



Universidad de Concepción  
Dirección de Postgrado  
Facultad de Ciencias Biológicas - Programa de Magíster en Bioquímica y  
Bioinformática

**“Evaluación funcional de miRNA implicados en el proceso infeccioso y reproductivo de *Caligus rogercresseyi*”**



Tesis para optar al grado de Magíster en Bioquímica y Bioinformática

CRISLERI DAYAN CARRERA NAIPIL  
CONCEPCIÓN-CHILE  
2019

Profesor Guía: Dr. Cristian Gallardo Escárte  
Dpto. de Oceanografía, Facultad de ciencias naturales y oceanográficas  
Centro interdisciplinario para la investigación acuícola (INCAR)  
Profesor Co-Guía: Dra. Violeta Morín  
Dpto. de Bioquímica y Biología Molecular  
Facultad de Ciencias Biológicas  
Universidad de Concepción

Esta tesis ha sido realizada en el Laboratorio de biotecnología genómica acuícola del Centro de Biotecnología, Universidad de Concepción.

Profesor Guía

---

Dr. Cristian Gallardo Escárate  
Facultad de Ciencias Biológicas  
Universidad de Concepción

Profesor Co-Guía

---

Dra. Violeta Morin Muñoz  
Facultad de Ciencias Biológicas  
Universidad de Concepción

Comisión Evaluadora:



---

Dr. José Leonardo Guzman González  
Facultad de Ciencias Biológicas  
Universidad de Concepción

---

Dra. Valentina Valenzuela  
Centro Interdisciplinario para la investigación  
Acuícola  
Universidad de Concepción

Director de Programa

---

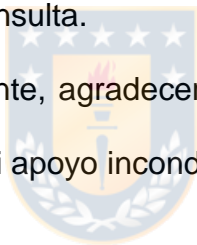
Dra. Elena Amparo Uribe  
Facultad de Ciencias Biológicas  
Universidad de Concepción

## **AGRADECIEMITOS**

Quiero agradecer en primer lugar a Becas CONICYT y al centro Interdisciplinario INCAR por proveer el apoyo de financiamiento para el desarrollo de la tesis.

En segundo lugar, a mis profesores guías Dr. Cristian Gallardo y la Dra. Violeta Morin por el apoyo durante este periodo y por darme la oportunidad de pertenecer a su laboratorio. También quiero agradecer al equipo del laboratorio principalmente a Andrea, Rayen y Paulina por todo su apoyo, gracias por sus palabras de aliento y los buenos momentos compartidos. También quisiera agradecer a Dra. Estefania Tarifeño por ayudarme a resolver mis dudas durante mi trabajo y por estar de manera excepcional para cada consulta.

Por ultimo y lo más importante, agradecer a mis padres y hermanos y a mi familia Alfredo y Emma, por ser mi apoyo incondicional y mi máxima motivación.



## TABLA DE CONTENIDOS

<b>LISTA DE TABLAS.....</b>	<b>VII</b>
<b>LISTA DE FIGURAS.....</b>	<b>VII</b>
<b>RESUMEN.....</b>	<b>IX</b>
<b>ABSTRACT.....</b>	<b>XI</b>
<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>1. CALIGUS ROGERCRESSEYI.....</b>	<b>1</b>
<b>2. PROCESO INFECTIVO .....</b>	<b>3</b>
2.1 SERPINAS .....	5
2.2 CATEPSINA .....	7
2.3 TRIPSINA.....	9
<b>3. REPRODUCCIÓN EN CALIGUS ROGERCRESSEYI.....</b>	<b>12</b>
3.1 VITELÓGENINA.....	12
<b>4. MICRORNA Y BIOGENESIS.....</b>	<b>15</b>
<b>HIPÓTESIS DE TRABAJO.....</b>	<b>19</b>
<b>OBJETIVOS.....</b>	<b>19</b>
OBJETIVO GENERAL .....	19
<b>OBJETIVOS ESPECÍFICOS.....</b>	<b>19</b>
<b>METODOLOGÍA .....</b>	<b>20</b>
1. IDENTIFICACIÓN DE GENES ASOCIADOS A PROCESO INFECTIVO Y REPRODUCCIÓN EN <i>CALIGUS ROGERCRESSEYI</i> .....	20
1.1 ALINEAMIENTO DE SECUENCIAS AMINOACÍDICAS Y ANÁLISIS FILOGENÉTICO .....	21
1.2 SECUENCIACIÓN ILLUMINA DE SMALL RNAs. ....	21
1.3 IDENTIFICACIÓN DE MIRNAS EN <i>C. ROGERCRESSEYI</i> . ....	22
1.4 PREDICCIÓN DE GENES DIRIGIDOS POR MIRNAS DE <i>C. ROGERCRESSEYI</i> . ....	23
2. EVALUACIÓN DE LOS NIVELES DE EXPRESIÓN DE MIRNAS CANDIDATOS Y SUS GENES TARGET DURANTE LOS DIFERENTES ESTADOS DE DESARROLLO DE <i>C. ROGERCRESSEYI</i> .....	24
2.1 EXTRACCIÓN DE RNA Y SÍNTESIS DE ADNC.....	24
2.2 ANÁLISIS DE EXPRESIÓN RELATIVA POR RT-PCR PARA GENES TARGET.....	25
2.3 HOUSEKEEPING PARA MIRNA DE <i>CALIGUS ROGERCRESSEYI</i> .....	25
2.4 ANÁLISIS DE EXPRESIÓN DE MIRNA.....	26

3. VALIDACIÓN A NIVEL FUNCIONAL LA INTERACCIÓN DE MRNA/MIRNAS CANDIDATOS RELACIONADOS AL PROCESO INFECTIVO Y REPRODUCTIVO EN <i>C. ROGERCRESSEYI</i> .....	27
3.1 AMPLIFICACIÓN DEL ORF PARCIAL PRÓXIMO AL 3'UTR DE VITELOGENINA I.....	28
3.2 PURIFICACIÓN DE FRAGMENTOS DE DNA MEDIANTE GELES DE AGAROSA.....	29
3.3 CLONAMIENTO Y PURIFICACIÓN DE PLÁSMIDO RECOMBINANTE.....	31
3.4 SECUENCIACIÓN DE PLÁSMIDOS .....	31
3.5 AMPLIFICACIÓN DEL 3'UTR POR PCR .....	32
3.6 SUBCLONAMIENTO DE LA SECUENCIA DEL 3'UTR QUE CONTIENE EL SITIO DE UNIÓN AL MICRORNA EN EL VECTOR PMIRGLO. ....	33
3.7 VALIDACIÓN FUNCIONAL DE MIRNA, ENSAYO DE LUCIFERASA.....	35
3.8 LIPOFECCIÓN DE CÉLULAS HELA .....	36
3.9 CULTIVO DE CÉLULAS HELA .....	37
<b>RESULTADOS.....</b>	<b>39</b>
1. IDENTIFICACIÓN DE PROTEASAS A PARTIR DE DATOS DE SECUENCIACIÓN MASIVA DE <i>C. ROGERCRESSEYI</i> .....	39
1.1 CATEPSINA.....	39
<i>Análisis filogenético</i> .....	40
1.2 SERPINA.....	43
<i>Análisis filogenético</i> .....	44
1.3 TRIPSINA.....	47
<i>Análisis filogenético</i> .....	48
1.4 VITELOGENINA I.....	50
1.4.1 Clonamiento y secuenciación del extremo 3' de la ORF y RACE de Vitelogenina I.....	51
1.5 PREDICCIÓN DE TARGET PARA MIRNA.....	57
2. EVALUACIÓN DE LOS NIVELES DE EXPRESIÓN DE MIRNAS CANDIDATOS Y SUS GENES TARGET DURANTE LOS DIFERENTES ESTADOS DE DESARROLLO DE <i>C. ROGERCRESSEYI</i> .....	59
2.1 HOUSEKEEPING PARA MIRNA DE <i>CALIGUS ROGERCRESSEYI</i> .....	59
2.2 ANÁLISIS DE EXPRESIÓN RELATIVA POR RT-PCR PARA BANTAM Y MIR-996 EN LOS DIFERENTES ESTADOS DE DESARROLLO DE <i>CALIGUS ROGERCRESSEYI</i> .....	61
2.3 ANÁLISIS DE EXPRESIÓN RELATIVA POR RT-PCR PARA CATEPSINA, SERPINA, TRIPSINA Y VITELOGENINA I EN LOS DIFERENTES ESTADOS DE DESARROLLO DE <i>C. ROGERCRESSEYI</i> .....	62
3. VALIDACIÓN A NIVEL FUNCIONAL LA INTERACCIÓN DE MRNA/MIRNAS CANDIDATOS RELACIONADOS AL PROCESO INFECTIVO Y REPRODUCTIVO EN <i>C. ROGERCRESSEYI</i> .....	66
3.1 CLONAMIENTO REGIONES 3'UTR Y SECUENCIACIÓN DE GENES.....	66
3.2. SUBCLONAMIENTO DE LAS REGIONES 3'UTR EN EL VECTOR PMIRGLO.....	71
3.3 ENSAYO LUCIFERASA.....	75

<b>DISCUSIÒN.....</b>	<b>78</b>
CATEPSINA.....	78
SERPINAS.....	83
TRIPSINA.....	85
VITELOGENINA I.....	88
<b>CONCLUSIÒNES.....</b>	<b>91</b>
<b>BIBLIOGRAFIA.....</b>	<b>93</b>
<b>MATERIAL SUPLEMENTARIO.....</b>	<b>98</b>



## LISTA DE TABLAS

<b>Tabla 1.</b> Partidores empleados en los análisis de expresión relativa para los genes de estudio.....	26
<b>Tabla 2.</b> Partidores empleados en los análisis de expresión relativa para los miRNA de estudio.....	27
<b>Tabla 3.</b> Partidores empleados en el clonamiento de la ORF parcial próxima al 3'UTR y RACE para Vitelogenina I.....	28
<b>Tabla 3 (Continuación).</b> Partidores empleados en el clonamiento de la ORF parcial próxima al 3'UTR y RACE para Vitelogenina I.....	29
<b>Tabla 4.</b> Partidores empleados para el clonamiento del cDNA de Cr_Catepsina, Cr_Serpina y Cr_Tripsina y Cr_Vitelogenina I de <i>C. rogercresseyi</i> que contienen el sitio de unión a Bantam y mir-996.....	34
<b>Tabla 5.</b> Predicción de target para los miRNA bantam y mir-996.....	58
<b>Tabla 6.</b> Resumen de los datos originales para RPKM y Conteo de la base de datos de miRNA para <i>Caligus rogercresseyi</i> .....	60
<b>Tabla 7.</b> Análisis del set de datos para valores de Conteo y RPKM para los miRNA de <i>Caligus rogercresseyi</i> .....	60
<b>Tabla 8.</b> Resultados entregados por NormFinder para la identificación de genes de normalización para miRNA de <i>Caligus rogercresseyi</i> .....	61

## LISTA DE FIGURAS

<b>Figura 1.</b> Ciclo de vida de <i>C. rogercresseyi</i> . Diferentes estados de desarrollo del parásito incluyendo las etapas de vida libre y las parasíticas.....	2
<b>Figura 2.</b> Representación esquemática de la regulación transcripcional regulado por miRNA en <i>C. rogercresseyi</i> .....	17
<b>Figura 3.</b> Diseño experimental para la obtención del gen Vitelogenina I y el 3'UTR mediante RACE.....	30
<b>Figura 4.</b> Plásmidos utilizados en el diseño experimental.....	33
<b>Figura 5.</b> Modulación de la función de miRNA.....	36
<b>Figura 6.</b> ARN mensajero de Cr_Catepsina de <i>C. rogercresseyi</i> .....	41
<b>Figura 7.</b> Alineamiento de la proteína Cr_Catepsina L de <i>C. rogercresseyi</i> .....	42
<b>Figura 8.</b> Relación filogenética de Cr_Catepsina de <i>C. rogercresseyi</i> .....	43
<b>Figura 9.</b> ARN mensajero de Cr_Serpina de <i>C. rogercresseyi</i> .....	45
<b>Figura 10.</b> Alineamiento de Cr_Serpina de <i>C. rogercresseyi</i> .....	46
<b>Figura 11.</b> Relación filogenética de Cr_Serpina de <i>C. rogercresseyi</i> .....	47
<b>Figura 12.</b> ARN mensajero de Cr_Tripsina de <i>C. rogercresseyi</i> .....	48
<b>Figura 13.</b> Alineamiento de Cr_Tripsina de <i>C. rogercresseyi</i> .....	49
<b>Figura 14.</b> Relación filogenética de Cr_Tripsina de <i>C. rogercresseyi</i> .....	50
<b>Figura 15.</b> Fragmentos secuenciados de la ORF parcial de Cr_Vitelogenina I de <i>C. rogercresseyi</i> clonados en el vector pGEM-T.....	52

<b>Figura 16.</b> Ensamble obtenido a partir de los cuatro fragmentos clonados y secuenciados para la ORF de Cr_Vitelogenina I. ....	53
<b>Figura 17.</b> Representación esquemática de los resultados obtenidos por secuenciación de los fragmentos del gen de Cr_Vitelogenina I de <i>C. rogercresseyi</i> .....	53
<b>Figura 18.</b> Contig obtenido mediante clonamiento y secuenciación para Cr_Vitelogenina I de <i>C. rogercresseyi</i> .....	54
<b>Figura 19.</b> Alineamiento de Cr_Vitelogenina I (ORF parcial) de <i>C. rogercresseyi</i> . ....	55
<b>Figura 20.</b> Relación filogenética de Cr_Vitelogenina I parcial de <i>C. rogercresseyi</i> ..	56
<b>Figura 21.</b> Sitio de unión predichos para los miRNA bantam y mir-996 en los extremos 3'UTR de catepsina, serpina, tripsina y vitelogenina I de <i>C. rogercresseyi</i> . ....	58
<b>Figura 22.</b> Valores de Ct para los candidatos a housekeeping para miRNA de <i>C. rogercresseyi</i> .....	62
<b>Figura 23.</b> Valores de Ct para miRNA de <i>C. rogercresseyi</i> relacionados a infección y reproducción.....	63
<b>Figura 24.</b> Expresión relativa de miRNA Bantam y mir-996 en los diferentes estados de desarrollo de <i>C. rogercresseyi</i> . ....	64
<b>Figura 25.</b> Expresión relativa de proteínas del secretoma en los diferentes estados de desarrollo de <i>C. rogercresseyi</i> . ....	65
<b>Figura 26.</b> Expresión relativa para Cr_Vitelogenina I en los diferentes estados de desarrollo de <i>C. rogercresseyi</i> .....	66
<b>Figura 27.</b> Resultados del clonamiento de los insertos de mRNA 3'UTR de catepsina, serpina y tripsina.. ....	68
<b>Figura 28.</b> Resultados del clonamiento de los insertos de mRNA 3'UTR de Cr_Vitelogenina I. ....	69
<b>Figura 29.</b> Inserto del mRNA clonado en el vector pGEM-T easy.....	70
<b>Figura 30.</b> Alineamiento de la secuencia del extremo 3'UTR de Cr_Serpina por secuenciación masiva y el fragmento clonado en el vector pGEM-T. ....	71
<b>Figura 31.</b> Resultados de clonamiento de los insertos del mRNA 3'UTR de catepsina, serpina y tripsina en el vector pmirGLO.....	72
<b>Figura 32.</b> Resultados del clonamiento del inserto del mRNA 3'UTR de Cr_Vitelogenina I en el vector pmirGLO.....	73
<b>Figura 33.</b> Alineamiento de la secuencia del extremo 3'UTR de Cr_Serpina clonado en el vector pGEM-T y el fragmento subclonado en el vector pmirGLO.....	74
<b>Figura 34.</b> Regulación de la expresión génica por miRNA mediante la unión directa al 3'UTR de la transcripción diana.....	74
<b>Figura 35.</b> Niveles de luminiscencia, ensayo de luciferasa en la línea celular HeLa.....	76



## RESUMEN

*Caligus rogercresseyi* es uno de los ectoparásitos más prevalentes en la industria del salmón en Chile. Uno de los aspectos más importantes de un patógeno es una capacidad de infección y reproducción. Existe evidencia que proteínas del secretoma como serpinas, tripsinas y catepsinas juegan un rol importante en la infección de un patógeno y en la interacción patógeno-hospedero. Así como otras proteínas son de importancia durante los procesos reproductivos como vitelogenina. Los avances en genómica funcional ha permitido la caracterización de genomas en especies no-modelos y específicamente de su fracción transcrita. En este punto, RNAs no-codificantes han emergido como los principales componentes moleculares con roles regulatorios de la expresión génica. El objetivo de este estudio fue determinar la modulación de miRNAs sobre genes de relevancia en el proceso infectivo y reproductivo de caligus. Esto se llevó a cabo mediante el estudio de predicción de interacciones entre serpina, catepsina y tripsina con el miRNA Bantam y para vitelogenina I con mir-996. Se realizó una búsqueda de secuencias en librería de cDNA específica para caligus, utilizando herramientas bioinformáticas. Para predecir los genes dirigidos por miRNAs se utilizaron los algoritmos: PITA, StarMir, RNAHyd. La expresión diferencial entre estadios de desarrollo se realizó por qPCR. Las secuencias 3'UTR diana para la unión al miRNA fueron clonadas y secuenciadas en el vector pmirGLO. A partir de esta construcción se realizó el ensayo funcional para medir los niveles de unión del miRNAs con su secuencia diana mediante la inhibición de la expresión del gen reportero. En la evaluación de expresión por qPCR se observó para Cr\_Catepsina, Cr\_Tripsina y Cr\_Vitelogenina, una mayor expresión en hembra, mientras que Cr\_Serpina se encuentra altamente

expresado en estado de nauplius. Según los análisis de predicción Bantam se uniría al mRNA de catepsina, serpina y tripsina con una energía de hibridación de -15,6, -22,7 y -20,7 kcal/mol respectivamente. Para vitelogenina I fue de -10,9 kcal/mol para mir-996. Estos resultados fueron verificados *in vitro* en el cual no se observó ningún efecto de bantam en la actividad de la luciferasa para Cr\_Serpina y Cr\_Catepsina mientras que para tripsina bantam logra disminuir la actividad de la luciferasa en un 50%. Lo mismo ocurrió para vitelogenina I en el cual se observó una disminución en un 30%. De acuerdo a estos resultados se puede concluir que bantam estaría regulando de manera post-transcripcional proteínas del secretoma de *C. rogercresseyi*, las cuales podrían tener una función en el proceso infectivo. Además, mir-996 también podría estar participando en la regulación post-transcripcional de vitelogenina I, un gen clave en la reproducción de *C. rogercresseyi*.



## ABSTRACT

*Caligus rogercresseyi* is one of the most prevalent ectoparasites in the salmon industry in Chile. One of the most important aspects of a pathogen is a capacity for infection and reproduction. There is evidence that secretome proteins such as serpins, trypsins and cathepsins play an important role in the infection of a pathogen and in the pathogen-host interaction. As well as other proteins are important during reproductive processes such as vitellogenin. Advances in functional genomics have allowed the characterization of genomes in non-model species and specifically their transcribed fraction. At this point, non-coding RNAs have emerged as the main molecular components with regulatory roles of gene expression. The objective of this study was to determine the modulation of miRNAs on relevant genes in the caligus infective and reproductive process. This was carried out by means of the prediction study of interactions between serpin, cathepsin and trypsin with the Bantam miRNA and for vitellogenin I with mir-996. A sequence search was performed on a specific cDNA library for caligus, using bioinformatics tools. To predict the genes directed by miRNAs, the algorithms were used: PITA, StarMir, RNAHyd. The differential expression between stages of development was made by qPCR. The 3'UTR target sequences for miRNA binding were cloned and sequenced in the pmirGLO vector. From this construction, the functional assay was performed to measure the binding levels of the miRNAs with their target sequence by inhibiting the expression of the reporter gene. In the evaluation of expression by qPCR was observed for Cr\_Catepsina, Cr\_Tripsina and Cr\_Vitellogenina, a greater expression in female, whereas Cr\_Serpina is highly expressed in nauplius state. According to the

prediction analysis, Bantam would bind to the mRNA of cathepsin, serpin and trypsin with a hybridization energy of -15.6, -22.7 and -20.7 kcal / mol respectively. For vitellogenin I it was -10.9 kcal / mol for mir-996. These results were verified in vitro in which no effect of bantam on the activity of luciferase was observed for Cr\_Serpina and Cr\_Catepsina whereas for trypsin bantam manages to decrease the activity of luciferase by 50%. The same happened for vitellogenin I in which a decrease of 30% was observed. According to these results, it can be concluded that bantam would be regulating post-transcriptional proteins of the secret of *C. rogercresseyi*, which could have a role in the infective process. In addition, mir-996 could also be involved in the post-transcriptional regulation of vitellogenin I, a key gene in the reproduction of *C. rogercresseyi*.



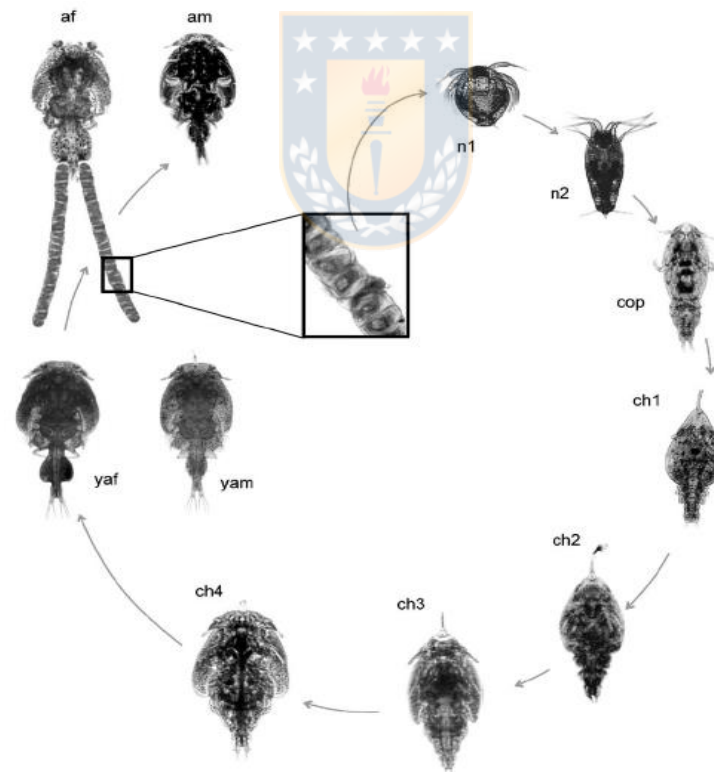
## INTRODUCCIÓN

### 1. *CALIGUS ROGERCRESSEYI*

*Caligus rogercresseyi* es uno de los patógenos más importantes y de mayor impacto económico que afecta a la industria del salmón en Chile, afectando principalmente a especies como el salmón del Atlántico (*Salmo salar*) y la trucha arcoíris (*Oncorhynchus mykiss*) (Bravo et al., 2013). El ciclo de vida de *C. rogercresseyi* se divide en ocho estados de desarrollo de los cuales tres corresponden a una fase planctónica, con dos estados nauplius y un estado copepodito, los otros cinco estados corresponden a la denominada fase adherida al hospedador como parásito, con cuatro estados de chalimus (I-IV), terminando con la etapa de machos y hembras maduros (González & Carvajal, 2003; Rozas, M. & Asencio, 2006; Rozas, M & Asencio, 2007). Los daños provocados por caligus están relacionados con la pérdida de calidad del producto final, un crecimiento retardado de los peces parasitados, un incremento de la susceptibilidad frente a otros patógenos, además de los costos generados por los tratamientos (Pino-Marambio et al., 2007). Este ectoparásito se alimenta de mucus y epidermis, generando en el pez la pérdida de las funciones físicas y protección microbiana y un aumento del estrés del hospedador (Pino-Marambio et al., 2007). La infección por caligus genera en el pez un estrés crónico, pérdida del apetito, pérdida del valor comercial por las lesiones de piel y musculatura (González & Carvajal, 2003). El ectoparásito incrementa los costos de producción en cerca de US \$ 480 millones al año en todo el mundo (Costello, 2009). Las estrategias para controlar las infestaciones de piojos en la acuicultura del salmón incluyen actualmente la administración de productos

químicos o medicamentos a través de baños o dietas formuladas (Igboeli et al., 2013), sin embargo, la sobreexposición a estos agentes químicos tiende a promover la resistencia a los fármacos en poblaciones silvestres de parásitos (Daborn & Le Goff, 2004).

Dado las problemáticas asociadas a *C. rogercresseyi* es necesario encontrar otras alternativas al control del patógeno, existiendo la necesidad de comprender parámetros importantes del piojo como lo es su ciclo de vida. Es por este motivo que este estudio se enfocó en su capacidad infectiva y reproductiva. Ambas etapas en el ciclo de vida del patógeno son esenciales para llevar con éxito su infección al huésped y su permanencia y reproducción en el hospedero.



**Figura 1.** Ciclo de vida de *C. rogercresseyi*. Diferentes estados de desarrollo del parásito incluyendo las etapas de vida libre y las parasíticas. af: Hembra, am: Machos, n1: Nauplius I, n2: Nauplius 2, cop: Copepodito, Ch1: Chalimus I, Ch2: Chalimus II, Ch3: Chalimus III, Ch4: Chalimus IV, yaf: Preadulto hembra, yam: Preadulto macho (González & Carvajal, 2003).

## 2. PROCESO INFECTIVO

Estudios recientes en análisis del secretoma revelaron que proteínas secretadas juegan un rol importante en la infección de un patógeno y en la interacción patógeno-hospedero (Ranganathan & Garg, 2009b). Se define como infección a la invasión de un anfitrión por un microorganismo patógeno, su multiplicación en los tejidos y la respuesta del huésped a su presencia. Durante el proceso infectivo de patógenos las proteasas, constituyen uno de los principales compuestos secretados. Los cuales, son utilizadas por el parásito con el fin de facilitar los procesos de alimentación, mantención de la homeostasis y la modulación de la respuesta inmune del huésped. Las proteínas excretoras/secretoras (ES) de los patógenos cambian el entorno de la célula huésped al suprimir el sistema inmune, para ayudar a la proliferación de la infección (Ranganathan & Garg, 2009b).

En *Fasciola hepática* han integrado un enfoque transcriptómico y proteómico de proteínas secretoras que están involucradas en las interacciones huésped-patógeno y correlacionan los cambios en su expresión con la migración del parásito (Robinson 2009). Dentro de los principales componentes de las secreciones de los parásitos adultos se encuentran las enzimas proteolíticas incluyendo catepsina L, catepsina B, proteasas de serina y carboxipeptidasas del tipo tripsina (Robinson et al., 2009). En *Ancylostoma caninum*, un parasito intestinal que se alimenta de sangre de mamíferos, se cree que proteínas del secretoma tienen un alto grado de participación en la interacción parásito y huésped. Se han identificado 105 proteínas denominadas ES en la etapa de alimentación del parásito, dentro de las cuales

destacan lectinas del tipo C, galactinas tipo S y varias proteasas principalmente aspártico, cisteína y metaloproteasas (Mulvenna et al., 2009).

En los últimos años la identificación de proteínas secretoras implicadas en la infección de diversos patógenos está destinado a la detección de posibles dianas farmacológicas y biomarcadores para aplicaciones de diagnóstico.

Las proteínas ES incluyen clases de moléculas funcionalmente diversas, tales como citocinas, quimiocinas, hormonas, enzimas digestivas, anticuerpos, proteinasas extracelulares, morfógenos, toxinas y péptidos antimicrobianos. Se sabe que algunas de estas proteínas están involucradas en procesos biológicos vitales, incluida la adhesión celular, la migración celular, la comunicación célula-célula, la diferenciación, la proliferación, la morfogénesis y la regulación de las respuestas inmunitarias (Maizels & Yazdanbakhsh, 2003; Nagaraj et al., 2008). Además, circulan por el cuerpo de un organismo, se localizan o liberan de la superficie celular, lo que las hace fácilmente accesibles a los medicamentos y / o al sistema inmunitario (Bonin-Debs et al., 2004; Ranganathan & Garg, 2009a). En *Fasciola hepática* se han identificado dos enzimas antioxidantes, en etapas larvarias apoyando la importancia de la desintoxicación y su importancia como efectores de la evasión inmune en las etapas del ciclo de desarrollo de trematodos (Gourbal et al., 2008).

Las proteínas secretoras entregan una información valiosa sobre la capacidad del parásito para establecer la interacción huésped-parásito, reflejando diferentes procesos relacionados con sus etapas de desarrollo y diferentes estrategias de evasión inmunológica (Moreno & Geary, 2008). En los últimos años se han analizado secretomas en diversos organismos, tipos de células y patologías,



ofreciendo una descripción general de diferentes bases de datos y software bioinformáticos que se utilizan para analizar el significado biológico de los datos del secretoma (Caccia et al., 2013).

Los mecanismos por el cual *C. rogercresseyi* es capaz de infectar al hospedero aún no se han descritos. Sin embargo, se ha observado en resultados de expresión génica que existe un patrón diferencial en la expresión de una variedad de genes relacionados a proteínas del secretoma en los diferentes estados de desarrollo del ectoparásito (Maldonado-Aguayo, Waleska et al., 2015; Maldonado-Aguayo, W & Gallardo-Escárate, 2014; Valenzuela-Miranda & Gallardo-Escárate, 2016).

## 2.1 SERPINAS

Las serpinas son una superfamilia de proteínas de alrededor de 45kDa que generalmente funcionan como inhibidores de proteasas de serina (Meekins et al., 2017). En vertebrados, inhibidores de proteasas de serina se han estudiado desde hace ya muchos años por lo que se conoce que están envueltos en fagocitosis, coagulación, activación del complemento, fibrinólisis, regulación de la presión de la sangre, entre otros. Por otra parte, en los últimos años se ha evidenciado que en invertebrados las proteasas de serina y los inhibidores están participando en procesos fisiológicos paralelos así como también en respuesta inmune innata (Clynen et al., 2005). En Artrópodos las serpinas que se producen y secretan en la hemolinfa regulan la cascada de señalización de las proteinasas amplificando la señal, resultando en la detección de patógenos, provocando una respuesta inmune innata (Meekins, Kanost, & Michel, 2017). Las serpinas son la familia más grande

conocida de inhibidores de serina proteasas y realizan una variedad de funciones fisiológicas en artrópodos. Las serpinas regulan la inmunidad innata de los insectos a través de la inhibición de las cascadas de proteinasas de serina que inician respuestas inmunes como la melanización y la producción de péptidos antimicrobianos (Clynen, Schoofs, & Salzet, 2005; Meekins, Kanost, & Michel, 2017). Además, varias serpinas con actividad antimicrobiana se expresan como serpinas de fase aguda en insectos tras la infección. Las avispas parasitoides pueden regular la expresión de la serpina del huésped para modular el sistema inmune del huésped. Además, también se han descubierto ejemplos de actividad de serpina en el desarrollo y reproducción en *Drosophila*.

Las serpinas también funcionan en interacciones huésped-patógeno más allá de la inmunidad, como constituyentes del veneno en avispas parasitoides y saliva de garrapatas y mosquitos que se alimentan de sangre. Estas serpinas tienen distintos efectos sobre la inmunosupresión y la anticoagulación y son de interés para el desarrollo de vacunas (Meekins, Kanost, & Michel, 2017). La mayoría de las serpinas de artrópodos se producen y secretan en la hemolinfa para regular procesos fisiológicos internos. Sin embargo, existen serpinas en los fluidos extracelulares como saliva o veneno producidos específicamente para ser transferidos a otros organismos (Colinet et al., 2009; Dorémus et al., 2013).

El piojo de mar *C. rogercresseyi* es un ectoparásito que afecta a la industria del salmón en Chile y su capacidad de infección al hospedero aún no se ha descrito, sin embargo, se ha observado en resultados de expresión génica que existe un patrón diferencial en la expresión de serpinas en el estado infectivo de copepodito. Sugiriendo que Cr-Serpina 10 podría estar participando en la infección y

permanencia del caligus en el hospedero (Maldonado-Aguayo, W & Gallardo-Escárate, 2014).

## **2.2 CATEPSINA**

En especies de invertebrados las proteasas juegan un papel importante en la penetración de tejidos, digestión, muda y evasión de las respuestas inmunes del huésped. Las catepsinas han sido una de las proteasas mayormente encontradas en el secretoma de parásitos. Catepsinas participan en la degradación de proteínas en el lisosoma, núcleo y citosol. Las cisteíno proteasas están conformadas por un grupo de 11 proteasas denominadas catepsinas B, H, L, S, C, K, O, F, V, X y W (Mach et al., 1994). En humanos existen dos tipos de catepsinas L, denominadas catepsina L o L1 y L2, ambas provienen de un ancestro en común presentando un alto porcentaje de identidad entre ellas (Brömme & Wilson, 2011). Catepsina L2 no posee sitios de glicosilación y su localización es tejido específico, mientras que catepsina L1 es expresada en todos los tejidos y presenta una alta glicosilación (Itoh et al., 1999). Catepsina L madura presenta un peso molecular de 29kDa (Mach, Mort, & Glössl, 1994) y es una cisteíno proteasa capaz de degradar proteínas de la matriz extracelular, presenta residuos de cisteínas importantes para la actividad catalítica y pertenece a la superfamilia de papaína (Groves et al., 1998).

En cáncer, se han encontrado catepsinas secretadas por células tumorales que participan en la degradación de componentes de la membrana basal como colágeno, elastina y proteínas de adhesión celular como E-caderina, favorecen procesos de invasión y metástasis tumoral (Tan et al., 2013). Catepsina L puede

proteger a las células cancerígenas sometidas a agentes quimioterapéuticos al degradar a catepsina D, una proteasa principalmente proapoptótica (Zheng et al., 2008).

En *Taenia solium* un hemilto aplanado responsable de la teniasis y de la cisticercosis humana, catepsina L cumple un rol preponderante para el desarrollo y supervivencia del parásito, participando en la invasión tisular, la evasión de la respuesta inmune, el desenquistamiento y el enquistamiento (León et al., 2013). La expresión de las catepsinas en cada parásito es similar, pero su expresión dentro de cada parásito es considerablemente variable, por ejemplo, en *F. hepática* la catepsina L y B son expresadas en estadios invasivos tempranos (Stack et al., 2011). Catepsina L fue observada en los productos secretorios/excretorios de *Lepeophtheirus salmonis* indicando un rol de la proteasa en inmunoevasión y estabilidad del parásito en el hospedero. En los último estudios se han identificado en la etapa infectiva de *L. salmonis* a 3 cisteína proteasas, incluida la catepsina L, lo que sugiere un papel facilitador importante en el aumento de la alimentación (Hamilton et al., 2018). Además, en *L. salmonis* se cree que los compuestos que se producen en las secreciones son responsables de la inmunomodulación de las respuestas del huésped, así como otros aspectos de las interacciones cruciales entre el huésped y el parásito. Se han identificado 187 proteínas ES individuales en los piojos adultos y pre-adultos, entre ellas endopeptidasas de tipo serina, peroxidasa y proteínas de vitelogeninas, endopeptidasas de tipo cisteína, esfingomielinasa y calreticulina (Hamilton et al., 2018). En garrapatas (*Ixodes ricinus*) también se ha demostrado que catepsina L es vital en la digestión de la sangre (Franta et al., 2011) y también se ha demostrado que la catepsina L de

*Fasciola hepatica* tiene propiedades inmunomoduladoras al suprimir la proliferación de linfocitos T ovinos y la expresión de CD4 durante el parasitismo por pulgas hepáticas (Prowse et al., 2002). Dada la importancia de Catepsina en la infección de patógenos, ha sido estudiada como vacuna teniendo resultados prometedores en el caso del trematodo *Fasciola hepatica*. La vacuna de Catepsina L mostró un alto efecto anti-embriónico / anti-fecundidad en parásitos vacunados, por lo que los autores sugieren que podría tener un impacto importante en la transmisión de la enfermedad al huésped (Mulcahy & Dalton, 2001).

En *C. rogercresseyi* se han identificado usando la tecnología de secuenciación de transcriptoma de alto rendimiento, 56 secuencias para catepsina-like distribuidas en cinco grupos de cisteína proteasas (B, F, L, Z y S) y una para el grupo aspártico proteasas (D) (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015). En caligus catepsina L se expresa abundantemente en todos los estados del ciclo de vida, mientras que catepsina B y K se expresan en los estados larvales y hembras adultas, sugiriendo su participación en el proceso de muda y desarrollo embrionario (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015).

### **2.3 TRIPSINA**

Una de las moléculas con mayor abundancia en los productos secretorios de parásitos son las proteasas de serina tales como tripsinas (Lehiy & Drolet, 2014). Las tripsinas hidrolizan uniones peptídicas en el grupo carboxilo de los aminoácidos básicos como arginina o lisina. Tripsina es una proteasa almacenada como

zimógeno, la que luego es activada durante la alimentación, siendo depositada en la saliva a través de autólisis o la acción de proteasas proporcionadas por el huésped (Blow, 1997). Las proenzimas son generalmente monómeros con un dominio catalítico C-terminal y uno o más dominios N-terminales, y se activan por escisión en el extremo amino del dominio catalítico. El dominio catalítico activado permanece atado al (a los) dominio (s) N-terminal a través de un enlace disulfuro. La función catalítica se logra mediante la acción de una tríada catalítica (histidina reactiva, ácido aspártico y serina), y el grado y el tipo de especificidad del sustrato están determinados por muchos residuos dentro de la hendidura de unión al sustrato (Perona & Craik, 1995). La función de las proteasas de serina ha sido asociada principalmente a los procesos de digestión, coagulación y transducción de señales (LeMosy et al., 1999; Zabel et al., 2005). Sin embargo, en los últimos años numerosos estudios han asociado la participación de estas proteasas en procesos como respuesta inmune, mecanismo de infección, desarrollo embrionario, incluso en la resistencia a fármacos.

Un ejemplo de participación en desarrollo embrionario ocurre en *L. salmonis* en el cual, se identificó y caracterizó LsTryp10 que se distribuye uniformemente en los ovarios y ovocitos, la cual se encuentra a lo largo de la periferia de los óvulos. La proteína LsTryp10 se deposita en los ovocitos y en todas las células embrionarias, suponiendo que la proteína LsTryp10 está involucrada en la regulación del degradoma de la yema (Skern-Mauritzen et al., 2009).

La tripsina y la quimiotripsina son las serinas proteasas más abundantes en los crustáceos, y éstas normalmente cumplen funciones en la digestión de proteínas en la dieta y la respuesta inmune. En *Marsupenaeus japonicus* se detectó una

quimiotripsina-like en el hepatopáncreas, el estómago y el intestino, la cual mostró una mayor expresión en los tejidos relacionados con la defensa (hemocitos, órganos linfoides y hepatopáncreas) después de la estimulación con peptidoglicano (Danwattananusorn et al., 2009). Con respecto a la resistencia a fármacos en *Sitophilus zeamais* se han estudiado las proteinasas de serina de tres cepas, una susceptible y dos resistentes a insecticidas. Las proteinasas purificadas caracterizadas como tripsina de todas las cepas fueron inhibidas por fenilmetil sulfonil fluoruro (PMSF), N-alfa-tosil-L-lisina clorometil cetona (TLCK), aprotinina, benzamidina e inhibidor de tripsina de soja (SBTI), según los resultados las proteinasas de tipo tripsina de las cepas resistentes mostraron una mayor afinidad por L-BApNA y N-alfa-p-tosil-L-Arg metil éster (L-TAME). Estos resultados sugieren que la actividad aumentada de la serina proteinasa puede jugar un papel en el mantenimiento de los mecanismos de resistencia a los insecticidas en algunas cepas del parásito del maíz (Silva et al., 2010).

En *Caligus rogercresseyi* se identificaron 44 secuencias para tripsina y 7 para quimiotripsina las cuales mostraron una modulación transcripcional diferenciada después de la exposición a los fármacos utilizados para el control del piojo como deltametrina y azametifos. Se evidenció una respuesta transcripcional sesgada por el sexo para Cr-Tryp1-6-10-36 y 41 bajo exposición a cualquiera de los agentes antiparasitarios, con una regulación de 100 veces mayor que el grupo control. Estos resultados sugieren que tripsina en *C. rogercresseyi* podría desempeñar un papel en la metabolización de los fármacos utilizados para el control de caligus (Valenzuela-Miranda & Gallardo-Escárte, 2016).

Dado las problemáticas asociadas a *C. rogercresseyi*, comprender los fundamentos moleculares implicados en el proceso infectivo del piojo del salmón en Chile es fundamental para el diseño de nuevas estrategias de manejo de plagas para este ectoparásito. Por esta razón, la presente tesis aborda unos de los principales mecanismos de regulación post-transcripcionales sobre genes relacionados con el proceso infectivo de *Caligus rogercresseyi*.

### **3. REPRODUCCIÓN EN *CALIGUS ROGERCRESSEYI***

#### **3.1 VITELOGENINA**

Vitelogenina es una fosfoglicoproteína compleja de alto peso molecular que liga calcio. Es producida en los hepatocitos y su síntesis es estimulada por estrógeno E2 (Tufail & Takeda, 2008). Las vitelogeninas (Vgs) son precursoras de la proteína principal de almacenamiento del huevo, la vitelina. Son moléculas grandes (~200-kD) sintetizadas en el cuerpo graso en un proceso que involucra modificaciones estructurales sustanciales (por ejemplo; glicosilación, lipidación, fosforilación y escisión proteolítica) de la proteína naciente antes de su secreción y transporte (Tufail & Takeda, 2008). Esta proteína se sintetiza generalmente extra ováricamente y se transporta por el sistema de circulación al ovario, en donde se internaliza en los ovocitos en crecimiento a través de endocitosis mediada por receptor, para luego ser escindida proteolíticamente para generar 2 proteínas de la yema, lipovitelin y fosvitina que luego son usados como nutrientes para el desarrollo de embriones (Arukwe & Goksøyr, 2003; Zhang et al., 2011). A pesar de que en sus inicios vitelogenina era asociado específicamente al sexo femenino ahora existen



múltiples antecedentes que la relacionan con funciones no sexo específicas. En abejas se cree que vitelogenina es un factor importante que afecta en la señalización hormonal, comportamiento en la alimentación, inmunidad, resistencia al estrés y en la esperanza de vida. También actúa como un receptor de reconocimiento de patrones multivalentes capaz de unirse a lipopolisacáridos, ácido lipoteicoico, peptidoglicano y viriones. Se ha demostrado que Vg es un receptor de reconocimiento de patrones multivalentes capaz de reconocer microbios por interacción con PAMPs (Zhang, Wang, Li, & Li, 2011). En *Lepeophtheirus salmonis* se han caracterizado 2 Vitelogeninas (LsVit1 y LsVit2), asociando su función en reproducción como proveedores de la nutrición embrionaria y larvaria (Dalvin et al., 2011).

Los estudios relacionados con reproducción en *C. rogercresseyi* son escasos pero se han reportado transcritos putativamente relacionados a procesos de diferenciación sexual y reproducción describiendo genes de vitelogeninas (Farlora et al., 2014a). Además, también se han caracterizado a nivel molecular genes ligado al sexo como Prohibitin 2 (Farlora et al., 2015). A pesar que en *C. rogercresseyi* se desconoce las funciones que podría tener vitelogenina I, se asocia a la formación del huevo dada su alta expresión en estado hembra del parásito.

Comprender los mecanismos implicados en la reproducción del piojo del salmón es fundamental para el diseño de nuevas estrategias de control. Es por este motivo que en la presente tesis se pretende estudiar los mecanismos de regulación post-transcripcionales sobre genes relacionados con el proceso reproductivo de *Caligus rogercresseyi* como vitelogenina I.

Existen antecedentes de genes en caligus relacionados con infestación, diferenciación sexual y sistema inmune (Chávez-Mardones & Gallardo-Escárate, 2015; Chávez-Mardones et al., 2015; Farlora, Araya-Garay, & Gallardo-Escárate, 2014a; Vera-Bizama et al., 2015). Adicionalmente, estudios recientes lograron la caracterización molecular y la modulación transcripcional de la vía de biogénesis de los microRNA en el piojo del salmón, los resultados revelaron nuevos conocimientos sobre la naturaleza de la regulación por miRNA los cuales podrían estar desempeñando funciones en la regulación del desarrollo y la respuesta a los fármacos en *C. rogercresseyi* (Valenzuela-Miranda et al., 2015). El descubrimiento de la presencia de miRNA en el piojo del salmón genera un nuevo desafío para comprender la regulación de genes tan importantes en el caligus, los cuales podrían estar participando en la regulación de procesos cruciales en su ciclo de vida como lo es el proceso infectivo y reproductivo del piojo de mar.

Teniendo en cuenta estos antecedentes la propuesta de la presente tesis pretende aportar información genómica valiosa, centrada en estudiar la importancia de genes asociados al secretoma relacionados al proceso de infección en *C. rogercresseyi*, y del gen relacionado al proceso reproductivo. Por otra parte, conocer los mecanismos de regulación post-transcripcional mediado por miRNA, nos ayudaría a entender como este sistema de regulación permite al ectoparásito cumplir funciones vitales para su capacidad infectiva y reproductiva.

#### 4. microRNA y BIOGENESIS

Los microRNA (miRNA) son una clase de ARN pequeños con capacidad de reconocer por complementariedad de bases regiones del ARN mensajero blanco. Esta unión resulta en la inhibición de la expresión de dicho mensajero, como un mecanismo de regulación de la expresión génica post-transcripcional. El primer miRNA descrito fue let-7, que ayuda a controlar la duración de los diferentes estadios de desarrollo de la larva *Caenorhabditis elegans* (Lee et al., 1993) y desde entonces, han sido reportadas secuencias de miRNA para una amplia variedad de organismos (Griffiths-Jones et al., 2008; Kozomara & Griffiths-Jones, 2014).

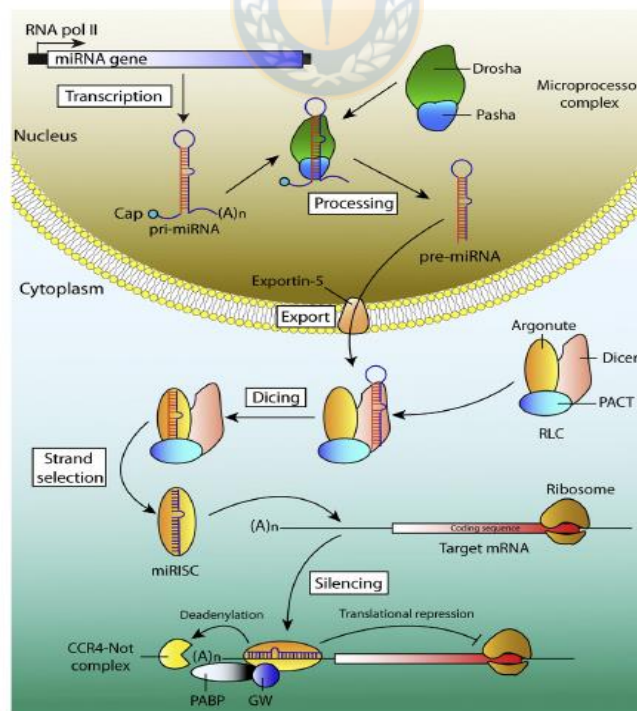
Los genes de los miRNA se transcriben por las RNA polimerasa II o III. Al inicio, los miRNA forman parte de una estructura de tallo y loop de aproximadamente 80 nucleótidos, que forma parte del pri-miRNA (Xiao & Rajewsky, 2009). Esta estructura pri-miRNA es procesada por un complejo proteico compuesto por dos proteínas; la enzima Nuclear RNasa III denominada Drosha y su cofactor DGCR8 en Vertebrados o Pasha en invertebrados (proteína con dominios de unión al RNA). El microprocesador corta el pri-miRNA para generar una horquilla más corta de unos 70 nt de longitud denominada pre-miRNA. El pre-miRNA, se exporta desde los poros del núcleo al citoplasma a través de la proteína Exportina-5, un miembro de la familia de las karioferinas que son factores de transporte núcleo-citoplasma dependientes de Ran (GTPasa). En el citoplasma los pre-miRNA son reconocidos por la enzima RNasa II Dicer. Esta enzima corta el tallo y el bucle del pre-miRNA para generar un RNA dúplex formado por una cadena de miRNA madura y una cadena de miRNA complementaria de aproximadamente 22 nucleótidos de longitud. Después de la

digestión, Dicer permanece asociado al miRNA dúplex, que posteriormente es liberado por acción de la RNA helicasa A (Células humanas). El miRNA maduro es incorporado al complejo silenciador inducido por RNA (RISC), constituido por varias proteínas como Dicer, TRBP y principalmente Argonauta. En esta etapa el RNA dúplex es separado, una cadena (miRNA sentido) es degradada, mientras que la otra cadena (miRNA, antisentido o guía) es incorporada al complejo RISC (Lugo Trampe & Trujillo Murillo, 2009). El miRNA guía actúa como tal para identificar a los ARNm blanco, con el fin de permitir que RISC actúe sobre ellos, bloqueando la síntesis de proteínas, por desestabilización del ARNm, y represión traduccional.

Otro aspecto importante de los miRNAs es que la regulación de la expresión del ARNm diana depende del grado de complementariedad entre ellos, principalmente a la región seed de los miRNA que consta de 8 nucleótidos ubicados en el extremo 5' del miRNA (Takacs & Giraldez, 2010). Una complementariedad imperfecta conduce a la represión de la traducción, mientras que una complementariedad perfecta o casi perfecta de la región seed, conduce a la escisión del ARNm diana. Los miRNAs descritos en animales son a menudo imperfectamente complementarios con sus sitios de destino, por lo que la represión de la traducción es el mecanismo predominante de inhibición de la expresión de ARNm en células animales (Zhuo et al., 2013). En contraste, los miRNAs de plantas generalmente tienen una complementariedad perfecta o casi perfecta con los sitios objetivo, provocando la escisión de los ARNm target (Duan et al., 2006). La tercera manera en la que los miRNAs regulan la expresión del ARNm implica la deadenilación del ARNm. Cuando un miRNA se une de manera imperfectamente complementaria al sitio diana del ARNm, el miRNA acelera la deadenilación de la

cola poli (A) y la degradación de los ARNm (Takacs & Giraldez, 2010). Así, modulando la estabilidad y traducción de los ARNm objetivos, los miRNAs desempeñan un papel clave en la regulación de la expresión génica, influyendo en una serie de procesos fisiológicos, incluyendo el metabolismo, respuesta inmune y reproducción (Patil et al., 2014).

En *Caligus rogercresseyi* se ha observado una expresión diferencial de los miRNA en los diferentes estados de desarrollo (Gallardo-Escárte, C et al., 2017). Se encontraron 40 familias de miRNA de 12 familias de especies de artrópodos. Dentro de los miRNA que destacan son el miRNA-996-4 y mirRNA-124 los cuales se expresarían ligado al sexo, por otra parte, se encontraron isoformas de bantam altamente expresadas durante la etapa infectiva del caligus (Gallardo-Escárte, C et al., 2017).



**Figura 2.** Representación esquemática de la regulación transcripcional regulado por miRNA en *C. rogercresseyi*. Diferentes grupos de genes identificados en *C. rogercresseyi* y su papel en el

procesamiento, exportación, maduración y silenciamiento del ARNm mediado por miRNAs (Valenzuela-Miranda, Nuñez-Acuña, Valenzuela-Muñoz, Asgari, & Gallardo-Escárdate, 2015).

En la presente tesis se pretende estudiar la modulación de la expresión de genes asociados a la etapa infectiva como catepsina, serpina y tripsina mediante la regulación transcripcional de bantam, y para el proceso reproductivo estudiar la modulación de la expresión de vitelogenina I gen asociado a la etapa reproductiva por parte de la regulación transcripcional de mir-996.



## **HIPOTESIS DE TRABAJO**

Los miRNAs bantam y mir-996 poseen un rol modulador sobre la expresión de genes claves para el proceso infectivo y reproductivo de *Caligus rogercresseyi* a través de mecanismos post-regulatorios.

## **OBJETIVOS**

### **OBJETIVO GENERAL**

Determinación de la modulación de miRNAs sobre genes asociados a la etapa infectiva y reproductiva del copépodo ectoparásito *Caligus rogercresseyi*.

### **Objetivos Específicos**

1. Identificar a nivel bioinformático interacciones mRNA/miRNAs de proteínas asociadas al secretoma y reproducción en *Caligus rogercresseyi*.
2. Evaluar los niveles de expresión de miRNAs candidatos y sus genes target durante los diferentes estados de desarrollo de *C. rogercresseyi*.
3. Validar a nivel funcional la interacción de mRNA/miRNAs candidatos relacionados al proceso infectivo y reproductivo en *C. rogercresseyi*.

## METODOLOGIA

### 1. Identificación de genes asociados a proceso infectivo y reproducción en *Caligus rogercresseyi*.

Los análisis bioinformáticos se realizaron utilizando el software CLC Genomics Workbench (Versión 6.0.1, CLC Bio, Dinamarca), a partir de una base de datos para *C. rogercresseyi* (Gallardo-Escárate, Cristian et al., 2014) previamente obtenida en el Laboratorio de Biotecnología y Genómica Acuícola (LBGA) de la Universidad de Concepción. El RNA total se extrajo de los grupos que contenían diez individuos de cada estadio larval (copépodos y chalimus) y adultos (hembras y machos) utilizando el kit Ribopure™ (Ambion®, Life Technologies™, EE.UU.) de acuerdo con las instrucciones del fabricante. La cantidad, pureza y calidad del RNA aislado se midieron en TapeStation 2200 (Agilent Technologies Inc., Santa Clara, CA, EE.UU.) usando el kit de re-agente R6K de acuerdo con las instrucciones del fabricante. Posteriormente, las bibliotecas de cDNA de doble cadena se construyeron usando el kit TruSeq RNA Sample Preparation v2 (Illumina®, San Diego, CA, USA). Se secuenciaron dos repeticiones biológicas para cada grupo de muestras (n = 8) mediante la plataforma MiSeq (Illumina®) utilizando secuencias de 2 x 250 lecturas en pares en el LBGA, Universidad de Concepción.

La anotación transcriptómica se basó en un análisis tBLASTx contra una base de datos EST de GenBank de proteínas catepsina, serpinas y tripsinas para el caso del estudio de proteínas involucradas en el proceso infectivo (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015; Maldonado-



Aguayo, W & Gallardo-Escárate, 2014; Valenzuela-Miranda & Gallardo-Escárate, 2016) y vitelogenina I para el proceso de reproducción (Farlora, Araya-Garay, & Gallardo-Escárate, 2014a). Para determinar las secuencias a analizar se considerará un valor de corte de E-value de  $\leq 1E-05$ .

### **1.1 Alineamiento de secuencias aminoacídicas y análisis filogenético**

Las secuencias nucleotídicas identificadas, fueron comparadas con secuencias proteicas de otras especies de artrópodos para ello se realizó un alineamiento múltiple con una matriz BLOSUM62, el cual está incluido en el software Geneious (version 9, Biomatter Ltda., New Zealand). Posteriormente, para la generación del árbol filogenético se utilizó como modelo de distancia genética a Jukes-Cantor, el método de construcción se basó en Neighbor-Joining. Por último, la información obtenida fue bootstrapped 1000 veces.

### **1.2 Secuenciación Illumina de Small RNAs.**

Los análisis bioinformáticos se realizaron utilizando el software CLC Genomics Workbench (Versión 6.0.1, CLC Bio, Dinamarca), a partir de una base de datos de small RNAs para *C. rogercresseyi* (Gallardo-Escárate, C et al., 2017) previamente creada en el LBGA. La cual fue obtenida mediante secuenciación de veinte individuos de cada estado de desarrollo de caligus. Los individuos fueron recolectados por separado e inmediatamente después del muestreo, cada etapa de desarrollo se agrupó en tres repeticiones biológicas en 1 mL de RNAlater Stabilization Solution (Ambion®, USA) y fueron almacenados a  $-80^{\circ}\text{C}$ . El RNA total

se extrajo de cada grupo usando el Trizol™ Kit 120 (Ambion®) siguiendo las instrucciones del fabricante. La cantidad, pureza y calidad del RNA aislado se midieron en el TapeStation 2200 (Agilent Inc., Santa Clara, CA, EE.UU.) utilizando el kit de reactivos R6K de acuerdo a las instrucciones del fabricante. Las muestras con un RIN > 8,0 se agruparon y se usaron para la preparación de las bibliotecas. Posteriormente, a partir de pools de RNA de cada estado de desarrollo se construyeron bibliotecas de cDNA de doble cadena utilizando el kit TruSeq Small RNA (Illumina®, San Diego, CA, EE.UU). Dos repeticiones biológicas para cada grupo de muestras fueron secuenciadas por la plataforma MiSeq (Illumina®) utilizando secuencias de 41 ciclos en LBGA de la Universidad de Concepción.



### **1.3 Identificación de miRNAs en *C. rogercresseyi*.**

A partir de la base de datos generada por secuenciación de small RNA (Gallardo-Escárate, C et al., 2017), los datos de Illumina de baja calidad, con un puntaje de calidad inferior a 0,05 en la escala de Phred, con una longitud corta, o con tres o más nucleótidos ambiguos fueron estrictamente eliminados utilizando el software CLC Genomics Workbench (Versión 9.1, CLC Bio, Dinamarca). Además, cualquier secuencia limpia que coincidiera con el ARNm de metazoo, ARNr, ARNt, ARNsn, secuencias repetitivas, u otros ncRNAs depositados en las bases de datos NCBI (<http://www.ncbi.nlm.nih.gov/>), RFam (<http://rfam.janelia.org/>), o Rfam (<http://www.girinst.org/rebase/>) fueron descartados. Posteriormente, las lecturas de nucleótidos retenidas fueron mapeados en las secuencias EST de cangrejos y camarones (<http://www.ncbi.nlm.nih.gov/> y <http://www.marinegenomics.org/>) y

sobre el genoma *Daphnia pulex* (<http://genome.jgi.Psf.org/Dappu1/>) basado en las versiones públicas. A continuación, las diversas secuencias mapeadas y limpiadas fueron contadas, y se generó una lista de secuencias únicas. Estas secuencias fueron alineadas contra (Griffiths-Jones, Saini, van Dongen, & Enright, 2008; Hackenberg et al., 2011) Pre-miRNA y ARNm maduro (5' y 3') secuencias enumeradas en el miRBase 21 (Griffiths-Jones, Saini, van Dongen, & Enright, 2008; Hackenberg, Rodríguez-Ezpeleta, & Aransay, 2011). Luego se verificaron las secuencias de miRNA para las predicciones de la estructura secundaria utilizando CLC Genomics Workbench Software con condiciones de plegado predeterminadas.

#### **1.4 Predicción de genes dirigidos por miRNAs de *C. rogercresseyi*.**

Para predecir los genes dirigidos por miRNAs, se utilizaron tres algoritmos: PITA (Kertesz et al., 2007), STarMir (Kanoria et al., 2016) y RNAHydrid.2.1.2. STarMir se utilizó para la búsqueda de las secuencias seed de los miRNA que constan en secuencias de 2-8 nucleótidos ubicados en el extremo 5' del miRNA, en las secuencias 3'UTR. STarMir es una herramienta para predecir sitios potenciales de unión para uno o múltiples miRNAs en una secuencia de RNA mensajero diana. Se basa en los modelos de predicción logística desarrollados con datos de unión a miRNA. A cada uno de los sitios candidatos se le asigna una probabilidad logística como medida de confianza en el sitio predicho. Trabaja con 3 modelos para hacer predicciones; modelo para *H. sapiens*, *M. musculus* y *C. Elegans*, los cuales pueden ser aplicables a otras especies de mamíferos. El resultado del análisis entrega parámetros de medida de estabilidad de unión (DG hybrid) para el miRNA y la

secuencia diana calculado como RNA híbrido. Una medida del potencial de nucleación (DG nucl) para miRNA y la hibridación objetivo y una medida del cambio energético total de la hibridación (DG total) (Kanoria et al., 2016). Por otra parte, PITA se centra en la accesibilidad de la secuencia target que está estrictamente vinculada a la estructura secundaria del transcrito. La suposición principal se basa en el hecho de que la estructura del mRNA juega un papel en el reconocimiento del miRNA promoviendo o desfavoreciendo termodinámicamente la interacción. PITA primero predice las secuencias target utilizando el análisis de complementariedad dentro de las regiones seed del miRNA y luego compara la energía libre obtenida de la formación del duplex entre el miARN-target y el costo energético de separar la secuencia target para hacer accesible al miRNA (M Witkos et al., 2011). Para los análisis se utilizaron las regiones 3'UTR de cada gen candidato. Para determinar las secuencias a analizar se consideró un valor de corte de energía de hibridación de  $<-10$  kcal/mol y un score positivo.

## **2. Evaluación de los niveles de expresión de miRNAs candidatos y sus genes target durante los diferentes estados de desarrollo de *C. rogercresseyi***

### **2.1 Extracción de RNA y síntesis de ADNc.**

Para la extracción de RNA se utilizó el reactivo Trizol (Invitrogen) de acuerdo a las instrucciones del fabricante. La cuantificación del RNA se realizó por medio de un espectrofotómetro ND 1000 (NanoDrop Technologies). Para la síntesis de cDNA

se utilizó el kit RevertAid H Minus Firts Strabd cDNA Synthesis (Fermentas), según las instrucciones del fabricante. Como control endógeno se utilizó  $\beta$ - tubulina II de *C. rogercresseyi* diseñado y estandarizado previamente en el laboratorio.

## **2.2 Análisis de expresión relativa por RT-PCR para genes target.**

La reacción q-PCR se llevó a cabo en el equipo StepOnePlus™ (Applied Biosystems). Para la cuantificación de la expresión génica de los genes de interés se utilizó el método  $\Delta\Delta C_t$  comparativo utilizando el gen  $\beta$ -tubulina II como housekeeping para la normalización de los datos de expresión. Para la reacción se utilizó el kit SYBGRen con ROX (Fermentas) en un volumen final de reacción de 10 $\mu$ L. Las condiciones de amplificación utilizadas fueron: 95°C por 10 min, 40 ciclos de 95°C por 15 s, 30 s a T° de alineamiento (de acuerdo a cada gen), 30 s seguido de curva de melting (95°C por 15s, 60°C por 1 min y 95°C por 15s).

## **2.3 Housekeeping para miRNA de *Caligus rogercresseyi***

Para la búsqueda de un housekeeping para miRNA de *C. rogercresseyi* se utilizaron los valores de conteo y RPKM para los miRNA de caligus. De ambos sets de datos se seleccionaron los miRNA que tenían valores en todos los estados de desarrollo y los que presentaban menor varianza (Tabla 6). Para el diseño de partidores de los miRNA seleccionados se buscó su secuencia madura en miRBASE (Tabla 2). Con los datos de Ct obtenido mediante RT-PCR para miRNA se analizaron mediante la plataforma R el cual es un software libre para

computación estadística y gráficos. En la cual utilizamos el software NormFinder, la cual es una herramienta para la identificación de genes de normalización.

## 2.4 Análisis de expresión de miRNA.

Para el caso de los análisis de RT-qPCR para los miRNA se llevaron a cabo de acuerdo a Mei et al. (2012), y se utilizó el kit miScript II RT y miScript SYBR Green PCR (Qiagen) de acuerdo a las instrucciones del fabricante. Finalmente, a los datos de RT-qPCR se le realizaron análisis estadísticos en el software JMP v9.0 (SAS Institute, Inc., Cary, NC, EE.UU.). Se aceptaron diferencias estadísticamente significativas con  $p < 0,005$ .

**Tabla 1.** Partidores empleados en los análisis de expresión relativa para los genes de estudio (Fuente: Elaboración propia).

Nombre	Gen	Secuencia 5'-3'	%GC	Tm (°C)	Largo
qCr_Cat_F1		TTGGTGGCTGTTCTCTCCGTTTCCT			25
qCr_Cat_R1	Catepsina L	TCTTGGGCACATTGATGTTGGCTGG	52	60	25
qCr_Ser_F1		CCTTTCAGGCCATTCGTGAGCGAAA			25
qCr_Ser_R1	Serpina	TGCAGCATTGACGAGAGCAACCTTC	52	60	25
qCr_Trip_F1		GCTGCTACAGGCAGTTGGCTGAAAA			25
qCr_Trip_R1	Tripsina	AGTCCTTTCCTTCAGTGTAGGCGCA	52	60	25
qCr_Vit1_F1		TCCCCCGAATACGTTCTCTGCAACA			25
qCr_Vit1_R1	Vitelogenina I	TGCGAATTCTGAAGGAGTCGACGGA	52	60	25

**Tabla 2.** Partidores empleados en los análisis de expresión relativa para los miRNA de estudio (Fuente: Elaboración propia).

MiRNA	Secuencia <i>Caligus rogercresseyi</i>	mirBASE	Secuencia madura
mir-2505	CCACACACACACACACACACACAC	<i>Mus musculus</i>	ugugugugugugugugugug
mir-263 a 1	TTCCCGTGAATTCTTCCAGTGCCATTT	<i>Drosophila melanogaster</i>	aauggcacuggaagaauucacggg
mir-8 2	TACATCTTTACCTGACAGTATTA	<i>Drosophila melanogaster</i>	caucuuaaccgggcagcauuaga
mir-276	TAGGAACTTCATACCGTGCTTTTGGAAAT	<i>Polistes canadiensis</i>	uaggaacuucacuaccgugcucu
mir-996	AGATGAGTTTGTATCTAGTCA	<i>Apis mellifera</i>	ugacuagauacauacucgucu
Bantam	TCAGCGTTCACAATGATCTCA	<i>Apis mellifera</i>	ugagaucauugugaaagcugauu

### 3. Validación a nivel funcional la interacción de mRNA/miRNAs candidatos relacionados al proceso infectivo y reproductivo en *C. Rogercresseyi*

De acuerdo a los análisis bioinformáticos de las secuencias del ARNm de los genes de interés para el caso de Cr\_Catepsina, Cr\_Tripsina y Cr\_Serpina, se identificó que estaba completa la región no traducible del extremo 3' denominada 3'UTR. Sin embargo, para Cr\_Vitelogenina I no fue posible la obtención del extremo 3'UTR de acuerdo a los datos entregados por secuenciación masiva por lo que fue necesario aplicar una técnica previamente al clonamiento del 3'UTR denominada RACE de la sigla Rapid Amplification of cDNAs ends (Fig. 3).

### 3.1 Amplificación del ORF parcial próximo al 3'UTR de Vitelogenina I

De acuerdo a los datos de secuenciación masiva del LBGA solo se obtuvo una secuencia parcial del gen Vitelogenina I la cual se usó como una plantilla para el diseño de primers con la herramienta Primer3 (Rozen y Skaletsky, 2000) (Tabla 3) incluida en el software Geneious Pro (Versión9, Biomatters, Nueva Zelandia), con el objetivo de amplificar la ORF parcial del gen de vitelogenina I para posteriormente realizar la amplificación del extremo 3'UTR de vitelogenina I de *C. rogercresseyi*. La secuencia parcial resultante se utilizó como plantilla para el diseño de nuevos primers específicos (Tabla 3) para la amplificación del extremo 3'UTR utilizando el Kit de amplificación de cDNA SMARTer™ RACE (Clontech) de acuerdo con la Instrucciones del fabricante.



**Tabla 3.** Partidores empleados en el clonamiento de la ORF parcial próxima al 3'UTR y RACE para Vitelogenina I (Fuente: Elaboración propia).

Primer	Amplificación	Secuencia 5'- 3'	% GC	Tm (°C)	Producto (pb)
Cr_Vit1_R_F1	ORF	CTTCTCACGTGGTTATTGCTGATTCCAA	43	60	613
Cr_Vit1_R_R1	ORF	CGGAAGCTCCATGGAAGGACGTGTT	56		
Cr_Vit1_R_F2	ORF	GTCCTACATCCGCAACTTCCTCGAAAC	52	58	655
Cr_Vit1_R_R2	ORF	GAAAGAAGAGGTCCAGGACAAGAAGA	46,2		
Cr_Vit1_R_F3	ORF	CCATGGAAGTCAAGATGGTTGTTGGA	46,2	60	643
Cr_Vit1_R_R3	ORF	GGGATACCTTGGCAAGAATCCTTAATCA T	41,4		
Cr_Vit1_R_F4	ORF	CTCCCCAAGATGTGCATCTTCGA	54,2	60	588
Cr_Vit1_R_F4	ORF	CCTGTGCTGGGTTGAGCATGATGAG	54,2		



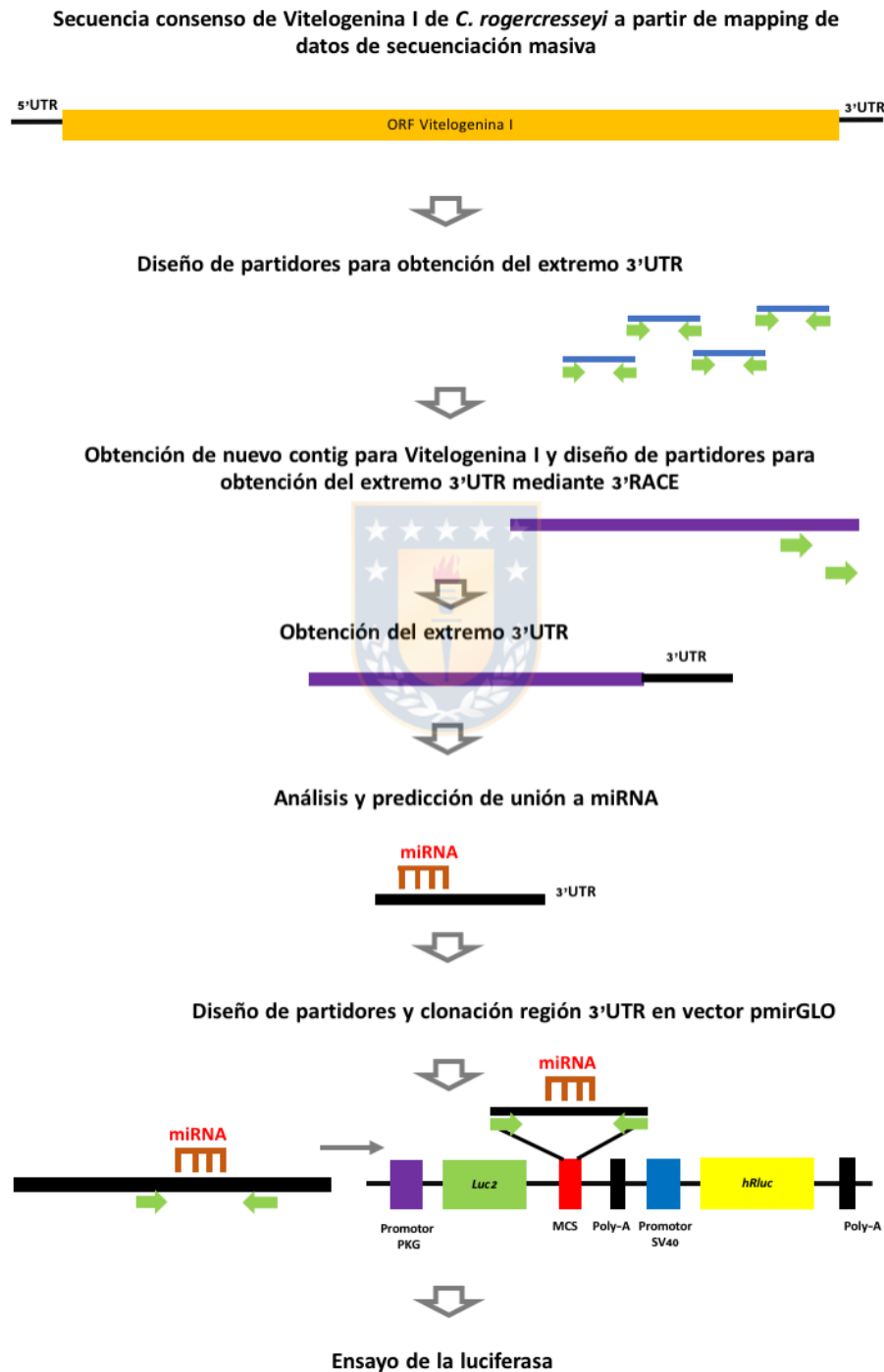
**Tabla 3 (Continuación).** Partidores empleados en el clonamiento de la ORF parcial próxima al 3'UTR y RACE para Vitelogenina I (Fuente: Elaboración propia).

Primer	Amplificación	Secuencia 5'- 3'	% GC	Tm (°C)	Producto (pb)
Cr_Vit1_FR1	Outer 1 (RACE)	CACCAAGCACTTGGTCGAGTAC	54,4	56	-
Cr_Vit1_FR2	Inner 1 (RACE)	CATCGGACACATGGCCTCCAC	61	56	-
Cr_Vit1_FR3	Outer 2 (RACE)	CGTACTACACGCCAACGAAGAGATC	52	58	-
Cr_Vit1_FR4	Inner 2 (RACE)	CGAGGAAGCCTCCACCTGCC	70	58	-

### 3.2 Purificación de fragmentos de DNA mediante geles de agarosa

Los fragmentos de DNA de interés fueron purificados desde geles de agarosa utilizando el kit comercial “E.Z.N.A Gel Extraction Kit” (Omega Bio- Tek). La muestra de DNA fue aplicada sobre un gel preparativo de agarosa al 1%. Luego de obtener una buena separación de la banda de interés, visualizada al UV a una longitud de onda de 300 nm, se cortó el trozo de gel que contenía la banda de DNA de interés. Luego se determinó el peso del trozo del gel extraído. Asumiendo que un gel de agarosa 1% tiene una densidad de 1 g/mL. Luego se agregó un volumen de Binding Buffer XP2 igual al volumen de gel y se incubó en baño termostático a 60 °C. La mezcla anterior se puso en una columna de purificación DNA-HiBind montada en un tubo de colección de 2 mL, se centrifugó a 10.000 g por 1 minuto y se descartó el eluido, posteriormente se lavó con 300 µl de Binding Buffer XP2, se centrifugó a 13.000 g por 1 minuto y se descartó el eluido. Se adicionó 700 µL de SPW Wash Buffer en la columna, se centrifugó a 13.000 g por 1 minuto y se descartó el eluido. En seguida se centrifugó la columna vacía a 14.000 g por 2 minutos, la cual se

montó sobre un tubo vacío de 1,5 mL para recibir el ADN. Se adicionó agua nanopure, la columna con agua montada en el colector fue incubada a 60 °C por 2 minutos, se centrifugó a 14.000 g por 1 minuto.



**Figura 3.** Diseño experimental para la obtención del gen Vitelogenina I y el 3'UTR mediante RACE

(Fuente: Elaboración propia).

### **3.3 Clonamiento y purificación de plásmido recombinante.**

Los productos PCR obtenidos anteriormente se clonaron en el vector pGEM-T Easy Vector System I (Promega), y se transformaron en bacterias electrocompetentes *E. coli* JM109 las cuales se dejaron crecer en placas agar LB / amp / Xgal durante toda la noche a 37 ° C. Las Bacterias de la cepa JM109 que contenían el plásmido pGEM-T easy con el inserto fueron crecidas en medio Luria-Bertani (Peptona 10g/L, extracto de levadura 5 g/L, NaCl 10 g/L y ampicilina 100 µg/mL) bajo agitación a 37°C durante toda la noche. Para la purificación del plásmido a mediana escala (50 mL de cultivo bacterial) se siguieron las indicaciones del proveedor del sistema comercial E.Z.N.A. Plasmid DNA Mini Kit II (OMEGA, bio.tek), basado en el método de lisis alcalina y filtración con columnas. La concentración de DNA fue determinada usando un espectrofotómetro de luz UV. La confirmación de la presencia del inserto se realizó por PCR convencional utilizando los partidores diseñados (Tabla 3).

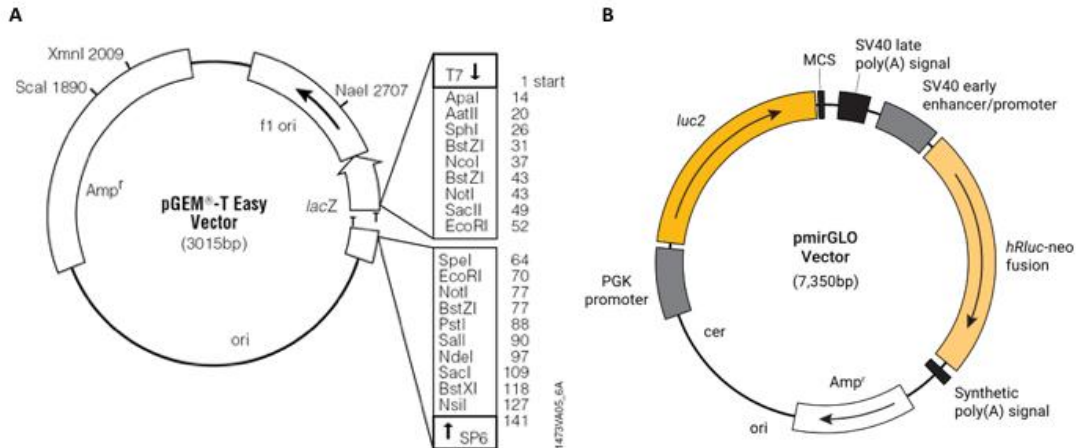
### **3.4 Secuenciación de plásmidos**

Los plásmidos adquiridos se secuenciaron en MACROGEN en ambas direcciones, y sus secuencias se ensamblaron luego usando el software Geneious Pro (Versión 9, Biomatters Ltda.). Se utilizó como referencia las secuencias obtenidas por secuenciación masiva.

### 3.5 Amplificación del 3'UTR por PCR

De acuerdo a los análisis bioinformáticos de las secuencias del ARNm de Cr\_Catepsina, Cr\_Tripsina y Cr\_Serpina, se identificó que estaba completa la región 3'UTR. La región 3'UTR para Cr\_Vitelogenina I fue obtenida mediante RACE. En primera instancia se diseñó una primera pareja de partidores "mismatch" que tuvieran el sitio de corte reconocido por la enzima de restricción *PmeI* y *XbaI* (Tabla 4). Los genes Cr\_Catepsina, Cr\_Serpina, Cr\_Tripsina y Cr\_Vitelogenina I fueron amplificados por reacción en cadena de la polimerasa (PCR, por sus siglas en inglés *Polymerase Chain Reaction*) empleando como plantilla DNA genómico de *C. rogercresseyi*. El ciclo de amplificación consistió de una desnaturalización inicial a 95°C por 5 minutos, seguida de 35 ciclo a 95 °C por 30 segundos, temperatura de hibridación (depende de cada primer) por 30 segundos, extensión a 72 °C por 30 segundos y una extensión final a 72 °C por 7 minutos. Los productos de reacción fueron separados en geles de agarosa al 1% preparados en buffer TAE (Tris acetato 40mM, EDTA 1 mM, pH 8). Posteriormente, los productos obtenidos del tamaño esperado, fueron purificados con el kit de purificación de banda de E.Z.N.A. Gel Extraction (OMEGA bio-tek) y cuantificados midiendo la absorbancia a 260 nm en Nanodrop ND1000 espectrofotómetro (Thermo Fisher Scientific, EE.UU.).

Los plásmidos adquiridos se secuenciaron en MACROGEN en ambas direcciones, como se mencionó en el punto 3.4.



**Figura 4.** Plásmidos utilizados en el diseño experimental. **A.** Vector pGEM-T Easy utilizado para el clonamiento de los insertos y su posterior secuenciación. **B.** Vector pmirGLO para el ensayo de la luciferasa (Fuente: Elaboración propia).

### 3.6 Subclonamiento de la secuencia del 3'UTR que contiene el sitio de unión al microRNA en el vector pmirGLO.

El cDNA del 3'UTR que contiene el sitio de unión al miRNA contenida en el plásmido pGEM-T easy se amplificó utilizando cepas de la bacteria *E. Coli* BL-21 para la clonación, propagación y mantenimiento del plásmido. *E. coli* BL-21 fue cultivada en medio Luria Bertani (LB) líquido con agitación constante a 250 rpm a 37°C suplementado con ampicilina (Amp) 100 µg/mL. El plásmido purificado se trató con enzimas de restricción PmeI (GTTTAAAC) y XbaI (TCTAGA) para asegurarnos de direccionalidad de inserción del constructo en el vector. La digestión consistió en 1µL de Enzima de Restricción (10U/µl), 1µg de DNA, 5µL de Buffer 10X y completado a 50µL con Agua, la reacción fue incubada a 37°C durante 15 minutos. La muestra de DNA fue aplicada sobre un gel de agarosa al 1%. Luego de obtener una buena separación de la banda de interés, visualizada al UV, se extrajo el trozo

de gel que contenía la banda de DNA de interés y fueron purificados utilizando el kit comercial “E.Z.N.A Gel Extraction” (Omega Bio- Tek) como se menciona en el punto 3.2.

El vector pmirGLO Dual-Luciferase miRNA Target Expression (PROMEGA) fue previamente linealizado con las enzimas de restricción PmeI y XbaI para insertar el fragmento de interés, utilizando la enzima DNA ligasa del bacteriófago T4, siguiendo las instrucciones del fabricante. Se transformaron en bacterias *E. coli* BL.21 para la clonación, propagación y mantenimiento del plásmido.

**Tabla 4.** Partidores empleados para el clonamiento del cDNA de Cr\_Catepsina, Cr\_Serpina y Cr\_Tripsina y Cr\_Vitelogenina I de *C. rogercresseyi* que contienen el sitio de unión a Bantam y mir-996 (Fuente: Elaboración propia).

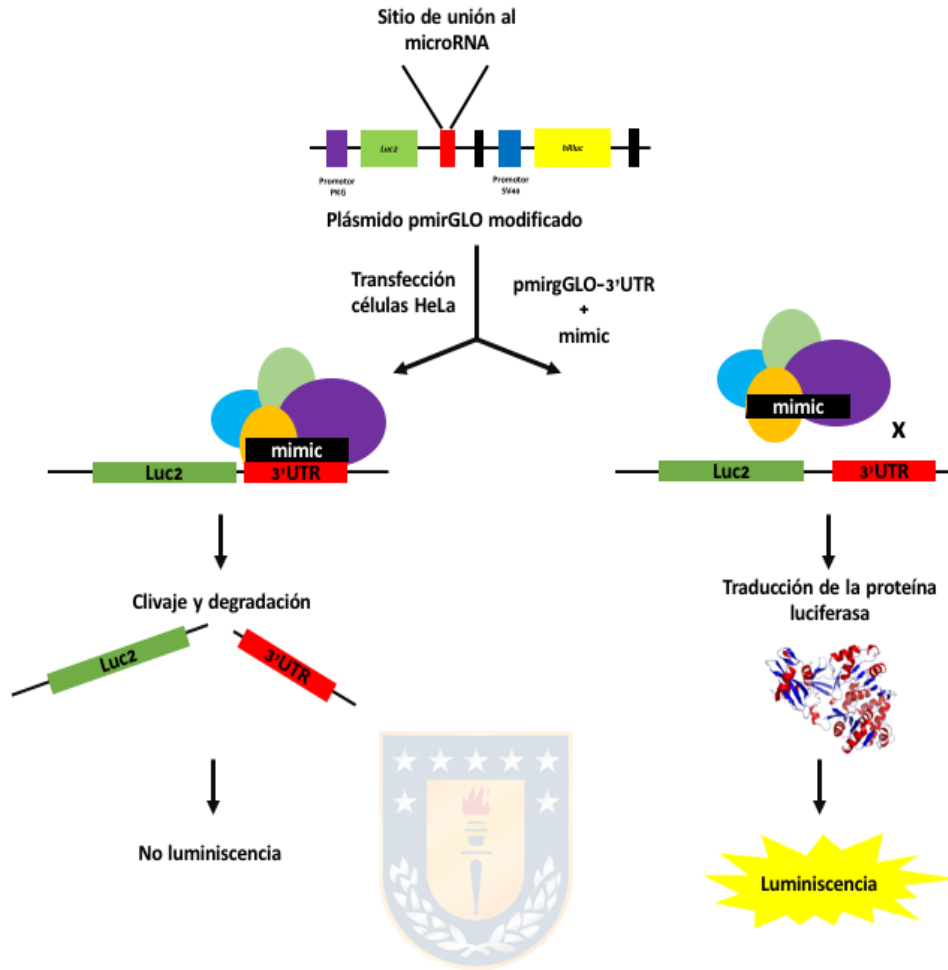
Primer	Secuencia 5'-3'	%GC	Tm (°C)	Producto (pb)
Cr_Catp_Bantam_F1	GTTTAAACCTTGACATTTTCAATTAGGCTCCTTACCTC	36,8	56	633
Cr_Catp_Bantam_R1	TCTAGAGGATCAGTCGATCAATATAACCTTTATGGAGCAACTCTGG	41,3		
Cr_Serp_Bantam_F1	GTTTAAACAAATGCTTGGCCGTATGCGGAAATAAATCA	36,8	60	685
Cr_Serp_Bantam_R1	TCTAGAGAGAAAAAGACATCCGAATTGTGTCCCTCCT	43,2		
Cr_Trip_Bantam_F1	GTTTAAACCCAGTAAAAAAGATCGGAAGAGCGTCGTGTAG	40,5	50	293
Cr_Trip_Bantam_R1	TCTAGAGAGCGTCACTTGACAGGAGAGTTCAAAAAGA	43,2		
Cr_Vit1_utr_mir996_F1	GTTTAAACGAAATCACCCGGAACCATACTGAA	59	68	221
Cr_Vit1_utr_mir996_R1	TCTAGAGCGAATTGGGCCCGACGTCCG	62		

GTTTAAAC Sitio de corte por la enzima de restricción PmeI.

TCTAGA: Sitio de corte por la enzima de restricción XbaI.

### **3.7 Validación funcional de miRNA, ensayo de luciferasa.**

La secuencia 3' UTR del gen de interés se subclonó en el vector pmir-GLO aguas abajo del gen de la luciferasa de la luciérnaga. El nivel de transcripción de la luciferasa se regula por su interacción con el miRNA, lo que da como resultado una actividad variada de la luciferasa cuantificable mediante ensayo de luminiscencia, en otras palabras, si el mimic (miRNA) se hibrida a la zona blanco (ARNm), se bloquea la traducción de luciferasa y no hay señal (Fig. 5). Para la determinación de la actividad de la luciferasa se empleó el kit “Dual-Glo Luciferase Assay System” (Promega), siguiendo el protocolo del fabricante. Los plásmidos pmirGLO\_3'UTR se co-transfectaron posteriormente en células HeLa (Adenocarcinoma) de acuerdo al protocolo descrito por Jin et al. (2013b). La transfección de células eucariontes se realizó siguiendo el método basado en la introducción del DNA plasmídico mediante la utilización de lípidos catiónicos que debido a su carga positiva se unen cargas negativas de los fosfatos del DNA y facilitan la fusión del DNA con la membrana celular, en este caso se utilizó Lipofectamina 3000 según las instrucciones del fabricante. Se analizarán cuatro repeticiones biológicas con tres repeticiones técnicas.



**Figura 5.** Modulación de la función de miRNA. Se insertó un fragmento de ARNm que contiene el sitio de unión a un miRNA en el sitio de clonación múltiple del vector pmirGLO (Promega) en el extremo 3' del gen de la luciferasa. Tras la transfección en células HeLa, miRNA se une a su sitio de unión dentro del plásmido pmirGLO\_3'UTR, causando la degradación del ARNm de luciferasa y, por lo tanto, evitando la luminiscencia. La ausencia de interacción del miRNA y el fragmento 3'UTR clonado, permite la traducción del ARNm de luciferasa de Renilla y se observa luminiscencia (Fuente: Elaboración propia).

### 3.8 Lipofección de células HELA

Para los experimentos de ganancia de función de miRNA Bantam y mir-996 en células HELA, se utilizó el método de Lipofectamina según las recomendaciones del proveedor. Se utilizó como control de transfección un vector que contiene la



proteína verde fluorescente (GFP), la cual, al ser iluminada con luz ultravioleta, produce una brillante fluorescencia verde. Brevemente, las células fueron crecidas en multiplacas de 96 pocillos por 24 h hasta un 90% de confluencia, fueron lavadas con PBS 1X estéril para luego adicionar 50 uL de medio de cultivo. Luego, se prepararon de forma separada la lipofectamina y la mezcla del plásmido pmirGLO\_3'UTR (200ng por pocillo) y los *mimic* (30 y 100mM) en 10 µL de medio de cultivo. Luego se formaron los liposomas al mezclar la lipofectamina y *mimic* (*Bantam* y *mir-996*) durante 15 min a temperatura ambiente. Los 10 µL de liposomas fueron agregados e incubados durante 24 y 48 h a 37°C en estufa de cultivo. Al cabo de este período, se reemplazó el medio con DMEN fresco. El efecto de la presencia de bantam y mir-996 fue evaluado 24 y 48 h posterior a la transfección por medio del equipo Synergy H1 (BioTek). Como control se adicionó a los liposomas un plásmido que incluye la secuencia de la proteína fluorescente verde (GFP 200ng por pocillo).

### **3.9 Cultivo de células HeLa**

Las células HeLa (células epiteliales de adenocarcinoma cervicouterino, No. ATCC: CCL-2) fueron cultivadas en placas de 100 x 20 mm con 8 mL de medio de cultivo DMEM alto en glucosa (Dulbecco's Modified Eagle's Medium, HyClone) suministrado como DMEM, 13,4 g/L; glucosa, 4500 mg/L, NaHCO<sub>3</sub>, 3,7 g/L) suplementado con suero bovino fetal (SBF, HyClone) al 5% v/v; Penicilina 100 U/mL-Estreptomicina 100 ug/mL (HyClone) y L-glutamina (HyClone), 2 mM. La incubación fue realizada en estufa de cultivo a 37°C en atmósfera con un 5% de

CO<sub>2</sub>. Para realizar los pasajes o siembras en pocillos, el medio de cultivo fue aspirado y descartado desde la placa, la que fue lavada dos veces con PBS 1X pH 7,4 (NaCl, 137 mM; KCl, 2,7 mM; Na<sub>2</sub>HPO<sub>4</sub>, 4,3 mM; KH<sub>2</sub>PO<sub>4</sub>, 1,4 mM; pH 7,4). Luego, se incubó con 500 uL de tripsina al 0,25% p/v (HyClone) por 2 minutos a 37°C.



## RESULTADOS

### 1. Identificación de proteasas a partir de datos de secuenciación masiva de *C. rogercresseyi*.

A partir de las secuencias predichas para el secretoma de *C. rogercresseyi* se identificaron las proteasas Cr\_Catepsina, Cr\_Serpina y Cr\_Tripsina previamente reportadas como proteasas altamente expresadas durante el estado infectivo del ectoparásito correspondiente al estado de copepodito (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015; Maldonado-Aguayo, W & Gallardo-Escárate, 2014; Valenzuela-Miranda & Gallardo-Escárate, 2016). De acuerdo a los análisis bioinformáticos de las secuencias del ARNm de Cr\_Catepsina, Cr\_Tripsina y Cr\_Serpina, se identificó que estaba completa la región 3'UTR. La región 3'UTR para Cr\_Vitelogenina I fue obtenida mediante RACE.

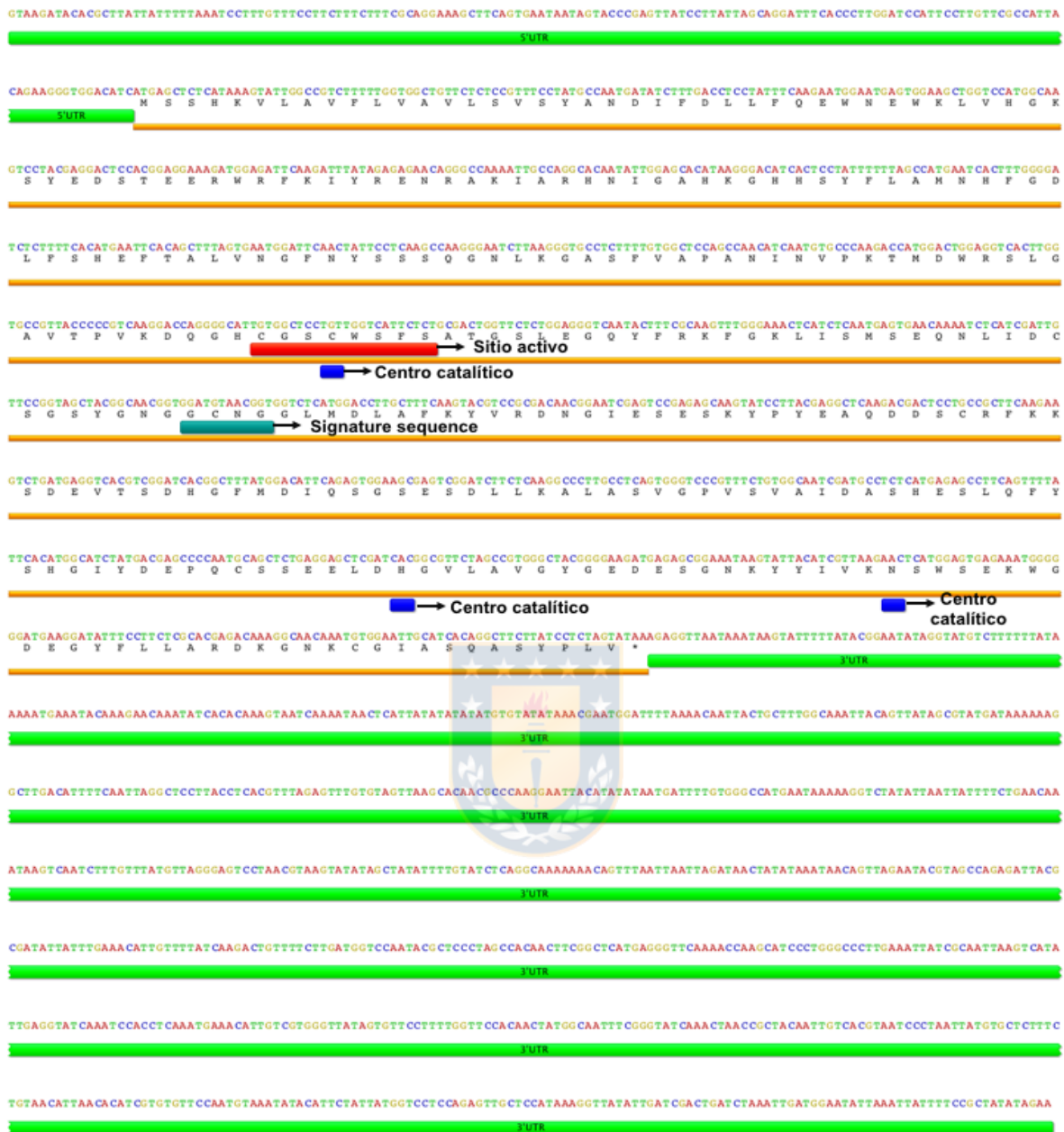
#### 1.1 CATEPSINA

A partir de las secuencias anotadas para Cr\_Catepsina, se encontró que el Contig 0017956 para Catepsina de *C. rogercresseyi* contiene una región 5'UTR de 151 nucleótidos, un extremo 3'UTR de 862 nucleótidos y una ORF de 1011 nucleótidos que codifican para una proteína de 376 aminoácidos, correspondiente a Cr\_Catepsina L (Fig. 6). Los análisis estructurales para la secuencia se llevaron a cabo mediante el programa SMART, encontrando que la secuencia perteneciente al grupo L corresponden a la familia de papaína de las cisteín proteasas con un dominio Pept\_C1, además de presentar un motivo de inhibición (Inhibitor I29).

## **Análisis filogenético**

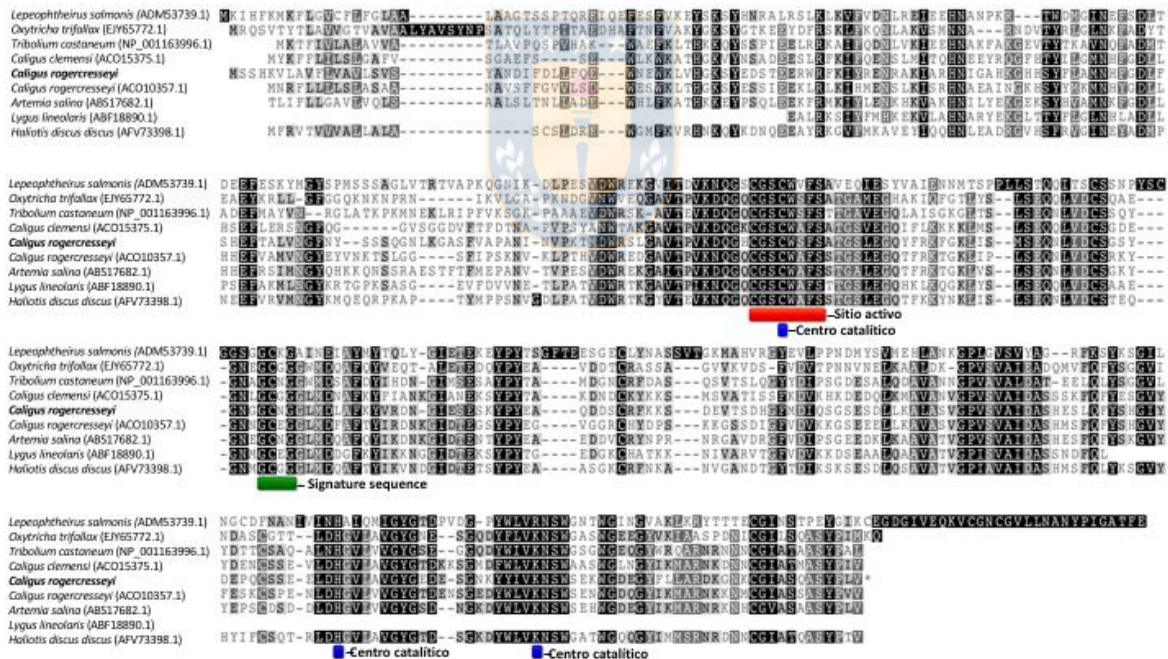
El análisis filogenético pretende investigar la relación evolutiva de las secuencias encontradas en este estudio. Para esto, se realizó un árbol filogenético, donde la robustez de los nodos fue soportada por un alto valor de bootstrap (Fig. 8), cada secuencia de catepsina, fue agrupada diferencialmente, donde en el clado de las catepsinas L, se encuentra Cr\_Catepsina de *C. rogercresseyi* marcada en color naranja, con *Artemia salina* y *Caligus clemensi*. Además, se formaron otros clados donde se agruparon otras catepsinas como la C, B y la K.



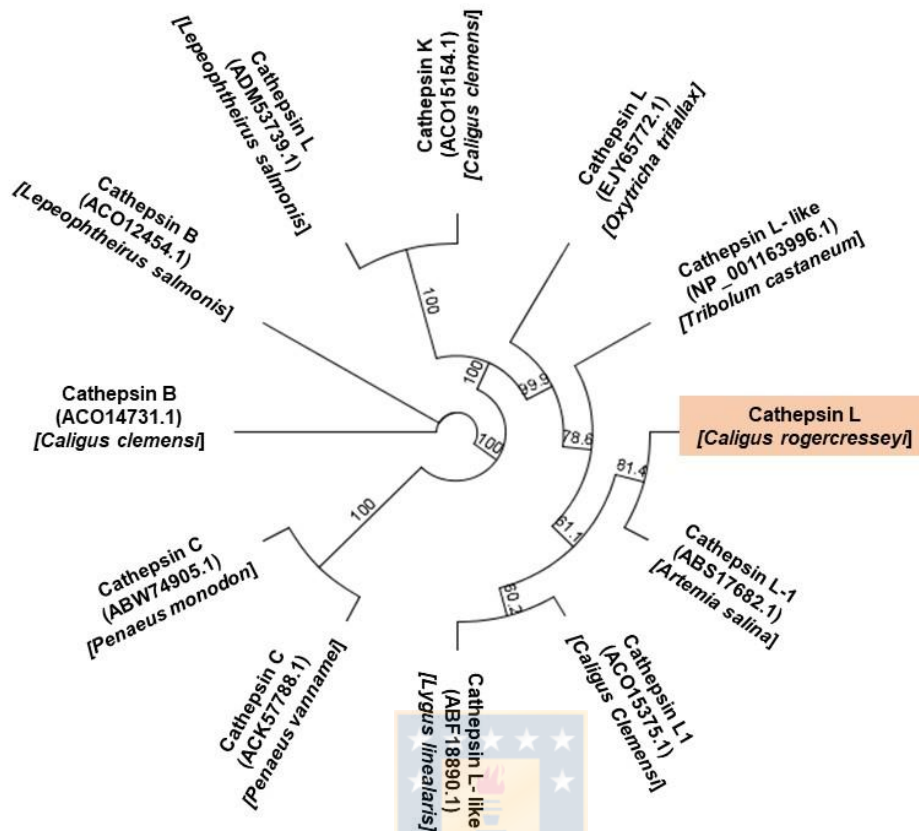


**Figura 6.** ARN mensajero de Cr\_Catepsina de *C. rogercresseyi*. La región codificante para Cr\_Catepsina, está representada de color amarillo. Los extremos 5 y 3' UTR están representados en color verde. El sitio activo se encuentra representado por un cuadro de color rojo, en cuadros azules los aminoácidos que forman el centro catalítico y en rectángulo verde la signature sequence (Fuente: Elaboración propia).

Por otra parte, el alineamiento de secuencias de catepsina L con otros organismos artrópodos, muestra que el transcrito denotado como Cr\_Catepsina L de *C. rogercresseyi* presenta altos niveles de conservación, identificando la tríada catalítica correspondiente a cisteína, histidina, asparragina, lo cual fue altamente conservado entre las especies analizadas (CGSCWSFS) y la secuencia señal (GCNGG) (Fig. 7). De los resultados del alineamiento Cr\_Catpesina\_L de *C. rogercresseyi* presentó un 53,571% de identidad con Catepsina L de *Artemia salina*, un 46,849% con Catepsina L de *Caligus clemensi* y un 24,165% con *Lepeophtheirus salmonis*.



**Figura 7.** Alineamiento de la proteína Cr\_Catepsina L de *C. rogercresseyi*. *Lepeophtheirus salmonis* (ADM53739.1), *Oxytricha trifallax* (EJY65772.1), *Tribolium castaneum* (NP\_001163996.1), *Caligus clemensi* (ACO15375.1), *Caligus rogercresseyi* (ACO10357.1), *Artemia salina* (ABS17682.1), *Lygus lineolaris* (ABF18890.1), *Haliotis discus discus* (AFV73398.1). En rojo se representa el sitio activo, en cuadros azules los aminoácidos que forman el centro catalítico y en rectángulo verde la Signature sequence (Fuente: Elaboración propia).



**Figura 8.** Relación filogenética de Cr\_Catepsina de *C. rogercresseyi*. El análisis se realizó con otras catepsinas (Catepsina B, L, K, C) de diferentes organismos: *Tribolium castaneum* (NP\_001163996.1), *Lepeophtheirus salmonis* (ADM53739.1, ACO12454.1), *Caligus clemensi* (ACO15375.1, ACO15154.1, ACO14731.1), *Oxytricha trifallax* (EJY65772.1), *Penaeus vannamei* (ACK57788.1), *Penaeus monodon* (ABW74905.1), *Artemia salina* (ABS17682.1) y *Lygus lineolaris* (ABF18890.1). Catepsina de *C. rogercresseyi* se encuentra demarcada en color naranja. El árbol fue construido por el método de Neighbor-Joining (Fuente: Elaboración propia).

## 1.2 SERPINA

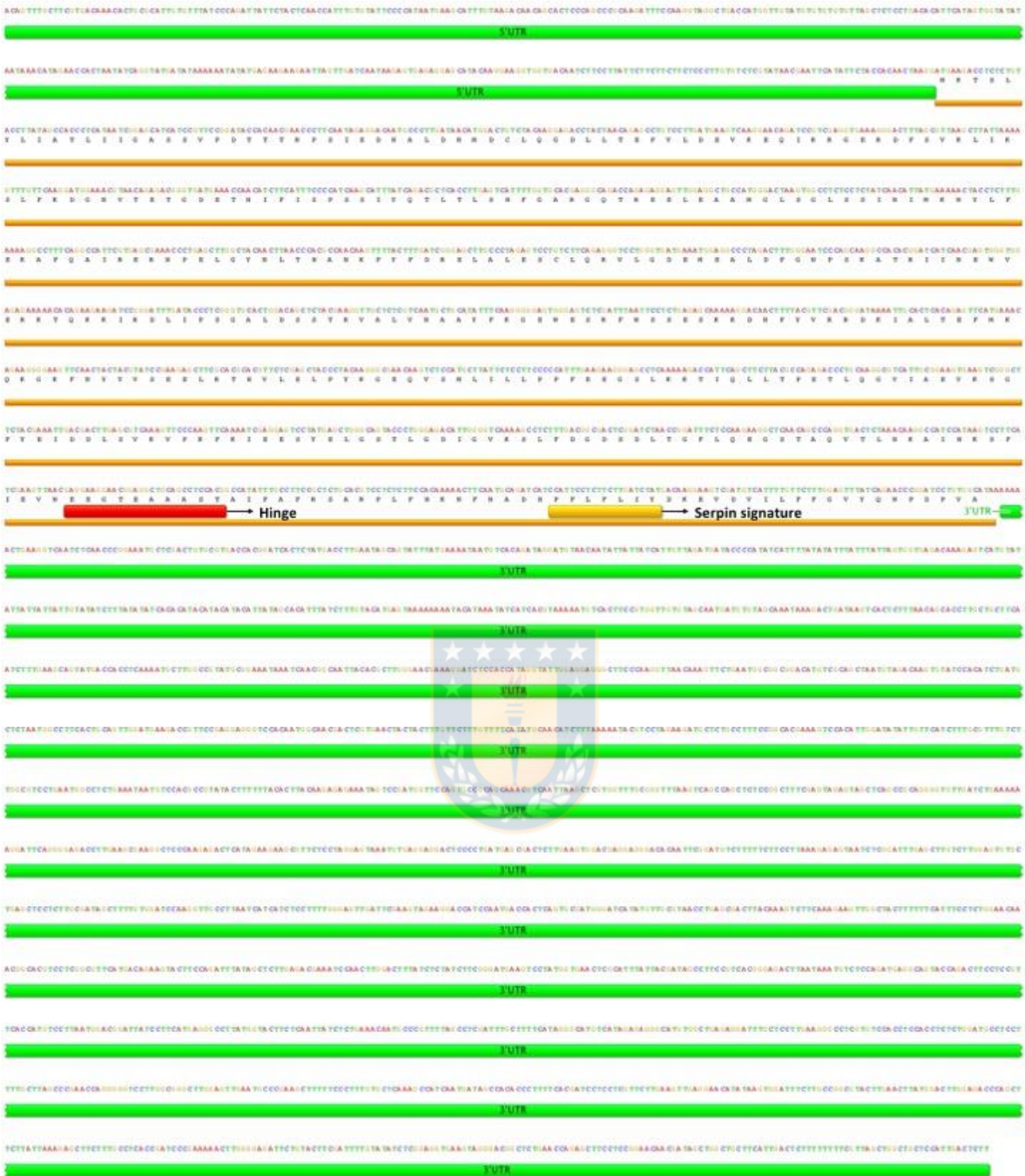
Cr\_Serpina (Contig 0008452) contiene una región 5'UTR de 356 nucleótidos, un extremo 3'UTR de 2077 nucleótidos y una ORF de 1334 nucleótidos que codifican para una proteína de 444 aminoácidos correspondiente a Cr\_Serpina 10 (Fig. 9). El análisis estructural a través del programa de SMART mostró que serpina

presentaban el motivo serpin característico. El alineamiento de secuencias aminoacídicas entre serpinas confirma la presencia del motivo serpin característico de estas proteínas, incluyendo sitios conservados como el bucle de centro reactivo (RCL) y el péptido señal (Fig. 10).

### **Análisis filogenético**

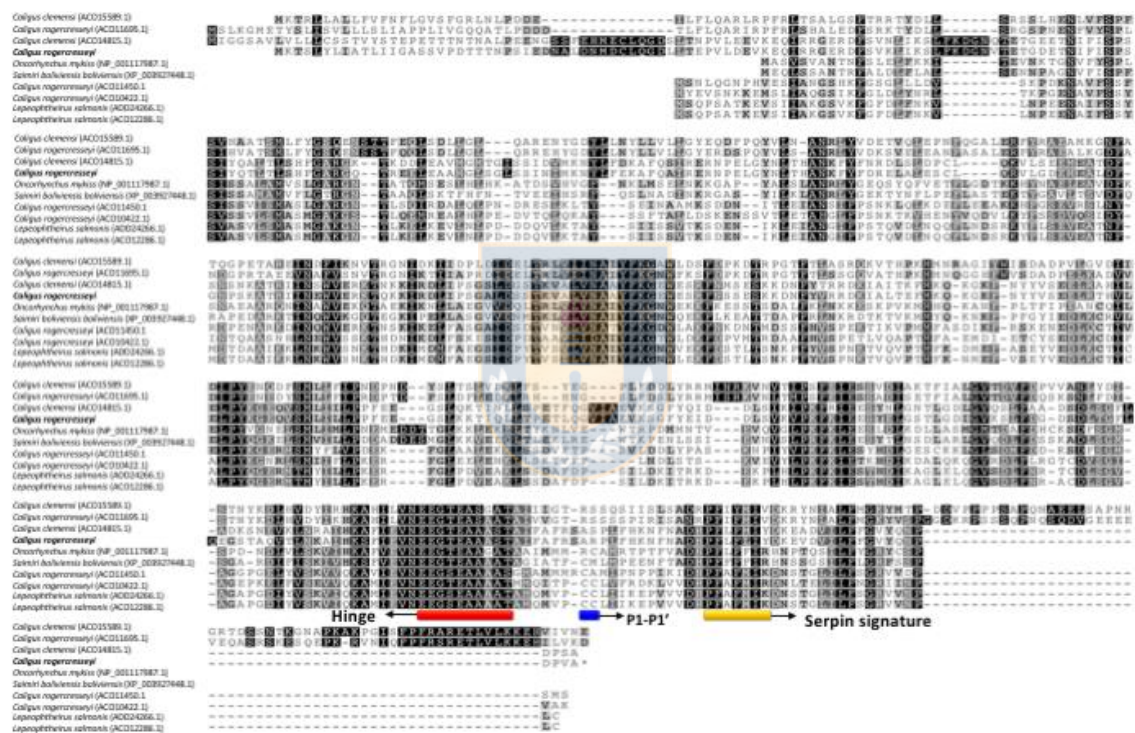
La relación evolutiva para Cr\_Serpina 10 encontrada en el transcriptoma de *C. rogercresseyi* se evaluó mediante un alineamiento de secuencias aminoacídicas y la realización de un análisis filogenético, utilizando el método del vecino más cercano (NJ) con el modelo distancia genética de Jukes-Cantor. Por otro lado, el análisis de homología, a partir del alineamiento con otras secuencias reveló que Cr\_Serpina 10 de *C. rogercresseyi* presentó un 77,33% de identidad con Serpina B4 de *Caligus clemensi*, un 31,633% con Leukocyte elastase inhibitor de *Saimiri bolivianis bolivianis*, y un 21,586% con Serpina B8 de *Caligus rogercresseyi* (Fig. 10).



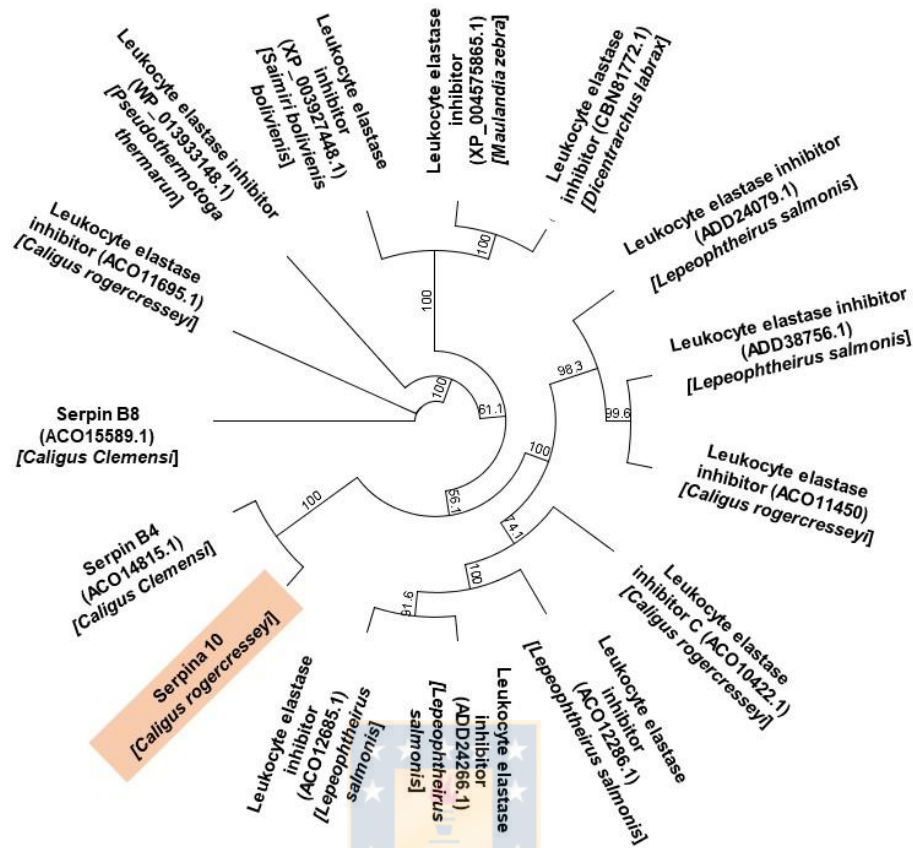


**Figura 9.** ARN mensajero de *Cr\_Serpina* de *C. rogercresseyi*. La región codificante para *Cr\_Serpina*, está representada de color amarillo. Los extremos 5 y 3' UTR están representados en color verde. El sitio hinge se encuentra representado en color rojo y la secuencia serpin signature en color amarillo (Fuente: Elaboración propia).

Los resultados del análisis filogenético, mostraron nodos soportados un bootstrap de 1000 iteraciones (Fig. 11). El árbol filogenético, muestra que las secuencias anotadas para inhibidores de elastasa de leucocitos forman un clado con otras secuencias anotadas para el mismo inhibidor (Fig. 11). Cr\_Serpina de caligus se agrupa en el mismo clado que Serpina B4 de *Caligus clemensi* soportado con un valor nudo de 100%.



**Figura 10.** Alineamiento de Cr\_Serpina de *C. rogercresseyi*. *Caligus clemensi* (ACO15589.1, ACO14815.1), *Caligus rogercresseyi* (ACO11695.1, ACO11450.1, ACO10422.1), *Oncorhynchus mykiss* (NP\_001117987.1), *Saimiri boliviensis boliviensis* (XP\_003927448.1), *Lepeophtheirus salmonis* (ADD24266.1, ACO12286.1). El sitio hinge se encuentra representado en color rojo y la secuencia serpin signature en color amarillo (Fuente: Elaboración propia).



**Figura 11.** Relación filogenética de Cr\_Serpina de *C. rogercresseyi*. El análisis se realizó con otros miembros de la superfamilia serpina (Leucocyte elastase inhibitor, SERPINs B8 y B4) de diferentes organismos: *Maylandia zebra* (XP\_004574865.1), *Saimiri boliviensis boliviensis* (XP\_003927448.1), *Pseudothermotga thermanun* (WP\_013933148.1), *Dicentrarchus labrax* (CBN81772.1), *Lepeophtheirus salmonis* (ADD38756.1, ADD24266.1, ADD24079.1, ACO12685.1, ACO12286.1), *Caligus clemensi* (ACO15589.1, ACO14815.1), *Caligus rogercresseyi* (ACO11695.1, ACO11450.1, ACO10422.1). Cr\_Serpina 10 de *Caligus rogercresseyi* se encuentra marcada en color. El árbol fue construido por el método de Neighbor-Joining (Fuente: Elaboración propia).

### 1.3 TRIPSINA

El ARNm de Cr\_Tripsina (Contig 0013229) contiene una región 5'UTR de 686 nucleótidos, un extremo 3'UTR de 477 nucleótidos y una ORF de 774 nucleótidos que codifican para una proteína de 257 aminoácidos (Fig. 12).

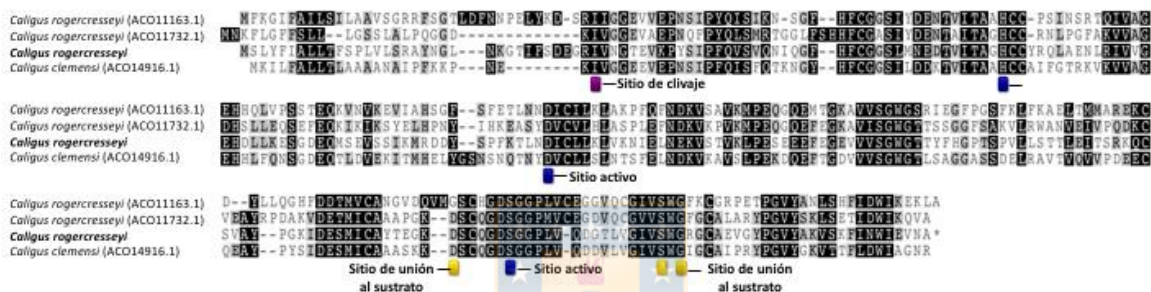


**Figura 12.** ARN mensajero de Cr\_Tripsina de *C. rogercresseyi*. La región codificante para *Cr\_Tripsina*, está representada de color amarillo. Los extremos 5 y 3' UTR están representados en color verde. El sitio de clivaje se encuentra representado en un cuadro de color azul, en color morado se representa el sitio de unión al sustrato y en cuadros de color azul el sitio activo (Fuente: Elaboración propia).

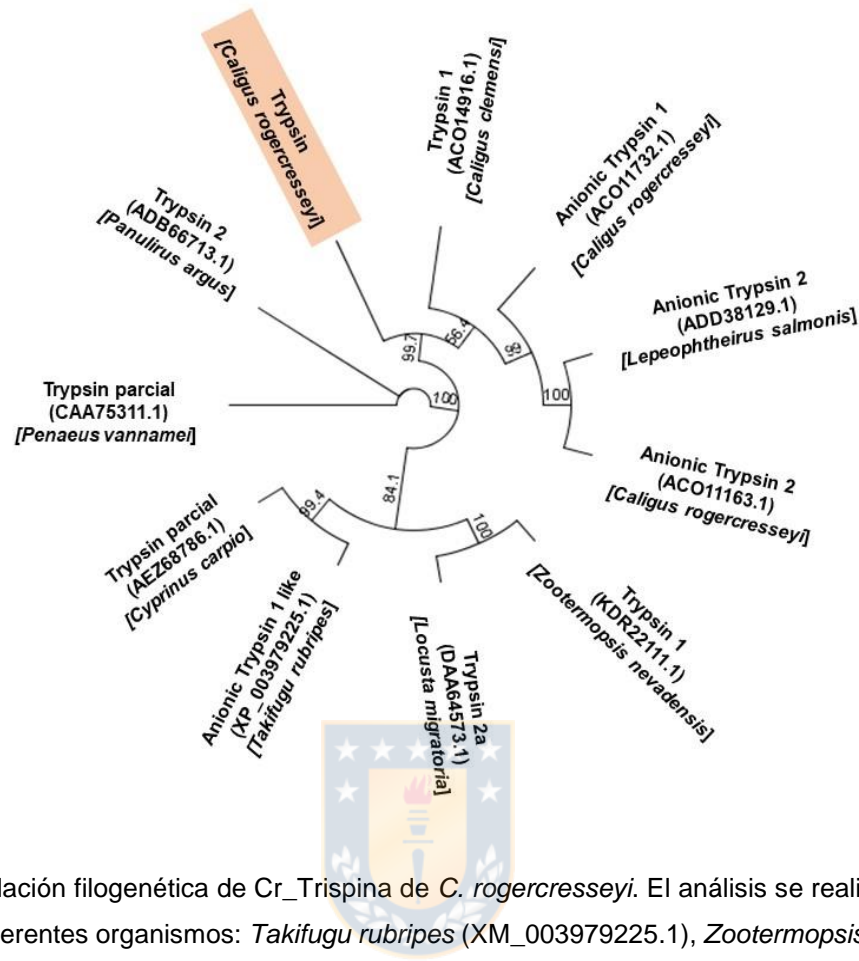
### Análisis filogenético

El análisis estructural a partir de SMART muestra que las secuencias de tripsina contienen el dominio de proteasa de serina, además de una actividad endopeptidasa putativa. El análisis de BLASTx muestra que la secuencia de aminoácidos predicha es homóloga con tripsinas de otras especies de crustáceos

incluyendo a *C. clemensi* y otras secuencias de *C. rogercresseyi* (Fig. 13), con una identidad de 40,38 y 39,47% para otras tripsinas de *C. rogercresseyi* y un 47,86% para tripsina de *C. clemensi*. Además, el alineamiento múltiple entre secuencias muestra los residuos de la triada catalítica H, D y S característica de las proteasas de serina. El árbol filogenético, muestra que las secuencias anotadas tripsina forman un clado con tripsina 1 de *Caligus clemensi*, y tripsina anionica 1 y 2 de *C. rogercresseyi* (Fig. 14).



**Figura 13.** Alineamiento de Cr\_Tripsina de *C. rogercresseyi*. *C. rogercresseyi* (ACO11163.1), *C. rogercresseyi* (ACO11732.1), *C. clemensi* (ACO14916.1) El sitio de clivaje se encuentra representado en un cuadro de color morado, en color amarillo se representa el sitio de unión al sustrato y en cuadros de color azul el sitio activo (Fuente: Elaboración propia).



**Figura 14.** Relación filogenética de Cr\_Tripsina de *C. rogercresseyi*. El análisis se realizó con otras tripsinas de diferentes organismos: *Takifugu rubripes* (XM\_003979225.1), *Zootermopsis nevadensis* (KDR22111.1), *Lacusta migratoria* (DAA64573.1), *Penaeus vannamei* (CAA75311.1), *Cyprinus carpio* (AEZ68786.1), *Lepeophtheirus salmonis* (ADD38129.1), *Panulirus argus* (ADB66713.1), *C. clemensi* (ACO14916.1), *C. rogercresseyi* (ACO11732.1, ACO11163.1). Cr\_Tripsina de *C. rogercresseyi* esta marcada en color. El árbol fue construido por el método de Neighbor-Joining (Fuente: Elaboración propia).

#### 1.4 VITELOGENINA I

De acuerdo a los datos de secuenciación masiva del LBGA solo se obtuvo una secuencia parcial del gen Vitelogenina I la cual se usó como una plantilla para el diseño de primers que permitió la amplificación la ORF parcial del gen de vitelogenina I la cual se utilizó para realizar la amplificación del extremo 3'UTR de Vitelogenina I de *Caligus rogercresseyi* mediante RACE.

### 1.4.1 Clonamiento y secuenciación del extremo 3' de la ORF y RACE de Vitelogenina I

Para la amplificación del extremo próximo al 3'UTR del gen de la Cr\_Vitelogenina I de *C. rogercresseyi* se diseñaron 4 set de partidores, los cuales se solapaban en aproximadamente 100 nucleótidos para asegurar un mejor ensamble de las secuencias. De los resultados de secuenciación, se obtuvieron 4 fragmentos de 601, 566, 466 y 503 nucleótidos de longitud (Fig. 15, 16 y 17). El ensamble de los fragmentos dio un contig de 1698 nucleótidos (Fig. 18), que al ensamblarlo con el contig Contig10477 (734 nucleótidos) obtenido por secuenciación masiva se obtuvo una secuencia consenso de 2356 nucleótidos. El resultado del blast arrojó un 73% de similitud con *Lepeophtheirus salmonis* (Fig. 19).

A partir del nuevo contig para vitelogenina I de 2356 nucleótidos se diseñaron un nuevo set de partidores forward para realizar la amplificación del extremo 3'UTR mediante la técnica 3'RACE. De acuerdo a los resultados se obtuvo una región 3'UTR de 120 nucleótidos el cual presentaba la cola de Poli-Adenilación (Fig. 18).

**A**

```

1      10      20      30      40      50      60      70      80
CTTCTCAGTGGTTATTGCT GATTCCAAGACTCGCCTTCC TTCCATCCGC TCCAGATGGA ACTTGGAGGAGATGTTGAGT
90      100     110     120     130     140     150     160
GACGAGATC TCAATGGAACA CGACATGGT CGTTTACTACG GAAAGCGCTCCACTCCCGC GATGGAATGTCCAAGTTCGT
170     180     190     200     210     220     230     240
TGTCGAGGCATTTGCCTCCA AGTCTGAAGGACTCAAGACC TCCGTTATCGGGTCCCCCGA ATACGTTCTCTGCAACAAGG
250     260     270     280     290     300     310     320
AAATCGGAGAAGGCCGCACC CTCGCCCTGTTTGGGAGAA GCTTCGTCACTCGCCACCT CCGTCGACTCCTTCAGAAT
330     340     350     360     370     380     390     400
CGCACCAAGT TCCCATGGAG CAGCACTACC CAATCCACCT ATACCAAGTCTTACATCCGC AACTTCCTCGAAACCGTCTT
410     420     430     440     450     460     470     480
CTATCCCTACTTCACCCGAAC ACTACTTCGAGGCCACCTCT GCGGTTGAAGGAATCCCTC TGACGAGACCCCTTTTCGAAG
490     500     510     520     530     540     550     560
CCTCCGTTTCTCGCGAAGGT ACCTTGGCTCAAATGAAGTA CAAGGAGCATGGACTCGAAT GGAACCTTCCCAACGTTCCG
570     580     590     600     601
CTCCCAAGCCCTTGACCA ACACGTCCTTCCATGGAGCT T

```

**B**

```

1      10      20      30      40      50      60      70      80
GTCCTACATCCGCAACTTCC TCGAAACCGTCTTCTATCCC TACTTCACCGAACACTACTT CGAGGCCACCCTTGCCGTTG
90      100     110     120     130     140     150     160
AAGGAATCCCTCTGACGAG ACCCTTTTCGAAGCCCTCCGT TTCTCGCGAAGGTACCTTGG CTCAAATGAAGTACAAGGAG
170     180     190     200     210     220     230     240
CATGGACTCGAATGGAACCT TCCCAACGTT CGCTCCCA AGCCCTTGACCCAACACGTC CTTCCATGGAGCTTCCGCAA
250     260     270     280     290     300     310     320
CTCCCAAACTCCACCGGAT CCCGATGGGTCAAAAAGAT ACCTCTCAACAATCCCCCGC CTCCTGTACCGTTGAACCCA
330     340     350     360     370     380     390     400
ACTTTGTCTCCACCTTCGAC AACAAGACCTATCCCTTACAC CCTCAACGACTGCGAACACC TCGTCGTCAAGGACTGCTCC
410     420     430     440     450     460     470     480
GGTATGATCCCATGGCCGT CACCGCCCGCAAGAGCGGCC CATCCATGGAAGTCAAGATG GTTGTGGAGGAATGATCGT
490     500     510     520     530     540     550     560
TTACCTGACCCCATCCAGG GATCCGTCATCAACAACGTC ACCATCAACGGCAAGCAGGC TCTTCTGTCTTGGACCTCT
570     580     590     600
TCTTTC

```

**C**

```

1      10      20      30      40      50      60      70      80
CCATGGAAGTCAAGATGGTT GTTGGAGGAATGATCGTTT CCTGACCCCATCCAGGGAT CCGTCATCAACAACGTACC
90      100     110     120     130     140     150     160
ATCAACGGCAAGCAGGCTCC TCTTGTGCGAGGGTGGCCCTC TCAGATACACTGAGCCCTC CTCTCCGGCTCTGTAAGAC
170     180     190     200     210     220     230     240
CCAAGGACACCTCATTTGCC TCACTCGGACCCGTCGCCACT GAAGGCATAACCCTCTTAA GATCTGGAGCTACCTTGACG
250     260     270     280     290     300     310     320
GAACTGTTGTTGTCAAGCAC ATCAGAACC GGCTCACCGT CATCTTCGATGGAGAGCGTC TCGAGGTCAGCCACCACC
330     340     350     360     370     380     390     400
CTCGTGTCCACCCGGGCTG CGGTATTTGCGGCACATGA ACGGTGAGTCCACCCCGAT CTTGCCCTCCCAAGATGTG
410     420     430     440     450     460     468
CATCTTCGAGAAGCCCGCC ATGCCGCTTTCTTTATATG ATTAAGGATCTTGCCAAGG TATCCC

```

**D**

```

1      10      20      30      40      50      60      70      80
CTCCCCAAGATGTGCATCT TCGAGAAGCCCGCCATGCC GCTTCTCTTATATGATTA GGATTCTTGCCAAGGTATCC
90      100     110     120     130     140     150     160
CCGTCCCCGAGGAGAAGACT GCCTTCGAAAAGCCTCCGC CACCTGCACCCTCAAACC TCGTCGCCACTCCCTTGGCT
170     180     190     200     210     220     230     240
GACCTCCTTAAGCGACTCGT CATGGTCCGCTCCGGACCT C TGGCATCATCACCAGCAC TTGGTCGAGTACCGCGCTAA
250     260     270     280     290     300     310     320
CGGAAATGAAATTTGCTTCT CCCAAACCCCTCCGATC TGCAGTGGTTCGTTCCGTCC CGTCATCGGACACATGGCCT
330     340     350     360     370     380     390     400
CCACTCCCTTCACCTGCATG CAATCCTTCAACGGTTC TG CAAAGCACATCAAGGGACGCA TCTACGCCAACGAAGATC
410     420     430     440     450     460     470     480
CCTGAGCTCAGAAAGTACC CACCACCTACAACAGAATGA TCGAGGAAGCTCCACCTGC CAAACTGGAGGAGATGCCTC
490     500     510
ATCATGCTCAACCCAGCACA GGA

```

**E**

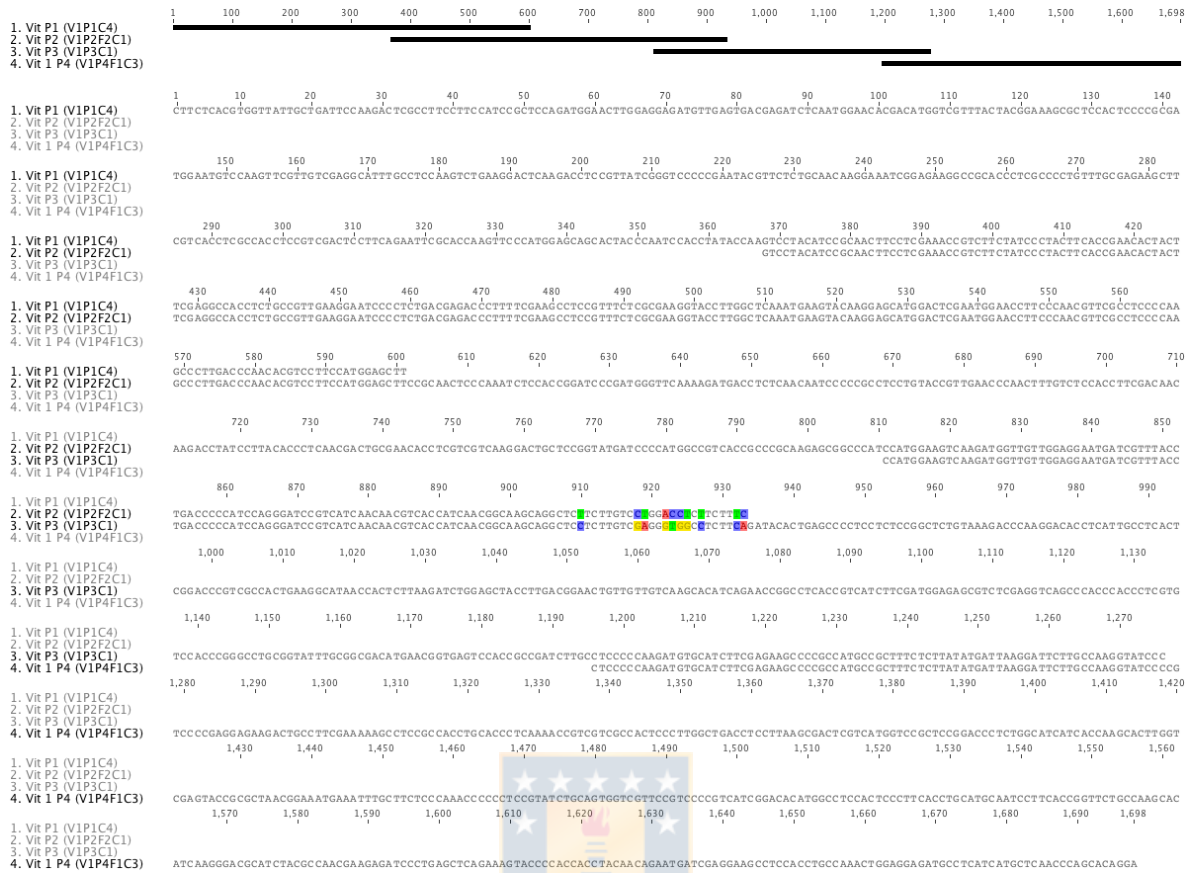
```

1      10      20      30      40      50      60      70      80
CATCGGACACATGGCCTCCA CTCCCTTCACTTCGATGCAA TCCTTACC GGTTCTGCCAA GCACATCAAGGGACGCATCT
90      100     110     120     130     140     150     160
ACGCCAACGAAAGAGATCCC GAGCTCAGAAAGTACCCAC CACCTACAACAGAATGATCG AGGAAGCTCCACTGCCAA
170     180     190     200     210     220     230     240
ACTGGAGGAGATGCCTCATC TGGCTCTCAATCCACATCA ACCCCACTGGAAGTGAAGA TTCCAACCCCAAAATGTATG
250     260     270     280     290     300     310     320
CTACTAAGGAAATCACCCGG AACCATACTGAAAACATTTA TTTTTGAAATGAAAGTACC TCTGCGGAAATTTATATAT
330     340     350     360     364
TAATAATAATAAAGTCATAT CAAATGAATAAAAAAAAAA AAAA

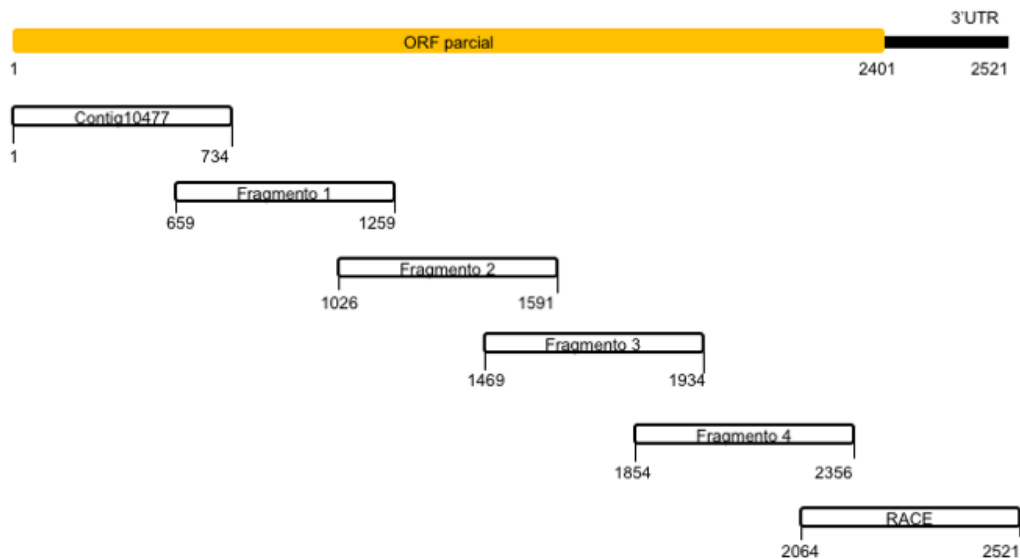
```

**Figura 15.** Fragmentos secuenciados de la ORF parcial de Cr\_Vitelogenina I de *C. rogercresseyi* clonados en el vector pGEM-T. **A.** Fragmento 1, **B.** Fragmento 2, **C.** Fragmento 3, **D.** Fragmento 4, **E.** Fragmento obtenido mediante RACE (Fuente: Elaboración propia).

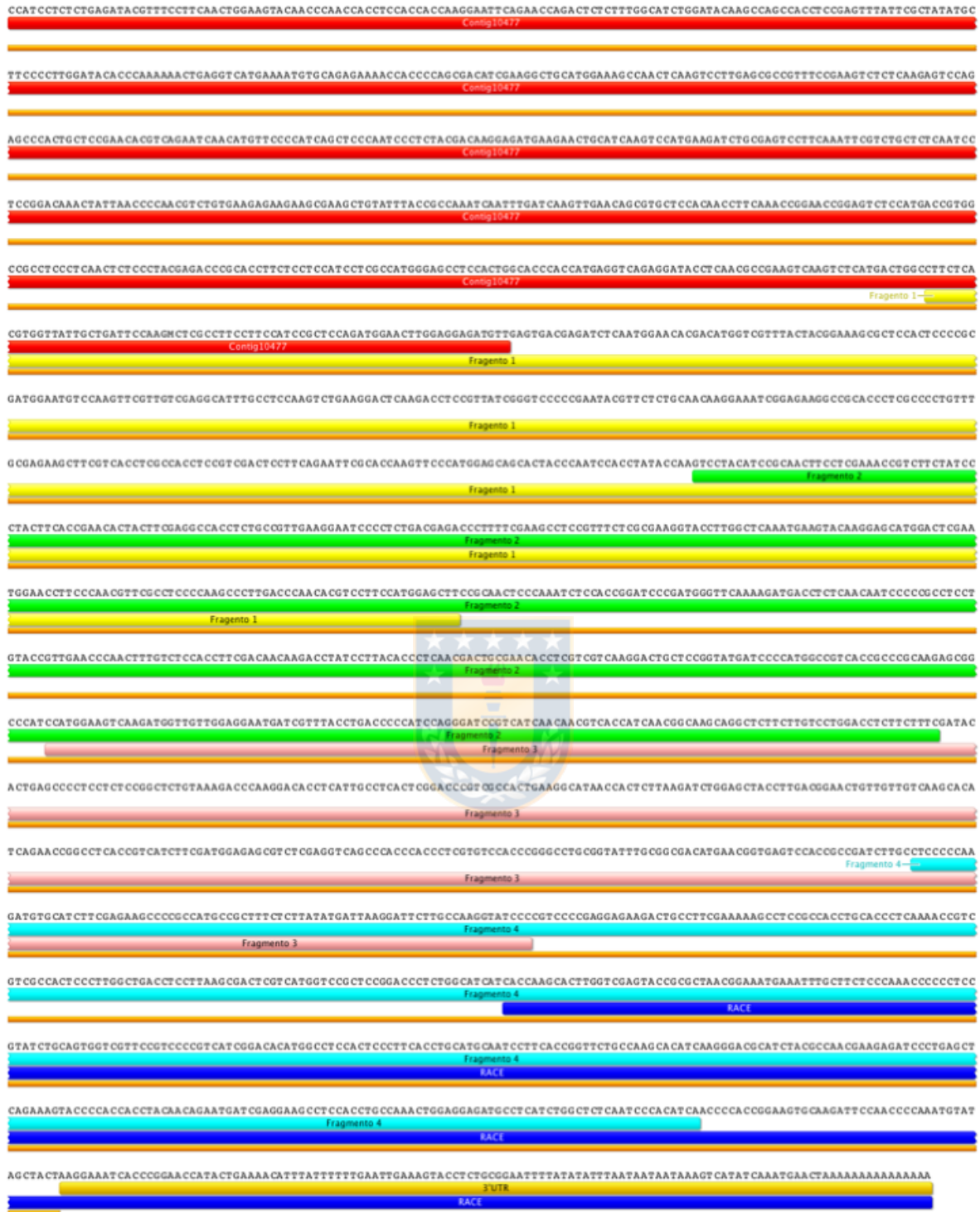




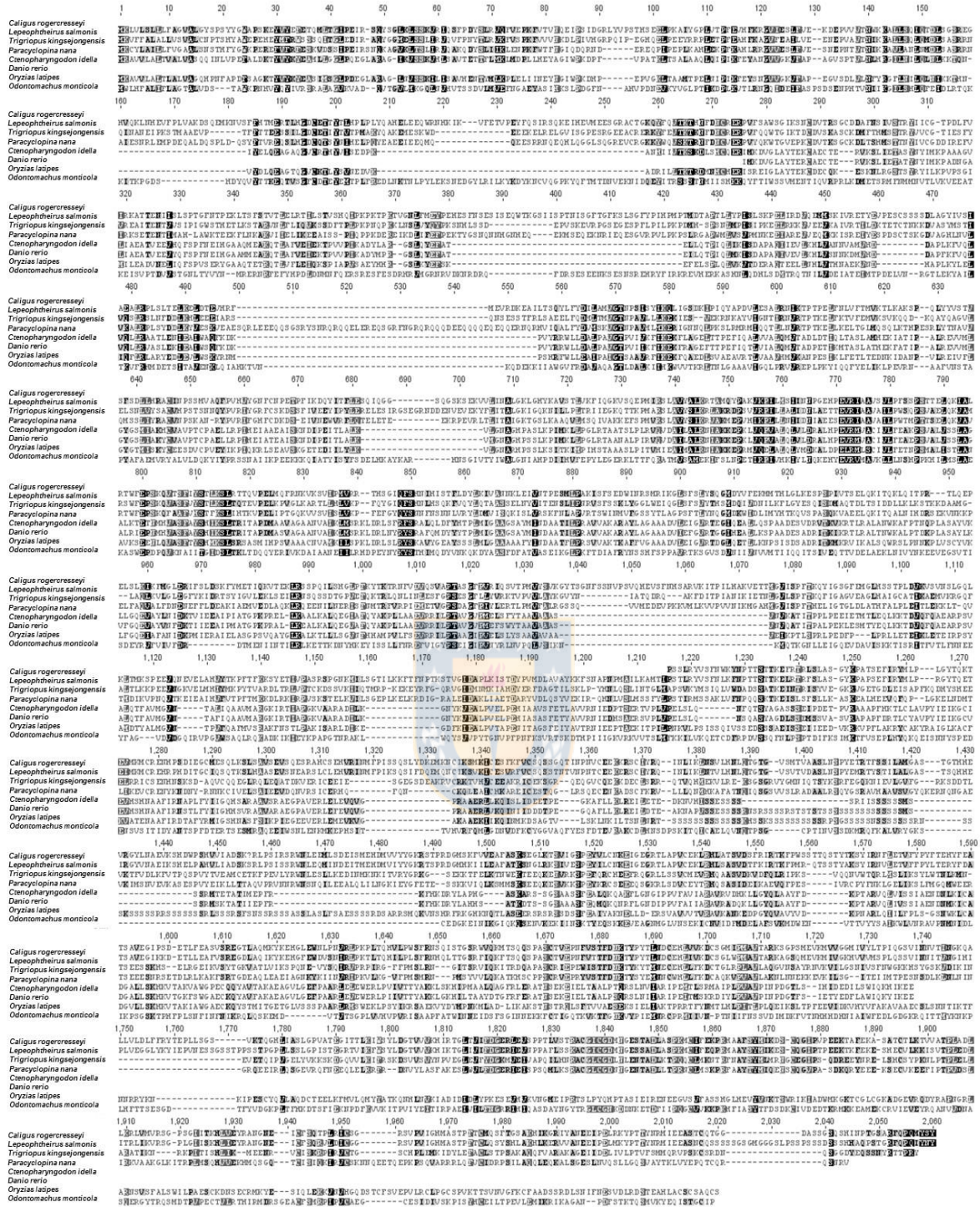
**Figura 16.** Ensamble obtenido a partir de los cuatro fragmentos clonados y secuenciados para la ORF de Cr\_Vitelogenina I (Fuente: Elaboración propia).



**Figura 17.** Representación esquemática de los resultados obtenidos por secuenciación de los fragmentos del gen de Cr\_Vitelogenina I de *C. rogercresseyi* (Fuente: Elaboración propia).

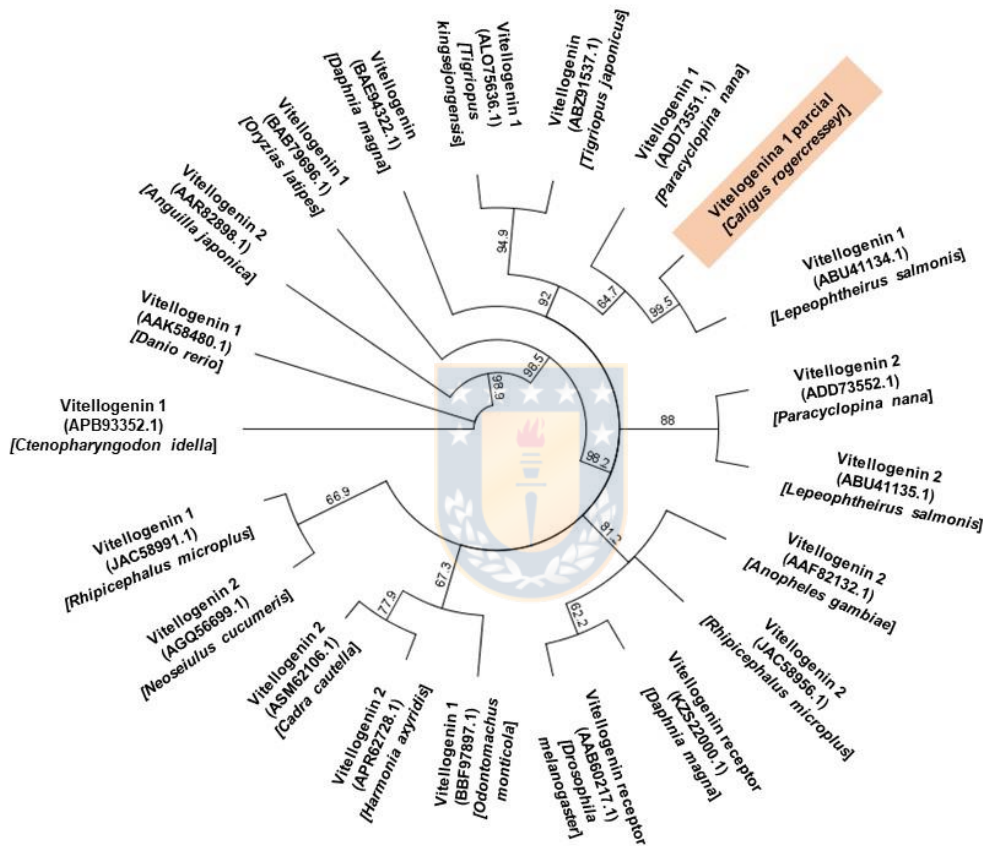


**Figura 18.** Contig obtenido mediante clonamiento y secuenciación para Cr\_Vitelogenina I de *C. rogercresseyi*. Contig10477 está representado en color rojo, los fragmentos 1, 2, 3 y 4 obtenidos mediante clonamiento y secuenciación se representa en color amarillo, verde, rosado y celeste respectivamente. El fragmento obtenido mediante 3'RACE está representado de color azul, la región 3'UTR se encuentra representada en color amarillo (Fuente: Elaboración propia).



**Figura 19.** Alineamiento de Cr\_Vitelogenina I (ORF parcial) de *C. rogercresseyi*. El alineamiento se realizó con Vitelogeninas de otras especies. *Lepeophtheirus salmonis* (ABV41134.1), *Trigriopus kingsejongensis* (ALO75636.1), *Paracyclops nana* (ADD73551.1), *Ctenopharyngodon idella* (APB93352.1), *Danio rerio* (VAAK58480.1), *Oryzias latipes* (BAB79696.1) y *Odontomachus monticola* (BBF97897.1) (Fuente: Elaboración propia).

El árbol filogenético, muestra que las secuencias anotadas Vitelogenina I forman un clado con otras secuencias anotadas para el mismo gen (Fig. 20). Cr\_Vitelogenina I de caligus se agrupa en el mismo clado que Vitelogenina I de *Lepeophtheirus salmonis* soportado con un valor nudo de 99,5%.



**Figura 20.** Relación filogenética de Cr\_Vitelogenina I parcial de *C. rogercresseyi*. *Daphnia magna* (KZS22000.1, BAE94322.1), *Rhipicephalus microplus* (JAC58991.1, JAC58956.1), *Odontomachus monticola* (BBF97897.1), *Oryzias latipes* (BAB79696.1), *Cadra cautella* (ASM62106.1), *Harmonia axyridis* (APR62728.1), *Ctenopharyngodon idella* (APB93352.1), *Tigriopus kingsejongensis* (AL075636.1), *Neoseiulus cucumeris* (AGQ56699.1), *Paracyclops nana* (ADD73552.1, ADD73551.1), *Lepeophtheirus salmonis* (ABU41135.1, ABU41134.1), *Trigriopus japonicus* (ABZ91537.1), *Anguilla japónica* (AAR82898.1), *Danio rerio* (AAK58480.1), *Anopheles gambiae* (AAF82132.1), *Drosophila melanogaster* (AAB60217.1). Cr\_Vitelogenina I de *Caligus rogercresseyi* esta demarcada en color. El árbol fue construido por el método de Neighbor-Joining (Fuente: Elaboración propia).

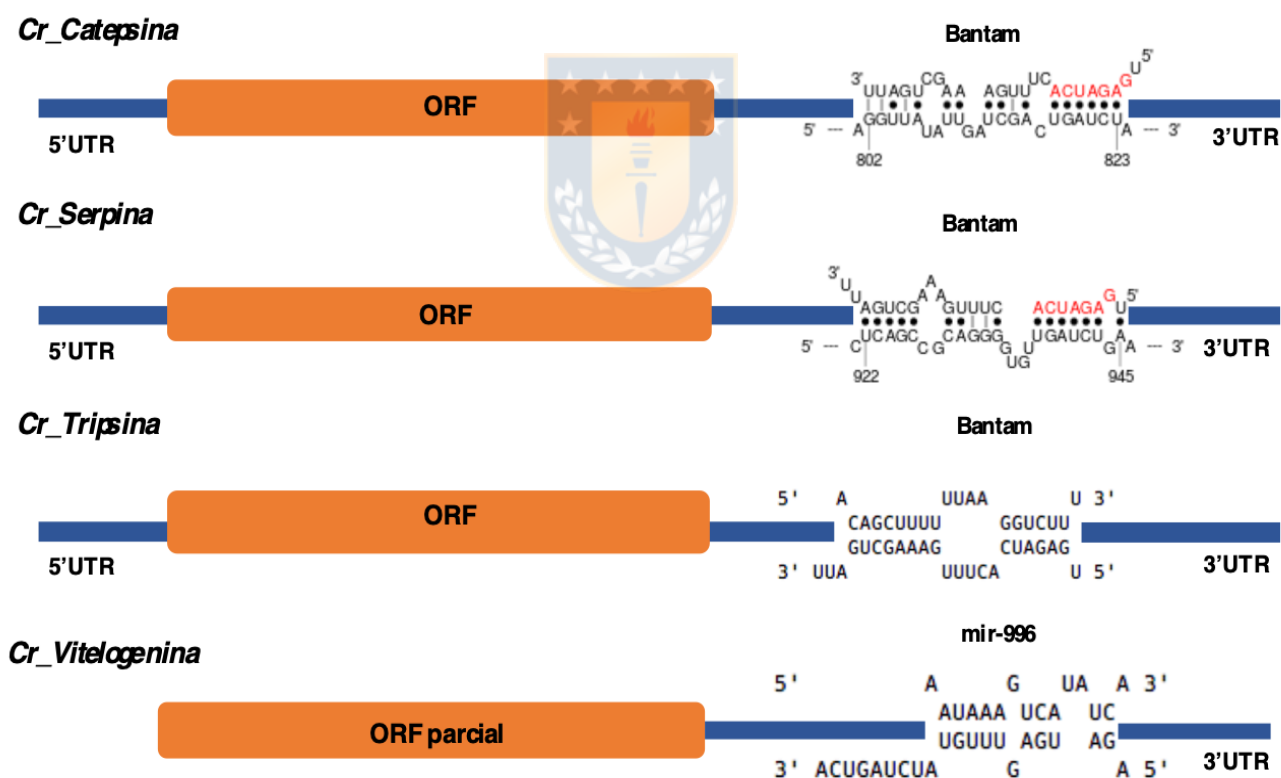
## 1.5 Predicción de target para miRNA

El análisis de predicción de target para los miRNA se realizó utilizando plataformas bioinformáticas, para los análisis de predicción de target para Bantam se identificaron secuencias de unión al miRNA en los tres genes de estudio, Cr\_Catepsina, Cr\_Serpina y Cr\_Tripsina, genes involucrados en la capacidad de infectar al hospedero. Para Cr\_Catepsina la secuencia de unión al miRNA bantam fue “5'-AGGTTATATTGATCGATGA-3'” ubicada en el extremo 3'UTR entre los nucleótidos 802 y 823 con un valor de energía de hibridación de -15,6 kcal/mol (Fig. 21). Para Cr\_Serpina la secuencia de unión al miRNA bantam fue “5'-CTCAGCCGACGGGGTGTGATCTAA-3'” ubicada en el extremo 3'UTR entre los nucleótidos 922 y 945 (Fig. 21) con un valor de energía de hibridación de -22,7 kcal/mol. Para Cr\_Tripsina la secuencia de unión al miRNA bantam fue “5'-ACAGCTTTTTTAAGGTCTT-3'” (Fig. 21) ubicada en el extremo 3'UTR entre los nucleótidos 243 y 260 con un valor de energía de hibridación de -20,7 kcal/mol (Tabla 5).

El análisis de predicción de target para Vitelogenina I se realizó utilizando el programa RNAhybrid.2.1.2. Para Vitelogenina I la secuencia de unión al miRNA mir-996 fue “5'-AATAAAGUCATATCA -3'” (Fig. 21) ubicada en el extremo 3'UTR entre los nucleótidos 86 y 106 con un valor de energía de hibridación de -10,9 kcal/mol (Tabla 5).

**Tabla 5.** Predicción de target para los miRNA bantam y mir-996 (Fuente: Elaboración propia).

Gen	miRNA	UTR	Posición	Unión seed	Energía de Hibridación
Cr_Catepsina	bantam	3'	802-823	818-823	-15.6
Cr_Serpina	bantam	3'	922-945	938-943	-22.7
Cr_Tripsina	bantam	3'	243-263	256-262	-20,7
Cr_Vitelogenina I	mir-996	3'	86-106	88-93	-10,9



**Figura 21.** Sitio de unión predichos para los miRNA bantam y mir-996 en los extremos 3'UTR de catepsina, serpina, tripsina y vitelogenina I de *C. rogercresseyi* (Fuente: Elaboración propia).

## **2. Evaluación de los niveles de expresión de miRNAs candidatos y sus genes target durante los diferentes estados de desarrollo de *C. rogercresseyi***

Para poder evaluar los niveles de expresión de los miRNA de *C. rogercresseyi* en los diferentes estados de desarrollo fue necesario encontrar un housekeeping para la normalización de los valores de Ct.

### **2.1 Housekeeping para miRNA de *Caligus rogercresseyi***

Para la búsqueda de un housekeeping para miRNA de *C. rogercresseyi* se utilizaron los valores de conteo y RPKM de la base de datos de miRNA de caligus (Tabla 6) previamente obtenidos en el LBGA publicados por Gallardo-Escárdate, C et al. (2017). De ambos sets de datos se seleccionaron los miRNA que tenían valores en todos los estados de desarrollo y con menores valores de varianza. El set de datos de valores de Conteo contenían un total de 673 miRNA de los cuales solo 34 presentaban valores en todos los estados de desarrollo (Tabla 6). El set de datos para PRKM contenía un total de 170 miRNA de los cuales solo 5 cumplían con el requisito de presentar valores en todos los estados de desarrollo. Los miRNA seleccionados fueron mir-2505, mir-263-a-1, mir-276 y mir-8 2, los cuales presentaban menor valor de varianza (Tabla 7). Para los análisis de expresión relativa solo se utilizarón los miRNA, mir-2505, mir-263-a-1, mir-276 y mir-8 2, para los cuales se buscó su secuencia madura de miRNA en mirBASE para el diseño de primers. Los valores de Ct fueron analizados en la plataforma R utilizandó el software NormFinder. De acuerdo a los resultados entregados por NormFinder el

mejor candidato como housekeeping es el mir-276 con un valor de estabilidad de un 0,73 (Tabla 8).

**Tabla 6.** Resumen de los datos originales para RPKM y Conteo de la base de datos de miRNA para *Caligus rogercresseyi* (Fuente: Elaboración propia).

Análisis	Set datos Conteo	Set datos RPKM
Datos originales	673 miRNA	170 miRNA
Datos eliminando los promedios cero	546 miRNA	154 miRNA
Datos dejando solo estados de desarrollo con valores > 0	34 miRNA	5 miRNA

**Tabla 7.** Análisis del set de datos para valores de Conteo y RPKM para los miRNA de *Caligus rogercresseyi* (Fuente: Elaboración propia).

miRNA	Set de datos	Naupliu		Chalimus			Media	Promedio	Varianza
		s I-II	Copepodito	I-II	Hembra	Macho			
mir-2505	RPKM	15976,5	472813,239	60738,58	53282,2	7446,68	122051	140304,7	3,898E+10
mir-263a-1	Conteo	2	2	1	1	1	1,4	0,48	0,3
mir-276	Conteo	1	1	2	1	1	1,2	0,32	0,2
mir-8-2	Conteo	1	1	1	3	1	1,4	0,64	0,8

Los valores de Ct se muestran graficados en la figura 22, se observa menor variación entre los valores de Ct y los estados de desarrollo para mir-276 (Fig. 22C),



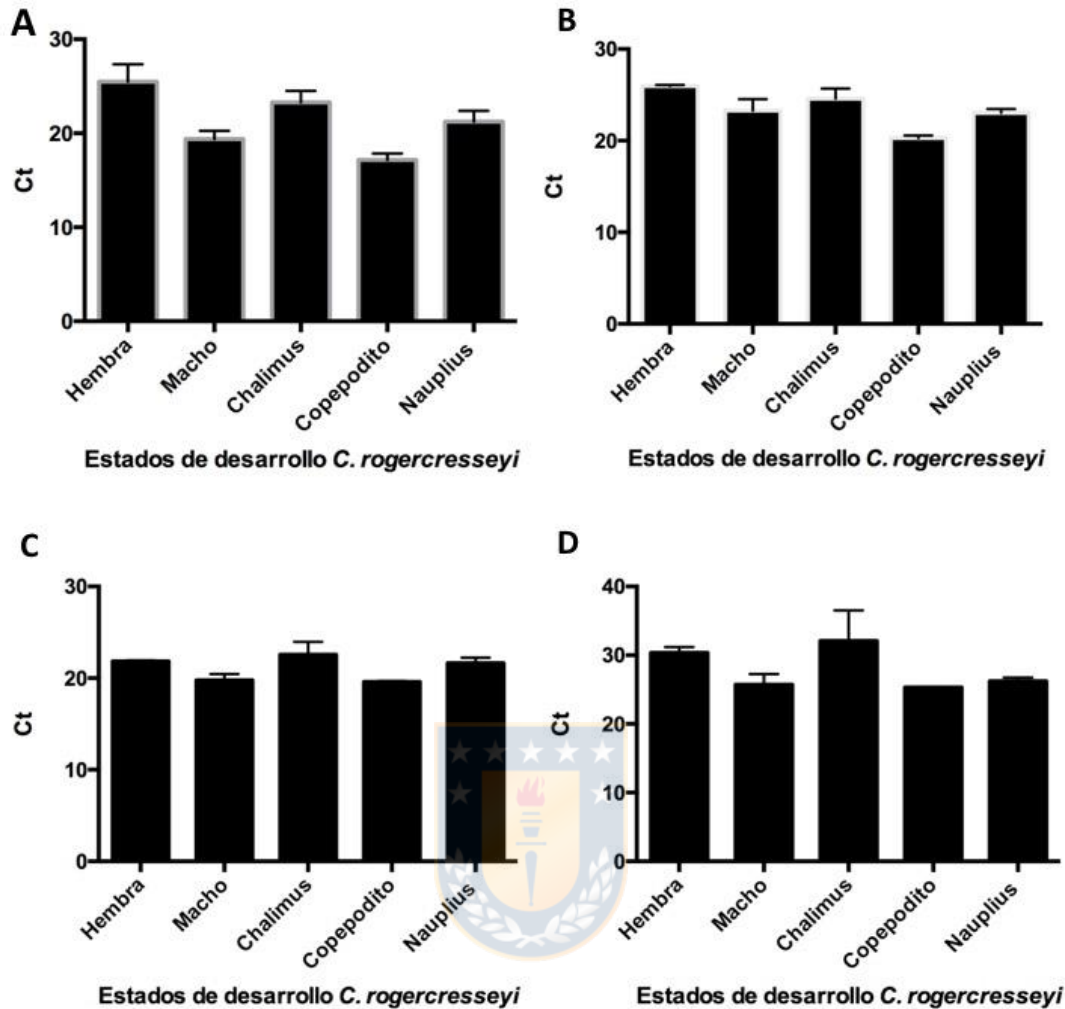
mientras que el mir-2505 (Fig. 22A) presenta mayor diferencia entre los valores de Ct para los diferentes estados de desarrollo de caligus.

**Tabla 8.** Resultados entregados por NormFinder para la identificación de genes de normalización para miRNA de *Caligus rogercresseyi* (Fuente: Elaboración propia).

miRNA	GroupDif	GroupSD	Stability
mir-276	2,33	0,16	0,73
mir-263a-1	2,54	1,01	1,34
mi-2505	3,61	1,21	1,41
mir-8-2	4,65	2,09	1,51

## 2.2 Análisis de expresión relativa por RT-PCR para Bantam y mir-996 en los diferentes estados de desarrollo de *Caligus rogercresseyi*.

En la figura 23 se muestran graficados los valores de Ct de los miRNA de estudio, los cuales son bantam un miRNA relacionado al proceso infeccioso y mir-996 un miRNA relacionado con reproducción (Gallardo-Escárdate, C et al., 2017). Los resultados del análisis de expresión de los miRNA de estudio ya normalizados, utilizando como housekeeping a mir-276 se muestran en la figura 24. Para bantam hay un aumento en los niveles de expresión que se correlaciona en los estados de desarrollo más tempranos de nauplius, copepodito y chalimus (Fig. 24A). En el caso de mir-996 se observa un mayor aumento en los niveles de expresión en los estados más tempranos del ciclo de vida como nauplius y copepodito, como también en el estado de hembra (Fig. 24B).



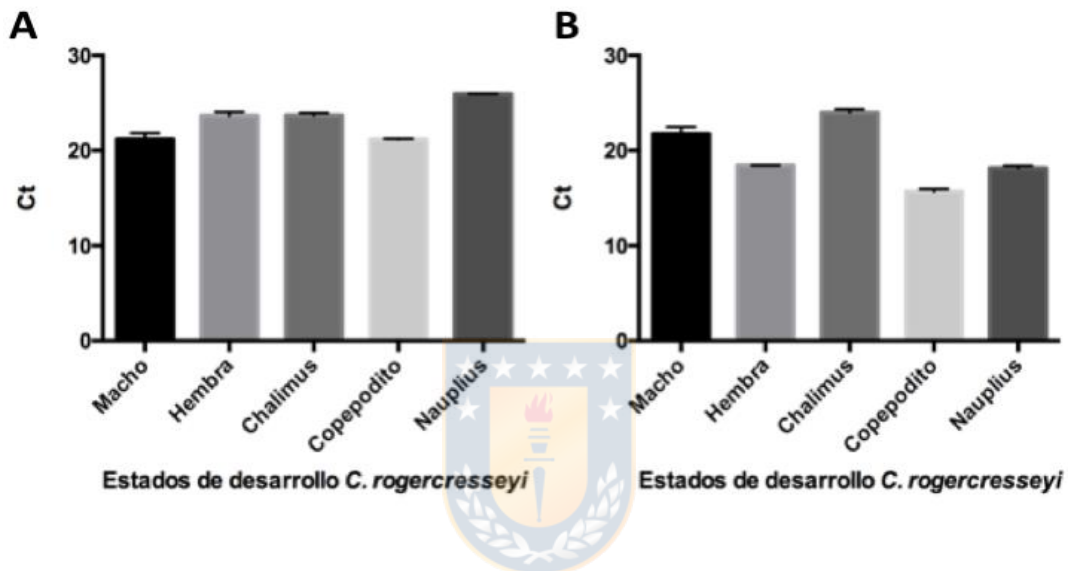
**Figura 22.** Valores de Ct para los candidatos a housekeeping para miRNA de *C. rogercresseyi*. **A.** mir-2505, **B.** mir-263a 1, **C.** mir-276, **D.** mir-8-2. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ) (Fuente: Elaboración propia).

### 2.3 Análisis de expresión relativa por RT-PCR para catepsina, serpina, tripsina y vitelogenina I en los diferentes estados de desarrollo de *C. rogercresseyi*

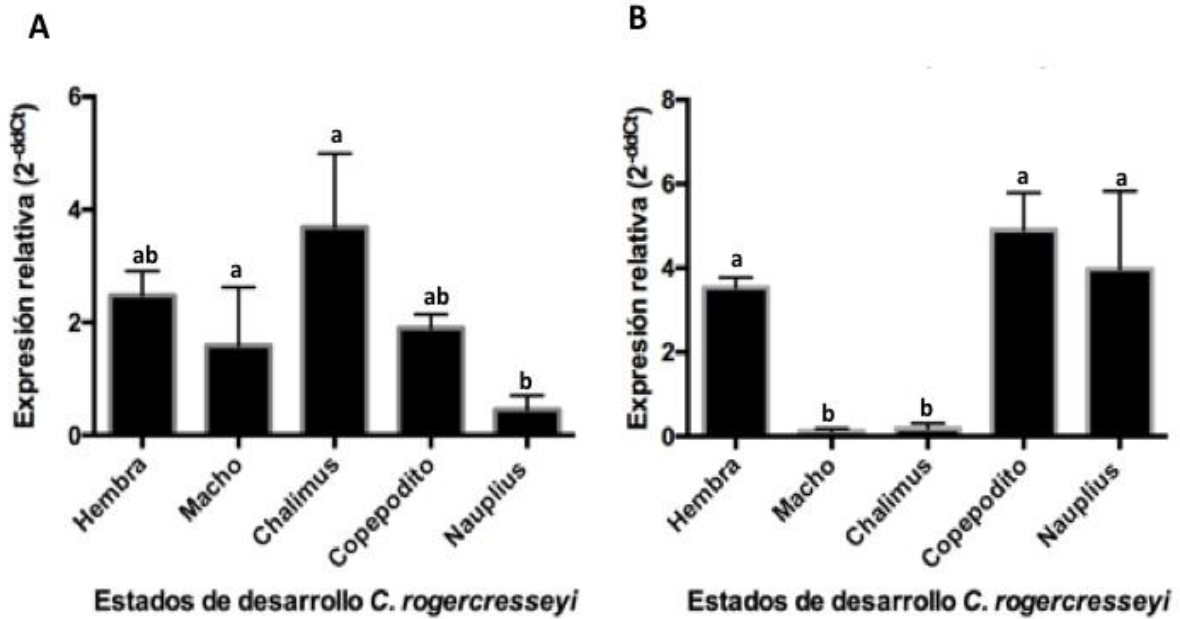
De acuerdo a los resultados de expresión relativa en Cr\_Catepsina se puede observar en hembras valores de expresión significativamente mayores a los otros estados de desarrollo evaluados (Fig. 25A). En el caso de Cr\_Serpina se encuentra altamente expresado en estado de nauplius (Fig. 25B). Para Cr\_tripsina se observa

una mayor expresión en el estado adulto de hembra seguido por chalimus (Fig. 25C).

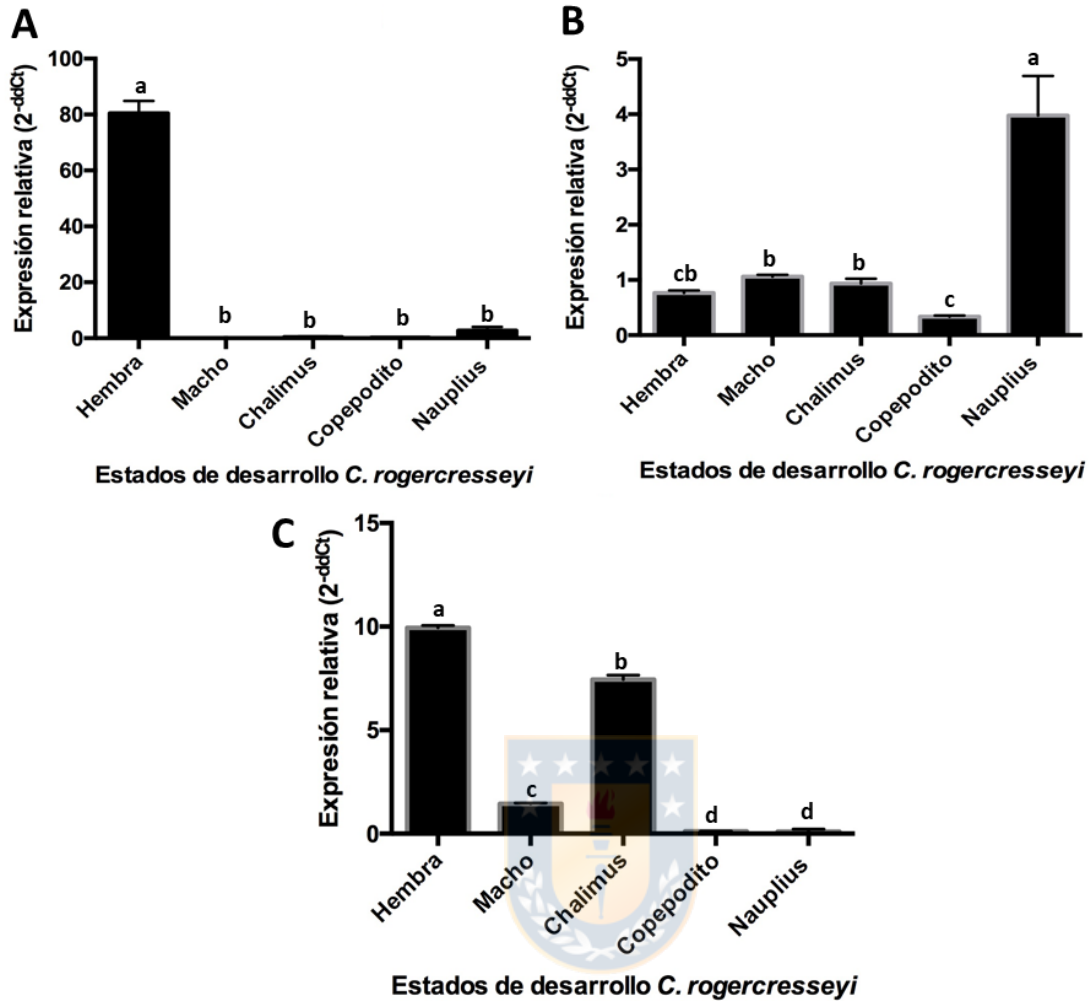
De acuerdo a los análisis de expresión relativa para vitelogenina I de *C. rogercresseyi*, se observó mayor expresión significativa del gen en el estado de hembra (Fig. 26).



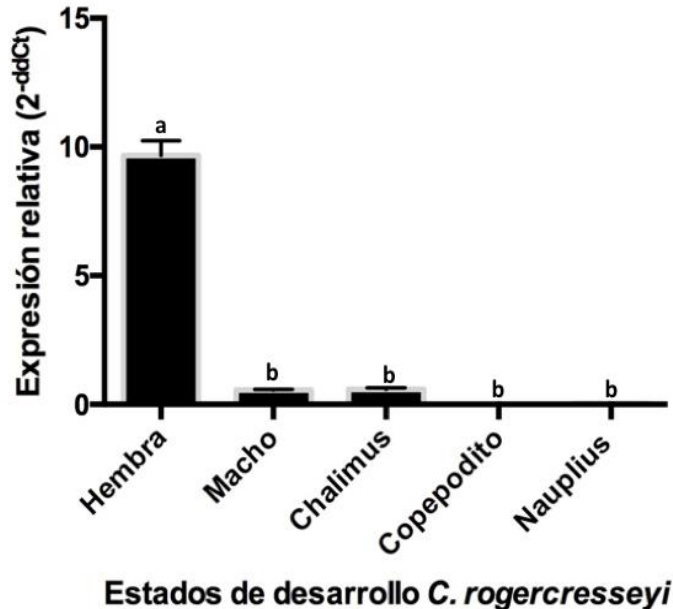
**Figura 23.** Valores de Ct para miRNA de *C. rogercresseyi* relacionados a infección y reproducción. **A.** Bantam, **B.** mir-996. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ) (Fuente: Elaboración propia).



**Figura 24.** Expresión relativa de miRNA Bantam y mir-996 en los diferentes estados de desarrollo de *C. rogercresseyi*. **A.** miRNA Bantam, **B.** mir-996. Los niveles de expresión fueron normalizados por el control endógeno mir-276. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ). Las diferencias significativas fueron consideradas con un valor  $p < 0.05$  (Fuente: Elaboración propia).



**Figura 25.** Expresión relativa de proteínas del secretoma en los diferentes estados de desarrollo de *C. rogercresseyi*. **A.** Cr\_Catapsina, **B.** Cr\_Serpina, **C.** Cr\_Tripsina. Los niveles de expresión fueron normalizados por el control endógeno  $\beta$ -tubulina. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ). Las diferencias significativas fueron consideradas con un valor  $p < 0.001$  (Fuente: Elaboración propia).



**Figura 26.** Expresión relativa para Cr\_Vitelogenina I en los diferentes estados de desarrollo de *C. rogercresseyi*. Los niveles de expresión fueron normalizados por el control endógeno  $\beta$ -tubulina. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ). Las diferencias significativas fueron consideradas con un valor  $p < 0.001$  (Fuente: Elaboración propia).

### 3. Validación a nivel funcional la interacción de mRNA/miRNAs candidatos relacionados al proceso infectivo y reproductivo en *C. Rogercresseyi*.

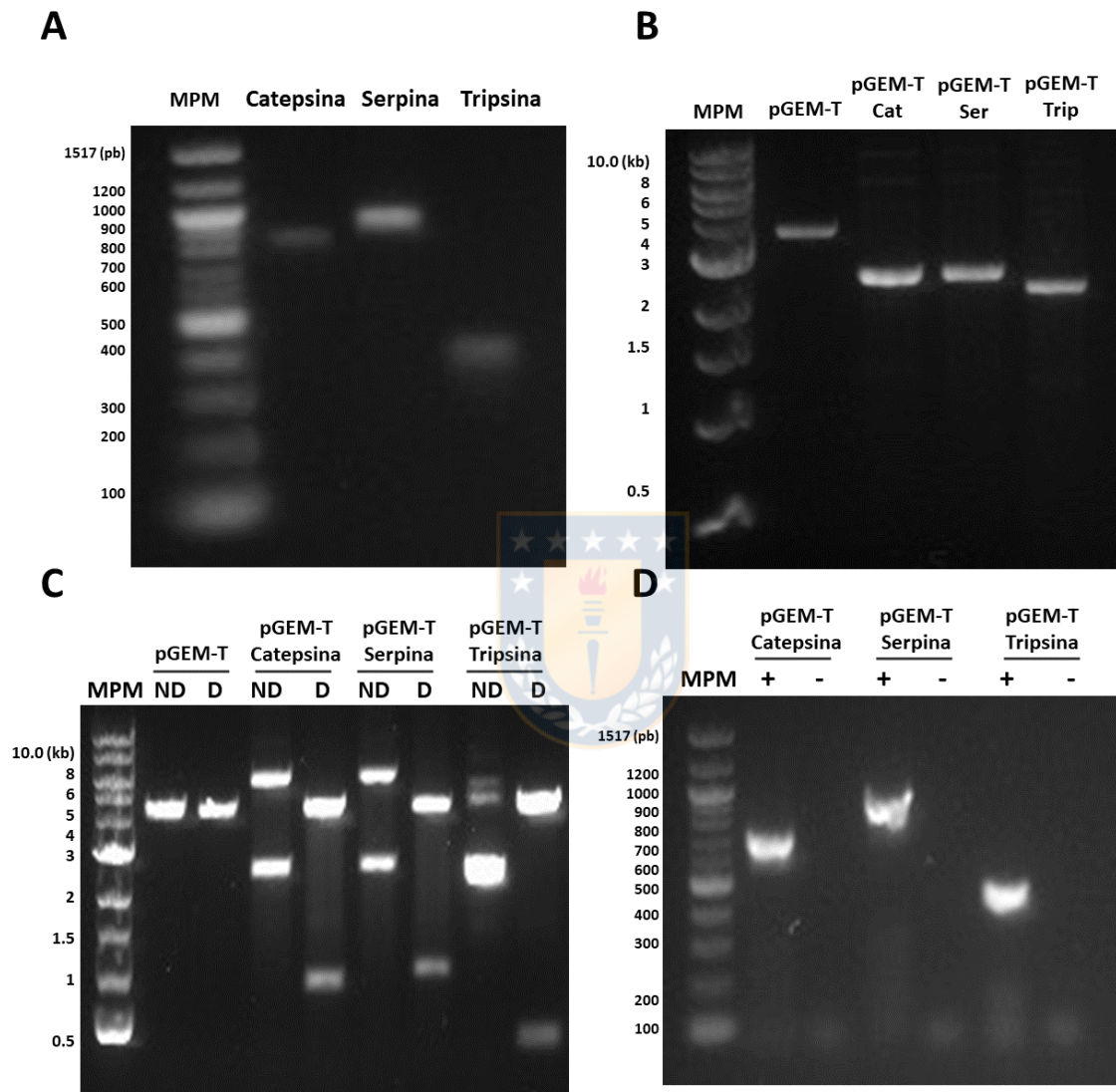
#### 3.1 Clonamiento regiones 3'UTR y secuenciación de genes

Se clonó el ADNc de los extremos 3'UTR en el vector pGEM-T easy system II que contienen el sitio de unión al microRNA como se indicó en la metodología. A continuación, se seleccionaron clones recombinantes y se purificarón los plásmidos. A partir del análisis de restricción con PmeI y XbaI y confirmación por PCR, 5 clones fueron enviados a secuenciación. De ello, se obtuvo que el plásmido 4, plásmido 2,

y plásmido 4, para serpina, tripsina y catepsina respectivamente (Fig. 27), contenía la secuencia del extremo 3'UTR del ARNm que presenta el sitio de unión al microRNA bantam con más de un 90% de identidad respecto de la secuencia nucleotídica obtenida mediante secuenciación masiva, correspondiente a los contig Contig0008452 para serpina, Contig0013229 para tripsina y Contig0017956 para catepsina. Para el caso de Vitelogenina I el clon 1 contenía la secuencia de unión al mir996 con un 100% de identidad con la secuencia obtenida mediante clonamiento, RACE y secuenciación.

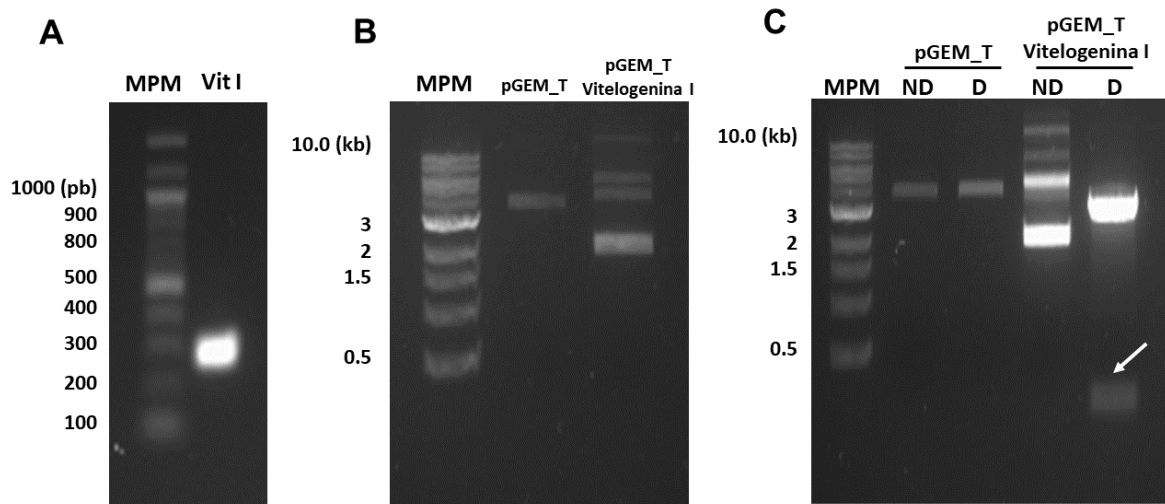
Para Cr\_Catepsina se obtuvo una secuencia de 645 nucleótidos (Fig. 27C) que contiene en sus extremos los sitios de corte para las enzimas de restricción PmeI y XbaI (GTTTAAAC y TCTAGA respectivamente), además de presentar el sitio de unión a Bantam "AGGTTATATTGATCGATGA" (Fig. 28A), el cual presentó un fragmento con un 98,420% de identidad con la secuencia del 3'UTR original. Para Cr\_Serpina se obtuvo una secuencia de 698 nucleótidos que contiene en sus extremos los sitios de corte para las enzimas de restricción (Fig. 27C). Además, de presentar el sitio de unión a Bantam "CTCAGCCGACGGGGTGTTGATCTAA" (Fig 28B). El fragmento obtenido presentó un 99,416% de identidad con el Contig0008452. Para Tripsina se obtuvo una secuencia de 291 nucleótidos que presentaba los sitios de corte para PmeI y XbaI (Fig. 27C), y un sitio de unión a Bantam "ACAGCTTTTTTAAGGTCTT" (Fig. 29C), con un 94,982% de identidad con la secuencia del 3'UTR original. Para Vitelogenina I se obtuvo una secuencia de 221 nucleótidos (Fig. 28) que contiene en sus extremos los sitios de corte para las enzimas de restricción PmeI y XbaI. También presentaba el sitio de unión a mir-996 "AAGTCATATCAAATGAACTA" (Fig. 29D). El fragmento presentó un 99% de

identidad con la secuencia del 3'UTR original. Los análisis de predicción de target para bantam y mir-996 se realizaron nuevamente para las secuencias clonadas de modo de corroborar la predicción de target.



**Figura 27.** Resultados del clonamiento de los insertos de mRNA 3'UTR de catepsina, serpina y tripsina. **A.** Producto PCR clonado en el vector pGEM-T easy system II, **B.** Plásmido pGEM\_UTR purificado con el inserto de ARNm 3'UTR de Catepsina, Serpina, Tripsina, **C.** Ensayo con enzimas de restricción para los plásmidos pGEM-T\_3'UTR\_catepsina, pGEM-T\_3'UTR\_Serpina y pGEM-T\_3'UTR\_Tripsina, **D.** PCR en tiempo final para el plásmido pGEM\_UTR para verificar la presencia de los insertos de Catepsina, Serpina y Tripsina. ND: No Digerido, D: Digerido, MPM: Marcador de peso molecular (Fuente: Elaboración propia).





**Figura 28.** Resultados del clonamiento de los insertos de mRNA 3'UTR de Cr\_Vitelogenina I. **A.** Producto PCR clonado en el vector pGEM-T esasy system II, **B.** Plásmido pGEM\_UTR purificado con el inserto de ARNm 3'UTR de Vitelogenina I. **C.** Ensayo con enzimas de restricción para los plásmidos pGEM\_3'UTR\_Vit\_I. ND: No Digerido, D: Digerido, MPM: Marcador de peso molecular (Fuente: Elaboración propia).

## A. Cr\_Catepsina

TTTAAACCTTGACATTTTCAATTAGGCTCCTTACCTCACGTTAGAGTTTGTGTAGTTAAGCACAAAGCCCAAGGAATTACATATAAAGATTTTGTGGCC  
PmeI  
CATGAATAAAAAGGTCTATATTAATTAATTTCTGAACAAAATAAGTCAATCTTTGTTATGTTAGGGAGTCCTAACGTAAGTATATAGCTATATTTGTAT  
CTCAGGCCAAAAAACAGTTTAATTAATTAGATAACTATATAAAGAACAGTTAGAAATACGTAGCCAGAGATTACGCCGATATTAATTTGAAACATTGTTTAT  
CAAGACTGTTTTCTTGATGGTCCAAATACGCTCCCTAGCCAACAACCTTGGCTCATGAGGGTTCAAAACGAAGCATCCCTGGGCCCTTGAAATATCGCAAT  
TAAGTCATATTTGAGGTATCAAAATCCACCTCAATGAACAATTGTCTGGGTTATAGTGTCTTTTGGTCCACAACATATGGCAATTTGGGTATCAAAAG  
TAAACCGCTACAAATGTCACATAATCCCTAAATTAATGTCCTTTCTGTAAGTAAACACATCGTGTGTTCCAATGTAATAATATTATATATGTTCTCTC  
CAGAGTTGCTCCATAAAGGTTATATTGATCGACTGATCCCTCTAGA  
Union a Bantam XbaI

## B. Cr\_Serpina

TTTAAACAAATGCTTGGCCGTAAGCGGAAATAATCAACGGCAATTACGCGCTTGGGAAAGAAAGGATCTCCACCATAGGTAATTGAGGAGGGCTTCCCA  
PmeI  
AGGTAAACAAAGTTCTGAATGGCGGCGGACATGTCGCGAGCTAATGTAGACGAGTGTATCCACATCTGATGCTCTAATGGCTTCACTGCAGTTGGATGAA  
GACCGTTCCGAGGAGGGTCCACAAATGGCAACGACTCGTGAACCTACTTGTCTTGTCTTTTTCATATGCAACATCTTTAAAAATACGCTCTAGAAAGATG  
CTCTGCCCTTTCGGCACGAAAGTCCACATTTGGATATATTTGTTCTCTTTGCGTTTGTCTTGGCGCTCGAATGGCCTCTGAAATAATGTCCACGCCGTAT  
ACTTTTTTACACTTACAAGAGAGAAAATAGTCCGATGGTCCAGTGCAGCGCAACGTCATTAAGCTCGTGGTTTGGGGTTAAGTCAGCCAGCTCTC  
CGGCTTTTCGAGTAGAGTAGCTCAGCCGCGGGGTTGTTGATCTGAAAAAGGATTCAGGGGAGACCTTGAAGCGAAGGCTCCCAAGAGACTCATAGAAGAA  
Sito de unión a Bantam  
GCGTTCTCCTAGGAGTAAATGTGAGGAGGACTTCCCTGATGAGCGACTCTTGAAGTGGACGAGGAGGGACAAATTCGGATGCTTTTTCTCTCTAGA  
XbaI

## C. Cr\_Tripsina

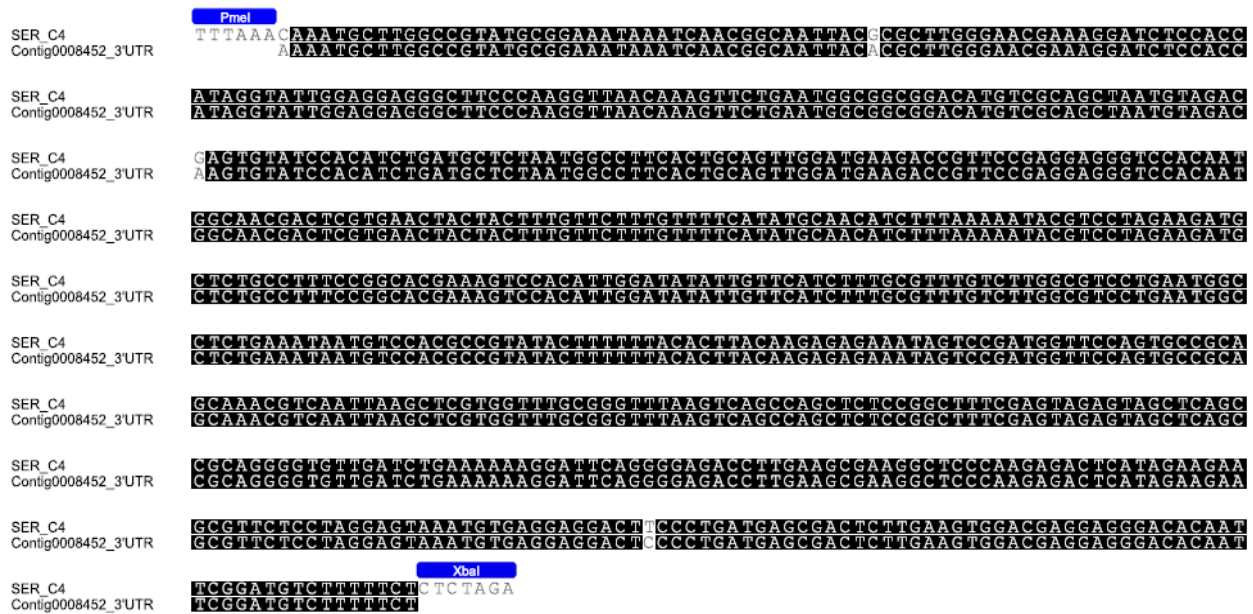
TTTAAACCCAGTAAAAAAGATCGGAAGAGCGTCGTAGAAAACATGAGGCATCTATGGAGGACTCCAGAAGTTCTCAACAATATCAAGGACAATA  
PmeI  
TCTTCGGATTCCTTCTTTTTAAGAGACCGAGGCGGATTTTTTCGCGTCTTCTTCTCAGACAGCTTTTTTAAGTCTTTTTACGATGGGCGCCGGATT  
Sito de unión a Bantam  
TTTTGCTCTGGAGGGCCACCTCTTTAGCGATTTGAAACATTTCACTCCTCGTTCGTCTTTTGAACCTCCTGTCAAGTGACGCTCTCTAGAAAT  
XbaI

## D. Cr\_Vitelogenina I

... TTTAAACGAAATCACCCGGAACCATACTGAAAACATTTATTTTTGAATGAAAGTACCTCTGCGGAATTTATATATTTAATAATAATAAAGTC  
PmeI  
... ATATCAAAATGAACATAAAAAAAAAAACCCTAGTGTAGTCTGATTAATTCGGATCCGCGAATCGAATTCGGCGGGCCCATGGCGGGCGGAG  
Sito de unión a mir-95  
... CATGCGACGTCGGGCCCAATTCGCTCTAGAA  
XbaI

**Figura 29.** Inserto del mRNA clonado en el vector pGEM-T easy. **A.** Cr\_Catepsina, **B.** Cr\_Serpina, **C.** Cr\_Tripsina, **D.** Cr\_Vitelogenina I. En cuadros azules se representan los sitios de corte para las enzimas de restricción PmeI y XbaI, en rojo se representan los nucleótidos que participan en la interacción con el microRNA bantam en el caso de Catepsina, Serpina y Tripsina. Para vitelogenina

I se representa el sitio de unión a mir-996 (Fuente: Elaboración propia).



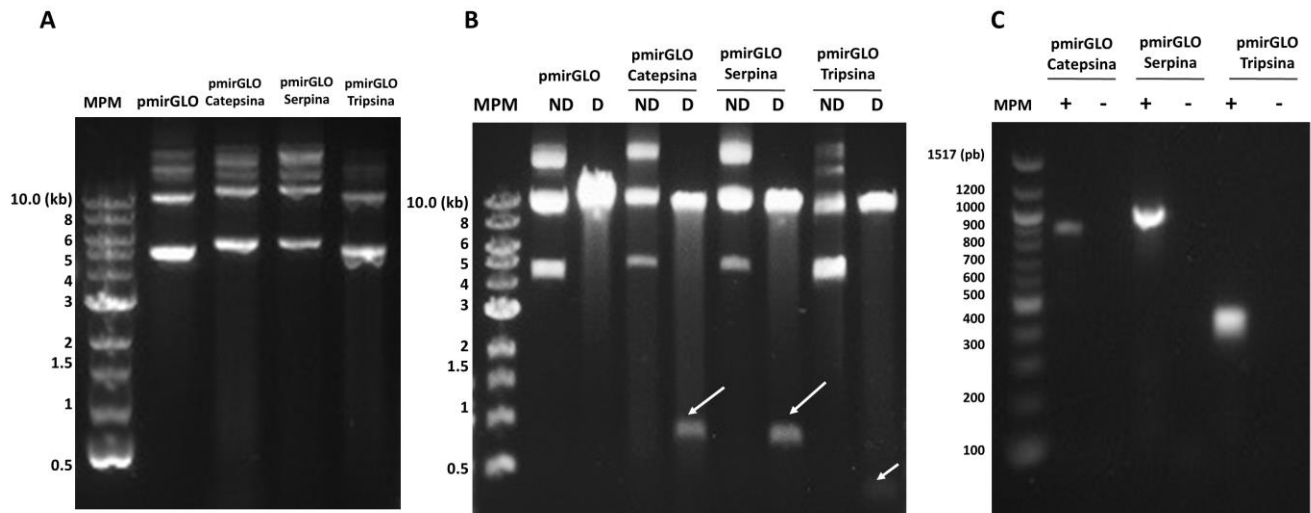
**Figura 30.** Alineamiento de la secuencia del extremo 3'UTR de Cr\_Serpina por secuenciación masiva y el fragmento clonado en el vector pGEM-T. Ser pGEM-T C4 corresponde a la secuencia 3'UTR de Serpina clonando en el vector pGEM-T y Ser Contig 0005452 corresponde a la secuencia obtenida por secuenciación masiva. En cuadros azules se indican los sitios de corte para PmeI y XbaI, en rojo se indica el sitio de unión a Bantam (Fuente: Elaboración propia).

### 3.2. Subclonamiento de las regiones 3'UTR en el vector pmirGLO.

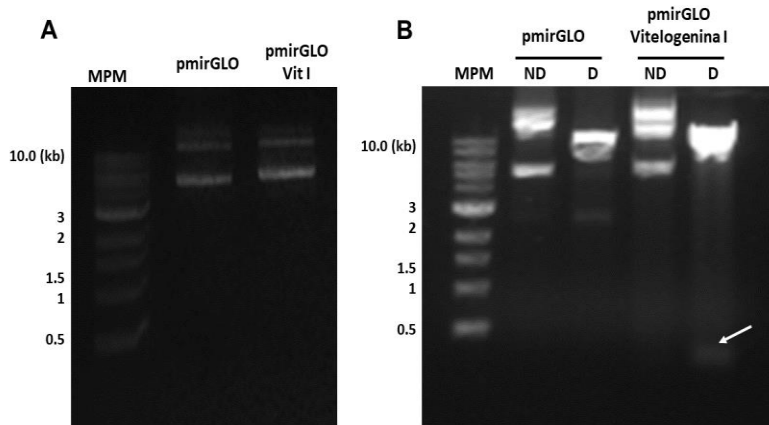
Los fragmentos clonados en el vector pGEM-T que contenían el sitio de unión al microRNA además, de los sitios de corte para las enzimas de restricción PmeI (extremos romos) y XbaI (extremos cohesivos) se subclonaron en el vector pmirGLO.

Los fragmentos para Cr\_Catepsina (645 nucleótidos), Cr\_Serpina (698 nucleótidos), Cr\_Tripsina (291 nucleótidos) y Cr\_Vitelogenina I (221 nucleótidos) fueron subclonados en el vector pmirGLO (Fig.31, Fig. 32). Los nuevos plásmidos fueron enviados a secuenciar a MACROGEN. Un ejemplo de alineamiento entre la

secuencia clonada en el vector pGEM-T y subclonada en el vector pmirGLO se muestran en la figura 33. Los resultados del diseño del vector pmirGLO para cada secuencia se representa en la figura 34.



**Figura 31.** Resultados de clonamiento de los insertos del mRNA 3'UTR de catepsina, serpina y tripsina en el vector pmirGLO. **A.** Plásmido pmirGLO\_3'UTR purificado con el inserto de ARNm 3'UTR de Catepsina, Serpina, Tripsina. **B.** Ensayo con enzimas de restricción para los plásmidos pmirGLO\_3'UTR\_catepsina, pmirGLO\_3'UTR\_Serpina, pmirGLO\_3'UTR\_Tripsina. **C.** PCR en tiempo final para el plásmido pGEM\_UTR para verificar la presencia de los insertos de Catepsina, Serpina y Tripsina (Fuente: Elaboración propia).



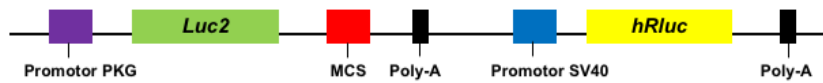
**Figura 32.** Resultados del clonamiento del inserto del mRNA 3'UTR de Cr\_Vitelogenina I en el vector pmirGLO. **A.** Plásmido pmirGLO\_3'UTR purificado con el inserto de ARNm 3'UTR de Vitelogenina I.

**B. Ensayo con enzimas de restricción para los plásmidos pmirGLO\_3'UTR\_Vit\_I (Fuente: Elaboración propia).**

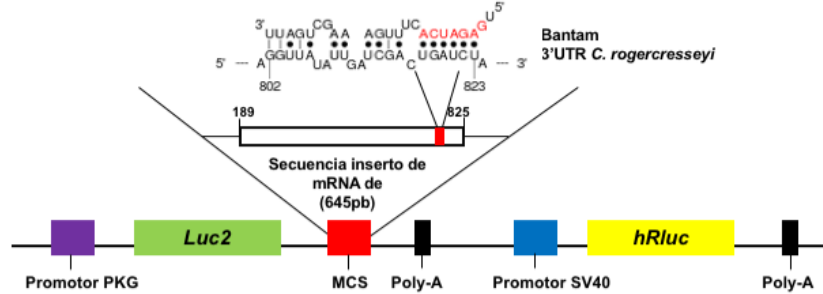


**Figura 33.** Alineamiento de la secuencia del extremo 3'UTR de Cr\_Serpina clonado en el vector pGEM-T y el fragmento subclonado en el vector pmirGLO. Ser pmirGLO corresponde a la secuencia 3'UTR de Serpina subclonado en el vector pmirGLO y SER\_C4 pGEM-T corresponde a la secuencia clonada en el vector pGEM-T. En cuadros azules se indican los sitios de corte para PmeI y XbaI, en rojo se indica el sitio de unión a Bantam (Fuente: Elaboración propia).

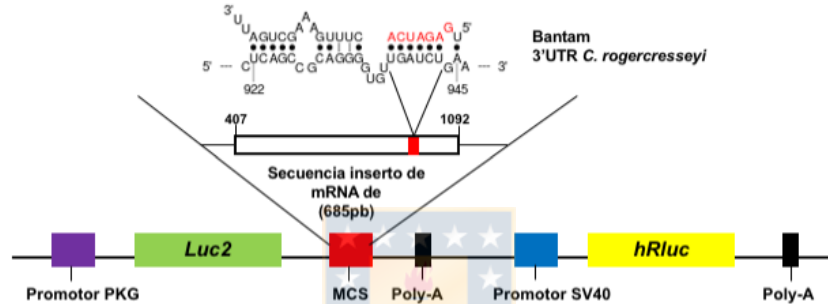
A) Vector pmigLO



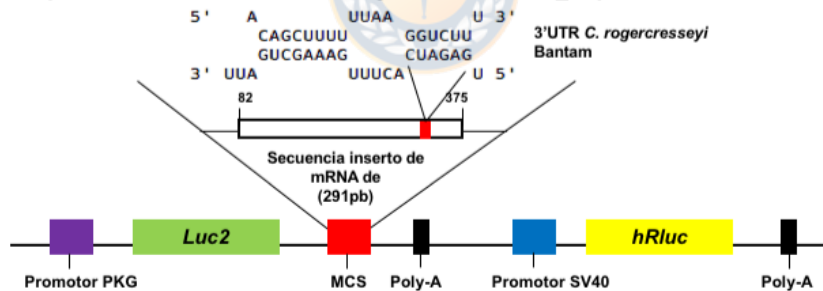
B) Vector pmigLO con el inserto de la región 3'UTR de Cr\_Catepsina



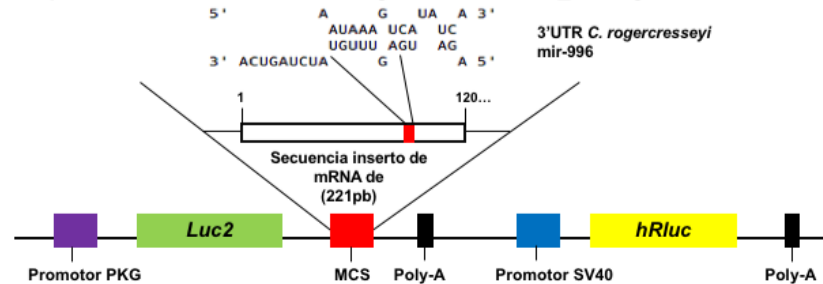
C) Vector pmigLO con el inserto de la región 3'UTR de Cr\_Serpina



D) Vector pmigLO con el inserto de la región 3'UTR de Cr\_Tripsina



E) Vector pmigLO con el inserto de la región 3'UTR de Cr\_Vitelogenina I



**Figura 34.** Regulación de la expresión génica por miRNA mediante la unión directa al 3'UTR de la transcripción diana. Diagramas del reportero de luciferasa, el vector pmirGLO construye la

codificación de **A.** Sin inserto, es decir, control, **B.** 3'UTR de Cr\_Catepsina, **C.** 3'UTR de Cr\_Serpina, **D.** 3'UTR de Cr\_Tripsina y **E.** 3'UTR de Cr\_Vitelogenina I (Fuente: Elaboración propia).

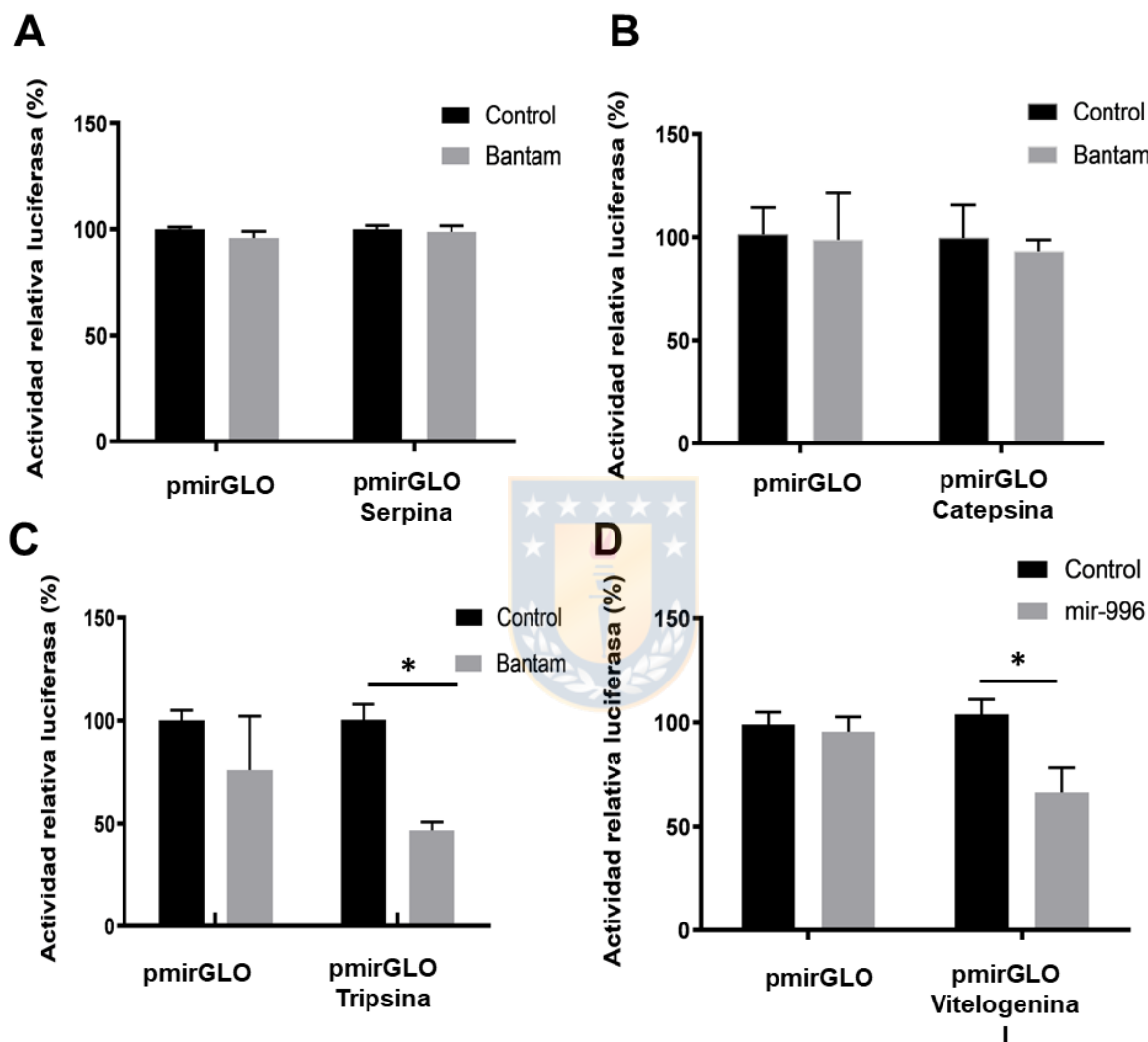
### 3.3 Ensayo luciferasa

Los miRNA son importantes reguladores de la expresión génica y se ha demostrado que juegan un papel importante en numerosos procesos biológicos. Uno de los métodos para comprender mejor las interacciones entre un miRNA y la UTR de su gen target es el ensayo de la luciferasa. El 3'UTR de un gen de interés se fusiona con el fin de un gen reportero de luciferasa. Si el 3'UTR es un objetivo de un miRNA, la estabilidad de la transcripción del reportero y / o la eficiencia de traducción va a cambiar con la variación en la concentración de miRNA funcional.

Según los análisis bioinformáticos utilizando plataformas de predicción de target para miRNA, bantam se uniría al extremo 3'UTR de Cr\_Catepsina, Cr\_Serpina y Cr\_Tripsina, mientras que el miRNA mir-996 se uniría al extremo 3'UTR de Cr\_Vitelogenina I. Basada en esta información se realizaron los ensayos de gen reportero para corroborar las predicciones entregadas por herramientas bioinformáticas.

Para Cr\_Serpina se estimó una energía de hibridación de -22,7 kcal/mol para la unión con bantam, sin embargo, la actividad de la luciferasa no se vio afectada por el miRNA bantam en las células co-tranfectadas con el vector pmirGLO como control ni con el vector pmirGLO\_Serpina el cual contenía corriente abajo del gen de la luciferasa la secuencia del extremo 3'UTR que contenía el sitio de unión al miRNA (Fig. 35A). Lo mismo ocurrió con Cr\_Catepsina, a pesar que se estimó una

energía de hibridación de -15,6 kcal/mol para la unión con bantam la actividad de la luciferasa no se vio afectada por la presencia del miRNA (Fig. 35B).



**Figura 35.** Niveles de luminiscencia, ensayo de luciferasa en la línea celular HeLa. **A.** Catepsina, **B.** Serpina, **C.** Tripsina, **D.** Vitelogenina I. Las barras muestran la desviación estándar  $\pm$  SD ( $n = 3$ ). Las diferencias significativas fueron consideradas con un valor  $p < 0.001$ , representadas con un (\*) (Fuente: Elaboración propia).



Al evaluar los niveles de luminiscencia en las células transfectadas con Cr\_Tripsina y Bantam se observó una reducción de un 50% (Fig. 35C).

Para Cr\_Vitelogenina I se estimó mediante plataformas bioinformáticas una energía de hibridación de -10,9 kcal/mol para la unión con el microRNA mir-996. El estudio *in vitro* de la interacción mRNA-miRNA se observó una disminución de la actividad de la luciferasa en más de un 30% en comparación con el control siendo esta diferencia estadísticamente significativa



## DISCUSIÓN

### CATEPSINA

Las proteasas juegan un papel importante en la penetración de tejidos, digestión, muda y evasión de la respuesta inmune del huésped. Unas de las proteasas mayormente encontradas en el secretoma de parásitos han sido las catepsinas. (Mach, Mort, & Glössl, 1994)

Las catepsinas pertenecen a la familia de las proteasas de cisteína, las cuales presentan una secuencia señal que permite la orientación del pro-péptido hacia diferentes lugares de almacenamiento. Los resultados obtenidos del tBLASTx y SMART son congruentes con esto, ya que se encontró la presencia de la secuencia inhibidor I29 en relación a un pro-péptido en catepsinas inactivas, el cual puede ser activado a través de la escisión proteolítica del pro-péptido N-terminal (Carmona et al., 1996). Estos resultados concuerdan con catepsina L descrita para *L. salmonis*, la cual presenta una secuencia de pro-péptido en el extremo N-terminal (McCarthy et al., 2012). Los miembros del grupo L de catepsinas han sido descritas como las proteasas más comunes entre especies de artrópodos (Sojka et al., 2011). En el piojo de mar *C. rogercresseyi*, no es la excepción, siendo las catepsinas L los transcritos más abundantes en todas las etapas de la vida analizadas de este copépodo (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015). Los resultados obtenidos, muestran los dominios conservados característicos de las catepsinas de tipo L, incluyendo las secuencias

de péptido señal MKALSVLACVVA, GCNGG, y CGSCWSFS. La secuencia GCNGG tenía un residuo de cisteína implicada en la formación de puentes disulfuro, indicando por lo tanto una función estructural. Mientras tanto, la secuencia de CGSCWSFS posee el sitio activo de cisteína y la triada catalítica CHN al igual que en catepsina L del piojo de mar *L. salmonis* en las secuencias de LsCL1 y LsCL2 (McCarthy et al., 2012). La función de proteasas de cisteína parece ser esencial en una variedad de procesos biológicos importantes de parásitos, tales como la muda, remodelación de cutícula, embriogénesis, alimentación y evasión inmune, haciéndolas un objetivo atractivo para el desarrollo de vacunas (Ford et al., 2009).

El análisis de expresión entre las diferentes etapas de desarrollo mostró una alta expresión de Cr\_Catepsina L en hembras adultas; lo que concuerda con estudios realizados en otras especies de parásitos, centrados en la correlación de proteasas de cisteína y el desarrollo embrionario (Dalton et al., 2003; Hashmi et al., 2004). En *Caligus rogercresseyi* se ha observado altos niveles de expresión Cr-Catk1 en estado hembra, asociando su función al desarrollo embrionario, por otra parte, *Cr-CatL1*, *Cr-CatL26* y *Cr-CatL27*, junto con *Cr-CatB1*, mostraron un aumento de expresión en los estadios de nauplius antes de ser regulado por disminución progresiva a lo largo del desarrollo, lo que sugiere su posible participación en los procesos de muda o remodelación de cutícula en *C. rogercresseyi* (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárate, 2015). Así, en *Caenorhabditis elegans*, la catepsina *Ce-CPL-1* demostró desarrollar un papel esencial en el procesamiento de la proteína de la yema durante el desarrollo embrionario; la pérdida de actividad, conduce a vesículas de yema en el citoplasma agrandados y la consecuente mortalidad embrionaria (Hashmi, Zhang, Oksov, &

Lustigman, 2004). En conjunto con esto último, se observa una dramática reducción de la eclosión en huevos, cuando el parásito *Fasciola hepática* es alimentada con huéspedes inmunizados contra catepsina L (Dalton et al., 2003). Del mismo modo, el parásito *Brugia malayi* presentó una cisteína proteasa, *Bm\_clp*, en la yema de huevo durante su desarrollo embrionario (Ford et al., 2009).

El descubrimiento de la presencia de miRNA en el piojo del salmón genera un nuevo desafío para comprender el método de regulación post-transcripcional en caligus, los cuales podrían estar participando en la regulación de procesos cruciales en su ciclo de vida como lo es el proceso infectivo y reproductivo del piojo de mar (Gallardo-Escárate, C et al., 2017). Mediante los enfoques de la bioinformática, diversos programas como PITA, RNAHíbrido, miRanda, Pictar, Targetscan, Tarbase entre otros, pueden predecir miles de genes que son potencialmente regulados por miRNAs (Shirdel et al., 2011) basados en la complementariedad de secuencias, si bien existe un gran número de falso positivos actualmente se está investigando para reducirlos mediante el refinamiento de dichas herramientas, y en base en estas, también se han generado base de datos robustas que contienen esta información y están a disposición de los investigadores para ser consultadas como miRbase, MirWalk, microRNA.org, miRDB, mir2disease.org y miRTarBase.

Los microARNs de animales parecen reconocer los ARN mensajeros objetivo mediante un emparejamiento de bases impreciso, y en general probablemente provocan la represión de traducción de sus objetivos, como se ha demostrado para lin-4 de *C. elegans* (Lee, Feinbaum, & Ambros, 1993). El gen diana del miRNA bantam en *Drosophila* es un gen pro-apoptótico HID (Head Involution Defective) el cual al ser regulado post-transcripcionalmente por bantam provoca la proliferación

celular y previene la apoptosis (Brennecke et al., 2003; Hipfner et al., 2002). Según los resultados de predicción de target de caligus para Cr\_Catepsina la secuencia de unión al microRNA Bantam fue “5'-AGGTTATATTGATCGATGA-3'” ubicada en el extremo 3'UTR con un valor de energía de hibridación de -15,6 kcal/mol, sin embargo, los resultados del ensayo de luciferasa no demostraron la regulación post-transcripcional de bantam para la Cr\_Catepsina de caligus que se estudió en la presente tesis. En el caligus catepsina L se expresa abundantemente en todos los estados del ciclo de vida de caligus, mientras que catepsina B y K se expresan en los estados larvales y hembras adultas, sugiriendo su participación en el proceso de muda y desarrollo embrionario (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárte, 2015). En *C. rogercresseyi* se han identificado usando la tecnología de secuenciación de transcriptoma de alto rendimiento, 56 secuencias para catepsina-like distribuidas en cinco grupos de cisteína proteasas (Maldonado-Aguayo, Waleska, Chávez-Mardones, Gonçalves, & Gallardo-Escárte, 2015). Estos antecedentes sugieren que otras isoformas de las 56 Catepsinas identificadas para caligus podrían estar siendo reguladas por bantam. Bantam desempeña un papel clave en la regulación del crecimiento a nivel sistémico en *Drosophila melanogaster* regulando la proliferación celular y la muerte celular (Brennecke, Hipfner, Stark, Russell, & Cohen, 2003). Sin embargo, en *C. rogercresseyi* sugieren una función para bantam relacionada con la evasión inmune en caligus durante la infestación en salmón atlántico (Gallardo-Escárte, C et al., 2017). La inhibición de la expresión de catepsina B1 mediante dsRNA en ratas infectados con *Schistosoma japonicum*, provocaron la supresión del gen SjCB1 en gusanos adultos, observando la supresión a largo plazo de tres genes funcionales

publicados (peroxiredoxin-1, mago nashi, receptor de insulina) en *S. Japonicum* (Li et al., 2018). Por otra parte, se ha demostrado que catepsina B participa en los procesos de crecimiento tumoral, vascularización e invasión de gliomas. En estudios utilizando ARNi para catepsina B se inhibió significativamente la expresión catepsina B y redujo el comportamiento invasivo de SNB19, la línea celular de glioblastoma. Reduciendo la interacción célula-célula de las células endoteliales microvasculares humanas, lo que resultó en la interrupción de la formación de la red capilar (Lakka et al., 2004). El silenciamiento exitoso de la cisteína proteasas catepsina B y L en la etapa infectiva de *Fasciola hepatica* mostro disminuciones en los niveles de transcripción diana y una disminución significativa en las proteínas codificadas en el estómago. El RNAi indujo fenotipos transitorios, anormales, y redujo significativamente la penetración de la pared intestinal de la rata (McGonigle et al., 2008). El estudio del efecto de la inhibición de catepsina también ha sido evaluada como alternativa de control en el parásito *Schistosoma mansoni* (Skelly et al., 2003).

La luciferasa de luciérnaga se usa comúnmente para evaluar la actividad transcripcional en células. La aplicación más común del ensayo del gen reportero de luciferasa es examinar la regulación de las actividades transcripcionales por los promotores y los factores de transcripción (Campos-Melo et al., 2014; Jin et al., 2013a). Recientemente, este ensayo también se ha adaptado para probar el efecto de la regulación postranscripcional mediada por miRNA en los genes diana. Esto se logra mediante una modificación del gen de luciferasa que contiene la secuencia de ARNm target predicha de unión al miRNA (a menudo ubicado en la 3'-UTR) (Jin, Chen, Liu, & Zhou, 2013a). Este método de evolución de interacción mRNA/miRNA

ha sido ampliamente usado en diversos organismos ya que permite evaluar *in vitro* las interacciones predichas mediante plataformas de predicción de target (Aldred et al., 2011; Campos-Melo, Droppelmann, Volkening, & Strong, 2014; Guo et al., 2013; Hao et al., 2017; Jayachandran et al., 2013b; Kaewkascholkul et al., 2016; Nottrott et al., 2006; Sun et al., 2018). El ensayo del gen reportero utilizando luciferasa se utilizó en un estudio asociados a enfermedad de Alzheimer un trastorno neurodegenerativo progresivo que se caracteriza por la acumulación de placas de amiloide  $\beta$ . En el estudio demostraron que el microRNA miR34a inhibe la eliminación mediada por autofagia de las fibrillas A $\beta$ . Además, de identificar como gen target a catepsina B (CTSB) demostrando que miR34a inhibe la remoción de A $\beta$  al inhibir a CTSB (Sun, Fan, Liao, & Zhang, 2018).



## **SERPINAS**

Los inhibidores de la serina proteasa (Serpins) son una superfamilia de proteínas ampliamente distribuida con un dominio SERPIN y participan en varias respuestas inmunitarias. Un análisis comparativo de las secuencias de aminoácidos para Cr\_Serpina con otros miembros conocidos de la superfamilia de la serpinas reveló que la secuencia del RCL compone una de las regiones más conservadas. Dado que el RCL es un dominio importante de serpinas, responsables de su especificidad, este alto grado de identidad de secuencias, implica una posible conservación de las serpinas en *C. rogercresseyi*. El análisis filogenético también sugirió que Cr\_Serpina estaba estrechamente relacionados con Serpina B4 del copépodo *C. clemensi*.

En *C. rogercresseyi* se ha observado una expresión diferencial en las distintas etapas de desarrollo para *Cr-serpin10*, la cual se expresa principalmente en la etapa de copepodito, lo que sugiere una función putativa en la evasión de la respuesta inmune (Maldonado-Aguayo, W & Gallardo-Escárate, 2014). Estos resultados no concuerdan con los obtenidos en este estudio ya que se observó una mayor expresión en el estado de nauplius, sin embargo, en el estudio de (Maldonado-Aguayo, W & Gallardo-Escárate, 2014) no se evaluó la expresión en nauplius.

Según los resultados de predicción de target, bantam de *C. rogercresseyi* tendría como target a Cr\_Serpina el cual se uniría a su ARNm en el extremo 3'UTR con un valor de energía de hibridación de -22,7 kcal/mol. Sin embargo, en los ensayos de luciferasa no se observó una inhibición de la expresión de la proteína en presencia del microRNA. En *Panaeus monodon* se identificó un miRNA hómologo a Bantam pmo-bantam que tiene como gen target a un inhibidor serina proteasa tipo Kunitz (KuSPI). La unión de pmo-bantam y KuSPI se evaluó utilizando programas de predicción de target el cual entregó una energía de hibridación de -19,7 kcal/mol (Kaewkascholkul et al., 2016). Los estudios de interacción mRNA/miRNA se realizaron mediante ensayo de gen reportero utilizando luciferasa, en el cual a diferencia de nuestros resultados se observó un 20% de inhibición de la expresión de luciferasa (pmirGLO-KuSPI) en presencia de pmo-bantam (Kaewkascholkul et al., 2016). En *Bombyx mori* se evaluó si Bmserpin-28 está involucrado en la inmunidad innata para lo cual se inyectó ARNi de Bmserpin-28 en larvas. El nivel de ARNm de Bmserpin-28 fue suprimido después de 24 y 48 h, por otra parte, los niveles de expresión de ARNm de AMP aumentaron



significativamente. Estos resultados indicaron que Bmserpin-28 parece inhibir significativamente la transcripción de los péptidos antimicrobianos (Gao et al., 2018).

En *Eriocheir sinensis* Serpina-2 participa en el mecanismo inmune innato en respuesta a la infección por *S. eriocheiris*. Al utilizar ARN de doble cadena, dsRNA Esserpin-2 los resultados indicaron que el silenciamiento de Esserpin-2 podría reducir efectivamente la susceptibilidad de *S. eriocheiris*. Por otra parte, el porcentaje de mortalidad acumulada de los cangrejos tuvo una significativa disminución en el grupo Esserpin-2 dsRNA lo que evidencia que la tasa de supervivencia de los cangrejos se indujo por la limitada actividad de Esserpin-2 (Yuan et al., 2018).



## **TRIPSINA**

Las proteasas de serina (SP) constituyen una de las familias de genes más grandes en el reino animal. Los SP están involucrados en una amplia gama de funciones fisiológicas importantes en los mamíferos, incluida la digestión de proteínas de la dieta, la coagulación sanguínea, las respuestas inmunitarias, la transducción de señales, la activación de hormonas y el desarrollo (Jiang & Mellors, 2004). Por otro lado, el alineamiento múltiple de proteínas reveló la tríada catalítica (His-Asp-Ser) que son características de la serina proteasa familia S1. Tripsina de *C. rogercresseyi* presenta una similitud de un 47,86% para tripsina de *C. clemensi* y un 40% para otras tripsinas de *C. rogercresseyi*. Estos resultados son consistentes con el estudio realizado por (Johnson et al., 2002), donde se informa de una tripsina

en *L. salmonis* con un 46 % de similitud con tripsinas de crustáceos e insectos. Las diferencias en el péptido señal y la tríada catalítica entre proteasas de serina sugieren que estas variantes podrían estar involucrados en diferentes procesos moleculares. En los últimos años numerosos estudios han asociado la participación de estas proteasas en procesos como respuesta inmune, mecanismo de infección, desarrollo embrionario, incluso en la resistencia a fármacos.

Según los resultados de expresión en caligus, Cr\_Tripsina se expresa mayormente en estado hembra, lo cual ya se había observado en *Caligus rogercresseyi* en respuesta a fármacos (Valenzuela-Miranda & Gallardo-Escárte, 2016). Por otra parte, en *L. salmónis* se identificó y caracterizó una serina peptidasa de tipo tripsina (LsTryp10) que se distribuye uniformemente en los ovarios y ovocitos. Los resultados de traducción y degradación del ARNm de LsTryp10, demostraron que LsTryp10 se produce tanto por el embrión como por las hembras adultas (Skern-Mauritzen et al., 2009). Estos resultados sugieren que Cr\_Tripsina de caligus podría tener una función a nivel embrionario.

Los miRNA son pequeños RNA no codificantes de ~21-nucleótidos (nt) que están codificados dentro de los genomas de organismos multicelulares y una expresión génica regulada. A nivel molecular, los miRNAs se complementan imperfectamente con la región no traducible 3' (3' UTR) de los ARNm diana, generando una inhibición y desestabilización de la traducción del ARNm (Griffiths-Jones, Saini, van Dongen, & Enright, 2008; Kozomara & Griffiths-Jones, 2014). Bantam es un miRNA de la mosca *Drosophila melanogaster* el cual se caracterizó originalmente a partir de genes que promueven la proliferación celular y suprimen la apoptosis durante el crecimiento del tejido (Hipfner, Weigmann, & Cohen, 2002).

El gen bantam de *Drosophila* codifica un miARN de 21 nucleótidos que promueve el crecimiento de los tejidos. La expresión bantam está regulada temporal y espacialmente en respuesta a ciertos patrones. En *C. rogercresseyi* se han identificado más de 200 isoformas de bantam que se expresan en los diferentes estados de desarrollo (Gallardo-Escárate, C et al., 2017). Según los resultados de predicción de target bantam en caligus regularía de manera post-transcripcional a los genes del secretoma catepsina, serpina y tripsina, por lo cual se observó mediante el ensayo de luciferasa que inhibe la expresión de tripsina en un 50%. En *Helicoverpa armígera* se identificó un miRNA, har-miR-2002b, el cual se expresa específicamente durante los estadios larvales, y el objetivo del miRNA es una serina proteasa similar a la tripsina (Ha-TLP). Los autores utilizaron un imitador de har-mir-2002 mediante alimentación oral, los resultados indicaron una reducción del 70% en la fecundidad y del 40% de mortalidad larvaria en presencia del imitador de har-mi-2002b. Además, los niveles de transcripción de Ha-TLP y la actividad enzimática se redujeron significativamente. En contraste, la inhibición de har-miR-2002b condujo a una importante regulación al alza de los niveles de transcripción de Ha-TLP. Este es el primer estudio que muestra el potencial de utilización de imitadores e inhibidores de miRNA en el control de plagas, al atacar genes de insectos específicos (Jayachandran et al., 2013a) por lo que se puede plantear la hipótesis que en *C. rogercresseyi* bantam puede ser utilizado para el control de caligus dado que regula la expresión de Cr\_tripsina de manera negativa. Sin embargo, en caligus no se hicieron ensayos *in vivo*.

## VITELOGENINA I

Ilumina es una plataforma líder en NGS (Next Generation Sequencing), que proporciona gran cantidad de datos, a bajo coste y con baja tasa de error, sin embargo, entre sus desventajas se encuentra que entrega lecturas cortas. Esta desventaja se pudo ver en la búsqueda de Vitelogenina I, en la base de datos de *C. rogercresseyi* (Farlora, Araya-Garay, & Gallardo-Escárate, 2014a; Farlora et al., 2014b; Gallardo-Escárate, Cristian, Valenzuela-Muñoz, & Nuñez-Acuña, 2014). La cual al ser de más de 5000 nucleótidos fue imposible obtener una secuencia completa y sin errores, sin embargo, nos permitió tener una secuencia consenso la cual nos permitió la obtención del extremo 3'UTR mediante la técnica RACE.

Para Cr\_Vitelogenina I se obtuvo una secuencia consenso de 2356 nucleótidos el cual mediante un blast se pudo predecir que presenta un 73% de identidad con *Lepeophtheirus salmonis*, lo cual también se observó en el árbol filogenético agrupándose en el mismo clado que Vitelogenina I de *Lepeophtheirus salmonis*.

De acuerdo a los análisis de expresión relativa para Vitelogenina I de *C. rogercresseyi* se observó mayor expresión del gen en el estado de hembra. Estos resultados son similares a los de (Farlora, Araya-Garay, & Gallardo-Escárate, 2014b). En *Lepeophtheirus salmonis* también se han caracterizado 2 Vitelogeninas (LsVit1 y LsVit2), asociando su función en reproducción como proveedores de la nutrición embrionaria y larvaria (Dalvin et al., 2011).

Para los análisis de predicción de target para Cr\_Vitelogenina I fue necesario obtener la secuencia de 3'UTR de Viteligenina I. El análisis de predicción de target para Vitelogenina I se realizó utilizando el programa RNAhybrid.2.1.2. Para Vitelogenina I la secuencia de unión al microRNA mir-996 fue "5'-AATAAAGUCATATCA -3'" ubicada en el extremo 3'UTR entre los nucleótidos 86 y 106 con un valor de energía de hibridación de  $-10,9$  kcal/mol. La unión predicha por las plataformas bioinformáticas fueron corroboradas *in vitro* observando una disminución en la actividad de la luciferasa en un 30%. En garrapatas mediante análisis de predicción de target utilizando RNAHíbrid se predijo que el microRNA mir-275 se unía con una energía de hibridación de  $-21,9$  y  $-22$  kcal/ mol a Vitelogenina 2 (Hao et al., 2017), a pesar que estos análisis se hicieron en la región ORF del gen donde el sitio de unión al microRNA se encontraban en la posición 1461 y 1569 (Hao et al., 2017), se observó mediante ensayo de luciferasa que mir-275 inhibe la expresión de la proteína de la luciferasa. Los miRNAs se suelen unir a la región 3' UTR del transcrito, donde inician la degradación del mRNA o inhiben la traducción. El silenciamiento mediado por miRNAs se llevó a cabo mediante inhibición de la traducción o por degradación dependiente del grado de complementariedad entre el miRNA y la región 3'UTR del transcrito (Duan, Wang, & Guo, 2006; Kertesz, Iovino, Unnerstall, Gaul, & Segal, 2007; Lugo Trampe & Trujillo Murillo, 2009). Sin embargo, existen varios estudios de evaluación de los sitios de unión de microARN mediante el ensayo del gen reportero de la luciferasa en el cual los ensayos de predicción de target se hacen en la región 5'UTR y ORF (Campos-Melo, Droppelmann, Volkening, & Strong, 2014; Jayachandran, Hussain, & Asgari, 2013b; Moretti et al., 2010; Ørom & Lund, 2010)

En garrapatas *Haemophysalis longicornis* se identificó Vg-2 como un gen target para miR-275, el cual redujo la actividad de la luciferasa en un 66,08%. Por otra parte, el silenciamiento de miR-275 dio lugar a problemas de digestión de la sangre, deterioro en el desarrollo del ovario y reducción significativa de la masa de los huevos. Además, el silenciamiento por ARN de interferencia de Vg-2 afectó la captación de sangre y la masa del huevo (Hao et al., 2017).



## CONCLUSIONES

El presente estudio, contempla el primer análisis de funcionalidad de miRNA en *Caligus rogercresseyi*. Los análisis de expresión diferencial indican que las proteasas del secretoma de caligus incluyendo tripsinas, catepsinas y serpina y vitelogenina fueron diferencialmente expresadas entre estadios. Se realizaron análisis de predicción de target de los miRNA bantam y mir-996 con 4 genes, los cuales fueron clonados aguas abajo del gen de la luciferasa.

- Los resultados presentados en este estudio indican que Cr\_Catepsina se encuentra diferencialmente expresada en estado hembra sugiriendo su participación en desarrollo embrionario. Además, es un gen target para miRNA bantam, el cual mediante su unión al mRNA inhibe la actividad de la luciferasa, demostrando la interacción mRNA/miRNA.
- La secuencia inhibidora de serin proteasa descrita en este estudio Cr\_Serpina 10 se encuentra mayormente expresado en los estadios temprano de caligus, sugiriendo su participación en mecanismos de evasión de la respuesta inmune proporcionada por el huésped. Cr\_Serpina 10 es un gen target del miRNA bantam, sin embargo, no fue demostrado mediante el ensayo de la luciferasa.
- En el presente estudio la proteasa de serina Cr\_Tripsina se encuentra diferencialmente expresada en estado hembra sugiriendo su

participación en desarrollo embrionario. Además, es un gen target para miRNA bantam, el cual mediante su unión al mRNA inhibe la actividad de la luciferasa, demostrando la interacción mRNA/miRNA.

- Cr\_Vitelogenina I se encuentra mayor mente expresado en hembra, sugiriendo su participación en reproducción, por otra parte, es un gen target para el miRNA mir-996. En el presente estudio se ha demostrado la interacción mRNA/miRNA, mediante la inhibición de la actividad de la luciferasa.

Finalmente, los resultados proporcionados en este estudio permitirán acelerar el entendimiento sobre las bases moleculares sobre el proceso infectivo y reproductivo en *C. rogercresseyi*, así como también el desarrollo de nuevas estrategias para el control de la caligidosis.





## BIBLIOGRAFIA

- Aldred, S. F., Collins, P., & Trinklein, N. (2011). Identifying targets of human microRNAs with the LightSwitch Luciferase Assay System using 3'UTR-reporter constructs and a microRNA mimic in adherent cells. *Journal of visualized experiments: JoVE*(55).
- Arukwe, A., & Goksøyr, A. (2003). Eggshell and egg yolk proteins in fish: hepatic proteins for the next generation: oogenetic, population, and evolutionary implications of endocrine disruption. *Comparative hepatology*, 2(1), 4.
- Blow, D. M. (1997). The tortuous story of Asp... His... Ser: Structural analysis of  $\alpha$ -chymotrypsin. *Trends in biochemical sciences*, 22(10), 405-408.
- Bonin-Debs, A. L., Boche, I., Gille, H., & Brinkmann, U. (2004). Development of secreted proteins as biotherapeutic agents. *Expert opinion on biological therapy*, 4(4), 551-558.
- Bravo, S., Nunez, M., & Silva, M. (2013). Efficacy of the treatments used for the control of *Caligus rogercresseyi* infecting Atlantic salmon, *Salmo salar* L., in a new fish-farming location in Region XI, Chile. *Journal of fish diseases*, 36(3), 221-228.
- Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., & Cohen, S. M. (2003). Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell*, 113(1), 25-36.
- Brömme, D., & Wilson, S. (2011). Role of cysteine cathepsins in extracellular proteolysis. In *Extracellular matrix degradation* (pp. 23-51): Springer.
- Caccia, D., Dugo, M., Callari, M., & Bongarzone, I. (2013). Bioinformatics tools for secretome analysis. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1834(11), 2442-2453.
- Campos-Melo, D., Droppelmann, C. A., Volkening, K., & Strong, M. J. (2014). Comprehensive luciferase-based reporter gene assay reveals previously masked up-regulatory effects of miRNAs. *International journal of molecular sciences*, 15(9), 15592-15602.
- Carmona, E., Dufour, É., Plouffe, C., Takebe, S., Mason, P., Mort, J. S., & Ménard, R. (1996). Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry*, 35(25), 8149-8157.
- 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.
- Chávez-Mardones, J., Asencio, G., Latuz, S., & Gallardo-Escárate, C. (2015). Hydrogen peroxide modulates antioxidant system transcription, evidencing sex-dependent responses in *Caligus rogercresseyi*. *Aquaculture Research*.
- Clynen, E., Schoofs, L., & Salzert, M. (2005). A review of the most important classes of serine protease inhibitors in insects and leeches. *Medicinal Chemistry Reviews-Online*, 2(3), 197-206.
- Colinet, D., Dubuffet, A., Cazes, D., Moreau, S., Drezen, J.-M., & Poirié, M. (2009). A serpin from the parasitoid wasp *Leptopilina boulardi* targets the *Drosophila* phenoloxidase cascade. *Developmental & Comparative Immunology*, 33(5), 681-689.
- Costello, M. J. (2009). The global economic cost of sea lice to the salmonid farming industry. *Journal of fish diseases*, 32(1), 115-118.
- Daborn, P. J., & Le Goff, G. (2004). The genetics and genomics of insecticide resistance. *TRENDS in Genetics*, 20(3), 163-170.
- Dalton, J. P., Neill, S. O., Stack, C., Collins, P., Walshe, A., Sekiya, M., . . . Khaznadji, E. (2003). *Fasciola hepatica* cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines. *International journal for parasitology*, 33(11), 1173-1181.
- Dalvin, S., Frost, P., Loeffen, P., Skern-Mauritzen, R., Baban, J., Rønnestad, I., & Nilsen, F. (2011). Characterisation of two vitellogenins in the salmon louse *Lepeophtheirus salmonis*: molecular, functional and evolutionary analysis. *Diseases of aquatic organisms*, 94(3), 211-224.
- Danwattananusorn, T., Kondo, H., Aoki, T., & Hirono, I. (2009). Molecular cloning, characterization and expression analysis of a chymotrypsin-like serine protease from kuruma shrimp *Marsupenaeus japonicus*. *Fisheries Science*, 75(5), 1231-1238.

- Dorémus, T., Urbach, S., Jouan, V., Cousserans, F., Ravallec, M., Demettré, E., . . . Darboux, I. (2013). Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins. *Insect biochemistry and molecular biology*, 43(3), 292-307.
- Duan, C., Wang, C., & Guo, H. (2006). Regulation of microRNA on plant development and viral infection. *Chinese Science Bulletin*, 51(3), 269-278.
- Farlora, R., Araya-Garay, J., & Gallardo-Escárate, C. (2014a). Discovery of sex-related genes through high-throughput transcriptome sequencing from the salmon louse *Caligus rogercresseyi*. *Marine genomics*, 15, 85-93.
- Farlora, R., Araya-Garay, J., & Gallardo-Escárate, C. J. M. g. (2014b). Discovery of sex-related genes through high-throughput transcriptome sequencing from the salmon louse *Caligus rogercresseyi*. 15, 85-93.
- Farlora, R., Nuñez-Acuña, G., & Gallardo-Escárate, C. J. G. (2015). Prohibitin-2 gene reveals sex-related differences in the salmon louse *Caligus rogercresseyi*. 564(1), 73-80.
- Ford, L., Zhang, J., Liu, J., Hashmi, S., Fuhrman, J. A., Oksov, Y., & Lustigman, S. (2009). Functional analysis of the cathepsin-like cysteine protease genes in adult *Brugia malayi* using RNA interference. *PLoS neglected tropical diseases*, 3(2), e377.
- Franta, Z., Sojka, D., Frantova, H., Dvorak, J., Horn, M., Srba, J., . . . Craik, C. S. (2011). IrCL1—The haemoglobinolytic cathepsin L of the hard tick, *Ixodes ricinus*. *International journal for parasitology*, 41(12), 1253-1262.
- Gallardo-Escárate, C., Valenzuela-Muñoz, V., Boltaña, S., Nuñez-Acuña, G., Valenzuela-Miranda, D., Gonçalves, A., . . . Roberts, S. J. A. G. (2017). The *Caligus rogercresseyi* miRNome: Discovery and transcriptome profiling during the sea lice ontogeny. 4, 8-22.
- 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.
- Gao, Q., Yang, L., Dai, J., Yuan, G., Wang, L., Qian, C., . . . Wei, G. (2018). Characterization and functional analysis of serpin-28 gene from silkworm, *Bombyx mori*. *Journal of Invertebrate Pathology*.
- Gonzalez, L., & Carvajal, J. (2003). Life cycle of *Caligus rogercresseyi*, (Copepoda: Caligidae) parasite of Chilean reared salmonids. *Aquaculture*, 220, 101-117. doi:[http://dx.doi.org/10.1016/S0044-8486\(02\)00512-4](http://dx.doi.org/10.1016/S0044-8486(02)00512-4)
- González, L., & Carvajal, J. (2003). Life cycle of *Caligus rogercresseyi*, (Copepoda: Caligidae) parasite of Chilean reared salmonids. *Aquaculture*, 220(1-4), 101-117.
- Gourbal, B. E., Guillou, F., Mitta, G., Sibille, P., Thèron, A., Pointier, J.-P., & Coustau, C. (2008). Excretory–secretory products of larval *Fasciola hepatica* investigated using a two-dimensional proteomic approach. *Molecular and biochemical parasitology*, 161(1), 63-66.
- Griffiths-Jones, S., Saini, H. K., van Dongen, S., & Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic acids research*, 36(suppl 1), D154-D158.
- Groves, M. R., Coulombe, R., Jenkins, J., & Cygler, M. (1998). Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. *Proteins: Structure, Function, and Bioinformatics*, 32(4), 504-514.
- Guo, R., Abdelmohsen, K., Morin, P. J., & Gorospe, M. (2013). Novel microRNA reporter uncovers repression of let-7 by GSK-3 $\beta$ . *PLoS one*, 8(6), e66330.
- Hackenberg, M., Rodríguez-Ezpeleta, N., & Aransay, A. M. (2011). miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. *Nucleic acids research*, 39(suppl 2), W132-W138.
- Hamilton, S., McLean, K., Monaghan, S. J., McNair, C., Inglis, N. F., McDonald, H., . . . vectors. (2018). Characterisation of proteins in excretory/secretory products collected from salmon lice, *Lepeophtheirus salmonis*. 11(1), 294.
- Hao, J., Luo, J., Chen, Z., Ren, Q., Guo, J., Liu, X., . . . Luo, J. (2017). MicroRNA-275 and its target Vitellogenin-2 are crucial in ovary development and blood digestion of *Haemaphysalis longicornis*. *Parasites & vectors*, 10(1), 253.

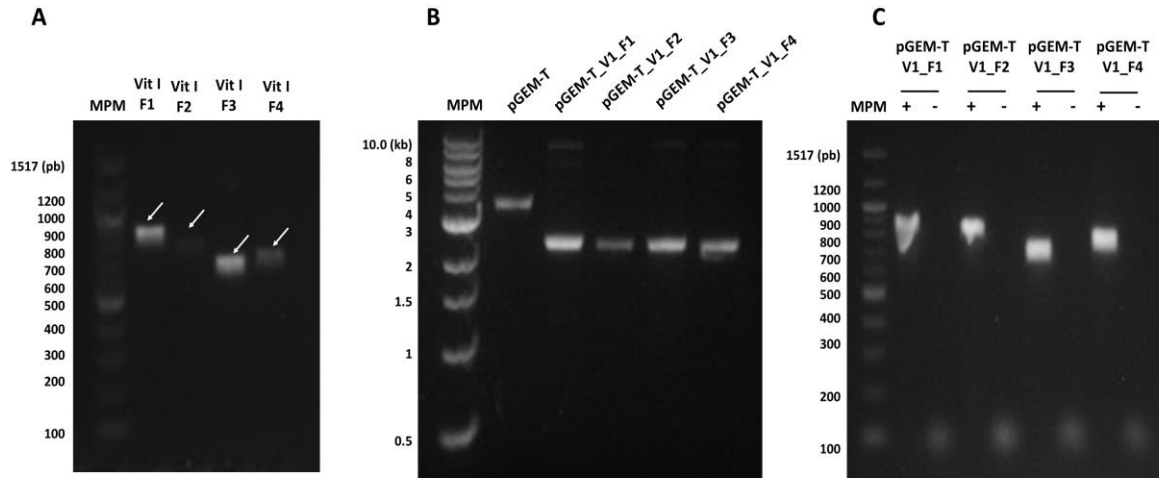
- Hashmi, S., Zhang, J., Oksov, Y., & Lustigman, S. (2004). The *Caenorhabditis elegans* cathepsin Z-like cysteine protease, Ce-CPZ-1, has a multifunctional role during the worms' development. *Journal of Biological Chemistry*, 279(7), 6035-6045.
- Hipfner, D. R., Weigmann, K., & Cohen, S. M. (2002). The bantam gene regulates *Drosophila* growth. *Genetics*, 161(4), 1527-1537.
- Igboeli, O., Purcell, S., Wotton, H., Poley, J., Burka, J., & Fast, M. (2013). Immunostimulation of *Salmo salar* L., and its effect on *Lepeophtheirus salmonis* (Krøyer) P-glycoprotein mRNA expression following subsequent emamectin benzoate exposure. *Journal of fish diseases*, 36(3), 339-351.
- Itoh, R., Kawamoto, S., Adachi, W., Kinoshita, S., & Okubo, K. (1999). Genomic organization and chromosomal localization of the human cathepsin L2 gene. *DNA Research*, 6(2), 137-140.
- Jayachandran, B., Hussain, M., & Asgari, S. (2013a). An insect trypsin-like serine protease as a target of microRNA: utilization of microRNA mimics and inhibitors by oral feeding. *Insect biochemistry and molecular biology*, 43(4), 398-406.
- Jayachandran, B., Hussain, M., & Asgari, S. (2013b). Regulation of *Helicoverpa armigera* ecdysone receptor by miR-14 and its potential link to baculovirus infection. *Journal of Invertebrate Pathology*, 114(2), 151-157.
- Jiang, P., & Mellors, A. (2004). Handbook of proteolytic enzymes. In: London: Elsevier.
- Jin, Y., Chen, Z., Liu, X., & Zhou, X. (2013a). Evaluating the microRNA targeting sites by luciferase reporter gene assay. In *MicroRNA Protocols* (pp. 117-127): Springer.
- Jin, Y., Chen, Z., Liu, X., & Zhou, X. (2013b). Evaluating the microRNA targeting sites by luciferase reporter gene assay. *MicroRNA Protocols*, 117-127.
- Johnson, S., Ewart, K., Osborne, J., Delage, D., Ross, N., & Murray, H. (2002). Molecular cloning of trypsin cDNAs and trypsin gene expression in the salmon louse *Lepeophtheirus salmonis* (Copepoda: Caligidae). *Parasitology research*, 88(9), 789-796.
- Kaewkascholkul, N., Somboonviwat, K., Asakawa, S., Hirono, I., Tassanakajon, A., & Somboonviwat, K. (2016). Shrimp miRNAs regulate innate immune response against white spot syndrome virus infection. *Developmental & Comparative Immunology*, 60, 191-201.
- Kanoria, S., Rennie, W., Liu, C., Carmack, C. S., Lu, J., & Ding, Y. (2016). STarMir Tools for Prediction of microRNA Binding Sites. *RNA Structure Determination: Methods and Protocols*, 73-82.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., & Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nature genetics*, 39(10), 1278-1284.
- Kozomara, A., & Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research*, 42(D1), D68-D73.
- Lakka, S. S., Gondi, C. S., Yanamandra, N., Olivero, W. C., Dinh, D. H., Gujrati, M., & Rao, J. S. (2004). Inhibition of cathepsin B and MMP-9 gene expression in glioblastoma cell line via RNA interference reduces tumor cell invasion, tumor growth and angiogenesis. *Oncogene*, 23(27), 4681.
- Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75(5), 843-854.
- Lehiy, C. J., & Drolet, B. S. (2014). The salivary secretome of the biting midge, *Culicoides sonorensis*. *PeerJ*, 2, e426.
- LeMosy, E. K., Hong, C. C., & Hashimoto, C. (1999). Signal transduction by a protease cascade. *Trends in cell biology*, 9(3), 102-107.
- León, N., Padilla, C., Pajuelo, M., Sheen, P., & Zimic, M. (2013). Cisteinoproteasas Catepsinas L de *Taenia solium*: Rol biológico en la infección y potencial uso para el inmunodiagnóstico de la neurocisticercosis. *Revista Peruana de Medicina Experimental y Salud Pública*, 30, 446-454.
- Li, J., Xiang, M., Zhang, R., Xu, B., & Hu, W. (2018). RNA interference in vivo in *Schistosoma japonicum*: Establishing and optimization of RNAi mediated suppression of gene expression by long dsRNA in the intra-mammalian life stages of worms. *Biochemical and Biophysical Research Communications*.
- Lugo Trampe, Á., & Trujillo Murillo, K. d. C. (2009). MicroRNAs: reguladores clave de la expresión génica. *Medicina universitaria*, 11(44).
- M Witkos, T., Koscianska, E., & J Krzyzosiak, W. (2011). Practical aspects of microRNA target prediction. *Current molecular medicine*, 11(2), 93-109.

- Mach, L., Mort, J. S., & Glössl, J. (1994). Maturation of human procathepsin B. Proenzyme activation and proteolytic processing of the precursor to the mature proteinase, in vitro, are primarily unimolecular processes. *Journal of Biological Chemistry*, 269(17), 13030-13035.
- Maizels, R. M., & Yazdanbakhsh, M. (2003). Immune regulation by helminth parasites: cellular and molecular mechanisms. *Nature Reviews Immunology*, 3(9), 733.
- Maldonado-Aguayo, W., Chávez-Mardones, J., Gonçalves, A. T., & Gallardo-Escárate, C. (2015). Cathepsin gene family reveals transcriptome patterns related to the infective stages of the salmon louse *Caligus rogercresseyi*. 10(4), e0123954.
- Maldonado-Aguayo, W., & Gallardo-Escárate, C. (2014). Increasing transcriptome response of serpins during the ontogenetic stages in the salmon louse *Caligus rogercresseyi* (Copepoda: Caligidae). 15, 55-64.
- McCarthy, E., Cunningham, E., Copley, L., Jackson, D., Johnston, D., Dalton, J. P., & Mulcahy, G. J. A. (2012). Cathepsin L proteases of the parasitic copepod, *Lepeophtheirus salmonis*. 356, 264-271.
- McGonigle, L., Mousley, A., Marks, N. J., Brennan, G. P., Dalton, J. P., Spithill, T. W., . . . Maule, A. G. (2008). The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *International journal for parasitology*, 38(2), 149-155.
- Meekins, D. A., Kanost, M. R., & Michel, K. (2017). *Serpins in arthropod biology*. Paper presented at the Seminars in cell & developmental biology.
- Mei, Q., Li, X., Meng, Y., Wu, Z., Guo, M., Zhao, Y., . . . Han, W. (2012). A facile and specific assay for quantifying microRNA by an optimized RT-qPCR approach. *PLoS one*, 7(10), e46890.
- Moreno, Y., & Geary, T. G. (2008). Stage-and gender-specific proteomic analysis of *Brugia malayi* excretory-secretory products. *PLoS neglected tropical diseases*, 2(10), e326.
- Moretti, F., Thermann, R., & Hentze, M. W. (2010). Mechanism of translational regulation by miR-2 from sites in the 5' untranslated region or the open reading frame. *Rna*.
- Mulcahy, G., & Dalton, J. (2001). Cathepsin L proteinases as vaccines against infection with *Fasciola hepatica* (liver fluke) in ruminants. *Research in veterinary science*, 70(1), 83-86.
- Mulvenna, J., Hamilton, B., Nagaraj, S. H., Smyth, D., Loukas, A., Gorman, J. J. J. M., & Proteomics, C. (2009). Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, *Ancylostoma caninum*. 8(1), 109-121.
- Nagaraj, S. H., Gasser, R. B., & Ranganathan, S. (2008). Needles in the EST haystack: large-scale identification and analysis of excretory-secretory (ES) proteins in parasitic nematodes using expressed sequence tags (ESTs). *PLoS neglected tropical diseases*, 2(9), e301.
- Nottrott, S., Simard, M. J., & Richter, J. D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nature Structural and Molecular Biology*, 13(12), 1108.
- Ørom, U. A., & Lund, A. H. (2010). Experimental identification of microRNA targets. *Gene*, 451(1-2), 1-5.
- Patil, V. S., Zhou, R., & Rana, T. M. (2014). Gene regulation by non-coding RNAs. *Critical reviews in biochemistry and molecular biology*, 49(1), 16-32.
- Perona, J. J., & Craik, C. S. (1995). Structural basis of substrate specificity in the serine proteases. *Protein Science*, 4(3), 337-360.
- Pino-Marambio, J., Mordue, A., Birkett, M., Carvajal, J., Asencio, G., Mellado, A., & Quiroz, A. (2007). Behavioural studies of host, non-host and mate location by the Sea Louse, *Caligus rogercresseyi* Boxshall & Bravo, 2000 (Copepoda: Caligidae). *Aquaculture*, 271(1), 70-76.
- Prowse, R. k., Chaplin, P., Robinson, H. C., & Spithill, T. W. (2002). *Fasciola hepatica* cathepsin L suppresses sheep lymphocyte proliferation in vitro and modulates surface CD4 expression on human and ovine T cells. *Parasite Immunology*, 24(2), 57-66.
- Ranganathan, S., & Garg, G. (2009a). Secretome: clues into pathogen infection and clinical applications. *Genome medicine*, 1(11), 113.
- Ranganathan, S., & Garg, G. J. G. m. (2009b). Secretome: clues into pathogen infection and clinical applications. 1(11), 113.
- Robinson, M. W., Menon, R., Donnelly, S. M., Dalton, J. P., Ranganathan, S. J. M., & Proteomics, C. (2009). An integrated transcriptomics and proteomics analysis of the secretome of the helminth pathogen *Fasciola hepatica* proteins associated with invasion and infection of the mammalian host. 8(8), 1891-1907.

- Rozas, M., & Asencio, G. (2006). <Evaluación de la Situación Epidemiológica de la Caligiasis en Chile: Hacia una Estrategia de Control Efectiva>. *SalmoCiencia*, 43-59.
- Rozas, M., & Asencio, G. (2007). Evaluación de la Situación Epidemiológica de la Caligiasis en Chile: Hacia una estrategia de control efectiva. *Salmociencia*, 2(1), 43-59.
- Shirdel, E. A., Xie, W., Mak, T. W., & Jurisica, I. (2011). NAViGaTing the microneome—using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs. *PLoS one*, 6(2), e17429.
- Silva, L., Reis, A., Pereira, E., Oliveira, M., & Guedes, R. (2010). Partial purification and characterization of trypsin-like proteinases from insecticide-resistant and-susceptible strains of the maize weevil, *Sitophilus zeamais*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 155(1), 12-19.
- Skelly, P. J., Da'dara, A., & Harn, D. A. (2003). Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference. *International journal for parasitology*, 33(4), 363-369.
- Skern-Mauritzen, R., Frost, P., Dalvin, S., Kvamme, B. O., Sommerset, I., & Nilsen, F. (2009). A trypsin-like protease with apparent dual function in early *Lepeophtheirus salmonis* (Krøyer) development. *BMC Molecular Biology*, 10(1), 44.
- Sojka, D., Francischetti, I. M., Calvo, E., & Kotsyfakis, M. (2011). Cysteine proteases from bloodfeeding arthropod ectoparasites. In *Cysteine Proteases of Pathogenic Organisms* (pp. 177-191): Springer.
- Stack, C., Dalton, J. P., & Robinson, M. W. (2011). The phylogeny, structure and function of trematode cysteine proteases, with particular emphasis on the *Fasciola hepatica* cathepsin L family. In *Cysteine proteases of pathogenic organisms* (pp. 116-135): Springer.
- Sun, B., Fan, P., Liao, M., & Zhang, Y. (2018). Modeling endophilin-mediated A $\beta$  disposal in glioma cells. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1865(10), 1385-1396.
- Takacs, C. M., & Giraldez, A. J. (2010). *MicroRNAs as genetic sculptors: fishing for clues*. Paper presented at the Seminars in cell & developmental biology.
- Tan, G.-J., Peng, Z.-K., Lu, J.-P., & Tang, F.-Q. (2013). Cathepsins mediate tumor metastasis. *World journal of biological chemistry*, 4(4), 91.
- Tufail, M., & Takeda, M. (2008). Molecular characteristics of insect vitellogenins. *Journal of insect physiology*, 54(12), 1447-1458.
- Valenzuela-Miranda, D., & Gallardo-Escárate, C. J. A. (2016). *Caligus rogercresseyi* serine proteases: Transcriptomic analysis in response to delousing drugs treatments. 465, 65-77.
- Valenzuela-Miranda, D., Nuñez-Acuña, G., Valenzuela-Muñoz, V., Asgari, S., & Gallardo-Escárate, C. (2015). MicroRNA biogenesis pathway from the salmon louse (*Caligus rogercresseyi*): Emerging role in delousing drug response. *Gene*, 555(2), 231-241.
- Vera-Bizama, F., Valenzuela-Muñoz, V., Gonçalves, A. T., Marambio, J. P., Hawes, C., Wadsworth, S., & Gallardo-Escárate, C. (2015). Transcription expression of immune-related genes from *Caligus rogercresseyi* evidences host-dependent patterns on Atlantic and coho salmon. *Fish & shellfish immunology*, 47(2), 725-731.
- Xiao, C., & Rajewsky, K. (2009). MicroRNA control in the immune system: basic principles. *Cell*, 136(1), 26-36.
- Yuan, M., Ning, M., Wei, P., Hao, W., Jing, Y., Gu, W., . . . Meng, Q. (2018). The function of serpin-2 from *Eriocheir sinensis* in *Spiroplasma eriocheiris* infection. *Fish & shellfish immunology*, 76, 21-26.
- Zabel, B. A., Allen, S. J., Kulig, P., Allen, J. A., Cichy, J., Handel, T. M., & Butcher, E. C. (2005). Chemerin activation by serine proteases of the coagulation, fibrinolytic, and inflammatory cascades. *Journal of Biological Chemistry*, 280(41), 34661-34666.
- Zhang, S., Wang, S., Li, H., & Li, L. (2011). Vitellogenin, a multivalent sensor and an antimicrobial effector. *The international journal of biochemistry & cell biology*, 43(3), 303-305.
- Zheng, X., Chu, F., Mirkin, B. L., Sudha, T., Mousa, S. A., & Rebbaa, A. (2008). Role of the proteolytic hierarchy between cathepsin L, cathepsin D and caspase-3 in regulation of cellular susceptibility to apoptosis and autophagy. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1783(12), 2294-2300.
- Zhuo, Y., Gao, G., Shi, J. a., Zhou, X., & Wang, X. (2013). miRNAs: biogenesis, origin and evolution, functions on virus-host interaction. *Cellular Physiology and Biochemistry*, 32(3), 499-510.

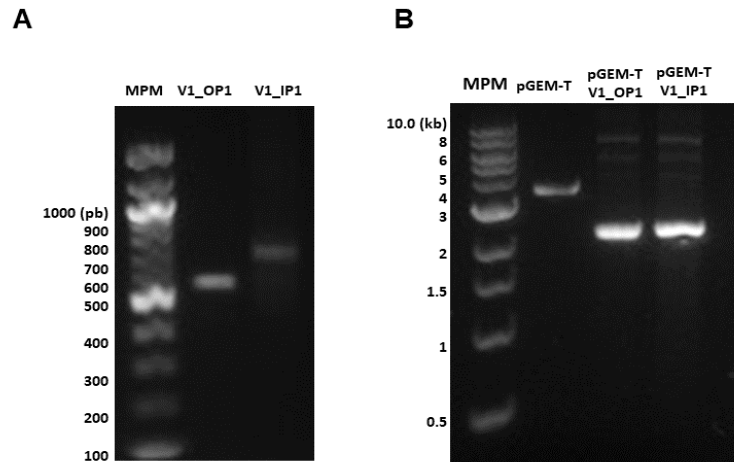
## MATERIAL SUPLEMENTARIO

### Clonamiento y secuenciación del extremo 3' de la ORF de Vitelogenina I



**Figura S36.** Resultados clonamiento de los fragmentos del gen de vitelogenina I. A: Producto pcr obtenido por pcr en tiempo final para ser clonado en el vector pGEM-T esasy system para los fragmentos 1, 2, 3 y 4 de la ORF parcial de Vitelogenina I. B: Plásmido pGEM-T purificado con el inserto de los fragmentos de la ORF parcial de Vitelogenina I. C: PCR en tiempo final para el plásmido pGEM\_fragments\_ORF para verificar la presencia de los insertos de la ORF parcial de Vitelogenina I (Fuente: Elaboración propia).

### 3'RACE para Vitelogenina I.



**Figura S37.** Resultados clonamiento de los fragmentos obtenidos mediante 3'RACE. A: Producto pcr obtenido por pcr en tiempo final para ser clonado en el vector pGEM-T esasy system II, B: Plásmido pGEM\_UTR purificado con los fragmentos OP1, IP1 obtenidos por 3'RACE (Fuente: Elaboración propia).

