**Levels of genetic diversity and inferred origins of *Penaeus vannamei* culture resources in China:
Implications for the production of a broad synthetic base population for genetic improvement**

Shengjie Ren^{a,*}, Peter B. Mather^b, Binguo Tang^c, David A. Hurwood^a

^a *Science and Engineering Faculty, Queensland University of Technology, QLD 4001, Australia*

^b *Australian Rivers Institute, Griffith University, QLD 4111, Australia*

^c *Beijing Shuishiji Biotechnology Co. Ltd., Daxing, Beijing 102600, China*

** Corresponding author: Shengjie Ren*

Faculty of Science and Engineering, Queensland University of Technology, 2 George St, Brisbane, QLD 4001, Australia. Tel: +61-(0)452610113, Email: shengjie.ren@hdr.qut.edu.au

Abstract

Knowledge about the origins and relative levels of genetic variation available in domesticated strains can be used to develop robust synthetic populations for use in stock improvement programs, in particular, where wild stocks are unavailable or the target species is alien. Here, we undertook a case study of Pacific white shrimp (*Penaeus vannamei*) where we utilised neutral molecular genetic variation data, population structure, and genetic relatedness as an initial step towards producing a base line for a long-term family-selection breeding program in China. The objective was to capture and maximise genetic variation in the base population while at the same-time to control inbreeding levels. Genetic diversity in 1162 individuals from 36 breeding lines sourced from 22 hatcheries with domesticated broodstock available in China was assessed using seven microsatellite loci. Genetic diversity levels in the sampled lines were considered to be healthy, with a mean number of alleles (A) 5.85, ranging from 4.00 in B4 to 12.43 in T1, and mean allelic richness (Ar) 4.46, ranging from 3.35 in B4 to 6.69 in T1 breeds. Mean inbreeding coefficient (F_{is}) was 0.07 among the 36 lines indicating acceptable levels of inbreeding. Genetic differentiation between lines from different hatcheries was moderate, with a mean F_{ST} estimate of 0.09 among all pairs of hatchery lines. Bayesian assignment and phylogenetic analyses indicated that hatchery lines in China could be broadly divided into four groups: i) a line from North America (NA1), originally sourced from Mexico (north Sinaloa) and Ecuador, ii) two representatives of the Kona line developed in Hawaii (Oceanic Institute) (NA2,3), iii) a line developed in Thailand (Southeast Asia 1 (SA1)), and iv) a mixed group containing stocks that originated in Latin America (predominantly from Panama and Colombia) (LA; SA2; CN1,2; NA4,5). From these data, we argue that a complete 4×4 diallel cross (16 crosses) of the four subpopulations should be trialled in China to develop a base line for a family selection breeding program with optimal levels of genetic variation and a relatively low inbreeding coefficient.

Keywords: Microsatellite, genetic diversity, population structure, prawn aquaculture, Bayesian assignment

1. Introduction

Capturing high levels of genetic variation in a base population in stock improvement programs is crucial for long-term success (Holtmark, et al., 2006; Holtmark, et al., 2008a; b). Achieving this however, can be problematical in places where the cultured species is not native and it is difficult to source wild germplasm. High levels of genetic variation is important because genetic gains, when selecting for specific traits, will depend on the levels of additive genetic variance present in the base population, the selection intensity applied and the relative level of heritability of the target traits (Falconer and Mackay, 1996; Gjedrem and Robinson, 2014). A number of well-organized breeding programs for aquatic species have ultimately failed because exploitable levels of genetic variation in the base population were low (Huang and Liao, 1990; Teichert-Coddington and Smitherman, 1988). In contrast, lessons learned in two well documented highly successful aquatic breeding programs that provide models for other species i.e. Atlantic salmon in Norway (Gjedrem, et al., 1991) and the Genetic Improvement of Farmed Tilapia (GIFT) (Eknath, et al., 1998) in Asia included both wild and domesticated strains when base populations were established initially. The importance of combining exotic and 'domestic' strains can contribute to a base population is clear in the GIFT case where change in growth rate is impressive with an average genetic gain of 10% per generation when compared with a control line (Hamzah, et al., 2014; Nguyen, 2016). Estimates of heritability for body weight ($h^2=0.28$) in the latest data (2012) indicate that exploitable genetic variation remains high in GIFT strains so genetic gains from the family selection approach should be ongoing (Nguyen, 2016). In general, results from incorporating domestic strains in aquatic animal stock improvement programs suggest that this approach is productive and can address shortages of high quality broodstock in non-native finfish aquaculture. Currently, farming of exotic aquatic species accounts for a significant component of aquaculture across the world and has contributed significant economic and great societal benefits, in particular in the Asia-Pacific region (Silva, et al., 2009; Singh and Lakra, 2011).

There can be a number of advantages associated with using domesticated strains to produce a base population in selective breeding programs of alien species. For example, much of the worldwide production of farmed Pacific white shrimp (*Penaeus vannamei*) and tilapia occurs in areas outside the natural geographic ranges of both species that make availability of wild strains challenging due to geographical distance, quarantine requirements for introductions and costs associated with sampling and transportation. Local domestic strains, where available, can provide a much easier and lower cost option for use in stock improvement programs. Secondly, domesticated strains have often accumulated favorable traits for artificial culture environments over many years that make them easier to handle and breed compared with wild strains (Olesen, et al., 2015). When developing a base population,

use of domestic or genetically improved strains, rather than wild animals as the starting resource, can provide the new breeding line with a competitive start, in particular for growth performance (Fernández, et al., 2017).

Farmed shrimp is a major commodity in the world seafood market, with annual production of more than 4.3 million tonnes worth more than \$USD 22 billion (FAO, 2016; Fernández de Alaiza García Madrigal, et al., 2017). Pacific white shrimp (*Penaeus vannamei*) that is native to southern North America, Central America and northern South America, now plays a substantial role in shrimp farming worldwide. In the Asia-Pacific region (where approximately 87% of global farmed shrimp is now produced), this exotic species contributed 78% to total production in 2014, with a biomass of 3.0 million tonnes. This represents the largest relocation of a single species in the planet's history (Saumena, 2015; Walker and Mohan, 2009). Pacific shrimp farming in Asia is now threatened however, by the quality and availability of seed for the industry. Almost all genetically improved broodstock available in this region come from Specific Pathogen Free (SPF) lines imported from the western hemisphere; current annual demand for *P. vannamei* SPF broodstock is approximately 1 million pairs (Giltterle and Diener, 2014). While SPF seed show improved growth rate performance, relatively low survival rates and high disease susceptibility are serious problems that have developed in recent times (Cock, et al., 2017; Cock, et al., 2009). Most SPF stocks are susceptible to many local pathogens and diseases (e.g. white spot disease (WSD), early mortality syndrome (EMS)) (Thitamadee, et al., 2016), and this problem has become a significant issue for the industry, particularly for the Asia-Pacific shrimp culture industry (Cock, et al., 2017; Moss, et al., 2012). One way to address these issues is to develop locally adapted breeding lines that can be subjected to selection for disease resistance, improved growth rate, etc. The first step in setting up these lines is to survey available genetic variation (GV).

Molecular markers (e.g. microsatellites) can provide effective tools for evaluating neutral genetic diversity levels, population structure and relatedness and kinship in domestic animal stocks. These characteristics have implications for both domesticated stock conservation and management of breeding programs (Carvalho, et al., 2015; Glowatzki - Mullis, et al., 2009; Loughnan, et al., 2016; Revidatti, et al., 2014; Wilkinson, et al., 2011; Wilkinson, et al., 2012). Applying information about relatedness among broodstock in artificial mating designs for stock improvement can provide an empirically based solution to reducing inbreeding levels when the initial generation is developed (Porta, et al., 2006; Rodzen, et al., 2004; Sekino, et al., 2004). This approach is highly relevant for genetic improvement in aquaculture, particularly when alien aquatic species are targeted. This is because there can be critical issues associated with high inbreeding levels in alien species due to difficulties

associated with sourcing new and genetically diverse broodstock. But more importantly, the high fecundity inherent in many cultured aquatic animals means that only a few broodstock are required to produce sufficient offspring to meet demand. The issues of inbreeding are twofold: i) inbreeding can lead to a decrease in fitness (inbreeding depression) through a reduction in heterozygosity and an associated increase in the expression of recessive deleterious alleles, and ii) low levels of genetic diversity can result in a poor response to selection (Li, et al., 2004; Schwartz and Beheregaray, 2008).

Despite the important role of *P. vannamei* in the shrimp farming industry, studies on genetic background of both global domesticated broodstock and wild populations are lacking or have only been limited in scope (Maggioni, et al., 2013; Mendoza-Cano, et al., 2013; Perez-Enriquez, et al., 2009; Valles-Jimenez, et al., 2004; Zhang, et al., 2014). The aim of the current study therefore, was to characterise genetic diversity levels and relationships among domesticated lines of the exotic farmed penaeid (*P. vannamei*) in China based on microsatellite DNA markers. The data generated will be used to inform decisions about the best approach to establishing a founding stock to capture the highest levels of genetic variation possible while controlling for inbreeding level from domesticated lines currently available in country.

2. Materials and methods

2.1 Sampling

Twenty-two commercial *P. vannamei* hatcheries from three provinces in southern China (Fujian, Guangzhou, and Hainan) were sampled to assess levels of genetic diversity, population structure and relatedness (Table 1). Standards for inclusion of specific hatcheries were as follows: (1) sampling from at least 10 hatcheries that were ranked among the top 20 performing *P. vannamei* hatcheries determined by the China Aquatic Products Processing and Marketing Association (CAPPMA) and where possible, (2) to sample broodstock with known different origins. The 22 hatcheries are identified but referred to anonymously as A to V (Table 1) to ensure that agreed confidentiality was addressed appropriately. Provenance of culture lines used to generate the original broodstock in each hatchery can be attributed to four geographical regions: i) North America (NA), ii) Latin America (LA), iii) Southern Asia (SA), and iv) China (CN). If multiple breeding lines have come from the same region, a unique number (1 to 5) was added after the name of the source region (Table 1). Details information of sources were provided on Table 1. Up to 50 individuals were sampled per line, with equal contribution by sex.

In total, 1162 individuals from 36 breeding lines from 22 hatcheries were sampled. Pleopods were taken from each individual and the tissue preserved in 95% ethanol prior to DNA extraction and genotyping.

Table 1. *Penaeus vannamei* sample information

2.2 DNA extraction and genotyping

Total genomic DNA was extracted from pleopod samples using a Marine Animal DNA Extraction Kit (TIANGEN BIOTECH CO. LTD.) following the manufacturers specified protocols. Seven microsatellite loci developed for the target species: *Lvan05* (Perez-Enriquez, et al., 2009), *Pvan1758* (Cruz, et al., 2002), and *TUMXLv5.45*, *TUMXLv7.121*, *TUMXLv8.256*, *TUMXLv9.103*, *TUMXLv10.312* (Meehan, et al., 2003) were used in the genetic analysis. Amplification volume was 15 μ l with the following conditions: 1 $\times$ Buffer I (Invitrogen), 0.2 mM dNTP (Invitrogen), 2.5 μ M of each primer (forward primers of seven loci were labelled fluorescently with HEX or FAM), 0.75 U Taq DNA polymerase (Invitrogen) and milliQ water to a final volume of 15 μ l. PCR cycling parameters included initial denaturing step of 95°C, then 35 cycles of 94°C for 30 s, specific annealing temperature for 45 s, 72°C for 1 min, and final extension at 72°C for 10 min. Genotyping of PCR products was undertaken on an ABI 3730xl genetic analysis system (Applied Biosystems) and allele size was generated using GeneMapper 4.0 software (Applied Biosystems) with a GenScan ROX-500 (Applied Biosystems) internal size standard.

2.3 Data analysis

2.3.1 Genetic diversity estimates

Number of alleles (*A*), expected heterozygosity (*He*), and observed heterozygosity (*Ho*) were estimated in GENALEX 6.502 (Peakall and Smouse, 2012). Allelic richness (*Ar*) and private allele richness (*PAr*) were calculated in HP-RARE 1.1 (Kalinowski, 2005), with parameters set for a minimum of 14 alleles per sample (7 diploid individuals). *Ar* is a weighted estimate of the number of alleles at a locus, independent of sample size, by applying a rarefaction process. *PAr* is a related estimator of the number of unique or rare alleles within a stock. Significant differences among genetic diversity measures for each breeding line were determined using a random complete block ANOVA that recognizes the inherent lack of independence of each microsatellite locus (block) between strains. Analyses were performed for mean number of alleles across loci ($\bar{A}$), and mean allelic richness ($\bar{Ar}$ – the number of alleles detected at a locus standardized on the number of individuals genotyped across all samples). *PAr* was not tested due to a large zero count. Where significant differences were detected, a

Tukey's HSD *posthoc* test was used to identify specific differences. Inbreeding coefficient (F) was estimated via three methods. Firstly, the widely used F_{is} index of the degree of random mating within populations was estimated with FSTAT 2.9.3.2 (Goudet, 2001) using the Weir and Cockerham method (Weir and Cockerham, 1984). A recent simulation has shown however, that F_{is} can underestimate real inbreeding levels for domestic breeding animals compared with a more recent Wang's TrioML (F_t) method (Doyle, 2016; Wang, 2014). Thus, a likelihood method that uses genotypes from a triad of individuals was used to estimate the inbreeding coefficient (F_t) (Wang, 2007; 2011). Finally, a third inbreeding coefficient index (DyadML (F_d) (Milligan, 2003)) was estimated to assess the correlation of F_{is} and F_t values. Deviations from Hardy-Weinberg (HWE) and Linkage Equilibrium (LE) were assessed using GENEPOP 4.1 (Rousset, 2008), applying a Markov Chain method with 10,000 dememorization steps followed by 20 batches (100 batches for LE) of 5000 iterations per batch. Sequential Bonferroni was used to adjust analysis-wide tests for the significance of HWE and LE (Rice, 1989).

2.3.2 Population genetic differentiation

Genetic differences among 36 stocks from the 22 sampled hatcheries were estimated using several methods. Firstly, pairwise F_{ST} (Weir and Cockerham, 1984), a statistic that partitions genetic diversity within and among each stock was calculated in ARLEQUIN 3.5.2.2 (Excoffier and Lischer, 2010). Also, a pairwise matrix assessing allele frequency heterogeneity among samples was constructed using Nei's D_A (Nei, et al., 1983), calculated with Populations 1.2.31 software (Langella, 1999). The PHYLIP program (Plotree and Plotgram, 1989) was used to construct an unrooted neighbour-joining cladogram based on the D_A matrix. Results were entered into FigTree 1.4.3 (Rambaut, 2012) to produce a high-quality phylogenetic tree representation. Grouping of individuals was performed using a Bayesian clustering method executed in the STRUCTURE 2.3.4 software (Pritchard, et al., 2000). Individuals were assigned to the most probable cluster out of k putative clusters with and without the use of sample location as a prior reference ('locprior'). Parameter settings included admixture and correlated allele frequencies (Falush, et al., 2003). For each k (1-35), 20 replicates were run, applying set parameters of 1,000,000 iterations of a Markov Chain Monte Carlo (MCMC) process with a burn-in length of 100,000 iterations (Gilbert, et al., 2012). Threshold q -value set for stock cluster was >0.9 for a single cluster and <0.9 for the detection of admixture. The most likely number of genetic populations (k) was estimated with STRUCTURE HARVESTER (Earl, 2012; Evanno, et al., 2005). Admixture proportions of each stock applying 20 replicates were averaged via CLUMPP 1.1.2 for the best k (Jakobsson and Rosenberg, 2007) and final barplots were generated with DISTRUCT 1.1 (Rosenberg, 2004).

2.3.3 Relatedness estimates

Estimated relatedness (r_{xy}) among stocks (36 stocks) and average r_{xy} of randomised mating between populations were estimated using the estimator, r_{QG} (Queller and Goodnight, 1989). r_{QG} is a widely used index of relatedness used for kinship studies of both captive and natural populations (Blouin, 2003). In the current study, 1139 individuals (genotyped with a minimum of four microsatellite loci) were used to generate a relatedness estimates among stocks. In addition, a moment estimator of pairwise relatedness from LynchRd (r_{LR}) (Lynch and Ritland, 1999) was calculated for comparison with the r_{QG} method. All relatedness estimators were generated with COANCESTRY 1.0.1.7 software (Wang, 2011). This software incorporates seven relatedness and four inbreeding estimators, to generate the most accurate estimates from multi-locus genotype data. Irrespective of the most appropriate relatedness estimator, accuracy depends on the dataset in each study; i.e. the number of microsatellite loci screened and the relative polymorphism information content (PIC) of the markers employed (Van de Castele, et al., 2001; Wang, 2011).

2.3.4 Effective population size

Effective population size (N_e) was estimated using a linkage mating model by assuming monogamous mating type as reported for both *P. vannamei* (Harris, et al., 2016) and other closely related penaeid species: *P. merguensis* (Knibb, et al., 2014) and *P. monodon* (Marsden, et al., 2013). A linkage mating model N_e from 36 hatchery populations was estimated in LDNE 1.31, applying a minimum allele frequency of 0.05 (Waples and Do, 2008). When considering the impacts of artificial mating design of populations for estimating N_e , a molecular co-ancestry method (N_{e2}) was used (Nomura, 2008). N_{e2} was estimated with NEESTIMATOR 2.01 that is considered to be a powerful method when there is missing data as it screens out rare alleles, and it is designed for large data sets (Do, et al., 2014).

3. Results

3.1 Genetic diversity and HWE estimates

Mean number of alleles ($\bar{A}$) per locus in the sampled stocks ranged from 4.00 in B4 to 12.43 in T1, with an overall average across all sampled 36 stocks of 5.85 (Table 2). A similar ranking pattern was observed for allele richness estimates ($\bar{A}r$) compared with $\bar{A}$, with a highest value of 6.69 in T1 to the lowest value of 3.35 in B4, and an overall average of 4.46 (Table 2). Highest private allele richness (PAr) (0.48) was observed in the V1 stocks, followed by T1 (0.39) and Q2 (0.28) and stocks from NA1, NA2 and SA1 possessed significantly lower

values of PAr except for the K2 stocks. Expected heterozygosity (He) across all sampled loci ranged from 0.58 in B4 to 0.81 in T1, while observed heterozygosity (Ho) ranged from 0.47 in B4 to 0.77 in T1 (Table 2). Mean estimates of He and Ho across all loci and stocks were 0.69 and 0.66, respectively.

Table 2. Genetic diversity measures for 36 batches *P. vannamei* broodstock

While the inbreeding coefficient estimate (Fis) was low at 0.07 in all stocks, this value was 3-fold greater where the other two calculation methods were applied, with 0.21 (Ft) and 0.22 (Fd), respectively. The RCB ANOVA identified significant differences among genetic diversity estimates for the 36 breeding lines (for A , $F_{35, 210} = 7.448$, $p < 0.001$; Ar , $F_{35, 210} = 3.738$, $p < 0.001$). In general, the Tukey's *posthoc* test demonstrated that breeding lines derived from CN2, NA4 and SA2 possessed the highest diversity (see Figure 1 for A , Ar (not presented) showed a similar pattern).

Figure 1: A bar plot showing mean number of alleles ($\bar{A}$) for 36 breeding lines. Error bars represent ± 1 SD. Letters above bars indicate where significant differences among breeding lines exist determined from Tukey's HSD test ($P < 0.05$). A1-V1 are 36 breeding lines ID from 22 hatcheries in China; and NA1-NA5, SA1, SA2, LA, CN1, CN2 are 10 group ID of the origins history of the above 36 breeding lines (details information see Table 1.).

Among 252 HWE tests for 7 loci among 36 stocks, only 3 tests showed significant deviation ($P < 0.05$) from HWE after standard Bonferroni correction. Among these three significant deviations, there was no obvious pattern for specific loci or population. After Bonferroni correction, 6 out 756 pairs of loci were found to deviate significantly from LE ($P < 0.05$), but again, no specific patterns were detected. As such, all loci were considered to be reliable, independent estimators of neutral diversity in each of the 36 breeding lines.

3.2 Population Genetic Differentiation

Genetic differentiation among stocks (pairwise F_{ST} estimates) from lowest to highest, ranged from 0.00 to 0.29. Mean F_{ST} estimates among the 36 stocks were moderate and ranged from 0.06 (D1 and I1 stocks) to 0.14 (B4

stocks). Overall mean genetic differentiation (F_{ST}) among stocks was 0.09. Genetic differentiation (F_{ST}) of stocks from different hatcheries that had originated from the same geographical region historically was lower compared with the overall F_{ST} , with a mean value of 0.05 within NA1, 0.01 within NA2, 0.03 within SA1, 0.00 within NA3 and NA4 (Table 3). Among 78 pairwise F_{ST} comparisons of 13 breeding lines within NA1 group, 53 tests were statistically significant, which indicated most breeding lines tested within the NA1 group were diverged. While all 28 pairwise F_{ST} tests within NA2 group were not significant, 13 pairwise F_{ST} estimates among 15 tests within SA1 group were also not significant. In addition, no significant differences were observed for pairwise F_{ST} comparisons within NA3, or NA4 respectively. It is interesting to be noted that 37 of 40 pairwise F_{ST} tests were not significant between breeding lines in China that originated from NA2, NA3 and NA4, suggesting a close genetic relationship among NA2, NA3, and NA4. Pairwise genetic distance estimates (D_A) ranged from 0.03 (I1 vs. J1) to 0.56 (B4 vs. R1). Average pairwise genetic distance estimates ranged from 0.17 for I1 stocks to 0.41 for R1 stocks (Table 3). Pairwise genetic distance estimates between hatcheries that had originated from the same geographical region historically also were lower compared with the overall average D_A (0.23), with average D_A 0.12 between NA1, 0.11 between NA2, 0.15 between SA1, 0.06 between NA3 and 0.09 between NA4.

Table 3. Population genetic differentiation amongst 36 *P. vannamei* stocks

A phylogenetic reconstruction of stock (Hatchery) relationships is represented in Figure 2. Stocks with a NA1 origin clustered to form a monophyletic clade, as did all stocks in the SA1 group. NA2 and NA3 stocks together formed a third clade, consistent with pairwise F_{ST} comparisons showing that NA2 and NA3 shared a close genetic relationship. Seven breeding lines originating from NA4, NA5, LA, SA2, CN1 and CN2 loosely clustered together to form a fourth, polyphyletic clade. However, due to a lack of multiple lines being available from the various source populations represented in this clade, we cannot discard the possibility that each of these lineages is genetically distinct and could potentially be readily identified through multi-locus analysis.

Figure 2: An unrooted neighbour joining tree for 36 *P. vannamei* breeding lines based on seven microsatellite loci using the Nei's D_A genetic distance method. Each breeding line label includes the abbreviations of origins and hatchery ID as per Table 1.

The most appropriate number of clusters applying the Bayesian method in STRUCTURE was, $k=2$ even though $k=4$ was also strongly supported (see Supplementary Figure S1). With $k=2$, all stocks that originated from NA1 clustered into a single group with stocks from LA forming a separate cluster. The other 22 stocks showed an admixture pattern after Bayesian clustering (Figure 3_a). When $k=4$, all stocks from NA1 and LA clustered into two discrete genetic populations, (Figure 3_b) but stocks from NA2 separated to form a separate cluster. Despite a pattern of minor admixture for some stocks from SA1, all SA1 stocks formed a single genetic cluster (Figure 3_b). Remaining stocks showed an admixture pattern. Some individuals from D2 stocks showed a different Bayesian cluster assignment within populations that suggested that they may constitute a mix of different genetic resources, historically. Overall, the Bayesian cluster results were consistent with the population structure analysis based on F_{ST} and the reconstructed phylogenetic tree.

Figure 3: Individual assignment based on Bayesian analysis for 36 breeding lines at: a) Structure plot for $K=2$; b) Structure plot for $K=4$.

3.3 Relatedness estimates

There was a high correlation between results from different relatedness estimate methods ($r=0.85$). Average relatedness estimates using R_{QG} within stocks ranged from 0.02 (R1 and T1) to 0.32 (B4), with an average value of 0.16. In general, average relatedness values using R_{DML} were larger than for R_{QG} , and ranged from 0.12 (T1) to 0.4 (B4) with average estimate of 0.26. Significantly higher values were found for relatedness comparisons within groups than between different groups (Table 4).

Table 4. Average relatedness estimates amongst 36 *P. vannamei* stock

3.4 Effective population size (N_e)

Based on the linkage model, effective population size (N_{el}) estimates ranged from 19.9 to 749.6, with four populations recorded as infinite. Applying the molecular co-ancestry method (N_{e2}), most stocks showed

estimates (N_{e2}) that were much lower than for N_{e1} and they ranged from 5.0 to 68.6. Nine populations for N_{e2} were infinite, indicating that their N_e size was very large (Table 2). Infinite N_e estimates generated with both N_{e1} or N_{e2} estimators may result from either mixing of stocks from different genetic stocks in history (Hartl and Clark, 1997) or analysis of small sample sizes per line. A parallel result has been reported for N_e estimates in captive barramundi stocks where lines with small sample sizes were estimated as infinite. (Loughnan, et al., 2016).

4. Discussion

4.1 Genetic variation levels within and among stocks

The current project provides the most comprehensive analysis available for the status of genetic diversity levels and inferred genetic origins of farmed *P. vannamei* broodstocks in China. The results presented here can also be used as a reference and foundation for assessing genetic resources of *P. vannamei* stocks in other countries in Asia where this species is becoming the major farmed penaeid.

Overall, levels of genetic diversity in 36 farm stocks assessed here were similar compared with another earlier, but less comprehensive study on genetic diversity in *P. vannamei* stocks in China (Zhang, et al., 2014). Minimal differences observed between results obtained in the two studies most likely reflect the different panel of microsatellite loci used to assess diversity in the stocks examined. In general, genetic diversity levels in farm stocks of *P. vannamei* in China were slightly higher when compared with stocks farmed in Mexico (Perez-Enriquez, et al., 2009; Vela Avitúa, et al., 2013), consistent with the fact that genetic resources in Mexican hatcheries are believed to have come from a single origin (i.e. the ‘Los Malago’ breeding line that forms the basis of the industry there (Perez-Enriquez, et al., 2009)). In contrast in China, farm stocks have been developed from a number of different sources that originated from different geographical locations (from both from the native and introduced ranges) before they were introduced to China.

Among the breeding lines examined here, estimates of genetic diversity in stocks that originated from NA1, NA2 and SA1 sources in general were lower compared with others (Table 2). Potentially, this difference may reflect the longer time span they have spent managed as closed breeding lines (Benzie, 2009) or differences in relative levels of genetic variation present in the initial founders for each breeding program. It is interesting to note that genetic diversity levels in the CN2 stocks were much higher than in other Chinese stocks tested here (Table 2). Broodstock management practices and strategies employed in the CN2 hatchery are different from many others in China. First, the number of broodstock maintained in the CN2 hatchery was significantly higher

than in other hatcheries, with approximately 15,000 to 25,000 pairs maintained each year for PL production. Assuming that mating success approaches 10% per night during the nauplii supply season, this implies that 1,500 to 2,500 pairs likely contribute to each larval batch. Assuming that breeders are unrelated and contributions from pairs are approximately equal, this number of breeders can maintain a high effective population size and avoid genetic erosion of the population due to genetic drift effects. Each year the CN2 hatchery also obtains a small proportion of new stock from other hatcheries and thoroughly mixes their diversity into their own stocks, introducing new alleles to their breeding program.

Inbreeding levels (F_{is}) were similar to those reported in an earlier study of *P. vannamei* in China (Zhang, et al., 2014). These estimates suggest that most broodstocks currently used in hatcheries in China likely apply effective inbreeding management and follow the recommendations proposed by Moss, et al. (2007) who stated that inbreeding levels should not exceed 10%. Chinese commercial broodstocks show much lower inbreeding levels compared with those reported in farmed *P. vannamei* stocks in Mexico (F_{is} 0.25) (Perez-Enriquez, et al., 2009) and Brazil (F_{is} 0.30) (Maggioni, et al., 2013). However, it should be noted that while high homozygosity in these two studies may reflect significant inbreeding, deviations from HWE more likely suggests that homozygote excess was a function of null alleles (a common problem with microsatellite studies) and that inbreeding in these studies may have been overestimated. Inbreeding coefficient F_{is} estimates in a number of the Chinese stocks tested here however, were much higher than that reported in the stocks imported originally. A possible explanation for this observation is that while they probably received a representative gene pool initially, insufficient management of broodstock genetic resources subsequently over time may account for observed high F_{is} estimates in these stocks. Notably, F_t (mean F_t = 0.21) and F_d (mean F_d = 0.22) estimates for inbreeding level were three times higher than the F_{is} estimates (mean F_{is} = 0.07). Doyle (2016) argued that F_{is} can be a poor indicator of inbreeding level in stocks. In our study, F_d and F_t estimates were much higher than for the more widely used F_{is} estimates, a difference that could suggest a more serious inbreeding issue in Chinese stocks. We must be cautious however, in concluding that stocks are inbred based solely on F_d or F_t estimates, as maximum likelihood methods can overestimate the real value, particularly for stocks with low relationship estimates (Milligan, 2003). The essential point about estimating inbreeding level is that it will always be a relative estimate and not an absolute value, so there needs to be some comparison point rather than a discussion based solely on the inbreeding estimation method that is employed. For example, in order to acknowledge any real decline in genetic variation in captive populations of *P. vannamei*, ideally we would need a comparison point with a reference wild population (Vela Avitúa, et al., 2013). While this was not possible in the current study, for

the purpose of genetic monitoring of broodstocks of *P. vannamei* in closed rearing systems in China, we can obtain a relative assessment of the success or otherwise of management practices employed to manage genetic diversity over time by sampling successive broodstock cohorts over years in the same hatchery using the same marker loci (De Lima, et al., 2008). Based on F_{is} comparisons with that in Mexico (Perez-Enriquez, et al., 2009) and Brazil (De Lima, et al., 2008), there is indication that management of genetic diversity levels in China may have been better than in hatcheries elsewhere.

4.2 Population differentiation and origins of genetic resources

The estimate of mean level of genetic differentiation (F_{ST}) among Chinese *P. vannamei* culture stocks was 0.09 (Table 3) indicating a moderate degree of allelic heterogeneity among stocks. This result is consistent with those reported in previous studies in China (F_{ST} , 0.08) (Zhang, et al., 2014) and Brazil (F_{ST} , 0.06) (Maggioni, et al., 2013), but were higher than estimates reported for stocks in Mexico (F_{ST} , 0.02) (Perez-Enriquez, et al., 2009). Differences reflect that Mexican *P. vannamei* culture stocks have had a common ancestry, while stocks in Brazil and China most likely have resulted from mixing of multiple lineages following the importation of exotic germplasm because no native wild stocks were available in China.

Genetic differentiation estimates among the 36 Chinese stocks showed pairwise F_{ST} estimates ranging from 0.00 to 0.29 (Table 3). Stocks with the same inferred wild origin as expected showed the lowest pairwise F_{ST} estimates, (average of 0.05 among NA1, 0.01 among NA2, 0.03 among SA1, 0.00 between NA3 and 0.00 between NA4). Most pairwise F_{ST} tests within NA2 or SA1 groups (Table 3, shadowed areas) were low and not significant, which indicates that breeding lines from different hatcheries in China with the same origins (either from NA2 or SA1) show limited genetic differentiation. In contrast, breeding lines in China with NA1 origins show moderate levels of genetic heterogeneity. This potentially results from several independent nucleus breeding centres utilising NA1 stocks. Of interest, pairwise F_{ST} estimates between NA2, NA3 and NA4 were low, indicating a common ancestry for these groups. When the founding information for NA2 and NA3 was traced back, we established that both had been developed independently from the Kona line in Hawaii.

Population NA4 was sourced from a Hawaiian population but information was not available as to whether this was from the Kona line or other lines in Hawaii. While results of the Structure analysis here provides little confidence for a significant contribution from the Kona line in NA4, F_{ST} analysis shows relatively low levels of differentiation among NA2, 3 and 4. This suggests that while NA2 and NA3 had a common Kona line ancestry, NA4 resulted from mixing of Kona line genes with those from other lines.

In the present study, the unrooted neighbour-joining dendrogram is consistent with the historical records of *P. vannamei* culture stocks origins in China (Figure 2). Four main clades were distinguished: i) NA1; ii) NA2 and NA3; iii) SA1; and iv) a clade representing all other stocks. It is reasonable to suggest that NA2 and NA3 grouped together in a single clade because they were both derived from a common founding population (the Kona line developed in Hawaii). China has been the major *P. vannamei* producer for decades, and there have been many records of importations since 1987 (Li, et al., 2006). When we combine our results from the Bayesian cluster analysis with those from a review paper of global exchange of penaeid shrimps by Benzie (2009) and information from shrimp companies and research institutes in China regarding the origins of their founding populations, we can synthesize a more detailed account of founding genetic resources for each group.

Founding populations of NA1 came from Mexico (north Sinaloa) and from Ecuador, while the NA2 and NA3 lines originated from genetic resources sourced from the Kona line in Hawaii. A recent study on seven culture populations of *P. vannamei* in China (Zhang et al., 2014) indicated the two clusters in their phylogenetic reconstruction of genetic relationships align with our clades i and ii. Origins of SA1 are more obscure, but potentially may have had contributions from stocks in the USA and South America. According to the records, the LA stock was developed from multiple founding populations that were sourced from wild populations in South America. The official records for the founding population for CN1 suggest that they are the result of the mixing of several cultured genetic lines from the USA and a wild population from Ecuador. SA2 was developed from germplasm sourced from Hawaii and a Kentucky culture population, confirmed by its mixed ancestry in the Structure analysis. The founding population for NA5 came from Texas but originated from South America. CN2, the stock with the highest diversity, was developed from a mixture of Chinese culture populations that includes genetic resources possibly from many stocks introduced to China, but predominantly displaying a Latin American ancestry in the Structure analysis. While no official information was available about the origins of the NA4 population, the Structure analysis suggests that it likely had a base population developed from several lines from the USA and South America. The only available information about the origin of the NA4 line was that it was constructed from a combination of Hawaiian resources and a special specific pathogen resistant (SPR) line. According to a history of genetic improvement of *P. vannamei* culture stocks (Cock, et al., 2009), SPR resources originated from survivors of stocks in commercial ponds in South America sourced from Ecuador and Colombia in the mid-1990s when Taura syndrome virus (TSV) was prevalent in the shrimp industry. The broodstock generated from the survivors apparently possessed useful genetic variation for TSV resistance and this was applied in a family selection program in Hawaii.

In our study, all of the stocks that have South American origins (i.e. NA4, LA, CN2, SA2, & NA5), either as a uniform or admixed population following Bayesian clustering ($k=4$), showed significantly higher numbers of private alleles (Table 2). Given the very limited information available on the genetic background of virtually all stocks sampled here and the broad geographical distribution of natural populations of the target species, the markers appear to have traced the ancestry of the lines quite successfully. Of interest, assignment of individuals from the D2 stock (for either $k = 2$ or 4), suggested that it consisted of two distinct breeding lines that have been kept isolated from each other (i.e. there is no pattern of admixture between groups). While all individuals sampled for this study were collected at the same time from breeding tanks from this hatchery, broodstock used for nauplii production can however, be between 12 to 16 weeks old after eyestalk ablation and sometimes hatchery managers keep healthy females from previous hatchery runs and mix them with broodstock that are younger. Theoretically, a pattern of two different lines without admixture could occur if the sets of genotypes in the two age groups differ significantly by chance. In the current situation however, there are records of importation of two different lines that clearly assign to the NA1 and LA groups. It is obvious that broodstock from these two lines have been kept separate within the hatchery.

4.3 Implications for forming a genetic foundation (base) population for genetic improvement

Our results provide a strong foundation for making decisions about how to form a synthetic base population of an exotic aquaculture species, *P. vannamei* with high levels of genetic variation for a stock improvement program in China. Assuming that a classical complete diallel cross (Gjedrem and Robinson, 2014; Hung, et al., 2014; Nguyen, 2016) will be conducted to establish the base population, the first question will be how to identify discrete subpopulations of the available stocks. Results from the genetic differentiation analyses (pairwise F_{ST} , D_a , phylogenetic tree and Bayesian clustering) of the available domesticated stocks of *P. vannamei* indicate four discrete subpopulations, NA1, Kona line, SA1 and Latin America (LA) closed group. While at present there is no available information on additive genetic variance of commercially important traits for these groups in China, such as experiment tested that the Kona line is known to have been selected to improve growth performance but performed poorly for resistance against WSSV (Cuéllar-Anjel, et al., 2012). All of the four subpopulations described here however, have been used for shrimp farming in China, widely. We suggest that the best approach will be a complete 4×4 diallel cross with equal offspring contributed from the four subpopulations. This approach will capture not only high levels of genetic variation, as the stocks are successful domesticated lines they will also carry copies of different QTL alleles that have been selected indirectly, potentially also including alleles influencing growth and survival traits (Sae-Lim, et al., 2016) as an

outcome of past domestication and farming processes. Stochastic simulation has suggested that using four subpopulations to form the base population will maximise genetic gains in future generations of selection while saving sampling costs associated with using a greater number of lines in the base population (Holtmark, et al., 2006).

Relatedness estimates generated in the current study can provide a guideline to avoid inbreeding risk across the proposed generations of selection. Based on R_{QG} relatedness index results (Table 4), the proposed diallel cross mating design among stocks should theoretically maintain R_{QG} near zero because none of the four subpopulations identified for inclusion in the base population have shared a recent common ancestry as a result of many generations of isolation during their domestication process. While initially there had been potential for higher rates of inbreeding for the mating pairs within subpopulations, this can be addressed by sourcing stocks of the same groups from different hatcheries with highly genetic diversity levels for mating. By incorporating relatedness information, and clear pedigree records, any risk of significant inbreeding potentially leading to inbreeding depression and any negative impacts of genetic drift can be minimised.

Conclusions

Here we have documented genetic diversity levels, genetic differentiation among, and relatedness patterns of, *P. vannamei* culture lines currently available in China. Our results confirm that modern *P. vannamei* hatchery lines have come from multiple genetic sources and the genetic differentiation patterns described in hatchery lines in China are consistent with the historical records of introductions. The data described here provide a foundation and guidelines for both hatchery stock management and for developing a base population with high levels of genetic diversity for a future breed improvement program.

Acknowledgements

We thank Mr. Huixiao Zhang and Mr. Tao Lu for providing samples of various shrimp strains at Hainan hatchery. Also thanks to Mr. Abing Gao from the Shrimp Hatchery Association of Xiamen (SHAX) for supplying broodstock and for sharing his valuable knowledge of the history on the origins of *Penaeus vannamei* broodstock in Fujian Province and Mr. Jian Tan, manager of the hatchery at Zhangjiang, generously assisting with broodstock sampling at Guangdong Province and providing the history of breeding line origins of broodstock there. Thanks to the Queensland University of Technology (QUT) for providing a QUT Postgraduate Research Award (QUTPRA) to Shengjie Ren to help undertake this research.

References

- Benzie, J.A., 2009. Use and exchange of genetic resources of penaeid shrimps for food and aquaculture. *Reviews in Aquaculture*. 1, 232-250.
- Blouin, M.S., 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution*. 18, 503-511.
- Carvalho, G., Paiva, S., Araújo, A., Mariante, A., Blackburn, H., 2015. Genetic structure of goat breeds from Brazil and the United States: Implications for conservation and breeding programs. *Journal of Animal Science*. 93, 4629-4636.
- Cock, J., Salazar, M., Rye, M., 2017. Strategies for managing diseases in non-native shrimp populations. *Reviews in Aquaculture*. 9, 211–226.
- Cock, J., Gitterle, T., Salazar, M., Rye, M., 2009. Breeding for disease resistance of Penaeid shrimps. *Aquaculture*. 286, 1-11.
- Cruz, P., Mejia-Ruiz, C., Perez-Enriquez, R., Ibarra, A., 2002. Isolation and characterization of microsatellites in Pacific white shrimp *Penaeus (Litopenaeus) vannamei*. *Molecular Ecology Notes*. 2, 239-241.
- Cuéllar-Anjel, J., White-Noble, B., Schofield, P., Chamorro, R., & Lightner, D. V. (2012). Report of significant WSSV-resistance in the Pacific white shrimp, *Litopenaeus vannamei*, from a Panamanian breeding program. *Aquaculture*, 368-369, 36-39. DOI: 10.1016/j.aquaculture.2012.08.048
- De Lima, A.P.S., Dos Santos, A.C.L., Dantas, H.L., Maggioni, R., Coimbra, M.R.M., 2008. Genetic monitoring of broodstocks of the marine shrimp *Litopenaeus vannamei* in a closed rearing system in Pernambuco, Brazil. *Aquaculture research*. 39, 1461-1466.
- Do, C., Waples, R.S., Peel, D., Macbeth, G., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*. 14, 209-214.
- Doyle, R.W., 2016. Inbreeding and disease in tropical shrimp aquaculture: a reappraisal and caution. *Aquaculture Research*. 47, 21-35.
- Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. 4, 359-361.
- Eknath, A., Dey, M., Rye, M., Gjerde, B., Abella, T., Sevilleja, R., Tayamen, M., Reyes, R., Bentsen, H., 1998. Selective breeding of Nile tilapia for Asia, 6th World Congress on Genetics Applied to Livestock Production. University of New England Armidale, Australia, pp. 89-96.

- 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.
- Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 10, 564-567.
- Falconer, Mackay, 1996. *Introduction to quantitative genetics*. New York, NY: Pearson/Prentice Hall. 463, 464.
- 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.
- FAO, 2016. *Aquaculture Production (Quantities and values) 1950-2013*. FishStatJ - Software for fishery statistical time series, FAO Fisheries and Aquaculture Department. Rome.
- Fernández de Alaiza García Madrigal, R., da Silva, U.d.A.T., Tavares, C.P.d.S., Ballester, E.L.C., 2017. Use of native and non-native shrimp (Penaeidae, Dendrobranchiata) in world shrimp farming. *Reviews in Aquaculture*, DOI: 10.1111/raq.12206
- Fernández, J., Toro, M.Á., Sonesson, A.K., Villanueva, B., 2014. Optimizing the creation of base populations for aquaculture breeding programs using phenotypic and genomic data and its consequences on genetic progress. *Frontiers in Genetics*. 5, 414.
- 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., 2012. Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program structure. *Molecular Ecology*. 21, 4925-4930.
- Giltterle, T., Diener, J., 2014. Genetic improvement programs for shrimp vary between Asia, Americas. *Global Aquaculture Advocate*, 44-47.
- Gjedrem, T., Robinson, N., 2014. Advances by selective breeding for aquatic species: a review. *Agricultural Sciences*. 5, 1152.
- Gjedrem, T., Gjøen, H.M., Gjerde, B., 1991. Genetic origin of Norwegian farmed Atlantic salmon. *Aquaculture*. 98, 41-50.
- Glowatzki-Mullis, M.L., Muntwyler, J., Bäumle, E., Gaillard, C., 2009. Genetic diversity of Swiss sheep breeds in the focus of conservation research. *Journal of Animal Breeding and Genetics*. 126, 164-175.
- Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3). Available from <http://www.unil.ch/izea/software/fstat.html>.

- Hamzah, A., Ponzone, R.W., Nguyen, N.H., Khaw, H.L., Yee, H.Y., Nor, M., Azizah, S., 2014. Performance of the Genetically Improved Farmed Tilapia (GIFT) Strain Over Ten Generations of Selection in Malaysia. *Pertanika Journal of Tropical Agricultural Science*. 37, 411-429.
- Harris, L., Sellars, M., Perez, F., 2016. Differential family-based survival of Pacific White shrimp during communal larval culture and implications for selective breeding programs. *Aquaculture International*. 24, 273-279.
- 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.
- Holtmark, M., Sonesson, A.K., Gjerde, B., Klemetsdal, G., 2006. Number of contributing subpopulations and mating design in the base population when establishing a selective breeding program for fish. *Aquaculture*. 258, 241-249.
- Holtmark, M., Klemetsdal, G., Sonesson, A.K., Woolliams, J.A., 2008a. Establishing a base population for a breeding program in aquaculture, from multiple subpopulations, differentiated by genetic drift: I. Effects of the number of subpopulations, heritability and mating strategies using optimum contribution selection. *Aquaculture*. 274, 232-240.
- Holtmark, M., Klemetsdal, G., Sonesson, A.K., Woolliams, J.A., 2008b. Establishing a base population for a breeding program in aquaculture, from multiple subpopulations, differentiated by genetic drift: II. Sensitivity to assumptions on the additive genetic relationships of base animals. *Aquaculture*. 274, 241-246.
- Huang, C.-M., Liao, I.-C., 1990. Response to mass selection for growth rate in *Oreochromis niloticus*. *Aquaculture*. 85, 199-205.
- Hung, D., Nguyen, N.H., Hurwood, D.A., Mather, P.B., 2014. Quantitative genetic parameters for body traits at different ages in a cultured stock of giant freshwater prawn (*Macrobrachium rosenbergii*) selected for fast growth. *Marine and Freshwater Research*. 65, 198-205.
- 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.
- 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.

- Knibb, W., Quinn, J., Lamont, R., Whatmore, P., Nguyen, N.H., Remilton, C., 2014. Reproductive behaviour of captive *Fenneropenaeus merguensis*: Evidence for monogamy and high between family variances for offspring number. *Aquaculture*. 426, 60-65.
- Langella, O., 1999. POPULATIONS 1.2. 30 Population genetic software. CNRS UPR9034.[Online.] Available at bioinformatics.org/~tryphon/populations.
- Li, F., Lin, J., Liu, C., 2006. Genetic diversity analysis of primary parent and the first filial generation of pacific white shrimp, *Litopenaeus vannamei*. *Marine Sciences (Chinese language)*. 30, 64.
- 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., Smith-Keune, C., Jerry, D.R., Beheregaray, L.B., Robinson, N.A., 2016. Genetic diversity and relatedness estimates for captive barramundi (*Lates calcarifer*, Bloch) broodstock informs efforts to form a base population for selective breeding. *Aquaculture Research*. 47, 3570-3584.
- Lynch, M., Ritland, K., 1999. Estimation of pairwise relatedness with molecular markers. *Genetics*. 152, 1753-1766.
- Maggioni, R., Coimbra, M.R.M., Costa, R.B.d., Diniz, F.M., Molina, W.F., Oliveira, D.M.d., Puchnick-Legat, A., 2013. Genetic variability of marine shrimp in the Brazilian industry. *Pesquisa Agropecuária Brasileira*. 48, 968-974.
- Marsden, G., Richardson, N., Mather, P., Knibb, W., 2013. Reproductive behavioural differences between wild-caught and pond-reared *Penaes monodon* prawn broodstock. *Aquaculture*. 402, 141-145.
- Meehan, D., Xu, Z., Zuniga, G., Alcivar-Warren, A., 2003. High frequency and large number of polymorphic microsatellites in cultured shrimp, *Penaes (Litopenaeus) vannamei* [Crustacea: Decapoda]. *Marine Biotechnology*. 5, 311-330.
- Mendoza-Cano, F., Grijalva-Chon, J.M., Perez-Enriquez, R., Ramos-Paredes, J., Varela-Romero, A., 2013. Genetic diversity of mitochondrial DNA from *Litopenaeus vannamei* broodstock used in northwestern Mexico. *Ceinc. Mar.* 39, 401-412.
- Milligan, B.G., 2003. Maximum-likelihood estimation of relatedness. *Genetics*. 163, 1153-1167.
- Moss, D.R., Arce, S.M., Otsoshi, C.A., Doyle, R.W., Moss, S.M., 2007. Effects of inbreeding on survival and growth of Pacific white shrimp *Penaes (Litopenaeus) vannamei*. *Aquaculture*. 272, S30-S37.

- Moss, S.M., Moss, D.R., Arce, S.M., Lightner, D.V., Lotz, J.M., 2012. The role of selective breeding and biosecurity in the prevention of disease in penaeid shrimp aquaculture. *Journal of Invertebrate Pathology*. 110, 247-250.
- Nei, M., Tajima, F., Tatenno, Y., 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*. 19, 153-170.
- Nguyen, N.H., 2016. Genetic improvement for important farmed aquaculture species with a reference to carp, tilapia and prawns in Asia: achievements, lessons and challenges. *Fish and Fisheries*. 17, 483-506.
- Nomura, T., 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*. 1, 462-474.
- Olesen, I., Bentsen, H.B., Phillips, M., Ponzoni, R.W., 2015. Can the global adoption of genetically improved farmed fish increase beyond 10%, and how? *Journal of Marine Science and Engineering*. 3, 240-266.
- Peakall, P., Smouse, R., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 28, 2537-2539.
- Perez-Enriquez, R., Hernández-Martínez, F., Cruz, P., 2009. Genetic diversity status of White shrimp *Penaeus (Litopenaeus) vannamei* broodstock in Mexico. *Aquaculture*. 297, 44-50.
- Plotree, D., Plotgram, D., 1989. PHYLIP-phylogeny inference package (version 3.2). *cladistics*. 5, 163-166.
- Porta, J., Porta, J.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*, 258-275.
- Rambaut, A., 2012. FigTree version 1.4. 0. Available at <http://tree.bio.ed.ac.uk/software/figtree>.
- Revidatti, M., Delgado Bermejo, J., Gama, L., Periati, V.L., Ginja, C., Alvarez, L., Vega-Pla, J., Martínez, A., 2014. Genetic characterization of local Criollo pig breeds from the Americas using microsatellite markers. *Journal of Animal Science*. 92, 4823-4832.
- Rice, W.R., 1989. 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.
- Sae-Lim, P., Gjerde, B., Nielsen, H.M., Mulder, H., Kause, A., 2016. A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. *Reviews in Aquaculture*. 8, 369-393.
- Saumena, 2015. Indonesia becomes second-largest vannamei producer in the world. *BFTECH.ORG - BIOFLOC NEWS*.
- Schwartz, T., Beheregaray, L., 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.
- Silva, S.S.D., Nguyen, T.T., Turchini, G.M., Amarasinghe, U.S., Abery, N.W., 2009. Alien species in aquaculture and biodiversity: a paradox in food production. *Ambio*. 38, 24-28.
- Singh, A.K., Lakra, W.S., 2011. Risk and benefit assessment of alien fish species of the aquaculture and aquarium trade into India. *Reviews in Aquaculture*. 3, 3-18.
- Teichert-Coddington, D.R., Smitherman, R.O., 1988. Lack of response by *Tilapia nilotica* to mass selection for rapid early growth. *Transactions of the American Fisheries Society*. 117, 297-300.
- Thitamadee, S., Prachumwat, A., Srisala, J., Jaroenlak, P., Salachan, P.V., Sritunyalucksana, K., Flegel, T.W., Itsathitphaisarn, O., 2016. Review of current disease threats for cultivated penaeid shrimp in Asia. *Aquaculture*. 452, 69-87.
- Valles-Jimenez, R., Cruz, P., Perez-Enriquez, R., 2004. Population Genetic Structure of Pacific White Shrimp (*Litopenaeus vannamei*) from Mexico to Panama: Microsatellite DNA Variation. *Marine Biotechnology*. 6, 475-484.
- Van de Castele, T., Galbusera, P., Matthysen, E., 2001. A comparison of microsatellite-based pairwise relatedness estimators. *Molecular Ecology*. 10, 1539-1549.

Vela Avitúa, S., Montaldo, H.H., Márquez Valdelamar, L., Campos Montes, G.R., Castillo Juárez, H., 2013.

Decline of genetic variability in a captive population of Pacific white shrimp *Penaeus (Litopenaeus) vannamei* using microsatellite and pedigree information. *Electronic Journal of Biotechnology*. 16, 9-9.

Walker, P.J., Mohan, C., 2009. Viral disease emergence in shrimp aquaculture: origins, impact and the effectiveness of health management strategies. *Reviews in Aquaculture*. 1, 125-154.

Wang, J., 2007. Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research*. 89, 135-153.

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

Wang, J., 2014. Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. *Journal of Evolutionary Biology*. 27, 518-530.

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.

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

Wilkinson, S., Haley, C., Alderson, L., Wiener, P., 2011. An empirical assessment of individual-based population genetic statistical techniques: application to British pig breeds. *Heredity*. 106, 261-269.

Wilkinson, S., Wiener, P., Teverson, D., Haley, C., Hocking, P., 2012. Characterization of the genetic diversity, structure and admixture of British chicken breeds. *Animal Genetics*. 43, 552-563.

Zhang, K., Wang, W., Li, W., Zhang, Q., Kong, J., 2014. Analysis of genetic diversity and differentiation of seven stocks of *Litopenaeus vannamei* using microsatellite markers. *Journal of Ocean University of China*. 13, 647-656.

Table 1 *Penaeus vannamei* sample information. (^a Year of sampling; ^b Number used for microsatellite markers genotyping; NA1, North America 1; NA2, North America 2; SA1, South Asia 1; NA3, North America 3; NA4, North America 4; LA, Latin America; CN1, China 1; CN2, China 2; SA2, South Asia 2; NA5, North America 5; ^c Nucleus line located in USA, source of stock from Mexico and Ecuador; ^d Nucleus line located in USA, originated from The Kona line; ^e Nucleus line located in Thailand, source of stock from USA and South America; ^f Nucleus line located in USA, originated from The Kona line; ^g Nucleus line located in USA, originated from a population in Hawaii and a SPR line; ^h From Colombia; ⁱ Developed in China, source of stock from USA and South America; ^j Developed in China, source of stock from several local populations in China; ^k Nucleus line located in Thailand, source of stock from two populations of USA, ^l Nucleus line located in USA, source of stock from two population from USA)

Year ^a	Location (Province)	Hatchery	Population ID	Sample size ^b	Source
2013	Hainan	A	A1	50	NA1 ^c
2013	Fujian	B	B1	50	NA1 ^c
2013	Guangdong	C	C1	51	NA1 ^c
2013	Hainan	D	D1	43	NA1 ^c
2013	Hainan	D	D2	50	NA1 ^c
2013	Guangdong	E	E1	50	NA1 ^c
2014	Fujian	B	B2	20	NA1 ^c
2014	Guangdong	E	E2	20	NA1 ^c
2014	Hainan	F	F1	20	NA1 ^c
2014	Fujian	B	B3	15	NA1 ^c
2014	Fujian	G	G1	20	NA1 ^c
2015	Hainan	B	B4	25	NA1 ^c
2015	Hainan	H	H1	19	NA1 ^c
2013	Guangdong	I	I1	50	NA2 ^d
2013	Guangdong	E	E3	50	NA2 ^d
2013	Guangdong	J	J1	50	NA2 ^d
2014	Fujian	K	K1	15	NA2 ^d
2014	Fujian	L	L1	15	NA2 ^d
2014	Guangdong	I	I2	15	NA2 ^d
2015	Fujian	K	K2	25	NA2 ^d
2015	Hainan	M	M1	20	NA2 ^d
2013	Hainan	N	N1	50	SA1 ^e
2014	Hainan	N	N2	15	SA1 ^e
2014	Hainan	N	N3	20	SA1 ^e
2014	Hainan	O	O1	20	SA1 ^e
2014	Hainan	O	O2	30	SA1 ^e
2015	Hainan	N	N4	20	SA1 ^e
2014	Guangdong	P	P1	30	NA3 ^f
2014	Guangdong	P	P2	30	NA3 ^f
2014	Fujian	Q	Q1	30	NA4 ^g
2014	Fujian	Q	Q2	23	NA4 ^g
2013	Hainan	R	R1	50	LA ^h
2013	Hainan	S	S1	50	CN1 ⁱ
2013	Fujian	T	T1	50	CN2 ^j
2015	Fujian	U	U1	50	SA2 ^k
2015	Hainan	V	V1	21	NA5 ^l

Table 2. Genetic diversity measures for 36 batches *P. vannamei* broodstock ($N=1162$) from 22 hatcheries in China based on 7 microsatellite loci: North America (NA), South Asia (SA), Latin America (LA), and China (CN). Sample size (N), average number of alleles (A), allele richness (Ar), private allele richness (PAr), observed heterozygosity (Ho), expected heterozygosity (He), inbreeding coefficient of the Weir and Cockerham method (Fis), inbreeding coefficient of Wang's TrioML method (Ft), inbreeding coefficient of DyadML method (Fd), effective population size of the linkage mating model (Ne_1), and effective population size of the molecular co-ancestry method (Ne_2).

Source	Pop ID	N	A	Ar	PAr	Ho	He	Fis	Ft	Fd	Ne_1	Ne_2
NA1	A1	50	4.00	3.58	0	0.59	0.64	0.10	0.28	0.29	29.4	14.4
	B1	50	4.57	3.63	0	0.65	0.66	0.02	0.24	0.25	54.4	8.0
	C1	51	4.71	3.7	0	0.63	0.67	0.07	0.23	0.24	48.9	5.0
	D1	43	5.00	3.85	0	0.70	0.68	-0.02	0.18	0.19	65.7	32.6
	D2	50	7.00	4.95	0.01	0.69	0.75	0.09	0.18	0.19	21.9	6.8
	E1	50	5.00	3.86	0	0.62	0.66	0.08	0.25	0.26	59.9	8.5
	B2	20	4.71	4.11	0	0.67	0.68	0.04	0.21	0.22	75.4	20.4
	E2	20	4.57	3.77	0	0.60	0.64	0.08	0.25	0.27	61.2	68.6
	F1	20	4.57	4.15	0.01	0.69	0.70	0.04	0.15	0.16	Infinite	Infinite
	B3	15	4.57	4.06	0.08	0.64	0.65	0.07	0.25	0.26	37.7	5.4
	G1	20	5.43	4.49	0	0.68	0.74	0.10	0.22	0.23	Infinite	Infinite
	B4	25	4.00	3.35	0	0.47	0.58	0.21	0.41	0.43	749.6	10.4
NA2	H1	19	4.86	3.92	0.02	0.59	0.62	0.07	0.26	0.28	Infinite	Infinite
	I1	50	6.00	4.66	0.01	0.71	0.72	0.02	0.14	0.15	55.8	15.9
	E3	50	5.86	4.59	0	0.70	0.72	0.05	0.17	0.19	76.5	18.8
	J1	50	5.29	4.28	0	0.71	0.71	0.02	0.17	0.18	66.3	13.8
	K1	15	4.57	4.17	0	0.71	0.67	-0.03	0.16	0.17	115.2	7.1
	L1	15	5.29	4.63	0	0.68	0.71	0.08	0.21	0.22	137.6	Infinite
	I2	15	5.43	4.58	0.01	0.67	0.69	0.07	0.19	0.20	51.9	9.3
	K2	25	6.71	4.5	0.21	0.53	0.61	0.15	0.29	0.31	Infinite	7.8
SA1	M1	20	4.86	4.11	0.02	0.62	0.66	0.09	0.23	0.24	163.4	26.2
	N1	50	4.71	3.69	0.01	0.62	0.63	0.03	0.26	0.27	96.7	9.3
	N2	15	5.57	4.51	0.03	0.69	0.69	0.04	0.19	0.20	231.7	30.7
	N3	20	5.29	4.25	0.06	0.63	0.66	0.07	0.22	0.23	34.6	11.9
	O1	20	4.57	3.81	0	0.61	0.61	0.03	0.22	0.24	Infinite	33.6
	O2	30	6.71	4.85	0	0.67	0.73	0.11	0.21	0.23	28.5	6.6
NA3	N4	20	6.29	4.74	0.05	0.59	0.68	0.16	0.25	0.26	298.9	Infinite
	P1	30	6.00	4.91	0.03	0.72	0.75	0.05	0.17	0.18	53	14.6
NA4	P2	30	5.57	4.47	0.04	0.71	0.73	0.04	0.18	0.19	73.3	Infinite
	Q1	30	9.00	6.11	0.25	0.71	0.77	0.11	0.21	0.22	134	12.0
LA	Q2	23	8.00	5.74	0.28	0.71	0.78	0.11	0.17	0.18	276.3	Infinite
	R1	50	6.00	4.54	0.23	0.67	0.72	0.08	0.23	0.24	114.1	Infinite
CN1	S1	50	6.43	4.75	0.04	0.64	0.74	0.14	0.23	0.25	92.5	19.2
CN2	T1	50	12.43	6.69	0.39	0.77	0.81	0.05	0.13	0.13	361.1	Infinite
SA2	U1	50	10.00	5.67	0.2	0.70	0.75	0.08	0.17	0.18	132.1	20.5
NA5	V1	21	6.86	4.81	0.48	0.62	0.69	0.12	0.27	0.28	19.8	5.0
Mean		---	5.85	4.46	0.07	0.66	0.69	0.07	0.21	0.22	---	---

Table 4 Average relatedness estimates amongst 36 *P. vannamei* stocks. The pairwise relatedness estimates amongst 36 breeds calculated using R_{QG} (Queller & Goodnight, 1989) (above diagonal). The pairwise relatedness estimates of R_{DML} (Milligan, 2003) between 36 breeds (below diagonal). Tables in shadow areas were the comparison between same groups of origin. The last two rows on the table represent average R_{QG} and R_{DML} of 36 batches breeds.

I D	A 1	B 1	C 1	D 1	D 2	E 1	B 2	E 2	F 1	B 3	G 1	B 4	H 1	I 1	E 3	J 1	K 1	L 1	I 2	K 2	M 1	N 1	N 2	N 3	O 1	O 2	N 4	P 1	P 2	Q 1	Q 2	R 1	S 1	T 1	U 1	V 1
A 1		0 2 0	0 0 9	0 0 7	0 0 4	0 1 0	0 0 6	0 0 5	0 0 3	0 1 9	0 0 5	0 0 5	0 0 0	0 0 3	0 0 9	0 0 3	0 0 8	0 1 0	0 0 4	0 1 0	0 0 8	0 0 1	0 0 1	0 0 4	0 0 0	0 0 5	0 0 2	0 0 5	0 1 3	0 0 9	0 1 5	0 0 8	0 1 1	0 1 2	0 1 0	
B 1	0 3 1		0 1 1	0 0 8	0 0 4	0 1 2	0 1 0	0 0 8	0 0 6	0 2 4	0 0 5	0 0 8	0 0 4	0 0 2	0 0 4	0 0 0	0 1 0	0 0 8	0 0 5	0 0 8	0 0 2	0 0 0	0 0 1	0 0 3	0 0 1	0 0 4	0 0 2	0 0 6	0 0 2	0 0 9	0 1 3	0 0 9	0 1 1	0 0 9	0 0 9	
C 1	0 2 2	0 2 4		0 1 2	0 0 5	0 1 1	0 0 1	0 1 1	0 1 0	0 0 9	0 0 4	0 0 1	0 1 1	0 0 1	0 0 2	0 0 0	0 0 0	0 0 3	0 0 3	0 0 5	0 0 2	0 0 0	0 0 1	0 0 3	0 0 0	0 0 3	0 0 2	0 0 0	0 0 1	0 0 6	0 0 6	0 1 3	0 0 3	0 0 7	0 0 2	0 0 3
D 1	0 2 2	0 2 2	0 2 4		0 0 7	0 1 3	0 0 8	0 0 2	0 1 0	0 0 6	0 0 4	0 0 9	0 0 3	0 0 2	0 0 3	0 0 0	0 0 0	0 0 4	0 0 1	0 0 4	0 0 3	0 0 7	0 0 6	0 0 2	0 0 1	0 0 6	0 0 1	0 0 3	0 0 4	0 0 5	0 0 5	0 1 0	0 0 2	0 0 4	0 0 7	
D 2	0 1 9	0 1 1	0 1 8	0 2 0		0 0 6	0 0 2	0 0 2	0 0 0	0 0 1	0 0 1	0 0 0	0 0 0	0 0 0	0 0 1	0 0 0	0 0 3	0 0 0	0 0 2	0 0 1	0 0 6	0 0 6	0 0 2	0 0 0	0 0 1	0 0 4	0 0 1	0 0 0	0 0 4	0 0 5	0 0 3	0 0 2	0 0 6	0 0 3	0 0 3	
E 1	0 2 3	0 2 5	0 2 5	0 2 6	0 1 9		0 1 2	0 1 0	0 1 1	0 0 4	0 0 7	0 0 6	0 0 5	0 0 4	0 0 5	0 0 4	0 0 5	0 0 0	0 0 2	0 0 6	0 0 7	0 0 6	0 0 3	0 0 4	0 0 0	0 0 6	0 0 1	0 0 0	0 0 4	0 0 4	0 0 7	0 1 5	0 0 4	0 1 0	0 0 4	0 0 9
B 2	0 2 0	0 2 4	0 2 9	0 1 5	0 2 3	0 0 0		0 1 8	0 1 4	0 2 9	0 0 5	0 0 5	0 0 1	0 0 2	0 0 0	0 0 5	0 0 4	0 0 5	0 0 3	0 0 2	0 0 3	0 0 1	0 0 1	0 0 4	0 0 0	0 0 3	0 0 3	0 0 5	0 0 2	0 0 4	0 0 7	0 0 0	0 1 4	0 0 9	0 0 2	0 0 9
E 2	0 1 9	0 2 2	0 2 3	0 2 6	0 2 9	0 0 0	0 0 0		0 1 4	0 1 7	0 1 3	0 1 4	0 0 3	0 0 0	0 0 1	0 0 4	0 0 1	0 0 4	0 0 3	0 0 3	0 0 3	0 0 1	0 0 5	0 0 8	0 0 0	0 0 4	0 0 3	0 0 2	0 0 5	0 0 5	0 1 2	0 0 2	0 0 9	0 0 0	0 0 2	
F 1	0 1 8	0 2 2	0 2 2	0 2 1	0 2 2	0 0 0	0 0 2	0 0 6		0 0 6	0 0 7	0 0 1	0 0 2	0 0 3	0 0 2	0 0 1	0 0 1	0 0 2	0 0 5	0 0 5	0 0 0	0 0 2	0 0 2	0 0 7	0 0 3	0 0 5	0 0 2	0 0 2	0 0 3	0 0 5	0 1 1	0 0 3	0 0 8	0 0 1	0 0 8	
B 3	0 3 0	0 3 6	0 2 1	0 2 6	0 2 5	0 0 0	0 0 1	0 0 9	0 0 0		0 0 7	0 0 1	0 0 6	0 0 2	0 0 3	0 0 5	0 0 8	0 0 0	0 0 7	0 0 4	0 0 0	0 0 2	0 0 3	0 0 1	0 0 6	0 0 3	0 0 7	0 0 1	0 0 8	0 0 6	0 0 4	0 0 1	0 0 1	0 0 1	0 0 2	
G 1	0 1 9	0 2 0	0 1 7	0 1 5	0 1 9	0 0 0	0 0 2	0 0 2	0 0 2	0 0 1		0 0 4	0 0 6	0 0 1	0 0 1	0 0 3	0 0 2	0 0 0	0 0 2	0 0 1	0 0 2	0 0 1	0 0 2	0 0 1	0 0 3	0 0 3	0 0 2	0 0 3	0 0 2	0 0 5	0 0 5	0 1 2	0 0 5	0 0 9	0 0 5	0 0 7
B 4	0 1 8	0 2 2	0 2 2	0 2 6	0 2 7	0 0 0	0 0 5	0 0 4	0 0 2	0 0 7	0 0 2		0 0 8	0 0 3	0 0 0	0 0 1	0 0 4	0 0 4	0 0 4	0 0 0	0 0 3	0 0 5	0 0 2	0 0 5	0 0 0	0 0 5	0 0 6	0 0 2	0 0 9	0 0 9	0 1 7	0 0 8	0 0 3	0 0 8	0 0 5	
H 1	0 5 1	0 1 7	0 2 0	0 2 5	0 2 4	0 0 0	0 0 2	0 0 4	0 0 2	0 1 7	0 1 9	0 1 3		0 0 1	0 0 3	0 0 0	0 0 2	0 0 0	0 0 5	0 0 1	0 0 3	0 0 2	0 0 1	0 0 3	0 0 3	0 0 0	0 0 6	0 0 3	0 0 2	0 0 7	0 0 7	0 1 0	0 0 1	0 1 3	0 0 3	
I 1	0 3 4	0 1 1	0 1 3	0 1 5	0 2 7	0 0 0	0 0 2	0 0 2	0 0 4	0 0 5	0 0 2	0 0 2	0 0 1		0 0 8	0 0 0	0 0 8	0 0 5	0 0 2	0 0 4	0 0 9	0 0 3	0 0 3	0 0 3	0 0 5	0 0 3	0 0 0	0 0 4	0 0 4	0 0 2	0 0 1	0 0 5	0 0 1	0 0 3	0 0 1	0 0 2
E 3	0 8 2	0 0 1	0 1 1	0 1 3	0 1 7	0 0 0	0 0 1	0 0 1	0 0 3	0 0 3	0 0 3	0 0 1	0 0 1	0 0 1		0 0 8	0 0 0	0 0 9	0 0 6	0 0 1	0 0 5	0 0 3	0 0 3	0 0 3	0 0 3	0 0 2	0 0 4	0 0 3	0 0 1	0 0 3	0 0 1	0 0 2	0 0 2	0 0 2	0 0 2	
J 1	0 1 3	0 1 6	0 1 5	0 1 6	0 1 3	0 0 0	0 0 3	0 0 3	0 0 4	0 0 8	0 0 4	0 0 5	0 0 3	0 0 0	0 0 9		0 0 8	0 0 8	0 0 0	0 0 1	0 0 1	0 0 3	0 0 2	0 0 3	0 0 1	0 0 5	0 0 2	0 0 3	0 0 7	0 0 1	0 0 4	0 0 3	0 0 4	0 0 1	0 0 3	
K 1	0 0 9	0 0 3	0 1 4	0 1 8	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 2	0 0 0	0 0 0		0 0 9	0 0 1	0 0 8	0 0 2	0 0 2	0 0 1	0 0 6	0 0 4	0 0 1	0 0 5	0 0 4	0 0 0	0 0 2	0 0 0	0 1 0	0 0 4	0 0 4	0 0 5	
L 1	0 0 9	0 1 1	0 1 3	0 1 6	0 1 3	0 0 0	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 2	0 0 0	0 0 8	0 0 5	0 0 3	0 0 9	0 0 5	0 0 4	0 0 1	0 0 3	0 0 4	0 0 3	0 0 6	0 0 8	0 0 1	0 0 1	0 0 4	0 0 0	0 0 3	0 0 2	0 0 0

Fig 1

Fig 2

Fig 3(a&b)

Highlights

- Despite the most important role of *Penaeus vannamei* in shrimp aquaculture, population genetics information for broodstock from either domesticated or wild population is limited; in this paper we report on genetic (microsatellite) variation for this important culture species captured in 36 Chinese breeding lines from 22 hatcheries which is the most comprehensive sampling design to our knowledge.
- Microsatellite data analysis successfully resolved different origins of *P. vannamei* culture stocks used in China that resulted in the detection of four lineages with genetically diverse origins.
- Results of the analysis here has provided the foundation and guidelines to develop a base line for genetic improvement program with optimal levels of genetic variation and a relatively low inbreeding coefficient.
- We believe that the data presented here clearly demonstrate the utility of accessing genetic resources from a variety of sources (particularly when natural populations are not available) and show this as a viable and necessary first step for developing a base population for pursuing subsequent genetic improvement in alien aquaculture species.