Coffs Harbour    0.05495  -0.04598   0.09135
Wallis L.       -0.00424   0.03051  -0.01608   0.06620
L. Tuggerah     -0.01472   0.04351  -0.03055   0.08084  -0.02164
L. Illawarra     0.04647  -0.05020   0.07428  -0.03875   0.05021   0.06977
Mallacoota      -0.03445  -0.00237  -0.02583   0.02830  -0.02016  -0.01953   0.01486


------------
FST P values
------------

Number of permutations : 100

                 Swains  Mor. I.  Mor. B.  Coffs H.  Wall. L.  L. Tugg.  L. Ill.

Moreton I.       0.18812
Moreton Bay      0.53465  0.16832
Coffs Harbour    0.06931  0.99010  0.10891
Wallis L.        0.38614  0.19802  0.48515  0.01980*
L. Tuggerah      0.60396  0.09901  0.63366  0.03960* 0.70297
L. Illawarra     0.05941  0.99010  0.09901  0.99010  0.12871  0.07921
Mallacoota       0.93069  0.39604  0.62376  0.13861  0.72277  0.59406  0.30693

------------------------
Population pairwise FSTs
------------------------


Distance method: Jukes & Cantor

                 Swains  Moreton I.  Moreton B. Coffs H.  Wallis L.  L. Tuggerah  L. Illawarra

Moreton I.      -0.01337
Moreton Bay      0.00723   0.05026
Coffs Harbour    0.02221  -0.04151   0.07208
Wallis L.        0.00842   0.04029   0.00072   0.08341
L. Tuggerah      0.00106   0.02943   0.00822   0.07581  -0.00611
L. Illawarra     0.00780  -0.04514   0.03485  -0.03465   0.05351   0.06049
Mallacoota      -0.03628  -0.02489  -0.00687   0.01278  -0.01177  -0.01576  -0.00309


------------
FST P values
------------

Number of permutations : 100

                 Swains  Moreton I.  More. B.  Coffs H.  Wallis L.  L. Tuggerah  L. Illawarra

Moreton I.       0.51485
Moreton Bay      0.32673   0.19802
Coffs Harbour    0.12871   0.99010   0.10891
Wallis L.        0.24752   0.14851   0.37624   0.01980*
L. Tuggerah      0.30693   0.10891   0.29703   0.00000** 0.45545
L. Illawarra     0.31683   0.86139   0.33663   0.94059   0.06931   0.06931
Mallacoota       0.97030   0.76238   0.43564   0.20792   0.52475   0.51485   0.38614
```

Table 5.18 - Matrix of net COI sequence divergences among *P. plebejus* populations using PCR-RFLP data.

SW	MI	MB	CH	WL	LT	LI	MALL
0.00000000							
0.00000000	0.00000000						
0.00002613	0.00007333	0.00000000					
0.00003856	0.00000000	0.00009435	0.00000000				
0.00000695	0.00006585	0.00000000	0.00013414	0.00000000			
0.00000202	0.00006800	0.00002133	0.00014413	0.00000000	0.00000000		
0.00001752	0.00000000	0.00003047	0.00000000	0.00008487	0.00011885	0.00000000	
0.00000000	0.00000000	0.00000000	0.00001909	0.00000000	0.00000000	0.00000000	0.00000000

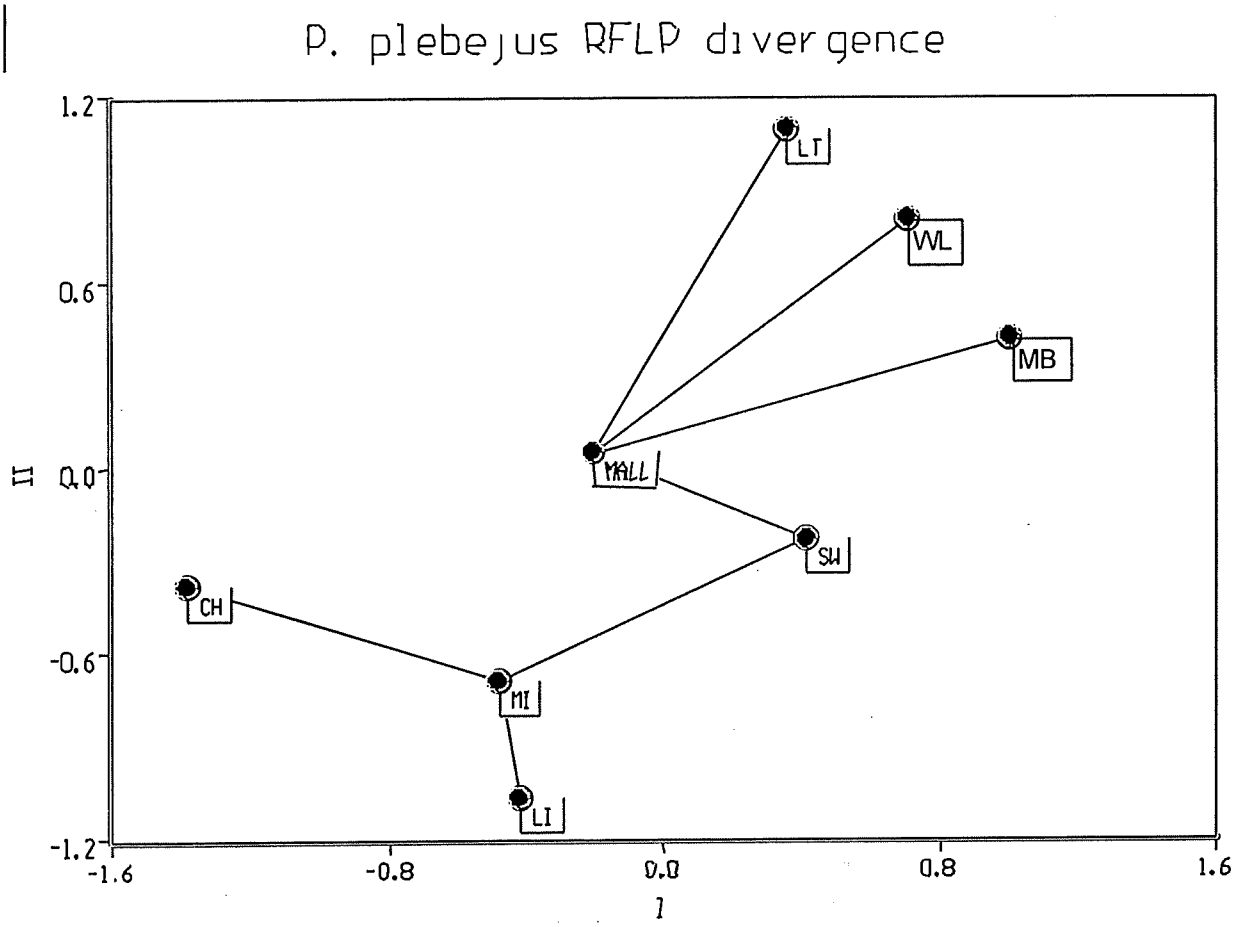


Figure 5.23 - Multidimensional Scaling (MDS) pattern of net COI sequence divergences among *P. plebejus* populations using PCR-RFLP data.

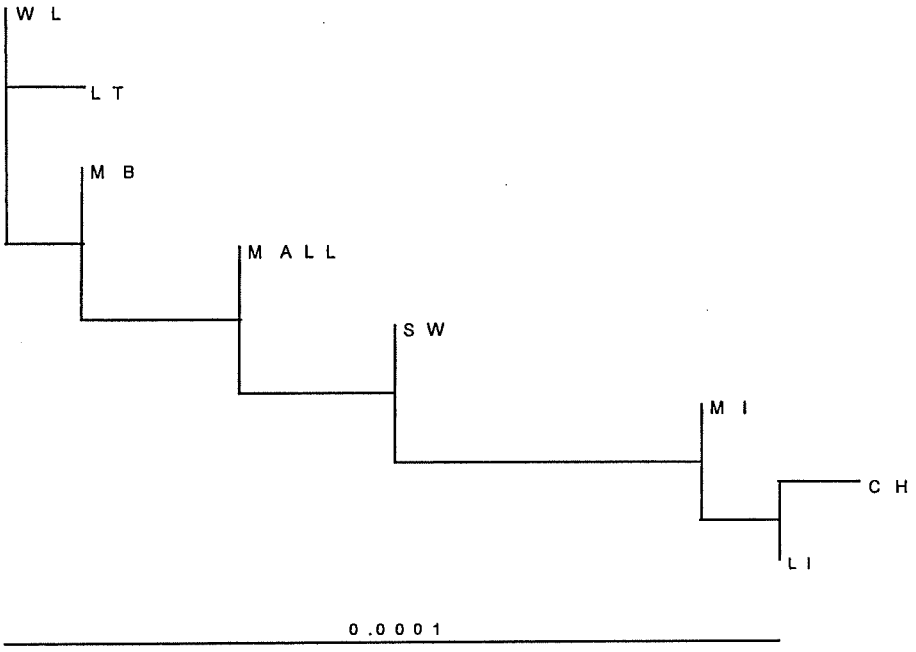


Figure 5.24 - Neighbour-joining tree of net COI sequence divergences among *P. plebejus* populations using PCR-RFLP data.

5.2.2 *P esculentus*

5.2.2.1 16s rDNA

5.2.2.1.1 Sequences

The geographic distribution of 16s haplotypes, as determined by complete sequencing , is shown in Table 5.19. The relatively low level of diversity did not suggest that this marker would prove particularly effective in analysing population structure within this species, and only a small number of individuals were analysed as a result. Contingency table analysis of the distribution of haplotypes shows a significant difference among regions (P = 0.002), however, this is mostly due to a large difference between the samples from Queensland and Western Australia. Among Qld. samples, the overall contingency chi-square P = 0.41. F-statistic analyses, both excluding and including the information on sequence divergence among haplotypes, support this conclusion (Table 5.20). The relatively low levels of variation detected in this segment meant that further samples or analyses were not deemed particularly useful to pursue.

Table 5.19 - Geographic distribution of *P. esculentus* 16s haplotypes, as determined by complete sequencing.

Haplo	SB	W	TS	C	Mac	MB
A	0	1	1	2	3	1
B	0	0	2	0	0	1
C	0	0	0	1	0	1
D	0	4	2	3	3	0
E	4	0	0	0	0	0
F	0	1	0	1	0	0

Table 5.20 - F-statistics among *P. esculentus* samples calculated from 16s sequences. The two regions considered were WA and all Qld. samples.

(F_{ST} : variance among subpopulations relative to total.
F_{SC} : variance among subpopulations relative to regions.
F_{CT} : variance among regions relative to total.)

Data Type	F_{ST}	P	F_{SC}	P	F_{CT}	P
Haplotype	0.467	(0.007)	0.003	(0.52)	0.465	(0.17)
Sequence	0.563	(0.001)	0.049	(0.37)	0.54	(0.16)

Table 5.21 - Matrices of pairwise F_{ST}'s and P's for *P. esculentus* 16s haplotypes.

```
------------------------
Population pairwise FSTs
------------------------

Computing conventional F-Statistics from haplotype frequencies

                 Shark Bay       Weipa      Torres Strait      Cairns        Mackay

Weipa              0.64179
Torres Strait      0.55801       0.00535
Cairns             0.50339      -0.09948      -0.04356
Mackay             0.64179      -0.02857       0.00535      -0.09948
Moreton Bay        0.57895       0.20354      -0.10795      -0.03279       0.09091


------------
FST P values
------------

Number of permutations : 100

                Shark Bay       Weipa     Torres Strait      Cairns         Mackay

Weipa           0.00000***
Torres Strait   0.00990***     0.29703
Cairns          0.00000***     0.73267       0.59406
Mackay          0.00000***     0.51485       0.34653       0.80198
Moreton Bay     0.01980*       0.29703       0.82178       0.71287       0.45545


------------------------
Population pairwise FSTs
------------------------


Distance method: Jukes & Cantor

                 Shark Bay         Weipa      Torres Strait     Cairns        Mackay

Weipa              0.71518
Torres Strait      0.62974       0.06708
Cairns             0.58532      -0.08506       0.00162
Mackay             0.75779      -0.10017       0.01267      -0.11551
Moreton Bay        0.66681       0.30787      -0.04727       0.05421       0.25032
```

```
------------
FST P values
------------

Number of permutations : 100

                Shark Bay       Weipa       Torres Strait      Cairns          Mackay

Weipa           0.00990**
Torres Strait   0.00000***      0.34653
Cairns          0.00990**       0.70297         0.40594
Mackay          0.00990**       0.99010         0.44554         0.99010
Moreton Bay     0.04950*        0.06931         0.53465         0.17822         0.02970*
```

Table 5.22 - Matrix of net sequence divergence among populations of *P. esculentus* for 16s.

SB	0.00000000					
W	0.00294000	0.00000000				
TS	0.00252000	0.00014000	0.00000000			
C	0.00240000	0.00000000	0.00000000	0.00000000		
Mac	0.00252000	0.00000000	0.00000000	0.00000000	0.00000000	
MB	0.00210000	0.00084000	0.00000000	0.00010000	0.00042000	0.00000000

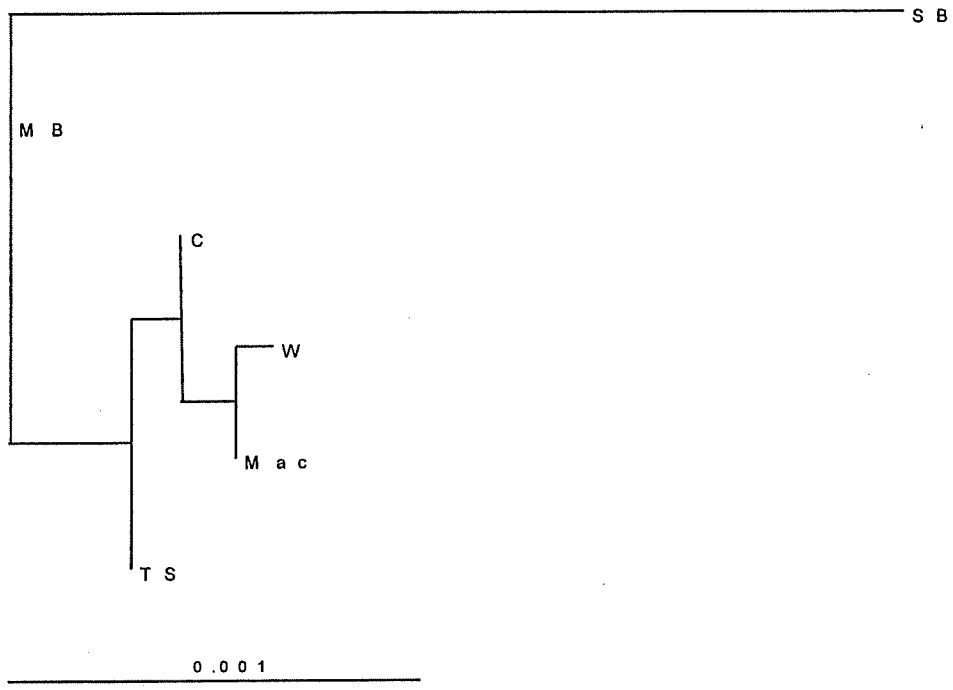


Figure 5.25 - Neighbour-joining tree of net 16s sequence divergences among *P. esculentus* populations.

5.2.2.2 COI

5.2.2.2.1 Sequences

Diversity within this fragment is much greater than that within 16s. The geographic distribution of haplotypes from complete sequencing is shown in Table 5.23. A larger data set was analysed than for the 16s fragment, and there was a clear pattern of population subdivision within the species. Contingency table analyses of haplotype frequency showed significant differences among all

samples ($P < 0.0001$). Considering Qld samples only, the chi-square $P = 0.16$. F_{ST} analyses again supported these conclusions.

Table 5.23 - Geographic distribution of *P. esculentus* COI haplotypes, as determined by complete sequencing.

Haplos	WA	W	TS	C	Mac	MB
A	0	4	4	5	2	2
B	0	0	0	1	0	1
C	4	0	0	0	0	0
D	1	0	0	0	0	0
E	0	0	1	0	0	1
F	0	0	0	0	0	1
G	0	1	0	0	3	0
H	0	1	0	0	0	0
I	0	0	1	0	0	0
J	0	0	0	0	1	0
K	0	0	0	0	0	1

Table 5.24 - F-statistics among *P. esculentus* samples calculated from COI sequences. The two regions considered were WA and all Qld. samples.

($\mathbf{F_{ST}}$: variance among subpopulations relative to total.
$\mathbf{F_{SC}}$: variance among subpopulations relative to regions.
$\mathbf{F_{CT}}$: variance among regions relative to total.)

Data Type	F_{ST}	P	F_{SC}	P	F_{CT}	P
Haplotype	0.21	(0.00001)	0.06	(0. 16)	0.39	(0.16)
Sequence	0.42	(0.00001)	0.03	(0.44)	0.69	(0.17)

Table 5.25 - Matrices of pairwise F_{ST} 's and P's for *P. esculentus* COI haplotypes from sequencing data.

```
------------------------
Population pairwise FSTs
------------------------

Computing conventional F-Statistics from haplotype frequencies

                Shark Bay        Weipa   Torres Strait       Cairns       Mackay

Weipa             0.49265
Torres Strait     0.49265      -0.08000
Cairns            0.63614      -0.05000      -0.05000
Mackay            0.42195       0.04000       0.14286      0.26154
Moreton Bay       0.31719       0.01429      -0.02222      0.08800      0.06250
```

```
------------
FST P values
------------

Number of permutations : 100

                Shark Bay        Weipa      Torres Strait     Cairns          Mackay

Weipa            0.03960*
Torres Strait    0.01980*       0.99010
Cairns           0.00990**      0.99010        0.99010
Mackay           0.00000***     0.13861        0.04950*       0.08911
Moreton Bay      0.00000***     0.06931        0.71287        0.17822        0.22772

------------------------
Population pairwise FSTs
------------------------


Distance method: Jukes & Cantor

                   Shark Bay        Weipa Torres Strait        Cairns        Mackay

Weipa                0.78494
Torres Strait        0.73110       0.00023
Cairns               0.84680       0.00013      -0.09115
Mackay               0.75934       0.00682       0.17163       0.23995
Moreton Bay          0.64234      -0.05031      -0.14301      -0.05934       0.02477


------------
FST P values
------------

Number of permutations : 100

                Shark Bay        Weipa      Torres Strait     Cairns          Mackay

Weipa            0.00990**
Torres Strait    0.00000***     0.42574
Cairns           0.00000***     0.38614        0.99010
Mackay           0.00000***     0.27723        0.00990**      0.18812
Moreton Bay      0.00000***     0.81188        0.89109        0.90099        0.44554
```

Table 5.26 - Matrix of net COI sequence divergences among *P. esculentus* populations using sequence data.

```
SB
W          0.00429666
TS         0.00428666  0.00000222
C          0.00430000  0.00000222  0.00000000
Mac        0.00471000  0.00000944  0.00042111  0.00042111
MB         0.00427666  0.00000000  0.00000000  0.00000000  0.00007500
```

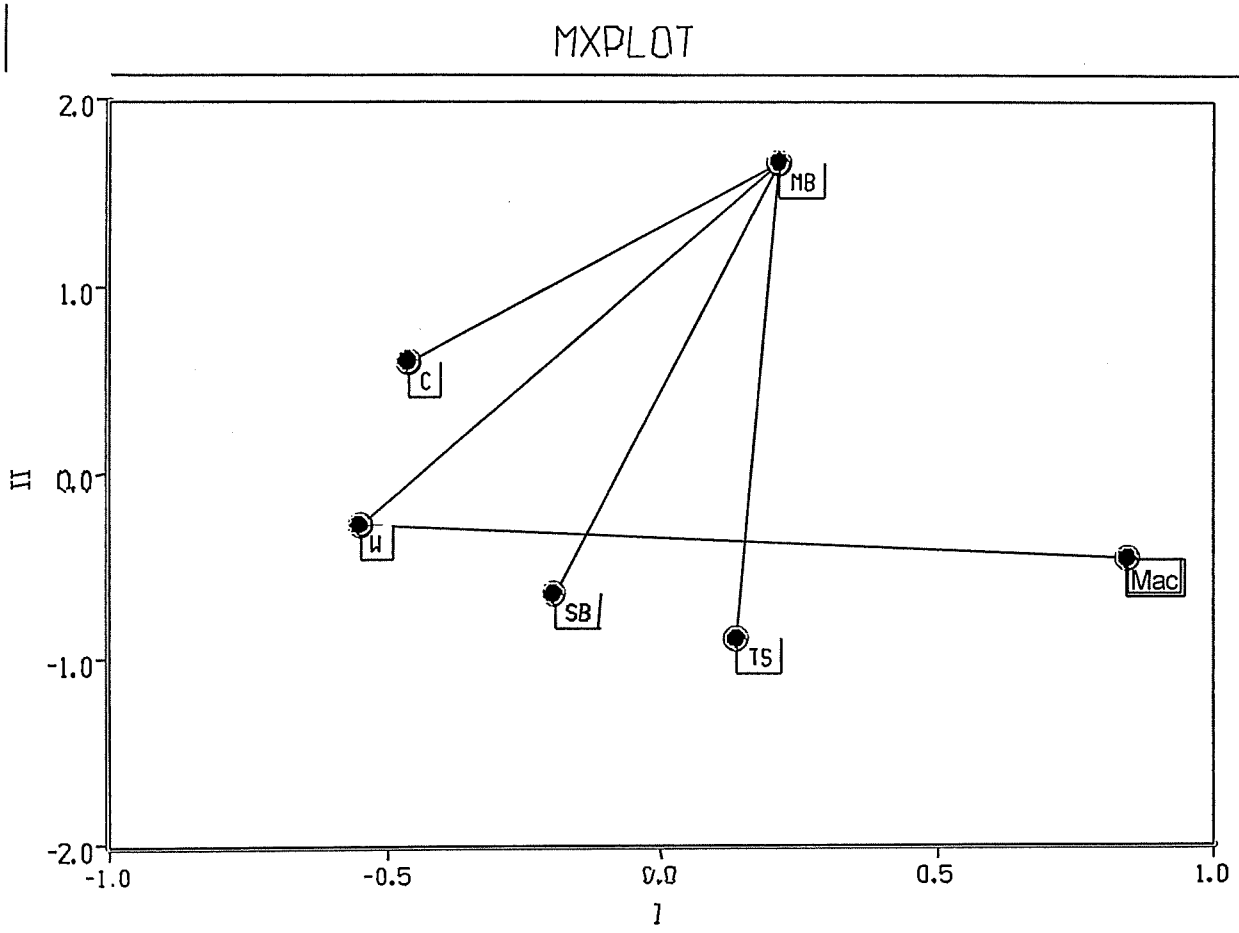


Figure 5.26 - Multidimensional Scaling (MDS) pattern of net COI sequence divergences among *P. esculentus* populations using sequence data.

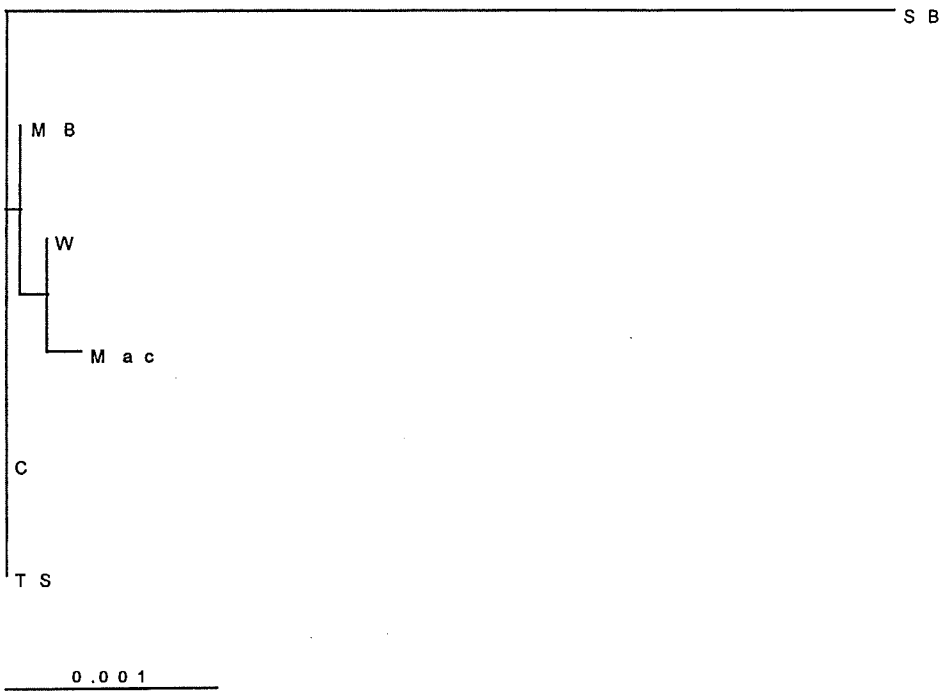


Figure 5.27 - Neighbour-joining tree of net COI sequence divergences among *P.Êesculentus* populations using sequencing data.

The high frequency of single haplotypes found in locations meant that tests of population subdivision were not very powerful without additional samples being analysed. This was undertaken using PCR-RFLPs.

5.2.2.2.2 RFLPs

Again, a reduced set of haplotypes could be distinguished using this method, however all major clades could be distinguished. An additional 136 individuals and four sample sites were genotyped. The total distribution of COI haplotypes (from both sequencing and RFLPs) among locations is shown in Table 5.27. Contingency table analyses of haplotype frequency showed significant differences among all samples ($P < 0.0001$). Considering Qld samples only, the chi-square $P = 0.16$. F_{ST} analyses again supported these conclusions.

Table 5.27 - Geographic distribution of *P. esculentus* COI haplotypes, as determined by PCR-RFLP's.

Haplos	A'	B	C'	E	G'	I	K
SB	1	0	9	0	0	0	0
NB	12	0	0	0	1	0	1
FB	1	0	0	0	1	0	1
G	9	0	0	0	5	0	0
W	13	0	0	0	2	0	2
GOC	4	0	0	0	0	0	0
TS	21	0	0	1	0	1	3
C	39	1	0	0	6	0	2
Mac	13	0	0	0	2	0	0
MB	6	0	0	1	2	0	1

Table 5.28 - F-statistics among *P. esculentus* samples calculated from COI PCR-RFLP's. The two regions considered were WA and all Qld. samples.

Data Type	F_{ST}	P	F_{SC}	P	F_{CT}	P
Haplotype	0.65	(0.0001)	0.003	(0.44)	0.65	(0.10)
Sequence	0.81	(0.0001)	0.015	(0.19)	0.81	(0.09)

Table 5.29 - Matrices of pairwise F_{ST}'s and P's for *P. esculentus* COI haplotypes from PCR-RFLP data.

```
------------------------
Population pairwise FSTs
------------------------

Computing conventional F-Statistics from haplotype frequencies

              Shark Bay  Nichol B.  Fog Bay  Groote    Weipa      GOC   Torres St.  Cairns  Mackay

Nichol Bay    0.73460
Fog Bay       0.72399  -0.07975
Groote        0.60961   0.09175   0.07711
Weipa         0.64631  -0.04913  -0.05721   0.02981
GOC           0.83936  -0.10149  -0.09674   0.13551  -0.03591
Torres Strait0.67709  -0.03686  -0.04132   0.13868  -0.02406  -0.06122
Cairns        0.67807  -0.03776  -0.04267   0.06556  -0.02840  -0.05239  0.00690
Mackay        0.75047  -0.05456  -0.05723   0.06335  -0.02693  -0.07660  0.00794 -0.03671
Moreton Bay   0.55083   0.02344   0.00847  -0.04587  -0.03605   0.04996  0.03410  0.02410 0.03405


------------
```

FST P values

Number of permutations : 100

	Shark Bay	Nichol B.	Fog B.	Groote	Weipa	GOC	Torres St.	Cairns	Mackay
Nichol Bay	0.00000*								
Fog Bay	0.00000*	0.99010							
Groote	0.00000*	0.10891	0.09901						
Weipa	0.00000*	0.77228	0.77228	0.21782					
GOC	0.00000*	0.99010	0.99010	0.04950*	0.23762				
Torres Strait	0.00000*	0.73267	0.89109	0.00990**	0.68317	0.66337			
Cairns	0.00000*	0.83168	0.99010	0.12871	0.79208	0.22772	0.27723		
Mackay	0.00000*	0.83168	0.99010	0.04950*	0.63366	0.99010	0.31683	0.81188	
Moreton Bay	0.00000*	0.38614	0.45545	0.58416	0.57426	0.32673	0.12871	0.21782	0.21782

Population pairwise FSTs

Distance method: Jukes & Cantor

	Shark B.	Nichol B.	Fog Bay	Groote	Weipa	GOC	Torres St.	Cairns	Mackay
Nichol B.	0.82827								
Fog Bay	0.82194	-0.07988							
Groote	0.78970	0.13122	0.11859						
Weipa	0.79127	-0.05668	-0.06298	0.08098					
GOC	0.83936	-0.13228	-0.13053	0.13551	-0.08698				
Torres St.	0.79485	-0.01574	-0.01976	0.22432	-0.00661	-0.07997			
Cairns	0.82007	-0.03472	-0.03852	0.09256	-0.02346	-0.08800	0.04831		
Mackay	0.84008	-0.03671	-0.04026	0.06335	-0.02311	-0.07660	0.06594	-0.03825	
More. Bay	0.73772	-0.02947	-0.03861	-0.01143	-0.05110	-0.07459	0.02423	-0.01350	-0.029*

FST P values

Number of permutations : 100

	Shark B.	Nichol B.	Fog B.	Groote	Weipa	GOC	Torres St.	Cairns	Mackay
Nichol Bay	0.00000*								
Fog Bay	0.00000*	0.99010							
Groote	0.00000*	0.05941	0.10891						
Weipa	0.00000*	0.99010	0.99010	0.09901					
GOC	0.00000*	0.42574	0.99010	0.21782	0.77228				
Torres Strait	0.00000*	0.35644	0.51485	0.00000*	0.46535	0.76238			
Cairns	0.00000*	0.76238	0.99010	0.03960*	0.77228	0.99010	0.04950*		
Mackay	0.00000*	0.83168	0.74257	0.21782	0.64356	0.40594	0.03960*	0.99010	
Moreton Bay	0.00000*	0.56436	0.72277	0.31683	0.65347	0.74257	0.21782	0.48515	0.53465

Table 5.30 - Matrix of net COI sequence divergences among *P. esculentus* populations using PCR-RFLP data.

SB									
NB	0.003440								
FB	0.003440	0.0000000							
G	0.003670	0.0001150	0.0001057						
W	0.003460	0.0000000	0.0000000	0.000050					
GOC	0.003440	0.0000000	0.0000000	0.000230	0.000020				
TS	0.003440	0.0000003	0.0000004	0.000232	0.000020	0.000000			
C	0.003470	0.0000000	0.0000000	0.000066	0.000000	0.000030	0.000030		
Mac	0.003460	0.0000000	0.0000000	0.000050	0.000000	0.000020	0.000020	0.000000	
MB	0.003495	0.0000000	0.0000000	0.000000	0.000000	0.000055	0.000035	0.000000	0.000000

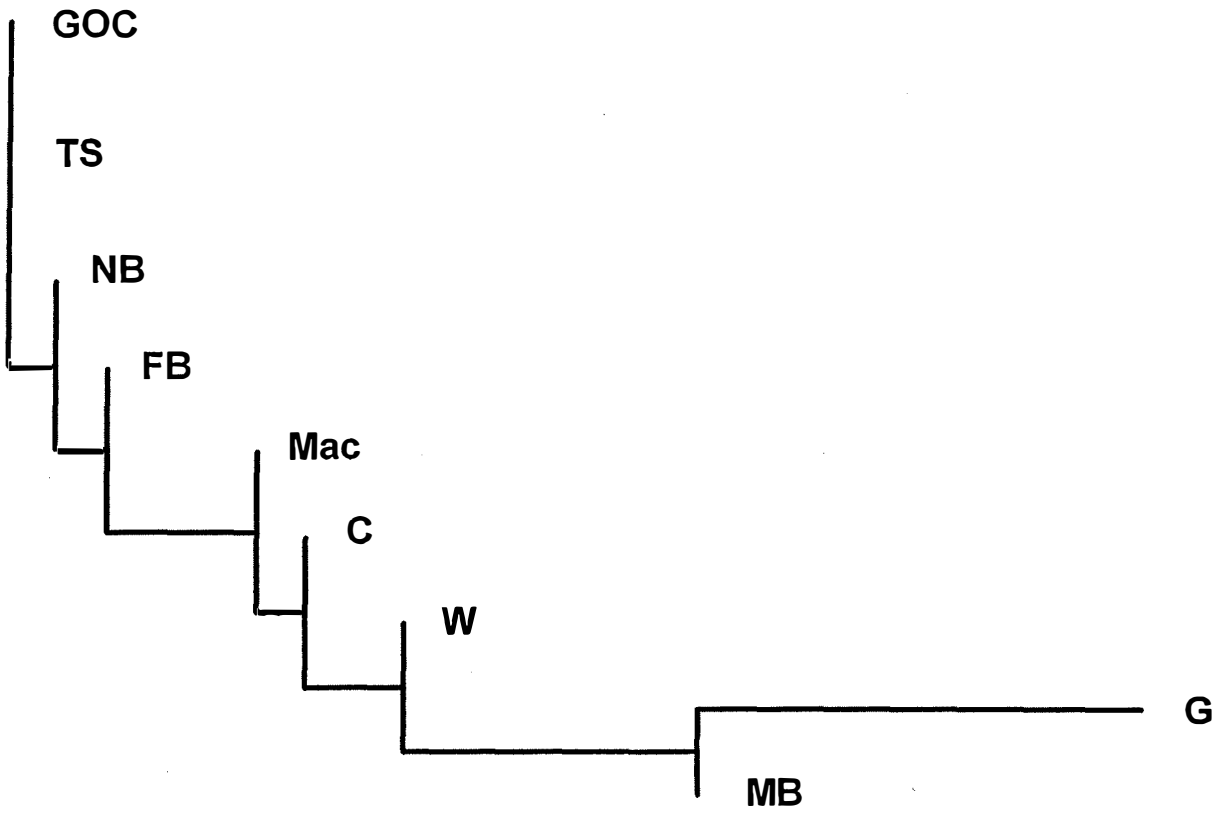


Figure 5.28 - Neighbour-joining tree of net COI sequence divergences among *P. esculentus* populations using PCR-RFLP data.

6. Discussion

6.1 Technique Assessment & Development

6.1.1 Mitochondrial DNA

The analysis of mitochondrial DNA variation in penaeid prawns has proved to be very possible, relatively uncomplicated, and entirely feasible for projects such as this investigating population-level variation, be it of wild or cultured animals. Universal degenerate PCR primers have proved to be suitable for accessing the initial sequence information from the mitochondria of these species, something which has certainly not always been the case with invertebrate marine organisms. This has proved to be the case with three regions of mtDNA, those being the 12s rDNA, 16s rDNA and CO I genes. This did not prove to be the case with the control region however, and it was only with perseverance and the application of new data that any sequence variation data was obtained from this region at all.

The development of *Penaeus* - specific primers for the CO I gene has been particularly productive in allowing more rapid and accurate assessment of mtDNA variation possible in these species. These primers appear to work extremely well, and the development of different primers for the different sub-genera of the genus has now made it possible to very easily assess mtDNA variation in additional species of *Penaeus* .

Furthermore, there was considerable development of PCR-RFLP markers for mtDNA. This has allowed large numbers of individuals to be rapidly screened for mtDNA variation, something that is not possible with mtDNA sequencing. Importantly, because the specific RFLP markers were developed from sequence information, this rapid screening technique was accomplished without missing a large portion of the natural variation, a very common consequence of most mtDNA studies employing RFLP techniques, where randomly-chosen restriction enzymes are used. Instead, here, the restriction enzymes were specifically chosen to target known highly variable sites in the DNA sequence. Using this approach also enabled the development of RFLP markers for variable DNA sites that would not normally be possible to detect using RFLPs. Because the variable sites were known, it was possible to create restriction sites at variable positions where such sites did not exist naturally. This was done using mismatch PCR primers, and permitted much more DNA variation to be quickly assessed than if a sequencing approach was not undertaken.

The most exciting has been the development of *Penaeus* - specific PCR primers for amplifying the mitochondrial control region. This has not previously been done successfully for decapod crustacea, and it now allows an enormous amount of hitherto inaccessible mtDNA variation to be accessed. Once again, this means that mtDNA variation can now be quickly assessed in additional species of penaeid prawn in the future. Unfortunately, however, the development of these primers took considerably longer than was originally anticipated. This has meant that, because these primers could not be developed until the end of the project, that there was insufficient time in the project to allow sufficient numbers of individuals to be screened for mtDNA control region variation to enable a worthwhile population-level analysis of that variation. However, it has now been shown that there is considerable potential for the analysis of this segment of DNA in both wild populations of *Penaeus* as well as, in the future, cultured populations of these species.

6.1.2 Microsatellite DNA

This has also been an exciting development in the analysis of DNA variation in penaeid prawns. It was not previously known if good microsatellite variation existed in penaeid prawns, how easily it would be to find, or to screen, and whether or not microsatellite PCR primers developed for one species of prawn could potentially be used on other species. This study has been the first to examine these questions in penaeid prawns, and has provided a number of valuable initial answers to these questions.

Firstly, it does appear that there are abundant microsatellites in prawn nuclear genomes. They are relatively easily accessible for analysis using standard microsatellite detection techniques, but they do appear to be relatively long and complicated (imperfect) repeats, which does tend to make their use for population analyses somewhat problematic. Specifically, because such a high proportion of prawn microsatellites appear to be so long and complicated, that it is therefore substantially more difficult to easily find enough microsatellites to undertake good population-level analyses. From the initial twenty positive clones, only three excellent microsatellites could be easily assessed for variation. However, these loci were certainly highly variable (see below), and were of far greater use than allozyme loci, the only other nuclear loci previously analysed for population level genetic variation in prawns.

The other complication with the microsatellite PCR primers, is that appears that they are largely, or at least substantially, species specific. This makes it much less worthwhile to develop a large number of primers for one species, if a number of them cannot be used for other species as well.

Once again, however, the development of these primers took considerably longer than was originally anticipated. This has meant that, because these primers could not be developed until the end of the project, that there was insufficient time in the project to allow sufficient numbers of individuals to be screened for microsatellite variation to enable a worthwhile population-level analysis of that variation. However, it has now been shown that there is considerable potential for the analysis of this microsatellite DNA in both wild populations of *Penaeus* as well as, in the future, cultured populations of these species.

6.2 Levels of Genetic Variation

Five different DNA markers of genetic variation have been examined in penaeid prawns during this study. These have been microsatellite DNA and four different mtDNA regions: 12s rDNA, 16s rDNA, CO I and the control region. Each of these types of DNA marker have provided different levels of DNA variation for analysis in these species, and have been of varying degrees of usefulness in carrying out further analyses of population structure and within- and among-population levels of variation.

The 12s rDNA gene was sequenced only in *P. esculentus* . Among 3 individuals from 2 populations, there were found to be only 3 different haplotypes(only 4 sites, 1.1%, were polymorphic), giving a haplotype diversity of 1.0 and a sequence diversity of only 0.7%. This was clearly insufficient variation to be worth pursuing as a population marker, and was therefore not pursued.

The 16s rDNA gene was sequenced in both *P. esculentus* and *P. plebejus* . Among 10 *P. plebejus* individuals from 3 populations, there were found to be only 4 different haplotypes (only 4 sites,

0.8%, were polymorphic), giving a haplotype diversity of 0.56 and an average sequence diversity of only 0.18%. There was a slightly higher level of variation among 31 *P. esculentus* individuals from 6 populations, where there were found to be only 4 different haplotypes (5 sites, 1.0%, were polymorphic), giving a haplotype diversity of 0.63 and an average sequence diversity of 0.18%. Although this is useful genetic variation, which can be put to good use in certain applications, this level of variation was also not particularly suitable for making population genetic inferences, as it did not provide sufficient statistical power for population-level analyses. Therefore this gene was also not pursued further for these analyses, beyond the first initial screening of DNA sequencing of this gene.

The mitochondrial cytochrome oxidase I (COI) gene was the only remaining mitochondrial region for which primers had been developed that may work in the penaeid species. This gene proved to be considerably more variable in the target species, as suspected . Among 34 *P. plebejus* individuals from 8 populations, there were found to be 12 different haplotypes (11 sites, 2.6%, were polymorphic), giving a haplotype diversity of 0.74 and an average sequence diversity of 0.28%. Among 35 *P. esculentus* individuals from 6 populations, there were found to be 11 different haplotypes (9 sites, 1.9%, were polymorphic), giving a haplotype diversity of 0.60 and an average sequence diversity of 0.18%. given the generally higher levels of diversity present, and the lack of other available DNA regions for analysis at the time, this was the gene concentrated on for population analyses, using both DNA sequences and PCR-RFLPs. This gene has provided a great deal of useful information about DNA diversity at both the intra-and inter-population levels. As penaeid-specific primers have also been developed for these species and other penaeid species, it is also now a very valuable gene for quickly analysing variation in penaeid and crustacean species. These primers have now been utilised in other crustacean studies both in Australia and overseas, examining DNA variation in 23 other prawn species (in collaboration with Dr. K. H. Chu at the Chinese University of Hong Kong), 4 species of mud crab (in collaboration with Dr. C. Keenan at the QDPI Bribie Island Aquaculture Centre) and three species of bay lobster (in collaboration with Dr. T. Burton, Zoology Dept., Univ. of Qld.).

By far the most nucleotide sequence variability found in the mitochondrial genome was that in the control region. Among 6 *P. plebejus* individuals from 3 populations, there were found to be 6 different haplotypes (52 sites, 5.3%, were polymorphic), giving a haplotype diversity of 1.0 and an average sequence diversity of 2.2%. Among 4 *P. esculentus* individuals from 2 populations, there were found to be 4 different haplotypes (26 sites, 2.7%, were polymorphic), giving a haplotype diversity of 1.0 and an average sequence diversity of 3.5%. These levels of diversity are 10 or more times greater than those found in other parts of mitochondrial genome, and suggest that this region may be particularly suitable for analysing for within-and among-population variation in penaeid prawns and perhaps other crustaceans as well. The fact that penaeid-specific primers have now also been developed for these species and other penaeid species, makes this perhaps the best mitochondrial region for quickly analysing variation in penaeid and crustacean species. These primers have now been utilised in other crustacean studies both in Australia and overseas, examining DNA variation in other prawn species (in collaboration with Dr. K. H. Chu at the Chinese University of Hong Kong and Dr. C. Keenan at the QDPI Bribie Island Aquaculture Centre).

The levels of nuclear genetic variation previously detected (using allozymes) in penaeid prawns have been notoriously low in comparison to many other organisms (Lavery et al., 1993). In *P. plebejus* , the maximum heterozygosity (H) detected at any one allozyme locus has been only approximately 0.35, while the average over all loci is only 0.033. In *P. esculentus* the maximum heterozygosity (H) detected at any one allozyme locus has been only approximately 0.47, while the

average over all loci is only 0.022. (Mulled & Latter). In extreme contrast, the levels of variation detected in penaeid prawns using microsatellites have been very high. In *P. esculentus* , the maximum heterozygosity detected in one microsatellite locus was 0.92, while the average over the four loci screened for a sufficient number of individuals to give reasonable estimates, was 0.88. Although there may be other difficulties associated with the ease of use of microsatellite loci in penaeid prawns (see above), it certainly does appear that there is a wealth of genetic variation that can be detected in these species when microsatellites are used. These markers would prove particularly useful in examining any aspects of paternity and inheritance, especially in aquaculture questions.

Overall, there has been very good success in finding highly variable genetic markers for penaeid prawns. This has previously been a difficult task, but it now apparent that at the DNA level, unlike what is found at the protein level, there is a wealth of genetic variation that can be detected. And there are now a number of choices as to which markers can be employed for any particular question. For haploid, uniparentally inherited genetic data (mtDNA), there are two very good alternatives: the COI gene for moderate levels of diversity (which for some purposes can be more useful than extremely high levels of variation), and the control region for high levels of diversity. For diploid, biparentally inherited markers (nDNA), it is now known that microsatellite loci offer a potentially very large amount of DNA variation for examination in these species.

6.3 Stock Structure

6.3.1 *P. plebejus*

The results from 16s rDNA sequencing revealed little population variation throughout the species' distribution. The diversity detected using this marker was low, and there were few haplotypes distinguished. The geographic distribution of those haplotypes (Table 5.11), suggests that there is little evidence of significant population structuring shown in this marker, and the statistical tests support this.

The results from the mitochondrial COI sequencing are much more extensive in the number of individuals as well as the number of locations examined. However, the limited number of individuals that could be analysed using the full sequencing approach meant that there was insufficient power to detect any possible population structure. This is reflected in the patterns of results shown in section 5.2.1.2.1. In the F_{ST}'s calculated from either haplotype frequencies or sequence differences, most values are relatively small, but even where large F_{ST}'s are observed (e.g., Coffs Harbour v. L. Tuggerah, $F_{ST} = 0.18$), the probability levels are not at all significant ($P = 0.3 - 0.4$). The net sequence divergences between samples are also very low, ranging from zero to 0.1%. The pattern of relationships among populations is shown in both the MDS pattern and the NJ tree. These both show that there is no clear pattern of genetic geographic relationships among samples. There are no obvious geographic groupings which may be indicative of the existence of distinct stocks, and samples are not obviously closely related to nearby samples, which may indicate gradual isolation by distance throughout the east coast distribution.

When the numbers of individuals were boosted dramatically by the use of PCR-RFLPs, the power of the statistical tests was improved. However, these tests were still somewhat inconclusive. The levels of genetic differentiation among sampling locations were, overall, just bordering on statistical significance. The contingency table analysis of haplotype frequencies showed marginally significant differences among samples ($P = 0.04$), whereas overall F_{ST} was 0.015 which was not significantly

different from zero (P = 0.15). Individual pairwise comparisons showed where the greatest departures from panmixia occurred, but again were largely only marginally statistically significant. The two samples which were most distinct were those from Wallis Lake and Lake Tuggerah. On both haplotype and sequence based F_{ST}'s, they were found to be significantly different from their nearest sample to the north, that of Coffs Harbour. This was primarily due to the higher frequency of the 'E' clade alleles in the two lake samples, compared to other locations (Table 5.16). These alleles had a slightly lower frequency in the Coffs Harbour sample. This indicates that this difference is only bordering on statistical detection. In addition, a "private" allele was found in two individuals from Lake Tuggerah and nowhere else. The pattern of relationships among samples is shown in the net sequence divergences, which are all quite low, and in the MDS pattern and NJ tree. These tend to show once again that there is no clear genetic geographic grouping of the samples, but that there is a continuum of relationships among them, with Wallis L. and L. Tuggerah at one end (being the only real geographic grouping) and Coffs Harbour at the other. Overall, these results suggest that these lake populations may have a slightly different genetic composition, perhaps due to some level of reduced dispersal from the lake. However, unless there is some likelihood of possible maturation and spawning of this species within the lakes or in oceanic waters nearby, these results may also be an artifact.

A similar situation occurs with the most northern sample, from the Swain Reefs. Depending on the test, this population is marginally statistically significantly different from its nearest samples of Moreton Island and Moreton Bay. This appears to be due largely to the presence of the D allele in three individuals. This may indicate some level of genetic subdivision, however this allele is also found in one individual of the most southern sample in Mallacoota.

The final conclusions from this study regarding possible population subdivision and stock structure in *P. plebejus* are that it is clear that there are no large genetic differences between samples from any of the locations sampled. It is possible that there are small genetic differences among some locations, but, given the levels and patterns of variation found, and the time required to analyse that variation, no firm conclusions could be drawn about these small differences. It would require a larger number of individuals to be analysed to determine more accurately if such small differences actually exist. If this was deemed of importance, this could now be done very quickly, as all the tools and techniques are in place, as well as a considerable number of additional samples for which there was insufficient time to analyse genetically.

6.3.2 *P. esculentus*

The patterns of genetic variation shown in the 16s rDNA are far more clear than those shown in *P. plebejus*. It is immediately apparent, despite the relatively low levels of variation detected, that there are large genetic differences between the east coast and west coast samples, but much smaller differences, if any, among the east coast samples. This highlights the value of including additional samples beyond the region of most interest, so as to be able to gain a better comparative perspective of the levels of among-population genetic variation. It can be seen that the Western Australian sample (from Shark Bay) has only a mitochondrial haplotype (E) that is not found on the east coast, and which is also genetically distinct from the east coast haplotypes. The overall F_{ST} values thus show that, of all the genetic variation found in the species using this marker, only a small proportion distinguishes samples from the east coast (0.3-5%, depending on whether sequence differences are accounted for), whereas 46-54% is partitioned among the east and west coasts. Pairwise comparisons indicate that one F_{ST} value between two east coast samples is significant, that of 0.25 (sequence-based) between the Moreton Bay and Mackay samples, which has a probability of

P = 0.029. However, this is not a sufficient difference to distinguish either of these samples as distinct stocks. It can be seen that the net sequence divergences among the east coast samples are still quite low (zero - 0.1%). The NJ tree shows that the Shark Bay sample is dramatically different from the other samples, much more so than any of the east coast samples are different from each other, and that the Moreton Bay and Mackay samples are at either end of the spectrum of differences among east coast samples, which does not provide any firm understanding of these differences in terms of stock structure.

The COI sequencing results are largely in concordance with those from 16s. Once again the WA sample is extremely different genetically from the east coast samples, possessing unique, and genetically divergent mitochondrial haplotypes. Among the east coast samples, however, there is not a clear-cut pattern of genetic differences. There are no overall significant genetic differences among samples when analysed either by haplotype frequency contingency table or by haplotype or sequence F_{ST}'s. In examining the pairwise differences between samples, the only east coast comparison that is significant is that between Mackay and the Torres Strait samples. When the pattern of relationships is examined in the NJ tree, it can be seen that the Mackay sample seems to be the most divergent of the east coast samples (but that this difference is substantially less than the observed differences with the Shark Bay sample). However, the fact that it is significantly divergent only from the most distant of the other east coast samples, and that the pattern of relationships makes little geographic sense, suggests that it may not be a real result. the haplotype frequency table shows that the difference is largely due to the occurrence of three G haplotypes in Mackay, and this may simply be a sampling artifact.

The difficulties with sample numbers is alleviated considerably in the COI PCR-RFLP data set. A much larger number of individuals were analysed, and this substantially helps the statistical analyses. There are highly significant genetic differences between the WA and eastern samples. However, once again, there is not a clear pattern of relationships evident among the eastern samples, as many of the genetic differences among samples are just bordering on the level of significance. In the pairwise comparisons, the two eastern samples which appear significantly different from the others are from Groote and the Torres Strait. However, examining the haplotype frequency table, and the NJ tree, show that these two samples do not form a distinct group at all, but in fact lie at the ends of the spectrum of genetic composition exhibited among these samples. While the Groote sample has a high frequency of the G' haplotype and a low frequency of the K haplotype, the Torres Strait sample is characterised by a low frequency of the G' and a high frequency of the K. Thus, at the most, it appears that *P. esculentus* from these two regions may each be somewhat distinct from other eastern samples of these species. The other east coast samples from Cairns, Mackay and Moreton Bay do not appear to be at all significantly genetically distinct. there is no evidence from this data set to support the tentative result from the COI sequencing data set that the Mackay sample is significantly different from the others, suggesting that this was a sampling artifact. However, it is apparent that larger sample sizes would undoubtedly improve the ability to discern patterns of relationships among these east coast samples.

6.3.3 Stock Structure Conclusions

It is worthwhile to summarise here the patterns of stock structure found in the two target species.

In *P. plebejus*, the distribution of genetic variation was relatively uniform throughout the samples collected from the east coast. There was a large amount of variation detected overall, which should have provided sufficient polymorphism to make the tests powerful enough to detect very significant

stock differences with only small sample sizes. However, no samples were clearly different genetically from the others, although there were suggestions that the lake populations from Wallis L. and L. Tuggerah may be marginally different. These differences would require confirmation with larger sample sizes to be sure of their true nature.

In *P. esculentus* , much greater genetic variation was detected among samples, but this was almost exclusively due to the inclusion of an outlying population from WA for comparison. This valuable comparison certainly indicates the levels of differentiation that could be expected from stocks of penaeid species if they had been completely isolated from each other, with no migration for a considerable time. The results show that the levels of differentiation are many times greater than those found among the eastern samples. Even among the eastern samples, though, there was greater evidence of stock differentiation in *P. esculentus* than in *P. plebejus*. Both the Groote and Torres Strait samples appeared to each be significantly differentiated from other eastern populations. However, once again, the true meaning of these differences, as well as the smaller ones between other east coast populations, can not be unequivocally determined without the analysis of larger sample sizes, due to the patterns of variation, and the statistically marginal levels of divergence.

Although it appears that few definitive conclusions can be made about the stock structure of east coast populations of either *P. plebejus* or *P. esculentus* , this is largely due to the fact that it is apparent that there must be at least moderate levels of gene flow among all the east coast populations of these species which is preventing substantial divergence of mtDNA haplotype frequencies. Thus the very strong conclusion that can be made from these studies is that there are definitely no major barriers to gene flow or effective migration among east coast stocks of either of these two species. The DNA analyses, unlike previous allozyme studies, have undoubtedly provided enough genetic variation as raw material for the analyses so that this conclusion of no major stock differences in either species can be made with confidence. However, this is not the entire answer to the question of stock differentiation in these species. Major genetic differentiation among stocks can occur only when gene flow (or effective migration, i.e., individuals both migrating from one location to another and effectively breeding there) is very low or absent. Even low levels of effective migration among populations can reduce genetic differentiation among those populations to a level that makes it difficult to detect without large sample sizes. For example, the minimum levels of migration between locations per generation that would account for the observed levels of east coast population differentiation in *P. plebejus* and *P. esculentus* may be as low as 2.1 and 1.5 respectively. That is, even though little or no east coast stock discrimination has been detected in these species genetically, it may still mean that only small numbers of individuals are moving from one region to another each year. Taking into consideration the very large population sizes of each of these species, this means that the percentage of each stock that could be migrating to result in the observed patterns of genetic variation could still easily be well below 1% of the total. In genetic terms, this is still considerable migration. In fisheries management terms, this may be totally insignificant migration. Genetic techniques can most easily detect where there are major restrictions to migration, such that only a few individuals or less per generation move between stocks. This can usually be accomplished with small sample sizes, and in a relatively short time, if suitably variable genetic markers are available. Unfortunately for crustacean, and particularly penaeid prawn, fisheries, these markers have previously not been available. As such, even the most basic of stock structure questions (i.e., are there major restrictions to migration resulting in highly differentiated stocks?) have previously been unanswerable. This study has certainly provided the tools and results to answer these basic questions for *P. plebejus* and *P. esculentus* , with the unequivocal conclusion that no, there are no major restrictions to migration along the east coast in these species. Now, in order to pursue more fine-scale questions about whether minor restrictions to migration occur among some east coast populations, it is apparent that additional data is required. It was reasonable

to expect that some progress could be made towards these goals in this study, and indeed, potential minor restrictions to migration have been identified in both species. However, the nature and patterns of variation that have been uncovered in these species (in particular, the large numbers of rare haplotypes which are spread widely throughout the species' distributions), has unfortunately meant that the task of minor stock discrimination in these species is not at all easy. What this means is that large sample sizes will be required in order to adequately answer these questions. If this is required in the future, the existing patterns of variation can be used to give good estimates of the sample sizes required to answer such questions. In addition, if it is important that the precautionary principle be enforced, (i.e., it is considered important enough from a sustainable management viewpoint to ensure that distinctly different stocks are not ignored simply due to the stringency of statistical testing), it would also be important to evaluate the power of the statistical tests employed, and if necessary, adjust upwards the critical alpha significance levels. Such considerations are often found to be important in other fisheries under management. Ultimately such decisions lie in the hands of the fisheries managers. This study has provided managers with answers to the basic stock discrimination questions in these species, as well as the basic tools necessary for answering additional, more fine-scale, questions in the future.

7. Benefits

A number of practical benefits emanate from this research:

- The identification along the east coast of single unit stocks of the two species examined.
  - That is, the determination that, on genetic criteria at least, no east coast fishing areas require separate management.

- The quantification of moderate to large minimum potential rates of migration between fishing regions.
  - That is, the determination that no fishing areas appear genetically distinct in any way, and that prawn migration between them is likely to be moderate to large in genetic terms (more than 2 individuals per generation), although not necessarily large in management terms.

- The improved, more efficient management of the nation's prawn resources, enabling the maximum yield sustainable in the long term.
  For example:
  - there appear to be no previously unidentified distinct stocks which may be vulnerable to over-fishing. Thus, there are no genetic grounds for protecting individual fishing areas on a zonal or regional basis.
  - potential fishing closures to areas may be found to be unnecessary on genetic grounds, due to the mixing of stocks being apparently adequate to maintain genetic homogeneity.
  - the identification of single east coast unit stocks will permit more accurate determination of stock-recruitment relationships.

- The provision of a suite of refined genetic techniques which can be readily applied to examining the stock structure of additional prawn or crustacean species, and also to examining genetic variation and identifying individuals and families in cultured prawns.

The research will most directly benefit fisheries managers in Queensland, by providing them with detailed information about the genetic stock structure of the species examined, allowing them to make sound management decisions. The expected long-term impacts of this research will be to ensure the continued conservation of Australia's valuable prawn genetic resources, and to assist managers in maintaining the maximum sustainable harvest of those resources.

The study's results will also prove extremely valuable to other researchers, by providing an up-to-date assessment of the best genetic techniques with which to undertake further research on Australia's commercial penaeid prawns. A number of these species support valuable fisheries throughout Australia, and there is considerable scope and need for additional detailed research to be undertaken on these species in each fishery.

8. Further Development

This research has identified and developed a number of DNA analysis techniques for examining genetic variation in prawns and crustaceans, and has used these techniques to look for major genetic discontinuities among populations of *P. plebejus* and *P. esculentus* . However the amounts, types and patterns of variation detected in these species (particularly the large number of dispersed rare haplotypes) has meant that it is only major genetic differences that could be detected using moderate sample sizes of individuals. The results have shown that no major genetic differences among populations have been detected along the east coast for either species. However, the results have suggested that minor differences may occur in some populations (e.g., the Wallis Lake and Lake Tuggerah samples of *P. plebejus* , and the Torres Strait sample of *P. esculentus*). To confirm the existence of such potential minor genetic differences among populations would require additional analyses of larger sample sizes. It needs to be remembered that even minor genetic differences among stocks would represent dramatic restrictions to migration and movement of individuals in fisheries terms, as genetic techniques are conservative estimators of stock subdivision.

If such additional analyses were required, then, because the techniques have now been so well developed, it would be a relatively easy task to accomplish, particularly as additional samples are already in hand, and as the mitochondrial control region and microsatellite markers have now also been developed. Furthermore, particularly for *P. esculentus*, only a relatively small number of sample locations could be included in these analyses. Therefore it may be potentially very useful to examine more populations of this species to ensure that cryptic stock subdivisions along the east coast distribution have not been overlooked.

9. Conclusion

Objectives and Outcomes:

- Assess the utility of a range of DNA techniques for examining genetic variation within species of penaeid prawn.

The DNA techniques of mitochondrial DNA (mtDNA) sequencing (examining 12s rDNA, 16s rDNA, the cytochrome oxidase I (COI) gene and the control region), mtDNA PCR-RFLP's (polymerase chain reaction - restriction fragment length polymorphisms) and microsatellite DNA VNTR's (variable number of tandem repeats) have proved extremely useful in detecting large amounts of genetic variation in species of penaeid prawn. This is in strong contrast to previous allozyme methods of detecting genetic variation in these species.

- Refine those techniques for examining east coast prawn stocks.

All these techniques have been refined to allow rapid screening of genetic variation in populations of *P. plebejus* and *P. esculentus*. Unfortunately, the development of techniques for examining the mitochondrial control region and microsatellite DNA proved problematic, and was not completed until near the end of the project, precluding their use in detailed population analyses.

- Describe the genetic stock structure of two commercial species (*Penaeus esculentus* and *P. plebejus*) throughout their East Coast distribution.

P. plebejus appears to have a relatively panmictic population structure along the east coast of Qld. After examining genetic variation in 295 individuals from 13 locations throughout the species' distribution, considerable variation was found within populations, but little difference was detected among populations. The only population samples that may potentially be genetically different from the others were those from Wallis Lake and Lake Tuggerah. However, these differences were only minor, and would require further verification using larger sample sizes.

P. esculentus appears to have considerably greater population structure throughout its Australian distribution, but very little along its east coast distribution. After examining genetic variation in 175 individuals from 10 locations throughout the species' distribution (all but two sites in Qld waters), considerable variation was found within populations, but, once again, little difference was detected among east coast populations, although it did appear to be greater than that found in *P. plebejus* . The Torres Strait sample may potentially be genetically different from the others. However, once again, this difference was only minor, and would require further verification using larger sample sizes. In contrast, the outlying Western Australian sample (from Shark Bay), which was used for comparison, is substantially genetically differentiated from all east coast populations.

- Determine whether genetically distinct stocks of these species exist in the east coast fishery and, if so, locate the boundaries between such stocks.

It therefore appears that genetically distinct stocks of prawn of either species do not exist along the Qld east coast. However, the nature of the genetic variation detected in these species has meant that

only major genetic differences could be easily detected using moderate sample sizes of individuals. To confirm that no minor genetic differences occur among populations would require additional analyses of larger samples.

- Quantify the potential rates of effective prawn migration between regions for each species.

The observed lack of genetic subdivision along the east coast does not necessarily suggest that the proportions of the stock of either species that successfully migrate and breed are a considerable proportion of the species, but may merely mean that there is a small proportion that does so. Because of the very large population sizes, this proportion could easily be less than 1%, but is difficult to quantify more accurately. Genetic techniques are best suited to estimating levels of migration that are lower than what appears to be occurring among east coast prawn populations.

- Provide specific advice on the management implications of the results.

These results suggest that there are no genetic reasons why the east coast fisheries of both species should not be managed as unit stocks. In the case of *P. plebejus* , this means that the fisheries in both NSW and Qld are likely to be very interdependent, and largely sharing the same resource. In the case of *P esculentus*, this means that there appears, on this evidence at least, to be no genetic justification for segregating the stock into different management units. Although these are "negative" results, they are still of considerable importance to the fisheries.

The high levels of genetic variation detected in these species, and the refinement of the techniques for doing so, provides a very important base for further genetic research in these and other species of penaeid prawn, be that in stock discrimination on a larger scale, or in the rapidly-expanding field of prawn aquaculture.

10. References

1. Mulley, J. & Latter, B. (1981a) Geographic differentiation of eastern Australian Penaeid prawn populations. *Aust J. Mar. Freshw. Res.* 32:889-95.
2. Mulley, J. & Latter, B. (1981b) Geographic differentiation of tropical Australian penaeid prawn populations. *Aust. J. Mar. Freshw. Res.* 32:897-906.
3. Dall, W. *et al.* (1990) The Biology of the Penaeidae. *Adv. Mar. Biol.* 27.
4. Dall, W. (1985) A review of penaeid prawn biological research in Australia. *In* P. Rothlisberg et al. (eds.) *2nd Australian National Prawn Seminar.* NPS2, Cleveland Aust.
5. Salini, J. *et al.* (1986) An electrophoretic analysis of stock structure in the ornate tropical rock lobster, Panulirus ornatus. *In* A. Haines *et al.*(eds) *Torres Strait Fisheries Seminar.* Aust. Govt. Publ. Serv., Canberra. 218-232.
6. Agerberg, A. (1990) Genetic variation in three species of freshwater crayfish; *Astacus astacus*, *A. leptodactylus* and *Pacifastacus leniusculus*, revealed by isozyme electrophoresis. *Hereditas* 113:101-108.
7. Sunden, S. and Davis, S. (1991) Evaluation of genetic variation in a domestic population of *Penaeus vannamei*: a comparison with three natural populations. *Aquaculture* 97:131-142.
8. Lavery, S. (1993) Low allozyme variation in the coconut crab *Birgus latro. Comp. Biochem. Physiol.* 104B: 353-359.
9. Tam, Y.K. and Chu, K.H. (1993) Electrophoretic study on the phylogenetic relationships of some species of *Penaeus* and *Metapenaeus* (Decapoda: Penaeidae) from the South China Sea. *J. Crust. Biol.* 13: 697-705.
10. Moritz, C., Dowling, T. & Brown, W. (1987) Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* 18:269-92.
11. Lavery, S. (1991) Genetic population subdivision in the coconut crab, Birgus latro. *Mem. Qld Mus.* 31:210.
12. Brasher, D.J., Ovenden, J.R., and White, R.W. (1992) Mitochondrial DNA variation and phylogenetic relationships of *Jasus* spp. (Decapoda: Palinuridae) *J. Zool. Lond.* 27: 1-16.
13. Benzie, J.A.H., Ballment, E. and Frusher, S. (1993) Genetic structure of *Penaeus monodon* in Australia. *Aquaculture* 111: 1-4.
14. Lavery, S. - unpublished electrophoretic data from preliminary study of population differentiation in CSIRO prawn samples from the Gulf of Carpentaria.
15. Rothlisberg, P.C., Church, J.A. and Fandry, C.B. (in press) A mechanism for nearshore concentration and estuarine recruitment of postlarval *Penaeus plebejus* Hess (Decapoda, Penaeidae). *Estuar. Coast. Shelf Sci.*
16. Lavery, S. - unpublished mtDNA data from present preliminary study.
17. Batuecas, B. *et al.* (1988) Genome organisation of Artemia mtDNA. *Nucl. Acids Res.* 16:6515-6529.
18. Ballard, J.W.O. et al. (1992) Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. *Science* 258: 1345-1348.
19. Cunningham, C., Blackstone, N. & Buss, L. (1992) Evolution of king crabs from hermit crab ancestors. Nature 355:539-542.
20. Palumbi, S. R. and Benzie, J. (1991) Large mitochondrial DNA differences between morphologically similar Penaeid shrimp. *Molec. Mar. Biol. Biotech.* 1: 27-34.
21. Williams, J.G.K. *et al.* (1991) DNA polymorphisms amplified by arbitrary primers are useful genetic markers. *Nucl. Acids Res.* 18: 6531-6535.
22. Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic markers. *Nucl. Acids Res.* 17: 6463-6471.

11. Appendixes

11.1 *Appendix 1: Valuable Information Arising from the Research*

Prawn Samples

- Tissue samples from 661 *P. plebejus* individuals from 13 locations throughout the species' distribution
- Tissue samples from 211 *P. esculentus* individuals from 10 locations throughout the species' distribution
- DNA extracts from 295 *P. plebejus* individuals from 13 locations throughout the species' distribution
- DNA extracts from 175 *P. esculentus* individuals from 10 locations throughout the species' distribution
-

DNA Techniques for Prawn Analyses:

- Optimised sample storage protocols for DNA extraction
- Optimised DNA extraction methods
- Optimised mtDNA PCR amplification protocols
- Penaeid specific primers for amplification of mt COI gene and Control Region
- Series of identified PCR-RFLP markers for rapidly analysing variation in 16s rDNA and COI for *P. plebejus* and *P. esculentus*
- Clone isolation, DNA sequences, PCR primers and amplification protocols for *P. esculentus* microsatellites

Analyses of DNA Variation in Prawns:

- DNA sequences of mt 12s rDNA gene, 16s rDNA gene, COI gene and control region for both *P. plebejus* and *P. esculentus*
- DNA sequences of 10 nuclear microsatellites from *P. esculentus*
- Measures of DNA variation within and among all population samples of *P. plebejus* and *P. esculentus* for two genes

11.2 Appendix 2: Staff Engaged on the Project

NAME	*POSITION*	*CONTRIBUTION*	*TIME*
Dr Shane Lavery	ARC Postdoctoral Research Fellow	Principal Investigator	100%
Elizabeth Gaskin	Temp. Laboratory Technicians	Technical laboratory assistance	25%
Anna Murrell			10%
Bradley Smith			25%
Dr Clive Keenan	Senior Fisheries Biologist	Allozyme genetics	5%
Dr Ken Reed	Director, QABC	Biotechnology applications	2%
Dr Greg Smith	Principal Biotechnologist	DNA techniques	5%
Ms Darelle Thomson	Senior Technical Officer	Laboratory assistance	10%
Mr Dave Trama	Master	Master, R.V. Deep Tempest	10%
Mr Peter Pardoe	Boatman	Deckhand, R.V. Deep Tempest	10%