



**FACULTADE DE CIENCIAS**

**GRAO EN BIOQUÍMICA**

**D. Jacobo Fariñas Varela**

**BASE XENÉTICA DA RESISTENCIA A  
PERKINSOSIS EN AMEIXA XAPONESA  
(*RUDITAPES PHILIPINARUM*)**

Traballo Fin de Grao  
*06/2024*

# Índice

0. Índice de acrónimos .....	3
1. Resumen/Resumo/Abstract .....	4
2. Introducción.....	7
2.1 Almeja japónica ( <i>Ruditapes philipinarum</i> ).....	8
2.2 <i>Perkinsus olseni</i> .....	11
2.3 Cría y selección .....	14
2.4 Tolerancia/Resistencia a <i>P. olseni</i> .....	15
2.5 Polimorfismos de nucleótido sencillo (SNPs) .....	16
3. Objetivos del estudio .....	18
4. Materiales y métodos .....	19
4.1 Obtención, cría e infección de individuos para el estudio.....	19
4.2 Extracción de ADN.....	21
4.3 Realización de qPCR.....	22
4.4 Genotipado de SNPs con el chip de Axiom .....	23
4.5 Asignación de familias .....	25
4.6 GWAS.....	26
4.7 Minería genómica.....	27
5. Resultados y discusión .....	29
5.1 Extracción de muestras de ADN .....	29
5.2 qPCR .....	30
5.3 Genotipado.....	31
5.4 Asignación familiar (análisis de paternidad) .....	32
5.5 GWAS.....	36
5.6 Minería genómica.....	40
6. Conclusiones .....	45
7. Anexos .....	47
8. Referencias bibliográficas.....	52

## 0. Índice de acrónimos

---

ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico complementario
ARN	Ácido ribonucleico
ARNm	Ácido ribonucleico mensajero
bp	Del inglés “ <i>base pair</i> ”
CIMA	Centro de Investigaciones Marinas de Vilanova de Arousa
dNTP	Del inglés “ <i>Deoxynucleotide Triphosphates</i> ”
DOS	Del inglés “ <i>Disk Operating System</i> ”
ECIMAT	Estación Científica para el Medio Marino de Toralla
FAP	Del inglés “ <i>Family Analysis Program</i> ”
GCTA	Del inglés “ <i>Genome-wide Complex Trait Analysis</i> ”
GO	Del inglés “ <i>Gene Ontology</i> ”
GRM	Del inglés “ <i>Genetic Relationship Matrix</i> ”
GWAS	Del inglés “ <i>Genome Wide Association Study</i> ”
InDel	Del inglés “ <i>Insertion and Deletion</i> ”
ITS	Del inglés “ <i>Internal Transcribed Spacers</i> ”
MAF	Del inglés “ <i>Minimum Allele Frequency</i> ”
OMSA	Organización Mundial de Sanidad Animal
OsHV	Del inglés “ <i>Ostreid Herpes variant</i> ”
PCR	Del inglés “ <i>Polymerase Chain Reaction</i> ”
POMS	Del inglés “ <i>Pacific Oyster Mortality Syndrom</i> ”
qPCR	Del inglés “ <i>Quantitative Polymerase Chain Reaction</i> ”
QTL	Del inglés “ <i>Quantitative Trait Locus</i> ”
RFTM	Del inglés “ <i>Ray’s fluid thioglycollate medium</i> ”
SNP	Del inglés “ <i>Single Nucleotide Polymorphism</i> ”

# 1. Resumen/Resumo/Abstract

---

## **Español: BASE GENÉTICA DE LA RESISTENCIA A PERKINSOSIS EN ALMEJA JAPONESA (RUDITAPES PHILIPPINARUM)**

Este estudio intenta comprender los componentes genómicos asociados a la resistencia a *Perkinsus olseni* en poblaciones de almeja japónica (*Ruditapes philippinarum*). Para ello, se estableció una población genéticamente heterogénea compuesta por 1000 individuos y se sometió a un desafío controlado con el parásito protozoario. La respuesta de las almejas fue evaluada, posteriormente, mediante la estimación precisa de la carga parasitaria de cada individuo mediante qPCR donde se observó una correlación entre el tamaño de la almeja y el nivel de infección. Simultáneamente, los individuos retados fueron genotipados utilizando un chip de SNPs Axiom multiespecie desarrollado por el grupo ACUIGEN, que incluye 8019 SNPs de almeja japónica. Esta información fue utilizada para realizar la asignación familiar de los individuos utilizando un programa informático de asignación paternal denominado FAP, concretamente 31 familias fueron determinadas a partir de cruces entre 1 macho x 1 hembra. Posteriormente, a través de un estudio de asociación de genoma completo, se identificaron marcadores genéticos asociados a la resistencia a *P. olseni*, y se investigó, mediante minería genética, las regiones genómicas próximas asociadas con la resistencia/tolerancia al parásito. Concretamente, se identificaron tres genes posiblemente relacionados con fenotipos asociados a la resistencia a este parásito intracelular: RPHI1A040444, RPHI1A016912 y RPHI1A020365; los dos primeros tienen asociados términos GO relacionados con la integridad de la membrana plasmática y transporte transmembrana y el último, términos Go vinculados a procesos proteolíticos. Como resultado de la identificación de estos QTLs, genes y variantes alélicas potencialmente responsables de esta resistencia, se pretende ofrecer un punto de partida para el desarrollo de estrategias de cría selectiva en almeja japónica.

**PALABRAS CLAVES:** *Perkinsus olseni*; *Ruditapes philippinarum*; resistencia; tolerancia; acuicultura; SNP; GWAS; minería genómica

## **Galego: BASE XENÉTICA DA RESISTENCIA A PERKINSOSIS EN AMEIXA XAPONESA (RUDITAPES PHILIPPINARUM)**

Este estudo intenta comprender os compoñentes xenómicos asociados á resistencia a *Perkinsus olseni* en poboacións de ameixa xapónica (*Ruditapes philippinarum*). Para iso, estableceuse unha poboación xeneticamente heteroxénea composta por 1000 individuos e someteuse a un desafío controlado co parásito protozoario. A resposta das ameixas foi avaliada, posteriormente, mediante a estimación precisa da carga parasitaria de cada individuo mediante qPCR e se observou unha correlación entre o tamaño da ameixa e o nivel de infección. Simultaneamente, os individuos retados foron xenotipados utilizando un chip de SNPs Axiom multiespecie desenvolvido polo grupo ACUIGEN, que inclúe 8019 SNPs de ameixa xapónica. Esta información foi utilizada para realizar a asignación familiar dos individuos utilizando un programa informático de asignación paternal denominado FAP, concretamente 31 familias foron determinadas a partir de cruces entre 1 macho x 1 femia. Posteriormente, a través dun estudo de asociación de xenoma completo, identificáronse marcadores xenéticos asociados á resistencia a *P. olseni*, e investigouse, mediante minería xenética, as rexións xenómicas próximas asociadas coa resistencia/tolerancia ao parásito. Concretamente, identificáronse tres xenes posiblemente relacionados con fenotipos asociados á resistencia a este parásito intracelular: RPHI1A040444, RPHI1A016912 e RPHI1A020365; os dous primeiros teñen asociados termos GO relacionados coa integridade da membrana plasmática e transporte transmembrana e o último, termos Go vinculados a procesos proteolíticos. Como resultado da identificación destes QTLs, xenes e variantes alélicas potencialmente responsables desta resistencia, preténdese ofrecer un punto de partida para o desenvolvemento de estratexias de cría selectiva en ameixa xapónica.

**PALABRAS CLAVES:** *Perkinsus olseni*; *Ruditapes philippinarum*; resistencia; tolerancia; acuicultura; SNP; GWAS; minería xenómica

## **English: GENETIC BASIS OF RESISTANCE TO PERKINSOSIS IN JAPANESE CLAM (RUDITAPES PHILIPPINARUM)**

This study aims to understand the genomic components associated with the resistance to *Perkinsus olseni* in japanese clam (*Ruditapes philippinarum*) populations. For this purpose, a genetically heterogeneous population of 1000 individuals was established and subjected to a controlled challenge with the protozoan parasite. The response of the clams was subsequently assessed by accurately estimating the parasite load of each individual by qPCR where a correlation between clam size and level of infection was observed. Simultaneously, challenged individuals were genotyped using a multispecies Axiom SNP-chip developed by the ACUIGEN group, which includes 8019 japanese clam SNPs. This information was used to perform the family assignment of individuals using a paternal assignment software called FAP, specifically 31 families were determined from crosses between 1 male x 1 female. Subsequently, through a genome-wide association study, genetic markers associated with resistance to *P. olseni* were identified and nearby genomic regions associated with resistance/tolerance to the parasite were investigated by genetic mining. Specifically, three genes possibly related to phenotypes associated with resistance to this intracellular parasite were identified: RPHI1A040444, RPHI1A016912 and RPHI1A020365, the first two of which have associated GO with plasma membrane integrity and transmembrane transport and the last one, Go terms linked to proteolytic processes. As a result of the identification of these QTLs, genes and allelic variants potentially responsible for this resistance, it is intended to provide a starting point for the development of selective breeding strategies in japanese clams.

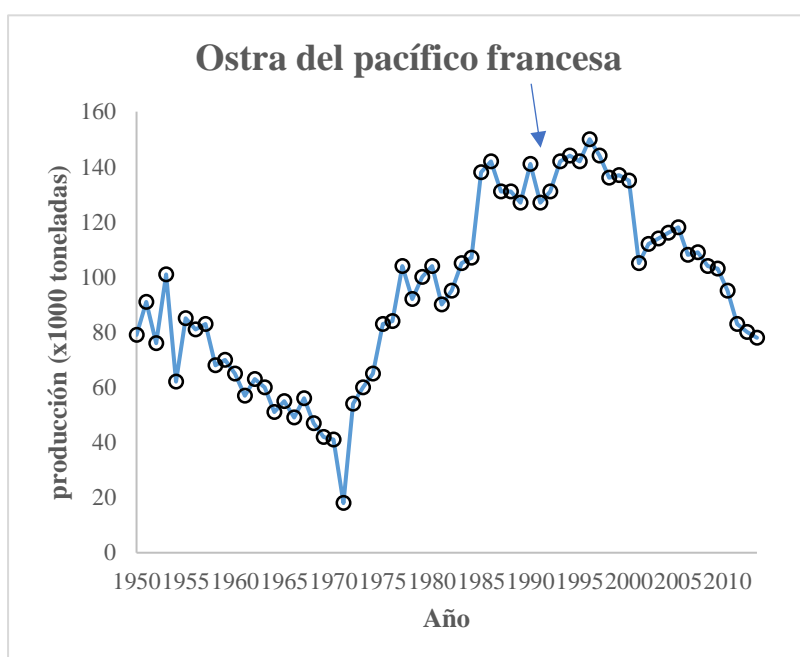
**KEYWORDS:** *Perkinsus olseni*; *Ruditapes philippinarum*; resistance; tolerance; aquaculture; SNP; GWAS; genomic mining

## 2. Introducción

Actualmente, la producción acuícola mundial representa una industria de 157.000 millones de dólares anuales, de los cuales, 20.500 millones proceden exclusivamente de la producción de moluscos para la alimentación (FAO, 2015).

La gran mayoría de las especies de bivalvos importantes comercialmente están expuestas a múltiples enfermedades víricas, bacterianas y parasitarias en su ambiente natural que pueden provocar su muerte o afectar a su crecimiento y, por lo tanto, determinar grandes pérdidas económicas a las empresas y mariscadores encargados de su explotación.

Un ejemplo representativo lo tenemos en el año 2008, cuando se observaron en las costas francesas episodios de mortalidades masivas de entre el 40 y el 100% de ostras juveniles del Pacífico (*Crasostrea gigas*) (Segarra *et al.*, 2010). Esta incidencia se asoció a una nueva variante de herpesvirus ostrícola tipo 1 (OsHV-1  $\mu$ var) que, junto con algunas infecciones causadas por bacterias oportunistas, dan lugar al síndrome de mortalidad de la ostra del pacífico (POMS: *Pacific Oyster Mortality Syndrom*), causante en gran medida de la dramática reducción en la producción de este molusco en las dos últimas décadas, tal y como se muestra en la Figura 1 (Martenot *et al.*, 2011).



[Figura 1] Producción de ostra del pacífico francesa (*C.gigas*) a lo largo de los últimos años. La flecha indica los primeros informes sobre muertes causadas por OsHV-1.

Valores adaptados de Guo & Ford (2016).

En los últimos 40 años, la distribución a nivel mundial de estas enfermedades y su mayor prevalencia e intensidad en bivalvos son una de las principales amenazas para la producción acuícola mundial. Esta situación está propiciada principalmente por dos motivos: los nuevos sistemas de cría de alta densidad, para el incremento de la producción y su valor económico, y el desarrollo de un mercado internacional que favorece la transferencia y exposición entre individuos de distintas regiones geográficas (Naylor *et al.*, 2021).

Ambos factores aumentan el riesgo de introducir y transmitir microorganismos dañinos en nuevas áreas, lo que puede dar lugar a consecuencias terribles al provocar una fuerte presión evolutiva sobre las especies y poblaciones autóctonas. También hay indicios que parecen indicar que los cambios en las corrientes y condiciones ambientales oceánicas ocasionadas por el cambio climático provocan la aparición de enfermedades que afectan de forma más frecuente y agresiva a los organismos marinos (Harvell *et al.*, 1999; Burge *et al.*, 2014).

En general, estas enfermedades bacterianas y víricas están asociadas con altas mortalidades en las primeras etapas de vida durante la etapa larvaria. Por el contrario, los parásitos protozoarios, entre los que se encuentra *P. olseni*, afectan a todas las etapas de la vida de los bivalvos, pero su impacto (principalmente pérdida de tejido y muerte) se manifiesta especialmente en los adultos, justo antes de la comercialización del producto (Proestou *et al.*, 2023).

## 2.1 Almeja japónica (*Ruditapes philipinarum*)

*R. philipinarum*, también conocida como almeja japónica, es una especie de almeja (perteneciente a la familia Veneridae) ampliamente distribuida en zonas intermareales de las costas del océano Pacífico occidental, sobre fondos de arena gruesa y fango (Secretaría General de Pesca & Ministerio de Agricultura, Pesca y Alimentación, 2023a). Tiene una gran importancia ecológica en las zonas costeras filtrando grandes cantidades de agua y fitoplancton, además de proporcionar alimento a una gran cantidad de organismos marinos y amortiguar el efecto del incremento de CO<sub>2</sub> incorporándolo como carbonato a sus conchas (Carnegie *et al.*, 2016).

Estructuralmente, presenta una concha ovalada con ambas valvas lisas, simétricas y con el lado anterior más corto y estrecho que el posterior. La superficie externa es relativamente brillante, de un color tostado o gris oscuro, y presenta costillas radiales juntas y aserradas más marcadas en los extremos. Estas valvas le permiten aislarse del exterior al producir un cierre hermético que mantiene, a su vez, el agua necesaria en el interior para su supervivencia (Secretaría General de Pesca & Ministerio de Agricultura, Pesca y Alimentación, 2023a). En particular esta especie de almeja tiene la capacidad de resistir importantes variaciones de salinidad y temperatura en el medio que habita, además de presentar las respuestas innatas altamente eficaces típicas de los moluscos (Guo *et al.*, 2015).

La almeja japónica es una especie sedimentívora que se alimenta mediante filtración de microorganismos y plancton movidos por las corrientes marinas usando dos sifones que se proyectan hacia el agua circundante. Su genoma tiene un tamaño aproximado de 1,6 Gb, distribuido en 19 cromosomas (Martínez P., datos no publicados). A nivel reproductivo, es una especie dioica y con fecundación externa. El desarrollo de los gametos se produce en ambos sexos entre diciembre y febrero y su maduración coincide con el final de la primavera para, a mediados de verano, ser expulsados al exterior en forma de óvulos y espermatozoides maduros (Chung *et al.*, 2013). Tras la fecundación producida en el medio y la eclosión del huevo nace el nuevo individuo en forma larvaria que, durante las primeras semanas vive suspendido en el agua y se desplaza a favor de las corrientes marinas. Al cabo de 15 o 20 días, se fija mediante un órgano denominado biso al sustrato y da comienzo la metamorfosis. Durante este proceso, las nuevas almejas pierden el órgano natatorio y desarrollan las branquias, a su vez, las valvas larvarias comienzan un proceso de transformación que terminará dando las futuras conchas adultas (Vogeler *et al.*, 2020).

A nivel económico, es uno de los principales motores comerciales de las zonas costeras del sudeste asiático, de donde es originaria, gracias a su rápido crecimiento y elevada tasa de fecundidad. Estas características la convierten en una de las principales especies de almeja comercializadas a nivel mundial, llegando a ser uno de los bivalvos de mayor importancia en el sector marisquero, con una producción de alrededor de tres millones de toneladas anuales tan solo en China (Zheng *et al.*, 2023).

En particular, la introducción de esta especie de almeja fue especialmente efectiva en la costa oeste de Norteamérica, en la costa francesa, en el mar Adriático y en las costas de la península Ibérica, como se puede observar en la figura 2.



[Figura 2] Representación gráfica de la evolución del mercado de la almeja japónica (*R. philippinarum*) en el territorio español (Secretaría General de Pesca & Ministerio de Agricultura, Pesca y Alimentación, 2023b).

A su vez, la almeja japónica es uno de los muchos moluscos que son parasitados por *P. olseni*, esta interacción entre huésped y hospedador determina un proceso coevolutivo entre ambas especies, desarrollando los mecanismos de defensa de la almeja contra el parásito, que a su vez genera nuevos mecanismos de infección.

## 2.2 *Perkinsus olseni*

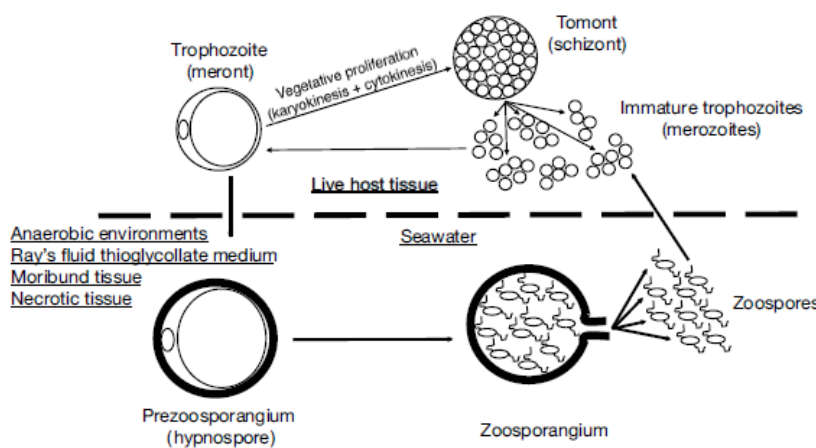
La mayoría de las enfermedades en moluscos que tienen consecuencias importantes a nivel socioeconómico están causadas por protozoos (Carnegie *et al.*, 2016). En particular, el conjunto de endoparásitos protozoarios del género *Perkinsus* es responsable de enfermedades que, colectivamente, se denominan perkinsosis. Esta infección protozoaria afecta a una gran cantidad de moluscos a nivel mundial causando mortalidades masivas en las poblaciones hospedadoras, derivando en grandes pérdidas económicas en los sectores de la pesca y la acuicultura.

Dentro de este grupo de protozoos destacan por su implicación en infecciones parasitarias en moluscos dos de las siete especies incluidas en el género *Perkinsus*: *P. marinus* y *P. olseni* (= *P. atlanticus*). Son organismos dinoflagelados pertenecientes al phylum Apicomplexa y están reconocidas en la lista de patógenos de declaración obligatoria de la Organización Mundial de Sanidad Animal (OMSA) (World Organisation for Animal Health, 2021), al provocar enfermedades transmisibles consideradas de gran importancia socioeconómica y/o sanitaria a nivel internacional.

En la actualidad, *P. olseni* es una de las principales causas de mortalidad en distintas especies de moluscos bivalvos, entre las que se encuentra *R. philippinarum*, siendo responsable de la caída de la producción de almeja en Europa y Asia. En concreto, este organismo provoca la retracción en el manto y un retraso en el crecimiento, además de estar asociado a mortalidades de hasta el 40% en cualquier fase del ciclo vital del hospedador (Fernández-Boo *et al.*, 2015). Entre las enfermedades de bivalvos, es posible que la perkinsosis haya provocado las pérdidas económicas más severas; es por ello que entre la comunidad científica internacional hay un gran interés en su estudio desde hace años (Villalba *et al.*, 2005). Desde el punto de vista médico, es interesante por ofrecer un sistema de expresión heterólogo para genes implicados en enfermedades humanas como por ejemplo es el caso de la infección por *Plasmodium falciparum*, responsable de la malaria (Yadavalli *et al.*, 2020).

En cuanto al proceso de infección, este parásito protozoario es intracelular facultativo y se caracteriza por su capacidad para infectar las células de la hemolinfa de distintas especies de moluscos, dando lugar a lesiones que pueden interferir en la reproducción y otros procesos fisiológicos como el crecimiento, la respiración y/o la supervivencia (Choi & Park, 1997). En los hemocitos del hospedador comienza un proceso de replicación que tiene la capacidad de eludir las respuestas inmunitarias específicas y, consecuentemente, perpetuar el proceso infeccioso dentro del organismo. Los procesos de infección y evolución de esta parasitosis están a su vez muy relacionados con la temperatura; diversos estudios apuntan a una relación directa entre la mortalidad y los ciclos estacionales, siendo esta máxima en períodos cálidos y secos (Chu & La Peyre, 1993).

La forma de infección de este parásito es directa ya que prescinde de hospedadores intermediarios y, por otra parte, no parece depender de la densidad del patógeno (Moullac *et al.*, 2008). Una vez dentro del propio individuo infectado, podemos diferenciar distintas localizaciones de la infección según la fase de desarrollo en la que se encuentre este protozoo.



[Figura 3] Esquema representativo del ciclo vital de *P. olseni*. El ciclo de vida de este parásito es variable según el escenario en el que se encuentre, pudiéndose diferenciar hasta 5 tipos asociados a diferentes estadios de desarrollo. En el agua del mar se encuentra en forma de zoosporangios y zoosporas; en ambientes anaeróbicos en forma de hiposporas y, finalmente, en el tejido vivo del hospedador en forma de trofozoítos y esquizontes (Hanrio *et al.*, 2021).

Tras la absorción del parásito por filtración, la célula infectiva inicial (denominada trofozoíto) comienza un proceso de división mediante múltiples cariocinesis y citocinesis sucesivas, tal como se observa en la figura 3. Estas divisiones dan lugar a una agregación de hasta 32 células filiales envueltas en una única membrana generando una “célula progenitora” de gran tamaño que recibe el nombre de tomont o esquizonte (Hanrio *et al.*, 2021). Esta última, debido al crecimiento de las diversas células filiales localizadas en su interior, rompe su envoltura para dar paso a la liberación de los trofozoítos inmaduros (o merozoítos) en el tejido del hospedador.

Posteriormente, estos trofozoítos inmaduros experimentan un incremento en su tamaño por la formación de vacuolas, transformándose de esta forma en trofozoítos maduros. Este ciclo se repite varias veces durante la permanencia del parásito en el hospedador hasta su muerte, desembocando en la liberación del mismo al ambiente marino y en la infección de nuevos individuos por los trofozoítos maduros filiales. No obstante, es relevante resaltar que ciertos trofozoítos (formas más interesantes a la hora de realizar los estudios) son expulsados al medio durante el transcurso de vida del hospedador mediante un proceso conocido como diapédesis, manifestándose en fluidos corporales como la sangre o en heces (Bushek *et al.*, 2002).

Al exponerse a las condiciones externas de pocos nutrientes y niveles bajos de pH y oxígeno, los trofozoítos maduros comienzan a agrandarse y desarrollar una pared celular gruesa y duradera que les permite sobrevivir en condiciones desfavorables, este estadio vital se denomina hipnospora o prezoosporangio (Bushek *et al.*, 2002).

Al localizarse el prezoosporangio libre en el medio puede transformarse a zoosporangio que, a su vez, presenta una proliferación interna de hasta 100 zoosporas contenidas en su interior. Estas zoosporas tienen una estructura elipsoidal biflagelada, que les permite movilidad para su liberación a través de un túbulo de descarga al ambiente externo (Perkins & Menzel, 1966).

En la actualidad, aún no se ha estudiado suficiente el impacto que pueda tener el consumo de moluscos infectados en humanos. No obstante, se ha observado que en ciertos ratones humanizados con HLA-DR4, la ingestión de moluscos infectados por *P. marinus* (especie del mismo género que *P. olseni*) induce una respuesta inmune sistémica pese a no manifestar ninguna patología significativa (Wijayalath *et al.*, 2014).

## 2.3 Cría y selección

La industria acuícola afronta importantes desafíos en la actualidad debido a enfermedades como la perkinsosis que afectan a las especies de bivalvos. En respuesta a esta situación, las empresas han implementado prácticas de gestión y protocolos de bioseguridad adecuados para controlar el desarrollo y expansión de estas enfermedades. Sin embargo, estos enfoques de contención no son lo suficientemente eficaces, especialmente en el medio natural donde se crían la mayoría de los moluscos, lo cual indica la necesidad de estrategias complementarias para favorecer la resistencia de las poblaciones de bivalvos (Proestou *et al.*, 2023).

Una estrategia interesante a la hora de afrontar este desafío es la cría selectiva; esta aprovecha la variación genética natural propia de las especies para seleccionar individuos y/o poblaciones con una determinada característica de interés. En este caso, la cría selectiva se refiere a la selección de individuos con rasgos de resistencia específicos, lo que permite mejorar el control de ciertas enfermedades a la vez que mejora la rentabilidad general de industria acuícola.

Relacionado con este aspecto, varios ensayos de campo y estudios de heredabilidad han demostrado que las poblaciones de bivalvos responden favorablemente a la presión ejercida por las enfermedades que los infectan a través de la cría selectiva. Por ejemplo, a través de programas de seguimiento a largo plazo se observó un aumento en la resistencia al virus del herpes ostrícola (OsHV) en la ostra del Pacífico como consecuencia del proceso de selección (Proestou *et al.*, 2023). Otro caso relevante, es la rápida adquisición de resistencia a la marteiliosis (*Marteilia cochillia*) en el berberecho común (*Cerastoderma edule*) en la Ría de Arousa en Galicia (Pampín *et al.*, 2023; Vera *et al.*, 2023).

Además, estos estudios relevaron la alta heredabilidad para la resistencia a enfermedades en individuos infectados en laboratorio. Los valores de heredabilidad en sentido estricto estimados oscilan entre 0,4 y 0,99, lo cual respalda la eficacia de la cría selectiva como una estrategia de mejora poblacional a nivel de resistencia a patógenos en bivalvos (Azéma *et al.*, 2017).

En definitiva, durante el proceso de cría selectiva se explota la variación genética natural de las especies de bivalvos para dar lugar a poblaciones enteras con un mismo rasgo de resistencia. De esta forma se pueden aumentar las medidas de control de enfermedades ya conocidas a la vez que se mejora la rentabilidad económica de esta industria.

## 2.4 Tolerancia/Resistencia a *P. olsenii*

La respuesta del hospedador frente a cualquier enfermedad puede ser de dos tipos: resistencia y tolerancia, los cuales no son necesariamente excluyentes.

La resistencia se define como la capacidad de un individuo hospedador de minimizar los efectos de la infección por el impedimento activo de su entrada; no obstante, una elevada resistencia o tolerancia no necesariamente conlleva una alta tasa de supervivencia, especialmente si el costo asociado a la eliminación del patógeno es muy alto.

Por otro lado, mientras la resistencia se define como el impedimento activo por parte del hospedador a la entrada de los patógenos en su organismo, la tolerancia tiene que ver con la capacidad de sobrevivir al patógeno o no ver afectada su tasa de crecimiento (Råberg *et al.*, 2008). Los individuos que presentan estos fenotipos tolerantes muestran la capacidad de mantener una aptitud funcional a nivel de reproducción y supervivencia (Råberg *et al.*, 2007). Esto es consecuencia, principalmente, del mantenimiento de la infección por debajo del umbral establecido como infectivo, incluso estando bajo la exposición del patógeno. La forma de estudiar la tolerancia en el laboratorio consiste en el análisis del desempeño del hospedador expuesto a diferentes niveles de patógenos (Proestou *et al.*, 2023).

Análisis recientes en poblaciones de ostras sugieren que los fenotipos tolerantes podrían tener relación con la reducción de la densidad del patógeno en el medio, así como en la disminución de la mortalidad asociada al mismo (Ben-Horin *et al.*, 2018). Por este motivo, algunos expertos afirman que el promover la tolerancia en poblaciones expuestas al patógeno en vez de promover la resistencia, puede ser más efectivo a la hora de reducir la cantidad de parásitos como *P. olsenii* (Proestou *et al.*, 2023).

Dado que la resistencia y tolerancia anteriormente citadas presentan distintas estrategias a la hora de enfrentar al patógeno, deben tenerse en cuenta de forma independiente en el momento de realizar la selección genética.

## 2.5 Polimorfismos de nucleótido sencillo (SNPs)

Los SNP (*Single Nucleotide Polymorphism*) representan la forma más simple de variación de ADN (o ácido desoxirribonucleico) en las poblaciones, concretamente son lugares del genoma en el que los cromosomas homólogos difieren en la secuencia de ADN en un único nucleótido dentro de una población y, como consecuencia, los individuos podrán presentar distintos genotipos. Se caracterizan por la presencia de más de un alelo en el mismo locus, pudiendo llegar a establecerse distintos criterios de polimorfismo a la hora de identificarlos, como por ejemplo considerar el alelo con menor frecuencia (MAF: *Minimum Allele Frequency*)  $\geq 0,01$  (Nava-Cedeño *et al.*, 2022). Son resultado de múltiples eventos de mutación a lo largo del tiempo y del genoma, con tasas de mutación específicas en cada especie y que pueden llegar a cambiar la pauta de lectura en las regiones codificadoras, como es el caso de las inserciones y deleciones (o *InDel*) (García-Díaz & Kunkel, 2006). A su vez, las mutaciones que no conllevan un cambio en la pauta de lectura pueden ser de dos tipos: transiciones, donde se sustituyen bases nitrogenadas del mismo tipo (purina por purina o pirimidina por pirimidina) y transversiones, donde se sustituye un nucleótido por otro de distinto tipo (pirimidina por purina o viceversa). Consecuentemente, estos polimorfismos pueden ser responsables de la diversidad genética entre los individuos y la evolución del genoma.

En general podemos diferenciar sus consecuencias en: cambios en la codificación de los aminoácidos (SNPs no sinónimos); no producir cambios en la secuencia de aminoácidos (SNPs sinónimos o silenciosos); o encontrarse en regiones no codificantes. Estas últimas pueden llegar a afectar a la expresión génica en el caso de encontrarse en elementos reguladores como *enhancers* o promotores. Los más interesantes a la hora de abordar este estudio son los primeros ya que son los responsables de las diferencias individuales en la respuesta. Estos polimorfismos de nucleótido sencillo pueden, a su vez, influir en diversos procesos orgánicos como la propia expresión génica (si el SNP se encuentra en una región reguladora) o en la conformación y estabilidad del ARN mensajero (ARNm) (Shastry, 2009).

En consecuencia, la identificación y el análisis de estas variaciones génicas pueden mejorar la comprensión de su efecto sobre las características fenotípicas en el propio individuo. Sin embargo, en la actualidad, la mayoría de los SNPs localizados a lo largo del genoma de *R.philippinarum* no se han caracterizado, aunque se intuye una posible asociación entre estos polimorfismos de una única base con proteínas alteradas y otras diferencias fenotípicas. Su estudio a nivel resistencia de agentes infectivos pretende detectar la variación o el conjunto de variaciones nucleotídicas que favorezcan el fenotipo de resistencia o tolerancia en una población. Precisamente para el éxito de esta estrategia, se necesita el estudio de un gran número de individuos y controles, posibilitado gracias a nuevas tecnologías de alto rendimiento.

A grandes rasgos, marcos generales como la Organización Mundial de Sanidad Animal junto con métodos de diagnóstico precisos y una amplia base de conocimientos científicos sobre estas enfermedades son la base para controlar los patógenos perjudiciales para la producción de moluscos (Carnegie *et al.*, 2016).

### 3. Objetivos del estudio

---

El objetivo general del presente estudio está relacionado con averiguar la base genética de la resistencia a perkinsosis en *R. philippinarum* y sentar las bases para su aplicación en un programa de selección en criadero para obtener una semilla resistente o tolerante a esta parasitosis. Para alcanzar este propósito se plantean una serie de objetivos específicos que guiarán el desarrollo del estudio:

- Fundar una población genéticamente heterogénea compuesta por una importante cantidad de familias de hermanos.
- Someter dicha población a un reto con el parásito *P. olsenii* en criadero.
- Evaluar mediante qPCR (*Quantitative Polymerase Chain Reaction*) la carga parasitaria de cada individuo.
- Genotipar dicha población con un chip de SNPs desarrollado en el laboratorio del grupo ACUIGEN que incluye marcadores para *R. philippinarum*.
- Realizar un análisis genético de asignación familiar para determinar las familias de los individuos analizados.
- Identificar y caracterizar los marcadores genéticos asociados a la resistencia a este protozoo mediante un estudio de asociación de genoma completo (GWAS: *Genome Wide Association Study*), además de comprender su papel en la respuesta inmune.
- Realizar minería genética en las regiones genómicas asociadas con la resistencia/tolerancia al parásito para tratar de identificar genes candidatos.
- Proponer soluciones para obtener una línea resistente a *P. olsenii* por medio de la cría selectiva.

Con estos resultados se pretende profundizar en la comprensión de los mecanismos genéticos implicados en la resistencia a *P. olsenii* en *R. philippinarum* (almeja japónica), lo cual podría tener importantes implicaciones tanto en el sector de la acuicultura como en la propia conservación de esta especie en los ecosistemas naturales.

## 4. Materiales y métodos

---

Las primeras etapas de este estudio se realizaron en el año 2022, y se corresponden con la cría y obtención de familias de almejas japónicas utilizadas para este estudio.

### 4.1 Obtención, cría e infección de individuos para el estudio

En noviembre de 2022 se enviaron al ECIMAT (Estación Científica para el Medio Marino de Toralla, Vigo) una población de reproductores de almeja japónica del banco de Placeres, naturales de la ría de Pontevedra. En este centro de investigación se fundaron hasta 40 familias mediante cruzamientos de 1 macho con 1 hembra y se realizó la cría y mantenimiento de las progenies hasta que alcanzaron la talla mínima para el desarrollo del reto con *P. olseni*. En ese momento se trasladaron al Centro de Investigaciones Marinas de Vilanova de Arousa (CIMA) donde las semillas se mantuvieron en circuito abierto una semana entera para, posteriormente, traspasarse a circuito cerrado con exposición al parásito.

El reto se realizó también en el CIMA, a partir de células de *P. olseni* obtenidas de una muestra de 115 individuos de *Ruditapes decussatus* (almeja fina) procedentes también de la ría de Pontevedra. Para su realización, se seleccionaron e introdujeron estos individuos en tanques de aclimatación en circuito cerrado y se sometieron a luz ultravioleta para eliminar todas las posibles bacterias y organismos que pudiesen interferir en el aislamiento del parásito. Posteriormente, se realizó el diagnóstico de las 115 almejas finas, concretamente fragmentos branquiales de estos individuos para analizar la carga parasitaria. Para llevar a cabo este proceso, se comenzó abriendo y cultivando las almejas en *Ray's fluid thioglycollate medium* (RFTM), que produce la liberación del parásito de las branquias al resto de tejidos. Tras la incubación de los tejidos, se tiñeron con Lugol y se analizaron individualmente, según la escala de Mackin, para determinar el nivel de infección de entre 0 (ausencia de parásito) a 5 (carga máxima de parásito) utilizando microscopía óptica.

Tras este análisis se seleccionaron los 12 individuos con un nivel de infección por *P. olseni* máximo (grado de infección 5). Se recogieron muestras de branquia, vianda y sifón de cada individuo para realizar el aislamiento del parásito. El proceso consistió en la degradación de los tejidos con tripsina y a continuación su filtrado a través de 4 filtros: 300, 150, 100 y 20. Lo que queda sin filtrar se tiñe con rojo neutro y, posteriormente, se cuantifica empleando una cámara de Neubauer para estimar el número de prezoosporangios y el porcentaje de viabilidad (Liste, 2002).

En este punto, se prepararon tres tanques junto con los recipientes flotantes, estos contenían las almejas para el reto en unas condiciones determinadas: el agua se estabilizó a una temperatura constante de 20 °C gracias a un calentador situado en la zona de los tanques, a la vez que un sifón permitía la circulación del agua por un tubo situado en la parte baja. Después, se procedió a la introducción de individuos de almeja japónica (*R. philippinarum*) en cada tanque, manteniéndose en circuito cerrado (la entrada de agua se limitaba únicamente para alimentar por jarreo). Concretamente, estos individuos son los que se utilizaron para el estudio.

Se añadieron alrededor de 35 millones de zoosporangios de *P. olseni* en dos de los tres tanques utilizados en este experimento, ya que el tercero se utilizó como tanque control (individuos no expuestos al parásito del que se pretende estudiar resistencia). Cada uno de los tanques, control e infectados, incluyó un total de aproximadamente 1500 almejas. Durante ese periodo se mantuvieron una semana en circuito cerrado para evitar la salida de parásitos del sistema, al tiempo que se iban liberando las zoosporas biflageladas que tienen como finalidad infectar a las almejas. A los tanques se les añadió de forma constante fitoplancton para estimular la filtración y evitar los efectos indeseados de falta de alimento.

Posteriormente, las almejas japónicas expuestas al reto con el parásito se traspasaron a otros tanques de circuito abierto. Después de tres semanas de exposición al parásito, se tomaron las muestras de los tanques control e infectado; se extrajo la vianda de cada una de ellas y se introdujo en tubos Eppendorf con la cantidad apropiada de etanol para su fijación y posterior extracción de ADN en el laboratorio de ACUIGEN de la Facultad de Veterinaria, en Lugo.

Se midieron el alto y ancho de las almejas analizadas con un calibre, concretamente 1000 individuos; las de más de 3mm de ancho se abrieron directamente empleando un cuchillo y se introdujeron en un tubo con etanol. Los individuos de menor tamaño se midieron mediante un programa específico integrado en la lupa binocular, debido a la mayor dificultad para su medición con el calibre. Previamente al traspaso de tejido a los tubos de etanol, se realizó un proceso de descalcificación en las almejas de menor tamaño ya que no se podía separar la vianda de la concha directamente.

Los tejidos blandos preservados en etanol se utilizaron para la extracción de ADN; esta, a su vez, fue utilizada posteriormente para evaluar la carga parasitaria mediante qPCR, la realización del genotipado con un chip de SNPs y los análisis genómicos estadísticos y bioinformáticos oportunos.

## 4.2 Extracción de ADN

La extracción del ADN para los análisis posteriores se realizó a partir del tejido branquial de diferentes individuos machos y hembras (correspondientes con las almejas estudiadas); por otro lado, se extrajo también el ADN de la generación parental para el posterior análisis de paternidad e identificación de las familias.

Todas las almejas analizadas se clasificaron por tamaño para la extracción de ADN: almejas “grandes”, extraídas en placa utilizando el kit comercial *E-Z 96™ Plant DNA DS Kit* de la casa comercial Omega Bio-tek (la clase más abundante); almejas “medianas” y “pequeñas”, que se extrajeron en tubo con el kit *E.Z.N.A.® Plant DNA DS Kit* y con el kit *E.Z.N.A.® MicroElute® Genomic DNA Kit*, respectivamente (ambos kits de la casa comercial Omega Bio-tek). Algunas de las almejas más pequeñas necesitaron un proceso de descalcificación previo y, por ese motivo, precisaron una purificación previa a la extracción, pues los componentes de la concha que quedan disueltos en el etanol pueden interferir en las reacciones enzimáticas realizadas posteriormente sobre el ADN.

Desde que se tuvo el tejido ya pesado e identificado, se comenzó a realizar las extracciones de ADN con el kit correspondiente. El kit de extracción de ADN de plantas es el más parecido al que previamente ofrecía la casa comercial para la extracción de ADN en moluscos, previsiblemente por la complejidad del genoma de los moluscos y la

gran cantidad de variantes estructurales que presenta. Prácticamente todos los reactivos utilizados durante este proceso están incluidos en los respectivos kits y, por otra parte, el material empleado pertenece al propio laboratorio del grupo ACUIGEN:

- Microcentrífuga: Thermo Scientific™ Microcentrífuga Sorvall™ Legend™ Micro 17
- Espectrofotómetro de microvolumen: Thermo Scientific NanoDrop™ 2000
- Reactivos comunes en los distintos kits: Homogenizer Mini Columns, 2 mL Collection Tubes, CSPL Buffer, Proteinase K Solution (enzima encargada de romper enlaces peptídicos y que presenta una amplia especificidad), RBB Buffer, XP2 Binding Buffer, HBC Buffer (previamente fue diluido con isopropanol 100%), DNA Wash Buffer, Elution Buffer, CSPL Buffer y RNase A (endonucleasa que escinde específicamente el ARN monocatenario de los residuos de pirimidina) de esta forma se obtiene ADN libre de ARN y se eliminan las regiones no hibridadas de los híbridos de ARN-ADN.
  - E-Z 96™ Plant DNA DS Kit: E-Z 96® DNA plate. Este sistema elimina la necesidad de cloroformo tradicionalmente asociado con métodos de lisis basados en CTAB (Bromuro de hexadeciltrimetilamonio), un detergente catiónico.
  - E.Z.N.A.® Plant DNA DS Kit: HiBind® DNA Mini Columns (encargada de la asociación con el ADN). Este kit es idóneo para la recuperación eficiente de ADN genómico de hasta 30 kb de tamaño.
  - E.Z.N.A.® MicroElute® Genomic DNA Kit: HiBind MicroElute® Column. Este kit permite el aislamiento rápido y sencillo del ADN genómico a partir de pequeños volúmenes y tamaños de muestras.

### 4.3 Realización de qPCR

Esta técnica cuantitativa es una variante de la reacción en cadena de la polimerasa (PCR: *Polymerase Chain Reaction*) utilizada para amplificar y cuantificar de forma simultánea y absoluta un amplicón específico utilizado como molde el ADN de la muestra. Se pueden utilizar la propia molécula de ADN que se pretende cuantificar o hebras de ADN complementario (ADNc), obtenido a partir de la retrotranscripción del ARN.

Para la cuantificación del ADN del parásito se emplearon un par de cebadores específicos denominados *Perk-ITS-q F1* y *Perk-ITS-q R2*, desoxinucleótidos (dNTPs: *Deoxynucleotide Triphosphates*), un tampón de reacción adecuado, cloruro de magnesio ( $MgCl_2$ ) y una ADN polimerasa estable a altas temperaturas. La secuencia amplificada se corresponde con las regiones espaciadoras intergénicas transcritas 1 y 2 (*Internal Transcribed Spacers: ITS-1* e *ITS-2*) del ARN ribosómico, siguiendo los criterios descritos por Ríos *et al.* (2020). Por otro lado, la cuantificación realizada por medio de esta técnica puede ser de dos tipos: absoluta y relativa; esta última es la más común y utiliza la expresión de otro gen de referencia con expresión constante para la normalización de las medidas. Estos genes suelen estar involucrados en funciones básicas del organismo y presentan una expresión constitutiva relativamente estable (*housekeeping*), de esta forma la cuantificación relativa del gen estudiado con respecto al gen de referencia nos permite conocer el nivel de expresión y compararla con otras muestras, en el caso de emplear el mismo gen de referencia.

Durante varios meses en las instalaciones del CIMA se cultivó el parásito en laboratorio hasta alcanzar una muestra de 100 millones de células. A partir de ella, en el laboratorio de ACUIGEN se hicieron diluciones para crear una recta de calibrado que nos sirvió para cuantificar la cantidad de parásito encontrada en cada individuo de forma cuantitativa a partir del número de ciclos que necesitaba la PCR para detectar el ADN del parásito. Además, verificamos la precisión de la recta utilizando los individuos establecidos como control negativo (no infectados) en el experimento.

#### 4.4 Genotipado de SNPs con el chip de Axiom

Este tipo de estudios se centran en la diversidad genética y pretenden obtener las variaciones presentes entre individuos y poblaciones y su asociación con caracteres de interés. Utilizando diversos métodos se puede realizar un análisis del genoma completo o centrarse en el genotipado específico de unos determinados SNPs de interés.

En este caso particular, el genotipado se realizó mediante un chip denominado *Axiom Affymetrix multispecies UI33 Plus 2.0 SNP array*, fabricado por Thermo Fisher SCIENTIFIC y representado en la figura 4. La información genómica para el diseño de este chip fue suministrada por el grupo ACUIGEN, perteneciente al departamento de

Zoología, Genética y Antropología Física de la Universidad de Santiago de Compostela. Se identificaron los SNPs más polimórficos distribuidos homogéneamente a lo largo del genoma a partir de la secuenciación del genoma completo de 10 individuos machos y 10 hembras de cada una de las especies integradas en el chip. Se incluyeron, además, SNPs relacionados con características de interés tales como: el sexo, el origen geográfico o la resistencia a determinadas enfermedades. En concreto, se incluyeron 8019 SNPs para almeja japónica que, a su vez, fueron clasificados durante este proceso en distintas categorías según la robustez del genotipado y sus características: SNPs polimórficos de alta resolución (*PolyHighResolution*), SNPs de individuos genotipados por debajo del umbral de significación (*CallRateBelowThreshold*), SNPs con ausencia del homocigoto correspondiente al alelo de menor frecuencia (*NoMinorHom*), SNPs monomórficos de alta resolución (*MonoHighResolution*), variantes nucleotídicas fuera de rango (*Off Target Variants*) y otros (*Other*).



[Figura 4] Imagen del chip *Axiom Affymetrix multispecies SNP array*. En este chip multiespecie de 57 mil SNPs se incluyen marcadores para las cinco especies más importantes de moluscos en la acuicultura gallega: *Ostrea edulis*, *Ruditapes decussatus*, *Cerastoderma edule*, *Mytilus galloprovincialis* y *Ruditapes philippinarum*.

Los chips o *arrays* de SNPs incluyen entre miles y millones de sondas de ADN específicas para cada variante particular de un SNP. Estas sondas son pequeñas secuencias de oligonucleótidos de 25 pares de bases (bp: *base pair*) de longitud que se unen de forma específica a secuencias complementarias de ADN (ADNc) generadas a partir de ARN de la muestra (Lipshutz *et al.*, 1999). Cuando el ADN de la muestra hibrida con las sondas complementarias en el chip de forma específica se lleva a cabo un lavado que asegura que solo uniones específicas permanezcan ligadas a este. Por último, se utiliza un escáner para medir la fluorescencia en cada punto del chip. La fluorescencia indica la presencia de un fragmento específico de ADN unido a una sonda, y su intensidad se correlaciona con la cantidad de ADN presente (Irizarry, 2003). Posteriormente, el

patrón de fluorescencia del chip se compara con una referencia para identificar las variantes genéticas.

Este método permite la identificación precisa y rápida de miles de polimorfismos a la vez, haciendo posible un estudio genético representativo de todo el genoma en un elevado número de individuos, ideal para este tipo de estudios.

#### 4.5 Asignación de familias

Los cruces entre parejas de padres y madres tuvieron lugar en el ECIMAT dando lugar a todos los individuos analizados en este estudio. Posteriormente, se realizó un análisis de parentesco para determinar familias de padre e hijos a partir de los datos obtenidos del genotipado de los individuos. En este tipo de análisis basados en datos moleculares es crucial la precisión a la hora de realizar predicciones, especialmente el poder de resolución que se pueda llegar a conseguir de un conjunto de loci determinados.

El software utilizado en este caso para inferir los individuos progenitores de las muestras fue FAP (Family Analysis Program). Este programa informático basado en DOS (*Disk Operating System* o "Sistema Operativo en Disco") realiza la asignación familiar de mayor probabilidad a partir de los genotipos parentales y filiales, utilizando la herencia mendeliana como parámetro para calcular la probabilidad de todas las relaciones candidatas en un subgrupo de individuos (Taggart, 2006).

FAP utiliza un enfoque basado en la exclusión y genera tres tipos de información: la predicción de un panel de distintos loci para asignar descendientes a una determinada pareja de progenitores; la asignación de descendientes a las parejas de progenitores; y la identificación de loci que podrían ser problemáticos por errores de genotipado. Este programa es particularmente interesante cuando se conocen todos los progenitores en la progeñe estudiada.

## 4.6 GWAS

Es un tipo de estudio en el que se comparan marcadores de ADN a lo largo del genoma utilizando individuos que presentan diferencias fenotípicas para un determinado rasgo; de este modo se pretende hallar marcadores presuntamente asociados a este carácter. Este tipo de pruebas se centran usualmente en asociaciones entre SNPs y un determinado rasgo de interés.

En este estudio concretamente, se utilizaron 8.019 SPNs y el GWAS se realizó con el software GCTA (*Genome-wide Complex Trait Analysis*), programa que realiza el análisis genético de marcador por marcador siguiendo un modelo lineal mixto. La correlación para la estructura de este modelo se basa en una matriz de relaciones genómicas (GRM: *Genome Relationship Matrix*), que a su vez se describe mediante la ecuación (Eq.1);

$$y_i = \mu + T_i + a_j x_{ij} + u_i + \varepsilon_i \quad (\text{Eq. 1})$$

donde  $y_i$  es el fenotipo observado del individuo,  $\mu$  es la media fenotípica de la población (al ser el tamaño una característica cuantitativa),  $T_i$  se corresponde a efectos fijos (tanques en nuestro estudio),  $a_j$  es el efecto aditivo del alelo de referencia del candidato (SNP $_j$ ),  $x_{ij}$  es el recuento de los alelos de referencia (0,1 o 2) para el SNP $_j$  para el individuo ( $i$ ),  $u_i$  es el valor aditivo poligénico aleatorio del individuo ( $i$ ) y finalmente,  $\varepsilon_i$  es el efecto residual para el individuo ( $i$ ) (Fraslin *et al.*, 2020).

El vector de efectos residuales se distribuye normal e independientemente  $\varepsilon \sim N(0, I\sigma_\varepsilon^2)$  siendo  $\sigma^2$  la varianza residual. Por otro lado, el vector aleatorio de efectos poligénicos sigue una distribución normal  $u \sim N(0, G\sigma_g^2)$ , en este caso  $\sigma^2$  es la varianza genética estimada y G una matriz de relaciones genómicas construída a partir de la información de los SNPs (Yang *et al.*, 2011).

Posteriormente, se realizó un segundo análisis GWAS considerando únicamente el subconjunto de SNPs más robusto en cuanto a genotipado, representando aproximadamente la mitad de SNPs (3904 concretamente). Esta segunda selección de SNPs permite un análisis estadístico más fiable, sin embargo, la potencia del análisis se

ve reducida al disminuir la densidad de SNPs en el genoma; precisamente por este motivo es necesario tener en cuenta la información de ambas aproximaciones. Además, se realizaron GWAS posteriores aplicando una corrección por tamaño para corregir posibles correlaciones que pudieran interferir en las asociaciones detectadas, ya que se encontró un sesgo en los valores de infección relacionado con el tamaño.

Conceptualmente, este tipo de análisis relaciona los SNPs y su posición en los cromosomas con el valor estadístico de la asociación (*p-values*) generado. Estos, a su vez, se definen como la probabilidad de la desviación de la hipótesis nula (no asociación), por lo que valores bajos se asocian con significatividad. En general, este tipo de análisis de asociación genética utilizan un *p-value* límite de  $\alpha = 0,05$  corregido por múltiples test, teniendo en cuenta la totalidad del genoma (*Genome-wide*) o cada cromosoma particular (*Chromosome-wide*) (Qu *et al.*, 2010).

#### 4.7 Minería genómica

Tras la identificación de los SNPs con mayor asociación a la resistencia, obtenidos por medio del GWAS (menor *p-value*), se procedió a la inspección del genoma en esa región (minería genética) para tratar de identificar genes candidatos relacionados con la resistencia al parásito. Concretamente, este estudio se centró en aquellos SNPs que sobrepasaron en los Manhattan plot generados en el GWAS los umbrales de *genome-wide* y *chromosome-wide* (Calderón-Chagoya *et al.*, 2023). Además, se estudian otros SNPs con menor significatividad pero asociados a cromosomas con una importante representación en los Manhattan plot.

Después de haber seleccionado estos SNPs se identificaron con su código correspondiente y se comenzó el estudio del genoma en esas regiones. Durante el desarrollo de este estudio se produjo una actualización del genoma de *R. philippinarum*, derivando en el reposicionamiento de los SNPs seleccionados en el nuevo genoma. Este proceso se realizó mediante el alineamiento del SNP específico junto con una secuencia contigua de 100 bp frente al genoma actualizado de esta especie de almeja. Se utilizó un software denominado “Bowtie2”, programa que combina diagramas de causa y consecuencia y los fusiona con barreras u obstáculos a nivel individual en un solo diagrama para alinear la secuencia correspondiente (De Ruijter & Guldenmund, 2016).

Finalmente, se procedió a la búsqueda de genes. Se diseñaron ventanas genómicas de  $\pm 200$  kb (*upstream* y *downstream*) alrededor del polimorfismo identificado, correspondientes a un total de 400 kb. Acto seguido, tras la identificación de los genes localizados en esas regiones, se determinaron las proteínas para las cuales codifican y sus respectivos términos GO (*Gene Ontology*).

Estos últimos son identificadores alfanuméricos universales que asocian una proteína con una determinada función mediante la interacción entre múltiples bases de datos y la clasifican funcionalmente en un sistema semi-jerarquizado.

Muchas de estas proteínas codificantes participan en procesos celulares muy conservados entre organismos, lo que facilita la transferencia de anotaciones biológicas entre organismos diferentes, debido a la similitud entre secuencias presentada (Du *et al.*, 2010).

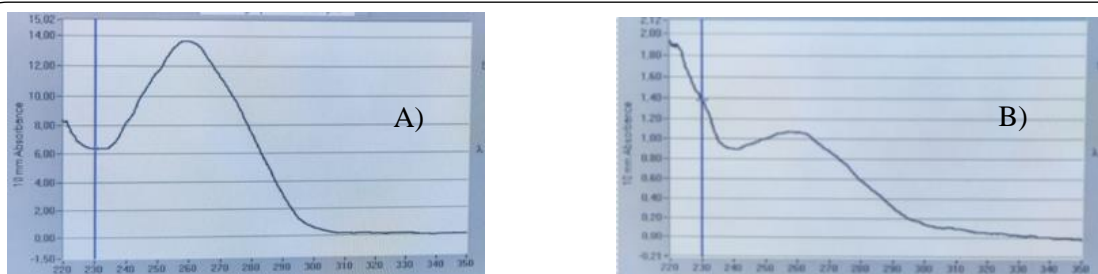
## 5. Resultados y discusión

En esta sección se presentan los resultados obtenidos a partir de los experimentos realizados sobre el estudio de genes de la especie de almeja *R. philippinarum* involucrados en la resistencia al parásito protozoario *P. olseni*. La organización de los datos obtenidos se realiza en función de los objetivos individuales previamente citados en apartados anteriores.

### 5.1 Extracción de muestras de ADN

El análisis del ADN extraído comienza inicialmente con la valoración de la pureza y concentración de las muestras de ADN. Para evaluar la pureza, se midieron unos cocientes de absorbancia a dos longitudes de onda distintas, en nuestro caso analizamos dos cocientes: 260nm / 280nm y 260nm / 230nm. Para asegurar su correcta calidad, los valores deberían ser próximos a 1,8 en el caso del cociente 260/280 y próximos a 2,0 en el caso del cociente 260/230.

Los valores de la relación de absorbancias a 260 y 280 se mantuvieron todos próximos a 1,8, lo que *a priori* validaría la calidad del ADN; en el caso de la relación 260/230, se obtuvieron valores en una franja numérica más amplia (desde 1,75 a 2,15) que igualmente corroborarían la correcta pureza de las muestras. Sin embargo, en esta última medición aparecieron 3 valores próximos a 0,6, representados en la figura 5b, que tuvieron que ser desechados al implicar presencia de fenol y otros contaminantes orgánicos en las muestras.



[Figura 5] Comparación gráfica Abs/longitud de onda: a) Muestra de ADN sin contaminantes; b) Muestra con contaminantes orgánicos o fenoles

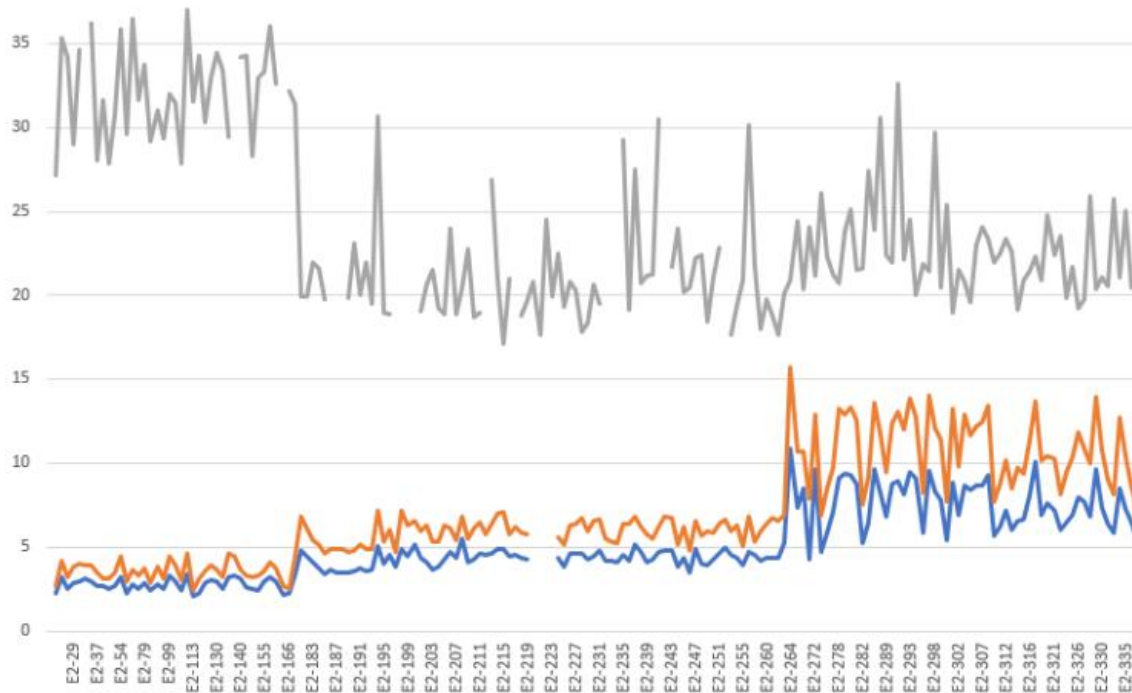
Posteriormente, se realizó una extracción de ADN de los padres candidatos utilizados para la fundación de las familias. Los valores obtenidos en la cuantificación de concentración y pureza del ADN de estos son muy similares a los obtenidos en la extracción anterior.

## 5.2 qPCR

A partir del ADN extraído previamente se realizó un análisis de la carga parasitaria de cada individuo para determinar la presencia o ausencia del ADN de *P. olsenii*. En este estudio, se analizó la posible correlación entre los valores obtenidos en la qPCR y el tamaño de los individuos; en concreto, se compararon los ciclos de la qPCR en los que se detectó el incremento exponencial de la curva del amplicón correspondiente al ADN del parásito con el tamaño de los individuos.

Los valores de tamaño se toman en relación al alto y ancho de los individuos. Ambas medidas mostraron una correlación muy elevada y significativa ( $r = 0,996$ ;  $p < 0,001$ ) por lo que el estudio se realiza teniendo en cuenta únicamente el ancho de la concha, ya que el empleo de las dos medidas resultaba estadísticamente redundante y acarrea un cierto error. En cuanto a los valores absolutos, oscilaron entre 2 mm y 15 mm; por otro lado, los valores de los ciclos de la qPCR oscilaron entre 17 y 38. Además, las medidas de tamaño, a su vez, mantienen también una fuerte correlación con el número de ciclos de PCR (carga parasitaria) ( $r = 0,689$ ;  $p < 0,001$ ).

Estos resultados obtenidos demostraron la existencia de una correlación significativa entre el tamaño y el nivel de infección de los individuos. Tras la representación de estos resultados, en la figura 6, se observa que los individuos con menores valores de alto y ancho mostraron un mayor número de ciclos en la PCR para ser observada la infección (similar a la observada en los controles negativos), lo que indica la prácticamente total ausencia de *P. olsenii* en las muestras.



[Figura 6] Gráfica realizada a partir de los valores obtenidos en la qPCR. En el eje de las abscisas encontramos los individuos y en el eje de ordenadas los valores de: altura de las conchas en mm (línea azul), anchura de las conchas en mm (línea naranja) y ciclos de la PCR a la que se detectó la presencia de *P. olsenii* en los individuos (línea gris). Gráfica completa con todos los resultados obtenidos en ANEXO I.

Esta ausencia del parásito en los individuos de menor tamaño se debe probablemente a la incapacidad de infectar a estas almejas a través del tejido branquial, bien porque aún no presentan unas branquias totalmente desarrolladas o simplemente por su pequeño tamaño y menor capacidad de filtración.

### 5.3 Genotipado

La realización del genotipado se llevó a cabo utilizando un panel compuesto por 8019 SNPs. De estos, se determinó por el mismo chip, que 3259 cumplían los parámetros óptimos preestablecidos para el estudio (especificados en el ANEXO II). En general, la validación de los SNPs se realizó siguiendo unos criterios rigurosos que incluían parámetros como la propia calidad del genotipado (suministrados por Thermo Fisher SCIENTIFIC) y los errores mendelianos detectados en las familias tras la reconstrucción genealógica.

Los SNPs validados se encontraban distribuidos uniformemente a lo largo del genoma, permitiendo una cobertura adecuada para estudios de asociación genética y, a su vez, la identificación de variantes genéticas asociadas a determinados fenotipos. Los 3259 SNPs utilizados, corresponden a la suma de los SNPs identificados como *CallRateBelowThreshold*, *PolyHighResolution*, *NoMinorHom* y *MonoHighResolution* (Howe *et al.*, 2013); mostrados en la tabla 1:

Tipo de SNP	Cantidad	Porcentaje
Other	4055	50.567 %
CallRateBelowThreshold	1657	20.663 %
PolyHighResolution	1458	18.182 %
OTV	705	8.792 %
NoMinorHom	106	1.322 %
MonoHighResolution	38	0.474 %

Tabla 1: Clasificación de la calidad del genotipado de los SNPs realizada por el software de Axiom.

Otra posible forma de elección de este subgrupo de SNPs sería directamente los seleccionados como mejores y recomendados en el propio archivo de salida. Serían 1602 SNPs, que corresponderían con el 19,98% de los SNPs genotipados en el panel inicial. Estos se compondrían por los SNPs determinados como *PolyHighResolution* junto con los *NoMinorHom* y los *MonoHighResolution*. Se desechó esta opción por representar un panel de SNPs de reducido número que limitaba la búsqueda significativa de asociaciones a nivel del genoma completo.

Por otra parte, con afán de aumentar el número de SNPs analizados en el GWAS se redujo el valor umbral predefinido de error desde un 98% a un 90%, aumentando el número de SNPs asignados como *CallRateBelowThreshold*. De esta forma, y junto con algunos otros SNPs añadidos por los buenos resultados obtenidos en el análisis FAP, se logra un panel total de 3904 SNPs definitivos.

#### 5.4 Asignación familiar (análisis de paternidad)

A partir de los genomas analizados mediante el programa FAP se obtuvieron una serie de archivos de salida con diferentes datos. Destacan en particular dos: un archivo de salida en formato de texto plano que presenta todas las posibles combinaciones de padres

para cada individuo del estudio con su respectiva probabilidad y un segundo archivo que presenta las mismas características que el anterior, pero que incluyen únicamente los cruces de padres más probables para cada descendiente utilizado en el estudio junto con su probabilidad asociada.

Dentro de todas las familias parentales que obtuvo el programa como más probables se excluyeron algunas por no tener un valor suficientemente significativo. El límite de significación en este apartado del estudio fue de 0,05, siguiendo la recomendación de los autores del programa (Taggart, 2006). Las familias que presentan valores de asignación por encima de este límite es probable que no sean reales, y su agrupación por el programa se debe a la falta de alguno de los progenitores usados en el estudio (tanto machos como hembras) que corresponderían justamente con los padres reales. Al no haber podido extraer el ADN de todos los progenitores quedaron algunos individuos que no pudieron asignarse correctamente por el programa. En concreto, 14 familias identificadas por este programa no fueron tomadas en cuenta por este motivo.

Por otro lado, las familias que se identificaron consistentemente (error < 0,05) son las representadas en la tabla 2 mostrada a continuación:

<b>Hembra (H)</b>	<b>Macho (M)</b>	<b>Nº hijos</b>
Hembra 1	Macho 3	74
Hembra 2	Macho 4	15
Hembra 3	Macho 5	4
Hembra 4	Macho 6	17
Hembra 6	Macho 8	137*
Hembra 8	Macho 10	1
Hembra 9	Macho 11	16
Hembra 12	Macho 15	1
Hembra 16	Macho 20	39*
Hembra 18	Macho 21	9
Hembra 19	Macho 23	10
Hembra 26	Macho 30	1
Hembra 27	Macho 32	6
Hembra 30	Macho 36	1
Hembra 33	Macho 41	6
Hembra 34	Macho 42	15

[Tabla 2a] Número de individuos de estudio asignación por cada pareja de progenitores significativamente posibles según FAP. Los valores que presentan asterisco (\*) corresponden con uno o más individuos descartados por superar un porcentaje de error del 5% con FAP.

Hembra (H)	Macho (M)	Nº hijos
Hembra 35	Macho 43	22
Hembra 36	Macho 44	14
Hembra 40	Macho 46	5
Hembra 41	Macho 47	1
Hembra 42	Macho 48	13
Hembra 45	Macho 49	15
Hembra 46	Macho 50	20
Hembra 48	Macho 59	22*
Hembra 49	Macho 60	12
Hembra 50	Macho 61	4
Hembra 51	Macho 62	8
Hembra 52	Macho 63	23
Hembra 55	Macho 66	29
Hembra 56	Macho 67	26
Hembra 58	Macho 68	25

[Tabla 2b] Continuación de la tabla 2a. Número de individuos de estudio asignación por cada pareja de progenitores significativamente posibles según FAP. Los valores que presentan asterisco (\*) corresponden con uno o más individuos descartados por superar un porcentaje de error del 5% con FAP.

Analizando los ovocitos que fueron expulsados al medio por cada hembra, se observó que había una discrepancia importante entre hembras con mucha producción de ovocitos y contribuciones a las progenies. *A priori* estos valores deberían estar correlacionados en la mayoría de los casos, ya que cuantas más células germinales se produzcan mayor probabilidad de fecundación habría. Sin embargo, no es el caso para todas las hembras analizadas tal como observamos en las figuras 7 y 8.



[Figura 7] Cantidad de ovocitos expulsados al medio por cada hembra. La distribución en el esquema es porcentual y los valores corresponden con datos del orden de más de 100.000 ovocitos por hembra.

En estas gráficas observamos como individuos con mucho peso en la generación de ovocitos, como las hembras 46 y 55, apenas tienen hijos. Sin embargo, hembras con poca producción de ovocitos presentan un gran número de hijos, como es el caso de la hembra 6. Esta discordancia podría explicarse tanto por una deficiente calidad de las puestas como por la ausencia de progenitores masculinos con calidad y cantidad de esperma apropiados. Además, se propone que las hembras que produjeron un gran número de ovocitos no tienen un peso importante en la figura 8 porque fueron fecundadas por otro macho que puede no está representado en la asignación parental (no representados en la gráfica), principalmente, por la falta de tejido a la hora de la extracción.



[Figura 8] Cantidad de almejas japónicas hijas por cada almeja madre. El primer valor corresponde con el número de la madre correspondiente y el segundo con el número de hijos que se le asignaron a través del programa FAP.

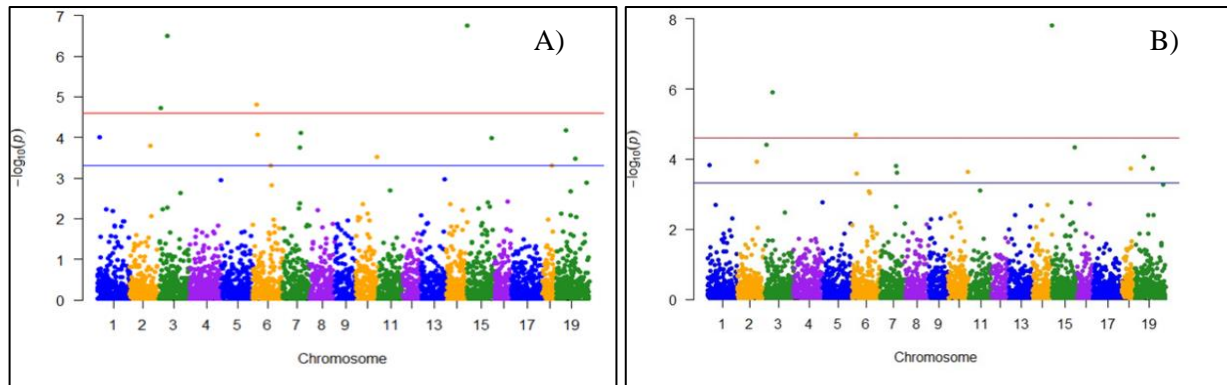
Dentro de los valores obtenidos por medio de este programa se diferencian tres combinaciones parentales diferentes: cuando supera el 10% de errores mendelianos en el análisis FAP de una pareja de macho y hembra se consideró que los dos progenitores son incorrectos para el hijo asignado, valores menores al 10% se interpretaron como la asignación correcta de al menos uno de los dos padres para ese determinado individuo y, finalmente en el caso de que el valor FAP sea menor al 5%, se consideró que ambos padres seleccionados se corresponden con son los padres reales de la almeja descendiente asignada.

Los SNPs iniciales localizados en estas familias generadas con errores inferiores al 5% se emplearon para aumentar el número de SNPs del panel seleccionado mediante el genotipado y, de esta forma, disponer de una mayor densidad de marcadores para el GWAS y consecuentemente incrementar la robustez del análisis.

## 5.5 GWAS

Para la realización de este análisis, se planteó el estudio utilizando diferentes paneles de SNPs seleccionados. Este tipo de análisis muestra resultados más robustos cuanto mayor es el número de SNPs analizados, al realizarse mayor número de test y representar de forma más amplia la totalidad del genoma. Debido a este motivo, se utilizaron los siguientes paneles de SNPs: un panel con el número total de SNPs genotipados con el chip de Affymetrix (8019), que a su vez fueron seleccionados previamente para este estudio; y un panel de SNPs reducido (3904), consecuencia de la selección de SNPs realizada por el propio chip de genotipado y los SNPs asociados a buenos valores en el análisis familiar. Además, se diferenciaron paneles con y sin corrección por tamaño, realizada por el propio software de GTCA, donde se incluía la corrección por tamaño de las almejas como un factor fijo. Finalmente, se realizaron los análisis con y sin los individuos pequeños, debido a que las almejas más pequeñas no mostraron capacidad de infectarse con el parásito, tal como mostró la qPCR.

De esta forma se realizaron ocho análisis GWAS con distintas combinaciones al haber tres variables en la realización de este estudio de asociación con dos opciones cada una ( $2^3$ ). Tras la comparación de los distintos Manhattan plot generados por este programa, se observaron diferencias mínimas entre los diferentes paneles de SNPs que presentan corrección por tamaño (o “fixed”) y los que no, como podemos ver en la figura 9. Por estos motivos, el estudio posterior se realizó utilizando el panel más robusto de SNPs (3904) y teniendo en cuenta la corrección por talla de las almejas generada por el propio programa.



[Figura 9] Comparativa GWAS a partir del panel de 3904 SNPs con corrección (A) y sin corrección (B).

En el caso del Manhattan plot de 3904 SNP con corrección observado en la figura 9A, se identificaron cuatro SNPs que sobrepasan el umbral de significación *genome-wide*, lo que implica una corrección más estricta del *p-value*; a su vez, otros diez SNPs mostraron *p-values* lo suficientemente bajos como para situarse por encima del umbral *chromosome-wide*, implicando asociaciones genéticas sugerentes y con una corrección para múltiples test menos estrictas. A grandes rasgos, la distribución de los SNPs con mayores valores de asociación a la resistencia contra *P. olsenii* se concentraron en los cromosomas 3, 6, 7, 15 y 19, como se observa en la tabla 2.

Chr	SNP	bp	A1	A2	Freq	b	se	<i>p-value</i>
1	AX-685330191	2705158	T	C	0.394434	-1,81249	0.465674	9,93471E-05
2	AX-685332421	36988789	T	G	0.0103383	416.009	110.197	0,000159913
3	AX-685333608	1933468	T	A	0.127753	1,87253	0.438271	1,93256E-05
3	AX-685333266	14358795	A	G	0.224706	1,46305	0.286299	3,21803E-07
6	AX-685338389	6985961	A	G	0.037037	2,77136	0.641691	1,56856E-05
6	AX-685338416	8072357	T	G	0.0859873	1,79522	0.456853	8,51102E-05
6	AX-685338331	32113105	G	T	0.10801	136.451	0.391812	0,000496659
7	AX-685352450	30531655	G	T	0.472619	-137.644	0.366946	0,00017608
7	AX-685340557	32056285	C	A	0.282864	-0,976091	0.247202	7,86243E-05
10	AX-685344335	38797696	T	C	0.224311	114.319	0.316871	0,000308867
15	AX-685348163	906197	A	G	0.0168539	4,87876	0.93484	1,80067E-07
15	AX-685349350	44413459	G	A	0.0269231	287.329	0.739563	0,000102277
18	AX-685352382	17893725	A	T	0.00267857	734.167	21.039	0,000483813
19	AX-685353058	16935047	A	T	0.0706751	1,9038	0.477547	6,70236E-05
19	AX-685353562	34899085	A	G	0.0821782	157.455	0.438748	0,000332289

[Tabla 2] Selección de *p-values* bajos en el panel de 3904 SNPs con corrección por tamaño del estudio por GWAS. Valores sombreados en rojo corresponden con *p-values* por encima del umbral *genome-wide* y sombreados en azul por encima del umbral de significación *chromosome-wide*.

Posteriormente, se analizó el Manhattan plot realizado a partir del panel de 3904 SNP con corrección de tamaño, eliminando los individuos muy pequeños que no presentaban infección ya que, como se explica previamente, se concluye que al no tener el tamaño del tejido branquial de un tamaño suficiente no pueden adquirir la infección.

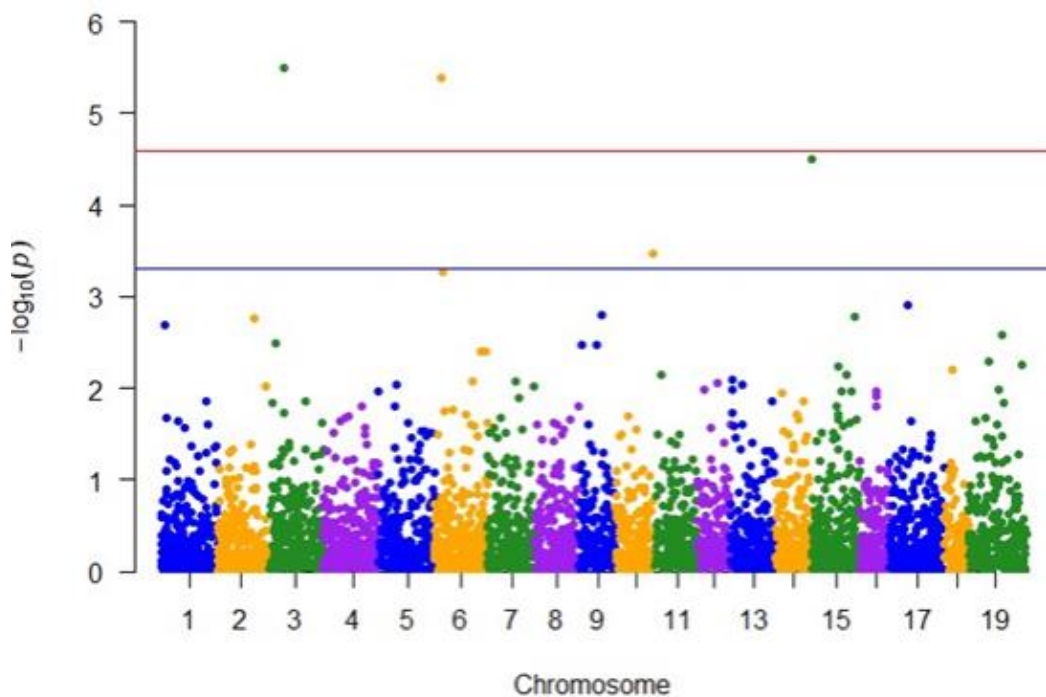
Chr	SNP	bp	A1	A2	Freq	b	se	p-value
3	AX-685333266	14358795	A	G	0.207071	115.044	0.246941	3,18117E-06
3	AX-685333736	6776682	A	G	0.249505	-0.67282	0.228685	0,00325969
6	AX-685338416	8072357	T	G	0.0733945	138.094	0.398801	0,000534738
6	AX-685338389	6985961	A	G	0.0239521	313.659	0.681261	4,14272E-06
6	AX-685338416	8072357	T	G	0.0733945	138.094	0.398801	0,000534738
6	AX-685339279	39963518	G	T	0.27453	-0.544587	0.206411	0,0083309
6	AX-685339476	47305118	G	T	0.362475	0.551031	0.191463	0,004002
6	AX-685339618	52797233	T	G	0.480926	0.735187	0.255042	0,0039439
10	AX-685344335	38797696	T	C	0.201613	103.747	0.28925	0,000334824
15	AX-685348163	906197	A	G	0.00302419	660.528	158.826	3,19909E-05
15	AX-685348867	26919981	C	G	0.203463	-0.613952	0.222866	0,00587283
15	AX-685349830	36270404	A	T	0.15102	-0.810445	0.300695	0,0070338
15	AX-685349350	44413459	G	A	0.0163265	237.351	0.753249	0,00162694
19	AX-685353154	20394039	C	T	0.446289	-0.809734	0.288851	0,00505831
19	AX-685353562	34899085	A	G	0.0667373	115.385	0.383118	0,00259765
19	AX-685354044	53913338	T	A	0.143443	-104.832	0.377375	0,00547069

[Tabla 3] Selección de *p-values* bajos en el panel de 3904 SNPs con corrección por tamaño y sin individuos pequeños. Valores sombreados en rojo corresponden con *p-values* por encima del umbral *genome-wide*, los sombreados en azul corresponden con valores por encima del umbral de significación *chromosome-wide* y los sombreados en gris con valores inferiores a ambos umbrales que presentan interés (por su *p-value* relativamente bajo o por el número de asociaciones significativas que tiene el cromosoma en el que se encuentran). Obtenida del archivo de salida del estudio por GWAS.

En este análisis detectamos una menor cantidad de SNPs significativos, tal y como se observa en la tabla 3. Además, se observó una variación en los resultados obtenidos con respecto a los Manhattan plot anteriormente señalados en la figura 9. En cuanto al peso de SNPs significativos por cromosoma, algunos siguen manteniendo una importante representación, como es el caso del cromosoma 6 y 15. *A priori*, la asociación mostrada por marcadores localizados en estos dos cromosomas en los diferentes GWAS realizados indican una asociación consistente entre la resistencia al parásito y los propios marcadores localizados en estos cromosomas.

En el caso de los cromosomas 3 y 19, siguen teniendo un peso importante, como se observa en la figura 10. Sin embargo, no presentan valores tan significativos como en los GWAS realizados con individuos de pequeño tamaño. Esto podría indicar una potencia insuficiente del test realizado, lo cual podría mejorarse mediante el incremento del número de individuos genotipados. De igual forma, una representación más equilibrada de las familias contribuiría a incrementar la potencia del test y a mejorar la significación.

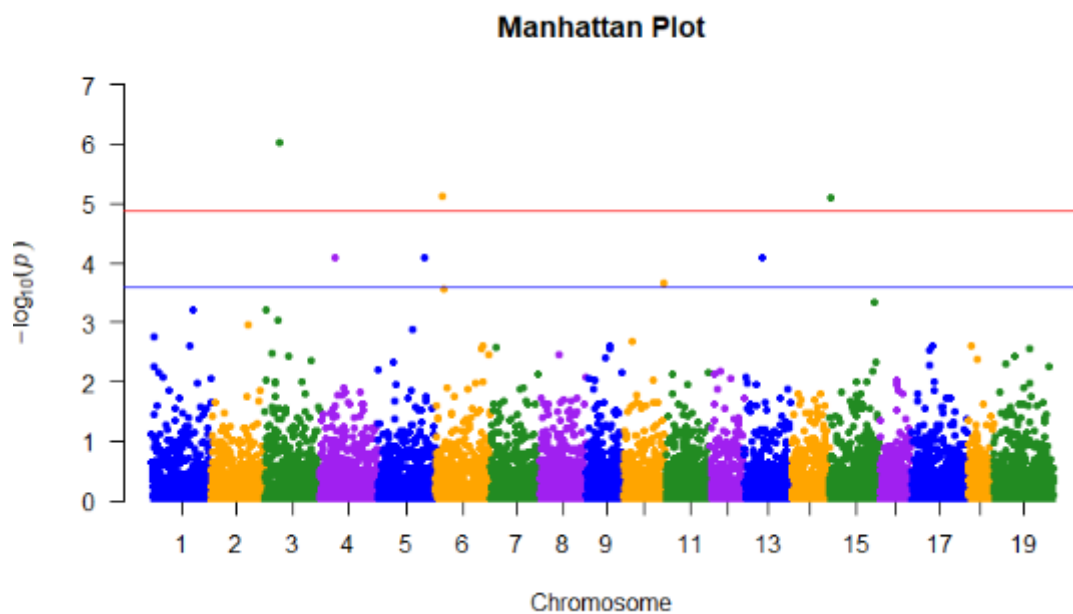
Por último, no se puede descartar la presencia de genes relacionado con el tamaño de las almejas en esos cromosomas, lo cual implicaría un menor tamaño en los individuos y, en consecuencia, una menor capacidad de infección.



[Figura 10] Manhattan plot generado a partir del panel de 3904 SNP corregido y sin individuos pequeños.

Por otro lado, cromosomas como el 9 muestran de una mayor representación y otros, como en el caso del cromosoma 7 reducen significativamente su asociación, por la posible presencia de genes relacionados con el tamaño. Probablemente ambas situaciones también se deban al aumento de robustez en los resultados ya que, aunque haya menor cantidad de SNPs, se entiende que todos estos pertenecen a individuos con un nivel de infección importante, lo que a su vez favorece estadísticamente su significatividad a la hora de asociarlos al concentrar una población idónea para el estudio.

Finalmente, se analizaron los resultados de los GWAS realizados a partir de los paneles de SNPs iniciales (8019) a modo de comparación y corroboración de resultados. En estos se observó una relación directa de SNPs significativos con los análisis realizados anteriormente, manteniendo una concentración alta en los cromosomas 3, 6, 15 y 19 como se observa en la figura 11. Los SNPs situados en estos cromosomas mostraron un nivel de significatividad alto y constante a lo largo de todos los estudios de asociación realizados, por lo que podemos considerarlas como las zonas de genoma más interesantes a la hora de realizar la minería de genes.



[Figura 11] Manhattan plot generado a partir del panel de 8019 SNP corregido y sin individuos pequeños.

## 5.6 Minería genómica

Posteriormente al GWAS, se localizan los SNPs que se consideran de mayor interés siguiendo principalmente dos ideas: su bajo  $p$ -value y la distribución que presentan a nivel cromosómico, siendo más interesante a la hora del estudio aquellos que se encuentren en cromosomas con varios SNPs significativos. Concretamente, se seleccionan 42 SNPs que en la mayoría de los casos eran comunes entre los GWAS realizados a partir de distintos paneles y muestras de individuos.

A la hora de identificarlos en el genoma, se tuvo que realizar un alineamiento previo contra el genoma actualizado para realizar la minería a partir de este último, tal como se explica en el apartado de materiales y métodos. Tras este proceso, únicamente 33 SNPs alinearon correctamente con el nuevo genoma por lo que son los que se analizaron en este estudio (identificados en el ANEXO III).

Acto seguido, se buscaron los genes situados en ventanas genómicas de 400 kb generadas rodeando a cada SNP estudiado. Tras este análisis, se localizaron genes en las regiones adyacentes a 33 SNPs. En estas regiones pueden estar presentes uno o más genes que, a su vez, pueden codificar para la síntesis de una o más proteínas como es el caso del SNP nombrado AX-685333608, situado en el cromosoma 3. En la ventana generada alrededor de este SNP se identifican 3 genes distintos: RPHI1A008242, RPHI1A022465 y RPHI1A011818 que sintetizan cada uno para una única proteína; en el caso del primero, codifica a su vez para dos isoformas diferentes. Normalmente, estas proteínas codificadas por un único gen presentan una estructura y funciones muy similares por lo que a la hora del estudio de sus términos GO asociados se pueden tomar, en la mayoría de los casos, como una función única.

A continuación, se clasificaron funcionalmente los genes codificantes de proteínas en esas regiones del genoma con sus respectivos términos GO. En este caso, únicamente los genes situados en las ventanas genómicas generadas por 15 de los 33 SNPs presentaban genes con un término GO asociado, ya que el genoma de esta especie de almeja aún no está suficientemente anotado.

Para este estudio de resistencia a *P. olseni* se pretende buscar genes con proteínas codificadas que presenten un término GO que pueda estar relacionado con la resistencia a este parásito.

Tras relacionar los SNPs con *p-values* más significativos en el GWAS con las proteínas cercanas se observó que apenas 3 de ellos tenían algún término GO asociado, hecho que limita el estudio, ya que algunos SNP situados en cromosomas con alta tasa de asociación y con *p-values* muy bajos no fueron identificados.

SNP	CHR	POSICION	GEN	PROTEINA	ONTOLOGÍA	TÉRMINO GO	DESCRIPCION
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	MF	0015293	symporter activity
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	BP	0055085	transmembrane transport
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	CC	0016020	membrane
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	BP	0006814	Sodium ion transport
AX-685338416	6	8072357	RPHI1A020365	RPHI1A020365P1	MF	0008234	cysteine-type peptidase activity
AX-685338416	6	8072357	RPHI1A020365	RPHI1A020365P1	BP	0006508	proteolysis
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	MF	0004930	G protein-coupled receptor activity
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	BP	0007186	G protein-coupled receptor signaling pathway
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	CC	0016020	membrane

[Tabla 4] Identificación de los SNPs con mayor grado de significatividad en el GWAS previo con los genes situados en sus inmediaciones. A su vez se relacionan estos genes con las proteínas para las cuales codifican y sus respectivos términos GO.

En la tabla 4 se muestran los 3 SNPs anteriormente citados, cada uno de ellos presentan numerosos genes identificados en esas regiones nucleotídicas; sin embargo, únicamente uno de estos genes codifica para proteínas con términos GO conocidos en cada una. A su vez, estas proteínas se asocian a varios términos GO.

El gen denominado RPHI1A040444 codifica una proteína relacionada con cuatro funciones: actividad simportadora, transporte transmembrana, transporte de iones sodio y actividad de membrana. Todas de ellas están relacionadas con la propia membrana celular y su integridad, estos términos GO se corresponden con fenotipos que pueden tener relación con la resistencia a *P. olsenii* ya que para infectar intracelularmente al hospedador necesita previamente atravesar la membrana celular. Por este motivo, distintas alteraciones en esta, pueden inhibir la capacidad de entrada de este parásito y por consecuencia generar resistencia.

En el mismo cromosoma en el que se encuentra el gen previamente descrito (6) se localiza otro gen, nombrado como RPHI1A020365. Este sintetiza una proteína que presenta dos términos GO asociados: proteólisis y actividad cisteína-proteasa. Estas funciones presentan de igual manera interés en su estudio ya que pueden estar relacionadas con mecanismos de defensa del hospedador contra el parásito, pudiendo

llegar a degradar proteínas fundamentales para su supervivencia una vez infecten la célula hospedadora o degradando proteínas que favorezcan el proceso de infección de *P. olseni*. En ambos casos, individuos que presenten estos fenotipos superarían la infección, en el primer caso generando una tolerancia hacia el patógeno y en el segundo una resistencia, al impedir directamente su entrada.

Por otro lado, se identificó un gen con un nivel de asociación a la resistencia en el cromosoma 10. Este gen, identificado como RPHI1A016912, sintetiza una proteína asociada a tres funciones: actividad del receptor acoplado a proteína G, vía de señalización del receptor acoplado a proteína G y a la actividad de la membrana. Como en el primer caso, estas diferencias a nivel nucleotídicas pueden provocar alteraciones en las proteínas de membrana, dando como resultado fenotipos que favorezcan la resistencia al patógeno al dificultar la interacción de las proteínas de reconocimiento del patógeno y, por ende, su infección.

Al ver la posible asociación de genes localizados en el cromosoma 6 con la resistencia a *P. olseni* se realizó un análisis más exhaustivo de SNPs localizados en esas regiones. Tras su búsqueda, únicamente se encontraron 2 genes situados en esas ventanas genómicas y que a su vez presenten términos GO conocidos, concretamente los genes RPHI1A038107 y RPHI1A040339. Sin embargo, los *p-values* obtenidos por medio del GWAS mostraron valores que no superaban el umbral de *chromosome-wide* y las funciones indicadas por los términos GO no parecen estar directamente relacionadas a la resistencia. Estas funciones son principalmente a nivel nuclear, como la organización de la cromatina o la unión al ADN, aunque también presentan relación con funciones como la fosforilización o la actividad dioxigenasa.

Para finalizar, cabe aclarar que de la misma forma que se actualiza la secuencia de *R.philippinarum*, se actualiza también la distribución de genes a nivel cromosómico en el último genoma publicado. Como resultado, los genes que se identifican en este trabajo en un determinado cromosoma se relacionaron a la hora del alineamiento con el genoma actualizado con cromosomas distintos, incluso genes que se encontraban en distintos cromosomas pasaron a estar situados en el mismo, y viceversa.

En el caso de los genes citados anteriormente, el primero (RPHI1A040444) se localiza en el cromosoma 16 en cambio, el segundo (RPHI1A020365), se localiza en el cromosoma 18, cuando en el genoma antiguo se situaba en el mismo cromosoma. Por último, RPHI1A016912 localizado en el cromosoma 10 pasa a estar en el cromosoma 7 y RPHI1A038107 y RPHI1A040339 se pasan de identificar del cromosoma 6 al 16 (manteniéndose en el mismo cromosoma que RPHI1A040444). Todos estos cambios se pueden observar en la tabla 5, que identifica el SNP inicial con el cromosoma en el que se encuentra según el genoma actualizado, a la vez que los genes localizados en la ventana genómica de 400.000 bp generada alrededor del SNP.

SNP	CHR	POSICIÓN	GENES (+200K)
AX-685344335	SUPER_7	43394045	RPHI1A016912
AX-685338389	SUPER_16	57264893	RPHI1A040444
AX-685339279	SUPER_16	6246965	RPHI1A038107
AX-685339618	SUPER_16	3307345	RPHI1A040339
AX-685338416	SUPER_18	38651429	RPHI1A020365
AX-685352382	SUPER_18	40353546	RPHI1A004898

[Tabla 5] SNPs con mejores valores de significatividad obtenidos del GWAS con su cromosoma correspondiente según el alineamiento realizado frente al genoma actualizado de *R.philippinarum*. Además, aporta información sobre los genes que se sitúan 400.000 bp alrededor del SNP. Para ver la tabla completa acudir al ANEXO IV.

Finalmente, se observa una correlación entre los resultados obtenidos y otros estudios relacionados con la resistencia a perkinsosis en distintas especies de bivalvos: análisis diferenciales entre individuos control e infectados con *P. marinus* demostraron un posible papel de la regulación de las proteasas en la respuesta rápida y organizada contra el parásito (Proestou & Sullivan, 2020).

Por otra parte, estudios en larvas de ostra del Pacífico expuestas a OsHV-1 indicaron una regulación positiva de las vías de señalización asociadas a proteínas G con la resistencia al virus (He *et al.*, 2015; Guo & Ford, 2016). Posteriormente, se demostró esta relación en larvas de ostra oriental infectadas con *P. marinus* (Proestou & Sullivan, 2020). Ambos trabajos, respaldan la asociación de este tipo de genes asociados a proteínas G, como RPHI1A16912, con la resistencia a patógenos comunes entre los bivalvos, previsiblemente por su papel en la respuesta a señales extracelulares y la regulación de procesos como la migración celular, la inflamación y la fagocitosis (Li *et al.*, 2019; Gao *et al.*, 2021).

## 6. Conclusiones

---

El objetivo principal del presente estudio era investigar las diversas variables genómicas asociadas con la resistencia al parásito *Perkinsus olseni* en familias de almeja japónica (*Ruditapes philippinarum*). Para alcanzar este objetivo, se emplearon enfoques metodológicos avanzados que incluyen el uso de tecnologías de genotipado de mediana densidad y estudios de asociación de genoma completo (GWAS). Concretamente:

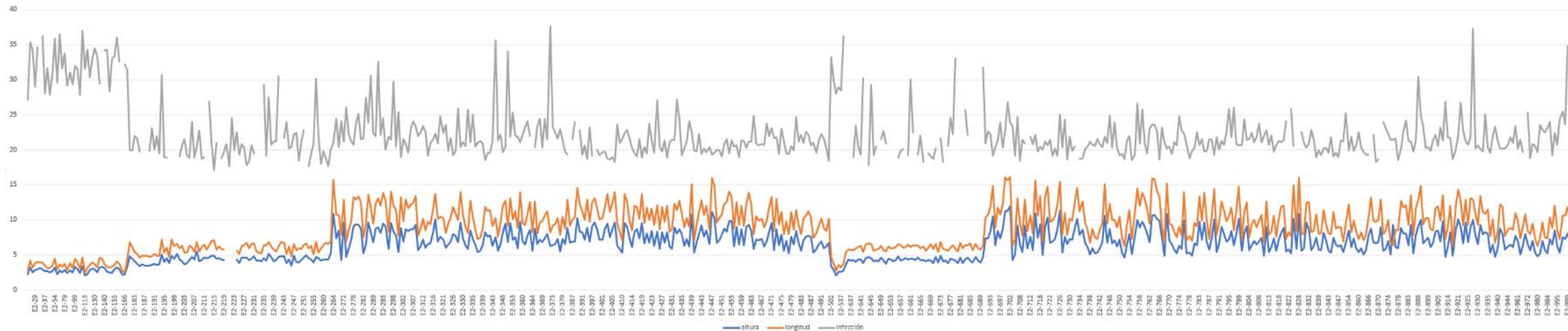
- Se generó una población genéticamente diversa mediante 40 cruzamientos de 1 macho x 1 hembra y se realizó posteriormente la cría y mantenimiento de la progenie. Esta variabilidad genética fue crucial a la hora de asegurar representativamente diversas variantes genéticas a lo largo del genoma con capacidad de influir en la resistencia a *P. olseni*.
- La población estudiada fue sometida a un desafío controlado con el parásito *P. olseni*. Este reto permitió evaluar la respuesta de cada individuo frente a la infección parasitaria.
- Se desarrolló una PCR cuantitativa (qPCR) con la intención de medir la carga parasitaria en cada individuo y se observó una relación directa entre tamaño de las almejas y el nivel de infección, consecuentemente se expuso la necesidad de más ciclos de PCR para identificar el ADN de *P. olseni* en los individuos de menor tamaño de almeja japónica.
- El ADN extraído de los individuos se empleó para genotipar, utilizando un chip de SNPs desarrollado en el laboratorio del grupo ACUIGEN que incluye un total de 8019 marcadores distribuidos uniformemente en el genoma de *R.philippinarum*. Posteriormente, se seleccionó un panel de 3259 SNPs que presentaban una calidad de genotipado óptima (98%).
- Gracias a un programa de asignación familiar (FAP) se identificaron 31 familias de individuos. Mediante este análisis se consiguieron rescatar SNPs no seleccionados en el genotipado como óptimos, dando lugar a un panel definitivo de 3904 SNPs.

- La realización de un estudio de asociación de genoma completo (GWAS) permitió identificar y caracterizar SNPs asociados a la resistencia al parásito *P. olseni*. Tras el análisis de varios paneles de SNPs se identificaron un total de 42 SNPs con *p-values* significativamente bajos y que en la mayoría de los casos eran comunes entre los GWAS realizados.
- A partir de los SNPs previamente citados se realizó una minería de genes en regiones genómicas de 200 kb a ambos lados de cada marcador y se localizaron genes posiblemente relacionados con la resistencia/tolerancia al parásito. Concretamente 3 genes fueron identificados en regiones próximas a los SNPs anteriores: RPHI1A040444, RPHI1A020365 y RPHI1A016912, y se estudiaron funciones de las proteínas para las que codificaban. Como consecuencia, se encontró una relación entre proteínas relacionadas con la membrana celular y la entrada del parásito (*P. olseni*), además de otras proteínas relacionadas con procesos proteolíticos.

Finalmente, se propone la obtención de una línea resistente a *P. olseni*, mediante la cría selectiva de individuos con marcadores genéticos asociados a la resistencia utilizando la información genotípica de los SNPs más relevantes. Esta estrategia no solo sería prometedora a la hora de mejorar la productividad en las poblaciones de *R.philippinarum* sino también proporcionando una herramienta efectiva para el manejo y control de patologías parasitarias a la vez que proporcionar herramientas prácticas para la selección genética.

## 7. Anexos

- **ANEXO I: Gráfica realizada a partir de los valores obtenidos en la qPCR:** En el eje de las abscisas encontramos los individuos y en el eje de ordenadas los valores de: altura (línea azul), anchura (línea naranja), ciclos de la PCR a la que se detectó la presencia de *P.olseni* en los individuos (línea gris).



➤ **ANEXO II: Umbrales de los parámetros en el control de calidad de SNPs (SNPs QC Thresholds).** Workflow Type: Genotyping

### SNPs QC Thresholds

- Species-type: Diploid
- cr-cutoff:  $\geq 97$
- fld-cutoff:  $\geq 3.6$
- het-so-cutoff:  $\geq -0.1$
- het-so-XChr-cutoff:  $\geq -0.1$
- het-so-ZChr-cutoff:  $\geq -0.1$
- het-so-otv-cutoff:  $\geq -0.3$
- hom-ro-1-cutoff:  $\geq 0.6$
- hom-ro-2-cutoff:  $\geq 0.3$
- hom-ro-3-cutoff:  $\geq -0.9$
- hom-ro: true
- num-minor-allele-cutoff:  $\geq 2$
- hom-ro-hap-1-cutoff:  $\geq 0.4$
- hom-ro-hap-1-XChr-cutoff:  $\geq 0.1$
- hom-ro-hap-1-MTChr-cutoff:  $\geq 0.4$
- hom-ro-hap-1-ZChr-cutoff:  $\geq 0.1$
- hom-ro-hap-2-cutoff:  $\geq 0.2$
- hom-ro-hap-2-XChr-cutoff:  $\geq 0.05$
- hom-ro-hap-2-MTChr-cutoff:  $\geq 0.2$
- hom-ro-hap-2-ZChr-cutoff:  $\geq 0.05$
- hom-hap-X-cutoff:  $\geq -0.2$
- hom-hap-Y-lower-cutoff:  $\leq 1$
- hom-hap-Y-upper-cutoff:  $\leq 0.2$
- CN0-hap-X-cutoff:  $\geq -0.2$
- CN0-hap-Y-cutoff:  $\geq 0$

- CN0-dip-X-cutoff:  $\geq 0.2$
- CN0-dip-Y-cutoff:  $\geq 0$
- aaf-XChr-cut:  $< 0.36$
- aaf-ZChr-cut:  $< 0.36$
- fld-XChr-cut:  $\geq 4$
- fld-ZChr-cut:  $\geq 4$
- homfld-XChr-cut:  $\geq 6.5$
- homfld-ZChr-cut:  $\geq 6.5$
- homfld-YChr-cut:  $\geq 6.5$
- homfld-WChr-cut:  $\geq 6.5$
- min-YChr-samples-cut:  $\geq 5$
- min-WChr-samples-cut:  $\geq 5$
- y-restrict:  $\leq 0.2$
- min-genotype-freq-samples:  $\geq 20$
- genotype-p-value-cutoff:  $\geq 1E-06$
- priority-order: PolyHighResolution, NoMinorHom, MonoHighResolution, OTV, UnexpectedGenotypeFreq, CallRateBelowThreshold, Other, OtherMA
- recommended: PolyHighResolution, NoMinorHom, MonoHighResolution, Hemizygous

### Multiplate Normalization Thresholds

- Samples per plate:  $\geq 8$
- fld-cutoff:  $\geq 4.9$
- fld-XChr-cut:  $\geq 4.9$

- **ANEXO III: Tabla con SNPs alineados correctamente por “Bowtie2” en el genoma actualizado:** se establece posición en el genoma y un nuevo cromosoma asignado, además de todos los genes localizados en las regiones periféricas al SNP ( $\pm 200k$  pb).

### CÓDIGO DE COLORES:

-Verde: al menos un gen en la región periférica al SNP ( $\pm 200k$  pb) que codifica para una proteína con término GO asociado. P-value en el GWAS significativo

-Naranja: al menos un gen en la región periférica al SNP ( $\pm 200k$  pb) que codifica para una proteína con término GO asociado. P-value en el GWAS no significativo, pero en chr de interés.

-Rojo: no se encontró ningún gen en la región periférica al SNP ( $\pm 200k$  pb) que codifica para una proteína con término GO asociado.

-Sin color: SNPs con p-values poco significativos y situados en cromosomas sin mucha representación en los análisis GWAS anteriores

SNP	CHR	POS	GENES ( $\pm 200K$ )
AX-685330505	SUPER_3	5147263	RPHI1A042632,RPHI1A000408,RPHI1A015952,RPHI1A037181,RPHI1A033911,RPHI1A004498,RPHI1A027464,RPHI1A018291,RPHI1A034679,RPHI1A031585,RPHI1A021496,RPHI1A029663,RPHI1A011702,RPHI1A017505,RPHI1A024206,RPHI1A016274,RPHI1A033456,RPHI1A038125,RPHI1A020681,RPHI1A025537,RPHI1A030081
AX-685330191	SUPER_3	9654358	RPHI1A008491,RPHI1A038510,RPHI1A020723,RPHI1A010407,RPHI1A028432,RPHI1A006094,RPHI1A013820,RPHI1A041752,RPHI1A018418,RPHI1A019652,RPHI1A014808,RPHI1A008740,RPHI1A026804,RPHI1A028147,RPHI1A034520,RPHI1A030105,RPHI1A024299,RPHI1A026288,RPHI1A046049,RPHI1A036804,RPHI1A007249,RPHI1A004050,RPHI1A020404,RPHI1A022790,RPHI1A039532,RPHI1A009997,RPHI1A018954,RPHI1A024863
AX-685330907	SUPER_3	57021043	RPHI1A016827,RPHI1A020479,RPHI1A004092,RPHI1A024535,RPHI1A008050
AX-685332421	SUPER_8	49812868	RPHI1A038430,RPHI1A009097,RPHI1A002766,RPHI1A019185,RPHI1A006878,RPHI1A023509,RPHI1A040374,RPHI1A033688,RPHI1A010420,RPHI1A003825,RPHI1A044049,RPHI1A014778
AX-685332921	SUPER_6	11934013	RPHI1A021499,RPHI1A028129,RPHI1A038104,RPHI1A044326,RPHI1A041877,RPHI1A026201,RPHI1A021065,RPHI1A030577,RPHI1A001675,RPHI1A042424,RPHI1A034336,RPHI1A029599,RPHI1A003341
AX-685333608	SUPER_6	9694809	RPHI1A022303,RPHI1A039059,RPHI1A010995,RPHI1A034353,RPHI1A006416,RPHI1A001603,RPHI1A019571,RPHI1A042651,RPHI1A037378,RPHI1A009795,RPHI1A004685,RPHI1A006816,RPHI1A011818,RPHI1A022465,RPHI1A027280,RPHI1A015967,RPHI1A020943,RPHI1A008242,RPHI1A013214
AX-685333736	SUPER_9	67621009	RPHI1A016489,RPHI1A010161,RPHI1A005965,RPHI1A002357,RPHI1A041975,RPHI1A009652,RPHI1A031738,RPHI1A018226,RPHI1A023344,RPHI1A031510
AX-685333921	SUPER_6_unloc_1	934443	RPHI1A037730,RPHI1A001599,RPHI1A004972,RPHI1A045410,RPHI1A011427,RPHI1A006474,RPHI1A018760,RPHI1A025551,RPHI1A031860,RPHI1A038172,RPHI1A044328,RPHI1A029334,RPHI1A001344,RPHI1A007439,RPHI1A014069,RPHI1A020285,RPHI1A018002,RPHI1A000310,RPHI1A031213,RPHI1A013058,RPHI1A022452
AX-685333266	SUPER_6	22685339	RPHI1A017957,RPHI1A005642,RPHI1A009025
AX-685333270	SUPER_6	22921501	RPHI1A005642,RPHI1A009025,RPHI1A014509,RPHI1A031368,RPHI1A036121,RPHI1A041330,RPHI1A028161
AX-685334621	SUPER_6	56283413	RPHI1A008143,RPHI1A011314,RPHI1A027854,RPHI1A044014,RPHI1A014644,RPHI1A031171,RPHI1A001548,RPHI1A017893,RPHI1A035667,RPHI1A008831,RPHI1A026163,RPHI1A024183,RPHI1A019273,RPHI1A036370,RPHI1A034220,RPHI1A017661,RPHI1A021190,RPHI1A016686,RPHI1A014853,RPHI1A040979,RPHI1A022608
AX-685334780	SUPER_13	45541610	RPHI1A033369,RPHI1A044134,RPHI1A026392,RPHI1A001556,RPHI1A015247,RPHI1A017496,RPHI1A003593

AX-685335442	SUPER_12	20725268	RPHI1A002222,RPHI1A030542,RPHI1A022075,RPHI1A028681,RPHI1A000084,RPHI1A029588,RPHI1A001294,RPHI1A040382
AX-685337864	SUPER_10	17193284	--
AX-685338389	SUPER_16	57264893	RPHI1A038967,RPHI1A011355,RPHI1A018083,RPHI1A000311,RPHI1A029387,RPHI1A035769,RPHI1A044121,RPHI1A026389,RPHI1A033928,RPHI1A016290,RPHI1A039607,RPHI1A022861,RPHI1A040444,RPHI1A012670
AX-685338416	SUPER_18	38651429	RPHI1A019289,RPHI1A024157,RPHI1A034198,RPHI1A004944,RPHI1A020365,RPHI1A037597,RPHI1A034330,RPHI1A010951,RPHI1A027934
AX-685339114	SUPER_13	27952640	RPHI1A001207,RPHI1A019143,RPHI1A022478,RPHI1A029166,RPHI1A010940,RPHI1A040219,RPHI1A007843
AX-685339279	SUPER_16	6246965	RPHI1A045608,RPHI1A038107,RPHI1A015216,RPHI1A043208,RPHI1A025576,RPHI1A007602,RPHI1A007715,RPHI1A035577,RPHI1A028016,RPHI1A010065
AX-685339618	SUPER_16	3307345	RPHI1A023947,RPHI1A023948,RPHI1A023945,RPHI1A023944,RPHI1A042993,RPHI1A040014,RPHI1A036479,RPHI1A030284,RPHI1A000900,RPHI1A040339,RPHI1A025140,RPHI1A019189,RPHI1A035287,RPHI1A044794
AX-685340557	SUPER_11	34820228	RPHI1A015732,RPHI1A046010,RPHI1A028394,RPHI1A010446,RPHI1A035200,RPHI1A017452,RPHI1A026453,RPHI1A030261,RPHI1A033835,RPHI1A037295,RPHI1A035334,RPHI1A004743,RPHI1A015814
AX-685344335	SUPER_7	43394045	RPHI1A000773,RPHI1A017357,RPHI1A010895,RPHI1A004176,RPHI1A044777,RPHI1A016912,RPHI1A005850,RPHI1A023724,RPHI1A018669,RPHI1A009136,RPHI1A012501,RPHI1A015140,RPHI1A030741,RPHI1A002440,RPHI1A017230,RPHI1A033385,RPHI1A005438,RPHI1A020576,RPHI1A038537
AX-685347074	SUPER_10	9945634	RPHI1A011315,RPHI1A018198,RPHI1A024346,RPHI1A040946,RPHI1A038783,RPHI1A033713,RPHI1A005985,RPHI1A023851,RPHI1A019416,RPHI1A011395,RPHI1A032231,RPHI1A024494,RPHI1A021165,RPHI1A013879,RPHI1A013686,RPHI1A030329,RPHI1A008635,RPHI1A042793,RPHI1A044657,RPHI1A038340,RPHI1A009564
AX-685348163	SUPER_8	426633	RPHI1A006482
AX-685349435	SUPER_9	26036062	RPHI1A020548,RPHI1A026956,RPHI1A033606,RPHI1A029798,RPHI1A000238,RPHI1A045528,RPHI1A016072,RPHI1A013469,RPHI1A027035,RPHI1A002163,RPHI1A042844,RPHI1A004245,RPHI1A043698,RPHI1A026563,RPHI1A008586,RPHI1A014071,RPHI1A042107,RPHI1A044605,RPHI1A029145,RPHI1A042136,RPHI1A026564,RPHI1A029126,RPHI1A014041,RPHI1A038584
AX-685348867	SUPER_8	49482368	RPHI1A026026,RPHI1A023941,RPHI1A018626,RPHI1A001457,RPHI1A040719,RPHI1A014155,RPHI1A030687,RPHI1A003999,RPHI1A020368,RPHI1A008769,RPHI1A038408,RPHI1A044413,RPHI1A028164,RPHI1A017533,RPHI1A001175,RPHI1A007716,RPHI1A036519,RPHI1A043019,RPHI1A026696,RPHI1A038430
AX-685349853	SUPER_18	10855373	RPHI1A045548,RPHI1A015896,RPHI1A010722
AX-685349984	SUPER_9_u nloc_1	22210	RPHI1A004333,RPHI1A043729,RPHI1A031012,RPHI1A001671,RPHI1A041001,RPHI1A012561,RPHI1A045780
AX-685352382	SUPER_18	40353546	RPHI1A038770,RPHI1A030036,RPHI1A000644,RPHI1A041911,RPHI1A032824,RPHI1A004898,RPHI1A034308,RPHI1A037581,RPHI1A020992,RPHI1A031033,RPHI1A020125,RPHI1A042149,RPHI1A011244,RPHI1A002662,RPHI1A034841,RPHI1A019723
AX-685352790	SUPER_2	74048818	RPHI1A026851,RPHI1A031683,RPHI1A015157,RPHI1A022058,RPHI1A027720,RPHI1A010268,RPHI1A039585,RPHI1A000185
AX-685353470	SUPER_2	82998770	RPHI1A014007,RPHI1A029423,RPHI1A001279,RPHI1A016627,RPHI1A029397,RPHI1A001244,RPHI1A016556,RPHI1A000999,RPHI1A040266
AX-685353154	SUPER_2	49844735	RPHI1A039218,RPHI1A009877,RPHI1A026330,RPHI1A019426,RPHI1A035892,RPHI1A045738,RPHI1A000251,RPHI1A039156,RPHI1A002830,RPHI1A022280,RPHI1A034605,RPHI1A033806,RPHI1A026633,RPHI1A042937,RPHI1A014872,RPHI1A002605,RPHI1A030620
AX-685354224	SUPER_3	84265954	RPHI1A031615,RPHI1A015086,RPHI1A021744,RPHI1A005234,RPHI1A027129,RPHI1A043454,RPHI1A020698,RPHI1A007053,RPHI1A000791,RPHI1A017376,RPHI1A029087,RPHI1A028257
AX-685354044	SUPER_7	21737638	RPHI1A000985,RPHI1A030241,RPHI1A013782,RPHI1A027303,RPHI1A010775,RPHI1A044497,RPHI1A016231,RPHI1A004138,RPHI1A041478,RPHI1A039672,RPHI1A031835,RPHI1A029659,RPHI1A001149,RPHI1A018947,RPHI1A037080,RPHI1A010584,RPHI1A038255,RPHI1A040096,RPHI1A028135,RPHI1A017594,RPHI1A005177,RPHI1A007294

➤ **ANEXO IV: Tabla relación genes con términos GO:** SNPs sobre umbral *genome-wide* (verde) y SNPs en cromosomas de interés (naranja).

AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	MF	0015293	symporter activity
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	BP	0055085	transmembrane transport
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	CC	0016020	membrane
AX-685338389	6	6985961	RPHI1A040444	RPHI1A040444P1	BP	0006814	sodium ion transport
AX-685338416	6	8072357	RPHI1A020365	RPHI1A020365P1	MF	0008234	cysteine-type peptidase activity
AX-685338416	6	8072357	RPHI1A020365	RPHI1A020365P1	BP	0006508	proteolysis
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	CC	0005929	cilium
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	MF	0004674	protein serine/threonine kinase activity
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	BP	0016310	phosphorylation
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	BP	0060271	cilium assembly
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	CC	0005634	nucleus
AX-685339279	6	39963518	RPHI1A038107	RPHI1A038107P1	MF	0005524	ATP binding
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P1	MF	0051213	dioxygenase activity
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P1	BP	0006325	chromatin organization
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P1	CC	0005634	nucleus
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P1	MF	0003677	DNA binding
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P1	MF	0046872	metal ion binding
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P2	MF	0051213	dioxygenase activity
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P2	BP	0006325	chromatin organization
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P2	CC	0005634	nucleus
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P2	MF	0003677	DNA binding
AX-685339618	6	52797233	RPHI1A040339	RPHI1A040339P2	MF	0046872	metal ion binding
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	MF	0004930	G protein-coupled receptor activity
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	BP	0007186	G protein-coupled receptor signaling pathway
AX-685344335	10	38797696	RPHI1A016912	RPHI1A016912P1	CC	0016020	membrane
AX-685352382	18	17893725	RPHI1A004898	RPHI1A004898P1	CC	0046540	U4/U6 x U5 tri-snRNP complex
AX-685352382	18	17893725	RPHI1A004898	RPHI1A004898P1	BP	0000398	mRNA splicing, via spliceosome
AX-685352382	18	17893725	RPHI1A004898	RPHI1A004898P2	CC	0046540	U4/U6 x U5 tri-snRNP complex
AX-685352382	18	17893725	RPHI1A004898	RPHI1A004898P2	BP	0000398	mRNA splicing, via spliceosome

## 8. Referencias bibliográficas

- Azéma, P., Lamy, J., Boudry, P., Renault, T., Travers, M., & Dégremont, L. (2017).** Genetic parameters of resistance to *Vibrio aestuarianus*, and OsHV-1 infections in the Pacific oyster, *Crassostrea gigas*, at three different life stages. *Genetics Selection Evolution*, 49(1). <https://doi.org/10.1186/s12711-017-0297-2>
- Ben-Horin, T., Burge, C. A., Bushek, D., Groner, M. L., Proestou, D. A., Huey, L. I., Bidegain, G., & Carnegie, R. B. (2018).** Intensive oyster aquaculture can reduce disease impacts on sympatric wild oysters. *Aquaculture Environment Interactions*, 10, 557-567. <https://doi.org/10.3354/aei00290>
- Burge, C. A., Eakin, C. M., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., Petes, L. E., Prager, K. C., Weil, E., Willis, B. L., Ford, S. E., & Harvell, C. D. (2014).** Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society. *Annual Review Of Marine Science*, 6(1), 249-277. <https://doi.org/10.1146/annurev-marine-010213-135029>
- Bushek, D., Ford, S., & Chintala, M. (2002).** Comparison of in vitro-cultured and wild-type *Perkinsus marinus*. III. Fecal elimination and its role in transmission. *Diseases Of Aquatic Organisms*, 51(3), 217-225. <https://doi.org/10.3354/dao051217>
- Calderón-Chagoya, R., Vega-Murillo, V. E., García-Ruiz, A., Ríos-Utrera, Á., Martínez-Velázquez, G., & Montañón-Bermúdez, M. (2023).** Genome and chromosome wide association studies for growth traits in Simmental and Simbrah cattle. *Animal Bioscience*, 36(1), 19-28. <https://doi.org/10.5713/ab.21.0517>
- Carnegie, R. B., Arzul, I., & Bushek, D. (2016).** Managing marine mollusc diseases in the context of regional and international commerce: policy issues and emerging concerns. *Philosophical Transactions - Royal Society. Biological Sciences*, 371(1689), 20150215. <https://doi.org/10.1098/rstb.2015.0215>
- Choi, K., & Park, K. (1997).** Report on the Occurrence of *Perkinsus* sp. in the Manila Clams, *Ruditapes philippinarum* in Korean. *Journal Of Aquaculture*, 10(3), 227-237. [http://www.koreascience.or.kr/article/ArticleFullRecord.jsp?cn=GSHHBD\\_1997\\_v10n3\\_227](http://www.koreascience.or.kr/article/ArticleFullRecord.jsp?cn=GSHHBD_1997_v10n3_227)
- Chu, F., & La Peyre, J. (1993).** *Perkinsus marinus* susceptibility and defense-related activities in eastern oysters *Crassostrea virginica*: temperature effects. *Diseases Of Aquatic Organisms*, 16(1), 223-234. <https://doi.org/10.3354/dao016223>
- Chung, E., Chung, J., & Lee K. (2013).** Gametogenic Cycle, the Spawning Season, First Sexual Maturity, and the Biological Minimum Size in Male *Ruditapes philippinarum* (Bivalvia: Veneridae) in Western Korea. *Journal Of Life Sciences*. <https://doi.org/10.17265/1934-7391/2013.06.008>
- De Ruijter, A., & Guldenmund, F. (2016).** The bowtie method: A review. *Safety Science*, 88(1), 211-218. <https://doi.org/10.1016/j.ssci.2016.03.001>
- Du, Z., Zhou, X., Ling, Y., Zhang, Z., & Su, Z. (2010).** agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research*, 38(suppl\_2), W64-W70. <https://doi.org/10.1093/nar/gkq310>
- FAO. (2015).** FAO Global Aquaculture Production database updated to 2013 –

- Summary information. See <http://www.fao.org/3/a-i4899e.pdf>.
- Fernández-Boo, S., Villalba, A., & Cao, A. (2015).** Cell proteome variability of protistan mollusc parasite *Perkinsus olseni* among regions of the Spanish coast. *Diseases Of Aquatic Organisms*, *113*(3), 245-256. <https://doi.org/10.3354/dao02835>
- Fraslin, C., Phocas, F., Bestin, A., Charles, M., Bernard, M., Krieg, F., Dechamp, N., Ciobotaru, C., Hozé, C., Petitprez, F., Milhes, M., Lluch, J., Bouchez, O., Poncet, C., Hocdé, P., Haffray, P., Guiguen, Y., & Quillet, E. (2020).** Genetic determinism of spontaneous masculinisation in XX female rainbow trout: new insights using medium throughput genotyping and whole-genome sequencing. *Scientific Reports*, *10*(1). <https://doi.org/10.1038/s41598-020-74757-8>
- Gao, H., Li, Y., Wang, M., Song, X., Tang, J., Feng, F., & Li, B. (2021).** Identification and Expression Analysis of G Protein-Coupled Receptors in the Miridae Insect *Apolygus lucorum*. *Frontiers In Endocrinology*, *12*. <https://doi.org/10.3389/fendo.2021.773669>
- Garcia-Diaz, M., & Kunkel, T. A. (2006).** Mechanism of a genetic glissando\*: structural biology of indel mutations. *Trends In Biochemical Sciences*, *31*(4), 206-214. <https://doi.org/10.1016/j.tibs.2006.02.004>
- Guo, X., & Ford, S. E. (2016).** Infectious diseases of marine molluscs and host responses as revealed by genomic tools. *Philosophical Transactions - Royal Society. Biological Sciences*, *371*(1689), 20150206. <https://doi.org/10.1098/rstb.2015.0206>
- Guo, X., He, Y., Zhang, L., Lelong, C., & Jouaux, A. (2015).** Immune and stress responses in oysters with insights on adaptation. *Fish & Shellfish Immunology*, *46*(1), 107-119. <https://doi.org/10.1016/j.fsi.2015.05.018>
- Hanrio, E., Batley, J., Dungan, C., & Dang, C. (2021).** Immunoassays and diagnostic antibodies for *Perkinsus spp.* pathogens of marine molluscs. *Diseases Of Aquatic Organisms*, *147*(1), 13-23. <https://doi.org/10.3354/dao03635>
- Harvell, C. D., Kim, K., Burkholder, J. M., Colwell, R. R., Epstein, P. R., Grimes, D. J., Hofmann, E. E., Lipp, E. K., Osterhaus, A. D. M. E., Overstreet, R. M., Porter, J. W., Smith, G. W., & Vasta, G. R. (1999).** Emerging Marine Diseases--Climate Links and Anthropogenic Factors. *Science*, *285*(5433), 1505-1510. <https://doi.org/10.1126/science.285.5433.1505>
- He, Y., Jouaux, A., Ford, S. E., Lelong, C., Sourdain, P., Mathieu, M., & Guo, X. (2015).** Transcriptome analysis reveals strong and complex antiviral response in a mollusc. *Fish & Shellfish Immunology*, *46*(1), 131-144. <https://doi.org/10.1016/j.fsi.2015.05.023>
- Howe, G. T., Jayawickrama, K., Kolpak, S. E., Kling, J., Trappe, M., Hipkins, V., Ye, T., Guida, S., Cronn, R., Cushman, S. A., & McEvoy, S. (2020).** An Axiom SNP genotyping array for Douglas-fir. *BMC Genomics*, *21*(1). <https://doi.org/10.1186/s12864-019-6383-9>
- Howe, G. T., Yu, J., Knaus, B., Cronn, R., Kolpak, S., Dolan, P., Lorenz, W. W., & Dean, J. F. (2013).** A SNP resource for Douglas-fir: de novo transcriptome assembly and SNP detection and validation. *BMC Genomics*, *14*(1), 137. <https://doi.org/10.1186/1471-2164-14-137>
- Irizarry, R. A. (2003).** Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, *31*(4), 15e-115. <https://doi.org/10.1093/nar/gng015>
- Li, H., Kong, N., Sun, J., Wang, W., Li, M., Gong, C., Dong, M., Wang, M., Wang, L., & Song, L. (2019).** A C1qDC (CgC1qDC-6) with a collagen-like domain mediates hemocyte phagocytosis and migration in oysters. *Developmental And Comparative Immunology/Developmental & Comparative Immunology*

- Immunology*, 98, 157-165.  
<https://doi.org/10.1016/j.dci.2019.04.010>
- Lipshutz, R. J., Fodor, S. P., Gingeras, T. R., & Lockhart, D. J. (1999).** High density synthetic oligonucleotide arrays. *Nature Genetics*, 21(S1), 20-24.  
<https://doi.org/10.1038/4447>
- Liste, S. M. C. (2002).** Estudio de la perkinsosis en la almeja fina, *Tapes decussatus* (Linnaeus, 1758), de Galicia. Dialnet.  
<https://dialnet.unirioja.es/servlet/tesis?codigo=109604>
- Martenot, C., Oden, E., Travaillé, E., Malas, J., & Houssin, M. (2011).** Detection of different variants of Ostreid Herpesvirus 1 in the Pacific oyster, *Crassostrea gigas* between 2008 and 2010. *Virus Research*, 160(1-2), 25-31.  
<https://doi.org/10.1016/j.virusres.2011.04.012>
- Moullac, G. L., Cheize, M., Gastineau, O., Daniel, J., Coz, J. R. L., Huvet, A., Moal, J., Pouvreau, S., Van Wormhoudt, A., & Samain, J. (2008).** Ecophysiological and Metabolic Adaptations to Sulphide Exposure of the Oyster *Crassostrea gigas*. *Journal of Shellfish Research*, 27(2), 355-363.  
[https://doi.org/10.2983/0730-8000\(2008\)27\[355:EAMATS\]2.0.CO;2](https://doi.org/10.2983/0730-8000(2008)27[355:EAMATS]2.0.CO;2)
- Nava-Cedeño, D. J., Alonso-Cerezo, M. C., Sanz-García, A., Vega-Zelaya, L., Gordillo-Perdomo, J. J., Toledo-Heras, M., Pastor-Gómez, J., Torres-Díaz, C., Pulido-Rivas, P., & García-Sola, R. (2022).** Asociación entre los polimorfismos genéticos de nucleótido único en genes transportadores ABC con la epilepsia farmacorresistente en la población española. *Revista de Neurología/Revista de Neurología Electrónica*, 75(09), 251.  
<https://doi.org/10.33588/rn.7509.2022133>
- Naylor, R. L., Hardy, R. W., Buschmann, A. H., Bush, S. R., Cao, L., Klinger, D. H., Little, D. C., Lubchenco, J., Shumway, S. E., & Troell, M. (2021).** A 20-year retrospective review of global aquaculture. *Nature*, 591(7851), 551-563.  
<https://doi.org/10.1038/s41586-021-03308-6>
- Pampín, M., Casanova, A., Fernández, C., Blanco, A., Hermida, M., Vera, M., Pardo, B. G., Coimbra, R. M., Cao, A., Iglesias, D., Carballal, M., Villalba, A., & Martínez, P. (2023).** Genetic markers associated with divergent selection against the parasite *Marteilia cochillia* in common cockle (*Cerastoderma edule*) using transcriptomics and population genomics data. *Frontiers In Marine Science*, 10.  
<https://doi.org/10.3389/fmars.2023.1057206>
- Perkins, F. O., & Menzel, R. W. (1966).** Morphological and cultural studies of a motile stage in the life cycle of *Dermocystidium marinum*. *Proc Natl Shellfish Assoc*, 56, 23.  
<https://scholarworks.wm.edu/cgi/viewcontent.cgi?article=3046&context=vimsarticles>
- Proestou, D. A., & Sullivan, M. E. (2020).** Variation in global transcriptomic response to *Perkinsus marinus* infection among eastern oyster families highlights potential mechanisms of disease resistance. *Fish & Shellfish Immunology*, 96, 141-151.  
<https://doi.org/10.1016/j.fsi.2019.12.001>
- Proestou, D. A., Sullivan, M. E., Lundgren, K. M., Ben-Horin, T., Witkop, E. M., & Hart, K. M. (2023).** Understanding *Crassostrea virginica* tolerance of *Perkinsus marinus* through global gene expression analysis. *Frontiers in Genetics*, 14(1).  
<https://doi.org/10.3389/fgene.2023.1054558>
- Qu, H., Tien, M., & Polychronakos, C. (2010).** Statistical significance in genetic association studies. *Clinical And Investigative Medicine*, 33(5), 266.  
<https://doi.org/10.25011/cim.v33i5.14351>
- Råberg, L., Graham, A. L., & Read, A. F. (2008).** Decomposing health: tolerance and resistance to parasites in animals. *Philosophical Transactions - Royal Society. Biological Sciences*, 364(1513), 37-49.  
<https://doi.org/10.1098/rstb.2008.0184>

- Råberg, L., Sim, D., & Read, A. F. (2007). Disentangling Genetic Variation for Resistance and Tolerance to Infectious Diseases in Animals. *Science*, 318(5851), 812-814.  
<https://doi.org/10.1126/science.1148526>
- Ríos, R., Aranguren, R., Gastaldelli, M., Arcangeli, G., Novoa, B., & Figueras, A. (2020). Development and validation of a specific real-time PCR assay for the detection of the parasite *Perkinsus olseni*. *Journal Of Invertebrate Pathology*, 169(1), 107301.  
<https://doi.org/10.1016/j.jip.2019.107301>
- Secretaría General de Pesca & Ministerio de Agricultura, Pesca y Alimentación. Gobierno de España. (2023a). ALMEJAS: Almeja babosa, fina, japonesa y rubia. *EsAcui*.[https://www.mapa.gob.es/es/pesca/temas/acuicultura/almejas\\_tcm30-628430.pdf](https://www.mapa.gob.es/es/pesca/temas/acuicultura/almejas_tcm30-628430.pdf)
- Secretaría General de Pesca. Ministerio de agricultura, pesca y alimentación. Gobierno de España. (2023b). VALOR DE PRIMERA VENTA (EUROS) PRODUCCIÓN ACUÍCOLA NACIONAL DE ALMEJA JAPONESA. *EsAcui*.  
[https://www.mapa.gob.es/es/pesca/temas/acuicultura/pv\\_alm\\_japonesa\\_tcm30-628141.pdf](https://www.mapa.gob.es/es/pesca/temas/acuicultura/pv_alm_japonesa_tcm30-628141.pdf)
- Segarra, A., Pépin, J. F., Arzul, I., Morga, B., Faury, N., & Renault, T. (2010). Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. *Virus Research*, 153(1), 92-99.  
<https://doi.org/10.1016/j.virusres.2010.07.011>
- Shastri, B. S. (2009). SNPs: Impact on Gene Function and Phenotype. *Methods in molecular biology* (pp. 3-22).  
[https://doi.org/10.1007/978-1-60327-411-1\\_1](https://doi.org/10.1007/978-1-60327-411-1_1)
- Taggart, J. B. (2006). PROGRAM NOTE: FAP: an exclusion-based parental assignment program with enhanced predictive functions. *Molecular Ecology Notes*, 7(3), 412-415.  
<https://doi.org/10.1111/j.1471-8286.2006.01616.x>
- Vera, M., Wilmes, S. B., Maroso, F., Hermida, M., Blanco, A., Casanova, A., Iglesias, D., Cao, A., Culloty, S. C., Mahony, K., Orvain, F., Bouza, C., Robins, P. E., Malham, S. K., Lynch, S., Villalba, A., & Martínez, P. (2023). Heterogeneous microgeographic genetic structure of the common cockle (*Cerastoderma edule*) in the Northeast Atlantic Ocean: biogeographic barriers and environmental factors. *Heredity*, 131(4), 292-305.  
<https://doi.org/10.1038/s41437-023-00646-1>
- Villalba, A., Casas, S., López, C., & Carballal, M. (2005). Study of perkinsosis in the carpet shell clam *Tapes decussatus* in Galicia (NW Spain). II. Temporal pattern of disease dynamics and association with clam mortality. *Diseases Of Aquatic Organisms*, 65(1), 257-267.  
<https://doi.org/10.3354/dao065257>
- Vogeler, S., Carboni, S., Li, X., Nevejan, N., Monaghan, S. J., Ireland, J. H., & Joyce, A. (2020). Bivalves are NO different: nitric oxide as negative regulator of metamorphosis in the Pacific oyster, *Crassostrea gigas*. *BMC Developmental Biology*, 20(1).  
<https://doi.org/10.1186/s12861-020-00232-2>
- Wijayalath, W., Majji, S., Kleschenko, Y., Pow-Sang, L., Brumeanu, T. D., Villasante, E. F., Vasta, G. R., Fernández-Robledo, J., & Casares, S. (2014). Humanized HLA-DR4 Mice Fed with the Protozoan Pathogen of Oysters *Perkinsus Marinus* (Dermo) Do Not Develop Noticeable Pathology but Elicit Systemic Immunity. *PloS One*, 9(1), e87435.  
<https://doi.org/10.1371/journal.pone.0087435>
- World Organisation for Animal Health. (2021, 27 marzo). Clasificación de enfermedades de declaración obligatoria a la OIE – Lista B - OMSA - Organización Mundial de Sanidad Animal.  
<https://www.woah.org/es/que-hacemos/sanidad-y-bienestar->

[animal/enfermedades-animales/antigua-clasificacion-de-enfermedades-de-declaracion-obligatoria-a-la-oie-lista-b/](#)

**Yadavalli, R., Umeda, K., & Robledo, J. A. F. (2020).** *Perkinsus marinus*. *Trends In Parasitology*, 36(12), 1013-1014.  
<https://doi.org/10.1016/j.pt.2020.05.002>

**Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011).** GCTA: A Tool for Genome-wide Complex Trait Analysis. *American Journal Of Human Genetics*, 88(1), 76-82.  
<https://doi.org/10.1016/j.ajhg.2010.11.011>

**Zheng, S., Zhang, T., Tu, K., Li, L., Liu, Z., Wu, B., Zhou, L., & Sun, X. (2023).** Population Genetics of Manila Clam (*Ruditapes philippinarum*) in China Inferred from Microsatellite Markers. *Biology*, 12(4).  
<https://doi.org/10.3390/biology12040557>

