

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

Genetic characterization of aquaculture species by low-density SNP panels

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.

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 *Bcgl* 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 genética.

*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. Teléfono: (612) 123 8504.

E-mail: rperez@cibnor.mx

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 *BcgI* 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.

Introducción

La acuicultura es una actividad de gran importancia en el mundo con una producción superior a 87.5 millones de toneladas de animales acuáticos en el 2020 y una proyección a alcanzar 106 millones de toneladas en el 2030 (FAO, 2022). Para lograr este incremento, se requiere la instrumentación de estrategias que fomenten el mejoramiento mediante manejo y selección genética (FAO, 2019).

En México, cultivo de especies marinas se centra en el camarón *Penaeus (Litopenaeus) vannamei* y el ostión *Crassostrea gigas* con producciones promedio anuales (2018-2021) de 181,637 y 21,626 toneladas, respectivamente (CONAPESCA, 2021). El cultivo de peces marinos como el huachinango, el pargo lunarejo y la totoaba, es aún incipiente, pero con potencial de desarrollo (Escárcega Rodríguez, 2018; Amparo Venegas, 2019).

La producción de larvas y juveniles que son capaces de sobrevivir a las diversas condiciones de cultivo, y que presenten un adecuado crecimiento durante la engorda, depende de lotes de reproductores basados en planes de manejo que aseguren el mantenimiento de una alta variabilidad genética que evite los problemas asociados con la endogamia (Gjerde,

2010). Para ello, se requiere del seguimiento de las líneas de cultivo, enfocándose de manera particular en la determinación del pedigrí de los ejemplares (Vandeputte & Haffray, 2014). El uso de Polimorfismos de Nucleótido Simple (SNPs por sus siglas en inglés) como marcadores genéticos es una herramienta útil para estos fines y ha sido demostrada en especies acuícolas (Maqsood & Ahmad, 2017).

El desarrollo de las tecnologías de secuenciación masiva o de “próxima generación” (NGS por sus siglas en inglés) ha abierto la posibilidad de ampliar el estudio de diversas especies, utilizando modificaciones de la técnica conocida como Secuenciación de DNA asociada a sitios de restricción (Restriction-site Associated DNA Sequencing o RAD-Seq por sus siglas en inglés) (Robledo *et al.*, 2018). Esta estrategia consiste en la reducción del tamaño y complejidad del genoma mediante la digestión con enzimas de restricción, secuenciación masiva y la caracterización de SNPs, sin la necesidad de un genoma de referencia (Robledo *et al.*, 2018). A manera de ejemplo, se han reportado estudios en el camarón blanco con 2bRAD (Lyu *et al.*, 2021), SLAF-seq (Peng *et al.*, 2020) y NextRAD (Perez-Enriquez *et al.*, 2018).

La técnica 2bRAD se basa en la digestión del genoma con una enzima de restricción del tipo 2B que identifica una secuencia específica y corta hacia ambos lados de la misma, dando como resultado fragmentos de 30 a 36 bases (Wang *et al.*, 2012), lo que ofrece la ventaja de obtener fragmentos de tamaño específico minimizando una selección sesgada de fragmentos por tamaño (Hernandez-Castro *et al.*, 2017). A estos fragmentos se unen adaptadores y códigos de barras para la posterior identificación de individuos después de la secuenciación. Barbanti *et al.* (2020) reportaron la inclusión de adaptadores semi-selectivos, que redujeron la cantidad de SNPs identificados y les permitió corroborar su factibilidad en dos especies para las cuales no existía genoma de referencia. Lo anterior significa que la adición de adaptadores altamente selectivos puede ser utilizada para desarrollar paneles de marcadores de baja densidad (< 200 SNPs), con los cuales es posible determinar con alta certidumbre el parentesco como, por ejemplo, en camarón (Perez-Enriquez & Max-Aguilar, 2016) y en ostiones (Lapègue *et al.*, 2014). Bajo esta consideración se propone que, modificando el protocolo de Barbanti *et al.* (2020) para adaptadores altamente selectivos, se puede reducir el número de marcadores para que se ajusten a laboratorios con facilidades de secuenciación de bajo-medio rendimiento (MiniSeq o MiSeq, Illumina).

En el presente estudio se implementó una metodología de caracterización genética aplicable a tres especies acuícolas de importancia comercial en el noroeste de México mediante la técnica 2bRAD.

Material y Métodos

La implementación de la metodología 2bRAD consistió de tres etapas: un análisis *in silico* para la definición de la enzima de restricción más adecuada, la estandarización de la técnica con la enzima seleccionada y la validación la misma para análisis de variabilidad genética y/o parentesco. Una premisa para el análisis consistió en la obtención de un balance entre un número

adecuado de marcadores genéticos para la caracterización genética, y la mayor cantidad de muestras para analizarse en una corrida de secuenciación.

Análisis bioinformáticos *in silico* para definición de la enzima tipo 2B.

Se realizó la identificación *in silico* de los sitios de reconocimiento de cuatro enzimas de restricción tipo 2B (*Bcgl*, *BsaXI*, *AflI* y *CspCI*; mat.compl1, mat.compl2) en los genomas de tres especies de importancia acuícola: el camarón blanco *Penaeus (Litopenaeus) vannamei* (No. acceso GenBank: ASM378908v1; Zhang *et al.*, 2019), ostión del Pacífico *Crassostrea gigas* (No. acceso GenBank: cgigas_uk_roslin_v1 NC_047559-NC_047568; Peñaloza *et al.*, 2021) y jurel o pez fuerte *Seriola rivoliana* (No. acceso GenBank: GCA_002994505.1 ASM299450v1; Seetharam, 2018). Los genomas de las especies se obtuvieron de la base de datos GenBank (National Center for Biotechnology Information <https://www.ncbi.nlm.nih.gov/>) y se descargaron en el servidor del Laboratorio de Genómica y Bioinformática del Centro de Investigaciones Biológicas del Noroeste, S.C. (CIBNOR).

Se contabilizó el número de sitios de reconocimiento de cada enzima sobre la secuencia genómica en dirección sentido y anti-sentido con los comandos de búsqueda y conteo de línea (mat.compl3). El siguiente paso consistió en estimar el número de sitios remanentes dependiendo del número de bases selectivas a incluir en los adaptadores.

Se realizó una estimación *in silico* de la proporción de sitios variables respecto al número de sitios de reconocimiento en el genoma del camarón blanco. Se recurrió a una base de datos previa de un análisis genómico en 88 ejemplares de diversos orígenes (Perez-Enriquez *et al.*, 2018). Dicha base de datos se depuró para contener sólo fragmentos con sitios de reconocimiento de la enzima *Bcgl* (comando *zgrep*) y se alineó con el genoma completo del camarón de Zhang *et al.* (2019) siguiendo el procedimiento de Perez-Enriquez *et al.* (2018).

Estandarización de la técnica 2bRAD.

La estandarización de la técnica 2bRAD se realizó con un número reducido de muestras de las especies de interés para verificar el número de marcadores a obtener con la enzima elegida y con adaptadores con cuatro bases selectivas. Esta etapa se dividió en tres secciones: Obtención de muestras, construcción de genotecas y secuenciación y análisis de datos.

Obtención de muestras y extracción de DNA. Se obtuvieron muestras de tejido muscular y aleta (según la especie) de cada una de las especies acuícolas de interés (camarón, ostión y jurel; n = 8 de cada una). Las extracciones de DNA se obtuvieron con el Qiagen DNA extraction kit (QIAGEN), siguiendo el protocolo del fabricante. La concentración de DNA de cada muestra se midió en espectrofotómetro Nanodrop (Thermoscientific) y se ajustó a 25 ng/μl. Se prepararon grupos de 4 muestras de DNA cada uno: dos de ostión (Cg-pool-1 y Cg-pool-2), dos de jurel (Sr-pool-1 y Sr-pool-2) y cuatro de camarón (Lv-sel2-pool-1, Lv-sel2-pool-2, Lv-sel4-pool-1, Lv-sel4-pool-2).

Construcción de genotecas. Las genotecas 2bRAD se prepararon siguiendo el protocolo de Matz *et al.* (2018) con modificaciones (ver protocolo detallado en mat.compl4). Se realizó la digestión del DNA con la enzima *Bcgl* (definida en la etapa anterior), la ligación de adaptadores (dos de ostión con adaptadores sin bases selectivas, dos de jurel con adaptadores sin bases selectivas, dos de camarón con un adaptador con dos bases selectivas y dos de camarón con dos adaptadores con dos bases selectivas), la amplificación y adición de códigos de barra, agrupación de librerías y purificación del fragmento objetivo de 180 pb.

Secuenciación y análisis de datos. La genoteca obtenida se preparó para ser secuenciada (mat.compl4) en el equipo MiSeq (Illumina) con la celda MiSeq Reagent Kit v3 (150-cycle) en CIBNOR. La calidad de las lecturas se analizó con el programa FastQC version 0.11.2 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). El programa Trimmomatic (Bolger *et al.*, 2014) se utilizó para eliminar secuencias de baja calidad y contaminantes (adaptadores), así como para separar las secuencias por especie de acuerdo con los códigos de barra asignados. Las secuencias se alinearon al genoma de referencia de cada especie, con el programa BWA (Li & Durbin, 2009) y se verificó el porcentaje de alineamiento con el programa Samtools (Li *et al.*, 2009). Los archivos *bam* de cada pool se utilizaron para el descubrimiento de SNPs con los programas Samtools, Bcftools, Vcftools (Li *et al.*, 2009; Danecek *et al.*, 2011). Para ostión y jurel, se realizaron análisis *in silico* para estimar el número de SNPs remanentes en caso de utilizar adaptadores con bases selectivas.

Validación de la técnica 2bRAD.

La validación de la técnica 2bRAD y el panel de SNPs obtenido para los análisis de caracterización genética y parentesco se llevó a cabo con el análisis de 96 muestras de jurel *S. rivoliana*, ostión *C. gigas*, huachinango *Lutjanus peru*, pargo lunarejo *L. guttatus* (n = 8) y totoaba *Totoaba macdonaldi* (mat.compl5). Se utilizaron tres familias conocidas de ostión de cultivo (n = 24, 6 reproductores y 18 progenies) y dos lotes de desoves de jurel de dos tanques de reproducción (n = 45: 9 reproductores y 13 progenies de un tanque, y 8 reproductores y 15 progenies del segundo tanque). Además, se incluyeron muestras silvestres de otras especies de importancia acuícola como el huachinango *Lutjanus peru* (n = 8) de Sinaloa (Reguera-Rouzaud *et al.*, 2021), *L. guttatus* (n = 8) de Panamá (Rivera-Lucero, 2020) y totoaba *Totoaba macdonaldi* (n = 8) del Golfo de California (Valenzuela-Quíñonez *et al.*, 2016).

La construcción de genotecas y secuenciación se llevó a cabo mediante el procedimiento del material suplementario (mat.compl4), utilizando los dos adaptadores con dos bases selectivas cada uno. En el material suplementario (mat.compl6) se muestran las secuencias de los adaptadores e iniciadores utilizados en la construcción de las genotecas, así como la ubicación de los códigos de barra en la placa de 96 pozos para la identificación de cada uno de los ejemplares. La secuenciación se llevó a cabo en un equipo MiSeq (Illumina) con una celda MiSeq Reagent Kit v3 (150-cycle). Los archivos de secuencias se analizaron de manera similar a lo reportado en la sección anterior desde el análisis de calidad y *demultiplexado* hasta el alineamiento y descubrimiento de SNPs. En el caso de las especies de *Lutjanus* y *Totoaba*, dado que no existe en el GenBank genomas completos de las mismas, el análisis de las secuencias se realizó con

el programa Stacks (Catchen *et al.*, 2011), el cual permite la creación de un pseudo-genoma de referencia mediante las propias secuencias de cada una de las especies. Para la creación de un catálogo de SNPs se utilizó la función *denovo_map.pl* con parámetros de profundidad mínima para formar un apilamiento $m = 3$, número máximo de diferencias entre apilamientos $M = 3$ y número máximo de diferencias entre loci putativos $n = 3$.

Los archivos tipo *vcf* con genotipos de los SNPs de cada especie, se filtraron con el programa *vcftools* (Danecek *et al.*, 2011) con los siguientes parámetros: eliminación de *indels*, retención de SNPs con dos alelos, SNPs representados en al menos el 80% de los individuos (*max-miss* = 0.8) y una calidad mínima de lectura de $Q = 30$. Para ostión, el parámetro *max-miss* fue de 0.9 y se retuvo sólo un SNP por fragmento. Los loci con dos alelos en al menos uno de los lotes o difiriendo con el genoma de referencia se consideró como polimórfico, por lo cual no fue necesario estimar la frecuencia mínima del alelo menor (MAF). Los archivos de SNPs filtrados se convirtieron a formato GenePop utilizando el programa PGDSpider2 (Lischer & Excoffier, 2012) y éstos se importaron desde GenAlEx (Peakall & Smouse, 2012) para los análisis de equilibrio de Hardy-Weinberg y variabilidad genética (No. de alelos por locus, heterocigosidad) por especie y por lote (i.e. progenitores y progenie en ostión y jurel). Los loci desviados del equilibrio de Hardy-Weinberg ($p < 0.05$) fueron excluidos. En ostión y jurel se realizaron análisis de parentesco con los programas Colony (Jones & Wang, 2010) y Cervus (Kalinowski *et al.*, 2007).

Sobre la base de implementar una estrategia de análisis de genotipificación estandarizado para las especies en cuestión y los SNPs obtenidos, se estimó el número de ejemplares (N_{ej}) que pueden ser analizados en una corrida de secuenciación en una plataforma MiSeq (Illumina), mediante la siguiente ecuación:

$$N_{ej} = N_{lps} * (1 - Mer) / Prof / Nsr \quad (\text{eq. 1})$$

Donde:

Nlps: Número de lecturas por celda de secuenciación. *Mer*: Porcentaje de merma en lecturas derivadas del control PhiX y lecturas de mala calidad; *Prof*: Profundidad de lectura deseada; *Nsr*: Número aproximado de sitios de reconocimiento de la enzima 2B en el genoma de la especie en cuestión reducida debido al uso de bases selectivas.

Resultados y Discusión

Análisis bioinformáticos *in silico* para la selección de la enzima tipo 2B.

El número promedio de sitios de reconocimiento por enzima entre especies correspondió en orden ascendente a *CspCI*, *Bcgl*, *AflI* y *BsaXI* con 53,328, 116,354, 118,873 y 278,549, respectivamente (Tabla 1). Se observa amplia heterogeneidad entre el número de sitios de

reconocimiento por millón de bases para cada enzima y cada especie, lo que sugiere que no hay una distribución aleatoria de bases nucleotídicas en estas especies. Por ejemplo, el número de sitios de reconocimiento por millón de bases con *Bcgl* es 30 % mayor en camarón que en ostión, y 27 % que en jurel, sin embargo, con las enzimas *BsaXI*, *AflI* y *CspCI* sucede lo contrario. En un análisis de 2bRAD de Hernandez-Castro *et al.* (2017) con el genoma de 702 Mb de *Rhodnius prolixus* (insecto vector de la enfermedad de Chagas), se obtuvieron 147, 292 y 100 sitios por millón de bases para cada una de las enzimas *Bcgl*, *AflI* y *CspCI*, respectivamente. Lo anterior indica que no hay un número de sitios de reconocimiento asociado a diferentes niveles taxonómicos.

Considerando que el tamaño del genoma entre jurel y ostión es similar y que la enzima *Bcgl* da como resultado un número similar de sitios de reconocimiento en estas dos especies (entre 65 K y 70 K, respectivamente, Tabla 1), *Bcgl* fue la enzima seleccionada para los análisis subsecuentes. Aun cuando la enzima *CspCI* también sería una enzima adecuada debido a una variabilidad más restringida, la variación aleatoria en el número de bases de corte a los costados de la secuencia (New England Biolabs, com. pers.), resulta en una incertidumbre no deseada durante la digestión.

Tabla 1. Número de sitios de reconocimiento de las enzimas de restricción 2B *Bcgl*, *BsaXI*, *AflI* y *CspCI* de los análisis *in silico* en los genomas de acceso público de camarón, ostión y jurel.

Especie y tamaño del genoma (Gb)		Enzima			
		<i>Bcgl</i>	<i>BsaXI</i>	<i>AflI</i> ¹	<i>CspCI</i>
Camarón blanco <i>Penaeus (Litopenaeus) vannamei</i> (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
Ostión del Pacífico <i>Crassostrea gigas</i> (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
Jurel <i>Seriola rivoliana</i> (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
Promedio		116,354	278,549	118,873	53,328

Fwd: resultado en dirección sentido; Rev: resultado en dirección anti-sentido; Total: sumatoria de ambas; Prop: proporción de sitios de reconocimiento por millón de bases nucleotídicas; Gb: miles de millones de pares de bases.

¹ La secuencia de reconocimiento en sentido y anti-sentido es la misma por lo que solo se dan los valores de Fwd y Total.

Fuente: Elaboración propia.

La reducción promedio en el número de sitios de reconocimiento en función del uso de adaptadores con una a cuatro bases selectivas es bastante homogéneo independientemente de la especie y la enzima utilizada en la digestión, con un promedio de 22.9 %, 5.4 %, 1.3 % y 0.3 %, respectivamente. En el caso de *BsaXI*, la reducción es 0.1 % con cinco sitios selectivos. Cabe señalar que cuando se analiza el número de sitios remanente en función de las bases selectivas, se observa un mayor número de sitios cuando las bases son A o T (datos no mostrados). Esto implica que dependiendo de la aplicación para la cual se pretenda realizar la caracterización genética, se puede estimar el número de bases selectivas más adecuada. Es decir, si el estudio implica un análisis de asociación de marcadores a caracteres de interés (GWAS por sus siglas en inglés), el número de bases selectivas podría ser 0, 1 o 2, pero si el estudio se refiere a análisis de parentesco con la posibilidad de analizar a una gran cantidad de individuos de manera simultánea, el uso de 3 o 4 bases selectivas sería la recomendada.

El análisis con la base de datos de camarón diseñada por Perez-Enriquez *et al.* (2018) para determinar la proporción de sitios variables (SNPs) dentro de los fragmentos con sitios de reconocimiento de la enzima *Bcgl*, indicó que, de una muestra aleatoria de 50 fragmentos, el 65 % de ellos tuvo al menos un SNP. De esta manera, el análisis *in silico* indica que utilizando la enzima *Bcgl* en camarón sin bases selectivas se obtendrían 214,518 sitios de reconocimiento y, por lo tanto, alrededor de 139,437 SNPs. Este es un valor que pudiera estar sobreestimado si se compara con los 94,113 SNPs (después del filtrado) reportados por Lyu *et al.* (2021). Si a la digestión se le agregan los adaptadores con dos o cuatro bases selectivas, el número de SNPs esperados sería de 7,530 y 418, respectivamente. Bajo estos parámetros, la utilización de *Bcgl* y adaptadores sin bases selectivas con el ostión y el jurel daría como resultado entre 42,000 y 45,000 SNPs para cada una.

Estandarización de la técnica 2bRAD.

La construcción de las genotecas con los pooles de camarón, ostión y jurel digeridas con *Bcgl* dio como resultado un número muy variable de lecturas de secuenciación (Tabla 2), probablemente debido al proceso de preparación y no a la naturaleza de los genomas de las especies. El porcentaje promedio de alineamiento de las secuencias contra los genomas de referencia para camarón y jurel (68 % y 95 %, respectivamente), es acorde a lo esperado de acuerdo con los tamaños de los genomas de dichas especies que están registrados en el GenBank (Seetharam, 2018; Zhang *et al.*, 2019). Para el ostión para el cual el genoma de referencia representa el genoma de la especie en su totalidad, no hubo diferencias en el porcentaje de alineamiento entre pooles (Peñaloza *et al.*, 2021).

En camarón, el número de SNPs filtrados obtenidos con el uso de cuatro bases selectivas en dos adaptadores es de 2,874 (Tabla 2). Para ostión y jurel, se obtuvieron cantidades similares de SNPs antes del filtrado (aprox. 35,000) que se redujeron en 25 % y 60 %, respectivamente, después del filtrado (Tabla 2). La menor reducción de SNPs en el ostión en comparación con el jurel se considera que es debido a la alta heterocigosidad que se reporta en el ostión (Peñaloza *et al.*, 2021). El análisis *in silico* para estas dos especies indica que, utilizando dos adaptadores con dos bases selectivas cada uno, se obtendrían entre 300 y 400 SNPs. En función del

objetivo de generar un panel de SNPs de baja densidad (cientos de SNPs) se determinó que la técnica 2bRAD con cuatro bases selectivas es adecuada para ostión y jurel. Sin embargo, para el camarón esta técnica no es viable por la elevada cantidad de marcadores (2,874); una estrategia alternativa para el desarrollo de un panel de baja densidad (100 a 200 SNPs) para esta especie es la técnica denominada GT-Seq (Genotipificación por miles mediante secuenciación; Campbell *et al.*, 2015) que utiliza información del genoma y de SNPs especie-específico previamente identificados.

Validación de la plataforma 2bRAD.

La secuenciación de las cinco especies (jurel, ostión, huachinango, pargo lunarejo y totoaba) dio como resultado un promedio de 165,428 lecturas por individuo y con un elevado porcentaje de alineamiento a los genomas de referencia (jurel y ostión) (Tabla 3). En jurel se obtuvieron 3,320 y 160 SNPs después del alineamiento y después del filtrado, respectivamente; en ostión, los SNPs obtenidos fueron 3,731 y 114, respectivamente (Tabla 3). El análisis comparativo de réplicas (jurel, huachinango y ostión) para las dos primeras especies resultó en un porcentaje bajo de error de genotipificación (0.6 % y 1.03 %, respectivamente). Para ostión el porcentaje de error de genotipificación, después de la depuración de SNPs fue de 2.1 %, que, si bien es relativamente bajo, pudiera tener implicaciones en la precisión de los análisis de asignación de parentesco.

El análisis de desviación del equilibrio de Hardy-Weinberg indicó que varios loci fueron significativos ($p < 0.05$), los cuales fueron excluidos para los análisis subsecuentes (jurel: $n = 1$; ostión $n = 3$; totoaba $n = 7$). En huachinango y en pargo lunarejo, los loci con potencial desequilibrio no se excluyeron debido al reducido tamaño de muestra. Sin embargo, no debe descartarse la posibilidad de la presencia de loci sujetos a presión selectiva, tal como se reportó en el bagre *Rhamdia quelen* (Ríos *et al.*, 2020), para lo cual se requerirán análisis posteriores con tamaños de muestra adecuados en cada una de las especies en cuestión.

Tabla 2. Número de lecturas de secuenciación (L), porcentaje de alineamiento (% P) de las secuencias filtradas al genoma de referencia para cada pool y número de SNPs obtenidos antes y después del filtrado (SNPs-F).

Especie y tipo de genoteca	Número identificación de Pool	L	% P	SNPs	SNPs-F ¹
Camarón con dos bases selectivas	Lv-sel2-pool-1	1,110,928	64.4	26,535	19,105
	Lv-sel2-pool-2	4,532,602	71.9		
Camarón con cuatro bases selectivas	Lv-sel4-pool-1	1,893,650	63.4	4,377	2,874
	Lv-sel4-pool-2	1,341,554	72.0		
Ostión sin bases selectivas	Cg-pool-1	4,274,049	76.8	34,387	13,664 (403 ²)
	Cg-pool-2	4,618,046	76.5		
Jurel sin bases selectivas	Sr-pool-1	1,125,540	93.9	35,142	26,251 (309 ²)
	Sr-pool-2	1,339,927	96.1		

¹ Criterios de filtrado: Eliminación de indels, SNPs con dos alelos, Q = 30

² SNPs remanentes inferidos por análisis *in silico* si se utilizaran cuatro bases selectivas

Fuente: Elaboración propia

Tabla 3. Número de lecturas de secuenciación promedio por individuo (NL), porcentaje de alineamiento al genoma de referencia (% AI) y SNPs para cada especie, totales y retenidos después del filtrado.

Especie	NL	% AI	SNPs totales	SNPs retenidos
Jurel <i>Seriola rivoliana</i>	208,062	94.5	3,320	160
Ostión del Pacífico <i>Crassostrea gigas</i>	150,990	82.5	3,731	114
Huachinango <i>Lutjanus peru</i>	86,917	NA	953	434
Pargo lunarejo <i>Lutjanus guttatus</i>	180,775	NA	1,631	543
Totoaba <i>Totoaba macdonaldi</i>	230,397	NA	335	151

NA: No aplica debido a que el alineamiento se realizó respecto a un genoma *de novo* generado con Stacks.

Fuente: Elaboración propia

En el jurel *S. rivoliana*, la variabilidad genética indica un valor alto del número de alelos por locus (N_a) en el grupo de reproductores, sin embargo, tanto los valores del número efectivo (N_e) de alelos como de la heterocigosidad (H_o , H_e) muestran valores de medios a bajos (Tabla 4). Sin embargo, el valor negativo de F indica que no hay endogamia en este grupo. No existen datos de variabilidad genética en poblaciones de cultivo de *S. rivoliana*; sin embargo, en la especie cercana *S. lalandi*, la variabilidad genética de lotes de cultivo, estimada con microsatélites, indicó valores intermedios de H_e (0.652 – 0.810) (Martínez Matus, 2016).

Cuando se compara la variabilidad de reproductores y progenie de *S. rivoliana* se observa una reducción en N_a , H_o y H_e de aproximadamente un 7 %, con una mínima variación en N_e y F (Tabla 4). Esta reducción de variabilidad genética de una generación a otra es común en lotes de cultivo, debido a que no todos los reproductores participan en los desoves (e.g. Perez-Enriquez *et al.*, 1999). El análisis de parentesco basado en 159 SNPs, permitió inferir con alta probabilidad la estructura de familias de hermanos completos y medios hermanos en los tanques de cultivo de jurel muestreados (probabilidad de no exclusión de la pareja de padres fue de 9×10^{-9}). Adicionalmente, se identificó que, de las 10 hembras, tres estuvieron representadas en las progenes, y de nueve machos (uno de ellos no analizado genéticamente), seis participaron en los desoves (Tabla 5). Esto permitió determinar que la estructura familiar de la siguiente generación fue conformada por una mayor cantidad de machos que de hembras, como se presenta en otras especies [e.g. *Seriola lalandi* (Martínez Matus, 2016), *Lutjanus guttatus* (Perez-Enriquez *et al.*, 2020)]. Las secuencias del panel de 159 SNPs para *S. rivoliana* se presentan en el material suplementario (mat.compl7).

Tabla 4. Valores promedio de variabilidad genética de juveniles y reproductores de cultivo de jurel y ostión, y de silvestres de huachinango, pargo lunarejo y totoaba.

Especie	Grupo	N	N_a	n_e	H_o	uH_e	F
<i>Seriola rivoliana</i>	Progenitores	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
	Variación de progenitores a juveniles		-7%	0%	-6%	-7%	NA
Ostión del Pacífico <i>Crassostrea gigas</i>	Progenitores	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
	Variación de progenitores a juveniles		5.1%	-0.4%	-4.6%	-7.5%	NA
Huachinango <i>Lutjanus peru</i>	Lperu-1	7.7	1.99	1.36	0.230	0.253	0.007

Continuación

Tabla 4. Valores promedio de variabilidad genética de juveniles y reproductores de cultivo de jurel y ostión, y de silvestres de huachinango, pargo lunarejo y totoaba.

Especie	Grupo	N	Na	n_e	Ho	uHe	F
Pargo lunarejo	Lgut-1	4.0	1.78	1.38	0.249	0.274	-0.061
<i>L. guttatus</i>	Lgut-2	3.5	1.59	1.32	0.218	0.233	-0.109
Totoaba	Tot-1	6.7	1.99	1.46	0.327	0.308	-0.123
<i>Totoaba macdonaldi</i>							

N: promedio de individuos genotipificados; Na: número de alelos por locus; n_e : número efectivo de alelos por locus; Ho: heterocigosidad observada; uHe: heterocigosidad esperada ajustada al tamaño de muestra; F: índice de endogamia. NA: No aplica

Fuente: Elaboración propia

En el ostión, de acuerdo con los criterios de Perez-Enriquez *et al.* (2024), la variabilidad genética mostró valores medios a altos en el número de alelos por locus pero medios a bajos (y menores en juveniles que en reproductores) en He (Tabla 4). Sin embargo, se incrementó el Na de una generación a otra, debido a alelos que están en la progenie, pero ausentes los reproductores; esto pudiera ser por la baja profundidad de secuenciación en algunos reproductores o a mutaciones particulares, lo cual ha sido reportado en *C. gigas* como un proceso necesario para contrarrestar los efectos de alelos deletéreos (Durland *et al.*, 2021). El panel de SNPs presentó una probabilidad de no exclusión de la pareja de padres de 1×10^{-15} . El análisis de parentesco de ostión con el programa Colony con 114 SNPs, mostró concordancia de agrupamiento genético a nivel de hermanos completos en la mayor parte de los ejemplares de acuerdo a lo esperado (Tabla 5). Las secuencias del panel de 114 SNPs para *C. gigas* se presentan en el material suplementario (mat.comp18).

Tabla 5. Asignación de parentesco de grupos familiares de jurel (Sr) y ostión (Cg) obtenidos con los programas Colony y Cervus.

Especie	ID de Progenie	Pareja esperada de progenitores	Asignación con Colony			
			GF	P	ID de Macho	ID de Hembra
Jurel <i>Seriola rivoliana</i>	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, Sr11-J4, Sr12-J4	ND	2	1	#1	Sr9-TR1
Ostión del Pacífico <i>Crassostrea gigas</i>	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: Grupo familiar o de desove; P: Probabilidad; ND: No determinada. Los símbolos #1 y #2 indican progenitores potenciales no encontrados en los candidatos. GA: grupo de asignación: (A): coincidencia de parentesco esperado y obtenido con ambos métodos; (B): coincidencia de parentesco esperado y obtenido, pero un progenitor no fue significativo; (C): asignación distinta a lo esperado con al menos un método. El superíndice en las asignaciones con Cervus indican significancia (*; $\alpha < 0.95$) o no significativo (ns).

Fuente: Elaboración propia

Continuación

Tabla 5. Asignación de parentesco de grupos familiares de jurel (Sr) y ostión (Cg) obtenidos con los programas Colony y Cervus.

Especie	ID de Progenie	Pareja esperada de progenitores	Asignación con Cervus		GA
			ID de Macho	ID de Hembra	
Jurel <i>Seriola rivoliana</i>	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-TR1 ^{ns}	Sr9-TR1*	B
	Sr3-J4, Sr5-J4, Sr6-J4, Sr8-J4, Sr10-J4, Sr11-J4, Sr12-J4	ND	Sr5-TR1*	Sr9-TR1*	B
	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
Ostión del Pacífico <i>Crassostrea gigas</i>	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: Grupo familiar o de desove; P: Probabilidad; ND: No determinada. Los símbolos #1 y #2 indican progenitores potenciales no encontrados en los candidatos. GA: grupo de asignación: (A): coincidencia de parentesco esperado y obtenido con ambos métodos; (B): coincidencia de parentesco esperado y obtenido, pero un progenitor no fue significativo; (C): asignación distinta a lo esperado con al menos un método. El superíndice en las asignaciones con Cervus indican significancia (*; $\alpha < 0.95$) o no significativo (ns).

Fuente: Elaboración propia

Los valores de variabilidad para las otras tres especies (huachinango, pargo lunarejo y totoaba) presentaron valores relativamente elevados de n_e , aunque valores intermedios de H_e (Tabla 4), los cuales se pueden deber, al reducido tamaño de muestra y a la falta de un genoma de referencia. La disponibilidad de genomas de referencia para estas especies será de relevancia para una mejor descripción de la diversidad. En este sentido, si bien Chambers *et al.* (2023) reconocen que la falta de genomas de referencia en especies no modelo y el pequeño tamaño de los fragmentos de restricción del 2bRAD son problemáticos para la genotipificación, la reconstrucción filogenética y la proporción de genotipos faltantes de esta técnica, genera resultados comparables y con mayor repetitividad que con otras técnicas como ddRAD. De hecho, la caracterización genética de especies acuáticas no modelo mediante la técnica de 2bRAD, con un adaptador con una base selectiva, para la reducción del tamaño del genoma, y con ello la optimización de costos de secuenciación, ya había sido reportada por Barbanti *et al.* (2020). En dicho estudio, los autores indican que el número de muestras para secuenciación es un balance entre número de loci, número de lecturas por individuo derivado de la plataforma de secuenciación (en ese caso HiSeq, Illumina) y el presupuesto disponible, sugiriendo una profundidad de lectura de $20\times$.

En el presente trabajo, se demostró la factibilidad de reducir el tamaño del genoma en varias especies acuícolas mediante dos adaptadores con dos bases selectivas cada uno para la obtención de un número suficiente de loci en una plataforma MiSeq (Illumina). Con ello, una estimación del número de ejemplares que pueden ser analizados en una corrida de secuenciación (ver eq. 1) es de aproximadamente $Ne_j = 270$, considerando una celda de 25 millones de lecturas (N_{lps}), 20% de merma (Mer), una profundidad de secuenciación de $20\times$ ($Prof$) y un número aproximado de sitios de reconocimiento de la enzima 2b en el genoma de la especie en cuestión reducida debido al uso de bases selectivas de 3700 (N_{sr}) (Tabla 3). Se requerirá el costeo de un análisis de estas características para determinar su aplicabilidad por empresas acuícolas; sin embargo, éste tenderá a tener amplias variaciones derivadas de las condiciones de trabajo del laboratorio de secuenciación y de los costos de adquisición de los diversos insumos.

Conclusiones

Se implementó la técnica 2bRAD para la obtención de paneles de marcadores genéticos de baja densidad (100 a 500 SNPs), cuyo protocolo es aplicable de manera indistinta para la caracterización genética de diversas especies acuícolas del noroeste de México (osti6n, jurel, huachinango, pargo lunarejo y totoaba). Se demostr6 que los paneles para jurel y osti6n permiten la asignaci6n de parentesco y paternidad de lotes de cultivo. Estos paneles ser6n de utilidad para el manejo de lotes de reproductores de cultivos experimentales y comerciales de las especies del presente estudio y para otras especies de inter6s acu6cola, lo que contribuir6 al desarrollo de sus cultivos. Asimismo, la estrategia es aplicable para la caracterizaci6n de recursos pesqueros de inter6s comercial.

Contribución de los autores

Conceptualización del trabajo, RPE, AMA.; desarrollo de la metodología, RPE, AMA, GMC; manejo de software, RPE, CEF.; validación experimental, RPE, AMA, GMC, CEF; análisis de resultados, RPE, CEF; Manejo de datos, RPE, AMA, GMC, CEF; escritura y preparación del manuscrito, RPE, AMA; redacción, revisión y edición, RPE, AMA, GMC, CEF; administrador de proyectos, RPE; adquisición de fondos, RPE.

Todos los autores de este manuscrito han leído y aceptado la versión publicada del mismo.

Financiamiento

Esta investigación fue financiada por Consejo Sudcaliforniano de Ciencia y Tecnología (COSCYT), número de convenio CAR/14/2021.

Agradecimientos

Se agradece la donación de muestras a Kampachi Farms de México (jurel), Acuícola Robles (ostiön), Fausto Valenzuela (totoaba) y el apoyo técnico de Susana Ávila Álvarez.

Conflicto de interés

Los autores declaran no tener conflicto de interés.

Referencias

- 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 ddRAD for phylogenomics. *Ecology and Evolution*, 13(3), e9842. <https://doi.org/10.1002/ece3.9842>
- CONAPESCA. (2021). Anuario Estadístico de Acuicultura y Pesca, Edición 2021. Comisión Nacional de Acuicultura y Pesca, México. <https://www.gob.mx/conapesca/documentos/anuario-estadistico-de-acuicultura-y-pesca>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., De Pisto, 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., 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. *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>
- 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., Tremblay, 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-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., Palaikostas, 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>

Material Suplementario

Mat.compl1. Sitios de reconocimiento y corte de las enzimas 2B y ejemplo de ligación de adaptadores.

Mat.compl2. Ejemplo de ligación de adaptadores (secuencias en azul) con diferentes números de bases selectivas (bases nucleotídicas marcadas en rojo) de fragmentos digeridos con la enzima *Bcgl* (secuencias en amarillo y naranja).

Mat.compl3. Comandos utilizados para búsqueda de sitios de reconocimiento de enzimas en los análisis *in silico*.

Mat.compl4. Protocolo 2bRAD desde la preparación de genotecas hasta la secuenciación.

Mat.compl5. Ubicación de las muestras de cada uno de los ejemplares en una placa de 96 pocillos. Muestras ubicadas en las posiciones F6, G6 y H6 se utilizaron como réplicas.

Mat.compl6. Secuencias de los adaptadores e iniciadores utilizados en la construcción de las genotecas, así como la ubicación de los códigos de barra en la placa de 96 pozos para la identificación de cada uno de los ejemplares.

Mat.compl7. Listado de los 159 SNPs identificados en jurel, con su ubicación respecto al genoma de referencia *Seriola rivoliana* (No. acceso GenBank: GCA_002994505.1 ASM299450v1; Seetharam, 2018). CHR: Cromosoma o grupo de ligamiento; POS: posición; Bases: variantes para cada SNP.

Mat.compl8. Listado de los 114 SNPs identificados en ostión con su ubicación respecto al genoma de referencia *Crassostrea gigas* (No. acceso GenBank: cgigas_uk_roslin_v1 NC_047559-NC_047568; Peñaloza *et al.* 2021). CHR: Cromosoma o grupo de ligamiento; POS: posición; Bases: variantes para cada 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/12>