



Universidad  
de Concepción

TIENDA VIRTUAL  
UDEC

DOCUMENTO DE CAJA

**12702861**

FECHA DE PAGO

**11-04-2022**

Corr	Descripción	Persona	Cantidad	Monto
1372344	FORMACIÓN DEL EXPEDIENTE, PRIMER AÑO	19.808.791-2 HAUENSTEIN DAROCH, JAIME ANTONIO	1	\$5.000
1372343	CERTIFICADO DE TÍTULO, PRIMER AÑO	19.808.791-2 HAUENSTEIN DAROCH, JAIME ANTONIO	1	\$6.500
			<b>TOTAL</b>	<b>\$11.500</b>



Universidad de Concepción

Carrera Biología Marina

Departamento de Oceanografía

Facultad de Ciencias Naturales y Oceanográficas

**CARACTERIZACIÓN DE RNAs LARGOS NO CODIFICANTES (lncRNAs) EN  
CEPAS DE *Caligus rogercresseyi* MEDIANTE TECNOLOGÍAS DE  
SECUENCIACIÓN Y EVALUACIÓN DE LA REGULACIÓN  
TRANSCRIPCIONAL ANTE LA EXPOSICIÓN FARMACOLÓGICA.**

Por Jaime Antonio Hauenstein Daroch

Tesis presentada a la facultad de Ciencias Naturales y Oceanográficas de la  
Universidad de Concepción para optar al título profesional de Biólogo Marino

Profesor patrocinante: Dr. Cristian Gallardo Escárate

Director de tesis: Dr. Diego Valenzuela Miranda

Marzo de 2022, Concepción

## **AGRADECIMIENTOS**

Agradezco a mi familia por todo el cariño y apoyo incondicional que me han dado a lo largo de mi vida, en especial a mi madre Gladys y a mi abuela Maria Nemesia, que siempre se han preocupado de que nunca nos falte nada a mi y a mis hermanos. También agradezco a mis amigos y amigas que siempre han estado ahí y han sido una grata compañía a través de este proceso, además del apoyo mutuo y la retroalimentación generada, en especial Karina y Manuel que han sido una grata compañía a través de todo este proceso. A todos aquellos profesores con vocación que conocí a lo largo de mis estudios. Al Dr. Cristian Gallardo-Escárte que me brindó su apoyo en esta última etapa de mi carrera, inspirándome a profundizar en el campo de la genética y darme la oportunidad de poner en práctica mis conocimientos teóricos.

Finalmente, agradezco al Centro INCAR del Departamento de Oceanografía UdeC por abrirme sus puertas y facilitar el material de laboratorio y financiamiento para la realización de esta tesis, al igual que todos los miembros del equipo de investigación del Laboratorio de Biotecnología y Genómica Acuática, los cuales me han acogido y aconsejado, ayudándome a desarrollar y potenciar mi desempeño en el uso de herramientas moleculares, sobretodo a mis mentoras de laboratorio Bárbara Benavente y Constanza Sáez, así como también a mi director de tesis Diego Valenzuela.

Este trabajo fue realizado en el Laboratorio de Biotecnología y Genómica Acuática (LBGA) perteneciente al Centro de Biotecnología de la Universidad de Concepción (UdeC).

## TABLA DE CONTENIDOS

TABLA DE CONTENIDOS .....	3
ÍNDICE DE TABLAS .....	5
ÍNDICE DE FIGURAS .....	6
<b>1. RESUMEN.....</b>	<b>7</b>
1.1 ABSTRACT .....	8
<b>2. INTRODUCCIÓN.....</b>	<b>9</b>
2.1 PARASITISMO EN AMBIENTES ACUÁTICOS Y SU RELACIÓN CON LA ACUICULTURA .....	9
2.2 TRATAMIENTOS FARMACOLÓGICOS CONTRA <i>Caligus</i> .....	10
2.3 MECANISMOS DE RESISTENCIA FARMACOLÓGICA .....	11
2.4 RNAs NO CODIFICANTES: ACTORES CLAVE EN LA REGULACIÓN DEL TRANSCRIPTOMA .....	14
2.5 TECNOLOGÍAS DE SECUENCIACIÓN EN EL ESTUDIO DE TRANSCRIPTOMAS .....	16
2.6 HIPÓTESIS .....	20
2.7 OBJETIVOS .....	20
<b>3. MATERIALES Y MÉTODOS .....</b>	<b>21</b>
3.1 EXTRACCIÓN DE RNA, OBTENCIÓN DE BASES DE DATOS Y SECUENCIACIÓN .....	21
3.1.1 OBTENCIÓN BASE DE DATOS NANOPORE .....	21
3.1.2 OBTENCIÓN BASES DE DATOS DE FÁRMACOS .....	21
3.2 CARACTERIZACIÓN <i>IN SILICO</i> DE lncRNAs .....	23
3.2.1 FILTRADO DE lncRNAs .....	23
3.2.2 MAPEADO Y CLUSTERIZACIÓN .....	23
3.2.3 COMPARACIÓN ILLUMINA - NANOPORE.....	25
3.2.4 ANÁLISIS DE EXPRESIÓN DIFERENCIAL .....	25
3.3 ESTIMACIÓN FUNCIONALIDAD DE lncRNAs DIFERENCIALMENTE REGULADOS.....	26
3.3.1 CO-LOCALIZACIÓN Y POSICIÓN DENTRO DEL GENOMA .....	26
3.3.2 ANÁLISIS DE ENRIQUECIMIENTO GO.....	27
3.4 VALIDACIÓN MOLECULAR .....	27
3.4.1 DISEÑO Y ESTANDARIZACIÓN DE PARTIDORES .....	27
3.4.2 SÍNTESIS DE cDNA Y qPCR .....	28

3.4.3 ESTIMACIÓN DE EXPRESION RELATIVA MEDIANTE DELTA CT COMPARATIVO .....	28
<b>4. RESULTADOS .....</b>	<b>29</b>
4.1 CARACTERIZACIÓN DE lncRNAs Y COMPARACIÓN DE TECNOLOGÍAS NGS .....	29
4.1.1 IDENTIFICACIÓN DE lncRNAs .....	29
4.1.2 COMPARACIÓN ILLUMINA – NANOPORE .....	30
4.2 CARACTERIZACIÓN DE LA MODULACIÓN TRNASCRIPTOMICA DE lncRNAs EN POBLACIONES DE <i>C. rogercresseyi</i> EXPUESTAS A FÁRMACOS .....	32
4.2.1 RNA-SEQ Y EXPRESIÓN DIFERENCIAL .....	32
4.2.2 EXPRESIÓN DIFERENCIAL DE lncRNAs Y DISEÑO DE PARTIDORES PARA GENES CANDIDATOS .....	34
4.2.3 EXPRESIÓN DIFERENCIAL DE mRNAs .....	37
4.3 ASIGNACIÓN DE FUNCIONES BIOLÓGICAS A lncRNAs EXPRESADOS DIFERENCIALMENTE .....	39
4.3.1 GENE ONTOLOGY (GO) ANÁLISIS Y CO-LOCALIZACIÓN .....	39
4.4 VALIDACIÓN MOLECULAR DE LA EXPRESIÓN DE lncRNAs .....	41
4.4.1 ESTANDARIZACIÓN DE PARTIDORES Y VALIDACIÓN POR qPCR .....	41
4.4.2 CÁLCULO DE EXPRESIÓN RELATIVA PARA GENES CANDIDATOS .....	42
<b>5. DISCUSIÓN .....</b>	<b>44</b>
<b>6. CONCLUSIÓN .....</b>	<b>57</b>
<b>7. GLOSARIO .....</b>	<b>59</b>
<b>8. BIBLIOGRAFÍA .....</b>	<b>60</b>

## ÍNDICE DE TABLAS

<b>TABLA 1.</b> Temperaturas de melting ( $T_m$ ) óptimas para los partidores diseñados .....	41
<b>TABLA 2.</b> LncRNAs y sus genes cercanos identificados en el genoma anotado de <i>Caligus</i> .....	43

## ÍNDICE DE FIGURAS

<b>FIGURA 1.</b> Mecanismos moleculares por los que pueden actuar los lncRNAs .....	19
<b>FIGURA 2.</b> Pipeline para la identificación de lncRNAs en reads obtenidos desde secuenciación nanopore .....	24
<b>FIGURA 3.</b> Diagrama de venn realizado con las bases de datos de lncRNAs illumina-nanopore para <i>C. rogercresseyi</i> .....	30
<b>FIGURA 4.</b> Gráficos comparativos de RNAs no codificantes largos (lncRNAs) obtenidos por ambos métodos de secuenciación (illumina/nanopore) y RNAs codificantes (mRNAs) obtenidos por secuenciación illumina .....	31
<b>FIGURA 5.</b> Heat maps de expresión para lncRNAs en cepas de <i>Caligus rogercresseyi</i> resistentes y susceptibles a fármacos (Azametifos, deltametrina y cipermetrina) .....	33
<b>FIGURA 6.</b> Diagramas de venn con sus respectivos heat maps de las secuencias de lncRNAs expresados diferencialmente entre cepas de <i>Caligus rogercresseyi</i> resistentes y susceptibles a fármacos .....	35
<b>FIGURA 7.</b> Diagrama de venn realizado para lncRNAs expresados diferencialmente entre cepas de <i>caligus</i> expuestas a fármacos y respectivo heat map de genes expresados diferencialmente entre los tres fármacos .....	36
<b>FIGURA 8.</b> Heat maps de expresión para mRNAs en cepas de <i>Caligus rogercresseyi</i> resistentes y susceptibles a fármacos .....	37
<b>FIGURA 9.</b> Diagramas de venn de las secuencias de mRNAs expresados diferencialmente entre ambas cepas de <i>Caligus rogercresseyi</i> .....	38
<b>FIGURA 10.</b> Plots de Gene Ontology (GO) para RNAs no codificantes largos (lncRNAs) por cada fármaco evaluado en ambas poblaciones de <i>Caligus rogercresseyi</i> .....	40
<b>FIGURA 11.</b> Gráficos de expresión relativa estimada para evaluar la expresión de los genes candidatos en cada pool de muestras .....	43

## 1. Resumen

El piojo de mar *Caligus rogercresseyi* es un copépodo ectoparásito que afecta principalmente a especies de salmónidos. En la actualidad, los tratamientos farmacológicos corresponden a su principal mecanismo de control. Sin embargo, esto genera grandes pérdidas económicas para la industria, impactos ambientales, y desarrollo de resistencia a fármacos. En este contexto, conocer las respuestas moleculares del piojo de mar ante la exposición a estos fármacos puede contribuir al desarrollo de nuevas estrategias de control para *Caligus*. Estudios previos han vinculado estas respuestas con RNAs no codificantes como los long non-coding RNAs (lncRNAs), siendo moléculas clave en la regulación del transcriptoma al intervenir en diversos procesos biológicos como la respuesta inmune, reproducción, renovación celular y desarrollo, entre otros. El objetivo de este trabajo fue caracterizar lncRNAs utilizando tecnologías de secuenciación de reads cortos (Illumina) y largos (Nanopore), y evaluar su regulación transcripcional en cepas de *C. rogercresseyi* susceptibles y resistentes a la exposición de Azametifos, Deltametrina y Cipermetrina. Se caracterizaron nuevos lncRNAs con ambas tecnologías de secuenciación, además, se evidenciaron patrones de expresión específicos asociados a cada condición evaluada. El análisis de co-localización sugiere una asociación entre lncRNAs y funciones como regulación transcripcional, reparación de tejidos, respuesta toxicológica, entre otros. Se encontraron lncRNAs diferencialmente expresados en cada uno de los tratamientos farmacológicos, los cuales podrían ser testeados como nuevos marcadores de resistencia farmacológica. Los resultados además sugieren que los estudios previos de caracterización de lncRNAs mediante secuenciación illumina, podrían estar subestimando la cantidad de lncRNAs expresadas por el genoma. Debido a esto, el estudio del transcriptoma podría beneficiarse con la incorporación de tecnologías de secuenciación de última generación como nanopore, ampliando la identificación de RNAs codificantes/no codificantes, así como también su asociación con la respuesta a fármacos en *Caligus*. La información generada en este trabajo puede servir para el desarrollo de nuevos marcadores de resistencia basados en lncRNAs, así como también ayudar a comprender los mecanismos moleculares que pueden contribuir a la resistencia farmacológica en *Caligus rogercresseyi*.

## 1.1. Abstract

The sea lice *Caligus rogercresseyi* is an ectoparasitic copepod that mainly affects salmonid species. Currently, pharmacological treatments correspond to its main control mechanism. However, this generates economic losses for the industry, environmental impacts and the development of drug resistance. In this context, knowing the molecular responses of sea lice to exposure to these drugs can contribute to the development of new control strategies for *Caligus*. Previous studies have linked these responses with non-coding RNAs such as long non-coding RNAs (lncRNAs), which are key molecules in the regulation of the transcriptome by intervening in various biological processes such as the immune response, reproduction, cell renewal and development, among others. The aims of this work was to characterize lncRNAs using short-read (Illumina) and long-read (Nanopore) sequencing technologies, and evaluate their transcriptional regulation in *C. rogercresseyi* strains susceptible and resistant to exposure Azamethiphos, Deltamethrin and Cypermethrin. New lncRNAs were characterized with both sequencing technologies, in addition, specific expression patterns associated with each evaluated condition were evidenced. The co-localization analysis suggests an association between lncRNAs and functions such as transcriptional regulation, tissue repair and toxicological response, among others. Differentially expressed lncRNAs were found in each of the pharmacological treatments, which could be tested as new markers of pharmacological resistance. The results also suggest that previous studies characterizing lncRNAs by illumina sequencing could be underestimating the amount of lncRNAs expressed by the genome. Due to this, the study of the transcriptome could improve with the incorporation of state-of-the-art sequencing technologies such as nanopore, expanding the identification of coding/non-coding RNAs, as well as their association with drug response in *Caligus*. The information generated in this work can contribute for the development of new resistance markers based on lncRNAs, as well as help to understand the molecular mechanisms that can contribute to drug resistance in *Caligus rogercresseyi*.

## 2. Introducción:

### 2.1. Parasitismo en ambientes acuáticos y su relación con la acuicultura.

El parasitismo en ambientes acuáticos cumple una importante función al regular las poblaciones de organismos, contribuyendo a mantener el equilibrio en los ecosistemas. Un ejemplo es el *Caligus rogercresseyi*, copépodo ectoparásito (parasita externamente) endémico de Chile. Éste se adhiere a la piel de los peces que parasita, generando heridas cutáneas que incrementan el estrés, debilitando el sistema inmune del hospedador y aumentando la probabilidad de contraer infecciones secundarias por virus o bacterias (Costello, 2004). Si bien las especies de peces hospedadoras de *Caligus* son variadas, las principales corresponden a aquellos de la familia Salmonidae, única familia del orden Salmoniformes. Entre ellos se encuentra la trucha arcoíris, el salmón del Atlántico y salmón coho, siendo las especies de cultivo más recurrentes en la acuicultura mundial. *Caligus rogercresseyi* tiene un amplio potencial reproductivo, pero el número de individuos está limitado al tamaño de los peces hospederos y a su densidad poblacional, por lo que, naturalmente, una baja densidad de hospedadores restringe la dispersión y el tamaño poblacional de este parásito. Sin embargo, el cultivo de salmónidos ha entregado las condiciones perfectas para favorecer su proliferación (Aaen, Helgesen, Bakke, Kaur, & Horsberg, 2015), debido principalmente al alto volumen de peces que se tiene en las jaulas de cultivo, lo cual es un factor de estrés adicional en los peces que contribuye a desequilibrar los procesos biológicos en el organismo.

En Chile, la salmonicultura se ha establecido como una de las principales actividades económicas del país. En este contexto, la caligidosis (enfermedad producida por el *Caligus*) se ha vuelto uno de los desafíos más importantes para la industria, generando grandes pérdidas económicas y efectos negativos en el ecosistema, incluyendo el impacto del *Caligus* sobre otros potenciales hospederos que habitan en zonas aledañas a los centros de cultivo. En Chile, los costos unitarios de producción aumentan en promedio \$1.4 US/Kg, pudiendo

variar según las condiciones ambientales y espaciales del cultivo (Dresdner et al., 2019; Núñez-Acuña et al., 2020). Estas pérdidas están explicadas principalmente por los costos de tratamiento, el aumento de la predisposición a infecciones secundarias, el alargamiento del ciclo de cultivo producto de la disminución de la eficiencia de conversión alimenticia, y la limpieza de carcasas durante el procesamiento (Carvajal, Gonzalez, & George-Nascimento, 1998; Costello, 2006; Johnson, Treasurer, Bravo, Nagasawa, & Kabata, 2004). Esto cobra gran relevancia en un escenario donde la producción de acuicultura está superando a la pesca, mientras que las concesiones que se entregan a lo largo las costas chilenas cada vez son mayores, y la carga parasitaria por *Caligus* aumenta en todos los centros de cultivo, por ejemplo, solo entre 2004-2007 la carga parasitaria promedio por pez aumento de 5 a 20 piojos en el caso de la trucha arcoíris y de 10 a 34 en el salmón del atlántico (Rozas & Asencio, 2007).

## 2.2. Tratamientos farmacológicos contra *Caligus*.

Diferentes medidas se han tomado para controlar la caligidosis, siendo la más común el uso de tratamientos farmacológicos sobre los cultivos de peces. Entre ellos se encuentran piretroides como cipermetrina y deltametrina, que afectan los canales iónicos de sodio-potasio en la sinapsis neuronal provocando parálisis muscular (Sevatdal, Copley, Wallace, Jackson, & Horsberg, 2005), u organofosforados como azametifos que actúa inhibiendo la acetilcolinesterasa (AChE), lo que altera el metabolismo de neurotransmisores presentes en la unión neuromuscular y en sitios del sistema nervioso central (Jones, Sommerville, & Wootten, 1992). Sin embargo, su eficacia y fácil aplicación han llevado al uso intensivo de estos compuestos. En este contexto, se ha visto que a mayor uso de fármacos se reduce la eficacia promedio de los tratamientos, lo que hace pensar que emerge una resistencia farmacológica por parte de estos organismos, la cual ha aumentado considerablemente en menos de una década (Helgesen, Bravo, Sevatdal, Mendoza, & Horsberg, 2014; Núñez-Acuña et al., 2020). Ante

este tipo de problemática, diversos estudios han sido destinados a comprender las respuestas moleculares del parásito frente a la exposición farmacológica. El sistema profenoloxidasa (proPO) es esencial para inducir la melanización, teniendo un importante rol en la inmunidad innata de los crustáceos, principalmente en la cicatrización de heridas y defensa contra patógenos, siendo clave en la respuesta farmacológica. Se ha documentado que hay patrones de expresión génica similares de factor activador de profenoloxidasa I y II, homólogos de la serina proteasa I, en respuesta a cipermetrina y deltametrina. Fenoloxidasas y transportadores Na-K presentan un patrón similar en respuesta a cipermetrina y azametifos, habiendo una sobreexpresión en los individuos expuestos. Gran cantidad de proteínas están ligadas a la acción de fármacos particulares, por ejemplo, la carboxilesterasa es sobreexpresada en *Caligus* expuestos a deltametrina, pero suprimida tanto para la acción de cipermetrina y azametifos, mientras que P-gp es sobreexpresada en respuesta a estos últimos y suprimida para azametifos. De forma similar, el receptor neuronal de acetilcolina (nACh) es suprimido por la acción de azametifos, habiendo una sobreexpresión de tropomiosina. También, glutatión-s-transferasa es sobreexpresado frente a los tres fármacos, mientras que la quitinasa es suprimida. La acetilcolinesterasa (AChE) es suprimida frente a azametifos y cipermetrina, alterando la acción de neurotransmisores (Vera Bizama, 2017). Además de la expresión diferencial de proteínas asociadas a la acción farmacológica, esta expresión también puede estar fuertemente ligada a la población donde ejerce acción el fármaco, pudiéndose evidenciar una notoria diferencia.

### 2.3. Mecanismos de resistencia farmacológica.

La reducción en la biodisponibilidad del fármaco es uno de los mecanismos de resistencia más frecuentes, ya sea aumentando el metabolismo intracelular del fármaco o expulsándolo desde las células diana en el momento de su interacción con la membrana plasmática o después de su metabolización (Shepherd &

Carney, 2000). En la célula, el metabolismo de xenobióticos consta de cuatro fases: Absorción, Modificación, Conjugación y Expulsión (Gueguen et al., 2006), habiendo una serie de transformaciones bioquímicas en cada una de estas fases. La familia de proteínas de unión a ATP, conocidas como proteínas ATP-Binding Cassette (ABC), utilizan ATP para el transporte de moléculas a través de la membrana plasmática, teniendo varios dominios transmembrana (TMD) y de unión a nucleótidos (NBD), los cuales se unen e hidrolizan el ATP para producir la energía necesaria para la expulsión de compuestos tóxicos como lo son los fármacos. Esta familia, además, está compuesta de subfamilias diferenciadas en sus secuencias y dominios estructurales, encontrándose por todo el reino animal y vegetal. En artrópodos, han sido caracterizadas en especies como *Drosophila melanogaster*, *Anopheles gambiae*, *Apis mellifera*, *Bombix mori* y *Daphnia pulex* (Sturm, Cunningham, & Dean, 2009).

P-gp es una glicoproteína presente en la membrana plasmática, producida en humanos por el gen *mdr1* (Borst, Evers, Kool, & Wijnholds, 1999), es responsable del transporte de diferentes compuestos hacia el exterior de la célula, actuando como una bomba de eflujo dependiente de ATP, que lleva a la disminución intracelular de moléculas como drogas, esteroides, xenobióticos y péptidos (Gottesman & Pastan, 1993). La hidrólisis del ATP provee la energía necesaria para la expulsión de los fármacos. Por otro lado, MRP1 también es una glicoproteína con funciones asociadas a la desintoxicación de compuestos. Codificada por el gen ABCC1, se expresa en la mayoría de los tejidos, principalmente en pulmones, testículos, riñones, hígado, músculo esquelético y células sanguíneas (Cole et al., 1992; Flens et al., 1996; Stride et al., 1996). Su principal localización es la membrana plasmática, teniendo preferencia por el transporte de compuestos aniónicos como glutatión (GSH) y sus conjugados mediante co-transporte. Su sobreexpresión estaría asociada con niveles bajos y altos de GSH intracelular y extracelular respectivamente (Zaman et al., 1995). Debido a esto, diversos autores proponen que la expulsión del fármaco fuera de

la célula está condicionada por la acción de proteínas ABC transportadoras que actúan como bombas de expulsión de compuestos, entre ellas P-glycoprotein (P-gp) y Proteínas de resistencia múltiple a fármacos (MRP) (Verrelle et al., 1991), siendo relevantes como mecanismos de defensa bioquímica contra xenobióticos, así como targets de toxicidad. Estudios realizados desde bioensayos de *C. rogercresseyi* expuestos a Deltrametrina y Azametifos, reportan expresión diferencial de mensajeros para proteínas ABC, entre las que destacan ABCB (P-gp) y ABCC (MRP1) (Valenzuela-Muñoz, Sturm, & Gallardo-Escárate, 2015). La eliminación de xenobióticos mediante transportadores, no solo reduce la biodisponibilidad y la acción terapéutica, sino que también puede implicar procesos de protección celular contra compuestos tóxicos y conferir resistencia multixenobiotica (Buss & Callaghan, 2008). Este proceso se explica por la adaptación evolutiva, observable a corto plazo en organismos con ciclos de vida cortos como *Caligus*, donde la resistencia incrementa probablemente por mutación. En general, los genes de resistencia son poco frecuentes en poblaciones de organismos, sin embargo, frente a estímulos (por ejemplo, fármacos) que gatillen la transcripción de estos genes, los individuos que los posean se verán favorecidos y los transmitirán a su descendencia, generando poblaciones cada vez más resistentes (Dermauw & Van Leeuwen, 2014; LEÓN, 2016). En la actualidad, la creciente profundización de estudios transcriptómicos, está comenzando a revelar que las respuestas moleculares frente a la exposición de fármacos, no solo estarían ligadas a la expresión de genes codificantes, sino que, además, tendrían un componente importante asociado a la respuesta de RNAs no codificantes. En este contexto, La biotecnología, guiada por la mano de la biología molecular, surge como una nueva carta a jugar en la lucha contra el uso de agentes químicos.

#### 2.4. RNAs no codificantes: actores clave en la regulación del transcriptoma.

En sus inicios, el dogma central de la biología molecular establecía que la información genética contenida en el DNA era transcrita (leída) a RNA, y que luego, dicha información era “traducida” para sintetizar proteínas como producto final. Como todo proceso y función dentro de la célula esta mediado por proteínas, las hace fundamentales para la vida (Trovero & Geisinger, 2019), lo cual lógicamente llevó a pensar que la mayor parte del genoma se expresa como RNA codificante de proteínas (mRNA). Sin embargo, con el paso de los años, el desarrollo de tecnologías de secuenciación masiva (NGS) junto a nuevas técnicas y herramientas moleculares, revelaron que la porción del transcriptoma (expresión del genoma) que no codifica proteínas es mucho mayor de lo que se pensaba (Collins, Lander, Rogers, Waterston, & Conso, 2004). En humanos, por ejemplo, un 98% del transcriptoma corresponde a RNA no codificante (ncRNA) (Mattick, 2011), por lo que se espera que sean muy variados. Entre ellos se pueden encontrar los tRNA y rRNA, fundamentales para síntesis proteica, los snRNA que actúan en el corte y empalme de intrones, piRNA que interactúan con proteínas piwi de la familia argonauta con un importante rol en el silenciamiento de transposones (Mercer, Dinger, & Mattick, 2009). Los micro RNAs (miRNAs) cumplen funciones fundamentales regulando la traducción de proteínas, estos pueden actuar uniéndose al mRNA para cortar su cola poly A característica, que le permite resistir a la degradación citoplasmática en su viaje desde el núcleo celular hasta el ribosoma, por lo que la acción del miRNA lo desestabiliza para permitir su degradación y así reprimir la traducción (Laouressergues et al., 2015). Por otro lado, los long non-coding RNAs (lncRNAs) son un tipo de RNA no codificante caracterizado por tener un potencial de codificación nulo o mínimo, longitud mayor a 200 nt, bajo contenido del % G-C, así como también niveles de expresión más bajos que los evidenciados en transcritos codificantes, (Valenzuela-Miranda, Etebari, Asgari, & Gallardo-Escárate, 2017), pudiendo

regular la expresión de genes mediante una amplia variedad de mecanismos (fig. 1), entre ellos modificación de cromatina, ensamble de complejos proteicos, y la inhibición o activación de miRNAs o proteínas, pudiendo también actuar como precursores de moléculas más pequeñas como miRNAs (Rogoyski, Pueyo, Couso, & Newbury, 2017; Trovero & Geisinger, 2019) y péptidos funcionales (Laouressergues et al., 2015; Li, Leng, Fan, Pan, & Ye, 2017; Magny et al., 2013). Según la review de Schmitz et al. (2016) y trabajos realizados por Mercer et al. (2009) los lncRNAs están involucrados en múltiples procesos de regulación molecular a nivel de transcripción, post-transcripción, y actividad proteica, como, por ejemplo, modificación de histonas, metilación del DNA, compactación de la cromatina, organización del genoma, e interacción con factores de transcripción. Una de las características más relevantes de los lncRNAs es su patrón de expresión altamente específico, ya sea asociado a un tipo de célula, proceso, función, tejido o etapa del desarrollo determinada, con una especificidad mucho mayor que los RNAs codificantes (Cavagnari, 2012; Deng, Liu, Nie, Weining, & Wu, 2018; Jiang et al., 2016; Trovero & Geisinger, 2019), teniendo roles fundamentales en todos los procesos biológicos, como la inmunidad, detoxificación, renovación celular, reproducción y desarrollo, entre otros, por lo que también, tienen una relación directa con el desarrollo de una amplia gama de enfermedades (Boltana, Valenzuela-Miranda, Aguilar, Mackenzie, & Gallardo-Escarate, 2016; Carpenter et al., 2013; Imamura & Akimitsu, 2014; Murillo-Maldonado & Riesgo-Escovar, 2019; Schmitz, Grote, & Herrmann, 2016; Tarifeno-Saldivia, Valenzuela-Miranda, & Gallardo-Escarate, 2017; Taylor, Chu, Spektor, & Soloway, 2015; Valenzuela-Munoz, Valenzuela-Miranda, & Gallardo-Escarate, 2018; Zhao & Lin, 2015).

En este contexto, gran parte de los trabajos realizados en la actualidad, se centran en estudiar el genoma de *Caligus* y su respuesta transcriptómica ante la exposición farmacológica, revelando genes y patrones de expresión altamente específicos, donde moléculas como lncRNAs y miRNAs son los principales

actores involucrados (Valenzuela-Miranda, Etebari, Asgari, & Gallardo-Escárte, 2017; Valenzuela-Munoz, Valenzuela-Miranda, & Gallardo-Escarate, 2018; Núñez-Acuña et al., 2020; Tarifeno-Saldivia, Valenzuela-Miranda, & Gallardo-Escarate, 2017). Algunos trabajos realizados en este ámbito evidencian una expresión diferencial de lncRNA en *Caligus* ante la respuesta a fármacos como azametifos y deltametrina (Valenzuela-Miranda et al., 2017), estando además estrechamente relacionados con la expresión diferencial de genes codificantes, pudiendo evidenciarse un patrón de co-expresión específico para cada fármaco o población que haya sido expuesta, siendo de suma relevancia la interacción entre RNAs codificantes y no codificantes (Núñez-Acuña et al., 2020). Se ha descrito que al menos 163 miRNAs maduros son expresados diferencialmente en respuesta a fármacos utilizados para el control de *Caligus* (Azametifos, Deltametrina y Cipermetrina), encontrándose sitios de unión para miRNAs en genes ADME, asociados a la absorción de fármacos, distribución, metabolismo y excreción. Además, hay evidencia de interacción entre miRNAs y lncRNAs, teniendo una correlación con la modulación RNA codificante/ no codificante, sugiriendo nuevos mecanismos moleculares implicados en la regulación de genes (Núñez-Acuña, Valenzuela-Muñoz, Valenzuela-Miranda, & Gallardo-Escárte, 2021), los cuales estarían asociados a la resistencia farmacológica en piojos de mar.

#### 2.5. Tecnologías de secuenciación en el estudio de transcriptomas:

Con el pasar de los años, el desarrollo de las tecnologías de secuenciación y herramientas moleculares han permitido realizar numerosos estudios de transcriptómica comparativa en busca de identificar genes responsables o involucrados en el desarrollo de una característica específica. Estos estudios son llevados a cabo mediante la síntesis de secuencias cortas (short reads) de cDNA (DNA copia o complementario) generadas a partir de RNA celular extraído desde el organismo de interés, donde generalmente se compara la expresión genética

de un individuo o grupo de individuos, bajo diferentes tratamientos (Trapnell et al., 2012).

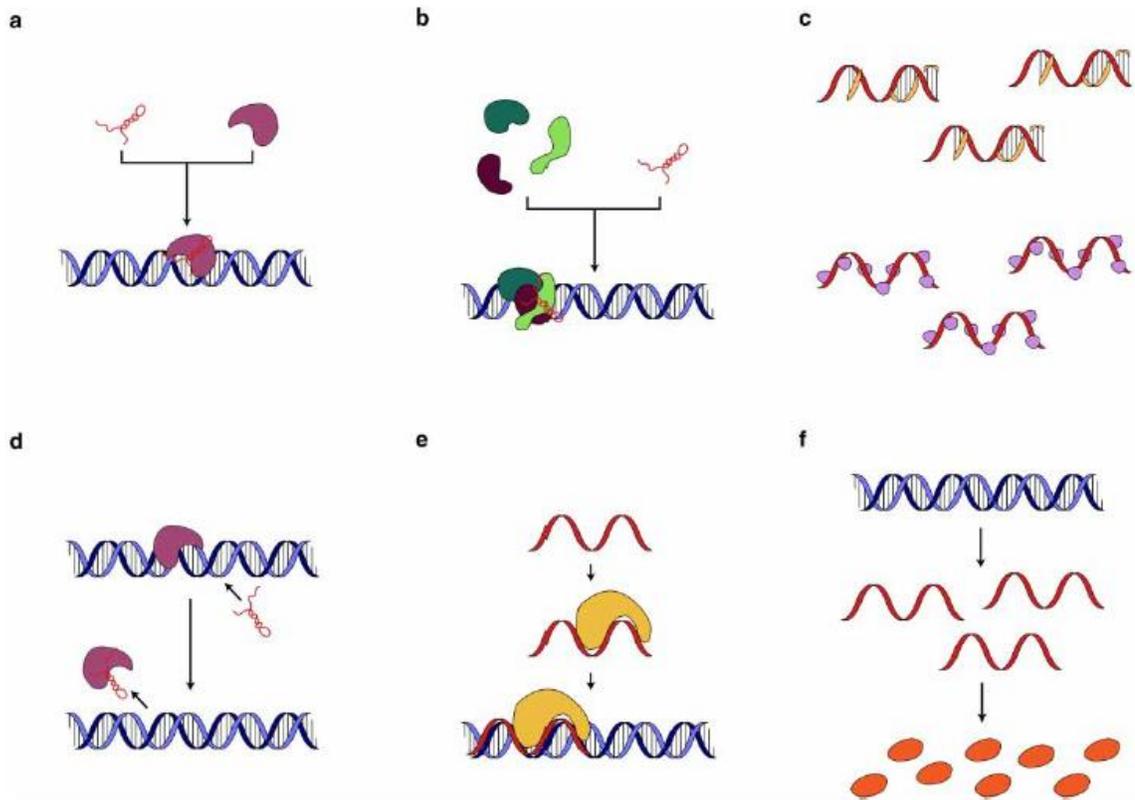
Las tecnologías de Next Generation Sequencing (NGS) han permitido profundizar el conocimiento respecto de la interacción de las moléculas de RNA no codificante dentro del genoma de los organismos, pero como todo método estas tienen sus ventajas y desventajas. La secuenciación Illumina es una tecnología de NGS de segunda generación, siendo una de las más utilizadas, esta fragmenta el DNA en múltiples partes, entregando gran cantidad de lecturas (reads) con longitudes entre los 100-150 pb (short reads), las cuales se ligan a adaptadores genéricos y se anillan a un portaobjetos para amplificar cada fragmento mediante PCR, creando un spot de muchas copias de un mismo fragmento. Luego, son separados en una única hebra y sometidos a secuenciación. El cartridge de secuenciación contiene nucleótidos marcados por fluorescencia, DNA polimerasa, y un nucleótido terminador, gracias al cual se agrega una base a la vez. En cada ciclo el terminador es removido y permite agregar la siguiente base en el sitio, además, debido a las señales fluorescentes el ordenador detecta la base añadida en cada ciclo, obteniendo una secuencia de nucleótidos. Finalmente, es frecuente realizar un ensamble *de novo* con los reads obtenidos desde la secuenciación para poder armar e identificar las secuencias a las que corresponden (Kircher, Heyn, & Kelso, 2011).

Por otro lado, Nanopore es una tecnología NGS de tercera generación y de secuenciación de molécula única, el cual es capaz de entregar reads mayores a 2000 pb (Rollano Peñaloza & Mollinedo Portugal, 2017; Tosco Herrera, 2019), lo cual elimina la necesidad del ensamble, ya que el DNA es guiado a través de un nanoporo proteico, el cual mide la diferencia de voltaje entre cada base nitrogenada (A, G, C, T), entregando la secuencia completa del gen. La mayoría de los equipos, como los manufacturados por Oxford Nanopore technologies, utilizan Flow cells, las cuales tienen varios nanoporos incrustados en una

membrana de sílice electroresistente, donde cada uno tiene su propio electrodo que mide la corriente eléctrica que fluye a través de él (Heikema et al., 2020).

Mediante el análisis bioinformático de las secuencias, lo cual requiere un poder computacional significativo, así como también de herramientas y softwares específicos, es posible cuantificar el número de secuencias que coinciden con los genes o transcritos de una condición o individuo en particular, y a través de comparaciones identificar los genes con expresión diferencial significativa. En este tipo de análisis con frecuencia se usa un genoma o transcriptoma conocido como punto de referencia, sin embargo también es posible realizarlo sin tener información conocida de la especie que se estudia (especies no modelo) (Riascos-Arcos, Espitia Navarro, & López Gerena, 2015).

A pesar de las diferencias existentes entre ambas tecnologías de secuenciación, la caracterización de lncRNAs en *Caligus* se ha desarrollado utilizando tecnología illumina. De este modo, se desconoce como el uso de tecnologías de secuenciación de reads largos (Nanopore) puede favorecer el descubrimiento de nuevos lncRNAs que podrían estar involucrados en la resistencia farmacológica de *Caligus*. Debido a esto, el objetivo de este trabajo fue identificar y caracterizar el repertorio de lncRNAs mediante secuenciación nanopore en cepas resistentes y susceptibles de *Caligus rogercresseyi* expuestas a los principales fármacos utilizados para tratar la caligidosis en la industria salmonera: Azametifos, deltametrina y cipermetrina, evaluando también los patrones de expresión de lncRNAs asociados a las distintas condiciones, y determinando también su co-localización con genes funcionales dentro del genoma de *C. rogercresseyi*. De manera general, se comparan dos tecnologías de secuenciación distintas utilizadas en el estudio del transcriptoma: Nanopore & Illumina, con la finalidad de comparar y complementar los puntos fuertes y débiles de cada método.



**Figura 1.** Mecanismos moleculares por los cuales pueden funcionar los lncRNAs. a) Algunos como *Xist* y *roX1* pueden modular la expresión genética uniéndose a modificadores de transcripción o cromatina. b) Otros como *HOTAIR* pueden actuar como scaffolds moleculares, formando complejos proteicos con roles regulatorios. Ej: Polycomb, Trithorax. c) Señuelos moleculares secuestrando miRNAs o proteínas. d) Atenuar o remover factores de transcripción, proteínas o RNAs de su localización funcional. e) pueden actuar como guías moleculares, permitiendo la formación de complejos en sitios específicos. f) Algunos pueden ser activamente traducidos a péptidos funcionales (smORFs), como los péptidos SclA y SclB (Rogoyski et al., 2017).

## 2.6. Hipótesis:

- La utilización de tecnologías de secuenciación de reads largos (Nanopore) puede ser usada para el descubrimiento de nuevos lncRNAs en *Caligus rogercresseyi* antes inexplorados por tecnología de reads cortos
- La expresión de lncRNAs en *Caligus rogercresseyi* muestra patrones específicos asociados a la exposición farmacológica en poblaciones resistentes y susceptibles a fármacos.

## 2.7. Objetivos:

General:

- Caracterizar lncRNAs mediante secuenciación Nanopore y evaluar su expresión transcriptómica en poblaciones resistentes y susceptibles (R/S) a fármacos utilizados para el control de caligidosis.

Específicos:

- Comparar la caracterización de lncRNAs utilizando tecnologías de secuenciación de reads largos (Nanopore) y reads cortos (Illumina).
- Caracterizar la modulación transcriptómica de lncRNAs en poblaciones de *C. rogercresseyi* R/S, expuestas y no expuestas a azametifos, deltametrina y cipermetrina.
- Asignar posibles funciones biológicas a los lncRNAs diferencialmente modulados en poblaciones R/S.
- Validar la expresión de lncRNAs como posibles nuevos marcadores de resistencia farmacológica en *Caligus*.

### 3. Materiales y métodos:

#### 3.1. Extracción de RNA, obtención de bases de datos y secuenciación:

##### 3.1.1. Obtención base de datos nanopore:

Se extrajo tejido de *Caligus rogercresseyi* obtenido desde un cultivo de *Salmo salar* (salmón del atlántico) realizado en Dichato (estación de biología marina, UdeC), realizando un macerado del organismo desde su estado infeccioso de copepodito y en ambos géneros del adulto (hembra y macho), siendo fijado y almacenado a -80°C para su posterior extracción de RNA total por método de TRIzol (Valenzuela-Muñoz et al., 2018). Para medir la pureza y concentración del RNA extraído se utilizó un espectrofotómetro Nanodrop ND-1000 (ThermoFisher Scientific, Waltham, MA, USA), y su integridad fue determinada mediante electroforesis en gel de agarosa al 1.2%.

La data del transcriptoma se obtuvo mediante secuenciación nanopore en la plataforma MINION, desde el RNA total aislado. Posteriormente, se construyeron librerías de secuenciación para cada muestra usando el cDNA-PCR Barcoding Kit® de NT Oxford nanotechnology (ONT, UK).

##### 3.1.2. Obtención bases de datos de fármacos:

Desde poblaciones previamente identificadas como resistentes y susceptibles a Azametifos, Deltametrina y Cipermetrina, se realizó un bioensayo de *Caligus rogercresseyi* expuestos, eligiendo para secuenciar aquellos expuestos a concentraciones de 100, 3 y 5 ppb, respectivamente. Por cada fármaco se expusieron 2 poblaciones de piojos: una susceptible y otra resistente, obteniéndose un total de 12 muestras, considerando los respectivos grupos control. Posterior a la exposición, las muestras fueron fijadas en RNAlater y almacenadas a -80°C para su posterior procesamiento. Para la extracción de RNA por cada pool de muestras se preparó 1ml de TRIzol en un tubo eppendorf de 2 ml con una esfera de cerámica autoclavada. Cada pool fue realizado con

muestras de 2 machos y 2 hembras de *C. rogercresseyi* expuestos al fármaco correspondiente, siendo cortadas/maceradas con un bisturí en una placa Petri, agregando el tejido en los tubos preparados con TRIzol, homogeneizando en un TissueLyser a una frecuencia de 24/s por 5 min. A continuación, se agregaron 200 µl de cloroformo para posteriormente incubar en hielo por 3 min. Luego de la incubación, las muestras fueron puestas en centrifuga a 12.000g por 15 min a 4 °C, pudiéndose luego observar la fase acuosa en el tubo, característica del método TRIzol, en donde quedan retenidos los ácidos nucleicos para su posterior recuperación, traspasando 500 µl de RNA a tubos eppendorf de 1.5 ml. Se agregó un volumen de isopropanol similar al de la muestra (500 µl) mezclando mediante inversión e incubando por 20 min a -20 °C. Luego de la incubación, se hizo un centrifugado a 12.000 g durante 10 min a 4 °C, resultando en la visualización de un pellet blanco, el cual corresponde al RNA extraído. El sobrenadante fue removido para luego lavar el pellet con 1 ml de etanol al 75%, centrifugando a 7.500g por 5 min a 4 °C, repitiendo este paso 2 veces. Hecho esto, se descartó el sobrenadante, se eliminó el excedente con micro pipeta, y se resuspendió el pellet de las muestras en 50 µl de agua para biología molecular (libre de nucleasas), midiendo la concentración y pureza mediante un espectrofotómetro Nanodrop ND-1000 (ThermoFisher Scientific, Waltham, MA, USA) y determinando la integridad mediante electroforesis en gel MOPS al 1.2%. Las muestras que cumplieron con los criterios de integridad y pureza fueron utilizadas para los análisis siguientes. Mediante secuenciación en la plataforma Mi-seq (illumina) se construyeron librerías de las muestras utilizando el TruSeq RNA Sample Preparation Kit V2 (Illumina®, USA), desde donde se obtuvieron sus bases de datos, las cuales fueron trimmeadas con la herramienta Trim Reads del software CLC Genomics Workbench 20, para remover los adaptadores y continuar con análisis posteriores.

### 3.2. Caracterización *in silico* de lncRNAs:

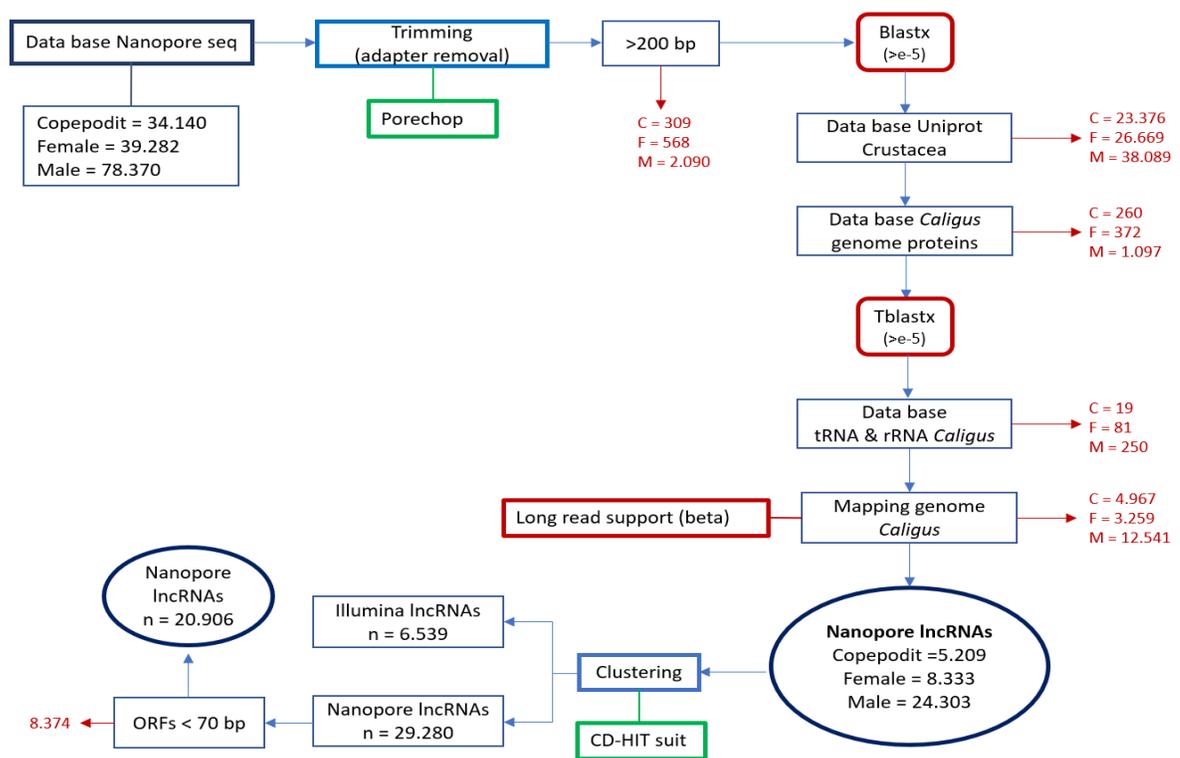
#### 3.2.1. Filtrado de lncRNAs:

Las lecturas obtenidas por nanopore se agruparon en 3 conjuntos de datos, uno de copepodito y 2 de adultos (macho y hembra), desarrollando posteriormente un pipeline (fig. 2) personalizado para la identificación de los transcritos de lncRNAs. El primer paso fue remover los adaptadores realizando un trimming con el software Porechop. Posteriormente se removieron todas las secuencias con una longitud menor a 200 bp. Hecho esto, se utilizó la herramienta Blastx para comparar las secuencias de nanopore con una base de datos de proteínas de crustáceos obtenida desde uniprot, identificando las secuencias con potencial codificante para descartarlas (e value < -4). El mismo paso fue realizado nuevamente para las secuencias no codificantes restantes, pero esta vez comparándolas con una base de datos de proteínas de *Caligus rogercresseyi* obtenida desde su genoma (NCBI, Assembly ASM1338718v1, ID 38161). Las secuencias no codificantes restantes fueron comparadas mediante TBlastx con una base de datos de tRNA y rRNA de *Caligus rogercresseyi* extraída desde su genoma, para descartar secuencias RNAs no codificantes que no corresponden a lncRNA, considerando finalmente a las secuencias restantes como lncRNAs potenciales.

#### 3.2.2. Mapeado y clusterización:

Los transcritos de lncRNAs obtenidos fueron mapeados contra el genoma anotado de *C. rogercresseyi* extraído desde la base de datos del NCBI (Assembly ASM1338718v1, ID 38161) utilizando los parámetros estándar, para identificar las secuencias que coinciden con el genoma. Las secuencias restantes de los 3 sets de datos se clusterizaron utilizando el software CD-HIT suit, con la finalidad de agrupar aquellas secuencias que comparten un grado de identidad mayor al

90% bajo una única secuencia representativa, para formar una única base de datos de nanopore, que luego fue comparada con una base de datos ya existente de lncRNAs, obtenida por secuenciación illumina (Gallardo-Escárte, Valenzuela-Muñoz, & Nuñez-Acuña, 2014), la cual también fue mapeada y clusterizada, para de esta forma comparar ambos métodos de secuenciación. Finalmente se evaluaron los ORFs de la base de datos nanopore para descartar aquellas secuencias con ORFs que codifiquen para más de 70 aminoácidos (fig. 2).



**Fig. 2:** Pipeline para la identificación de lncRNAs en reads obtenidos desde secuenciación nanopore. La mayor parte del proceso de filtrado se realizó con el software CLC genomics workbench 20, las casillas de colores rojo y verde indican las herramientas utilizadas dentro y fuera del software CLC respectivamente, los números en rojo indican las secuencias descartadas en cada paso. Adicionalmente se construyó una base de datos única de lncRNAs para *C. rogercresseyi* utilizando la data de nanopore.

### 3.2.3. Comparación Illumina-Nanopore:

Para la comparación entre ambos métodos de secuenciación se consideró el contenido GC, longitud de secuencias, energía libre mínima (MFE) y la cantidad de secuencias, esta última fue realizada mediante Blastn entre la base de datos nanopore e illumina, considerando como compartidas aquellas secuencias con e value entre 0 & -50. El contenido GC, longitud y MFE fueron calculados y graficados utilizando el software R.

### 3.2.4. Análisis de expresión diferencial:

Las bases de datos de fármacos fueron utilizadas para evaluar los patrones de expresión de los lncRNAs de la base de datos nanopore. Para ello se realizó un análisis de RNA-seq para las bases de datos de fármacos tomando la de nanopore como referencia, para luego construir por cada fármaco heat maps comparativos entre la expresión de lncRNAs del grupo control y el expuesto para ambas cepas. Utilizando el software CLC Genomics Workbench 20 (CLC bio, Denmark), por cada fármaco se realizó una comparación estadística entre los datos del control y el expuesto de cada cepa, mediante la herramienta Differential expression for RNA seq, construyendo diagramas de venn para encontrar los genes que se expresan diferencialmente (FDR-p value = 0,01 & fold change = 4) entre ambas cepas, creando expression browsers para obtener los valores de expresión de los genes, estimados como transcritos por millón (TPM). Adicionalmente se realizaron estadísticos solo entre los datos de las cepas expuestas por cada fármaco, para encontrar los genes expresados diferencialmente entre los 3 fármacos. Los transcritos de lncRNAs con valores de fold change más contrastantes fueron seleccionados desde los genes expresados diferencialmente entre cepas por cada fármaco y entre los 3 fármacos, para posteriormente diseñar sets de primers con el software geneious

pro y validar su expresión mediante la técnica de PCR a partir de la síntesis de cDNA obtenido desde RNA total de las muestras expuestas a los diferentes fármacos evaluados.

### 3.3. Estimación funcionalidad de lncRNAs diferencialmente regulados

#### 3.3.1. Co-localización y posición dentro del genoma:

Para conocer la co-localización entre lncRNAs y genes funcionales, estos fueron mapeados contra el genoma de *Caligus* para conocer su ubicación (coordenadas) y posteriormente anotarlos, extrayendo todas aquellas secuencias codificantes (CDS) en una ventana de 10.000 pb por cada flanco de los lncRNAs anotados, para su posterior análisis e identificación. De esta forma, usando las anotaciones del genoma de *C. rogercresseyi* se identificaron los genes codificantes que están solapados con regiones genómicas de lncRNA, evaluando los niveles de expresión utilizando el software CLC genomics workbench (CLC bio, Denmark). Este procedimiento se realizó para las secuencias expresadas diferencialmente en las distintas cepas expuestas a fármacos, obtenidas desde los venn diagrams que se construyeron previamente, mapeándolas y anotándolas en el genoma, para extraer todas aquellas secuencias codificantes (CDS) en una ventana de 10.000 pb por cada flanco de lncRNA anotado. La identificación de los genes solapados a secuencias de lncRNA se llevó a cabo mediante un Blastx realizado contra la base de datos Arthropoda que fue obtenida desde uniprot, obteniendo los ID de los CDS (Coding Sequences) de fármacos para realizar un análisis GO en uniprotID.

### 3.3.2. Análisis de enriquecimiento GO:

Este análisis se realizó para identificar los procesos biológicos más representados entre los genes codificantes dentro del vecindario de genes cercano a lncRNAs. Los ID de las secuencias codificantes (CDS) identificadas por el análisis de co-localización, se analizaron en la base de datos UniprotID, identificando proteínas y funciones asociadas. La información entregada se descargó y ordenó, estimando cantidad y abundancia relativa de los genes, para posteriormente graficar los resultados usando el software R.

### 3.4. Validación molecular:

#### 3.4.1. Diseño y estandarización de partidores:

Para el diseño de partidores se utilizó el software geneious pro, estableciendo los siguientes parámetros: Tm = 50-65, %GC = 50-80, Size = 19-27, amplicon = 80-120. Una vez obtenidos los pares de partidores entregados por el programa, se ingresaron sus secuencias en el software Oligo Analyzer para discriminar entre los diferentes sets de partidores y conservar aquellos con valores de Delta G > -7, con la finalidad de reducir la probabilidad de que se formen dímeros de primer al momento de realizar la PCR para validar la expresión de los lncRNAs seleccionados. Se hicieron alícuotas de los partidores llevándolos a una concentración de 10 ng/μl, para posteriormente estandarizar la temperatura mediante la técnica de PCR evaluando su amplificación a 58, 60, 62 y 64 °C, para encontrar la temperatura a la cual tienen una mejor amplificación y descartar aquellos que se dimerizan o no amplifican. Estos fueron probados en cDNA obtenidos desde pool de muestras de *Caligus rogercresseyi* expuestos a Azametiphos, Deltamethrin y Cypermethin, utilizando como control negativo agua para biología molecular. Las Tm óptimas para los partidores, fueron determinadas mediante la visualización en gel de agarosa (1,8%) del producto de PCR obtenido en las distintas temperaturas. La concentración de los partidores

fue estandarizada mediante qPCR realizando 5 diluciones seriadas de cDNA de *Caligus* expuesto a deltametrina, dejándolo a concentraciones de 400, 80, 16, 3.2 y 0.64 ng/μl respectivamente, preparando un mix de reacción por cada par de partidores. La eficiencia optima de los partidores consideró valores de entre 80-100.

#### 3.4.2. Síntesis de cDNA y qPCR:

Se realizó extracción de RNA para 12 pools de muestras compuestos cada uno por 4 individuos adultos de *Caligus rogercresseyi*, correspondientes a dos machos y dos hembras, habiendo obtenido los pools desde poblaciones de *Caligus* resistentes y susceptibles expuestas Azametifos, Deltametrina y Cipermetrina, más sus respectivos controles. Para la síntesis de cDNA se utilizó el kit RevertAid H Minus First Strand cDNA Synthesis – ThermoScientific®, más 200 ng/μl de RNA, utilizando en la reacción 5 μl de RNA y 6 de H<sub>2</sub>O por cada pool al cual se le realizo la extracción, amplificándolo mediante PCR. Una vez terminada la reacción, el cDNA fue diluido en agua de biología molecular en una razón de 1:5 para llevarlo a una concentración de 400 ng/μl. Teniendo el cDNA listo, se realizaron reacciones de qPCR en cada muestra, utilizando de los partidores diseñados para los genes de lncRNAs identificados en el análisis bioinformático, solo aquellos que obtuvieron una estandarización optima, usando en la reacción una concentración de partidores de 0,5 μM.

#### 3.4.3. Estimación de expresión relativa mediante delta Ct comparativo:

Los datos de Delta Ct de los genes evaluados, fueron obtenidos mediante qPCR y registrados en un ordenador, siendo descargados para su posterior ordenamiento en Excel. A partir de estos datos se calculó la expresión relativa de los genes en los diferentes tratamientos, utilizando Excel y posteriormente graficados mediante el software GraphPad Prism 8.

## 4. Resultados:

### 4.1. Caracterización de lncRNAs y comparación de tecnologías NGS.

#### 4.1.1. Identificación de lncRNAs:

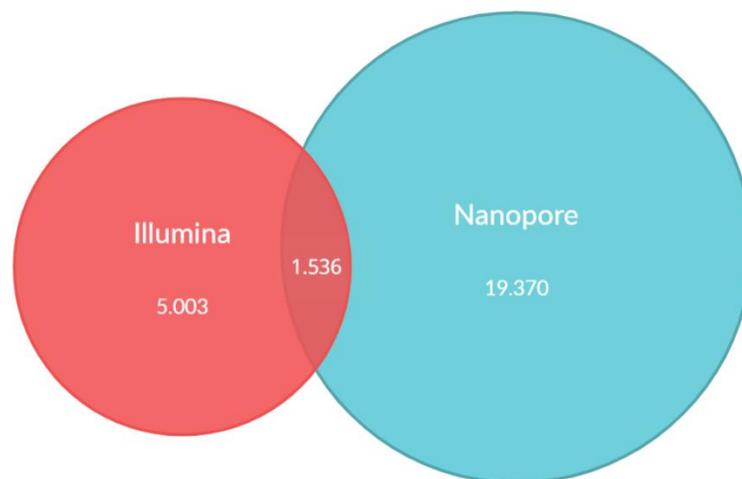
Un pipeline personalizado (Fig. 2) fue creado específicamente para la identificación de lncRNAs desde reads obtenidos por secuenciación nanopore. Un total de 34.140, 39.282 y 78.370 reads fueron obtenidos desde la secuenciación del copepodito, hembra y macho respectivamente, los cuales fueron sometidos a una serie de filtros para remover cualquier transcripto con potencial codificante. El primer filtro descartó todos aquellos transcritos menores a 200 pb, para descartar RNAs no codificantes de menor tamaño.

Los filtros siguientes realizados mediante Blastx y Tblastx removieron los transcritos con potencial codificante y los posibles tRNAs y rRNAs. Las secuencias restantes se mapearon contra el genoma anotado de *C. rogercresseyi*, identificando los transcritos que efectivamente se encuentran en el genoma, obteniendo luego de estos filtros, un total de 5.209, 8.333 y 24.303 lncRNAs para copepodito, hembra y macho respectivamente.

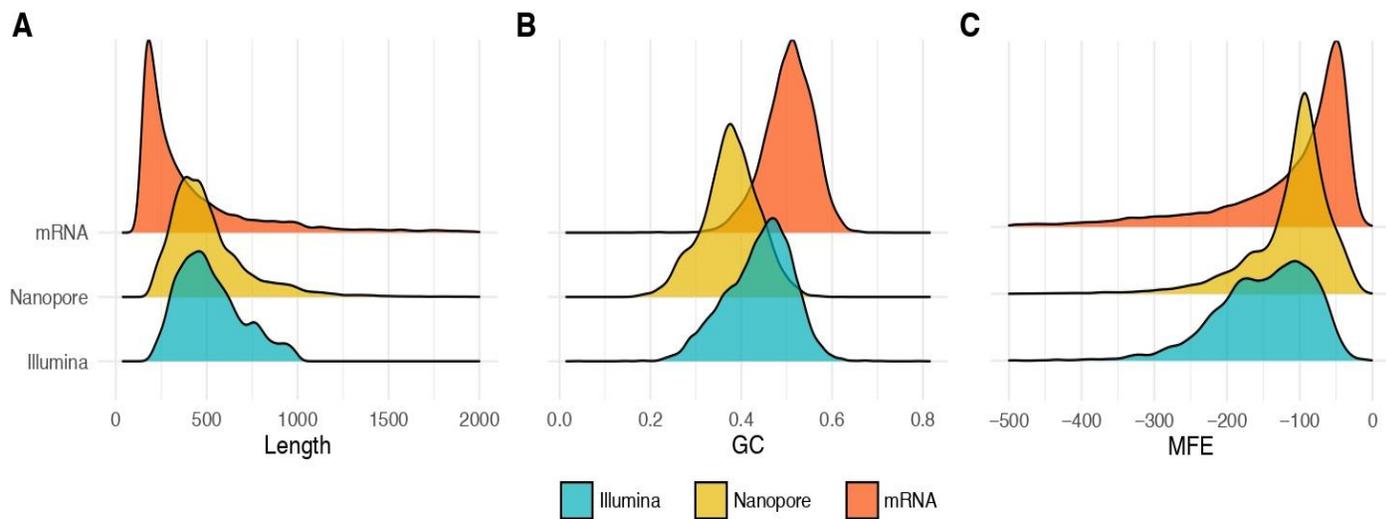
Las secuencias de lncRNAs de los 3 sets de datos fueron clusterizadas, agrupando aquellas secuencias que comparten entre sí un grado de identidad mayor al 90% bajo una única secuencia representativa de ese conjunto, formando así una única base de datos de lncRNAs para *Caligus rogercresseyi* a partir de secuenciación nanopore, obteniendo un total de 29.280 lncRNAs. Finalmente se aplicó un filtro para evaluar el potencial codificante a través de la presencia de Open Reading Frames (ORFs), descartando aquellas secuencias con ORFs que codifican para más de 70 aminoácidos, obteniendo finalmente una base de datos depurada con un total de 20.906 secuencias de lncRNAs.

#### 4.1.2. Comparación illumina – nanopore:

Las 20.906 secuencias obtenidas de la base de datos depurada de nanopore fueron comparadas con una base de datos de lncRNAs del trabajo realizado por Valenzuela-Miranda et al. (2017) en adición a otras secuencias obtenidas previamente, las cuales fueron obtenidas mediante secuenciación illumina. Esta base de datos también fue mapeada y clusterizada, quedando un total de 6.539 secuencias. A través del diagrama de venn realizado con ambas bases de datos se observa que la cantidad de data final obtenida por nanopore es mucho mayor a la obtenida por illumina. Mediante un análisis en Blastn se identificaron 1.536 secuencias compartidas entre ambos métodos de secuenciación, junto a 19.370 que serían identificables exclusivamente mediante nanopore y 5.003 exclusivas de illumina (Fig. 3).



**Fig. 3:** Diagrama de venn realizado con las bases de datos de lncRNAs Illumina y Nanopore para *C. rogercresseyi*. Las secuencias compartidas entre ambas bases de datos se identificaron mediante Blastn entre ambas, considerando como compartidas aquellas secuencias con valores de e value entre 0 & -50.



**Fig. 4:** Gráficos comparativos de RNAs no codificantes largos (lncRNAs) obtenidos por dos métodos de secuenciación diferentes (illumina y nanopore) y RNAs codificantes (mRNA) obtenidos por secuenciación illumina. En la figura se comparan la longitud de secuencias, el % Guanina-Citocina (GC), y la energía libre mínima (MFE).

La longitud de secuencias, %GC, y de energía libre mínima (MFE) fueron comparados entre los lncRNAs obtenidos por ambos métodos de secuenciación, y adicionalmente los mRNAs del genoma anotado de *Caligus* obtenidos por illumina, los cuales fueron usados como referencia (Fig. 4). En la comparación de longitud, los mRNAs con tamaños de 200 pb concentran la mayor frecuencia, mientras que los lncRNAs de tamaños entre 300-450 pb fueron los más frecuentes. Nanopore detectó levemente secuencias más largas de no codificantes (>1000 pb). Sin embargo, al igual que en illumina, la mayor frecuencia está concentrada en secuencias con tamaños entre 300-450 pb. En illumina, otros peaks de frecuencias más bajos fueron detectados en secuencias de 750 y 950 pb. Por otro lado, los mRNAs alcanzaron longitudes máximas similares, e incluso mayores que las de los lncRNAs de nanopore. En el caso de

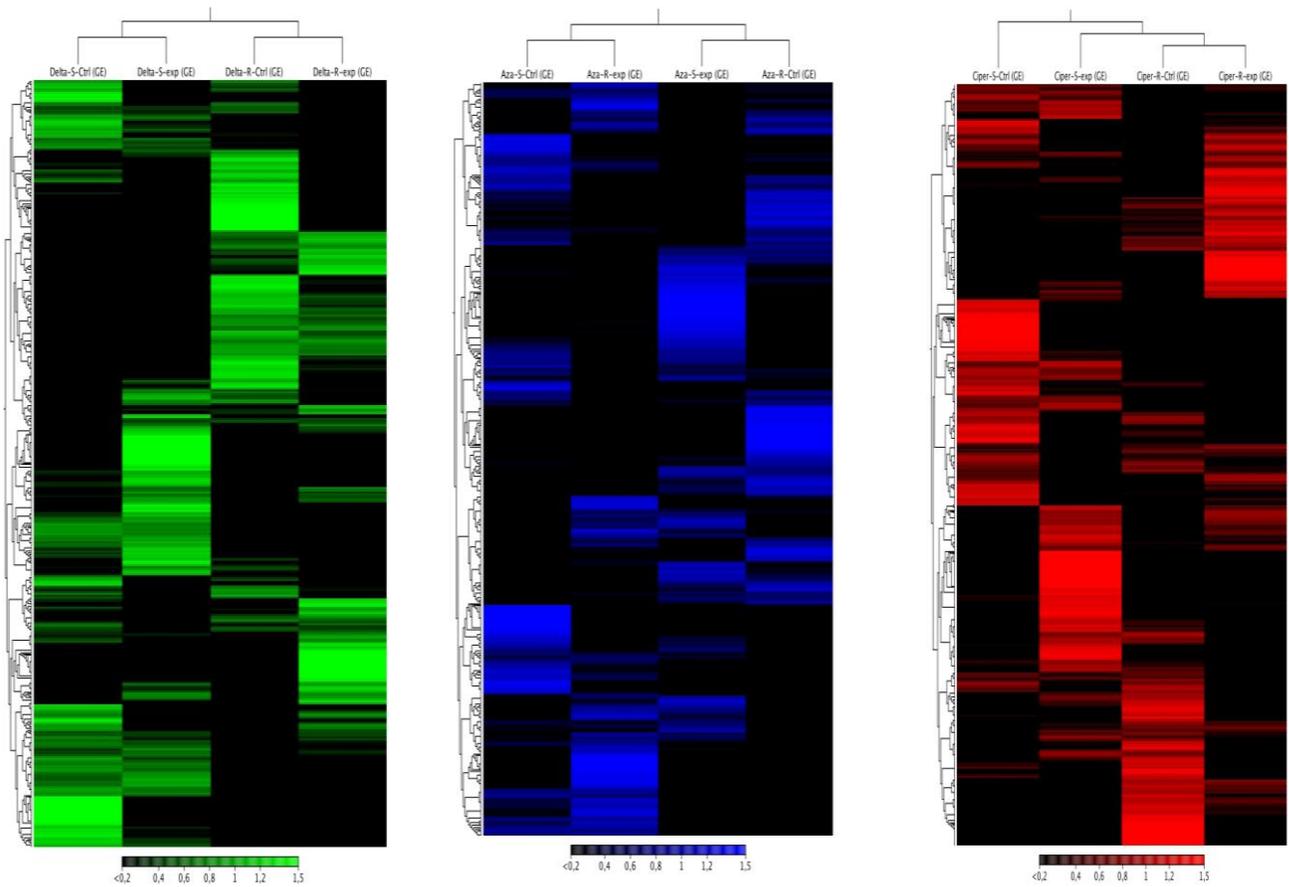
la comparación GC, las secuencias de mRNAs arrojaron un mayor %GC que el de las secuencias de no codificantes, sin embargo, los rangos de %GC para los no codificantes detectados con ambos métodos de secuenciación son más amplios que para las secuencias codificantes. En el caso de nanopore, este alcanza un rango levemente más amplio que illumina, detectando una mayor cantidad de secuencias con contenido GC relativamente bajos en comparación a las secuencias de lncRNA y mRNA obtenidas por illumina, las cuales presentan rangos de detección similares en ambos tipos de moléculas, y secuencias con contenido GC más altos. Por otro lado, el rango de energía libre mínima (MFE) detectada en los mRNA es mayor que la de los lncRNAs, teniendo un peak de máxima frecuencia en aquellas secuencias con MFE de -50, teniendo un amplio rango de detección. En contraste, nanopore detecta un peak de frecuencia concentrado en secuencias con MFE de -100, es decir, el doble de baja que la de los mRNA, teniendo un rango de detección más acotado, pero más robusto, mientras que en illumina hubo una mayor homogeneidad entre la MFE detectada en las secuencias de lncRNAs (Fig. 4).

#### 4.2. Caracterización de la modulación transcriptómica de lncRNAs en poblaciones de *C. rogercresseyi* expuestas a fármacos.

##### 4.2.1. RNA-seq y expresión diferencial:

Los sets de datos de fármacos fueron utilizados para evaluar la respuesta transcripcional de lncRNAs ante la exposición a fármacos (Azametifos, Deltametrina y Cipermetrina) en cepas de piojos susceptibles y resistentes. Mediante la construcción de heat maps (fig. 5) para cada fármaco, se compararon los patrones de expresión de lncRNAs entre el grupo control y el expuesto en cada cepa, donde se observó un contraste entre los clusters de genes de los diferentes tratamientos, así como también entre cepas (Fig. 5). Por ejemplo, para cada fármaco se observa expresión al alza en ciertos clusters de genes en el

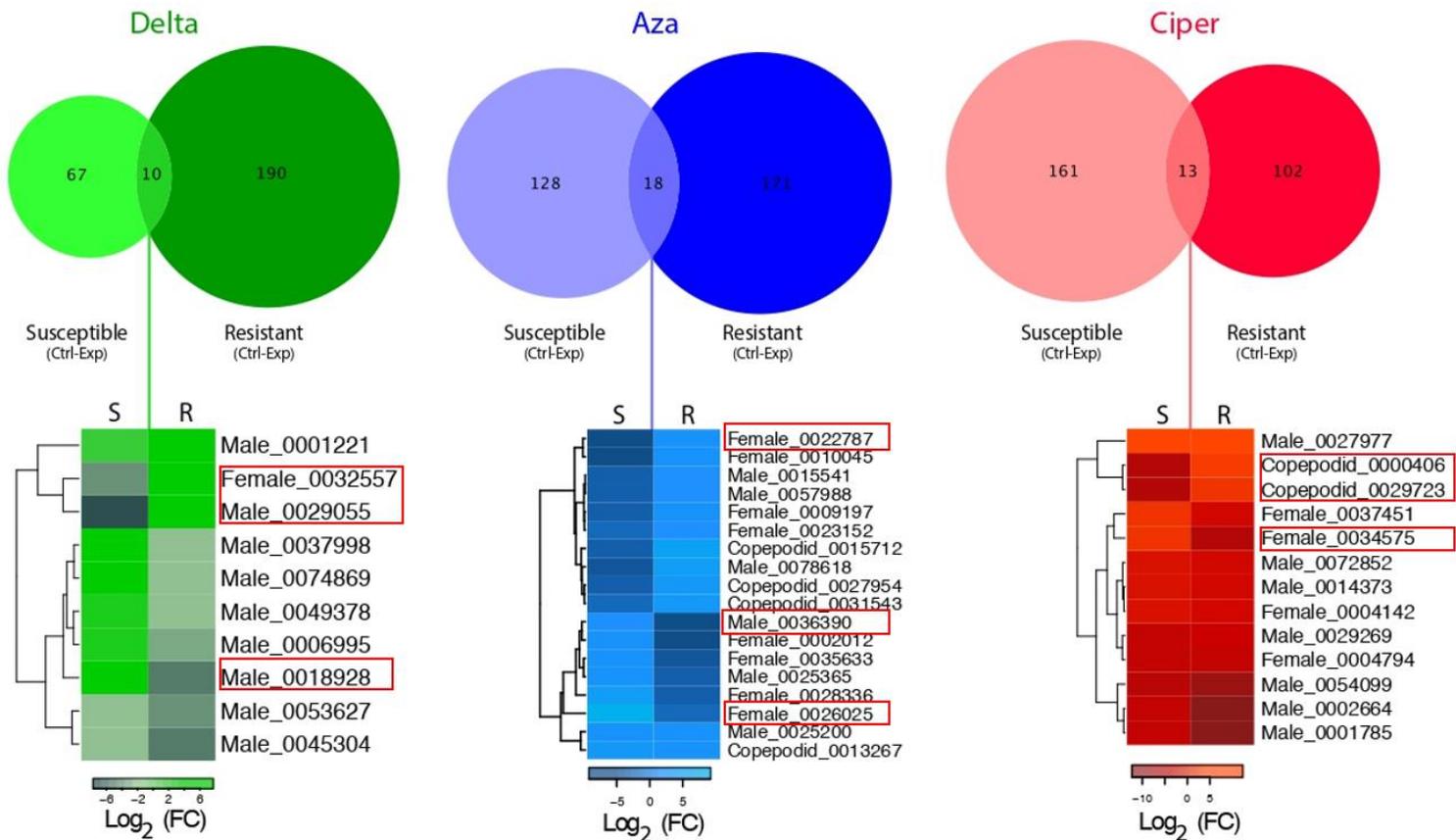
grupo control, mientras que en el grupo expuesto disminuye, obteniéndose también resultados similares al comparar los patrones de expresión entre cepas.



**Fig. 5:** Heat maps de expresión para lncRNAs en cepas de *Caligus rogercresseyi* resistentes y susceptibles a fármacos (Azul = Azametifos, verde = deltametrina, rojo = cipermetrina). En las figuras se puede observar una comparación entre las cepas que fueron expuestas a fármacos y su respectivo grupo control. Los valores de fold change se desplazan entre 0,2 – 1,5 (baja y alta regulación respectivamente).

#### 4.2.2. Expresión diferencial de lncRNAs y diseño de partidores para genes candidatos.

Posteriormente, se procedió a la identificación de lncRNAs diferencialmente regulados (FDR-p value = 0,01 & Fold change = 4) para cada una de las condiciones. Los resultados son representados mediante un diagrama de venn (Fig. 6). De esta manera, se encontraron un total de 267 lncRNAs en *Caligus* expuestos a Deltametrina, donde 67 fueron regulados exclusivamente en la población susceptible, 190 en la población resistente, y 10 compartidos entre ambas poblaciones. Ante la exposición a Azametifos, fueron registrados 128 lncRNAs regulados exclusivamente en la población susceptible y 171 en la resistente, además de 18 compartidos entre ambas poblaciones. Mientras que, en el caso de los expuestos a Cipermetrina, fueron detectados 161 lncRNAs regulados en la población susceptible, 102 en la resistente, y 13 compartidos entre ambas. En general, se detectó una mayor cantidad de lncRNAs expresados diferencialmente en las cepas resistentes de delta y aza, y en las cepas susceptibles de ciper.

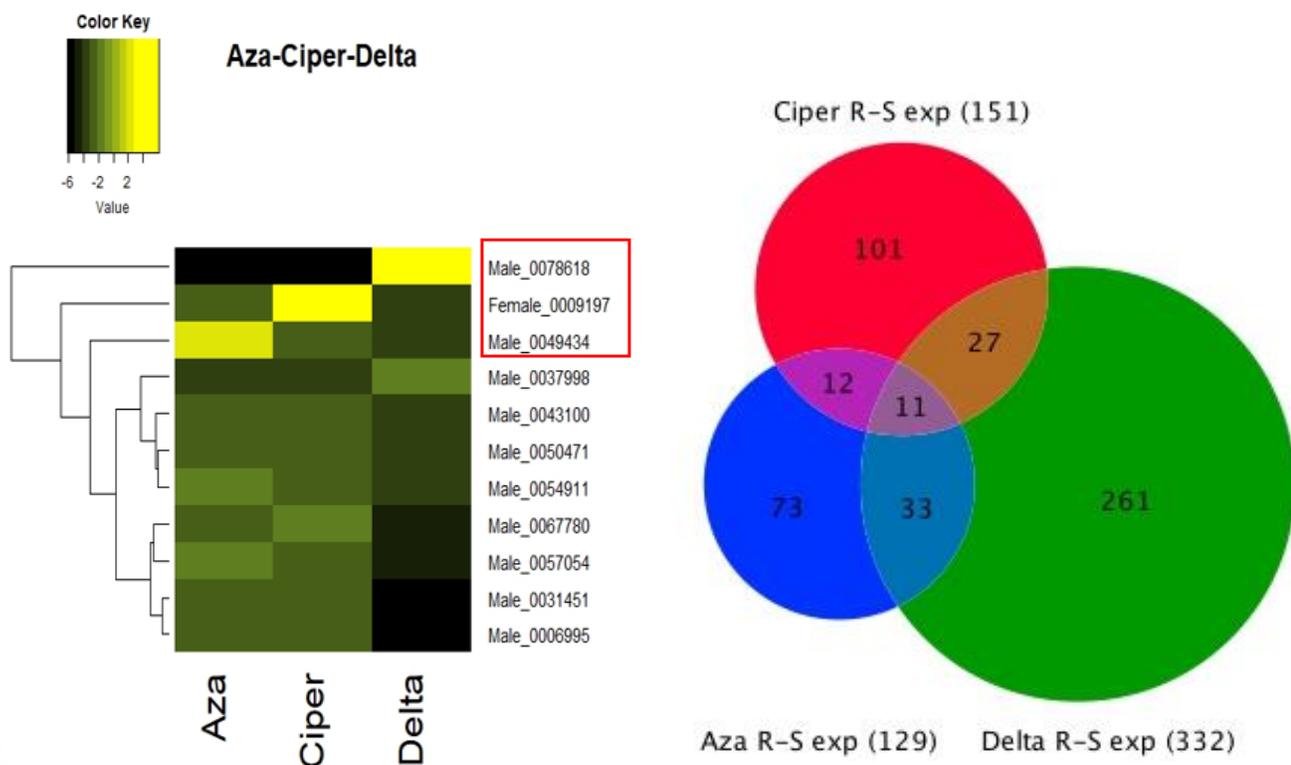


**Fig. 6:** Diagramas de venn (FDR-p value = 0,01 & fold change = 4) con sus respectivos heat map de las secuencias de lncRNAs expresados diferencialmente entre ambas cepas de *Caligus rogercresseyi*: resistentes y susceptibles a fármacos (Azametifos, deltametrina, cipermetrina). En cada cepa (R & S) están los lncRNA expresados diferencialmente entre el grupo que fue expuesto a los fármacos y su respectivo grupo control.

Un total de 10, 18 y 13 genes diferencialmente expresados entre ambas cepas fueron encontrados para los fármacos deltametrina, azametifos y cipermetrina respectivamente, realizándose un heat map por cada fármaco, para observar el contraste entre la respuesta transcripcional de ambas cepas. Sets de primers fueron diseñados con el software Geneious pro para los genes Female\_0032557, Male\_0029055 y Male\_0018928 de deltametrina; Female\_0022787,

Male\_0036390 y Female\_0026025 de azametifos; Copepodid\_0000406, Copepodid\_0029723 y Female\_0034575 de cipermetrina, para posteriormente validar su expresión mediante la técnica de qPCR. Globalmente, azametifos arrojó la mayor cantidad de genes expresados diferencialmente, mientras delta es el más contrastante al tener en la cepa susceptible y resistente la menor y mayor cantidad respectivamente, de genes expresados diferencialmente. En cambio, cipermetrina, a diferencia de los otros dos fármacos, es el único en donde la cepa resistente obtuvo la menor cantidad de genes con expresión diferencial.

Adicionalmente se realizaron los mismos pasos anteriores pero esta vez solo comparando entre cepas expuestas, para luego evaluar los genes expresados diferencialmente entre los 3 fármacos y elegir aquellos con valores de fold change más contrastantes para el diseño de partidores.

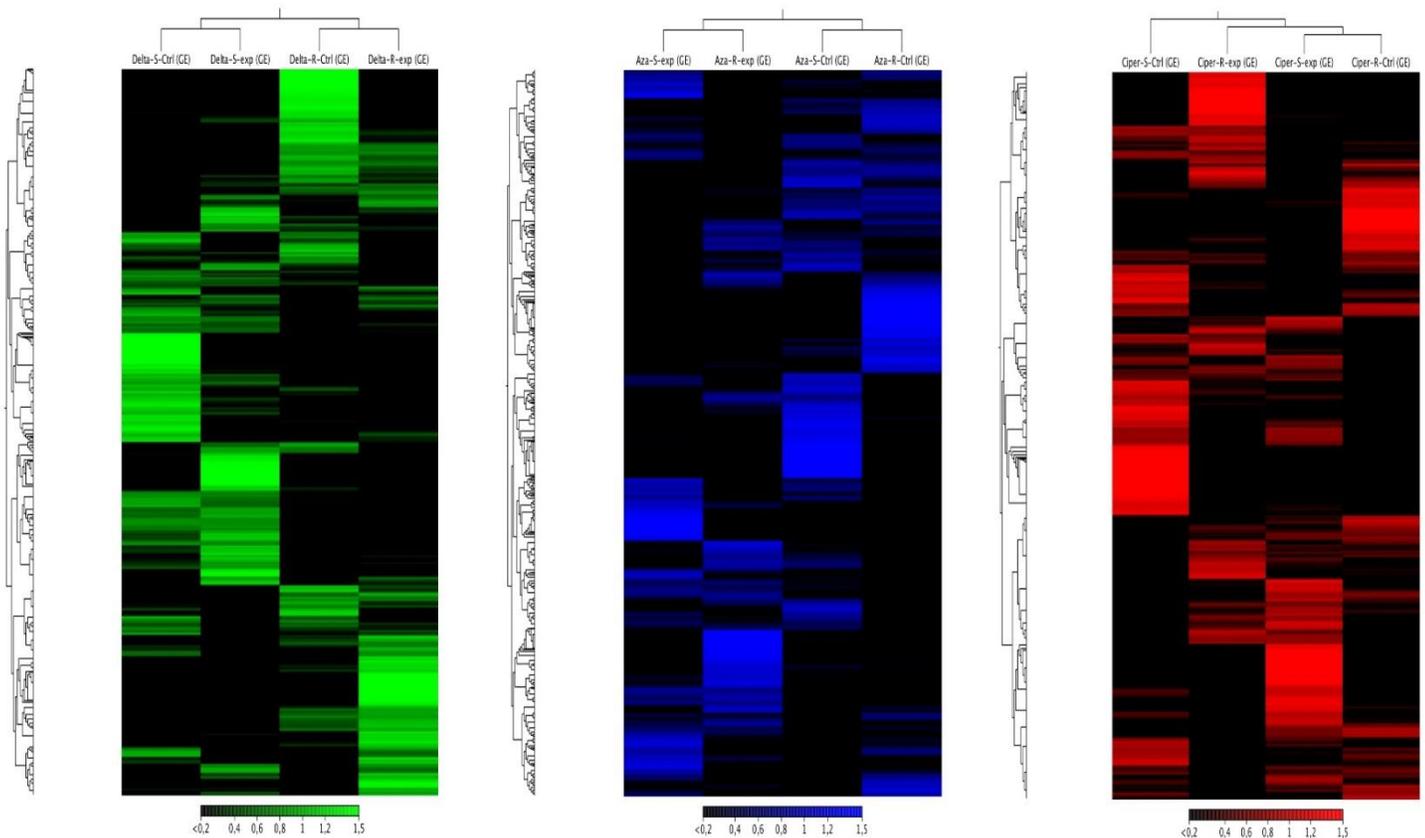


**Fig. 7:** A la derecha, diagrama de venn realizado para los lncRNAs expresados diferencialmente (FDR-p value = 0,01 & fold change = 4) entre las cepas susceptibles y resistentes que fueron

expuestas a los fármacos, y a la izquierda, su respectivo heat map de los genes expresados diferencialmente entre los 3 fármacos.

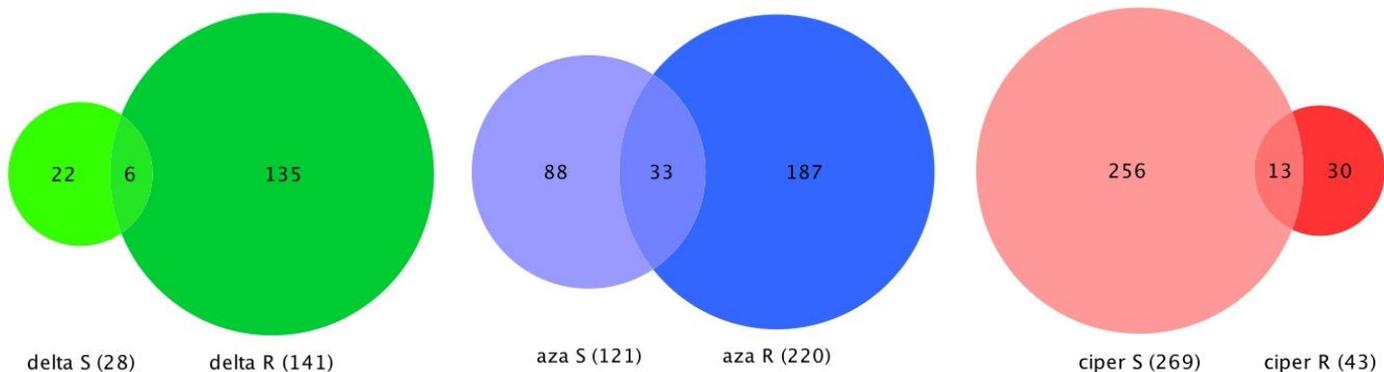
La mayor cantidad de lncRNAs expresados diferencialmente está representada en deltametrina con 261 genes, seguido de cipermetrina con 101, mientras que la menor cantidad está representada por azametifos con 73. Solo se detectaron 11 genes con expresión diferencial entre los 3 fármacos, visualizando la expresión de estos genes a través de un heat map, de los cuales se tomaron Male\_0078618, Female\_0009197 y Male\_0049434 para el diseño de partidores y su validación por qPCR, debido a encontrarse ligados cada uno a un fármaco en particular (Fig. 7).

#### 4.2.3. Expresión diferencial de mRNAs.



**Fig. 8:** Heat maps de expresión para mRNAs en cepas de *Caligus rogercresseyi* resistentes y susceptibles a fármacos (Azul = Azametifos, verde = deltametrina, rojo = cipermetrina). En las figuras se puede observar una comparación entre las cepas que fueron expuestas a fármacos y su respectivo grupo control. Los valores de fold change se desplazan entre 0,2 – 1,5 (baja y alta regulación respectivamente).

Análisis y procedimientos similares a los realizados para la creación de heat maps y venn diagrams de los lncRNAs fueron aplicados para los RNAs codificantes (mRNA) de *C. rogercresseyi* (fig. 8 & 9), evaluando su expresión diferencial ante la exposición a fármacos, con la finalidad de hacer una comparación general entre la respuesta transcripcional de ambos tipos de moléculas. Tanto los heat maps como los venn diagrams reflejan patrones de expresión de genes similares al de los lncRNAs, por ejemplo, la mayor cantidad de genes expresados diferencialmente en deltametrina y azametifos se encuentran en las poblaciones resistentes (135 & 187 respectivamente), mientras que en cipermetrina están en la población susceptible (256), siendo un patrón común entre RNAs codificantes y no codificantes (Fig. 8).



**Fig. 9:** Diagramas de venn (FDR-p value = 0,01 & fold change = 4) de las secuencias de mRNAs expresados diferencialmente entre ambas cepas de *Caligus rogercresseyi*: resistentes y susceptibles a fármacos (Azametifos, deltametrina, cipermetrina). En cada cepa (R & S) están los mRNA expresados diferencialmente entre el grupo que fue expuesto a los fármacos y su respectivo grupo control.

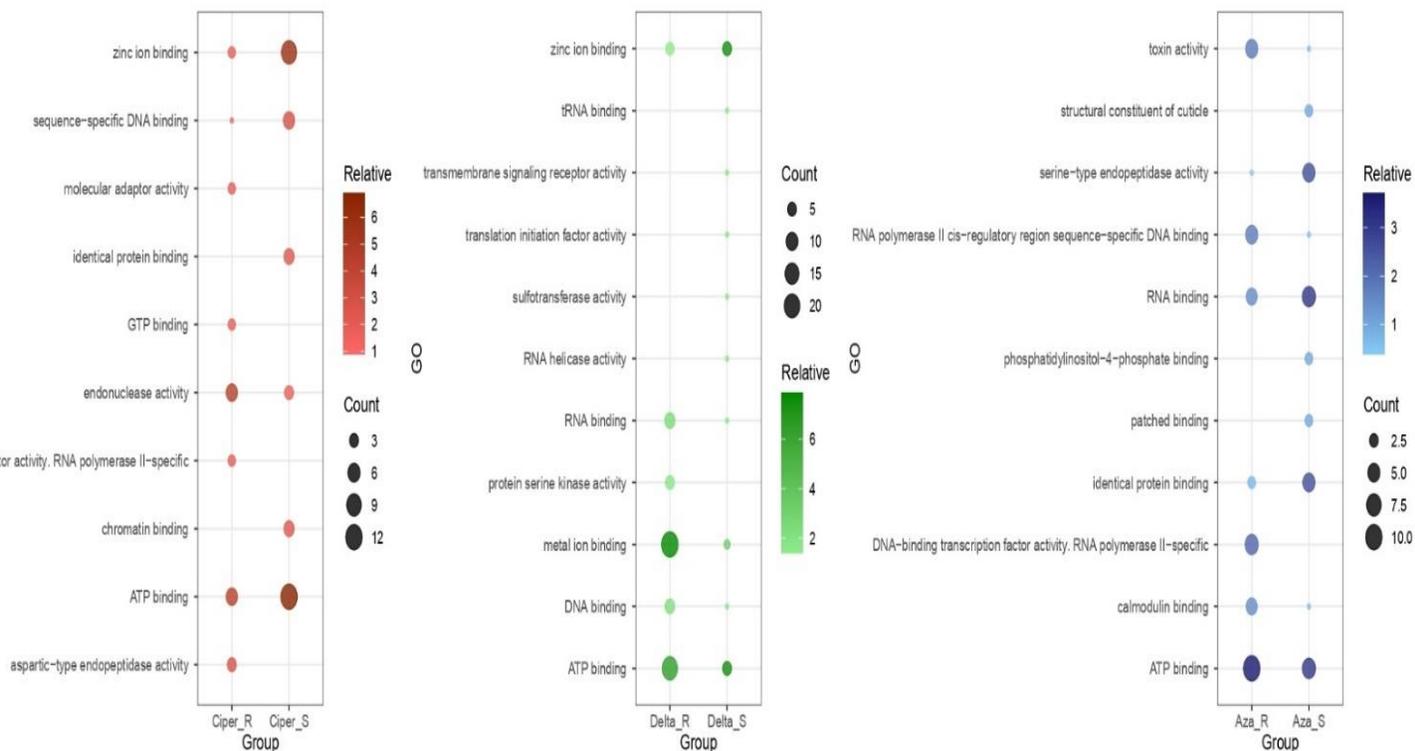
De la misma forma, azametifos contiene la mayor cantidad de genes expresados diferencialmente entre ambas cepas (33), deltametrina la menor (6), y cipermetrina contiene la misma cantidad (13) tanto para lncRNAs como para mRNAs. La mayor cantidad de genes no codificantes con expresión diferencial se encuentra en las cepas resistentes de deltametrina (190), mientras que la mayor cantidad para genes codificantes se encuentra en las cepas susceptibles de cipermetrina (256). Los mRNAs evidencian una mayor regulación ante la acción de cipermetrina y azametifos en cepas susceptibles (256) y resistentes (187) respectivamente, mientras que los lncRNAs serían mayormente afectados por cipermetrina y deltametrina en cepas susceptibles (161) y resistentes (190) respectivamente. En cepas resistentes la acción de deltametrina produce una mayor expresión diferencial de lncRNAs (190), mientras que en mRNA (187) es gatillada por azametifos.

#### 4.3. Asignación de funciones biológicas a lncRNAs expresados diferencialmente.

##### 4.3.1. Gene Ontology (GO) análisis y co-localización:

Los resultados obtenidos del análisis GO (Gene Ontology) muestran proteínas y funciones de genes codificantes cercanos a los lncRNAs que tienen gran relevancia tanto en procesos moleculares, como también de mantención celular y actividad enzimática. Entre las proteínas encontradas destacan la Serina que muestra una mayor regulación en cepas susceptibles, a diferencia de las cepas resistentes en las cuales esto no se observa, y la regulación de no codificantes cercanos es escasa, también, las cepas resistentes mostraron una mayor regulación de Calmodulina. ATP binding proteins tienen una baja actividad en cepas susceptibles, a diferencia de cepas resistentes en donde hay una mayor regulación. Tanto factores de transcripción como de post-transcripción fueron encontrados, entre ellos tRNAs, zinc ion bindings, la enzima helicasa, factores de

unión a cromatina, adaptadores moleculares. La sulfotransferasa está regulada solo en cepas susceptibles, mientras que la Kinasa tiene una mayor actividad de transcripción en cepas resistentes, las cuales también muestran regulación de variados factores de transcripción como los ya mencionados. Las cepas susceptibles también muestran expresión de fosfatidilinositol en respuesta a azametiphos, lípido que generalmente se encuentra en muy bajos niveles. En cepas resistentes expuestas a cipermetrina, se encontró expresión de GTP. En cepas resistentes expuestas a cipermetrina hay una mayor actividad endonucleasa que en cepas susceptibles (Fig. 10).



**Fig.10:** Plots de Gene Ontology para RNAs no codificantes largos (lncRNAs) por cada fármaco evaluado en ambas poblaciones de *C. rogercresseyi*.

#### 4.4. Validación molecular de la expresión de lncRNAs

##### 4.4.1. Estandarización de partidores y validación por qPCR:

Desde la estandarización de T<sub>m</sub>, se obtuvieron temperaturas de 58, 60, 62 y 64 °C en 4, 2, 1 y 1 par de partidores respectivamente, de un total de 12 pares, habiendo 3 de ellos que fueron descartados debido a que no hubo amplificación (N.A) (Tabla 1). Las T<sub>m</sub> óptimas para los partidores, fueron determinadas mediante la visualización en gel de agarosa del producto de PCR obtenido en las distintas temperaturas.

**Tabla 1.** Temperaturas de melting (T<sub>m</sub>) óptimas para los partidores diseñados, obtenidas desde la estandarización realizada por PCR. El número entre paréntesis corresponde a una identificación simplificada de cada partidor.

Partidor	M49434(1)	M36390(2)	M78618(3)	M18928(4)	M29055(5)	F22787(6)	F32557(7)	F9197(8)	F26025(9)	F34575(10)	C29723(11)	C406(12)
T <sub>m</sub> (°C)	58	60	58	N.A	58	60	58	60	N.A	64	62	N.A
Amplicón (bp)	83	103	108	97	88	114	95	86	108	80	96	97

El partidor número 2 mostró dimerización excesiva tanto en las muestras como en el control negativo, por lo que la concentración del primer reverse utilizada en el mix de PCR fue ajustada para una amplificación óptima, probando una concentración de 0,4 µM y una de 0,6 µM en las diferentes temperaturas, siendo la de 0,6 la que mostró mejor amplificación.

La eficiencia obtenida de la estandarización de concentraciones por qPCR, mostraron eficiencias fuera de los rangos aceptables (80-100) en la mayoría de los casos, excepto en el caso del partidor N° 1, 2 y 8. Debido a esto, se incluyó el uso de un partidor previamente estandarizado y diseñado para el lncRNAs 3544, identificado mediante secuenciación illumina, el cual también fue encontrado dentro de los expresados diferencialmente de Nanopore. Los

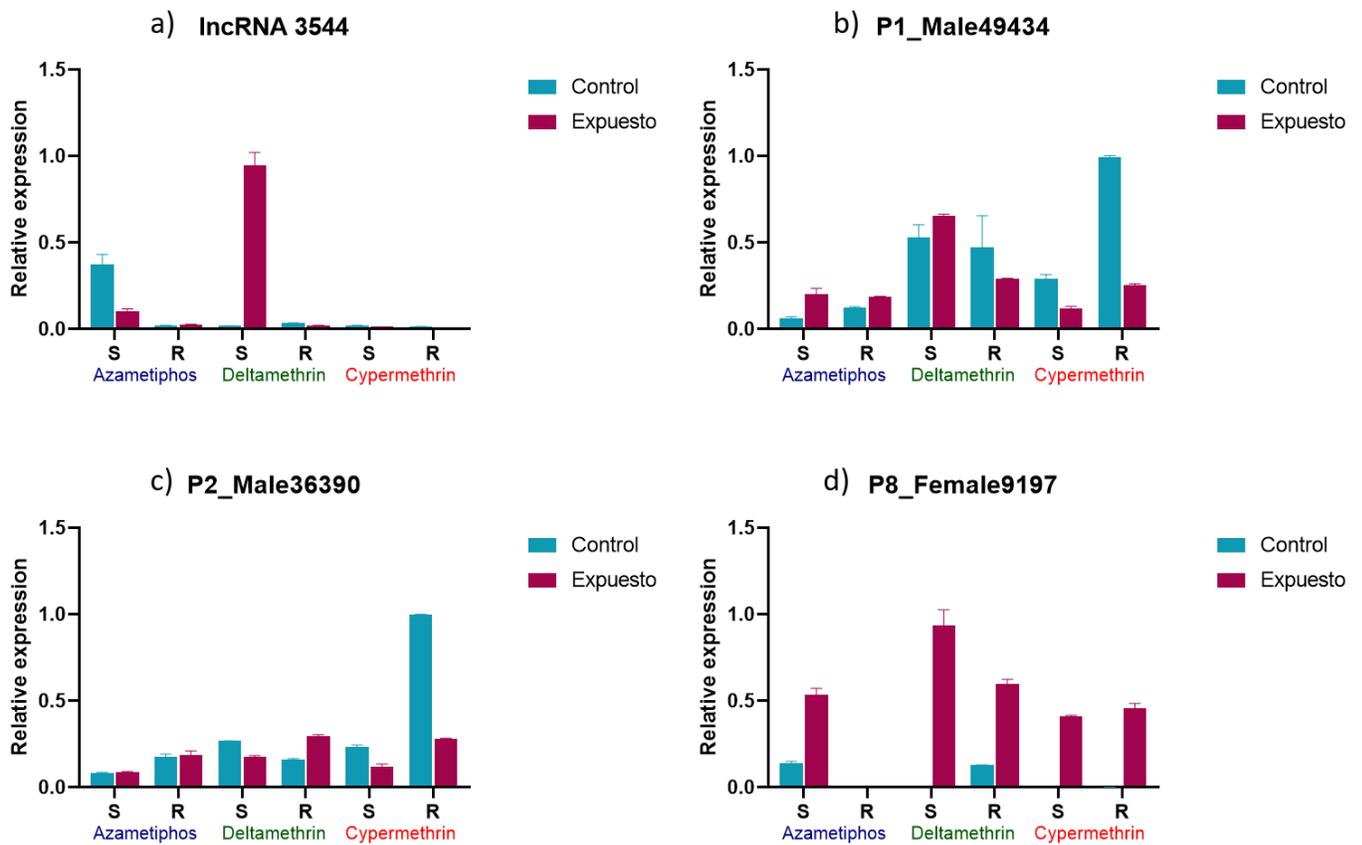
partidores para los lncRNAs Male\_49434 (1), Male\_36390 (2), Female\_9197 (8), y el lncRNA 3544, fueron utilizados en los cDNAs obtenidos desde los 12 pools de muestras, para mediante qPCR obtener la data necesaria para calcular su expresión relativa.

#### 4.4.2. Cálculo de expresión relativa para genes candidatos:

Los resultados indican que las cepas resistentes reducen la expresión del lncRNA Male\_49434 (Fig. 11b) y Male\_36690 (Fig. 11c) ante la exposición a cipermetrina, en relación al control en donde está sobreexpresado. Las cepas susceptibles conservan el mismo comportamiento, pero hay una expresión significativamente más baja. Sin embargo, cuando esta cepa es expuesta a deltametrina se observa una mayor expresión de Male\_49434. En cepas resistentes expuestas a deltametrina, la expresión de este no codificante es opuesta a la de Male\_36390. Por otro lado, el lncRNA 3544 (Fig. 11a) es escasamente transcrito en la mayoría de los tratamientos, excepto en las cepas susceptibles expuestas a deltametrina en donde se encuentra sobreexpresado, pero en respuesta a azametifos ocurre lo contrario, reduciendo su expresión en relación al control. El lncRNA Female\_9197 (Fig. 11d) evidencia un alto nivel de expresión en todos los tratamientos, siendo más evidente en las cepas susceptibles en respuesta a deltametrina. Ante la acción de azametifos las cepas susceptibles también evidencian expresión de este lncRNA, sin embargo, en las cepas resistentes su transcripción es nula. De forma similar, los controles de la mayoría de los tratamientos reprimen la transcripción de Female\_9197. Dentro de una ventana de 10.000 bp por ambos flancos de estos lncRNAs, se identificaron lncRNAs y proteínas no caracterizadas cercanas, sin embargo, entre los genes cercanos a Female\_9197 se identificó el gen Glutaryl-CoA deshidrogenasa mitocondrial. Además, tanto para el lncRNAs 3544 y Male\_36390 se registró la presencia de transposasas aledañas a una distancia mayor de 10.000 bp.

**Tabla 2.** Tabla de lncRNAs y sus genes cercanos identificados en el genoma anotado de *Caligus*.

LncRNA	Genes cercanos
lncRNA 3544	lncRNA nanopore, Transposasa
Male_49434	lncRNAs nanopore
Male_36390	Transposable element tcb2 transposasa, proteínas no caracterizadas
Female_9197	Glutaryl-CoA dehidrogenase_mitocondrial, proteínas hipotéticas



**Fig. 11:** Gráficos de expresión relativa estimada para evaluar la expresión de a) lncRNAs 3544, b) Male\_49434, c) Male\_36390, y d) Female\_9197 en cada pool de muestras.

## 5. Discusión:

El desarrollo de nuevas tecnologías de secuenciación para la obtención de información genómica, ha permitido expandir nuestro conocimiento sobre los mecanismos moleculares que rigen determinados procesos biológicos. En esta tesis, el objetivo fue utilizar nuevas tecnologías de secuenciación para la caracterización de lncRNAs que puedan estar involucrados en la respuesta farmacológica de *Caligus rogercresseyi* expuestos a diferentes drogas utilizadas para su control. Nuestros resultados indican que la secuenciación nanoporos permite la caracterización de nuevos lncRNAs que no habían sido descritos para esta especie.

Este hecho puede deberse a dos razones: la primera es debido a la naturaleza misma del método de secuenciación, el cual puede entregar reads más largos que illumina teniendo un rango de acción más dinámico (Rollano Peñaloza & Mollinedo Portugal, 2017; Tosco Herrera, 2019); la segunda, muy relacionada con la anterior, se debe a que nanopore obtiene la secuencia completa del gen al pasar el DNA por el nanoporo, eliminando la necesidad de hacer un ensamblaje *de novo*, a diferencia de illumina en donde el DNA debe fragmentarse en múltiples reads pequeños que luego deben agruparse en contigs mediante un ensamblaje *de novo* (Kircher, Heyn, & Kelso, 2011; Heikema et al., 2020). De esta forma, un read en nanopore equivaldría a la secuencia completa de un transcrito, a diferencia de illumina en donde múltiples reads ensamblados forman un contig equivalente a la secuencia del gen. Al agrupar tantas secuencias, estas pueden perderse dentro del contig, pudiendo no detectar todos los genes que realmente se encuentran en un transcriptoma, por lo que los trabajos realizados con illumina podrían estar subestimando la cantidad de secuencias de lncRNAs. Sin embargo, también es necesario considerar que la cantidad de secuencias obtenidas desde el posterior análisis bioinformático, siempre tenderá a variar según el pipeline de filtrado que se utilice. Por otro lado, tanto la longitud, el contenido GC, y la energía libre mínima (MFE) varían levemente entre las secuencias obtenidas por illumina

y nanopore, por ejemplo, el rango de GC detectado por nanopore es levemente más amplio que illumina, pero esa leve variación detecta secuencias con un % GC menor, que con illumina probablemente no se hubiesen detectado. Esto concuerda con lo descrito para la especie en trabajos previos, ya que los lncRNAs contienen % GC más bajos que los RNAs codificantes (Valenzuela-Miranda, & Gallardo-Escarate, 2017). De igual forma, si bien no hay una gran diferencia entre las longitudes de secuencias no codificantes entregadas por ambos métodos de secuenciación, nanopore detecta secuencias más largas que illumina, sin embargo, este último puede detectar fragmentos más largos cuando se trata de RNAs codificantes, aunque obviamente eso debe ser verificado en trabajos posteriores, ya que este trabajo está enfocado principalmente en evaluar lncRNAs mediante secuenciación nanopore, y los mRNAs solo fueron utilizados como referencia. La longitud media de las secuencias codificantes detectadas corresponde a 200 pb, mientras que para lncRNAs corresponde a 500 pb, lo cual contrasta con la literatura ya que los RNAs codificantes tienden a ser más grandes que los no codificantes (Valenzuela-Miranda, Etebari, Asgari, & Gallardo-Escárate, 2017). Esto puede deberse a que los mRNAs fueron obtenidos desde la primera versión del genoma de *Caligus*, la cual aún necesita ser refinada. Sin embargo, es importante considerar que tanto los lncRNAs y mRNAs pueden tener estructuras y longitudes similares (Murillo-Maldonado & Riesgo-Escovar, 2019). La energía libre mínima (MFE) está relacionada con la capacidad de formar estructuras secundarias y mantener su estabilidad, siendo las moléculas con secuencias más largas aquellas más propensas a esto. En este contexto, los resultados obtenidos son concordantes, ya que, si bien los mRNAs obtenidos son en su mayoría cortos, hay una porción de ellos que supera en longitud incluso a los lncRNAs de nanopore, lo que coincide con los niveles de MFE detectados tanto para lncRNAs como para mRNAs. Además, los lncRNAs estarían vinculados a formar estructuras secundarias menos estables que los mRNAs (Wang et al., 2014). Sin embargo, nuestros resultados contrastan

con los obtenidos en trabajos previos (Valenzuela-Miranda et al., 2017), donde los lncRNAs tienen valores de MFE mayores a los mRNAs, pero esto se debe al tamaño de los mRNAs que están anotados en el genoma. En resumen, ambos métodos de secuenciación son complementarios, y el uso de uno u otro también dependerá del objetivo de la investigación.

La expresión diferencial entre los clusters de genes observados en los heat maps (fig.5) realizados para los sets de datos, evidencia una regulación a nivel de transcriptoma tanto para poblaciones (R/S) como para tratamientos (control-expuesto) en cada fármaco. Esto nos entrega una aproximación de cómo el transcriptoma responde ante la exposición farmacológica y la diferencia que hay entre los patrones de expresión de las cepas tratadas. De igual forma, al observar la expresión de lncRNAs junto a la de mRNAs en cada set de datos, se puede observar un contraste en la respuesta transcriptómica de ambos tipos de moléculas, pudiendo tener roles específicos asociados a la respuesta farmacológica o incluso interactuar entre ellas.

Los lncRNAs son moléculas involucradas en múltiples procesos y funciones biológicas, teniendo un nivel de especificidad incluso mayor al mRNA (Schmitz, Grote, & Herrmann, 2016). Debido a esto, es probable que Male\_78618, Female\_9197 y Male\_49434, genes candidatos elegidos de aquellos expresados diferencialmente entre los 3 fármacos, tengan afinidad por Deltametrina, Cipermetrina y Azametifos respectivamente, siendo gatillados por la exposición a estos. Si esto es así, podrían ser utilizados como biomarcadores para medir la respuesta transcripcional y la eficiencia de los fármacos utilizados para tratar la caligidosis. Sin embargo, la validación molecular evidenció la expresión de Female\_9197 ante Azametifos y de Male\_49434 ante Cipermetrina.

Tanto los lncRNAs como los mRNAs mostraron patrones de expresión similares, lo cual podría estar relacionado con la respuesta transcriptómica empleada por el organismo para enfrentar la acción del fármaco, debido a que ambas moléculas

son capaces de interactuar mutuamente. Además, es frecuente que los RNAs codificantes dependan de la acción de RNAs no codificantes como los lncRNAs para su correcta expresión, funcionamiento y desempeño (Cajal & Hümmer, 2018). Esto, sugiere que ambas moléculas podrían estar co-expresándose, siendo más evidente el caso de cipermetrina en donde hubo la misma cantidad de lncRNAs y mRNAs expresados diferencialmente entre cepas. Por otro lado, los RNAs codificantes de cepas susceptibles evidencian una mayor regulación ante la acción de cipermetrina, y azametifos en el caso de cepas susceptibles y resistentes, mientras que los lncRNAs serían mayormente afectados por cipermetrina y deltametrina en cepas susceptibles y resistentes respectivamente. Además, la acción de deltametrina produce una mayor expresión diferencial sobre los lncRNAs, mientras que en mRNA es gatillada por azametifos, por lo que en cepas resistentes el efecto del fármaco podría actuar sobre un tipo específico de molécula. Por otro lado, el transcriptoma de las cepas susceptibles sería más sensible a la acción de cipermetrina, mientras que en las cepas resistentes hay una menor cantidad de genes diferencialmente expresados, lo que puede deberse a que el transcriptoma de los piojos resistentes se vea menos afectado por este fármaco. La posible co-expresión de RNAs codificantes/no codificantes concuerda con el trabajo de Núñez-Acuña et al. (2020), en donde se registraron patrones específicos del transcriptoma en piojos de mar expuestos a estos fármacos (azametifos, deltametrina y cipermetrina), así como también, lncRNAs fueron vinculados con genes codificantes expresados diferencialmente en cada tratamiento.

Los resultados obtenidos del análisis GO y co-localización evidencian la relevancia de los lncRNAs expresados en ambas cepas ante la exposición farmacológica, debido a que estas moléculas son capaces de interactuar con genes aledaños mediante una amplia variedad de mecanismos moleculares (Rogoyski et al., 2017). Por esta razón, los lncRNAs de cepas susceptibles estarían mayormente interactuando o regulando la expresión de genes

codificantes asociados a proteínas como las serinas, que están asociadas al mantenimiento celular, reparación, metabolismo, sistema nervioso, inmunidad, etc. (Ceña, Galindo, Jordán, & Fuster Lluch, 2004). En el sistema nervioso, las proteínas de la familia de las serinas participan en procesos degenerativos, y junto a las caspasas regulan funciones celulares como el desarrollo (Ceña et al., 2004). Las granzimas junto a la catepsina G también se relacionan con procesos degenerativos. La familia de proteasas tipo TR es considerada como un modulador temprano en los tejidos dañados, sirviendo como señal extracelular de muerte que conduce a la activación de mecanismos intracelulares mediados por calcio. Además, como la actividad de las serina proteasas es regulada por inhibidores endógenos, alteraciones en el equilibrio proteasa-inhibidor resultan cruciales en los procesos implicados en plasticidad y muerte neuronal inducida (Ceña et al., 2004). Por otro lado, las tripsinas están entre las serinas proteasas más abundantes en crustáceos, y algunos estudios las han relacionado con la resistencia a Azametifos, registrando una asociación entre tripsinas y la degradación de este fármaco en *Caligus rogercresseyi* (Chávez-Mardones & Gallardo-Escárate, 2015; Sáez-Vera, Núñez-Acuña, & Gallardo-Escárate, 2022). Esto concuerda con los resultados de este trabajo, donde se evidencia una mayor regulación de serinas en cepas susceptibles expuestas a Azametifos, lo cual puede estar asociado a la regulación de constituyentes estructurales de la cutícula que también fueron registrados en este tratamiento. En la cepa resistente esto no se observa, y la regulación de no codificantes cercanos es escasa. Otra proteína asociada a los no codificantes correspondería a calmodulina, la cual se encuentra en todas las células eucariotas, principalmente el cerebro y corazón, regulando la transducción de la señal de calcio a través de la activación de la ATPasa de calcio (Tinjacá Espinel, 1995). Debido a que hay una mayor regulación de esta proteína en cepas resistentes expuestas a azametifos, es posible que el calcio ayude a la metabolización del fármaco, probablemente mediante el transporte de moléculas. De hecho, la calcineurina es una enzima

dependiente de calcio, estimulada por la acción de calmodulina. Estudios realizados por Enan & Matsumura (1992), indican que la calcineurina es inhibida por los piretroides tipo 2, como Cipermetrina y Deltametrina, pudiendo tener un impacto directo sobre calmodulina. Los resultados obtenidos concuerdan con estos estudios previos, ya que calmodulina solo fue detectada en las cepas tratadas con Azametifos, siendo reprimida en respuesta a los piretroides tipo 2 (Enan & Matsumura, 1992). Además, el hecho de que en cepas resistentes haya una reducción en la regulación de serinas, y un aumento en la regulación de calmodulina, sugiere un posible mecanismo de resistencia en donde se le da prioridad a la regulación de calmodulina, debido tal vez, a que la regulación de serinas resulte energéticamente desfavorable. También se encontró la presencia de sulfotransferasas, enzimas presentes en la raíz del folículo piloso, hígado y plaquetas. Esta es capaz de activar las moléculas de algunos medicamentos para que puedan ejercer su función, ya que cuando un fármaco penetra en el cuerpo, gran parte de las enzimas encargadas de transformar el agente químico para su excreción se encuentran en el hígado (Del Arco, 2008), estando esta enzima regulada solo en cepas susceptibles expuestas a deltametrina, lo cual es lógico ya que indica que el fármaco está haciendo efecto. El fosfatidilinositol fue identificado entre uno de los genes codificantes evaluados, este es un componente fosfolipídico de la membrana celular, así como también de orgánulos intracelulares, y pueden regular la liberación de neurotransmisores mediante la activación e inhibición de diversos canales iónicos (Burgos & Alamilla, 2010), generalmente se encuentra en muy bajos niveles así que es posible que en cepas susceptibles haya una sobreexpresión en respuesta al daño producido por el fármaco. Es una molécula que participa en la comunicación celular promoviendo cambios dentro de la célula. Las kinasas son otras de las proteínas asociadas a estos codificantes, estas son enzimas que modifican a otras moléculas mediante fosforilación, activando o desactivándolas, lo cual es fundamental en diversos procesos del metabolismo como el desarrollo, la síntesis de fosfolípidos,

biogénesis de los ribosomas, incremento en la síntesis de mRNA (proteínas) y la inhibición de la autofagia, entre otras (Aguilar, Sánchez, Bocardo, & Pardo, 2018), pudiendo intervenir en cepas resistentes expuestas a deltametrina, las cuales muestran variados factores de liberación de energía como los ATP binding, necesaria para el transporte de moléculas dentro y fuera de la célula, habiendo una menor regulación en cepas susceptibles, debido probablemente a que al ser susceptible al fármaco hay menos ATP libre, y por lo tanto, menor transporte de xenobióticos. Por otro lado, en cepas resistentes hay una mayor regulación, pudiendo tener más disponibilidad de ATP y moléculas energéticas que induzcan el transporte de moléculas para detoxificar la célula. Otros de estos factores corresponderían a zinc ion bindings, helicasa, tRNAs, adaptadores moleculares, factores de unión a cromatina, entre otros, los cuales pueden actuar tanto a nivel de transcripción, como de post-transcripción, como por ejemplo los tRNAs que están involucrados directamente en la síntesis proteica en los ribosomas, al igual que los GTP bindings, o la helicasa que actúa en la replicación y reparación del ADN, renovación celular, biogénesis de ribosomas, así como también factores de unión a cromatina, que regulan la compactación de la cromatina, eucromatina y heterocromatina (Murillo-Maldonado & Riesgo-Escovar, 2019; Schmitz, Grote, & Herrmann, 2016; Trovero & Geisinger, 2019). Factores de unión a cromatina están regulados en cepas susceptibles expuestas a cipermetrina, probablemente para modular la actividad de transcripción en respuesta al fármaco. Proteasas y endonucleasas fueron las principales proteínas asociadas a los genes codificantes cercanos a los lncRNAs, las primeras se encuentran en toda clase de organismos y tienen una amplia gama de actividades biológicas (Ceña et al., 2004). Intervienen en la digestión de proteínas en el alimento. También participan en la señalización celular al descomponer otras proteínas, como anticuerpos u hormonas, pudiendo activar o desactivar vías. En cepas resistentes hay regulación, la cual puede deberse a la degradación del alimento, es decir, el organismo está en buen estado y el efecto del fármaco es poco o nulo.

Generalmente se producen primero como una molécula más grande que está inactiva. Esto protege de daños a la célula que lo sintetiza. Por otro lado, las endonucleasas catalizan la ruptura de enlaces fosfodiéster dentro de una cadena nucleotídica, aquí se incluyen las enzimas de restricción (Expósito\_Martínez, 2018). Es posible que haya reparación del DNA en cepas resistentes debido a una mayor actividad endonucleasa. Estos resultados concuerdan con la naturaleza de los fármacos evaluados, ya que al ser piretroides tipo 2 como es el caso de cipermetrina y deltametrina, estos inhiben la actividad neuronal y modifican la transición de los canales de sodio en estados inactivados y abiertos (Farag et al., 2021). Para el caso de azametifos, este corresponde a un organofosforado, el cual también tiene efectos neurotóxicos inhibiendo la actividad de acetil colinesterasa, lo que resulta en la parálisis y posterior muerte del individuo objetivo (Parsons, Escobar-Lux, Sævik, Samuelsen, & Agnalt, 2020).

Por otro lado, la expresión relativa de lncRNAs, sugiere que las cepas resistentes reducen la expresión del lncRNA Male\_49434 y Male\_36690 ante la exposición a cipermetrina. En ambos fueron identificados lncRNAs y genes codificantes aledaños aún no caracterizados, en ventanas de 10.000 bp en cada flanco, por lo que, al interactuar entre sí, podrían actuar como mecanismos moleculares de regulación, implicados probablemente en la metabolización del fármaco, pudiendo regular la expresión incluso de proteínas transportadoras de la familia ABC (ATP-Binding Cassette) (Gottesman & Pastan, 1993). Esto contrasta con lo observado en cepas susceptibles, donde hay una expresión significativamente más baja, además, en los resultados obtenidos desde los análisis GO, las cepas resistentes expuestas a este fármaco mostraron una disminución en factores de transporte como zinc ion bindings, adaptadores moleculares, GTP y ATP bindings, así como también factores de unión específicos al ADN y proteínas, y reguladores de transcripción como los chromatin bindings. Sin embargo, para obtener conclusiones más precisas, es necesario realizar más análisis para la

caracterización de los genes cercanos a estos no codificantes. En respuesta a deltametrina, las cepas susceptibles registran una mayor expresión de Male\_49434, mientras que, en cepas resistentes, la expresión de este no codificante disminuye respecto al control, teniendo un comportamiento opuesto al de Male\_36390. Sin embargo, en cepas resistentes los niveles de expresión de ambos lncRNAs son similares y se mantienen tanto para la acción de deltametrina y cipermetrina, por lo que es probable que las cepas resistentes tengan una mayor regulación de estos no codificantes, estabilizando sus niveles de transcripción en respuesta a la acción farmacológica. Por esto, se sugiere que para la acción de deltametrina, estos lncRNAs podrían estar regulando funciones asociadas a los factores identificados en el análisis GO, donde las cepas resistentes evidencian una mayor expresión de factores de ATP, RNA, y DNA bindings, unión a iones metálicos como el zinc y actividad proteica de serinas, pudiendo estos lncRNAs estar asociados al mantenimiento celular, metabolismo, sistema nervioso, detoxificación, entre otros (Sturm, Cunningham, & Dean, 2009; Schmitz et al., 2016). Mientras que, en comparación a las cepas susceptibles, hay una reducción en la actividad de receptores de señales transmembrana, sulfotransferasas, factores de transcripción como helicasas y de traducción como tRNA bindings e iniciadores de traducción. Esto puede deberse a que, ante la acción del fármaco, las cepas resistentes no necesitan sobreexpresar estas moléculas. En el caso del lncRNA 3544, en la mayoría de los tratamientos su transcripción es escasa o nula, excepto en cepas susceptibles expuestas a deltametrina en donde se encuentra sobreexpresado, posiblemente intentando aumentar la actividad de señalizadores y transportadores transmembrana en respuesta al fármaco, o también podría regular la actividad de sulfotransferasas involucradas en la transformación química del fármaco (Del Arco, 2008), sin embargo, en algunos fármacos la acción de esta enzima puede actuar como un activador. Ante azametifos ocurre lo contrario, reduciendo la expresión de este lncRNA en relación al control, por lo que, en este contexto, azametifos podría

tener una mayor efectividad sobre el parásito. El lncRNA Female\_9197 se encuentra altamente expresado en todos los tratamientos, siendo más evidente en las cepas susceptibles expuestas a deltametrina. En respuesta a azametifos, las cepas susceptibles también evidencian expresión de este lncRNA, sin embargo, en las cepas resistentes su transcripción es nula, al igual que en los controles de la mayoría de los tratamientos, donde su transcripción también se ve reprimida. Al ser expresado altamente en todos los tratamientos expuestos a fármacos, los cuales alteran el funcionamiento de los canales iónicos, es probable que este no codificante contribuya a la eliminación de toxinas regulando la acción de ATP bindings o calmodulina, teniendo implicancias en el transporte iónico, al igual que el fosfatidilinositol que es un componente fosfolípido de la membrana celular (Burgos & Alamilla, 2010; Tinjacá Espinel, 1995). Además, en respuesta a azametifos, el análisis GO también identificó factores de transcripción como RNA bindings y polimerasas específicas. Deltametrina retarda el transporte de iones como el Na (Sevatdal et al., 2005), por lo que la sobreexpresión de Female\_9197 en cepas susceptibles puede estar relacionada con la actividad del transporte iónico, aumentándolo para compensar el efecto del fármaco. Azametifos inhibe la síntesis de acetilcolinesterasa, enzima que hidroliza al neurotransmisor acetilcolina (Vera Bizama, 2017). Debido a esto, es posible que la nula expresión de Female\_9197 en cepas resistentes, sea utilizado como un tipo de mecanismo para protegerse de la acción del fármaco, pudiendo estar asociada la expresión de este lncRNA con la represión de otros genes, probablemente con roles reguladores de transcripción. Por esto, la nula transcripción de este no codificante podría aumentar la transcripción de genes asociados a la resistencia contra azametifos.

El gen Glutaryl-CoA deshidrogenasa mitocondrial fue identificado cercano a Female\_9197. Además, tanto el lncRNA 3544 y Male\_36390 registran presencia de transposasas aledañas a una distancia mayor de 10.000 bp, pudiendo ser

importantes en la regulación de transposones, probablemente involucrados en procesos de activación o desactivación de genes (Murillo-Maldonado & Riesgo-Escovar, 2019; Schmitz, Grote, & Herrmann, 2016). En este sentido, en respuesta a la acción farmacológica, la regulación de transposones sería energéticamente favorable en comparación a la regulación de RNAs codificantes, obteniéndose incluso mecanismos de resistencia más eficientes, debido a la alta especificidad que presentan los lncRNAs (Núñez-Acuña et al., 2022). La Glutaril-CoA deshidrogenasa (GCDH) es una enzima mitocondrial clave en la degradación de aminoácidos, evitando la acumulación de sustancias dañinas para la célula. La sustancia Glutaril-CoA se produce cuando se procesan los aminoácidos lisina, hidroxilisina y triptófano. Cuando la sustancia Glutaril-CoA no puede eliminarse, provoca que el ácido glutárico y otras sustancias dañinas se acumulen en la sangre y causen problemas (Rojas & Fortini, 2020). La regulación de este gen y la transcripción de Female\_9197 podría ser una importante relación RNA codificante/no codificante, actuando como un mecanismo de respuesta ante la exposición farmacológica, ya que este lncRNA se encuentra expresado en todos los tratamientos expuestos, excepto en cepas resistentes expuestas a azametifos, en donde su transcripción es reprimida debido a que, probablemente esto favorece la regulación de genes que ayudan a resistir el efecto de este fármaco. Los long non-coding RNAs (lncRNAs) identificados en este trabajo se sugieren como posibles bioindicadores de resistencia farmacológica. Como bioindicador para la resistencia a azametifos, Female\_9197 surge como un buen candidato, ya que es reprimido en cepas resistentes pero expresado en susceptibles. Para deltametrina, la mejor opción sería el lncRNA 3544, detectándose una sobreexpresión en cepas susceptibles y represión en las resistentes. Por último, ante la acción de cipermetrina, tanto el lncRNA Male\_49434 y Male\_36390 podrían ser usados, ya que ambos tienen un comportamiento similar ante este fármaco, y se podrían detectar sin la necesidad de exponer organismos ante el fármaco, habiendo una sobreexpresión en los

controles de la cepa resistente, y una disminución significativa en cepas susceptibles. Actualmente, la norma ambiental de sernapesca incorpora la evaluación de RNAs codificantes como biomarcadores de resistencia en *Caligus* ante la exposición a Deltametrina. Entre estos biomarcadores se encuentran genes como el de tripsina-2, que exhibe un aumento de expresión ante la exposición, y tripsina-5 que muestra una reducción en su expresión ante la acción del fármaco (Sáez-Vera et al., 2022; Nuñez-Acuña et al., 2022). Por su parte, los transportadores ABCB también podrían ser buenos biomarcadores al presentar altos niveles de transcripción en poblaciones resistentes a Azametifos (Valenzuela-Muñoz, Chavez-Mardones, & Gallardo-Escárte, 2015; Valenzuela-Muñoz, Sturm, et al., 2015). Todos estos genes están involucrados en el metabolismo de xenobióticos y detoxificación, por lo que tienen relación directa con la generación de resistencia a los fármacos. Sin embargo, en este contexto, los lncRNAs también han sido vinculados con procesos biológicos relacionados con mecanismos de resistencia farmacológica en especies de artrópodos, como receptores del sistema nervioso y mecanismos de detoxificación (Etebari, Furlong, & Asgari, 2015). Debido a esto, los lncRNAs podrían ser buenos candidatos para evaluar la sensibilidad a los fármacos, además, al ser moléculas funcionales por sí mismas, como biomarcadores tienen ventaja por sobre los mRNAs, ya que estos solo son funcionales cuando se traducen a proteínas, por lo que en los ensayos de monitoreo sus niveles de expresión no siempre son congruentes con los niveles de proteína. En trabajos previos, se proponen como biomarcadores el gen Cytochrome p450, y los lncRNAs 8967 y 532 para el monitoreo de la resistencia a Azametifos, teniendo valores de expresión bajos en poblaciones susceptibles (Sáez-Vera et al., 2022). Por esto, los lncRNAs propuestos por este trabajo podrían ser de gran ayuda para el monitoreo de la sensibilidad a los fármacos en *Caligus rogercresseyi*.

Desde la perspectiva de la ecología, el impacto generado por los fármacos afecta directamente a las poblaciones de organismos que habitan las zonas expuestas,

pudiendo generar problemas como neurotoxicidad, genotoxicidad e inmunotoxicidad (Farag et al., 2021), lo cual desestabiliza los ecosistemas reduciendo los servicios ecosistémicos. Considerando todo lo mencionado en este trabajo y aquellos realizados previamente, no cabe duda la relevancia que tienen los ácidos nucleicos en los procesos biológicos, en especial los RNAs no codificantes, así como su respuesta ante la interacción a agentes tóxicos como los fármacos, por lo que es evidente la necesidad de buscar nuevas alternativas a su uso para tratar enfermedades y plagas en centros de cultivo acuícolas. También, a la hora de obtener conclusiones, siempre es necesario considerar que una gran cantidad de factores están involucrados en la expresión del genoma, entre ellos alteraciones genéticas, epigenéticas, bioquímicas, mecanismos de interacción RNA codificante/no codificante, entre otros (Cajal & Hümmer, 2018). Aunque aún es poco lo que se sabe sobre lncRNAs, hay evidencia creciente y robusta en múltiples estudios que respaldan ciertas características de estas moléculas, como precisamente su especificidad a tejidos, células, proteínas y funciones biológicas. En conclusión, es difícil explicar o analizar con precisión la respuesta del genotipo, pero múltiples herramientas como la bioinformática y la secuenciación masiva pueden ayudar a revelar y comprender la importancia de moléculas que antes eran poco relevantes o desconocidas. En la actualidad, los lncRNAs emergen como moléculas clave para el futuro de la biotecnología y el desarrollo de tratamientos para una amplia gama de enfermedades, pudiendo llegar a reemplazar el uso de fármacos que deterioran la vida y el ecosistema.

## 6. Conclusión:

Desde sus inicios hasta la actualidad, el estudio del transcriptoma se ha ido profundizando, tanto que en las últimas décadas ha tenido un aumento considerable. Esto debido principalmente al desarrollo de nuevas tecnologías de secuenciación como Nanopore e Illumina, y herramientas moleculares como la PCR y la retrotranscripción. Sin embargo, en la actualidad aún es mucho lo que se desconoce, dada la complejidad del transcriptoma y sus mecanismos de acción, incluyendo la gran cantidad de interacciones entre moléculas, que además son altamente específicas. No obstante, estas tecnologías han permitido avanzar en el entendimiento del rol que tienen moléculas como los long non-coding RNAs (lncRNAs) en múltiples procesos biológicos, así como también en la respuesta del transcriptoma ante estímulos ambientales y el impacto que tienen sobre el fenotipo.

En este sentido, las hipótesis que se plantea en este trabajo, en relación al uso de tecnologías de reads largos para la identificación de nuevos lncRNAs en *C. rogercresseyi* es aceptada, al registrarse una gran cantidad de nuevos lncRNAs antes no descritos por tecnologías de reads cortos, evidenciando características un tanto diferentes a las descritas previamente. De igual forma, también es aceptada la hipótesis planteada en relación a la expresión de lncRNAs en *C. rogercresseyi*, donde se identificaron patrones de expresión específicos asociados al tipo de fármaco y a las poblaciones expuestas. Por otro lado, se cumplen los objetivos propuestos, al comparar la caracterización de lncRNAs realizada mediante tecnologías de secuenciación de reads largos (Nanopore) y cortos (Illumina). Además, fue caracterizada la modulación transcriptómica de lncRNAs en poblaciones de *C. rogercresseyi* resistentes y susceptibles, registrándose una mayor expresión diferencial en cepas resistentes expuestas a Azametifos y Deltametrina, y en susceptibles expuestas a Cipermetrina. También, fueron asignadas funciones biológicas a los lncRNAs diferencialmente

modulados en poblaciones R/S de *Caligus*, asociándolos a funciones como regulación transcripcional, detoxificación, inmunidad y metabolismo. Finalmente, se realizó la validación molecular de los lncRNAs candidatos, sugiriendo nuevos marcadores de resistencia farmacológica en *Caligus*, pudiendo ser de utilidad en el desarrollo de estrategias para el monitoreo de la resistencia farmacológica.

## 7. Glosario

**Clusterización:** Agrupación de secuencias con un cierto grado de identidad entre sí bajo una única secuencia representativa del conjunto.

**Mapeado:** Identificación de secuencias de desde una base de datos de consulta contra una base de datos de referencia.

**Ontología génica:** Del inglés Gene Ontology (GO), hace referencia a la caracterización de los atributos del producto génico en cuanto a funciones, procesos y localización, entre otros.

**Organofosforados:** Compuesto orgánico degradable con enlaces fósforo-carbono capaz de inhibir enzimas con actividad acetilcolinesterasa, produciendo una acumulación de acetilcolina, lo que genera una alteración del impulso nervioso. Se utilizan para el control de plagas en reemplazo de hidrocarburos clorados.

**Piojo de mar:** Copépodo ectoparásito *Caligus rogercresseyi*.

**Pipeline:** También llamado workflow, es la secuencia de pasos y procesos realizados con un fin específico, en este caso, la identificación de lncRNAs.

**Piretroides:** Moléculas sintetizadas artificialmente que se utilizan generalmente para el control de insectos plaga en cosechas. Estos prolongan la corriente que fluye por los canales de sodio al ralentizar o mantener abierto el canal, alterando la transmisión del impulso nervioso.

**RNA largo no codificante:** del inglés long non-coding RNA (lncRNA), son transcritos asociados a múltiples procesos biológicos.

**Transcriptoma:** Porción del genoma que es expresada o transcrita a RNA, ya sea codificante o no codificante.

**Transcrito:** RNA derivado de un gen en transcripción.

**Trimmed:** Remoción de adaptadores utilizados en la secuenciación, estos son removidos para limpiar las secuencias.

## 8. Bibliografía:

- Aaen, S. M., Helgesen, K. O., Bakke, M. J., Kaur, K., & Horsberg, T. E. (2015). Drug resistance in sea lice: a threat to salmonid aquaculture. *Trends in parasitology*, 31(2), 72-81.
- Aguilar, L. R., Sánchez, G. G., Bocardo, O. I. L., & Pardo, J. P. (2018). La cinasa de residuos de serina y treonina blanco de la rapamicina (TORC1) es esencial en el metabolismo celular de los eucariontes. *Revista de Educación Bioquímica*, 37(1), 4-13.
- Boltana, S., Valenzuela-Miranda, D., Aguilar, A., Mackenzie, S., & Gallardo-Escarate, C. (2016). Long noncoding RNAs (lncRNAs) dynamics evidence immunomodulation during ISAV-Infected Atlantic salmon (*Salmo salar*). *Sci Rep*, 6, 22698. doi:10.1038/srep22698
- Borst, P., Evers, R., Kool, M., & Wijnholds, J. (1999). The multidrug resistance protein family. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1461(2), 347-357.
- Burgos, A. P., & Alamilla, J. (2010). El fosfatidilinositol-4, 5-bifosfato y sus acciones sobre los canales iónicos. *Revista Biomédica*, 21(2), 97-107.
- Buss, D. S., & Callaghan, A. (2008). Interaction of pesticides with p-glycoprotein and other ABC proteins: A survey of the possible importance to insecticide, herbicide and fungicide resistance. *Pesticide biochemistry and physiology*, 90(3), 141-153.
- Cajal, S. R. y., & Hümmer, S. (2018). BEYOND GENES: UNDERSTANDING THE FUNCTION OF NON CODING DNA. *Real Academia Nacional de Medicina de España*, 230-236.
- Carpenter, S., Aiello, D., Atianand, M. K., Ricci, E. P., Gandhi, P., Hall, L. L., . . . Fitzgerald, K. A. (2013). A Long Noncoding RNA Mediates Both Activation and Repression of Immune Response Genes. *Science*, 341(6147), 789-792. doi:10.1126/science.1240925
- Carvajal, J., Gonzalez, L., & George-Nascimento, M. (1998). Native sea lice (Copepoda: Caligidae) infestation of salmonids reared in netpen systems in southern Chile. *Aquaculture*, 166((3-4)), 241-246.
- Cavagnari, B. M. (2012). [Regulation of gene expression: how do epigenetic mechanisms work]. *Arch Argent Pediatr*, 110(2), 132-136. doi:10.1590/S0325-00752012000200009
- 10.5546/aap.2012.132
- Ceña, V., Galindo, M., Jordán, J., & Fuster Lluch, O. (2004). Las serina proteasas y su función en los procesos de muerte neuronal. *Rev. neurol.(Ed. impr.)*, 449-457.
- Chávez-Mardones, J., & Gallardo-Escárate, C. (2015). Next-generation transcriptome profiling of the salmon louse caligus rogercresseyi exposed to deltamethrin (alphamax™): Discovery of relevant genes and sex-related differences. *Marine Biotechnology*, 17(6), 793-810.

- Cole, S., Bhardwaj, G., Gerlach, J., Mackie, J., Grant, C., Almquist, K., . . . Deeley, R. G. (1992). Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science*, *258*(5088), 1650-1654.
- Collins, F. S., Lander, E. S., Rogers, J., Waterston, R. H., & Conso, I. H. G. S. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, *431*(7011), 931-945. doi:10.1038/nature03001
- Costello, M. J. (2006). Ecology of sea lice parasitic on farmed and wild fish. *Trends in Parasitology*, *22*(10), 475-483. doi:10.1016/j.pt.2006.08.006
- Del Arco, C. (2008). Metabolismo de los fármacos. *Farmacología humana*, *5*, 73-85.
- Deng, P., Liu, S., Nie, X., Weining, S., & Wu, L. (2018). Conservation analysis of long non-coding RNAs in plants. *Sci China Life Sci*, *61*(2), 190-198. doi:10.1007/s11427-017-9174-9
- Dermauw, W., & Van Leeuwen, T. (2014). The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect biochemistry and molecular biology*, *45*, 89-110.
- Dresdner, J., Chávez, C., Quiroga, M., Jiménez, D., Artacho, P., & Tello, A. (2019). Impact of Caligus treatments on unit costs of heterogeneous salmon farms in Chile. *Aquaculture Economics & Management*, *23*(1), 1-27.
- Enan, E., & Matsumura, F. (1992). Specific inhibition of calcineurin by type II synthetic pyrethroid insecticides. *Biochemical pharmacology*, *43*(8), 1777-1784.
- Etebari, K., Furlong, M. J., & Asgari, S. (2015). Genome wide discovery of long intergenic non-coding RNAs in Diamondback moth (*Plutella xylostella*) and their expression in insecticide resistant strains. *Scientific reports*, *5*(1), 1-14.
- Expósito\_Martínez, M. (2018). Sistema CRISPR/Cas y su aplicación a la edición dirigida de genomas.
- Farag, M. R., Alagawany, M., Bilal, R. M., Gewida, A. G. A., Dhama, K., Abdel-Latif, H. M. R., . . . Naiel, M. A. E. (2021). An Overview on the Potential Hazards of Pyrethroid Insecticides in Fish, with Special Emphasis on Cypermethrin Toxicity. *Animals*, *11*(7), 1880. Retrieved from <https://www.mdpi.com/2076-2615/11/7/1880>
- Flens, M. J., Zaman, G., van der Valk, P., Izquierdo, M. A., Schroeijers, A. B., Scheffer, G. L., . . . Scheper, R. J. (1996). Tissue distribution of the multidrug resistance protein. *The American journal of pathology*, *148*(4), 1237.
- Gallardo-Escárate, C., Valenzuela-Muñoz, V., & Nuñez-Acuña, G. (2014). RNA-Seq analysis using de novo transcriptome assembly as a reference for the salmon louse *Caligus rogercresseyi*. *Plos One*, *9*(4), e92239.
- Gottesman, M. M., & Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual review of biochemistry*, *62*(1), 385-427.

- Gueguen, Y., Mouzat, K., Ferrari, L., Tissandie, E., Lobaccaro, J., Batt, A., . . . Gourmelon, P. (2006). *Cytochromes P450: xenobiotic metabolism, regulation and clinical importance*. Paper presented at the Annales de biologie clinique.
- Heikema, A. P., Horst-Kreft, D., Boers, S. A., Jansen, R., Hiltemann, S. D., de Koning, W., . . . Bont, L. J. (2020). Comparison of Illumina versus nanopore 16S rRNA gene sequencing of the human nasal microbiota. *Genes*, *11*(9), 1105.
- Helgesen, K., Bravo, S., Sevatdal, S., Mendoza, J., & Horsberg, T. (2014). Deltamethrin resistance in the sea louse *C. aligus rogercresseyi* (B oxford and B ravo) in Chile: bioassay results and usage data for antiparasitic agents with references to Norwegian conditions. *Journal of fish diseases*, *37*(10), 877-890.
- Imamura, K., & Akimitsu, N. (2014). Long Non-Coding RNAs Involved in Immune Responses. *Front Immunol*, *5*, 573. doi:10.3389/fimmu.2014.00573
- Jiang, L., Liu, W., Zhu, A., Zhang, J., Zhou, J., & Wu, C. (2016). Transcriptome analysis demonstrate widespread differential expression of long noncoding RNAs involve in *Larimichthys crocea* immune response. *Fish Shellfish Immunol*, *51*, 1-8. doi:10.1016/j.fsi.2016.02.001
- Johnson, S. C., Treasurer, J. W., Bravo, S., Nagasawa, K., & Kabata, Z. (2004). A review of the impact of parasitic copepods on marine aquaculture. *Zoological Studies*, *43*(2), 229-243. Retrieved from <Go to ISI>://WOS:000222853500010
- Jones, M., Sommerville, C., & Wootten, R. (1992). Reduced sensitivity of the salmon louse, *Lepeophtheirus salmonis*, to the organophosphate dichlorvos. *Journal of fish diseases*, *15*(2), 197-202.
- Kircher, M., Heyn, P., & Kelso, J. (2011). Addressing challenges in the production and analysis of illumina sequencing data. *BMC genomics*, *12*(1), 1-14.
- Laressergues, D., Couzigou, J. M., Clemente, H. S., Martinez, Y., Dunand, C., Becard, G., & Combiere, J. P. (2015). Primary transcripts of microRNAs encode regulatory peptides. *Nature*, *520*(7545), 90-U205. doi:10.1038/nature14346
- LEÓN, H. R. P. (2016). DISTRIBUCIÓN TISULAR DE PROTEÍNAS DE RESISTENCIA MÚLTIPLE A FÁRMACOS EN CALIGUS ROGERCRESSEYI.
- Li, L. J., Leng, R. X., Fan, Y. G., Pan, H. F., & Ye, D. Q. (2017). Translation of noncoding RNAs: Focus on lncRNAs, pri-miRNAs, and circRNAs. *Experimental Cell Research*, *361*(1), 1-8. doi:10.1016/j.yexcr.2017.10.010
- Magny, E. G., Pueyo, J. I., Pearl, F. M. G., Cespedes, M. A., Niven, J. E., Bishop, S. A., & Couso, J. P. (2013). Conserved Regulation of Cardiac Calcium Uptake by Peptides Encoded in Small Open Reading Frames. *Science*, *341*(6150), 1116-1120. doi:10.1126/science.1238802

- Mattick, J. S. (2011). Challenging the dogma: the central role of RNA in human development and cognition. *Febs Journal*, 278, 2-2. Retrieved from <Go to ISI>://WOS:000292333100007
- Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*, 10(3), 155-159. doi:10.1038/nrg2521
- Murillo-Maldonado, J. M., & Riesgo-Escovar, J. R. (2019). The various and shared roles of lncRNAs during development. *Developmental Dynamics*, 248(11), 1059-1069. doi:10.1002/dvdy.108
- Núñez-Acuña, G., Sáez-Vera, C., Valenzuela-Muñoz, V., Valenzuela-Miranda, D., Arriagada, G., & Gallardo-Escárate, C. (2020). Tackling the molecular drug sensitivity in the sea louse *Caligus rogercresseyi* based on mRNA and lncRNA interactions. *Genes*, 11(8), 857.
- Núñez-Acuña, G., Valenzuela-Muñoz, V., Valenzuela-Miranda, D., & Gallardo-Escárate, C. (2021). Comprehensive Transcriptome Analyses in Sea Louse Reveal Novel Delousing Drug Responses Through MicroRNA regulation. *Marine Biotechnology*, 23(5), 710-723.
- Núñez-Acuña, G., Sáez-Vera, C., Valenzuela-Muñoz, V., Valenzuela-Miranda, D., Gallardo-Escárate, C. (2022). NOVEL GENOMIC MARKERS ASSOCIATED WITH PHARMACOLOGICAL RESISTANCE IDENTIFIED THROUGH WHOLE-GENOME RESEQUENCING.
- Parsons, A. E., Escobar-Lux, R. H., Sævik, P. N., Samuelsen, O. B., & Agnalt, A.-L. (2020). The impact of anti-sea lice pesticides, azamethiphos and deltamethrin, on European lobster (*Homarus gammarus*) larvae in the Norwegian marine environment. *Environmental Pollution*, 264, 114725.
- Riascos-Arcos, J. J., Espitia Navarro, H. F., & López Gerena, J. (2015). Evaluación de las herramientas de secuenciación masiva (NGS) para identificar genes asociados con tolerancia al estrés hídrico en caña de azúcar. *Acta Agronómica*, 64(4), 355-362.
- Rogoyski, O. M., Pueyo, J. I., Couso, J. P., & Newbury, S. F. (2017). Functions of long non-coding RNAs in human disease and their conservation in *Drosophila* development. *Biochemical Society Transactions*, 45, 895-904. doi:10.1042/Bst20160428
- Rojas, A. P. M., & Fortini, Y. V. (2020). Glutaric aciduria type-1. *Revista Cubana de Medicina Intensiva y Emergencias*, 19(4), 1-9.
- Rollano Peñaloza, O. M., & Mollinedo Portugal, P. (2017). Análisis bioinformático de Arn-Seq con una perspectiva para Bolivia. *Revista Boliviana de Química*, 34(2), 50-55.
- Sáez-Vera, C., Núñez-Acuña, G., & Gallardo-Escárate, C. (2022). Sensitivity assessment to azamethiphos by time-to-response bioassay and biomarkers in the sea louse *Caligus rogercresseyi*. *Aquaculture*, 546, 737340.
- Schmitz, S. U., Grote, P., & Herrmann, B. G. (2016). Mechanisms of long noncoding RNA function in development and disease. *Cellular and*

- Molecular Life Sciences*, 73(13), 2491-2509. doi:10.1007/s00018-016-2174-5
- Sevatdal, S., Copley, L., Wallace, C., Jackson, D., & Horsberg, T. E. (2005). Monitoring of the sensitivity of sea lice (*Lepeophtheirus salmonis*) to pyrethroids in Norway, Ireland and Scotland using bioassays and probit modelling. *Aquaculture*, 244(1-4), 19-27.
- Shepherd, F., & Carney, D. (2000). Treatment of NSCLC: chemotherapy. *Textbook of Lung Cancer. Editado por Heine H Hansen. Martín Dunitz Ltd, London*, 213-242.
- Stride, B. D., Valdimarsson, G., Gerlach, J. H., Wilson, G. M., Cole, S., & Deeley, R. G. (1996). Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter. *Molecular pharmacology*, 49(6), 962-971.
- Sturm, A., Cunningham, P., & Dean, M. (2009). The ABC transporter gene family of *Daphnia pulex*. *BMC genomics*, 10(1), 1-18.
- Tarifeno-Saldivia, E., Valenzuela-Miranda, D., & Gallardo-Escarate, C. (2017). In the shadow: The emerging role of long non-coding RNAs in the immune response of Atlantic salmon. *Developmental and Comparative Immunology*, 73, 193-205. doi:10.1016/j.dci.2017.03.024
- Taylor, D. H., Chu, E. T. J., Spektor, R., & Soloway, P. D. (2015). Long Non-Coding RNA Regulation of Reproduction and Development. *Molecular Reproduction and Development*, 82(12), 932-956. doi:10.1002/mrd.22581
- Tinjacá Espinel, C. P. (1995). Caracterización parcial de la calmodulina de *Plasmodium falciparum*.
- Tosco Herrera, E. (2019). Análisis de datos en Biocomputación: tecnología MinION y herramientas bioinformáticas.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., . . . Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, 7(3), 562-578.
- Trovero, M. F., & Geisinger, A. (2019). Los ARNs no codificantes largos y su vinculación con las patologías testiculares *In Anales de la Facultad de Medicina*, 6, 12-47. doi:10.25184/anfamed2019v6n1a8
- Valenzuela-Miranda, D., Etebari, K., Asgari, S., & Gallardo-Escárate, C. (2017). Long noncoding RNAs: Unexplored players in the drug response of the sea louse *Caligus rogercresseyi*. *Agri Gene*, 4, 1-7. doi:10.1016/j.aggene.2017.03.001
- Valenzuela-Munoz, V., Valenzuela-Miranda, D., & Gallardo-Escarate, C. (2018). Comparative analysis of long non-coding RNAs in Atlantic and Coho salmon reveals divergent transcriptome responses associated with immunity and tissue repair during sea lice infestation. *Dev Comp Immunol*, 87, 36-50. doi:10.1016/j.dci.2018.05.016
- Valenzuela-Muñoz, V., Chavez-Mardones, J., & Gallardo-Escárate, C. (2015). RNA-seq analysis evidences multiple gene responses in *Caligus*

- rogercresseyi exposed to the anti-salmon lice drug azamethiphos. *Aquaculture*, 446, 156-166.
- Valenzuela-Muñoz, V., Sturm, A., & Gallardo-Escárate, C. (2015). Transcriptomic insights on the ABC transporter gene family in the salmon louse *Caligus rogercresseyi*. *Parasites & vectors*, 8(1), 1-14.
- Vera Bizama, F. O. (2017). Identificación y evaluación de genes asociados al sistema profenoloxidasa de *Caligus rogercresseyi* y su respuesta transcriptómica frente a los antiparasitarios deltametrina, cipermetrina y azametifos.
- Verrelle, P., Frédéric, M., Fonck, Y., Feillel, V., Dionet, C., Kwiatkowski, F., . . . Chassagne, J. (1991). Clinical relevance of immunohistochemical detection of multidrug resistance P-glycoprotein in breast carcinoma. *JNCI: Journal of the National Cancer Institute*, 83(2), 111-116.
- Wang, Y., Li, Y., Wang, Q., Lv, Y., Wang, S., Chen, X., . . . Li, X. (2014). Computational identification of human long intergenic non-coding RNAs using a GA-SVM algorithm. *Gene*, 533(1), 94-99.
- Zaman, G., Lankelma, J., vAN Tellingen, O., Beijnen, J., Dekker, H., Paulusma, C., . . . Borst, P. (1995). Role of glutathione in the export of compounds from cells by the multidrug-resistance-associated protein. *Proceedings of the National Academy of Sciences*, 92(17), 7690-7694.
- Zhao, X. Y., & Lin, J. D. (2015). Long Noncoding RNAs: A New Regulatory Code in Metabolic Control. *Trends in Biochemical Sciences*, 40(10), 586-596. doi:10.1016/j.tibs.2015.08.002