Chapter 5: The relationship between population genetic structure and pelagic larval duration in coral reef fishes on the Great Barrier Reef.

Publication: Bay LK Crozier RH and Caley MJ (In Press) The relationship between population genetic structure and pelagic larval duration in coral reef fishes on the Great Barrier Reef. Marine Biology

Abstract

Pelagic larval duration (PLD) is a commonly used proxy for dispersal potential in coral reef fishes. Here I examine the relationship between PLD, genetic structure and genetic variability in geographically widespread and ecological generalist species from one coral reef fish family (Pomacentridae) that differ in mean larval duration by more than a month. Genetic structure was estimated in eight species using a mitochondrial molecular marker (control region) and in a sub-set of five species using nuclear molecular markers (ISSRs). Estimates of genetic differentiation were similar among species with pelagic larvae, but differed between molecular markers. The mtDNA indicated no structure while the ISSR indicated some structure between the sampling locations. I found a relationship between PLD and genetic structure using both markers. These relationships, however, were caused by a single species, *Acanthochromis polyacanthus*, which differed from all the other species examined here in lacking a larval phase. With this species excluded, there was no relationship between PLD and genetic structure using either marker despite a range of PLDs of more than 20 days. Genetic diversities were generally high in all species and did not differ significantly among species and locations. Nucleotide diversity and total heterozygosity were negatively related to maximum PLD but again these relationships were caused by *A. polyacanthus* and disappeared when this species was excluded. These genetic patterns are consistent with moderate gene flow among well-connected locations and indicate that at this phylogenetic level (i.e., within family) the duration of the pelagic larval phase is unrelated to patterns of genetic differentiation.

Introduction

Identifying patterns of connectivity among populations and the extent of self-recruitment is important to the study of the population biology of marine species (e.g., Jones et al. 1999; Swearer et al. 1999; Taylor and Hellberg 2003) and to the effective management and conservation of marine resources (Palumbi et al. 2003; Palumbi 2004). Studies using genetic markers have become vital in this context due to the difficulties associated with directly observing movement among populations in the marine environment (e.g., Hedgecock 1986; Bohonak 1999; Hellberg et al. 2002). Although more genetic estimates of connectivity among marine populations are becoming available, generalisations about which species are likely to have genetically structured populations, and the processes driving such differentiation, are still hard to draw (Bohonak 1999). Comparative studies may be particularly important in filling this gap in our current knowledge since they allow for factors affecting dispersal to be isolated and controlled (Bohonak 1999).

Many marine organisms are relatively sedentary as adults, dispersing primarily during the larval phase (Leis 1991; Bonhomme and Planes 2000). The dispersal potential of these species may be related to traits such as egg type (e.g., pelagic vs. benthic: Shulman and Bermingham 1995; Shulman 1998) and larval development (direct vs. pelagic: Hunt 1993; Hellberg 1996; Arndt and Smith 1998; Ayre and Hughes 2000), length of the pelagic larval phase (PLD: Waples 1987; Doherty et al. 1995; Shulman and Bermingham 1995; Riginos and Victor 2001) and pelagic larval environment (inshore vs. offshore: Riginos and Victor 2001). In previous reports, species with shorter larval durations have generally displayed stronger genetic differentiation than species with longer larval durations (e.g., Waples 1987; Doherty et al. 1995; Riginos and Victor 2001). However, many exceptions, such as genetically differentiated populations of species with extensive larval durations have also been documented (e.g., Planes et al. 1998; Barber et al. 2000; Taylor and Hellberg 2003). Consequently, the relationship between potential and realised dispersal remains unclear (Shulman 1998; Bohonak 1999) and further study is warranted.

Behavioural, physiological and ecological factors may affect dispersal abilities, and these may vary among taxonomic groups (Bohonak 1999). For example, larval swimming ability, a potentially important determinant of dispersal distance varies considerably among reef-fish families (Stobutzki 1998; Leis 2002; Fisher 2005). Consequently, the distances that particular species may disperse during pelagic larval

phases of the same length may differ among families with different swimming abilities. Likewise, the spawning mode and early developmental patterns of species may also affect larval dispersal (Bohonak 1999). Directly developing species generally display strong genetic structure across small geographical distances (Doherty et al. 1994; Bernardi 2000; Planes et al. 2001) and the inclusion of such species may greatly affect comparative investigations of PLD and gene flow (Bohonak 1999; Riginos and Victor 2001). Conversely, species that broadcast spawn pelagic eggs may display lower genetic structure, but it is currently unclear whether this occurs because of the early developmental state at which they enter the plankton, the distribution of these larvae in the water column or because of the longer larval duration generally associated with this suite of life history traits (Leis 1991; Victor 1991; Shulman 1998). Because examinations of the relationship between PLD and genetic structure in marine fishes to date have incorporated a range of distantly related species, displaying different spawning characteristics and adult ecologies (e.g., Waples 1987 (10 species, 9 families); Doherty et al. 1995 (7 species, 3 families); Shulman and Bermingham 1995 (8 species, 6 families); but see Riginos and Victor 2001 (3 species, 1 family)) current understanding of the relationship between genetic structure and pelagic larval duration in species that display little variation in their biology and ecology is poor.

Pelagic larval durations within species may vary greatly both spatially and temporally (Chapter 4). For example, the PLD of *Amphiprion melanopus* was reported as 16.7 days ($\pm$SE = 1.5) in a sample of 3 individuals from Palau (Wellington and Victor 1989) but 10.96 days ($\pm$SE = 0.32) in a sample of 15 individuals from One Tree Island on the Great Barrier Reef (Chapter 4). At present it is unclear whether such intra-specific variation in PLDs may affect the relationship with genetic structure because only mean larval duration has previously been used to predict genetic structure (e.g., Doherty et al. 1995; Shulman and Bermingham 1995; Riginos and Victor 2001). Consequently, investigations that incorporate information about the mean pelagic larval duration and its variation have the potential to increase our understanding of the relationship between gene flow and larval duration further (Leis 1991; Victor 1991).

In order to minimise the potential effects of spawning strategy, adult ecology and phylogeny on any relationship between pelagic larval duration and gene flow, I examined the relationship between mean, minimum and maximum pelagic larval duration, genetic structure and genetic diversities among common, widespread and generalised species of the family Pomacentridae on the Great Barrier Reef (GBR). I

estimated population genetic structure between northern and southern locations of the Great Barrier Reef, Australia, in eight species using a rapidly evolving mtDNA sequence marker (control region) and, in a sub-set of five species, using nuclear genetic fingerprints (ISSRs). In this technique semi-arbitrary banding profiles with bands corresponding to a DNA sequence deliniated by two microsatellites are amplified by PCR (Zietkiewicz et al. 1994) and compared among individuals. I examined the relationship between mean, minimum and maximum pelagic larval duration and genetic differentiation. I then examined patterns of genetic variability (heterozygosity, haplotype and nucleotide diversity) among species and locations and their relationships to mean, minimum and maximum pelagic larval duration. Lastly I re-analysed data from Doherty et al (1995) to examine the potential roles of using different estimates of PLD on the relationship between PLD and genetic structure.

Materials and Methods

Specimen collection and estimation of pelagic larval duration

Eight species that span a broad range of pomacentrid genera were selected. These species differ in their pelagic larval duration, but share characteristics such as benthic spawning, a generalised ecology in terms of habitat use, high local abundances and wide geographical distributions (Table 1). Based on their relationship between PLD and gene flow, Doherty et al. (1995) suggested that the genetic structure of species with larval durations of less than 9 days are likely to be primarily influenced by drift, and hence, display strong genetic structure, whereas, the population genetic structure of species with larval durations between 9 and 39 days are likely to be influenced by both drift and migration and species with larval durations longer than 39 days are likely to display low genetic structure due to high migration rates and minimal importance of drift (Doherty et al. 1995). Because I was interested in genetic structure that was under the influence of both drift and migration, I investigated species displaying pelagic larval durations of less than 39 days. No pomacentrid species with a larval duration of less than 9 days, other than *A. polyacanthus* (Table 1), was available for inclusion in this study.

Approximately 25 individuals per species were collected from each of the same two locations: Northern GBR (Lizard Island, 14^{o}40S; 145^{o}28E) and southern GBR (One Tree Island 23^{o}30S; 152^{o}05E). These locations are separated by more than 1200 km with no obvious habitat discontinuities, or other hard barriers to dispersal. Collections were made during 2000 and 2001 and all individuals from each species were collected

in the same year. Fish were collected using hand spears, and clove oil and hand nets. Animals were placed in an ice slurry following capture, and then transported to shore where a tissue sample (fin clip preserved in 100% EtOH) was taken.

Table 1: Study species and their geographical, biological and ecological attributes. Markers used and sample sizes, latitudinal spread (Lat. spread), pelagic larval duration (PLD) (min – max), diet (P = planktivore, H = herbivore), reproductive mode (B = benthic spawning) and habitat use (L = live, D= dead).

Species	Sample sizes		Geography, Biology and Ecology				
	Control region	ISSR	Lat. spread[1]	PLD	Diet[1]	Rep.[1]	Habitat use[1]
Acanthochromis polyacanthus	41	46	40	0[2]	P	B	L/D coral
Amphiprion melanopus	42	46	40	11 (8 – 22)[3]	P	B	Anemones (3 spp.)
Pomacentrus moluccensis	46	48	55	15 (14 – 21) [3]	P/ H	B	L/D branching corals
Pomacentrus amboinensis	41		55	17 (15 – 32)[3]	P/ H	B	L/D coral, sand
Chromis atripectoralis	41	44	62	16 (10 – 24)[4]	P	B	L/D branching corals
Chrysiptera rex	46	48	57	18.2 (16 – 25)[3]	H	B	D coral
Amblyglyphidodon curacao	42		45	17 (15 – 22)[3]	P	B	L/D branching corals
Stegastes nigricans	39		60	28 (16 – 32) [5]	P/ H	B	L/D branching corals

Sources: [1]www.fishbase.org; [2]Kavanagh (2000); [3]Chapter 4; [4]Thorrold and Milicich (1990) and Murdoch (1995); [5]Thresher et al. (1989) and Wellington and Victor (1989).

All species were included in the mtDNA analysis and a sub-set of five of these species was included in the ISSR analysis (Table 1). In five species, published estimates of mean larval duration from counts of pre-settlement rings on otoliths of fishes from the GBR were used (Chapter 4) (Table 1). Estimates of mean larval duration were obtained for two additional species, *Chromis atripectoralis* and *Stegastes nigricans* by averaging published estimates from Thresher et al. (1989), Wellington and Victor (1989), Thorrold and Milicich (1990) and Murdoch (1995) (Table 1). Because point estimates of larval durations often display little variation (Chapter 4), the minimum and maximum larval durations used here were estimated by using the extremes of the range of pre-settlement ring counts published so far for each species (Table 1).

DNA extraction and amplification

Genomic DNA was extracted from approx 0.25 cm^2 of fin tissue (rehydrated by several TE washes) by a modified Phenol-Chloroform extraction procedure (Sambrook and Russell 2001, excluding the phenol-chloroform step) and resuspended in 50µl of TE. Concentrated DNA stock was diluted 1:50 yielding a final DNA concentration of approximately 5 ng/µL. A 400 base-pair region of the mitochondrial control region (hyper variable region I, HVR I) was amplified using the universal primers CR-A L15995 (5'-AATTCTCACCCCTAGCTCCCAAAG-3') and CR-E H16498 (5'-CCTGAAGTAGGAACCAGATG-3') (Lee et al. 1995). After a representative sample of all species had been sequenced, a specific forward primer was designed (dLoopF 5'-CATATATGTRTTATCAACATTA-3'), and this was used with CR-E in all further PCR and sequencing reactions. PCR reactions were carried out on a PE Applied Biosystems 9700 in 25µl containing 1x PCR Buffer (Promega), 3.5 mM $MgCl_2$, 200 µM each dNTP, 0.4 µM each primer, 10 ng template DNA and 0.1 unit of *Taq Polymerase* (Promega). Amplification using the polymerase chain reaction (PCR) was conducted with a cycling profile of 30 s at 94°C, 45 s at 48°C and 60 s at 72°C for 30 cycles. The cycling profile was flanked by an initial 2 min denaturing step (94°C) and a 10 min terminal extension phase (72°C). Using this procedure, only a single product ranging in size from 335 – 555 base pairs was amplified in most species, although two products (approximately 550 and 350 base pairs) were occasionally amplified in *Pomacentrus amboinensis*. Analysis of these fragments in *P. amboinensis* confirmed that the presence of the smaller fragment was due to a repeat in the t-RNA Pro end of

the Control region as previously described for a pomacentrid species (Bernardi et al. 2002). Only the larger fragment was sequenced. PCR products were cleaned up using PCR clean up columns (Qiagen) and resuspended in 20μL of ddH2O. Two μL of the cleaned product was sequenced in the forward and reverse direction using a dyenamic ET dye terminator kit (Megabase) chemistry (Amersham Biosciences). Sequence products were cleaned using sephadex G-50 columns. Labelled extension products were sequenced on a Megabase 1000 (Amersham Biosciences) at the Genetic Analysis Facility of James Cook University. Representative sequences have been deposited in GenBank under accession numbers DQ199708 – DQ199726, DQ199879 – DQ199899, DQ212199 – DQ212495.

Genetic fingerprints were obtained from 21 – 24 individuals per location using Inter Simple Sequence Repeats (ISSR). In this technique, semi-arbitrary banding profiles (where each banding position corresponds to a DNA sequence delimited by two microsatellites) are amplified by PCR (Zietkiewicz et al. 1994; Bornet and Branchard 2001). Because the PCR reaction is primed by a specific but universal primer, it allows highly reproducible fingerprints to be rapidly obtained across distantly related taxa. Five 5 fluorolabelled universal primers were used (809: $(AG)_8G$, 834: $(AG)_8YT$, 841:$(AG)_8YC$, 864: $(ATG)_6$ and 880: $(GGAGA)_3$). Reactions were carried out in 15μL reactions containing 1x High Fidelity PCR Buffer, 2 mM $MgSO_4$, 200 μM each dNTP, 0.8 μM primer, 1 ng template DNA (diluted in water) and 0.3 unit of *Taq Polymerase* (Invitrogen Life Technologies). All amplifications were conducted on the same Peltier thermal cycler (DNA Engine Tetrad 2) following a cycling profile with an initial step of 96°C for 5 min then 35 cycles of denaturing at 96°C (1 min.), primer annealing at 50°C (30 sec.), then extension at 68°C (1 min.) and a final extension step of 68°C (10 min). PCR products were cleaned using sephadex G-50 plates (Whatman) and multiplexed into 10μL volumes containing 0.5μl of each product and 0.25μl standard (Amersham ET 900). Labelled extension products were genotyped on a Megabase 1000 (Amersham Biosciences) at the Genetic Analysis Facility in the Advanced Analytical Center (James Cook University). Presence and absence of bands between 50 and 850 base pairs were scored using MegaBACE Fragment Profiler 1.2 (Amersham Biosciences), then converted into binary data matrices and concatenated. Raw binary data matrices are available from the author upon request.

Statistical Analyses

mtDNA: Forward and reverse sequences were aligned using Sequencher 4.2 (GeneCodes Corp. Michigan USA). The best fitting substitution model and associated rate heterogeneity were estimated separately for each species using Modeltest 3.5 (Posada and Crandall 1998) and PAUP* 4.0b10 (Swofford 1998) and these, where possible, were implemented in all subsequent analyses. Genetic diversity measures (haplotype and nucleotide diversity) (Tajima 1983; Nei 1987) and their associated standard deviations were calculated using Arlequin 2.000 (Schneider et al. 2000). Standard deviations were converted to 95% confidence intervals (95% CI = $\pm$1.96*(SD / $\sqrt{}$ (n))) and genetic diversities were interpreted as statistically different when 95% confidence intervals did not overlap. Estimates of genetic structure were calculated as pairwise Φ_{ST} values following the methods implemented in Arlequin and significance levels were corrected for multiple comparisons following the Dunn-Sidak method (Sokal and Rohlf 1995). Transition–transversion ratios indicated that saturation may be occurring in some of the species, and therefore, all analyses were repeated using transversions only.

ISSR data: Due to the large number of fragments amplified by the 5 primers, only bands with minimum frequencies of 0.25 within species were analysed. No differences were found among analyses of 0.05, 0.1 and 0.25 minimum frequency data sets conducted for a subset of species (Bay, unpublished data). Mendelian segregation of fragments with a single dominant (amplified) and recessive (absent) allele at each banding position was assumed. Because dominant data do not allow within-individual heterozygosities to be estimated, Hardy-Weinberg equilibrium was assumed in the analysis of these data. Given the large population sizes and pelagic larval dispersal in all species except *A. polyacanthus,* this assumption was deemed reasonable in this study. Analogues of codominant genetic diversity measures including within-population diversity H_e (here H_j), total heterozygosity H_t and mean heterozygosity across populations H_s (here H_w) were calculated following the methods of Lynch and Milligan (1994) using AFLP-Surv 1 (Vekemans et al. 2002). Their associated standard errors were converted to 95% confidence intervals as above. Genetic differentiation (Φ_{PT}) was estimated using Genalex 5 (Peakall and Smouse 2001) and significance levels were corrected for multiple comparisons (Sokal and Rohlf 1995).

The relationships between genetic differentiation, genetic diversities and larval duration were explored independently for each marker using mean, minimum and

maximum PLD both including and excluding *A. polyacanthus,* by fitting least-squares regressions. Estimates of fixation (Φ_{ST} and Φ_{PT}) showed some departure from normality, however, transformations failed to rectify these and untransformed data were analysed. Genetic diversity estimates between locations for each species separately were interpreted as statistically different when confidence intervals did not overlap. The non-parametric Mann Whitney U test was used to test for differences in genetic diversities between locations and the Kruskal-Wallis test was used to test for differences among species. The results of Doherty et al. (1995) were reanalysed using log transformed mean F_{ST} values from their Table 5, mean, minimum and maximum PLD values from their Table 1 and the mean, minimum and maximum PLD values used here (Table 1).

Results

Genetic differentiation and PLD

Control region: 335 – 555 base pairs at the 5' end of the control region were resolved in 39 – 46 individuals from each of eight species of pomacentrid fishes on the Great Barrier Reef (Table 2). Pairwise genetic distances were generally low and insignificant in most species (although near significant in *A. curacao* and *C. rex*), but high and significant in *A. polyacanthus* (Table 3). A significant negative relationship between mean, minimum and maximum PLD (Mean PLD: $F_{(1,6)} = 17.23$, $P = 0.006$; Minimum PLD: $F_{(1,6)} = 11.21$, $P = 0.015$; Maximum PLD: ($F_{(1,6)} = 15.85$, $P = 0.007$) and Φ_{ST} was evident (Fig. 1). Mean, minimum and maximum PLD explained 70, 59 and 68% of the variation in Φ_{ST}, respectively. These relationships were not significantly different with the 95% confidence intervals of all slopes and intercepts overlapping (mean PLD: y = -0.037 (-0.06 – -0.02) x + 0.679 (0.304 – 1.05); Min PLD: y = -0.038 (-0.07 – -0.01) x + 0.611 (0.203 – 1.019); Max PLD: y = -0.027 (-0.04 – -0.01) x + 0.689 (0.299 – 1.079). Consequently, only the relationship between mean PLD and Φ_{ST} is presented (Fig. 1). All three relationships, however, were driven primarily by *A. polyacanthus.* With this species removed, little variation in genetic structure was explained (Mean PLD: $F_{(1,5)} = 2.89$, $P = 0.15$; Min PLD: $F_{(1,5)} = 1.804$, $P = 0.237$; Max PLD:$F_{(1,5)} = 2.76$, $P = 0.16$).

Table 2: Sample sizes, Number of transitions (ts) vs. number of transversions (tv), ts/tv ratio, substitution model selected and among rates variation (gamma) estimated by Modeltest and model implemented for the eight species.

Species	No. base pairs	No.Ts/ No.Tv	ts - tv ratio	Model	-ln Likelihood	Gamma	Model[1]
Acanthochromis polyacanthus	351	34/24	8.5	TN(93)	1220.65	0.3012	TN(93) + G
Amphiprion melanopus	335	17/3	6.3	HKY	576.73	0	TN(93)
Pomacentrus moluccensis	349	21/4	5.9	HKY	652.70	0	TN(93)
Pomacentrus amboinensis	555	26/7	3.4	HKY	983.98	0	TN(93)
Chromis atripectoralis	349	42/7	7.8	HKY	924.56	0	TN(93)
Amblyglyphidodon curacao	398	73/5	12.5	HKY	989.57	0.1687	TN(93) + G
Chrysiptera rex	348	31/3	14.4	HKY	751.84	0.0069	TN(93) + G
Stegastes nigricans	423	51/5	16.1	HKY	1065.54	0.0146	TN(93) + G

[1] Implemented in Arlequin.

Transitions accounted for the majority of substitutions in all species (Table 2) although they accounted for less of the variation in *A. polyacanthus*. Therefore, the population genetic structure of these species may be affected by saturation. Estimates of genetic differentiation based on transversions alone were low and insignificant in all species except *A. polyacanthus* (Table 3). The exclusion of transitions did not appear to affect the relationship with PLD. Using the transversion data alone, PLD explained less variation in Φ_{ST} (Adj. $R^2 = 0.54$) and was not significantly different to the relationship between PLD and Φ_{ST} obtained with both transitions and transversions because the 95% confidence intervals of both slopes and intercepts overlapped (Mean PLD: $F_{(1,6)} = 9.128$, $P = 0.02$, $y = -0.033$ (-0.059 – -0.006)x + 0.64 (0.18 – 1.10)). This relationship also became statistically insignificant with *A. polyacanthus* removed ($F_{(1,5)} = 0.76$, $P = 0.42$).

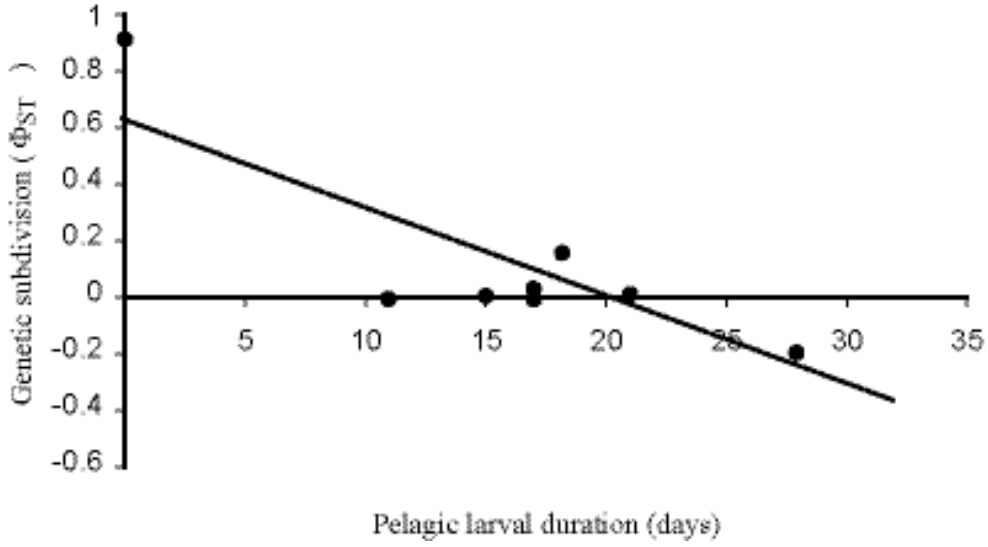

Fig. 1: The relationships between mean pelagic larval duration and genetic structure of the mtDNA (Φ_{ST}) in eight species of coral reef fishes.

Estimates of genetic differentiation based on the ISSR markers were higher compared to estimates based on the mtDNA control region and indicated significant genetic structure between sampling locations in all species (Table 3). There was a significant negative relationship between minimum PLD and genetic subdivision explaining 85% of the variation ($F_{(1,3)} = 22.81$, $P = 0.01$) (Fig. 2).

Table 3: Estimates of genetic differentiation based on MtDNA control region (Φ_{ST}) including and excluding transitions and ISSRs (Φ_{PT}) and their significance levels.

Species	Transitions and transversions		Transversions		ISSR	
	Φ_{ST}	p	Φ_{ST}	P	Φ_{PT}	p
Acanthochromis polyacanthus	0.91	<0.001[1]	0.89	<0.001[1]	0.426	0.001[2]
Amphiprion melanopus	-0.005	0.47	-0.001	0.42	0.188	0.001[3]
Pomacentrus moluccensis	0.006	0.25	0.01	0.18	0.018	0.022[4]
Pomacentrus amboinensis	-0.008	0.60	-0.002	0.45		
Chromis atripectoralis	0.012	0.23	0.011	0.25	0.272	0.001[5]
Amblyglyphidodon curacao	0.030	0.06	0.024	0.07		
Chrysiptera rex	0.158	0.28	0.044	0.07	0.027	0.007[6]
Stegastes nigricans	-0.195	0.97	0.014	0.17		

Bonferroni corrected α: [1] 0.006, [2] 0.01, [3] 0.02, [4] 0.05, [5] 0.01, [6] 0.03.

However, neither mean nor maximum PLD was significantly related to Φ_{PT} (Mean PLD: $F_{(1,3)} = 5.76$, P = 0.1; Max PLD: $F_{(1,3)} = 4.00$, P = 0.14). The relationship between minimum PLD and Φ_{PT} became statistically insignificant upon exclusion of *A. polyacanthus* ($F_{(1,2)} = 4.98$, P = 0.16).

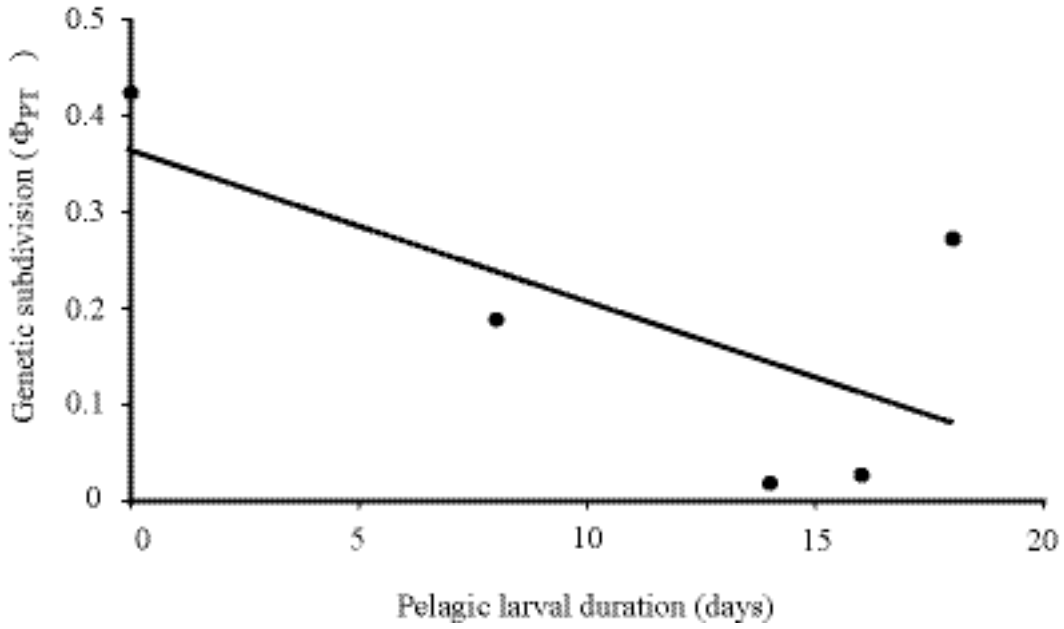

Fig. 2: The relationships between minimum pelagic larval duration and genetic structure of the nucDNA (Φ_{PT}) in five species of coral reef fishes.

A re-analysis of the results of Doherty et al. (1995) found that their relationship between PLD and gene flow remained significant when using their estimate of minimum PLD ($F_{(1,5)} = 22.26$, P = 0.005, Adj. $R^2 = 0.78$) but became insignificant when maximum PLD was used ($F_{(1,5)} = 4.95$, P = 0.07, Adj. $R^2 = 0.39$). The relationships between mean and minimum PLD and genetic structure were not significantly different and the 95% confidence intervals of both slopes and intercepts overlapped (Mean PLD: y = -0.04 (-0.06 – -0.02)x -- 0.313 (-0.95 – 0.32); Min PLD: y = -0.04 (-0.07 – -0.02)x - 0.343 (-1.04 – 0.35)). The relationship between mean PLD and genetic structure of the benthic spawning species (including the brooding species) became insignificant when mean PLD values estimated in Chapter 4 were used (($F_{(1,3)} = 4.07$, P = 0.14). Maximum PLD values (Chapter 4) produced a significant relationship with genetic differentiation, whereas minimum PLD did not (Max PLD: $F_{(1,3)} = 13.51$, P = 0.035; Min PLD: $F_{(1,3)} = 2.08$, p = 0.245).

Genetic diversity and PLD

None of mean, minimum or maximum PLD could predict haplotype diversity (Mean PLD: $F_{(1,6)} = 3.51$, P = 0.11; Min PLD: $F_{(1,6)} = 1.74$, P = 0.24; Max PLD: $F_{(1,6)} = 3.14$, P = 0.13). Mean and maximum PLD were negatively related to nucleotide diversity (Mean PLD: $F_{(1,6)} = 6.49$, P = 0.04, Adj. R^2 0.44; Max PLD: $F_{(1,6)} = 18.18$, P = 0.005, Adj. R^2 = 0.53) but again this relationship was reliant on *A. polyacanthus* (with *A. polyacanthus* excluded: Mean PLD: $F_{(1,5)} = 0.58$, P = 0.48; Max PLD: $F_{(1,5)} = 0.02$, P = 0.89). Minimum PLD did not predict nucleotide diversities (Min PLD: $F_{(1,6)} = 5.89$, P = 0.051). Minimum PLD could predict total heterozygosity (H_t) ($F_{(1,3)} = 12.39$, P = 0.039, Adj. $R^2 = 0.74$), but neither mean nor maximum PLD could predict H_t (Mean PLD: $F_{(1,3)}$ = 6.97, P = 0.078; Max PLD: $F_{(1,3)} = 2.57$, P = 0.21). Again, this relationship became insignificant upon exclusion of *A. polyacanthus* ($F_{(1,2)} = 4.59$, P = 0.17). Neither mean, minimum nor maximum PLD could predict mean heterozygosity among locations (H_j) (Mean PLD: $F_{(1,3)} = 1.99$, P = 0.25; Min PLD: $F_{(1,3)} = 3.77$, P = 0.15; Max PLD: $F_{(1,3)}$ = 2.89, P = 0.19).

Genetic diversities among species and locations

Haplotype diversities ranged from 0.8 – 1, other than *A. polyacanthus* at One Tree Island which was 0.49. Haplotype diversities were similar between sampling locations for all species except two (*A. polyacanthus* and *C. rex*). In these species, the northern populations (Lizard Island) were more diverse than the southern ones (One Tree Island) (Fig. 3a). Nucleotide diversities ranged from 0.1 – 3.2% (although most species were less than 2.5%) and were similar between sampling locations of all species except two. In *A. polyacanthus,* the northern location was more diverse, and in *P. amboinensis* the southern location was more diverse (Fig. 3b). Heterozygosities ranged from 0.3 – 0.4 among species and locations, and were significantly different between sampling locations in two species (*A. melanopus* and *C. atripectoralis*) (Fig. 4). There were no significant differences between haplotype diversity (Fig. 3a Mann-Whitney U = 23.5, Z = -0.89, P = 0.37), nucleotide diversity (Fig. 3b Mann-Whitney U = 29.5, Z = -0.26, P = 0.79) or heterozygosity (Fig. 4 Mann-Whitney U = 12, Z = -0.104, P = 0.92) and geographical location, nor were there any significant differences in genetic diversities among species (Haplotype diversity: Kruskal-Wallis $H_{(7,\ 16)} = 13.08$, P = 0.07;

Nucleotide diversity Kruskal-Wallis $H_{(7, 16)} = 9.26$, $P = 0.23$; H_w: Kruskal-Wallis $H_{(4, 10)} = 5.345$, $P = 0.25$).

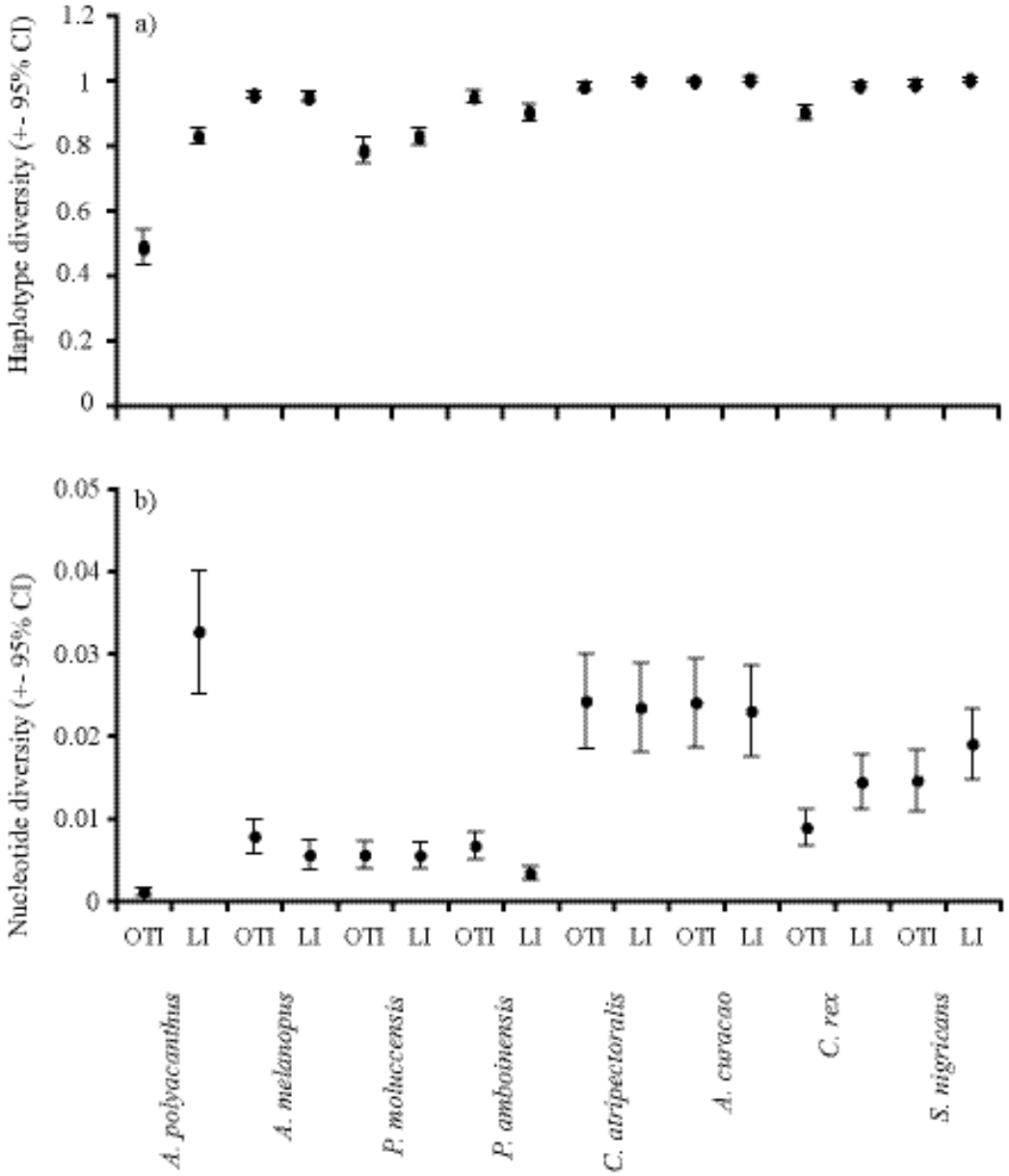

Fig. 3: Genetic diversities of the mtDNA in southern (One Tree Island, OTI) and northern (Lizard Island, LI) populations of eight species of coral reef fishes.

Discussion

Genetic structure on the GBR

Estimates of genetic differentiation differed between molecular markers, but were consistent among species. In both datasets, *A. polyacanthus* displayed much stronger genetic differentiation than all other species, although less structure was indicated by

the ISSRs (Table 3). This result was expected given the strong genetic structure often recorded in this species (Doherty et al. 1994; Doherty et al. 1995; Chapter 2). It also adds support to the notion that directly developing marine species generally display more genetic structure than ones with pelagic larvae (e.g., Hunt 1993; Hellberg 1996; Arndt and Smith 1998).

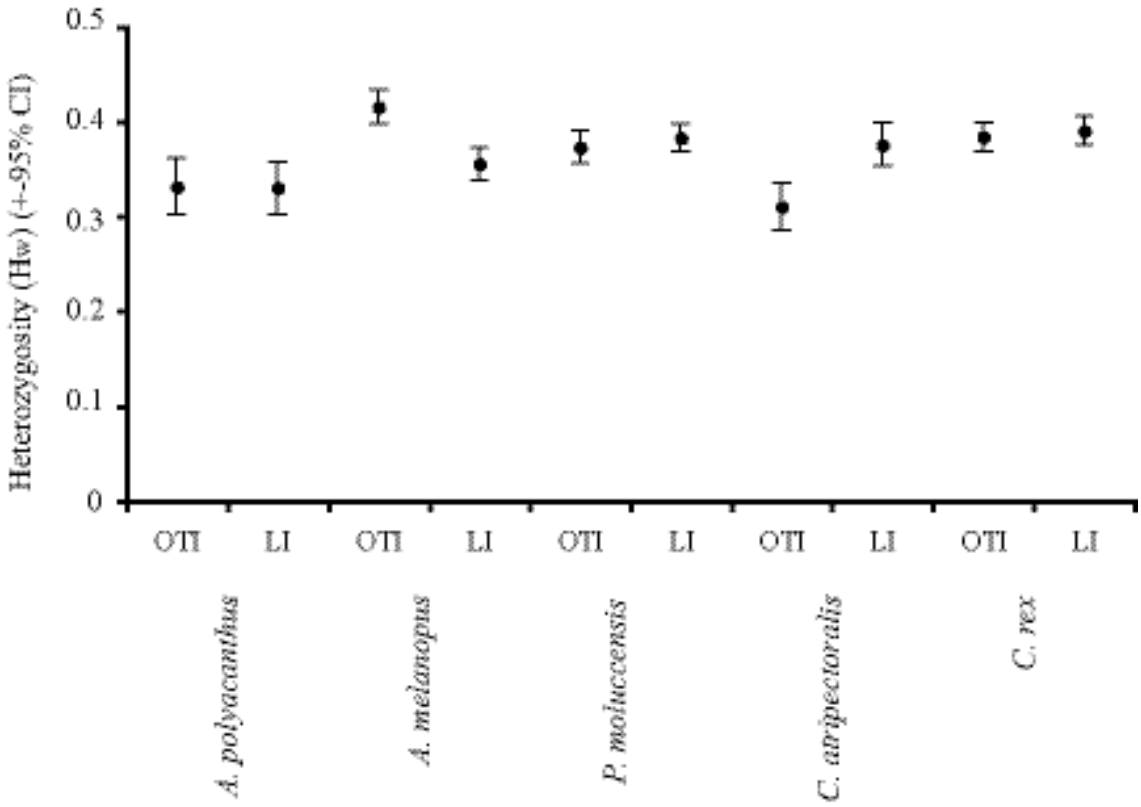

Fig. 4: Within population heterozygosity (H_w) in southern (One Tree Island, OTI) and northern (Lizard Island, LI) populations of five species of coral reef fishes.

The mtDNA marker revealed no genetic structure between northern and southern locations in all species with pelagic larvae. These species all displayed very small and insignificant Φ_{ST} estimates (Table 3). In contrast, the ISSRs revealed significant structure between the sampling locations with Φ_{PT} values ranging from 0.02 – 0.27 (Table 3). Such differences in the population structure indicated by the two markers may arise because of the large effective population size (reducing the effect of genetic drift) and the higher variability of the nucDNA marker compared to the mtDNA marker. The higher level of fixation indicated by the mtDNA may also be indicative of

sex-biased dispersal. Sex-biased dispersal may evolve when there are sex-biased fitness consequences associated with acquiring and defending reproductive resources (Greenwood 1980; Clarke et al. 1997). Where sex-biased dispersal has been reported in fishes, it tends to be male biased (Hutchings and Gerber 2002; Fraser et al. 2004). Sex-biased dispersal can be identified directly by a comparison of the population genetic structure of males and females (Mossman and Waser 1999), or indirectly, by a comparison of maternally inherited molecular markers and those with bi-parental inheritance (Avise 2000). If males migrate at higher rates than females, maternally inherited mitochondrial genes should show more genetic structure, compared to bi-parentally inherited nuclear genes. This is the pattern observed in *A. polyacanthus* (Table 3). Conversely, if dispersal is female biased, mitochondrial markers would be expected to be less differentiated than nuclear markers. This pattern was observed in all species with pelagic larvae (Table 3). Male-biased dispersal has recently been suggested for *A. polyacanthus* on the GBR (Chapter 2) and the higher genetic structure of the mtDNA marker compared with the nucDNA marker reported here is consistent with this hypothesis (Φ_{ST} = 0.90; Φ_{PT} = 0.43). The other species included in this study are all thought to disperse primarily during the pelagic larval phase (Leis 1991; Bonhomme and Planes 2000). It is unclear whether these larvae are sex differentiated (Fishelson 1998) and if so, how female larvae would realise greater dispersal than male larvae. Furthermore, the adults of many of the species included here change sex (e.g., *A. melanopus* Godwin and Thomas 1993) and reproductive resources are defended by males in most of the species (but females in *A. melanopus*). Therefore, I consider that the differences in genetic differentiation between the mtDNA and nucDNA markers for these species with pelagic larvae are unlikely to be caused by sex-biased dispersal.

It is also possible that the differences in population structure indicated by the mtDNA and nucDNA markers were caused by differences in mutation rates between markers. The ISSRs record microsatellite variation (Zietkiewicz et al. 1994; Bornet and Branchard 2001) and should therefore evolve at a rapid rate. If so, under low levels of migration, such rapidly evolving markers should show local genetic structure because of a high mutation rate, whereas markers evolving at a slower rate might not (Avise 2000). Consequently, all species with pelagic larvae included in this study display genetic structures consistent with some migration between populations in the northern and southern parts of the GBR. This conclusion is compatible with previous results based on

allozyme markers (Doherty et al. 1995) where significant genetic structure was found in five of the same species among locations on the GBR.

Genetic diversities may differ within and among species due to differences in their evolutionary histories (Fauvelot et al. 2003). However, genetic diversities recorded here were similar within and among species. Haplotype diversities were generally high and comparable with those observed in other coral reef fishes on the GBR (e.g., *Chlorurus sordidus h* = 0.98 Bay et al. 2004; *Gobiodon histrio h* = 0.87 Munday et al. 2004; *Pseudochromis fuscus h* = 0.79 Messmer et al. 2005). The very low haplotype diversity in the southern population of *A. polyacanthus* is most likely due to a recent mitochondrial or population bottleneck (van Herwerden and Doherty 2006; Chapter 3). Nucleotide diversities were also generally high and comparable to estimates from other coral reef fish species on the GBR (e.g., *Chlorurus sordidus* % π = 3.0 Bay et al. 2004; *Gobiodon histrio* % π = 1.23 Munday et al. 2004; *Pseudochromis fuscus* % π = 0.36 Messmer et al. 2005). The very high nucleotide diversity from the northern population of *A. polyacanthus* appears to be the result of the presence of two differentiated mtDNA lineages in this region (Chapter 3). Estimates of heterozygosity were similar between sampling locations within species (except in *A. melanopus* and *C. atripectoralis*) and among species (Fig. 4). ISSRs have not previously been used to investigate the population genetic structure of fishes and could therefore not be compared to previous estimates. Consequently, the estimates of genetic variability of both mtDNA and nucDNA markers indicated that recent evolutionary histories did not differ substantially among locations or species.

PLD and population genetic structure

I identified a relationship between PLD and genetic structure using both the mtDNA and the nucDNA markers (Fig. 1 and 2). However, this relationship was strongly influenced by the inclusion of the directly developing species (*A. polyacanthus*) that displayed strong genetic differentiation between locations. Once this species was removed from the analysis, little variation was explained. Riginos and Victor (2001) found a significant relationship between PLD and gene flow in three species of blennies with mean larval durations spanning 32 days (range = 28 – 50 days). It is possible that the range of larval durations exhibited by the species included in the present study (28 days) was not great enough to reveal measurable differences in genetic differentiation.

However, given the significant genetic structure revealed by the ISSRs here and the analysis by Doherty et al. (1995) indicating that the genetic structure of species with larval durations between 9 – 39 days should be determined by the opposing forces of drift and migration, such a possibility appears unlikely. Instead, the species with pelagic larvae included in this study more likely displayed very similar levels of genetic structure unrelated to the length of their larval life. Because these species are closely related, and have relatively generalised ecologies and similar spawning behaviours (Table 1), this result may indicate that biological and ecological traits, some of which may be associated with phylogeny, may be important in producing differences in gene flow, independent of the duration of the pelagic larval phase.

Population genetic structure may result from a range of causes that were not explored here. For example, ecological specialisation may play an important role in shaping patterns of genetic structure even in species with extensive larval durations (e.g., Taylor and Hellberg 2003; Rocha et al. 2005). The inclusion of such species may have affected previous investigations of the relationship between genetic structure and spawning strategy. For example, Shulman and Bermingham (1995) reported high genetic structure in *Halichoeres bivittatus* indicating low migration rates contrary to expectations in a species with pelagic eggs and a relatively long larval duration (i.e., 24.1 days). The high population genetic structure of this species was more recently attributed to its level of ecological specialisation (Rocha et al. 2005). With the removal of this species, the data of Shulman and Bermingham (1995) were consistent with the idea of higher gene flow in pelagically spawning species. Consequently, examinations of the relationship between gene flow and PLD or spawning mode need to account, where possible, for such biological and ecological factors. Future investigations comparing the relationship between PLD and genetic differentiation among groups of species with different biological, ecological or biogeographical attributes have the potential to illuminate this issue further.

PLDs may vary substantially within species (e.g., Chapter 4). Such variation could weaken any relationship between PLD and genetic differentiation (Leis 1991; Victor 1991) and may vary among datasets generated using different molecular markers that evolve at different rates. For example, the population genetic structure recorded by more slowly evolving molecular markers may be greatly affected by low levels of migration (Wright 1931). Individuals displaying a longer PLD may, therefore, influence the population genetic structure of species, even if such maximum PLDs are only rarely

expressed. In contrast, allele frequencies of rapidly evolving molecular markers should not be as strongly affected by rare long distance dispersers because of the rate at which new genetic variation is generated by mutation within local populations. My analyses, and the re-analysis of the results of Doherty et al. (1995), indicated that the relationship between PLD and genetic structure may differ depending on the estimate of PLD used and its variation. Using mean, minimum or maximum PLD did not affect the relationship between PLD and genetic structure estimated using mtDNA. Only minimum PLD could predict gene flow and total heterozygosity using the more rapidly evolving nucDNA marker. The estimates of genetic structure based on allozymes could be predicted from their estimates of mean and minimum PLD, but not from my mean and maximum PLD estimates. The low number of species included in the ISSR analysis could potentially have restricted my ability to detect such relationships using this marker. However, it is also likely that the population genetic structure of the species investigated here based on the nucDNA, and presumably faster evolving marker, were more strongly influenced by individuals displaying a shorter PLD, whereas the relationships based on the mtDNA, and presumably slower evolving marker were more strongly affected by mean and maximum larval durations.

Conclusion

I identified a relationships between PLD and population genetic structure using both mtDNA and nucDNA markers. These relationships, however, were dependent on the inclusion of a directly developing species with high genetic structure. With this species removed PLD could not predict population genetic structure in the species examined here. The relationship between PLD and genetic structure varied depending on the estimates of PLD and genetic structure used. These results suggest that observed relationships between the population genetic structure and pelagic larval duration may be highly dependent on the molecular marker, estimator of PLD and species used. Further consideration of such variation has the potential to provide additional insights into the relationship between population genetic structure and dispersal potential.