Aquaculture Research

Aquaculture Research, 2015, 1–15 doi: 10.1111/are.12807

Genetic diversity and relatedness estimates for captive barramundi (*Lates calcarifer*, Bloch) broodstock informs efforts to form a base population for selective breeding

Shannon R Loughnan[1], Carolyn Smith-Keune[2], Dean R Jerry[2], Luciano B Beheregaray[1] & Nicholas A Robinson[1,3]

[1]School of Biological Sciences, Flinders University, Adelaide, South Australia, Australia

[2]Centre for Sustainable Tropical Fisheries and Aquaculture, College of Marine and Environmental Sciences, James Cook University, Townsville, Queensland, Australia

[3]Nofima, Ås, Norway

Correspondence: S R Loughnan, School of Biological Sciences, Flinders University, PO Box 2100, Adelaide 5001, South Australia. E-mail: shannon.loughnan@flinders.edu.au

Abstract

Aquaculture of barramundi or Asian seabass (*Lates calcarifer*) is growing in both Australia and South-east Asia and there is substantial interest to improve production efficiency through selective breeding. The establishment of a large and genetically diverse base population is a prerequisite for a sustainable and long-term productive breeding program. Before selective breeding programs can begin for Australian barramundi it is important to assess the overall genetic diversity of current captive broodstock populations. To address this question, 407 captive barramundi broodstock from eight separate Australian broodstock populations were genotyped using 16 polymorphic microsatellite DNA markers. A Bayesian STRUCTURE analysis indicated that captive Australian broodstock are broadly divided into two genetic stocks. Multivariate analysis between broodstock individuals and pairwise F_{ST} between broodstock populations also supported the existence of two stocks. Comparisons with data obtained from natural stocks suggested that hatchery individuals were either sourced from the two stocks or represented an admixture between them. Genetic diversity was low within each broodstock population (allelic richness ranged from 2.67 to 3.42 and heterozygosity ranged from 0.453 to 0.537) and relatedness estimates within hatcheries were generally low (average r was equal to 0.141). We recommend sourcing captive individuals according to high levels of neutral genetic diversity and low levels of relatedness for the establishment of a base population. We also make recommendations about including genetically diverse wild individuals.

Keywords: Asian seabass, captive breeding, genetic diversity, microsatellite, relatedness

Introduction

The long-term success of closed selective breeding programs is contingent on the extent of additive genetic variance captured in the base population. This is because the response to selection for any particular trait depends on the intensity of selection, the heritability of the trait and the additive genetic variance existing in the population for the trait (Falconer & Mackay 1996). The naturally high levels of genetic diversity observed in marine fish species (DeWoody & Avise 2000) is thought to be what makes the rate of genetic improvement so rapid for many fish species compared to livestock or plants (Gjedrem & Robinson 2014). There will be greater scope for genetic improvement of important production traits if high genetic diversity for those traits can be captured in the base population, and if loss of genetic diversity can be limited after the population is closed to new entrants. This is why the broad range of genetic variation captured in the base population at the commencement of some selective

breeding programs, has been partly attributed to the success of genetic improvement (e.g. Nile Tilapia, *Oreochromis niloticus* Linnaeus 1758, Eknath, Bentsen, Ponzoni, Rye, Nguyen, Thodesen & Gjerde 2007; Atlantic salmon *Salmo salar* L. 1758, Gjedrem, Gjøen & Gjerde 1991).

Before controlled crosses can be used to assess the breeding value of individuals or the extent of additive genetic variance for particular traits, decisions need to be made about which wild or existing captive populations should be used as sources for establishing the base population. At the time that animals are sourced for the base population it is usually not possible to assess the level of additive genetic variability for particular traits, or to make breeding value predictions, because the fish have not been grown together for the same length of time in a common environment. However, it is possible to measure and compare overall levels of genetic variability across the genome at selectively neutral loci. Populations that are more genetically diverse with respect to random selectively neutral loci are generally less likely to have been exposed to genetic drift and other processes reducing genetic diversity, and consequently will be more likely to contain high genetic diversity for genes affecting traits, and therefore more likely to contain particular genetic variants that benefit the on-farm performance of stock.

Typically, a loss of genetic diversity occurs in all closed populations through genetic drift and this loss is increased with each generation of breeding if the genetically effective population size (N_e) is low (Frankham, Ballou & Briscoe 2002) and if the breeding of close relatives is not avoided. Inbreeding (ΔF) is known to lead to depression of fitness in fish (Wang, Hard & Utter 2002) due to exposure of deleterious recessive genes and it can also reduce the potential for achieving genetic gain. Breeding programs without an adequate base population and/or with poorly managed 'selective breeding' (Li, Park, Endo & Kijima 2004; Schwartz & Beheregaray 2008), could therefore negatively affect performance traits. Such populations would also require regular supplementation with new animals to limit inbreeding depression of fitness and control the loss of genetic variation to acceptable levels. Limiting inbreeding becomes more difficult with successive generations of selective breeding, although it is generally accepted that a N_e greater than 100, resulting in ΔF less than 0.5%, is sufficient in each generation to avoid serious problems in captive populations (Fjalestad 2005; Sonesson, Wooliams & Meuwissen 2005).

By utilizing molecular DNA markers such as microsatellites, neutral genetic diversity and relatedness of broodstock candidates can be estimated prior to establishing the base population. The relatedness or kinship between individuals x and y (r_{xy}) is a measure of the fraction of alleles that are identical by descent (IBD) and a pair of individuals are deemed related if they share one or more alleles inherited from a common ancestor. By sourcing parents based on low r_{xy}, it is possible to reduce the chance that inbreeding will occur during the initial generations of selection (Doyle, Perez-Enriquez, Takagi & Taniguchi 2001; Rodzen, Famula & May 2004; Sekino, Sugaya, Hara & Taniguchi 2004). It is recommended that individuals with lower average r_{xy} (when compared to all other broodstock candidates) should be given priority for inclusion in the base population (Porta, Porta, Martínez-Rodríguez & Alvarez 2006).

Asian seabass (*Lates calcarifer*, Bloch, 1790) also known as barramundi in Australia, has potential as a candidate species for genetic improvement as its production in aquaculture is growing (Skirtun, Sahlqvist & Viera 2013). Barramundi also has high fecundity (Palmer, Blackshaw & Garrett 1993) and has moderate heritability for economically valuable traits such as growth rate (Wang, Lo, Zhu, Lin, Feng, Li, Yang, Tan, Chou, Lim, Orban & Yue 2008; Domingos, Smith-Keune, Robinson, Loughnan, Harrison & Jerry 2013). Barramundi readily spawn in captive culture and naturally breed in groups (mass spawn) providing the opportunity for creating numerous parent pair families. However, the mass spawning nature of this species means that there is little control over the contribution of individual broodstock to a particular spawning event (Frost, Evans & Jerry 2006; Wang *et al.* 2008; Loughnan, Domingos, Smith-Keune, Forrester, Jerry, Beheregaray & Robinson 2013) and N_e is therefore typically much less than the census size (N_c). As a result, a substantial number of unrelated broodstock are required to control ΔF and to provide a $N_e > 100$.

This study investigates how founding parents should be best sourced so that the genetic diversity in this base stock (and, consequently all subsequent generations) is maximized. This is the first consideration that should be made when establishing a closed breeding program (Ballou & Lacy

1995). In this study we use general indicators of overall allelic variability (16 microsatellite loci) to guide the choice of founding parents such that general allelic diversity is maximized and relatedness minimized in the establishment of a base population of barramundi for selective breeding in Australia. This is one of the first practical examples of the use of population genetic data in this way for guiding the establishment of a selective breeding program.

Materials and methods

Sampling, DNA extraction and genotyping

Barramundi broodstock samples (N_c = 407) were collected from eight commercial Australian hatcheries; one in Western Australia (WA, N_c = 48) and one in the Northern Territory (NT, N_c = 71), and six in Queensland (QLD1, N_c = 58; QLD2, N_c = 14; QLD3, N_c = 111; QLD4, N_c = 80; QLD5, N_c = 9; QLD6, N_c = 16). The broodstock sampled represented more than 90% of the total population in the Australian industry at the time of this study. Within each hatchery, all broodstock made accessible were sampled regardless of whether they were under current use, were listed as backup broodstock, or had not yet reached sexual maturity. At the time of sampling there were 136 females, 180 males and 91 fish of undetermined sex, and the majority of broodstock (N = 349) were captive bred fish (the result of one or more generations of breeding in captivity and some were acquired from other hatcheries). Only the NT and QLD4 captive populations contained wild caught broodstock, with 51 and 7 individuals, respectively, while the remaining captive populations did not contain any wild-caught broodstock. Samples from three natural barramundi localities were included for comparison to the broodstock populations; St George Basin (STG) N = 30, Liverpool Creek (LVP) N = 32 and the Burdekin River (BUR) N = 24. Jerry and Smith-Keune (2014) detected three distinct genetic regions among samples from the Australian distribution of barramundi and the three natural populations in this study were chosen to represent each of these; STG for the western stock, LVP for the central admixed region and BUR for the eastern stock.

All broodstock were sedated in a saltwater bath containing 40 ppm AQUI-S (Aquatic Diagnostic Services International, NSW, Australia) and a small segment of caudal fin (*ca.* 1 cm^2) was removed and preserved in either 80% ethanol or DMSO-salt solution (20% DMSO, 0.25 M disodium-EDTA and NaCl to saturation at pH 8) (Seutin, White & Boag 1991). Passive integrated transponder (PIT) tags implanted in each individual were scanned to provide unique identification. While sedated, broodstock were cannulated with a 2.16 mm outside diameter (OD) catheter tube and the sex confirmed via observation of eggs or sperm under a microscope. Broodstock were then recovered from anaesthesia and placed back into their holding tanks as per standard industry practice.

Methods of DNA extraction were described in Loughnan *et al.* (2013) following the CTAB (cetyl trimethylammonium bromide) protocol described by Adamkewicz and Harasewych (1996). As for Loughnan *et al.* (2013), the same 17 microsatellite markers were amplified in two multiplex reactions using the polymerase chain reaction (PCR) procedures described therein, however, due to the detection of null alleles for marker *Lca287*, this locus was excluded from multiplex one. Multiplex one included markers *LcaM03* (Yue, Li & Orban 2001), *LcaM16*, *LcaM40* (Yue, Li, Chao, Chou & Orban 2002), *Lca57* (Zhu, Lin, Lo, Xu, Feng, Chou & Yue 2006a), *Lca154*, *Lca178* (Zhu, Wang, Lo, Feng, Lin & Yue 2006b) and *Lca371* (Wang, Zhu, Lo, Feng, Lin, Yang, Li & Yue 2007). Multiplex two included *LcaM08*, *LcaM20, LcaM21* (Yue *et al.* 2002), *Lca58*, *Lca64*, *Lca69*, *Lca70*, *Lca74* and *Lca98* (Zhu *et al.* 2006a). Genotyping was performed on a MegaBACE® 1000 DNA Analysis System (GE Healthcare, Silverwater, NSW, Australia) and MegaBACE® software Fragment Profiler® was used for fragment analysis. The genotyping error rate was calculated at 1.8% and the proportion of samples requiring re-genotyping to correct for potential genotyping errors was 3.5%.

Population analysis

To test for the presence of null alleles, large allele dropout and scoring errors MICRO-CHECKER 2.2.3 was utilized (Van Oosterhout, Hutchinson, Wills & Shipley 2004), applying 95% confidence intervals for Monte Carlo simulations. Null alleles were not accounted for when scoring genotypes. Following this, the average numbers of alleles (A), plus expected (H_e) and observed (H_o) heterozygosity were estimated in GENALEX 6.5 (Peakall & Smouse 2012). Allelic richness (A_r) and private

allelic richness (PA_r) were estimated in HP-RARE 1.1 (Kalinowski 2005), incorporating a rarefaction approach for a minimum of 14 alleles per sample (7 diploid individuals). A_r is a measure of the number of alleles independent of sample size and PA_r is a measure of unique or rare alleles within a population. The inbreeding coefficient (F_{is}), which measures the degree of random mating within populations and associated significance tests were calculated in FSTAT 2.9.3.2 (Goudet 2002) using the Weir and Cockerham (1984) method followed by sequential Bonferroni correction for multiple comparisons (Rice 1988). Tests for Hardy–Weinberg Equilibrium (HWE) and linkage disequilibrium (LD) were calculated in GENEPOP 4.1 (Rousset 2008). Significance was determined with sequential Bonferroni correction. Exact P values under the Markov Chain method were implemented with a dememorization step of 10,000 followed by 20 batches (100 batches for LD) of 5000 iterations per batch. Kruskal–Wallis tests were performed in IBM SPSS 20.0 for assessing whether broodstock populations differed statistically for two measures of genetic diversity; A_r and PA_r.

BOTTLENECK 1.2.02 (Cornuet & Luikart 1996; Piry, Luikart & Cornuet 1999) was utilized to check for signatures of recently reduced N_e within each broodstock population. The stepwise mutation model (SMM) and the two-phase model (TPM) were selected in BOTTLENECK and run for 1000 iterations, as recommended for microsatellite applications (Luikart & Cornuet 1998). The variance for TPM was set at 30 and the proportion of SMM in TPM was 70%. The mode-shift option was also applied to observe the distribution of allele frequencies (Luikart, Allendorf, Cornuet & Sherwin 1998). A mode-shift is often found in populations that have experienced a recent bottleneck. Due to the relatively small number of markers available for bottleneck analysis (<20) the more appropriate Wilcoxon's test was applied to the data (Piry *et al.* 1999).

Population genetic structure was assessed across the 407 captive broodstock and 86 wild samples to determine the number of genetic stocks represented across the industry and to aid in sourcing candidates for a base population in a selective breeding program. A range of methods were utilized in the analysis of population structure. Firstly, the Bayesian method of individual clustering applied in STRUCTURE 2.3.3 was used (Pritchard, Stephens & Donnelly 2000) and accessed at the Bioportal computing resource (https://www.bioportal.uio.no/; Kumar, Skjæveland, Orr, Enger, Ruden, Mevik, Burki, Botnen & Shalchian-Tabrizi 2009). The most probable individuals were assigned to k populations with and without the use of sample location as a prior reference ('locprior'), a protocol designed to assess weak population structure. Admixture and correlated allele frequencies were applied for both models (Falush, Stephens & Pritchard 2003). Twenty replicate runs at each k (1–11) were performed (Gilbert, Andrew, Bock, Franklin, Kane, Moore, Moyers, Renaut, Rennison, Veen & Vines 2012). A burn in length of 100,000 iterations and one million Markov Chain Monte Carlo (MCMC) repetitions were applied for each run. The q-value threshold for stock assignment was >0.90 to a single cluster and <0.90 for the detection of admixture. STRUCTURE HARVESTER (Earl & vanHoldt 2012) was used to assess the most likely number of genetic populations (k) represented in the dataset (Evanno, Regnaut & Goudet 2005). The admixture proportions of each individual over the 20 replicates were averaged for the best k using CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and barplots were designed in DISTRUCT 1.1 (Rosenberg 2004).

A multivariate method for discriminant analysis of principal components (DAPC) was also used to test for population structure between captive broodstock (Jombart, Devillard & Balloux 2010) using the R package *adegenet* (Jombart 2008; R Development Core Team 2013). Finally, pairwise F_{ST} (Weir & Cockerham 1984), which is a measure of population differentiation due to genetic structure and associated P values were estimated in GENALEX 6.5 and incorporated 999 permutations (Peakall & Smouse 2012).

Relatedness estimates

The software COANCESTRY 1.0.1.2 (Wang 2011) was utilized to estimate relatedness (r_{xy}) between each dyad (e.g. pairs of individuals) within each of the eight broodstock populations and three natural localities. The program incorporates seven relatedness and three inbreeding estimators, to enable selection of the most appropriate estimator for the data set. The best performing relatedness estimator depends on the dataset of each study and more specifically on the number of microsatellite markers and the levels of variation detected (Van de Casteele, Galbusera & Matthysen 2001; Wang

2011). The COANCESTRY software incorporates Monte Carlo simulations, which were run with known allele frequencies calculated from the observed genotypes from all populations. True relationship classifications, which provide specific genealogical relations were set at $r_{xy} = 0.5$ for parent–offspring (PO) and full sib (FS), $r_{xy} = 0.25$ for half sib (HS) and $r_{xy} = 0$ for unrelated (U), simulating 1000 dyads for each relationship type and with 1000 bootstraps to calculate 95% confidence intervals. Following the simulation, the best estimator that yielded a strong correlation between the true and estimated values was selected. This was the Queller and Goodnight (1989) estimator (r_{QG}), with a correlation coefficient of $R = 0.79$, $P < 0.05$. Ranging from −1 to 1, the Queller and Goodnight (1989) relatedness estimator is one of the most widely chosen estimators for studies of kinship in both captive and natural populations (see Blouin 2003 for a review) and was applied to the empirical genotype dataset in this study to calculate r_{xy} estimates between all possible dyads. One-way ANOVA incorporating Tukey's post hoc tests were performed in IBM SPSS 20.0 to test for differences in r_{QG} between the broodstock populations. To infer FS and HS relationships within broodstock populations and natural stocks and also between all populations, the full likelihood method was implemented in COLONY 2.0.5.7 (Jones & Wang 2010).

Pedigree reconstruction and effective population size

PEDIGREE 2.2 (Herbinger, O'Reilly & Verspoor 2006) was utilized to determine the number of partitioned kin groups or FS groups involved in each of the eight captive broodstock populations and three natural localities. Full-sib constraint (FSC) was tested at both 0 (off) and 1 (on). Generating a kin partition (FSC = 0) suggests that individuals within a captive broodstock population could be related in an undetermined way. For the majority of the captive broodstock, pedigree records were limited, although hatcheries have generally aimed to collect potential broodstock from putatively unrelated individuals (author's personal observations). Setting FSC = 1 suggests that individuals within a group were derived from a single FS family, which was highly unlikely for both the captive and wild populations included in this study. In order to estimate the relative contributions of individuals to kin or FS groups without pedigree information, weights of W = 1, W = 5 and W = 10 were tested with one million iterations at temperatures of T = 10 and T = 30 for both FSC = 0 and FSC = 1.

The effective population size (N_e) for each captive broodstock population and natural locality was calculated using two methods; the linkage disequilibrium method utilizing LDNE 1.31 (Waples & Do 2008) with a minimum allele frequency of 0.05, and the molecular co-ancestry method of Nomura (2008), as implemented in NEESTIMATOR 2.01 (Do, Waples, Peel, Macbeth, Tillett & Ovenden 2014).

Results

Genetic diversity and HWE

The average number of alleles per locus was highest in the NT hatchery ($A = 5.6$), however, A_r and PA_r was highest in the natural localities representing the eastern stock, the Burdekin River ($A_r = 3.46$) and representing the western stock, the St George Basin ($PA_r = 0.53$) (Table 1). Excluding the natural stocks, A_r (3.42) and PA_r (0.32) estimates were highest in the NT broodstock population, and this was the population with the greatest number of reportedly wild caught individuals (72%). The lowest value of A_r was recorded for QLD6 (2.67) and PA_r was the lowest for QLD5 (0.02). Kruskal–Wallis tests revealed no significant difference in levels of A_r between the broodstock populations ($P = 0.827$), between the natural stocks ($P = 0.794$), or between the broodstock populations and natural stocks ($P = 0.911$). There was a significant difference for PA_r between the captive broodstock populations ($P < 0.001$), the natural stocks ($P < 0.05$) and between the broodstock populations and natural stocks ($P < 0.001$). F_{is} was not significantly different from zero for any captive populations or natural stocks. There were significant deviations from HWE estimates at four loci ($P < 0.05$) for a selection of the captive populations; *Lca070* in the NT broodstock population, *LcaM040* in the WA population, *Lca058* for QLD3 and WA and *Lca074* for QLD1. MICRO-CHECKER detected null alleles at five loci; *LcaM16* for NT, *LcaM040* and *Lca058* for WA, *Lca069* for QLD3 and *Lca178* for QLD4. All three natural localities conformed to HW expectations and no null alleles were detected. Exact tests

Table 1 Measures of genetic diversity for eight captive broodstock ($N = 407$) and three natural barramundi localities ($N = 86$) based on 16 microsatellite DNA markers: Western Australia (WA), Northern Territory (NT), Queensland (QLD) and natural populations representing a western stock, St George Basin (STG), a central admixed region, Liverpool Creek (LVP) and an eastern stock, Burdekin River (BUR). Population size (N), average number of alleles (A), mean allelic richness (A_r) and private allelic richness (PA_r), mean observed (H_o) and expected (H_e) heterozygosity, and the average inbreeding coefficient (F_{is})

Sample group	*N*	*A*	A_r	PA_r	H_o	H_e	F_{is}
WA	48	4.2	3.18	0.07	0.469	0.497	0.069
NT	71	5.6	3.42	0.32	0.503	0.509	0.020
QLD1	58	4.3	3.11	0.05	0.514	0.491	−0.038
QLD2	14	3.9	3.35	0.20	0.513	0.537	0.082
QLD3	111	5.5	3.23	0.05	0.506	0.506	0.005
QLD4	80	4.4	3.19	0.05	0.513	0.518	0.016
QLD5	9	3.2	3.04	0.02	0.532	0.482	−0.042
QLD6	16	3.1	2.67	0.05	0.475	0.453	−0.014
All captive samples*		4.3	3.15	0.10	0.503	0.499	0.071
Western (STG)	30	4.5	3.38	0.53	0.536	0.528	0.001
Central (LVP)	32	4.1	3.15	0.22	0.498	0.515	0.049
Eastern (BUR)	24	4.6	3.46	0.16	0.581	0.549	−0.036
All wild samples*		4.4	3.33	0.30	0.538	0.531	0.074

*Overall average of parameters.

for the non-random association of alleles at different loci (linkage disequilibrium) revealed 6% of loci pairs from five broodstock populations (WA, NT, QLD1, 3 and 4) presented significant P values ($P < 0.001$), following sequential Bonferroni correction (Rice 1988). No loci pairs from the natural localities demonstrated significant linkage disequilibrium. No bottleneck signatures were detected for the SMM or TPM mutation models within both the captive populations or natural localities, and the allele frequency distribution tests remained in a normal L-shaped distribution (Luikart *et al.* 1998).

Population structure

Discriminant analysis of principal components analysis revealed three main genetic populations (Fig. 1); one including individuals sampled from the QLD populations, which also included Burdekin River samples (eastern stock), NT and WA broodstock were grouped with wild individuals from the western stock (St George Basin), while wild individuals from the central admixed region (Liverpool Creek) were not grouped with any other population. The F_{ST} for most pairwise hatchery comparisons, except for a selection of QLD comparisons was significantly greater than zero ($P < 0.05$), such that the F_{ST} across all hatcheries was high (0.071, $P < 0.001$). Departure of F_{ST} from zero was significant ($P < 0.05$) for the comparison between WA and NT broodstock populations ($F_{ST} = 0.036$), and the highest F_{ST} was between these two populations and the six hatcheries from QLD (F_{ST} ranged from 0.050 to 0.115).

There was little difference in the STRUCTURE output when incorporating either the 'no locprior' or 'locprior' models and as a result the output from 'no locprior' is presented. As determined by the Evanno *et al.* (2005) method, the most appropriate number of genetic populations (k) was two, although there was also a minor indication of $k = 3$ (Fig. 2). Only two genetic clusters are clear in the barplot, although levels of admixture were detected for some individuals within populations (Fig. 3a). The barplot shows that the majority of WA and NT broodstock were allocated to stock one (yellow bars) while the majority of broodstock from QLD populations represented stock two (blue bars). Natural populations indicating the western stock (St George Basin) and the central admixed region (Liverpool Creek) were allocated to stock one and the eastern stock (Burdekin River) to stock two. Discriminant analysis of principal components analysis divided the natural localities into three populations, two stocks and a region of admixture, and only when they were screened independently from the captive populations did STRUCTURE analysis also confirm $k = 3$ (Fig. 3b). Upon observing the average population threshold q values of the broodstock populations, QLD1, 2, 3, 4 and 6 were allocated to stock two ($q > 0.90$), while WA, NT and QLD5 were assigned to stock one (Table 2). Eighteen individuals were allocated to stock one from broodstock population QLD3 and a level of admixture was detected by q values <0.90 (stock one 0.17, stock two 0.83). Similarly, the NT broodstock population also had q values <0.90 (stock one 0.74, stock two 0.23) but was mostly assigned to stock one. No individuals were admixed within the WA broodstock, although 12 were assigned to stock two. One to four individuals

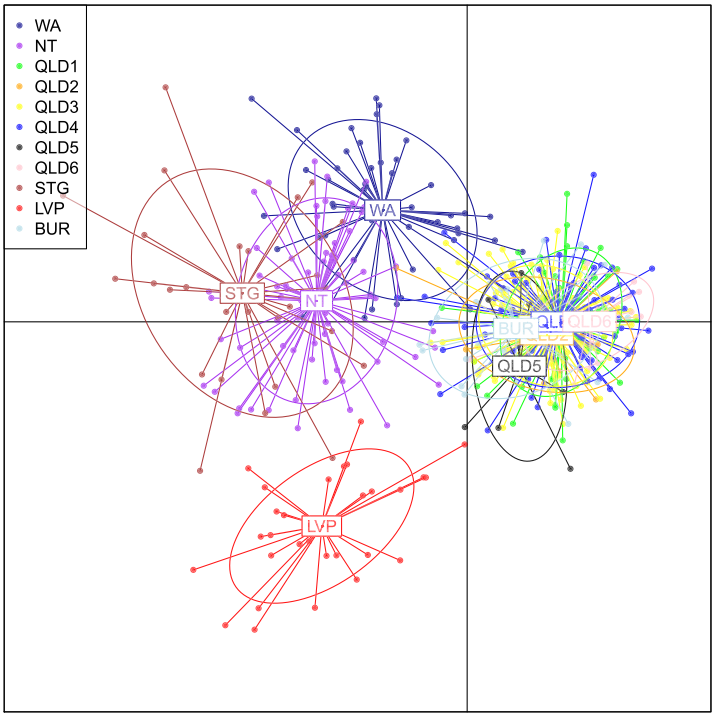


Figure 1 Scatterplots of the discriminant analysis of principal components (DAPC) for 407 individuals from eight *Lates calcarifer* broodstock populations and three natural localities ($N = 86$). Plots represent individual genotypes and colours represent populations. The first two principal components are represented by X and Y axes respectively.

were detected for admixture within five broodstock populations, however, admixture only accounted for 2.5% of total broodstock individuals.

Broodstock relatedness and relationships

Relatedness was estimated separately within each broodstock population and the average within population r_{QG} was 0.151 (Fig. 4). The lowest level of r_{QG} was recorded for QLD2 ($r_{QG} = 0.015 \pm 0.020$ SE) and the highest for QLD6 ($r_{QG} = 0.314 \pm 0.022$ SE). The average level of r_{QG} for the natural localities was estimated at 0.114. Following tests of statistical significance, the r_{QG} of QLD6 was significantly higher ($P < 0.05$) than every other population estimate of r_{QG}. The lowest percentage of FS relationships was estimated within the three natural localities (0.2–0.7%) and the highest was detected within broodstock population QLD5 (11.1%). The range of HS relationships within QLD broodstock populations were the highest and ranged from 6.9% to 88.9%. The proportion of dyads between broodstock populations estimated as having a FS relationship was no greater than 2.22% (between QLD4 and QLD5) (Table 3). The proportion of dyads with a HS relationship was greatest between all of the QLD populations, such as between QLD2 and QLD5 with 11.11%. Half-sib relationships were detected between broodstock populations and their putative natural localities, such as between all QLD broodstock populations and the natural population from the eastern stock (Burdekin River; ranging from 2.19% to 3.91%), between WA broodstock and the western stock (St George basin; 2.15%) and NT broodstock and the central admixed region (Liverpool Creek; 1.80%).

Pedigree reconstruction

Generating a kin partition (parameters; FSC = 0, W = 1, T = 10) rather than applying a full-sib constraint (FSC) resulted in the highest partition scores, ranging from 359 to 30,036 between each population (Appendix 1a). The number of kinship groups estimated for the captive broodstock populations ranged from 7 (WA) to 32 (NT). The WA broodstock population displayed the highest ratio of contribution (0.38). The ratio of contribution did not exceed 0.10 across the three natural localities and the number of estimated kin groups reached a total of 19, 23 and 20 for the western stock (St George Basin, $N = 30$), the central admixed region (Liverpool Creek, $N = 32$) and the eastern stock (Burdekin River, $N = 24$) respectively (Appendix 1b).

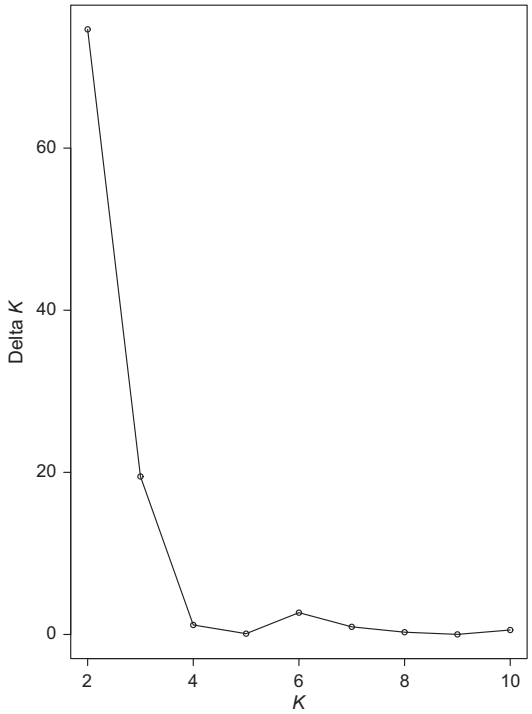


Figure 2 The most appropriate number of genetic groups (k), as determined by the Evanno *et al.* (2005) method, for eight captive barramundi populations (Nc = 407).

Effective population size

Utilizing the heterozygosity excess method, estimations of N_e could only be calculated for three broodstock populations and one natural locality. The remaining populations recorded infinite (∞) N_e estimates, indicating that the population was too large to calculate (Do *et al.* 2014), and as a result the N_e estimates are not presented. When implementing the LDNE method, N_e for the broodstock populations ranged from 3.7 to 44 and overall N_e for all captive broodstock was equal to 16.8 (Table 4). However, N_e estimates from broodstock populations with a small sample size ($N < 15$) such as QLD2 and QLD5 may not be reliable. An accurate measure of N_e could not be calculated for each of the three natural localities (infinite) and this may be attributed to the mixing of individuals between the two natural genetic stocks and the central admixed region (Wahlund effect, Hartl & Clark 1997).

Discussion

Overall, barramundi broodstock representing the captive breeding population of the Australian industry contained slightly lower levels of neutral genetic diversity (A_r = 3.15, PA_r = 0.10), when compared to the natural localities surveyed in this study (A_r = 3.33, PA_r = 0.30) (Table 1). In addition, relatedness estimates were higher within captive broodstock populations (r_{QG} = 0.151) when compared to natural localities (r_{QG} = 0.114) (Fig. 4). According to the results from Senanan, Pechsiri, Sonkaew, Na-Nakorn, Sean-In and Yashiro (2014) and Yue, Zhu, Lo, Wang, Lin, Feng, Pang, Li, Gong, Liu, Tan, Chou, Lim and Orban (2009), captive stocks from Southeast Asia displayed higher levels of allelic genetic diversity when compared to the Australian populations

Figure 3 (a) STRUCTURE barplot for eight captive barramundi populations (Nc = 407) and three natural localities (N = 86). The inferred number of populations (k) was equal to two; stock one represented in yellow and stock two in blue. Samples are separated by a black line and each bar represents one individual. (b) STRUCTURE barplot for three natural localities of barramundi; a western stock (St George Basin, N = 30), a central admixed region (Liverpool Creek, N = 32) and an eastern stock (Burdekin River, N = 24). The inferred number of populations (k) was equal to three.

Table 2 Assignment values from eight broodstock groups and three natural barramundi localities to stock one or two ($q > 0.90$) according to average admixture proportions in STRUCTURE. *N* represents the sample size of each population and the column of admixture represents individuals that recorded *q* values <0.90. The percentages of individuals assigned to the three clusters are displayed on the bottom row

Group	*N*	Stock one	Stock two	Admixture
WA	48	36	12	0
NT	71	67	0	4
QLD1	58	0	56	2
QLD2	14	1	13	0
QLD3	111	18	92	1
QLD4	80	3	75	2
QLD5	9	8	0	1
QLD6	16	0	16	0
Western (STG)	30	29	0	1
Central (LVP)	32	28	1	3
Eastern (BUR)	24	0	19	5
Total*	493	190	284	19
		38.5%	57.6%	3.9%

*The total count of individuals.

presented in this study. This could be due to differences in the evolutionary history of regional populations, sample size or the use of different genetic markers by our studies. Three of the loci used in our study were also utilized by Senanan *et al.* (2014), and 10 loci were used by Yue *et al.* (2009). Previous DNA barcoding analysis has revealed that Australian and Myanmar barramundi samples might represent two different species (Ward, Holmes & Yearsley 2008), and this could also account for the observed differences between this study and those of individuals from Southeast Asia.

Our results provide a foundation for the choice of individuals to source for inclusion into a selective breeding program. The results indicate that all eight existing captive broodstock populations contain some broodfish that do not share recent common ancestry to any other broodfish. This includes individuals WA_09, NT_50, QLD1_21, QLD2_11, QLD3_35, QLD4_21, QLD5_5 and QLD6_6 from each of the captive broodstock populations, which should be given high priority for inclusion in the base population. In total, 243 broodfish (WA = 27, NT = 31, QLD1 = 46, QLD2 = 8, QLD3 = 80, QLD4 = 46, QLD5 = 4, QLD6 = 1) should be excluded from the base population because they have a high likelihood of sharing recent common ancestry with several other fish. In total, 164 out of 407 captive fish were deemed suitable for inclusion based on low overall relatedness levels to other fish. This suggests that it could be possible to source broodstock for the base population from existing captive stocks, without any need to source wild individuals. The advantage of using existing captive stock is that these animals have already undergone some degree of domestication which can make them less stressed and more productive than their wild counterparts (Gjedrem 2000). However, our results demonstrate that current captive populations lack alleles from the central admixed region, and wild individuals from this area should also be included in the founding population. In addition, overall N_e for the captive populations is well under the desired number of individuals ($N_e > 100$), and from these results additional broodstock demonstrating a broad range of genetic diversity should be included in the founding population.

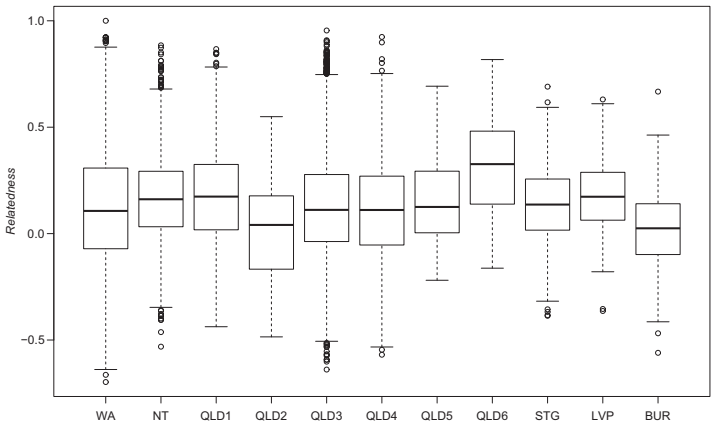


Figure 4 Relatedness estimates for eight captive broodstock populations and three natural barramundi localities as determined by the Queller and Goodnight (1989) estimator. Plots represent median values of relatedness and the box signifies upper and lower quartiles around the median. Data points outside of the vertical lines are classified as outliers.

Table 3 Proportion of dyads between eight captive broodstock groups and three natural localities of barramundi estimated as having a full-sib (above diagonal) and half-sib (below diagonal) relationship

	WA	NT	QLD1	QLD2	QLD3	QLD4	QLD5	QLD6	Western	Central	Eastern
WA		0.03	–	–	–	0.18	–	–	–	–	–
NT	1.76		–	–	–	–	–	–	0.05	0.04	–
QLD1	0.29	0.17		–	0.50	0.41	0.38	–	–	–	–
QLD2	–	0.40	5.42		0.39	0.80	0.79	1.34	–	–	–
QLD3	0.84	0.33	4.07	5.08		0.39	–	1.86	–	–	0.08
QLD4	1.48	0.63	5.28	5.98	2.27		2.22	–	–	0.04	0.05
QLD5	–	0.31	4.60	11.11	2.10	3.75		–	–	–	–
QLD6	0.13	–	1.19	5.80	10.70	0.55	–		–	–	–
Western	2.15	1.13	0.23	–	0.36	0.25	–	–		1.04	0.14
Central	1.04	1.80	0.32	0.22	0.17	0.43	–	0.20	–		–
Eastern	0.78	0.41	2.80	2.38	2.59	2.19	3.70	3.91	–	1.43	

Approximately 2.5% of all broodstock were identified as admixed stock (Table 2). These could be animals that have been directly sourced from sites in the natural environment where admixture naturally occurs or might be the result of crosses between the two major natural stocks (eastern and western stocks) that have occurred in captivity. From STRUCTURE analysis, broodstock allocated to stock one (such as the individuals from NT) demonstrated the highest levels of A_r and PA_r, with 3.42 and 0.32 respectively (Table 1). All but two broodstock populations contained at least one individual from stock one, which revealed that these individuals were either originally sourced from a common wild region, or in some cases there was an exchange of captive bred broodstock between hatcheries. The history of the sourcing and distribution of broodstock seems to have therefore resulted in a predominance of genetic variation from stock two and sourcing additional individuals that represent stock one and the central admixed region is recommended. Therefore, emphasis should be put on including animals for the base population who were assigned to stock one, such as individuals WA_07 and NT_01, and animals NT_07 and QLD1_20 who were unassigned to the two main stocks possibly as a result of recent admixture. Equalization of the representation of animals from these two stocks in the base population will ensure that variation in the base population is maximized.

Table 4 Estimates of the effective population size (N_e) for eight broodstock populations and three natural barramundi localities, implementing linkage disequilibrium (LDNE) with a minimum allele frequency of 0.05. *N* is the population size and CI is the 95% confidence interval range

Sample population	*N*	LDNE	
		N_e	95% CI
WA	48	3.7	3.2–5.1
NT	71	26.8	21.3–34.2
QLD1	58	22.0	17.4–28.0
QLD2	14	9.3	4.6–19.4
QLD3	111	7.4	6.0–8.8
QLD4	80	16	13.2–19.2
QLD5	9	44.0	8.0–∞
QLD6	16	7.0	2.9–15.4
All captive samples	407	16.8	15.0–18.8
Western (STG)	30	∞	120.2–∞
Central (LVP)	32	∞	102.3–∞
Eastern (BUR)	24	∞	67.4–∞
All wild samples	86	41.5	31.3–56.7

The genetic structure observed in the captive broodstock populations could be caused by a Wahlund effect. In this case, broodstock populations would be composed of animals sourced directly from two or more discrete wild stocks (Hartl & Clark 1997). This hypothesis is supported by the findings of previous natural barramundi population genetic studies that suggest the existence of marked stock structure (Keenan 1994, 2000; Chenoweth, Hughes, Keenan & Lavery 1998a,b; Doupé, Horwitz & Lymbery 1999). When developing breeding strategies for the selective breeding program, knowledge about the source of individuals used for breeding needs to be taken into account. The two main stocks detected from the STRUCTURE analysis must have evolved in isolation for some time for this genetic structure to exist. Accordingly, the inferred area of admixture

probably reflects secondary contact due to contemporary gene flow between the two stocks – a hypothesis previously proposed for barramundi in the Indo – West Pacific interface (Chenoweth *et al.* 1998b). It is possible that these stocks contain unique adaptive diversity, or have diverged in such a way that their performance in aquaculture as purebred or crossbred stock differs. Including individuals from the two stocks into a breeding population is expected to cover the suite of neutral genetic diversity detected in this study while accounting for the capture of locally adapted alleles. This strategy should promote adaptive potential in captive conditions, which are expected to quickly promote the selection of high fitness genotypes.

Decisions about which founding animals should be sourced to establish the base population will affect the ultimate success of the selective breeding program, as all subsequent generations will be derived from these animals. As the breeding program progresses, the emphasis put on different traits will vary and new (initially unexpected) traits may become evident. Genetic evaluations are made after the initial base stock has been chosen and the population has been closed to new entrants. It is not possible to make estimates of breeding values until the founding parents have been bred and their progeny tested in the aquaculture environment (Gjerde 2005a). Therefore, it is important to try to capture as much allelic variability as possible when the initial founders for the breeding program are chosen (as this is the only opportunity before the breeding population is closed), and it is important to maintain general levels of genetic variability while selective pressure is applied to traits of current importance (Gjerde 2005b). Accordingly, indications of relatedness and genetic diversity should be more widely used in the future as a guide for choosing animals to form the base population for selective breeding programs and for ongoing monitoring of levels of inbreeding and genetic drift.

Conclusion

This study has focused on captive barramundi broodstock and discusses ways in which individuals could be utilized to establish a genetically diverse and unrelated base population for selective breeding, which could also provide valuable information for other aquaculture candidate species. The eight broodstock populations sampled supply a large proportion of broodstock under production in Australia, and as a result the reported levels of neutral genetic diversity and relatedness were a thorough representation of those existing in the industry at the time of sampling. A high proportion of captive dyads tested in this study did not share any recent common ancestry, as such, it would be possible to utilize existing captive broodfish in a way that avoids inbreeding in the initial generations of the breeding program. We also discovered that measures of neutral genetic diversity of captive barramundi stocks were similar to levels within natural populations, reducing the necessity for including wild sourced broodstock into the base population. However, to further maximize levels of genetic diversity and N_e, plus reduce long-term inbreeding rates, it is recommended that a mixture of both captive bred and wild broodstock be included in the base population, particularly from the central admixed region. Further work to investigate genetic diversity and structure among the widespread natural populations across northern Australia and to simulate different scenarios for establishing the base population, is needed so that efficient strategies for capturing new genetic diversity for barramundi selective breeding can be devised.

The Australian industry has access to natural populations of barramundi spanning coastal and river systems from Western Australia to central Queensland, and additional broodstock should be sourced from the most genetically diverse of these natural populations. Therefore, we recommend that the industry composes the base population using captive individuals in a way that high A_r and low relatedness is achieved, and using wild individuals such that allelic variation of natural genetic stocks is well represented.

Acknowledgments

The authors would like to thank the Australian Seafood CRC, the Fisheries Research and Development Corporation and the Australian Barramundi Farmers Association for funding this study (project 2009/730). This research was conducted as part of S.L. PhD dissertation research (animal ethics approval project number E345), supported by an Australian Seafood CRC bursary and Flinders University AJ & IM Naylon PhD scholarship. All molecular work was carried out within the

Molecular Ecology and Evolution Laboratory (MEEL) at James Cook University and utilized multiplex marker conditions developed or modified from published conditions within the Aquaculture Genetics Research Group (Smith-Keune and Jerry). We are grateful to the hatcheries involved in this study, for allowing access to the samples and providing background information on the broodstock.

References

Adamkewicz S.L. & Harasewych M.G. (1996) Systematics and biogeography of the genus *Donax* (Bivalvia: Donacidae) in eastern North America. *American Malacological Bulletin* **13**, 97–103.

Ballou J.D. & Lacy R.C. (1995) Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: *Population Management for Survival and Recovery* (ed. by J.D. Ballou, M. Gilpin & T.J. Foose), pp. 76–111. Columbia University Press, New York, USA.

Blouin M.S. (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution* **18**, 503–511.

Chenoweth S.F., Hughes J.M., Keenan C.P. & Lavery S. (1998a) Concordance between dispersal and mitochondrial gene flow: isolation by distance in a tropical teleost, *Lates calcarifer*. *Heredity* **80**, 187–197.

Chenoweth S.F., Hughes J.M., Keenan C.P. & Lavery S. (1998b) When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. *Proceedings of the Royal Society of London B* **265**, 415–420.

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

DeWoody J.A. & Avise J.C. (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* **56**, 461–473.

Do C., Waples R.S., Peel D., Macbeth G.M., Tillett B.J. & Ovenden J.R. (2014) NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* **14**, 209–214.

Domingos J.A., Smith-Keune C., Robinson N., Loughnan S., Harrison P. & Jerry D.R. (2013) Heritability of harvest growth traits and genotype-environment interactions in barramundi, *Lates calcarifer* (Bloch). *Aquaculture* **402–403**, 66–75.

Doupé R.G., Horwitz P. & Lymbery A.J. (1999) Mitochondrial genealogy of Western Australian barramundi: applications of inbreeding coefficients and coalescent analysis for separating temporal population processes. *Journal of Fish Biology* **54**, 1197–1209.

Doyle R.W., Perez-Enriquez R., Takagi M. & Taniguchi N. (2001) Selective recovery of founder genetic diversity in aquacultural broodstocks and captive, endangered fish populations. *Genetica* **111**, 291–304.

Earl D.A. & vanHoldt B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359–361.

Eknath A.E., Bentsen H.B., Ponzoni R.W., Rye M., Nguyen N.H., Thodesen J. & Gjerde B. (2007) Genetic improvement of farmed tilapias: composition and genetic parameters of a synthetic base population of *Oreochromis niloticus* for selective breeding. *Aquaculture* **273**, 1–14.

Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Falconer D.S. & Mackay T.F.C. (1996) Selection: I. The response and its prediction. In: *Introduction to Quantitative Genetics Ed 4* (ed. by D.S. Falconer & T.F.C. Mackay), pp. 184–204. Pearson Education Limited, Harlow, Essex, UK.

Falush D., Stephens M. & Pritchard J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587.

Fjalestad K.T. (2005) Breeding strategies. In: *Selection and Breeding Programs in Aquaculture* (ed. by T. Gjedrem), pp. 145–158. Springer, Dordrecht, the Netherlands.

Frankham R., Ballou J.D. & Briscoe D.A. (2002) Loss of genetic diversity in small populations. In: *Introduction to Conservation Genetics* (ed. by R. Frankham, J.D. Ballou & D.A. Briscoe), pp. 227–253. University Press, Cambridge, UK.

Frost L.A., Evans B.S. & Jerry D.R. (2006) Loss of genetic diversity due to hatchery culture practices in barramundi (*Lates calcarifer*). *Aquaculture* **261**, 1056–1064.

Gilbert K.J., Andrew R.L., Bock D.G., Franklin M.T., Kane N.C., Moore J.-S., Moyers B.T., Renaut S., Rennison D.J., Veen T. & Vines T.H. (2012) Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology* **21**, 4925–4930.

Gjedrem T. (2000) Genetic improvement of cold-water fish species. *Aquaculture Research* **31**, 25–33.

Gjedrem T. & Robinson N. (2014) Advances by selective breeding for aquatic species: a review. *Agricultural Sciences* **5**, 1152–1158.

Gjedrem T., Gjøen H.M. & Gjerde B. (1991) Genetic origin of Norwegian farmed Atlantic salmon. *Aquaculture* **98**, 41–50.

Gjerde B. (2005a) Prediction of breeding values. In: *Selection and Breeding Programs in Aquaculture* (ed. by T.

Gjedrem), pp. 197–231. Springer, Dordrecht, the Netherlands.

Gjerde B. (2005b) Design of breeding programs. In: *Selection and Breeding Programs in Aquaculture* (ed. by T. Gjedrem), pp. 173–195. Springer, Dordrecht, the Netherlands.

Goudet J. (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Hartl D.L. & Clark A.G. (1997) Population substructure. In: *Principles of Population Genetics* (ed. by D.L. Hartl & A.G. Clark), pp. 111–162. Sinauer Associates, Sunderland, USA.

Herbinger C.M., O'Reilly P.T. & Verspoor E. (2006) Unravelling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. *Molecular Ecology* **15**, 2261–2275.

Jakobsson M. & Rosenberg N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806.

Jerry D.R. & Smith-Keune C. (2014) The Genetics of Asian Seabass. In: *Biology and Culture of Asian Seabass Lates calcarifer* (ed. by D. Jerry), pp. 137–177. CRC Press, New York, USA.

Jombart T. (2008) adegenet: an R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405.

Jombart T., Devillard S. & Balloux F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* **11**, 94.

Jones O.R. & Wang J. (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**, 551–555.

Kalinowski S.T. (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187–189.

Keenan C.P. (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**, 1123–1148.

Keenan C.P. (2000) Should we allow human-induced migration of the Indo-West Pacific fish, barramundi *Lates calcarifer* (Bloch) within Australia? *Aquaculture Research* **31**, 121–131.

Kumar S., Skjæveland Å., Orr R.J.S., Enger P., Ruden T., Mevik B.H., Burki F., Botnen A. & Shalchian-Tabrizi K. (2009) AIR: a batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* **10**, 357–363.

Li Q., Park C., Endo T. & Kijima A. (2004) Loss of genetic variation at microsatellite loci in hatchery strains of the Pacific abalone (*Haliotis discus hannai*). *Aquaculture* **235**, 207–222.

Loughnan S.R., Domingos J.A., Smith-Keune C., Forrester J.P., Jerry D.R., Beheregaray L.B. & Robinson N.A. (2013) Broodstock contribution after mass spawning and size grading in barramundi (*Lates calcarifer*, Bloch). *Aquaculture* **404–405**, 139–149.

Luikart G. & Cornuet J.-M. (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology* **12**, 228–237.

Luikart G., Allendorf F.W., Cornuet J.-M. & Sherwin W.B. (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *The Journal of Heredity* **89**, 238–247.

Nomura T. (2008) Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications* **1**, 462–474.

Palmer P.J., Blackshaw A.W. & Garrett R.N. (1993) Successful fertility experiments with cryopreserved spermatozoa of barramundi, *Lates calcarifer* (Bloch), using dimethylsulfoxide and glycerol as cryoprotectants. *Reproduction, Fertility and Development* **5**, 285–293.

Peakall R. & Smouse P.E. (2012) GenALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics Applications Note* **28**, 2537–2539.

Piry S., Luikart G. & Cornuet J.-M. (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *The Journal of Heredity* **90**, 502–503.

Porta J., Porta M., Martínez-Rodríguez G. & Alvarez M.C. (2006) Genetic structure and genetic relatedness of a hatchery stock of Senegal sole (*Solea senegalensis*) inferred by microsatellites. *Aquaculture* **251**, 46–55.

Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Queller D.C. & Goodnight K.F. (1989) Estimating relatedness using genetic markers. *Evolution* **43**, 258–275.

R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rice W.R. (1988) Analyzing tables of statistical tests. *Evolution* **43**, 223–225.

Rodzen J.A., Famula T.R. & May B. (2004) Estimation of parentage and relatedness in the polyploid white sturgeon (*Acipenser transmontanus*) using a dominant marker approach for duplicated microsatellite loci. *Aquaculture* **232**, 165–182.

Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137–138.

Rousset F. (2008) GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–106.

Schwartz T.S. & Beheregaray L.B. (2008) Using genotype simulations and Bayesian analyses to identify individuals of hybrid origin in Australian bass: lessons for fisheries management. *Journal of Fish Biology* **72**, 435–450.

Sekino M., Sugaya T., Hara M. & Taniguchi N. (2004) Relatedness inferred from microsatellite genotypes as a tool for broodstock management of Japanese flounder *Paralichthys olivaceus*. *Aquaculture* **233**, 163–172.

Senanan W., Pechsiri J., Sonkaew S., Na-Nakorn U., Sean-In N. & Yashiro R. (2014) Genetic relatedness and differentiation of hatchery populations of Asian seabass (*Lates calcarifer*) (Bloch, 1790) broodstock in Thailand inferred from microsatellite genetic markers. *Aquaculture Research* [Epub ahead of print].

Seutin G., White B.N. & Boag P.T. (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69**, 82–90.

Skirtun M., Sahlqvist P. & Viera S. (2013) *Australian Fisheries Statistics 2012*. FRDC project 2010/208. ABARES, Canberra, Australia. November, CC BY 3.0.

Sonesson A.K., Wooliams J.A. & Meuwissen T.H.E. (2005) Kinship, relationship and inbreeding. In: *Selection and Breeding Programs in Aquaculture* (ed. by T. Gjedrem), pp. 73–87. Springer, Dordrecht, the Netherlands.

Van de Casteele T., Galbusera P. & Matthysen E. (2001) A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology* **10**, 1539–1549.

Van Oosterhout C., Hutchinson W.F., Wills D.P.M. & Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538.

Wang J. (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11**, 141–145.

Wang S., Hard J.J. & Utter F. (2002) Salmonid inbreeding: a review. *Reviews in Fish Biology and Fisheries* **11**, 301–319.

Wang C.M., Zhu Z.Y., Lo L.C., Feng F., Lin G., Yang W.T., Li J. & Yue G.H. (2007) A microsatellite linkage map of Barramundi, *Lates calcarifer*. *Genetics* **175**, 907–915.

Wang C.M., Lo L.C., Zhu Z.Y., Lin G., Feng F., Li J., Yang W.T., Tan J., Chou R., Lim H.S., Orban L. & Yue G.H. (2008) Estimating reproductive success of brooders and heritability of growth traits in Asian sea bass (*Lates calcarifer*) using microsatellites. *Aquaculture Research* **39**, 1612–1619.

Waples R.S. & Do C. (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* **8**, 753–756.

Ward R.D., Holmes B.H. & Yearsley G.K. (2008) DNA barcoding reveals a likely second species of Asian sea bass (barramundi) (*Lates calcarifer*). *Journal of Fish Biology* **72**, 458–463.

Weir B.S. & Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Yue G., Li Y. & Orban L. (2001) Characterization of microsatellites in the IGF-2 and GH genes of Asian seabass (*Lates calcarifer*). *Marine Biotechnology* **3**, 1–3.

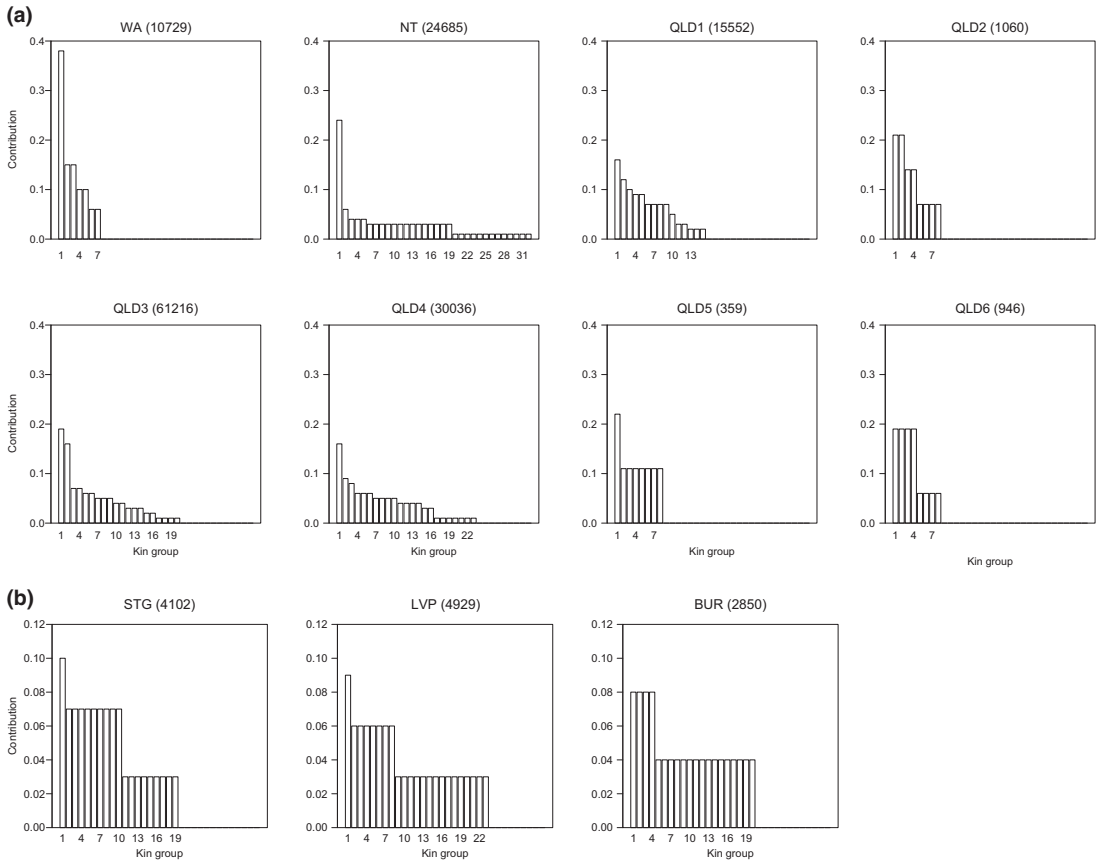
Yue G.H., Li Y., Chao T.M., Chou R. & Orban L. (2002) Novel microsatellites from Asian sea bass (*Lates calcarifer*) and their application to broodstock analysis. *Marine Biotechnology* **4**, 503–511.

Yue G.H., Zhu Z.Y., Lo L.C., Wang C.M., Lin G., Feng F., Pang H.Y., Li J., Gong P., Liu H.M., Tan J., Chou R., Lim H. & Orban L. (2009) Genetic variation and population structure of Asian seabass (*Lates calcarifer*) in the Asia-Pacific region. *Aquaculture* **293**, 22–28.

Zhu Z.Y., Lin G., Lo L.C., Xu Y.X., Feng F., Chou R. & Yue G.H. (2006a) Genetic analyses of Asian seabass stocks using novel polymorphic microsatellites. *Aquaculture* **256**, 167–173.

Zhu Z.Y., Wang C.M., Lo L.C., Feng F., Lin G. & Yue G.H. (2006b) Isolation, characterization, and linkage analyses of 74 novel microsatellites in Barramundi (*Lates calcarifer*). *Genome* **49**, 969–976.

Appendix



The allocation of barramundi individuals to kin groups from eight captive broodstock populations (1a) and three natural localities (1b), as reconstructed using PEDIGREE 2.2. Each vertical bar represents a kin group and values in parenthesis are partition scores.