REVISTA BIO CIENCIAS
http://revistabiociencias.uan.edu.mx
https://doi.org/10.15741/revbio.11.e1534

ISSN 2007-3380

Original article / Artículo original

Genetic characterization of aquaculture species by low-density SNP panels

Caracterización genética de especies acuícolas mediante paneles de SNPs de baja densidad

Max-Aguilar, A. [1], *Mendoza-Carrión, G.* [1], *Escobedo-Fregoso, C.* [1,2], *Pérez-Enríquez, R.*[1]*

[1] *Centro de Investigaciones Biológicas del Noroeste, S.C. Instituto Politécnico Nacional, 195, Col. Playa Palo de Santa Rita Sur. 23096, La Paz, B.C.S., México.*
[2] *Consejo Nacional de Humanidades, Ciencia y Tecnología, Av. Insurgentes Sur 1582, Col. Crédito Constructor, Benito Juárez, C.P. 03940, CDMX, México.*

Please cite this article as/Como citar este artículo: Max-Aguilar, A., Mendoza-Carrión, G., Escobedo-Fregoso, C., Pérez-Enríquez, R. (2024). Genetic characterization of aquaculture species by low-density SNP panels. *Revista Bio Ciencias*, 11, e1534. https://doi.org/10.15741/revbio.11.e1534

Article Info/Información del artículo
Received/Recibido: June 17th 2023.
Accepted/Aceptado: March 22th 2024.
Available on line/Publicado: April 22th 2024.

ABSTRACT

The aquaculture industry has high growth with a projection of reaching a worldwide production of 106 million tons of fish and shellfish in 2030. For this, the implementation of genetic management and selection programs, based on the monitoring of the genetic variability, inbreeding, and pedigree of the cultivated stocks are needed. In this study, the 2bRAD methodology was implemented to genetically characterize aquaculture species through low-density SNP (Single Nucleotide Polymorphisms) panels of 100-500 genetic markers. The 2bRAD technique, implemented with the enzyme *Bcg*I and using four selective bases in the adaptors, resulted in low-density panels of 114 and 159 SNPs for the Pacific oyster *Crassostrea gigas* and almaco jack *Seriola rivoliana*, respectively. The panels for these species were validated with parentage and paternity assignment tests, and thus, they are suitable for studying genetic variability and pedigree follow-up in cultivated stocks. For the whiteleg shrimp *Penaeus* (*Litopenaeus*) *vannamei*, the medium- and high-density panels (2,874 and 19,105 SNPs, respectively) have potential applications for studies of marker-associations to traits of interest. The developed 2bRAD platform potentially applies to other cultivated marine fish species such as red snapper, rose spotted snapper, and totoaba.

KEY WORDS: Genetic diversity, Genetic markers, Parentage analysis, Genetic selection.

***Corresponding Author:**

Ricardo Pérez-Enríquez. Programa de Acuicultura. Centro de Investigaciones Biológicas del Noroeste, S.C. Instituto Politécnico Nacional, 195, Col. Playa Palo de Santa Rita Sur. 23096, La Paz, B.C.S., México. Telefono: (612) 123 8504.
E-mail: rperez@cibnor.mx

Revista Bio ciencias
ISSN 2007-3380

RESUMEN

El sector acuícola presenta un elevado crecimiento con una proyección a alcanzar una producción mundial de 106 millones de toneladas de pescados y mariscos en el 2030. Para ello, se requiere la implementación de programas de manejo y selección genética basados en el monitoreo de la variabilidad genética, la endogamia y el pedigrí de los lotes de cultivo. En este estudio se implementó la metodología 2bRAD para caracterizar genéticamente especies de cultivo acuícola con paneles de baja densidad de 150 a 500 marcadores genéticos tipo SNPs (Polimorfismos de Nucleótido Simple). La implementación de la técnica 2bRAD con corte del DNA con la enzima *Bcg*I y el uso de adaptadores con cuatro bases selectivas, generó paneles de baja densidad de 114 y 159 SNPs para el ostión del Pacífico *Crassostrea gigas* y el jurel *Seriola rivoliana*, respectivamente. Los paneles para estas especies se validaron con pruebas de parentesco, por lo que son adecuados para estudios de variabilidad genética y seguimiento del pedigrí de lotes de cultivo. Para el camarón *Penaeus* (*Litopenaeus*) *vannamei* los paneles de mediana y alta densidad (2,874 y 19,105 SNPs, respectivamente), tienen aplicación potencial para estudios de asociación amplia al genoma de marcadores a caracteres de interés. La plataforma 2bRAD desarrollada es potencialmente aplicable a otras especies de peces marinos de cultivo como huachinango, pargo lunarejo y totoaba.

PALABRAS CLAVE: Variabilidad genética, Marcadores genéticos, Análisis de parentesco, Selección.

Introduction

Aquaculture is an activity of great importance, with the world's production of aquatic animals of more than 87.5 million tons in 2020 and a projection to reach 106 million tons in 2030 (FAO, 2022). To achieve this increment, strategies that promote improvement through management and genetic selection are required (FAO, 2019).

In Mexico, the cultivation of marine species is centered on shrimp *Penaeus* (*Litopenaeus*) *vannamei* and oyster *Crassostrea gigas* with average annual productions (2018-2021) of 181,637 and 21,626 tons, respectively (CONAPESCA, 2021). Culturing marine fishes such as red snapper, rose spotted snapper, and totoaba is still incipient but has potential development (Escárcega Rodríguez, 2018; Amparo Venegas, 2019).

The production of larvae and juveniles that are capable of surviving over diverse culture conditions and that show adequate growth during grow-out depends on breeding stocks that

Revista Bio ciencias
ISSN 2007-3380

are based on management plans that ensure the maintenance of genetic diversity that avoid inbreeding associated problems (Gjerde, 2010). For that purpose, the follow-up of the genetic lines is required, focusing in particular on the determination of the pedigree of the individuals (Vandeputte & Haffray, 2014). Single Nucleotide Polymorphisms (SNPs) as genetic markers are a helpful tool for these objectives, which has been demonstrated in aquaculture species (Maqsood & Ahmad, 2017).

The development of technologies of massive sequencing or “next generation sequencing” (NGS) has opened the possibility of widening the study of diverse species, using modifications with the technique known as Restriction-site Associated DNA Sequencing or RAD-Seq (Robledo *et al*., 2018). This strategy consists of reducing the size and complexity of the genome by digestion by restriction enzymes, massive sequencing, and the characterization of SNPs without the need for a reference genome (Robledo *et al*., 2018). For example, studies in the whiteleg shrimp have been reported with 2bRAD (Lyu *et al*., 2021), SLAF-seq (Peng *et al*., 2020), and NextRAD (Perez-Enriquez *et al*., 2018).

The 2bRAD technique is based on genome digestion with a 2B-type restriction enzyme that identifies a specific sequence and cuts on both sides of it, resulting in fragments of 30-36 bases long (Wang *et al*., 2012), offering the advantage of obtaining fragments of specific size and minimizing a biased selection of fragments due to their size (Hernandez-Castro *et al*., 2017). Adaptors and barcodes are joined to these fragments to identify the individuals after sequencing. Barbanti *et al*. (2020) reported that the inclusion of semi-selective adaptors decreased the quantity of identified SNPs and allowed the corroboration of their feasibility in two species for which no reference genome existed before. This result means that the addition of highly selective adaptors can be used for the development of low-density (< 200 SNPs) panels, with which it is possible to determine parentage with high confidence in, for example, shrimp (Perez-Enriquez & Max-Aguilar, 2016) and oysters (Lapègue *et al*., 2014). In this regard, it is proposed that by modifying the protocol of Barbanti *et al*. (2020) by using highly selective adaptors, the number of markers can be adjusted to fit medium or low-yield sequencing facilities (MiniSeq o MiSeq, Illumina).

In the present study, a methodology for the genetic characterization of three species of aquaculture relevance in northwestern Mexico employing the 2bRAD technique is implemented.

Materials y Methods

The implementation of the 2bRAD methodology consisted of three steps: an *in silico* analysis for the definition of the most adequate restriction enzyme, the standardization of the technique with the selected enzyme, and its validation by the analysis of genetic diversity and/or parentage. The premise for the analysis consisted of obtaining a balance between an adequate number of markers for the genetic characterization and the highest amount of individuals that can be analyzed in a single sequencing run.

Revista Bio ciencias
ISSN 2007-3380

In silico bioinformatic analyses for 2B-type enzyme definition.

The *in silico* identification of four 2B-type restriction enzymes (*Bcg*I, *Bsa*XI, *Alf*I, and *Csp*CI; mat.compl1, mat.compl2) was done on the genomes of three aquaculture important species: whiteleg shrimp *Penaeus* (*Litopenaeus*) *vannamei* (GenBank accession no.: ASM378908v1; Zhang *et al*., 2019), Pacific oyster *Crassostrea gigas* (GenBank accession no.: cgigas_uk_roslin_v1 NC_047559-NC_047568; Peñaloza *et al*., 2021), and almaco jack *Seriola rivoliana* (GenBank accession no.: GCA_002994505.1 ASM299450v1; Seetharam, 2018). The genomes of the species were obtained from the GenBank database (National Center for Biotechnology Information https://www.ncbi.nlm.nih.gov/). They were downloaded to the local server at the Genomics and Bioinformatics Laboratory of Centro de Investigaciones Biológicas del Noroeste, S.C. (CIBNOR).

The number of recognition sites of each enzyme on the genomic sequencing on sense- and anti-sense directions were counted with search and line-count commands (mat.compl3). The next step consisted of estimating the remaining sites depending on the number of selective bases included in the adaptors.

The proportion of variable sites relative to the number of recognition sites in the whiteleg shrimp was estimated. A previously generated database of genomic analysis of 88 individuals of several origins was used (Perez-Enriquez *et al*., 2018). This database was depurated only to contain fragments with the recognition sites of the enzyme *Bcg*I (zgrep command) and aligned to the complete genome of Zhang *et al*. (2019) following the procedure in Perez-Enriquez *et al*. (2018).

2bRAD technique standardization.

The standardization of the 2bRAD technique was done on a reduced number of samples of the species of interest to verify the number of markers to be obtained with the selected enzyme and with adaptors with four selective bases. This step was divided into three sections: sample obtaining, library construction and sequencing, and data analysis.

Sample obtaining and DNA extraction. Tissue and fin samples (depending on the species) of each aquaculture species (shrimp, oyster, almaco jack, n = 8 each) were obtained. DNA extractions were done with the Qiagen DNA extraction kit (QIAgen), following manufacturer instructions. DNA concentration of each sample was measured in a Nanodrop (Thermo Scientific) spectrophotometer and adjusted to 25 ng/µl. Pools of 4 DNA samples each were made: two from oyster (Cg-pool-1 and Cg-pool-2), two from almaco jack (Sr-pool-1 and Sr-pool-2), and four for shrimp (Lv-sel2-pool-1, Lv-sel2-pool-2, Lv-sel4-pool-1, and Lv-sel4-pool-2).

Library construction. 2bRAD libraries were prepared following the protocol of Matz *et al*. (2018) with modifications (see detailed protocol in mat.compl4). After the DNA was digested with the enzyme *Bcg*I (defined in the previous step), the adaptors were ligated (adaptors without selective bases for the two oyster pools, adaptors without selective bases for the two almaco

Revista Bio ciencias
ISSN 2007-3380

jack pools, one adaptor with two selective bases for two shrimp pools and two adaptors with two selective bases for the other two shrimp pools), PCR amplification and barcode adding, libraries pooling and purification of a 180 bp fragment.

Sequencing and data analysis. The library was prepared for sequencing (mat.compl4) in a MiSeq (Illumina) using a MiSeq Reagent Kit v3 (150-cycle) sequencing cell at CIBNOR. Sequencing reads quality was analyzed with the software FastQC version 0.11.2 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The Trimmomatic software (Bolger *et al*., 2014) was used to delete low-quality and contaminating sequences (adaptors) and separate the sequences for each species according to the assigned barcodes. The sequences were aligned with BWA (Li & Durbin, 2009) to each species' reference genome, and the alignment percentage was calculated with Samtools (Li *et al*., 2009). The bam files of each pool were used for SNP discovery with samtools, Bcftools, and Vcftools (Li *et al*., 2009; Danecek *et al*., 2011). For oyster and almaco jack, *in silico* analyses were done to estimate the number of remaining SNPs in the case of using adaptors with selective bases.

2bRAD technique validation.

The validation of the 2bRAD technique and the SNP panel obtained for the genetic characterization and parentage analyses were done by analyzing 96 samples of almaco jack *S. rivoliana*, oyster *C. gigas*, red snapper *Lutjanus peru*, rose spotted snapper *L. guttatus* (n = 8) and totoaba *Totoaba macdonaldi* (mat.compl5). Three known families of cultivated oyster (n = 24, 6 broodstock and 18 progenies), two almaco jack spawning groups from two breeding tanks (n = 45, 9 broodstock and 13 progenies from one culture tank, and eight broodstock and 15 progenies from the second tank) were used. In addition, wild samples from other aquaculture-important species, such as red snapper *Lutjanus peru* (n = 8) from the state of Sinaloa (Reguera-Rouzaud *et al*., 2021), *L. guttatus* (n = 8) from Panamá (Rivera-Lucero, 2020), and totoaba *Totoaba macdonaldi* (n = 8) from the Gulf of California (Valenzuela-Quiñonez *et al*., 2016), were included.

Library construction and sequencing were done with the procedure described in the supplementary material (mat.compl4), using the two adaptors with two selective bases each. The supplementary material (mat.compl6) contains the adaptor sequences and primers used for library construction and the position in a 96-well plate of the individual identification barcodes. Sequencing was done in a MiSeq (Illumina) using a MiSeq Reagent Kit v3 (150-cycle) sequencing cell. Sequence files were analyzed as described above throughout the demultiplexing, alignment, and SNP discovery process. In the case of *Lutjanus* and *Totoaba*, the sequence analysis was done with Stacks (Catchen *et al*., 2011) because there are no complete genomes of these species in the GenBank. Stacks software allows the creation of pseudo-reference genomes using the sequences of each species. The SNP catalog creation step was using the *denovo_map.pl* function, with the following parameters: minimum depth to form a stack m = 3, maximum number of differences among stacks M = 3, and maximum number of differences among putative loci n =3.

Revista Bio ciencias
ISSN 2007-3380

SNP genotypes vcf files of each species were filtered with vcftools (Danecek *et al*., 2011) with the following parameters: removal of *indels*, retention of two alleles, SNPs represented in at least 80% of the individuals (*max-miss* = 0.8), and a minimum read quality of Q = 30. For oyster, the *max-miss* parameter was 0.9, and only one SNP per fragment was retained. Loci with two alleles in at least one lot or varying relative to the reference genome were considered polymorphic. Thus, it was not necessary to calculate the minimum allele frequency of the minor allele (MAF). The filtered SNP files were converted into GenePop format with the software PGDSpider2 (Lischer & Excoffier, 2012). They were imported from GenAlEx (Peakall & Smouse, 2012) to analyze Hardy-Weinberg equilibrium and genetic variability (number of alleles per locus, heterozygosity) by species and lot (i.e., broodstock and progenies in oyster and almaco jack). Deviating loci from Hardy-Weinberg equilibrium ($p < 0.05$) were excluded. In oyster and almaco jack, parentage analyses were done with the software Colony (Jones & Wang, 2010) and Cervus (Kalinowski *et al*., 2007).

To implement a standardized genotyping analysis strategy for the subject species and the obtained SNPs, the number of individuals (*Nej*) that can be analyzed in a single sequencing run of a MiSeq (Illumina) platform was estimated with the following equation:

$$Nej = Nlps * (1 - Mer) / Prof / Nsr \quad \text{(eq. 1)}$$

Where:

Nlps: Number of reads per sequencing cell; *Mer*: Reads depletion percentage derived from the PhiX control and low-quality reads; *Prof*: Desired read sequencing depth; *Nsr*: Approximate number of recognition sites of the 2B-type enzyme in the species' genome after using selective bases.

Results and Discussion

In silico bioinformatic analysis for 2B-type enzyme selection.

The average number of recognition sites for each enzyme among species corresponded, in ascending order, to *Csp*CI*, Bcg*I*, Alf*I, and *Bsa*XI with 53,328, 116,354, 118,873, and 278,549, respectively (Table 1). A wide heterogeneity in the number of recognition sites per million bases for each enzyme and species is observed, suggesting that there is no random distribution of the nucleotidic bases in these species. For instance, the number of recognition sites per million bases of *Bcg*I is 30% higher in shrimp than in oysters and 27% than in almaco jack; however, the contrary is true for *Bsa*XI*, Alf*I, and *Csp*CI. In the 2bRAD study of Hernandez-Castro *et al*. (2017) with the 702 Mb genome of *Rhodnius prolixus* (vector insect the Chagas disease), 146, 292, 100 sites per million bases for each of the enzymes *Bcg*I*, Alf*I, and *Csp*CI, were respectively obtained.

Revista Bio ciencias
ISSN 2007-3380

This result indicates that there is not an associated number of recognition sites dependent on the taxonomy level.

Under the consideration that the genome size is similar between almaco jack and oyster and that the BcgI enzyme results in a similar number of recognition sites in both species (between 65 K and 70 K, respectively, Table 1), BcgI was the selected enzyme for subsequent analyses. Even though the CspCI enzyme would also be an adequate enzyme due to a more restricted variability, the random variation in the number of bases on both sides of the sequence (New England Biolabs, pers. comm.) derives in an undesired uncertainty during the digestion step.

Table 1. Number of recognition sites of the 2B restriction enzymes *Bcg*I, *Bsa*XI, *Alf*I, and *Csp*CI for the *in silico* analysis of the public-access genomes of whiteleg shrimp, Pacific oyster, and almaco jack.

Species and genome size (Gb)		Enzyme			
		BcgI	***BsaXI***	***AlfI*** [1]	***CspCI***
Whiteleg shrimp *Penaeus* (*Litopenaeus*) *vannamei* (1.6)	Fwd	106,530	205,589	136,798	31,678
	Rev	107,988	138,791	NA	29,129
	Total	214,518	344,380	136,798	60,807
	Prop	134	215	85	38
Pacific oyster *Crassostrea gigas* (0.68)	Fwd	31,891	75,448	62,887	19,631
	Rev	33,394	99,988	NA	17,420
	Total	65,285	175,436	62,887	37,051
	Prop	96	257	92	54
Almaco jack *Seriola rivoliana* (0.66)	Fwd	32,835	190,101	156,933	32,751
	Rev	36,425	125,729	NA	29,376
	Total	69,260	315,830	156,933	62,127
	Prop	105	479	238	94
Mean		116,354	278,549	118,873	53,328

Fwd: result in forward direction; Rev: result in the reverse direction; Total: sum of both; Prop: proportion of recognition sites per million nucleotidic bases; Gb: billion pair bases.

[1] The forward and reverse sequence of the recognition site is the same, thus only Fwd and total values are given

Source: Own elaboration

The average reduction in the number of recognition sites as a function of the adaptors with one to four selective bases is relatively homogenous, independently of the species and the enzyme, with 22.9 %, 5.4 %, 1.3 %, and 0.3 %, respectively. In the case of *Bsa*XI, the reduction

Revista Bio ciencias
ISSN 2007-3380

with five selective bases is 0.1 %. However, when the number of remaining sites, as a function of the selective bases, the number is higher when the bases are A or T (data not shown). This result implies that, depending on the application for which the genetic characterization will be applied, an adequate number of selective bases can be estimated. That is to say that if the study implies an association analysis of markers to a relevant genetic trait (known as GWAS), the number of bases could be 0, 1, or 2, but if the study is about parentage testing to simultaneously analyze a large number of individuals, 3 or 4 selective bases would be recommended.

The analysis with the shrimp database designed by Perez-Enriquez *et al*. (2018) to determine proportion variable sites (SNPs) within the fragments with *Bcg*I recognition sites indicated that, in a random sample of 50 fragments, 65% showed at least one SNP. Thus, the *in silico* analysis indicates that using the *Bcg*I enzyme without selective bases would result in 214,518 recognition sites and, thus, approximately 139,437 SNPs. This number is an overestimation compared to the 94,113 SNPs (after filtering) reported by Lyu *et al*. (2021). If, during the digestion step, the adaptors are added with two or four selective bases, the expected number of SNPs would be 7,530 and 418, respectively. Under the same parameters, using *Bcg*I and adaptors without selective bases in oyster and almaco jack would result in 42,000 to 45,000 SNPs for each species.

2bRAD technique standardization.

*Bcg*I digested libraries of shrimp, oyster, and almaco Jack pools resulted in a highly variable number of sequencing reads (Table 2), probably due to the preparation process rather than the species' genome nature. The mean sequencing reads alignment against the shrimp's, and almaco jack's reference genomes (68 % and 95 %, respectively) is coincident with expectations according to the size of the genomes of those species in the GenBank (Seetharam, 2018; Zhang *et al*., 2019). For oysters, a species for which the reference genome is complete, there were no differences in the alignment percentages between pools (Peñaloza *et al*., 2021).

In shrimp, the number of filtered SNPs obtained using four selective bases in two adaptors is 2,874 (Table 2). The number of SNPs in oyster and almaco jack was similar before filtering (approx. 35,000), which were reduced by 25 % and 60 % after filtering, respectively (Table 2). The lower reduction in oysters compared to almaco jack might be due to the high heterozygosity reported in oysters (Peñaloza *et al*., 2021). The *in silico* analysis for these two species indicates that using two adaptors with two selective bases each, 300 to 400 SNPs would be obtained. Depending on the objective of a low-density SNP panel (hundreds of SNPs), it was determined that the 2bRAD technique with four selective bases is adequate for oyster and almaco jack. Nevertheless, for shrimp, this technique is not viable due to the high number of markers (2,874); an alternative strategy for the development of a low-density panel (100 to 200 SNPs) in this species is the technique known as GT-Seq (Genotyping-in-thousands by sequencing; Campbell *et al*., 2015) that uses the genomic information and previously identified species-specific SNPs.

Revista Bio ciencias
ISSN 2007-3380

2bRAD technique validation.

Sequencing of five species (almaco Jack, oyster, red snapper, rose spotted snapper, and totoaba) resulted in an average of 165,428 reads per individual and a high alignment percentage against references genomes (almaco jack and oyster) (Table 3). In almaco jack, 3,320 and 160 SNPs were obtained after alignment and filtering, respectively; in oysters, the number of SNPs was 3,731 and 114, respectively (Table 3). The comparative analysis of replicates (almaco jack, red snapper, and oyster) showed a low genotyping rate in the first two (0.6 % and 1.03 %, respectively). In oysters, the genotyping rate after SNP depuration was 2.1 %, although relatively low, which may have implications for the precision of parentage assignment analyses.

The analysis of deviation from Hardy-Weinberg equilibrium indicated that several loci were significant ($p < 0.05$), which were excluded for subsequent analyses (almaco jack: n = 1; oyster n = 3; totoaba n= 7). In red snapper and rose spotted snapper, the loci potentially in disequilibriu m were not excluded due to a low sample size. Nevertheless, the presence of loci under selective pressure cannot be ruled out, as reported in the catfish *Rhamdia quelen* (Ríos *et al*., 2020), for which further analyses with adequate sample sizes for each species will be required.

Table 2. Number of sequencing reads (L), alignment percentage (% P) of the filtered sequences to the reference genome of each pool, and number of obtained SNPs before and after filtering (SNPs-F).

Species and genomic library type	Pool ID number	L	% P	SNPs	SNPs-F [1]
Shrimp with two selective bases	Lv-sel2-pool-1	1,110,928	64.4	26,535	19,105
	Lv-sel2-pool-2	4,532,602	71.9		
Shrimp with four selective bases	Lv-sel4-pool-1	1,893,650	63.4	4,377	2,874
	Lv-sel4-pool-2	1,341,554	72.0		
Oyster without selective bases	Cg-pool-1	4,274,049	76.8	34,387	13,664 (403 [2])
	Cg-pool-2	4,618,046	76.5		
Almaco jack without selective bases	Sr-pool-1	1,125,540	93.9	35,142	26,251 (309 [2])
	Sr-pool-2	1,339,927	96.1		

[1] Filtering criteria: Indels deletion, two-alleles SNPs, Q = 30

[2] Remaining SNPs inferred from the *in silico* analysis if four selective bases were used

Source: Own elaboration

Revista Bio ciencias
ISSN 2007-3380

Table 3. Average number of sequencing reads per individual (NL), percentage of alignment to the reference genome (% Al), and SNPs for each species, totals and retained after filtering.

Species	NL	% Al	Total SNPs	Retained SNPs
Almaco jack *Seriola rivoliana*	208,062	94.5	3,320	160
Pacific oyster *Crassostrea gigas*	150,990	82.5	3,731	114
Red snapper *Lutjanus peru*	86,917	NA	953	434
Rose spotted snapper *Lutjanus guttatus*	180,775	NA	1,631	543
Totoaba *Totoaba macdonaldi*	230,397	NA	335	151

NA: Not applicable because the alignment was done relative to a *de novo* genome generated by Stacks.
Source: Own elaboration

The genetic variability in almaco Jack *S. rivoliana* indicates a high number of alleles per locus (Na) in the broodstock. However, both the effective number of alleles (Ne) and heterozygosity (Ho, He) were medium to low (Table 4). Nevertheless, the negative value of F indicates a lack of inbreeding in this group. There are no data on the genetic variability in cultivated stocks of *S. rivoliana*; however, in the closely related species *S. lalandi*, the genetic variability of cultivated populations estimated with microsatellites indicated medium He values (0.652 − 0.810) (Martínez Matus, 2016).

When the variability between broodstock and progeny of *S. rivoliana* is compared, a reduction in Na, Ho, and He of approximately 7 % is observed, with a minimum variation in n_e and F (Table 4). This genetic variability reduction from one generation to the other is commonly seen in cultivated stocks because not all broodstock individuals participate during the spawning events (e.g., Perez-Enriquez *et al*., 1999). The parentage analysis in the almaco jack cultivation tanks based on 159 SNPs allowed the inference of the familiar structure of full- and half-sibs with high certainty. (the non-exclusion probability of parent pairs was 9×10^{-9}). In addition, it can be observed that, among ten females, three were represented in the progenies, and of nine males (one not genetically analyzed), six participated in the spawning events (Table 5). This result indicates a familiar structure in which more males than females are represented in the following generation, as in other species [e.g., *Seriola lalandi* (Martínez Matus, 2016), *Lutjanus guttatus* (Perez-Enriquez *et al*., 2020)]. The sequences of the 159 SNPs panel for *S. rivoliana* are shown in the supplementary material (mat.compl.7).

Revista Bio ciencias
ISSN 2007-3380

Table 4. Average values of genetic variability of juveniles and broodstock of almaco jack and Pacific oyster, and wild samples of red snapper, rose spotted snapper, and totoaba.

Species	Group	N	Na	n_e	Ho	uHe	F
Almaco jack *Seriola rivoliana*	Broodstock	15.9	1.72	1.23	0.167	0.159	-0.072
	Juveniles	26.5	1.60	1.23	0.157	0.147	-0.073
	Broodstock to juveniles variation		-7%	0%	-6%	-7%	NA
Pacific oyster *Crassotrea gigas*	Broodstock	5.9	1.79	1.36	0.235	0.250	-0.045
	Juveniles	16.7	1.89	1.35	0.224	0.233	0.019
	Broodstock to juveniles variation		5.1%	-0.4%	-4.6%	-7.5%	NA
Red snapper *Lutjanus peru*	Lperu-1	7.7	1.99	1.36	0.230	0.253	0.007
Rose spotted snapper *L. guttatus*	Lgut-1	4.0	1.78	1.38	0.249	0.274	-0.061
	Lgut-2	3.5	1.59	1.32	0.218	0.233	-0.109
Totoaba *Totoaba macdonaldi*	Tot-1	6.7	1.99	1.46	0.327	0.308	-0.123

N: mean number of genotyped individuals; Na: number of alleles per locus; ne: effective number of alleles per locus; Ho: observed heterozygosity; uHe: expected heterozygosity adjusted to sample size; F: inbreeding coefficient. NA: Not applicable.

Source: Own elaboration

In oysters, following the criteria of Perez-Enriquez *et al*. (2024), the genetic variability showed medium to high values in the number of alleles per locus, but medium to low (and lower in progeny than in broodstock) in He (Table 4). However, there was an increment in Na from one generation to the other due to alleles present in the progeny but not in the broodstock; this can be a problem associated with a low sequencing depth in some of the parents in the broodstock or particular mutations, which has been reported in *C. gigas* as a necessary process to counteract against the effect of deleterious alleles (Durland *et al*., 2021). The SNP panel had a non-exclusion probability for the parent pair of 1×10^{-15}. The parentage assessment in oysters with Colony using 114 SNPs showed an agreement of the genetic grouping at the fullsib level in most individuals as expected (Table 5). The sequences of the 114 SNPs panel for *C. gigas* are shown in the supplementary material (mat.compl.8).

Revista Bio ciencias
ISSN 2007-3380

Table 5. Parentage assignment of familiar groups of almaco jack (Sr) and oyster (Cg) obtained with software Colony and Cervus.

Species	Progey ID	Expected breeding couple	Colony assignment			
			GF	P	Sire ID	Dam ID
Almaco jack *Seriola rivoliana*	Sr6-J1, Sr10-J1	ND	1	1	Sr10-TR2	Sr12-TR2
	Sr9-J1	ND	1	1		
	Sr12-J1	ND	1	1	Sr10-TR2	Sr15-TR2
	Sr2-J1, Sr4-J1, Sr5-J1	ND	1	1	Sr11-TR2	Sr15-TR2
	Sr13-J1	ND	1	1		
	Sr1-J1	ND	1	1	Sr14-TR2	Sr12-TR2
	Sr11-J1	ND	1	1	Sr14-TR2	Sr15-TR2
	Sr3-J1, Sr8-J1	ND	1	1	Sr16-TR2	Sr15-TR2
	Sr7-J1	ND	1	1		
	Sr1-J4, Sr2-J4, Sr4-J4, Sr14-J4	ND	2	1	Sr1-TR1	Sr9-TR1
	Sr15-J4	ND	2	1	Sr2-TR1	Sr9-TR1
	Sr7-J4, Sr9-J4, Sr13-J4	ND	2	1	#1	Sr9-TR1
	Sr3-J4, Sr5-J4, Sr6-J4, Sr8-J4, Sr10-J4, Sri11-J4, Sr12-J4	ND	2	1	#1	Sr9-TR1
Pacific Oyster *Crassostrea gigas*	Cg4.1 , Cg4.2, Cg4.3, Cg10.1 , Cg10.2, Cg10.3	Cg.M.P5 - Cg.H.P6	1	0.99	Cg.M.P5	#1
	Cg2.1 , Cg2.2, Cg2.3, Cg14.1 , Cg14.2, Cg14.3	Cg.M.P10 - Cg.H.P9	2	1	Cg.M.P10	Cg.H.P9
	Cg3.1, Cg15.1	Cg.M.P7 - Cg.H.P8	3	1	Cg.M.P7	Cg.H.P8
	Cg3.3	Cg.M.P7 - Cg.H.P8	3	1	Cg.M.P7	Cg.H.P8
	Cg15.3	Cg.M.P7 - Cg.H.P8	3	1	Cg.M.P7	Cg.H.P8
	Cg9.3	Cg.M.P7 - Cg.H.P8	3	1	Cg.M.P7	Cg.H.P8
	Cg9.1	Cg.M.P7 - Cg.H.P8	4	0.99	#1	#2

GF: familiar or spawning group; P: assignment probability; ND: Not determined. Symbols #1 and #2 indicate potential breeders not found in candidates.

Source: Own elaboration

Revista Bio ciencias
ISSN 2007-3380

Continuation

Table 5. Parentage assignment of familiar groups of almaco jack (Sr) and oyster (Cg) obtained with software Colony and Cervus.

Species	Progey ID	Expected breeding couple	Cervus assignment		GA
			Sire ID	Dam ID	
Almaco jack *Seriola rivoliana*	Sr6-J1, Sr10-J1	ND	Sr10-TR2*	Sr12-TR2*	A
	Sr9-J1	ND	Sr10-TR2^{ns}	Sr12-TR2*	B
	Sr12-J1	ND	Sr10-TR2*	Sr15-TR2*	A
	Sr2-J1, Sr4-J1, Sr5-J1	ND	Sr11-TR2*	Sr15-TR2*	A
	Sr13-J1	ND	Sr11-TR2*	Sr15-TR2^{ns}	B
	Sr1-J1	ND	Sr14-TR2*	Sr12-TR2*	A
	Sr11-J1	ND	Sr14-TR2*	Sr15-TR2*	A
	Sr3-J1, Sr8-J1	ND	Sr16-TR2*	Sr15-TR2*	A
	Sr7-J1	ND	Sr16-TR2*	Sr15-TR2^{ns}	B
	Sr1-J4, Sr2-J4, Sr4-J4, Sr14-J4	ND	Sr1-TR1*	Sr9-TR1*	A
	Sr15-J4	ND	Sr2-TR1*	Sr9-TR1*	A
	Sr7-J4, Sr9-J4, Sr13-J4	ND	Sr4-TR1ns	Sr9-TR1*	B
	Sr3-J4, Sr5-J4, Sr6-J4, Sr8-J4, Sr10-J4, Sri11-J4, Sr12-J4	ND	Sr5-TR1*	Sr9-TR1*	B
Pacific Oyster *Crassostrea gigas*	Cg4.1 , Cg4.2, Cg4.3, Cg10.1 , Cg10.2, Cg10.3	Cg.M.P5 - Cg.H.P6	Cg.M.P5*	Cg.H.P6*	B
	Cg2.1 , Cg2.2, Cg2.3, Cg14.1 , Cg14.2, Cg14.3	Cg.M.P10 - Cg.H.P9	Cg.M.P10*	Cg.H.P9*	A
	Cg3.1, Cg15.1	Cg.M.P7 - Cg.H.P8	Cg.M.P7*	Cg.H.P8*	A
	Cg3.3	Cg.M.P7 - Cg.H.P8	Cg.M.P7*	Cg.H.P8^{ns}	B
	Cg15.3	Cg.M.P7 - Cg.H.P8	Cg.M.P7*	Cg.H.P9*	C
	Cg9.3	Cg.M.P7 - Cg.H.P8	Cg.M.P7*	Cg.H.P9^{ns}	C
	Cg9.1	Cg.M.P7 - Cg.H.P8	Cg.M.P7*	Cg.H.P9^{ns}	C

GF: familiar or spawning group; P: assignment probability; ND: Not determined. Symbols #1 and #2 indicate potential breeders not found in candidates.

Source: Own elaboration

Revista Bio ciencias
ISSN 2007-3380

The variability values for the other three species (red snapper, rose spotted snapper, totoaba) showed relatively high for n_e but medium for He (Table 4), which can be due to a small sample size and a lack of a reference genome. The availability of reference genomes for these species will be relevant for a better description of the diversity. Even though Chambers *et al*. (2023) recognize that the lack of reference genomes in non-model species and the small size of the 2bRAD restriction fragments are problematic for genotyping, the phylogenetic reconstruction and the proportion of missing genotypes derived from the technique generate comparable results, and higher repeatability than other ddRAD approaches. The genetic characterization of non-model aquatic species by 2bRAD with one adaptor with one selective base was reported by Barbanti *et al*. (2020). These authors reported that the number of samples for sequencing is balanced between the number of loci, the number of reads per individual derived from the sequencing platform (in that case, HiSeq, Illumina), and the available budget at a suggested depth coverage of 20×.

The present study demonstrated the feasibility of reducing the genome size of several aquaculture species by using two adaptors with two selective bases in each to obtain a sufficient number of loci in a MiSeq (Illumina) sequencing platform. In this regard, the estimated number of individuals that can be analyzed in a single sequencing run is approximately *Nej* = 270 (see eq. 1), considering a sequencing cell of 25 million reads (*Nlps*), 20% of reads depletion (*Mer*), a sequencing depth of 20× (*Prof*), and an estimated number of recognition sites of the 2B enzyme due to the use of selective bases of 3700 (*Nsr*) (Table 3). A costing analysis of such characteristics to determine the applicability in aquaculture companies will be required; nevertheless, this estimation will still have high variability depending on the conditions of the sequencing laboratory and the acquisition costs of the supplies.

Conclusions

The 2bRAD technique was implemented to obtain low-density genetic marker panels (100 to 500 SNPs), which protocol is indistinctly applicable for the genetic characterization of diverse aquaculture species of northwestern Mexico (oyster, almaco jack, red snapper, rose spotted snapper, and totoaba). The study demonstrated that the almaco jack and oyster panels are suitable for parentage and paternity assignment in aquaculture stocks. These panels will be helpful for experimental and commercial broodstock management for both the studied and other aquaculture-relevant species, contributing to their aquaculture development. In addition, the strategy can be applied to the genetic characterization of fisheries resources of commercial interest.

Authors contribution

Conceptualization, RPE, AMA.; methodology development, RPE, AMA, GMC; software management, RPE, CEF.; experimental validation, RPE, AMA, GMC, CEF; analysis of results, RPE, CEF; data management, RPE, AMA, GMC, CEF; manuscript preparation and writing, RPE,

Revista Bio ciencias
ISSN 2007-3380

AMA; drafting, revision, and edition, RPE, AMA, GMC, CEF; project administrator, RPE; funds acquisition, RPE.

All manuscript authors have read and accepted its published version.

Funding

This research was funded by Consejo Sudcaliforniano de Ciencia y Tecnología (COSCYT) number CAR/14/2021.

Acknowledgements

We are grateful for samples donation to Kampachi Farms de México (almaco jack), Acuícola Robles (oyster), Fausto Valenzuela (totoaba), and the technical support of Susana Ávila Álvarez.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

Amparo Venegas, A. O. (2019). Aplicación de la tecnología de biofloc (BFT) al cultivo de *Totoaba macdonaldi*. [Tesis de Maestría en Ciencias, Centro de Investigación Científica y de Educación Superior de Ensenada]. http://cicese.repositorioinstitucional.mx/jspui/handle/1007/2762

Barbanti, A., Torrado, H., Macpherson, E., Bargelloni, L., Franch, R., Carreras, C., & Pascual, M. (2020). Helping decision making for reliable and cost-effective 2b-RAD sequencing and genotyping analyses in non-model species. *Molecular Ecology Resources*, 20(3), 795–806. https://doi.org/10.1111/1755-0998.13144

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015). Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, 15(4), 855-867. https://doi.org/10.1111/1755-0998.12357

Catchen, M. J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, H. J. (2011). Stacks: building and genotyping loci *de novo* from short-reads sequences. *G3: Genes Genomes Genetics*, 1(3), 171–182. https://doi.org/10.1534/g3.111.000240

Chambers, E. A., Tarvin, R. D., Santos, J. C., Ron, S. R., Betancourth-Cundar, M., Hillis, D. M., Matz, M. V., & Cannatella, D. C. (2023). 2b or not 2b? 2bRAD is an effective alternative to

Revista Bio ciencias
ISSN 2007-3380

ddRAD for phylogenomics. *Ecology and Evolution*, 13(3), e9842. https://doi.org/10.1002/ece3.9842

CONAPESCA. (2021). Anuario Estadístico de Acuacultura y Pesca, Edición 2021. Comisión Nacional de Acuacultura y Pesca, México. https://www.gob.mx/conapesca/documentos/anuario-estadistico-de-acuacultura-y-pesca

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., De Pristo, M. A., Handsaker, R., Lunter, G., Marth, G., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools, *Bioinformatics*, 27(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Durland, E., De Wit, P., & Langdon, C. (2021). Temporally balanced selection during development of larval Pacific oysters (*Crassostrea gigas*) inherently preserves genetic diversity within offspring. *Proceedings of the Royal Society B*, 288(1958), 20203223. https://doi.org/10.1098/rspb.2020.3223

Escárcega Rodríguez, S. (2018). Preselección de especies para la piscicultura marina en el Pacífico Sur de México. *Ciencia Ergo Sum*, 25(1), 1-31. https://www.redalyc.org/journal/104/10453975009/html/

FAO. (2019). The State of the World's Aquatic Genetic Resources for Food and Agriculture. FAO Commission on Genetic Resources for Food and Agriculture assessments. Rome. https://www.fao.org/aquatic-genetic-resources/activities/sow/en/

FAO. (2022). The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. Rome, FAO. https://doi.org/10.4060/cc0461en

Gjerde, B. (2010). Design of breeding plans. In Gjedrem, T. Selection and breeding programs in aquaculture. (pp. 173-194). Ed. Springer.

Hernandez-Castro, L.E., Paterno, M., Villacís, A.G., Andersson B., Costales, J.A., De Noia, M., Ocaña-Mayorga, S., Yumiseva, C.A., Grijalva, M.J. & Llewellyn. M.S. (2017). 2b-RAD genotyping for population genomic studies of Chagas disease vectors: *Rhodnius ecuadoriensis* in Ecuador. *PLOS Neglected Tropical Diseases*, 11(7), e0005710. https://doi.org/10.1371/journal.pntd.0005710

Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3), 551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16(5), 1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Lapègue, S., Harrang, E., Heurtebise, S., Flahauw, E., Donnadieu, 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. *Molecular Ecology Resources*, 14(4), 820–830. https://doi.org/10.1111/1755-0998.12230

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Revista Bio ciencias
ISSN 2007-3380

Lischer, H. E. L., & Excoffier L. (2012). PGDSpider: An automated data conversion tools for connecting population genetics and genomics programs. Bioinformatics, 28(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Lyu, D., Yu, Y., Wang, Q., Luo, Z., Zhang, Q., Zhang, X., Xiang, J., & Li, F. (2021). Identification of growth-associated genes by Genome-Wide Association Study and their potential application in the breeding of Pacific white shrimp (*Litopenaeus vannamei*). *Frontiers in Genetics*, 12, 611570. https://doi.org/10.3389/fgene.2021.611570

Maqsood, H.M. & Ahmad, S.M. (2017). Advances in Molecular Markers and Their Applications in Aquaculture and Fisheries. *Genetics of Aquatic Organisms*, 1, 27–41. https://doi.org/10.4194/2459-1831-v1_1_05

Martínez Matus, E. (2016). Diversidad genética y éxito reproductivo del Jurel de Castilla *Seriola lalandi* (Valenciennes, 1833) en cautiverio. [Tesis de Maestría en Ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada]. http://cicese.repositorioinstitucional.mx/jspui/handle/1007/674

Matz, M. V., Treml, E. A., Aglyamova, G. V,. & Bay L. K. (2018). Potential and limits for rapid genetic adaptation to warming in a Great Barrier Reef coral. *PLoS Genetics*, 14(4), e1007220. https://doi.org/10.1371/journal.pgen.1007220

Peakall, R., & Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, 28(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peng, M., Zeng, D., Zhu, W., Chen, X., Yang, C., Liu, Q., Li, Q., Wang, H., Liu, H., Liang, J., Lin, Y., Chen, X., & Zhao, Y. (2020). Construction of a high-density genetic map and identification of quantitative trait loci for nitrite tolerance in the pacific white shrimp (*Litopenaeus vannamei*). *Frontiers in Genetics*, 11, 571880. https://doi.org/10.3389/fgene.2020.571880

Peñaloza, C., Gutierrez, A. P., Eöry, L., Wang, S., Guo, X., Archibald, A. L., Bean, T. P., & Houston, R. D. (2021). A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. *GigaScience*, 10(3), giab020. https://doi.org/10.1093/gigascience/giab020

Perez-Enriquez, R., & Max-Aguilar A. (2016). Pedigree traceability in whiteleg shrimp (*Litopenaeus vannamei*) using genetic markers: A comparison between microsatellites and SNPs. *Ciencias Marinas*, 42(4), 227–235. https://doi.org/10.7773/cm.v42i4.2662

Perez-Enriquez, R., Magallón-Barajas, F.J., & Llera-Herrera, R. (2024). Bases to inform a genetic line of whiteleg shrimp *Penaeus vannamei* of Mexican origin. *Latin American Journal of Aquatic Research*, 52(1)., 150-162. https://doi.org/10.3856/vol52-issue1-fulltext-3066

Perez-Enriquez, R., Robledo, D., Houston, R. D., & Llera-Herrera, R. (2018). SNP markers for the genetic characterization of Mexican shrimp broodstocks. *Genomics*, 110(6), 423–429. https://doi.org/10.1016/j.ygeno.2018.10.001

Perez-Enriquez, R., Takagi, M., & Taniguchi, N. (1999). Genetic variability and pedigree tracing of a hatchery reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. Aquaculture, 173(1-4), 411–421. https://doi.org/10.1016/S0044-8486(98)00469-4

Perez-Enriquez, R., Valadez-Rodríguez, J. A., Max-Aguilar, A., Dumas, S., & Diaz-Viloria, N. (2020). Parental contribution in a cultivated stock for the spotted rose snapper *Lutjanus guttatus* (Steindachner, 1869) estimated by newly developed microsatellite markers. *Latin American Journal of Aquatic Research*, 48(2), 247–256. https://doi.org/10.3856/vol48-

Revista Bio ciencias
ISSN 2007-3380

issue2-fulltext-2424

Reguera-Rouzaud, N., Díaz-Viloria, N., Pérez-Enríquez, R., Espino-Barr, E., Rivera-Lucero, M. I., & Munguía-Vega, A. (2021). Drivers for genetic structure at different geographic scales for Pacific red snapper (*Lutjanus peru*) and yellow snapper (*Lutjanus argentiventris*) in the tropical Eastern Pacific. *Journal of Fish Biology*, 98(5), 1267–1280. https://doi.org/10.1111/jfb.14656

Ríos, N., Casanova, A., Hermida, M., Pardo, B. G., Martínez, P., Bouza, C., & García, G. (2020). Population genomics in *Rhamdia quelen* (Heptapteridae, Siluriformes) reveals deep divergence and adaptation in the neotropical region. *Genes*, 11(1), 109. https://doi.org/10.3390/genes11010109

Rivera-Lucero, M. I. (2020). Análisis comparativo de la composición genética en el pacífico panameño de poblaciones de *Lutjanus peru*, *L. guttatus* y *L. argentiventris*. [Tesis de Licenciatura en Biología Marina. Universidad Marítima Internacional de Panamá].

Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P., & Houston R. (2018). Applications of genotyping by sequencing in aquaculture breeding and genetics. *Reviews in Aquaculture*, 10(3), 670–682. https://doi.org/10.1111/raq.12193.

Seetharam, A. S. (2018). Genome assembly and annotation for Almaco Jack (*Seriola rivoliana*). Abstracts on line. Plant and Animal Genome XXVI. P0271. https://plan.core-apps.com/pag_2018/abstract/ea4431cd14b6570e9d2f2a543c1d4d0a

Valenzuela-Quiñonez, F., De-Anda-Montañez, J. A., Gilbert-Horvath, E., Garza, J. C., & García-De León F. J. (2016). Panmixia in an endangered species: the totoaba (*Totoaba macdonaldi*) in the Gulf of California. *Journal of Heredity*, 107(6), 469–503. https://doi.org/10.1093/jhered/esw046

Vandeputte, M., & Haffray, P. (2014). Parentage assignment with genomic markers: a major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. *Frontiers in Genetics*, 5, 432. https://doi.org/10.3389/fgene.2014.00432

Wang, S., Meyer, E., McKay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, 9(8), 808–810. https://doi.org/10.1038/nmeth.2023

Zhang, X., Yuan, J., Sun, Y., Li, S., Gao, Y., Yu, Y., Liu, C., Wang, Q., Lv, X., Zhang, X., Ma, K.Y., Wang, X., Lin, W., Wang, L., Zhu, X., Zhang, C., Zhang, J., Jin, S., Yu, K., … Xiang, J. (2019). Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nature Communications*, 10(1), 356. https://doi.org/10.1038/s41467-018-08197-4

Revista Bio ciencias
ISSN 2007-3380

Supplementary Material

Mat.compl1. Recognition and cutting sites of the 2B enzymes and an example of the adaptor's ligation.

Mat.compl2. Example of the ligation of adaptors (sequences in blue) with different numbers of selective bases (nucleotidic bases marked in red) of *Bcg*I enzyme-digested fragments (sequences in yellow and orange).

Mat.compl3. Commands used to search for enzyme recognition sites in the *in silico* analyses.

Mat.compl4. 2bRAD protocol from library preparation to sequencing.

Mat.compl5. Sample location of each individual in a 96-well plate. The samples located in positions F6, G6, anf H6 were used as replicates.

Mat.compl6. Adaptors and primer sequences used in library construction, as well as location of the barcodes in the 96-well plate for the identification of each individual.

Mat.compl7. List of the 159 SNPs identified in almaco jack, with their location relative to the reference genome of *Seriola rivoliana* (GenBank accession no.: GCA_002994505.1 ASM299450v1; Seetharam 2018). CHR: Chromosome or linkage group; POS: position; Bases: variants for each SNP.

Mat.compl8. List of the 114 SNPs identified in oyster, with their location relative to the reference genome of *Crassostrea gigas* (GenBank accession no.: cgigas_uk_roslin_v1 NC_047559-NC_047568; Peñaloza *et al*. 2021). CHR: Chromosome or linkage group; POS: position; Bases: variants for each SNP.

Material 1-8: https://revistabiociencias.uan.edu.mx/index.php/BIOCIENCIAS/libraryFiles/downloadPublic/13

Material 4: https://revistabiociencias.uan.edu.mx/index.php/BIOCIENCIAS/libraryFiles/downloadPublic/11