Chapter 1: General Introduction

The genetic structure of populations is determined by complex interactions among many genetic, ecological and evolutionary processes (Hartl and Clark 1997; Avise 2000). Ecological and demographic factors including population size, generation time, reproductive behaviour and patterns of migration among suitable habitats may affect the distribution of genetic variation within and among populations. Natural selection may shape allele frequencies in response to local conditions and genetic drift may be a powerful evolutionary force in isolated populations (Slatkin 1985). Genetic factors such as variation in mutation rates and recombination may further affect the genetic composition of populations (Hartl and Clark 1997). An understanding of the processes responsible for the genetic structure of populations therefore requires a detailed appreciation of how these factors interact, and how they vary spatially and temporally.

The physical characteristics of the marine environment and the biological attributes of marine species present a number of evolutionary paradoxes to geneticists seeking to understand the processes determining population structure and speciation in the sea (Knowlton 1993; Palumbi 1994; Grosberg and Cunningham 2001). The marine environment is fluid and there is a general absence of physical barriers to dispersal (Vermeij 1987). Despite this, many marine communities such as those inhabiting coral reefs are very speciose (Sale 1991) and many widespread species comprise groups of cryptic species (Knowlton 1993). Fishes on coral reefs live in a naturally fragmented ecosystem where the isolation of suitable habitat patches may facilitate genetic isolation of populations. However, the pelagic larval stage exhibited by many of these species (Leis 1991; Leis and Carson-Ewart 2000) allows them to disperse widely and may homogenise population genetic structures (Palumbi 1994). Coral reef fishes are generally characterised by large, local populations and have high reproductive outputs, which may decrease the importance of population bottlenecks and genetic drift associated with founder effects in isolated populations (Birky et al. 1989; Hellberg et al. 2002; Kritzer and Sale 2004). However, coral reefs have a dynamic evolutionary history where Pleistocene sea level fluctuations greatly affected the distribution of habitats as well the size and connectivity of populations (Benzie 1999). Therefore, the population genetic structure of many species may be affected by historical isolation, genetic bottlenecks and founder events associated with the colonisation of new habitats. Present

patterns of genetic variation within and among populations of many coral reef organisms is therefore likely to be the result of complex interactions among historical and present day factors (Grant and Bowen 1998; Benzie 1999).

Estimates of genetic variability within and among marine populations are rapidly accumulating in the literature, but generalisations about which species are likely to have genetically structured populations, and the processes driving such differentiation, are still hard to draw (Planes 2002). Many studies have used vast trans-oceanographic sampling strategies of species with long pelagic larval durations and presumably high dispersal potentials (e.g., Pacific: Planes and Fauvelot 2002; Bay et al. 2004; Caribbean: Taylor and Hellberg 2003; Geertje et al. 2004; Atlantic: Muss et al. 2001; Rocha et al. 2002). In general, these studies have detected significant genetic structuring among regions (but see Geertje et al. 2004) although some gene flow commonly occurs (but see Taylor and Hellberg 2003). Consequently, a positive relationship between genetic differentiation and geographical distance (isolation-by-distance) has often been found (e.g., Bay et al. 2004). Investigation of population genetic structure of marine species at smaller spatial scales have tended to investigate species with either a short or no pelagic larval duration and hence, low dispersal potentials (Doherty et al. 1994; Nelson et al. 2000; Planes et al. 2001; Bernardi and Vagelli 2004; Hoffman et al. 2005), or species with longer larval durations (Doherty et al. 1995; Planes et al. 1996; Planes et al. 1998; Dudgeon et al. 2000; Bernardi et al. 2001; Messmer et al. 2005). Investigations of species with a short or no pelagic larval stage have generally identified strong population genetic structure over quite short distances (e.g., 5 – 10 km) implying that spatial isolation and genetic drift are important ecological and evolutionary factors in such species. In contrast, species with longer larval durations may (Doherty et al. 1995; Planes et al. 1996; Planes et al. 1998; Messmer et al. 2005) or may not (Doherty et al. 1995; Dudgeon et al. 2000; Bernardi et al. 2001) display strong local structure suggesting that the factors that influence population genetic structure at local scales may also be complex.

The processes that determine the population genetic structure of coral reef fishes have also been investigated in a comparative framework (Doherty et al. 1995; Shulman and Bermingham 1995; Planes et al. 1998; Dudgeon et al. 2000; Riginos and Victor 2001; Fauvelot and Planes 2002; Fauvelot et al. 2003; Rocha et al. 2005). Such comparative studies may be particularly important in elucidating the mechanisms determining genetic structure in marine species because they allow for factors affecting

dispersal to be isolated and controlled (Bohonak 1999). Indeed, some of the more general conclusions about the roles of ecological specialisation (Rocha et al. 2005), larval behaviour (Riginos and Victor 2001) and historical habitat stability (Fauvelot et al. 2003) on the population genetic structure of coral reef fishes have emerged from such studies.

While all these studies have made valuable contributions to our understanding of the population genetic structure of coral reef fishes, very few have allowed for genetic variation to be partitioned among local scales within regions of any species, regardless of their presumed dispersal potential (but see Doherty et al. 1995). Consequently, we do not currently have a good appreciation of local-scale genetic patterns and their potential effects on regional patterns. Studies that allow genetic variation to be partitioned among local and regional scales, as well as controlled comparative studies among closely related species, have the potential to greatly increase our understanding of the general mechanisms that determine the population genetic structure of coral reef fishes. The general aim of this thesis was, therefore, to understand how the spatial and temporal complexity of coral reefs can influence the genetic structure of species occupying such environments. To this end, I examined the spatial genetic structure of one species in detail to elucidate the potential for local and regional scale variation in its population genetic structure. I then used a comparative approach to examine the roles of dispersal potential on gene flow, and how the population genetic structure can differ within the species range. In order to achieve these aims it was necessary to consider the major biological, historical and environmental factors that may influence the interpretation of population genetic data of coral reef fishes.

The genetic structure of populations is interpreted using several spatial models that vary in their complexity and the assumptions they make about the biological characteristics of the system under investigation. The complexity of these models and the degree to which assumption may be violated can greatly affect the resolution and interpretation of population genetic data. The island model, originally proposed by Sewall Wright (1931), estimates genetic differentiation (F_{ST}) by assuming that all populations are of equal size and have an equal probability of exchanging migrants regardless of their relative positions. The spatial position and the size of populations are likely to affect both patterns of emigration (i.e., larger populations may produce more emigrants) and the local effects of immigration (i.e., the effects of immigration may be greater in a smaller compared to larger population). Consequently, this model does not

describe the spatial structure of many real populations very well, except when the migration rate is low (Palumbi et al. 2003). This is because the migration rate is inversely related to the log of genetic structure (F_{ST}), so that even moderate migration rates will produce very small F_{ST} estimates associated with relatively high error (Waples 1998). Differences in F_{ST} estimates among populations or species are therefore difficult to distinguish statistically, even where the migration rates producing them are different (Neigel 1997; Waples 1998). Most genetic investigations on coral reef fishes to date have used this island model.

The stepping-stone, or isolation-by-distance model incorporates spatial variation by assuming that populations in closer proximity are more likely to exchange migrants than more distantly separated ones (Wright 1943; Kimura 1955; Kimura and Weiss 1964; Weiss and Kimura 1964). Migration rates can be estimated by correlating the genetic differentiation of populations with the geographical distance separating them, and therefore, allow migration rates to be estimated more precisely than under the island model, especially when migration rates are high (Palumbi et al. 2003). The isolation-by-distance model has increasingly been applied to the population genetic structure of coral reef fishes (e.g., Planes et al. 1996; Planes and Fauvelot 2002; Bay et al. 2004) and has indicated that patterns of gene flow may differ among species with high dispersal potential.

Metapopulation genetic models consider differences in effective population sizes, colonisation patterns and extinction rates in the interpretation of migration and concomitant genetic structure of populations (Slatkin 1977, 1985, 1987; Wade and McCauley 1988; Whitlock and McCauley 1990). A metapopulation is composed of a number of spatially structured ephemeral populations that interact and persist through time via migration (Hanski 1991; Hanski and Gilpin 1997). The patterns of migration, extinction and re-colonisation of these populations can have profound effects on the distribution of genetic variation within and among the populations of a metapopulation (Pannell and Charlesworth 1999, 2000; Pannell 2003). For example, extinctions may decrease genetic variation within local populations but increase genetic structure among populations depending on the pattern of colonisation (Wade and McCauley 1988; Whitlock and McCauley 1990). Migration may reduce genetic differentiation among populations over time so that younger populations display stronger genetic differentiation than older populations (Giles and Goudet 1997). The physical structure of coral reefs suggests that the application of metapopulation theory holds much

promise for understanding the spatial genetic structure of many coral reef species (Swearer et al. 2002). Despite this, we currently have a poor appreciation of the presence, spatial extent and genetic consequences of metapopulation dynamics in marine systems.

The vast majority of coral reef fishes have a bipartite life history where dispersal occurs primarily during the pelagic larval phase (Sale et al. 1980; Leis 1991; Leis and Carson-Ewart 2000). As such, most species have a potential for large-distance dispersal and characteristics of the larval phase have commonly been used to predict the genetic structure of populations. The dispersal potential of coral reef fishes have been investigated with respect to a range of larval traits including egg type (Shulman and Bermingham 1995; Shulman 1998), pelagic larval environment (Riginos and Victor 2001) and most commonly the length of the pelagic larval phase (PLD: Waples 1987; Doherty et al. 1995; Shulman and Bermingham 1995; Riginos and Victor 2001). When examined previously, a relationship between mean larval duration and genetic differentiation has generally been found (Waples 1987; Doherty et al. 1995; Shulman and Bermingham 1995; Riginos and Victor 2001). This relationship may be greatly influenced by the inclusion of highly genetically structured, directly developing species (Bohonak 1999; Riginos and Victor 2001). Furthermore, behavioural (e.g., Taylor and Hellberg 2003), physiological (e.g., Shulman 1998) and ecological (e.g., Rocha et al. 2005) factors may also affect dispersal abilities, and these may vary among taxonomic groups (Bohonak 1999). Despite the potential importance of such characteristics in determining variation in dispersal rates, examinations of the relationship between PLD and gene flow in marine fishes to date have incorporated a range of distantly related species, displaying different spawning characteristics and adult ecologies (Waples 1987; Doherty et al. 1995; Shulman and Bermingham 1995; Riginos and Victor 2001). Consequently, we do not have a good understanding of how PLD relates to dispersal in species that display little variation in their biology and ecology.

Under the neutral theory of molecular evolution, genetic variation in unlinked markers is generated by mutation, subsequently modified by random genetic drift in isolated populations and homogenised among populations via migration (Hartl and Clark 1997). An explicit assumption of many population genetic analyses is, therefore, that populations are in migration-drift equilibrium; their current genetic structure is the result of the opposing effects of genetic drift and migration (Hartl and Clark 1997). The distribution of genetic variation within and among populations can be greatly affected

by historical effects such as genetic bottlenecks and founder events (Avise 2000) and such signatures may be retained in populations over many generations depending on their effective population sizes and rates of migration (Crow and Aoki 1984). Consequently, many species may not be in migration-drift equilibrium and their genetic structure may reflect historical as well as current patterns of gene flow (Benzie 1999). Under this scenario, the use of F_{ST} as an indicator of genetic isolation becomes problematic (Neigel 1997, 2002); recent coalescence-based maximum likelihood methods (Kuhner et al. 1998; Beerli and Felsenstein 1999, 2001) may ameliorate some of the problems associated with estimating population genetic structure in non-equilibrium species (Neigel 2002).

Sea level fluctuations associated with Pleistocene glacial events greatly affected the presence and distribution of coral reefs and are likely to have had profound effects on resident faunas (Paulay 1990; Benzie 1999). The historical effects on the present day patterns of genetic structure and genetic variability in coral reef fishes have increasingly been considered (Doherty et al. 1994; Shulman and Bermingham 1995; Dudgeon, 2000; Nelson et al. 2000; Planes et al. 2001; Fauvelot et al. 2003). Strong genetic differentiation among closely spaced populations has been interpreted in the context of historical isolation of populations (Nelson et al. 2000), or founder events associated with the colonisation of new habitats (Doherty et al. 1994; Planes et al. 2001). In species with low genetic structure, reduced genetic diversities have been associated with historical habitat stability (Fauvelot et al. 2003). While many coral reef fishes display low genetic variability indicating shallow coalescent histories (Grant and Bowen 1998), other species display high levels of genetic diversity (Planes 1998). Consequently, the current genetic structure of many coral reef fishes may be strongly influenced by historical events, however, the importance of such historical factors may vary among locations and species.

The factors determining the extent of a species geographic range have long interested biologists (Darwin 1859; Mayr 1963). The distributional range of a species is determined by spatial and temporal variation in demographic parameters such as births, deaths and dispersal (Holt et al. 2005). Biological and environmental conditions are generally assumed to be optimal in the centre of a species range and to decline towards its periphery (Hoffmann and Parsons 1991). Populations should therefore become smaller, more fragmented and experience increased extinction rates towards the edge of the range (Levins 1970; Lennon et al. 1997). These effects should be evident in the

genetic structure of such populations. For example, gene flow may be reduced towards the species margin because of increased fragmentation and smaller population sizes (Levins 1970). Genetic isolation, smaller population sizes and increased extinction rates should, therefore, reduce genetic diversities in peripheral populations (Holt 1987). Many coral reef fish species have borders that are not associated with any obvious habitat discontinuities or barriers to dispersal. They, therefore, constitute a good system to test predictions from species border theory, however, variation in the population genetic structure among central and peripheral populations have rarely been considered (but see Planes and Fauvelot 2002).

In this thesis I examine the processes that may determine the population genetic structure in one family of coral reef fishes, the Pomacentridae, on the Great Barrier Reef (GBR). I capitalise on the unique attributes of the coral reef fish assemblages and the physical structure of the coral reefs on the GBR. The GBR is unique among many coral reef systems of the world in being a largely linear band of highly interconnected, though spatially separated, reefs of relatively recent origin (Hopley and Thom 1983; Larcombe 2001). Environmental conditions change along latitudinal and longitudinal gradients and concomitant effects on fish species distribution, abundance and demographic patterns are evident (Russ 1984; Gust et al. 2001; Gust 2004). The large number of individual reefs facilitates a detailed examination of local-scale genetic structure of coral reef organisms and its latitudinal and longitudinal variation. There is an absence of obvious dispersal barriers on the GBR, and species currently occupying the GBR are likely to have been affected by recent sea level changes. It is, therefore likely that the population genetic structure of coral reef fishes on the GBR may be affected by this disturbance history (Doherty et al. 1994; Planes et al. 2001), but there is no *a priori* reason why this should have affected some species differently from others. The high species richness of coral reef fishes on the GBR also enables comparative investigations of closely related and co-occurring species to be undertaken. Such a design can allow potential confounding factors such as ecological specificity, spawning characteristics and distributional range effects to be controlled.

To examine the processes driving patterns of gene flow and genetic variability in coral reef fishes on the Great Barrier Reef, I addressed four specific issues:

1. The potential for local spatial genetic structure of a low dispersal species and the utility of metapopulation theory to describe the population genetic structure in this species

2. The role of local extinctions on metapopulation dynamics of a low dispersal species, and how this may vary spatially among regions located in the centre and on the peripheral of the distributional range.
3. The relationship between dispersal ability, gene flow and genetic diversity in ecologically generalised and widespread species.
4. How metapopulation processes affect patterns of gene flow and genetic diversities on the species margin in ecologically generalised species.

This thesis is constructed as a series of stand-alone, but conceptually interconnected publications. **Chapter 2** examines the genetic structure of a common direct developing coral reef fish, *Acanthochromis polyacanthus* within and among regions on the Great Barrier Reef using a mitochondrial sequence marker and three microsatellite loci. I examine patterns of gene flow and reciprocal migration rates (i.e., migration from a to b, and vice versa) within and among regions and evaluate the conformation to different spatial genetic models at local and regional scales. The role of genetic bottlenecks and founder effects in *A. polyacanthus* were examined among the same locations in **Chapter 3**. I used frequentist and Bayesian maximum likelihood analyses to evaluate the roles of local extinctions and founder events on genetic diversities at local and metapopulation levels. I further examined if there was a difference in extinction dynamics towards the distributional range edge of this species. In **Chapter 4,** I examine the potential for intraspecific variation in the pelagic larval duration (PLD) of twelve coral reef fish species. Point estimates of mean PLD from the literature are commonly employed in a variety of applications including the prediction of genetic differentiation among populations (e.g., Doherty et al. 1995; Shulman and Bermingham 1995). Because genetic structure is greatly affected by even low levels of migration (Wright 1943), it is possible that maximum rather than mean PLD may better predict gene flow and emerging evidence suggests that PLDs may vary considerably temporally and spatially within species (Leis 1991; Cowen and Sponaugle 1997). Consequently a characterisation of intraspecific variation in this trait was necessary before it could be used to predict gene flow here. In **Chapter 5,** I examine the relationship between dispersal potential (mean and maximum PLD) and gene flow in eight pomacentrid species. To control for potentially confounding factors, this examination was conducted using closely related species that display similar spawning behaviours and generalised ecologies. The potential role of demographic processes in determining species' borders

is examined in **Chapter 6**. Patterns of gene flow and genetic diversities were examined in three congeneric species pairs that displayed very similar biological and ecological attributes. Each species pair consisted of one species sampled at two locations in the centre of its range (central species) and another species sampled at a central and marginal location in its range (peripheral species). This design allowed for the genetic consequences of range margins to be elucidated.