

## Pedigree traceability in whiteleg shrimp (*Litopenaeus vannamei*) using genetic markers: A comparison between microsatellites and SNPs

### Rastreabilidad del pedigrí en camarón blanco (*Litopenaeus vannamei*) mediante marcadores genéticos: Una comparación entre microsatélites y SNP

Ricardo Perez-Enriquez\*, Adriana Max-Aguilar

Centro de Investigaciones Biológicas del Noroeste (CIBNOR), Calle IPN 195, La Paz, Baja California Sur, México.

\* Corresponding author: rperez@cibnor.mx

**ABSTRACT.** To increase productivity, genetic improvement in cultivated shrimp is of much interest. Evaluation of genetic parameters (e.g., heritability, genotype–environment interaction, inbreeding) and the design of appropriate breeding plans are necessary steps towards genetic improvement. Pedigree traceability by genetic markers is relevant to these objectives. The aim of this study was to compare the performance of 2 genetic marker panels (microsatellites and single nucleotide polymorphisms [SNPs]) for pedigree traceability of a cultivated stock of whiteleg shrimp (*Litopenaeus vannamei*). Pedigree of a stock from 81 full-sib families reared in a common environment was assessed with microsatellites and SNPs as genetic markers. The panel of 76 SNPs performed better than microsatellites, allowing 94–96% of parentage assignment of the tested progeny ( $n = 192$ ). A minimum number of 50 SNPs with a proportion of 60% loci with a minimum allele frequency of 0.3 is suitable for successful pedigree assignment. SNP markers are suggested for confidently testing the pedigree of shrimp from known parental broodstock.

**Key words:** shrimp farming, parentage assignment, genetic identification, assignment likelihood, exclusion probability.

**RESUMEN.** El mejoramiento genético en el camarón cultivado es de amplio interés para incrementar la productividad. La evaluación de parámetros genéticos (e.g., heredabilidad, interacción entre genotipo y ambiente, endogamia) y el diseño de planes de apareamiento son elementos necesarios para el mejoramiento genético. La rastreabilidad del pedigrí con marcadores genéticos es importante para estos objetivos. El objetivo de este trabajo fue comparar el desempeño de 2 paneles de marcadores genéticos (microsatélites y polimorfismos de nucleótido sencillo [SNP]) para rastrear el pedigrí en una población de camarón blanco (*Litopenaeus vannamei*) de cultivo. Se determinó el pedigrí de la progenie de 81 familias de hermanos completos mediante el uso de microsatélites y SNP como marcadores genéticos. Un panel de 76 SNP probó tener un mejor desempeño que los microsatélites para la asignación de parentesco, obteniéndose un 94–96% de asignación en una muestra de la progenie ( $n = 192$ ). Un número mínimo de 50 SNP con una proporción de 60% de loci con una frecuencia alélica mínima de 0.3 es adecuado para una asignación exitosa de pedigrí. Se sugiere utilizar los marcadores SNP para determinar el parentesco de la progenie de camarón proveniente de padres conocidos.

**Palabras clave:** cultivo de camarón, asignación de parentesco, identificación genética, probabilidad de asignación, probabilidad de exclusión.

## INTRODUCTION

It is estimated that only 10% of the world aquaculture production is associated with genetic improvement programs; thus, there exists a large potential for increasing production by this approach (Gjedrem et al. 2012). The whiteleg shrimp, *Litopenaeus vannamei*, is one of the species for which there is growing effort to implement genetic selection programs (Andriantahina et al. 2013). This species is the most important worldwide in shrimp farming, contributing more than 71% of the total production (FAO 2014).

Pedigree tracking is necessary in genetic breeding programs to estimate genetic parameters, including heritability, genotype-by-environment ( $G \times E$ ) interactions, and inbreeding (Gjedrem 2010). The determination of parentage relationships among broodstock is also useful when designing

## INTRODUCCIÓN

Se estima que sólo el 10% de la producción acuícola mundial está asociada a programas de mejoramiento genético, por lo que existe un alto potencial para el incremento de la producción a través de esta metodología (Gjedrem et al. 2012). El camarón blanco, *Litopenaeus vannamei*, es una de las especies para la cual existe un esfuerzo creciente en la instrumentación de programas de selección genética (Andriantahina et al. 2013). Ésta es la especie más importante en la acuicultura de camarones marinos en el mundo, ya que contribuye con más del 71% del total del volumen de producción (FAO 2014).

Rastrear el pedigrí en los núcleos genéticos es necesario para la estimación de parámetros genéticos como la heredabilidad, la interacción entre genotipo y ambiente ( $G \times A$ ) y la

breeding plans in order to minimize loss of genetic diversity and inbreeding and maximize productivity (Gjedrem 2010).

Pedigree traceability in genetic breeding programs is usually done by family-based grow-out and/or physical tagging with colored elastomers. This strategy has limitations for individual identification and parentage assignment; hence, genetic tagging has been proposed (Gjedrem 2010). Yue and Xia (2014) summarize several practical considerations for this type of analysis focusing on microsatellites and single nucleotide polymorphisms (SNPs).

The feasibility of microsatellites in parentage assessment has been studied in several shrimp species, such as *Marsupenaeus japonicus* (Jerry et al. 2004), *Penaeus monodon* (Jerry et al. 2006), and *Fenneropenaeus chinensis* (Dong et al. 2006). However, it has been reported that the putative parents of more than 10% of the individuals are usually not accurately assigned (Jerry et al. 2004, Dong et al. 2006, Jerry et al. 2006). This non-assignment proportion has also been reported for fish (e.g., Perez-Enriquez et al. 1999, Christie et al. 2010).

Next-generation sequencing (high-throughput sequencing) has opened the possibility of using SNPs to assess parentage. Lapègue et al. (2014) used SNP markers for pedigree identification of 2 species of cultivated oysters, and suggested using a panel of at least 150 SNPs for a success rate over 95%. For the giant tiger shrimp, *P. monodon*, Sellars et al. (2014) obtained adequate parentage assignments with panels of 56 and 63 SNPs. In salmon, a 60-loci SNP panel is feasible to genetically differentiate between wild and cultivated stocks (Karlsson et al. 2011).

In the whiteleg shrimp, *L. vannamei*, genetic markers have mostly been used to characterize genetic diversity and composition of wild and cultured stocks (Cruz et al. 2004, Freitas and Galetti 2005, Perez-Enriquez et al. 2009, Vela-Avitúa et al. 2013). A few studies have focused on individual identification or parentage assignment of breeding stocks (Perez-Enriquez R, unpublished data; Perez F, pers. comm.). The objective of this study was to compare the performance of 2 genetic marker panels (microsatellites and SNPs) for pedigree traceability of a cultivated stock of whiteleg shrimp.

## MATERIALS AND METHODS

### Sampling

Broodstock from the fifth generation of a commercial breeding program (Acuicultura Mahr hatchery) were selected to obtain 81 full-sib families by artificial insemination. A homogenous number of nauplii per family were placed in a single 20-t tank; larval and juvenile rearing was done following the hatchery's standard procedures until the juveniles weighed 20 g, when a random sample (192 individuals) was collected. Pleopods from broodstock ( $n = 162$ ) and progeny were preserved in 80% ethanol.

endogamia (Gjedrem 2010). Además, la determinación de las relaciones de parentesco entre reproductores es útil para el diseño de las cruza que minimicen la pérdida de diversidad genética y el incremento de la endogamia y maximicen la productividad de la siguiente generación (Gjedrem 2010).

La rastreabilidad del pedigrí en los programas de mejoramiento genético se lleva a cabo mediante la crianza de familias por separado y/o mediante el marcaje físico con elastómeros de color. Sin embargo, dado que esta estrategia tiene limitaciones respecto a la identificación individual de los ejemplares y la asignación de parentesco, el marcaje genético se ha sugerido como una alternativa (Gjedrem 2010). Yue y Xia (2014) resumen varias consideraciones prácticas para este tipo de análisis particularizando en microsatélites y polimorfismos de nucleótido sencillo (SNP).

La factibilidad de los microsatélites para el análisis de parentesco se ha estudiado en algunas especies de camarón, tales como *Marsupenaeus japonicus* (Jerry et al. 2004), *Penaeus monodon* (Jerry et al. 2006) y *Fenneropenaeus chinensis* (Dong et al. 2006). Sin embargo, se ha reportado que 10% o más de los ejemplares no son asignados adecuadamente a sus padres putativos (Jerry et al. 2004, Dong et al. 2006, Jerry et al. 2006). El índice de falta de asignaciones con el uso de microsatélites ha sido también reportado para peces (e.g., Perez-Enriquez et al. 1999, Christie et al. 2010).

Con el desarrollo de la secuenciación masiva de nueva generación, se abrió la posibilidad de utilizar los SNP en estudios de parentesco. Lapègue et al. (2014) utilizaron marcadores SNP para identificar el pedigrí de 2 especies de ostión cultivado y sugirieron utilizar un panel de al menos 150 SNP para poder identificar el parentesco de más del 95% de la progenie. Para el camarón tigre, *P. monodon*, Sellars et al. (2013) obtuvieron asignaciones adecuadas con paneles de 56 y 63 SNP. En salmón, se ha demostrado que con 60 loci de SNP es posible diferenciar genéticamente entre salmón silvestre y salmón de cultivo (Karlsson et al. 2011).

En el camarón blanco, *L. vannamei*, los marcadores genéticos se han utilizado principalmente para la caracterización de la diversidad y composición genética de poblaciones silvestres o lotes de cultivo (Cruz et al. 2004, Freitas y Galetti 2005, Perez-Enriquez et al. 2009, Vela-Avitúa et al. 2013). Son limitados los estudios dirigidos a la identificación individual o asignación de parentesco de individuos cultivados (Perez-Enriquez R, datos no publicados; Perez F com. pers.). El objetivo de este trabajo fue probar y comparar el desempeño de 2 paneles de marcadores genéticos (microsatélites y SNP) para rastrear el pedigrí en una población de camarón blanco de cultivo.

## MATERIALES Y MÉTODOS

### Obtención de las muestras

Se seleccionó un lote de quinta generación del núcleo genético de la empresa Acuicultura Mahr para la obtención

## Genotyping

DNA from pleopods was obtained with the Wizard SV 96 Genomic DNA Purification System (A2371, Promega). Microsatellite and SNP genetic profiles were obtained for each individual. A set of 5 published microsatellite loci was used: Pvan1758 and Pvan1815 (Cruz et al. 2002), Lvan05 (Perez-Enriquez et al. 2009), and TumxLv8.256 and TumxLv10.312 (Meehan et al. 2003). Microsatellites Pvan1758, Lvan05, and TumxLv8.26 were amplified by PCR in 11- $\mu$ L reactions (1  $\mu$ L DNA at 10–95 ng  $\mu$ L<sup>-1</sup>, 1 $\times$  PCR buffer, 3.5 mM MgCl<sub>2</sub>, 0.25 mM dNTPs, 0.39  $\mu$ M each of the forward and reverse primers, and 0.025 U  $\mu$ L<sup>-1</sup> Taq polymerase) under the authors' reported conditions. For Pvan1815 locus, a touchdown protocol was used (95 °C for 3 min, 30 cycles at 94 °C for 45 s, annealing temperature for 45 s [starting at 60 °C and decreasing each cycle 0.3 °C down to 51 °C], 65 °C for 1 min, and final extension at 65 °C for 20 min). TumxLv10.312 locus was run using the modified protocol proposed by Yoshida and Awaji (2000) (94 °C for 2 min; 5 cycles at 94 °C for 5 s, 60 °C for 1 min, and 72 °C for 1 min; 20 cycles at 94 °C for 1 s, 60 °C for 1 min, and 72 °C for 1 min; and final extension at 72 °C for 10 min). An additional set of 5 loci (Livan04, Livan13, Livan44, Livan51, and Livan60; Table 1) was used. They contained 1.5 mM MgCl<sub>2</sub> and the same amount of the other components described above was used with the following amplification conditions: 95 °C for 3 min; 30 cycles at 94 °C for 35 s, 57 °C for 40 s, and 72 °C for 40 s; and final extension at 72 °C for 5 min. PCR products were analyzed in an automatic sequencer (ABI3130, Applied Biosystems). Genotypes were read with the software GeneMapper 4.0.

SNP genotyping with a 76-loci panel was performed at the Center for Aquaculture Technologies (San Diego, California, <http://aquatechcenter.com/>) using allele-specific PCR assays on an EP1 instrument platform (Fluidigm

de 81 familias de hermanos completos mediante inseminación artificial. Se sembraron cantidades homogéneas de nauplios por familia en un tanque común, y se siguieron los procedimientos estándar de engorda hasta que los organismos alcanzaron un peso promedio de 20 g. Luego, se obtuvo una muestra aleatoria de 192 individuos de la progenie a este peso. Los pleópodos de los reproductores ( $n = 162$ ) y la progenie se preservaron en etanol al 80%.

## Obtención de los genotipos

El ADN se obtuvo de los pleópodos utilizando el Wizard SV 96 Genomic DNA Purification System (A2371, Promega). Se obtuvieron perfiles genéticos individuales con los microsatélites y SNP. Se utilizó un lote de 5 microsatélites previamente publicados: Pvan1758 y Pvan1815 (Cruz et al. 2002), Lvan05 (Perez-Enriquez et al. 2009) y TumxLv8.256 y TumxLv10.312 (Meehan et al. 2003). Los microsatélites Pvan1758, Lvan05 y TumxLv8.26 se amplificaron mediante PCR en un volumen de reacción final de 11  $\mu$ L (1  $\mu$ L de ADN a 10–95 ng  $\mu$ L<sup>-1</sup>, amortiguador 1 $\times$ , 3.5 mM de MgCl<sub>2</sub>, 0.25 mM de dNTP, 0.39  $\mu$ M de cada iniciador sentido y antisentido, 0.025 U  $\mu$ L<sup>-1</sup> de Taq polimerasa) bajo las condiciones reportadas por los respectivos autores. Para el *locus* Pvan1815, se utilizó un protocolo de descenso gradual de temperatura (95 °C por 3 min, 30 ciclos a 94 °C por 45 s, temperatura de acoplamiento por 45 s [iniciando en 60 °C y disminuyendo cada ciclo 0.3 °C hasta alcanzar 51 °C], 65 °C por 1 min y extensión final a 65 °C por 20 min). Para el *locus* TumxLv10.312, se utilizó el protocolo modificado de Yoshida y Awaji (2000) (94 °C por 2 min; 5 ciclos a 94 °C por 5 s, 60 °C por 1 min y 72 °C por 1 min; 20 ciclos a 94 °C por 1 s, 60 °C por 1 min y 72 °C por 1 min; y extensión final a 72 °C por 10 min). Además, se utilizó un lote adicional de 5 *loci* (Livan04, Livan13, Livan44, Livan51, Livan60; Tabla 1), para los cuales se usaron 1.5 mM de MgCl<sub>2</sub> y la misma cantidad de los otros componentes mencionados arriba con las condiciones de amplificación siguientes: 95 °C por 3 min; 30 ciclos a 94 °C por 35 s, 57 °C por 40 s y 72 °C por 40 s; y extensión final a 72 °C por 5 min. Los productos de la PCR fueron analizados en un secuenciador (ABI3130, Applied Biosystems). Los genotipos se obtuvieron con el programa GeneMapper 4.0.

Los genotipos de SNP de un panel de 76 *loci* se obtuvieron en el Center for Aquaculture Technologies (San Diego, California, <http://aquatechcenter.com/>) utilizando ensayos de PCR específicos para cada alelo en una plataforma instrumental EP1 (Fluidigm Corporation, San Francisco, CA). Los genotipos se analizaron en un archivo Excel.

## Análisis del pedigrí

Para ambos grupos (reproductores y progenie), se estimó la diversidad genética (número de alelos por *locus* y heterocigosidad observada y esperada) con ambos marcadores

**Table 1.** Primer sequences of *Litopenaeus vannamei* microsatellites.

**Tabla 1.** Secuencias de los iniciadores de microsatélites de *Litopenaeus vannamei*.

Locus	Forward and reverse primer sequences (5'–3')
Livan04	ATTCTTGGAGTATGCGGTGG TGATTTGAGAACGAGACGGA
Livan13	AAAGTGC GGATATTAGTGT TTTTGT TCTGCACGTCTTCCTTTCT
Livan44	ACCCTCTCATCAAGCAGTGG TCCACAGAAGAGCGTGT TTTG
Livan51	CAATTACTCCGGCCTCAAGA AACCGTACACAGGCCAATTC
Livan60	TGGCCGTAGATACTGACCCCT CATGCAGGCTTGAAGAGTGA

Corporation, San Francisco, CA). Genotypes were analyzed in an Excel file.

### Pedigree analysis

Genetic diversity (number of alleles per locus, observed and expected heterozygosity) was estimated for each group (broodstock and progeny) from both marker panels (microsatellites and SNPs). The probability of identity (likelihood that 2 individuals have exactly the same multilocus genotype) and the combined probability of non-exclusion (likelihood of not excluding a non-true parent) were calculated using Cervus 3.0.3 (Kalinowski et al. 2007). The presence of null alleles and other genotyping artifacts in microsatellites was analyzed with the Microchecker software (Van Oosterhout et al. 2004).

Both panels were used to perform paternity tests to determine the pedigree (family of origin) of each progeny. The analysis was run only for progeny and broodstock with at least 80% of the multilocus genotype complete (i.e., 8 microsatellite loci or 60 SNP loci).

The parentage assignment tests were done by direct and probabilistic exclusion methods. Direct exclusion, which implies the exclusion of potential breeders with more than 2 mismatching loci, was done with Vitassign software (Vandeputte et al. 2006). This program also performs a simulation to determine the assignment probability from the percentage of unique assignments. Cervus 3.0.3 software (Kalinowski et al. 2007) was used for probabilistic assignment tests. This program gives a minimum assignment value (named LOD) to the pair with the highest likelihood of being the true couple, as a function of the proportion of sampled parents (0.90), genotyping error (0.01), minimum number of genotyped loci, and confidence level (95%). Simulation and real data assignments were compared between microsatellites and SNP markers.

Rarefaction curves were made to estimate the minimum number of SNP loci needed to obtain more than 98% of correct assignments. SNP loci were classified into 2 groups: those with minimum allele frequency (MAF)  $\geq 0.3$  (both alleles had a frequency between 0.3 and 0.7) and those with MAF  $< 0.3$  (one of the alleles had a frequency of less than 0.3). Four rarefaction curves were drawn based on the proportion (20%, 40%, 50%, and 60%) of loci with MAF  $\geq 0.3$ . Paternity analyses were done with Vitassign software (allowing 2 mismatches) and Cervus software (1.0 proportion of sampled breeders, 0.001 genotyping error, 95% confidence level). The curves were built by obtaining the assignments with 10, 20, 30, 40, 50, 60, and 70 loci, randomly selected from the original set of 76 loci. Only the correctly-assigned progeny from the pedigree test were used for the analyses.

## RESULTS

Of the 192 individuals analyzed, complete multilocus genotypes were obtained for 168 using microsatellites

genéticos (microsatélites y SNP). Además, se calculó la probabilidad de identidad (probabilidad de que 2 individuos tengan exactamente el mismo genotipo múltiple) y la probabilidad combinada de no exclusión (probabilidad de no excluir a un padre no verdadero) mediante el programa Cervus 3.0.3 (Kalinowski et al. 2007). La presencia de alelos nulos y otras irregularidades de genotipificación se analizaron con el programa Microchecker 2.2.3 (Van Oosterhout et al. 2004).

Con ambos paneles, se realizaron análisis de paternidad para determinar el pedigrí (la familia de origen) de cada individuo de la progenie. Para estos análisis sólo se utilizaron ejemplares con un mínimo de 80% de genotipos completos (i.e., 8 microsatélites o 60 SNP).

Los análisis de asignación de parentesco se realizaron mediante exclusión directa y probabilística. La exclusión directa, que implica la eliminación de progenitores potenciales que tengan más de 2 *loci* sin coincidencia con los alelos de la progenie, se llevó a cabo con el programa Vitassign (Vandeputte et al. 2006). Este programa también realiza una simulación para determinar el poder de la asignación calculando el porcentaje de asignaciones únicas. Para la asignación probabilística se utilizó el programa Cervus 3.0.3 (Kalinowski et al. 2007), el cual otorga un valor mínimo de asignación (denominado LOD) a la pareja más probable en función de parámetros definidos (0.90 de proporción de padres muestreados, 0.01 de error para cálculos de probabilidad, número mínimo de *loci* genotipificados y 95% de nivel de confianza). Se realizó un análisis comparativo del porcentaje de asignación de paternidad obtenido con los microsatélites y SNP tanto a partir de simulaciones como de los datos reales.

Se elaboraron curvas de rarefacción para estimar el número mínimo de *loci* de SNP necesario para obtener más del 98% de asignaciones correctas. Para ello, los *loci* de los SNP se clasificaron en 2 grupos: el grupo con frecuencia alélica mínima (FAM)  $\geq 0.3$  (ambos alelos presentaron una frecuencia entre 0.3 y 0.7) y el grupo con FAM  $< 0.3$  (uno de los alelos presentó una frecuencia menor que 0.3). Se elaboraron 4 curvas de rarefacción en función de la proporción (20%, 40%, 50% y 60%) de *loci* con FAM  $\geq 0.3$ . Para cada curva, se realizaron análisis de paternidad con el programa Vitassign (permitiendo 2 no coincidencias) y con el programa Cervus (1.0 de proporción de padres muestreados, 0.001 de error para cálculos de probabilidad y 95% de nivel de confianza). Las curvas se construyeron obteniendo las asignaciones con 10, 20, 30, 40, 50, 60 y 70 *loci*, seleccionados al azar del lote original de 76 *loci*. Únicamente se utilizaron los ejemplares que tuvieron asignación correcta de acuerdo con el análisis de pedigrí descrito arriba.

## RESULTADOS

De los 192 individuos de la progenie analizados, se obtuvieron lecturas completas de 168 con los microsatélites

(87.5%) and for 186 using SNPs (96.9%). Individuals with incomplete genotypes were not used for the parentage assignment tests.

Higher genetic diversity values (number of alleles per locus, heterozygosity) were obtained with microsatellites than with SNPs; however, the non-exclusion probability was 3 orders of magnitude higher (i.e., a lower probability) with SNPs (Table 2). This characteristic led to a higher percentage of paternity assignment using the SNP panel with both exclusion methods (direct and probabilistic) and with both simulated and real data (Table 2). The use of the 76-loci SNP panel resulted in 94–96% progeny assignment to a unique parental couple, much higher than with microsatellites. The other 4–6% does not correspond to progeny with multiple potential parents, but to a lack of assignment probably due to individuals that originated in a different spawning lot. Null alleles were detected in 3 microsatellite loci (Lvan05, Livan13, and Livan60), which may explain the reduced assignment with these markers.

The rarefaction curves indicate that paternity tests by direct exclusion using 50 loci reached 100% assignments when 60% of loci had a  $MAF \geq 0.3$ , and between 98.3% and 99.4% of assignments for  $MAF \geq 0.3$  proportions of 20%, 40%, and 50% (Fig. 1). The probabilistic assignments at 50 loci reached 100% in 3 of the  $MAF \geq 0.3$  proportions (40%, 50%, and 60%; Fig. 1).

Considering 50 loci as the minimum number for a parentage assignment analysis, the estimated error (the difference between the assigned parental couple and the true parental couple) showed a decreasing trend as a function of the

(87.5%) y 186 con los SNP (96.9%). Los individuos con genotipos incompletos no se consideraron para los análisis de parentesco.

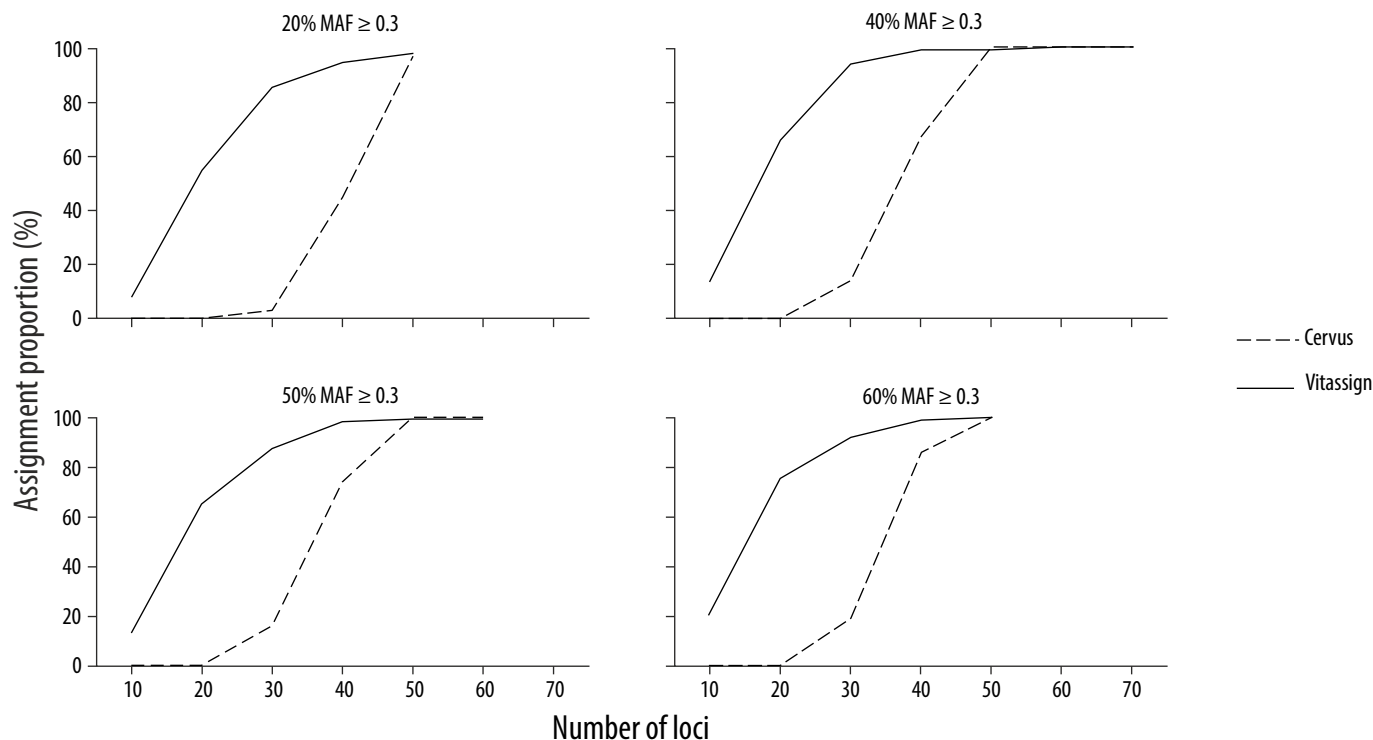
Los valores de diversidad genética (número de alelos por *locus*, heterocigosidad) fueron mayores con los microsatélites que con los SNP; sin embargo, la probabilidad de no exclusión fue 3 órdenes de magnitud más potente (i.e., un valor de probabilidad más bajo) con los SNP (Tabla 2). Esta característica se reflejó en un porcentaje mayor de asignación de paternidad al utilizar el panel de SNP con los 2 métodos de exclusión (directa y probabilística), tanto en la simulación como con los datos reales (Tabla 2). El uso del panel de 76 SNP resultó en una asignación de 94% a 96% de la progenie a una pareja única de padres, ampliamente mayor que lo obtenido con los microsatélites. Cabe mencionar que el 4–6% restante no se refiere a ejemplares con asignaciones múltiples, sino a una falta de asignación, lo que indica que probablemente esos ejemplares pertenecen a otros lotes de desove. Se detectaron alelos nulos en 3 *loci* (Lvan05, Livan13 y Livan60), que pueden explicar la reducida proporción de asignaciones con estos marcadores.

Las curvas de rarefacción indican que las pruebas de paternidad por exclusión directa usando 50 *loci* alcanzaron un 100% de asignaciones cuando el 60% de los *loci* tuvieron una  $FAM \geq 0.3$ , y entre 98.3% y 99.4% de asignaciones para las proporciones de  $FAM \geq 0.3$  de 20%, 40% y 50% (Fig. 1). Las asignaciones probabilísticas con 50 *loci* alcanzaron el 100% en 3 proporciones de  $FAM \geq 0.3$  (40%, 50% y 60%; Fig. 1).

**Table 2.** Mean values of genetic diversity (standard deviation in parentheses), probability of identity, combined non-exclusion probability for both parents, and percentage of parentage assignments of progeny (simulated and real data) obtained with microsatellite and SNP panels by direct exclusion and likelihood.

**Tabla 2.** Valores promedio de diversidad genética (desviación estándar en paréntesis), probabilidad de identidad, probabilidad combinada de no exclusión de ambos padres y porcentaje de asignaciones de parentesco de la progenie (datos simulados y reales) obtenidos con los paneles de microsatélites y SNP por exclusión directa y probabilística.

	Microsatellites	SNPs
Number of alleles per locus	6.6 (1.8)	1.97 (0.16)
Observed heterozygosity	0.58 (0.16)	0.34 (0.14)
Expected heterozygosity	0.69 (0.08)	0.34 (0.14)
Probability of identity	$1.4 \times 10^{-9}$	$2.3 \times 10^{-23}$
Combined non-exclusion probability	$1.6 \times 10^{-5}$	$1.0 \times 10^{-8}$
Parentage of assignment to both parents (simulated data)		
Percentage of assignment by direct exclusion (Vitassign)	83.3%	97.7%
Percentage of assignment by probabilistic exclusion (Cervus)	23.0%	85.0%
Parentage of assignment to both parents (real data)		
Sample size	168	186
Percentage of assignment by direct exclusion (Vitassign)	76.8%	94.1%
Percentage of assignment by probabilistic exclusion (Cervus)	15.0%	96.0%



**Figure 1.** Rarefaction curves of parentage assignment with the SNP panel by direct exclusion (Vitassign software) and probabilistic exclusion (Cervus software), according to the proportion of loci with minimum allele frequency (MAF)  $\geq 0.3$ .

**Figura 1.** Curvas de rarefacción de las asignaciones de parentesco con el panel de SNP por exclusión directa (programa Vitassign) y exclusión probabilística (programa Cervus), de acuerdo con la proporción de *loci* con frecuencia alélica mínima (MAF)  $\geq 0.3$ .

proportion of loci with MAF  $\geq 0.3$ , the lowest value (2.3%; Fig. 2) being obtained with the 60% proportion.

## DISCUSSION

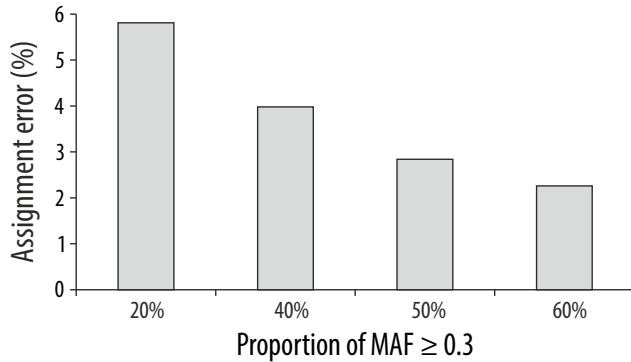
This is the first report to the authors' knowledge of the usefulness of a SNP panel for *L. vannamei*. It has been demonstrated that a panel of 76 SNPs is suitable to identify the pedigree of cultivated whiteleg shrimp. Despite the lower genetic diversity compared to microsatellites, the certainty of the SNP panel in parentage assignment is higher. The lower resolution of microsatellites than SNPs for this type of application has also been described for tilapia (Trọng et al. 2013) and the giant tiger shrimp, *P. monodon* (Sellars et al. 2014). Genotyping errors derived from null alleles, allele drop-outs, and mutations (due to a high mutation rate) may explain the lower performance of microsatellites (Ellegren 2000, Pompanon et al. 2005, Trọng et al. 2013, Lapègue et al. 2014). Nevertheless, as previously reported (Trọng et al. 2013), the deletion of loci with high null allele frequency did not contribute to the increase in assignment proportions.

Direct exclusion analysis, allowing a 2.6% of SNP genotyping error (2 mismatches), resulted in >96% of the progeny being assigned, a high assignment certainty. This precision

Considerando 50 *loci* como el número mínimo para un análisis de asignación de paternidad, el error estimado (la diferencia en las parejas de padres obtenidas entre métodos) mostró una tendencia de disminución del error en función de la proporción de *loci* con FAM  $\geq 0.3$ , y la proporción del 60% fue la proporción con la que se obtuvo el valor más bajo (2.3%; Fig. 2).

## DISCUSIÓN

Éste es el primer trabajo del conocimiento de los autores acerca de la utilidad de un panel de SNP en *L. vannamei*. Se demostró que la utilización del panel de 76 SNP es adecuado para la identificación de la genealogía o pedigrí del camarón blanco de cultivo. A pesar de una menor variabilidad en comparación con los microsatélites, el panel de SNP presenta mayor certidumbre en las asignaciones de parentesco. La menor resolución de los microsatélites en comparación con los SNP ha sido descrita también para otras especies acuáticas como la tilapia (Trọng et al. 2013) y el camarón tigre, *P. monodon* (Sellars et al. 2014). Las diferencias estriban en errores de genotificación como resultado de la presencia de alelos nulos, la supresión de alelos y las mutaciones, estas últimas derivadas de una relativa alta tasa



**Figure 2.** Percentage of assignment error by probabilistic analysis with 50 loci, as a function of the proportion of loci with minimum allele frequency (MAF)  $\geq 0.3$ .

**Figure 2.** Porcentaje de error de asignación mediante el análisis probabilístico con 50 *loci*, en función de la proporción de *loci* con frecuencia alélica mínima (MAF)  $\geq 0.3$ .

requires previous knowledge of the mating plan used to obtain the progeny, that is, knowing which male mated with which female. Nevertheless, in those cases where the mating plan is unknown, use of the SNP panel is still feasible because likelihood analysis is capable of identifying the families with an error of less than 3% and 95% confidence. Trøng et al. (2013) reported similar assignment levels with both analytical methods (direct and probabilistic exclusion) for a cultivated tilapia population.

Even though the number of loci of the SNP panel (76) is slightly higher than the number of loci used by Sellars et al. (2014) for *P. monodon* (53 and 63 SNPs), the results indicate that, when the proportion of loci with MAF  $> 0.3$  in the panel is 60%, 50 loci would be enough for the correct assignment of more than 98% of the progeny. These results contrast with those obtained by Lapègue et al. (2014) with oysters and Trøng et al. (2013) with tilapia who used panels of between 122 and 384 loci. In fact, Lapègue et al. (2014) recommends a minimum of 150 loci with a mean MAF of 0.3 to obtain a robust assignment.

Assignment probability is a function of the population allele frequencies (Kalinowski et al. 2007); therefore, related individuals within the broodstock, which share more alleles than unrelated individuals, will decrease the probability of correct assignment. In this regard, Trøng et al. (2013) obtained a low to moderate precise assignment in tilapia as a result of high genetic similarity among individuals, due to more than 8 generations of genetic selection. In this study, the broodstock belonged to a fifth generation of selection for which inbreeding has been maintained at a low level (1.1–2.4%; Perez-Enriquez et al., unpublished data). Although this low inbreeding seems to have no effect on the assignment precision when using SNPs, some microsatellite loci might have been affected, increasing the frequency of null alleles. Information on broodstock management history is always recommended when genetically tracing pedigree.

de mutación de los microsatélites (Ellegren 2000, Pompanon et al. 2005, Trøng et al. 2013, Lapègue et al. 2014). Sin embargo, de manera similar a lo reportado por Trøng et al. (2013), la eliminación de *loci* con frecuencias elevadas de alelos nulos no contribuyó a incrementar la proporción de asignación.

El análisis de exclusión directa con un error permitido de genotipificación de los SNP del 2.6% (2 no coincidencias) resultó en más del 96% de la progenie con asignación a una sola pareja, lo cual brinda una alta certidumbre en la asignación. Este nivel de precisión requiere del conocimiento del plan de apareamiento bajo el cual se obtuvo la progenie, es decir, qué hembra fue fecundada por qué macho. Sin embargo, aun si se desconoce el plan de apareamiento, el uso del panel de SNP es factible ya que el análisis probabilístico es capaz de identificar las familias con un 95% de confianza y un error menor que 3%. Trøng et al. (2013) obtuvieron niveles similares de asignación de parentesco con ambos métodos analíticos (exclusión directa y probabilística) para una población de cultivo de tilapia.

Aun cuando el número de *loci* del panel de SNP (76) es ligeramente mayor que el número utilizado por Sellars et al. (2014) para *P. monodon* (56 y 63 SNP), los resultados indican que al utilizar un panel con una proporción de *loci* con FAM  $\geq 0.3$  de 60%, 50 *loci* serían suficientes para obtener una asignación correcta para más del 98% de la progenie. Estos resultados difieren de lo obtenido por Lapègue et al. (2014) para ostras y Trøng et al. (2013) para tilapia, quienes utilizaron paneles de entre 122 y 384 *loci*. De hecho, Lapègue et al. (2014) recomiendan un mínimo de 150 *loci* con un FAM promedio de 0.3 para obtener una asignación suficientemente potente.

La probabilidad de asignación está en función de las frecuencias alélicas de la población (Kalinowski et al. 2007); por lo tanto, los individuos emparentados dentro del lote de reproductores, los cuales comparten más alelos que individuos no emparentados, disminuirán la probabilidad de asignación correcta. En este sentido, Trøng et al. (2013) señalan que una de las razones de obtener una baja a moderada asignación en la población de tilapia que analizaron se debe a los altos niveles de parentesco entre individuos derivados de un programa de selección por talla por más de 8 generaciones. En el presente estudio, el lote se encontraba en la quinta generación de selección, para la cual la endogamia se ha mantenido en un nivel reducido de 1.1–2.4% (Perez-Enriquez et al., datos no publicados). Aunque esta baja endogamia pudiera no tener efecto en la precisión de asignación con los SNP, algunos microsatélites parecen haber sido afectados, ya que incrementó la frecuencia de alelos nulos. Cuando se realizan análisis de pedigrí es recomendable mantener registros adecuados del manejo histórico del lote de reproductores.

La certidumbre en la asignación de pedigrí con SNP abre posibilidades para realizar estimaciones de heredabilidad, interacción  $G \times A$  y manejo reproductivo en condiciones de cultivo en ambientes comunes y bajo condiciones en las

The certainty in pedigree assignment using SNP markers opens the possibility for experimental studies in a common environment for estimating heritability,  $G \times E$  interactions, and broodstock management when physical tagging is not possible or efficient. With the analysis platforms presented here, the use of SNPs in shrimp is suggested as a higher benefit/cost alternative (cost per sample, genotyping rate, analysis time) than microsatellites.

## ACKNOWLEDGMENTS

This study was funded by CIBNOR (grant 985-0). Ira Fogel at CIBNOR provided editorial services. The authors are grateful to Acuacultura Mahr SA de CV and Blue Genetics Mexico SA de CV for providing broodstock and progeny.

## REFERENCES

- Andriantahina F, Liu X, Feng T, Xiang J. 2013. Current status of genetics and genomics of reared penaeid shrimp: Information relevant to access and benefit sharing. *Mar. Biotechnol.* 15(4): 399–412.  
<http://dx.doi.org/10.1007/s10126-013-9500-9>
- Christie MR, Johnson DW, Stallings CD, Hixon MA. 2010. Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish. *Mol. Ecol.* 19(5):1042–1057.  
<http://dx.doi.org/10.1111/j.1365-294X.2010.04524.x>
- Cruz P, Ibarra AM, Mejía-Ruiz H, Gaffney PM, Perez-Enriquez R. 2004. Genetic variability assessed by microsatellites in a breeding program of Pacific white shrimp (*Litopenaeus vannamei*). *Mar. Biotechnol.* 6(2): 157–164.  
<http://dx.doi.org/10.1007/s10126-003-0017-5>
- Cruz P, Mejía-Ruiz CH, Perez-Enriquez R, Ibarra AM. 2002. Isolation and characterization of microsatellites in Pacific white shrimp *Penaeus (Litopenaeus) vannamei*. *Mol. Ecol. Notes* 2(3): 239–241.  
<http://dx.doi.org/10.1046/j.1471-8286.2002.00211.x-i2>
- Dong S, Kong J, Zhang T, Meng X, Wang R. 2006. Parentage determination of Chinese shrimp (*Fenneropenaeus chinensis*) based on microsatellite DNA markers. *Aquaculture* 258(1–4): 283–288.  
<http://dx.doi.org/10.1016/j.aquaculture.2006.04.044>
- Ellegren H. 2000. Microsatellite mutations in the germline: Implications for evolutionary inference. *Trends Genet.* 16(12): 551–558.  
[http://dx.doi.org/10.1016/S0168-9525\(00\)02139-9](http://dx.doi.org/10.1016/S0168-9525(00)02139-9)
- [FAO] Food and Agriculture Organization of the United Nations. 2014. The State of World Fisheries and Aquaculture 2014. Rome, 223 pp.
- Freitas PD, Galetti Jr PM. 2005. Assessment of the genetic diversity in five generations of a commercial broodstock line of *Litopenaeus vannamei* shrimp. *Afr. J. Biotechnol.* 4: 1362–1367.
- Gjedrem T. 2010. Selection and Breeding Programs in Aquaculture. Springer, Dordrecht, The Netherlands.
- Gjedrem T, Robinson N, Rye M. 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture* 350–353: 117–129.  
<http://dx.doi.org/10.1016/j.aquaculture.2012.04.008>
- Jerry DR, Evans BS, Kenway M, Wilson K. 2006. Development of a microsatellite DNA parentage marker suite for black tiger shrimp *Penaeus monodon*. *Aquaculture* 255(1–4): 542–547.  
<http://dx.doi.org/10.1016/j.aquaculture.2005.11.026>
- Jerry DR, Preston NP, Crocos PJ, Keys S, Meadows JRS, Li Y. 2004. Parentage determination of Kuruma shrimp *Penaeus (Marsupenaeus) japonicus* using microsatellite markers (Bate). *Aquaculture* 235(1–4): 237–247.  
<http://dx.doi.org/10.1016/j.aquaculture.2004.01.019>
- Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16(5): 1099–1106.  
<http://dx.doi.org/10.1111/j.1365-294x.2007.03089.x>
- Karlsson S, Moen T, Lien S, Glover KA, Hindar K. 2011. Genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Mol. Ecol. Resources* 11(S1): 247–253.  
<http://dx.doi.org/10.1111/j.1755-0998.2010.02959.x>
- Lapègue S, Harrang E, Heurtebise S, Flahauw E, Donnadiou C, Gayral P, Ballenghien M, Genestout L, Barbotte L, Mahla R, Haffray P, Klopp C. 2014. Development of SNP genotyping arrays in two shellfish species. *Mol. Ecol. Resources* 14(4): 820–830.  
<http://dx.doi.org/10.1111/1755-0998.12230>
- Meehan D, Xu Z, Zuniga G, Alcivar-Warren A. 2003. High frequency and large number of polymorphic microsatellites in cultured shrimp, *Penaeus (Litopenaeus) vannamei* (Crustacea: Decapoda). *Mar. Biotechnol.* 5(4): 311–330.  
<http://dx.doi.org/10.1007/s10126-002-0092-z>
- Perez-Enriquez R, Hernández-Martínez F, Cruz P. 2009. Genetic diversity status of white shrimp *Penaeus (Litopenaeus) vannamei* broodstock in Mexico. *Aquaculture* 297(1–4): 44–50.  
<http://dx.doi.org/10.1016/j.aquaculture.2009.08.038>
- Perez-Enriquez R, Takagi, M, Taniguchi N. 1999. Genetic variability of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. *Aquaculture* 173(1–4): 413–423.  
[http://dx.doi.org/10.1016/S0044-8486\(98\)00469-4](http://dx.doi.org/10.1016/S0044-8486(98)00469-4)
- Pompanon F, Bonin A, Bellemain E, Taberlet P. 2005. Genotyping errors: Causes, consequences and solutions. *Nat. Rev. Genet.* 6(11): 847–859.  
<http://dx.doi.org/10.1038/nrg1707>
- Sellers MJ, Dierens L, McWilliam S, Little B, Murphy B, Coman G, Barendse W, Henshall J. 2014. Comparison of microsatellite and SNP DNA markers for pedigree assignment in black tiger shrimp, *Penaeus monodon*. *Aquacult. Res.* 45(3): 417–426.  
<http://dx.doi.org/10.1111/j.1365-2109.2012.03243.x>



- Trọng TQ, van Bers N, Crooijmans R, Komen H. 2013. A comparison of microsatellites and SNPs in parental assignment in the GIFT strain of Nile tilapia (*Oreochromis niloticus*): The power of exclusion. *Aquaculture* 388–391: 14–23. <http://dx.doi.org/10.1016/j.aquaculture.2013.01.004>
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4(3): 535–538. <http://dx.doi.org/10.1111/j.1471-8286.2004.00684.x>
- Vandeputte M, Mauger S, Dupont-Nivet M. 2006. An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Mol. Ecol. Resources* 6(1): 265–267. <http://dx.doi.org/10.1111/j.1471-8286.2005.01167.x>
- Vela-Avitúa S, Montaldo HH, Márquez-Valdelamar L, Campos-Montes GR, 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 J. Biotechnol.* 16(4). <http://dx.doi.org/10.2225/vol16-issue4-fulltext-11>
- Yoshida K, Awaji M. 2000. Improvement of polymerase chain reaction condition to detect polymorphic dinucleotide repeat microsatellite DNA marker in the puffer fish *Fugu rubripes*. *Fish. Sci.* 66(2): 397–399. <http://dx.doi.org/10.1046/j.1444-2906.2000.00061.x>
- Yue GH, Xia JH. 2014. Practical considerations of molecular parentage analysis in fish. *J. World Aquacult. Soc.* 45(2): 89–103. <http://dx.doi.org/10.1111/jwas.12107>

*Received June 2016,  
accepted September 2016.*