[An image (superior view of a mud crab, *Scylla serrata*) has been removed from page ii of the digital version of this thesis for copyright reasons.]

Genetic structure within the distribution of the Indo-West Pacific mud crab *Scylla serrata* (Forskål, 1775).

David Gopurenko B.Sc.

Australian School of Environmental Studies
Faculty of Environmental Sciences
Griffith University

Submitted in fulfilment of the requirement of the degree of Doctor of Philosophy

October 2002

"When they had satisfied their hunger and thirst, their thoughts returned to their dear comrades whom Scylla had snatched from the hollow ship and devoured; and they wept till soothing sleep overtook them."

The Odyssey. Homer

[IMAGE REMOVED]

***Scylla serrata* (Forskål, 1775)**

(Image source: Keenan *et al*. 1998)

Summary

It is often hypothesised that marine species with mobile planktonic phases are capable of widespread dispersal and may therefore be genetically homogenous throughout their distribution. Studies that have demonstrated positive correlation between duration of plankton phase and levels of gene flow reinforce the prediction that life history characteristics of marine species determine the potential extent of genetic and demographic connectivity throughout their distributions. This prediction has however been challenged by studies that have employed genetic markers highly sensitive to both historical and contemporary demographic changes. Disparities between dispersal potential and measured levels of gene flow have been demonstrated both among historically disconnected ocean basins and within semi-enclosed areas of strong hydraulic connectivity. These studies and others highlight a need for greater focus on factors that may influence population structure and distribution for marine species.

In this thesis, I have examined genetic structure within and among populations of an estuarine species of mud crab *Scylla serrata* (Forskål, 1775) using a number of genetic markers and methods. The species is widely distributed throughout mangrove and estuarine habitats of the Indo - West Pacific (IWP); it is generally assumed that life-history characteristics of *S. serrata* promote high levels of population admixture and gene flow throughout its distribution. Alternatively, factors that have promoted population genetic structure for a variety of IWP marine species may also have affected *S. serrata* populations. By investigating genetic structure at several spatial scales of sampling, I was able to address a variety of hypotheses concerning the species distribution, dispersal, and genetic structure.

Episodic changes to marine habitat and conditions experienced within the IWP during the Pleistocene may have affected genetic structure for a broad variety of marine taxa. The relative strength of this hypothesis may be assayed by comparative genetic studies of widespread IWP taxa with high dispersal capacity. In order to ascertain levels of historical and contemporary gene flow for *S. serrata*, I investigated the phylogeographic distribution of mitochondrial DNA haplotypes sampled throughout the species range. Adults were sampled from three west Indian ocean locations (N=21), six west Pacific sites (N=68), and two sites from northern eastern Australia (N=35). Temperature gradient gel electrophoresis and sequencing of 549 base pairs of the mitochondrial cytochrome oxidase I (COI) coding gene identified 18 distinct haplotypes. Apart from

that seen in northern Australia, haplotype diversity was low ($h < 0.36$) at each of the locations. Total nucleotide diversity in the entire sample (excluding northern Australian locations) was also low ($\pi = 0.09$). Haplotypes clustered into two clades separated by approximately 2% sequence divergence. One clade was widespread throughout the IWP (clade 1) whereas the other was strictly confined to northern Australia (clade 2). Genealogical assessment of sequenced haplotypes relative to their distributions suggested that a historical radiation of clade 1 *S. serrata* throughout the IWP occurred rapidly and recently (<1Myr bp) from a west Pacific origin. The evidence of fixed unique haplotypes at the majority of locations suggested that contemporary maternal gene flow between trans-oceanic sites was limited. Contrary to reports for other widespread species of IWP taxa, there was no evidence of lengthy periods of regional separation between Indian from Pacific Ocean populations. However, results may indicate a separation of northern Australian crabs from other locations before and during the IWP radiation. I speculated that this isolation might have resulted in the formation of a new species of *Scylla*.

Additional sampling of mud crabs from the Australian coastline allowed an examination of the diversity and distribution of clade 1 and 2 haplotypes among recently formed shelf-connected coastal locations, and across a historical bio-geographic barrier. Over 300 individuals were sampled from multiple locations within coastal regions (western, northern and eastern) of Australia and analysed for mutational differences at the COI gene. Analysis of molecular variance partitioned by sampling scale (Among regions, within regions, and within all locations) indicated mitochondrial haplotypes were structured regionally ($P < 0.001$), which contrasted with evidence of genetic panmixia within regions. Regional genetic structure broadly correlated with hydrological circulation, supporting the contention that release and transport of propagules away from the estuary may allow genetic connectivity among widespread shelf-connected *S. serrata* populations. That similar patterns of maternal gene flow were absent among trans-oceanic populations may indicate that the spatial scale of effective dispersal for this species is generally limited to areas of coastal shelf. The two clades of haplotypes were geographically separated either side of the Torres Strait, a narrow sea channel connecting the northern and eastern regions of coastal Australia. This pattern of historical genetic separation was concordant with a number of other marine species across northern Australia, and might indicate a shared history of vicariance induced by eustasy. Alternatively, differences in diversity and distribution of the clades may be

evidence of two independent expansions of clade 1 and 2 crab populations into Australian regions following post-glacial estuary formation. Overall, despite evidence of genetic panmixia within extensive sections of the Australian distribution, there was also evidence of significant barriers to maternal gene flow with both shallow and deep regional phylogeographic assortment of mtDNA haplotypes. The presence of these barriers indicated both historical and contemporary factors have imposed limits to effective dispersal by this species among coastal habitats.

A subset of the Australian sample (8 locations, $N = 188$) was also examined for variation at five microsatellite loci developed specifically here for *S. serrata*. I examined variation among samples at each of the loci to: a) independently verify regional structure among crab populations previously detected using the mtDNA analysis; b) test for evidence of co-distributed non-interbreeding stocks of *S. serrata* within Australian waters by examining samples for segregation of alleles within microsatellite loci concordant with the two mtDNA clades. The frequency and distribution of alleles for each of the highly polymorphic microsatellites were homogenous at all levels of sample partitioning and contrasted sharply with the instances of both weak and strong regional phylogeographic assortment of mtDNA haplotypes. These contrasting results between different genomic markers were examined in relation to the species life history, and to differences in mutational rate and inheritance of the genetic markers. Several hypotheses may explain the disparity, however it is most likely that rampant homoplasy and high rates of mutation at the microsatellite in conjunction with large *Ne* at locations may be concerted to delay equilibrium between genetic drift and migration among populations at these highly polymorphic nuclear markers. There was also no evidence that alleles at microsatellite loci were co-segregated with mtDNA clades and therefore no evidence of segregated breeding between the clades of crabs. Whether or not this result was also driven by homoplasy at the microsatellites remains unknown.

Recently established mud crab populations (~ 3-4 years old) observed in a number of southwest Australian estuaries are almost 1000 kilometres south of their previously recorded distribution on the Western Australian coast. Colonisation of the southwest region may have occurred either by a natural range expansion from northwest Australian mud crab populations or by means of translocation from any number of mud crab sources within the Indo - West Pacific. I used mtDNA analysis to verify the species

and determine the potential source population(s) of the colonists, by comparing sampled genetic material from the southwest ($N = 32$) against that previously described for the genus. I also compared levels of diversity at mtDNA and two microsatellite loci between the colonist and suspected source population(s) to qualitatively determine if the southwest populations experienced reductions in genetic diversity as a result of the colonisation process. All colonist samples had *S. serrata* mtDNA COI sequences identical to one previously described as both prevalent and endemic to northwest Australia. High levels of genetic diversity among source and colonist populations at two microsatellite loci contrasted to the mitochondrial locus which displayed an absence of variation among colonists compared to moderately diverse source populations. I argued that the southwest was recently colonised by large numbers of *S. serrata* propagules derived from the northwest of Australia, possibly due to an enhanced recruitment event coinciding with the reported strengthening of the Leeuwin Current during 1999. Contrasting levels of diversity among nuclear and mitochondrial loci may be attributed to a difference in response by the two genomes to the colonisation process. I predict that such differences may be generally prevalent among plankton-dispersed species.

Finally, I discuss aspects of the species distribution and biogeography obtained as a composite of the various results and ideas expressed in this thesis. I propose that *S. serrata* populations in the IWP may have experienced several cycles of extinction and population retraction from temperate areas followed by subsequent periods of colonisation and rapid coastal expansion in response to the effects of glacial episodes on coastal habitats in the IWP. I propose that persistence of this species as remnant populations of clade 1 and 2 crabs at equatorial locations during low sea level stands provided source populations for later expansions by the species into a variety of coastal areas throughout the IWP. Further analysis is required to determine if mtDNA clade 1 and 2 crabs are non-interbreeding species of mud crab.

Acknowledgments

Professor Jane Hughes provided considerable levels of support as the principle supervisor in my study. I could not possibly list all those instances, as they were so numerous. More than this, Jane provided critical assessment of my work and motivated me into new areas of thought, for this I am fortunate and most grateful.

Co-supervisors, Clive Keenan and Stuart Bunn provided valuable support and insight during the early stages of this work. Clive provided free access to his crab collections and for this, and his trust, I thank him.

My gratitude also extends to the following people who either directly or indirectly assisted in procuring crab samples used in this study: Ted Allen, Lynda Bellchambers, Rollie Bowman, Rex Etchell, Bill Fisher, Ben Fraser, Misa Gopurenko, Tracy Hay, Don Heales, Sue Helmke, Arvid Hogstrom, Clive Keenan, Bill Kehoe, Chad Lunow, David Perkins, John Salini, Steve Sly, Ilona Stobutzki, Charles Tenakanai, Carolyn Williams. Without the efforts of these people, I would have been stranded. May your crab pots always be full (of crabs…) Special thanks to Tracy Hay from Northern Territory Fisheries for arranging a number of sampling opportunities and keeping my project in mind from time to time.

I thank Derek Kennedy, Jing Ma and Justine Smith for their professional guidance and teaching of laboratory protocols and procedures. In particular, I single out Jing Ma who has skills so far advanced that it may be considered inexplicable. Jing often demonstrated the virtue of persevering with a clear mind…thanks for the patience. To all other members of the laboratory who assisted me, thanks. A big thank you to Mia Hillyer, especially for maintaining an efficient working environment (and the lively discussions concerning methodology).

To the many and varied people who provided critical advise concerning theoretical aspects of my work, I am indebted and enlightened. In particular, I single out Michael Arthur, Stephen Chenoweth, James Holman, Jane Hughes, David Hurwood, Rachel King, Dugald McGlashin, Mark Ponniah, Alicia Schultheis, Jill Shephard and miscellaneous “Broken Head geeks”. A big thank you to Mark Ponniah for convincing me several years ago, to follow up the work on mud crabs.

I also thank Jane Hughes and Sonja Parsonage for proofreading drafts of this thesis; their scrutiny and recommendations were invaluable. Valuable comments concerning chapters or written work within the manuscript were also provided by Stephen Chenoweth, Alicia Schultheis and Jill Shephard. Craig Moritz and Stuart Bunn also provided useful comments on written work during the early stage of this thesis.

I express my deep gratitude to Jill Shephard for providing professional assistance, criticism and insightful advice. I thank you for the strength of your support and friendship.

To family and friends (including Leto), thanks for being aware and understanding – and for the support.

TABLE OF CONTENTS

List of Figures

List of Tables

Declaration

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

David Gopurenko

CHAPTER 1: **General introduction**

1 .1 Formation of population genetic structure

In his influential work of 1963 Mayr recognised that mechanisms of evolution, which stratify genetic diversity at the population level, are crucial in the formation of new phylogenetic lineages and eventually species. Populations then, are the least inclusive unit at which speciation may occur. Therefore, investigation of factors that partition population genetic structure provides insight to evolutionary process.

In the absence of selection, population genetic structure is determined by the quantity and frequency of allelic admixture among constituent demes (gene flow), relative to the rate of stochastic change in allele frequency within demes due to mutation and genetic drift (Slatkin 1987). Gene flow, mediated by migration or dispersal of gametes among populations, homogenises the geographic distribution of alleles whereas genetic drift and mutation lead to divergence. The time required to attain equilibrium in a population between the opposing effects of gene flow and genetic drift/mutation (assuming effects are constant) is correspondingly related to the effective number of individuals in the population contributing to the next generation (*Ne*) and the fraction of migrants per generation (*m*) (Wright 1931). Hence, formation of genetic structure can be directly influenced by demographic variables in a species life history such as population size, reproductive patterns and dispersal potential (Slatkin 1987).

1 .2 Measurement and modelling of population genetic structure

1.2.1 Genetic structure and estimates of population sub-division

Conventional estimates of genetic structure among populations based on proportional measures of shared alleles allow approximations of population sub-division as influenced by gene flow and genetic drift. Wright (1965, 1978) introduced measures (F-statistics) of sub-division based on the correlation of alleles within individuals that describe levels of inbreeding partitioned within and among samples. These F-statistics describe relatedness of individuals within sub-populations (F_{IS}), among sub-populations (F_{ST}) and throughout the entire population (F_{IT}). In the absence of gene flow, genetic drift and mutation lead to differentiation of allelic frequency and eventual fixation of

exclusive alleles in different sub-populations; and the fixation index among sub-populations (F_{ST}) approaches a maximum at one. Under conditions of strong gene flow with very high migration rates among subpopulations (as measured by the product of *Ne* and *m*), allelic frequencies among sub-populations become homogeneous and F_{ST} approaches zero.

1.2.2 Models of dispersal and population genetic structure

There is a considerable body of evidence indicating various forms of dispersal and or migration employed by species affects the spatial scale and distribution of genetic structure among populations in manifestly different ways (Slatkin and Barton 1989). The island model (Wright 1931) assumes an equal likelihood of migration among populations regardless of their spatial proximity to one another. Under this model, gene flow will occur at an equal rate among populations. In contrast, under a stepping-stone model (adapted from Wright's 1943 isolation by distance concept) migration occurs at a constant rate but only between neighbouring populations distributed in either one or two-dimensional arrays (Kimura and Weiss 1964). At equilibrium under a stepping-stone model, it is expected that dispersal (and therefore transfer of alleles) is more likely to occur among adjacent rather than distant populations. As a result, levels of genetic structure among populations may be positively correlated with geographic distance between populations (Slatkin 1993). Both island and stepping-stone models of migration employ restrictive assumptions (ie: constant and equal sized *Ne* among populations, symmetrical migration) that are seldom met in the real world. Nonetheless, specific knowledge of conditions and processes that affect a species distribution can lead to the generation of testable hypotheses of population structure expected from defined models of migration (Lavery *et al*. 1996-a, Chenoweth and Hughes 1998-a).

1.2.3 The use of varied markers for analysing population genetic structure

The rate of approach to genetic equilibrium can vary among classes of genetic loci due to differing mechanisms of mutation and modes of inheritance. As a result, patterns of population genetic structure detected using loci subject to various modes of evolution can differ (Wade *et al*. 1994). This is particularly apparent among genetic markers derived from haploid cytoplasmic and diploid nuclear systems. Because of its maternal mode of transmission, mitochondrial DNA (mtDNA) has an effective population size

one quarter that of the nuclear genome (Birky *et al*. 1989). Correspondingly, the rate of approach to genetic equilibrium occurs rapidly and amounts of among deme divergence due solely to genetic drift can be as much as four-fold greater for mtDNA than for nuclear loci (Birky *et al*. 1989). This feature of mtDNA makes it a particularly sensitive single locus for detecting genetic structure among populations.

Microsatellites are one class of diploid nuclear marker that provide fundamentally different but comparable information to mtDNA, for estimating population structure. Excepting conditions of functional or selective constraints, bi-parentally inherited neutrally evolving microsatellites generally have high rates of mutation resulting in rapid accumulation of co-dominant alleles within populations (Bowcock *et al*. 1994). The rapid turnover of new alleles at microsatellite loci make them susceptible within populations to shifts in allelic frequency induced by stochastic and demographic processes.

It can be expected that detected patterns of population genetic variation among microsatellite and mtDNA data sets may differ according to the response of each independent genome to the combined effects of mutation, migration, and effective population size. By exploring these differences, comparative multi genome surveys provide a means for testing competing hypotheses of population structure not otherwise possible using single loci (Hey and Harris 1999). This approach has been useful for identifying instances of sex-biased dispersal (Karl *et al*. 1992), testing hypotheses of historical demographic change (Avise 1995), estimating effective population size (Slade *et al*. 1998) and gene flow among populations and congenerics (Morrow *et al*. 2000).

1.2.4 Phylogeographic association of mtDNA haplotypes

As used for nuclear loci, the frequency of mtDNA alleles (haplotypes) assayed from a species distribution may be used to infer levels of gene flow between populations. MtDNA also provides a type of information concerning population structure that is generally unavailable from nuclear derived markers. Due to its maternal mode of inheritance and lack of recombination, mutational derivatives of mtDNA haplotypes maintain relationships of identity by descent. Hence, the variety of mtDNA haplotypes among individuals of a population is a reflection of not just allelic diversity but also the genealogical association of maternal lineages within the population. Because of this

facility to infer relationships of descent, mtDNA analysis provides a powerful means of reconstructing the phylogenetic history of matrilines in a population relative to their geographic distribution (Avise *et al*. 1987). This "phylogeographic" (*sensu* Avise *et al*. 1987) facility therefore provides a means of detecting historical as well as contemporary processes shaping population structure.

Avise *et al*. (1987) predicted that the genetic architectures of extant species could be categorized using mtDNA analysis as a series of five distinct phylogeographic outcomes. Each of these categories is defined by the extent of geographical spread of haplotype lineages relative to their genealogical association to one another and determined by extrinsic and intrinsic evolutionary processes that have affected populations over time. The large number of empirical studies that comply with these categorical outcomes appears to confirm this prediction (review Avise 2000). Methods that incorporate cladistic or genealogical analysis of geographically structured mtDNA haplotypes operate from the assumption that the relative age of each haplotype in a population (as determined by a measure of genealogical association) is intrinsically informative for separating historical from contemporary demographic events. Neigel *et al*. (1991) observed that in many cases where mtDNA was used for analysis of population structure, closely related haplotypes were often geographically clustered. A model was developed by Neigel *et al*., which predicted that the variance of a lineage's geographic distribution is proportional to its relative age (assuming constant dispersal over time). In this model, older haplotypes in a population are expected to be more geographically widespread than derived haplotypes as they have had more time to disperse from an initial point of origin.

Perhaps the simplest method of dealing with this type of data was that developed by Excoffier *et al*. (1992) to calculate levels of molecular variance among pre-determined hierarchically grouped population samples. This approach (referred to as AMOVA), which is based on analysing the proportional variance of haplotype frequencies within sample groups, can also incorporate the amount of divergence among haplotypes as a proxy for the time to coalescence between haplotype pairs within a group. AMOVA is very effective for testing among population models where there is some *a-priori* knowledge or expectation of geographical/ historical structure.

1 .3 Population structure among plankton dispersed marine species

1.3.1 The spatial scale of distribution and connectivity among marine populations

Species are geographically distributed primarily by their habitat requirements, environmental tolerances, life history characteristics, and ecological interactions. Colonisation of vacant patches is a key process by which a species may expand its geographical range (Levins 1969 & 1970). This process is particularly apparent among coastal marine species that maintain widespread populations through migratory or dispersive phases in their life history (Underwood and Fairweather 1989). For some species, colonisation of new habitat can result from the successful transport and establishment of propagules outside of their immediate range. Alternatively, range expansion by nearest neighbour (stepping - stone) dispersal may result in the progressive colonisation of vacant patches from adjacent source populations (MacArthur and Wilson 1967). In both instances, the persistence of a colonised patch is either dependent on source patches for renewal or is capable of self-recruitment beyond some local extinction threshold (Hanski 1982, Harrison 1991).

Areas impervious to demographic expansion, or range boundaries, define limits to the distribution of a species. Persistent range boundaries concordant among groups of co-distributed species mark junctions between biogeographical provinces and are apparent in both terrestrial and marine systems (Pielou 1979, Briggs 1995, Veron 1995). In the absence of any defined barriers, it is often assumed that populations of marine species can be highly connected within provinces due to the interaction of evolved life histories with broad-scale oceanographic processes (Scheltema 1971). Given that many marine species have evolved freely dispersed plankton phases, there is strong potential for physical processes to affect distribution patterns through their influence on the process of recruitment (Veron 1995, Gaylord and Gaines 2000). The extent to which this is a general phenomenon among plankton-dispersed species is contentious.

1.3.2 Historical fragmentation of marine populations

Range expansions and retractions of a species distribution, following historical climate shifts, can result in remnant populations scattered over a considerable area (Fleminger 1986, Dynesius and Jansson 2000, Hewitt 2000). Fragmentation of a species distribution may result if the extent of distance between remnants exceeds the dispersal capacity of a species (Palumbi 1994). Fragmentation may also occur by historical

disruption of dispersal routes by emergent physical barriers, resulting in historical separation (vicariance) of prior continuous populations (Avise 1994). Hypotheses of vicariance allow tests of comparative biogeography by proposing that similar geographic disjunct distributions observed for species with differing life-histories, occur by convergent historical range partitioning (Rosen 1978, Avise 1994). As an example, McCoy and Heck (1976) proposed that the co-distribution of distinct assemblages of taxa comprising the shallow water marine habitats (corals, seagrasses and mangroves) are more likely a consequence of historical modification and tectonic separation of previously widespread biotic assemblages rather than the accumulation of genera within such assemblages by numerous identical dispersal events.

1.3.3 Planktonic dispersal and population connectivity

Larval phases within the life history of many marine species develop as freely dispersed plankton. Release of larvae as plankton allows dispersal of the larvae to habitats beneficial to their development and so provides opportunities for recruitment to intermittently distributed populations and habitats. Transport of plankton entrained in current systems may potentially overcome great distances separating populations and therefore maintain continuity of population structure over vast geographic scales akin to an island model of migration (Scheltema 1971, Palumbi and Wilson 1990, Gaylord and Gaines 2000).

Inferential evidence of the potentially large scale of dispersal effected by plankton comes from observations that the majority of coastal marine species that have colonized isolated oceanic islands generally have dispersive plankton phases (Kay and Palumbi 1987). Direct evidence of extensive plankton dispersal is apparent by the rapid rate at which some species spread after colonization or introduction to new areas. Advection of invasive mussel larvae (*Perna perna*) into prevailing long-shore coastal currents from an initial site of translocation, subsequently resulted in colonisation of this species throughout large areas of the Gulf of Mexico over a four-year period (Holland 2001).

Although planktonic dispersal is thought of primarily as an agent for long-range dispersal and population continuity, it can for some species, potentially be limited between populations via retention of locally produced propagules (McConaugha 1992, Strathmann *et al*. 2002). There is a range of evidence indicating that physical

oceanographic and biological processes can limit planktonic dispersal resulting in population structures similar to that expected for stepping stone or isolation by distance models (Hellberg 1994). Entrainment and dispersal of plankton as passive particles in marine currents may be limited by the nature of the physical processes involved and the time spent by a species in the plankton phase (Day and McEdward 1984). Numerous species actively control their dispersal during the plankton phase by behaving in ways that enhance or minimise opportunities for entrainment into currents (Sastry 1983, Day and McEdward 1984). Depending upon their behavioural abilities and developmental requirements, plankton may either be retained in close proximity to source adult habitats, or exported elsewhere (Burton and Feldman 1982, McConaugha 1992, Carr and Reed 1993). In the former case, replenishment of adult populations results from recruitment of locally produced propagules. In the latter case, the potential for transfer and mixing of propagules derived from varied sources results in settlement into adult habitats by recruits not locally produced (Palmer and Strathman, 1981). These two opposing strategies represent the extremes of a range of possibilities employed by different species as adaptations for retention or broadcast (McConaugha 1992).

1 .4 Genetic structure among populations of plankton dispersed species

1.4.1 The paradigm of genetic connectivity via planktonic dispersal

Marine species with high dispersal potential at some or all stages of their life cycle are generally predicted to have strong genetic connectivity throughout their ranges similar to that expected for an island model of migration (Palumbi 1994). As argued in the preceding section, this may be particularly so for species with dispersive plankton phases. Positive correlation between the duration of a planktonic larval phase and levels of gene flow have been demonstrated for a broad range of marine taxa (Waples 1987, Williams and Benzie 1993, Doherty *et al*. 1995, Hellberg 1996), and so reinforce predictions that planktonic dispersal promotes genetic and demographic connectivity among populations (Scheltema 1971, Crisp 1978).

Results contrary to the predicted positive correlation between plankton duration and dispersal of have also been reported, highlighting a need for greater focus on factors influencing the distribution and dispersal of plankton (Shulman and Bermingham 1995).

Significant genetic structure among populations of marine species with high dispersal potential have been correlated with oceanographic features (Arnaud *et al*. 1999) instances of patchy recruitment (Johnson *et al*. 1993), self recruitment in conjunction with isolation by distance (Planes and Fauvelot 2002) and in some cases the role of selection on recruits (Koehn *et al*. 1980).

Similarly, species predicted to have poor dispersal potential via a short or absent planktonic phase, occasionally display unexpected levels of genetic homogeneity. This is particularly the case for a number of low dispersal species inhabiting deep-sea hydrothermal vent habitats, which paradoxically show genetic similarity among habitats separated by considerable distances. It has been argued that genetic homogeneity among populations of these species is maintained by a process of ongoing and frequent extinction - recolonisation of sub-populations within loosely connected meta-populations (Jollivet *et al*. 1999).

Limits to contemporary dispersal leading to reduction in contemporary gene flow among marine populations have been observed across semi-permeable barriers such as the complex of islands and associated current systems within the Indo-Malay archipelago (Duda and Palumbi 1999, Barber *et al*. 2000). As well as restricting contemporary gene flow, these barriers have also resulted in historical separations of high gene flow species into disconnected ocean basins.

1.4.2 Historical episodes of gene flow and fragmentation among marine populations within the Indo-West Pacific region

The effects of episodic climate change on the distribution of marine biota within species rich areas of the Indo-West Pacific region (IWP) warrants attention due to a number of emergent patterns of concordant genetic structure among co-distributed taxa within the region (Benzie 1999, Hewitt 2000). Over the last 1-3 million years, abrupt changes in climate driving glacial cycles have resulted in short and long term periods of fluctuation in global sea level (Chappel and Shackleton 1986, Voris 2000), sea surface temperatures (CLIMAP 1976, Lea *et al*. 2000) and oceanic current systems (Bush and Philander 1998). These fluctuations are purported to have shifted the distributional ranges of IWP species due to both habitat changes that resulted from climatic oscillation (Fleminger

1986) and also by the provision of new dispersal routes via changes in currents and oceanic circulation patterns (Benzie and Williams 1997; Palumbi *et al*. 1997).

Evidence of gene flow via shared allelic ancestry may occasionally be apparent in the absence of contemporary dispersal among marine populations and so confound interpretation of population connectivity (Benzie 1999). Populations founded in newly available habitat by historical episodes of long-distance peripheral range expansion can display genetic homogeneity over considerable geographic scale, despite an absence of contemporary connection (Hewitt 2000). High levels of genetic similarity among west Pacific populations observed for a number of marine invertebrates are counter to that expected for dispersal by contemporary current systems and may reflect Pleistocene dispersal and or stochastic expansion events (Benzie and Williams 1997, Palumbi *et al*. 1997, Williams and Benzie 1997). These examples demonstrate that considerable time lag may exist between historical genetic and contemporary demographic signatures of population structure (Lavery *et al*. 1996-b).

Fluctuations in sea level also resulted in many instances of vicariance within the IWP. For instance, periods of sea level regression resulting in the coalescence of islands and land masses within the Sunda and Sahul shelves areas of the Indo-Australian archipelago effectively blocked several avenues of dispersal between the tropical waters of the West Pacific and the East Indian Ocean (Fleminger 1986, Briggs 1995, Voris 2000). The cyclic emergence – subsidence of physical barriers to marine dispersal across the Indo-Australian archipelago have repeatedly sub-divided marine populations distributed either side of these coastal groups (McMillan and Palumbi 1995). The distributional ranges of a number of taxa within both the Siganid and Chaetodontid families of fish either adjoin or coincide either side of this region, suggesting that clusters of sister species arose in this area by episodes of allopatric separation (Woodland 1983, Blum 1989). Subsequent genetic analysis of the Chaetodontids indicated both species and population divisions co-occurring across the Indo - Malay Archipelago, date back to within the Plio - Pleistocene time frame (McMillan and Palumbi 1995). Genetic analyses for a broad selection of Indo-Pacific marine species indicate similar concordant divisional timeframes between the Indian and Pacific provinces (see review Benzie 1999). These collective results indicate emergence of these barriers during the Plio-Pleistocene epochs substantially affected the structure and distribution of many marine species in the IWP.

1.4.3 Marine translocations and genetic analysis

The geographic ranges of some marine species have expanded due to human induced translocations (review – Grosholz 2002). In these instances, exotic introduced species are observed to flourish at locations far removed from their normal observed ranges (Carlton 1996). Successful transfer of a species into new habitat supersedes natural impediments to colonisation of a species to an area and so confounds natural biogeographical patterns of species distribution. Arguably, the greatest cause of translocation occurs via the release of propagules contained in ship's ballast waters from one location to another. Given the high frequency of commercial international shipping, translocation of species (and pathogens) via these means may occur with high frequency and is therefore an area of concern.

Genetic analysis may help to distinguish between natural and exotic colonisations of a species by comparing the genetic profile of the colonist population relative to that detected for pre-existing populations (Geller *et al*. 1994). This approach is especially powerful when prior genetic analysis of the target species has demonstrated some degree of recognisable phylogeographic structure within the species distribution. With sufficient power of resolution, genetic analysis may not only identify the source of an exotic population (Geller *et al*. 1997), but also provide a qualitative estimate of the effective size of the immigrant gene pool relative to the source population (Bagley and Geller 2000, Holland 2001, Villablanca *et al*. 1998). Because genetic diversity and structure within newly colonised populations are representative of non-equilibrium conditions, examination of diversity levels at cytoplasmic and nuclear genetic markers may provide insight into the responses of varied genetic markers to the colonisation process.

1 .5 The mud crab *Scylla serrata*

1.5.1 Life history and dispersal

Scylla serrata (Forskål, 1775) is an estuarine adapted portunid crab found throughout the Indo-West Pacific (IWP) region. For most of its lifecycle, *S. serrata* is resident within sub-tidal and inter-tidal areas associated with estuarine habitats (Hyland *et al*. 1984). Typical of all marine decapods, *S. serrata* exhibits a multiphasic life cycle

(Figure 1.1) with distinct changes in morphology, physiology and behaviour apparent between each major successive stage (Hill 1982).

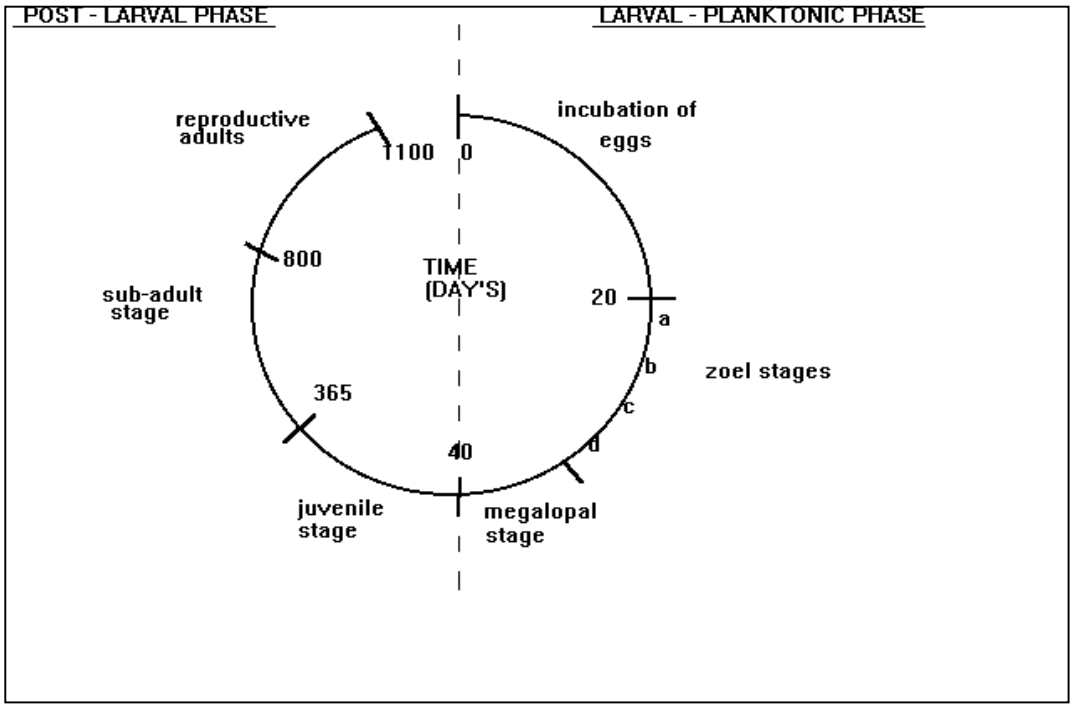


Figure 1.1 Diagrammatic display of stages within the life cycle of *Scylla serrata.*

Ovigerous female crabs release a clutch estimated to contain up to 2 million eggs offshore from adult habitats. As illustrated in figure 1.1, newly hatched *S. serrata* go through a succession of zoel-megalopal developmental stages, representing the species freely dispersed planktonic phase. This plankton phase typically lasts up to three weeks. Settlement of *S. serrata* to inshore habitats occurs during the megalopal stage of development (Hill 1982, Fielder and Heasman 1986). Successive growth stages from juvenile to sexually mature adults occur within estuarine habitats over a period ranging from 18-27 months depending upon regionally variable growth rates (Fielder and Heasman 1986). On average, mud crabs live for a period upwards of three years (Heasman 1980).

Studies of the potential range of dispersal by *S. serrata* indicate that adult movements are generally limited and confined to intertidal areas (Hill *et al*. 1982, Hyland *et al*. 1984). Results of mark-recapture sampling of adult crabs over a one year period in eastern Australia indicate approximately 70% of recaptures occur within 3 kilometres

from the site of release (Hyland *et al*. 1984). The exception for adults occurs during spawning periods when ovigerous females may migrate up to 95 kilometres offshore to release eggs (Hill 1994). The extent of dispersal during the plankton stages is unknown, however it has been speculated that considerable transfer is possible by entrainment of plankton into currents following the offshore release (Hill 1994). The consensus among researchers is that the propensity for widespread planktonic dispersal of this species coupled with its high fecundity and quick generation time may ensure admixture and homogeneity of population structure distributed over quite large areas. In this sense, dispersal among *S. serrata* populations may resemble Wright's (1931) Island model of migration.

1.5.2 Taxonomy of Scylla serrata and some inferences of population distribution

The contemporary distribution of the genus *Scylla* encompasses an enormous range of temperate-tropical coastal habitats located from east Africa, across the Asian sub-continent down to Australia and out to the Melanesian islands within the west Pacific. Ambiguity concerning the taxonomic status of *Scylla* morphs throughout this distribution resulted in debate as to whether the genus consisted of one or several species. Despite localised variation in meristic, morphological and behavioural traits noted by a host of researchers, all morphs were placed in synonymy by Stephenson and Campbell (1960) and classified as *Scylla serrata*. Using a combination of genetic and morphological characters, Keenan *et al*. (1998) subsequently revised the genus to include four distinct non-hybridising species (*serrata, olivacea, paramamosain* and *tranquebarica*). The type specimen originally described by Forskål (1775) is synonymous with *S. serrata* under this latter classification. Keenan *et al*. also delineated an approximate distribution for each of the four species by screening diagnostic character differences among specimens sampled from a large number of localities throughout the IWP. Their results indicate:

1. All species of *Scylla* except *S. serrata* are present and sympatric at numerous South-East Asian locations

2. *S. serrata* is the most widespread of the species in the IWP, present at numerous Indian and west Pacific coastal continental and island locations

3. *S. serrata* may be absent from equatorial tropical waters

1 .6 Aims

Genetic studies of widespread taxa within the IWP are providing controversial results counter to the paradigm that planktonic dispersal in the marine environment is generally sufficient to maintain genetic homogeneity over vast scales. Moreover, genetic analysis is disseminating new independent perspectives of biogeographic mechanisms affecting the evolution and ecology of marine taxa.

In this thesis, I will examine the genetic structure of *S. serrata* populations using molecular genetic methods to test hypotheses concerning biogeographical history and levels of population connectivity in the marine environment. There are several features concerning the life history of this species and its distribution in the IWP that makes it ideal for examining hypotheses of genetic structure specific to coastal marine species.

The widespread distribution of *S. serrata* in the IWP relative to other species within the genus may indicate a propensity by this species for dispersal across broad oceanic distances and among distant populations (Keenan *et al*. 1998). The potential mobility of *S. serrata*, mediated by offshore spawning and planktonic dispersal, may result in high levels of exchange among populations (Hill 1994). An examination of genetic structure among *S. serrata* populations throughout the IWP using mtDNA analysis may provide estimates of the frequency of trans-oceanic gene flow events and so provide an estimate of the broad geographic scale of demographic connection and the levels of dispersal of which the species is capable.

Because mtDNA can be used to infer genealogical relationships among samples, historical signatures of gene flow or sub-division within this species distribution may be evident in the contemporary geographic distribution of mtDNA haplotypes. Therefore, mtDNA analysis may determine if historical patterns of population expansion and vicariance reported for a variety of IWP marine species had similarly affected *S. serrata* populations (Chapter 3).

The extent to which gene flow can occur among populations connected by coastlines on a single continental shelf will also be examined by sampling multiple coastal locations within the Australian distribution. This distribution encompasses three semi-independent marine regions partially separated by contemporary broad-scale patterns of hydrology and totally separated during the episodes of historical sea level reduction.

Patterns of genetic structure within and among these regions should provide an indication of the type and extent of dispersal possible along these coastal areas. Examination of these populations using mtDNA (Chapter 4) and microsatellite markers (Chapter 5) will provide independent comparative means of testing for genetic structure.

I also had an opportunity to examine a recent colonisation by the species into a coastal region where the species had not previously been observed. I will use genetic methods to determine if the colonists are derived from a natural range expansion, or if they had been translocated from a foreign source (Chapter 6). Regardless of their source, by comparing genetic diversity among markers (mtDNA and microsatellites) I may determine the independent responses of the mtDNA and nuclear genes to the effects of the colonisation process.

CHAPTER 2: **Methods**

2 .1 Sampling

Specimens for analysis were provided by three independent sampling efforts. Specific details of sample sizes and locations for each of these efforts are described within relevant chapters. The following is a general summary of the three efforts.

Frozen leg tissue (-72° C) from individual mud crabs (morphologically identified as *S. serrata*) collected from various IWP locations, were derived from the collections of Dr. Clive Keenan (Southern fisheries: Queensland DPI). All samples derived from Keenan were originally collected before 1996.

Post - larval mud crab populations around the Australian coastline and a small number of South East Asian locations were sampled at various occasions from 1996 – 2000, via my own efforts and that by voluntary contributors. For these efforts, adults were captured in mangrove and estuarine habitats using standard mud crab traps, whereas juveniles and sub-adults sheltering under stony surfaces during low tide recess were captured by hand. A single cheliped was excised from each crab before its release. Excised chelipeds were catalogued and either frozen on site using dry ice or preserved with buffer designed to maintain tissue samples at room temperature (20 % DMSO, 0.5 M EDTA pH 8.0, saturated with NaCl). Although *S. serrata* were primarily targeted, other species of *Scylla* were sampled if present within a location. Samples were transferred for long-term refrigeration at -72° C within the Molecular Ecology laboratory, Griffith University.

Samples of southwest Australian mud crab populations were provided by the efforts of Dr. Lynda Bellchambers (West Australian Department of Fisheries) during the 2001 – 2002 summer period.

2 .2 Laboratory methods

Methodologies for deriving both mtDNA and nuclear DNA markers from samples are described here. General DNA extraction, PCR and sequencing protocols are shared among classes of genetic marker. Specific methodologies or conditions are stated as such.

2.2.1 DNA extraction

In the course of this study, three extraction protocols were used to separate DNA from sample tissue. The majority of samples were subjected to a total DNA extraction using protocols similar to that of Sambrook *et al.* (1989). In preparation, 500 mg tissue samples were homogenised with 500 µl of extraction buffer (20mM Tris- Hcl pH 8.0, 10mM EDTA, 0.5% SDS) and 5 µl of Proteinase K (20 mg/ml) using a stainless steel grinder. Samples were agitated for up to 12 hours at a temperature of 65°C, to encourage lysis of cell walls. After agitation, samples were centrifuged at 13000 rpm for 15 minutes to separate supernatant from sediments. Supernatant was siphoned off and stored at 4° C pending extraction. Proteins and lipids were removed via sequential extractions of the supernatant in equal volume (~500 µl) of phenol. After addition of the phenol, samples were vortexed for 1 minute and centrifuged for 15 minutes. The resulting top phase was siphoned from the lower phenol-sediment phases and transferred to a new 1.5ml tube. Additional phenol extractions were repeated to improve purity of the supernatant. A final extraction was performed using chloroform-isoamyl alcohol (24:1) to remove traces of phenol. Nucleic acids were precipitated out of the supernatant using an equal volume of chilled isopropanol and pelleted by spinning at 13000 rpm for 15 minutes. After residual fluids were removed, pellets were washed with 500 µL of 70% ethanol and spun at 13000 rpm for 5 minutes. Residual salts and ethanol were removed and pellets were then vacuum dried for 15 minutes and re-suspended with 50 µL of 0.1M Tris HCL buffer (pH 8.0). Extracted DNA was visualised by loading 3 µL of DNA template with a 1 µL mix of Bromophenol Blue/water into a 5 mm thickness 1.6% agarose gel (Bio-Rad ultrapure DNA grade agarose) containing 100 mg/l Ethidium Bromide. 10 µL of 30% Lambda BSTE II restriction phage was loaded to the first well and acted as a marker reference for estimation of band size. Gels were electrophoresed for 30 minutes at 100v (Bio-Rad Minisub operating with a Bio-Rad 200/2.0 power source) whilst submerged in 1 x TAE running buffer. Gels were photographed under an ultra – violet light source (UVP

Transilluminator). Estimated quantities of DNA extracted by this method were on average > 200 ng/ul.

Alternatively, DNA was extracted from samples using Chelex (Biorad) as an agent for separating organic products from nucleic acids (Zhang and Hewitt 1998). The relative ease and speed of this method justified its use in the latter stages of the project. In contrast to total DNA extraction, Chelex extractions were simple and rapid but provided less quantity of purified DNA per sample. Approximately 5 mg of sample tissue was ground and then incubated for 12 hours (55 º C) in a buffer containing 5 µl of Proteinase K (20 mg/ml) and 500 µl of 20 % chelex mix. Following a 15-minute spin at 13000 rpm, the top 250 µl fraction of supernatant was removed and mixed with an equal volume of 20 % chelex mix. The supernatant was immediately frozen at -72 C. The quantity of DNA extracted by this method is minimal (< 10 ng/µl) but sufficient for PCR amplification. Immediately prior to PCR (section 2.2.2), the supernatant was incubated at 100 º C for 15 minutes before a portion was introduced as the final component of a PCR mix.

A third extraction protocol followed the procedures of Tamura and Aotsuka (1988) designed to preferentially extract mitochondrial DNA from other nucleic acids. This protocol was applied to a small subset (N =10) of the samples (refer section 2.2.5) and provided an independent test for the presence of a nuclear insertion of the target mitochondrial gene.

2.2.2 PCR amplification of mtDNA products

Pilot attempts to target and amplify the hyper-variable non-coding region of mud crab mtDNA failed, despite screening a variety of universal primer combinations. As an alternative, approximately 650 bp from the 3' region of the mitochondrial cytochrome oxidase subunit I coding gene (COI) was targeted as a potentially informative marker for this study. This gene region is well characterised for invertebrate taxa and has been used extensively for both inter and intra specific phylogenetic analysis. COI sequences derived from Keenan *et al*. (1998) using universal primers COIa & COIf (Palumbi *et al*., 1991) provided a template to identify internal conserved priming sites specific for *Scylla* species. To this end, heavy strand primer Mtd-10 (Roehrdanz 1993) and light strand primer C/N 2769 were designed to amplify a 597 bp product internal to the COIa & f combination (Figure 2.1& Table 2.1). This primer combination was used for all

samples and provided strong PCR product devoid of secondary amplified products. A small number of individuals (N=10) were also amplified using the COIa & f primers to provide an independent check for nuclear insertion (refer section 2.2.5).

PCR reactions contained 0.5 µL of total template DNA (or 15 µL of chelax derived template), 1 unit of *taq* DNA polymerase (Promega), 0.125 mM dNTP's (Promega), 1mM polymerase reaction buffer, 0.5 M of each primer (5 pmol/µL), made up to a total volume of 25 µL with autoclaved distilled water. Reactions were contained in 150 µL tubes and topped with 30 µL of light mineral oil to prevent evaporation. Negative controls lacking template, were run simultaneously with sample reactions to test for contamination of reagents by foreign DNA. PCR amplifications were conducted using Minicycler thermocyclers (MJ Research) set for 35 cycles of a three-stage temperature profile. These cycles typically consisted of a 30 second denaturing stage set at 94° C, a 30 second annealing stage set at lowest temperature among primer pairs (table 2.1), and a 1 minute extension stage at 72°C. Upon completion of the 30 cycles, PCR products were given an extra 10-minute extension stage before reducing to a storage temperature of 4°C. To check for success and quality of the reaction, 3uL of PCR product were electrophoresed in agarose as per conditions described in section 2.2.1.

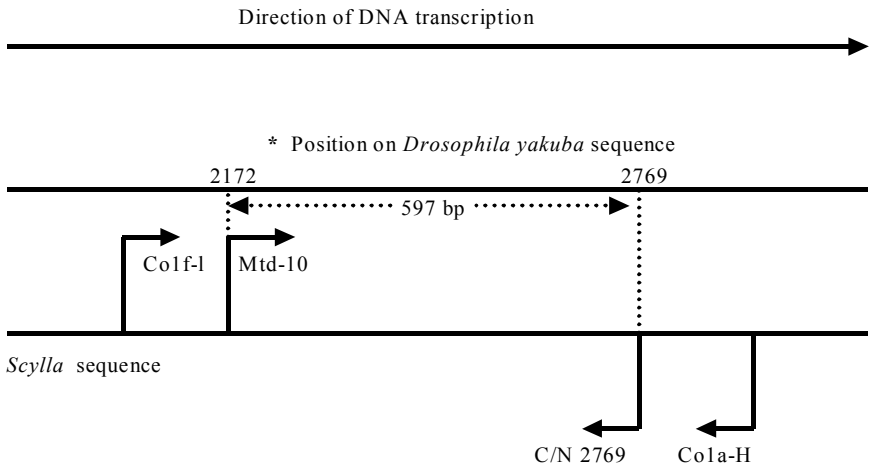


Figure 2.1 Schematic diagram of COI primer locations relative to *Drosophila yakuba* sequence (Clary and Wolstenholme 1985). * Positions given for *D. yakuba* as listed in GENBANK, Accession # X03240.

Table 2.1 Primer Sequences, annealing temperature in degrees Celsius (Ta°), length in base pairs (bp) and reference source. Nucleotide residues include standard residues (C,G,A,T), R = A or G (purine degenerate) and Y = C or T (pyrimidine degenerate).

Primer	Nucleotide bases (5' – 3')	Ta°	bp	Ref.
COIf – L	CCT GCA GGA GGA GGA GAY CC	63	20	1
Mtd-10	T TGA TTT TTT GGT CAT CCA GAA GT	59	24	2
C/N 2769	TT AAG TCC TAG AAA ATG TTG RGG GA	65	25	3
COIa – H	AGT ATA AGC GTC TGG GTA GTC	57	21	1

1: Palumbi *et al*. (1991)
2: Roehrdanz (1993)
3: This study

2.2.3 Temperature gradient gel electrophoresis of mtDNA

Temperature gradient gel electrophoresis (TGGE) provides an efficient means of screening large numbers of samples for the presence of allelic variants among homologous PCR fragments (Lessa and Applebaum, 1993). The method operates on the principle that the rate of electrophoretic migration of a DNA duplex through a denaturing gel slows down at a critical range of temperatures during which helical unwinding of the DNA occurs. The temperature range for such helical unwinding, or melting domain, is determined by the constituent nucleotide bases within the DNA fragment (Wartell *et al*. 1990). Ultimately complete DNA strand disassociation results in the near cessation of fragment migration during electrophoresis. Hence, wild type and mutant homologous PCR fragments may be discriminated by their relative rates of elecrophoretic migration through a denaturing gel with an increasing gradient of temperature. In theory, TGGE can identify DNA duplexes that differ by a single mutation (Wartell *et al*. 1990). However, empirical results using TGGE indicate that this expectation is not always met and that different haplotypes may sometimes be indistinguishable (Campbell *et al*. 1995). Employing heteroduplex analysis in conjunction with TGGE can compensate for this redundancy in TGGE (Lessa and Applebaum 1993, Campbell *et al*. 1995). The process of denaturing and renaturing a mix of wild type and mutant PCR products allows artificial hybrid duplexes (heteroduplexes) to form. Base pair mismatches within a hybrid heteroduplex effectively lower the melting profile of the fragment. Consequently, heteroduplexes exhibit reduced thermal stability relative to the original homoduplexes. The

electrophoretic migration rates of heteroduplex products over a temperature gradient are often retarded relative to that seen for the constituent homoduplexes. The resulting thermal phenotypes are then scored for the relative mobilities of both homo and heteroduplex migration. By this means, TGGE screening for variant PCR products is enhanced by heteroduplexing samples with a reference wild type. The ease of scanning allelic variants among large samples justifies the use of this method for population genetic analysis and as a precursor to sequencing.

The capacity of TGGE to detect differential melting properties among heterozygous DNA fragments is dependent upon the presence of a reversible melting domain(s) within the length of the investigated fragment (Wartell *et al*. 1990). The thermal melting profile for the targeted DNA fragment was determined using perpendicular TGGE. For this procedure 500 ng of un-purified PCR product mixed with 20 µL of 10 x ME buffered dye (200 mM MOPS, 10 mM EDTA, 0.05% bromophenol blue and xylene cyanol, pH 8.0) was made up to a total volume of 200 µL with distilled -autoclaved water and loaded into a well spanning the length of a 5% polyacrylamide gel (21.6 g urea, 2.25 mL glycerol, 7.5 mL of 30: 0.5 acrylamide: bis, 0.9 mL 50 x ME buffer, 100 µL 10% APS, 300 µL TEMED adjusted to a total volume of 44.6 mL with distilled water). Electrophoresis conducted on a horizontal Diagen TGGE system maintained temperature gradients using two external variable temperature water baths. Electrophoresis for 30 minutes at 300 volts with a uniform plate temperature of 20° C allowed migration of the sample away from the origin in the absence of a thermal gradient. A temperature gradient of 20° to 60° C perpendicular to the direction of electrophoretic migration was then equilibrated before an additional 60 minutes of electrophoretic run time. Following electrophoresis and in preparation for staining, DNA was fixed within the gel via two separate three-minute soaks in 0.5% acetic acid and 10% ethanol. The DNA was stained for 10 minutes in a 0.1% $AgNO_3$ bath, followed by a quick wash in distilled water to remove excess stain. Stains were developed for ~ 20 minutes in a solution of 1.5% NaOH, 0.01% $NaBH_4$ and 0.015% formaldehyde to allow visualisation of the stains. Fixation of the developed stains was ensured by a 30 - minute soak in 0.75% Na_2CO_3.

The melting profile for the target DNA confirmed the presence of a reversible melting region and allowed extrapolation of an optimal temperature gradient and run time for parallel TGGE. It can be seen in Figure 2.2 that a characteristic melting pattern occurs

within the temperature range of ~ 35° to 45° C, where electrophoretic migration of DNA is grossly retarded due to helical unwinding. The temperature at which DNA helices are halfway unwound (Tm) was extrapolated from the melting profile as 37° C. The perpendicular gel provided an estimate of the migration rate of the DNA as ~ 2.5 cm/hour prior to helical unwinding. Using this information, parallel TGGE gels were run for 3.5 hours at 300V using a temperature gradient 8° C either side of Tm (i.e.: T1=29° C, T2=45° C). A typical parallel run simultaneously screened 26 independent samples relative to one reference.

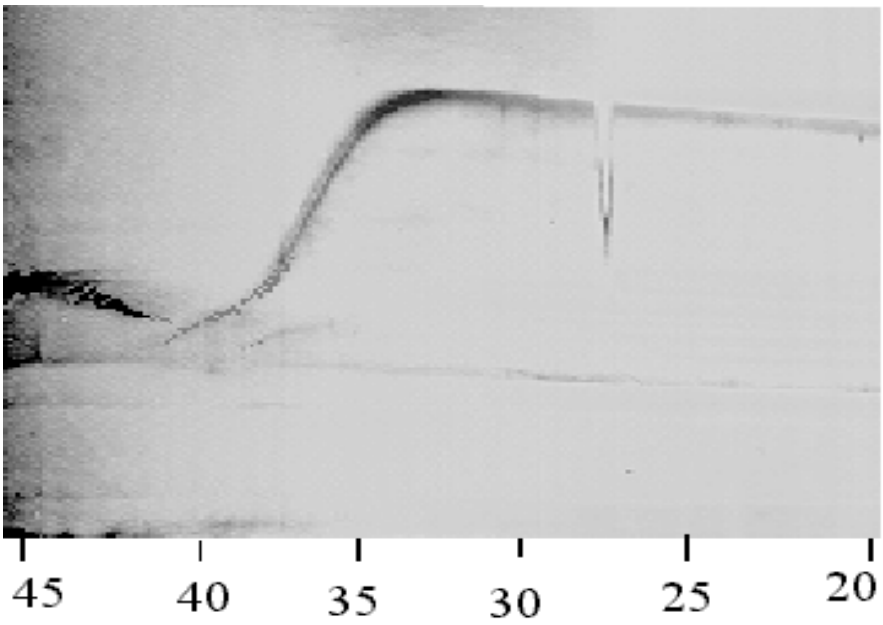


Figure 2.2 Perpendicular TGGE. Direction of DNA migration is orientated from bottom to top. A temperature gradient from high to low was established across the gel from left to right as per displayed 5° C gradations. Melting curve due to disassociation of DNA strands can be seen between temperatures 33° – 40° C.

To enhance the ability of parallel TGGE to detect variant haplotypes, samples were heteroduplexed with reference DNA. Campbell *et al*. (1995) advise rerunning samples using different heteroduplex references thereby validating detected mutant haplotypes. Pilot work identified several different individuals when used as heteroduplex references

that produced very different phenotypic signatures among samples. To this end, samples were independently run 2-3 times on optimised parallel TGGE and heteroduplexed with a different reference at each occasion. This approach allowed the detection of haplotypes that were occasionally scored as identical using a single heteroduplex reference. Prior to running optimised parallel TGGE, samples containing 20 ng of unpurified product from each of reference and sample to be screened, 4 μL of 8M urea, 0.8 μL of 10 x ME buffer and dye were placed into a thermocycler and subjected to a 10 minute denaturation period (94° C). This was followed by a gradual period of renaturation over 30 minutes to a temperature of 20° C. This process of denaturation and renaturation allowed the formation of homo/heteroduplex DNA complexes between the reference and sample DNA. Following electrophoresis and staining, individuals were scored for the relative mobility of their phenotypes. These conditions clearly identified a variety of homoduplex and heteroduplex phenotypes (Figure 2.3). As each sample was run at least twice with different references, haplotypes were scored as a composite of the independent mobility tests.

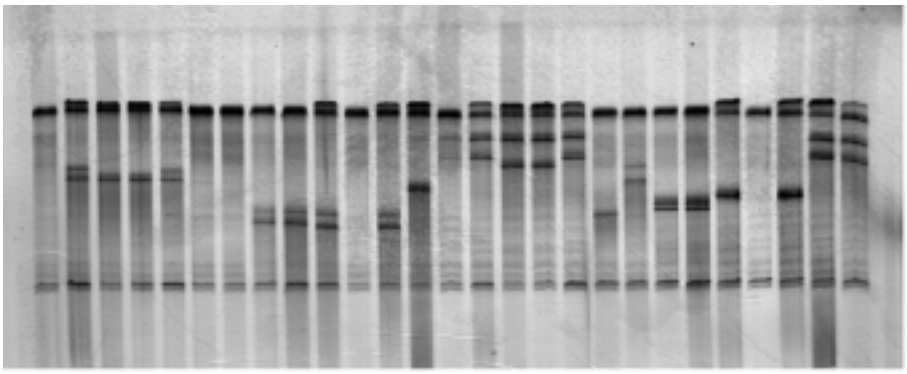


Figure 2.3 Parallel TGGE conducted for 27 samples, each heteroduplexed with a reference. Direction of DNA migration here is from bottom to top. Temperature gradient was parallel to the direction of migration and increased in the direction of migration. Heteroduplex bands clearly resolved from homoduplex bands by their retarded migration relative to the homoduplex bands.

2.2.4 Sequencing

All defined haplotypes were sequenced from both 5' and 3' directions, to identify and verify mutational differences within the sample. Prior to sequencing, PCR products were purified using a "Quiaquick PCR purification kit" (Quiagen). For each sample, 50 µL of PCR product was electrophoresed and excised from a 2% agarose gel. The excised product was mixed with an agarose solvent and spun at 13000 rpm for 1 minute through a spin column to remove agarose and solvents. The column was given an additional rinse and spin with recommended buffers to enhance purity of the product. Purified DNA was eluted in 25 µL of 0.1 M TE buffer (pH 8.0). Sequencing was conducted using the products and protocols of ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer). Approximately 20 ng of purified template DNA was mixed with 3.2 pMol of a primer, 8.0 µL of TRR mix made up to a total volume of 20 µL with double distilled water. Reactions were cycle sequenced in a thermocycler using the following profile: 25 cycles of 10 seconds at 95°C, 5 seconds at 50°C, 4 minutes at 60 ° C. Reaction products were then precipitated and pelleted as per the manufacturer's recommendations, before being vacuum dried for 20 minutes. Dried pellets were sequenced on an ABI 377 automated sequencer operated by the Science Faculty, Griffith University. Results of sequencing were aligned (refer section 2.3.3) against comparable *Scylla spp* COI sequence data. All scored mutations were checked for concordance among 5' and 3' sequencing efforts and subjected to re-sequencing when ambiguous.

2.2.5 Checks against nuclear insertion

Translocation of mtDNA-derived sequences to the nuclear genome (nuclear insertion) has been observed for a number of taxa (refer Sorenson and Fleischer 1996). Nuclear insertions generally evolve at a slower rate than mitochondrial counterparts due to the differences in effective population size between genomes. Unlike their mtDNA homologs, insertions to the nuclear genome are liable to recombination and reticulate evolution. As a result, phylogenetic analysis of mtDNA haplotypes may be confounded by inclusion of nuclear inserted homolog sequences within a dataset. For this study, two independent checks for nuclear insertion were conducted on ten samples (five from northern Australia, three from eastern Australia and two from South Africa). Both independent checks involved the comparison of sequences for the eight samples derived from conventional extraction/PCR steps (described above) with sequences derived using variant extraction and PCR steps. Any mutational differences between sequences from

the same sample would then be diagnostic of an insertion. The first method involved preferential extraction of mtDNA relative to nuclear DNA using a protocol described by Tamura and Aotsuka (1988). This method involves two major procedures, preparation of mitochondrial pellet by low speed centrifugation and mtDNA extraction by alkaline lysis - allowing circular mtDNA to separate from linear nuclear DNA. For each sample 50 mg of tissue was homogenised in 1ml of buffer containing 0.25 M sucrose, 10 mM EDTA, 30 mM Tris-HCl (pH 8.0). The homogenate was centrifuged at 1000g for 1 minute to pellet nuclei and cellular debris. Supernatant was recovered and re-centrifuged at 12000g for 10 minutes to pellet the mitochondria. Supernatant was discarded and the pellet was resuspended in 50 µL of buffer containing 10 mM Tris-EDTA (pH 8.0), 0.15 M NaCl and 10 mM EDTA. To this 100 µL of freshly prepared 0.18 M NaOH containing 1% SDS was added, vortexed and kept on ice for 10 minutes. 75 µL of 3 M potassium and 5 M acetate solution was added, mixed and kept on ice for 10 minutes. This mixture was centrifuged at 12000g for 5 minutes, after which the supernatant was siphoned off. The remaining pellet was then subjected to a standard phenol-chlorophorm extraction procedure as described in section 2.2.1. PCR products from this method were then compared to products derived from the total DNA extractions.

The second check for nuclear insertion involved comparison of sequences derived from universal COIa & f primers with that derived from Mtd10 & C/N2769 (primers internal to COIa & f). The premise of this test is that a set of conserved universal primers may be more likely to amplify a nuclear homolog than a set of primers specifically designed for the amplification of mtDNA product. As well as these checks, sequences for all identified haplotypes were inspected for signature indications of translocation such as multiple peaks at a given nucleotide position, unexpected ‘stop’ codons and evidence of inserted/deleted nucleotide positions or coding frame-shifts within the gene fragment.

2.2.6 Microsatellite library construction

A partial genomic library was constructed from *S. serrata* DNA and screened for simple -sequence repeat regions using a methodology similar to that of Rassmann *et al*. (1991). Approximately 10 µg of genomic DNA extracted from a single crab was digested for three hours with restriction enzyme *Sau*3A1. After separation on a 1.5 % agarose gel, DNA fragments in the size range of 200 – 800 base pairs were excised, purified and ligated to an equal volume of plasmid vector pUC18 (Amersham-Pharmacia). The

plasmids had previously been digested with *Bam*HI and dephosphorylated to create overhanging ends to match those resulting from the *Sau*3A1 digest. Recombinant plasmids were electroporated into competent *Escherichia coli* cells (strain JM109, Promega) and incubated for an hour at 37 º C. Cells were spread on to agar plates containing LB-Ampicillin and incubated overnight at 37 º C to promote selective growth of transformed colonies. A total of 2200 recombinant colonies were picked from plates and incubated overnight in a grid formation on new LB-Ampicillin agar plates and later stored at 4º C. Recombinant colonies were blotted from the plates on to filter membranes (Hybond-N, Amersham). DNA from this transfer was cross–linked with the membrane, denatured and probed with oligonucleotides [$(ACC)_8$; $(AAC)_8$; $(AAG)_8$; $(AGC)_8$; $(ACG)_8$; $(ACT)_8$; $(CA)_{15}$; $(AG)_{12}$] that had been end labelled with [γ^{32}P] – dATP (Perkin Elmer). Cross – linked single stranded DNA was hybridised with the probes overnight before being exposed onto X-ray film at – 80 º C for 12 hours. Autoradiographs revealed 63 positive clones that hybridised with probed repeats. Colonies containing repeats were thus identified and picked from the stored agar plates and cultured overnight at 37 º C. Plasmid DNA was extracted from cultures by an alkaline - lysis miniprep and sequenced using Big Dye Terminators (Perkin Elmer) and universal plasmid primers (M13 F & R, Amersham Pharmacia Biotech). Sequences were determined by electrophoresis on an ABI 377 automated sequencer. Sixty of the clones contained recognisable microsatellite arrays. Over 95 percent of microsatellites contained (CT/AG)*n* di-nucleotide repeat motifs. A single tri-nucleotide repeat sequence (CCA)*n* was also identified. Only eight of the 60 candidate microsatellites had sufficient or adequate flanking region for primer design. Primers were designed to maximise annealing temperature and minimise flanking regions.

2.2.7 Screening for Microsatellite variation

PCR amplification of targeted microsatellites was successful for five of the primer pairs (Table 2.2). PCR reactions contained 50 -100 ng of template DNA; 0.25 units of *Taq* DNA polymerase (GIBCO BRL); 0.25mM of dNTPs; 2.5mM $MgCl_2$; 0.5 µM of each primer (one primer end labelled with fluorescent HEX, refer Table 2.2); in 1 * proprietary reaction buffer (50mM KCl, 20mM Tris-HCl pH 8.4; GIBCO BRL) up to 12.5 µL total reaction volume. PCR reactions were cycled with the temperature profile: 94 º C denature for 1 minute, followed by 5 cycles of 94 º C for 28 seconds, annealing temperature less 4 º C (see Table 2.2) for 28 seconds and 72 º C for 40 seconds; 32 cycles of 94 º C for 28 seconds, annealing temperature for 28 seconds, 72 º C for 40

seconds followed by a final 72 º C extension of 7 minutes. Denatured PCR products and TAMRA size standard (ABI) were electrophoresed through 5% denaturing acrylamide gels using a GelScan 2000 rig (CORBETT RESEARCH). Detected bands were analysed for product size using ONE –Dscan (Scanalytics) software.

Table 2.2 Five microsatellite loci isolated from *S. serrata*, forward (F) and reverse (R) primer sequences, observed repeat number and motif of sequenced alleles. Optimal PCR annealing temperature (T_a) in º Celsius for each locus as indicated. Also shown are the observed numbers of alleles for each locus sampled from one Australian location (N = 36).

Locus (Genbank #)	Primer sequences 5' → 3'	Repeat Motif	T_a	# of alleles
Ss-101 (AF 508135)	F: HEX-ATT CAA CAC GCG CGC GTA CGC R: GCA GTT TAC CAT ATG CTT GGG	$(AG)_{36}$	55	26
Ss-103 (AF 508134)	F: HEX-GTT ATA TAA GAA ATA ATG TCC R: GTT CCT GCT ATG TAA TCC CG	$(GA)_{36}$	45	20
Ss-112 (AF 508133)	F: TCA TTC TCA GTA CCT TTA ATC R: HEX-GTT ATC GTC TGC TGG GAC C	$(GA)_{37}$	45	23
Ss-403 (AF 508132)	F: GAC AAA GGA GCA CTC AGC CAC R: HEX-GAA GGA TTC ACT TGT CCA CGC	$(CT)_{24}$	55	19
Ss-513 (AF 508131)	F: HEX-GGC CGG GTG AGG GAT GAG CC R: CGT TTC CGC AAC CAA CAG ATG	$(CT)_{14}$	55	5

2 .3 Data analysis

2.3.1 Sample diversity measures

Measures of mtDNA variation within spatial sample categories (population, region, etc.) were estimated using Nei's (1987) haplotype diversity (h) statistic, as implemented in the population genetics software Arlequin, versions 1.1 & 2.0 (Schneider *et al*. 2000). This statistic estimates the probability of randomly drawing two different haplotypes from a sample. Diversity levels at microsatellite loci were measured as the observed (direct count) and expected (unbiased estimate; Nei 1978) heterozygosity values within

populations using GENPOP ver. 3.1d (Raymond and Rousset 1995). Individual microsatellite loci were also tested for evidence of linkage disequilibrium and significant departures from Hardy – Weinberg equilibrium (HWE). Significant departures from HWE were examined for excessive or reduced homozygosity levels by calculating the inbreeding co-efficient (*Fis*). GENEPOP computed exact tests based on a Markov chain method to estimate probability values for each of the linkage disequilibria and HWE tests. For all exact tests, the de-memorization number, number of batches and number of iterations per batch were set at maximum default to increase precision of the *P* estimates and compensate for the large number of alleles usually found within microsatellite data sets.

2.3.2 Tests for genetic differentiation and population structure

Genetic differentiation between paired locations was estimated using two analogous forms of Wright's fixation index. Weir and Cockerham's (1984) *F*-statistics were estimated using Arlequin (Schneider *et al*. 2000) for mtDNA data and also in FSTAT, version 2.9.3 (Goudet 1995) for microsatellite data. Slatkin's (1995) *R*-statistics were also calculated for microsatellites using RST-CALC, version. 2.2 (Goodman, 1997). The calculation of *R* - statistics differs from *F*st primarily by assuming that microsatellites evolve via an accumulation of stepwise mutations rather than by random size shifts in allelic length. Loci evolving in a stepwise mode may therefore have a high propensity for allelic homoplasy, and so confound estimates of genetic distances estimated by conventional *F* - statistics. Slatkin's (1995) *R*st equivalent of the *F*st estimator incorporates the mutational bias inherent under a stepwise model of mutation by measuring average sum-squared differences in allele size from a sample as a proxy for the average time to coalescence among alleles within the sample. The two statistics were used on microsatellite loci to see if estimated differentiation among populations varied according to the mutational model selected. Significance of paired multi-locus fixation estimates were tested by random permutation (10,000 replicates) and adjusted for multiple comparisons by sequential Bonferroni correction (Rice, 1989). Single locus pairwise comparisons were examined for significant departure from allelic homogeneity using exact permutation test. All exact tests were run using GENEPOP, with de-memorisation number, number of batches and number of iterations per batch set to maximum default values to increase precision of the exact probability estimates.

AMOVA (Excoffier *et* al. 1992) was used to assess spatial scales of population subdivision by estimating the relative amounts of genetic variance partitioned among hierarchical sampling designs. *F*-statistic analogs (Φ) for each hierarchical sample were also generated in AMOVA and tested for significance by permutation test (10, 000 replicates). Data used by AMOVA was entered as either pure frequency information (all haplotypes/alleles assumed equi-distant, analogous to Weir and Cockerham's (1984) unbiased *F*st estimator); or incorporated a measure of genetic distance among haplotypes/alleles as well as their relative frequency in the sample (analogous to Lynch and Crease's (1990) *N*st).

2.3.3 MtDNA sequence comparisons and phylogenetic association

MtDNA sequences were aligned using either MEGA (Kumar *et al*. 1994) or BioEdit ver. 5.0.9 (Hall 1999) to identify positions of all variable and parsimonious sites relative to comparable consensus sequences. Variable sites were also examined for their degeneracy within amino acid codon groups. Sequences were assessed for conformity to neutral expectations of mutation accumulation among protein coding mtDNA genes by observing the ratio of synonymous (silent) to non-synonymous (amino acid-changing) mutational changes among a set of sequences. It is expected that the majority of mutations at mtDNA protein coding genes within intraspecific datasets are silent (Kocher *et al*. 1989).

Phylogenetic relationships among haplotypes were established by constructing either minimum spanning trees (Excoffier and Langeney 1989) or neighbour-joining trees (Saitou and Nei 1987) with bootstrapping (Felsenstein 1995). Both methods constructed topological relationships using measures of genetic distance between pairs of sampled haplotypes. Distance data for tree construction was computed within MEGA as a pairwise matrix of either the percent or absolute number of mutational differences between all putative haplotypes. Minimum spanning trees joined haplotypes connected by the least number of mutational differences as nodes in a network, progressively encompassing all haplotypes as a single network tree. Nucleotide homoplasy among haplotypes resulting in ambiguous connection in the network was accounted for by imposing two rules governing choice of connection (Excoffier and Langaney, 1989):

> 1: a link between two rare (<5%) haplotypes is less likely than a link between a rare and a frequent (>5%) haplotype

2: links between haplotypes in the same geographic location are favoured over links from different locations

As well as providing a visual representation of haplotype relationships, networks allow inferences of the coalescent history among haplotypes relative to their geographic and temporal proximity. Neighbour-joining trees were constructed using percent sequence differences among haplotype pairs. Bootstrap re-sampling of nucleotide sites among haplotypes (2000 replicates) was used to obtain percentage support for nodes within the tree topology. Tree construction and bootstrapping were performed using PHYLIP version 3.5c (Felsenstein 1995).

2.3.4 Assignment tests

Multi-locus genotypes from individuals can be used to build up genetic profiles to identify exclusive groups within a sample and assign individuals to the group from which it is most likely to be derived. The power of assignment using multiple independent microsatellite loci is positively correlated with the degree of genetic structure present among groups in a sample (Maudet *et al*. 2002). Assignment testing in this study was conducted using Cornuet *et al*.'s (1999) exclusion – simulation method to obtain likelihood probabilities for each individual assignment as implemented in the software GENECLASS. In this method, the individual to be assigned is removed from the dataset and the frequencies of all alleles at a locus for all remaining individuals in nominated groups are computed (assuming HWE in these groups). The likelihood of the removed individual's genotype is then computed in each group. Since each locus is assumed independent, the likelihood of an individual's multi-locus genotype occurring in a given population is calculated as the product of likelihood estimates measured at each of the loci. Assignment probabilities by this method are however confounded by the presence of alleles endemic to a single group (absence of the allele in a nominated group will result in zero probability of the multilocus assignment to that group). This problem was overcome by artificially inserting a single representative of endemic alleles to all samples where it is absent ("Add one in" procedure) as recommended by Cornuet *et al*. (1999).

2.3.5 Bottleneck tests

Populations were tested for loss of allelic diversity at microsatellite loci as might occur immediately following a severe bottleneck or founder event, using the software

BOTTLENECK version 1.2.02 (Piry *et al*. 1999). Populations that undergo a strong reduction in *Ne* also experience a stochastic loss of alleles (particularly rare or low frequency alleles). To a lesser extent, heterozygosity in these populations is also reduced. The reduction in allelic diversity is expected to be consistently greater than the loss of heterozygosity across independent loci. Cornuet and Luikart (1996) devised a means of testing for this difference by comparing the expected heterozygosity (*HE)* of observed alleles to the equilibrium heterozygosity (*Heq*) estimated from the observed number of alleles in a sample of size *N* under the assumptions of a constant size population and mutation - drift equilibrium. A shared pattern of significantly excessive *HE* compared to *Heq* among loci is indicative of the reduction in allelic diversity at loci following a recent bottleneck (Piry *et al*. 1999). This difference in magnitude of *HE* compared to *Heq* among loci may persist for several generations depending on *Ne*, until a new mutation –drift equilibrium is established. Accordingly, the test is conservative in that only the most recent bottleneck events may be detected.

For each population (or nominated group) tested, the distribution of *Heq* for individual loci was computed by simulating the coalescent process (1000 iterations) of the observed alleles under a two - phased mutation model (TPM) as recommended by Piry *et al*. (1999). The TPM assumes both one step and multi step mutations and may therefore model microsatellite evolution more closely than either the Infinite Allele (IAM) or the Stepwise mutation model (SMM). Each locus is tested for a significant excess or deficit of *HE* relative to *Heq*, with the probability of departure obtained from the simulations. If sufficient loci are examined (N> 5), the Wilcoxon sign-rank test can be used to determine if the magnitude of *HE* is significantly greater than *Heq* across all loci.

2.3.6 Test for neutrality and historical demographic changes

Patterns of nucleotide substitution among mtDNA haplotypes can be tested for concordance with that expected for a sample at mutation- drift equilibrium. Tajima (1989) demonstrated that at equilibrium the absolute number of segregating sites (*S*) and the average number of pair-wise nucleotide differences (π) observed among a sample of haplotypes are approximately equal to the parameter θ (where $\theta = 2N\mu$) and that the magnitude of difference (*D*) between *S* and π has a mean and variance of 0 and 1. Because estimates of *S* are not affected by the frequency of a particular segregating site

in a sample (as it is for π), violations to mutation - drift equilibria manifesting as an excess of recently derived low frequency mutations inflates S relative to π (D assumes a negative value). Conversely, an excess of older intermediate frequency mutations inflates π (D assumes a positive value). Because covariance is expected between the two estimates at equilibrium, significant shifts in the magnitude of D, provides evidence of process(es) affecting a sample. Both selection and changes in *Ne* operating within a population can shift the magnitude of D. For instance, a strongly negative value of D (where there is an excess of low frequency mutations in a sample) may arise by a selective sweep for closely related haplotypes containing an advantageous mutation (Maruyama and Birky 1991). Demographic expansion following either a founder or bottleneck event may also result in the accumulation of derived haplotypes of low frequency in a population and so mimic the effects of a selective sweep. Conversely, strong positive values of D resulting from a reduction of low frequency mutations may result from either balancing selection for older haplotypes or population admixture of previously separate populations. Because mtDNA is normally considered selectively neutral (but see caveats; Ballard and Kreitman 1994, Rand *et al*. 1994), significant shifts in the values of D may be used to infer the effects of historical demographic changes that may still be evident among the variety of haplotypes found in a sample. Fu (1997) developed a statistic (F_S) analogous to Tajima's D, that has greater sensitivity for detecting excesses of new mutations and therefore has more power than Tajima's test for rejecting neutrality due to excessive numbers of rare mutations. Fu's statistic is therefore especially useful for detecting departures from neutrality caused by either selective sweeps or historical demographic expansions (Excoffier and Schneider 1999).

Statistics for both Tajima's D and Fu's F_S were estimated from samples of haplotypes using Arlequin ver. 2.0 (Schneider *et al*. 2000). Statistics measured for pooled samples were only conducted if there was no prior evidence of population segregation among samples as admixture of formerly segregated alleles may inflate the number of intermediate frequency mutations in the sample. Significance of the statistics (D and F_S) was established by a simulation method implemented in Arlequin. Simulation statistics were recomputed from randomly generated samples within a population at equilibrium (with an estimated parameter $\theta = \pi$) and repeated for 5000 simulations to obtain the null distribution of the test statistic and its P value.

CHAPTER 3: **Analysis of genetic structure among Indo-West Pacific *Scylla serrata* populations**

3 .1 Introduction

It is hypothesised that perturbations to oceanographic process experienced within the Indo-West Pacific (IWP) during the Pleistocene affected both genetic and population structure for a broad variety of marine taxa (McMillan and Palumbi 1995, Lavery *et al*. 1996-a, Williams and Benzie 1998). The relevance of this hypothesis may be assessed by comparative genetic studies of other widespread IWP taxa.

In this chapter, I investigate the phylogeographic distribution of mitochondrial DNA (mtDNA) haplotypes sampled from *S. serrata* populations throughout the IWP. I hypothesise that the genetic structure of *S. serrata* populations may be concordant with that reported for other widespread IWP marine taxa. Specifically I expect to see regional genetic partitioning of populations either side of the Indo - Australian archipelago, symptomatic of vicariant separation of a former continuously distributed species. I also expect that partitions in the intraspecific gene tree would coalesce within a Pleistocene time frame. Alternatively, a lack of structure among regional populations may provide some evidence that levels of gene flow among trans-oceanic *S. serrata* populations have been sufficient to obscure the formation of genetic structure within the distribution of this species.

3 .2 Methods

3.2.1 Sampling

To test for evidence of regional genetic partitioning of mud crab populations, a total of 124 individual *S. serrata* were collected from 11 locations representing three regions (Indian, Pacific and northern Australia) within the IWP (Table 3.1 and Figure 3.1). As each of the three regions were potentially isolated from the others during the periods of sea level retreat, it may be expected that genetic variation will be partitioned mostly among regions rather than within. The sampling design used here allows comparisons of pooled population structure both within and among the regions.

Table 3.1 IWP sample regions, locations and sample size (*N*). Site locations as per figure 3.1

REGION	SITE	LOCATION	*N*
West Indian ocean	1	Red sea: Yemen	5
	2	Mauritius	5
	3	South Africa	11
Northern Australia	4	Northern Australia: Roper River	21
	5	Northern Australia: Karumba	14
West Pacific ocean	6	Papua New Guinea: Fly River	3
	7	Eastern Australia: Moreton bay	40
	8	New Caledonia	6
	9	Fiji	7
	10	Solomon Islands	7
	11	Japan: Okinawa	5

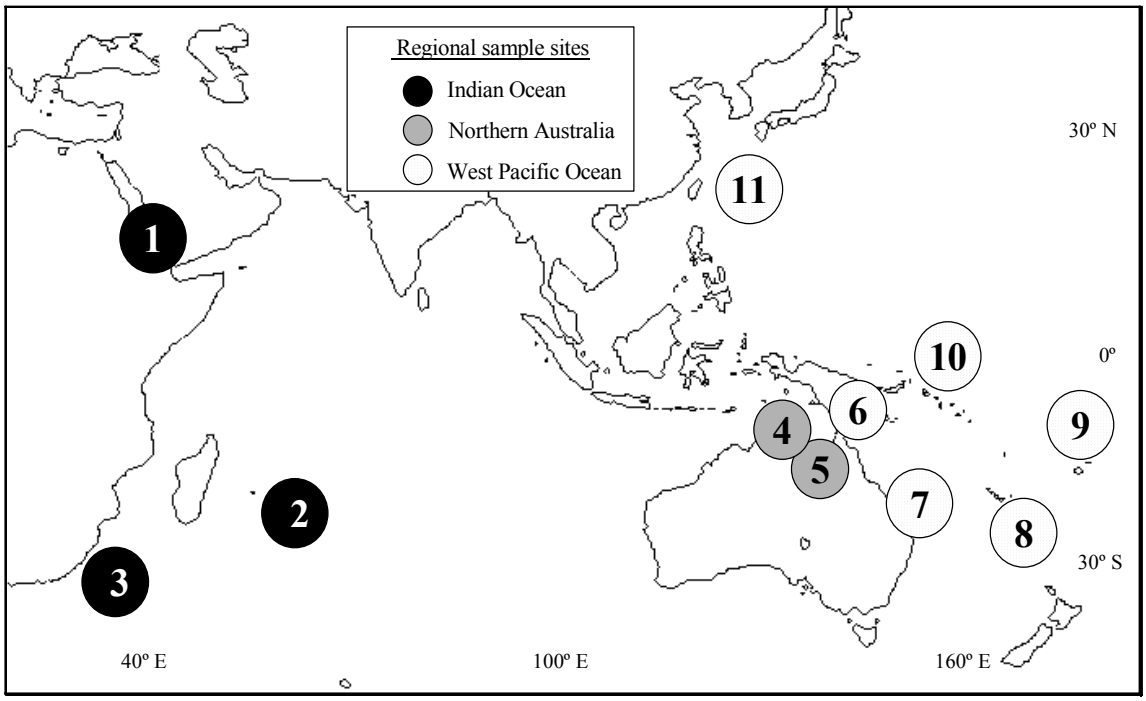


Figure 3.1 Sampling sites throughout the Indo-West Pacific. Regional Indian Ocean sites include: Red Sea (Jeddah) (1), Mauritius (2), South Africa (Durban) (3); northern Australian sites: Roper River (4), Karumba (5); west Pacific sites: PNG (Fly River) (6), eastern Australia (Moreton bay) (7), New Caledonia (8), Fiji (9), Solomon Islands (10), Okinawa (11).

Twenty four additional crab samples identified by morphology as sister species of *S. serrata* (as per Keenan *et al*. 1998) were also included for analysis. The additional 24 *Scylla* specimens were derived from a variety of locations in and near to the Indo-Australian archipelago (Table 3.2). These specimens provided additional analysis of genetic relationships among the species of *Scylla* and permitted a test for the presence of clock-like behaviour of the mtDNA COI gene used for this study.

Table 3.2 Locations and quantities of additional species of *Scylla* sampled for analysis.

Putative species	Sample location	N
S. olivacea	Bangkok	1
	East Papua New Guinea	10
	Taiwan	2
	North west Australia	1
	Northern Australia	3
	Bali	2
S. tranquebarica	Sabah	1
	Bali	2
S. paramamosain	Hong Kong	2

3.2.2 DNA methods

DNA was obtained using sequential phenol/ chlorophorm extractions and used in PCR amplification of the mtDNA COI gene as described in sections 2.2.1 and 2.2.2. Mutational differences between PCR samples were identified using temperature gradient gel electrophoresis - heteroduplex analysis, as described in section 2.2.3. All defined haplotypes were sequenced in two directions using protocols described in section 2.2.4. A subset of the sample ($N = 10$) was subjected to preferential mtDNA extraction using the alkaline lysis procedure and tested for evidence of nuclear inserted homologs as described in section 2.2.5.

3.2.3 Diversity and divergence estimates

Haplotype diversity (h) was calculated for each sample location to allow comparison of mtDNA diversity levels among sample locations. Regional h and nucleotide diversity (π) (Nei 1987) estimated for pooled locations both within and between the Indian,

Pacific and northern Australia regions allowed comparisons of these statistics at several spatial scales. Estimates of π within each of the three major geographical regions were calculated as the mean percent - sequence difference between all paired haplotypes pooled within a region (Nei 1987). Estimates of inter-regional π (or net nucleotide divergence among regions) were corrected for intra-regional polymorphism as described by Nei (1987).

3.2.4 Phylogenetic analysis

Inter-specific relationships within the *Scylla* genus were established by constructing a neighbour-joining tree using all identified *S. serrata* haplotypes and haplotypes for 14 crabs morphologically identified as sister species of *Scylla*. Proportion of base pair differences between haplotypes was used as the distance measure in tree construction. Probability of topology was tested with bootstrap resampling (2000 replicates). A test for clock-like behaviour of the data was conducted using Puzzle version 4.0.2. (Strimmer and von Haeseler 1996) by comparing the ratio of log-likelihood estimates of tip/root branch lengths between trees constructed with and without assumption of linear mutation rate (Felsenstein 1988).

Genealogical relationships among *S. serrata* haplotypes were estimated by building a minimum spanning tree (section 2.3.3) based on the minimum absolute number of mutational differences among pairs of haplotypes.

3 .3 Results

3.3.1 Distribution of Scylla serrata haplotypes

Eighteen distinct *S. serrata* haplotypes were identified from the 11 IWP locations (Table 3.3). Multiple haplotypes were not detected in samples of less than 15 individuals. For the three Australian locations, results suggest greater haplotype diversity in northern Australia than in eastern Australia. Three sites in both the Indian and Pacific regions are fixed for their own unique haplotype (Table 3.3). Haplotype A is the only widespread haplotype being shared not only among Australian sites but also with Papua New Guinea and Okinawa. This widespread haplotype is completely absent from the outer Pacific and Indian Ocean regions. Haplotype A is seen as the majority haplotype (> 80%) at the east Australian location and as a minority haplotype (< 9%) in

northern Australia. Each Australian location has its own array of private haplotypes in low frequency.

Table 3.3 Distribution of *Scylla serrata* haplotypes among 11 Indo-West Pacific sites (*N*: total sample size; *Nh*: number of haplotypes per location; *h*: haplotype diversity, Nei 1987)

LOCATION	HAPLOTYPES																					
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	*N*	*Nh*	*h*
Red Sea										5									*5*	*1*	*0*
Mauritius											5								*5*	*1*	*0*
South Africa												11							*11*	*1*	*0*
Roper River	3														14	2	1	1	*21*	*5*	*0.523*
Karumba													9	1	4				*14*	*3*	*0.500*
Pap. New Guinea	3																		*3*	*1*	*0*
eastern Australia	32	1	2	2	1	2													*40*	*6*	*0.352*
New Caledonia									6										*6*	*1*	*0*
Fiji								7											*7*	*1*	*0*
Solomon Islands							7												*7*	*1*	*0*
Okinawa	5																		*5*	*1*	*0*

3.3.2 Nucleotide composition of Scylla serrata haplotypes

Sequencing of PCR products (597 bp length) produced 549 bp of comparable unambiguous sequence for all haplotypes. Twenty-four sites were identified as variable among the 18 defined haplotypes (Table 3.4). Nucleotide sequences of all 18 haplotypes are deposited with GENBANK under accession numbers AF097002 - AF097019. All but one of these variable sites can be assigned as a 1st or 3rd codon position synonymous transition; hence, the majority of mutations are silent as is expected for intra-specific comparisons of a protein-coding gene (Kocher *et al*. 1989). The remaining variable site (position 523) is a 1st position transition resulting in a change of amino acid (in comparison to consensus sequence A) from alanine to threonine. This variant site/amino acid is only present in six northern Australian haplotypes and fixed for the consensus state in all other haplotypes. The incidence of homoplasy at sites 231 and 333 among the various haplotype sequences is observed. Sequences derived from procedures to check for nuclear insertions were identical to their homologous counterparts indicating there was no evidence of nuclear insertion within the sample.

Table 3.4 Variable nucleotide positions among *Scylla serrata* haplotypes (A to R). Haplotypes are compared with consensus haplotype A, identical sites are denoted by a dot and variant sites by their nucleotide substitution; *P* parsimonious sites, (*N*) number of each haplotype in sample.

Haplotype	Nucleotide position																										(*N*)
	10	15	36	78	81	93	111	153	156	189	198	210	231	234	273	288	318	333	450	475	480	519	523	549			
A	T	A	C	C	G	C	T	G	A	G	A	T	C	A	G	G	T	G	A	T	T	G	G	T			(43)
B	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			(1)
C	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.			(2)
D	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			(2)
E	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			(1)
F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.			(2)
G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.			(7)
H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.			(7)
I	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	C	.	.	.	.			(6)
J	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.			(5)
K	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	G	.	.	.	.	.			(5)
L	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	.			(11)
M	C	G	T	.	A	.	.	.	G	A	G	.	T	.	A	.	C	.	.	.	.	.	A	C			(9)
N	.	G	T	.	A	.	.	.	G	A	G	.	T	.	A	.	C	.	.	.	.	.	A	C			(1)
O	C	G	T	.	A	.	.	.	G	A	G	.	T	.	A	A	C	A	.	.	.	.	A	C			(18)
P	C	G	T	T	A	.	.	.	G	A	G	.	T	.	A	A	C	A	.	.	.	.	A	C			(2)
Q	C	G	T	.	A	.	.	.	G	A	G	.	T	G	A	A	C	A	.	.	.	.	A	C			(1)
R	C	G	T	.	A	.	.	.	G	A	G	?	.	.	A	A	C	A	.	.	.	.	A	C			(1)
	P	*P*	*P*		*P*				*P*	*P*	*P*		*P*		*P*	*P*	*P*	*P*		*P*			*P*	*P*			

3.3.3 Evolutionary relationships among species of Scylla

Reconstruction of genetic relationships among morphologically defined species of *Scylla* using a neighbour-joining tree are concordant with that reported by Keenan *et al*. (1998) and provide no evidence of mtDNA introgression among the species. It can be seen in the tree (figure 3.2) that the four species of *Scylla* are delineated as four well supported groups of haplotypes, with average levels of interspecific sequence difference (~ 12 %) more than six times greater than that observed at the intraspecific level (~ 2 %). Intraspecific relationships within both *S. olivacea* and *S. serrata* contain separate strongly supported clades of closely related haplotypes that respectively differ by a minimum of 1.7 and 2.2 percent sequence difference. It is interesting to note that there is a phylogeographic basis for the separation of clades in *S. serrata* (an IWP group and a

northern Australian group – refer next section), whereas no such geographic separation of clades is observed for *S. olivacea*.

Results of the log-likelihood ratio test could not reject the null hypothesis of clock-like evolution of the COI gene among species of *Scylla* (log-likelihood ratio test statistic delta = 27.31, $P = 0.607$). Temporal estimates of divergence among *Scylla spp*. based on measured sequence differences at COI may therefore not be confounded by unequal rates of evolution among haplotypes at this gene fragment.

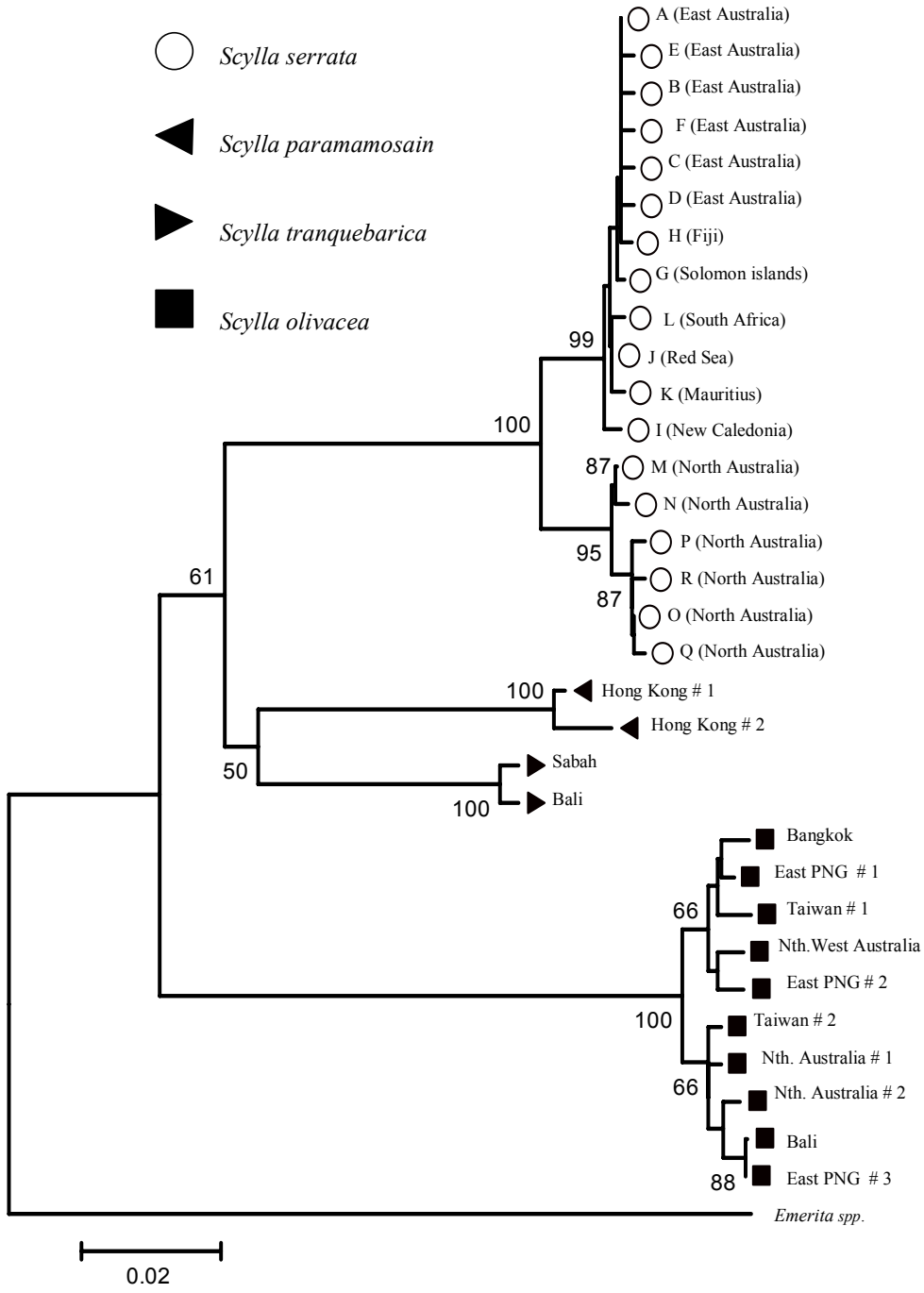


Figure 3.2 Neighbour-joining tree of genetic relationships among species of *Scylla*. Tree based on percentage sequence difference among pairs of mtDNA COI haplotypes. Scale bar equals 2 % sequence difference. Bootstrap *P* values > 50% support shown at nodes (2,000 replicates). Species of *Emerita* used as outgroup (Decapoda: Anomura. GENBANK accession # AF246159).

3.3.4 Evolutionary relationships among Scylla serrata haplotypes

Among the 18 putative haplotypes, 15 of the 24 variable nucleotide sites are parsimoniously shared. The resolved network of possible phylogenetic relationships among haplotypes (Figure 3.3) clearly distinguishes two clades separated by a minimum of 12 bp, which represents a 2.2% sequence difference between clades. Clade 1 contains haplotypes found in Indian and Pacific Ocean regions (and to a minor extent northern Australia), whereas clade 2 haplotypes are confined entirely to northern Australia. Both clades are radial in structure and the haplotypes within each clade are closely related - the average difference between lineages within clades was the same at 0.44% sequence difference. The combination of homoplasy and low sequence divergence within each clade made it difficult to identify a most recent common ancestor (MRCA) within clades. Adopting criteria of Crandall and Templeton (1993) that the most probable MRCA within a network is both geographically widespread and has the most number of haplotype connections then the most probable candidate within clade 1 is haplotype A and for clade 2 haplotype O.

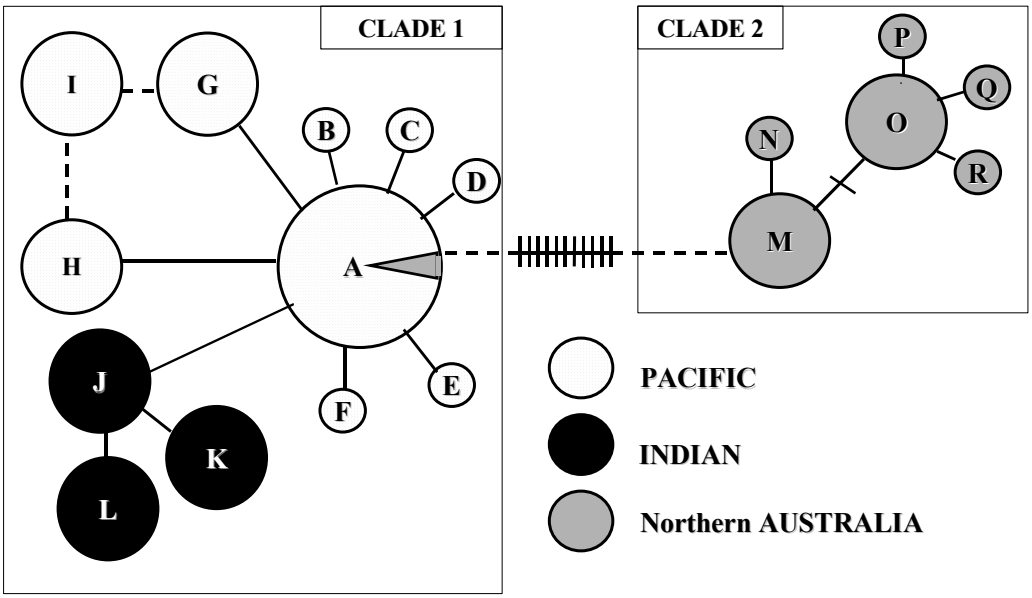


Figure 3.3 Network of phylogenetic relationships among *Scylla serrata* haplotypes. Haplotypes are joined by least number of substitutions between all possible sequence comparisons. Connection between two haplotypes = one mutation; additional mutations are indicated by cross bars, dashed lines = ambiguous connections between haplotypes. Haplotypes (A to R) are labelled as in Table 3.3 and shaded according to regional occurrence. Size of haplotypes illustrates relative frequency within sample (*small circles N* < 3; *intermediate circles N:* 5 to 20; *large circle N* > 20).

Comparisons of intra versus inter-regional nucleotide diversity for the Indian and Pacific haplotypes are similar and are an order of magnitude less than inter-regional comparisons with northern Australia (Table 3.5). Haplotype diversity estimates are similar among all regional comparisons.

Table 3.5 Regional nucleotide and haplotype diversity estimates. π mean percent sequence difference among all sampled *Scylla serrata* haplotypes. Inter-region π corrected for intra - region polymorphism (Nei 1987); *h* haplotype diversity, Nei 1987).

Region	π	*h*
Intra - regional comparison		
Indian Ocean	0.16	0.61
Pacific Ocean	0.21	0.62
northern Australia	0.56	0.69
Inter – regional comparison		
Indian vs Pacific	0.09	0.76
Indian vs northern Australia	1.12	0.80
Pacific vs northern Australia	1.08	0.72

Assessment of the coalescent histories of clades 1 and 2 was achieved by converting the observed maximum percent sequence difference (d_{MAX}) to the most recent common ancestor (MRCA) within a clade to units of time, thereby extrapolating the relative maximum timescale over which events may have occurred. Thus time since divergence (T) between two groups was calculated as: $T = d_{MAX}/2\mu$ where μ is the rate of nucleotide substitution. For this study, it is assumed that mitochondrial COI lineages accumulate neutral mutations at a linear rate of 1.15% per million years, which is the equivalent of 2.3% divergence between lineages per million years. Brower (1994) approximated this rate estimate from levels of intraspecific mtDNA divergence observed in five arthropod taxa with independently dated divergence times. Using this measure, the maximum age of coalescence to the MRCA within each clade was approximately 250,000 yr.'s bp, whereas the between clade estimate of time to coalescence was slightly greater than 1 Myr.'s bp (Table 3.6).

Table 3.6 Estimates of maximum coalescent events relative to the most recent common ancestor (MRCA) of clade 1 and 2. Where d_{MAX} = maximum percent sequence difference from a MRCA observed within a sample; T = estimated time to coalescence in years).

Coalescent event	d_{MAX}	T
MRCA clade 1	0.546	237 000
MRCA clade 2	0.546	237 000
MRCA clade 1 & 2	2.550	1 109 000

3 .4 Discussion

MtDNA analyses of sampled *Scylla serrata* populations throughout the IWP indicate several distinct patterns of genetic association within the species distribution. Perhaps the most intriguing result to arise within this study is the presence of two mtDNA lineages within this species (referred to here as clade 1 and clade 2 crabs, Figure 3.2 and 3.3). The two clades consist of distinct haplotype assemblages that differ by approximately 2 % sequence divergence (Table 3.6). This level of divergence suggests a coalescent ancestry between the two clades dating back approximately 1 Myr.'s before present (Table 3.6). The demographic histories of these two clades appear to have proceeded along independent trajectories resulting in markedly different patterns of distribution and population genetic structure.

I hypothesised that vicariance resulting in separation of Indian from Pacific Ocean populations as observed for several IWP marine species, would similarly have affected the population genetic structure of *S. serrata*. Results for *S. serrata* are contrary to this hypothesis and instead suggest that radiation throughout the IWP occurred as a single rapid wave of expansion possibly emanating from a west Pacific origin during the late Pleistocene. For example, mean corrected percent sequence difference between the Indian and Pacific regions are similar to uncorrected levels of differences seen within each of these two regions (Table 3.5). The network of haplotype relationships (Figure 3.3) indicates that all lineages from Indian and Pacific localities are nested within clade 1 and are ultimately derived from a single common ancestor - currently widespread in the West Pacific (haplotype A, Table 3.3). The low level of sequence divergence within this clade indicates a recent origin for the expansion of *S. serrata* populations throughout the IWP. If the clock calibrations accurately reflect the cumulative mutation rate for mtDNA, then the estimated timing of genealogical coalescence for all clade 1

haplotypes falls within the late Pleistocene (Table 3.6). Surprisingly the presence of fixed unique haplotypes in each of the populations peripheral to the West Pacific core, suggests that intervening maternal gene flow since the initial radiation has been insufficient to overcome the effects of local random genetic drift. The lack of effective maternal gene flow among populations appears to have been sufficient to allow the development of distinct lineages in discrete populations through mutation and random lineage assortment (Neigel and Avise 1986).

Estimated total haplotype diversity for *S. serrata* is comparable to that reported for other widespread IWP species (Table 3.7). However, apart from that seen in northern Australia, individual populations of *S. serrata* are characterised by low haplotype diversity (Table 3.3). The observed low diversity may be an artefact of the small sample size; seven of the eleven locations each had fewer than ten individuals sampled (Table 3.3). However, comparison of haplotype diversity estimates among the various taxa with comparable sample sizes in Table 3.7 suggest that average diversity within *S. serrata* populations is depauperate relative to that seen for other widespread plankton dispersed taxa. Periodic fluctuations of local effective population size resulting in population bottlenecks may have contributed to the reduction of haplotype diversity at these locations. These fluctuations may have allowed the chance fixation of locally derived haplotypes, especially under conditions where there is a lack of sustained gene flow from neighbouring populations (Neigel and Avise 1986).

Table 3.7 Comparison of mtDNA haplotype diversity estimates for seven species of invertebrates distributed throughout Indo-West Pacific (IWP) and Pacific locations (Pac). *N* and *N* (ave) total and average per site sample size; *h* total sample haplotype diversity, *h* (ave) average per site haplotype diversity; estimates of *h* for all species (except *B. latro*) calculated using raw data derived directly from cited studies.

Species	Sites	*N*	*h*	*N* (ave)	*h* (ave)	Source
Birgus latro	8 IWP	160	0.98	19	0.97	Lavery *et al*. 1996-a
Echinometra oblonga	4 Pac	37	0.79	9	0.62	Palumbi *et al*. 1997
Echinometra mathaei	8 Pac	60	0.84	7.5	0.57	" "
Echinometra sp.nov.A	7 Pac	67	0.97	9.6	0.82	" "
Echinometra sp.nov.C	4 Pac	39	0.92	9.8	0.81	" "
Linckia laevigata	17 IWP	370	0.89	21.8	0.74	Williams & Benzie 1997 & 98
Scylla serrata	11 IWP	124	0.79	11.3	0.13	Present study

Contrary to the insular distribution of haplotypes among peripheral populations, the distribution of the clade 1 most recent common ancestor (haplotype A) is immense, spanning a distance of approximately 6,500 km. from Okinawa to east Australia. This north - south distribution testifies to extensive levels of gene flow along the western margin of the Pacific. Similar patterns of extensive West Pacific gene flow are reported for several marine species including a number of giant clams (Benzie and Williams 1997) and a species of sea urchin (Palumbi *et al*. 1997). Remarkably these gene flow patterns are contrary to present day ocean currents, which suggests that episodes of population expansion and gene flow along the West Pacific margin may have been far greater during the Pleistocene than they are at present (Benzie and Williams 1997).
I suspect that conditions during the Pleistocene may have facilitated episodes of *S. serrata* population expansion resulting in the widespread distribution of the species within the West Pacific. Levels of gene flow may have been sufficient to provide both connectivity among populations via the rapid spread of the MRCA haplotype and hinder random fixation of alternative haplotypes among these locations. It may be surmised that fortuitous dispersal effected by clade 1 crabs during the Pleistocene was sufficient to found widely distributed IWP populations. Intervening levels of dispersal do not appear to have ensured genetic panmixia throughout the IWP suggesting that historically, trans-oceanic dispersal and colonisation have occurred in a sporadic context for this species.

The distribution and genetic structure of clade 2 crab populations contrasts with that seen for clade 1 crabs. Clade 2 crabs are strictly confined to populations within northern Australia and the high haplotype diversity estimates within these populations are markedly greater than that seen in locations outside of this region (Table 3.3). These features may be indicative of a long-term stability of clade 2 populations within the region. This second clade differs from the IWP clade by approximately 2% sequence divergence. The estimated coalescence time for these two clades (Table 3.6) falls within the early Pleistocene, which pre-dates that estimated for the most recent IWP expansion by *S. serrata* by approximately 850,000 years. These results are surprising given the apparent rapid radiation of clade 1 crabs throughout the IWP. Several scenarios may explain this unusual phylogeographic pattern. The first scenario supports my initial hypothesis; i.e., the geographically restricted clade 2 crabs represent the remnants of formerly widespread Indian Ocean populations that extended to northern Australia but were separate from the Pacific. Subsequent retractions of Indian Ocean populations

during periods of glacial activity may have led to extinction of the clade in all but the northern Australian locations. These retractions would necessarily have had to occur before the current radiation of the clade 1 *S. serrata* back into the Indian Ocean. Abundant fossil evidence of *S. serrata* reported from an early Pleistocene formation in South Africa (Cooper and Kensley 1991) clearly predates the estimated timing of the current clade 1 IWP radiation, but does lend support to the existence of prior Indian Ocean populations of *Scylla spp.* dating back at least 1 Myr.'s bp.

An alternative scenario to the Indian Ocean expansion-retraction hypothesis is that clade 2 crabs have evolved *in-situ* within or around the northern Australian region. Aspects of this region may have experienced periods of allopatry, isolated from both the Indian and Pacific Oceans because of Pleistocene sea level and tectonic fluctuation (Chappell 1976, McManus 1985). The marine environment of northern Australia has been noted as a potential biological province separate from the IWP for a variety of taxa (Briggs 1995). High levels of endemism within this region have been reported for gastropods (Laseron 1956), echinoderms (Endean 1957), sponges (Bergquist 1967), shore-fish (Wilson and Allen 1987), and estuarine brachyurans (Davie 1985). The biological provinciality of this region may have its origins in vicariance, although this issue has never been fully explored. The presence of haplotype A as a low frequency clade 1 haplotype within northern Australia (Table 3.3) suggests that gene flow into this region from the IWP has occurred. The absence of any lineage derived from haplotype A within northern Australia suggests that this gene flow has been recent. Given the high degree of potential isolation of biota within the northern Australian region, it may be pertinent to speculate as to whether clade 2 crabs represent a remnant race of *S. serrata* or whether they represent the descendents of an incipient speciation event. If clade 2 crabs represent a distinct species separate from clade 1 crabs, then it is expected that a lack of hybridisation between the species will result in a pattern of fixed nuclear DNA differences between the two clades. Comparison of nuclear DNA between clade 1 and 2 crabs derived from the same location may provide a means of testing this prediction; this will be investigated in Chapter 5.

CHAPTER 4: **MtDNA Analysis of Australian *Scylla serrata* populations**

4 .1 Introduction

Results presented in the previous chapter indicate contemporary IWP *S. serrata* populations are derived from a historical population expansion, which occurred during the Pleistocene under conditions that permitted widespread maternal gene flow. The absence of shared haplotypes among the majority of IWP locations indicates however that contemporary gene flow among these locations is infrequent and therefore dispersal of *S. serrata* over trans-oceanic scales is limited. It is unknown if levels of gene flow, as promoted by contemporary dispersal, are more prevalent among *S. serrata* populations located within the bounds of a continental shelf.

The Australian distribution of *S. serrata* is confined to tropical-temperate waters, spanning over 7000 kilometres of shelf-connected coastline. The structure of populations along this distribution may be considered recent given that contemporary Australian estuaries developed after the last marine transgression 8000 years ago (Kench 1999). As identified in chapter 3, two distinct clades of mtDNA haplotypes occur within Australian waters. One clade is associated with a widespread IWP lineage (clade 1 crabs), the other endemic to northern Australia (clade 2 crabs). It is possible that the Australian distribution contains regional *S. serrata* populations derived from historically independent lineages. Regional genetic structure has also been observed for several marine species distributed among tropical-temperate Australian coastal areas (Keenan 1994, Elliot 1996, Duke *et al*. 1998, Williams and Benzie 1997, Benzie 2000). These species generally show a concordant pattern of genetic separation between eastern and western populations, attributed to periods of population separation associated with Pleistocene sea level changes.

As suggested in Chapter 3, modifications of Australian *S. serrata* populations by episodes of eustatic variation in northern Australia may have led to the formation of regional phylogeographic structure similar to that reported for other co-distributed species. The extent of admixture among these regional crab populations following their formation may reflect contemporary patterns of dispersal for the species. Here I examine in detail, mtDNA genetic structure among adult *S. serrata* populations sampled from three Australian coastal shelf regions (western, northern and eastern). The regions are contiguous but differ by broad-scale patterns of hydrology. I use this information to

estimate the scale of genetic connection within and among shelf - connected populations of *S. serrata* and to identify factors that may have influenced the Australian distribution.

4 .2 Methods

4.2.1 Sampling and DNA methods

For the purpose of hierarchical comparisons, I sampled multiple locations within three coastal Australian regions (western, northern and eastern) (Table 4.1 and Figure 4.1). Sampling from 11 Australian locations provided over 300 post-larval *Scylla serrata* (Table 4.1). The collected *S. serrata* were distinguished in the field from the small brown mud crab *S. olivacea*, a sister species sympatric with *S. serrata* in the northern and western Australian regions (Taylor 1984, Keenan *et al.* 1998). I screened crabs for variation in a 549 base pair portion of the mitochondrial COI gene using temperature gradient gel electrophoresis (TGGE) in conjunction with heteroduplex analysis. Methods of DNA extraction, PCR, TGGE and sequencing are as described in Chapter 2.

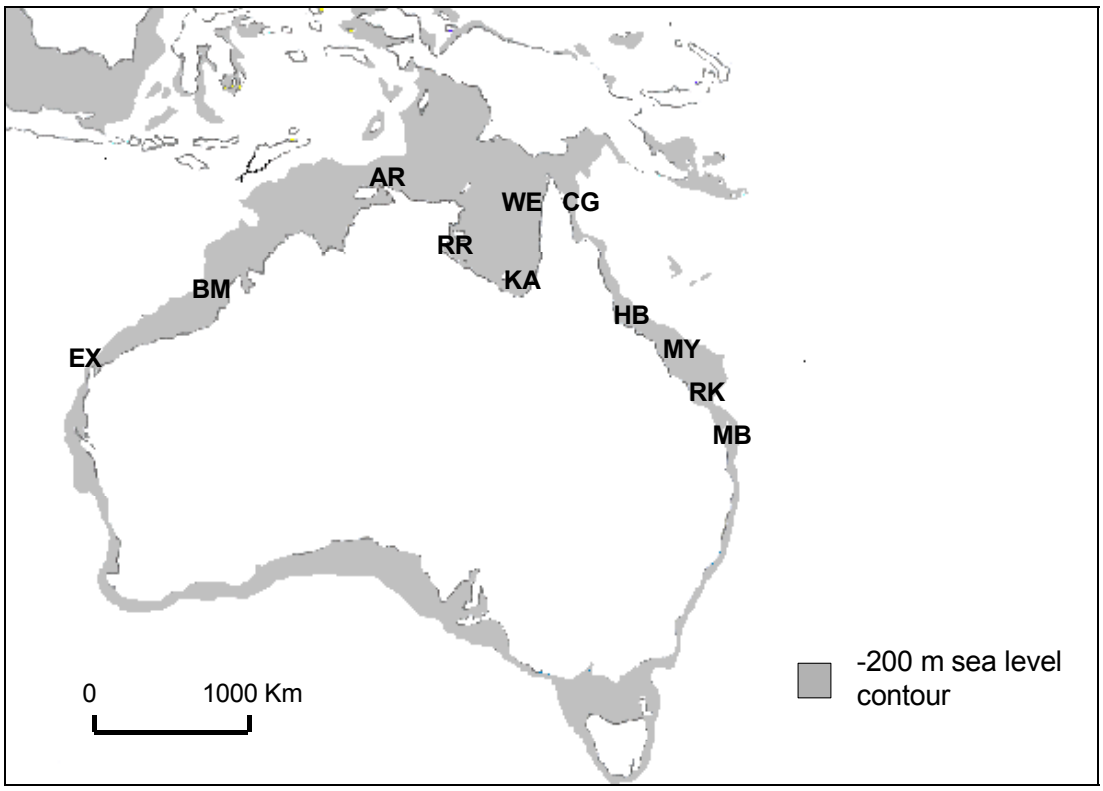


Figure 4.1 Australian locations sampled for *Scylla serrata*: EX, Exmouth; BM, Broome; AR, Adelaide River; RR, Roper River; KA, Karumba; WE, Weipa; CG, Cape Grenville; HB, Hinchinbrook; MY, Mackay; RK, Rockhampton; MB, Moreton bay.

Table 4.1 Sampled Australian locations, geographical co-ordinates, sample size (*N*) and regional grouping of locations.

Region	Location	Co-ordinates	(*N*)
Western	Exmouth	22°15S 114°15E	24
	Broome	18°0 S 122°15E	8
Northern	Adelaide River	12°14S 131°16E	16
	Roper River	14°43S 135°27E	31
	Karumba	17°31S 140°50E	14
	Weipa	12°45S 141°30E	35
Eastern	Cape Grenville	11°57S 143°12E	22
	Hinchinbrooke	18°32S 146°45E	21
	Mackay	21°36S 148°39E	34
	Rockhampton	23°22S 150°32E	22
	Moreton bay	27°25S 153°02E	76

4.2.2 Diversity and population structure

I estimated levels of mtDNA polymorphism among crabs within locations and regions using Nei's (1987) statistic of haplotype diversity (*h*) and associated sampling variance. Pair-wise F_{ST} estimates between all sample locations provided comparative measures of sub-division between locations, F_{ST} calculations incorporated the proportion of sequence divergence between haplotypes. Exact tests of population differentiation (Raymond and Rousset, 1995) were employed to test if haplotypes were randomly distributed (panmixia) among population pairs. Reported probability (*P*) values for the pair wise tests were estimated from 10,000 steps in a markov chain. I also examined the effect of spatial scale of sampling on population structure using an analysis of molecular variance (AMOVA) as described by Excoffier *et al*. (1992). Sample locations in the AMOVA were grouped at three hierarchical levels (among regions, among localities within regions, and within locations) and examined for the percentage of overall genetic variation apportioned to each level. Two AMOVA's were constructed - one used haplotype frequency data only (haplotype matrix), where as the other incorporated percent sequence difference between haplotypes as well as frequency data (nucleotide matrix).

4.2.3 Phylogenetic reconstruction and neutrality test

As in chapter 3, phylogenetic relationships among haplotypes were constructed using a neighbour-joining tree and a minimum spanning tree but using only *S. serrata* haplotypes for analysis. Tests for significance of topology were implemented for both methods. Distances generated between haplotypes and used as data in tree construction incorporated various assumptions and evolutionary rates - but provided negligible differences in topology.

Populations were tested for significant deviation from mutation – drift equilibrium by estimating Tajima's *D* and Fu's F_S statistics and associated probability estimates as described in section 2.3.6.

4 .3 Results

4.3.1 Haplotype abundance, distribution and composition

Twenty-three distinct haplotypes were identified among the sample of 306 *Scylla serrata* (Table 4.2). The most frequent haplotype (1A) was present within all three regions, though it was far more abundant within the eastern region. The frequency of this haplotype abruptly shifted between sample locations Cape Grenville and Hinchinbrook. Apart from 1A, none of the haplotypes detected here are also found outside of Australian waters. Haplotype 1A is synonymous with that previously described (Chapter 3, haplotype A) as the most recent common ancestor of clade 1 haplotypes in the IWP. There was no other sharing of haplotypes among regions, though some sharing was evident among locations within regions. Eleven haplotypes were endemic to single locations, and collectively represented 12 percent of all individuals sampled.

There was no evidence among sequenced haplotypes of any insertions, deletions, or mutations outside of those expected for a protein-coding gene surveyed within a species. Twenty-nine variable sites were detected among the 23 haplotypes. Apart from a single amino acid change resulting from a 1st position transition mutation (position 523 previously reported in chapter 3) all other variable sites are 1st and 3rd site silent transitions (Table 4.3).

Table 4.2 Distribution and count of *Scylla serrata* haplotypes among Australian locations.

LOCATION	HAPLOTYPES																				
	1A	1B	1C	1D	1E	1F	1G	1H	1I	1J	2A	2B	2C	2D	2E	2F	2G	2H	2I	2J	2K	2L
Moreton Bay	65	2	3	4	1	4																
Rockhampton	19						2	1														
Mackay	27		1	4			1		1													
Hinchinbrook	19		1			1																
Cape Grenville	1									21												
Weipa	1										9					18	3	3	1			
Karumba											4					9	1					
Roper River	3										20	5	1	1	1							
Adelaide River	1										10	2				3						
Broome	1																			5	2	
Exmouth	3																			18	2	1

Table 4.3 Variable nucleotide sites among *Scylla serrata* haplotypes for 549 base pairs of the mtDNA CO1 gene. Haplotypes compared to consensus sequence 1A (deposited GENBANK accession # AF097002). Synonymous sites denoted by a dot, variant sites as per type of nucleotide substitution: A, C, G and T.

Haplotypes	Nucleotide position																													
	0 0 3	0 1 0	0 1 5	0 3 6	0 5 7	0 7 8	0 8 1	0 9 3	1 1 1	1 3 5	1 5 3	1 5 6	1 6 8	1 8 9	1 9 8	2 1 0	2 3 1	2 3 4	2 4 0	2 7 3	2 8 8	3 1 8	3 3 3	4 8 0	4 8 9	5 2 3	5 2 8	5 3 1	5 4 9
1A	T	T	A	C	A	C	G	C	T	T	G	A	T	G	A	T	C	A	T	G	G	T	G	T	T	G	A	G	T
1B	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1D	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1E	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
1G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.
1H	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
1J	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
2A	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	A	C	A	.	.	A	.	.	C	
2B	.	C	G	T	.	T	A	.	.	.	.	G	.	A	G	.	T	.	.	A	A	C	A	.	.	A	.	.	C	
2C	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	G	.	A	A	C	A	.	.	A	.	.	C	
2D	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	.	.	.	A	A	C	A	.	.	A	.	.	C	
2E	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	C	A	A	C	A	.	.	A	.	.	C	
2F	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	.	C	.	.	.	A	.	.	C	
2G	.	.	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	.	C	.	.	.	A	.	.	C	
2H	.	C	G	T	G	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	A	C	A	.	.	A	.	.	C	
2I	.	C	G	T	.	.	A	.	.	.	.	.	.	A	G	.	T	.	.	A	A	C	A	.	.	A	.	.	C	
2J	.	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	.	C	A	.	.	A	.	.	C	
2K	.	C	G	T	.	.	A	.	.	.	.	G	C	A	G	.	T	.	.	A	.	C	A	.	.	A	.	.	C	
2L	C	C	G	T	.	.	A	.	.	.	.	G	.	A	G	.	T	.	.	A	.	C	A	.	.	A	.	.	C	

4.3.2 Haplotype diversity and neutrality test

Haplotype diversity within locations ranged from 0.091 to 0.672 biased towards lower diversity within eastern region locations (Table 4.4). Haplotype diversity differed significantly between the northern and eastern regions ($P < 0.01$ Mann-Whitney U test).

The few clade 1 haplotypes present within northern and western region populations were removed before neutrality analysis, as inclusion of these haplotypes would artificially inflate estimated values of π at those locations. Test's of neutrality (Tajima's *D* and Fu's F_S) generally indicated that eastern region populations had an excess of low frequency derived mutations among haplotypes (negative values of D and F_S) where as the reverse was apparent among the majority of western and northern populations (Table 4.4). Roper River was the exception among the northern populations, both

neutrality statistics were negative (but non-significant) due to the occurrence of several low frequency derived haplotypes. Significant deviations from neutrality occurred at the majority of eastern region populations for estimates of F_S. Most of the same populations were also significant for Tajima's *D*. Neutrality test's were not conducted for regions as there is some indication of subdivision within both the northern and eastern regions (refer 4.3.4).

Table 4.4 Haplotype diversity and neutrality estimates among Australian samples. Sample size (*N*), number of haplotypes detected (*haps*), diversity of haplotypes (*h*) and associated variance. Neutrality statistics include Tajima's *D* (*D*) and Fu's F_S (F_S) and associated probability values (*P*) of having a greater observed statistic than random generated from 5000 simulations.

Location	*N*	*haps*	*h*	*D*	P (*D*)	F_S	*P*(F_S)
Exmouth	24	4	0.431 ± 0.117	-0.710	0.258	3.925	0.948
Broome	8	3	0.543 ± 0.184	0.559	0.844	3.091	0.922
(*Western total*)	32	4	0.466 ± 0.097				
Adelaide River	16	4	0.592 ± 0.122	0.035	0.565	0.780	0.628
Roper River	31	6	0.563 ± 0.094	-0.129	0.623	-0.122	0.825
Karumba	14	3	0.539 ± 0.115	0.255	0.636	0.901	0.659
Weipa	35	6	0.672 ± 0.064	0.304	0.669	0.346	0.657
(*Northern total*)	96	10	0.698 ± 0.033				
Cape Grenville	22	2	0.091 ± 0.081	-1.162	0.150	-0.957	0.075
Hinchinbrook	21	3	0.186 ± 0.110	-1.514	**0.047**	-1.920	**0.012**
Mackay	34	5	0.364 ± 0.100	-1.493	**0.048**	-3.086	**0.001**
Rockhampton	22	3	0.255 ± 0.116	-1.175	0.085	-1.310	**0.020**
Moreton bay	79	6	0.320 ± 0.067	-1.464	**0.032**	-4.126	**0.001**
(*Eastern total*)	178	10	0.443 ± 0.044				

4.3.3 Haplotype phylogeny and phylogeography

Both neighbour-joining and minimum spanning trees resolved four haplotype groups nested within two distinct clades (Figures 4.2 and 4.3). These two clades are synonymous with those described in Chapter 3. Additional sampling employed in this study failed to detect any intermediate haplotypes between the two clades. The average level of sequence difference between clades (> 2.2 %) was an order of magnitude greater than within clades (clade 1 = 0.36%, clade 2 = 0.43%). Clade 1 haplotypes

clustered as a single monophyletic group, whereas clade 2 haplotypes fell into three closely related groups. Differences between and within the three internal groups of clade 2 were minimal (~ 0.25 %), with varied statistical support for the internal topology (figure 4.2). Regional phylogeographic association of haplotypes was evident in several instances. Clade 2 haplotypes were prevalent throughout the western and northern regions but totally absent from the eastern region (Figure 4.2 and 4.3). There was also regional assortment of clade 2 haplotype groups; the western and the northern regions each contained mutually exclusive groups of clade 2 haplotypes indicating some degree of regional lineage sorting. In contrast, Clade 1 haplotypes were present in all three regions - by virtue of the presence of haplotype 1A. All other Clade 1 haplotypes were restricted to eastern locations.

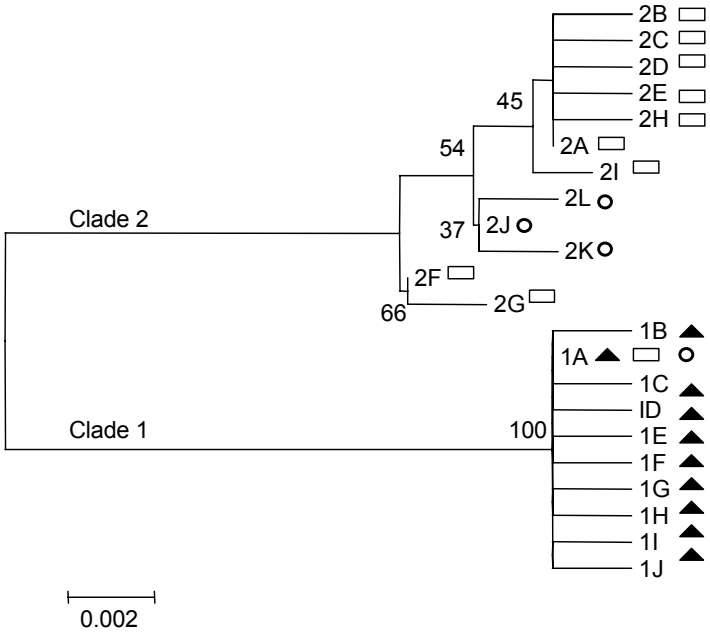
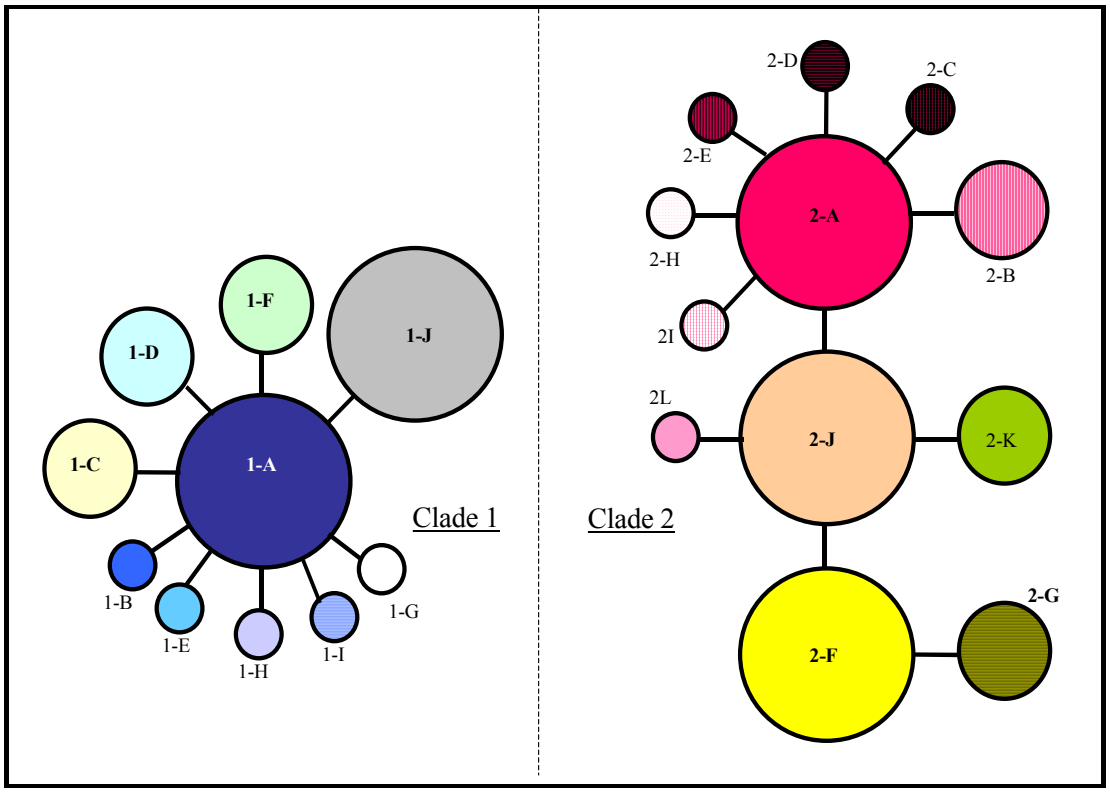


Figure 4.2 Unrooted Neighbour-joining tree of haplotype relationships among Australian *Scylla serrata*, based on proportion of nucleotide differences. Percent levels of bootstrap support indicated at nodes. Regional occurrence of haplotypes indicated (western = circle, northern = rectangle, eastern = filled triangle). Scale bar equals 0.2 percent sequence difference.

a.



b.

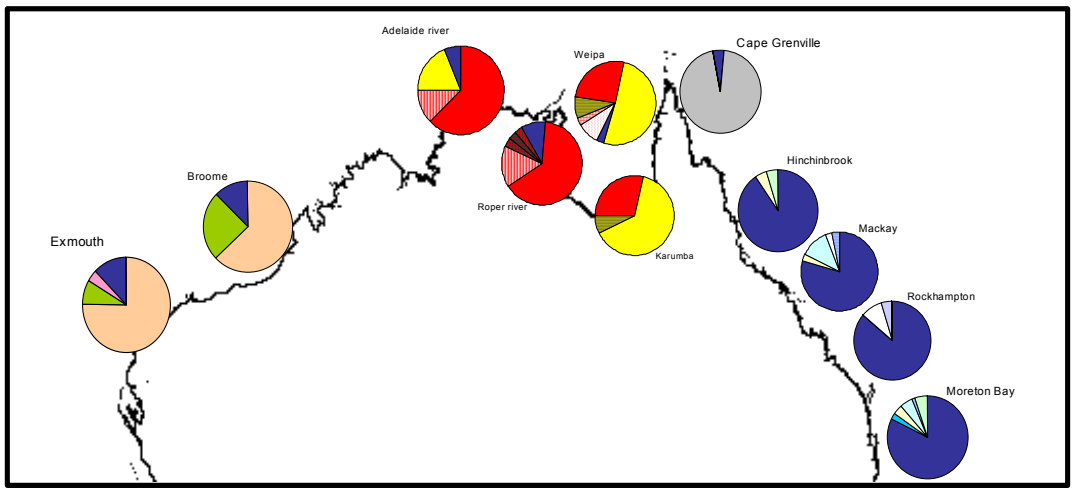


Figure 4.3 Network of haplotypes (a) within clade 1 and 2 and their distribution (b) among Australian locations. Size of haplotypes in (a) correlates with their absolute abundance (*N*) within the entire sample, where small circles $N < 4$, intermediate circles $N = 4 – 10$ and large circles $N > 10$.

4.3.4 Population structure

Fixation indices (Φ) were highly significant (P < 0.001) at all levels of hierarchical sampling in the AMOVA's indicating evidence of strong population structure within each of the levels of spatial sampling. AMOVA generally supported a prevailing pattern of regional structure using both the haplotype and nucleotide matrices (Table 4.5), though the effect was much stronger in the latter analysis where over 86% (as opposed to 37%) of genetic variation was due to the regional partitioning of locations. The within population variance component of the nucleotide matrix (~11%) exceeded that among populations within regions (~2.5%), which indicated that although moderate levels of haplotype variation exists within regions it is mostly shared among the constituent populations (although there is some evidence for significant structure occurring within regions). Similarly, the least amount of variation in the haplotype matrix occurred among populations within regions (< 20 %), however this matrix indicated variation within populations (> 42%) was more pronounced than among regions (> 37%). The difference in magnitude of percent variation among regions between the two matrices (87% Vs. 37%) was due to the segregation of distinct lineages of closely related haplotypes among the regions. This had the effect of inflating the "among region" component of variation within the nucleotide matrix but had no bearing within the haplotype matrix as all haplotypes in this matrix are treated as equidistant. Thus the AMOVA analysis indicated there was a strong phylogeographic basis for the regional partitioning of genetic structure.

Table 4.5 Hierarchical analysis of molecular variance among *Scylla serrata* populations within and among three regions of Australia. Table includes both haplotype frequency and nucleotide (proportion of pair wise sequence difference) matrices partitioned for Australian locations. Calculation of fixation indices (Φ) and partitioning of total genetic variance as per Excoffier *et al*., (1992). *P* is the probability of getting a value larger than the estimate calculated from 10,000 random permutation test.

Variance Component	Variance %	Φ statistic	*P*
Haplotype frequency matrix			
Among regions	37.37	$\Phi_{CT} = 0.374$	< 0.001
Among populations within regions	19.67	$\Phi_{SC} = 0.314$	< 0.001
Within populations	42.96	$\Phi_{ST} = 0.570$	< 0.001
Nucleotide matrix			
Among regions	86.65	$\Phi_{CT} = 0.866$	< 0.001
Among populations within regions	2.44	$\Phi_{SC} = 0.183$	< 0.001
Within populations	10.91	$\Phi_{ST} = 0.891$	< 0.001

The low levels of variation among populations within regions reported in AMOVA, is generally supported by the pair-wise test of population sub-division, however there are some exceptions. Pair-wise population F_{ST} estimates and exact tests of population differentiation (Table 4.6) indicated no evidence of sub-division between the two sample locations of the western region. Similarly within the eastern region, there was considerable sharing of haplotypes among four eastern locations from Hinchinbrook to Moreton bay. Pair-wise F_{ST} estimates indicated equally high levels of genetic similarity (average $F_{ST} < 0.015$) among these four eastern locations regardless of their geographic proximity to each other. Surprisingly, Cape Grenville at the northern periphery of the eastern region displayed significant genetic isolation ($P < 0.0005$) from all other eastern locations (Figure 4.3). Cape Grenville was dominated by an endemic haplotype (1J) that was part of the eastern clade, but absent from all other populations. In contrast to the eastern and western regions, there was significant genetic subdivision between two pairs of northern populations (Table 4.6). Populations from Adelaide River and Roper River were homogenous for haplotype composition but differed significantly ($P < 0.05$) from those at Karumba and Weipa. Although the two population pairs differed primarily by the relative frequency of haplotypes 2A and 2F, there was also evidence that other clade 2 haplotypes differed in frequency between these areas (Table 4.6, Figure 4.3).

Table 4.6 Pairwise F_{ST} estimates and exact test of population differentiation among Australian locations. F_{ST} (upper matrix) based on nucleotide content and haplotype frequencies. Exact test probabilities of non-differentiation (lower matrix) calculated from 10,000 markov steps, significance values as indicated except (* = $P < 0.05$) & (*** = $P < 0.0005$). See Figure 4.1 for sample location codes.

	LOCATIONS										
	EX	BM	AR	RR	KA	WE	CG	HB	MY	RK	MB
EX		-0.066	0.144	0.195	0.164	0.143	0.863	0.846	0.865	0.846	0.909
BM	0.648		0.086	0.136	0.165	0.118	0.911	0.895	0.908	0.894	0.940
AR	***	***		-0.024	0.167	0.094	0.919	0.907	0.917	0.906	0.944
RR	***	***	0.305		0.245	0.190	0.872	0.858	0.875	0.859	0.913
KA	***	***	*	***		-0.031	0.967	0.959	0.955	0.956	0.965
WE	***	***	*	***	0.921		0.906	0.895	0.904	0.894	0.931
CG	***	***	***	***	***	***		0.867	0.770	0.837	0.758
HB	***	***	***	***	***	***	***		0.019	0.018	-0.019
MY	***	***	***	***	***	***	***	0.325		0.025	0.006
RK	***	***	***	***	***	***	***	0.495	0.257		0.025
MB	***	***	***	***	***	***	***	0.957	0.255	0.123	

4 .4 Discussion

4.4.1 Genetic structure among shelf connected Scylla serrata populations.

MtDNA haplotypes are regionally structured among Australian *S. serrata* populations. It is evident from the analysis of molecular variance (Table 4.5) that there is more subdivision among regions than within them. That each of the three regions contains an insular assemblage of related haplotypes indicates the effect of historical phylogeographic process is still evident in the mtDNA profile of *S. serrata* populations (see 4.4.2). Persistence of these independent signatures demonstrates that effective levels of recent maternal gene flow among regions are minimal. This is particularly apparent among populations either side of the Torres Strait. The Strait forms a conduit between the Coral Sea and the Gulf of Carpentaria and provides a potential avenue for gene flow between the eastern and northern Australian regions (Marsh and Marshall 1983). The closest sampled locations on either side of this Strait (Weipa and Cape Grenville) share only haplotype 1A (Table 4.2). A similar pattern is present between the western and northern regions. Apart from 1A, there is complete regional segregation of all other haplotypes. The presence of 1A as a low frequency clade 1 haplotype within northern and western Australia may have resulted from secondary introgression into these regions during the Pleistocene expansion of the species (Chapter 3). The absence of lineages derived from haplotype 1A within northern and western Australia suggests that the introgression may be recent.

In contrast to that seen among regions, genetic homogeneity is evident among shelf-connected populations separated by over 1000 kilometres. Although there are instances of genetic sub-division within the northern and eastern regions, the scale of haplotype sharing among most locations is consistent with high levels of maternal genetic exchange. These levels of genetic exchange may also reflect historical mtDNA gene flow following the most recent colonisation of Australian shelf regions. Assuming sufficient time since colonisation, the instances of genetic subdivision both among and within regions may indicate that populations are approaching equilibrium between the effects of gene flow and drift.

Genetic connectivity among *S. serrata* populations may correspond with prevalent water circulation patterns within coastal-shelf regions. The data generally indicate that populations connected by coastal currents exhibit genetic panmixia, whereas abrupt

shifts in haplotype frequency and significant genetic subdivision coincide with zones of hydrological divergence. This correspondence is most evident within the eastern region. The East Australian Current (EAC) and long-shore winds induce movement and mixing of shelf water parallel to the eastern coast (Wolanski and Pickard 1985, Church 1987). Larval dispersal models indicate the northern component of the eastern region is potentially closed to recruitment from highly connected southern locations due to the reduced impact of the EAC on coastal shelf waters north of latitude 15 ° (Dight and James 1994). Results here indicate a clear genetic bifurcation within the eastern region generally concordant with this pattern. Mud crabs sampled from Cape Grenville at the top end of the eastern region are genetically distinct from populations south of 15 ° which display genetic panmixia at a scale over 1000 km's (Table 4.6, Figure 4.3). Cape Grenville is dominated by one haplotype (1-J) that may be derived from but is not seen among its southern counterparts.

Correlation between coastal hydrology and genetic connectivity is also apparent in the western region. Hydrological connectivity within this region is influenced by pole-ward flowing offshore circulation (Wilson and Allen 1987) and there is no indication of genetic subdivision among populations separated by over 1000 km's (Table 4.6). A different picture emerges in the northern region. These shelf waters consist of a shallow epi-continental sea in the Gulf of Carpentaria (GOC) that merges to the west with the Arafura Sea. Mixing of offshore waters within this semi-closed system by wind and tide driven currents may not follow a consistent and coherent pattern of circulation (Wolanski 1993). It is surprising then that results indicate the presence of genetic subdivision (Table 4.6) within this system. Populations sampled from the western half of this shelf (Adelaide River and Roper River) display near reciprocal differences in haplotype frequency from populations within the eastern half (Weipa and Karumba) (Table 4.6). These results suggest that there is limited maternal gene flow between the two coastal margins within the GOC. Wolanski (1993) described a coastal boundary layer within the GOC that allows extensive long-shore circulation of coastal waters but effectively limits the mixing of estuarine and offshore waters. This system may segregate adult *S. serrata* populations among the western and eastern halves of the GOC, however it would not preclude the mixing of propagules released offshore. Thus, substantial portions of the *S. serrata* larvae released into the GOC may be carried to areas close to their natal source. The cause of this larval movement may be due to an

interaction between life history and hydrology, as modelled for two species of prawn in the GOC (Condie *et al*. 1999).

The high levels of genetic exchange among shelf connected *S. serrata* populations separated by distances > 1000 km's contrasts strongly with that seen among trans-oceanic *S. serrata* populations separated by similar scales of distance. For example, there is no evidence of recent mitochondrial gene flow among any of the major Melanesian Island groups (New Caledonia, Fiji, Solomon Islands) despite connection by broad-scale oceanic currents between these groups (Chapter 3). It is feasible that larval duration is a limiting factor determining connection among trans-oceanic populations of *S. serrata*. Given that the species has a planktonic phase less than 28 days, propagules would have to travel upwards of 40 km's per day to scale the distances between the Melanesian islands. A similar pattern was reported by Doherty *et al*. (1995) for the anemone fish *Amphiprion melanopus* which has a free swimming plankton stage lasting ~ 18 days, whereby localised exchange among semi-connected habitats was more likely than across oceanic distances.

4.4.2 Historical connectivity and vicariance among Australian Scylla serrata populations.

This study identifies two distinct clades of mtDNA haplotypes that are for the most part mutually exclusive in locations either side of the Torres Strait. Observations of both sympatric and contiguous distributed sister species either side of the Torres Strait has prompted suggestions that the Strait serves as an intermittent biogeographical interface between the Indian and Pacific areas (Marsh and Marshall 1983, Davie 1985). Correspondence of this interface with a phylogeographic split among *S. serrata* populations may indicate that emergence of regionally independent *S. serrata* populations resulted from historical episodes of vicariance (Avise 1994). Indeed this argument has been invoked for a number of marine species that exhibit genetically heterogeneous populations either side of the Torres Strait (Keenan 1994, Elliot 1996, Begg *et al*. 1998, Chenoweth *et al*. 1998-b). Studies of the relationship between eustasy and the north Australian continental shelf indicate that this area experienced substantial habitat modification due to Pleistocene sea level changes (Chappell 1983, Kench 1999). Low sea levels lasting ~100,000 years resulted in land bridges between northern Australia and southern New Guinea. These bridges were breached by episodes of inundation lasting less than 20,000 years. In this context, episodes of submergence

equal to or greater than contemporary levels have been relatively rare. There have been many opportunities for genetic sub division between eastern and western marine regions interspersed by briefer periods of connection and potential gene flow (Keenan 1994). Under a vicariance scenario, it can be expected that in the absence of recent gene flow, the sundering of formerly widespread populations across the northern region would result in a reciprocal pattern of mtDNA monophyly either side of the disturbance (Avise 1994). Results here clearly indicate such a pattern however the estimated time to coalescence of the two *S. serrata* clades date back one million years before present (chapter 3). Under this time frame, cladogenesis occurred before several of the most recent periods of Pleistocene marine regression and transgression over the northern Australian region. Estimates of temporal divergence extrapolated from molecular data are subject to wide variances and may be confounded by accelerated rates of mutation at the lineages being examined (Avise 2000). Tests conducted in chapter 3 provided no evidence of accelerated mutation rate among lineages within the *Scylla* genus. Nevertheless, temporal estimates provided here can only be considered as an approximation of the true divergence times and therefore must be treated with some caution.

In contrast to my study, sampled haplotypes from the estuarine fish *Lates calcarifer* across northern Australia coalesce at approximately 350,000 years before present (Chenoweth *et al*. 1998-b). For this species, mtDNA cladogenesis corresponds with the start of the penultimate low sea level stand. Populations of *L. calcarifer* show strong evidence of secondary mixing of lineages across northern Australia, which would have occurred following prior periods of reconnection between the marine regions. These patterns are absent from the crab data indicating prior evidence of admixture among *S. serrata* clades was totally erased by the process of population retraction resulting from sea level retreats.

Alternatively, the zone of divergence resulting in the initial cladogenesis of *S. serrata* may have origins outside of the northern Australian region. The low levels of haplotype diversity in eastern Australia are consistent with that reported for a variety of IWP *S. serrata* populations (Chapter 3) where locations are dominated by the presence of a single most frequent haplotype. These widespread patterns of reduced diversity typify regions of recent range expansion (Hewitt 1996, Templeton 1998). I previously argued that eastern Australian *S. serrata* are derived from an expansion of *S. serrata* throughout

the IWP during the late Pleistocene (Chapter 3). Evidence from the neutrality test (Table 4.4) clearly demonstrates that eastern region locations have significant excesses of low frequency derived mutations among haplotypes. It may be argued that such excesses are signatures of population expansion following either a bottleneck or founding event (Excoffier and Schneider 1999) however the possibility that the excesses are due to a "selective sweep" among eastern region populations for a favoured haplotype cannot be ruled out (Maruyama and Birky 1991). Fossil evidence of *S. serrata* found among eastern Australian coastal Pleistocene deposits attest to the existence of this species in the region during the late Pleistocene (Hill *et al*. 1970). Similarly, fossil evidence at Japan dating back at least two hundred thousand years (Karasawa and Tanaka, 1994) and older fossils at South Africa (Chapter 3) indicate the presence of *S. serrata* populations at these locations that pre-date that estimated from the genetic data.
As argued for South African *S. serrata* populations in Chapter 3, the absence of a variety of older mtDNA lineages in contemporary eastern region populations may be an indication that earlier populations either went extinct or regressed to low numbers in response to the vicissitudes of the Pleistocene glacial periods.

In contrast, the general lack of deviation from neutrality and high diversity levels in northern and western Australian populations suggest a historical persistence of these populations relative to that seen at the eastern region. This is surprising given the recent age of northern coastal habitats (Kench 1999) and suggests that moderate levels of haplotype diversity existed within the founding clade 2 populations as they expanded into new coastal habitats of the northern Australian region following increased sea levels. The data here may indicate that clade 2 *S. serrata* populations have persisted in refugia (*sensu* Hewitt 1996) on shelf habitats of the northwestern Australian area during the lengthy episodes of lowered sea level. In contrast, the data for clade 1 crabs may confirm the contention of the previous chapter that clade 1 crabs colonised Australian waters from a west Pacific source following the most recent glacial retreat.

CHAPTER 5: **Microsatellite analysis of Australian *Scylla serrata* populations.**

5 .1 Introduction

Empirical surveys of population genetic structure using multi-locus genetic markers provide independent and comparable data for estimating the effects of gene flow and genetic drift throughout a sample distribution. This approach assumes greater power for hypothesis testing when the genetic markers of choice have intrinsically different rates of mutation and or mode of inheritance, such as that which exist between diploid nuclear and haploid cytoplasmic systems.

Here I present a genetic analysis of Australian *S. serrata* populations using microsatellite markers. Results from this analysis are directly comparable to mtDNA data in Chapter 4. I wish to know if the population structure detected using mtDNA is evident in nuclear encoded genes. Specifically I expected that the regional genetic partitioning of Australian mud crab populations evident in mtDNA analysis would also be present in microsatellite analysis. Given that long periods of allopatry among northern and eastern *S. serrata* populations have facilitated emergence of separate mtDNA clades either side of the Torres Strait, I expected the frequency and distribution of alleles at microsatellite loci to be especially divergent either side of this geographic junction. Furthermore, comparisons of regional diversity levels among the various classes of genetic marker can be used to test for concordant signatures of change in effective population size (*Ne*) among regions. Detected levels of regional mtDNA haplotype diversity suggest that eastern region populations have at some time, experienced reductions in population size not experienced at other Australian populations. These reductions may be in response to recent population bottlenecks or alternatively may reflect the historical process of colonisation and range expansion by mud crabs of the eastern region. Concordant multi-locus signatures of reduced allelic diversity among eastern region populations relative to that seen in the north and west would provide evidence that the reductions in *Ne* have occurred recently in response to bottlenecks. Finally, I wanted to examine the taxonomic relationship between the two clades of mtDNA present among *S. serrata*. The two clades which generally have different geographic distributions and mtDNA population genetic structure may represent recently diverged sister species within *S. serrata*. If so, I would expect no

interbreeding (or nuclear genetic exchange) between individuals representative of the mtDNA clades. Concordant patterns of segregation among independent genetic markers that have varied rates of evolution would provide a rigorous means of delineating genetically independent stocks of *S. serrata*.

5 .2 Methods

5.2.1 Sampling

A subset ($N = 188$) of *S. serrata* DNA samples previously scored for mtDNA variation in chapter 4, were screened for microsatellite variation in this study. Since the aims of this study are to compare microsatellite variation relative to that previously demonstrated for mtDNA, $N = 180$ crabs were analysed from eight coastal Australian locations representative of three distinct phylogeographic regions (western, northern and eastern) and two clades (clade 1 & 2) based on the previous mtDNA analysis (Table 5.1). Using this sampling scheme, patterns of population structure and diversity estimated from microsatellites could be directly compared to results of the mtDNA analysis. Microsatellite profiles of an additional eight crabs from the northern and western regions previously shown to have clade 1 mtDNA haplotypes (Chapter 4) were examined using assignment procedures to test the hypothesis of segregated breeding between clade 1 and 2 crabs.

Table 5.1 Sample regions, locations and size (*N*). Also shown is the mtDNA clade to which crabs are designated as per Chapter 4.

Region	Clade	Location	*N*
Western	2		
		Exmouth	22
		Broome	8
Northern	2		
		Roper River	27
		Karumba	13
		Weipa	25
Eastern	1		
		Hinchinbrook	21
		Mackay	28
		Moreton bay	36

5.2.2 Microsatellite amplification

Five sets of primers previously described in Chapter 2.2.7 (Table 2.2) were used in PCR reactions for the amplification of microsatellite loci. Methods for running and scoring microsatellites are as described in section 2.2.7.

5.2.3 Diversity measures and tests of spatial genetic structure

Tests for genotypic linkage disequilibrium between pairs of loci, departures from HWE at single loci and estimates of heterozygosity within sample locations were conducted as per chapter 2.3.1. Genetic differentiation and significance testing among paired locations at single and multiple loci was estimated using *F* and *R* statistics as described in Chapter 2.3.2. I tested for the presence of regional genetic structure by grouping samples into regions (as described in chapter 4) and conducting permutation tests for allelic homogeneity at single loci, and using Fisher's (1954) combined Chi-squared test for multiple loci. In these tests, individuals with Clade 1 mtDNA haplotypes found in the northern and western regions were excluded from the analysis. Tests for allelic homogeneity were also conducted between samples grouped by mtDNA clade identity to see if arrays of nuclear alleles were segregated among the mtDNA clades.

AMOVA (Schneider *et al*. 2000) was also used to investigate how multilocus microsatellite variation was partitioned among various spatial scales of sampling, as previously used in Chapter 4 (section 4.2.2). Locations were grouped at three hierarchical levels (among regions, among localities within regions, and within locations) and examined for the percentage of overall genetic variation apportioned to each level in the AMOVA. AMOVA tests were conducted using genetic distances calculated as both F_{ST} (allelic frequency only) and R_{ST} estimates. Replicate tests of these AMOVA's excluded clade 1 crab samples present within western and northern locations, to examine for any potentially confounding influence of these individuals on the perceived genetic structure. MtDNA haplotypes (using the sub-set sample employed in this chapter) were also analysed in AMOVA, so that a direct comparison could be made with the microsatellite results.

5.2.4 Tests for evidence of reduced allelic diversity and genetic bottlenecks

Tests were conducted for the effects of recent bottlenecks on microsatellite loci (refer Chapter 2.3.5) within two arbitrary population groupings. The first test considered two

sets of crab grouped by association as either mtDNA clades 1 or 2. This allowed the examination of potential recent differences in demographic history between the clades. The second test considered crabs grouped by association to geographic region (western, northern and eastern) to see if recent bottleneck effects were present and or confined to any of the regions. In both tests, microsatellite scores for eight crabs that were associated with clade 1 mtDNA haplotypes and geographically sympatric with populations of clade 2 crabs were removed from the analysis.

5.2.5 Assignment probability tests

Multilocus assignment tests were conducted using procedures described in Chapter 2.3.4. All samples were initially examined for the probability of being assigned (whether correctly or incorrectly) to each of the nominated groups identified in the previous chapter (i.e.: grouped by region; grouped by mtDNA clade). This self-assignment procedure, therefore determined if individuals could be reliably "self assigned" to their groupings based solely on relatedness at multilocus genotypes. For the next stage to have any relevance, it is necessary that the majority of individuals within each group have a greater probability of self- assignment than assignment to alternative groups. Individuals with null scores at any of the five loci ($N = 9$) were removed from analysis.

The second component of the assignment procedure tested if mtDNA clade 1 crabs sampled from the northern and western regions ($N = 8$) had greater genetic relatedness (as measured by the multilocus microsatellites) to clade 1 crabs found in the eastern region, than to clade 2 crabs from their immediate geographic proximity. Each of the eight crabs were scored for their multilocus microsatellite profile and compared, using assignment procedures, to those derived from the aforementioned group samples. By this means, I was able to determine the probabilities of assignment for each of these samples to the groups. The eight samples were removed from the preceding analysis (self-assignment procedure), to avoid artificially inflating the probability of assignment to their immediate geographic proximity. Genetic segregation at microsatellite loci concordant with the mtDNA division would provide evidence that nuclear genetic exchange among sympatric clade 1 and 2 crabs is absent and bolster arguments for the existence of co-distributed non-interbreeding stocks of *S. serrata* within Australian waters.

5 .3 Results

5.3.1 Allelic diversity and heterozygosity

All microsatellite loci had greater than 28 observed alleles apart from Ss-513, which had only seven alleles (Figures 5.2 a-e). Excluding Ss-513, the average number of observed alleles per locus was 31. The observed and expected heterozygosity at each location was generally greater than 80 % for all loci apart from Ss-513 (Table 5.2). The levels of per location heterozygosity at locus Ss-513 were generally half as much as that observed for the other four loci (Table 5.2). The reduced heterozygosity observed for Ss-513 was due to an excess of one allele (allele # 2) relative to others at that locus (Figure 5.1e).

Table 5.2 Summary of variation at five microsatellite loci among *N* samples of *Scylla serrata* from eight Australian locations. The number of observed alleles (*K*), inbreeding co-efficient (*Fis*), average observed and expected heterozygosity (H_O and H_E), and probability of deviation from Hardy-Weinberg equilibrium (P_{HW}) as described.

Locus		Exmouth	Broome	Roper r.	Karumba	Weipa	H'brook	Mackay	Moreton
Ss-101									
	N	22	5	27	12	24	21	28	27
	K	19	8	21	14	18	21	19	26
	Fis	+0.001	-0.053	-0.051	+0.134	+0.027	+0.012	+0.013	+0.035
	H_O	0.955	1.000	1.000	0.833	0.917	0.952	0.929	0.926
	H_E	0.956	0.956	0.953	0.967	0.942	0.964	0.940	0.963
	P_{HW}	0.119	1.000	0.885	0.226	0.658	0.818	0.355	0.311
Ss-103									
	N	21	3	26	10	21	18	24	30
	K	16	4	17	12	14	14	16	20
	Fis	-0.022	+0.273	+0.062	+0.041	+0.026	-0.018	-0.079	+0.009
	H_O	0.952	0.667	0.885	0.900	0.905	0.944	1.000	0.903
	H_E	0.933	0.933	0.947	0.942	0.932	0.929	0.929	0.947
	P_{HW}	0.351	0.467	0.193	0.677	0.657	0.249	0.836	0.644
Ss-112									
	N	22	6	26	10	19	19	24	31
	K	20	8	21	13	19	15	16	23
	Fis	-0.020	+0.107	+0.074	+0.172	+0.012	+107	-0.067	+0.007
	H_O	0.955	0.833	0.885	0.800	0.947	0.842	1.000	0.936
	H_E	0.938	0.924	0.959	0.968	0.960	0.945	0.939	0.945
	P_{HW}	0.618	0.492	**0.037**	**0.036**	0.728	0.479	0.900	0.829
Ss-403									
	N	22	6	27	13	25	21	26	35
	K	17	7	22	13	21	17	17	19
	Fis	+0.043	+0.074	+0.066	+0.070	-0.014	-0.014	-0.032	-0.047
	H_O	0.864	0.833	0.889	0.846	0.960	0.952	0.962	1.000
	H_E	0.903	0.924	0.955	0.908	0.947	0.940	0.934	0.929
	P_{HW}	0.564	0.676	0.112	0.264	0.583	0.083	**0.040**	0.367
Ss-513									
	N	21	8	27	13	25	21	28	36
	K	07	3	04	03	05	04	05	05
	Fis	-0.068	-0.105	-0.039	+0.178	-0.238	-0.240	+0.137	-0.039
	H_O	0.714	0.375	0.667	0.385	0.560	0.571	0.464	0.611
	H_E	0.695	0.442	0.642	0.517	0.455	0.463	0.542	0.605
	P_{HW}	0.532	1.000	0.941	0.420	0.888	0.733	0.258	0.956

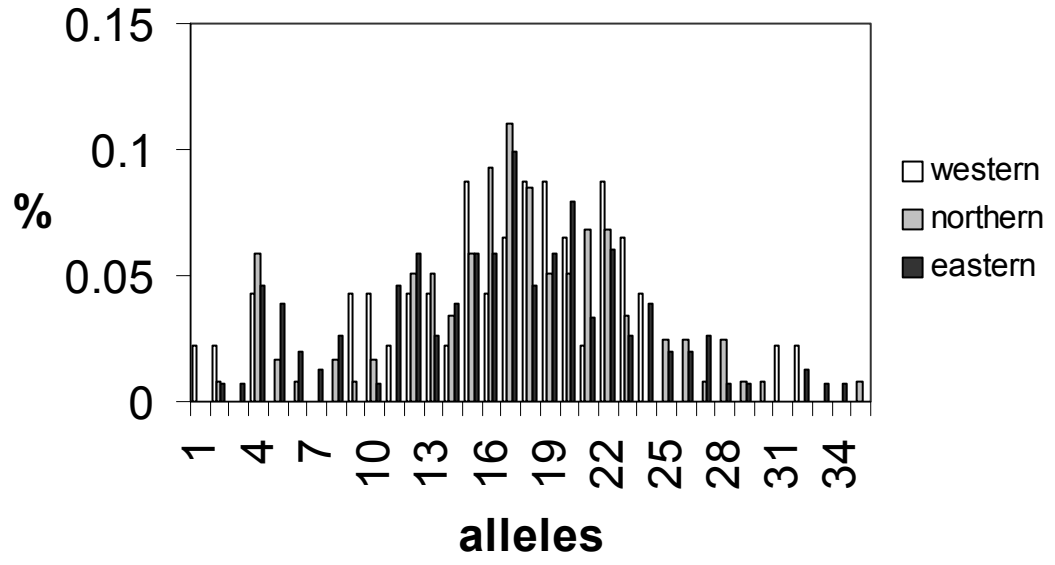


Figure 5.1a Allele frequencies at three regions for locus Ss-101

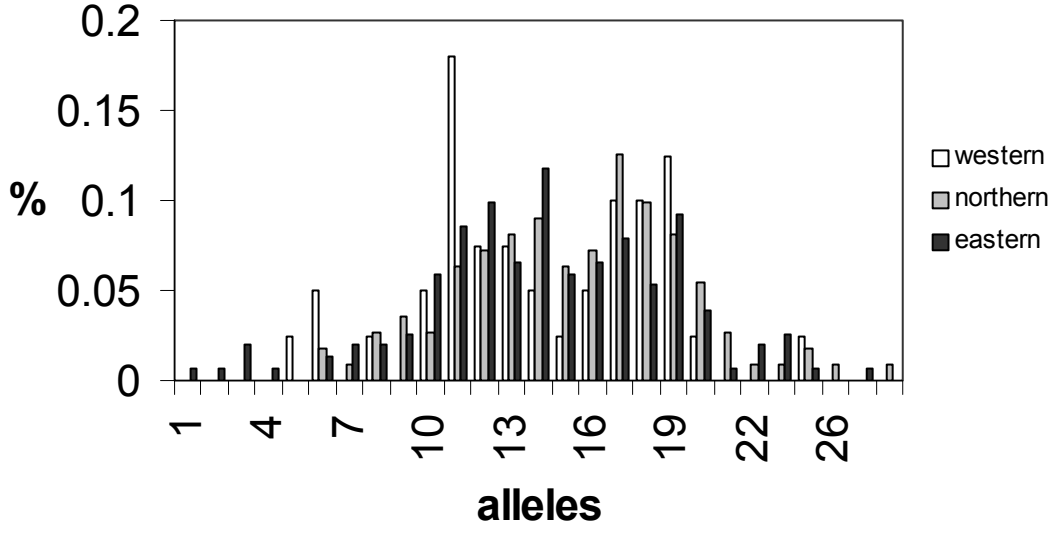


Figure 5.1b Allele frequencies at three regions for locus Ss-103

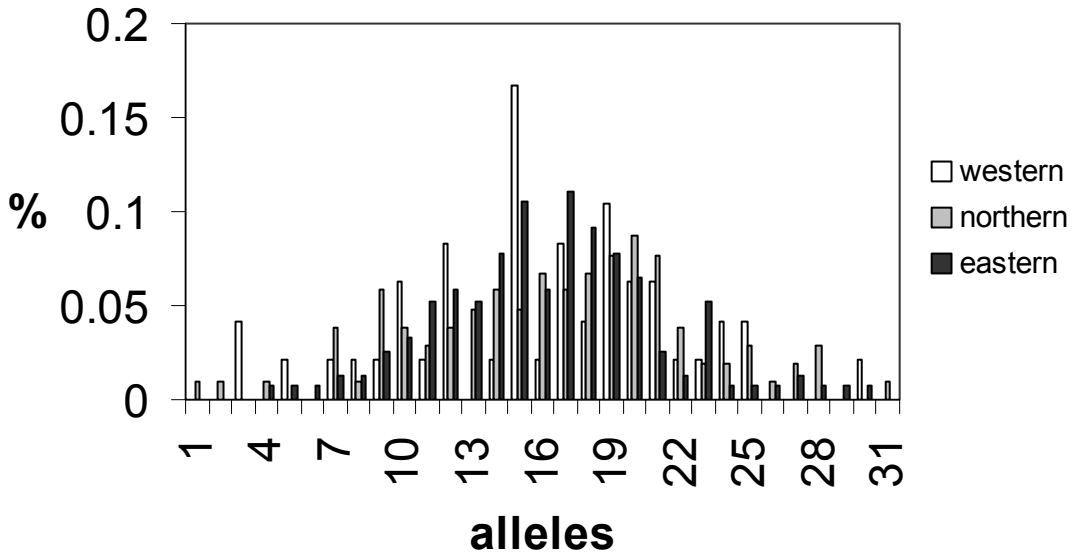


Figure 5.1c Allele frequencies at three regions for locus Ss-112

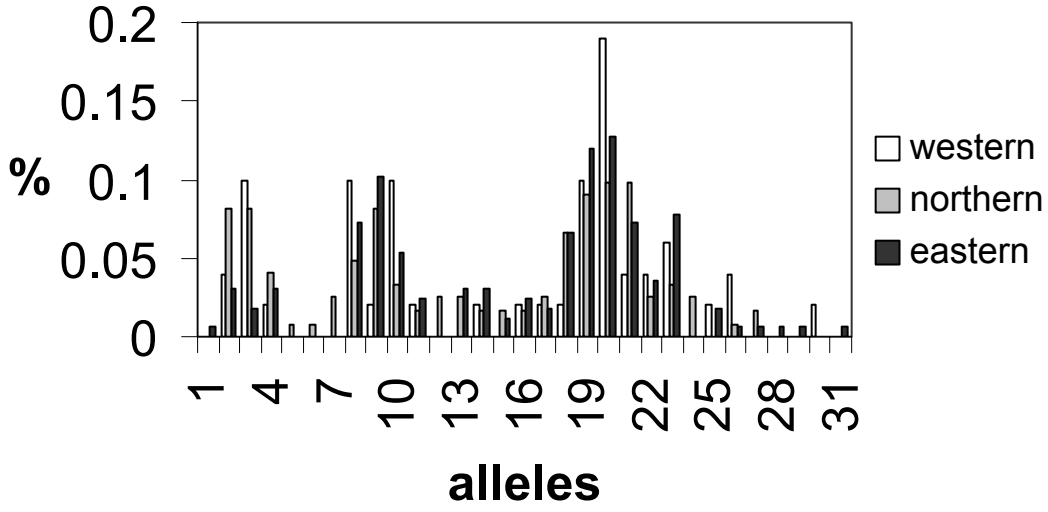


Figure 5.1d Allele frequencies at three regions for locus Ss-403

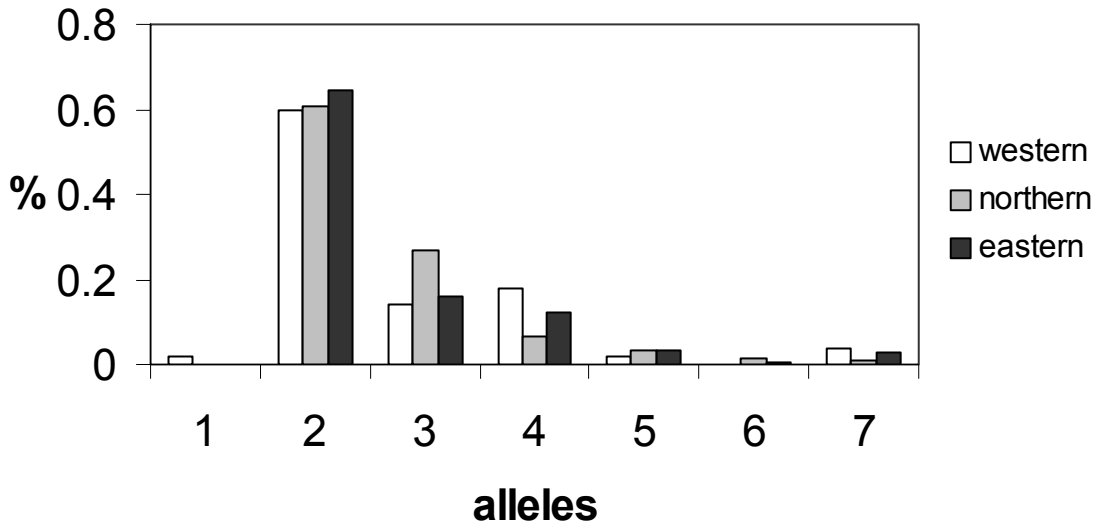


Figure 5.1e Allele frequencies at three regions for locus Ss-513

5.3.2 Tests for linkage disequilibrium and deviation from random mating

There was no evidence within locations, regions or clades of any significant ($\alpha = 0.05$) association of alleles among pairs of loci after tests for genotypic linkage disequilibrium. Three of the 40 single locus tests for deviations from HWE within locations were significant at $P < 0.05$ (Table 5.2). Significant deviation from HWE was due to marginal excesses of homozygotes (positive *Fis* values) at locus Ss-112, and an excess of heterozygotes (negative *Fis* value) in one instance at locus Ss-403. None of the tests remained significant following Bonferroni multiple test adjustment.

5.3.3 Fixation indices and tests for allelic homogeneity

Over 97 percent of all paired location F_{ST} and R_{ST} estimates at individual loci were less than 0.05, generally indicating low levels of genetic differentiation among sample locations (Appendix). Exact tests of homogeneity in allele frequencies for individual loci were significantly rejected ($P < 0.05$) at 18 of the 140 pairwise tests. In some

instances comparisons between locations were significant at a single locus only, however several paired comparisons that included either Roper River or Exmouth were congruently significant at two of the loci (Ss-403 and Ss-513).

Multi locus fixation indices among paired locations were also minimal, the majority of estimates either negative or below 0.01 for both F_{ST} and R_{ST} (Table 5.3). Levels of genetic differentiation between locations did not appear to correlate with geographic proximity; estimates of F_{ST} and R_{ST} were minimal between the two most distant sites (Moreton Bay and Exmouth). A few pairwise tests involving either Exmouth, Roper River or Karumba were significant at the 5 % level, but none of the tests remained significant following Bonferroni corrections. The magnitude of difference between F_{ST} and R_{ST} estimates for each of the individual pairwise location tests was minimal, however estimates between locations shown to be significant (before Bonferonni correction) were much larger for R_{ST} than for F_{ST} and may indicate that where present, genetic differentiation was better detected using estimates of R_{ST} than F_{ST}.

Table 5.3 Multilocus F_{ST} (lower matrix) and R_{ST} (upper matrix) estimates between all pairs of locations. Probabilities obtained by combined permutation tests (10,000 replicates). Significant estimates at $P < 0.05$ indicated in bold (before Bonferroni adjustments). Location abbreviations: EX, Exmouth; BM, Broome; RR, Roper River; KA, Karumba; WE, Weipa; HB, Hinchinbrook; MY, Mackay; MB, Moreton Bay.

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		-0.027	**0.018**	-0.030	0.006	-0.016	-0.001	-0.019
BM	0.025		-0.021	-0.076	-0.055	-0.002	-0.043	-0.042
RR	0.006	0.024		-0.020	**0.048**	**0.066**	**0.048**	**0.028**
KA	**0.014**	**0.013**	0.006		-0.005	-0.009	0.003	-0.017
WE	0.012	0.001	**0.013**	0.007		0.013	-0.007	0.002
HB	0.007	0.009	0.010	0.008	-0.002		0.013	-0.003
MY	0.008	0.007	**0.011**	0.007	-0.001	0.001		-0.011
MB	0.002	-0.003	0.008	0.008	0.001	0.003	0.002	

Apart from some minor differences, the frequency and distribution of alleles was similar among the three regions at each of the five loci (Figures 5.2a-e). Exact tests for differences in allelic distribution among sample locations grouped either by regions or clades were non significant ($P > 0.05$) in all cases except in several instances among regions at locus Ss-403 and Ss-513 (Table 5.4). The exact test at locus Ss-513 indicated the northern region significantly differed from the western ($P = 0.036$) region and was marginally different from the east ($P = 0.094$). The significant value among grouped regions seen at locus Ss-403 ($P = 0.042$) is due to marginal (but not significant) differences between northern from eastern ($P = 0.090$) and western regions ($P = 0.060$).

Table 5.4 Exact test probabilities (*P*) for allelic homogeneity among locations grouped by either clades or regions (Clade 1 crabs present within western or northern regions removed from analysis) for each of the five loci. Combined *P* values across all loci determined by Chi-squared test (Fisher's method).

		Regions *P* value for samples grouped by:			
Locus	Clades	W $_{VS}$ N	W $_{VS}$ E	N $_{VS}$ E	All
Ss-101	0.380	0.499	0.606	0.433	0.543
Ss-103	0.679	0.786	0.742	0.851	0.878
Ss-112	0.542	0.447	0.257	0.366	0.331
Ss-403	0.221	0.060	0.142	0.090	**0.042**
Ss-513	0.540	**0.036**	0.561	0.094	0.085
Combined	0.611	-	-	-	0.149

Analysis of molecular variance across various scales of pooled samples indicates that microsatellite allelic variation within populations was far greater (>98%) than that either among regions or among locations within regions (Table 5.5). Fixation statistics within AMOVA for each hierarchical sampling level were small (< 0.015) and non significant, regardless of whether fixation statistics were calculated by allelic frequency (F_{ST}) or also incorporated a model of stepwise mutation (R_{ST}). Exclusion of clade 1 crab samples from the northern and western regions resulted in very slight changes to percentage of

variation apportioned to sampling levels (Table 5.5) but did result in a significant among region effect for the R_{ST} AMOVA. In comparison to that seen for the microsatellite loci, AMOVA results of mtDNA data (using the same samples) indicate substantial and significant regional structure. Over 87 % of the variation in this dataset is due to regional partitioning of mtDNA haplotypes.

Table 5.5 Analysis of molecular variance within and among sample locations from three regions of Australia. Calculation of fixation indices (Φ) and partitioning of total genetic variance (Variance %) as per Excoffier *et al*. (1992). Probability (*P*) of getting a fixation statistic greater than the observed value by chance calculated from 15 000 random permutations.

A: MtDNA F_{ST} (incorporating nucleotide difference)

Variance Component	Variance %	Φ statistic	*P*
Among regions	87.17	$\Phi_{CT} = 0.872$	**< 0.001**
Among locations within regions	1.09	$\Phi_{SC} = 0.085$	**< 0.007**
Within locations	11.74	$\Phi_{ST} = 0.883$	**< 0.001**

B: Microsatellites F_{ST}

Variance Component	Variance %	Φ statistic	*P*
Among regions	0.45	$\Phi_{CT} = 0.005$	< 0.089
Among locations within regions	0	$\Phi_{SC} = 0.000$	< 0.980
Within locations	99.55	$\Phi_{ST} = -0.002$	< 0.835

C: Microsatellites R_{ST}

Variance Component	Variance %	Φ statistic	*P*
Among regions	1.24	$\Phi_{CT} = 0.012$	< 0.131
Among locations within regions	0	$\Phi_{SC} = 0.000$	< 0.890
Within locations	98.76	$\Phi_{ST} = 0.003$	< 0.558

D: Microsatellites F_{ST} (Clade 1 haplotypes removed from north and west)

Variance Component	Variance %	Φ statistic	*P*
Among regions	0.3	$\Phi_{CT} = 0.003$	< 0.278
Among locations within regions	0	$\Phi_{SC} = 0.000$	< 0.955
Within locations	99.7	$\Phi_{ST} = -0.002$	< 0.847

E: Microsatellites R_{ST} (Clade 1 haplotypes removed from north and west)

Variance Component	Variance %	Φ statistic	*P*
Among regions	1.42	$\Phi_{CT} = 0.014$	**< 0.049**
Among locations within regions	0	$\Phi_{SC} = -0.003$	< 0.647
Within locations	98.58	$\Phi_{ST} = 0.011$	< 0.197

5.3.4 Bottleneck test

Although most loci had marginally larger values of H_E compared to H_{eq} when samples were grouped by clade, there was no evidence that the differences in the magnitude were significant at any of the five loci (Table 5.6). This was reflected by a non-significant score (P= 0.156 for clade 1 and P = 0.625 for clade 2) averaged across loci using the Wilcoxon sign – rank test. Similar patterns and results were found when crabs were grouped by geographic region (Wilcoxon sign-rank test: P = 0.625 for western; P = 0.625 for northern; P = 0.156 for eastern region). Overall, values of H_E were not shown to be significantly greater than H_{eq} and therefore there was no evidence for the effects of recent bottlenecks upon samples grouped by either mtDNA clade or geographic region.

Table 5.6 Tests for differences between heterozygosity (*HE*) and expected equilibrium heterozygosity (*Heq*) for each of 5 microsatellite loci within (a) mtDNA clades and (b) geographic regions. Probability (*P*) estimation of magnitude difference between *HE* and *Heq* obtained from coalescent simulation (1000 iterations) of the observed number of alleles (*K*). * $P < 0.05$ (two tailed test).

(a) Group by mtDNA Clades

	Locus	*2N*	*K*	*HE*	*Heq*	P
Clade 1						
	Ss-101	152	32	0.957	0.947	0.150
	Ss-103	152	24	0.937	0.922	0.188
	Ss-112	153	27	0.939	0.934	0.439
	Ss-403	166	25	0.933	0.924	0.356
	Ss-513	170	06	0.542	0.626	0.200
Clade 2						
	Ss-101	164	21	0.950	0.939	0.345
	Ss-103	151	22	0.932	0.914	0.155
	Ss-112	152	29	0.953	0.940	0.117
	Ss-403	172	28	0.942	0.933	0.310
	Ss-513	172	07	0.573	0.681	0.133

(b) Group by geographic region

	Locus	*2N*	*K*	*HE*	*Heq*	P
Eastern						
	Ss-101	152	32	0.957	0.947	0.150
	Ss-103	152	24	0.937	0.922	0.188
	Ss-112	153	27	0.939	0.934	0.439
	Ss-403	166	25	0.933	0.924	0.356
	Ss-513	170	06	0.542	0.626	0.200
Northern						
	Ss-101	118	26	0.948	0.940	0.275
	Ss-103	111	21	0.936	0.923	0.224
	Ss-112	104	26	0.956	0.943	0.064
	Ss-403	122	26	0.947	0.940	0.303
	Ss-513	122	06	0.558	0.682	0.104
Western						
	Ss-101	46	21	0.961	0.948	0.082
	Ss-103	40	15	0.923	0.920	0.515
	Ss-112	48	21	0.944	0.947	0.321
	Ss-403	50	18	0.909	0.931	0.073
	Ss-513	50	06	0.598	0.700	0.073

5.3.5 Assignment test

Assignment tests had limited power for correctly assigning individuals to their sample origin (Table 5.7). Less than 10 % of samples could be correctly assigned to their group of origin (regions and clades), even with assignment probabilities set to a minimum of 80 % certainty. Similar proportions were incorrectly assigned (~ 10 %) and the vast majority (> 80 %) was ambiguously assigned. Ambiguously assigned individuals generally had probabilities of assignment well below the nominated 80 % minimum certainty level, and many of those individuals had approximately equal probability of assignment to the nominated groups. This is evident in Figure 5.2 where it can be seen that assignment probabilities for the majority of samples are equally apportioned between the two clades.

Table 5.7 Results of self-assignment tests. The exclusion-simulation method of assignment set to a minimum *P*-value of 0.80 for the probability of origin. Percentage of individual samples ($N = 171$) correctly assigned to their group of origin as indicated. Also shown are the percentages of individuals incorrectly and ambiguously assigned (assignment < 80 % or equal probability among groups).

	Percentage assigned		
Group	Correct	incorrect	ambiguous
Clades	0.08	0.07	0.85
Regions	0.07	0.10	0.83

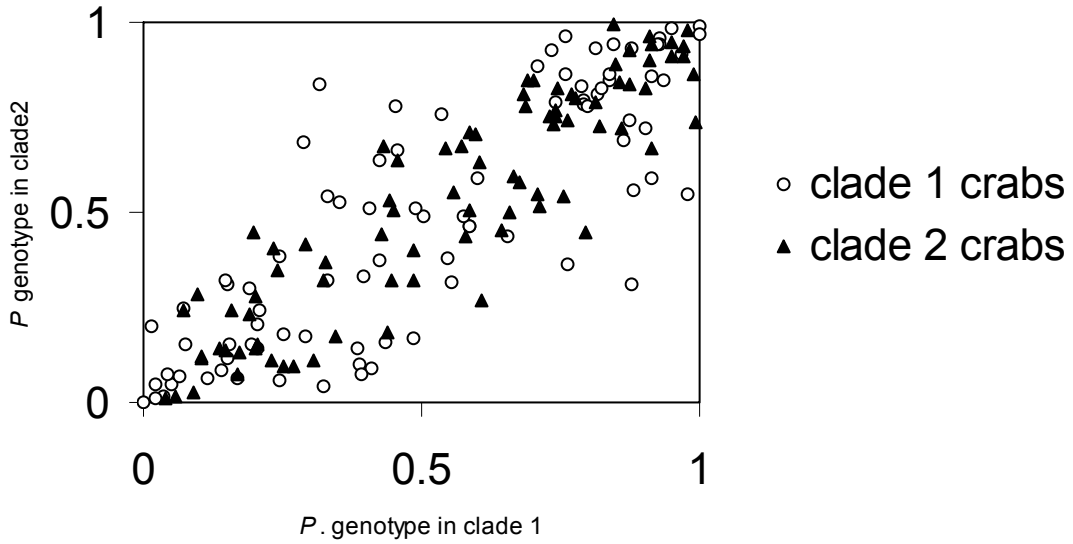


Figure 5.2 Likelihood probabilities of assignment to either clade 1 or clade 2 groups. Based on genotype profiles of 5 microsatellite loci, for clade 1 crabs (*N*= 84) and clade 2 crabs (*N*= 87). Likelihood of assignment to clades estimated from 10,000 simulations using the frequency method (Paetkau *et al*. 1995), and "Add one in Procedure" for rare alleles, as implemented in GeneClass (Cornuet *et al*. 1999).

Despite the low levels of confidence demonstrated in the self-assignment procedures, multilocus profiles for eight clade 1 samples from the western and northern regions were compared to the various groups and the probabilities of their assignment to these groups computed (Table 5.8). There is no consistent pattern of assignment of these individuals to either the regions or clades. Several of the samples from the northern region also have low probability of assignment (<10 %) to any of the nominated groups, indicating that they have a unique multilocus genotype and that one or more alleles in their profile are either rare or not seen among other individuals.

Table 5.8 Assignment probabilities (*P*) of eight clade 1 samples found in the western (samples: W 01-04) and northern (samples: N 01-04) regions among three regions and two clades.

	Regional *P*			Clade *P*	
Sample	Western	Northern	Eastern	Clade1	Clade 2
W-01	0.066	0.100	0.139	0.141	0.200
W-02	0.079	0.244	0.302	0.310	0.362
W-03	0.136	0.279	0.184	0.191	0.442
W-04	0.470	0.374	0.801	0.806	0.559
N-01	0.001	0.019	0.063	0.066	0.022
N-02	0.001	0.017	0.040	0.041	0.017
N-03	0.001	0.009	0.014	0.014	0.011
N-04	0.405	0.952	0.794	0.792	0.929

5 .4 Discussion

5.4.1 Microsatellite versus mtDNA structure

The geographic distribution of microsatellite alleles among Australian *S. serrata* populations is generally homogenous. For the majority of comparisons, there was far more allelic variation within than between the sampled locations. Furthermore observed microsatellite heterozygosity estimates for each of the sampled locations were high with little evidence of differences in magnitude of allelic diversity. These results contrast to those found for mtDNA analysis of the same locations, which indicated that closely related mtDNA lineages are sorted among coastal regions, with reduced heterozygosity among eastern relative to northern region populations.

The differences in population structure between the two classes of genetic marker are clearly identified in the way that AMOVA has apportioned genetic variation among the spatial sampling scales (Table 5.5). The vast majority of variation within all of the microsatellite matrices is present within locations (98 %) whereas for the mtDNA matrix, greater than 87 percent of the variation is apportioned to differences among regions. The large amount of variation among regions in the mtDNA data set is derived from the instances of both shallow and deep phylogeographic assortment of haplotypes to regions. The shallow assortment of mtDNA lineages between the western and northern regions reflects recent lineage sorting (~ 8000 years bp.) whereas the separation between the north and east is derived from a much older division (~ 1 Myr's bp.) within the species (Chapters 3 and 4). It is interesting to note that the distribution of microsatellite alleles among regions is relatively homogenous despite these instances of shallow and deep phylogeographic separation present in the mtDNA data. It is possible that both contemporary and historical factors that have promoted regional assortment of mtDNA lineages have either not affected or are not apparent in the microsatellite data.

Among the microsatellite matrices however, there is a slight but significant among region effect within the R_{ST} matrix when clade 1 haplotypes present in the north and west are removed from analysis. It is interesting to note that the among region component of the corresponding F_{ST} matrix is not significant. This suggests that R_{ST} may be slightly more sensitive than F_{ST} as an estimator of structure at these loci. If so, this also suggests that the loci are evolving by a stepwise mutation model. Because R_{ST} is based on variance of allele size rather than frequency (as for F_{ST}), it is expected to be

more suited for detecting historical rather than contemporary divergence among populations (see review Balloux and Lugon-Moulin 2002). It is tempting then to suggest that the significant R_{ST} seen among regions is a "whisper" of the deep divergence seen in the mtDNA dataset. There is some very limited evidence from exact tests of homogeneity among regions that the distribution of alleles in the northern region is different from the others (Table 5.4). This is not a consistent pattern across all loci and may indicate stochastic biases involving allele frequencies at two of the loci (Ss-403 and Ss 513) in the Northern region population, Roper River (see later).

There is partial agreement among the mtDNA and microsatellite data sets, as to levels of genetic structure among locations within regions. Both classes of marker indicate an absence of genetic structure among the three locations sampled within the eastern region and two locations in the western region. Likewise, the presence of population structure in the northern region as detected by mtDNA may also be apparent in the microsatellite data. Pair-wise tests between locations indicate significant fixation indices across microsatellite loci between Roper River and Weipa (Table 5.3), a result shared with mtDNA (F_{ST} = 0.190, $P < 0.0005$; Table 4.6). This interpretation must be treated with caution, as these differences are being driven by significant variation at only two of the five loci tested; as previously mentioned Roper River is also significantly different from a number of locations outside of the northern region at these two loci and differences may therefore reflect biases in the dataset.

Allelic diversity levels within the three Australian regions are equally high at all microsatellite loci. There is no evidence of any recent reductions in microsatellite diversity within the eastern region relative to the others. Therefore, it seems unlikely that reduced levels of mtDNA diversity in the eastern region populations are due to the effects of a recent population bottleneck. The reduced levels are more likely historical signatures of either the expansion of *S. serrata* into this region following its colonisation, or a much older bottleneck event that is evident in the mtDNA but not the nuclear data.

The lack of evidence for partitioning of geographic structure observed at microsatellites is also prevalent when individuals are grouped into the two mtDNA clades of *S. serrata* detected in Australian waters. There is no evidence that microsatellite alleles are concordantly segregated with the mtDNA clade structure seen in chapter 4. Nor is there

any evidence that clade 1 crabs found in the northern and western regions (which contain predominantly clade 2 crabs) are more similar to crabs within the eastern region than to those in their immediate geographic proximity. Therefore, based on the microsatellite data, there is no evidence of segregated breeding between sympatric clade 1 and 2 crabs.

5.4.2 Reasons for discordance among genetic markers

Empirical genetic surveys of marine species using microsatellites often confirm previously detected structure as determined by mtDNA analysis (Appleyard *et al*. 2001, Brooker *et al*. 2000, Gold and Turner 2002). In some instances, microsatellites have reportedly provided far greater resolution of present day structure compared to historical signatures determined by mtDNA analysis (Buonaccorsi *et al*. 2001, Shaw *et al*. 1999, Wirth and Bernatchez 2001), so providing a powerful means of delineating contemporary stock structure within a species distribution. There are few instances where profound geographic structure observed among populations using allozyme or mtDNA analysis is not also observed to some extent using microsatellite analysis (but see De Innocentiis *et al*. 2001). Under what circumstances can such contrasting patterns of geographic structure among the classes of genetic marker occur?

Estimates of population differentiation based on fixation indices may contrast among different markers due to inherent levels of diversity available at loci. At microsatellites, it is common for reported heterozygosity levels within populations to approach very high levels (> 0.8) in contrast to the moderate to low levels of mtDNA diversity. This is in part due to the greater susceptibility of mtDNA for loss of diversity by genetic drift. Hedrick (1999) demonstrated that maximum values of F_{ST} averaged over loci (as is often the case at microsatellite loci) are reduced when levels of within population homozygosity diminishes. Therefore, estimates of F_{ST} averaged over loci can be low among populations despite subdivision, because of high diversity of alleles present within the populations. This argument however does not hold for the current data since estimates of F_{ST} at individual microsatellite loci are equally as low as the averaged estimates. Furthermore, unlike the prior argument, rich allelic diversity enhances the ability of certain statistical tests that are based on the frequency and distribution of alleles among populations for detecting population structure. The majority of exact tests of allelic homogeneity used in this study, which are independent of the fixation indices, also failed to demonstrate differences among regions at the microsatellite loci.

Discordance of the microsatellites from the mtDNA does not appear to be influenced by the generally greater diversity of alleles present at the microsatellite loci.

Another possible explanation for the discordance between genetic markers may be due to differential levels of gene flow between sexes within a species, relative to patterns of inheritance of the genetic markers used to detect genetic structure. Under circumstances where females are philopatric and gene flow is primarily mediated by male dispersal, maternally inherited haploid genes can be segregated among populations whilst bi-parental inherited diploid genes are panmictic. Instances of male mediated gene flow and female philopatry although uncommon for marine species have been hypothesised from genetic evidence for green turtles (Karl *et al*. 1992) and humpback whales (Baker *et al*. 1998). In both of these cases, it is thought that male mediated gene flow results not from a greater propensity for physical dispersal by males but from sporadic matings which occur during periods of overlapping seasonal migration between otherwise discrete populations.

If male-biased gene flow is prevalent among *S. serrata* populations, then transfer of maternal genes is by some means restricted – either due to direct differences in dispersal ability between the sexes or indirectly by life history/ behavioural influences. Dispersal among *S. serrata* populations may occur at all stages of the species life cycle, therefore a physical basis for differences in dispersal ability between the sexes, although not documented, may occur at any of the stages. Mark-recapture studies indicate that post-larval *S. serrata* movement is generally limited and confined to inter-tidal habitats. There is however no indication of greater incidence of male dispersal relative to females among these inshore habitats (Helmke 2002, Hyland *et al*. 1984). On the contrary, there is evidence indicating that mature female *S. serrata* migrating offshore to release eggs, occasionally return to estuaries far removed from their point of origin (Hyland *et al*. 1984, Hill 1994). It could be argued that this migratory behaviour may result in a greater bias towards female mediated gene flow via the long distance transfer of sexually mature females among estuaries. Apart from incrementing the standing population, it is unlikely however that these vagrant females contribute significantly to the gene pool within a new location, as there is some evidence that the fecundity of female *S. serrata* is greatly reduced following the initial spawn (Ong 1966, also see Heasman *et al*. 1985).

Effective dispersal and genetic transfer among *S. serrata* populations is more likely to occur during the first four weeks of development when large numbers of crab zoea are released offshore away from adult habitat and are entrained into coastal current systems as plankton (Hill 1994). Given that entrainment, dispersal and recruitment of plankton is generally considered random and mostly subject to physical rather than biological processes, it seems unlikely that there would be differences in either dispersal or colonisation ability among the sexes at this stage.

The effects of kin-structured colonisation (Levin 1988) may also result in contrasting patterns of population structure among diploid and haploid genetic markers (Wade *et al*. 1994). This is especially pronounced when colonisation of patches occurs from the breeding efforts of few (or related) females that have been multiply inseminated from different males. Reproduction by successful multiple inseminations effectively skews the ratio of sexually transmitted genes in a population towards a greater male component. The effect is exaggerated for species that produce large numbers of progeny per reproductive effort. One outcome of this mode of reproduction is to reduce the *Ne* of the maternal relative to paternal genes contributed to ongoing generations. This would manifest as less genetic diversity within populations and greater genetic structure among populations for maternally inherited genes compared to bi-parentally inherited nuclear genes (Wade *et al*. 1994).

There is some evidence that female *S. serrata* can store packets of sperm in their oviducts for several months following copulation (Du Plessis 1971), therefore this form of kin-structured breeding may be possible for the species. Empirical evidence to address this hypothesis for the species is absent from the literature, however tentative results not supporting the above hypothesis came from my examination of a single gravid female *S. serrata* captured offshore in the Gulf of Carpentaria. I tested this sample for evidence of multiple mating by comparing microsatellites amplified from among multiple biopsy samples ($N = 5$) of the egg mass relative to the mother. Evidence of more than four different alleles in the egg mass at any given locus among the multiple samples would indicate evidence that successful mating with multiple partners had occurred. The microsatellite analysis did not provide evidence of multiple male genotypes among the replicate samples. A greater number of independent gravid female samples are required to test this hypothesis fully.

It has been postulated that under certain conditions, signatures of historical and contemporary population genetic structure may be obscured for microsatellites due to allelic size homoplasy combined with constraints on maximum allele size (Garza *et al*., 1995; Nauta and Weissing 1996, Estoup *et al*. 2002). Constraints on maximum allele size in rapidly mutating microsatellites can lead to a saturation of back mutations (homoplasy) that homogenise allele size distributions among populations. As a result, alleles sampled from discrete populations may be identical in state but not by descent and estimates of genetic differentiation between the populations may be artificially low, even in the absence of gene flow between populations (Orti *et al*., 1997).

Allelic size homoplasy at microsatellites has been detected at various levels of taxonomic survey (i.e. among species, among and within populations) and is therefore not a trivial problem for population genetic surveys (see review Estoup *et al*. 2002). Instances of homoplasy reported for aquatic species are rare, possibly due to the low number of surveys conducted on this group and perhaps the most useful demonstration of the effects of homoplasy on genetic structure was reported for a terrestrial species. Queney *et al*. (2001) observed severe size homoplasy at microsatellites between two well-supported phylogeographic groups of rabbit populations on the Iberian Peninsula. Results presented here for *S. serrata* are similar to that of Queney *et al*., in that there is almost complete overlap in allele distribution across a set of microsatellite loci, despite the instances of deep division between two mtDNA clades (in both instances mtDNA division is dated to 1 MYR's b.p.). Despite the overlap in allelic distribution, Queney *et al*. were able to show significant F_{ST} values between the groups based on an excess of low frequency private alleles particular to one region. Thus despite evidence of homoplasy, Queney *et al*. were still able to demonstrate some level of regional population structure using microsatellites. In contrast to Queney *et al*., I was not able to detect the instances of either historical or recent separation of regional *S. serrata* populations using microsatellite analysis.

Assuming homoplasy at microsatellites is present among the *S. serrata* populations, the question arises as to why it is so pronounced as to make most forms of genetic analysis using these markers on this species effectively redundant. Simulation studies reported by Estoup *et al.* (2002) demonstrate that the probability of occurrence of allelic size homoplasy between two historically separated populations increases due to a number of factors, including the degree of allelic overlap between the populations prior to

separation, the effective population sizes and most importantly mutation rate and mode at the loci. The simulations conducted by Estoup *et al*. predict that homoplasy will be frequent at those loci evolving at moderate to high rates of step-wise mutations (with constraints on maximum allelic size). The effect is particularly enhanced if *Ne* within the separate populations is large (Estoup *et al*. 2002). Although homoplasy can maintain shared allelic size range among isolated populations (or species), maintenance of shared allelic frequency distribution among isolated populations requires that *Ne* is of such magnitude that effects of genetic drift within these populations is negligible. In this sense, isolated populations yet to come to equilibrium between drift and migration, (as a result of high *Ne*), may take an excessive time to show differences in allelic frequency distribution at these loci. This pattern is reminiscent of the conflicting patterns of non-equilibrium genetic homogeneity at allozymes relative to mtDNA structure observed among populations of some marine invertebrates (Lavery *et* al. 1996-a, Williams and Benzie 1997). Clearly, the evidence of complete and incomplete mtDNA lineage sorting among Australian *S. serrata* locations (Chapter 4) is an indication that the mitochondrial genome is approaching a genetic equilibrium in this distribution. Although equilibrium is expected to occur much faster at mtDNA compared to nuclear genes (Birky *et al*. 1989), the rate of approach at the microsatellite loci in *S. serrata* populations may be retarded by a combination of high *Ne*, polymorphism and homoplasy.

There are certain features of the *Scylla* genus that may agree with the predictions of Estoup *et al*. (2002) and therefore suggest that homoplasy coupled with high *Ne* has homogenised microsatellite frequency distribution within the crab populations. First, there is some indication that allelic overlap at microsatellite loci is also present among other species of *Scylla*. When I initially screened loci for polymorphism among samples, I also included a small number of *S. olivacea* for analysis ($N = 10$) and found partial overlap in allelic distribution at all loci with *S. serrata*. This at least indicates the potential for allelic overlap at the loci exists beyond the species level and suggests that considerable allelic diversity at these loci may have been present within the species before the bifurcation into clade 1 and 2 *S. serrata*. Although I do not have any direct evidence, it may be tentatively inferred that both *Ne* and the rates of mutation at microsatellite loci are quite high within the populations of *S. serrata*. The levels of heterozygosity and allelic diversity present at the loci are equally high at all spatial scales of sampling (within locations, within regions) indicating the effects of genetic drift at locations are not observable. Therefore, either *Ne* or the rate of mutation (or

both), are of such high levels that shifts in frequency and or loss of alleles via genetic drift are minimal at these loci.

5.4.3 Conclusions

The problems associated at the geographic level of population analysis also apply to the question of taxonomy within the species. That is, because there is a very strong possibility that homoplasy in the microsatellite data set is rampant, their usefulness for determining if there is segregated genetic exchange between clade 1 and 2 is also limited. This issue remains unresolved. The results presented here provide a clear caveat in regards to the use of microsatellite markers for estimating nuclear gene flow, or as comparative loci to mtDNA analyses of population structure and species delineation. The potential high rates of mutation attainable at microsatellites are useful for paternity assays, however this same trait may (as was shown here) also limit the power of their resolution for addressing historical population or taxonomic evolutionary questions.

The high diversity of alleles observed within populations may however prove useful for monitoring extremely recent demographic events. The microsatellites may be effective for monitoring of potential crashes in *Ne* at populations subject to natural bottlenecks or intense fishing pressures. Similarly, estimates of microsatellite allelic diversity present at newly colonized areas by the species may provide useful information concerning the potential number of successful recruits entering an area. This will be explored in the next chapter.

CHAPTER 6: **Colonisation of the southwest Australian coastline by *Scylla serrata*: evidence for a recent range expansion, or human induced translocation?**

6 .1 Introduction

The reported distribution of *S. serrata* in Australian waters among estuarine habitats from Exmouth to Sydney (Heasman *et al*. 1985, Keenan *et al*. 1998) spans a large portion of the western, northern, and eastern margins of the continent. The absence of this species on the cooler southern Australian coastline, like that for a number of tropical marine species, may be due to temperature intolerance (Veron 1995). Incidental observations since 2000 of established adult mud crab populations inhabiting estuaries on the southwest corner of Australia are almost 1000 kilometres south of the limit of their previously observed distribution on the northwest Australian coastline.

Coastal shelf waters of southwest Australia mark a junction between tropical and temperate biogeographical zones (Wilson and Allen 1987; Hutchins 2001; Williams *et al*. 2001) and are strongly influenced by the Leeuwin Current. The Leeuwin consists of warm tropical waters driven south and parallel to the outer fringe of the Western Australian continental shelf in a southerly direction. This current is strongest during winter – autumn, providing a seasonal conduit for the transfer of low latitude tropical plankton along the western and southern Australian coastline (Hutchins 1991). Anomalous recruitment patterns for a variety of coastal marine species to the southwest region may be linked to variation in the intensity and duration of the Leeuwin Current due to changes in the Southern Oscillation Index (Caputi *et al*. 1996). It is possible that a range expansion of northwest Australian *S. serrata* populations into the southwest region has occurred in response to a reported record intensification of the Leeuwin current during the 1999 season. Alternatively, it is suggested that entry of *S. serrata* into southwest Australian estuaries was a result of artificial translocations, either via ballast release from international shipping or dumping of commercial harvest derived from the Gulf of Carpentaria, Northern Australia (the source of crabs for most of the region's consumption).

To investigate this recent colonisation, 'immigrant' southwest populations were sampled for genetic material to be compared with previously described genetic profiles

of IWP and Australian *S. serrata* populations. I argued in Chapter 4 that maternal gene flow among *S. serrata* populations is more frequent along coastal shelf areas connected by similar currents than areas not connected by either currents or coastal shelf. Therefore, it is expected that a natural range expansion by this species is more likely to be derived from a neighbouring shelf population than a trans - oceanic source. The aims here are to identify the genetic source of the southwest populations using mtDNA analysis. If the immigrants have extended south with the Leeuwin current then it is expected that they will have greater genetic affinity with northwest Australian populations than populations outside of this region. Alternatively, immigrants with genetic profiles derived from outside of the northwest Australian region would be symptomatic of human induced translocation. The mtDNA analysis will also determine if more than one species of *Scylla* are present among the immigrant populations.

Finally, I compare genetic diversity within the southwest populations relative to suspected source populations, using both haploid mtDNA and diploid microsatellite markers. The number of individuals founding a new population initially limits both the potential effective population size (*Ne*) and the amount of genetic diversity available to the colony (Nei *et al*. 1975). Therefore, comparison of genetic diversity among source and colonist populations may provide a qualitative means of comparing *Ne* among the populations (Shephard *et al*. 2002). If few individuals founded the immigrant populations, reduced levels of genetic diversity across all markers relative to the source populations can be expected. Alternatively, similar levels of diversity among source and immigrant populations would be indicative of substantial transfer of propagules from source populations.

6 .2 Methods

6.2.1 Sampling and laboratory analysis

Mud crabs were collected from six recently colonised estuarine locations of the south – west corner of Australia (Figure 6.1 & Table 6.1) during the summer transition of 2001 – 2002. DNA extraction and PCR protocols followed those previously described for amplifying mtDNA COI and microsatellite alleles. MtDNA samples were scored for haplotype identity relative to sequenced individuals using temperature gradient gel electrophoresis of heteroduplexed samples. Immigrant crabs were also screened for

variation at two co-dominant polymorphic microsatellite loci (Ss-101 & Ss-403) using methods described in Chapter 2.2.7.

Table 6.1 Sample size (*N*) from six southwest Australian locations and their co-ordinates. Carapace width size range in millimetre's (mm) and percentage of females sampled per location as indicated.

Location	Co-ordinates	*N*	Size (mm)	Sex (% ♀)
Moore River	31°21'S; 115°29'E	02	164-170	?
Swan River	32°03'S; 115°44'E	14	146-195	40
Leschenault estuary	33°13'S; 115°42'E	02	> 189	00
Busselton	33°38'S; 115°20'E	03	118-190	?
Wilson Inlet	34°59'S; 117°25'E	10	127-200	100
Kalgan River	34°57'S; 117°58'E	01	200	100

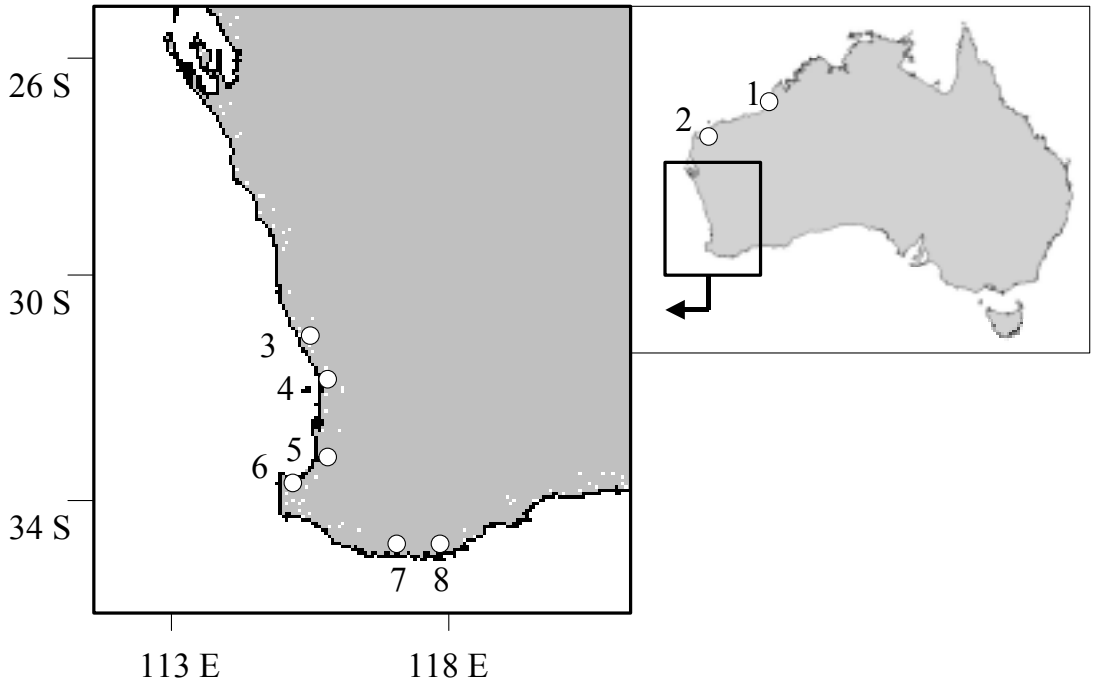


Figure 6.1 Sample locations from northwest Australia: (1) Broome and (2) Exmouth; and the southwest: (3) Moore River, (4) Swan River, (5) Leschenault estuary, (6) Bussleton, (7) Wilson Inlet and (8) Kalgan River.

6.2.2 Phylogenetic assessment

To identify the species and potential source population(s) of the immigrants I compared sequences derived from immigrant crabs with those for all other *Scylla* haplotypes. Sequences of the immigrant mtDNA haplotypes were aligned and compared against all previously identified homologous *Scylla* sequences.

6.2.3 Estimates of genetic diversity

Genetic diversity of both the immigrant and suspected source population(s) was estimated for all genetic markers employed. Because of low sample size at some locations, microsatellite analyses of immigrants were limited to Swan River ($N = 14$) and Wilson Inlet ($N = 10$). Summary diversity measures and exact tests for homogeneity of microsatellite allele frequency between source and immigrant populations were conducted as per Chapter 5.2.3.

Immigrant populations were grouped and examined for loss of allelic diversity at microsatellite loci as might occur immediately following a severe bottleneck or founder event, using methods described in Chapter 2.3.5. Each locus was tested for a significant excess or deficit of H_E relative to H_{eq}, with the probability of departure obtained from simulations. A shared pattern of significantly excessive H_E among loci would be indicative of a bottleneck effect. Samples with incomplete genotypes were excluded from analysis.

6 .3 Results

6.3.1 Species phylogeny and identification

A single mtDNA haplotype was detected in the entire sample of immigrant crabs. Assessment of the immigrant haplotype sequence by alignment with all other homologous *Scylla* haplotypes indicates that it is synonymous with the most frequent clade 2 haplotype previously observed from Western Australian *S. serrata* populations (haplotype designation 2-J; Chapter 4). As reported in Chapter 4, haplotype 2-J is endemic to Western Australian populations and has not been recorded from any other Australian or IWP sample locations. It is therefore highly likely that the immigrant

crabs are entirely derived from Western Australian populations of *S. serrata*. There was no evidence of other species of *Scylla* in the immigrant sample despite the sympatric co-distribution of *S. serrata* and *S. olivacea* within the northwest Australian region (Taylor 1984).

6.3.2 Genetic diversity among source and colonist populations

Despite similar sample sizes employed (Table 6.2), mtDNA diversity within the southwest was totally depauperate and contrasts with the moderate diversity levels (average $h = 0.466$) reported in the potential northwest Australian source populations (sampled populations, Exmouth and Broome; Chapter 4).

Table 6.2 Summary of variation at mtDNA COI locus and two microsatellite loci (Ss-101 & Ss-403) among *N* samples of *Scylla serrata.* Samples from northwest (Broome and Exmouth) and southwest (Swan River and Wilson Inlet) Australian regions. For the mtDNA locus, N_{Haps} refers to the number of detected haplotypes and *h* is estimated haplotype diversity (Nei, 1987). For microsatellites, *K* is the number of observed alleles, *Hobs* and *HE* refer to observed and expected heterozygosity (Nei, 1978), and *P (HW)* is the probability of deviation from that expected from Hardy-Weinberg equilibrium.

Locus		Broome	Exmouth	northwest (total)	Swan River	Wilson Inlet	southwest (total)
MtDNA							
	N	8	24	32	14	10	32
	N_{Haps}	3	4	4	1	1	1
	h	0.543	0.431	0.466	0	0	0
Ss-101							
	N	5	22	27	14	8	22
	K	8	19	21	12	11	18
	Hobs	1.0	0.955	0.963	0.857	1.0	0.909
	HE	0.956	0.956	0.955	0.884	0.925	0.907
	P(HW)	1.000	0.119	0.351	0.364	0.858	0.521
Ss-403							
	N	8	22	30	14	9	23
	K	7	17	19	13	11	19
	Hobs	0.625	0.864	0.801	0.786	0.889	0.826
	HE	0.858	0.903	0.899	0.921	0.915	0.923
	P(HW)	0.676	0.564	0.173	0.280	0.647	0.214

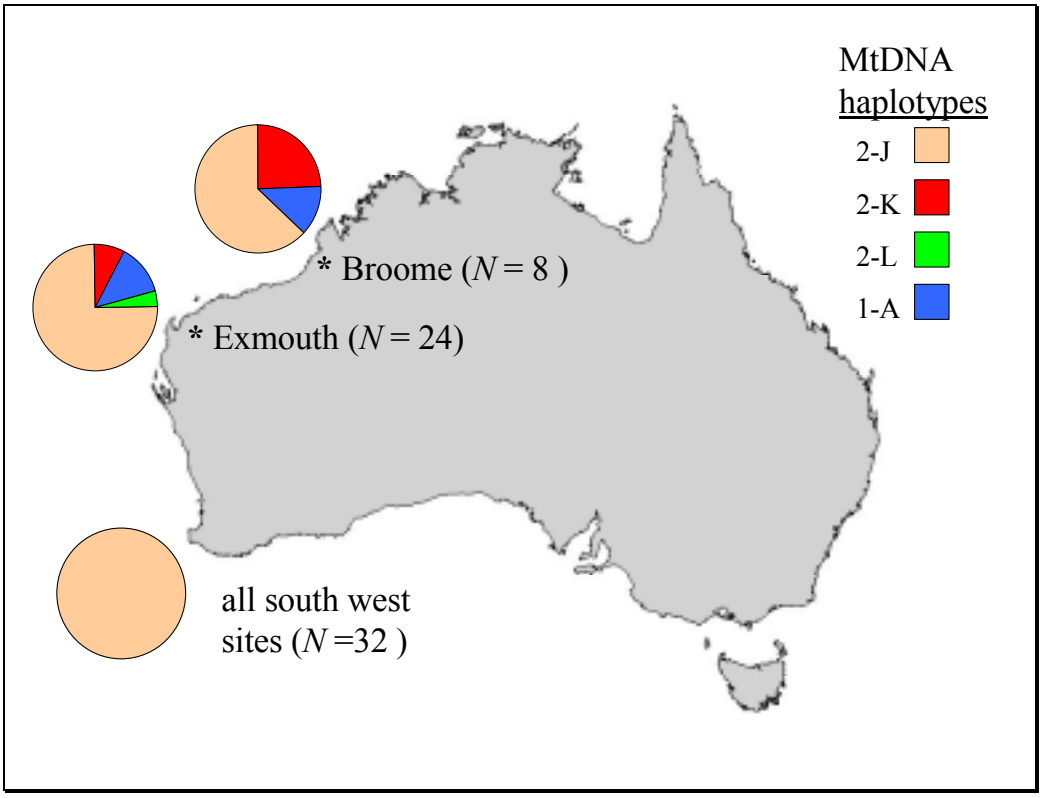


Figure 6.2 Distribution of all detected mtDNA haplotypes among northwest and southwest regions of Australia.

It is interesting to note the absence of three mtDNA haplotypes within the southwest region that collectively represent approximately 30 % of the sample from the northwest Australian populations (Figure 6.2). In contrast to the mtDNA results, microsatellite diversity levels were equally high among source and immigrant populations. Average observed and expected heterozygosity estimates at microsatellite loci were greater than 0.800 at all locations, due in part to the large number of alleles within locations observed for each locus (Table 6.2). Although some low frequency alleles were specific to each of the southwest and northwest region populations (Figure 6.3), exact tests for differences in distribution and frequency of alleles among these regionally grouped populations were not significant (Locus Ss-101 $P = 0.825$; Ss-403 $P = 0.311$).

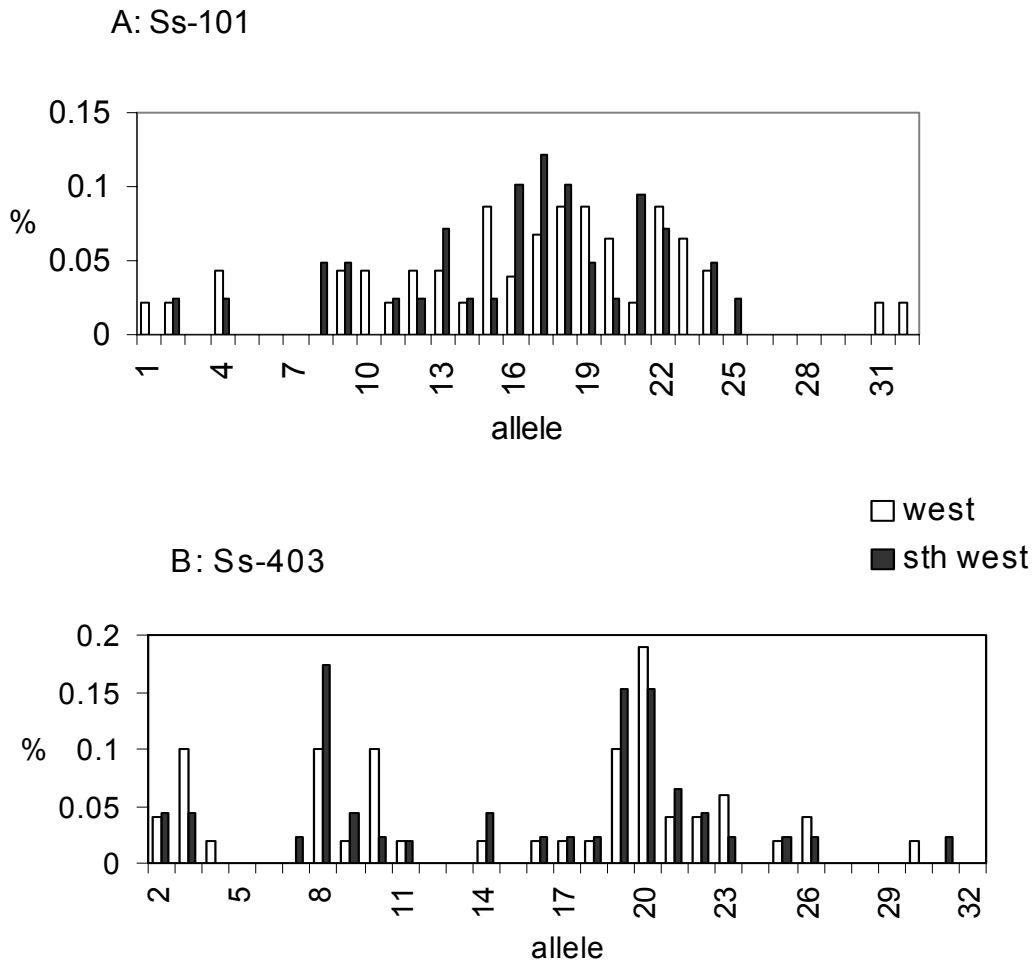


Figure 6.3 Distribution and comparison of alleles found at northwest and southwest regions for (A) microsatellite locus Ss-101 and (B) microsatellite locus Ss-403.

There was no evidence of a significant excess of *HE* relative to *Heq* at either of the microsatellite loci within the northwest or southwest population groups (Table 6.3) as might be expected if any of the groups had experienced a recent reduction in *Ne*. On the contrary, both loci in the southwest tended towards reduced *HE* and in one instance (Locus Ss-101), *HE* was significantly ($P < 0.05$) smaller than *Heq*. Reduced *HE* at these

loci was due to a slight excess of homozygotes. It is possible that grouping the two immigrant populations created an artificial Wahlund effect resulting in this excess. Exact tests for deviation from HWE were however non significant within the immigrant populations (Table 6.2).

Table 6.3 Tests for differences between average expected heterozygosity (*HE*) and equilibrium heterozygosity (*Heq*) for two microsatellite loci within regional northwest and southwest Australian populations. Probability (*P*) that *HE* differs from *Heq* obtained from coalescent simulation (1000 iterations) of the observed number of alleles (*K*) observed among *N* diploid crabs, * $P < 0.05$ (two tailed test).

	Locus	*2N*	*K*	*HE*	*Heq*	*P*
northwest						
	Ss-101	52	21	0.955	0.943	0.137
	Ss-403	54	19	0.899	0.927	0.056
southwest						
	Ss-101	42	18	0.907	0.931	0.049*
	Ss-403	46	19	0.923	0.940	0.079

6 .4 Discussion

6.4.1 Range expansion or translocation?

Expansions by *Scylla spp* into new territory have been reported on three prior occasions. Two of these expansions were deliberate translocations of adult crabs (to Hawaii and Florida) and in one case a lengthy period of translocation resulted in permanent self - perpetuating populations among the major Hawaiian Islands (Brock 1960). The means of transfer for a reported occurrence of a small number of *S. serrata* on New Zealand's north coast is equivocal. Seventeen adult crabs (mostly male) were found among several north coast estuaries during 1964 -5 (Dell 1964; Manikiam 1967). It was suggested by Dell (1964) that this appearance of *S. serrata* was possibly due to a single chance dispersal event by the species across the Tasman Sea from Australia. The apparent absence of crabs prior to 1964 and subsequent failure to detect crabs after 1965 led Manikiam (1967) to suggest that the occurrence of mud crabs in New Zealand was due to a single invasion by adult crabs that failed to establish a permanent population. Whether or not this occurrence resulted from a translocation or a natural process(es) is unknown.

The recent appearance of mud crabs on the southwest Australian coast since 2000 is extraordinary in terms of both the scale and location of the occurrence. Large numbers of crabs (juveniles and adults) have reportedly been caught at numerous major estuaries within the Southwest, spanning more than 650 kilometres of temperate coastline. My analysis of crabs sampled from a number of the estuaries indicates that *S. serrata* populations of northwest Australia are the most probable source of the southwest Australian mud crabs. This is based on the presence of a single mtDNA haplotype detected within the immigrant populations that is synonymous with one endemic to northwest Australian *S. serrata* populations. The absence of other haplotypes within the southwest rules out possible translocations of *Scylla spp.* from areas external to the northwest Australian region. Although I cannot categorically rule out the possibility of translocation from the northwest Australian region, I argue that the colonisation of the southwest by this species has likely occurred by a natural range expansion.

Though separated by a distance of approximately 1000 kilometres with scarce intervening estuarine habitats, it is not improbable for recruitment to have occurred from Exmouth (the southern most distribution of *S. serrata* in northwest Australian

waters) to the southwest. Studies of the potential range of dispersal by *S. serrata* indicate that adult movements are generally limited and confined to intertidal areas (Hyland *et al*. 1984). The exception occurs during spawning migrations when ovigerous females migrate up to 95 kilometres offshore to release eggs (Hill 1994). The extent of dispersal during the plankton stages is unknown, however it has been speculated that considerable dispersal is possible following the offshore release and entrainment into currents (Hill 1994). Previous genetic analysis of *S. serrata* populations indicated that dispersal by this species, though sufficient to maintain genetic homogeneity over thousands of kilometres, is limited to continental shelf areas connected by shared current systems (chapters 3 and 4). This suggests a strong association between current driven dispersal and population connectivity for this species.

The most obvious question about the colonisation of the southwest concerns the timing of the event – why has it not happened in the past? A number of established commercial fisheries exist within the southwest, including one for the blue swimmer crab *Portunis pelagicus*. Harvest of this species uses methods also adopted for the harvest of mud crabs in Australia; there are however no recorded reports of incidental by-catch of mud crabs from this fishery prior to 2000. Therefore this colonisation appears to be a recent phenomenon. *S. serrata* can tolerate a variety of marine environments, however there are critical salinity and temperature requirements for survival of the first stage zoeae (Hill 1974). Optimum survival conditions for this stage occur in marine saline waters ranging in temperature from 14 ° to 20 ° C. Peak spawning and recruitment of *S. serrata* in Australian waters occurs during the spring - summer months (Keenan *et al*. 1998), coincident with the period of weakest flow by the Leeuwin current. The distinctive seasonality in the strength of the Leeuwin and period of recruitment by this species may in part explain why *S. serrata* propagules normally fail to infiltrate areas on the western coast south of Latitude 22 °S. Strengthening of the Leeuwin system during summer months in response to positive Southern Oscillation Index / La - Nina events, allows tropically derived warm waters to penetrate well inshore along the southwest coastline (see Wilson *et al*. 2001). These events may provide a means for species that disperse during the warmer seasons to gain access to the southwest region. It may not be a coincidence that sightings of adult mud crabs in the southwest were reported approximately 2 years after the strongest summer Leeuwin flow seen in 80 years. Mud crabs can attain maturity within 18 months from settlement so it is entirely feasible that

S. serrata plankton recruited to and colonised the southwest estuaries as a result of and during the 1999 La Nina event.

In contrast to Western Australia, permanent *S. serrata* populations on the eastern Australian coastline are recorded as far south as Latitude 36 °S, where the East Australian current drives tropical waters south during the spring - summer period. Similarly, permanent *S. serrata* populations on the eastern African coastline that extend to Latitude 34 °S (Robertson 1996) are supplied with tropical waters during the spring – summer transition via the influence of the Agulhas current. Marginal differences in tolerance to marine conditions among the different *Scylla spp* may ultimately determine their broad distribution (Keenan *et al*. 1998). It was argued by Keenan *et al*. that the presence of *S. serrata* at a variety of widespread IWP locations not inhabited by other mud crab species may be due to greater tolerance by this species towards higher salinity waters. While salinity levels may be a limiting factor, it is likely that water temperature may also determine the broad distribution for each species of *Scylla*. The complete absence of other *Scylla spp* from all temperate coastal locations inhabited by *S. serrata* supports this idea. Patterns of latitudinal gradation of coastal marine species in the IWP are greatly influenced by temperature and boundary currents (Veron 1995). The distribution of another genus of Portunid crabs (*Thalamita spp*.) on the Western Australian coast exhibits a latitudinal reduction in the number of species present from north to south (Stephenson and Hudson 1957) that may be related to temperature tolerance. The results of my study clearly indicate the southward colonisation by *S. serrata* only, despite the presence of at least one other species of mud crab (*S. olivacea)* in the northwest region (Taylor 1984) and the remaining species further north within equatorial South - East Asia (Keenan *et al*. 1998).

6.4.2 Contrasting levels of genetic diversity

Estimated levels of genetic diversity among source and colonist populations varied according to the genetic marker used, providing markedly different interpretations concerning the potential founder population size. Measures of allelic diversity at microsatellite loci are equally high among source and immigrant populations and provide no evidence of inbreeding and or reductions of effective population size. These observed levels of nuclear diversity are consistent with that expected for a mass transfer of propagules from a genetically heterogenous source, rather than a founding by relatively few individuals. In contrast to these results, microsatellite analyses of

translocated green crab populations (*Carcinus spp*.) indicate invasions by that genus have been accompanied by large reductions in allelic diversity and evidence of genetic bottlenecks among colonist populations (Bagley and Geller 2000). This effect on the invasive populations of *Carcinus spp* reinforces my contention that colonisation of the southwest by *S. serrata* occurred by natural rather than induced means. To observe the levels of genetic diversity in my study by translocation would require large numbers of propagules or breeding adults released independently into each of the southwest estuaries. As a caveat, Holland (2001) found evidence among invasive mussels (*Perna perna*) that a substantial portion of a foreign gene pool can gradually be transferred among adjacent estuaries following an initial release of plankton contained in ballast waters.

In contrast to the microsatellite results, mtDNA diversity in the southwest is depauperate relative to the northwest populations, indicating the immigrant populations experienced a mtDNA bottleneck and was therefore founded by a small sample of the northwest maternal gene pool. Due to its maternal mode of transmission and lack of recombination, mtDNA has an effective population size one quarter that of the nuclear genome. As an outcome, patterns of cytoplasmic mtDNA diversity are very susceptible to genetic drift and may differ from that of autosomal nuclear diversity due to differences in response by these two genomes to both stochastic and population processes (Birky *et al*. 1989). In particular, there is propensity for rapid loss of mtDNA alleles within newly colonised populations (Nei *et al*. 1975, Villablanca *et al*. 1998), especially if founding occurred by relatively few females or by closely related kin (Wade *et al*. 1994). It may be predicted that gross differences in diversity among nuclear and cytoplasmic genetic markers within colonist populations is common for plankton dispersed marine species. These species are often extremely fecund and exhibit patterns of recruitment subject to chaotic patchiness (Underwood and Fairweather 1989). Potential differences in size between the effective and absolute numbers of individuals colonising a patch may be of an order of magnitude resulting in far greater stochastic loss of diversity at mtDNA than at nuclear markers. This prediction is expected to hold whether colonisation occurred via natural means or by translocation.

CHAPTER 7: **General discussion**

7 .1 The amphitropical distribution of *Scylla serrata* populations throughout the Indo -West Pacific: some inferences.

The current distribution of *S. serrata* populations throughout the IWP inferred from the work of Keenan *et al*. (1998) may be described as *amphitropical* (sensu Veron 1995, pp 38), in that populations although generally absent from tropical equatorial areas, occur at a large number of oceanic islands and at all the major temperate coastal margins of the IWP. In contrast, the remaining three species within the *Scylla* genus are exclusively found among a variety of equatorial and tropical locations from the Indian sub-continent to Papua New Guinea (Keenan *et al*. 1998).

Similar patterns of distribution to that of *S. serrata* have been noted for a variety of marine taxa that are absent from, but range broadly either side of the Indo-Malay region (refer Briggs, 1995; pp 224-231). Briggs (1987, 1999) contends that concordant patterns of disjunct species distribution in the IWP are symptomatic of an ongoing process of extinction and replacement of basal species by more derived species within a geographical centre of origin. Briggs argued these extinctions occur primarily by species interactions (ie: competitive exclusion) and are more likely to affect older lineages in centres of high species diversity such as within the tropical waters of the Indo-Malay Archipelago. The process eventually leads to a "centrifugal" pattern of peripheral distribution and fragmentation of relict populations for older species, either side of the Indo-Malay Archipelago.

Alternatively, distributions of many IWP species may have been fragmented as a result of conditions experienced in the IWP during periods of Pleistocene glacial activity, akin to that purported to have occurred for terrestrial species in the northern hemisphere (see review, Hewitt 2000). The effects of vicariance, disrupted dispersal avenues, changed habitat and environmental conditions due to glacial episodes would have promoted the emergence of provincialism (Rosen 1984, Fleminger 1986) and genetic structure (McMillan and Palumbi 1995, Chenoweth *et al*. 1998, Barber *et al*. 2000) between species and populations either side of the Indo-Australian archipelago. This glacial - eustasy model of dispersal / vicariance predicts that subdivision of marine taxa by

vicariance (and subsequent expansion) into separate geographic provinces leads to convergent allopatric patterns of segregated sister populations thus promoting the emergence of sister species either side of the Indo- Australian Archipelago (McMillan and Palumbi 1995, Lessios *et al*. 2001).

Although these two hypotheses differ in terms of causality, both predict that disjunct distributions seen for a variety of IWP taxa result from historical partitioning of formerly widespread populations into separate geographic regions. The two hypotheses differ in that the relative "age" of a species with a disjunct population structure in the IWP is expected to be relevant for the former centrifugal hypothesis but not for the latter. Under Briggs's centrifugal hypothesis, amphitropical distributions are relicts of equatorial extinctions of older lineages. Therefore, centrifugal distributions are expected to occur for species representative of the older lineages within a genus.

Investigation of phylogenetic relationships among species of *Scylla* using mtDNA analyses both by Keenan *et al*. 1998 and in my work (Figure 3.2, Chapter 3) demonstrate extant populations of *S. serrata*, *S. tanquebarica* and *S. paramamosain* share a common ancestor and are nested as a clade derived from the basal lineage leading to extant populations of *S. olivacea*. Clearly then, *S. serrata* is a derived species within the genus, which unlike its tropically confined sister species has colonized a large variety of temperate locations either side of the Indo-Malay Archipelago. Therefore, I reject the centrifugal hypothesis as a valid explanation of contemporary *S. serrata* distribution.

Surprisingly, mtDNA analysis in this thesis also demonstrated a lack of evidence for a deep phylogenetic schism between provincial Indian and West Pacific *S. serrata* populations. This result is contrary to that demonstrated for an increasing number of coastal marine species in the IWP that mostly appear to demonstrate similar patterns of genetic provinciality either side of the Indo-Australian Archipelago. The equally shallow levels of nucleotide diversity seen both within and between the provincial Indian and west Pacific *S. serrata* populations (Table 3.5) clearly demonstrated that the contemporary widespread amphitropical distribution of this species resulted from a historically recent radiation throughout the region.

Data from an independent genetic survey of East African *S. serrata* populations by Fratini *et al*. (2002) mostly reflects the results and patterns shown for clade 1 crabs. Fratini *et al*. sequenced homologous mtDNA COI for 77 individuals sampled from four locations spanning ~ 350 kilometres of the central east African coast (3 locations from Kenya, and 1 from Zanzibar). Fratini *et al*. also incorporated my data from locations in the Indian Ocean (Red Sea, Mauritius, South Africa) to expand the geographic scale of their analysis. Haplotypes from the four central coast locations are all from a single clade (equivalent to Clade 1), with sharing of two high frequency haplotypes (homologous to the Mauritius and South African haplotypes "K" and "L" in Table 3.3) and no evidence of population subdivision among these central African locations. Strong population subdivision exists between this central group and the Red sea, and moderate subdivision when either South Africa or Mauritius is compared to the central locations. There was no evidence within their data of any clade 2 haplotypes or haplotypes intermediate between clade 1 and 2. Thus, their analysis confirmed my contention for a recent origin for the African crabs, consistent with a radiation of clade 1 crabs into this region during the late Pleistocene.

Curiously, the number of haplotypes per location and consequently haplotype diversity detected by Fratini *et al*. is greater than that seen in Chapters 3 and 4 in my analyses of IWP and Australian locations. The low haplotype diversity at IWP locations reported in Chapter 3 might indeed be an artifact of low sample sizes employed for that component. Larger sample sizes used for the Australian study (Chapter 4) are similar to that used by Fratini *et al*. and also show a marked increase in haplotype diversity compared to the IWP study. However, levels of diversity at east African locations reported by Fratini *et al*. (range: 0.37 - 0.85) are greater than at Australian locations (0.091 - 0.672) and bears some attention. There is reason to believe that high estimates of haplotype diversity reported by Fratini *et al*. may erroneously be inflated by sequencing artifacts (the authors sequenced haplotypes in one direction only), I argue that at least eight of the 22 low frequency haplotypes reported by Fratini *et al*. are questionable. Six of the variable nucleotide sites among detected haplotypes reported by Fratini *et al*. occur at 1st and 2nd position codon sites and all result in amino acid changes relative to a consensus South African *S. serrata* haplotype. Comparison of homologous sequences from a wide range of decapods derived from GENBANK (including the alternative 3 species of *Scylla*) indicates these 6 amino acids are conserved within the genus *Scylla* and also among a variety of other crab species. Therefore, it is highly unlikely that these six sites reported

by Fratini *et al*. are valid mutations (if they are valid mutations, it would indicate a severe deviation from the expectations of intra-specific molecular evolution at protein genes and should therefore have been investigated by the authors). Removing these six sites from the data of Fratini *et al*. effectively removes eight of the reported twenty-four haplotypes in their study and so reduces diversity levels in their study to levels similar to that seen among Australian locations.

Apart from clade 2 crabs restricted to the northern and western Australian coastal shelves, genetic evidence of prior radiations by this species throughout the IWP region is absent. As argued in several of the chapters, fossil evidence of *S. serrata* at peripheral IWP locations (east Australia, south Africa and south Japan) appears to pre-date the age of contemporary populations estimated from genetic data. I suggest that the complete absence of older mitochondrial lineages among any of the peripheral IWP populations provides strong evidence that prior populations of *S. serrata* at these locations went extinct. This happened more likely in response to the breakdown of coastal current pathways along continental shelf areas and to the pronounced cooling of sea temperatures in the paleo-temperate regions (Veron 1995, McGowran *et al*. 1997). Persistence of relict populations of *S. serrata* during these cooler periods appears to have ensured the persistence of the species over time. Evidence that the relicts may have persisted near the Indo-Australian Archipelago is demonstrated by the contemporary presence of clade 2 crabs in the northern and western regions of Australia.

The question arises as to whether these events also affected other species of *Scylla* in a similar fashion. Limited sampling precludes any comments for *S. tranquebarica* and *S. paramamosain.* However some preliminary mtDNA data presented for the equatorially distributed *S. olivacea* in chapter 3 clearly demonstrates that genetic results for this species contrasts sharply to that detected for *S. serrata*. Evidence of high haplotype and nucleotide diversity among the limited *S. olivacea* samples suggests that multiple lineages have evolved and persisted in the Indo-Australian area and may have been resilient to the periods of glacially induced habitat changes. Furthermore, considerable sharing of diverse mtDNA lineages is apparent among locations, allowing the tentative suggestion that there has been considerable gene flow and introgression among these tropical populations of *S. olivacea*.

It is possible that contrasting patterns of genetic structure among *S. serrata* compared to *S. olivacea* is due to greater affinity of the latter species for temperate conditions. Selection for greater tolerance of temperate conditions by *S. serrata* relative to other species in the genus (Chapter 6) may have allowed the more widespread contemporary distribution of the species. As demonstrated in Chapters 4 and 6, populations of *S. serrata* can potentially maintain population homogeneity along vast stretches of shelf connected coastal habitat and potentially colonise large expanses of temperate coastal area quite rapidly.

I propose that *S. serrata* populations have been subjected to cycles of retraction and expansion throughout the IWP in response to the cycles of glacial activity during the Pleistocene. As the species is adapted to temperate conditions, strong dispersal potential among coastal connected populations has allowed *S. serrata* its contemporary distribution at all the temperate continental coastal margins of the IWP. Because the effects of glacial cycles were more pronounced along temperate compared to tropical coastal areas (McGowran *et al*. 1997), populations of *S. serrata* may have retreated from and have repeatedly gone extinct at peripheral locations, thus reducing the overall net genetic diversity of this species in the IWP. The coalescence of continental shelves and islands in the Indo-Australian Archipelago during the low sea level stands, although separating other more resilient IWP populations of marine species, may in contrast have allowed the retraction and dispersal of genetically depauperate coastal *S. serrata* populations to equatorial refuges. Persistence of *S. serrata* (as clade 1 and clade 2 groups) in separate equatorial refugia during the glacial cycles may have allowed the species to avoid complete extinction as well as providing source populations for future expansion events following the reversal of glacial cycles.

7 .2 Cryptic species and evidence of nuclear conservation within *Scylla*.

One of the main unresolved issues within this thesis concerns the taxonomic status of mtDNA clade 2 relative to clade 1 *S. serrata*. The question remains, do the two clades merely represent distinct maternal lineages within a species that have been subjected to different demographic histories; or are they recent maternal signatures of two independently evolving and reproductively segregated species. There was no reported phenotypic variation at morphological, meristic or allozyme characters among crabs

representative of the two clades within the analysis of Keenan *et al*. (1998). This may not be surprising given that the levels of observed phenotypic difference between the four recognized species are very subtle and are likely evolved during periods of recent allopatric separation (Keenan *et al*. 1998). My attempt to answer this question by testing for cyto-nuclear linkage disequilibria was inconclusive (Chapter 6). There was no evidence at any of the five-microsatellite loci of a pattern among individual samples of segregated nuclear alleles concordant with the mtDNA clade structure. However, the power of that test for detecting congruent patterns among mtDNA and nuclear genes was confounded by the likely presence of rampant homoplasy at the microsatellite loci.

Preliminary results of additional tests for cyto-nuclear disequilibria have also been inconclusive, but for different reasons to that seen at the microsatellite loci. Genealogical concordance among mtDNA and nuclear genes provides a powerful means of testing phylogenetic relationships and providing evidence for non-interbreeding groups (Avise 2000). Nuclear encoded introns can be useful for phylogenetic analysis of within genus relationships when the ratio of mtDNA nucleotide divergence among maternal lineages is at least three times greater than within the lineages (Palumbi *et al*. 2001). Clearly, this ratio is present in mtDNA lineages not just among the four species of *Scylla*, but also between the two clades of *S. serrata*. Therefore I would expect that nuclear intron analysis should resolve taxonomic relationships both among the four recognized species of *Scylla* and also provide an adequate test for the presence of co-segregated nuclear and mtDNA alleles within *S. serrata*.

I targeted a portion of the nuclear encoded elongation factor-1α (EF-1α) gene among a limited number of *Scylla* samples (a single sample from each of the two clades of *S. serrata*, *S. tranquebarica* and *S. olivacea*) previously sequenced for mtDNA COI. The portion of targeted EF-1α was homologous to that used by Williams *et al*. 2001 to identify species relationships among shrimp species within the genus *Alpheus*, and also encompassed the intron region targeted by Duda and Palumbi (1999) to detect population structure among Indian and Pacific populations of the tiger prawn *Penaeus monodon*. Levels of detected polymorphism at both of those prior studies were sufficient for phylogenetic analysis of historically separate groups of species and populations.

Results of analysis of *Scylla* samples using the nuclear exon / intron gene sequence, indicated mutation rate at this gene region is severely conserved relative to that seen at mtDNA, not only between the two clades of *S. serrata*, but also among three of the recognized species of *Scylla* (Table 7.1 and 7.2).

Table 7.1 Percent paired sequence difference among 4 samples representing species within the *Scylla* genus at 549 bp of the mtDNA COI gene (upper matrix) and 504 bp of the nuclear EF-1α gene (lower matrix). *S. serrata* # 1 and #2 refer to mtDNA clade 1 and 2 samples within that species.

	S. serrata #1	*S. serrata #2*	*S. tranquebarica*	*S. olivacea*
S. serrata #1		2.5	10.0	14.8
S. serrata #2	0		9.8	14.8
S. tranquebarica	1.0	1.0		13.9
S. olivacea	1.2	1.2	1.4	

Table 7.2 Summary of nucleotide variation among all four samples of *Scylla* at the mtDNA COI and nuclear EF-1α genes.

Gene	Total # of substitutions	# of sites (bp) compared	substitutions /site (%)	Substitution ratio *
COI	106	549	19.3	10.7
EF-1α	9	504	1.8	1.0

* relative to the EF-1α gene.

The substitution ratio between mtDNA COI and nuclear EF-1α gene for the *Scylla* genus (10.7 : 1) is more than twice that detected within the *Alpheus* shrimp genus (4.9 : 1) reported by Williams *et al*. (2001b) and the 3.8 : 1 ratio reported by Montiero and Pierce (2001) for the *Bicyclus* genus of moths. It can be expected that there will be a greater rate of neutral mutation accumulation among lineages at mtDNA compared to nuclear coding genes due to differences in *Ne* between the genomes (Birky *et al*. 1989),

however this rate appears to be far greater within the *Scylla* genus compared to other arthropod genera. It may therefore be surmised for the *Scylla* genus, that either the rate of nucleotide substitution is accelerated at the mitochondrial genome, or there is a selective constraint against the accumulation of mutations at the nuclear EF-1α locus. Observed rates of allozyme polymorphism and divergence are also reported to be abnormally low among species of *Scylla* compared to other marine genera (Sugama and Hutapea 1999, Fuseya and Watanabe 1996) and preclude their use for any meaningful intraspecific analyses. These results are at odds with the levels of polymorphism observed within the genus at the neutral microsatellite loci. It remains to be seen if levels of polymorphism at various classes of nuclear protein coding gene are also conserved within this genus.

Finis.

References

Appleyard SA, Grewe PM, Innes BH, and Ward RD (2001) Population structure of yellowfin tuna (*Thunnus albacares*) in the western Pacific Ocean, inferred from microsatellite loci. *Marine Biology* 139: 383-393

Arnaud S, Bonhomme F and Borsa P (1999) Mitochondrial DNA analysis of the genetic relationships among populations of scad mackeral (*Decapterus macarellus*, *D. macrosoma*, and *D. russelli*) in South-East Asia. *Marine Biology* 135: 699-707

Avise JC (1994) Molecular markers, natural history and evolution. Chapman and Hall, New York

Avise JC (1995) Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. *Conservation Biology* 9: 686-690

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University Press, Cambridge, Massachusetts

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA and Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522

Bagley MJ and Geller JB (2000) Microsatellite DNA analysis of native and invading populations of European green crabs. In: Pederson J (ed) *Marine Bioinvasions: Proceedings of the First National Conference.* MIT Sea Grant College Program, pp 241-243

Baker CS, Medrano-Gonzales L, Calambokidis J, Perry A, Pichler F, Rosenbaum H, Straley JM, Urban-Ramirez J, Yamaguchi M, and Von Ziegesar O (1998) Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology* 7: 695-707

Ballard JW and Kreitman M (1994) Unravelling selection in the mitochondrial genome of *Drosophila*. *Genetics 138*: 757-772

Balloux F and Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11: 155-165

Barber PH, Palumbi SR, Erdmann MV and Moosa MK (2000) A marine Wallace's line? *Nature* 406: 692-3

Begg GA, Keenan CP and Salini MJ (1998) Genetic variation and stock structure of school mackerel and spotted mackerel in northern Australian waters. *Journal of Fish Biology* 53 543-59

Benzie JAH (1999) Genetic structure of coral reef organisms: Ghosts of the dispersal past. *American Zoologist* 39: 131-145

Benzie JAH (2000) Population genetic structure in penaeid prawns. *Aquaculture Research* 31: 95-119

Benzie JAH and Williams ST (1997) Genetic structure of giant clam (*Tridacna maxima*) populations in the west Pacific is not consistent with dispersal by present - day ocean currents. *Evolution* 51: 768-783

Bergquist PR (1967) Australian intertidal sponges from the Darwin area. *Micronesica* 3: 175-202

Birky CW, Maruyama T and Fuerst P (1983) An approach to population genetic and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103: 513 –527

Birky CW(Jr), Maruyama T and Fuerst P (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121: 613-27

Blum SD (1989) Biogeography of the Chaetodontidae: an analysis of allopatry among closely related species. *Environmental Biology of Fishes* 25: 9-31

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, and Cavalli-Sforza LL (1994) High resolution of human evolution with polymorphic microsatellites. *Nature* 368: 455-457

Briggs JC (1987) Antitropical distribution and evolution in the Indo-West Pacific Ocean. *Systematic Zoology* 36: 237-247

Briggs JC (1995) Global Biogeography. Elsevier press, Amsterdam

Briggs JC (1999) Coincident biogeographic patterns: Indo-West Pacific ocean. *Evolution* 53: 326-335

Brock V (1960) The introduction of aquatic animals in Hawaiian waters. *Internationale Review der Gesamten Hydrobiologie* 45: 463-80

Brooker AL, Benzie JAH, Blair D and Versini JJ (2000) Population structure of the giant tiger prawn *Penaeus monodon* in Australian waters, determined using microsatellite markers. *Marine Biology* 136: 149-157

Brower AVZ (1994) Rapid morphilogical radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Science, USA.* 91: 6491-6495

BuonaccorsiVP McDowell JR and Graves JE (2001) Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Molecular Ecology* 10: 1179-1196

Burton RS and Feldman MW (1982) Population genetics of coastal and estuarine invertebrates: Does larval behaviour influence population structure? In: Kennedy VS (ed) *Estuarine Comparisons*. Academic Press, New York USA, pp 537-551

Bush ABG and Philander SGH (1998) The role of Ocean-Atmosphere interactions in tropical cooling during the last Glacial Maximum. *Science* 279: 1341-1344

Campbell NJH, Harriss FC, Elphinstone MS and Baverstock PR (1995) Outgroup heteroduplex analysis using temperature gradient gel electrophoresis: high resolution, large scale, screening of DNA variation in the mitochondrial control region. *Molecular Ecology* 4: 407-418

Caputi N, Fletcher WJ, Pearce A and Chubb CF (1996) Effect of the Leeuwin Current on the recruitment of fish and invertebrates along the West Australian coast. *Marine and Freshwater Research* 47: 147-55

Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology* 77: 1653-5

Carr MH, and Reed DC (1993) Conceptual issues relevant to marine harvest refuges: Examples from temperate reef fishes. *Canadian Journal of Fisheries and Aquatic Sciences* 50: 2019-2071

Chappell JMA (1976) Aspects of late Quaternary paleogeography of the Australian - East Indonesian region. In: Kirk RL, Thorne AG (eds) *The Origin of the Australians*. Human Biology Series No.6, Australian Institute of Aboriginal Studies, Canberra, pp 11-77

Chappell J (1983) A revised sea-level record for the last 300,000 years from Papua New Guinea. *Search* 14: 99-101

Chappell J and Shackleton NJ (1986) Oxygen isotopes and seal level. *Nature* 324: 137-140

Chenoweth SF and Hughes JM (1997) Genetic population structure of the catadromous Perciform: *Macquaria novemaculeata* (Percichthyidae). *Journal of Fish Biology* 50: 721-33

Chenoweth SF, Hughes JM, Keenan CP and Lavery S (1998-a) Concordance between dispersal and mitochondrial gene flow: isolation by distance in a tropical teleost, *Lates calcarifer* (Australian barramundi). *Heredity* 80: 187-197

Chenoweth SF, Hughes JM, Keenan CP and Lavery S (1998-b) When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. *Proceedings of the Royal Society of London .B* 265: 415-20

Church JA (1987) East Australian Current adjacent to the Great Barrier Reef. *Australian Journal of Marine and Freshwater Research* 38: 671-83

Clary DO and Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba* : nucleotide sequence, gene organisation, and genetic code. *Journal of Molecular Evolution* 22: 252-271

CLIMAP Project Memebers (1976) The surface of the ice age earth. *Science* 191: 1131-1141

Condie SA, Loneragan N and Die DJ (1999) Modelling the recruitment of tiger prawns *Penaeus esculentus* and *P. semisulcatus* to nursery grounds in the Gulf of Carpentaria, northern Australia: implications for assessing stock-recruitment relationships. *Marine Ecology Progress Series* 178: 55-68

Cornuet JM and Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144: 2001-2014

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 153: 1989-2000

Crandall KA and Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134: 959-969

Crisp DJ (1978) Genetic consequences of different reproductive strategies in marine invertebrates. In: Battaglia B, Beardmore J (eds) *Marine organisms, Genetics, Ecology, and Evolution*. Plenum Press, New York, pp 257-273

Davie PJF (1985) The biogeography of littoral crabs (Crustacea: Decapoda: Brachyura) associated with tidal-wetlands in tropical and sub-tropical Australia. In: Bardsley KN, Davie JDS, Woodroffe CD (eds) *Coasts and Tidal Wetlands of the Australian Monsoon Region*. Australian National University North Australia Research Unit Mangrove Monograph 1, pp 259-275

Day R and McEdward L (1984) Aspects of the physiology and ecology of pelagic larvae of marine benthic invertebrates. In: Steidinger KA, Walker LM (eds) *Marine Plankton Life Cycle Strategies*. CRC Press Inc., Florida

De Innocentiis S, Sola L, Cataudella S and Bentzen P (2001) Allozyme and microsatellite loci provide discordant estimates of population differentiation in the endangered dusky grouper (*Epinephelus marginatus*) within the Mediterranean Sea. *Molecular Ecology* 10: 2163-2175

Dell RK (1964) The large Indo – Pacific swimming crab Scylla serrata (Forskål) in northern New Zealand. *Records of the Dominion Museum, Wellington* 5(8): 59-62

Doherty PJ, Planes S and Mather P (1995) Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology* 76: 2373-91

Duda TF Jr. and Palumbi SR (1999) Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Marine Biology* 134: 705-710

Duke NC, Benzie JAH, Goodal JA and Ballment ER (1998) Genetic structure and evolution of species in the mangrove genus *Avicennia* (Avicenniaceae) in the Indo-West Pacific. *Evolution* 52: 1612-26

Du Plessis A (1971) A preliminary investigation into the morphological characteristics, feeding, growth, reproduction and larval rearing of *Scylla serrata* Forskål (Decapoda: Portunidae) held in captivity. Fisheries Development Corporation of South Africa

Dynesius M and Jansson R (2000) Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences USA* 97: 9115-9120

Elliot NG (1996) Allozyme and Mitochondrial DNA analysis of the tropical saddle-tail sea perch, *Lutjanus malabaricus* (Schneider), from Australian waters. *Marine and Freshwater Research* 47: 869-76

Endean R (1957) The biogeography of Queensland's shallow-water echinoderm fauna (excluding Crinoidea), with a rearrangement of the faunistic provinces of tropical Australia. *Australian Journal of Marine and Freshwater Resources* 8: 233-273

Estoup A Jarne P and Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11: 1591-1604

Excoffier L and Langaney A (1989) Origin and differentiation of human mitochondrial DNA. *American Journal of Human Genetics* 44: 73-85

Excoffier L and Schneider S (1999) Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proceedings of the National Academy of Science USA* 96: 10597-10602

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to Human mitochondrial DNA restriction data. *Genetics* 131: 479-491

Felsenstein J (1988) Phylogenies from molecular sequences: Inference and reliability. *Annual Review of Genetics* 22: 521-565

Felsenstein J (1995) PHYLIP (phylogeny inference package). Version 3.5c. Department of Genetics, University of Washington, Seattle

Fielder DF and Heasman MP (1986) The mud crab., Queensland Museum Booklet # 11.

Fisher R (1954) Statistical Methods for Research Workers. 12th Edition, Oliver and Boyd, Edinburgh

Fleminger A (1986) The Pleistocene equatorial barrier between the Indian and Pacific oceans and a likely cause for the Wallace's line. In: Pierrot-Bults AC, Van Der Spoel S, Aahuraner BJ, and Johnson RK (eds) *UNESCO Technical papers in Marine Science* 49: pp 84-97

Forskål P (1775) Descriptiones animalium, amphibiorum, piscium, insectorum, vermium; quae in itinere orientali obsevit Petrus Forskål. 1-19, I-XXXIV, 164 pp. Post mortem auctoris editit Carsten Niebuhr. Hauniae.

Fratini S and Vannini M (2002) Genetic differentiation in the mud crab Scylla serrata (Decapoda: Portunidae) within the Indian Ocean. *Journal of Experimental Marine Biology and Ecology* 272: 103-116

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925

Fuseya R and Watanabe S (1996) Genetic variability in the mud crab Genus *Scylla* (Brachyura: Portunidae). *Fisheries Science* 62: 705-709

Garza J, Slatkin M, Freimer N (1995) Microsatellite allele frequencies in humans and chimpanazees with implications for constraints on allele size. *Molecular Biology and Evolution* 12: 594-603

Gaylord B and Gaines SD (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *The American Naturalist* 155: 769-789

Geller JB, Carlton JT and Powers DA (1994) PCR – based detection of mtDNA haplotypes of native and invading mussels on the northeastern Pacific coast: latitudinal patterns of invasion. *Marine Biology* 119: 243-9

Geller JB, Walton ED, Grosholz ED and Ruiz GM (1997) Cryptic invasions of the crab *Carcinus* detected by molecular phylogeography. *Molecular Ecology* 6: 901-6

Gold JR and Turner TF (2002) Population structure of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico, as inferred from variation in nuclear-encoded microsatellites. *Marine Biology* 140: 249-265

Goodman SJ (1997) RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and gene flow from microsatellite data and determining their significance. *Molecular Ecology* 6: 881 – 885

Goudet J (1995) F-STAT (version 1.2): a computer program to calculate *F* – statistics. *Journal of Heredity* 86: 485 - 486

Goudet J (1999) An improved procedure for testing key innovations. *American Naturalist* 53: 549-555

Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology and Evolution* 17: 22-7

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41: 95-98

Hanski I (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos* 38: 210-21

Harrison S (1991) Local extinction in a metapopulation context: an empirical evaluation. *Biological Journal of the Linnean Society* 42: 73-88

Heasman MP (1980) Aspects of the general biology and fishery of the mud crab *Scylla serrata* (Forskål) in Moreton Bay, Queensland. PhD Thesis, Zoology Department, University of Queensland

Heasman MP, Fielder DR, and Shepherd RK (1985) Mating and spawning in the mud crab, *Scylla serrata* (Forskål) (Decapoda: Porunidae), in Moreton Bay, Queensland. *Australian Journal of Marine and Freshwater Research* 36: 773-783

Hedgecock D (1986) Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bulletin of Marine Science* 39: 550-564

Hedrick PW (1999) Perspective: Highly variable loci and their interpretataion in evolution and conservation. *Evolution* 53: 313-318

Hellberg ME (1994) Relationships between inferred levels of gene flow and geographic distance in a phylopatric coral, *Balanophyllia elegans*. *Evolution* 48: 1829-1854

Hellberg ME (1996) Dependence of gene flow on geographic distance in two solitary corals with different larval dispersal capabilities. *Evolution* 50: 1167-75

Helmke SA, Peverell SC, and Garrett RN (2002) Fisheries resources of Trinity Inlet. Queensland Information Series Publication. Queensland Department of Primary Resources. (In Press).

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247-76

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405: 907-913

Hey J and Harris E (1999) Population bottlenecks and patterns of human polymorphism. *Molecular Biology and Evolution* 16: 1423 – 1426

Hill BJ (1974) Salinity and temperature tolerance of zoeae of the portunid crab *Scylla serrata*. *Marine Biology* 25: 21-4

Hill BJ (1975) Abundance, breeding and growth of the crab *Scylla serrata* in two South African estuaries. *Marine Biology* 32: 119-26

Hill BJ (1982) The Queensland Mud crab Fishery. Queensland Fisheries Information Series FI 8201: 1-41

Hill BJ (1994) Offshore spawning by the portunid crab *Scylla serrata* (Crustacea: Decapoda). *Marine Biology* 120: 379-84

Hill BJ, Williams MJ and Dutton P (1982) Distribution of Juvenile, Subadult and Adult *Scylla serrata* on tidal flats in Australia. *Marine Biology* 69: 117-120

Hill D, Playford G and Woods JT (1970) Cainozoic fossils of Queensland. The Queensland Palaeontographical Society, Brisbane. The Wilke Group, Brisbane, Australia

Holland BS (2001) Invasion without a bottleneck: Microsatellite variation in natural and invasive populations of the brown mussel Perna perna (L). *Marine Biotechnology* 3: 407-15

Hutchins JB (1991) Dispersal of tropical fishes to temperate seas in the southern hemisphere. *Journal of the Royal Society of Western Australia* 74: 79-84

Hutchins JB (2001) Biodiversity of shallow reef fish assemblages in Western Australia using a rapid censusing technique. *Records of the Western Australian Museum* 20: 247-70

Hyland SJ, Hill BJ and Lee CP (1984) Movement within and between different habitats by the portunid crab *Scylla serrata*. *Marine Biology* 80: 57-61

Johnson MS, Holborn K and Black R (1993) Fine scale patchiness and genetic heterogeneity of recruits of the corallivorous gastropod *Drupella cornus*. *Marine Biology* 117: 91-93

Jollivet D, Chevaldonne P and Planque B (1999) Hydrothermal-vent alvinellid polychaete dispersal in the eastern Pacific.2. A metapopulation model based on habitat shifts. *Evolution* 53: 1128-1142

Karasawa H and Tanaka T (1994) Decapod Crustacea from the Atsumi Group (middle Pleistocene) of Aichi Prefecture, Central Japan. Scientific Reports of the Toyohashi Museum of Natural History 4: 11-19

Karl SA, Bowen BW and Avise JC (1992) Global population genetic structure and male – mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analysis of anonymous nuclear loci. *Genetics* 131: 163 - 173

Kay EA and Palumbi SR (1987) Endemism and evolution in Hawaiian marine invertebrates. *Trends in Ecology and Evolution* 2: 183-186

Keenan CP (1994) Recent evolution of population structure in Australian barramundi, *Lates calcarifer* (Bloch): An example of isolation by distance in one dimension. *Australian Journal of Marine and Freshwater Research* 45: 1123-48

Keenan CP (1999) Aquaculture of the Mud Crab, Genus *Scylla* – Past, present and future. In: Keenan CP, Blackshaw A (eds) *Mud crab aquaculture and biology. Proceedings of an international scientific forum held in Darwin, Australia, 21-24 April 1997*. ACAIR Proceedings No. 78, Ferguson and Co., Brisbane Australia, pp. 9-13

Keenan CP, Davie PJF and Mann DL (1998) A revision of the genus *Scylla* De Haan, 1833 (Crustacea: Decapoda: Brachyura: Portunidae). *Raffles Bulletin of Zoology* 46: 217-45

Kench PS (1999) Geomorphology of Australian estuaries: Review and Prospect. *Australian Journal of Ecology* 24: 367-80

Kimura M and Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49: 561-576

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX and Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Science, USA* 86: 6196-6200

Koehn RK, Newell RI and Immerman F (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. *Proceedings of the National Academy of Science USA* 77: 5385-5389

Kumar S, Tamura K and Nei G (1994) MEGA: molecular evolutionary genetic analysis software for microcomputers. *Cabios* 10: 189-191

Laseron CF (1956) The Families Rissoidae (Mollusca) from the Solanderian and Damperian provinces. *Australian Journal of Marine and Freshwater Resources* 7: 384-485

Lavery S, Moritz C and Fielder DR (1995) Changing patterns of population structure and gene flow at different spatial scales in the coconut crab (*Birgus latro*). *Heredity* 74: 531-541

Lavery S, Moritz C and Fielder DR (1996-a) Indo-Pacific population structure and evolutionary history of the coconut crab *Birgus latro*. *Molecular ecology* 5: 557-570

Lavery S, Moritz C and Fielder DR (1996-b) Genetic patterns suggest exponential population growth in a declining species. *Molecular Biology and Evolution* 13: 1106-1113

Lea DW, Pak DK and Spero HJ (2000) Climate impact of late Quaternary equatorial Pacific sea surface temperature variations. *Science* 289: 1719-1724

Lessa EP and Applebaum G (1993) Screening techniques for detecting allelic variation in DNA sequences. *Molecular Ecology* 2: 119-129

Lessios HA, Kessing BD, and Pearse JS (2001) Population structure and speciation in tropical seas: Global phylogeography of the sea urchin *Diadema*. *Evolution* 55: 955-975

Levin DA (1988) Consequences of stochastic elements in plant migration. *American Naturalist* 132: 643 –651

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological society of America* 15: 237-40

Levins R (1970) Extinction. In: Gerstenhaber M (ed) *Some Mathematical Questions in Biology.* Lectures on Mathematics in Life Sciences Vol. 2, American Mathematical Society, Providence, Rhode Island, pp.77-107

Luikart G and Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology* 12: 228-237

Lynch M and Crease TJ (1990) The analysis of population survey data on DNA sequence variations. *Molecular Biology and Evolution* 7: 377-394

MacArthur RH and Wilson EO (1967) The theory of Island Biogeography. Princeton University Press, Princeton NJ

McConaugha JR (1992) Decapod larvae: dispersal, mortality, and ecology. A working hypothesis. *American Zoologist* 32: 512-523

McCoy ED and Heck KLJr (1976) Biogeography of corals, seagrasses, and mangroves: an alternative to the centre of origin concept. *Systematic Zoology* 25: 201-210

McGowran B, Qianyu L, Cann J, Padley D, McKirdy DM and Shafik S (1997) Biogeographic impact of the Leeuwin current in southern Australia since the late middle Eocene. *Paleogeography, Paleoclimatology, Paleoecology* 136: 19-40

McManus JW (1985) Marine speciation, tectonics and sea-level changes in southeast Asia. *Proceedings of the 5th International Coral Reef Congress, Tahiti* 4: 133-138

McMillan WO and Palumbi SR (1995) Concordant evolutionary patterns among Indo - West Pacific Butterflyfishes. *Proceedings of the National Academy of Science, USA* 260: 229-236

Manikiam JS (1967) The occurrence of the mud crab, *Scylla serrata* in New Zealand. *New Zealand Marine Department Fisheries Technical Report* 24

Marsh LM and Marshall JI (1983) Some aspects of the zoogeography of northwestern Australian echinoderms (other than Holothurians). *Bulletin of Marine Science* 33: 671-87

Maruyama T and Birky CW (1991) Effects of periodic selection on gene diversity of organelle genomes and other systems without recombination. *Genetics* 127: 449-451

Maudet C, Miller C, Bassano B, Breitenmoser-Wursten C, Gauthier D, Obexer-Ruff G, Michallet J, Taberlet P, Luikart G (2002) Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex* (*ibex*)]. *Molecular Ecology* 11: 421-436

Mayr E (1963) Animal species and evolution. Belknap Press, Harvard University, Cambridge

Monteiro A and Pierce NA (2001) Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1α gene sequences. *Molecular Phylogenetics and Evolution* 18: 264-281

Moritz C and Lavery S (1996) Molecular ecology: Contributions from molecular genetics to population ecology. In: Sheppard AW and De Barro PJ (eds), *Frontiers of Population Ecology*. CSIRO Publishing, Melbourne, pp 433-450

Morrow J, Scott L, Congdon B, Yeates D, Frommer M and Sved J (2000) Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution* 54: 899 -910

Nauta M, and Weissing F (1996) Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics* 143: 1021-1032

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-90

Nei M (1987) Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nei M, Maruyama T and Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* 29: 1-10

Neigel JE and Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Nevo E, Karlin S (eds), *Evolutionary Process and Theory*. Academic Press, New York, pp 515-534

Neigel JE, Ball RM and Avise JC (1991) Estimation of single generation migration distances from geographic variation in animal mitochondrial DNA. *Evolution* 45: 423-432

Nichols RA and Hewitt GM (1994) The genetic consequences of long distance dispersal during colonization. *Heredity* 72: 312-317

Ong KS (1966) Observations of the post –larval life history of *Scylla serrata* (Forskål), reared in the laboratory. *Malaysian Agricultural Journal* 45: 429-443

Orti G, Pearse DE, and Avise JC (1997) Phylogenetic assessment of length variation at a microsatellite locus. *Proceedings of the National Academy of Sciences USA* 94 10745-10749

Paetkau D, Calvert W, Stirling I and Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* 4: 347 - 354

Palmer AR and Strathman RR (1981) Scale of dispersal in varying environments and its implications for life histories of marine invertebrates. *Oecologia* 43: 308-318

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics* 25: 547-572

Palumbi SR (1995) Using genetics as an indirect estimator of larval dispersal. In: McEdward L (ed), *Ecology of marine Invertebrate larvae*. CRC Press, New York, pp 369-87

Palumbi SR, Cipriaono F and Hare MP (2001) Predicting nuclear gene coalescence from mitochondrial data: The three times rule. *Evolution* 55: 859-868

Palumbi SR, Grabowsky G, Duda T, Geyer L and Tachino N (1997) Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution* 51: 1506-1517

Palumbi SR and Kessing B (1991) Population biology of the Trans-arctic exchange: MtDNA sequence similarity between Pacific and Atlantic sea urchins. *Evolution* 45: 1790-1805

Palumbi SR, Martin A, Romano S, McMillan WO, Stice L and Grabowski G (1991) The simple fool's guide to PCR, version 2.0. Department of Zoology, University of Hawaii, Honolulu

Palumbi SR and Wilson AC (1990) Mitochondrial DNA diversity in the sea urchins *Strongylocentrotus purpuratus* and *S. droebachiensis*. *Evolution* 44: 403-415

Pielou EC (1979) Biogeography. Wiley, New York

Piry S, Luikart G, and Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* 90: 502-503; (http://www.ensam.inra.fr/URLB)

Planes S and Fauvelot C (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution* 56: 378-399

Queney G, Ferrand N, Weiss S, Mougel F, and Monnerot M (2001) Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (Oryctolagus cuniculus). *Molecular Biology and Evolution* 18: 2169-2178

Rand DM, Dorfsman M, and Kann LM (1994) Neutral and non-neutral evolution of *Drosophila* mitochondrial DNA. *Genetics* 138: 741-756

Rassmann K, Schlötterer C, Tautz D (1991) Isolation of simple – sequence loci for use in polymerase chain reaction – based DNA fingerprinting. *Electrophoresis* 12: 113-118

Raymond M and Rousset F (1995) Genepop (Version 1.2) – Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248 -249

Rice WR (1989) Analyzing tables of statistical test. *Evolution* 43: 223-225

Robertson WD (1996) Abundance, population structure and size at maturity of *Scylla serrata* (Forskål) (Decapoda: Portunidae) in Eastern Cape estuaries, South Africa. *South African Journal of Zoology* 31: 177-85

Roehrdanz RL (1993) An improved primer for PCR amplification of mitochondrial amplification in a variety of insect species. *Insect Molecular Biology* 2: 89-91

Rosen BR (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology* 27: 159-188

Rosen BR (1984) Reef coral Biogeography and climate change through the late Cainozoic: just islands in the sun or a critical pattern of islands. In: Brenchley P (ed), *Fossils and Climate*. John Wiley and Sons, Chichester, UK, pp 201-260

Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406-25

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning, A Laboratory Manual, 2nd edition Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sastry AN (1983) Pelagic larval ecology and Development. In: Vernberg FJ, WJ (eds), *The Biology of the Crustacea*. Vol.7, Academic press, New York, pp 213-282

Scheltema RS (1971) Larval dispersal as a means of genetic exchange between geographically separated populations of shallow - water invertebrates. *Biological Bulletin of Marine Biological Laboratories* 140: 284-322

Schneider S, Roessli D and Excoffier L (2000) ARLEQUIN: *A software for population Genetics Data Analysis* (version 2.0). Dept. of Anthropology, Univ. of Geneva, Geneva.

Shaw PW, Turan C, Wright JM, O'Connell, and Carvalho GR (1999) Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity* 83: 490-499

Shephard JM, Hughes JM and Zalucki MP (2002) Genetic differentiation between Australian and North American populations of the monarch butterfly *Danaus plexippus* (L.) (Lepidoptera: Nymphalidae): an exploration using allozyme electrophoresis. *Biological Journal of the Linnean Society* 75: 437-52

Shulman MJ and Bermingham E (1995) Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution* 49: 897-910

Simon C, Frati F, Bechenbach A, Crespi B, Liu H and Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87: 651-701

Slade RW, Moritz C, Hoelzel AR and Burton HR (1998) Molecular population genetics of the southern elephant seal *Mirounga leonina*. *Genetics* 149: 1945 –1957

Slatkin M (1987) Gene Flow and the geographic structure of natural populations. *Science* 236: 787-792

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47: 264-279

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462

Slatkin M and Barton NH (1989) A comparison of three methods for estimating average levels of gene flow. *Evolution* 43: 1349-1368

Sorrenson MD and Fleischer RC (1996) Multiple independant transpositions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Science, USA* 93: 15239-15243

Stephenson W and Campbell B (1960) The Australian Portunids (Crustacea: Portunidae). IV. Remaining genera. *Australian Journal of Marine and Freshwater Research* 11: 73-122, plates 1-5

Stephenson W and Hudson JJ (1957) The Australian portunids (Crustacea; Portunidae). I. The genus *Thalamita*. *Australian Journal of Marine and Freshwater Research* 8(3): 312-68

Strathmann RR, Hughes TR, Kuris AM, Lindeman KC, Morgan SG, Pandolphi JM, and Warner RR (2002) Evolution of local recruitment and its consequences for marine populations. *Bulletin of Marine Science* 70 (1) Suppliment S: 377-396

Strimmer K and Von Haeseler A (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution* 13: 964-69

Sugama K and Hutapea JH (1999) Genetic characterisation in the mud crab *Scylla* (Brachyura: Portunidae). In: Keenan CP, Blackshaw A (eds), *Mud crab aquaculture and biology. Proceedings of an international scientific forum held in Darwin, Australia, 21-24 April 1997*. ACAIR Proceedings No. 78, Ferguson and Co., Brisbane, Australia, pp 43-47

Sulkin SD (1984) Behavioural basis of depth regulation in the larvae of brachyuran crabs. *Marine Ecological Progress Series* 15: 181-205

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595

Tajima F and Nei M (1984) Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution* 1: 269-285

Tamura K and Aotsuka T (1988) Rapid isolation method of animal mitochondrial DNA by the alkaline lysis procedure. *Biochemical Genetics* 26: 815-819

Taylor ML (1984) New species of mud crab found in Western Australia. *FINS* 17(2): 15-18

Templeton AR (1998) Nested clade analyses of phylogeographic data: Testing hypotheses about gene flow and population history. *Molecular Ecology* 7: 381-397

Underwood AJ and Fairweather PG (1989) Supply - side ecology and benthic marine assemblages. *Trends in Ecology and Evolution* 4: 16-20

Veron JEN (1995) Corals in space and time: biogeography and evolution of the Scleractinia. Southwood Press, Australia

Villablanca FX, Roderick GK and Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology* 7: 547-60

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* 27: 1153-1167

Wade MJ, McKnight ML, and Shaffer HB (1994) The effects of kin – structured colonization on nuclear and cytoplasmic genetic diversity. *Evolution* 48: 1114 - 1120

Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution* 41: 385-400

Wartell RM, Hosseini SH and Moran CPJr. (1990) Detecting base pair substitutions in DNA fragments by temperature-gradient gel electrophoresis. *Nucleic Acids Research* 18: 2699-2705

Weir BS and Cockerham CC (1984) Estimating *F*- statistics for the analysis of population structure. *Evolution* 38: 1358 – 1370

Williams A, Koslow JA and Last PR (2001) Diversity, density and community structure of the demersal fish fauna of the continental slope off western Australia (20 to 35 degrees S). *Marine Ecology Progress Series* 212: 247-63

Williams ST and Benzie JAH (1993) Genetic consequences of long larval life in the starfish *Linckia laevigata* (Echinodermata: Asteroidea) on the Great Barrier Reef. *Marine Biology* 117: 71-7

Williams ST and Benzie JAH (1997) Indo-west Pacific patterns of genetic differentiation in the high-dispersal starfish *Linckia laevigata*. *Molecular Ecology* 6: 559-73

Williams ST and Benzie AH (1998) Evidence of a biogeographic break between populations of a high dispersal starfish: Congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. *Evolution* 52: 87-99

Williams ST, Knowlton N, Weight LA and Jara JA (2001b) Evidence for three major clades within the snapping shrimp genus *Alpheus* inferred from nuclear and mitochondrial gene sequence data. *Molecular Phylogenetics and Evolution* 20: 375-389

Wilson BR and Allen GR (1987) Major components and distribution of marine fauna. In: Dyne GR, Walton DW (eds) *Fauna of Australia*. General Articles Vol 1A, Australian Govt. Publishing Service, Canberra, pp 43-68

Wilson SG, Taylor JG and Pearce AF (2001) The seasonal aggregation of whale sharks at Ningaloo Reef, Western Australia: currents, migrations and the El Nino/ Southern Oscillation. *Environmental Biology of Fishes* 61: 1-11

Wirth T and Bernatchez L (2001) Genetic evidence against panmixia in the European eel. *Nature* 409: 1037-1040

Wolanski E (1993) Water circulation in the Gulf of Carpentaria. *Journal of Marine Systems* 4: 401-20

Wolanski E and Pickard GL (1985) Long - term observations of currents on the Central Great Barrier Reef continental shelf. *Coral Reefs* 4: 47-57

Woodland DJ (1983) Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness. *Bulletin of Marine Science* 33: 713-717

Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97-159

Wright S (1943) Isolation by distance. *Genetics* 28: 114-138

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19: 395-420

Wright S (1978) Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago

Zhang DX and Hewitt GM (1998) Isolation of DNA from preserved specimens. In: Karp A, Isaac PG, Ingram DS (eds) *Molecular tools for screening biodiversity: Plants and Animals*. Chapman and Hall, London

Appendix

Estimates of F_{ST} (lower matrix) and R_{ST} (upper matrix) between paired locations for five loci (as per tables a-e). Location abbreviations: EX, Exmouth; BM, Broome; RR, Roper River; KA, Karumba; WE, Weipa; HB, Hinchinbrook; MY, Mackay; MB, Moreton Bay. Exact tests of allelic homogeneity between paired locations obtained by permutation tests (10,000 replicates). Significant deviation from homogeneity * P < 0.05, and ** P < 0.005 as indicated in the F_{ST} matrix.

a) Locus Ss-101

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		-0.053	-0.018	-0.031	0.012	-0.015	-0.018	-0.022
BM	0.001		-0.052	-0.127	-0.060	-0.092	-0.057	-0.075
RR	-0.002	-0.013		-0.022	0.034	0.001	-0.016	-0.023
KA	0.004	-0.010	-0.003		-0.024	-0.045	-0.027	-0.029
WE	-0.011	-0.006	-0.004	0.005		-0.025	0.014	0.023
HB	-0.003	-0.006	-0.003	0.004	-0.003		-0.009	-0.004
MY	-0.001	0.016	0.007	0.015*	0.001	0.006		-0.018
MB	-0.008	-0.003	-0.004	0.002	-0.007	0.007	0.005	

b) Locus Ss-103

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		0.056	0.038	-0.047	0.028	-0.010	0.052	-0.014
BM	0.042		-0.001	0.022	-0.059	0.065	-0.043	0.005
RR	0.003	0.025		-0.016	-0.009	0.158	-0.020	-0.009
KA	-0.022	0.039	-0.008		0.033	0.029	-0.007	-0.045
WE	0.017	-0.033	0.005	0.015*		0.072	-0.017	0.024
HB	0.002	0.041	-0.003	0.007	0.017*		0.019	0.012
MY	0.006	0.011	-0.010	0.003	0.007	0.007		-0.001
MB	0.007	-0.005	-0.002	0.012	-0.008	0.006	0.001	

c) Locus Ss-112

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		0.006	-0.001	-0.035	-0.022	-0.021	0.002	-0.023
BM	0.011		-0.053	-0.047	-0.058	-0.010	-0.006	0.011
RR	0.003	-0.001		-0.036	-0.028	-0.021	-0.019	-0.004
KA	-0.003	-0.020	-0.009		-0.049	-0.045	-0.028	-0.033
WE	0.007	0.009	0.006	-0.003		-0.035	-0.025	-0.021
HB	0.008	-0.028	-0.003	-0.004	-0.006		-0.017	-0.021
MY	0.004	-0.006	0.008*	-0.005	-0.001	-0.002		-0.003
MB	0.005	-0.026	0.005	-0.011	0.004	-0.007	0.001	

Appendix (cont.)

d) Locus Ss-403

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		-0.095	0.053	-0.024	-0.016	-0.017	-0.008	-0.019
BM	0.004		0.010	-0.081	-0.049	-0.065	-0.048	-0.061
RR	0.008	-0.002		-0.012	0.120	0.134	0.160	0.103
KA	0.058**	0.041	0.027		0.025	0.030	0.050	0.013
WE	0.011	0.028	0.009	0.022		-0.019	-0.014	-0.014
HB	0.007	0.024	0.010	0.033*	-0.013		-0.022	-0.016
MY	0.011	0.007	0.012*	0.025*	-0.004	-0.004		-0.011
MB	0.012*	-0.004	0.013*	0.022*	-0.001	-0.004	-0.001	

e) Locus Ss-513

	R_{ST}							
F_{ST}	EX	BM	RR	KA	WE	HB	MY	MB
EX		0.108	-0.020	0.094	0.028	0.050	-0.018	-0.002
BM	0.058		0.165	-0.043	0.013	0.006	0.049	0.096
RR	0.019	0.144		0.117	0.017	0.043	-0.018	-0.010
KA	0.037	0.021	0.023		-0.005	-0.016	0.035	0.058
WE	0.049*	0.009	0.068*	-0.017		-0.016	0.001	0.001
HB	0.034	-0.001	0.070*	-0.013	-0.012		0.011	0.015
MY	0.023	0.006	0.050*	-0.012	-0.009	-0.011		-0.013
MB	-0.008	0.042	0.038*	0.017	0.021*	0.007	0.010*	

Published papers arising from this thesis

Gopurenko D, Hughes JM & Ma J (2002) Identification of polymorphic microsatellite loci in the mud crab *Scylla serrata* (Brachyura: Portunidae). *Molecular Ecology Notes* 2(4): 481-3.

Gopurenko D & Hughes JM (2002) Regional patterns of genetic structure among Australian populations of the mud crab, *Scylla serrata* (Crustacea: Decapoda): evidence from mitochondrial DNA. *Marine and Freshwater Research* 53: 849-857.

Gopurenko D, Hughes JM, Keenan CP (1999) Mitochondrial DNA evidence for rapid colonisation of the Indo-West Pacific by the mud crab *Scylla serrata*. *Marine Biology* 134: 227-233.