



**UNIVERSIDAD NACIONAL DE LA PLATA**

**FACULTAD DE CIENCIAS EXACTAS**

**DEPARTAMENTO ÁREA CIENCIAS BIOLÓGICAS**

---

**Trabajo de Tesis Doctoral:**

Genómica del desarrollo embrionario de  
*Rhodnius prolixus*.

**Tesista:**

Pascual Agustina

**Director/a:**

Rivera Pomar Rolando

**Año:**

2018

# Resumen

El desarrollo de nuevos y asequibles métodos para la secuenciación de genomas completos, sumado al diseño de grandes proyectos de anotación genómica producen una gran cantidad de información que debe convertirse en datos funcionales. Dentro de este contexto, el análisis comparativo contribuye a generar información clave tanto para comprender la evolución como para la simple identificación génica en organismos desconocidos hasta el momento.

*Rhodnius prolixus* surgió como organismo modelo de estudio en un primer momento para estudios fisiológicos y bioquímicos, los cuales fueron llevados a cabo por Sir V. B. Wigglesworth. Pasados los años y con el advenimiento de las técnicas moleculares, así como también la disponibilidad y acceso a las tecnologías de secuenciación masiva de segunda generación, se re-direccionó el foco de investigación a estudios genéticos.

Contando como base con los estudios llevados a cabo en la década del 70 por Erwin Huebner sobre la morfología y citología de la estructura del ovario en *R. prolixus*, en este trabajo se propuso profundizar su descripción con técnicas moleculares modernas y analizar el perfil transcripcional de los ARN mensajeros durante la oogénesis y embriogénesis, para determinar la expresión génica en las diferentes etapas del desarrollo y validar la función de los genes relacionados con la oogénesis y con su contribución materna al huevo. Se generó información sobre el genoma expresado de *R. prolixus* desde la etapa de la oogénesis hasta los primeros estadios embrionarios. El material crudo obtenido tras la secuenciación se procesó para efectuar el posterior ensamble. Los transcriptos reconstruidos fueron constatados contra la base de datos generada a partir de los genes conocidos que se expresan durante el desarrollo embrionario de *Drosophila melanogaster*. De esta manera, se identificó un amplio repertorio de genes que se expresan durante los estadios estudiados, de los cuales se analizaron experimentalmente *Bicaudal C*, *Bicaudal D* y *cornichon*. Estos se validaron funcionalmente mediante la interferencia del ARN (ARNi) parental; en donde se evaluó analizando fertilidad y fenotipo a nivel de ovario, huevo y embrión.

Se observó que *Bicaudal C* presenta un patrón de expresión materno y folicular en el ovario. Actúa manteniendo la estructura del epitelio folicular de manera organizada durante todo el proceso de oogénesis, dando lugar a una acumulación controlada de vitelo y a un correcto establecimiento del corion. *Bicaudal D* se vio que es necesario para la producción de huevos embrionados. Su silenciamiento génico llevó a la oviposición de huevos en los cuales no fue

posible identificar ninguna estructura característica que permita inferir los ejes embrionarios que determinan el patrón corporal. Presenta una expresión génica materna y folicular a nivel del ovario; así como también expresión a nivel embrionario, exceptuando el estadio de cigoto, donde no se evidenció presencia alguna. *cornichon* no presentó función alguna durante la embriogénesis, aunque sí exhibió una activa expresión génica durante el desarrollo embrionario temprano, desde el estado de huevo no fertilizado hasta embrión gastrulante.

# **Agradecimientos**

A mi papá!

# Indice

<b>RESUMEN .....</b>	<b>2</b>
<b>AGRADECIMIENTOS .....</b>	<b>5</b>
<b>INDICE.....</b>	<b>7</b>
<b>ABREVIATURAS .....</b>	<b>11</b>
<b>INTRODUCCIÓN GENERAL.....</b>	<b>13</b>
1. EMBRIOLOGÍA DE INSECTOS .....	14
2. ESTUDIO DEL DESARROLLO EMBRIONARIO EN ORGANISMOS MODELOS NO CONVENCIONALES.....	19
3. <i>RHODNIUS PROLIXUS</i> COMO ORGANISMO MODELO DE ESTUDIO .....	20
<b>HIPÓTESIS Y OBJETIVOS .....</b>	<b>24</b>
1. HIPÓTESIS .....	25
2. OBJETIVO GENERAL.....	25
3. OBJETIVOS ESPECÍFICOS .....	26
<b>PARTE I .....</b>	<b>27</b>
<b>INTRODUCCIÓN .....</b>	<b>28</b>
1. TECNOLOGÍA DE SECUENCIACIÓN MASIVA DEL ARN (ARN-SEQ) .....	29
2. ANÁLISIS DE DATOS DE ARN-SEQ .....	31
3. DESCUBRIMIENTO DE GENES.....	34
<b>MATERIALES Y MÉTODOS .....</b>	<b>36</b>
1. MANTENIMIENTO DE LAS COLONIAS DE <i>R. PROLIXUS</i> . .....	37
2. TOMA DE MUESTRAS.....	37
3. EXTRACCIÓN DE ARN.....	38
4. SECUENCIACIÓN MASIVA.....	38
5. ANÁLISIS DE CALIDAD Y LIMPIEZA DE LAS <i>READS</i> .....	38
6. MAPEO DE LAS <i>READS</i> FRENTE A LA REFERENCIA.....	39
7. RECONSTRUCCIÓN DE LOS TRANSCRIPTOS.....	39
8. ANOTACIÓN FUNCIONAL .....	39
9. BÚSQUEDA POR SIMILITUD DE SECUENCIA E IDENTIFICACIÓN DE TRANSCRIPTOS. ....	40
<b>RESULTADOS.....</b>	<b>41</b>
1. TRANSCRIPTOMAS EMBRIONARIOS DE <i>RHODNIUS PROLIXUS</i> .....	42
1.1 <i>Análisis de calidad y limpieza de los transcriptomas embrionarios de R. prolixus</i> .....	42
1.2 <i>Alineamiento de reads contra el genoma anotado como referencia y comparación de métodos</i> .....	45
1.3 <i>Reconstrucción de los transcriptomas a partir del mapeo contra el genoma anotado de R. prolixus</i> ....	49
1.4 <i>Alineamiento de reads utilizando al genoma como referencia</i> .....	51
1.5 <i>Reconstrucción de los transcriptomas a partir del mapeo contra el Genoma de R. prolixus</i> .....	54
2. TRANSCRIPTOMAS DE OVARIO, HUEVOS NO FERTILIZADOS Y DE 0 HORAS POST OVOPOSICIÓN DE <i>RHODNIUS PROLIXUS</i> .....	54
2.1 <i>Análisis de calidad y limpieza de los transcriptomas generados mediante secuenciación paired-end</i> ... 54	
2.2 <i>Alineamiento de las reads al genoma anotado como referencia</i> .....	59
2.3 <i>Reconstrucción de los transcriptomas paired-end</i> .....	62
3. ANOTACIÓN FUNCIONAL DE TRANSCRIPTOMAS COMPLETOS .....	63
4. GENES DE DESARROLLO IDENTIFICADOS A PARTIR DE BÚSQUEDA POR SIMILITUD EN LOS TRANSCRIPTOMAS GENERADOS .....	71
<b>DISCUSIÓN.....</b>	<b>82</b>
1. DATOS CRUDOS DE SECUENCIACIÓN.....	83
2. ALINEAMIENTO DE LECTURAS DE ARN-SEQ A LA REFERENCIA Y ENSAMBLE .....	85
3. IDENTIFICACIÓN DE TRANSCRIPTOS.....	87

<b>PARTE II .....</b>	<b>92</b>
<b>INTRODUCCIÓN .....</b>	<b>93</b>
1. OOGÉNESIS EN INSECTOS .....	94
2. OOGENESIS EN <i>RHODNIUS PROLIXUS</i> .....	99
<b>MATERIALES Y MÉTODOS .....</b>	<b>102</b>
1. MANTENIMIENTO DE LAS COLONIAS DE <i>RHODNIUS PROLIXUS</i> .....	103
2. VALIDACIÓN POR LA TÉCNICA DE REACCIÓN EN CADENA DE LA POLIMERASA (PCR) .....	103
2.1 <i>Síntesis de ADN copia (ADNc)</i> .....	103
2.2 <i>Diseño de oligonucleótidos específicos</i> .....	103
2.3 <i>Amplificaciones mediante la reacción en cadena de la polimerasa con transcriptasa reversa (RT-PCR)</i> .....	104
3. HIBRIDACIÓN <i>IN SITU</i> .....	105
3.1 <i>Síntesis de las sondas de ARN</i> .....	105
3.2 <i>Hibridación</i> .....	106
4. INMUNOHISTOQUÍMICA CON FLUORESCENCIA .....	107
5. ARN DE INTERFERENCIA (ARNi) PARENTAL .....	108
6. MANIPULACIÓN DE OVARIOLAS .....	109
6.1 <i>Tinción con 4',6-diamino-2-fenilindol (DAPI) y faloindina</i> .....	110
6.2 <i>Análisis de secciones histológicas</i> .....	110
6.3 <i>Microscopía de transmisión (TEM)</i> .....	110
7. MANIPULACIÓN DE EMBRIONES .....	111
8. MANIPULACIÓN DE HUEVOS .....	111
8.1 <i>Microscopía electrónica de barrido</i> .....	111
9. ANÁLISIS DE DISTRIBUCIÓN DE LÍPIDOS Y MEMBRANAS .....	111
10. ANÁLISIS DEL TRASPASO DE NUTRIENTES Y MOLÉCULAS DURANTE LA VITELOGÉNESIS .....	112
<b>RESULTADOS.....</b>	<b>113</b>
1. DESCRIPCIÓN DE LA ESTRUCTURA Y MORFOLOGÍA DEL OVARIO CON MICROSCOPIA DE FLUORESCENCIA.....	114
1.1 <i>Re-descripción de la ovariola de <i>Rhodnius prolixus</i></i> .....	114
1.2 <i>Acumulación de vitelogenina/vitelina durante la vitelogénesis en la ovariola de <i>R. prolixus</i></i> .....	119
2. VALIDACIÓN DE LOS GENES IMPLICADOS EN LA OOGÉNESIS Y DESARROLLO TEMPRANO PREVIAMENTE IDENTIFICADOS.....	122
2.1 <i>Bicaudal C</i> .....	123
2.1.1 <i>Validación funcional por ARNi parental</i> .....	123
2.1.2 <i>Expresión génica de Bicaudal C en <i>R. prolixus</i></i> .....	136
2.2 <i>Cornichon</i> .....	137
2.2.1 <i>Validación funcional por ARNi parental</i> .....	137
2.2.2 <i>Expresión génica de cornichon en <i>R. prolixus</i></i> .....	141
2.3 <i>Bicaudal D</i> .....	142
2.3.1 <i>Validación funcional por ARNi parental</i> .....	142
2.3.2 <i>Expresión génica de Bicaudal D en <i>R. prolixus</i></i> .....	145
<b>DISCUSIÓN .....</b>	<b>148</b>
1. GENÉTICA DEL DESARROLLO EMBRIONARIO TEMPRANO EN <i>RHODNIUS PROLIXUS</i> .....	149
<b>CONCLUSIONES FINALES .....</b>	<b>157</b>
<b>SUPLEMENTARIO .....</b>	<b>160</b>
1. SUPLEMENTARIO 1 .....	161
2. SUPLEMENTARIO 2.....	161
3. SUPLEMENTARIO 3 .....	162

**BIBLIOGRAFÍA .....170**

# Abreviaturas

°C: grados Celsius

μ: micras

μl: microlitros

μm: micrómetros

aa: aminoácidos

ADN: ácido desoxirribonucleico

ADNc: ADN copia

ARN: ácido ribonucleico

ARNdc: ARN doble cadena

ARNi: ARN de interferencia

ARNm: ARN mensajeros

ARNr: ARN ribosomal

ARN-seq: secuenciación masiva del ARN

ARNt: ARN de transferencia

BG: Banda germinal

BrEt: Bromuro de etidio

EST: *Expressed sequence tags*

Fw: Forward

*hpo*: horas post ovipuesta

hs: horas

IgG: Inmunoglobulina G

kDa: Kilodaltons

Kv: Kilovoltios

min: minutos

ml: mililitros.

mM: miliMolar

Nº: Número

ng: nanogramos

NGS: secuenciación de próxima generación

nr: no redundante

Oligo: Oligodeoxiribonucleótido

ON: overnight

pb: pares de bases

PH: Logaritmo negativo de la concentración de hidrogeniones por litro

rpm: revoluciones por minuto

RT-PCR: Reacción en cadena de la polimerasa con transcriptasa reversa

Rv: Reverse

seg: segundos

sn: solución

T°a: Temperatura de hibridación

μM: microMolar

# **Introducción General**

## 1. Embriología de insectos

El plan corporal de los insectos se encuentra altamente conservado, siguiendo un patrón compuesto por una cabeza que consta de 6-7 segmentos, un tórax con 3 segmentos y un abdomen constituido por 8-11 segmentos. Pero este nivel de conservación morfológica esconde una gran diversidad subyacente en el desarrollo embrionario que da origen a este plan corporal (Davis and Patel, 2002; Krause, 1939).

En un comienzo, la embriogénesis se desarrolla de manera muy similar en todos los insectos hasta alcanzar el estadio de blastodermo. En primer lugar, se establecen repetidas divisiones nucleares, pero sin división citoplasmática, formándose un gran sincicio; luego estos núcleos migran hacia la periferia y dan lugar a la formación del blastodermo, el cual pasa de ser sincicial a celularizado, a medida que avanza el desarrollo. Más allá del estadio de blastodermo, la embriogénesis entre los insectos varía de forma sustancial respecto a cómo y cuándo el patrón corporal segmentado se modela.

Para describir las diferencias al respecto de cuándo se modela el patrón, se definieron tres grandes grupos de insectos: con embriogénesis de banda germinal (BG) corta, intermedia y larga (Davis and Patel, 2002; Krause, 1939). Esta denominación se basa en el número de segmentos especificados en el rudimento embrionario antes de la gastrulación (Krause, 1939; Tautz et al., 1994).

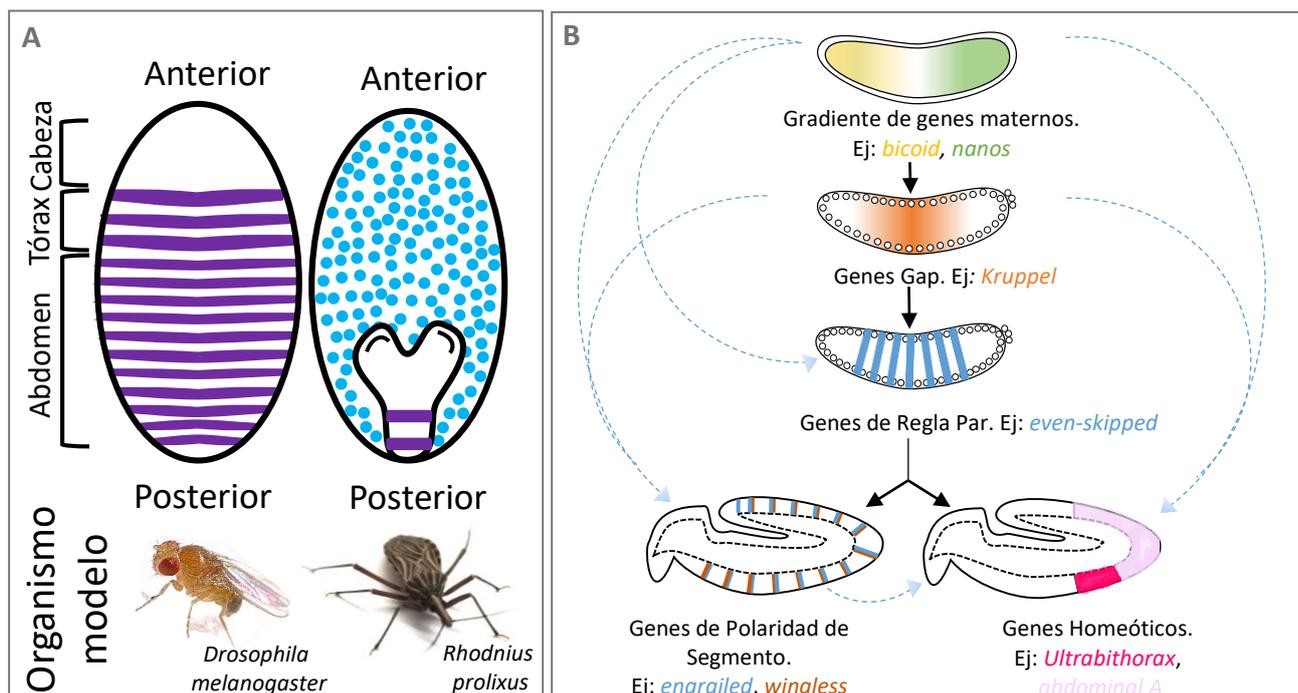
El tipo de desarrollo embrionario de los insectos de banda germinal larga es el más estudiado debido a que se encuentra representado en *Drosophila melanogaster*, donde el desarrollo sólo involucra procesos de subdivisión del blastodermo en segmentos individuales para producir el cuerpo entero. Es decir, que todos los segmentos del cuerpo son especificados simultáneamente en el estadio de blastodermo, y los tres tagmas (cabeza, tórax y abdomen) tienen proporciones muy similares tanto en la larva como en el adulto (Liu and Kaufman, 2005). En los insectos de banda germinal corta (corta/intermedia), sólo los segmentos anteriores son especificados en el blastodermo (el número varía según el tipo de banda germinal), mientras que los restantes son formados de manera progresiva durante una segunda fase de crecimiento de la banda germinal, luego de que la gastrulación tenga lugar. A partir de una región posterior del embrión, se producen divisiones mitóticas y reorganización celular para establecer el resto de los segmentos de manera secuencial y

progresiva, desde el extremo anterior al posterior del embrión (Davis and Patel, 2002; Sander, 1981). La única diferencia entre las especies de banda germinal intermedia y los de banda germinal corta reside en el número de segmentos anteriores que son especificados durante el estadio de blastodermo, por lo que por practicidad hoy en día se los denomina a ambos como banda germinal corta (Figura 1A).

El desarrollo embrionario del tipo de banda germinal larga se considera evolutivamente derivado, ya que este tipo de embriogénesis está restringida a los insectos holometábolos más derivados, como dípteros o himenópteros; en cambio el tipo de banda germinal corta e intermedia se encuentra altamente representado en los órdenes más basales de insectos (Liu and Kaufman, 2005). Paradójicamente, los procesos de desarrollo, incluido la segmentación, han sido mayormente estudiados en organismos de banda germinal larga, donde los segmentos del embrión se especifican en un entorno sincicial, permitiendo una rápida y fácil difusión de los distintos factores, estableciéndose gradientes que determinan la identidad de cada región en el embrión (Rivera-Pomar and Jackle, 1996). En cambio, en los insectos de banda germinal corta, la especificación se da principalmente por las interacciones de los factores con proteínas ancladas en la membrana que activan vías de señalización. De esta forma como producto de la expresión génica diferencial en cada segmento en formación, cada uno obtiene una identidad corporal determinada (Davis and Patel, 2002).

Como se mencionó previamente, en *D. melanogaster*, el desarrollo embrionario temprano ha sido estudiado de manera minuciosa, lo que permitió dilucidar la red génica que regula y dirige cada una de las etapas del proceso. La maquinaria molecular que gobierna la segmentación se basa en una estructura jerárquica de genes que conforman un circuito de activación-represión (Lawrence, 1992; Nusslein-Volhard et al., 1987; Nusslein-Volhard et al., 1980 ). Los patrones espacio-temporales de expresión génica que se ejecutan durante la segmentación fueron establecidos en este organismo modelo, permitiendo identificar distintas clases génicas que actúan a medida que avanza el desarrollo (Rivera-Pomar and Jackle, 1996) (Figura 1B). El primer control comienza con factores génicos que son depositados por la madre en el oocito durante la oogénesis, llamados genes maternos; los cuales establecen gradientes a lo largo del embrión que determinan la polaridad básica del mismo. Genes identificados en esta categoría son *bicoid* (*bcd*), *caudal* (*cad*), *nanos* (*nos*), *torso* (*tor*) (Driever and Nusslein-Volhard, 1988a; Driever and Nusslein-Volhard, 1988b; Nusslein-Volhard

et al., 1987). Un nivel por debajo de los genes maternos, y regulados por estos mismos, se encuentran los genes *gap*, los cuales presentan dominios de expresión solapados y subdividen al embrión en grandes áreas de expresión génica. Algunos ejemplos de esta clase de genes son: *giant (gt)*, *knirps (kni)*, *Krüppel (Kr)* (Arnosti et al., 1996; Gaul and Jackle, 1987; Gaul et al., 1987; Hülskamp and Tautz, 1991; Mohler et al., 1989; Pankratz et al., 1989). Estos últimos controlan a los de su misma clase y a las subsiguientes. En el tercer lugar de esta jerarquía encontramos a los genes de regla par, los cuales dividen al embrión en unidades periódicas de expresión. Algunos ejemplos de estos genes que establecen la periodicidad característica son *even-skipped (eve)*, *paired (prd)* (Clark, 2017; Frasch et al., 1987; Macdonald et al., 1986; Nusslein-Volhard and Wieschaus, 1980; Small et al., 1992). La transcripción de los diferentes genes de regla par resulta en un patrón de franjas de siete bandas verticales perpendiculares al eje anteroposterior. Las proteínas obtenidas a partir de estos genes activan los genes de polaridad de segmento, cuyos productos de ARN mensajeros y proteínas dividen al embrión en catorce unidades segmentarias extensas, estableciendo la periodicidad del embrión. Los productos proteicos de los genes *gap*, de regla par y polaridad de segmento interactúan para regular los genes selectores homeóticos, cuya transcripción determina el destino de desarrollo de cada segmento (Levine and Harding, 1989; Martínez-Arias and Lawrence, 1985). Esta red regulatoria subdivide el eje anteroposterior en 14 segmentos corporales: 3 segmentos en la cabeza, 3 segmentos torácicos y 8 segmentos abdominales (Lawrence, 1992).



**Figura 1:** Embriología de insectos. En la parte **A** se representan los tipos de embriogénesis correspondiente a banda germinal (BG) larga y BG corta, donde se visualizan la cantidad de segmentos especificados en estadio de blastodermo. Cada tipo de banda germinal está ejemplificado con un organismo modelo que la representa. En la parte **B** se grafica el esquema correspondiente a la red génica que regula la segmentación en *D. melanogaster*, donde se observa el patrón de expresión de cada tipo de clase génica. Las flechas, tanto las líneas completas como líneas de puntos, representan las relaciones regulatorias entre ellos. En la parte **A** la imagen de *D. melanogaster* fue tomada de Biopix.dk Director: Dag Westby. En la parte **B** la representación gráfica fue adaptada a partir de Sanson (2001)

Con el correr de los años, el advenimiento de estudios de desarrollo comparativo, permitió acortar la brecha que diferencia al tipo de banda germinal larga y corta. Pese a que las primeras etapas del desarrollo varían sustancialmente entre especies, se observó que la red génica que regula el proceso de segmentación, descrita en detalle para *D. melanogaster*, se conserva incluso entre especies filogenéticamente alejadas (Clark, 2017; Clark and Peel, 2017; Damen et al., 2005; Farzana and Brown, 2008; Janssen and Budd, 2013).

Por otra parte, fue posible determinar que los insectos de banda germinal corta y larga comparten la característica común de localizar de manera específica ARN mensajeros (ARNm), de origen materno, a lo largo del embrión, para dar lugar al comienzo de su desarrollo. En el primer organismo donde fue identificado este fenómeno fue en *D. melanogaster*, donde se reportó al ARNm del gen *bicoid* (*bcd*) secuestrado en la región anterior del oocito durante la ovogénesis (Ali-Murthy and Kornberg, 2016; Driever and Nusslein-Volhard, 1988a; Driever and

Nusslein-Volhard, 1988b; Frigerio et al., 1986). Fuera del orden diptera, en *Nasonia vitripennis* (orden Hymenoptera) se reportó la localización del ARNm *orthodenticle (otd)* en ambos polos, anterior y posterior, del embrión. En ambos organismos, estos ARNm maternos específicos se traducen y generan gradientes proteicos a lo largo del embrión que desempeñan un rol en el establecimiento del patrón corporal a lo largo del eje anteroposterior (Lynch et al., 2006). En los insectos de banda germinal corta, como *Tribolium castaneum*, existe esta localización específica de ARNm maternos, reportado para los genes *pangolin (pan)* y *eagle (eg)* durante el desarrollo embrionario (Bucher et al., 2005), pero hasta la fecha no se logró evidenciar una correlación entre localización de ARNm y establecimiento del patrón corporal (Brent et al., 2007).

Recientemente, se ha estudiado con mayor detalle el proceso de elongación de la banda germinal en *T. castaneum* observándose una gran similitud con lo determinado para *D. melanogaster*. Los eventos morfogénéticos que ocurren durante la extensión de la banda germinal en *D. melanogaster* están dirigidos por la intercalación mediolateral de las células ectodérmicas. En *T. castaneum* se había reportado que la elongación de la banda germinal era dirigida a partir de la llamada zona de crecimiento, pero con el transcurso de los años se logró identificar a la intercalación mediolateral celular como responsable de dicho proceso. Además, en ambos organismos, la intercalación se encuentra regulada por un grupo específico de genes *Toll* (Benton, 2018). Sumado a esto, se reportó la conservación de la función del gen *caudal (cad)* y la vía de señalización *Wnt* en el proceso de elongación en el organismo de banda germinal corta *T. castaneum*, en analogía a *D. melanogaster*. (Bolognesi et al., 2008; Copf et al., 2004; McGregor et al., 2008; Nunes da Fonseca et al., 2009; Olesnick et al., 2006).

Estas investigaciones reconcilian y acercan cada vez más las diferencias observadas respecto a cómo se modela el patrón corporal. La mayoría de los primeros avances en embriología se realizaron utilizando dibujos que comparaban y detallaban las diferentes etapas del desarrollo del embrión que, para la época, significaron hitos en la embriología y sentaron las bases para las futuras investigaciones. Con la posibilidad de la observación de embriones *in vivo*, así como también la oportunidad de teñir o marcar células específicas, fue posible rastrear cómo las células que componen al embrión temprano se diferencian a medida que avanza el desarrollo y permiten el modelado del plan corporal. El mapeo del destino celular permitió comparaciones más detalladas entre embriones y llevó al descubrimiento de

la importancia de la migración celular en el desarrollo embrionario. Por otro lado, la identificación de los actores que gobiernan dicho proceso, conservado entre las diferentes especies, hacen que los próximos estudios se centren en describir cómo estas redes génicas compartidas dan lugar a diferencias morfológicas.

## 2. Estudio del desarrollo embrionario en organismos modelos no convencionales.

La mosca de la fruta, *Drosophila melanogaster* es un organismo modelo increíblemente poderoso, “saltó a la fama” a principios del siglo XX, cuando el genetista T. H. Morgan comenzó a utilizarla como una herramienta de estudio para comprender la naturaleza de los genes y de las mutaciones génicas. Durante la mayor parte del siglo, *D. melanogaster* fue uno de los más importantes organismos de estudio en el laboratorio para llevar a cabo investigaciones de todo tipo, abarcando desde la identificación de la función de un gen hasta el entendimiento del comportamiento. Esta preferencia se debió, principalmente, a su fácil manejo y mantenimiento en el laboratorio. Ante este panorama, no resultó sorprendente que fuese elegido por Nüsslein-Volhard y sus colegas para desarrollar el primer *screening* genético para identificar los genes que participan durante el desarrollo embrionario (Nusslein-Volhard and Wieschaus, 1980; Nusslein-Volhard et al., 1984; Wanninger, 2015).

A partir de este avance, se convirtió en el organismo a elección para estudiar la genética del desarrollo en general, proporcionando una detallada comprensión de los mecanismos moleculares que modelan la mayoría de las características del patrón corporal de insectos, facilitando una base sólida para llevar a cabo estudios comparativos del desarrollo en insectos. Ante esto, se hizo evidente que muchos aspectos del desarrollo temprano de *D. melanogaster* se diferencian en lo que respecta al desarrollo de la mayoría de los insectos más basales en la filogenia, y a pesar de sus ventajas como organismo modelo, resultó ser una especie limitante para la comprensión del desarrollo embrionario temprano en otros insectos.

La posibilidad de conocer cuáles son los genes que conforman a los genomas de las diferentes especies hizo posible realizar la comparación entre estos de manera tal de lograr comprender la variedad de procesos biológicos (Tatusov et al., 1997), facilitando estudiar la evolución de los genomas. La comparación genómica permite transferir información génica desde un organismo modelo bien estudiado a especies no convencionales (Jiang et al., 2013),

con el objetivo de identificar los genes que se expresan en la especie de interés, y así comprender en más detalle el funcionamiento de los genomas y cómo esta información es interpretada para dar origen a una morfología particular.

El desarrollo y aplicación de técnicas moleculares en especies no convencionales, como la introducción de la interferencia del ARN (ARNi) como herramienta para realizar estudios funcionales simples y rápidos (Bucher et al., 2002; Fire et al., 1998; Hunter, 1999; Martinez and Tuschl, 2004), permitió abrir las puertas para conocer cómo esos genes que se expresan durante el desarrollo, establecen el patrón corporal en especies no estudiadas hasta ese momento (Bucher et al., 2002; Lynch et al., 2010). Es por esto que el análisis comparativo ha contribuido no sólo a generar información clave para comprender la evolución de desarrollo, sino también a introducir organismos no convencionales como organismos modelo de estudio de la biología del desarrollo (Wanninger, 2015).

### 3. *Rhodnius prolixus* como organismo modelo de estudio

Hemiptera es un orden numeroso de insectos que comprende cerca de 80.000 especies conocidas. La sinapomorfía de este grupo es la forma del aparato bucal suctopocador, con las particulares modificaciones de la mandíbula y maxila, dando lugar a la formación de dos canales abiertos: uno permite el traspaso de la saliva y el otro es utilizado para succionar savia o fluidos animales (Berenbaum, 1996).

Dentro de este orden se encuentra *Rhodnius prolixus*, especie hematófaga, uno de los vectores más importantes de la enfermedad de Chagas (Coura JR and B.-P.J., 2012; Chagas, 1909; Dias et al., 2002a; Dias et al., 2002b; Garcia and Azambuja, 2004; Kollien and Schaub, 2000). Debido tanto a su interés sanitario como biológico, se lo considera como organismo modelo de estudio, elegido en comparación con los otros vectores de la enfermedad (*Triatoma* y *Panstrongylus*), principalmente, porque su ciclo de vida es más corto que los demás lo que lo convierte en un modelo de estudio flexible y práctico para el trabajo de laboratorio (Figura 2).



**Figura 2:** *Rhodnius prolixus*. Se fotografiaron todos los estadios correspondientes al ciclo de vida del hemimetábolo, empleando una regla milimétrica como referencia. En la imagen tomada de Mesquita et al. (2015) se observa al triatomino nutriéndose con sangre, hábito conocido como hematofagia.

Desde comienzos de 1900 se han llevado a cabo investigaciones de la fisiología de *R. prolixus*.

La primera colonia en laboratorio de estos ejemplares se estableció en 1930, y fue fundamental para determinar el ciclo de vida de *R. prolixus* (Buxton, 1930; Wigglesworth, 1934). Las primeras investigaciones fueron llevadas a cabo por Sir V.B Wigglesworth, quién

reportó la existencia de un control hormonal durante la ovogénesis (Wigglesworth, 1936), así como también la acumulación de proteínas extra ováricas en oocitos (Wigglesworth, 1943); abriendo paso al estudio del transporte y deposición de vitelo en los oocitos en desarrollo en otras especies de insectos. En paralelo a esto, se describió el desarrollo embrionario de *R. prolixus* (Kelly and Huebner, 1989; Mellanby, 1935; Mellanby, 1936), identificando hitos claves dentro de la embriogénesis, como la formación de la banda germinal y los movimientos morfogenéticos de la anatrepsis y katatrepsis. Se concluyó además, que la ingesta de sangre es un paso necesario y obligatorio para la producción de huevos y por consecuencia para la fertilidad (Valle et al., 1987). Asimismo, se profundizó en el estudio de esta capacidad como un posible medio de control de la plaga; por lo que le siguieron varios trabajos en el que se describió el proceso de ovogénesis, se describió la unidad fundamental del ovario, la ovariola; conociéndose así en detalle la arquitectura y morfología del ovario en *R. prolixus* (Huebner and Anderson, 1972a; Huebner and Anderson, 1972b; Lutz and Huebner, 1980; Lutz and Huebner, 1981).; así como también se caracterizó la estructura y tamaño del huevo obtenido tras puesta (Wigglesworth, 1942).

El estudio de *R. prolixus* se vio opacado con el “boom” de *D. melanogaster* como organismo modelo. Pero tiempo después, con el advenimiento y disponibilidad de las herramientas moleculares y de las tecnologías de secuenciación masiva fue posible retomar el estudio en *R. prolixus* y re direccionar el foco de investigación, y no sólo estudiar la bioquímica y fisiología del desarrollo, sino que permitió comenzar a entender los mecanismos que dirigen el desarrollo del hemíptero.

En esta nueva era, contamos con el genoma de *R. prolixus* secuenciado (Mesquita et al., 2015) y con un número creciente de manera exponencial de secuencias curadas y accesibles *online* en Vectorbase (<https://www.vectorbase.org/>); así como también, con el reporte de algunos análisis a nivel transcriptómico (Medeiros et al., 2011; Ribeiro et al., 2014), proporcionando un marco firme tanto para la identificación y anotación génica como para estudios genómicos comparativos (Nunes-da-Fonseca et al., 2017). Por otro lado, la disponibilidad de valiosas herramientas moleculares para la identificación funcional de dichos genes, como por ejemplo ARN interferente parental (Lavore et al., 2012), da lugar a un estudio en detalle de la genética del desarrollo en *R. prolixus* permitiendo identificar algunos de los

principales actores que dirigen tal proceso (Berni et al., 2014; Lavore et al., 2015; Nunes-da-Fonseca et al., 2017).

Sumado a esto, se re-describió la embriología del triatomino utilizando microscopía de fluorescencia y confocal, lo que permitió registrar con detalle los fenómenos que ocurren durante el avance del desarrollo embrionario y correlacionar tiempo con evento en el desarrollo (Berni et al., 2014; Esponda-Behrens, 2014; Pagola, 2012).

Más allá de estos avances, aún quedan por determinar elementos genéticos que dirigen el desarrollo en *R. prolixus*, no sólo para responder cuestiones biológicas fundamentales, sino también para diseñar nuevas estrategias con el objetivo de reducir la difusión de la enfermedad de Chagas.

Por lo antes mencionado, en el presente trabajo se propone identificar los genes que se expresan a lo largo de las distintas etapas que comprende el desarrollo embrionario en *R. prolixus*. Para esto, se parte del estudio desde el estadio de oocito para comprender cuáles son los actores maternos y los propios del ovario, que cumplen un rol fundamental en el establecimiento de la polaridad presente en el huevo y desencadenar la activación de los genes cigóticos. Luego se continúa el análisis en estadios más avanzados, como blastodermo, gastrulación y formación de banda germinal para conocer los genes que dirigen tales procesos.

Para lograr tal propósito se llevó a cabo la investigación por dos vías de trabajo distintas pero complementarias; y para facilitar la comprensión y seguimiento a la hora de la lectura de este trabajo se lo dividió en dos partes, cada una con una estructura interna que consiste en introducción, materiales y métodos, resultados y discusión. En la parte I se describe el análisis a nivel bioinformático, en el que se incluye la generación de los datos de los genomas expresados; la limpieza, mapeo y ensamble de los mismos; y el posterior análisis por búsqueda de similitud de secuencia e identificación de dominios proteicos. En la parte II se detalla la identificación funcional de los genes activos durante la ovogénesis y desarrollo temprano. Para esto, en primer lugar, describe la morfología del ovario con imágenes de microscopía avanzada, y luego se realiza un análisis exhaustivo de la función y el patrón de expresión de los genes previamente identificados.

# **Hipótesis y Objetivos**

## 1. Hipótesis

El desarrollo embrionario es un proceso sumamente organizado dirigido por una jerarquía temporal de acción de los genes. La activación de los mismos es el resultado de interacciones de factores moduladores específicos presentes en las distintas etapas que dirigen el desarrollo del embrión; por lo que cualquier cambio en el patrón de expresión génica afecta la morfología final del insecto.

## 2. Objetivo general

Se propone generar información sobre el genoma expresado de *Rhodnius prolixus* desde la etapa de la oogénesis hasta los primeros estadios embrionarios para poder conocer los genes que se expresan a lo largo de todo el proceso e inferir mecanismos de regulación.

### 3. Objetivos específicos

#### **PARTE I**

- Secuenciar el transcriptoma embrionario de *R. prolixus*, comprendiendo desde los estadios referentes a la expresión de genes maternos hasta el inicio de la gastrulación.
- Secuenciar el transcriptoma del ovario de hembras adultas de *R. prolixus*.
- Realizar una anotación funcional de los genes identificados en los diferentes transcriptomas generados.
- Identificar y anotar en el genoma de *R. prolixus* genes vinculados a la oogénesis y al desarrollo embrionario temprano.

#### **PARTE II**

- Realizar una re-descripción morfológica del ovario de *R. prolixus*.
- Validar funcionalmente genes identificados, vinculados a la oogénesis y al desarrollo embrionario temprano.
- Determinar el patrón de expresión de dichos genes identificados.

# PARTE I

# Introducción

## 1. Tecnología de secuenciación masiva del ARN (ARN-seq)

La tecnología de secuenciación de próxima generación (NGS) hace referencia a la identificación masiva y paralela de las bases que conforman las secuencias de nucleótidos correspondientes a un organismo; su advenimiento ha revolucionado la manera de estudiar biología (Margulies et al., 2005; Wesolowski et al., 2013).

Una de sus aplicaciones es la transcriptómica, la secuenciación del *set* completo de ARN mensajeros (ARNm) de una muestra biológica, a nivel de, célula, órgano, tejido u organismo en particular. Es una aproximación que nos permite no sólo caracterizar el conjunto de genes expresados sino también, determinar su nivel de expresión. Este conjunto se lo conoce como transcriptoma, y a diferencia del genoma, complemento genético completo de un organismo, puede ser modulado por factores internos y externos (Velculescu et al., 1997). El análisis del transcriptoma es esencial para la interpretación de los elementos funcionales de un genoma, y para revelar los constituyentes moleculares de las células, los tejidos y órganos (Wang et al., 2009).

Tradicionalmente, los proyectos de transcriptómica se basaban en la secuenciación de EST (*Expressed sequence tags*), secuencias que representan los extremos 5' y 3' de ARNm. Este tipo de secuenciación permitió obtener información del nivel de la expresión génica, identificar casos de empalme o *splicing* alternativo y la presencia de genes, pero ha sido sustituido por la tecnología de NGS, ya que esta última permite obtener una cobertura mayor del transcriptoma (Liu et al., 2013). A su vez, tiene como ventajas que no se limita a la detección de transcritos que corresponden a una secuencia genómica ya conocida; presentan muy poco ruido de fondo, si es que lo hay; no presentan un límite máximo para la cuantificación; y muestras pequeñas de ARN son suficientes para ser analizadas por esta tecnología (Wang et al., 2009).

La información generada por la secuenciación masiva del ARN (ARN-seq) puede ser dividida en dos categorías: cualitativa y cuantitativa. Los datos cualitativos permiten identificar transcritos expresados y su arquitectura génica; aplicación conocida como anotación génica. Los datos cuantitativos, además de hacer posible la identificación génica, permiten medir cuantitativamente las diferencias de expresión, de los transcritos identificados, entre dos o más grupos o tratamientos, aplicación denominada análisis de la

expresión diferencial. Estos dos tipos de enfoques están relacionados y en cierto punto son inseparables. Una buena cobertura de anotación permite medir de manera más eficaz los niveles de expresión génica, así como también una buena anotación requiere cierto grado de profundidad de secuenciación ya que la simple existencia de una o dos lecturas no es suficiente para dar evidencia de la existencia de un gen. Más allá de esta relación, resulta difícil diseñar un experimento que pueda satisfacer los dos objetivos, una de las limitantes principales es la cuestión económica (Anders and Huber, 2010; McIntyre et al., 2011; Wang et al., 2011).

Los dos métodos de secuenciación masiva más usados son la pirosecuenciación y el método de PCR en puente (Bridge PCR) (Nagalakshmi et al., 2008). Independientemente de que tecnología sea usada, en líneas generales, el proceso comienza con la extracción de ARNm de una muestra celular específica, el que es utilizado para conformar la biblioteca de secuenciación (*library*), la cual está constituida por fragmentos o moléculas completas de ácidos nucleicos (Eid et al., 2009; Metzker, 2010). En la actualidad la mayoría de las plataformas utilizan la estrategia de fragmentación por lo que no es posible obtener la información de secuencias en la forma de transcritos completos (Fedurco et al., 2006).

El armado de la biblioteca es un paso clave a la hora de secuenciar ya que es el material de entrada, directo, que recibe el secuenciador, por lo que cada tecnología estableció protocolos plataforma/objetivo-específicos. En general, el proceso de preparación de la biblioteca se caracteriza por generar ADN copia (ADNc) simple cadena, incorporar adaptadores y amplificar el ADN generado. Una vez obtenida la biblioteca, un subconjunto de ella es utilizado como *input*, y tras el proceso en sí de secuenciación da lugar a la generación de las *reads* de secuenciación, lecturas cortas de ADN crudo. Existen dos alternativas a la hora de la obtención de las *reads*, realizar la secuenciación desde un único extremo, conocido como secuenciación *single-end*, o desde ambos, llamado *paired-end* (Wesolowski et al., 2013).

La reconstrucción de un transcripto completo a partir de las lecturas cortas generadas es un desafío técnico como consecuencia de la variabilidad técnica generada tras la preparación de la *library* y a la tasa de error propia del secuenciador, a la hora de darle identidad a cada nucleótido. Así como también debido a la variabilidad propia de la muestra, dada por la expresión génica diferencial, variabilidad de longitud de los transcritos, presencia de secuencias repetitivas, otros (Metzker, 2010; Wesolowski et al., 2013). Todas estas

dificultades deben ser superadas por un método eficiente, capaz de reconstruir transcriptos de buena calidad, variables en tamaño, nivel de expresión e isoformas (Haas and Zody, 2010). La reconstrucción de un transcriptoma puede realizarse por mapeo de las *reads* a un genoma o transcriptoma como referencia, o bien mediante un ensamble *de novo* (Conesa et al., 2016). Una vez obtenidos los datos transcriptómicos, se lleva a cabo el análisis dependiendo del diseño experimental planteado originalmente.

A pesar del requerimiento y desafío computacional para analizar los datos de ARN-seq, se han superado varios de estos obstáculos y se han abordado con éxito diferentes problemas biológicos como la cuantificación de *splicing* alternativo en tejidos, poblaciones y enfermedades; el descubrimiento de nuevos genes ligados al cáncer; el perfeccionamiento en los ensambles de los genomas e identificación de nuevos transcriptos (Berger et al., 2010; Blekhman et al., 2010; Denoeud et al., 2008; Guttman et al., 2010; Maher et al., 2009; Mortazavi et al., 2010; Trapnell et al., 2010; Wang et al., 2008; Wilhelm et al., 2011; Yassour et al., 2009).

La capacidad y posibilidad de secuenciar transcriptomas, así como también genomas, de forma masiva y rutinaria abre nuevas posibilidades en la investigación biológica y en aplicaciones biomédicas.

## 2. Análisis de datos de ARN-seq

Uno de los objetivos principales del análisis de los datos de secuenciación es reconstruir el conjunto completo de transcriptos correspondiente a los genes que se expresan en la célula de partida (Kim et al., 2013). Contando con un genoma o transcriptoma disponible, uno de los pasos críticos en el análisis de los datos de secuenciación es el mapeo de las *reads* a la referencia, el que proporciona la ubicación desde donde se originaron las lecturas (Kim et al., 2013).

Se desarrollaron varios métodos computacionales que permiten abordar los principales desafíos del análisis de ARN-seq, como la generación de lecturas cortas por el secuenciador (36-150 bases), tasas de error considerables, las lecturas que mapean en uniones exón-exón y el número creciente de lecturas obtenidas por experimento (Garber et al., 2011).

Hay dos enfoques algorítmicos principales para mapear las lecturas de secuenciación a una referencia. El primero, al que se lo conoce como alineadores de lecturas no empalmados o “*unspliced read aligners*”, alinea las *reads* a la referencia sin permitir grandes espacios entre las *reads* mapeadas. Se diferencian dos tipos de métodos: “*Seed methods*” (Homer et al., 2009; Smith et al., 2008) y “*Burrows-Wheeler transform methods*” (Smith et al., 2008), el primero se basa en encontrar coincidencias en pequeñas sub-secuencias, llamadas “*seeds*” (semilla), asumiendo que al menos una *seed* de una lectura mapea perfectamente con la referencia. Cada *seed* es utilizada como una región candidata estrecha donde métodos más sensibles (como *Smith-Waterman*) pueden ser utilizados para extender esa pequeña región alineada a un alineamiento completo; ejemplos de estos son *MAQ* (Li et al., 2008) y *Stampy* (Lunter and Goodson, 2011). En contraste, los métodos *Burrows-Wheeler*, como *Burrows-Wheeler alignment* (BWA) (Li and Durbin, 2009) y *Bowtie* (Langmead et al., 2009) compactan el genoma en datos estructurales, lo que resulta ser muy eficiente para llevar a cabo búsquedas con coincidencias exactas. Si se permiten alineamientos con alguna no coincidencia la eficiencia del método disminuye exponencialmente con el número de no coincidencias permitidas (Kim et al., 2013).

Estos tipos de programas son ideales para llevar a cabo mapeos contra un transcriptoma como referencia; si el transcriptoma es el propio de la especie que se está analizando, los métodos *Burrows-Wheeler* son más eficientes; ahora si se tiene solamente como referencia el transcriptoma de una especie cercana evolutivamente, el método de *seed* puede dar buenos resultados si se incrementen los parámetros de alineamiento para que este sea más sensible. La búsqueda de *unspliced read aligners* está limitada a identificar exones y uniones de exones conocidas, no permite la identificación de nuevos eventos de empalmes o exones.

Alternativamente, se encuentran los métodos “*spliced aligners*”, que permiten alinear de manera adecuada las lecturas de secuenciación a una referencia genómica, sin ser una dificultad la presencia de grandes regiones intrónicas para realizar el alineamiento. Estos métodos también se diferencian en dos categorías; “*exon first*” y “*seed and extend*”, el primero es característico de los softwares *MapSplice*, *SpliceMap* y *TopHat* (Au et al., 2010; Trapnell et al., 2009; Wang et al., 2010), y lleva a cabo el mapeo en dos pasos, en primer lugar, mapea las lecturas, al genoma utilizando “*unspliced read aligners*”, mapeos continuos. Luego,

las *reads* que no mapearon son fragmentadas en segmentos más pequeños y posteriormente alineados al genoma de manera independientes; esto permite identificar regiones que contengan sitios de empalmes que se encuentran rodeando a los segmentos de las *reads* que si mapearon. El segundo método *seed and extend*, primero fragmenta las lecturas en pequeñas *seed*, las cuales son ubicadas como guía en el genoma para localizar el alineamiento completo. Las regiones candidatas son luego analizadas con métodos más sensibles con el objetivo de determinar el sitio de empalme exacto presente en la lectura mapeada; este método es característico de los programas *GSNAP* y *QPALMA* (De Bona et al., 2008; Kent, 2002; Wu and Nacu, 2010). Este tipo de enfoque permite la identificación de nuevos transcritos y nuevos sitios de empalme alternativo (Trapnell et al., 2009).

Sea cual fuese el algoritmo utilizado, el alineamiento obtenido puede ser utilizado subsecuentemente para distintos tipos de análisis. Una de las vías más adoptadas es el ensamble de las “*reads*” mapeadas en unidades de transcripción o simplemente llamados transcritos, colectivamente este proceso se lo denomina reconstrucción del transcriptoma. Contando con un genoma o transcriptoma de referencia, uno de los programas más empleados para realizar la reconstrucción es *Cufflinks* (Garber et al., 2011; Trapnell et al., 2012); su metodología se basa en ensamblar transcritos individuales a partir de las lecturas que mapearon con la referencia. A su vez permite estimar la cantidad de *reads* mapeadas que dieron origen a dicho transcritos; esta es una aproximación cuantitativa cuando se cuenta con una única muestra secuenciada (Trapnell et al., 2010). En los casos donde existen réplicas biológicas para realizar análisis de expresión génica se utilizan otros softwares informáticos como por ejemplo *Cuffdiff*, perteneciente al programa *Cufflinks* (Trapnell et al., 2012).

### 3. Descubrimiento de genes

El genoma es la totalidad de la información genética que posee un organismo o una especie en particular; la genómica como tal, involucra el estudio del total de la información contenida en el ADN.

Una de las disciplinas que proporciona una vía para lograr conocer la genética de un organismo del que se desconocen los actores genéticos que lo conforman es la genómica comparada, que permite comparar las secuencias del complemento génico de distintas especies (Hardison, 2003; Jiang et al., 2013). Se basa en que las características comunes de dos organismos se codifican dentro de una región genómica que se conserva entre las especies; así como también, las secuencias de ADN que regulan la expresión génica (Hardison, 2003).

Uno de los programas bioinformáticos que permite llevar a cabo dicha tarea es BLAST, mediante un alineamiento local de secuencias. En base a un enfoque heurístico se determina si dichos genes se encuentran representados en ambas especies por similitud de secuencia (Altschul et al., 1990).

La similitud de secuencia es un resultado que proviene directamente de la observación. La homología de secuencia que es una interpretación basada en que una alta similitud entre dos secuencias se debe a que las mismas derivan de una misma secuencia ancestral (Pellegrini et al., 1999). Por homología de secuencia se entiende a la situación en la que las secuencias de dos o más proteínas o ácidos nucleicos son similares entre sí debido a que presentan un mismo origen evolutivo (Pellegrini et al., 1999).

Utilizando este enfoque comparativo fue posible conocer la genética del desarrollo embrionario de organismos modelos de estudio de los cuales se sabía relativamente poco hasta ese entonces, observándose que los componentes genéticos se conservan entre especies, incluso a las pertenecientes a distintos órdenes evolutivos (Behura et al., 2011; Dearden et al., 2006; Ewen-Campen et al., 2011; Harker et al., 2013; Lavore et al., 2015; Shigenobu et al., 2010). Para llevar a cabo dicho análisis, el organismo modelo por elección fue y es *D. melanogaster*, como resultado de los grandes avances que existen en la comprensión de la genética del desarrollo (Behura et al., 2011).

La información generada a partir de la similitud de secuencia entre genomas es un recurso, hoy en día, indispensable, para reconocer regiones funcionales y establecer las bases para predecir cuáles son las funciones correspondientes a tales inferencias (Hardison, 2003).

# **Materiales y Métodos**

## 1. Mantenimiento de las colonias de *R. prolixus*.

Los ejemplares machos y hembras de *R. prolixus* utilizados durante el transcurso de este trabajo fueron obtenidos a partir de una colonia establecida en el Centro de Bioinvestigaciones (UNNOBA-CIC). Son criados en recipientes de plástico de boca ancha y paredes lisas, cuyas tapas son reemplazadas por una malla de nylon muy fina que facilita su alimentación.

Los insectos se criaron a una temperatura de  $28 \pm 1^\circ\text{C}$ , una humedad relativa de 70%, y con fotoperiodo regulado (luz/oscuridad de 12:12 horas). Los ejemplares fueron alimentados *ad libitum*, con aves de corral, de manera regular cada 15 días. En estas condiciones la duración de la embriogénesis es de  $14 \pm 1$  días.

## 2. Toma de muestras.

Se colectaron los huevos a las 0, 12, 24 y 48 horas post ovipuesta (*hpo*), para que representen los estadios de interés; así como también huevos ovipuestos por hembras vírgenes (huevos no fertilizados). Por otro lado, se disectaron ovarios de hembras en plena oviposición. Para la oviposición es requerimiento *sine qua non* la alimentación. Una vez alimentados demoran aproximadamente una semana para iniciar la puesta.

Para la colecta de huevos, se vaciaron de huevos los frascos con insectos y se dejaron durante el tiempo de incubación estipulado. Al cabo de ese tiempo se colectaron los huevos depositados, representado estadísticamente una cohorte que continuó el desarrollo hasta el tiempo indicado en cada caso para permitir una sincronización de los huevos dentro de  $\pm 3$  horas. Se colectaron aproximadamente 200 huevos por tiempo secuenciado.

Para la disección de ovarios, se identificaron las hembras, se colocaron sobre placas de disección y se llevó a cabo la extracción manual de los mismos, eliminando todo aquello extra a la estructura tubular.

Hasta su procesamiento, las muestras fueron conservadas en tubos *Eppendorf* de tapa a rosca en nitrógeno líquido ( $-198,5^\circ\text{C}$ ).

### 3. Extracción de ARN.

La extracción de ARN se llevó a cabo siguiendo el protocolo del tiocianato de guanidinio-fenol usando el reactivo comercial TRIZOL. El mismo fue modificado de manera que se mantuvo el ARN a 4° C durante todo el procedimiento y todos los pasos de centrifugación se llevaron a cabo con la máxima velocidad de la centrífuga (13800 rpm) (*HealForce*). La precipitación en isopropanol se dejó ON (*overnight*) a -20° C.

El ARN obtenido se resuspendió en 20 µl. Para determinar la calidad se llevó a cabo una corrida electroforética, se sembró 1 µl de la muestra en un gel de agarosa 1% con 2 µl del agente intercalante, Bromuro de etidio (EtBr). Se sembró utilizando formamida para evitar contaminación o degradación del ácido nucleico, así como también para ayudar a que decante la muestra en la fosa del gel. Se realizó una corrida electroforética a 80V durante 40 min. Las muestras de ARN fueron almacenadas a -80°C hasta su secuenciación.

### 4. Secuenciación masiva.

La genoteca fue preparada en el servicio de secuenciación del *Transcriptome Analysis Lab* de la Universidad de Göttingen, fragmentando el ARN y sintetizando ADNc a partir de los fragmentos generados. Las muestras fueron secuenciadas utilizando la plataforma HiSeq-3000 de Illumina; el cual se caracteriza por la síntesis directa mediante PCR en puente (Bridge-PCR). Se establecieron dos corridas diferentes, por un lado, se secuenciaron los transcriptomas embrionarios en una corrida *single-end* (secuenciación de los fragmentos desde un extremo) y con lecturas de una longitud de 50 pares de bases (pb). Por otro, para los transcriptomas de ovario, huevos no fertilizados y el segundo de transcriptoma de tiempo *0hpo* se llevó a cabo una corrida *paired-end* (secuenciación de los fragmentos desde ambos extremos) y con lecturas de 150 pb.

### 5. Análisis de calidad y limpieza de las *reads*

Las lecturas crudas fueron pre-procesadas con el programa *FastQC toolkit* ([www.bioinformatics.babraham.ac.uk/projects/fastqc/](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)), se consideró como métrica de buena calidad valores iguales o superiores a un puntaje de calidad (Q) 30, el cual se determina a partir

de la predicción de la probabilidad de la incorporación de una base incorrecta. Para la eliminación de secuencias adaptadoras y secuencias de baja calidad, se utilizó *Fastx-toolkit* ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). La búsqueda de secuencias contaminantes, artefactos propios de la plataforma de secuenciación, se realizó mediante una búsqueda de similitud de secuencias entre las lecturas obtenidas y la base de datos de UniVec, compuesta por 5.456 secuencias génicas correspondientes a artefactos específicos de clonación y secuenciación, presente en NCBI (<ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec>), utilizando la herramienta BLAST (Altschul et al., 1990). Las lecturas que dieron *hit* utilizando un *e-value* menor o igual a  $1 \times 10^{-5}$  fueron eliminados del conjunto de datos por considerarse secuencias contaminantes utilizando un *script* codificado en el lenguaje *Python* (Suplementario 1).

## 6. Mapeo de las *reads* frente a la referencia

Las lecturas obtenidas tras el procesamiento de calidad fueron utilizadas para ser alineadas tanto al genoma de *R. prolixus* (versión RproC3) como a sus transcritos predichos *ab initio* (genoma anotado), ambos disponibles *online* en VectorBase (<https://www.vectorbase.org/>). Para el mapeo contra el genoma se utilizó el programa *Tophat2* (Kim et al., 2013) y en el caso del mapeo contra el genoma anotado como referencia se utilizaron *Tophat2* y *Bowtie1* (Kim et al., 2013; Langmead et al., 2009), con el fin de comparar los resultados. Para ambos programas se establecieron los parámetros de manera tal que reporten todos los alineamientos que se producen durante el proceso de mapeo. Se manipuló a los archivos de mapeos con la herramienta SAMTOOLS (Li et al., 2009), así como también se evaluó la calidad de los archivos de mapeo utilizando la plataforma *Qualimap* (Garcia-Alcalde et al., 2012).

## 7. Reconstrucción de los transcritos

A partir del archivo generado por los programas de mapeo se llevó a cabo el ensamble de las lecturas en transcritos utilizando el programa *Cufflinks* (Trapnell et al., 2012). Así como también se estimó la abundancia de los transcritos generados.

## 8. Anotación funcional

Se llevó a cabo una anotación de los genes utilizando los términos GO (provenientes del proyecto de ontología génica), el cual provee un vocabulario controlado para describir el gen

y los atributos del producto génico. El programa utilizado para tal propósito es el *Blast2GO* (Conesa et al., 2016), el cual realiza como paso previo a la anotación una búsqueda de similitud de secuencia contra la base de datos no redundante (nr) del NCBI (Altschul et al., 1990).

#### 9. Búsqueda por similitud de secuencia e identificación de transcritos.

A partir de una búsqueda *in silico* de los genes que se expresan durante el desarrollo embrionario y la oogénesis en *D. melanogaster* reportados en Flybase (<http://flybase.org/>) se generó una base de datos proteica (1.364 secuencias) que fue utilizada para realizar una búsqueda de similitud de secuencia con los distintos transcriptomas generados. Se llevó a cabo utilizando la herramienta BLASTX del NCBI (Altschul et al., 1990), estableciendo como parámetros de búsqueda un *e-value* de  $1 \times 10^{-5}$  y mostrar el mejor *hit* en formato tabular. Se corroboró la identidad de los transcritos de interés mediante *InterProScan* y *Clustal W* (Jones et al., 2014; Larkin et al., 2007).

# Resultados

## 1. Transcriptomas embrionarios de *Rhodnius prolixus*

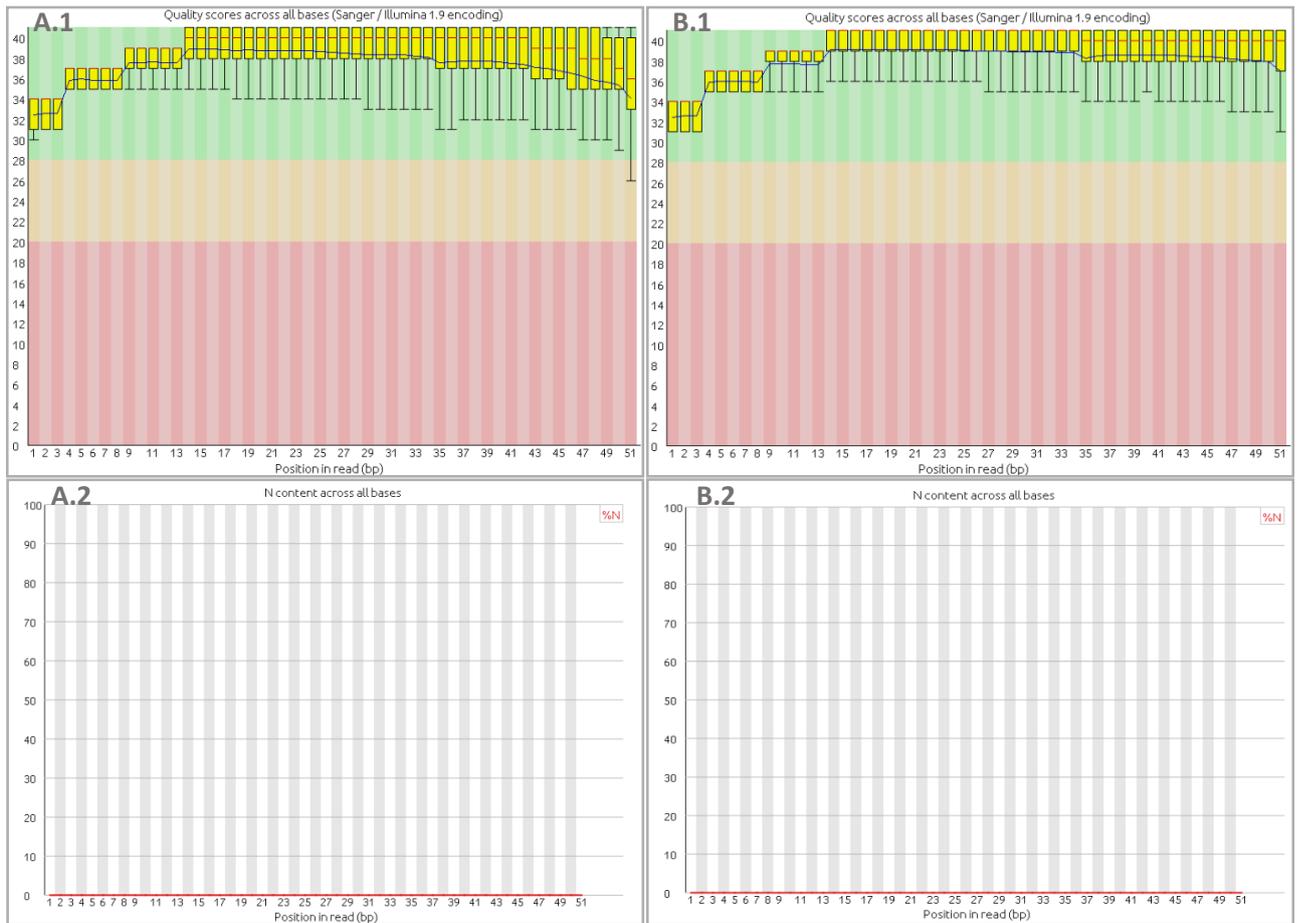
### 1.1 Análisis de calidad y limpieza de los transcriptomas embrionarios de *R. prolixus*

Como primera aproximación de la calidad y profundidad de los genomas expresados se llevó a cabo un análisis con el programa *FastQC*, lo que permitió obtener una exploración rápida de los datos crudos generados por la plataforma de secuenciación, y determinar los pasos sucesivos necesarios a seguir. Como primer paso se corroboraron los parámetros de corrida establecidos a la hora de efectuar la misma en la plataforma de secuenciación, se observó que la longitud final de las lecturas obtenidas fue de 51 pb, así como también que la secuenciación fue desde un único extremo, secuenciación *single-end*, resultados concordantes con lo establecido previamente. Para cada uno de los tiempos secuenciados se generaron 29.791.900 lecturas correspondientes al transcriptoma *0hpo*, 80.970.506 para el transcriptoma *12hpo*, 34.236.005 *reads* del transcriptoma *24hpo* y 48.825.374 lecturas correspondientes al transcriptoma *48hpo*.

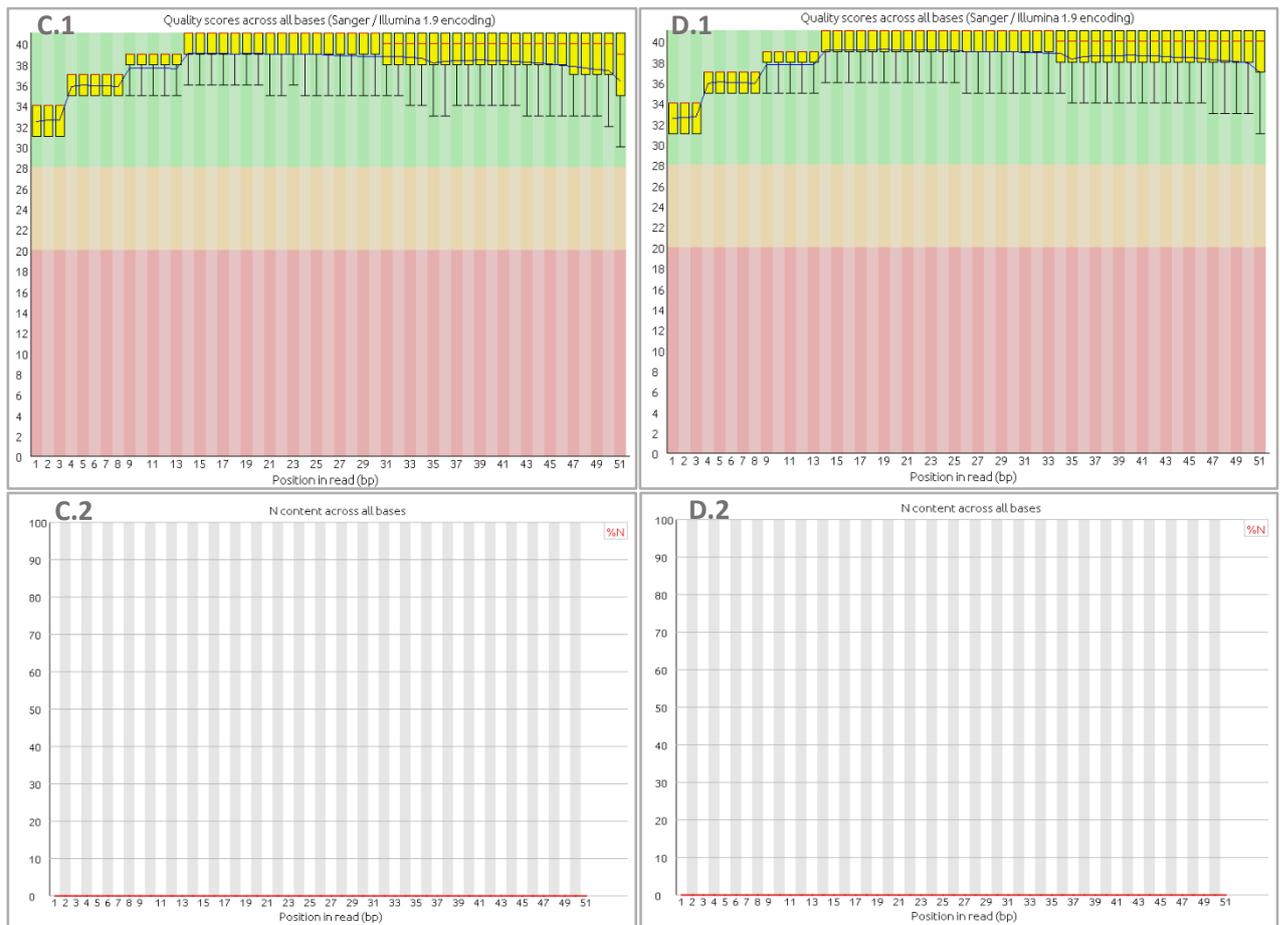
En cada tiempo secuenciado se observó el porcentaje correspondiente al contenido de GC, el valor Q, referente a la calidad de las secuencias generadas, y la presencia de bases no convencionales a lo largo de los 51 pb de las lecturas obtenidas. Respecto al porcentaje de contenido de GC se observó en tiempo *0hpo* 54%, en *12hpo* 37%, para *24hpo* 41% y en el tiempo *48hpo* 38%. Como se muestra en la Figura 3, los reportes de calidad de las secuencias, por base, fueron muy buenos, observándose para todos los casos valores superiores a Q30 a lo largo de toda la longitud de secuencia, lo que predice que la probabilidad de incorporar una base por error fue de 1 en 1.000. Por último, se descartó la presencia de N en las *reads*, como resultado de la no identificación en esa posición de una base convencional.

A pesar de que los datos crudos obtenidos para cada uno de los transcriptomas secuenciados mostraron métricas de calidad con valores óptimos para continuar con el paso de mapeo, se observó la presencia de secuencias sobre-representadas, lo que corresponderían a artefactos contaminantes propios de la secuenciación. Con la herramienta bioinformática *Fastx-toolkit*, el *set* de datos crudos se convirtió de formato FASTQ a FASTA, incorporando como parámetro para realizar dicha conversión que se descarten las secuencias que presentan alguna base no convencional (N) en toda su longitud, con el objetivo de

complementar el análisis con *FastQC*. Una vez realizada dicha conversión, se decidió llevar a cabo una búsqueda de similitud de secuencia. Para esto se utilizó, por un lado, el *set* de datos correspondiente a cada tiempo secuenciado y, por el otro, la base de datos de UniVec disponible *online* en NCBI. En todos los *sets* de datos analizados se encontraron la presencia de secuencias extras, por lo que se implementó un *script* de *Phyton* para su remoción. Como consecuencia de tal, el número de *reads* que conforman cada uno de los archivos de secuenciación se redujo, dejando un número de 29.689.359 lecturas en transcriptoma *0hpo*, 80.944.996 para el transcriptoma *12hpo*, 34.212.363 *reads* del transcriptoma *24hpo* y 48.677.380 lecturas correspondientes al transcriptoma *48hpo*. Estos archivos limpios fueron los materiales de partida para poder proseguir con el paso subsiguiente correspondiente al alineamiento de las secuencias crudas a la referencia.



**Figura 3:** Análisis de calidad de las lecturas que conforman a los transcriptomas secuenciados correspondientes a tiempo 0 (A), 12 (B), 24 (C) y 48 (D) horas post oviposición (*hpo*) mediante *FastQC*. (1) Calidad de secuenciación por base. En el eje Y se observan los valores de calidad posibles, abarcando valores de 0 a 40 (valor Q), y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. (2) Presencia de N a lo largo de la secuencia. En el eje Y se representa la presencia de N en valor porcentual (0 a 100%) y en el eje X se representan las diferentes posiciones de las lecturas en pares de bases. (Figura 3 continúa en la siguiente hoja)



**Figura 3 (continuación):** Análisis de calidad de las lecturas que conforman a los transcriptomas secuenciados correspondientes a tiempo 0 (A), 12 (B), 24 (C) y 48 (D) horas post oviposición (*hpo*) mediante *FastQC*. (1) Calidad de secuencia por base. En el eje y se observan los valores de calidad posibles, abarcando valores de 0 a 40 (valor Q), y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. (2) Presencia de N a lo largo de la secuencia. En el eje Y se representa la presencia de N mediadas en valor porcentual (0 a 100%) y en el eje X se representan las diferentes posiciones de las lecturas en pares de bases.

## 1.2 Alineamiento de *reads* contra el genoma anotado como referencia y comparación de métodos

Una vez llevado a cabo el análisis de calidad de las *reads*, junto con la respectiva depuración de las mismas, se procedió al alineamiento y mapeo de las lecturas limpias frente a la referencia, con el objetivo fundamental de conocer la ubicación de las correspondientes *reads* con respecto a dicha referencia. Como referencia se utilizó el genoma anotado de *R. prolixus*, disponible *online* en VectorBase, el cual está compuesto por 16.857 transcriptos predichos *ab initio*.

Como se mencionó previamente, se utilizaron dos estrategias diferentes para lograr este paso. Por un lado, se utilizó *Bowtie1* para realizar los alineamientos sin permitir la presencia de *gaps* entre las lecturas; y, por otro lado, se utilizó *Tophat2* que permite *gaps* durante el alineamiento.

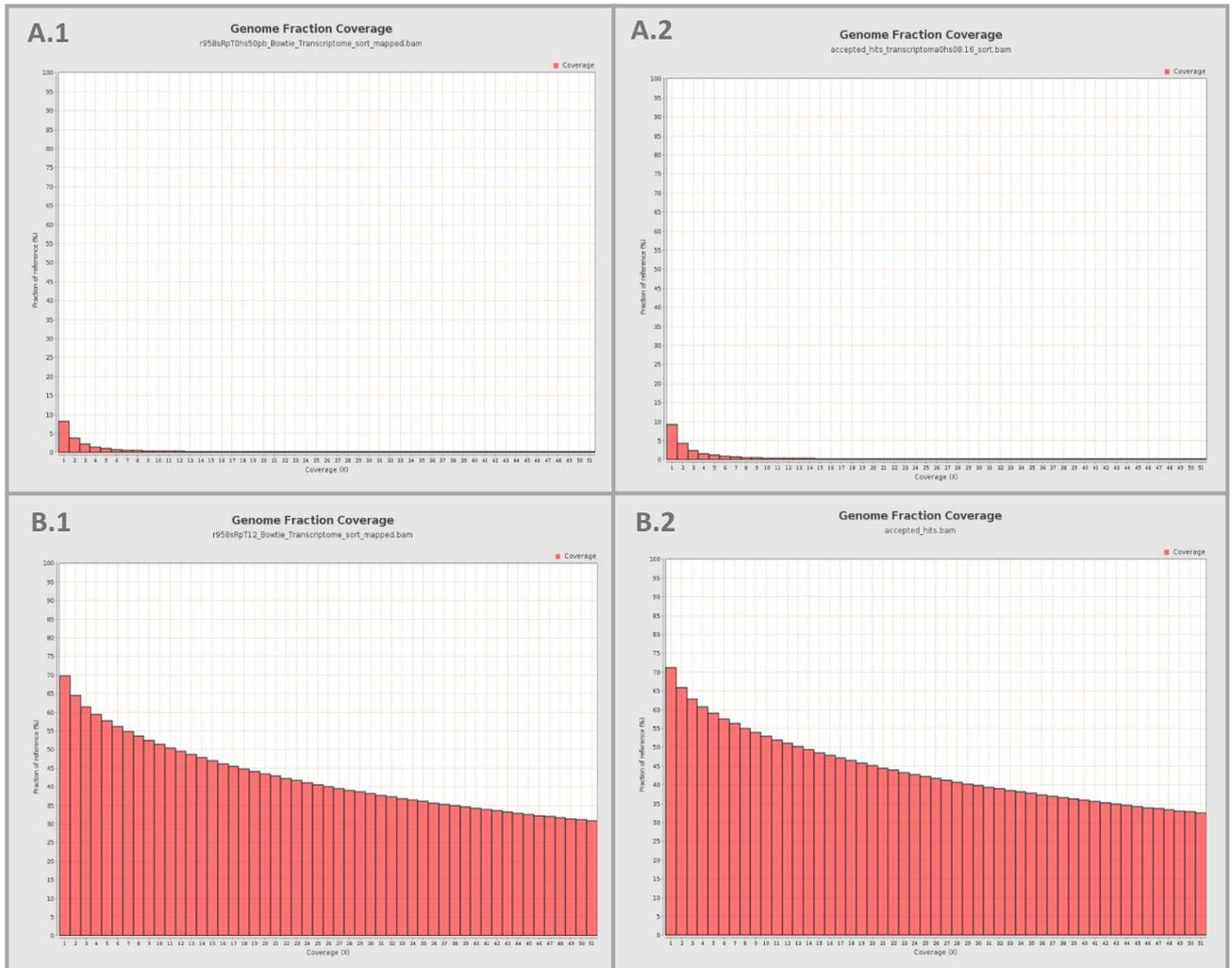
Utilizando el programa de evaluación de alineamiento, *Qualimap*, fue posible conocer las características de cada mapeo realizado (Tabla 1). Para cada tiempo analizado se obtuvo un porcentaje determinado de lecturas alineadas a la referencia; para el alineador *Bowtie1* se observó un porcentaje de *reads* alineadas que mapearon con la referencia de 0.75% en el transcriptoma de *0hpo*, 51.5% en el de *12hpo*, 43.3% en el transcriptoma *24hpo*, y 67.46% en el transcriptoma *48hpo*. Para el programa *Tophat2* el porcentaje de lecturas alineadas en el transcriptoma *0hpo* fue de 0.97%, en *12hpo* fue 57.5%, en *24hpo* 50.26% y en *48hpo* fue de 74.9%. A pesar de que se observó variación en el número de lecturas mapeadas, pudimos ver que la utilización de algoritmos diferentes de mapeo no llevó a variaciones significativas en las características del alineamiento respecto a la tasa de duplicación y el porcentaje correspondiente al contenido de GC. La tasa de duplicación, estimada a partir de la posición inicial correspondiente al alineamiento de las *reads* a la referencia, reportó valores que rondaron el 50%. Sí fue posible observar en todos los casos que el número de lecturas que se alinearon a la referencia fue mayor cuando se utilizó el programa *Tophat2*.

Transcriptoma	Número de lecturas limpias	<i>Bowtie1</i>				<i>Tophat2</i>			
		lecturas mapeadas	<i>reads</i> duplicadas n°/%	tasa de duplicación	%GC	lecturas mapeadas	<i>reads</i> duplicadas n°/%	tasa de duplicación	%GC
<b>0hpo</b>	29.689.359	222.109	156.156/ 70,31%	42,74%	43,06%	288.095	214.017/ 74,3%	42,63%	43,77%
<b>12hpo</b>	80.944.996	41.692.390	36.603.812/ 87,79%	66,80%	40,29%	46.575.687	41.205.132/ 84,7%	67,92%	40,33%
<b>24hpo</b>	34.212.363	14.814.182	11.943.016/ 80,62%	53,36%	40,27%	17.196.510	14.096.273/ 81,97%	54,55%	40,28%
<b>48hpo</b>	48.677.380	32.924.711	28.177.648/ 85,58%	56,46%	42,85%	36.584.916	31.512.684/ 86,14%	57,64%	42,73%

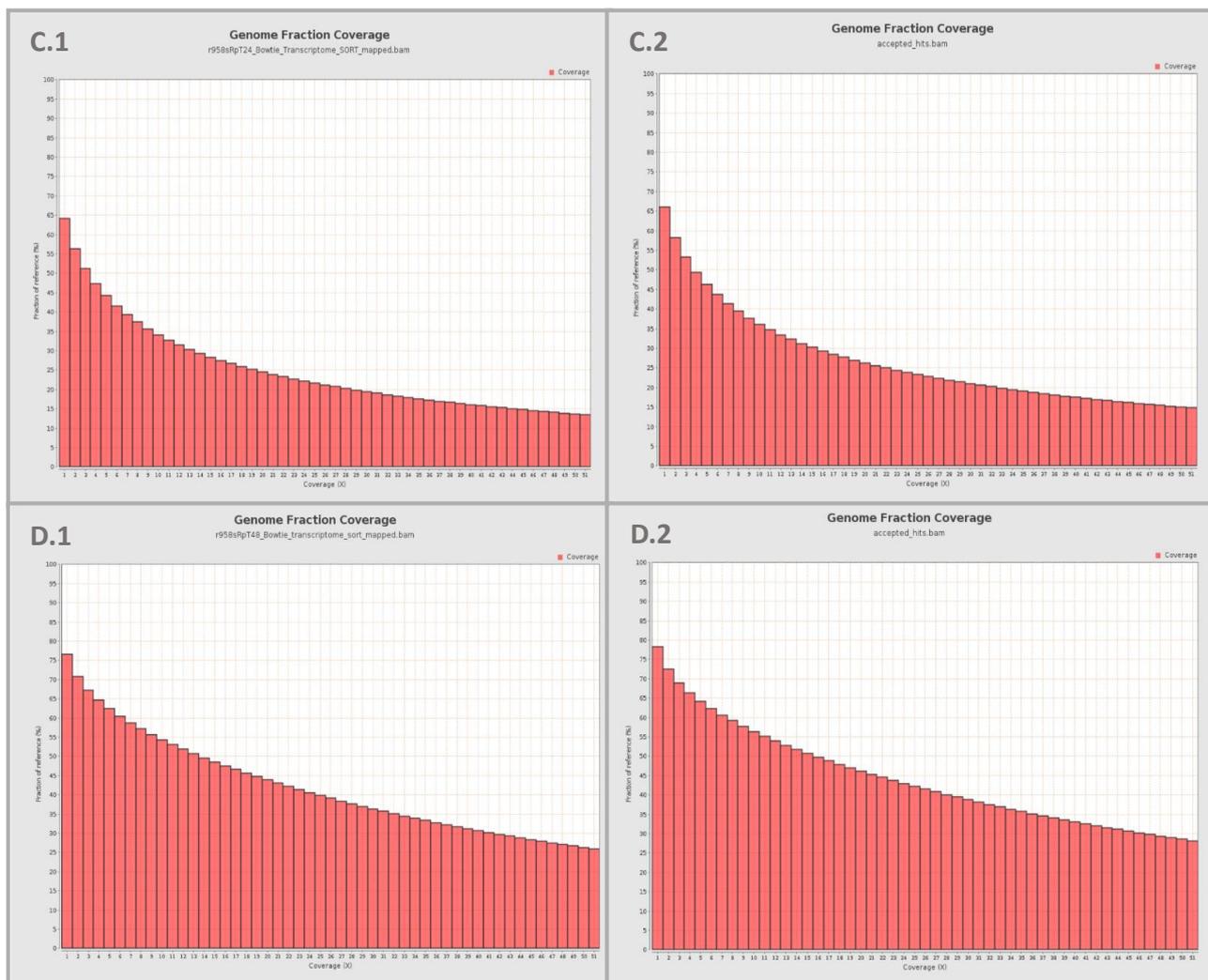
**Tabla 1:** Representación de la estadística obtenida tras el mapeo de los distintos transcriptomas generados utilizando dos programas diferentes para alinear, *Bowtie1* o *Tophat2*. Para cada programa utilizado se reporta el número de lecturas mapeadas, el número y porcentaje de *reads* duplicadas, la tasa de duplicación y el porcentaje del contenido GC.

Además de la estadística detallada (Tabla 1), fue posible inferir la fracción de cobertura respecto a la referencia para cada uno de los transcriptomas generados. Esto se entiende

como la cantidad de referencia que ha sido secuenciada con al menos una tasa determinada de cobertura. La Figura 4 muestra los gráficos correspondientes a la fracción de cobertura de la referencia, para cada uno de los tiempos analizados.



**Figura 4:** Gráficos correspondientes a la fracción de cobertura de la referencia para los transcriptomas embrionarios de *R. prolixus*. En el eje Y se representa el % de la fracción de la referencia y en el eje X la cobertura. Las barras rojas representan el porcentaje de cobertura presente. Transcriptomas de 0 (A), 12 (B), 24 (C) y 48 (D) horas post oviposición (*hpo*), mapeados utilizando *Bowtie1* (1) o *Tophat2* (2). (Figura 4 continúa en la siguiente hoja)



**Figura 4 (continuación):** Gráficos correspondientes a la fracción de cobertura de la referencia para los transcriptomas embrionarios de *R. prolixus*. En el eje Y se representa el % de la fracción de la referencia y en el eje X la cobertura. Las barras rojas representan el porcentaje de cobertura presente. Transcriptomas de 0 (A), 12 (B), 24 (C) y 48 (D) horas post oviposición (*hpo*), mapeados utilizando *Bowtie1* (1) o *Tophat2* (2).

En correlación al bajo número de *reads* mapeadas para el tiempo *0hpo*, se encontró una muy baja cobertura en la referencia, tanto para el mapeo con *Bowtie1* como con *Tophat2*, observándose que alrededor de un 5% de la referencia presentó una cobertura de 2X. En contraposición, para los tiempos *12hpo* y *24hpo* secuenciados se evidenció un alto número de lecturas mapeadas que se correlacionó con mayor cobertura de la referencia.

Respecto a los otros tiempos analizados, comparando ambos programas de mapeo, el tiempo *12hpo* generado con *Bowtie1* mostró una cobertura menor que con *Tophat2*, donde se observó que un 45% de la referencia con *Bowtie1* presentó una cobertura de 18X mientras

que con *Tophat2* una cobertura de 30X. Los restantes transcriptomas no mostraron valores indicativos diferentes al comparar la fracción de cobertura de la referencia.

### **1.3 Reconstrucción de los transcriptomas a partir del mapeo contra el genoma anotado de *R. prolixus***

Los archivos de mapeo generados en el paso anterior, fueron ordenados respecto a la posición de la referencia, determinada en pares de bases. Este ordenamiento se realizó con el paquete SAMTOOLS, y se generó un archivo de mapeo ordenado y uno indexado, correspondiente al mismo. Estos pasos fueron necesarios para llevar a cabo el proceso de ensamblado propiamente dicho, utilizando *Cufflinks*.

*Cufflinks* hace referencia tanto al programa como a una de las herramientas que forman parte del programa, y para la reconstrucción de los transcriptomas secuenciados en este trabajo se utilizaron ambos.

La herramienta *Cufflinks* generó tres archivos de salida, uno de los cuales se correspondió a las coordenadas de los transcritos generados con su respectivo valor de FPKM (*Fragments per kilobase of exon model per million reads mapped*); este es una unidad de expresión que representa una estimación normalizada de la expresión génica, calculado a partir del número de *reads* que mapearon a cada secuencia génica en particular y teniendo en cuenta la longitud de la secuencia y la profundidad de secuenciación. Como en este trabajo no se llevaron a cabo réplicas de secuenciación, este valor no es útil para un análisis de expresión génica diferencial, pero sí permite obtener un estimativo de la abundancia de cada transcrito en la muestra secuenciada. Los otros dos archivos generados mediante el ensamblador correspondieron, por un lado, a las coordenadas de las predicciones génicas y, por el otro, a las posibles isoformas; cada uno con sus respectivos valores de FPKM.

Más allá de transmitir la abundancia de cada transcrito dentro del transcriptoma generado (Suplementario 2), la reconstrucción del transcriptoma permite conocer la cantidad de transcritos ensamblados dentro de cada tiempo secuenciado. Así fue que se obtuvieron 1.412 transcritos para el tiempo *Ohpo* a partir de la utilización del programa *Bowtie1* y 1.339 transcritos luego de la utilización de *Tophat2*; 12.285 transcritos con *Bowtie1* y 11.324 con

*Tophat2* para el transcriptoma 12hpo; 12.264 transcritos tras utilizar el programa *Bowtie1* y 11.444 transcritos con *Tophat2* en el tiempo 24hpo; y, por último, 13.566 transcritos con *Bowtie1* y 12.317 con *Tophat2* para el transcriptoma correspondiente a 48hpo. Comparando ambos programas de mapeo es posible destacar que para todos los casos fue mayor el número de transcritos generados por transcriptoma utilizando el programa de mapeo *Bowtie1*. Ahora bien, observando en detalle las características de cada uno de los transcriptomas obtenidos, se reconoce que los transcritos generados a partir del programa de alineamiento *Bowtie1* presentan una menor longitud de secuencia en comparación con los obtenidos por *Tophat2*. Es importante remarcar que, según el algoritmo de mapeo utilizado se ensamblaron transcritos que fueron compartidos por ambos programas o fueron programa/método específico. Cuando se habla de transcritos compartidos se hace alusión a que el identificador correspondiente a cada una de las secuencias que componen a la referencia se encuentran presentes en ambos transcriptomas generados, sin importar la posición cromosómica ni la longitud de la predicción.

A la hora de explorar los transcriptomas generados, para evitar omitir alguna inferencia de identidad, se decidió utilizar la herramienta *Cuffmerge* de *Cufflinks*. Esta toma los fragmentos ensamblados por *Cufflinks* de la misma manera que *Cufflinks* trata a las lecturas mapeadas, las combina parsimoniosamente y produce un nuevo ensamble “merged” con el menor número de transcritos que explican de la mejor manera los datos totales. Se combinaron los transcritos ensamblados a partir del *unspliced aligner* (*Bowtie1*) con los generados a partir del *spliced aligner* (*Tophat2*), para cada tiempo secuenciado, y se generó un único archivo de ensamble. El número total de transcritos ensamblados varió según el transcriptoma analizado. En algunos casos, la mejor reconstrucción se explicó con un número mayor de transcritos que los iniciales generados de manera independiente por *Cufflinks*, y en otros, con un número promedio entre los dos ensambles obtenidos previamente (Tabla 2).

Transcriptoma	Método	Número de Transcritos reconstruidos	Número de Transcritos Método-específico	N° de Transcritos Compartidos	N° de Transcritos Cuffmerge
<b>0hpo</b>	<i>unspliced aligner</i>	1.412	641	360	1.589
	<i>spliced aligner</i>	1.339	548		
<b>12hpo</b>	<i>unspliced aligner</i>	12.285	6.556	4.902	12.193
	<i>spliced aligner</i>	11.324	5.595		
<b>24hpo</b>	<i>unspliced aligner</i>	12.264	6.133	4.631	11.444
	<i>spliced aligner</i>	11.444	5.313		
<b>48hpo</b>	<i>unspliced aligner</i>	13.566	7.234	5.148	13.512
	<i>spliced aligner</i>	12.317	5.983		

**Tabla 2:** Reconstrucción de los transcritos en cada transcriptoma secuenciado en *single-end*. Para cada programa de alineamiento utilizado se reporta el número de transcritos reconstruidos, los transcritos específicos a cada programa y los compartidos, y el número de transcritos generados por *Cuffmerge*.

Para descartar la presencia de remanentes de artefactos de la secuenciación, los transcriptomas generados fueron evaluados mediante una búsqueda de similitud de secuencia con el algoritmo BLASTX contra la misma base de datos que se utilizó para la limpieza de las lecturas crudas (UniVec). Para todos los casos el resultado fue negativo, desechando la posibilidad de contar con secuencias no deseadas en alguno de los transcriptomas.

#### 1.4 Alineamiento de *reads* utilizando al genoma como referencia

Además del alineamiento del *set* de datos limpios contra el genoma anotado, se prosiguió con el mapeo utilizando ahora al complemento genético total, genoma, de *R. prolixus* como referencia. El mismo se encuentra disponible *online* en VectorBase, y presenta 47.726 secuencias génicas, también denominadas *contigs*. Debido a la presencia de regiones intrónicas en el genoma sólo se llevó a cabo el alineamiento utilizando *Tophat2*, ya que utiliza un algoritmo el cual no considera como una dificultad la presencia de estas grandes regiones para realizar el mapeo.

Exceptuando el tiempo *0hpo*, en los restantes tiempos secuenciados, el número total de lecturas mapeadas fue mayor que el número de lecturas iniciales. Esto se debió a que las

lecturas mapearon más de una vez en la referencia, fenómeno denominado *multi-mapped*. El alineamiento generó un único archivo correspondiente a las lecturas totales que mapean a la referencia, en el cuál se diferencian las lecturas que mapean por primera vez a la referencia, (lecturas primarias) de las que corresponden a las que mapearon subsiguientemente (lecturas secundarias). En la tabla 3 se detalla la estadística del mapeo obtenida para cada tiempo analizado, utilizando el programa *Qualimap*. Comparando los cuatros tiempos secuenciados, se reportó que el porcentaje de mapeo correspondiente a *0hpo* fue muy bajo con respecto a los demás, siendo alrededor de un 0.8% a diferencia de los demás que presentan un porcentaje que ronda entre los 70-95%.

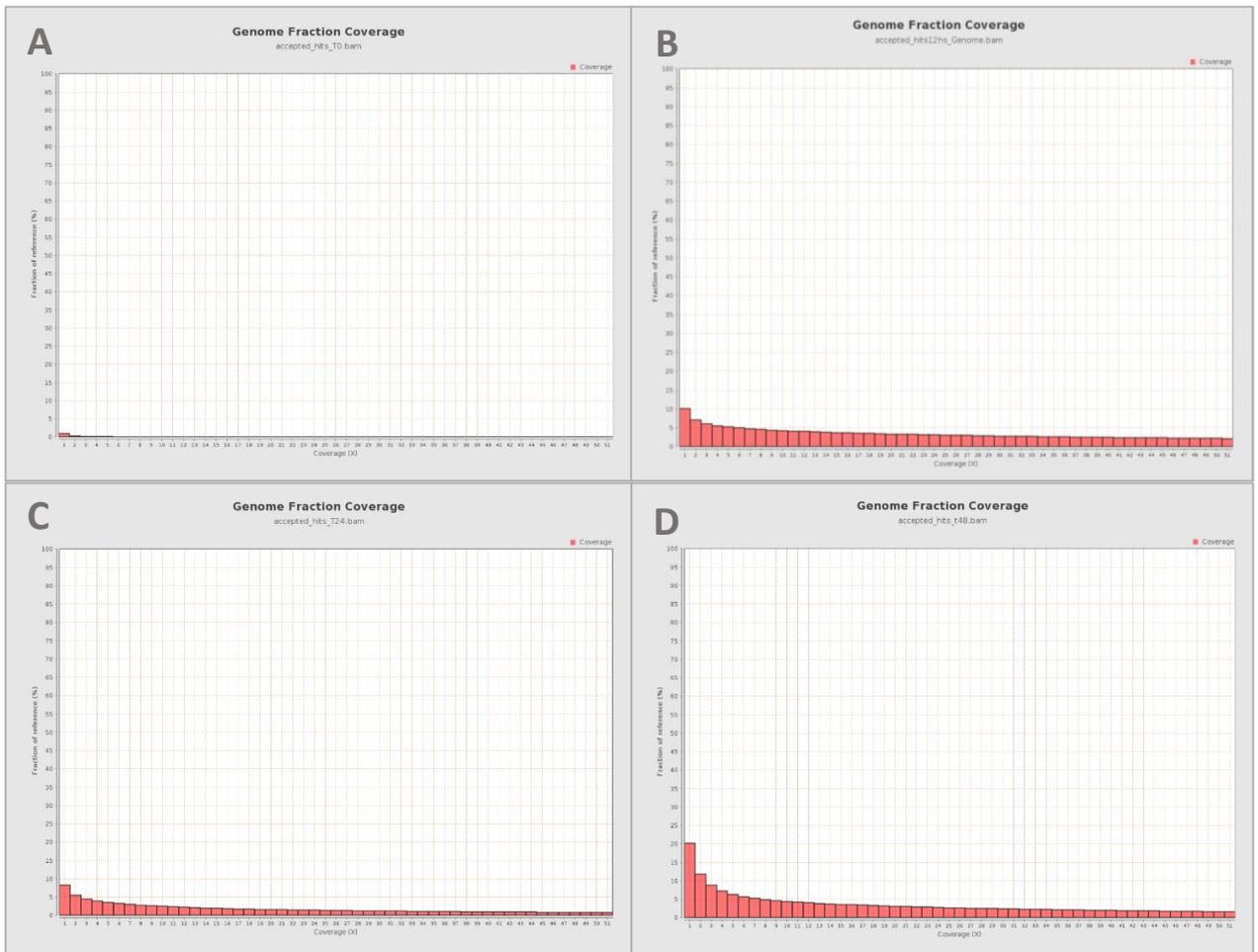
La ventaja que presenta el mapeo contra el genoma como referencia es que permite identificar lecturas que mapearon en regiones de empalme o en regiones previamente no reportadas como secuencias codificantes. Es por esto que, utilizando esta estrategia, se observó un mayor número de alineamientos en comparación al número obtenido utilizando al genoma anotado como referencia.

Transcriptomas	Número de lecturas limpias	<i>Tophat2</i>		
		Número de lecturas mapeadas totales "multi-mapped"	Número de lecturas primarias	Número de lecturas en región de empalme
<b>0hpo</b>	29.689.359	1.734.584	256.163	9564
<b>12hpo</b>	80.944.996	179.294.264	74.076.838	9.478.704
<b>24hpo</b>	34.212.363	84.636.003	23.994.004	3.028.523
<b>48hpo</b>	48.677.380	172.563.021	46.536.363	5.898.594

**Tabla 3:** Reconstrucción de los transcriptos en cada transcriptoma secuenciado en *single-end* y mapeado contra el genoma como referencia. Se reporta el número de lecturas iniciales, alineamientos totales realizados, las lecturas primarias y lecturas que se ubicaron en la región de empalme.

Por otro lado, se evaluó la fracción de cobertura de la referencia para cada uno de los transcriptomas analizados. Esta fue muy baja en comparación a los resultados obtenidos utilizando al transcriptoma como referencia (Figura 5). En correlación con el bajo número de *reads* mapeadas para el tiempo *0hpo* se observó una muy baja cobertura de la referencia,

observándose que menos de un 2% de la referencia presentó una cobertura de 1X. Para el tiempo 12hpo se observó que un 6% de la referencia presentó una cobertura de 6x; mientras que para el transcriptoma correspondiente a 24hpo aproximadamente un 6% de la referencia presentó una cobertura de 2x. En cambio, para 48hpo la cobertura, correspondiente al mismo porcentaje de la referencia, fue de 6x.



**Figura 5:** Gráficos correspondientes a la fracción de cobertura de la referencia para los transcriptomas embrionarios de *R. prolixus*, mapeados contra el genoma como referencia utilizando *Tophat2*. Transcriptomas a tiempo 0 (A), 12 (B), 24 (C) y 48 (D) horas post oviposición (hpo). En el eje Y se representa el % de la fracción de la referencia y en el eje X la cobertura. Las barras rojas representan el porcentaje de cobertura presente.

### **1.5 Reconstrucción de los transcriptomas a partir del mapeo contra el Genoma de *R. prolixus***

A partir del archivo de mapeo generado en el paso anterior, se llevó a cabo el proceso de ensamblado utilizando el programa *Cufflinks*, para cada uno de los tiempos analizados. Como sucedió para los ensamblajes que se originaron a partir del mapeo contra los transcriptos predichos *ab initio*, se generaron tres archivos: uno de los cuales se correspondió a las coordenadas de los transcriptos generados, otro a las coordenadas de las predicciones génicas y, el último a las posibles isoformas; cada uno con sus respectivos valores de FPKM (*Fragments per kilobase of exon model per million reads mapped*).

La reconstrucción del transcriptoma permite conocer la cantidad de transcriptos ensamblados dentro de cada tiempo secuenciado. Así fue que se obtuvieron 1.755 transcriptos para el tiempo *0hpo*, 27.189 transcriptos para *12hpo*; 28.013 transcriptos para el tiempo *24hpo* y, 33.132 transcriptos para el transcriptoma correspondiente a *48hpo*.

Para descartar presencia de remanentes de artefactos de la secuenciación, los transcriptomas generados fueron evaluados mediante una búsqueda de similitud con el algoritmo BLASTX contra UniVec. Para todos los casos el resultado fue negativo, desechando la posibilidad de contar con un transcriptoma con secuencias no deseadas.

## **2. Transcriptomas de ovario, huevos no fertilizados y de 0 horas post ovoposición de *Rhodnius prolixus***

### **2.1 Análisis de calidad y limpieza de los transcriptomas generados mediante secuenciación *paired-end***

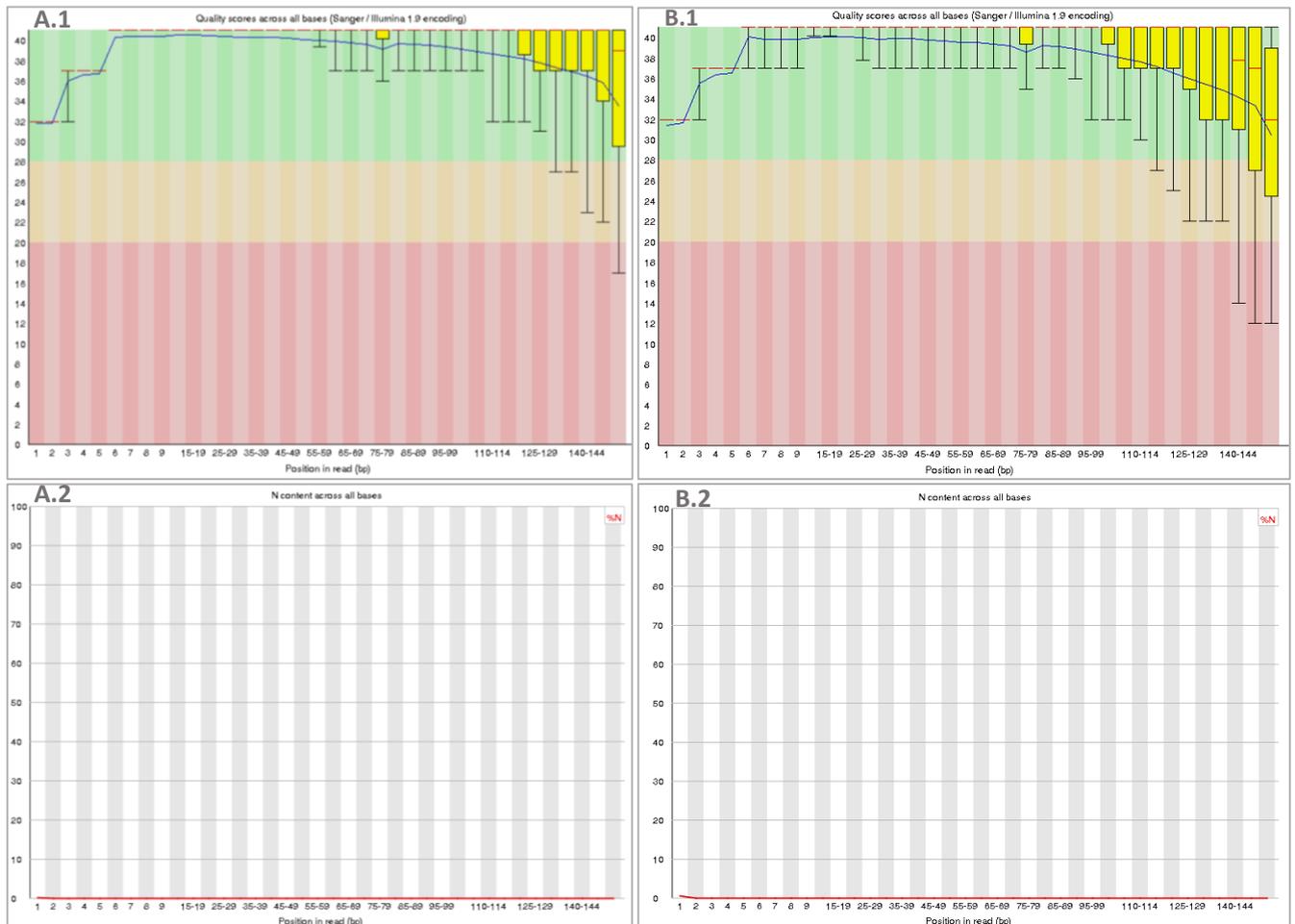
Así como se analizaron los transcriptomas embrionarios *single-end* previamente descritos, se llevó a cabo el mismo análisis para los transcriptomas correspondientes a ovariolas completas, huevos ovipuestos de hembras vírgenes (no fertilizados) y huevos correspondientes a *0hpo* todos generados mediante secuenciación *paired-end*. Se decidió realizar el análisis de estos tiempos con el objetivo de estudiar el desarrollo embrionario temprano desde la transición materna-cigótica, y lograr inferir los transcriptos de origen maternos. Asimismo, se decidió re-secuenciar el estadio de huevos a tiempo *0hpo* debido a la

baja cobertura y al bajo número de transcriptos reconstruidos que se generaron a partir de la secuenciación *single-end*.

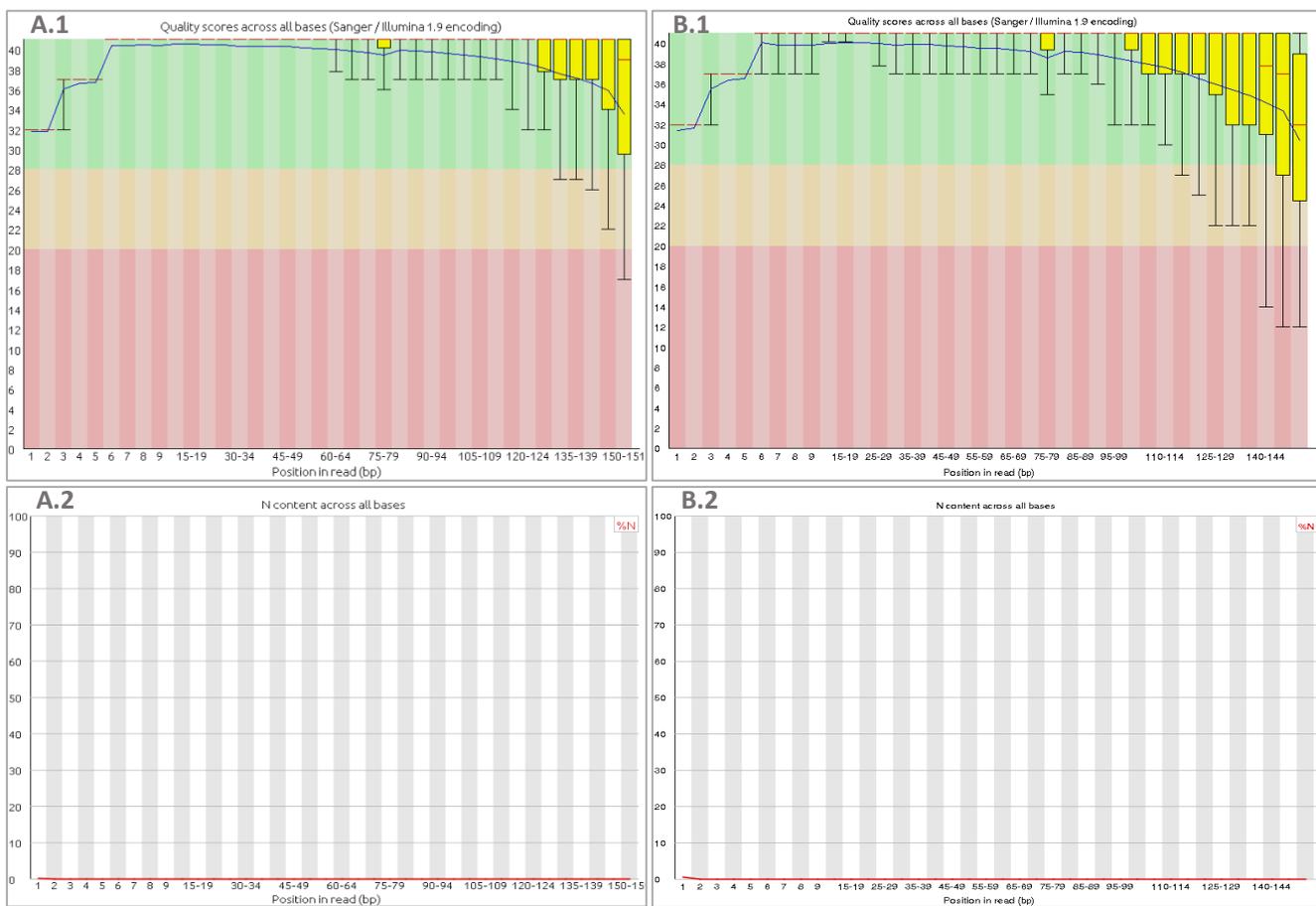
Debido al tipo de secuenciación llevada a cabo, estos transcriptomas muestran dos archivos de partida para analizar la calidad de secuencia por base (R1 y R2), cada uno correspondiente a cada sentido de secuenciación realizado (*paired-end*). En el caso del transcriptoma de ovario, como resultado de la salida cruda se obtuvieron 187.726.140 *reads* para la corrida R1 y 186.676.153 lecturas en la corrida R2. Para el transcriptoma de huevos no fertilizados se observó que el número de *reads* obtenido para cada una de las corridas R1 y R2 fue, 186.676.153 y 184.269.740, respectivamente. Por último, en lo que respecta al transcriptoma para *Ohpo* se generaron 127.066.601 lecturas correspondientes a la corrida R1 y 196.665.388 lecturas de la corrida R2. Estos *sets* de datos fueron analizados por el programa *FastQC*.

En los transcriptomas correspondiente a ovario (Figura 6) y huevos no fertilizados (Figura 7) se reportaron para todas las secuencias la misma longitud de 150 pb, coincidente con lo establecido en la plataforma de secuenciación. Además, se registraron valores de calidad superiores a Q40 en casi toda su longitud, lo que predice que la probabilidad de incorporar una base por error fue de 1 en 10.000. Así como también no mostraron presencia de N en ninguna de las lecturas presentes. Esta situación se observó tanto para las lecturas proporcionadas por la corrida R1 como para las de la corrida R2 (Figura 6 y 7). Sí se encontró la presencia de artefactos contaminantes, propios de la secuenciación, por lo que se prosiguió a la conversión del formato del conjunto de datos crudos de secuenciación desde FASTQ a FASTA utilizando la herramienta bioinformática *Fastx-toolkit*, para proseguir con la remoción de esas secuencias. Además, se evaluó la presencia de alguna base no convencional en las lecturas con el objetivo de complementar el análisis previo. La remoción de las secuencias con artefactos de secuenciación y/o secuencias contaminantes se realizó utilizando la estrategia de limpieza con el algoritmo BLASTX, realizando una búsqueda de similitud de secuencia contra la base de datos del UniVec. Las lecturas que dieron *hit* con la base de datos fueron eliminadas del *set* de datos crudos y, para realizar la depuración, se utilizó el *script* de *Phyton* antes mencionado. Tras este paso, el número de lecturas disminuyó, quedando el transcriptoma de ovario con 143.393.553 lecturas limpias de la corrida R1 y 144.957.560

lecturas limpias de la corrida R2; mientras que, en el transcriptoma correspondiente a huevos no fertilizados el número de lecturas presentes en cada archivo cambió a 130.084.446 *reads* en R1 y 130.728.390 *reads* en R2.



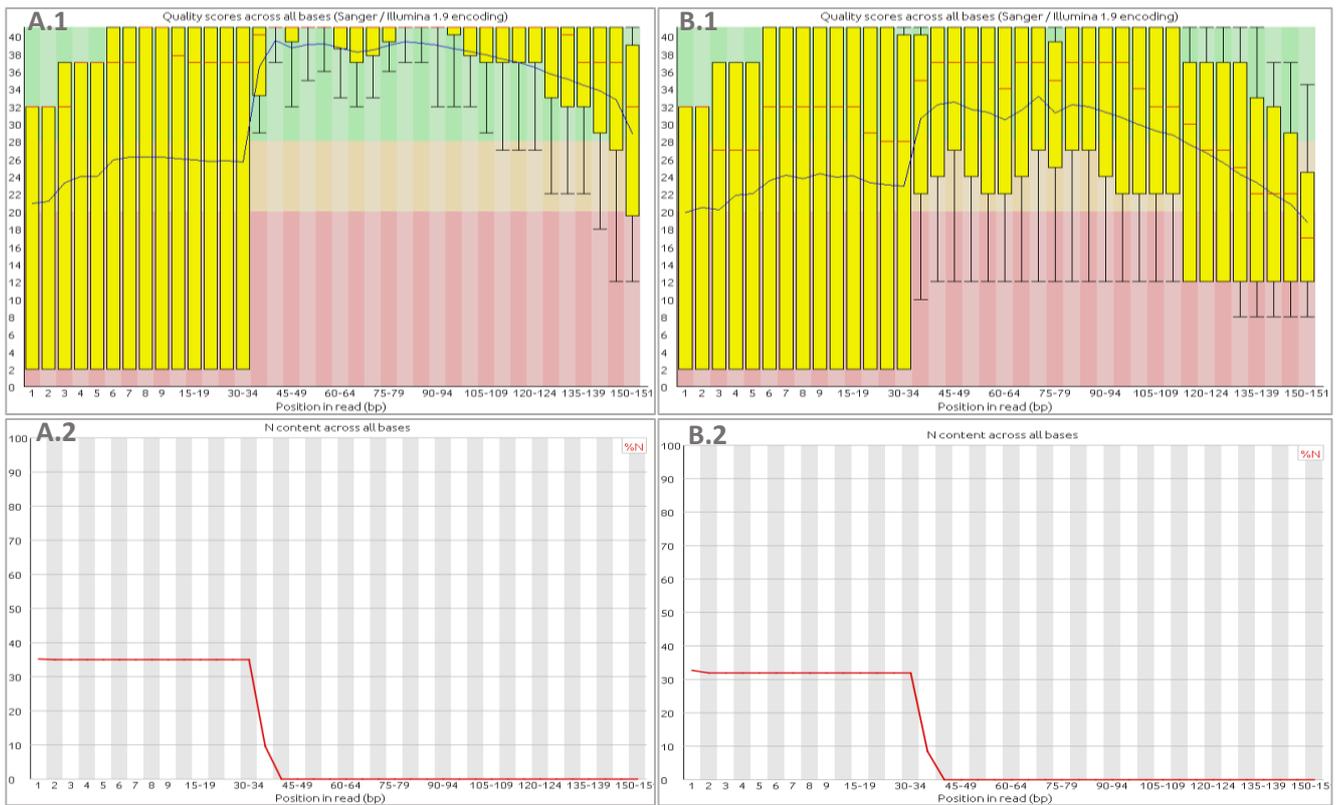
**Figura 6:** Análisis de calidad de las lecturas que conforman al transcriptoma correspondientes al ovario mediante *FastQC*. **(A)** Corrida R1, **(B)** Corrida R2. **(1)** Calidad de secuencia por base. En el eje y se observan los valores de calidad posibles, abarcando valores de 0 a 40 (valor Q), y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. **(2)** Presencia de N a lo largo de la secuencia. En el eje Y se representa la presencia de N mediadas en valor porcentual (0 a 100%) y en el eje X se representan las diferentes posiciones de las lecturas en pares de bases.



**Figura 7:** Análisis de calidad de las lecturas que conforman al transcriptoma correspondiente a huevos no fertilizados mediante *FastQC*. **(A)** Corrida R1, **(B)** Corrida R2. **(1)** Calidad de secuencia por base. En el eje y se observan los valores de calidad posibles, abarcando valores de 0 a 40 (valor Q), y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. **(2)** Presencia de N a lo largo de la secuencia. En el eje Y se representa la presencia de N mediadas en valor porcentual (0 a 100%) y en el eje X se representan las diferentes posiciones de las lecturas en pares de bases.

El transcriptoma *paired-end* para *Ohpo* presentó medidas adversas de calidad (Figura 8). En principio las corridas generaron una cantidad muy diferente de *reads*, obteniendo para la corrida R2 69.598.787 lecturas de más. Las lecturas crudas, tanto las R1 como las R2, presentaron muy mala calidad, con valores muy por debajo de Q30, principalmente desde el inicio de la *read* hasta los 30-34 pb, alcanzando valores tan bajos como Q2. En esta misma región de las lecturas se detectó la presencia de N con un porcentaje del 35%, en contraposición a los resultados anteriores (Figura 6, 7 y 8). Sumado a esto se encontraron secuencias sobrerrepresentadas, propias de artefactos contaminantes de la secuenciación. Para la limpieza de este *set* de datos se realizó un filtrado de calidad de las secuencias

utilizando como valor de corte un valor Q30, y luego se prosiguió a eliminar las lecturas con presencia de N, utilizando las herramientas bioinformáticas disponibles en *Fastx-toolkit*. Posteriormente, se realizó la conversión del formato correspondiente a los datos crudos de secuenciación desde FASTQ a FASTA, utilizando el mismo paquete de herramienta bioinformática. Una vez realizada la limpieza por calidad, se realizó la exploración en las lecturas de las secuencias con presencia de artefactos contaminantes. Para esto se realizó una búsqueda de similitud de secuencia contra la base de datos del UniVec, utilizando el algoritmo BLASTX. Una vez identificadas las secuencias contaminadas se implementó el *script* de *Python* para su remoción. Como resultado se obtuvieron 102.124.133 lecturas limpias correspondientes a la corrida R1 y 106.892.825 lecturas correspondientes a la corrida R2. Es notorio que tras la limpieza de los datos obtenidos en la corrida R2 se eliminó una gran cantidad lecturas, lo que coincide con el mayor número de lecturas generadas en la salida cruda de la plataforma de secuenciación.



**Figura 8:** Análisis de calidad de las lecturas que conforman al transcriptoma correspondiente al tiempo 0 horas pos oviposición (*hpo*) mediante FastQC. **(A)** Corrida R1, **(B)** Corrida R2. **(1)** Calidad de secuencia por base. En el eje y se observan los valores de calidad posibles, abarcando valores de 0 a 40 (valor Q), y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. **(2)** Presencia de N a lo largo de la secuencia. En el eje Y se representa la presencia de N mediadas en valor porcentual (0 a 100%) y en el eje X se representan las diferentes posiciones de las lecturas en pares de bases.

## 2.2 Alineamiento de las *reads* al genoma anotado como referencia

Siguiendo con la metodología descrita para el mapeo de las lecturas *single-end* correspondientes a los distintos tiempos del desarrollo, se procedió a alineamiento de las lecturas limpias *paired-end*. En este caso, se utilizó como referencia al genoma anotado de *R. prolixus*, compuesto por 16.857 transcritos predichos, disponible *online* en VectorBase.

En estos transcriptomas, también se utilizaron dos estrategias diferentes para lograr este paso: alineamientos con *Bowtie1* y con *Tophat2*. Este último, utiliza un algoritmo que hace que la presencia de regiones intrónicas no sea un inconveniente a la hora de realizar el alineamiento. Algo que merece ser aclarado es que el tiempo *0hpo paired-end* sólo se prosiguió a alinear frente a la referencia utilizando *Tophat2*, ya que la diferencia presente en

el número de lecturas de los archivos, R1 y R2, utilizados como *input*, impide que el algoritmo de *Bowtie1* genere un alineamiento de manera correcta.

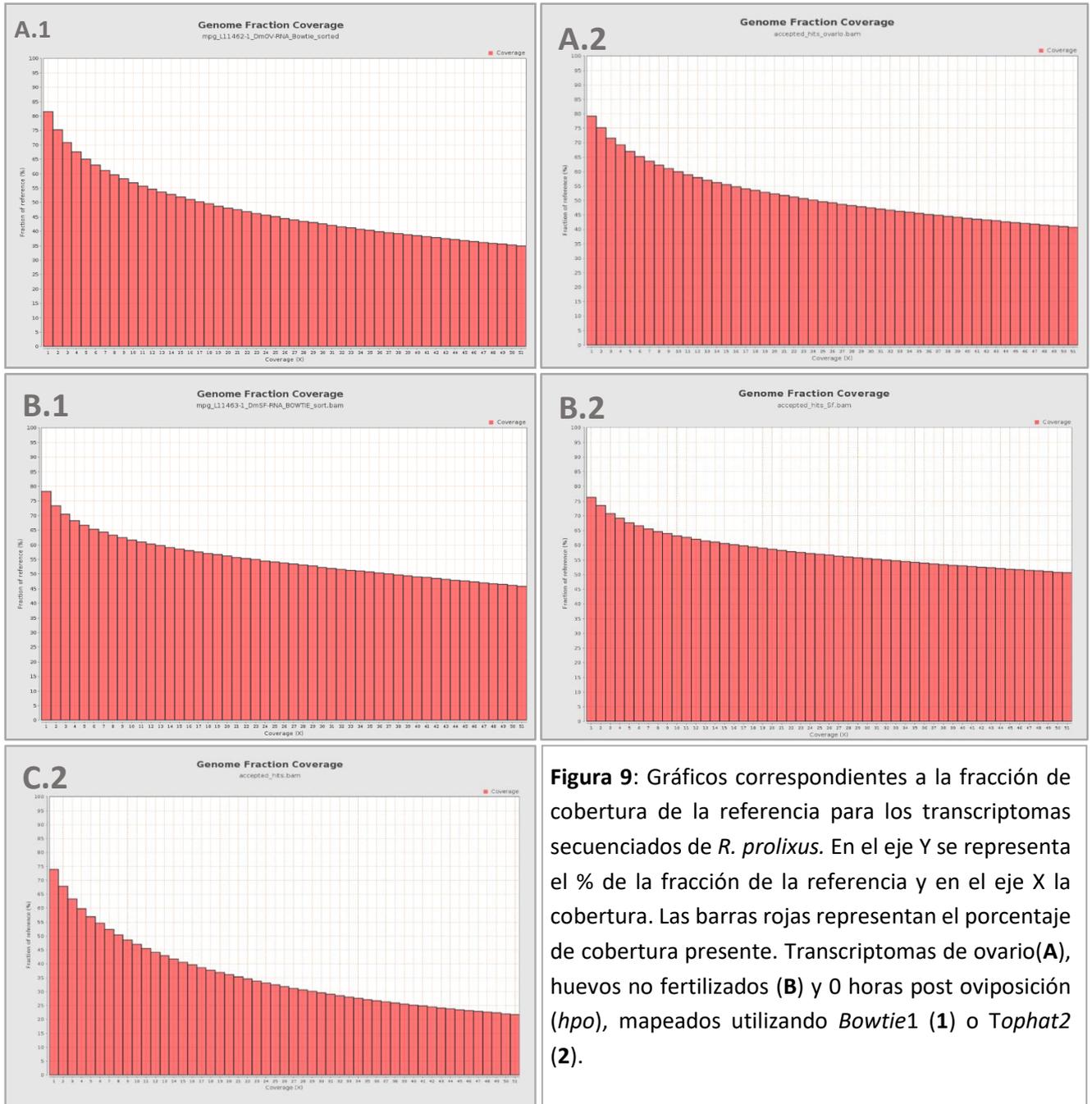
En la Tabla 4 se reporta la estadística para cada transcriptoma generado, donde se observó que el número de lecturas que mapearon a la referencia se diferenciaron según el programa utilizado. En el caso del transcriptoma de ovario, con el programa *Bowtie1* mapearon 150.351.626 lecturas y con *Tophat2* mapearon 94.915.589; para el transcriptoma de huevos no fertilizados, con *Bowtie1* mapearon 325.551.152 y con *Tophat2* 148.240.713. Para el transcriptoma de *Ohpo paired-end*, 33.630.090 *reads* mapearon a la referencia. Además de la cantidad de lecturas que mapearon a la referencia, fue posible conocer las características del mapeo utilizando *Qualimap*. Se identificó la tasa de duplicación, así como también el porcentaje de GC presente en cada mapeo realizado, observándose que las métricas del mapeo fueron muy buenas, y no reportaron diferencias significativas utilizando uno u otro algoritmo.

Transcriptoma	Número de lecturas limpias	<i>Bowtie1</i>				<i>Tophat2</i>			
		N° de lecturas mapeadas	N° de <i>reads</i> duplicadas número/%	tasa de duplicación	%GC	N° de lecturas mapeadas	N° de <i>reads</i> duplicadas número/%	tasa de duplicación	%GC
<b>Ovario</b>	R1 143.393.553	150.351.626	145.817.657/ 44,99%	71,59%	37,76%	94.915.589	91.405.783/ 96,3%	71,35%	37,30%
	R2 144.957.560								
<b>Huevos no fertilizados</b>	R1 130.084.446	325.551.152	320.244.654/ 71,43%	68,70%	52,99%	148.240.713	144.186.748/ 97,27%	67,77%	50,76%
	R2 130.728.390								
<b>Ohpo</b>	R1 102.124.113	*	*	*	*	33.630.090	311.57.279/ 92,56%	56,69%	42,26%
	R2 106.892.825								

**Tabla 4:** Representación de la estadística obtenida tras el mapeo de los distintos transcriptomas generados por secuenciación *paired-end*, utilizando dos programas diferentes para alinear, *Bowtie1* o *Tophat2*. Se reporta el número de lecturas mapeadas, el número y porcentaje de *reads* duplicadas, la tasa de duplicación y el porcentaje del contenido GC.

Por otro lado, se decidió analizar la cobertura de la referencia en cada uno de los transcriptomas secuenciados, representada como la cantidad de referencia que ha sido secuenciada con al menos una tasa determinada de cobertura. La fracción de cobertura de la referencia obtenida fue mayor cuando se utilizó el programa *Tophat2*, observándose que al menos del 50 % de la referencia se secuenció con una cobertura de 24X en el transcriptoma

correspondiente a ovario, 50X en el transcriptoma de huevos no fertilizados, y 8X en el transcriptoma *Ohpo* (Figura 9).



### 2.3 Reconstrucción de los transcriptomas *paired-end*

Los archivos de mapeo generados en el paso anterior fueron ordenados respecto a la posición en la referencia, en pares de bases; para luego llevar a cabo el proceso de ensamblado propiamente dicho. El proceso fue igual a lo descrito previamente para los estadios embrionarios *single-end*. Se obtuvieron como archivos de salida, el correspondiente a las coordenadas de los transcriptos generados, el correspondiente a las predicciones génicas y el que hace referencia a las posibles isoformas, cada uno con sus respectivos valores FPKM. Este último valor, permitió identificar la abundancia de cada transcripto dentro de los transcriptomas generados.

Además, se estableció el número de transcriptos ensamblados. Se determinó un total de 12.301 transcriptos correspondientes al transcriptoma de ovario, obtenidos a partir del alineamiento con *Tophat2* y 15.151 transcriptos obtenidos a partir del alineamiento con *Bowtie1*. Para el transcriptoma correspondiente a huevos no fertilizados, se generaron 11.272 transcriptos a partir del mapeo llevado a cabo con *Tophat2* y 13.558 para el archivo de mapeo obtenido por *Bowtie1*. Por último, para el transcriptoma de huevos *Ohpo* se generaron 11.663 transcriptos a partir del archivo de mapeo generado con *Tophat2*. Según el algoritmo de mapeo utilizado, se ensamblaron transcriptos que fueron compartidos por ambos programas (mismo identificador sin importar la posición ni longitud de la predicción) o fueron programa específico. Para no perder ninguna inferencia a la hora de identificar a los transcriptos generados se decidió utilizar la herramienta *Cuffmerge* de *Cufflinks*, la que combinó parsimoniosamente los fragmentos ensamblados por *Cufflinks* y permitió generar un nuevo ensamblado “*merged*” con el menor número de transcriptos que explican de la mejor manera los datos totales (Tabla 5).

Transcriptoma	Método	Número de Transcritos reconstruidos	Número de transcritos Método - específico	Número de transcritos compartidos	Transcritos <i>Cuffmerge</i>
<b>Ovario</b>	<i>unspliced aligner</i>	15.151	10.843	9.788	14.522
	<i>spliced aligner</i>	12.301	13.693		
<b>Huevos no fertilizados</b>	<i>unspliced aligner</i>	13.558	12.512	9.310	13.082
	<i>spliced aligner</i>	11.272	10.226		

**Tabla 5:** Reconstrucción de los transcritos en cada transcriptoma secuenciado en *paired-end*. Se reporta el número de transcritos reconstruidos, el número que se generaron según método, los compartidos por ambos métodos y el número de transcritos reconstruidos mediante *Cuffmerge*.

Con el objetivo de descartar la presencia de remanentes de artefactos de la secuenciación, se los evaluó mediante búsqueda de similitud de secuencia contra la base de datos del UniVec utilizando el algoritmo BLASTX. Para todos los casos el resultado fue negativo, desechando la posibilidad de contar con un transcriptoma con secuencias no deseadas.

### 3. Anotación funcional de transcriptomas completos

Contando con los transcriptomas reconstruidos, representando distintos momentos a lo largo del desarrollo de *R. prolixus*, se decidió realizar un análisis de ontología con el fin de brindarle identidad a los transcritos que lo componen. Como material de partida se utilizaron los transcriptomas generados a partir del mapeo contra el genoma anotado de *R. prolixus* como referencia, exceptuando a los transcriptomas *0hpo* y *48hpo* obtenidos por secuenciación *single-end*, que fueron excluidos de la anotación debido a una limitación computacional.

Se utilizó el programa Blast2GO para realizar la estructuración de los datos. Con el objetivo de incrementar la especificidad de la anotación funcional en primer lugar se realizó una búsqueda de similitud de secuencia mediante la implementación del algoritmo BLASTX contra la base de datos no redundante (nr) disponible en el repositorio del NCBI. Luego esos *hits* obtenidos por la búsqueda de similitud fueron mapeados contra la información funcional almacenada en las bases de datos de ontología génica, en base a los ID adquiridos en la

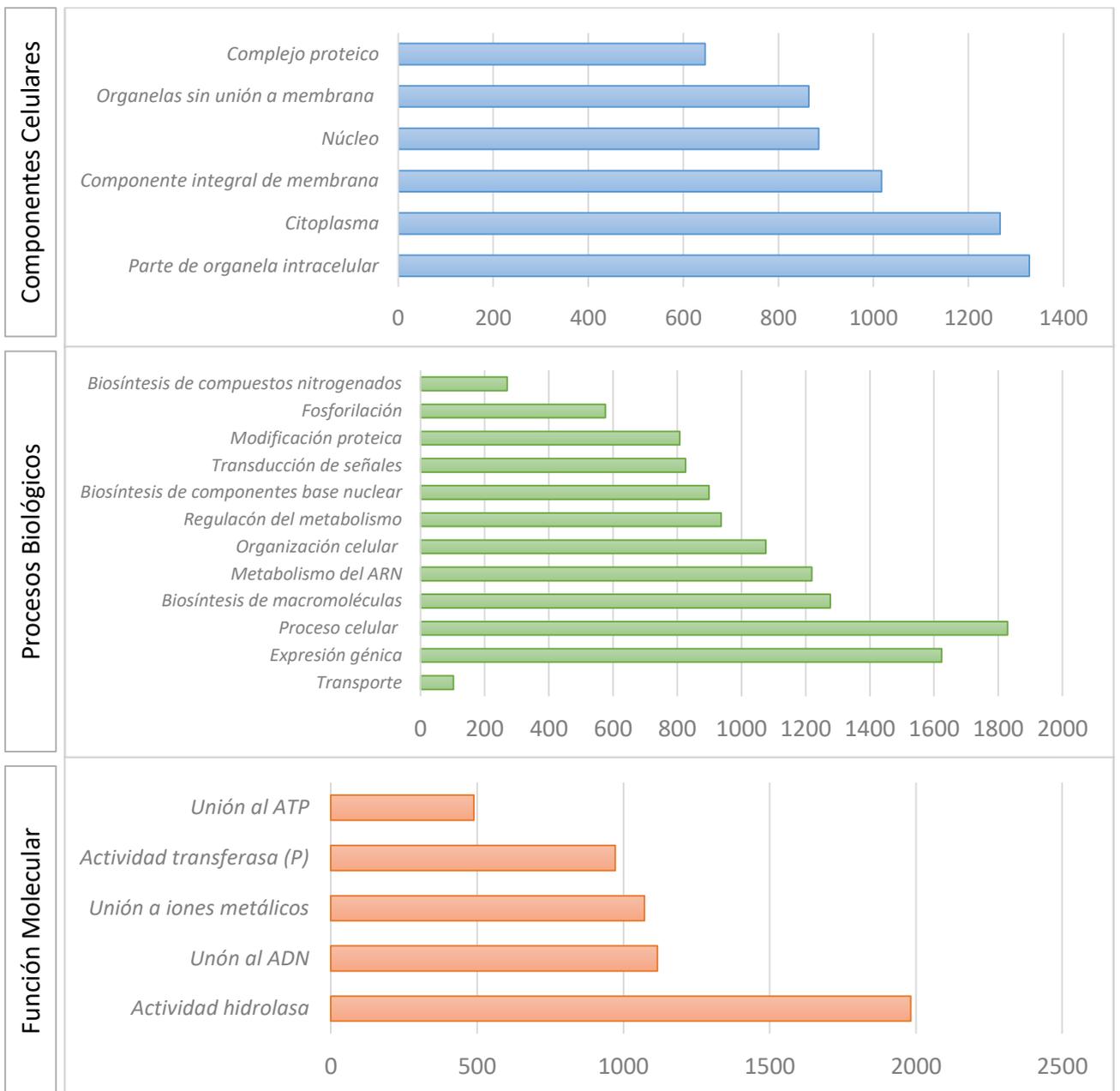
búsqueda de similitud. En dicha base de datos la información se presenta estructurada en tres categorías principales referentes a procesos biológicos, funciones moleculares y componentes celulares. Luego de efectuado el mapeo, se prosiguió a transferir la información de anotación a nuestras secuencias.

Como resultado de este análisis se obtuvo para cada uno de los transcritos distintos términos GO asociados, en donde se observó que un mismo transcripto fue representado en más de un término dentro de la misma categoría principal, e incluso en más de una de ellas. Por otro lado, no todos los transcritos reconstruidos en los transcriptomas generados mostraron identidad funcional asignada, esto se debió a que no todas las secuencias presentes dieron *hit* contra la base de datos nr y/o mapearon con la base de datos de ontología génica, por lo que no fueron incluidas en el proceso de anotación.

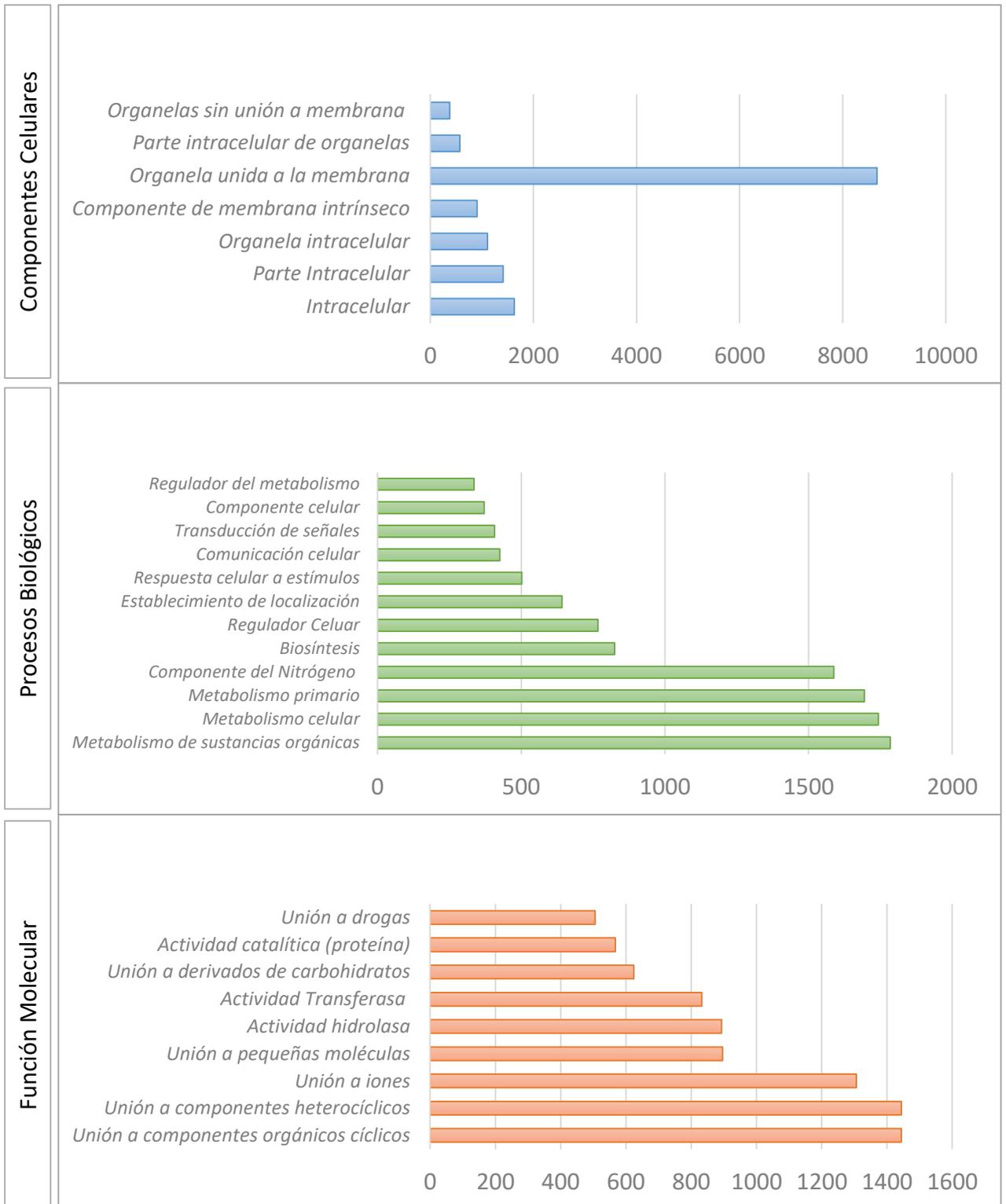
Los términos GO asociados obtenidos en este análisis se dividieron en base a la categoría principal asignada y que fueron representados en gráficos de barras que se muestran a continuación para cada uno de los transcriptomas secuenciados (Figuras 10-14).

Como resultado de la anotación funcional para el transcriptoma *0hpo* (Figura 10) se observó que aproximadamente 1.300 transcritos fueron identificados como componentes principales del citoplasma o como parte intracelular de organelas. Para la ontología referida a procesos biológicos se identificaron 1.800 transcritos en procesos biológicos y cerca de 1.600 identificados como expresión génica; y, se encontró a la actividad hidrolasa representada como la principal función molecular. En el transcriptoma *12hpo* (Figura 11) se identificaron aproximadamente 9.000 transcritos asociados a componentes de organelas unidos a la membrana. En cuanto a procesos biológicos, este estuvo principalmente representado por transcritos asociados al metabolismo celular, primario y de sustancias orgánicas. En la ontología función molecular se observó que aproximadamente 1.400 transcritos fueron anotados con el término referente a unión a componentes heterocíclicos y otros 1.400 transcritos como unión a componentes orgánicos cíclicos. En *24hpo* (Figura 12) se observó que las ontologías que hacen referencia a función molecular y procesos biológicos estuvieron representadas por las mismas anotaciones que se obtuvieron para el transcriptoma *12hpo*; y, en lo que respecta a componentes celulares, se observó que los transcritos reconstruidos se identificaron principalmente como componentes intracelulares. En el

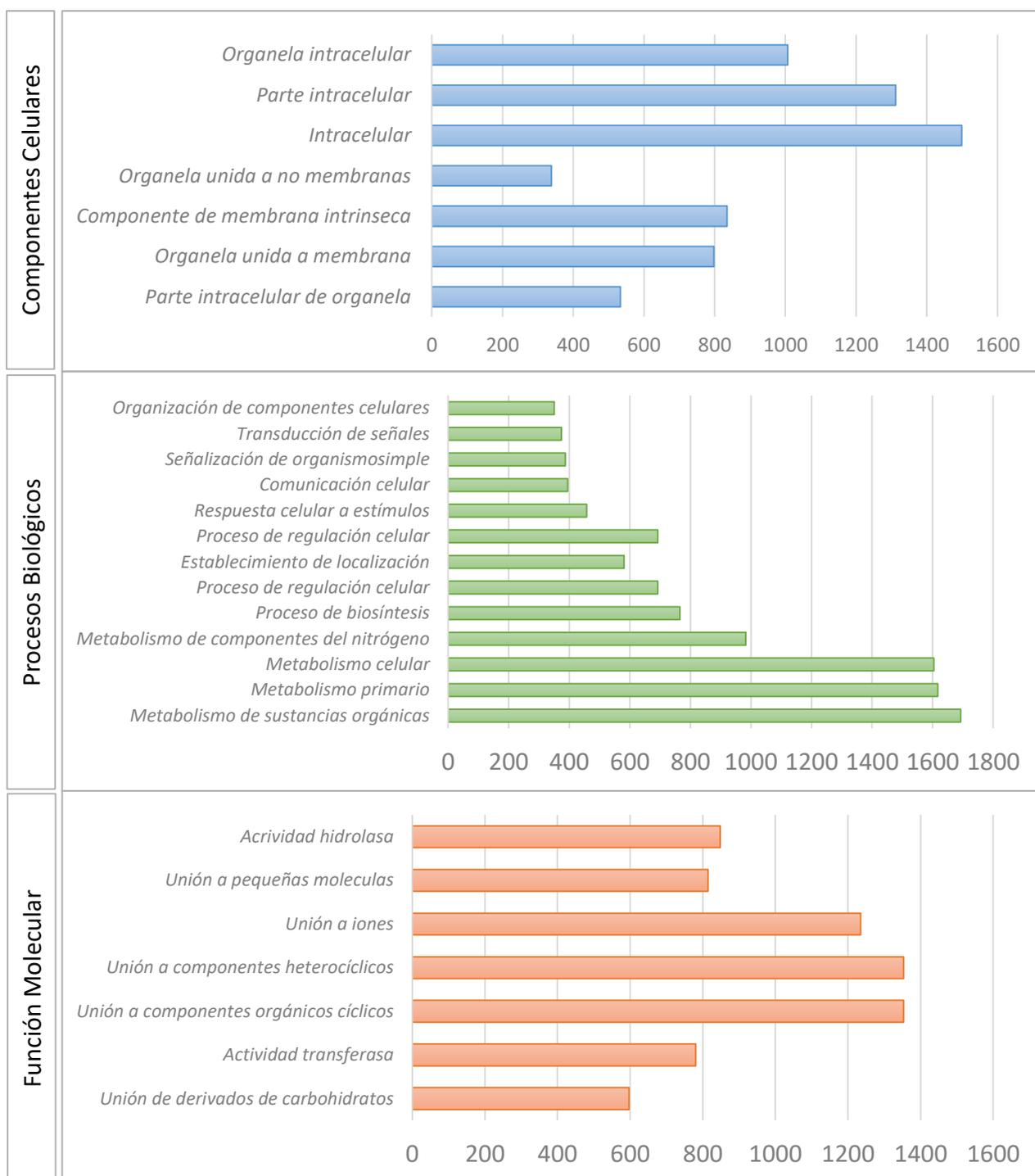
transcriptoma de huevos no fertilizados (Figura 13), se observó que aproximadamente 1.800 transcritos fueron identificados como componentes celulares o partes celulares. En la ontología referente a proceso biológico aproximadamente 2.300 transcritos se identificaron asociados a procesos celulares y 2.250 a procesos metabólicos; y, se encontró a la actividad hidrolasa representada como la principal función molecular. En el transcriptoma de ovario (Figura 14) se observó que la ontología componentes celulares estuvo principalmente representada por transcritos identificados como componentes intracelulares. En cuanto a la anotación correspondiente a la ontología de procesos biológicos, este estuvo principalmente representado por transcritos asociados al metabolismo celular, primario y de sustancias orgánicas, y en la ontología función molecular se observó que aproximadamente 225 transcritos fueron anotados con el término referente a unión a componentes heterocíclicos, otros 225 como unión a iones y otros 225 como unión a componentes orgánicos cíclicos. Observando el número de transcritos asociados a cada término GO, el transcriptoma de ovario fue el que presentó el menor número de transcritos anotados según ontología, en comparación a los restantes transcriptomas analizados.



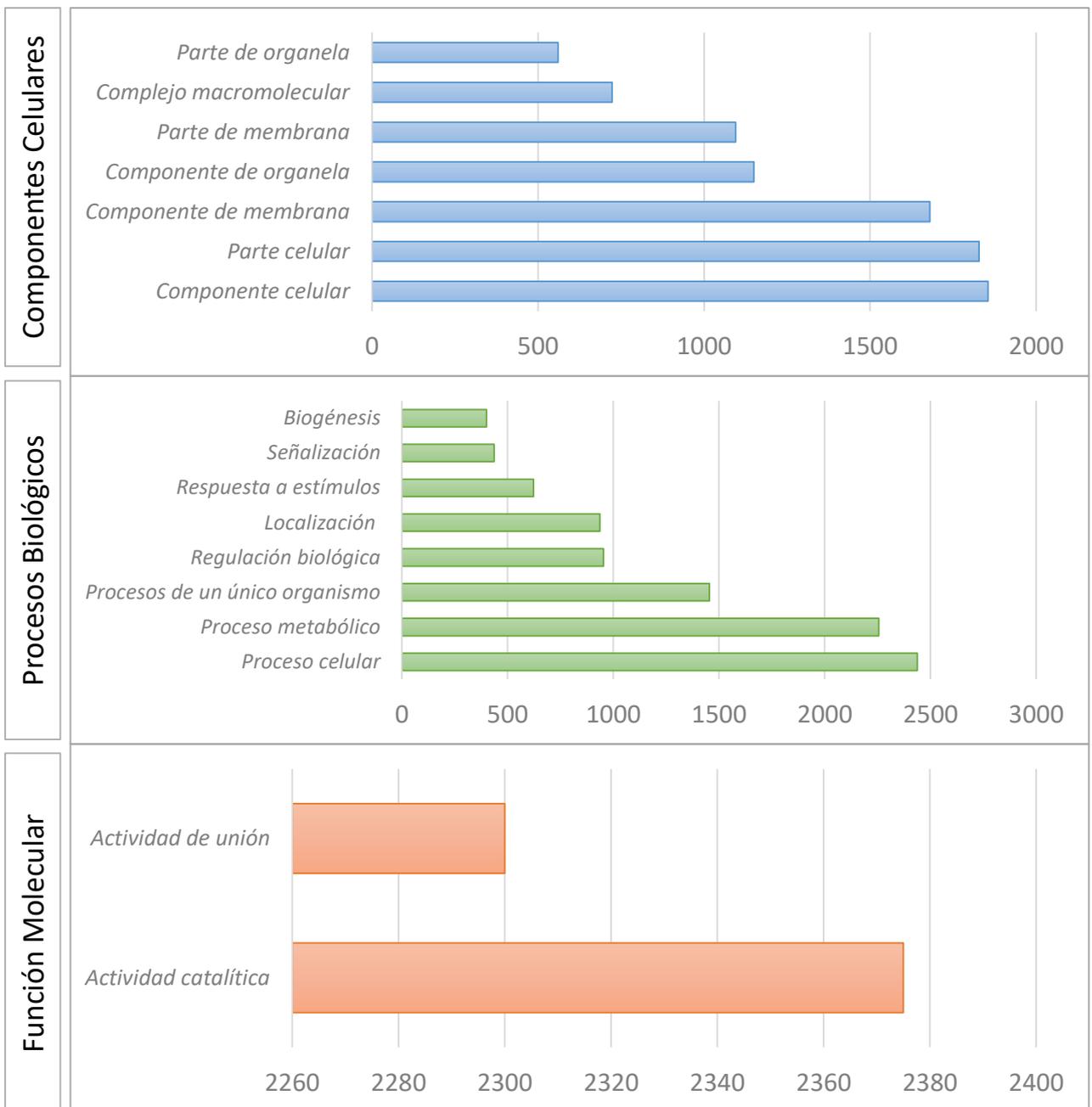
**Figura 10:** Gráfico de barras que representa la distribución de a anotación funcional correspondiente el transcriptoma de *Ohpo*. En el eje Y se muestran los términos (Ontología génica) GO y en el eje X el número de transcritos presentes en esa categoría. Ontologías: Componentes Celulares (celeste), Procesos Biológicos (verde), y Función Molecular (naranja).



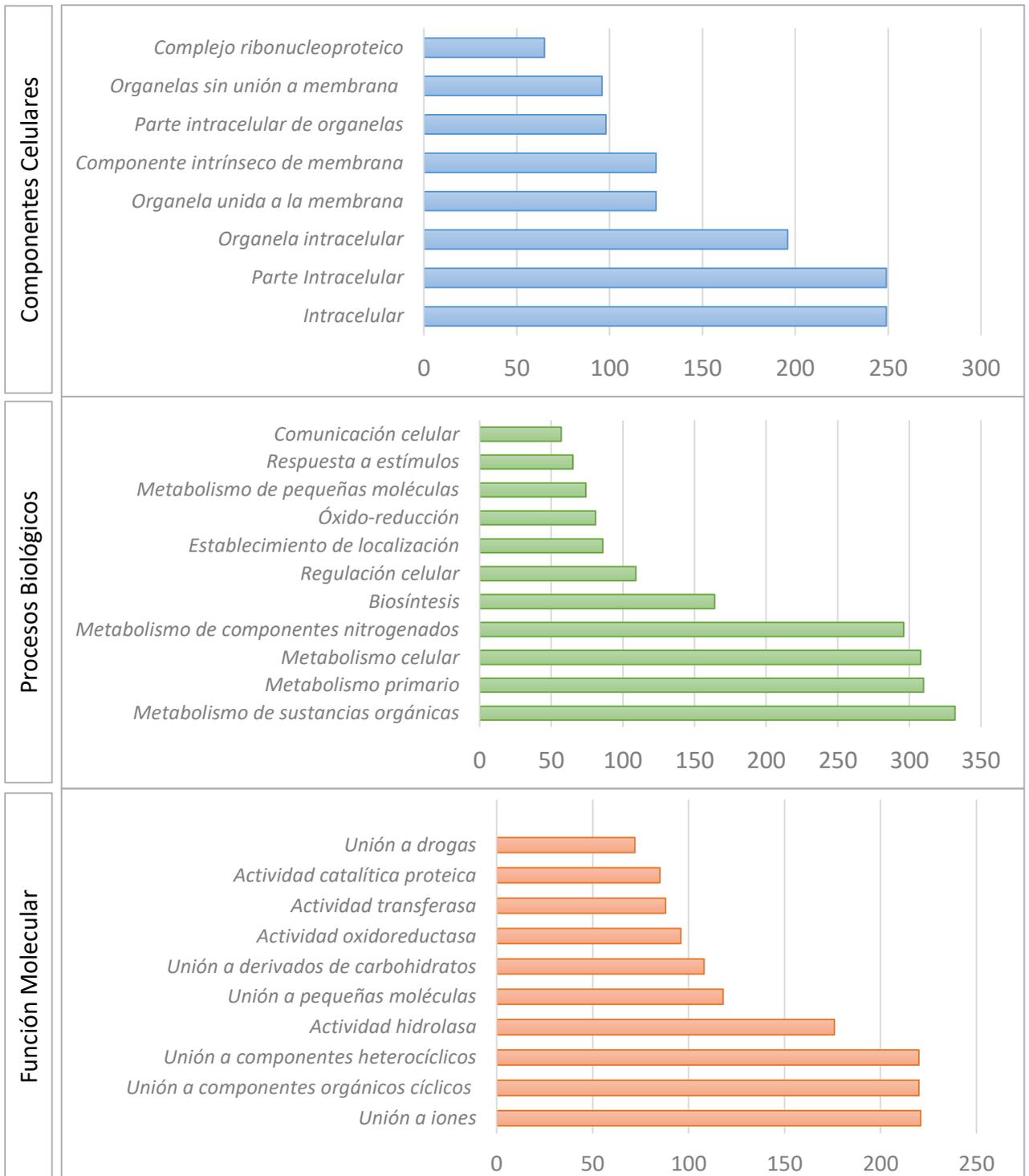
**Figura 11:** Gráfico de barras que representa la distribución de a anotación funcional correspondiente el transcriptoma de 12hpo. En el eje Y se muestran los términos (Ontología génica) GO y en el eje X el número de transcritos presentes en esa categoría. Ontologías: Componentes Celulares (celeste), Procesos Biológicos (verde), y Función Molecular (naranja).



**Figura 12:** Gráfico de barras que representa la distribución de a anotación funcional correspondiente el transcriptoma de 24hpo. En el eje Y se muestran los términos (Ontología génica) GO y en el eje X el número de transcritos presentes en esa categoría. Ontologías: Componentes Celulares (celeste), Procesos Biológicos (verde), y Función Molecular (naranja).



**Figura 13:** Gráfico de barras que representa la distribución de a anotación funcional correspondiente el transcriptoma de huevos no fertilizados. En el eje Y se muestran los términos (Ontología génica) GO y en el eje X el número de transcritos presentes en esa categoría. Ontologías: Componentes Celulares (celeste), Procesos Biológicos (verde), y Función Molecular (naranja).



**Figura 14:** Gráfico de barras que representa la distribución de a anotación funcional correspondiente el transcriptoma de ovario. En el eje Y se muestran los términos (Ontología génica) GO y en el eje X el número de transcritos presentes en esa categoría. Ontologías: Componentes Celulares (celeste), Procesos Biológicos (verde), y Función Molecular (naranja).

#### 4. Genes de desarrollo identificados a partir de búsqueda por similitud en los transcriptomas generados

Los transcriptomas ensamblados y analizados fueron comparados contra la base de datos generada a partir de los genes conocidos que se expresan durante la embriogénesis de *D. melanogaster*, con el objetivo de darle identidad mediante similitud de secuencias a los transcriptos. Como material de partida para realizar esta búsqueda se utilizaron los transcriptomas generados tras el mapeo al genoma anotado de *R. prolixus* como referencia; y como se mencionó previamente, para evitar pasar por alto alguna inferencia de identidad se utilizó el transcriptoma obtenido tras la utilización de *Cuffmerge*.

En el caso específico del tiempo *0hpo* como se utilizaron dos estrategias de secuenciación diferentes (*paired-end* y *single-end*) la generación del transcriptoma “maestro” se realizó para cada tipo de secuenciación realizada, luego como resultado final si se extrapolaron las identidades inferidas en uno u otro transcriptoma obtenido.

Como resultado se obtuvo un número considerablemente grande de secuencias identificadas en los transcriptomas de interés, observándose para la secuenciación *single-end* 206 *hits* en el transcriptoma *0hpo*, 2.979 *hits* en *12hpo*, 2.766 *hits* en *24hpo* y 3.424 *hits* para el transcriptoma de *48hpo*. En los generados por secuenciación *paired-end* se obtuvieron 3.307 *hits* para el transcriptoma de huevos no fertilizados, 3.200 *hits* para el transcriptoma de ovario, y 2.867 *hits* en el transcriptoma *0hpo*.

Como se mencionó en un inicio, debido a nuestro interés en los genes que se expresan en el desarrollo embrionario temprano, se decidió hacer foco en los que mostraron similitud de secuencia con genes que en *D. melanogaster* actúan durante la embriogénesis. Se los organizó según la función que cumplen en este organismo modelo, con el objetivo de poder inferir la presencia del gen, en un tiempo del desarrollo específico. Para esto se generaron tres grupos correspondientes a genes de segmentación, genes del grupo dorsal y genes relacionados con la oogénesis. El número total de genes analizados fue 74, de los cuales 38 componen los genes de segmentación, 16 los del grupo dorsal y 20 para los relacionados a la oogénesis. En las tablas 6-11 se muestran las identidades correspondientes a los *hits* obtenidos tanto para la secuenciación *single-end* como *paired-end*.

Genes de Segmentación	Gen ID	0hpo			12hpo			24hpo			48hpo		
		% ID	Longitud /hit	E-value	% ID	Longitud /hit	E-value	% ID	Longitud /hit	E-value	% ID	Longitud /hit	E-value
<i>Nanos</i>	RPRC002927	*	*	*	*	*	*	50.00	48	4,00E-09	48.00	50	4,00E-09
<i>Hunchback</i>	RPRC000230	*	*	*	66.67	159	5,00E-74	66.67	159	5,00E-74	66.67	159	7,00E-74
<i>Caudal</i>	RPRC000239	*	*	*	75.53	94	9,00E-42	75.53	94	5,00E-42	75.53	94	7,00E-42
<i>Kruppel</i>	RPRC000102	*	*	*	*	*	*	84.38	96	2,00E-56	79.70	133	2,00E-73
<i>Knirps</i>	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Giant</i>	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Tailless</i>	RPRC007025	*	*	*	80.00	80	1,00E-43	30.72	153	6,00E-13	43.03	416	2,00E-101
<i>Huckebein</i>	RPRC007615	*	*	*	*	*	*	50.52	97	2,00E-24	68.81	109	5,00E-53
<i>Empty spiracles</i>	RPRC010410	*	*	*	*	*	*	*	*	*	*	*	*
	RPRC014392	*	*	*	*	*	*	*	*	*	*	*	*
<i>Knot</i>	RPRC004534	*	*	*	96.39	83	3,00E-52	74.12	85	2,00E-37	96.25	80	7,00E-50
	RPRC013805	*	*	*	*	*	*	*	*	*	*	*	*
<i>Cap-n-collar</i>	RPRC011620	*	*	*	42.64	387	8,00E-65	42.64	387	6,00E-65	42.64	387	8,00E-65
<i>Tenascin major</i>	RPRC005422	*	*	*	*	*	*	64.64	280	9,00E-111	69.91	2316	0.0
<i>Hairy</i>	RPRC000496	*	*	*	69.35	124	1,00E-49	69.35	124	1,00E-49	68.46	130	8,00E-50
<i>Odd skipped</i>	RPRC011812	*	*	*	*	*	*	63.13	179	4,00E-76	63.13	179	3,00E-76
<i>Odd paired</i>	RPRC013047	*	*	*	*	*	*	*	*	*	*	*	*
<i>Runt</i>	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Sloppy paired 1</i>	RPRC000987	*	*	*	*	*	*	77.53	89	2,00E-47	77.53	89	2,00E-47
	RPRC003585	*	*	*	82.33	815	0.0	82.33	815	0.0	82.33	815	0.0
<i>Armadillo</i>	RPRC003617	*	*	*	*	*	*	*	*	*	*	*	*
<i>Decapentaplegic</i>	RPRC000401	*	*	*	77.66	94	3,00E-51	49.11	224	2,00E-65	46.01	363	7,00E-102
<i>Wingless</i>	RPRC005904	*	*	*	68.52	305	1,00E-136	70.75	294	1,00E-136	63.96	333	1,00E-136
<i>Cubitus interruptus</i>	RPRC007162	*	*	*	42.67	375	3,00E-62	44.32	88	8,00E-13	42.67	375	3,00E-62
<i>Hedgehog</i>	RPRC012384	*	*	*	49.21	315	2,00E-92	49.20	313	2,00E-91	49.21	315	2,00E-92
<i>Fused</i>	RPRC003744	*	*	*	52.45	286	5,00E-106	53.05	279	6,00E-105	52.65	283	8,00E-105
<i>Smoothed</i>	RPRC004966	*	*	*	52.40	626	0.0	52.40	626	0.0	52.40	626	0.0
<i>Engrailed</i>	RPRC003110	*	*	*	*	*	*	*	*	*	76.64	107	4,00E-50
	RPRC000282	*	*	*	*	*	*	*	*	*	81.63	98	4,00E-49
<i>Rasp</i>	RPRC014615	*	*	*	28.83	437	2,00E-50	28.83	437	2,00E-50	35.47	234	1,00E-34
<i>Pangolin</i>	RPRC010782	*	*	*	95.83	72	2,00E-43	95.83	72	6,00E-45	95.83	72	4,00E-44
<i>Gooseberry</i>	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Tout-velu</i>	RPRC000548	*	*	*	61.70	329	3,00E-131	61.35	326	3,00E-129	61.59	328	1,00E-130
<i>labial</i>	RPRC000004	*	*	*	*	*	*	*	*	*	*	*	*
<i>Deformed</i>	RPRC000437	*	*	*	*	*	*	*	*	*	81.98	111	1,00E-56
<i>Ultrathorax</i>	RPRC000565	*	*	*	*	*	*	*	*	*	100.00	53	4,00E-34
<i>Abdominal-B</i>	RPRC012974	*	*	*	58.82	68	4,00E-20	58.82	68	9,00E-20	*	*	*
<i>Proboscipedia</i>	RPRC002128	*	*	*	52.88	208	3,00E-48	52.88	208	3,00E-48	*	*	*
<i>Antennapedia</i>	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Abdominal-A</i>	RPRC000598	*	*	*	*	*	*	*	*	*	*	*	*
	RPRC006513	*	*	*	43.84	333	2,00E-92	43.84	333	3,00E-92	43.67	332	2,00E-91
<i>Torso-like</i>	RPRC006506	*	*	*	27.12	118	4,00E-08	*	*	*	*	*	*
<i>Pipsqueak</i>	RPRC012968	*	*	*	82.29	192	7,00E-111	82.11	190	2,00E-109	82.29	192	7,00E-111

Tabla 6: Genes involucrados con el proceso de segmentación, identificados en los transcriptomas embrionarios *single-end*. (\*) genes no identificados utilizando BLAST

Genes del grupo dorsal	Gen ID	olipo			12hpo			24hpo			48hpo		
		% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value
<i>Cappuccino</i>	RPRC003459	*	*	*	49.09	330	1,00E-99	49.20	313	2,00E-94	49.09	330	1,00E-99
	RPRC014581	*	*	*	85.23	88	2,00E-41	85.23	88	2,00E-41	85.23	88	3,00E-41
<i>Oo18 RNA-binding protein</i>	RPRC001021	*	*	*	43.98	191	9,00E-46	43.98	191	7,00E-46	43.98	191	9,00E-46
	RPRC001974	*	*	*	42.19	128	4,00E-22	42.19	128	4,00E-22	42.19	128	4,00E-22
<i>Spire</i>	RPRC001975	*	*	*	61.03	136	2,00E-48	61.03	136	2,00E-48	44.22	251	9,00E-49
	RPRC007924	81.63	49	1,00E-21	*	*	*	*	*	*	*	*	*
<i>Squid</i>	RPRC008474	*	*	*	58.52	229	8,00E-53	58.52	229	7,00E-53	58.52	229	8,00E-53
	RPRC000049	*	*	*	34.60	289	3,00E-40	34.60	289	3,00E-40	44.44	288	1,00E-74
<i>Nudel</i>	RPRC005278	*	*	*	*	*	*	*	*	*	47.66	128	2,00E-33
	RPRC007848	35.20	250	2,00E-44	35.20	250	4,00E-44	35.20	250	4,00E-44	35.20	250	4,00E-44
<i>Windbeutel</i>	RPRC003090	*	*	*	33.43	356	2,00E-50	33.43	356	3,00E-50	33.06	366	1,00E-50
	RPRC001191	*	*	*	31.48	270	2,00E-38	31.48	270	2,00E-38	31.48	270	2,00E-38
<i>Gastrulation defective</i>	RPRC002634	*	*	*	37.36	91	1,00E-17	37.36	88	9,00E-17	37.36	91	1,00E-17
	RPRC013110	*	*	*	25.64	156	2,00E-11	29.25	106	1,00E-10	25.64	156	2,00E-11
<i>Spätzle</i>	RPRC001090	*	*	*	53.06	196	2,00E-45	53.06	196	3,00E-48	53.06	196	2,00E-45
	RPRC017349	*	*	*	45.21	219	4,00E-49	45.21	219	4,00E-49	45.21	219	3,00E-49
<i>Capicua</i>	RPRC003790	*	*	*	76.08	209	7,00E-117	76.08	209	7,00E-117	76.08	209	7,00E-117
	RPRC009262	*	*	*	38.52	976	0.0	38.52	980	0.0	38.52	980	0.0
<i>Dorsal</i>	RPRC003790	*	*	*	76.08	209	7,00E-117	76.08	209	7,00E-117	76.08	209	7,00E-117
	RPRC009262	*	*	*	38.52	976	0.0	38.52	980	0.0	38.52	980	0.0
<i>Toll</i>	RPRC003790	*	*	*	76.08	209	7,00E-117	76.08	209	7,00E-117	76.08	209	7,00E-117
	RPRC009262	*	*	*	38.52	976	0.0	38.52	980	0.0	38.52	980	0.0

Tabla 7: Genes involucrados con el proceso de establecimiento del eje dorso ventral, identificados en los transcriptomas embrionarios *signle-end*. (\*) genes no identificados utilizando BLAST

Genes de oogenesis y ovario	Gen ID	olipo			12hpo			24hpo			48hpo		
		% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value
<i>Bazooka</i>	RPRC001227	*	*	*	45.49	488	4,00E-97	50.16	319	2,00E-74	45.49	488	4,00E-97
	RPRC000632	*	*	*	220	80	3,00E-79	61.36	220	2,00E-79	61.36	220	3,00E-79
<i>Bicaudal D</i>	RPRC000704	*	*	*	111	31	2,00E-39	72.07	111	2,00E-39	72.07	111	2,00E-39
	RPRC004076	*	*	*	257	100	8,00E-71	60.08	253	5,00E-69	59.92	257	8,00E-71
<i>Bicaudal C</i>	RPRC012518	*	*	*	50.00	66	3,00E-16	53.23	62	5,00E-16	50.00	66	3,00E-16
	RPRC001612	*	*	*	27.38	168	2,00E-15	27.38	168	2,00E-15	27.38	168	2,00E-15
<i>Egalitarian</i>	RPRC001613	*	*	*	51.14	395	9,00E-134	51.01	398	2,00E-134	51.14	395	9,00E-134
	RPRC012064	*	*	*	45.21	564	1,00E-128	45.21	564	6,00E-129	41.67	564	1,00E-116
<i>Egghead</i>	RPRC011094	*	*	*	73.01	452	0.0	73.01	452	0.0	73.01	452	0.0
	RPRC001415	*	*	*	42.33	163	4,00E-39	42.33	163	4,00E-39	42.33	163	4,00E-39
<i>Exuperantia</i>	RPRC012896	*	*	*	24.69	733	1,00E-34	23.97	317	5,00E-21	24.69	733	1,00E-34
	RPRC011695	*	*	*	88.89	108	6,00E-61	88.89	108	5,00E-61	88.89	108	4,00E-61
<i>Pointed</i>	RPRC006015	*	*	*	80.58	139	9,00E-76	80.29	137	4,00E-74	80.58	139	9,00E-76
	RPRC010019	*	*	*	60.99	141	7,00E-43	60.99	141	1,00E-42	60.99	141	4,00E-43
<i>Singed</i>	RPRC010985	91.03	145	2,00E-97	91.03	145	3,00E-97	91.03	145	3,00E-97	91.03	145	3,00E-97
	RPRC001891	*	*	*	823	407	0.0	42.95	752	0.0	42.95	759	0.0
<i>Staufen</i>	RPRC002460	*	*	*	659	359	2,00E-160	38.30	658	2,00E-160	38.24	659	2,00E-160
	RPRC013054	*	*	*	773	473	3,00E-163	35.83	773	3,00E-163	35.83	773	3,00E-163
<i>Mago nashi</i>	RPRC000252	*	*	*	*	*	*	*	*	*	27.03	111	2,00E-10
	RPRC000720	*	*	*	83.86	254	3,00E-150	83.79	253	2,00E-149	84.13	252	4,00E-149
<i>Pumilio</i>	RPRC005018	*	*	*	49.51	103	3,00E-21	49.51	103	3,00E-21	49.51	103	3,00E-21
	RPRC009661	63.29	79	4,00E-28	60.43	230	7,00E-93	61.23	227	1,00E-92	60.43	230	7,00E-93
<i>Vasa</i>	RPRC009663	51.06	47	3,00E-11	55.30	132	2,00E-46	55.30	132	2,00E-46	55.30	132	2,00E-46
	RPRC000606	*	*	*	27.34	289	2,00E-24	27.34	289	1,00E-24	27.34	289	5,00E-24
<i>Maelstrom</i>	RPRC002781	91.84	98	3,00E-60	85.90	156	2,00E-92	85.90	156	2,00E-92	85.90	156	2,00E-92
	RPRC014787	*	*	*	67.91	215	1,00E-95	67.91	215	1,00E-95	67.91	215	1,00E-95
<i>Ovo</i>	RPRC002781	91.84	98	3,00E-60	85.90	156	2,00E-92	85.90	156	2,00E-92	85.90	156	2,00E-92
	RPRC014787	*	*	*	67.91	215	1,00E-95	67.91	215	1,00E-95	67.91	215	1,00E-95
<i>Licorne</i>	RPRC002781	91.84	98	3,00E-60	85.90	156	2,00E-92	85.90	156	2,00E-92	85.90	156	2,00E-92
	RPRC014787	*	*	*	67.91	215	1,00E-95	67.91	215	1,00E-95	67.91	215	1,00E-95
<i>Maternal expression at 31B</i>	RPRC002781	91.84	98	3,00E-60	85.90	156	2,00E-92	85.90	156	2,00E-92	85.90	156	2,00E-92
	RPRC014787	*	*	*	67.91	215	1,00E-95	67.91	215	1,00E-95	67.91	215	1,00E-95
<i>Cyclin B</i>	RPRC000215	45.09	326	8,00E-68	*	*	*	*	*	*	*	*	*
	RPRC000215	45.09	326	8,00E-68	*	*	*	*	*	*	47.24	434	2,00E-102
<i>Armitage</i>	RPRC000215	45.09	326	8,00E-68	*	*	*	*	*	*	*	*	*
	RPRC000215	45.09	326	8,00E-68	*	*	*	*	*	*	47.24	434	2,00E-102

Tabla 8: Genes involucrados con el proceso de oogenesis, identificados en los transcriptomas embrionarios *signle-end*. (\*) genes no identificados utilizando BLAST

Genes de Segmentación	Gen ID	Ovario			No fertilizados			Ohpo		
		% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value
<i>Nanos</i>	RPRC002927	*	*	*	*	*	*	50.00	48	4,00E-09
<i>Hunchback</i>	RPRC000230	66.67	159	4,00E-74	66.67	159	6,00E-74	*	*	*
<i>Caudal</i>	RPRC000239	71.70	53	5,00E-19	75.53	94	8,00E-42	75.53	94	3,00E-42
<i>Kruppel</i>	RPRC000102	*	*	*	79.70	133	2,00E-73	*	*	*
<i>Knirps</i>	*	*	*	*	*	*	*	*	*	*
<i>Giant</i>	*	*	*	*	*	*	*	*	*	*
<i>Tailless</i>	RPRC007025	67.50	80	8,00E-35	42.79	416	445	*	*	*
<i>Huckebein</i>	RPRC007615	50.52	97	2,00E-24	50.52	97	2,00E-24	*	*	*
<i>Empty spiracles</i>	RPRC010410	86.54	52	8,00E-25	*	*	*	*	*	*
	RPRC014392	86.96	23	2,00E-08	87.50	24	3,00E-09	*	*	*
<i>Knot</i>	RPRC004534	70.77	65	1,00E-23	96.34	82	1,00E-51	*	*	*
	RPRC013805	95.83	72	4,00E-43	96.05	76	5,00E-46	*	*	*
<i>Cap-n-collar</i>	RPRC011620	42.64	387	2,00E-65	42.64	387	9,00E-65	58.06	93	2,00E-28
<i>Tenascin major</i>	RPRC005422	63.10	542	0.0	70.04	2333	0.0	40.58	69	2,00E-09
<i>Hairy</i>	RPRC000496	68.46	130	1,00E-50	68.46	130	8,00E-50	68.46	130	8,00E-50
<i>Odd skipped</i>	RPRC011812	*	*	*	63.13	179	3,00E-76	63.13	179	4,00E-76
<i>Odd paired</i>	RPRC013047	*	*	*	68.70	131	3,00E-53	60.61	99	9,00E-32
<i>Runt</i>	*	*	*	*	*	*	*	*	*	*
<i>Sloppy paired 1</i>	RPRC000987	*	*	*	77.53	89	2,00E-47	77.53	89	1,00E-47
<i>Armadillo</i>	RPRC003585	82.33	815	0.0	82.33	815	0.0	82.31	814	0.0
	RPRC003617	76.16	843	0.0	76.16	843	0.0	76.16	843	0.0
<i>Decapentaplegic</i>	RPRC000401	46.01	363	6,00E-102	46.01	363	6,00E-102	46.01	363	6,00E-102
<i>Wingless</i>	RPRC005904	67.73	220	3,00E-94	63.69	336	1,00E-136	*	*	*
<i>Cubitus interruptus</i>	RPRC007162	42.67	375	1,00E-62	42.67	375	3,00E-62	70.63	126	1,00E-47
<i>Hedgehog</i>	RPRC012384	47.81	297	2,00E-81	49.36	314	3,00E-92	49.21	315	2,00E-92
<i>Fused</i>	RPRC003744	52.43	288	2,00E-106	52.43	288	9,00E-107	52.45	286	5,00E-106
<i>Smoothed</i>	RPRC004966	52.40	626	0.0	52.40	626	0.0	52.40	626	0.0
<i>Engrailed</i>	RPRC003110	*	*	*	*	*	*	*	*	*
	RPRC000282	*	*	*	85.39	89	2,00E-48	85.06	87	5,00E-47
<i>Rasp</i>	RPRC014615	28.83	437	2,00E-50	28.83	437	2,00E-50	28.77	438	2,00E-50
<i>Pangolin</i>	RPRC010782	74.56	114	1,00E-44	74.56	114	7,00E-45	*	*	*
<i>Gooseberry</i>	*	*	*	*	*	*	*	*	*	*
<i>Tout-velu</i>	RPRC000548	61.59	328	1,00E-130	61.82	330	3,00E-132	65.55	119	4,00E-43
<i>labial</i>	RPRC000004	*	*	*	*	*	*	88.24	17	9,00E-06
<i>Deformed</i>	RPRC000437	*	*	*	83.12	77	9,00E-38	*	*	*
<i>Ultrabithorax</i>	RPRC000565	100.00	54	7,00E-35	100.00	54	9,00E-35	*	*	*
<i>Abdominal-B</i>	RPRC012974	58.82	68	4,00E-20	58.82	68	5,00E-20	*	*	*
<i>Proboscipedia</i>	RPRC002128	52.88	208	3,00E-48	52.88	208	4,00E-48	*	*	*
<i>Antennapedia</i>	*	*	*	*	*	*	*	*	*	*
<i>Abdominal-A</i>	RPRC000598	*	*	*	100.00	62	5,00E-39	*	*	*
	RPRC006513	43.84	333	2,00E-92	43.84	333	2,00E-92	43.33	330	5,00E-90
<i>Torso-like</i>	RPRC006506	*	*	*	*	*	*	27.91	129	4,00E-09
<i>Pipsqueak</i>	RPRC012968	82.11	190	6,00E-107	81.96	194	4,00E-112	81.42	183	3,00E-103

Tabla 9: Genes involucrados con el proceso de segmentación, identificados en los transcriptomas embrionarios *paired-end*. (\*) genes no identificados utilizando BLAST

Genes del grupo dorsal	Gen ID	Ovario			No fertilizados			Ohpo		
		% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value
<i>Cappuccino</i>	RPRC003459	49.24	331	3,00E-100	49.24	331	3,00E-100	49.12	283	4,00E-85
<i>Cornichon</i>	RPRC014581	82.61	92	4,00E-42	85.23	88	2,00E-41	86.05	86	1,00E-40
<i>Oo18 RNA-binding protein</i>	RPRC001021	45.11	184	1,00E-45	43.98	191	9,00E-46	43.98	191	7,00E-46
<i>Spire</i>	RPRC001974	42.19	128	4,00E-22	42.19	128	4,00E-22	42.52	127	8,00E-22
	RPRC001975	44.22	251	1,00E-48	44.22	251	9,00E-49	49.02	204	8,00E-49
<i>Squid</i>	RPRC007924	*	*	*	*	*	*	*	*	*
<i>Rhomboid</i>	RPRC008474	62.00	150	6,00E-31	58.52	229	8,00E-53			
<i>Nudel</i>	RPRC000049	34.51	284	8,00E-41	37.77	564	3,00E-111	48.18	137	1,00E-34
<i>Pipe</i>	RPRC005278	50.54	93	2,00E-27	*	*	*	*	*	*
<i>Windbeutel</i>	RPRC007848	35.20	250	4,00E-44	35.20	250	3,00E-44	35.20	250	3,00E-44
<i>Easter</i>	RPRC003090	33.06	366	6,00E-51	33.06	366	1,00E-50	33.06	366	4,00E-51
<i>Gastrulation defective</i>	RPRC001191	31.48	270	2,00E-38	31.48	270	2,00E-38	31.48	270	2,00E-39
<i>Spätzle</i>	RPRC002634	37.36	91	1,00E-17	37.36	91	1,00E-17	37.78	90	1,00E-17
	RPRC013110	*	*	*	25.64	156	2,00E-11	29.81	104	6,00E-11
<i>Capicua</i>	RPRC001090	53.06	196	2,00E-45	53.06	196	2,00E-45	84.21	19	4,00E-07
<i>Cactus</i>	RPRC017349	45.21	219	3,00E-49	45.21	219	4,00E-49	45.21	219	5,00E-49
<i>Dorsal</i>	RPRC003790	76.08	209	1,00E-116	76.08	209	7,00E-117	77.11	201	4,00E-113
<i>Toll</i>	RPRC009262	38.57	980	0.0	38.57	980	0.0	38.57	980	0.0

Tabla 10: Genes involucrados con el proceso de establecimiento del eje dorso ventral, identificados en los transcriptomas embrionarios *paired-end*. (\*) genes no identificados utilizando BLAST

Genes de oogenesis y ovario	Gen ID	Ovario			No fertilizados			Ohpo		
		% ID	Longitud hit	E-value	% ID	Longitud hit	E-value	% ID	Longitud hit	E-value
<i>Bazooka</i>	RPRC001227	45.49	488	3,00E-97	45.49	488	4,00E-97	37.17	269	4,00E-25
	RPRC000632	61.36	220	6,00E-79	61.36	220	6,00E-79	63.59	206	2,00E-79
<i>Bicaudal D</i>	RPRC000704	72.07	111	2,00E-39	72.07	111	2,00E-39	72.22	90	1,00E-30
	RPRC004076	59.92	257	7,00E-71	59.92	257	7,00E-71	61.80	178	9,00E-49
	RPRC012518	50.00	66	3,00E-16	50.00	66	3,00E-16	55.56	54	8,00E-14
<i>Bicaudal C</i>	RPRC001612	27.38	168	2,00E-15	27.38	168	2,00E-15	*	*	*
	RPRC001613	51.01	398	2,00E-134	51.01	398	2,00E-134	51.15	393	3,00E-133
<i>Egglitarian</i>	RPRC012064	45.21	564	1,00E-128	45.21	564	1,00E-128	51.84	380	2,00E-111
<i>Exuperantia</i>	RPRC011094	73.60	447	0.0	73.01	452	0.0	73.17	451	0.0
<i>Tudor</i>	RPRC001415	42.33	163	4,00E-39	42.33	163	4,00E-39	42.33	163	4,00E-39
<i>Pointed</i>	RPRC012896	24.69	733	1,00E-34	24.69	733	1,00E-34	24.69	733	8,00E-35
<i>Singed</i>	RPRC011695	88.89	108	4,00E-61	88.89	108	3,00E-61	88.89	108	4,00E-61
<i>Staufen</i>	RPRC006015	80.29	137	4,00E-74	80.58	139	9,00E-76	80.00	135	1,00E-72
<i>Mago nashi</i>	RPRC010019	60.99	141	1,00E-42	60.99	141	1,00E-42	60.99	141	4,00E-43
	RPRC010985	91.03	145	3,00E-97	91.03	145	3,00E-97	91.03	145	3,00E-97
	RPRC001891	39.90	822	0.0	43.03	760	0.0	42.86	749	0.0
	RPRC002460	38.24	659	2,00E-160	38.24	659	2,00E-160	38.30	658	2,00E-160
	RPRC013054	35.83	773	3,00E-163	35.83	773	1,00E-163	33.33	378	6,00E-71
	RPRC000252	*	*	*	*	*	*	35.87	315	6,00E-69
<i>Pumilio</i>	RPRC000720	83.86	254	3,00E-150	83.86	254	3,00E-150	84.06	251	4,00E-146
	RPRC005018	49.51	103	4,00E-21	49.51	103	4,00E-21	49.51	103	2,00E-21
<i>Vasa</i>	RPRC009661	60.43	230	7,00E-93	60.43	230	7,00E-93	60.47	215	1,00E-84
	RPRC009663	54.96	131	9,00E-46	55.30	132	2,00E-46	54.96	131	9,00E-46
<i>Maelstrom</i>	RPRC000606	27.34	289	5,00E-24	27.34	289	5,00E-24	27.34	289	4,00E-24
<i>Ovo</i>	RPRC002781	85.90	156	1,00E-92	85.90	156	2,00E-92	85.92	142	2,00E-84
<i>Licorne</i>	RPRC014787	67.91	215	1,00E-95	67.91	215	1,00E-95	67.76	214	4,00E-95
<i>Cyclin B</i>	*	*	*	*	*	*	*	*	*	*
<i>Armitage</i>	RPRC000215	*	*	*	41.67	132	5,00E-21	46.60	427	2,00E-104

Tabla 11: Genes involucrados con el proceso de oogenesis, identificados en los transcriptomas embrionarios *paired-end*. (\*) genes no identificados utilizando BLAST

Como era de esperar, debido a la baja cobertura de secuenciación que presentó el transcriptoma *Ohpo single-end* la cantidad de identidades inferidas fue muy bajas, observándose 5 genes representados en este tiempo secuenciado de un total de 74 genes analizados. Con la secuenciación *paired-end* de este mismo tiempo, se concluyó que la baja cantidad de *hits* obtenidos se debió al bajo número de transcritos reconstruidos tras la secuenciación.

Por otra parte, cabría esperar que genes con características relacionadas a la oogénesis no se encuentren presentes en huevos de tiempos referentes a gastrulación y formación de banda germinal (*48hpo*), así como tampoco que genes cigóticos se encuentren presentes en huevos no fertilizados. Esto no fue lo que sucedió en la identificación génica de los transcriptomas analizados. Se obtuvo una representación prácticamente completa en todos los tiempos estudiados, no permitiendo discriminar presencia o ausencia de expresión génica según momento del desarrollo específico.

Luego de la identificación génica mediante similitud de secuencia, se decidió realizar una búsqueda puntual de cada uno de ellos de manera manual, esto permitió corroborar las identidades asignadas, así como también la ausencia de los genes no identificados. Para corroborar la identificación, cada uno de los transcritos fueron consultados de manera individual mediante una nueva búsqueda por similitud de secuencia, en este caso, contra la base de datos nr proteica disponible *online* en NCBI.

Esta búsqueda no sólo permitió corroborar la identidad de cada transcripto, sino que también ayudó a resolver las ambigüedades en las cuales diferentes transcritos presentaron similitud con un mismo gen. En algunos casos estos correspondieron a falsos positivos, por lo que fueron descartados de la lista de genes identificados; y en otros, fue posible identificar genes que presentaron más de un transcripto en su inferencia. Lo que se pudo ver es que en la mayoría de los casos los identificadores de cada uno de las unidades transcripcionales que componen a la referencia, y que dan *hit* con un mismo gen, corresponden a fragmentos de un mismo transcripto. Esto se debe a que las predicciones en VectorBase son consideradas como dos unidades transcripcionales distintas y se le asignaron diferentes identificadores.

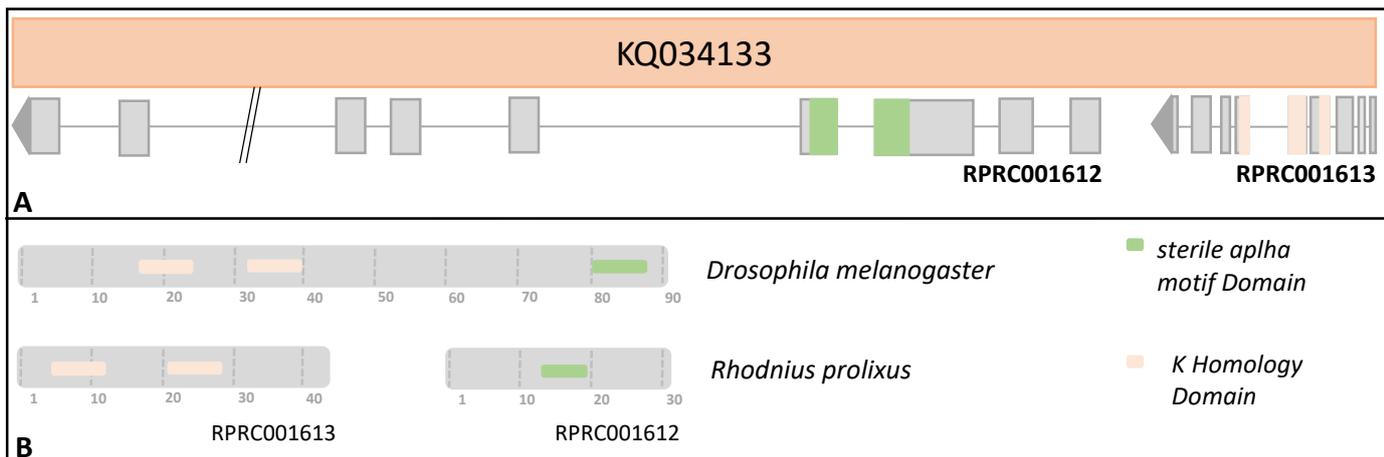
Para los genes que no se lograron identificar mediante estas búsquedas, se decidió, en primer lugar, corroborar la presencia de los mismos en la base de datos generada a partir de FlyBase, y se observó que la única predicción que no estaba presente fue la correspondiente al gen *maternal expression at 31B (me31B)*. Ante esto, se decidió hacer una búsqueda manual de la secuencia proteica en FlyBase y realizar el BLASTX en cada uno de los transcriptomas generados para ver si se lograba identificarlo. *me31B* se asignó al identificador RPRC009336, correspondiente al genoma anotado de *R. prolixus*, el cual se encontró en todos los transcriptomas analizados, exceptuando al tiempo *0hpo* correspondiente a la corrida *single-end*. En la tabla 12 se observan las características de los alineamientos obtenidos en cada tiempo secuenciado.

Por otro lado, a los que sí fueron identificados en la base de datos generada a partir del FlyBase, pero no en los transcriptomas reconstruidos a partir del mapeo contra el genoma anotado, se decidió volver a realizar una búsqueda de similitud de secuencia, pero utilizando a los transcriptomas reconstruidos a partir del mapeo contra el genoma de *R. prolixus* como referencia. Se pudo identificar al gen *knirps (kni)* en el *contig* ACPB03006526.1, a *giant (gt)* en el ACPB03018756.1, al gen *Ciclina B (CycB)* en el *contig* ACPB03008824.1 y *Antenapedia (Antp)* en el ACPB03004290.1. En el único tiempo secuenciado donde sólo se identificaron dos identidades fue en *0hpo*, donde se hallaron *gt* y *CycB*; en los demás transcriptomas se encontraron los 4 genes nombrados previamente. Por otro lado, no se observó la presencia del gen *Runt (run)* en los 4 tiempos del desarrollo secuenciados.

<b>Maternal expression at 31B</b> RPRC009336		<b>% ID</b>	<b>Longitud hit</b>	<b>E-value</b>
	<b>0hpo_SE</b>	*	*	*
	<b>12hpo</b>	81.35	429	0.0
	<b>24hpo</b>	81.35	429	0.0
	<b>48hpo</b>	81.35	429	0.0
	<b>0hpo_PE</b>	81.35	429	0.0
	<b>No fertilizados</b>	81.35	429	0.0
	<b>Ovario</b>	81.35	429	0.0

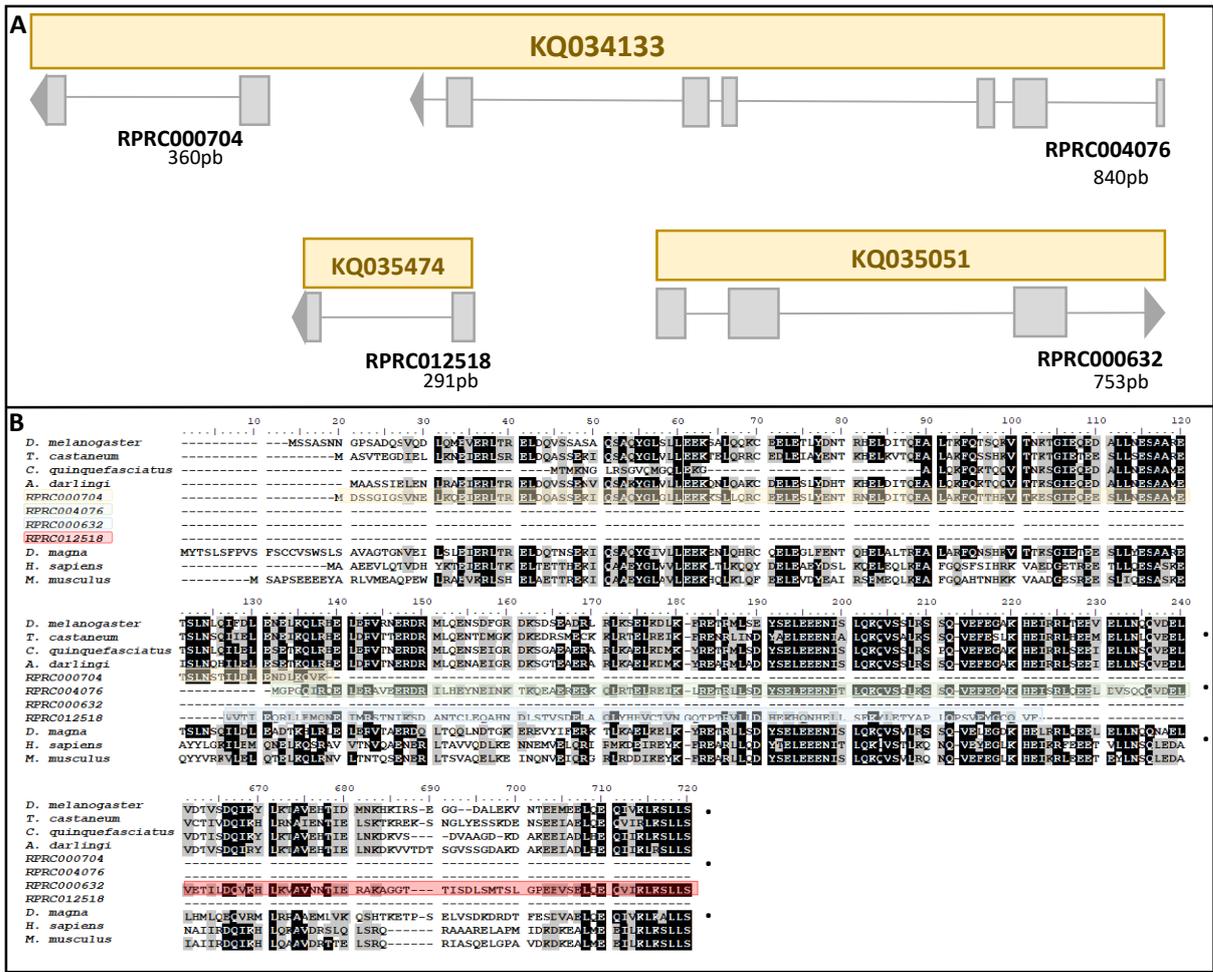
**Tabla 12:** Gen *me31B* identificado en todos los transcriptomas generados.  
(\* ) genes no identificados utilizando BLAST

Como paso final de identificación de los transcritos, se llevó a cabo un análisis más minucioso de aquellos en los que se iban a realizar validación experimental. Para esto se efectuó la búsqueda de dominios mediante *InterProScan*, y utilizando *Clustal W* se llevó a cabo un análisis de secuencia a nivel evolutivo. Uno de los genes que se decidió anotar con mayor profundidad fue *Bicaudal C* (*BicC*), se identificó en los transcriptomas generados como dos unidades transcripcionales diferentes (RPRC001613 de 1284pb y RPRC001612 de 954pb) (Figura 15). Ambos se encontraron dentro del mismo *SuperContig* KQ034133 en la cadena reversa (información disponible en VectorBase). Al analizar cada transcripto por separado se pudo determinar que el transcripto anotado *ab initio* RPRC001613 codifica para un fragmento proteico que presenta dos dominios KH (*K Homology*) y que el transcripto RPRC001612 contiene al dominio SAM (*sterile alpha motif*) ambos característicos en *BicC* (Mahone et al., 1995); en la parte B de la Figura 15 se esquematizan los dominios identificados. Para finalizar el análisis se efectuó un alineamiento múltiple con secuencias proteicas correspondientes a *BicC* de especies conocidas. Como era de esperar cada uno de los transcritos identificados se alineó con las demás secuencias en el segmento correspondiente a cada dominio identificado (Suplementario 3). Para continuar con el análisis de la parte II se decidió trabajar con el transcripto RPRC001613.



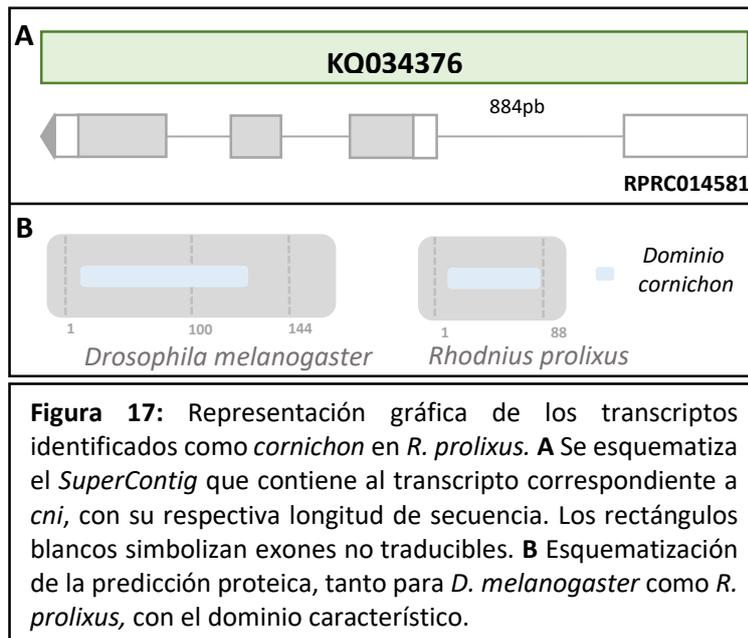
**Figura 15:** Representación gráfica de los transcritos identificados como Bicaudal C en *R. prolixus*. **A** se esquematiza el *SuperContig* que contiene a los transcritos identificados. **B** Comparación de las predicciones proteicas para *D. melanogaster* y *R. prolixus*, con la representación de sus dominios característicos.

Otro de los genes que decidimos validar en profundidad fue *Bicaudal D (BicD)*, el cual se identificó en nuestros transcriptomas en cuatro unidades transcripcionales diferentes correspondientes al genoma anotado de *R. prolixus*, identificados como RPRC000704, RPRC004076, RPRC012518 y RPRC000632. Los dos primeros transcritos *ab initio* se ubicaron en el mismo *SuperContig* registrado como KQ034133 en la cadena reversa del mismo. Los otros dos transcritos *ab initio* fueron ubicados, cada uno, en dos *SuperContis* diferentes. En la Figura 16A se muestra la distribución de los transcritos y su longitud característica; en VectorBase es posible explorar más información de estos transcritos. Para confirmar si correspondían a distintas partes del gen se prosiguió a realizar un alineamiento múltiple de las predicciones proteicas junto a secuencias de especies conocidas. La parte B de la Figura 16 muestra que las predicciones corresponden a distintas partes del gen, exceptuando a la predicción correspondiente al transcritos RPRC012518 la cual alineó con baja similitud de secuencia y en la misma región que el transcritos RPRC004076. Algo para remarcar es que las 4 unidades transcripcionales presentaron como dominio característico único al dominio Bicaudal-D. En la parte II de este trabajo se continuó con el análisis del transcritos RPRC004076.



**Figura 10:** Representación gráfica de los transcritos identificados como *Bicaudal D* en *R. prolixus*. En la parte **A** se esquematiza el *Supercontig* que contiene a los transcritos identificados para *BicD*, cada uno con el reporte de su longitud característica. En la parte **B** se esquematiza el alineamiento múltiple con cada transcritos identificados como *BicD*, representados con un color distintivo a lo largo del alineamiento. Los puntos en negro en los extremos marcan discontinuidad del alineamiento. La conservación de aminoácidos se visualiza con bloques negros, la conservación a nivel de grupo aminoacídico con bloques en color gris. Los guiones muestran la ausencia de secuencia alineada a lo largo del alineamiento.

Por último, se identificó a *cornichon* (*cni*) el cual fue reconocido en un único transcripto, RPRC014581, dentro de *SuperContig* KQ034376 con dirección en la cadena *reverse*. La Figura 17A muestra la estructura característica del transcripto. Además, se determinó que la predicción proteica presenta una longitud de 88 aminoácidos (aa) que presenta sitios *cornichon* conservados (Figura 17B). En este gen no encontramos ambigüedades con respecto a la anotación genómica presente en VectorBase.



# Discusión

## 1. Datos Crudos de Secuenciación

El surgimiento y acceso a las técnicas de secuenciación de nueva generación (NGS, *Next-generation sequencing*) han llevado a la producción de datos a gran escala. Junto a esto, la reducción de los costos de secuenciación ha disminuido drásticamente, entre dos a tres órdenes de magnitud en los últimos diez años (Shendure and Ji, 2008; Tucker et al., 2009). Este escenario hizo posible estudiar el campo transcriptómico de forma más flexible y global que lo aportado por técnicas anteriores. La secuenciación masiva y profunda de ARNs (ARN-Seq), permite catalogar los transcriptos expresados en una condición específica, siendo una técnica de alto rendimiento que ha encontrado diversas aplicaciones en la actualidad (Chen et al., 2011; McClure et al., 2013; Oszolak and Milos, 2011; t Hoen et al., 2008; Van Verk et al., 2013; Wang et al., 2009).

Aunque el alcance del ARN-Seq es mucho mayor a técnicas anteriormente usadas, su poder informativo se ve limitado por cuatro factores importantes: el diseño experimental, la información requerida sobre los datos secuenciados, la calidad de la secuenciación y el análisis bioinformático de los datos generados (Agarwal et al., 2010; Bradford et al., 2010; Oshlack et al., 2010; Sonesson and Delorenzi, 2013).

El diseño experimental es un paso de gran relevancia, ya que cualquier error en este sólo se puede corregir reemplazando las muestras con una nueva secuenciación; lo cual implica altos costos y extiende el tiempo de la investigación; en este proyecto se decidió realizar un diseño para generar datos cualitativos, con el fin de identificar los transcriptos que se expresan en un momento específico. El parámetro más importante para alcanzar dicho objetivo es lograr una buena cobertura de secuenciación, lo que permita que las *reads* generadas cubran de manera uniforme a cada transcripto. En este trabajo, las condiciones propuestas para llevar a cabo los objetivos planteados hicieron posible obtener una buena calidad de secuencia y de secuenciación. Por un lado, se utilizó la plataforma de secuenciación Illumina, que es considerada como una de las tecnologías más dominantes en el mercado (Metzker, 2010). Por otro, la manipulación de las muestras en condiciones de esterilidad y bajas temperaturas permitieron obtener ARN para secuenciar de buena calidad.

El principal parámetro que se debió mejorar fue la presencia de secuencias sobrerrepresentadas correspondientes a artefactos contaminantes propios de la secuenciación. La limpieza de las lecturas por búsqueda de similitud de secuencia contra la base de datos de UniVec, permitió eliminar en su totalidad las secuencias no deseadas, debido principalmente a que está compuesta exclusivamente por secuencias referentes a vectores, adaptadores, oligonucleótidos, *indels*, que son parte de los procesos de clonación y secuenciación (<https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/>). Tras este paso de limpieza, se eliminaron las secuencias que daban *hit* con las secuencias presentes en la base de datos, el número total de *reads* eliminadas fue muy pequeño en comparación al número de *reads* crudas obtenidas, por lo que no se consideró como un cambio significativo que influyera en el análisis bioinformático siguiente. El único transcriptoma que sí presentó un cambio significativo fue el correspondiente a la corrida *paired-end* del transcriptoma *Ohpo* (Figura 8), esto se debió a que las lecturas además de presentar artefactos propios de la secuenciación, mostraron valores de calidad muy bajos que son indicativos de una probabilidad de error alta, lo que lleva a la incorporación de base errónea. Esta mala calidad también se evidenció por la presencia de bases no convencionales a lo largo de las lecturas. Por esto, era necesario realizar una buena limpieza del *set* de datos crudos para continuar de manera eficiente con los pasos posteriores del análisis de datos. La calidad mala que mostraron los datos de partida se debió principal y únicamente a errores propios del servicio de secuenciación.

Si se hace referencia al porcentaje de secuencias contaminantes presentes en los *sets* de datos generados, en los transcriptomas obtenidos por secuenciación *paired-end* correspondiente a ovario y huevos no fertilizados, se obtuvo una alta presencia de secuencias contaminantes, las cuales representaban aproximadamente el 50% del *set* de datos inicial. Es probable que esta contaminación sea consecuencia de un ineficiente pre-proceso de limpieza, que por lo general es llevado a cabo por el servicio de secuenciación, donde se encargan de eliminar los *indels* y adaptadores incorporados para realizar la secuenciación en sí.

Según el consorcio de investigación ENCODE (Instituto Nacional de Investigación del Genoma Humano (NHGRI)), la profundidad de secuenciación adecuada varía dependiendo del objetivo del proyecto. Para experimentos cuyo propósito es evaluar la similitud entre dos

perfiles transcriptómicos, se recomienda tener 30 millones de lecturas pareadas de una longitud mayor a treinta nucleótidos. Si el experimento busca descubrir nuevos transcritos, o cuantificarlos de forma robusta, un mínimo de 100-200 millones de lecturas mayores a setenta pares de bases es lo recomendado. En nuestro caso, el número mínimo de lecturas obtenidas fue en el transcriptoma *single-end* de tiempo *Ohpo*, aproximadamente 29 millones de *reads*, con longitud de 50pb, profundidad razonable para el objetivo planteado.

Una vez obtenidos los datos de secuenciación, la elección correcta de las herramientas para reconstruir los transcriptomas es fundamental para alcanzar el objetivo planteado. Así, en este trabajo, se prosiguió a emplear las herramientas bioinformáticas más eficientes reportadas en la bibliografía para llevar a cabo los pasos subsiguientes (Kim et al., 2013; Langmead et al., 2009; Trapnell et al., 2012)

## 2. Alineamiento de lecturas de ARN-seq a la referencia y ensamble

El objetivo general al mapear o alinear una colección de lecturas obtenidas tras la secuenciación a una referencia es descubrir la verdadera ubicación de origen de cada una de las lecturas con respecto a esa referencia. Por lo que, la disponibilidad de una referencia que contenga regiones repetitivas, errores de ensamble o falta de información genética llevará a que el objetivo propuesto no sea posible para un subconjunto del total de las lecturas. Del mismo modo, la presencia de errores de secuenciación, así como también la presencia de polimorfismos dificultaran la alineación. A la luz de estos factores, se han desarrollado programas que son flexibles a la hora de establecer los parámetros del mapeo (Langmead et al., 2009; Trapnell et al., 2009; Trapnell et al., 2012). En este trabajo, se llevó a cabo el mapeo tanto contra los transcritos anotados *ab initio* como contra el genoma como referencia, ambos disponibles en VectorBase. En el caso donde se utilizó el genoma anotado como referencia, los límites exón/intrón no fueron una limitante, por lo que se pudo realizar el análisis utilizando los dos algoritmos propuestos en un inicio. Cabe resaltar que, como mostraron los resultados, se obtuvo una mayor cobertura cuando se utilizó la estrategia que permite la presencia de *gaps* al alinear (Figura 4). Esto se explica principalmente por la forma de actuar que presenta *Tophat2*; primero utiliza a *Bowtie* como “motor” de alineamiento de las *reads* a la referencia de manera continua, luego toma las lecturas que no pudieron alinearse a la referencia y se les aplica un proceso de fragmentación generando segmentos

más cortos, los cuales son alineados nuevamente de manera independiente, a la referencia. Esta estrategia hace que existan más posibilidades de que las lecturas encuentren su ubicación en la referencia. Independientemente del algoritmo que sea utilizado, el mapeo contra el transcriptoma como referencia se encuentra restringido a la reconstrucción únicamente de transcriptos conocidos y predichos (Trapnell et al., 2012).

Cuando se analizó el porcentaje de lecturas alineadas al genoma expresado como referencia, se observó que este fue mayor al 50% en todos los transcriptomas generados, a excepción del transcriptoma que corresponde a tiempo *Ohpo single-end*. Esto está correlacionado con el bajo número de lecturas utilizadas como material de partida para mapear. En los transcriptomas obtenidos por secuenciación *paired-end* y con longitud para cada *read* de 150 pb se esperaba encontrar una mayor cobertura, debido principalmente a la estrategia de secuenciación utilizada, y en consecuencia un mayor número de lecturas mapeadas. Llamativamente, esto no fue así debido a que el porcentaje de mapeo rondó el 50%, al igual que la *single-end*. Esta situación se podría llegar a revertir realizando el mapeo de estos transcriptomas contra la referencia genómica, ya que permitiría identificar lecturas que mapean en regiones de empalme o en regiones previamente no reportadas como secuencias codificantes. Sumado a que se encontró que la anotación *ab initio* de sus predicciones presenta limitaciones, por lo que mapear contra el genoma evitaría encontrar ambigüedades presentes en la anotación genómica.

Cuando se mapeó contra el genoma como referencia, el algoritmo más apropiado para llevar a cabo dicha tarea fue el correspondiente al programa *Tophat2*, ya que el alineamiento de las lecturas a la referencia puede abarcar uniones de empalme, y este programa presenta las características para resolver este desafío computacional de manera correcta. Es importante señalar que la baja cobertura de los transcriptomas generados, cuando se los mapeó a la referencia genómica, se debe principalmente a la estrategia de secuenciación planteada, lecturas cortas, de 50pb y *single-end*. Más allá de esta limitación técnica, mapear contra la referencia genómica tiene la ventaja de que permite identificar nuevos genes o transcriptos; sin anotaciones previas, por lo que no demanda de un conocimiento previo del conjunto de regiones transcriptas (Conesa et al., 2016). Particularmente, la identificación de los genes sin anotaciones previas se ejemplifica cuando, en el presente trabajo, se logró

identificar en los distintos tiempos del desarrollo genes como *giant (gt)*, *Ciclina B (CycB)*, *knirps (kni)* y *Antennapedia (Antp)*.

El programa que se utilizó para ensamblar (Trapnell et al., 2012) presenta un algoritmo basado en grafos acíclicos dirigidos que buscan el conjunto mínimo de transcritos independientes que pueden explicar los *reads* observados en cada experimento secuenciado. La metodología utilizada consiste en ir agrupando las lecturas en grupos que se mapean en la misma región de la referencia. Cuantas más *reads* se identifican en una región más confianza presenta el transcritos ensamblado.

A la hora de la reconstrucción de los transcriptomas se utilizó la herramienta *Cuffmerge*. Esta herramienta realiza un ensamblado unificado en base a los ensamblados generados a partir de los distintos algoritmos de mapeo utilizados (Trapnell et al., 2012; Trapnell et al., 2010). Se obtuvo un número similar de transcritos por tiempo, aproximadamente 12000 transcritos por transcriptoma, exceptuando el tiempo *Dhpo single-end* donde el número de transcritos obtenidos rondó los 1000. Esta diferencia en este tiempo analizado se observó desde el momento del análisis de calidad y la limpieza de las lecturas crudas, lo que muestra la importancia que tiene el *set* de datos inicial para proseguir con el análisis.

Por otro lado, es importante remarcar que, no siempre un mayor número de *reads* que mapean a la referencia lleva a un mayor número de transcritos reconstruidos. Esto se explica debido a que, en esas situaciones, no mapea la cantidad mínima necesaria para recrear un transcritos independiente (Conesa et al., 2016), situación que vimos ejemplificada en el transcriptoma de huevos no fertilizados.

### 3. Identificación de transcritos

Debido a que nuestro objetivo era identificar genes que se expresan durante el desarrollo embrionario, la búsqueda de similitud de secuencia nos permitió tener una aproximación mucho más específica de la información que contábamos en nuestros transcriptomas generados. La comparación entre genomas se convirtió, desde hace unos años, en un componente indispensable para comprender una gran variedad de procesos biológicos (Tatusov et al., 1997). Permite transferir información génica desde un organismo modelo bien

estudiado a especies no tan convencionales, lo que la convierte en una herramienta poderosa para el estudio de la evolución de los genomas (Jiang et al., 2013).

Cabe recordar que las especies comparadas en este trabajo se clasifican en distintos grupos de acuerdo a su proceso de desarrollo embrionario. *D. melanogaster* presenta embriogénesis de banda germinal larga, es un organismo más derivado, evolutivamente hablando, y con metamorfosis completa (Holometábolo); y por el otro lado, *R. prolixus* presenta metamorfosis incompleta (Hemimetábolo), con embriogénesis de banda germinal corta/intermedia que caracteriza a los órdenes basales de insectos. A pesar de estas diferencias embriológicas, ambos establecen un plan corporal compuesto por cabeza, tórax y abdomen (Davis et al., 2002); y en este establecimiento participan diferentes vías de señalización y se activan diferentes procesos de desarrollo para cumplir con tal objetivo. En este trabajo se decidió llevar a cabo este abordaje de inferencia de identidad debido a que se han realizado análisis comparativos entre el genoma del organismo modelo *D. melanogaster* y especies de insectos, como *Apis mellifera*, *Tribolium castaneum*, *Bombyx mori*, *Oncopeltus fasciatus*, *Acyrtosiphon pisum*, *Aedes aegypti*, *Culex quinquefasciatus*, *Anopheles gambiae* y, *R. prolixus* (Behura et al., 2011; Dearden et al., 2006; Ewen-Campen et al., 2011; Harker et al., 2013; Lavore et al., 2015; Shigenobu et al., 2010). En estos se demostró la conservación de varios de los genes que componen los procesos que se encuentran activos durante el desarrollo embrionario, lo que permite darle soporte a las inferencias génicas encontradas en los distintos tiempos analizados durante el desarrollo, los cuales abarcaron desde el proceso de oogénesis hasta la gastrulación y formación de banda germinal.

Resulta dificultoso conocer si la imposibilidad de identificar un gen dado refleja la ausencia del mismo en el momento específico analizado, si es debido a una divergencia evolutiva del gen, o si se debió a problemas, en la secuenciación y ensamblado, generados durante el proceso de obtención de los transcriptomas embrionarios.

En nuestros resultados, existieron genes, como *gurken* (*grk*), *bicoid* (*bcd*) y *oskar* (*osk*), que no lograron ser identificados en ninguno de los transcriptomas reconstruidos y fueron excluidos del análisis. La decisión de excluirlos se basó en los análisis previos realizados en el genoma anotado de *R. prolixus* (Lavore et al., 2015). La causa de esto se debió pura y exclusivamente a una divergencia evolutiva, ya que, estudios previos reportaron que la

expresión de los genes *grk* y *bcd* se identificó únicamente en la especie *D. melanogaster* (Dearden et al., 2006); asimismo, el gen *osk* se reportó exclusivamente dentro de orden diptera, en las especies *D. melanogaster*, *A. aegypti*, *C. quinquefasciatus* y *A. gambiae* (Behura et al., 2011).

Por otro lado, la presencia de problemas en la secuenciación se evidenció en la identificación prácticamente nula de genes del desarrollo en el transcriptoma correspondiente al tiempo *Ohpo single-end*, donde únicamente cinco genes fueron identificados mediante la similitud de secuencia (Tabla 6-8). Resultado que se revirtió al re-secuenciar el estadio correspondiente a cigoto (*Ohpo*) (Tabla 9-11), descartando de esta manera la posibilidad de que la ausencia sea específica del tiempo analizado.

Por otra parte, en este análisis evidenciaron las limitaciones que presentan los transcritos anotados *ab initio* en VectorBase, ya que observamos ciertas ambigüedades a la hora de asignarle identidad a los transcritos que componen los transcriptomas secuenciados en este trabajo.

Los resultados obtenidos indican la presencia de genes expresados a lo largo de los distintos tiempos del desarrollo secuenciados, impidiendo la identificación de expresión génica asociada a tiempo del desarrollo específico. Sin embargo, si se pudo inferir que los genes analizados presentan una expresión materna, debido a que desde el estado de ovario fue posible identificar su expresión. Esto tiene sustento en algunas validaciones experimentales en los cuales se observó expresión materna pero funcionalidad génica específica a un momento del desarrollo más avanzado, como fue el caso del gen *giant (gt)* y *Toll (Tl)* para *R. prolixus* (Berni et al., 2014; Lavore et al., 2012).

Desde hace unos años se empezó a ser más evidente que la expresión de los genomas en eucariotas es mucho más compleja de lo que se creía en un inicio; principalmente gracias al advenimiento de tecnologías de alta resolución que permitieron cambiar el paradigma de que el transcriptoma exclusivamente deriva de genes codificantes de proteínas y de ciertos ARNs no codificantes específicos como, ARN de transferencias (ARNt), ARN ribosomales (ARNr), pequeños ARN, otros. Esta situación dio un vuelco inesperado cuando numerosos estudios demostraron que la ARN polimerasa II se puede encontrar prácticamente en cualquier

ubicación del genoma (Kim et al., 2005; Steinmetz et al., 2006), demostrando de que existe una transcripción generalizada en los organismos eucariotas, la cual implica la expresión de ARNm no codificantes, largos y cortos, producidos a partir de regiones intergénicas, intragénicas y repetitivas. Estos ARN son transcritos en la orientación contraria (antisentido) respecto a las secuencias codificantes vecinas, y parecerían proporcionar un nivel adicional de regulación de la expresión génica; la cual sería necesaria para un ajuste fino durante una condición específica dada (Jacquier, 2009; Tisseur et al., 2011). En este trabajo, la expresión generalizada encontrada en los distintos tiempos del desarrollo estudiados podría ser un claro ejemplo de la transcripción extensiva previamente descrita, aunque es necesario desarrollar validaciones experimentales para afirmar tal hecho.

Por otro lado, en la embriogénesis un evento clave es la transición materno-cigótica, en la cual el desarrollo embrionario pasa de un control estrictamente materno a un control por productos cigóticos. Un factor clave en esta transición es la degradación de los transcritos maternos. Se evidenció que esta eliminación de ARNm distribuidos en el embrión de manera uniforme podría tener un rol tanto instructivo como permisivo. Este último, supone que la degradación materna podría permitir la transcripción cigótica en un patrón de expresión modelado en tiempo y espacio (Aanes et al., 2014; Walser and Lipshitz, 2011). Un claro ejemplo de esta situación se describió en *D. melanogaster* para un homólogo de Cdc25, el cual se elimina del embrión al finalizar la transición materno-cigótica; y luego, avanzado el desarrollo embrionario, es reemplazado por un patrón de expresión exclusivamente cigótico, el cual se correlaciona y regula al embrión gastrulante (Edgar and O'Farrell, 1990; Foe, 1989). Asimismo, un ejemplo de cómo la eliminación de ARNm puede ser instructivo se evidencia en la disminución del número de ciclos celulares, que tienen lugar previo a la celularización del blastodermo, como consecuencia de la disminución de la dosis materna correspondiente a los genes *string (stg)* y *twine (twe)* (Benoit et al., 2009; Lu et al., 2009).

Además de este cambio de expresión debido a la degradación de los ARNm maternos, se demostró que la dinámica transcripcional se caracteriza también por la activación de transcritos latentes así como por el inicio de la transcripción *de novo* (Aanes et al., 2014).

Todos estos reportes dan sustento a nuestro análisis transcriptómico en los cuales observamos una expresión génica generalizada a lo largo de todos los tiempos analizados. Abriendo un nuevo panorama de estudio de la transcripción durante el desarrollo

embrionario, planteando diferentes interrogantes para ser abordados en un futuro cercano. Para hacerle frente a estos interrogantes, decidimos proseguir con un análisis más minucioso y funcional en la Parte II para corroborar la expresión génica *in silico* de algunas identidades obtenidas y determinar su función durante el desarrollo embrionario.

# PARTE II

# Introducción

## 1. Oogénesis en insectos

La formación y maduración del oocito es un proceso altamente controlado, ya que todos los acontecimientos deben ocurrir en el momento indicado y en la secuencia correcta para que la posterior embriogénesis se realice adecuadamente.

Los ovarios son órganos internos compuestos de tejido somático que contienen a las células germinales a medida que se desarrollan en oocitos y finalmente en huevos (de Cuevas, 2005).

La estructura del ovario en los insectos es variable, correspondiéndose a que son un grupo altamente diversificado. Más allá de esto, en general se observa que los ovarios son dos y simétricos respecto del eje longitudinal del cuerpo, y que cada uno de ellos está formado por estructuras tubulares denominadas ovariolas, cuyo número es especie-específico. Están caracterizados por la presencia de distintos tipos celulares que acompañan y nutren al oocito en desarrollo, y que permiten el traspaso de la información genética a la siguiente generación (Lynch and Roth, 2011).

El oocito se establece a partir de las células madres de la línea germinal. La vía de desarrollo para alcanzar este estado se diferencia según el tipo de insecto que se trate. En algunos, cada citoblasto (la célula hija obtenida de una célula madre germinal) se diferencia directamente en un oocito; mientras que, en otros, pasa por una o más rondas de mitosis para formar un *cluster* celular también denominado quiste; en el cual las células permanecen conectadas unas a otras como consecuencia de una citocinesis incompleta durante la mitosis. Esta interconexión se establece por la formación de canales en forma de anillo, estructuras tubulares de citoplasma rodeados por filamentos de actina. En insectos que forman estos quistes, a su vez es posible diferenciar la cantidad de oocitos maduros que alcanza. En algunos, después de la formación del quiste, se rompen las uniones anulares y cada citocito (célula que compone al quiste) se diferencia en un oocito independiente. En otros insectos, estos canales anulares permanecen, y sólo una de las células que componen el quiste se diferencia como oocito mientras que las otras se convierten en células nodrizas, las cuales proveen nutrientes y citoplasma al oocito mediante los canales, hasta que el oocito alcanza la madurez. Una vez que esto ocurre, las células nodrizas degeneran y desaparecen (de Cuevas, 2005).

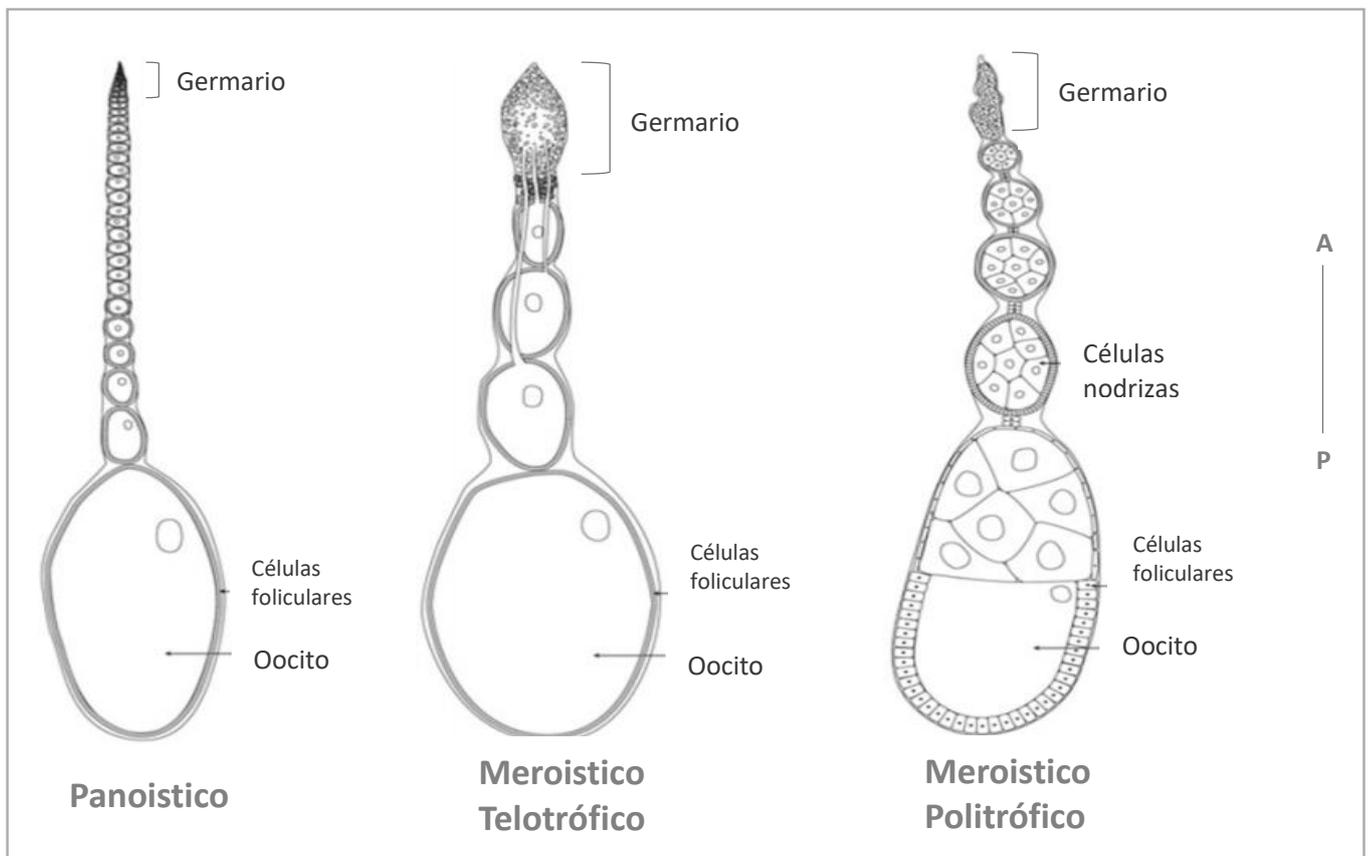
Más allá de la vía del desarrollo del oocito, en correspondencia a su característica citológica, los ovarios se agrupan en dos clases, panoístico o meroístico, que se diferencian por la

ausencia o presencia, respectivamente, de células nodrizas en su estructura (Figura 18) (Lynch and Roth, 2011).

Dentro del tipo de ovario meroístico, los cuales presentan células nodrizas, podemos diferenciar los subtipos politróficos y telotróficos. El subtipo politrófico lo encontramos principalmente en insectos holometábolos, y asociado a insectos de banda germinal larga. En estos, las células nodrizas acompañan el oocito en desarrollo a lo largo de la ovariola, y son rodeados por una capa de células foliculares que envuelven por completo a ambos conjuntos celulares. Por otra parte, el subtipo telotrófico se caracteriza por la presencia de las células nodrizas en el extremo anterior de cada ovariola, actuando como fuente común de ARNm y componentes citoplasmáticos para todos los oocitos. Al igual que el subtipo politrófico, presenta una cubierta de células foliculares somáticas. Este tipo de ovogénesis se la puede encontrar tanto en insectos hemimetábolos como holometábolos, y está asociada principalmente con insectos de banda germinal corta (Lynch et al., 2011).

Por otro lado, se encuentra al ovario del tipo panoístico, el cual se caracteriza por no presentar células nodrizas especializadas. Se compone de células germinales, las cuales se convierten todas en oocitos. El núcleo del oocito es el encargado de sintetizar todos los ARNm maternos y componentes citoplasmáticos necesarios para el desarrollo embrionario. Esta estructura también se encuentra rodeada por una capa de células foliculares somáticas. Este tipo de ovogénesis se la encuentra asociada exclusivamente con insectos de banda germinal corta (Lynch et al., 2011).

Sea cual fuese el tipo de estructura que caracteriza a cada ovario, todos comparten la presencia de un epitelio folicular que rodea a los oocitos en desarrollo, el cual establece una conexión entre estos y el medio extra-ovárico.



**Figura 18:** Estructura de las ovariolas correspondientes a tres organizaciones distintas de ovarios de insectos. De izquierda a derecha se esquematizan las ovariolas derivadas de ovarios panoísticos, telotróficos y politrófico, indicando para cada uno las distintas partes que lo componen (germario, células foliculares, células nodrizas, oocitos en desarrollo). Referencias: A, Polo anterior; P, Polo posterior. Figura adaptada de McLaughlin and Bratu (2015)

Los eventos del desarrollo claves durante la ovogénesis son el establecimiento de los ejes de simetría (Coutelis et al., 2008; Nusslein-Volhard and Roth, 1989; Schupbach and Roth, 1994; Speder et al., 2006), la acumulación de vitelo en el oocito (Atella et al., 2005), y la formación del corion y membrana vitelina (Papantonis et al., 2015).

Todos los organismos durante su desarrollo establecen la polaridad, es decir crean diferencias entre los polos opuestos en el embrión. Esta polaridad es necesaria principalmente para el establecimiento de los tres ejes corporales que definen la anatomía del embrión, el eje anteroposterior (A-P), el eje dorso-ventral (D-V) y el eje de simetría bilateral de izquierda-derecha (I-D). Esta polaridad, y por lo tanto el establecimiento de los ejes embrionarios, ocurre antes de la fecundación (Roth, 2003). En el organismo donde se lo ha estudiado con mayor detalle es en la mosca de la fruta, *D. melanogaster*, en el cuál la disposición de ciertas

proteínas e incluso ARNm en el huevo presentan una polarización, debido a que son depositados de manera desigual entre un polo y el opuesto durante la formación del huevo. Este gradiente se ejemplifica con el gen *bicoid* (*bcd*), el cual regula la expresión de genes que establecen el eje A-P (St Johnston and Nusslein-Volhard, 1992); y con el gen *gurken* (*grk*), el cual regula la expresión de genes que establecen el eje D-V (van Eeden and St Johnston, 1999).

Como se mencionó previamente, el oocito recorre la ovariola asociado a células formadoras de nutrientes (células nodrizas) y al epitelio folicular, que también aporta nutrientes; aunque además recibe nutrientes de tejido extra-ovárico, como el cuerpo grasoso, actuando como mediador del traspaso de los distintos componentes que conforman el vitelo; el componente principal es la vitelogenina que al ser secuestrada por el oocito es modificada co- y post- transcripcionalmente por la adición de glicosil, residuos de fosfatos y sulfatos. Tras la modificación, se la conoce como vitelina y es la forma en que se la almacena en el oocito. También son parte importante del vitelo proteínas, lípidos y azúcares, los cuales son almacenados en el huevo de una manera altamente organizada. Una vez producida la fertilización se dispara el programa que da lugar a la formación de un organismo completamente nuevo; es aquí donde el vitelo, previamente acumulado, comienza a ser utilizado, de manera gradual, de acuerdo con las necesidades específicas celulares reguladas por el programa genético del embrión (Atella et al., 2005).

El oocito se encuentra rodeado por dos estructuras protectoras, del lado más externo el corion y del lado interno la membrana vitelina (Margaritis et al., 1980). El corion se caracteriza por ser una estructura rígida, que permite el traspaso de gases y agua, mientras que la membrana vitelina es una estructura más resistente a este paso de moléculas (Konrad et al., 1993).

La membrana vitelina se forma por la acumulación, y posterior fusión, de compuestos en el espacio existente entre el oocito y las células foliculares. Se reportó en algunas especies como *D.melanogaster*, *Gryllus bimaculatus* y *Aeschna Sp.*, que la fuente de las secreciones que forman la membrana vitelina son las células foliculares (Cummings, 1972; Cummings et al., 1971; King and Koch, 1963). Sin embargo, en otros como el abejorro *Bombus terrestris*, se reportó que tanto las células foliculares como el oocito participan en la formación de la membrana (Hopkins and King, 1966).

Las células foliculares forman parte del epitelio del oocito y secretan principalmente las capas sucesivas que dan lugar a la formación del corion, en las últimas etapas de la ovogénesis. Se dividen y se separan en subpoblaciones que se mueven a posiciones específicas, adaptan sus superficies e intervienen así en la polarización del oocito. Es durante la vitelogénesis, depósito del vitelo en el oocito, que las células foliculares exhiben su mayor actividad catabólica, con el objetivo final del establecimiento del corion que ofrece protección mecánica al oocito en desarrollo. Esta estructura supramolecular es compleja, con moléculas grandes altamente organizadas regional y radialmente. Al finalizar la formación del huevo, las células foliculares se someten a muerte celular por apoptosis (Papantonis et al., 2015).

Todo este proceso es altamente regulado por distintos factores génicos que dan lugar a la correcta formación del huevo; y en el único organismo donde se estudió en detalle el proceso y sus actores genéticos fue en *D. melanogaster*. Se identificó un gran repertorio de genes que se clasificaron según el tipo de función descrita durante el proceso de ovogénesis (Brody, 1999):

- Genes que actúan bien temprano en el mantenimiento y división de las células madres que conforman la línea germinal, como armadillo (*arm*), pumilio (*pum*), vasa (*vas*), piwi (*piwi*), otros.
- Genes que actúan bien temprano en el huevo estableciendo la determinación y formación del eje A-P del oocito, Bicaudal C (*BicC*), Bicaudal D (*BicD*), gurken (*grk*), egghead (*egh*), otros.
- Genes que actúan temprano en el huevo para establecer la polaridad D-V, cappuccino (*capu*), cornichon (*cni*), squid (*sqd*), otros.
- Genes que afectan el citoesqueleto como, homeless (*hls*), karst (*kst*), singed (*sn*), abnormal spindle (*asp*), otros.

Esta pequeña clasificación permite conocer según función asociada los genes activos durante la ovogénesis, pero es importante tener en cuenta que hay genes que cumplen más de una función durante el proceso, así como también hay genes que su función no presenta una categorización funcional por lo que escapan de la clasificación antes mencionada.

## 2. Oogenesis en *Rhodnius prolixus*

Las hembras adultas de *R. prolixus* presentan dos ovarios cada uno con siete ovariolas. En cuanto a la disposición celular (correspondiente a ovario telotrófico), se caracteriza por presentar células nodrizas confinadas en el extremo anterior de la ovariola, región denominada trofario, y posterior a esta los oocitos en desarrollo, rodeados por células foliculares, zona denominada vitelario. En esta última, los oocitos a medida que progresan en el desarrollo no sólo aumentan el tamaño celular, sino que también aumentan en dimensión nuclear. Este último también conocido como vesícula germinal, sufre modificaciones en la compactación de la cromatina con el objetivo de acompañar y promover los cambios necesarios para la formación del huevo (Huebner and Anderson, 1972c).

El trofario o germario, se encuentra dividido en tres zonas en base al tamaño nuclear de las células nodrizas. En la zona I los núcleos son pequeños y experimentan división mitótica, mientras que en la zona II y III hay un aumento del tamaño nuclear y diferenciación de nucléolos. La zona II es más bien una zona de transición entre las otras dos. El rol que se lleva a cabo en el trofario es producir y proveer al oocito en desarrollo de nutrientes y ARNs, a través de cordones citoplasmáticos especializados denominados cordones tróficos, mediante un transporte unipolar mediado por una red de microtúbulos que favorecen el pasaje (Harrison and Huebner, 1997; Huebner and Anderson, 1972a; Nunes-da-Fonseca et al., 2017; Valdimarsson and Huebner, 1989).

La vitelogénesis es un proceso fundamental para el desarrollo del huevo; los insectos hematófagos requieren de una ingesta de sangre para activar numerosos genes esenciales para llevar a cabo la digestión de sangre, la síntesis de precursores proteicos del vitelo y la producción del huevo (Raikhel, 2005). Aunque el apareamiento es fundamental para estimular la vitelogénesis, se ha observado tanto para *R. prolixus* como *T. infestans*, la capacidad de dar origen a huevos utilizando la sangre ingesta en el último estadio larval, facultad denominada autogenia (Davey, 1987; Mundall and Engelmann, 1977; Noriega, 1992; Stoka et al., 1978), aunque cabe remarcar que la tasa de oviposición es menor que la obtenida tras la ingesta.

Durante la vitelogénesis, los precursores proteicos del vitelo son producidos en grandes cantidades por tejidos extraováricos, principalmente por el cuerpo graso. Dentro de los precursores generados, el primordial en el vitelo es la vitelogenina (Vg), secretada a la

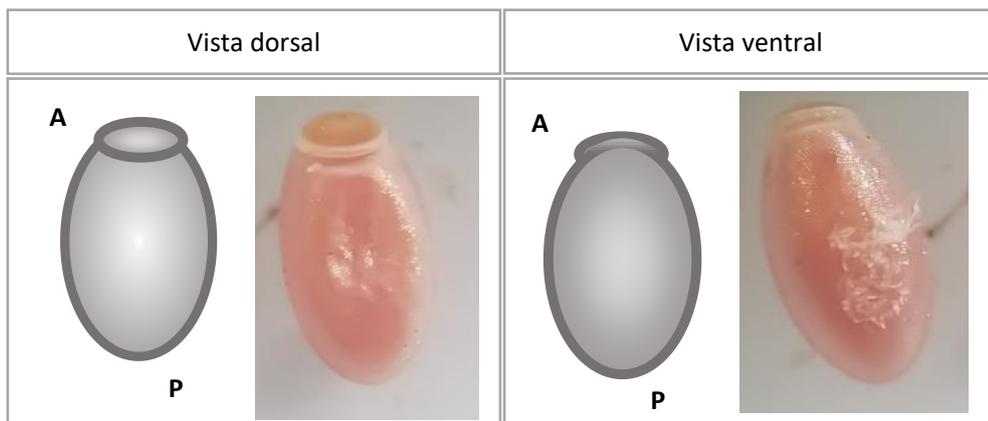
hemolinfa y captada por los oocitos en desarrollo a través de endocitosis mediada por receptor (Bellés, 1998; Giorgi et al., 2005). Una fuente complementaria de vitelogenina está a cargo de las células foliculares que rodean a los oocitos en desarrollo (Melo et al., 2000). La vitelogenina es almacenada en los oocitos en desarrollo en organelas especializadas o cuerpos vitelínicos como vitelina (Vn), siendo así el material crudo utilizado por los embriones para completar su desarrollo (Kunkel and Nordin, 1985).

Las células foliculares acompañantes sufren diferentes cambios citológicos durante los diferentes estadios que atraviesa el oocito, en correlación a la vitelogénesis, para alcanzar la maduración del huevo en el transcurso de la oogénesis. En el estado previtelogénico, correspondiente a los oocitos menos desarrollados ubicados en la parte más anterior de la ovariola, las células foliculares están organizadas en un epitelio columnar, con aproximadamente un alto de 40-50  $\mu$ . Se caracterizan por presentar complejos de Golgi y mitocondrias prominentes. Ya durante la fase de vitelogénesis temprana, las células foliculares se transforman en binucleadas y se organizan de manera tal que disminuyen los espacios intercelulares entre las células foliculares y entre las células foliculares y el oocito en desarrollo. Más allá de esta disposición, continúan siendo permeables, permitiendo el ingreso de vitelo y proteínas acompañantes. A medida que la etapa de vitelogénesis prosigue, tiene lugar la diferenciación del epitelio, en el cual las células apicales se diferencian morfológicamente de las laterales. Las células apicales, las que darán lugar al opérculo, alcanzan una longitud de 80-90  $\mu$  y se disponen de manera estrecha una al lado de la otra mientras que en los laterales las células se ubican de manera tal que generan grandes espacios intercelulares. A pesar de estos espacios, las células se mantienen asociadas a través de proyecciones tipo *gap* (Huebner and Anderson, 1972b). Todos estos descubrimientos indican que las células foliculares atraviesan cambios morfológicos en correlación al traspaso del estado previtelogénico a vitelogénico, cambios indispensables que no alteran la organización y disposición de las células que rodean al oocito en desarrollo (Huebner, 1981).

Para el fin de la oogénesis, cuando el huevo alcanzó el tamaño completo, se da lugar al estado post-vitelogénico o coriogénico, en el cual las células foliculares producen una segunda capa celular que da origen a la capa más externa del huevo, el corion. Sus funciones se caracterizan por proteger al huevo de la deshidratación, y regular la toma de oxígeno y fertilización (Beament, 1946; Huebner and Anderson, 1972b; Nunes-da-Fonseca et al., 2017).

Cuando el huevo ya está formado penetra en el oviducto para ser ovipuesto (Huebner and Anderson, 1972b). En este punto, el huevo es de color rosado, debido a la pigmentación del vitelo, con 2 mm de largo y 1 mm de diámetro, de forma ovalada con la cara dorsal cóncava y con la presencia del opérculo en la región anterior, el cual permite emerger a la larva durante su eclosión. Esta disposición muestra la constitución de la polaridad del huevo dentro del ovario (Figura 19).

La capa más externa del huevo, corion, es transparente lo que hace posible observar los cambios en la apariencia externa del embrión -morfología y pigmentación-, y permite el monitoreo del desarrollo embrionario. Toda su superficie presenta un patrón poligonal-hexagonal (ocasionalmente pentagonal) uniforme, cada uno con una fosa folicular central. Además, posee sustancia cementante en su superficie que le permite adherirse a las superficies donde son depositados (Beament, 1946).



**Figura 19:** Morfología de la estructura del huevo de *R. prolixus*. En el panel izquierdo se representa una esquematización gráfica del huevo junto a la fotografía tomada bajo lupa estereoscópica de vista dorsal de huevo y en el panel derecho se observa lo mismo correspondiente a la vista ventral. Referencias: A, Polo anterior; P, Polo posterior.

# **Materiales y Métodos**

## 1. Mantenimiento de las colonias de *Rhodnius prolixus*

Los ejemplares machos y hembras de *R. prolixus* utilizados durante el transcurso de este trabajo fueron obtenidos a partir de una colonia establecida en el Centro de Bioinvestigaciones (UNNOBA-CIC). Son criados en recipientes de plástico de boca ancha y paredes lisas, cuyas tapas son reemplazadas por una malla de nylon muy fina que facilita su alimentación. Los insectos se criaron a una temperatura de 28 $\pm$  1°C, una humedad relativa de 70%, y con fotoperiodo regulado (luz/oscuridad de 12:12 horas (hs)). Los ejemplares fueron alimentados *ad libitum*, con aves de corral, de manera regular cada 15 días. En estas condiciones la duración de la embriogénesis es de 14 $\pm$ 1 días.

## 2. Validación por la técnica de Reacción en Cadena de la Polimerasa (PCR)

### 2.1 Síntesis de ADN copia (ADNc)

Se extrajo ARN total a partir de huevos recolectados a distintos tiempos post ovipuesta, abarcando todos los estadios del desarrollo embrionario. El proceso se llevó a cabo utilizando el reactivo TRIZOL y se procedió según recomendaciones del fabricante (*ThermoFisher*). La calidad de la extracción se analizó mediante electroforesis en geles de agarosa al 1% teñidos con bromuro de etidio (BrEt). La cuantificación de su concentración fue a partir de la utilización el paquete de procesamiento de imágenes FIJI (Schindelin et al., 2012); con el cual se determinó la concentración óptima del ARN total extraído para llevar a cabo la producción de ADNc utilizando el kit *SuperScript™ VILO™ MasterMix (Invitrogen)*.

### 2.2 Diseño de oligonucleótidos específicos

Se diseñaron oligonucleótidos (oligos) específicos con la ayuda de los programas disponibles *online Primer3Plus* (Rozen and Skaletsky, 2000), *PCR Primer Stats* (Stothard, 2000) y a través de corrección manual, para poder validar mediante la técnica de Reacción en Cadena de la Polimerasa (PCR) (Mullis, 1990) los genes activos durante el desarrollo embrionario temprano y oogénesis identificados previamente en los transcriptomas generados.

Se diseñaron dos poblaciones de oligos distintas, las cuales se diferenciaron por la adición al oligo de parte de la secuencia promotora del Fago T7 (CGACTCACTATAGGG). La presencia del promotor T7 sirvió para posteriormente llevar a cabo experimentos de ARN interferente e hibridaciones *in situ*. A su vez, en los transcritos donde fue posible, se diseñaron dos pares de *primers* específicos para diferentes regiones del mismo transcripto con el fin de corroborar que el fenotipo obtenido corresponda al gen de interés y para reconocer la presencia de isoformas.

Sentido	Nombre	Secuencia	Uso
Forward	PrFw1BicC-T7	TAATACGACTCACTATAGGGCAAGGCACGTCAACAGCTAA	RNAi
Reverse	PrRv1BicC-T7	TAATACGACTCACTATAGGGGGATCGTTAGGAGCGATCAA	RNAi + <i>in situ</i>
Forward	PrFw1BicC	CAAGGCACGTCAACAGCTAA	<i>in situ</i>
Forward	PrFw2BicC-T7	TAATACGACTCACTATAGGGCGACTCAAACCTGGTGCAAA	RNAi
Reverse	PrRv2BicC-T7	TAATACGACTCACTATAGGGAACTTCGCCAGCGATAGAAA	RNAi + <i>in situ</i>
Forward	PrFw2BicC	CGACTCAAACCTGGTGCAAA	<i>in situ</i> + RT-PCR
Reverse	PrRv2BicC	AACTTCGCCAGCGATAGAAA	RT-PCR
Forward	PrFw2BicD-T7	TAATACGACTCACTATAGGGGCCGAAAAACAGATGGAAGA	RNAi
Reverse	PrRv2BicD-T7	TAATACGACTCACTATAGGGTCGACATGCTTTGTCTCTGG	RNAi + <i>in situ</i>
Reverse	PrRv2BicD	TCGACATGCTTTGTCTCTGG	RT-PCR
Forward	PrFw2BicD	GCCGAAAAACAGATGGAAGA	<i>in situ</i> + RT-PCR
Forward	PrFwcni-T7	TAATACGACTCACTATAGGGGAGCTTTCGATGAACTGAAAACA	RNAi
Reverse	PrRvcni-T7	TAATACGACTCACTATAGGGTGGAATGTTTAGGCACAATGA	RNAi + <i>in situ</i>
Forward	PrFwcni	AGCTTTCGATGAACTGAAAACA	<i>in situ</i> + RT-PCR
Reverse	PrRvcni	TGGAATGTTTAGGCACAATGA	RT-PCR
Forward	PrFwactina	GACTTGGCTGGTCGTGATCT	RT-PCR
Reverse	PrRvactina	ACCATCAGGCAATTCGTAGG	RT-PCR

**Tabla 13:** Lista de oligonucleótidos diseñados y utilizados durante este proyecto, con su respectivo uso. La secuencia se encuentra escrita con sentido 5' a 3'.

### 2.3 Amplificaciones mediante la reacción en cadena de la polimerasa con transcriptasa reversa (RT-PCR)

Como molde se utilizaron los ADNc generados, como previamente se detalló, abarcando diferentes momentos en el desarrollo; estado de cigoto, blastodermo, embrión gastrulante, banda germinal y huevo no fertilizado. Estadios coincidentes con los distintos tiempos del desarrollo secuenciados con tecnología de próxima generación, expuesto en la parte I de este trabajo.

Para realizar las amplificaciones por PCR la mezcla de reacción se compuso por: Buffer de reacción (1X), Cl<sub>2</sub>Mg (3 mM), DNTPs (25 μM), *primer forward* (250 μM), *primer reverse* (250 μM), Taq pol 1U, 2-10 ng ADNc como molde y agua estéril y ultrapura. La enzima para llevar a cabo dicha amplificación corresponde al *kit* T-PLUS ADN polimerasa (*INBIO HIGHWAY*).

El programa posee un gradiente de temperaturas de hibridación (T°a) para así determinar la temperatura óptima. El programa de amplificación que se utilizó fue: un ciclo de 5 seg a 94°C; 35 -- ciclos (30 seg a 94°C-- 30 seg a 52°C [gradiente 7°] -- 35 seg a 74°C) --; 5 min a 74°C. El termociclador que se utilizó para llevar a cabo las reacciones fue de la marca ESCO modelo Aeris™. Como control positivo de la reacción se utilizó un par de oligos específicos para el gen actina, conocido como un gen constitutivo durante el desarrollo.

Todos los productos de PCR se analizaron mediante electroforesis en geles de agarosa al 1% teñidos con bromuro de etidio (BrEt), los productos se sembraron conjuntamente al marcador de peso molecular Ladder 100 pb (*Productos Bio-Lógicos (PB-L)*). La cuantificación de su concentración fue a partir de la utilización el paquete de procesamiento de imágenes FIJI (Schindelin et al., 2012).

Los amplicones que se obtuvieron se enviaron a secuenciar a MACROGEN (Macrogen Inc., Corea del Sur) para confirmar su identidad.

### 3. Hibridación *in situ*

#### 3.1 Síntesis de las sondas de ARN

Las sondas de hibridación *in situ* fueron preparadas con digoxigenin-UTP utilizando el *kit* *RNA-Dig Labeling (Roche)*; el molde para dicho protocolo se hizo a partir de una PCR (5'RACE) usando el adaptador T7 sólo en el oligo *reverse*. Luego de la amplificación se hizo una extracción fenólica y una precipitación isopropanólica para obtener un molde limpio.

El análisis de calidad de la sonda generada se realizó mediante un *dotblot*; antes de detallar los pasos del protocolo es necesario hacer referencia al significado de las letras R y W.

R: (*rinse*): enjuagar: sólo enjuagar los embriones con 1ml de solución nueva

W: (*wash*): lavar: agitar el tubo de modo que los embriones se muevan y mezclen y todas las superficies del tubo se laven. Luego dejar que precipiten e invertir el tubo brevemente para que caigan los embriones u ovarios que quedaron pegados a la tapa y paredes del tubo.

Pasos a seguir:

- 1) Realizar una serie de diluciones (1:10; 1:100; 1:1000; 1:10000) de la sonda obtenida y de la muestra control proporcionada por el *kit*.
- 2) R R W5 en PBT (en *ependorf* de 2ml)
- 3) Incubar con 1ml de anti-Dig 1:2000 en PBT por 10 min
- 4) R R W10 en PBT
- 5) R R con *Buffer* alcalino (1M Tris (PH 9.5), NaCl, 20%tween 20)
- 6) Teñir con NBT-BCIP

Una vez establecida la concentración de uso de la sonda se procede a llevar a cabo la reacción de hibridación que se detalla a continuación.

### 3.2 Hibridación

Como material de partida se utilizaron tanto ovariolos como embriones de *R. prolixus*, previamente fijados, los cuales se encontraban en tubos *Eppendorf* de 2 ml fondo cónico con PBT.

En los ítems 6 y 7, que se detallan a continuación, se describe como se procesaron dichas muestras.

El protocolo de hibridación propiamente dicho fue establecido en el laboratorio en un trabajo conjunto con la Lic. Nazar. A continuación, se detallan los pasos a seguir:

- 1- descongelar solución (sn) fijadora 10 min a 70°C en baño o estufa de secado
- 2- fijar en 500 µl PBT + 500 µl de sn. Fijadora (PH:7) x 20 min en agitación a temperatura ambiente (T°amb).
- 3-R R W5 R con PBT.
- 4- Digerir con Proteinasa K; incubar 15 min en 500 µl PBT + 0.5 µl de proteínasa K [10mg/ml]. (T°amb=25°C) en agitación 50 rpm
- 5- R R con PBT
- 6- pos-fijar de nuevo con 500 µl PBT + 500 µl de sn. fijadora (PH:7) x 20 min en agitación 50 rpm.
- 7- R R W5 R con PBT
- 8-Pre-hibridar en 500-1000 µl de Hybe. 2hs (horas) a 60° C.

9- Sacar todo el Hybe y reemplazar x nuevo Hybe + X  $\mu$ l sonda. Incubar ON (*overnight*) - aprox 18hs- a 60° C o 52 °C baño húmedo. (X hace referencia a la concentración de uso establecida como se detalló en el punto 3.1)

10- R x3 con Hybe-B a 60°C o 52°C (dependiendo de la temperatura que se utilizó para la incubación)

11- R con Hybe-C tantos como para cubrir al menos 2hs, a 52/60° C. Las incubaciones se hacen como se detalla en el punto anterior. Dejando Hybe-C a la temperatura de incubación

12- R R W5 W10 con PBT. Temperatura ambiente 50 rpm

13- Incubar en sn de Hibridación-Anticuerpo durante 2/3hs a 50 rpm temperatura ambiente

14- Revelar con Anti-Dig 1:2000 en PBT 4°C ON en agitación 50 rpm

15- W10 W10 W10 con PBT temperatura ambiente

16- R R W5 en *Buffer* alcalino

17- 2, 5- 3  $\mu$ L DE NBT/BCIP en *Buffer* alcalino. (en oscuridad y agitación suave 50 rpm). El tiempo varía según experimento, desde 30' hasta ON. A partir de los 30 min ya se evidencia tinción y luego hay que ir controlándolos hasta observar que no haya cambios en la coloración, esto se lleva a cabo bajo lupa ya que permite ver con mayor detalle.

17- detener la reacción con R R W20 con PBT.

18- Deshidratación seriada (50% - 75% - 100%) con metanol para almacenarlos.

Para realizar tomas fotográficas hay que rehidratarlos, ya que en metanol se mueven, dejándolos en PBT. Los ovarios y embriones que presentaron marcación diferencial se visualizaron en un microscopio ZEISS A1 y se realizó la captura de imágenes con una cámara acoplada al mismo.

#### 4. Inmunohistoquímica con fluorescencia

Como material de partida se utilizaron ovariolas de *R. prolixus*, previamente fijadas, los cuales se encontraban en tubos *Eppendorf* de 2 ml fondo cónico con PBT.

En los ítems 6 y 7, que se detallan a continuación, se describe como se procesaron dichas muestras hasta ser utilizadas.

Se utilizó un anticuerpo policlonal de conejo anti-vitelina (fracción gamma-globulina) generado contra la principal subunidad de la apoproteína de la vitelina (Vn) del triatómido *Dipetalogaster máxima* (Blariza et al., 2016). El anticuerpo fue cedido por la Doctora Lilián Canavoso del Departamento de Bioquímica Clínica, Centro de Investigaciones en Bioquímica Clínica e Inmunología (CIBICI-CONICET), Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Córdoba, Argentina.

El protocolo de tinción con anticuerpo fue puesto a punto durante mi trabajo de doctorado, los pasos requeridos fueron los siguientes:

- 1- W10 W10 con PBX (PBS 1X - 0.1% Triton X-100)
- 2- Se incubó durante 2hs con PBX + NGS 5% (solución bloqueo) en agitación 50 rpm
- 3- Se removió la solución bloqueo y agregar anticuerpo primario anti-vitelina (anti-Vn) con una dilución 1:200 en PBX. Se incubó *overnight* a 4°C en agitación 50 rpm.
- 4- Se recuperó la dilución del anticuerpo y practicó R R R con PBX
- 5- W30 W30 W30 con PBX
- 6- Se incubó durante 2hs con PBX + NGS 5% (solución bloqueo) en agitación 50 rpm
- 7- Se removió la solución de bloqueo y se incubó con el anticuerpo secundario anti-conejo del subtipo IgG acoplado al fluoróforo Alexa 568, en una dilución 1:500. La incubación fue en agitación 50 rpm en oscuridad y a temperatura ambiente.
- 8- R R R con PBX
- 9- W15 120 con PBX en agitación 50 rpm.
- 10- Agregar glicerol 50% con 0.1 µg/ml DAPI e incubar durante 15 min en oscuridad.
- 11- Remover el glicerol 50% y agregar glicerol 100% para proseguir con la toma de imágenes de las ovariolas.
- 12- . Las imágenes de los ovarios se adquirieron utilizando el microscopio confocal LSM 800 ZEISS.

## 5. ARN de interferencia (ARNi) parental

Se prepararon moldes mediante PCR usando oligos específicos con adaptadores conteniendo la parte principal del promotor T7. Los productos de PCR se usaron como molde de transcripción, de modo que se generaron las cadenas *sense* y *anti-sense* en una misma

reacción. La reacción de transcripción fue llevada a cabo por la T7 RNA polymerase de *New England Biolabs*, según protocolo establecido por la empresa. Una vez finalizada la reacción se prosiguió con una digestión con DNAsa I (*QIAGEN*). Por último, se realizó una extracción con solventes orgánicos y precipitación alcohólica. Se hicieron controles de calidad del amplicon generado en gel de agarosa 1% con BrEt y se prosiguió a la cuantificación usando el paquete FIJI (Schindelin et al., 2012).

Una vez generado el ARN doble cadena (ARNdc), inmediatamente se procedió a la microinyección de hembras adultas vírgenes, sin previa alimentación, siguiendo el protocolo establecido en (Lavore et al., 2012). Es preferible inyectar el menor volumen posible, entre 1 y 5µl. La microinyección se realiza en el abdomen, entre el 5° y 6° esternito abdominal, en la región medio-lateral para no dañar órganos, la aguja se coloca en un ángulo muy agudo y orientada desde atrás hacia adelante, paralela y cercana a la zona pleural, para disminuir la tasa de mortalidad como consecuencia de la perforación abdominal. A los dos días post-inyección la herida suele estar cicatrizada y ha disminuido el estrés, de modo que se realiza la alimentación y se las coloca con machos para estimular la copula y dar inicio a la oviposición de huevos fecundados. Como control negativo, se usó ARNdc de  $\beta$ -lactamasa (Araujo et al., 2007), gen de resistencia a ampicilina ausente en eucariotas, generado de la misma manera como se describió para los genes blanco a estudiar. La determinación génica funcional se evaluó a nivel del ovario, analizando la estructura del mismo, así como también a nivel de la ovipuesta, analizando la puesta de huevos embrionados y la forma y estructura de los huevos generados.

## 6. Manipulación de ovariolas

Las hembras adultas fueron alimentadas y aproximadamente 5 días después se prosiguió a disectar los ovarios. Considerando que el momento de remoción coincide con un estado de oogénesis activa, estimulada por la alimentación, los ovarios se aislaron utilizando pinzas de disección *Dumont* N°5, bajo lupa de micro-disección y sumergidos en *Buffer* fosfato salino 1X (PBS 1X), eliminando todo tejido extra ovárico presente. Luego se transfirieron a un portaobjeto excavado para continuar con la limpieza de manera más cómoda hasta proceder

a la fijación de los mismos. Se utilizaron diferentes estrategias de fijación dependiendo del subsecuente análisis.

### **6.1 Tinción con 4',6-diamino-2-fenilindol (DAPI) y faloidina**

Para el procesamiento de ovariolas que van a ser analizadas mediante el uso de fluoróforos, se procedió a fijar las mismas en paraformaldehído (PFA) 4% durante 30 min en hielo. Luego del periodo de fijación, los ovarios permanecieron en hielo y fueron lavados, 3 veces, con PBT [PBS 1X + 0.1% Tween-20]. Se almacenaron a 4°C en PBT hasta ser procesadas.

Para la tinción específica del ADN y actina presente en las ovariolas disectadas, se prosiguió a la incubación de las muestras en una solución de glicerol 50% con 0.1 µg/ml DAPI junto a una dilución 1:400 de sonda de faloidina conjugada al colorante fluorescente Texas Red de *ThermoFisher*. Las imágenes se adquirieron con un microscopio confocal LSM 800 ZEISS.

### **6.2 Análisis de secciones histológicas**

Las muestras fueron fijadas con PFA 4%. Luego se prosiguió con lavados en buffer fosfato y una deshidratación en serie graduada de etanol (76%, 96%, 100%); por último, se hizo un pasaje de etanol 100% a Xilol 100% y se procedió a embeber la muestra en parafina (Wanderley-Teixeira et al., 2006). Los ovarios fueron teñidos con hematoxilina-eosina y se procesaron con un micrótopo (*Leica*) a un grosor de 5 µm. Los preparados obtenidos se visualizaron en un microscopio ZEISS A1 y se llevó a cabo la captura de imágenes con cámara acoplada al mismo.

### **6.3 Microscopia de transmisión (TEM)**

Los ovarios fueron fijados con 2.5 % de glutaraldehído. Post fijación fueron incubados durante 60 min con tetróxido de osmio al 1% a 4°C. Luego se prosiguió a una deshidratación seriada de etanol (25%, 30%, 50%, 80%, 90%) y por último un pasaje final de acetona 100%. Luego fueron pre-infiltrados *overnight* (ON) en resina Durcupan, para luego si realizar la infiltración en la resina durante 24hs a 56°C, protocolo modificado a partir de Huebner and Anderson (1970). Los cortes ultra finos se realizaron con un grosor de aproximadamente 60nm, los cuales se contrastaron con acetato de uranilo y citrato de plomo. Por último, se observaron y fotografiaron utilizando MET JEOL 1200 EXII a 80Kv.

## 7. Manipulación de embriones

Para la manipulación de los embriones tempranos, previos al cierre dorsal se prosiguió con el siguiente protocolo de fijación establecido por la Lic. Nazar durante su trabajo final de grado, realizado en nuestro laboratorio.

Como primer paso se colocaron a los embriones en un tupo *Eppendorf* tapa a rosca. Luego se los inmergió en nitrógeno líquido (-196°C) durante 15 min. Se retiraron los tubos del nitrógeno líquido, se los colocaron en hielo y se prosiguió con la adición de 200 µl de agua destilada a cada tubo, asegurándose de que queden todos los huevos sumergidos. Se hirvieron durante 3 min e inmediatamente se los transfirió al hielo hasta aguardar a que se enfríen.

Bajo lupa de disección y de manera manual se procedió a la remoción del opérculo y corion de los embriones utilizando pinzas de disección *Dumont* N°5. La fijación se llevó a cabo en PFA 4% durante 15 min, a la que le siguieron 3 lavados en PBT para eliminar el remanente de la solución fijadora. Se continuó con una deshidratación seriada (50%-75%-100%) en metanol para almacenarlos hasta su análisis.

## 8. Manipulación de huevos

### 8.1 Microscopía electrónica de barrido

Para llevar a cabo un análisis por microscopía de barrido los huevos colectados fueron fijados en glutaraldehído 2.5%, luego se lavaron con *Buffer Millonig*. Se prosiguió con una deshidratación en serie en etanol (Sorrivas de Lozano et al., 2014). Por último, se realizó la metalización en oro de los huevos y se visualizaron en MEB Quanta 250 (FEI).

## 9. Análisis de distribución de lípidos y membranas

Para evaluar la integridad de las membranas celulares de las ovariolas se decidió utilizar una sonda lipofílica denominada FM4-64FX (*ThermoFisher*). Para llevar a cabo el experimento, se utilizaron tanto hembras control como hembras que habían experimentado silenciamiento génico. En el quinto día post-alimentación, las hembras se ubicaron en placas de disección sostenidas con dos alfileres desde sus extremos con el fin de mantenerlas inmovilizadas y al mismo tiempo evitar afectar algún órgano vital. Se les removió parte de su cutícula del lado ventral en la parte más terminal del adulto, y se le inyectó una dilución 1:500 de la sonda en

PBS 1X durante 20 min. Inmediatamente se realizaron lavados con PBS 1x para remover restos de la sonda y se procedió a disectar los ovarios bajo lupa con pinzas de disección *Dumont* N°5 y comenzar el proceso de fijación como previamente se detalló en el ítem 6.1, exceptuando la tinción de actina con faloidina. Las imágenes se adquirieron con un microscopio confocal LSM 800 ZEISS.

#### 10. Análisis del traspaso de nutrientes y moléculas durante la vitelogenénesis

Se utilizaron dos dextranos de diferente tamaño para evaluar la relación entre espacio intercelular y pasaje de moléculas. Por un lado, se utilizó un dextrano de un peso molecular de 10 kiloDaltons (kDa) marcado con el colorante Texas Red, y otro con un peso molecular de 70 kDa conjugado con fluoresceína.

Para llevar a cabo el experimento, se utilizaron tanto hembras control como hembras que habían experimentado silenciamiento génico. En el cuarto día post-alimentación se les inyectó en un volumen final de 2  $\mu$ l correspondiente a una mezcla de dextranos, 1  $\mu$ l del dextrano de 10kDa y 1  $\mu$ l del dextrano de 70kDa. La microinyección se realizó en el abdomen, entre el 5° y 6° esternito abdominal, en la región medio-lateral para no dañar órganos. Se aguardó 24hs y se procedió a la disección de las ovarias como se describió previamente en el punto 6. Debido a que los dextranos están acoplados a fluoróforos todo el proceso de fijación se llevó a cabo en oscuridad y como se detalló en el ítem 6.1. Las imágenes se adquirieron con un microscopio confocal LSM 800 ZEISS.

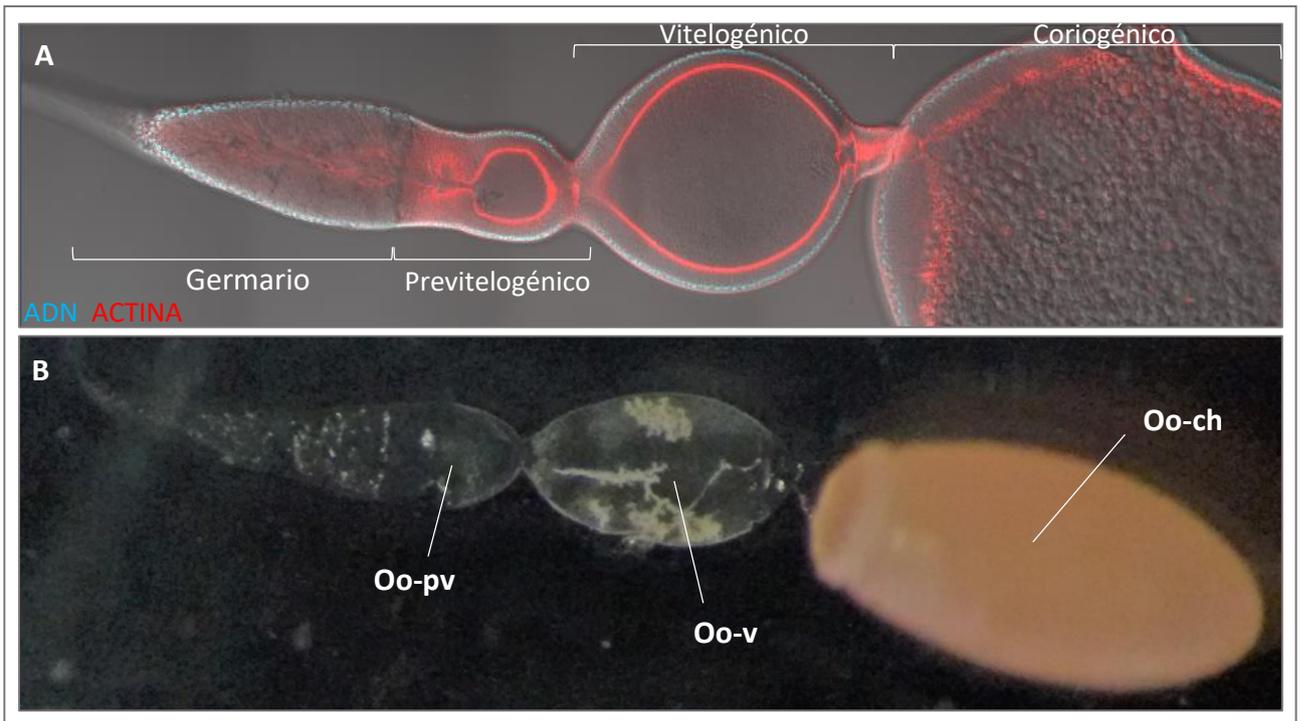
# Resultados

## 1. Descripción de la estructura y morfología del ovario con microscopía de fluorescencia

### 1.1 Re-descripción de la ovariola de *Rhodnius prolixus*

El ovario de *R. prolixus* es telotrófico. En él se distinguen células nodrizas en la zona del germario, las cuales se encuentran en conexión con los oocitos en desarrollo a través de cordones citoplasmáticos cumpliendo una función principalmente nutritiva, y rodeando a cada oocito en desarrollo se encuentran las células foliculares.

Como se mencionó previamente, durante las décadas de los 70 y 80, principalmente Erwin Huebner, realizó un variado estudio descriptivo sobre la morfología y estructura del ovario en nuestro organismo modelo de estudio, por lo que en este trabajo se decidió tomar todo su aporte y llevar a cabo un registro fotográfico con microscopía confocal y tinciones con fluorescencia para complementar su descripción, utilizando DAPI, que es un marcador fluorescente que se una fuertemente a regiones enriquecidas de adenina y timina en secuencias de ADN y faloidina, que es una sonda de actina F conjugada al colorante Texas Red (Figura 20).

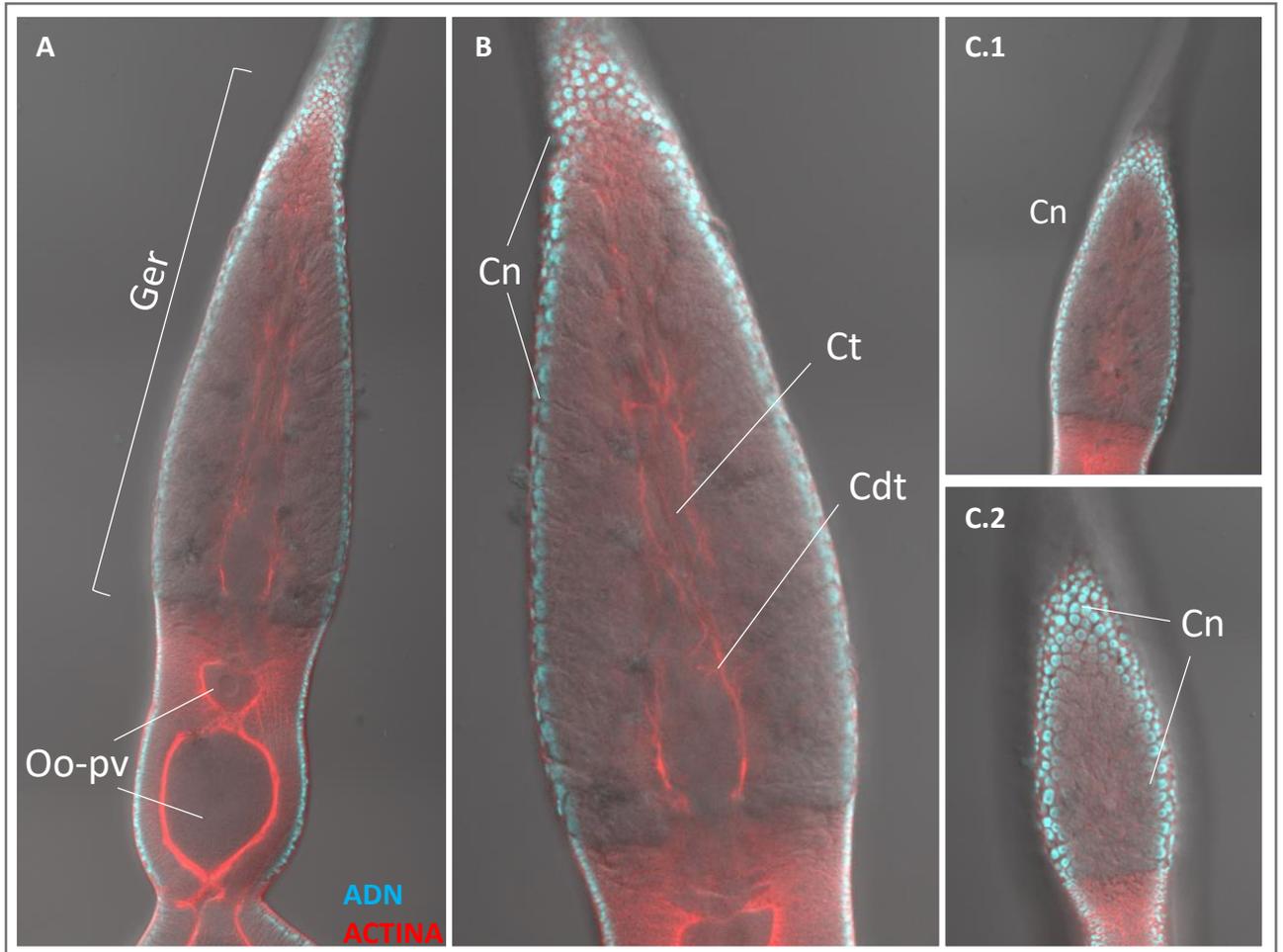


**Figura 20:** Ovariola de *Rhodnius prolixus*. En la parte **A** se visualiza la ovariola teñida con DAPI (ADN) y faloidina (actina) registrada con microscopía confocal. En la parte **B** se fotografía la ovariola registrada con una cámara acoplada a microscopio óptico. Referencias: Oocito previtelotogénico (Oo-pv), Oocito vitelotogénico (Oo-v), Oocito coriolotogénico (Oo-ch).

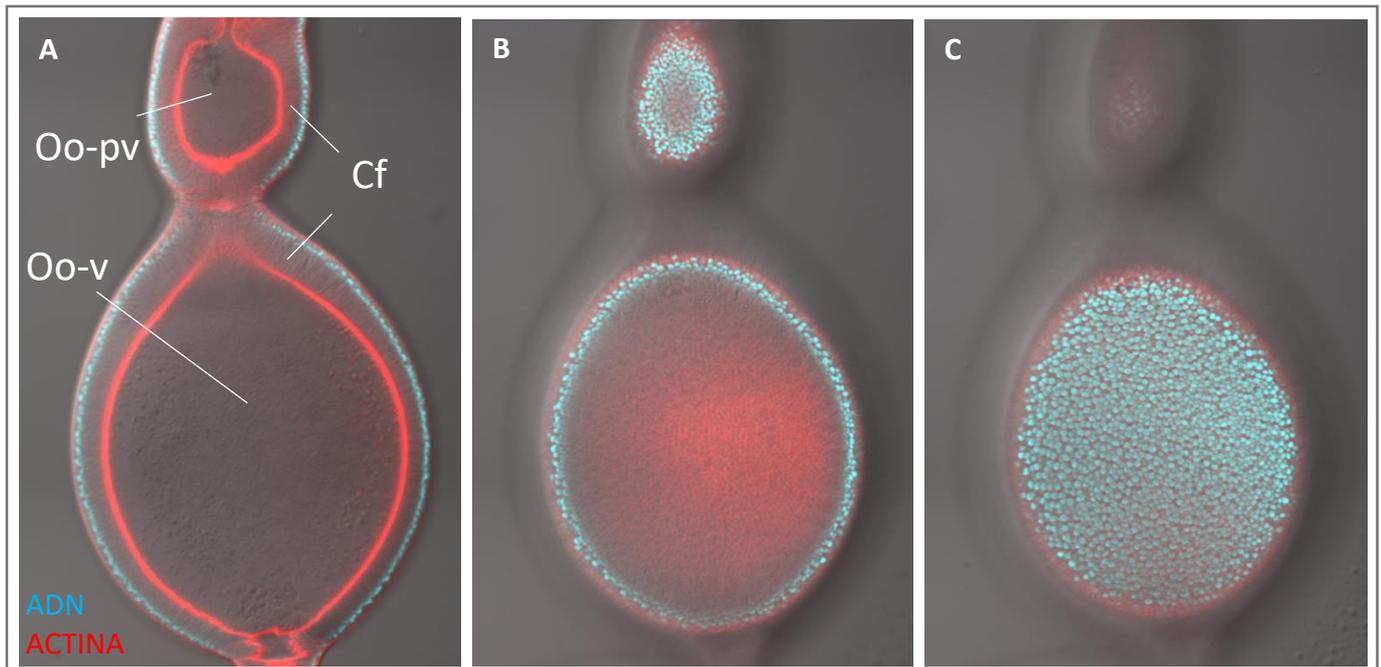
Comenzando desde el germario, la estructura más anterior de la ovariola, se pudo observar de manera indiscutible, una región central sin la presencia de núcleos, evidenciado por la ausencia de tinción nuclear dada por el DAPI. Esta región central si es rica en actina, donde se observa un centro o corazón trófico desde donde se proyectan puentes citoplasmáticos que se encuentran en conexión con los oocitos en desarrollo. A estos puentes se los conoce como cordones tróficos, y mantienen esta comunicación hasta que se completa el desarrollo del oocito (Figura 21).

A su vez, se diferenció en la zona apical y superficial del germario, distintos tamaños nucleares correspondientes a las células nodrizas, observándose que los que se encuentran en el extremo apical presentan un menor tamaño. De manera gradual el tamaño nuclear aumenta hacia la región más posterior del germario, observándose una zona media de transición, encontrando un tamaño nuclear variado (Figura 21).

En el vitelario, ya es posible observar a los oocitos en desarrollo, donde se diferencian los distintos estadios correspondientes a su progreso en el proceso en relación a su tamaño celular, ubicándose los menos desarrollados y más pequeños en la parte más próxima al germario, denominados oocitos previtelogénicos (Figura 22). Los oocitos más avanzados en el desarrollo presentan un mayor tamaño ubicados en la zona más terminal de la ovariola. Acompañando a estos oocitos en desarrollo, se pueden observar los núcleos correspondientes a las células foliculares que los rodean en toda su superficie, en una disposición organizada desde estadios inmaduros y manteniéndose durante todo el desarrollo del oocito hasta que tiene lugar la formación del corion (Figura 22). En la figura 23 se observa ya el oocito coriogénico donde es posible diferenciar dos patrones nucleares correspondientes a las células foliculares laterales y apicales, esta disposición especial en la parte apical del oocito son las que dan lugar al opérculo, por donde posteriormente la larva va a emerger cuando complete su desarrollo embrionario.

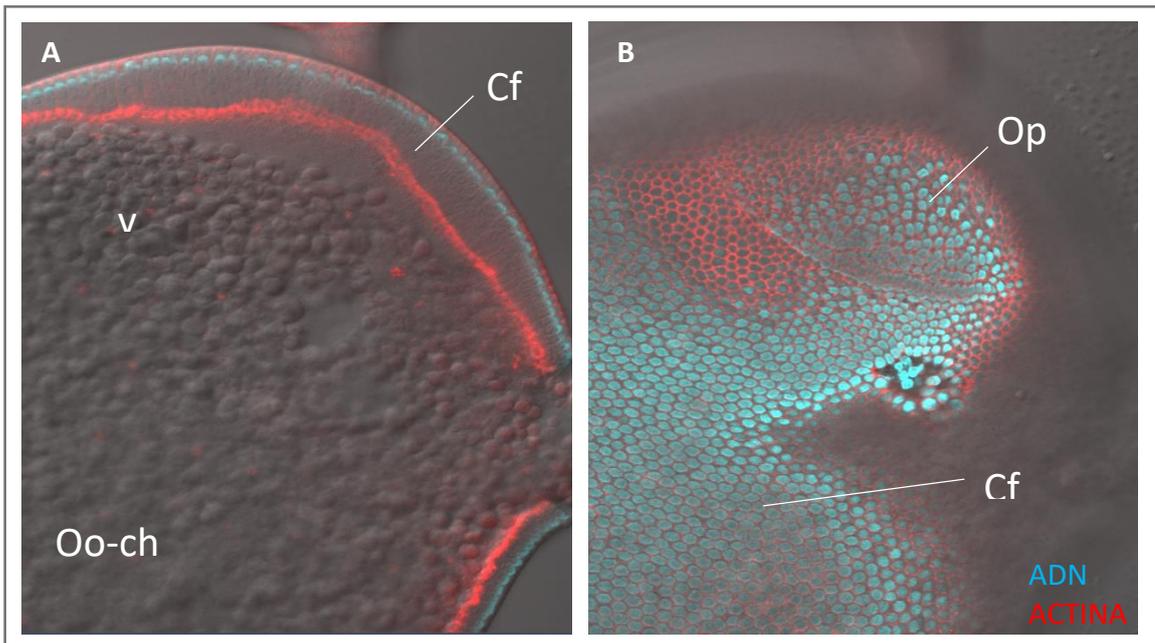


**Figura 21:** Germario del ovario de *Rhodnius prolixus*. Se observa en detalle la estructura del germario (A-B), logrando identificar a las células nodrizas (Cn), al corazón trófico (Ct) y a los cordones tróficos (Cdt). En las imágenes C.1 y C.2, se visualizan los distintos tamaños nucleares correspondientes a las células nodrizas. La tinción con DAPI hace visible el ADN y la faloidina marca regiones ricas en actina. Los oocitos no forman parte del germario, pero se señalan en la figura A, en su estadio previtelogénico (Oo-pv).



**Figura 22:** Oocitos en desarrollo en el ovario de *Rhodnius prolixus*. En las tres imágenes **(A-B-C)** se observa en distintos niveles de cortes longitudinales la estructura de los oocitos que conforman la ovariola de *R. prolixus*. La tinción con DAPI hace visible el ADN y la faloidina marca regiones ricas en actina. En los distintos niveles es posible identificar a las células foliculares que rodean al oocito en desarrollo desde un estado previtelogénico (Oo-pv) hasta vitelogénico (Oo-v). Referencia: Células foliculares (Cf)

Es importante remarcar que el patrón folicular que se imprime desde los estadios tempranos del desarrollo del oocito, se mantienen y acentúan en el estado coriogénico del mismo, donde se reproduce un patrón hexagonal, a veces pentagonal, característico de la superficie externa del corion del huevo ovipuesto.



**Figura 23:** Oocito coriogénico del ovario de *Rhodnius prolixus*. **A.** Se observa en detalle la estructura del oocito coriogénico (Oo-ch), logrando visualizar un alto contenido de vitelo (v). En la parte **B** es posible diferenciar las células foliculares (Cf) que se encuentran en la parte lateral del huevo, como las que conforman el opérculo (Op). La tinción con DAPI hace visible el ADN y la faloidina marca regiones ricas en actina.

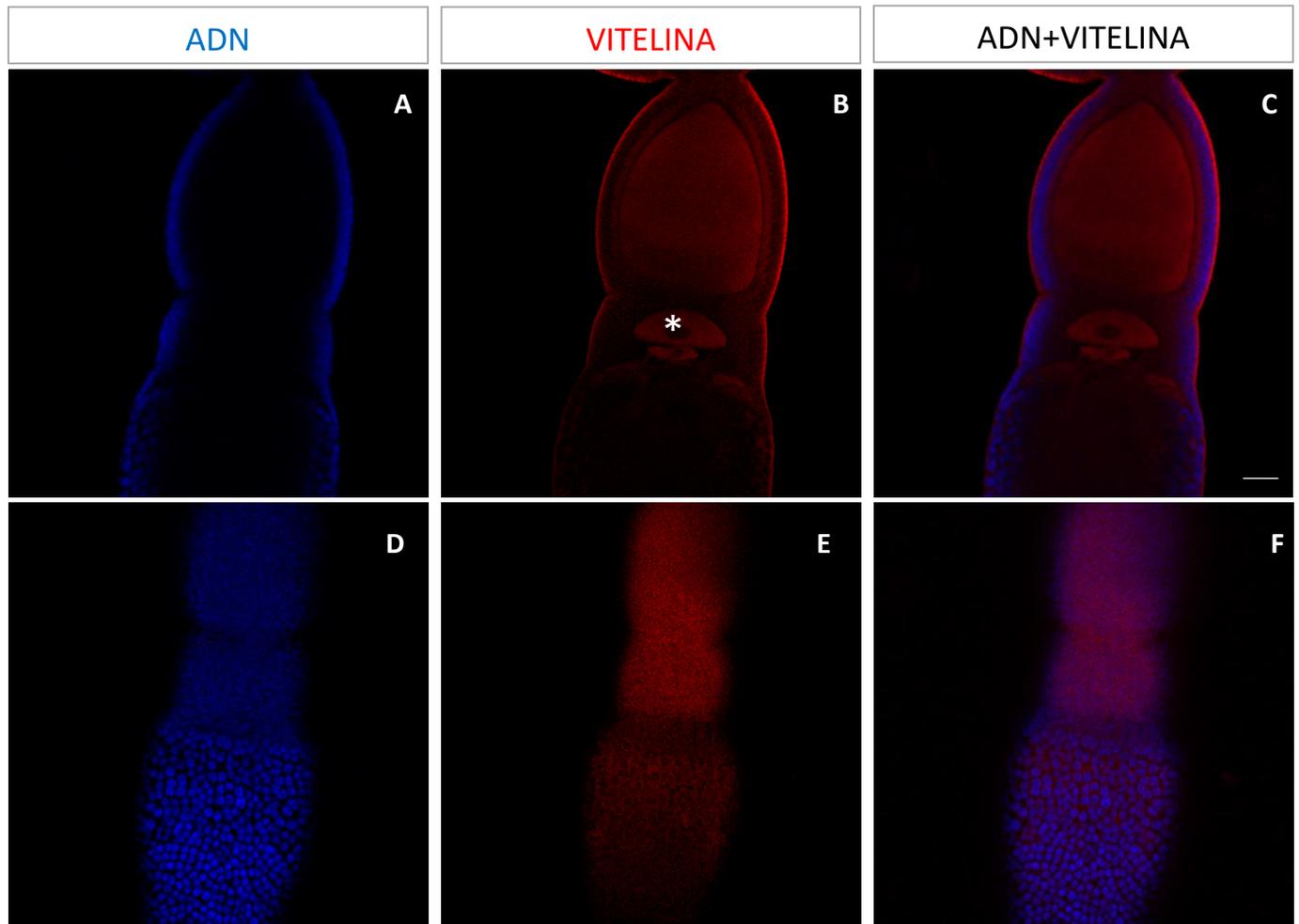
### 1.2 Acumulación de vitelogenina/vitelina durante la vitelogénesis en la ovariola de *R. prolixus*

Para analizar con mayor detalle el proceso de ovogénesis, se decidió estudiar la toma de nutrientes de origen extra ovárico, por el oocito en desarrollo. Ante esto, se utilizó un anticuerpo contra vitelina para la identificación de su localización en el ovario de hembras adultas de *R. prolixus*.

El anticuerpo reconoce con especificidad, tanto al precursor proteico denominado vitelogenina como al producto ya procesado acumulado en el oocito denominado vitelina. Para contextualizar de manera más sencilla la localización de dicha proteína se utilizó el marcador fluorescente DAPI como guía de referencia.

Con lo que respecta a la presencia de vitelogenina/vitelina en los oocitos previtelogénicos se evidenció la presencia de la proteína en los oocitos desde muy temprano en el desarrollo, con una marcada localización de vitelogenina en el lado apical de la membrana de las células foliculares, así como también en la membrana del oocito en desarrollo con una disposición

que permitió discernir entre lado apical y basal de la mismo. Además, se localizó a la proteína vitelina con una señal mucho más débil en el vitelo de los oocitos. En la Figura 24 se hace evidente dicha disposición previamente descrita.

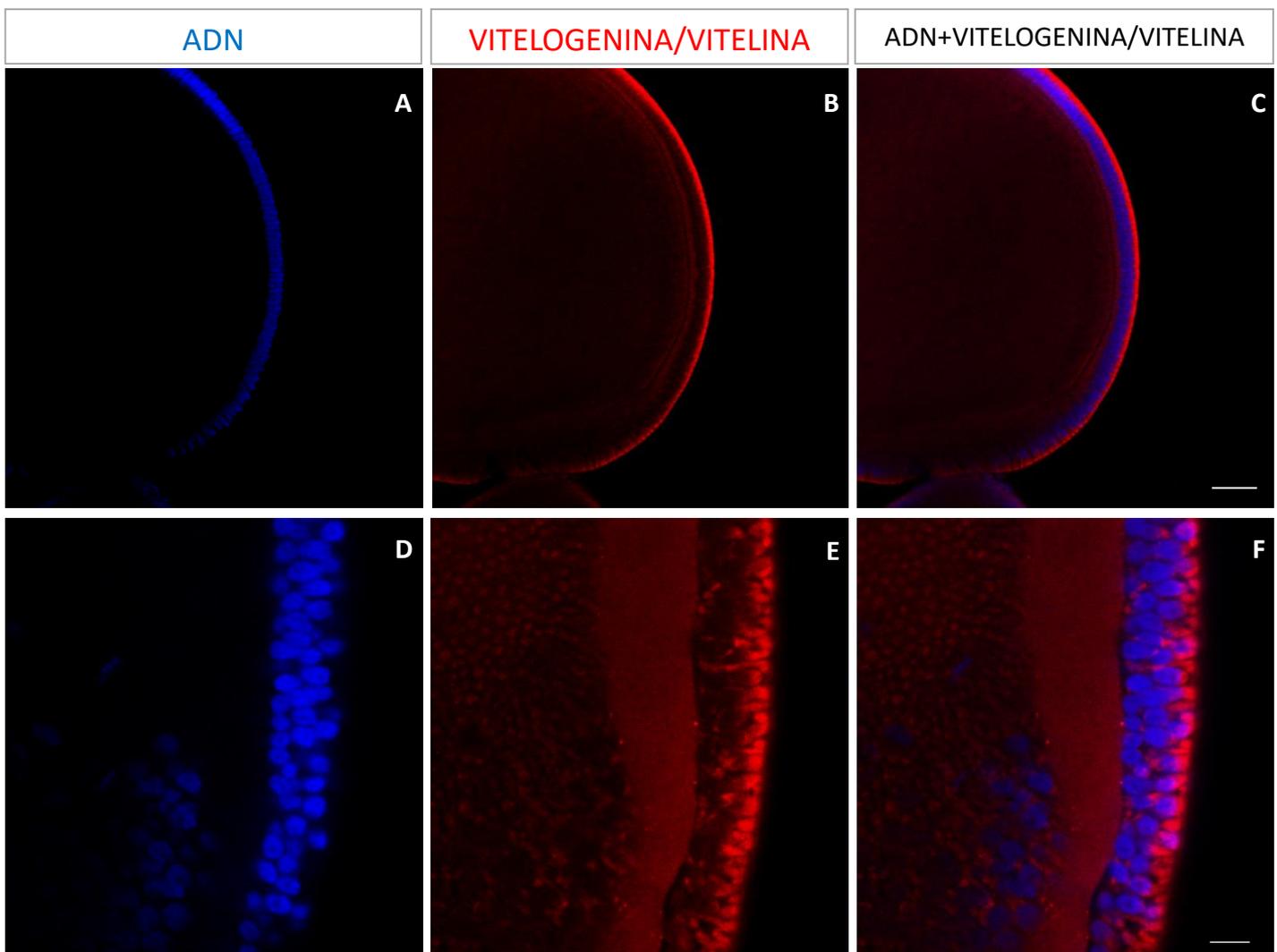


**Figura 24:** Oocitos en estado previtelogénico en ovariola de *Rhodnius prolixus*. En las imágenes **A** y **D** se visualiza el oocito en estado pre-vitelogénico teñido con DAPI (ADN), mostrando los núcleos correspondientes a las células foliculares que rodean a los oocitos en el vitelario, así como los núcleos de las células nodrizas que conforman el germario. En las imágenes **B** y **E** se visualiza la localización de la proteína vitelina, observando una marcada presencia en el lado apical de las células foliculares, así como también una localización en la membrana y en el vitelo de los oocitos en desarrollo. En **C** y **F** se observa la tinción tanto para ADN como para la proteína vitelina. La diferencia entre el panel superior (**A-B-C**) y el inferior (**D-E-F**) se debe al distinto plano focal utilizado para la captura de las imágenes utilizando microscopía confocal. Escala: 50  $\mu\text{m}$ . El \* indica la localización de la vesícula germinal o núcleo en el oocito previtelogénico.

Más avanzado el proceso de oogénesis, se observó al oocito en estado vitelogénico con el mismo patrón de localización observado en la fase previtelogénica pero con una señal más fuerte (Figura 25). Se observó a la proteína vitelogenina en la parte apical de la membrana de

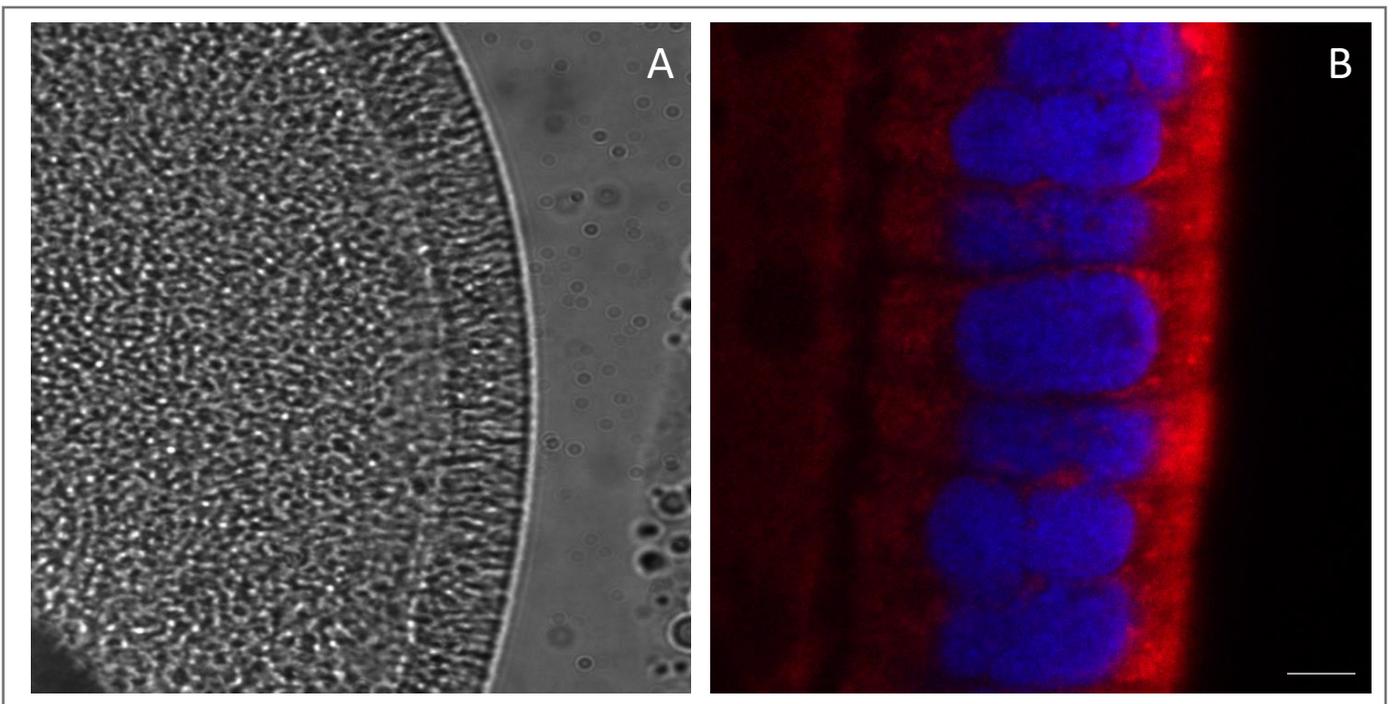
las células foliculares, así como también en la membrana oocítica; además se hizo evidente la presencia de vitelina en el vitelo del oocito localizada en los cuerpos vitelínicos característicos del oocito en fase vitelogénica (Figura 25E-F).

Por otro lado, cabe remarcar que no existió superposición en la localización de la proteína y la presencia del material genético correspondiente al núcleo de las células foliculares.



**Figura 25:** Oocitos en estado vitelogénico en ovariola de *Rhodnius prolixus*. En las imágenes **A** y **D** se visualiza el oocito en estado vitelogénico teñido con DAPI (ADN), mostrando los núcleos correspondientes a las células foliculares que rodean a los oocitos en el vitelario. En las imágenes **B** y **E** se visualiza la localización de la proteína vitelina, observando una marcada presencia en el lado apical de las células foliculares, así como también una localización en la membrana y en el vitelo de los oocitos en desarrollo. En la imagen **E** se observa la presencia de la vitelina en los cuerpos vitelínicos de los oocitos vitelogénicos. En **C** y **F** se observa la tinción tanto para ADN como para la proteína. La diferencia entre el panel superior (**A-B-C**) y el inferior (**D-E-F**) se debe al distinto objetivo de magnificación utilizado. Escala panel superior 50  $\mu$ m. Escala panel inferior: 10  $\mu$ m. La captura de las imágenes fue utilizando microscopía confocal.

Observando con mayor detalle el epitelio constituido por las células foliculares que rodea al oocito vitelogénico se observó la presencia de células foliculares binucleadas con un tamaño nuclear prominente dentro de la célula. Por otro lado, se registró una distribución desigual de la proteína de interés, con lo que respecta al lado basal-apical de la célula folicular, observándose una fuerte señal en el lado apical correspondiente con la localización de vitelogenina en la ovariola (Figura 26).



**Figura 26:** Células foliculares de oocito vitelogénico de *Rhodnius prolixus*. En la parte **A** se visualiza el oocito con microscopía de contraste diferencial interferencial (DIC) y en la parte **B** la imagen del oocito es obtenida por microscopía confocal, donde se observa la localización de la proteína vitelina con rojo y la presencia de material genético con azul marcado con DAPI. Escala: 5  $\mu$ m

## 2. Validación de los genes implicados en la oogénesis y desarrollo temprano previamente identificados

A partir del amplio repertorio de genes identificados en los transcriptomas embrionarios analizados en la parte I de este trabajo, se decidió realizar una búsqueda funcional de esas identidades inferidas y comenzar a trazar las relaciones génicas que dirigirían el proceso de oogénesis en *R. prolixus*. Debido a que no fue posible inferir alguna relación entre la presencia del transcrito y tiempo del desarrollo la elección de los genes a analizar se basó plenamente en el azar.

Los genes que se decidieron estudiar fueron *Bicaudal C (BicC)*, *Bicaudal D (BicD)* y *cornichon (cni)*. Para cada uno se diseñaron oligonucleótidos específicos en los transcritos correspondientes a las identidades inferidas.

Dependiendo del experimento con el que se prosiguió el análisis génico, hubo casos en los que fue necesario adicionarle al amplicon específico del gen de interés la secuencia correspondiente al promotor de la ARN T7 polimerasa, en uno o ambos extremos como se detalló en el apartado correspondiente a materiales y métodos.

## **2.1 *Bicaudal C***

### **2.1.1 Validación funcional por ARNi parental**

Para determinar el rol de *Bicaudal C (BicC)* en *Rhodinus prolixus (Rp- BicC)*, se generó ARNdc específico del gen a partir de dos amplicones diferentes generados. Por un lado, se partió de un amplicon de 237 pb generado utilizando el par de *primer* nombrados como PrFw1BicC y PrRv1BicC; y por el otro un amplicon de 291 pb generado utilizando el par de oligos PrFw2BicC y PrRv2BicC. Debido a que el amplicon generado con el par de *primers* 2 presentaba una longitud mayor se decidió iniciar el silenciamiento génico a partir de este fragmento obtenido. Se inyectaron 12 hembras vírgenes con diferentes dosis de ARN doble cadena (ARNdc) específico para el gen ARNdc<sup>BicC2</sup>. Como control negativo de la técnica se inyectaron hembras vírgenes con  $\beta$ - lactamasa de *Escherichia coli* (ARNdc <sup>$\beta$ -lac</sup>). En la tabla 14, en el panel de color rosa pálido se describe la dosis específica para cada inyección y los fenotipos observados; se evaluó tanto fertilidad como fenotipos en el ovario, en el huevo y en el embrión. Para lograr tal objetivo se decidió proseguir con dos estrategias de trabajo, por un lado, se tomaron algunas hembras inyectadas con ARNdc<sup>BicC</sup> y se le procedió a disectar sus ovarios (durante su estado ovipleno: 5 días post alimentación) y por el otro, a las remanentes se las dejaron para que continúen con el proceso de ovipuesta normal y se observó la descendencia.

	Dosis de ARNdc	Número de hembras inyectadas	Oviposición (N° de huevos)	Fenotipo a nivel del ovario	Fenotipo a nivel del huevo	Fenotipo embrional
ARNdc <i>BicC2</i>	0,8 µg	4	5	4/4 (100%)	5/5 (100%)	5/5 (100%)
	1,5 µg	3	50	3/3 (100%)	50/50 (100%)	50/50 (100%)
	2 µg	4	67	4/4 (100%)	56/67 (83.5%)	67/67 (100%)
	2,5 µg	1	24	1/1 (100%)	24/24 (100%)	24/24 (100%)
ARNdc <i>BicC1</i>	0,77 µg	2	16	2/2 (100%)	16/16 (100%)	16/16 (100%)
	1,2 µg	1	6	1/1(100%)	6/6 (100%)	6/6 (100%)
	2 µg	2	37	2/2 (100%)	32/37 (86,48%)	37/37 (100%)
ARNdc <i>βlac</i>	1 µg	2	72	0/2	0/72	5/72 (6,94%)

**Tabla 14:** Representación en números de los resultados obtenidos tras el silenciamiento génico de *Bicaudal C* (*BicC*). El panel de color rosa pálido hace referencia al ARNdc generado a partir del uso del primer par de oligos (*BicC2*) y el panel de color violeta claro hace referencia al uso del segundo par de oligos (*BicC1*). Las columnas muestran la cantidad en µg de ARNdc específico inyectado en hembras vírgenes y los resultados obtenidos a partir de cada oviposición. Las características del silenciamiento génico para ARNdc *BicC* se comparan con el control (ARNdc *βlac*)

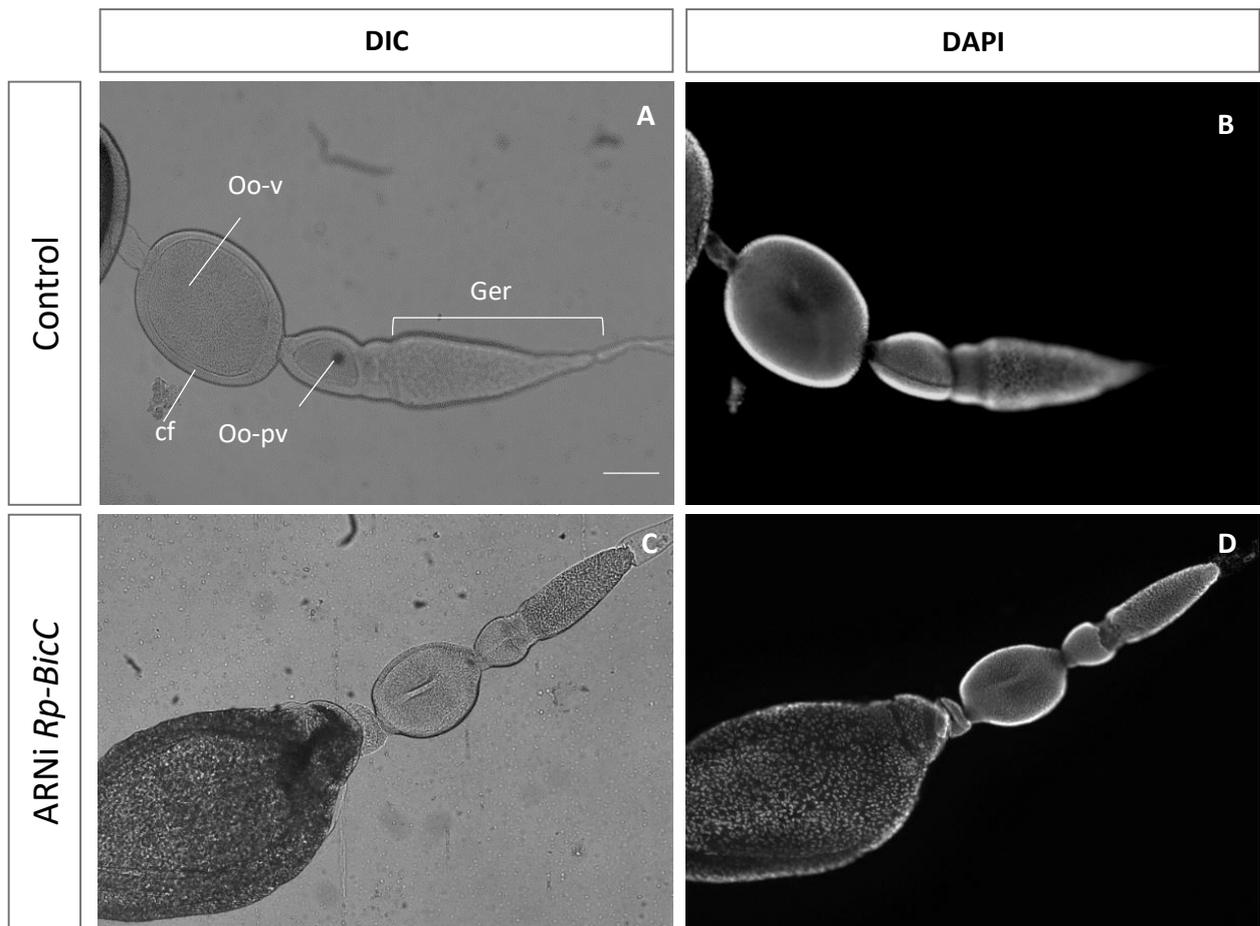
A pesar de utilizar diferentes dosis de ARNdc<sup>*BicC*</sup> se observó el mismo fenotipo en todas las concentraciones analizadas. De los 146 huevos puestos por el total de hembras interferidas sólo 11 no mostraron un fenotipo a nivel de la estructura del huevo. Estos huevos coinciden con la primer oviposición de las hembras vírgenes que se explica por su característica de autogenia descrita previamente. Respecto a la fertilidad, todas las hembras fueron infértiles luego de la inyección con ARNdc<sup>*BicC2*</sup>, ya que ovipusieron huevos, pero no viables para poder continuar con normalidad el desarrollo, además comenzaron la ovipuesta aproximadamente 10 días luego de la alimentación siendo que en el control *Rp-βlac* se observó puesta de huevos en aproximadamente 5 días post alimentación. A simple vista los huevos presentaron un tamaño y coloración diferente a los obtenidos por el control, perdieron su color rosa característico y se los observó de color blanquecino y más pequeños. Al disectar las ovariolas de las hembras interferidas se logró identificar grandes diferencias morfológicas en comparación al control, principalmente se observó un menor tamaño de los oocitos que componen a la ovariola de las hembras silenciadas genéticamente (Figura 27). Para abordar

estas diferencias con mayor detalle se decidió analizar la morfología y estructura tanto de los ovarios como de los huevos ovipuestos con técnicas específicas.



**Figura 27:** Ovarios y huevos de *Rhodnius prolixus* observados en la lupa estereoscópica y fotografiados con cámara acoplada a la misma. **A.** Ovario disectado de hembra control **B.** Ovario disectado de hembra interferida con ARNdc<sup>Bic2</sup>. **C.** Huevos ovipuestos de hembras interferidas y control.

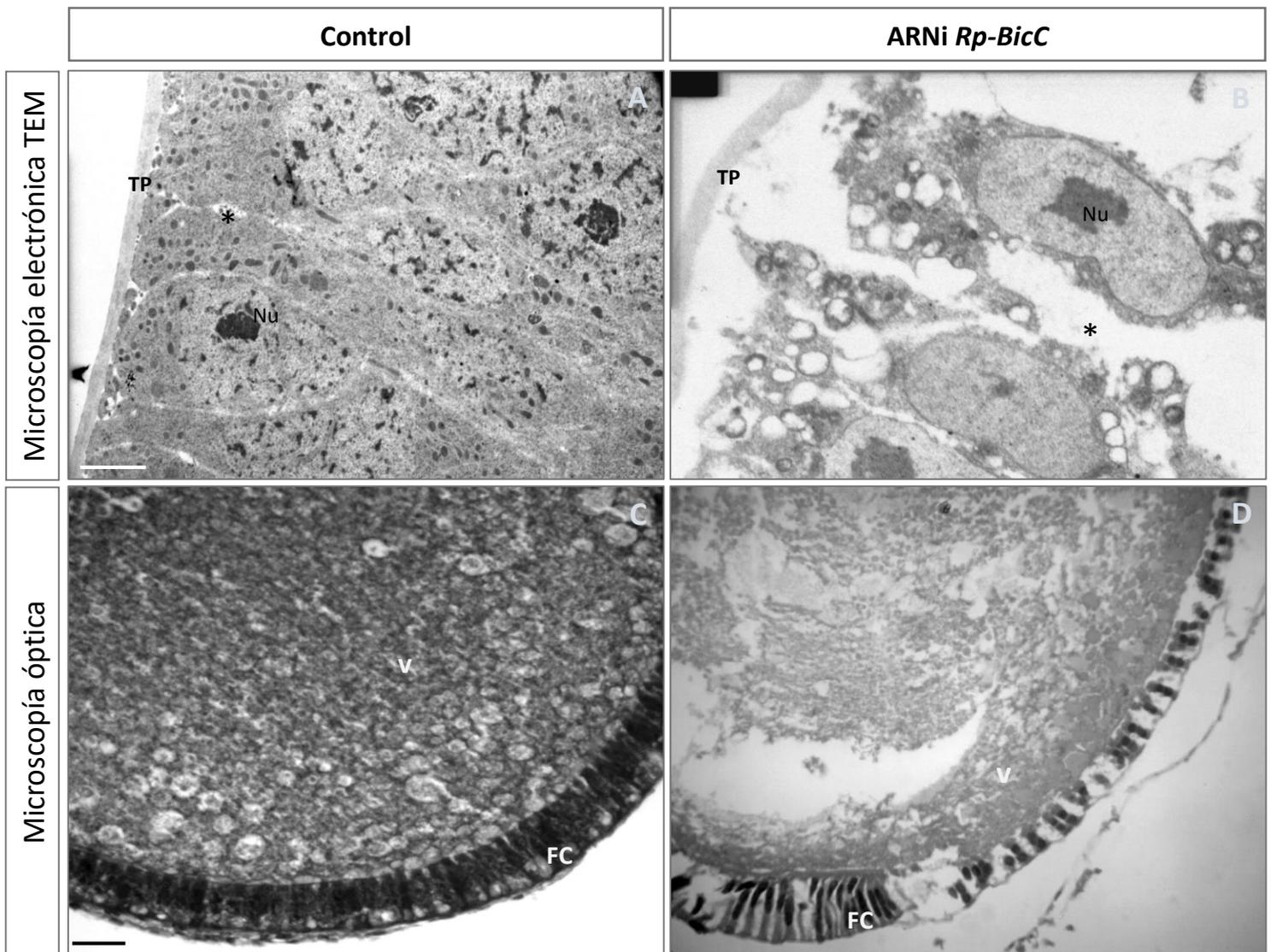
Para analizar la estructura del ovario *Rp-BicC*, en primer lugar, se decidió realizar tinciones con DAPI tanto en ovarios *Rp-βlac* como *Rp-BicC* para identificar el patrón nuclear en ambas situaciones. Como se puede observar en la Figura 28, el ovario correspondiente a *Rp-βlac* mostró una disposición organizada de los núcleos correspondientes a las células foliculares, rodeando a cada uno de los diferentes estadios de los oocitos en desarrollo. En cambio, para el ovario de *Rp-BicC* se observó una disposición desorganizada de estos, mostrando una especie de invaginación dentro del oocito en desarrollo, imitando a un globo desinflado. En lo que respecta al germario, la tinción con DAPI no permitió encontrar diferencias entre las dos condiciones observadas. En la misma figura se observan las ovariolas fotografiadas utilizando microscopía de contraste diferencial interferencial, que permite brindar información morfológica sin la utilización de tinte alguno. Estas imágenes permitieron corroborar las diferencias observadas a partir del patrón de marcación nuclear descrito previamente, observándose un oocito vitelogénico de menor tamaño con una disposición de las células foliculares de manera irregular en toda la superficie.



**Figura 28:** El silenciamiento de *Bicaudal C* afecta la morfología normal de los ovarios de *R. prolixus*. Ovariolas disectadas de hembras previamente inyectadas con ARNdc<sup>*β-lac*</sup> (control) o con ARNdc<sup>*BicC*</sup>. La disección fue llevada a cabo 5 días post alimentación. La figura **A/C** muestran a la ovariola obtenida utilizando microscopía de contraste diferencial interferencial (DIC) y en **B/D** la ovariola obtenida por microscopía confocal, utilizando DAPI para marcar el ADN presente. Escala: 200 μm. Referencias: Germario (Ger), Células foliculares (cf), Oocito- vitelogénico (Oo-v), Oocito-previtelogénico (Oo-pv).

Para evaluar esa disposición celular irregular en el ovario *Rp-BicC* se decidió realizar una introspección con microscopía electrónica de transmisión, así como también se efectuaron cortes histológicos de ovariolas teñidas con H&E para percibir en detalle esas diferencias foliculares. Como resultado se evidenció la irregularidad predicha con la tinción con DAPI, los oocitos de ovarios *Rp-BicC* en la fase de previtelogénesis fueron analizados con microscopía electrónica de transmisión y fue posible observar un incremento en el espacio intercelular entre las células foliculares y, entre las células foliculares y la tunica propia. Se observó que los espacios generados en la mayoría de las situaciones superaron en dimensión al tamaño celular (Figura 29). En lo que respecta a la forma de las células foliculares previtelogénicas,

esta se vio considerablemente afectada, así como también el tamaño de su núcleo celular. En el panel inferior de la Figura 29 se observa con detalle las características de los ovarios obtenidos tras silenciamiento en comparación con los ovarios control. Los oocitos de ovarios *Rp-BicC* en la fase de vitelogénesis se analizaron efectuando cortes histológicos de la corteza y la membrana plasmática de ovariolas teñidas con H&E. Como resultado de este análisis se pudo darle aún más peso a la desorganización que se anticipó previamente utilizando las distintas metodologías. Las células foliculares en los ovarios *Rp-BicC* mostraron una disposición no homogénea, repitiéndose esos espacios intercelulares que se evidenciaron en el estadio previtelogénico analizado con TEM. Sumado a esto, la forma celular se caracterizó por ser irregular, algunas presentaron la forma alargada característica en este estadio y otras mostraron una reducción del tamaño tanto en lo largo como en lo ancho de su superficie. A su vez, tanto el lado basal como apical de estas células mostraron una disposición irregular a lo largo de todo el oocito que envuelven.

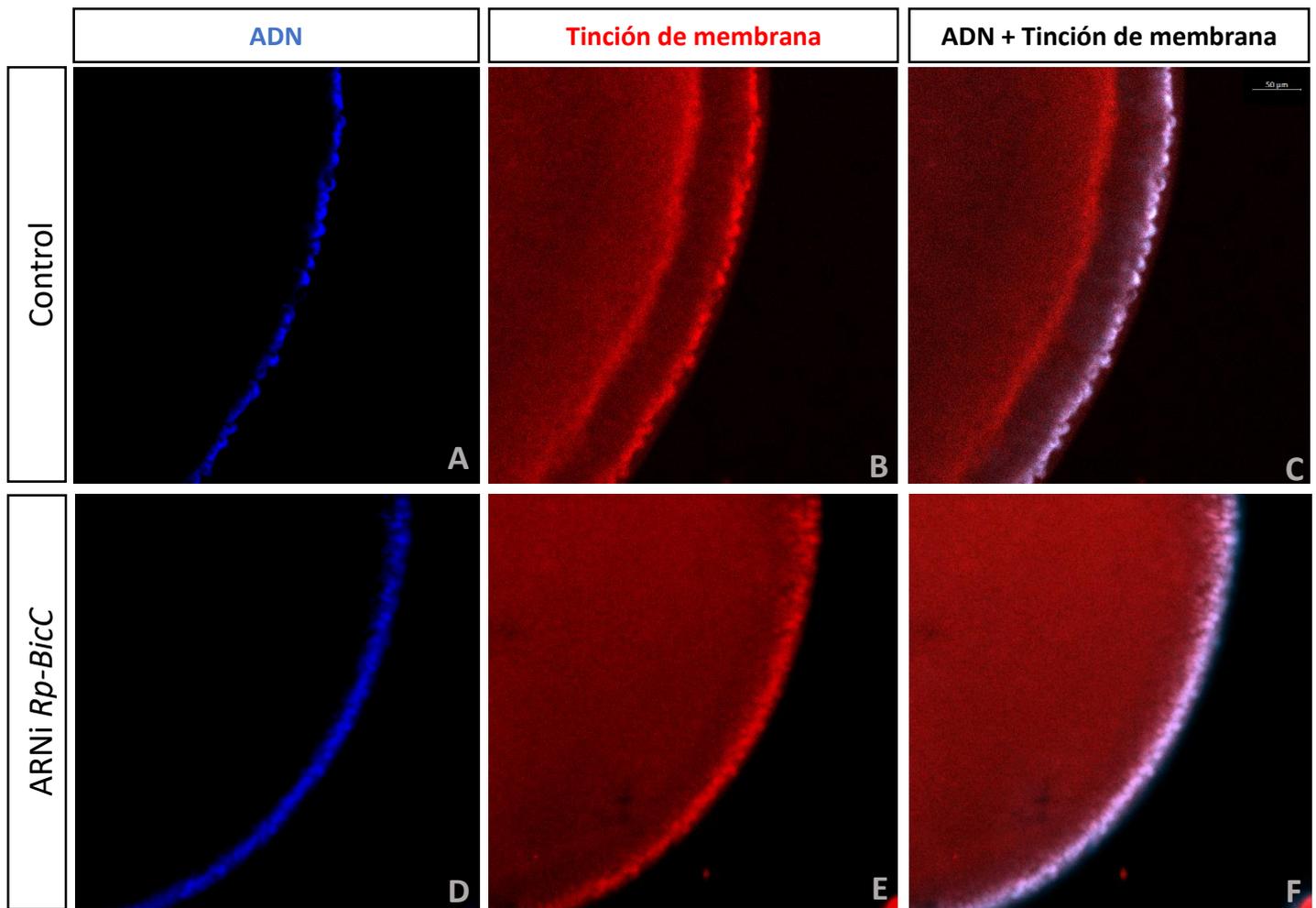


**Figura 29:** La interferencia de *Bicaudal C* produce un incremento en el espacio intercelular de las células foliculares. Imágenes representativas de cortes histológicos de la membrana plasmática y corteza de oocitos previtelogénicos y vitelogénicos de hembras control (panel izquierdo) y hembras silenciadas (panel derecho). En el panel superior (A-B) se observan oocitos de ovarios fotografiados con TEM. En el panel inferior (C-D) se observan oocitos de ovarios teñidos con H&E visualizados a través de microscopía óptica. Referencias: FC Células foliculares, V vitelo, TP tunica propia, Nu nucléolo, \* Espacios intercelulares. Escala TEM: 2  $\mu$ m. Escala microscopía óptica: 10  $\mu$ m.

Ante este cambio morfológico estructural en las ovarias y la observación de la puesta de huevos blancos y pequeños, se decidió analizar si se veía afectada la integridad de las membranas de las células foliculares, así como también la toma de vitelo durante la vitelogénesis.

Para el análisis de la distribución de los lípidos de membrana se utilizó una sonda lipofílica denominada FM4-64FX, a las hembras control y silenciadas se les inyectó una dilución 1:500 de la sonda en PBS 1X a los 5 días post alimentación.

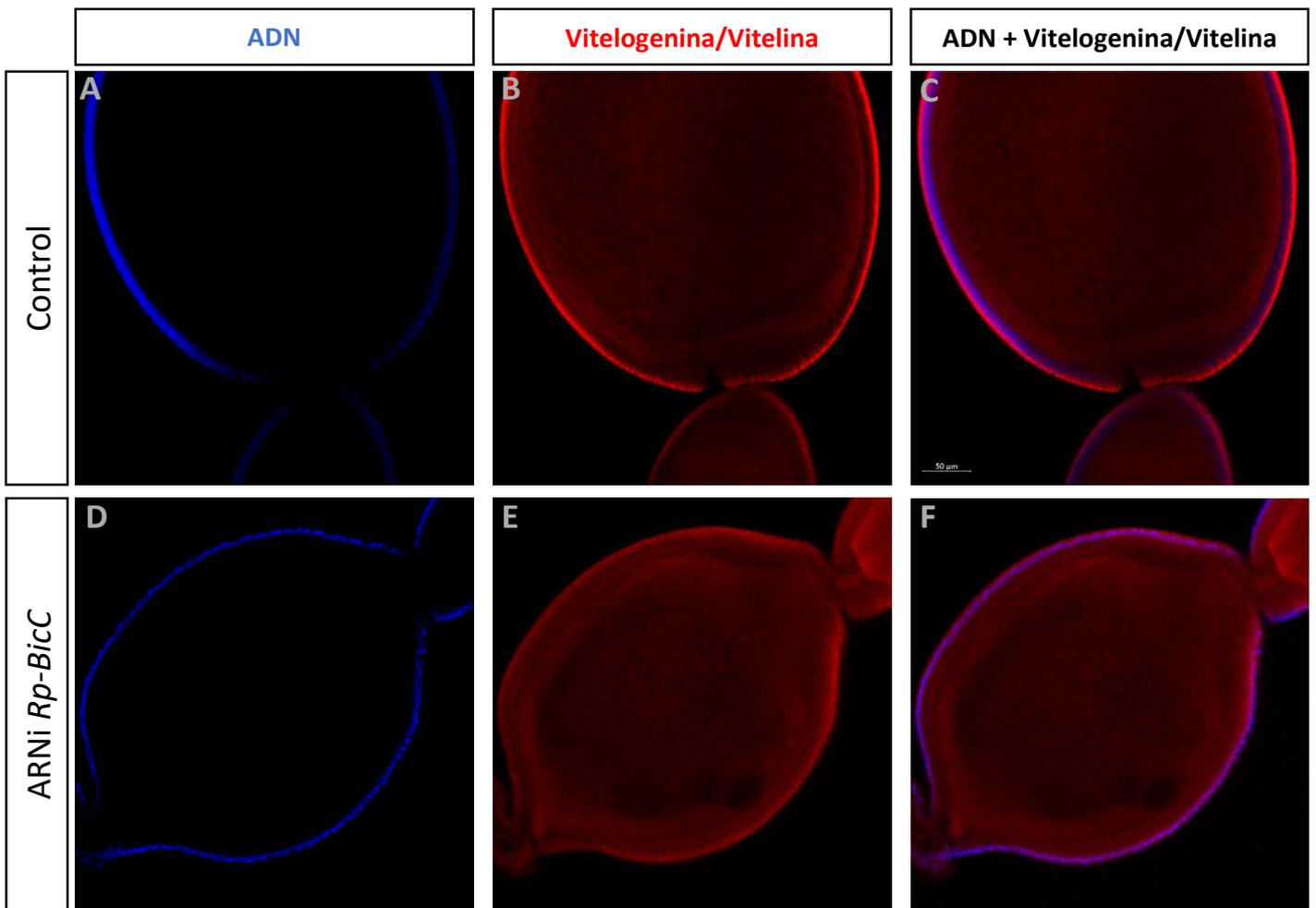
Como resultado, las ovariolas de hembras interferidas para *Rp-BicC* mostraron un patrón de disposición de lípidos diferentes al control, en donde la cara apical de las células foliculares presentó una distribución lineal y homogénea, pero en la cara basal de las mismas el contenido de lípidos fue insuficiente como para realizar una marcación visible. En el control sí se identificó la distribución de los lípidos de membrana, tanto en la cara apical como basal, del epitelio folicular que rodea al oocito en desarrollo, en estado vitelogénico. En la figura 30 se visualiza lo antes explicado, así como también se observa la presencia de marcación dentro del oocito de manera homogénea correspondiente al contenido lipídico del vitelo del mismo.



**Figura 30:** El silenciamiento de *Bicaudal C* produce oocitos vitelogénicos que presentan alterada la distribución lipídica en la membrana plasmática folicular. Hembras control e interferidas fueron inyectadas con la sonda FM4-64FX (1:500), los ovarios de las mismas fueron disectados y observados bajo microscopía confocal. El panel superior (A-C) se observa el oocito obtenido a partir de la hembra control y en el panel inferior (D-F) se observa el oocito obtenido de la hembra interferida. Con DAPI se visualiza la presencia de ADN y con rojo se visualiza la tinción de la membrana y la distribución lipídica marcada con FM4-64FX.

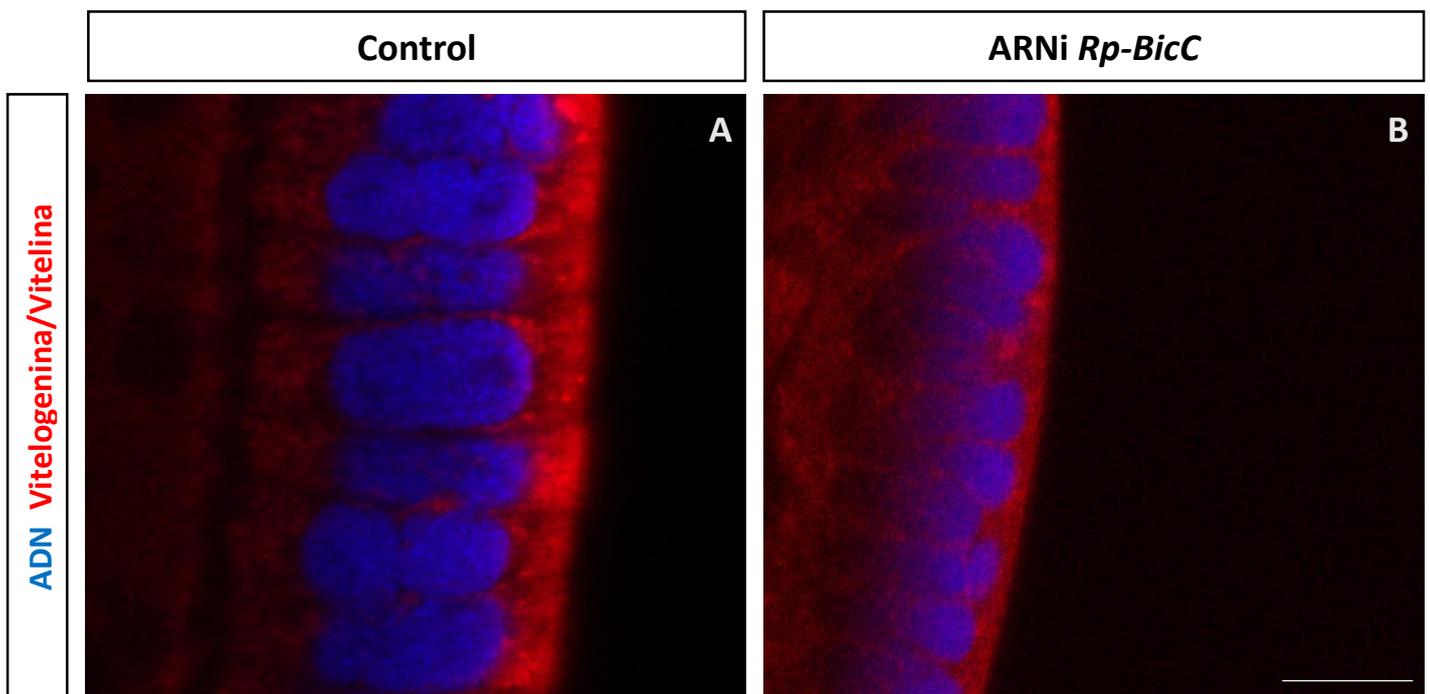
Para el análisis de la toma de vitelo durante la vitelogénesis se decidió utilizar el anticuerpo que previamente se utilizó para describir dicho proceso en la oogénesis de un ovario silvestre de *R. prolixus*. La metodología empleada fue llevada a cabo exactamente como se describió anteriormente, la única diferencia fue que como se utilizaron hembras silenciadas para *BicC* se efectuó la inmunohistoquímica en coordinación con tal experimento. Fue así que tanto ovarios de hembras control como interferidas se disectaron 5 días post alimentación y se los fijaron para proceder con la inmunohistoquímica.

Los oocitos obtenidos de hembras interferidas mostraron una disminución de la presencia de vitelogenina en el lado apical de las células foliculares que rodean al oocito vitelogénico. Así como también revelaron una distribución no homogénea de vitelina en el interior del oocito, observando una señal más fuerte en la superficie basal de la membrana oocítica que disminuye de manera gradual hacia el centro del oocito. Por otro lado, la membrana oocítica no se encontró definida de manera tan precisa como en el oocito vitelogénico del ovario control. A su vez, se observó, en concordancia con lo anteriormente registrado, que los oocitos pertenecientes a las ovariolas de hembras interferidas presentaron un menor tamaño en comparación con las hembras control (Figura 31)



**Figura 31:** Oocitos en estado vitelogénico en ovariola de *Rhodnius prolixus*. En el panel superior se visualizan oocitos correspondientes a las hembras control (A-C) y en el panel inferior se visualizan oocitos de hembras silenciadas (D-F), ambas imágenes obtenidas utilizando microscopía confocal. En las imágenes A y D se visualiza el oocito en estado vitelogénico teñido con DAPI (ADN), mostrando los núcleos correspondientes a las células foliculares que rodean a los oocitos en el vitelario. En las imágenes B y E se visualiza la localización de la proteína vitelogenina/vitelina, observando vitelogenina en el lado apical de la membrana folicular y vitelina en el interior de los oocitos. En C y F se observa la tinción tanto para ADN como para la proteína.

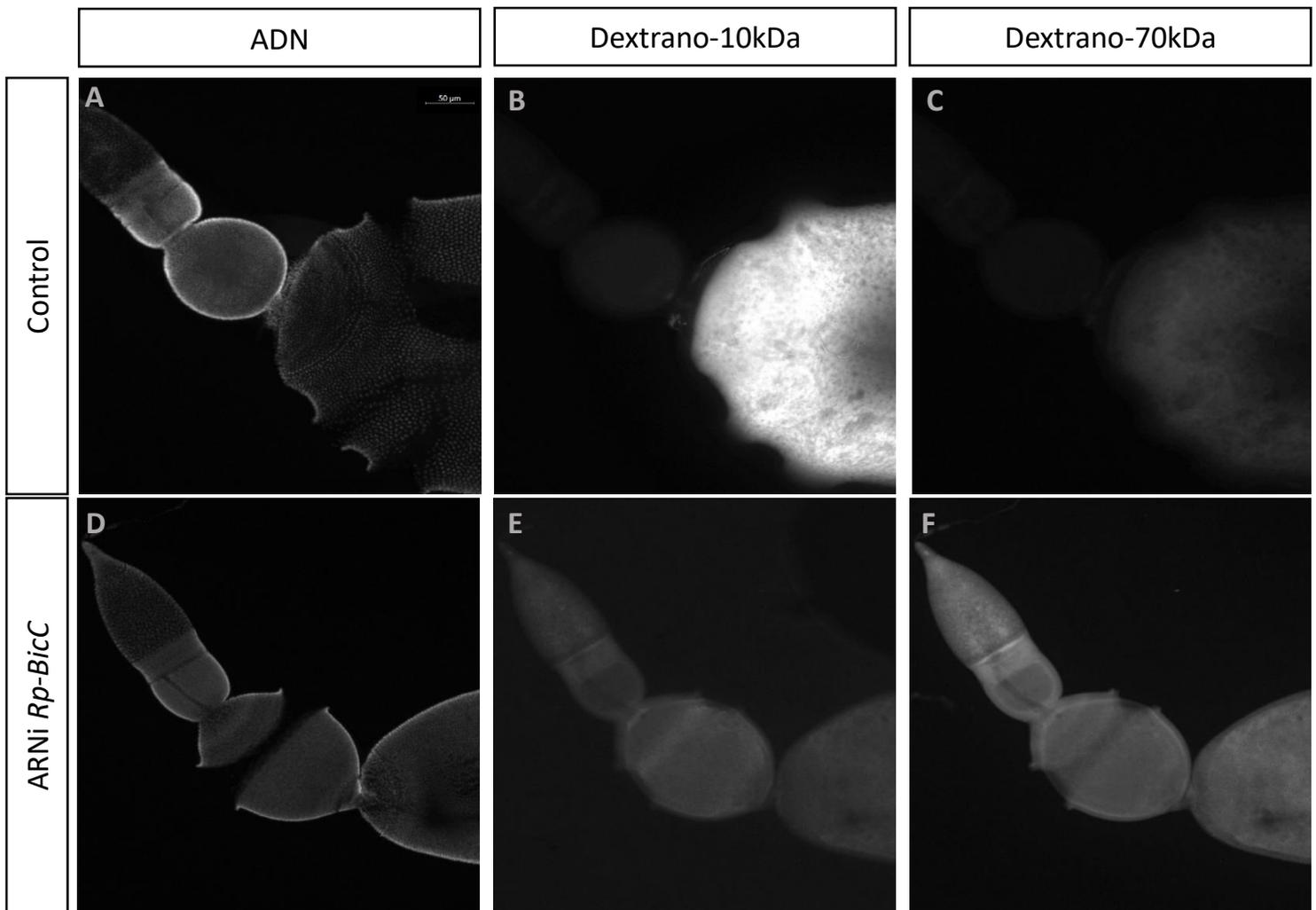
Observando en detalle las células foliculares y la membrana oocítica que rodean al oocito vitelogénico de la hembra *Rp-BicC* se pudo identificar una disminución significativa de la señal dada por la presencia de vitelogenina en el lado apical de las células foliculares, así como también un cambio de forma y tamaño de los núcleos que componen dichas células. Además, se observó un patrón de superposición entre la localización de los núcleos y la proteína. Por otro lado, no se percibió a la membrana oocítica de manera definida en los oocitos interferidos. En la Figura 32 se puede observar lo previamente descrito.



**Figura 32:** El silenciamiento de *Bicaudal C* produce un cambio notorio en la localización de la proteína vitelogenina en la membrana plasmática de los oocitos vitelogénicos. Los ovarios disectados de hembras control y silenciadas para *BicC* fueron teñidas con el anticuerpo anti-vn (rojo) junto a DAPI (ADN). Escala: 10  $\mu$ m.

Como se observaron diferencias con lo que respecta a la toma de vitelo durante la vitelogénesis, así como también en la disposición en las células foliculares que regulan dicho proceso, se decidió explorar aún más esto. Se prosiguió a evaluar si se veía afectado el traspaso de moléculas a través de los espacios existentes entre las células foliculares. Para esto, se inyectó una mezcla de dextranos de distinto peso molecular conjugados cada uno con distintas moléculas fluorescentes, por un lado, 10kDa-TexasRed y por el otro 70kDa-fluoresceína, para estudiar la permeabilidad de las ovariolas en hembras previamente

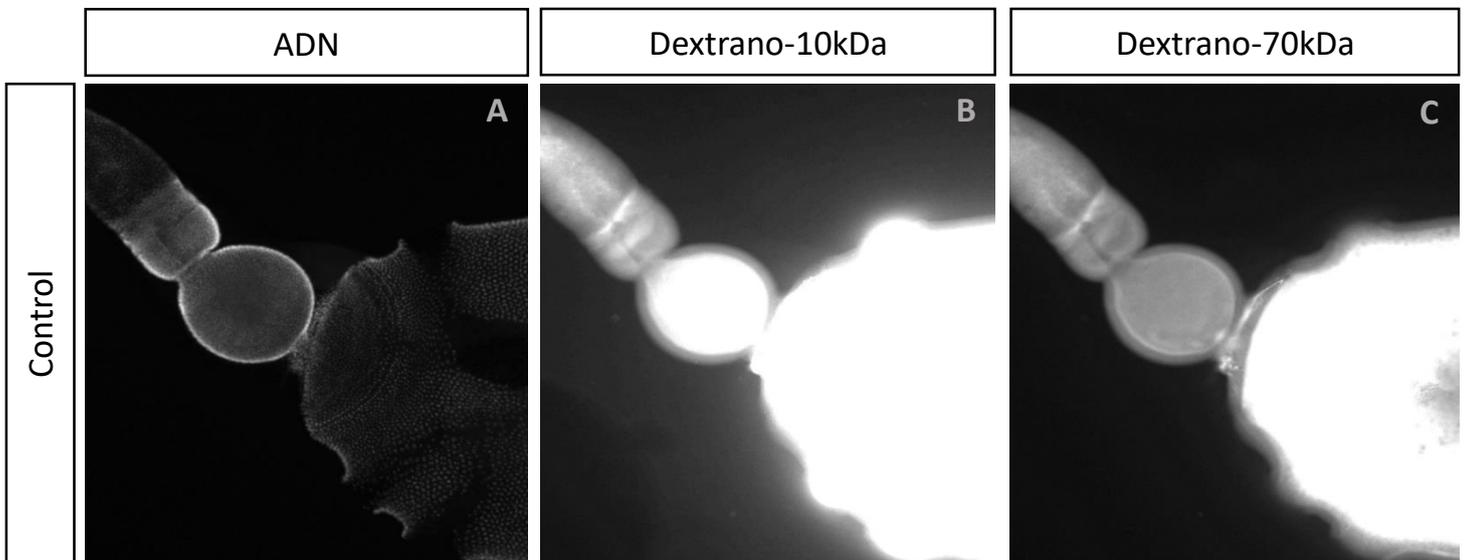
silenciadas y en hembras control. Utilizando el mismo tiempo de exposición (50 milisegundos) se capturaron imágenes de ambos ovarios disectados para evaluar su permeabilidad (Figura 33).



**Figura 33:** Estudio *in vivo* del ingreso de dextrano en oocitos en fase vitelogénica. Ovariolas de hembras control (A-C) e interferidas (D-F) inyectadas con mezcla de dextranos fueron disectadas y observadas bajo microscopio de fluorescencia. En el panel superior muestra las ovariolas correspondientes a las hembras control, inyectadas con *dcβ-lac* y en el panel inferior los ovarios de hembras silenciadas genéticamente inyectadas con *dcBicC*. El tiempo de exposición utilizado para ambas capturas fue de 50ms.

En los ovarios control fue posible detectar una fuerte localización de fluorescencia únicamente en el oocito en estado vitelogénico correspondiente al dextrano de peso molecular de 10kDa; y para el dextrano de 70kDa se reportó una localización en el mismo oocito pero con una señal muchísimo más débil. En cambio, las ovariolas de las hembras *Rp-BicC* no mostraron una localización específica de ninguno de los dos dextranos utilizados,

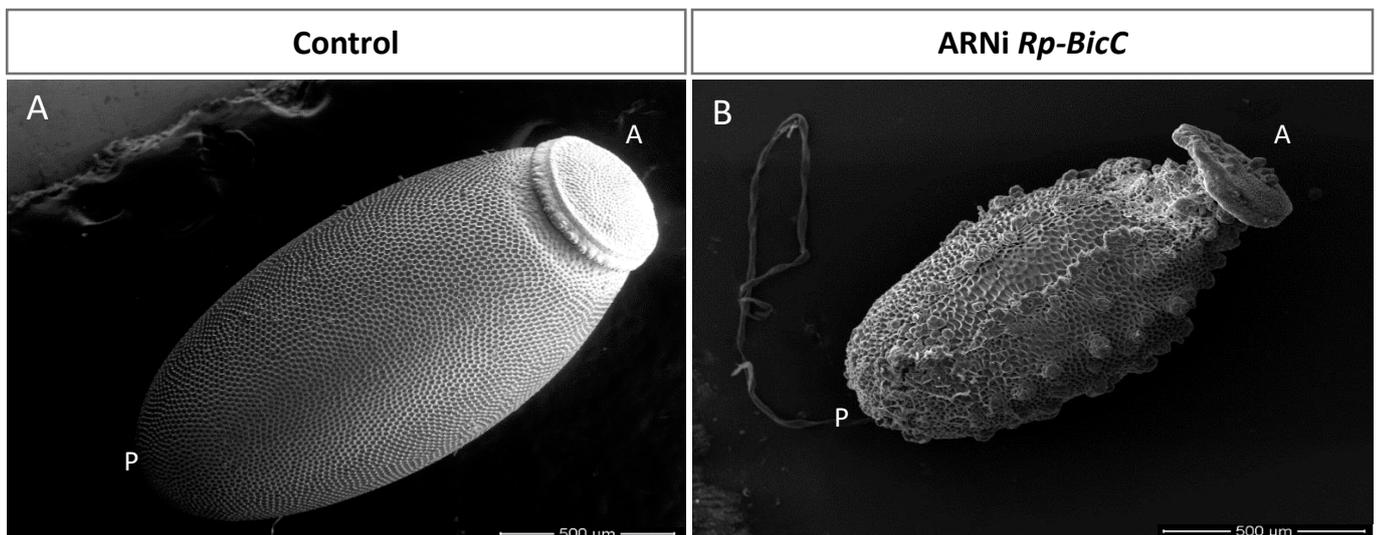
donde se observó una señal débil y uniforme en toda la dimensión de la ovariola. Comparando control con interferido se observó una marcada disminución de la presencia de señal correspondiente al dextrano de 10kDa en el oocito en vitelogénesis del ovario ARNdc *BicC2*. Aumentando el tiempo de exposición, en los ovarios control se observaron diferencias con lo que respecta a la permeabilidad de los distintos tipos de dextransos (Figura 34). Para el de 10kDa se siguió observando una fuerte localización en el oocito en estado vitelogénico, así como también se comenzó a observar presencia de señal intensa en el oocito previtelogénico. Esto no sucedió para el caso del dextrano de 70 kDa donde sólo se reportó localización del mismo en el oocito ya en fase vitelogénica.



**Figura 34:** Estudio *in vivo* del ingreso de dextrano en oocitos de ovariola control. Ovariolas de hembras control (A-C) inyectadas con mezcla de dextransos fueron disectadas y observadas bajo microscopio de fluorescencia. La captura de las imágenes fue realizada utilizando un tiempo de exposición de 500ms.

Para abordar el cambio registrado en los huevos obtenidos, se decidió utilizar la tecnología del microscopio electrónico de barrido para comprender en detalle el porqué de su cambio en el tamaño y coloración. Los huevos ovipuestos por las hembras *Rp-BicC* expusieron defectos evidentes en la estructura del corion, mostrando claras desviaciones del patrón poligonal-hexagonal uniforme que se observa en los huevos obtenidos a partir de la hembra control (Figura 35). Estos defectos en el corion se observaron tanto en la parte lateral del huevo como en el lado apical del mismo, y en ambas partes mostraron una clara disminución del tamaño de las fosas foliculares. Para darle sustento a las diferencias observadas se prosiguió a calcular

el área promedio correspondiente al patrón hexagonal (análisis de 30 fosas foliculares laterales) impreso en el corion de los huevos provenientes de las hembras silenciadas y de los huevos provenientes de las hembras control, los valores obtenidos coincidieron con lo observado, el área promedio del patrón hexagonal del huevo control presentó un valor de  $446.5 \mu\text{m}^2$ , y el huevo *BicC* mostró un valor de  $357.9 \mu\text{m}^2$ . Sumado a eso, se evaluó la longitud de ambos tipos de huevos observados y se pudo ver que el huevo control presentó una mayor longitud en comparación al huevo obtenido tras el silenciamiento génico, mostrando un valor aproximado  $1587.373 \mu\text{m}$  frente a  $1303.916 \mu\text{m}$ . Más allá de estas diferencias observadas, es importante remarcar que no se modificó la polaridad de los huevos ovipuestos por hembras *Rp-BicC*.



**Figura 35:** El silenciamiento de *Bicaudal C* produce un desorden en la disposición de las células foliculares que resulta en una malformación del corion. La figura **A** muestra a huevos ovipuestos por hembras control tomados por microscopía electrónica de barrido, y la figura **B** muestra a los huevos ovipuestos por hembras interferidas para *BicC*. Referencias: Polo anterior (A), Polo posterior (P).

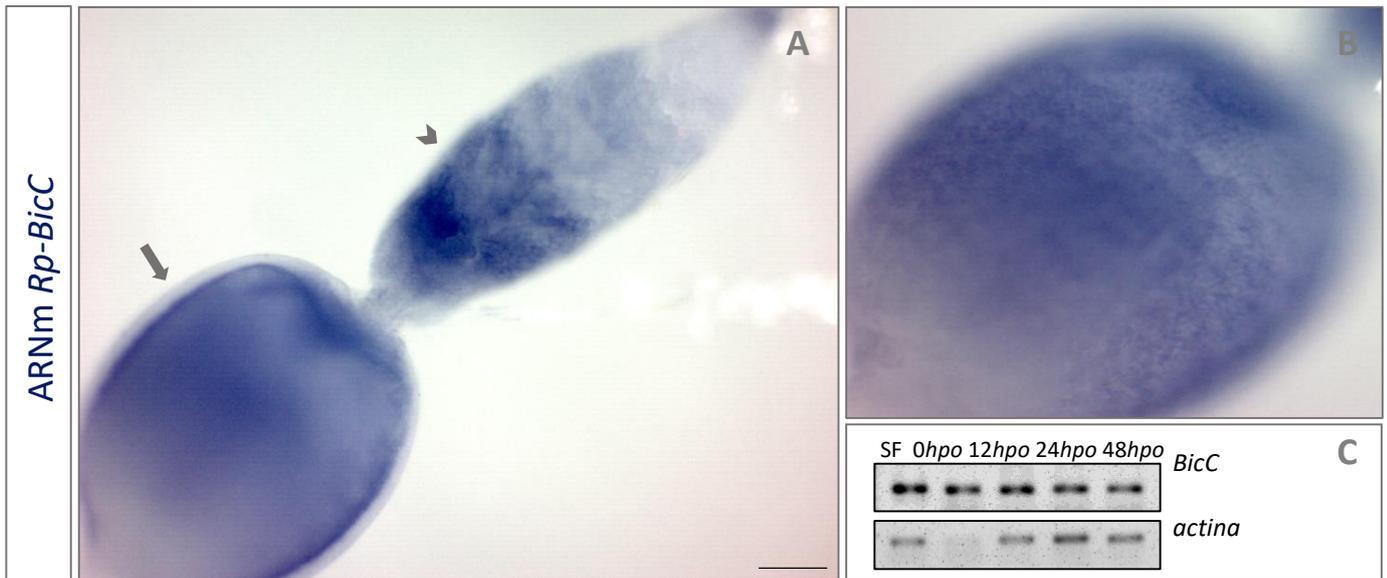
Para confirmar que el fenotipo observado no se debiera a un efecto inespecífico obtenido mediante la interferencia, se decidió repetir la inyección utilizando el otro amplicon de 231 pb utilizando los oligos específicos *Fw1* y *Rv1*. De la misma manera que se prosiguió anteriormente se inyectaron 5 hembras vírgenes y se evaluó tanto fertilidad como fenotipo a nivel ovario y embrional. Los resultados que se obtuvieron fueron 100% coincidentes con los descritos previamente, si cabe aclarar que la disminución en el número total de huevos

ovipuestos se debió, principalmente, a que como el fenotipo fue coincidente, se prosiguió rápidamente a disectar los ovarios para corroborar la existencia del mismo a nivel ovárico.

### **2.1.2 Expresión génica de *Bicaudal C* en *R. prolixus***

Para determinar la expresión del gen *BicC* en los ovarios de *R. prolixus*, se utilizó la sonda antisentido específica para *Rp-BicC*, generada como se detalló previamente, y se estudió su expresión durante todo el proceso de la gametogénesis. La hibridación *in situ* reveló que *BicC* en este organismo modelo tiene expresión citoplasmática en el epitelio folicular, tanto en estadio previtelogénico como vitelogénico, pero no se encuentra expresada en el interior de los oocitos en desarrollo. Por otra parte, en el germario se vio marcación confinada en la región citoplasmática y con una fuerte expresión en la región más próxima al vitelario, lo que correspondería a la zona III, región donde las células nodrizas se concentran con un gran tamaño nuclear (Figura 36).

Además del análisis de expresión en ovarios, se evaluó la expresión durante el desarrollo embrionario, específicamente en tiempos correspondientes al estado de cigoto, blastodermo, embrión gastrulante y banda germinal, así como también huevos no fertilizados. Para esto se efectuó una retro transcripción para posteriormente realizar una amplificación del gen de interés *BicC*, utilizando los oligos específicos, llevando a cabo la reacción en cadena de la polimerasa con transcriptasa reversa (RT-PCR). Se reportó que el amplicon obtenido correspondiente a *BicC* se expresa en todos los tiempos del desarrollo analizados (Figura 36).



**Figura 36:** *Bicaudal C* se expresa en el ovario y durante las distintas etapas del desarrollo de *R. prolixus*. **A** La hibridación in situ con la sonda específica para *BicC*. La punta de flecha indica la expresión en la zona III del trofario. La flecha señala al oocito vitelogénico con la expresión folicular. En la Figura **B** es una magnificación de la región para ver en detalle esa expresión citoplasmática. En la figura **C** se observa el producto de RT-PCR para *BicC* y actina (control) sembrado en un gel de agarosa 1%. Escala: 10  $\mu$ m.

## 2.2 Conrichon

### 2.2.1 Validación funcional por ARNi parental

Para descubrir el rol de *cornichon (cni)* en *Rhodinus prolixus (Rp-cni)*, se inyectaron 6 hembras vírgenes con diferentes dosis de ARN doble cadena (ARNdc) específico para el gen, ARNdc<sup>cni</sup>. Al igual de cómo se realizó par *BicC* se prosiguió a evaluar fertilidad, y fenotipo a nivel de ovario, huevo y embrión. A diferencia de *BicC*, donde se utilizaron dos amplicones diferentes para corroborar la identidad al fenotipo obtenido, para este gen se generó un único amplicon para trabajar. La razón principal de utilizar un solo par de oligos específicos es que el transcripto identificado como *cni* exhibió varias regiones ricas en timinas, lo que dificultaba la tarea de diseñar más de un buen par de oligos para trabajar. El amplicon que se utilizó para generar el ARNdc<sup>cni</sup> presentó una longitud total de 154 pb.

En la tabla 15 se describe la dosis específica para cada inyección y los fenotipos observados; para el análisis fenotípico se decidió proseguir con dos estrategias de trabajo, por un lado, se tomaron algunas hembras inyectadas con ARNdc<sup>cni</sup> y se le procedió a disectar sus ovarios (durante su estado ovipleno) y por el otro, a las remanentes se las dejaron para que continúen con el proceso de ovipuesta normal y observar la descendencia.

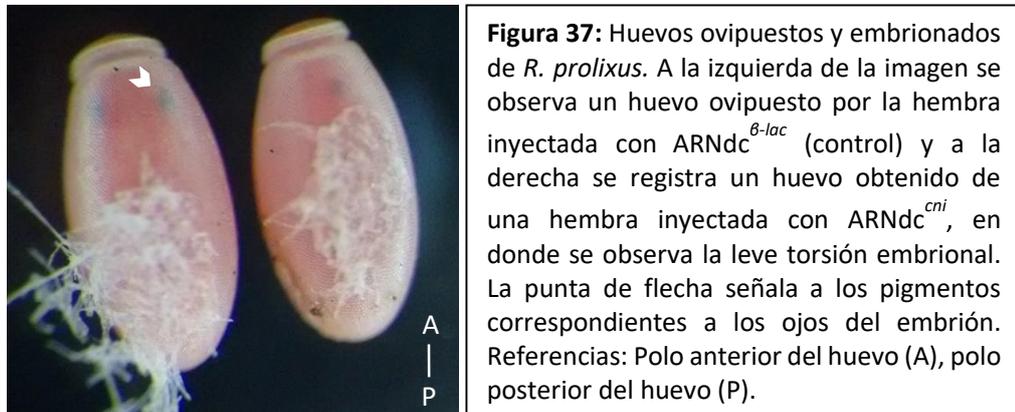
	Dosis de ARNdc	Número de hembras inyectadas	Oviposición (N° de huevos)	Fenotipo a nivel del ovario	Fenotipo a nivel del huevo	Fenotipo embrional
ARNdc <sup>cni</sup>	1.2 µg	3	33	0/3	0/33	15/33 (46%)
	0.8 µg	3	22	0/3	0/22	07/22 (32%)
ARNdc <sup>βlac</sup>	1 µg	2	72	0/2	0/72	5/72 (6,94%)

**Tabla 15:** Representación en números de los resultados obtenidos tras el silenciamiento génico de *cornichon (cni)*. El panel de color verde claro hace referencia al ARNdc generado a partir del uso de los oligos específicos. Las columnas muestran la cantidad en µg de ARNdc específico inyectado en hembras vírgenes y los resultados obtenidos a partir de cada oviposición. Las características del silenciamiento génico para ARNdc<sup>cni</sup> se comparan con el control (ARNdc<sup>βlac</sup>).

A simple vista no se observó ninguna diferencia estructural a nivel del ovario y del huevo, así como tampoco se vio afectada la capacidad de producir descendencia, ya que todos los huevos ovipuestos estaban embrionados; presentaban la coloración característica de rosa claro, así como también bien establecido el patrón hexagonal impreso por las fosas foliculares en el corion del mismo.

Si se observaba en detalle, bajo lupa, la disposición del embrión dentro del huevo, antes de llegar al momento de la eclosión, se podía observar que presentaban una leve diferencia con respecto a los huevos control. Esta diferencia se observó en los embriones avanzados en el desarrollo donde ya era posible visualizar los ojos a través del corion. Cerca de la mitad de los huevos ovipuestos exhibieron una leve basculación con respecto al eje dorso-ventral del huevo; un giro no usual en los embriones ARNdc<sup>βlac</sup>. Esta sutil diferencia fue considerada dentro de la categoría fenotipo embrional en la tabla antes presentada. En la Figura 37 se observa lo previamente explicado.

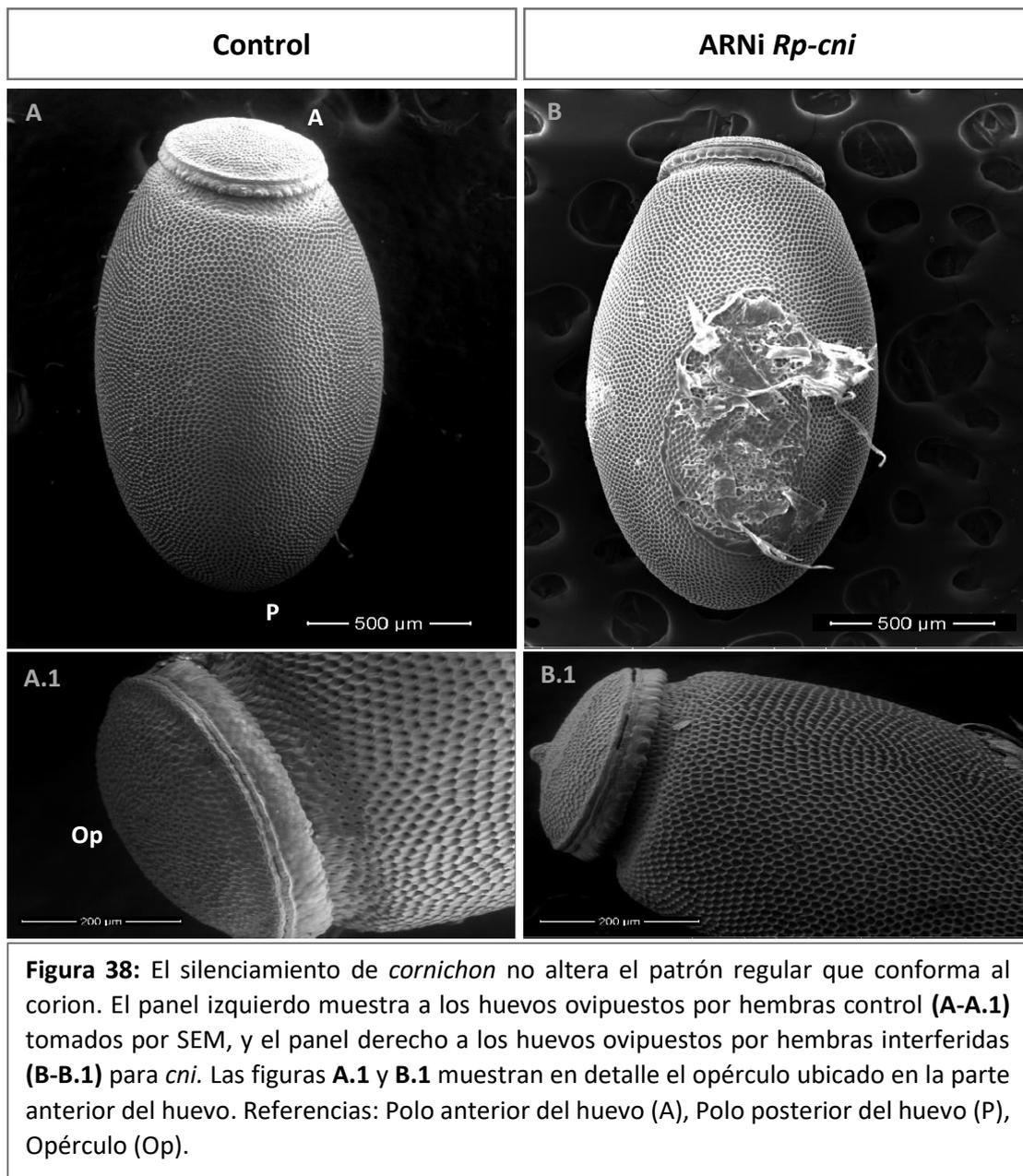
Cabe aclarar que aunque existía una disposición diferente del embrión dentro del huevo no afectó la continuidad del desarrollo ni la eclosión del mismo, todos los embriones emergieron del huevo y se desarrollaron en el primer estadio larval.



Para determinar si existía una correlación entre la torsión que se observó del embrión y la estructura del huevo que lo contiene se decidió utilizar la tecnología del microscopio electrónico de barrido. Se procedió a preparar a los huevos ovipuestos que presentaron la torsión, así como también huevos control para comparar ambas condiciones.

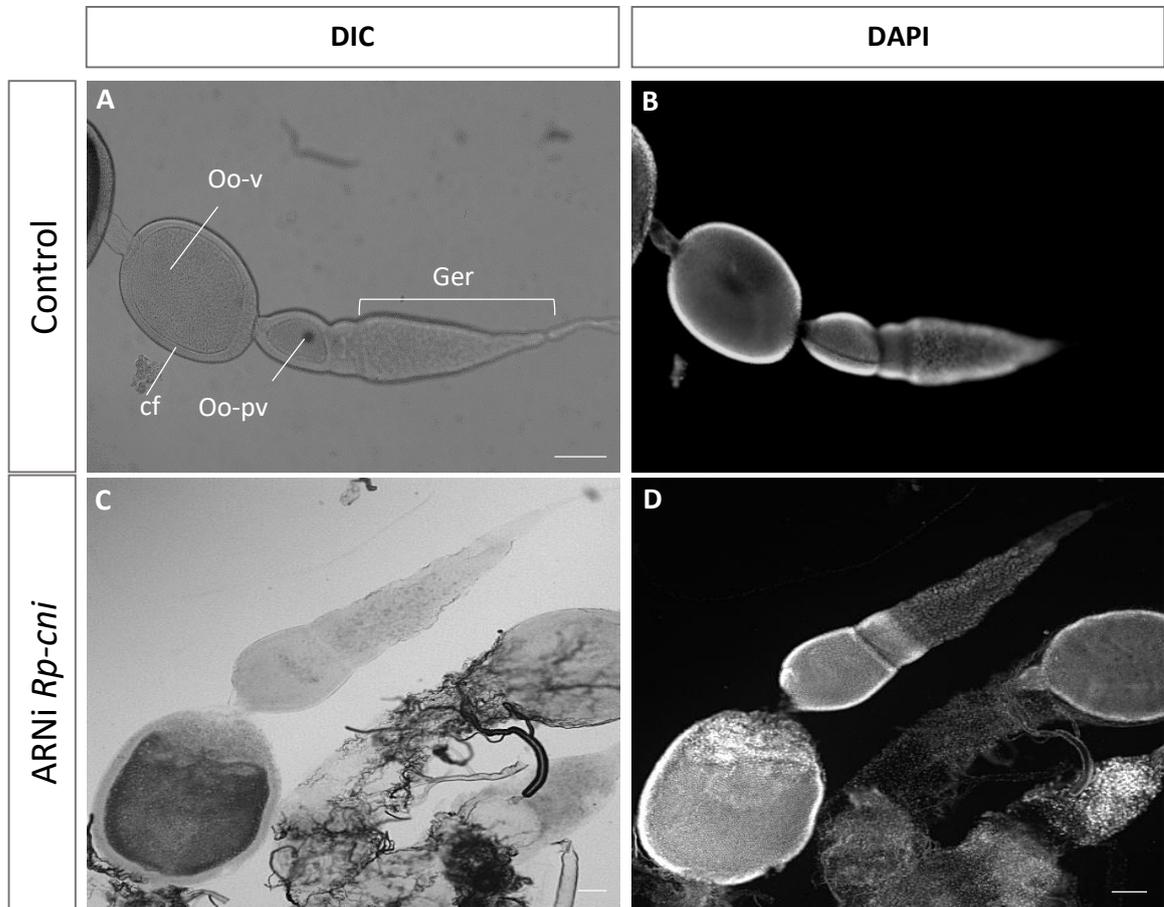
El patrón regular característico que se imprime en la superficie exterior del corion en el huevo control se observó y repitió en los huevos obtenidos a partir de hembras silenciadas. Las fosas foliculares de los huevos *Rp-cni* mostraron una forma hexagonal con un tamaño promedio coincidente con lo reportado para los huevos control.

Por otro lado, se evaluó la longitud de ambos tipos de huevos observados y se pudo ver que el huevo control presentó una menor longitud en comparación al huevo obtenido tras el silenciamiento génico, mostrando un valor aproximado 1587.373 μm frente a 1786.382 μm. Lo previamente descrito se visualiza con detalle en la Figura 38.



Para observar si existía alguna diferencia en la estructura del ovario, la cual no había podido ser registrada a simple vista bajo lupa, se prosiguió a analizar el patrón nuclear de las células que componen a las ovariolas; para esto, a las ovariolas previamente fijadas y disectadas tanto de hembras control como silenciadas genéticamente, se las trató con el marcador fluorescente DAPI y se las analizó por microscopía de fluorescencia. No observamos ninguna discrepancia en cuanto a la morfología y estructura del ovario de *Rp-cni*, resultado concordante con lo observado bajo lupa estereoscópica. El patrón nuclear fue coincidente en

ambos ejemplos de ovarios, encontrando analogías en todas las partes que componen a la ovariola; en el germario y en el vitelario (Figura 39).

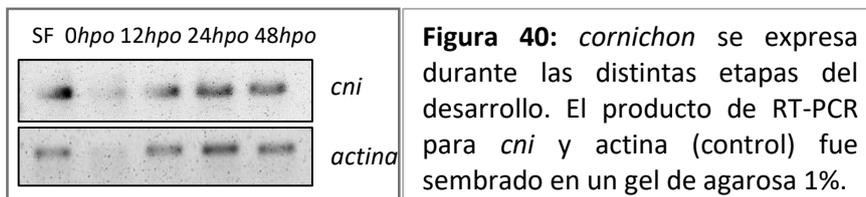


**Figura 39:** El silenciamiento de *cornichon* no afecta la morfología normal de los ovarios de *R. prolixus*. Ovariolas disectadas de hembras previamente inyectadas con ARNdc <sup>$\beta$ -lac</sup> (control) o con ARNdc<sup>*cni*</sup>. La disección fue llevada a cabo 5 días post alimentación. Las figuras (A-C) muestran a la ovariola obtenida utilizando microscopía de contraste diferencial interferencial (DIC) y en (B-D) la ovariola obtenida por microscopía confocal, utilizando DAPI para marcar el ADN presente. Referencias: Germario (Ger), Células foliculares (cf), Oocito- vitelogénico (Oo-v), Oocito-previtelogénico (Oo-pv). Escala panel superior: 200  $\mu$ m. Escala panel inferior: 100  $\mu$ m.

### 2.2.2 Expresión génica de *cornichon* en *R. prolixus*

El análisis de la expresión de *cni* fue realizado en los distintos tiempos del desarrollo, correspondientes al estado de cigoto, blastodermo, embrión gastrulante y banda germinal, así como también huevos no fertilizados. Se efectuó una retro transcripción para posteriormente realizar una amplificación del gen de interés *cni*, utilizando los oligos específicos. Como se

observa en la Figura 40 *cornichon* se encontró representado en todos los tiempos analizados, con una menor expresión en el estadio correspondiente al estado de cigoto (0hpo).



## 2.3 *Bicaudal D*

### 2.3.1 Validación funcional por ARNi parental

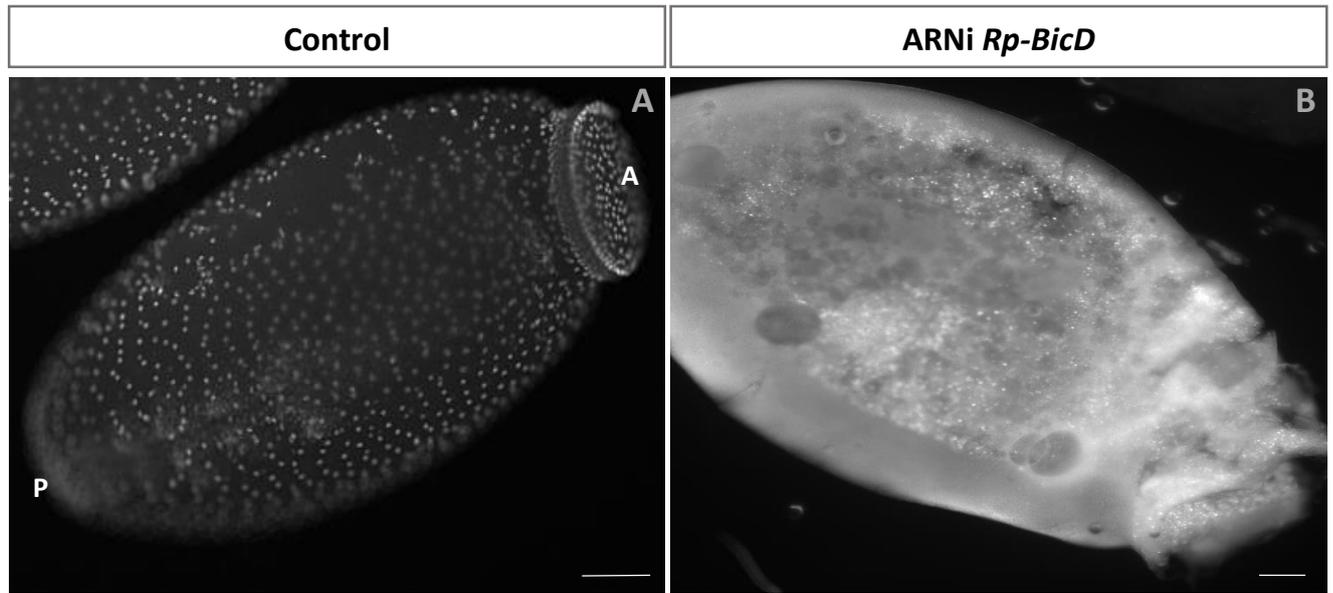
Para determinar el rol de *Bicaudal D* (*BicD*) en *Rhodinus prolixus* (*Rp- BicD*), se generó ARNdc específico del gen a partir de la utilización de un par de oligos específicos denominados PrFw2BicD y PrRv2BicD, los cuales amplificaron un fragmento de 245 pb, que se utilizó para proseguir con la interferencia.

Se inyectaron 20 hembras vírgenes con diferentes dosis de ARN doble cadena (ARNdc) específico para el gen ARNdc<sup>BicD2</sup>. Del total de hembras inyectadas, 11 no sobrevivieron como para poder ser analizadas tras el silenciamiento génico; en algunas circunstancias murieron previo a la alimentación y en otras, post alimentación, pero en ningún caso lograron alcanzar una fase de oviposición. Del restante que, sí prosiguió, se observó que el número de huevos ovipuestos fue relativamente bajo en comparación al control, 9 hembras ovipusieron 43 huevos en total (Tabla 16). De estos, sólo el 9% se caracterizó por presentar embrión en su interior. Las hembras que fueron silenciadas con la menor dosis de ARNdc específico fueron las únicas que ovipusieron huevos embrionados, las restantes dosis se caracterizaron por una oviposición de huevos no embrionados en su totalidad. Más allá de la ausencia de embrión, todos los huevos presentaron el color rosa característico, así como el tamaño y forma equivalente al huevo control. Como en las interferencias anteriores, el control negativo de la técnica fue con  $\beta$ - lactamasa de *Escherichia coli* (ARNdc <sup>$\beta$ -lac</sup>).

	Dosis de ARNdc	Número de hembras inyectadas	Oviposición (N° de huevos)	Fenotipo a nivel del ovario	Fenotipo a nivel del huevo	Fenotipo embrional
ARNdc <i>BicD2</i>	0.5 µg	5	22	0/5	0/22	18/22 (81,8%)
	1.3 µg	5	9	0/4	0/9	9/9 (100%)
	2 µg	6	2	0/5	0/2	2/2 (100%)
	2.5 µg	4	10	0/2	0/10	10/10 (100%)
ARNdc <i>βlac</i>	1 µg	2	72	0/2	0/72	5/72 (6,94%)

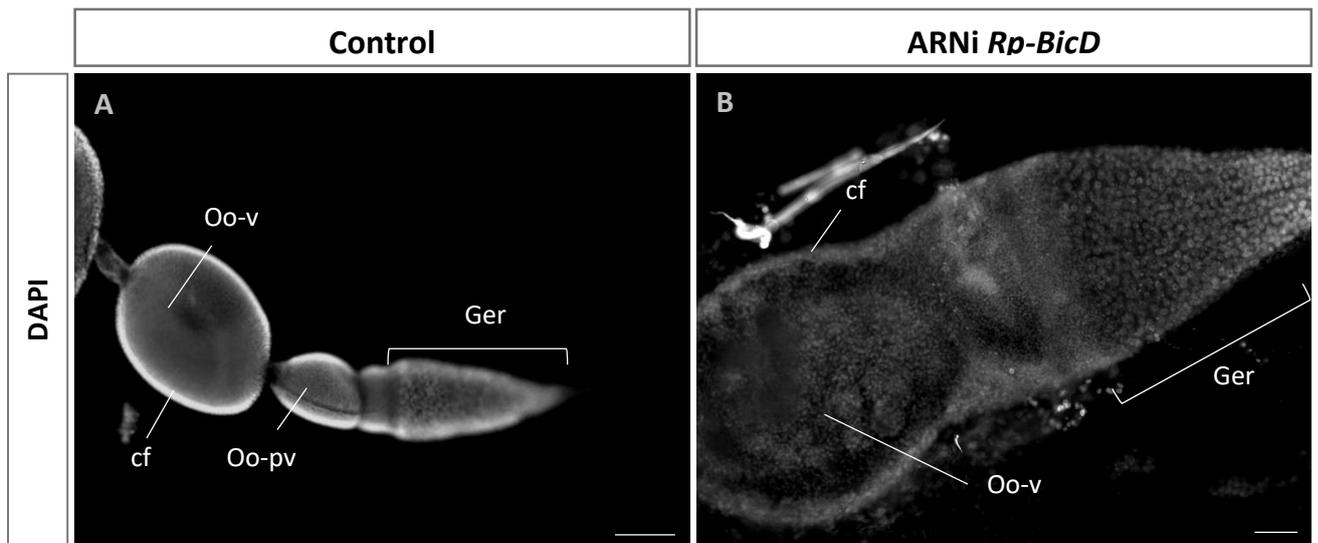
**Tabla 16:** Representación en números de los resultados obtenidos tras el silenciamiento génico de *Bicaudal D (BicD)*. El panel de color amarillo hace referencia al ARNdc generado a partir del uso de los oligos específicos. Las columnas muestran la cantidad en µg de ARNdc específico inyectado en hembras vírgenes y los resultados obtenidos a partir de cada oviposición. Las características del silenciamiento génico para ARNdc *BicD* se comparan con el control (ARNdc *βlac*).

La ausencia de embrión en una primera instancia se observó por la falta de continuidad del desarrollo. En comparación con los huevos control, nunca fue posible observar apéndices y pigmentación embrionaria, a través del corion, cuando se los examinó bajo lupa estereoscópica. Para corroborar si realmente no estaban embrionados o sí se encontraba detenido su desarrollo en etapas tempranas del mismo, se prosiguió a efectuar un análisis de tinción del patrón nuclear. Como se observa en la Figura 41, no se observó embrión en ninguno de los huevos ovipuestos por las hembras silenciadas.



**Figura 41:** El silenciamiento de *Bicaudal D* afecta la producción de huevos embrionados. La figura **A** muestra a huevos ovipuestos por hembras control teñidos con DAPI tomados por microscopía confocal, y la figura **B** muestra a los huevos ovipuestos por hembras interferidas para *BicD*. Los huevos correspondientes al control presentan un estado del desarrollo correspondiente a banda germinal. La escala **A**: 200  $\mu\text{m}$ . La escala en **B**: 100  $\mu\text{m}$ . Referencias: Polo anterior (A), Polo posterior (P).

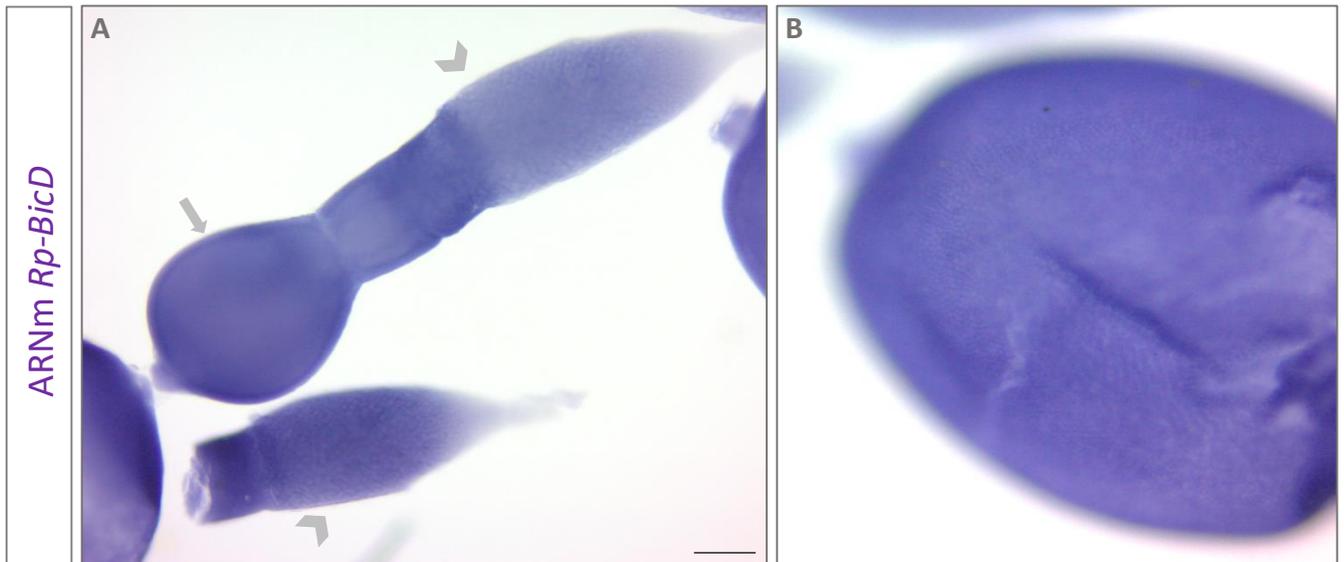
Por otro lado, se analizó la estructura del ovario para ver si existía alguna discrepancia en cuanto la morfología normal. En primer lugar, bajo lupa estereoscópica se disectaron y fijaron las ovariolas de hembras control y silenciadas genéticamente; se compararon a grandes rasgos y no se observó diferencia alguna. Entonces para evaluar con mayor detalle, se decidió analizar el patrón nuclear de las células que componen a las ovariolas. Como resultado se observó que el silenciamiento de *BicD* no alteró la morfología normal del ovario del triatomino (Figura 42).



**Figura 42:** El silenciamiento de *Bicaudal D* no afecta la morfología normal de los ovarios de *R. prolixus*. Ovariolas disectadas de hembras previamente inyectadas con ARNdc<sup>*β-lac*</sup> (control) o con ARNdc<sup>*BicD*</sup>. La disección fue llevada a cabo 5 días post alimentación. La figura **A** muestran a la ovariola control obtenida utilizando microscopía confocal y en **B** la ovariola de hembra interferida obtenida por microscopía confocal, en ambos se utilizó DAPI para marcar el ADN presente. Referencias: Germario (Ger), Células foliculares (cf), Oocito- vitelogénico (Oo-v), Oocito-previtelogénico (Oo-pv). Escala panel izquierdo: 200  $\mu$ m. Escala panel derecho: 100  $\mu$ m.

### 2.3.2 Expresión génica de *Bicaudal D* en *R. prolixus*

Para determinar la expresión del gen *BicD* en los ovarios de *R. prolixus*, se utilizó la sonda antisentido específica para *Rp-BicD*, generada como se detalló previamente, y se estudió su expresión durante todo el proceso de la gametogénesis. La hibridación *in situ* reveló que *BicD* en este organismo modelo tiene expresión citoplasmática en el epitelio folicular, tanto en estadio previtelogénico como vitelogénico, pero no se encuentra expresada en el interior de los oocitos en desarrollo. Por otra parte, en el germario se vio marcación confinada en la región citoplasmática en su totalidad, con una mayor expresión en la región próxima al vitelario, lo que correspondería a la zona III, región donde las células nodrizas se concentran con un mayor tamaño nuclear (Figura 43).

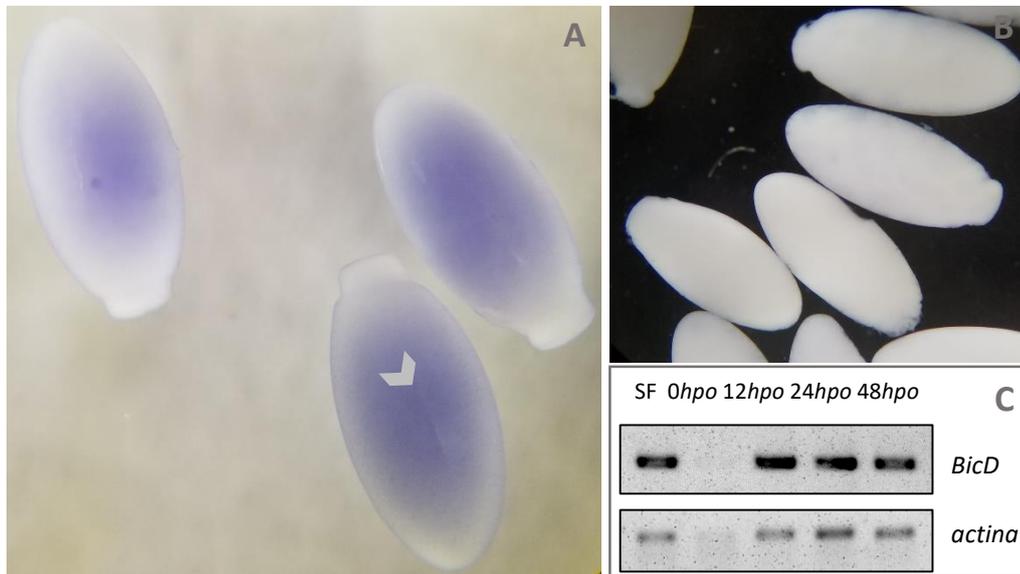


**Figura 43:** *Bicaudal D* se expresa en el ovario de *R. prolixus*. La hibridación *in situ* con la sonda específica para *BicD*. La punta de flecha indica la expresión confinada en el citoplasma del trofario. La flecha señala al oocito vitelogénico con la expresión folicular. En la Figura **B** es una magnificación de la región para ver en detalle esa expresión citoplasmática. La línea negra hace referencia a la escala que equivale a 10  $\mu\text{m}$ .

Además del análisis de expresión en ovarios, se evaluó la expresión durante el desarrollo embrionario. En primer lugar, se llevó a cabo una hibridación *in situ*, con la misma sonda antisentido que se diseñó para la tinción en los ovarios, en huevos disectados desde dentro del ovario, que ya estaban listos para ser ovipuestos, y en huevos muy tempranos del desarrollo, ya ovipuestos, que comprendían desde estado de cigoto hasta previo blastodermo celularizado. Como resultado se observó que en huevos colectados dentro del ovario *BicD* mostró un patrón de expresión que difunde a lo largo de la longitud del embrión con una mayor concentración en la región central del mismo (Figura 44A). En cambio, en embriones más avanzados (hasta 12hpo), no se registró marcación alguna con respecto a los transcritos correspondientes a *BicD* (Figura 44B).

Por otro lado, se evaluó la expresión del gen a través de una RT-PCR, utilizando como molde distintos tiempos del desarrollo, equivalentes al estado de cigoto, blastodermo, embrión gastrulante y banda germinal, así como también huevos no fertilizados. Para esto se efectuó una retro transcripción para posteriormente realizar una amplificación del gen de interés *BicD*. Se reportó a *BicD* presente en todos los tiempos del desarrollo analizados, exceptuando al

tiempo 0hpo (cigoto), resultado coincidente con lo reportado mediante la hibridación *in situ* (Figura 44C).



**Figura 44:** *Bicaudal D* se expresa de manera diferencial durante el desarrollo embrionario. La hibridación *in situ* con la sonda específica para *BicD* muestra en **A** la expresión localizada en el centro del embrión, indicada con la punta de flecha. En la Figura **B** se observa a los embriones de 0-10hpo sin marcación alguna. En la parte **C** se visualiza el producto de RT-PCR para *BicD* y actina (control) que fue sembrado en un gel de agarosa 1%, teñido con BrEt.

# Discusión

## 1. Genética del desarrollo embrionario temprano en *Rhodnius prolixus*

La formación de un organismo comienza a partir del cigoto, en el humilde proceso de formación del huevo; es durante la ovogénesis que se determinan los ejes de las coordenadas del cuerpo (anterior-posterior y dorsal-ventral) por acción de genes maternos; se forma el corion, estructura que permite su desarrollo una vez ovipuesto; y se carga el vitelo. La formación del huevo en sí y la determinación de los ejes ortogonales dependen de la acción coordinada de genes durante el desarrollo (Atella et al., 2005; Nusslein-Volhard and Roth, 1989; Papantonis et al., 2015; Schupbach and Roth, 1994; Speder et al., 2006).

El desarrollo del oocito comienza con la división asimétrica de una célula madre de la línea germinal, que da lugar a la formación del citoblasto y a una célula hija (Spradling, 1993; Telfer, 1975). El citoblasto produce eventualmente uno o más oocitos según el tipo de ovario. En *R. prolixus* las células madres de la línea germinal están restringidas al estadio larval 5 (Anderson, 1974; Telfer, 1975), por lo que las hembras adultas ya llevan un número discreto de oocitos detenidos en la profase de meiosis I, que se acumulan en la parte posterior del germario del ovario maduro (Huebner and Anderson, 1972c). Luego de la ingesta de sangre (Buxton, 1930), el proceso de ovogénesis tiene inicio y los oocitos comienzan a migrar de manera ordenada hacia la región posterior de la ovariola, donde comienzan a rodearse de células foliculares (Anderson, 1974; Huebner, 1981; Lutz and Huebner, 1980; Lutz and Huebner, 1981).

Los ARNm y los nutrientes producidos en el germario son transportados unipolarmente a través de puentes citoplasmáticos especializados a los oocitos en crecimiento, los cuales aumentan de tamaño en correlación a la toma de vitelogenina, que soporta el crecimiento del oocito. Este permanece conectado al germario hasta alcanzar un diámetro de 1000  $\mu\text{m}$  de diámetro y desarrollar el corion (Nunes-da-Fonseca et al., 2017).

Las bases sólidas del estudio de la morfología y fisiología de ovario en *R. prolixus* permitieron comenzar a explorar el proceso ya desde un punto de vista genético e intentar dar inicio al descubrimiento de la red génica que dirige el desarrollo embrionario, la cual aún no está clara. Ante esto, en este trabajo se propuso resurgir los trabajos llevados a cabo principalmente por Huebner, realizando la re-descripción de la estructura de la ovariola con registros fotográficos llevadas a cabo con tecnologías modernas.

Asimismo, se decidió estudiar uno de los procesos más importante para el desarrollo del embrión, la vitelogénesis, proceso fundamental debido a que el embrión se desarrolla por fuera del cuerpo de la madre, por lo tanto, para su supervivencia necesita utilizar algún material previamente almacenado. Este material se denomina vitelo, está compuesto por proteínas, lípidos, azúcares, otros. La principal fuente productora de las proteínas del vitelo es el tejido extra ovárico denominado cuerpo graso; y la proteína esencial del vitelo es la vitelogenina, que sule de aminoácidos, fosfatos y en menor medida lípidos, al embrión en desarrollo (Atella et al., 2005; Melo et al., 2000). Las moléculas de vitelogenina son modificadas para permitir la subsecuente secreción a la hemolinfa, a partir de donde son captadas por los oocitos en desarrollo a través de endocitosis mediada por receptores. Estos receptores se encuentran en la membrana oocítica, por lo que es necesario que las células foliculares disminuyan su tamaño y abran canales entre ellas para permitir el traspaso de la vitelogenina, fenómeno denominado patencia. Una vez en el interior, es almacenada en gránulos vitelínicos para su posterior consumo (Davey, 2007; Davey, 1987; Davey and Huebner, 1974). En el presente trabajo, mediante la utilización de anticuerpo anti-vitelina, se pudo constatar y dar sustento a la actividad del ovario como órgano vitelogénico complementario, ya que se observó una mayor presencia de la proteína en los oocitos en estado vitelogénico. Esto coincide con la captura de vitelogenina extra ovárica por el oocito, y a la propia producción del precursor en las células foliculares, en correlación a lo reportado a lo que respecta a la activación de la traducción de la vitelogenina con el cese de la absorción de la proteína desde la hemolinfa (Melo et al., 2000). A su vez, concuerda con observaciones realizadas en otros insectos, en los que se ha demostrado que el epitelio folicular también está involucrado en la síntesis de proteínas del vitelo (Bellés, 1998; Giorgi et al., 2005). Por otra parte, en el estado previtelogénico, se observó una fuerte acumulación de la proteína vitelogenina/vitelina en la región basal del epitelio folicular, con un patrón débil en el interior del oocito (Figura 24). Esto sería indicativo de que durante la previtelogénesis la proteína sintetizada en el ovario es almacenada en el epitelio folicular hasta ser tomada por el oocito durante la vitelogénesis, siendo una fuente propia del ovario que permite suplir las necesidades nutricionales del embrión. Esta acumulación fue reportada en *Triatoma infestans*, triatomino muy emparentado a *R. prolixus*, en el que además se reportó que la expresión génica de vitelogenina es constitutiva en el ovario previtelogénico (Blariza et al.,

2016). Por último, la detección de una fuerte señal de vitelina en el interior de los oocitos en desarrollo avanzado, localizada de manera organizada dentro de los gránulos de vitelo, fue coincidente con la disposición reportada en otros triatominos (Blariza et al., 2016)

Una vez revisada la morfología del ovario y el proceso de vitelogénesis, se continuó estudiando los genes activos durante la oogénesis, en base a la homología de secuencia con otras especies, y describir cuales se encuentran activos cumpliendo un rol específico durante el desarrollo.

Se comenzó con el estudio de *Bicaudal C*, el cual mostró desempeñar un papel fundamental en la oogénesis durante la formación del huevo. Los huevos en hembras mutantes de *R. prolixus* exhibieron intacta la polaridad antero-posterior, pero mostraron un perfil irregular de las células foliculares a lo largo de toda su superficie tanto en la región apical como lateral, desarrollando protuberancias en la superficie más externa del corion de tamaño significativo. Este gen fue estudiado en el hemiptera *Nilaparvata lugens*, especie evolutivamente más cercana al triatomino. Se reportó que la polaridad del huevo no se encuentra afectada por el silenciamiento génico, aunque sí se produce un detenimiento en la absorción de la proteína vitelogenina en el oocito en desarrollo, inhibiendo el proceso de vitelogénesis y desencadenando la producción de huevos no embrionados. Por otro lado, no se reportó la presencia de alteración alguna en las células foliculares, como si se encontró en el triatomino (Zhang et al., 2015). En comparación con *D. melanogaster*, se reportó que los huevos de hembras mutantes para *BicC* pierden la polaridad antero-posterior del huevo y exhiben estructura de copa debido a la falla en la migración de las células foliculares para rodear en su totalidad al oocito (Mahone et al., 1995); A pesar de *D. melanogaster* y *R. prolixus* son organismos evolutivamente más alejados y por lo tanto presentar distintos mecanismos para dirigir la oogénesis, comparten la función de *BicC* a nivel folicular.

En *R. prolixus* se conoce que el epitelio folicular atraviesa por una secuencia cronometrada de cambios en la forma y en el contacto celular, comprendiendo desde un estado herméticamente cerrado a uno abierto o permeable (Huebner y Anderson, 1972; Huebner e Injeyan, 1980). Durante las fases tempranas de la oogénesis, previtelogénesis y oogénesis temprana, las células foliculares aumentan su longitud, comenzando con un tamaño entre 100-200  $\mu\text{m}$  hasta una longitud de 400  $\mu\text{m}$ , y forman uniones intercelulares septadas de

manera regular a lo largo de toda su superficie lateral, estableciendo un patrón curvo característico. Adicional a estas uniones celulares, se establecen uniones tipo *gaps* y se forman los desmosomas, que son uniones basolaterales que también colaboran para mantener adheridas células vecinas. Frente a esta situación de estrecho contacto celular, es necesaria una modulación de las uniones para la generación de la permeabilidad epitelial, y así dar inicio a la vitelogénesis, con la toma de los precursores de vitelo por el oocito en desarrollo. El desarrollo de la permeabilidad, regulada hormonalmente, se correlaciona con la reducción del volumen celular, la modificación del citoesqueleto celular y la reorganización de las uniones celulares específicas de cada etapa (Abu-Hakima y Davey, 1977, Abu-Hakima and Davey, 1979; Huebner and Injeyan, 1981; Irlles et al., 2013). Además de las alteraciones a nivel celular, para potenciar el ingreso de precursores, exhiben canales interfoliculares que también sufren cambios en su permeabilidad correlacionados con el proceso de vitelogénesis (Huebner and Injeyan, 1981). Estos cambios indispensables no alteran la organización y disposición de las células que rodean al oocito en desarrollo (Huebner, 1981).

En hembras *Rp-BicC*, se observaron defectos a nivel folicular desde la fase previtelogénica, en la cual las células se disponen de manera tal que generan grandes espacios intercelulares, a diferencia de las condiciones normales donde se observa estrecha asociación (Figura 29). Este aumento en los espacios interfoliculares y desorganización celular en el epitelio folicular se mantienen durante toda la oogénesis, registrado también en la fase vitelogénica en los cortes histológicos teñidos con H&E (Figura 29). Esta desorganización celular no sólo afecta a la morfología del oocito en desarrollo, y por lo tanto a la formación del huevo, sino que perturba el traspaso de moléculas a través del epitelio. En primer lugar, se observaron diferencias importantes en cuanto a la intensidad de señal proporcionada por la localización de vitelogenina/vitelina en la ovariola correspondiente a la hembra silenciada genéticamente (Figura 31). La acumulación de vitelogenina en el epitelio folicular, dada tanto por la propia producción del ovario como por la captura desde la hemolinfa, se vio disminuida a lo largo de toda la ovariola, sugiriendo que la presencia alterada podría deberse a una desregulación en la producción proteica del epitelio folicular, en correlación a la alteración que presentan las células foliculares en cuanto a tamaño y forma. Otra hipótesis que podría estar explicando lo ocurrido es una falla en la modulación de la permeabilidad folicular durante la vitelogénesis

dada por la disposición irregular que muestran las células del epitelio. Aunque no fue posible adjudicar una causa a esta disminución de vitelina, es importante remarcar que la toma de componentes del vitelo se vio afectada significativamente durante la ovogénesis, y esto refleja principalmente la falta de la coloración rosa característica de los oocitos en desarrollo y los huevos ovipuestos de hembras silenciadas para *BicC* (Figura 27). El color rosa se debe a la presencia de una hemoproteína denominada proteína de unión al grupo hemo en *Rhodnius* (RHBP) que es capturada por el oocito vitelogénico a través de endocitosis mediada por receptores, proceso análogo a la toma de vitelogenina (Braz et al., 2002). De toda la descendencia derivada de las hembras silenciadas genéticamente ninguna mostró la coloración característica rosa. como fue observada en la descendencia control.

En paralelismo a la menor señal, no se pudo determinar la presencia, de manera diferencial, de la proteína, en relación al lado basal de la membrana del epitelio folicular y al lado apical de la membrana oocítica. De hecho, cuando se lo analizó en detalle, no fue posible identificar de manera fehaciente el límite que envuelve al oocito en desarrollo, infiriendo alguna falla en su producción y mantenimiento. Esta hipótesis o inferencia es apoyada por el análisis de distribución de lípidos con la sonda FM4-64FX, donde no se pudo obtener una marcación diferencial y nítida de la membrana plasmática basal, indicando una alteración en su constitución (Figura 30).

Sumado a todo esto, y en correlación a la alteración presente en el epitelio folicular, posteriormente se decidió evaluar el traspaso de moléculas inespecíficas, y de tamaño variado, a través de los canales que se establecen durante la vitelogénesis. La utilización de moléculas de dextrano de distinto tamaño conjugadas a fluoróforos permitió observar en los ovarios control, el fenómeno de modulación folicular durante la vitelogénesis, ya que únicamente fue posible observar el traspaso de moléculas de gran tamaño (70kDa) en el oocito vitelogénico (Figura 34), lo que es coincidente con la captura de la vitelogenina (~200kDa) por el oocito (Tufail et al., 2014). Por otro lado, en los ovarios interferidos se observó alteración de esta modulación folicular ya que se evidenció un flujo bidireccional establecido en el epitelio folicular; explicado por la relación negativa existente entre el aumento del espacio intercelular folicular y la disminución de fluorescencia, correspondiente a los dextranos, en los oocitos en desarrollo (Figura 33).

Todos estos resultados sugieren que la mal formación del corion y la oviposición de huevos no pigmentados y más pequeños, sería consecuencia de una mal formación y desorganización de las células foliculares, lo que genera un traspaso de moléculas a través del epitelio folicular de manera bidireccional y no unidireccional, como ocurre en condiciones normales. Este flujo bidireccional explicaría la disminución de la presencia de componentes del vitelo en los oocitos silenciados, como ocurre con la proteína HRBP y la vitelogenina/vitelina, así como también con el pasaje de moléculas de dextrano. Cabe aclarar que no se pudo determinar si la captura por endocitosis mediada por receptores también se ve afectada. En base a los resultados y al conocimiento previo, se plantea que este proceso se ve afectado, debido a que observamos una alteración en la formación de la membrana basal epitelial y en el lado apical de la membrana oocítica, lugar donde se encuentran los receptores que median la endocitosis (Tufail and Takeda, 2008). Por otro lado, es importante aclarar que no se pudo identificar con certeza alteración funcional en las células foliculares, asociada a la merma de vitelogenina, más allá de poder intuir desregulación en la síntesis de proteínas debido a su morfología inusual y disminución del tamaño nuclear.

Continuando con el análisis genético, se decidió analizar al gen *cornichon (cni)*. El mismo fue descrito en *D. melanogaster* como miembro fundador de una familia de proteínas conservadas que median la exportación de moléculas desde el retículo endoplasmático al citoplasma (Roth et al., 1995). Durante la oogénesis de *D. melanogaster*, *cni* está involucrado en el establecimiento de los ejes antero-posterior y dorso-ventral del oocito en desarrollo. Su acción está dada a través de la mediación del transporte del factor de crecimiento (TFG), *gurken (grk)*, a la superficie del oocito. El ARNm correspondiente a *grk* se localiza alrededor el núcleo del oocito en desarrollo, y una vez activada la traducción, el producto proteico se secreta hacia la capa folicular para dar inicio principalmente a la activación de la cascada de señalización que establece el eje dorso-ventral (Bökel et al., 2005). Algo importante de remarcar es que *grk* únicamente fue reportado en *D. melanogaster* (Dearden et al., 2006). La ausencia de un gen ortólogo generó intriga sobre la función que podría llegar a desempeñar *cni* en *R. prolixus*. Utilizando la técnica de ARNi parental no fue posible observar función alguna durante el desarrollo embrionario, sólo se observó un cambio muy sutil con lo que respecta a

la disposición del embrión dentro del huevo, pero no fue una alteración representativa de toda la descendencia obtenida (Figura 37).

Reportado en *D. melanogaster*, se encontró que, además de las funciones embrionarias, *cni* presenta funciones pleiotrópicas somáticas. El silenciamiento génico de *cni*, en el díptero, produce adultos que presentan ojos rugosos, pérdida de setas ocelares y modificación de la venación alar (Bökel et al., 2005). Aunque resulta interesante esta función reportada, esto escapa de los objetivos planteados en este trabajo por lo que se decidió no abordar dicha vía.

Por otro lado, es importante no pasar por alto que el amplicon utilizado para llevar a cabo el análisis funcional del gen presentó una longitud pequeña, debido a como se mencionó previamente resultó dificultoso encontrar otra región óptima para lograr diseñar otro par de oligos específicos. La opción de plantear uno nuevos oligos podría corroborar la ausencia de función durante la embriogénesis temprana de *R. prolixus*, o indicar la presencia de isoformas del transcripto con distintas actividades funcionales durante el ciclo de vida del triatomino.

Más allá de no reconocer una función de *cni* durante el desarrollo embrionario sí se pudo identificar expresión génica de manera constitutiva durante los distintos tiempos del desarrollo estudiados, resultados coincidentes con lo reportado en la parte I de este trabajo en relación al análisis de datos obtenidos por ARN-seq.

Otros de los genes que se prosiguió a analizar fue *Bicaudal D*. Al igual que los estudiados anteriormente, fue identificado por primera vez en la mosca de la fruta, *D. melanogaster*. En esta, se reportó que codifica para una proteína que conforma la maquinaria de localización de los ARNms en el oocito en desarrollo, para posteriormente determinar la polaridad de los ejes del embrión. Esta maquinaria interactúa con los microtúbulos citoplasmáticos para permitir el movimiento de los ARNms a los distintos compartimientos celulares (Claussen and Suter, 2005; Huynh and St Johnston, 2000). Se conoce que está en contacto con un amplio repertorio de genes sobre los que actúa durante la oogénesis y el desarrollo embrionario temprano, para organizar su localización final y permitir el establecimiento de la polaridad embrional (Vazquez-Pianzola et al., 2017).

En el presente trabajo se observó que el silenciamiento génico desencadenó la producción de huevos no embrionados por las hembras de *R. prolixus* tratadas con el ARNdc. Se observó

un número muy bajo de huevos embrionados, atribuibles a la capacidad de autogenia característica de la especie (Davey, 1987). Los huevos ovipuestos por las hembras silenciadas mostraron los atributos característicos presentes en los huevos control, polaridad antero-posterior del huevo, pigmentación rosa correspondiente a la captura de la proteína de unión al grupo hemo por el oocito y tamaño coincidente con las medidas estándares generales. A pesar de esto, no se visualizó esbozo de embrión alguno, ya que no se observó ninguna estructura característica del desarrollo embrionario, indicando que la interferencia impidió el desarrollo embrionario desde etapas muy tempranas del proceso. Estos resultados permiten suponer una función conservada del gen entre el triatomino y *D. melanogaster*, quedando por identificar genes blancos ortólogos y determinar la relación existente. Para sustentar a esta inferencia funcional se logró identificar que *BicD* se expresa desde muy temprano; se expresa tanto en el ovario como en los estadios embrionarios más avanzados. A pesar de esta expresión temprana, no se identificó transcripto alguno en huevos correspondientes al tiempo post oviposicional de 0-5 hs, ni por análisis mediante hibridación *in situ* ni por RT-PCR (Figura 43 - 44). Esta ausencia de expresión activa es coincidente con la función reportada para *D. melanogaster*. La función transportadora de ARNm en el díptero se da en los tiempos del desarrollo que comprenden las 0 a 8hs post ovipuesta, permitiendo inferir correlación funcional con el triatomino.

Por otro lado, es importante remarcar es que la inyección de ARNdc<sup>*BicD*</sup> produce una alta letalidad en las hembras interferidas, por lo que se puede hipotetizar plantear la existencia de un rol de *BicD* durante la adultez; aunque esta función aún no ha sido reportada en ningún otro insecto estudiado.

# **Conclusiones Finales**

- La buena calidad y óptima profundidad obtenida tras la secuenciación masiva de los diferentes estadios estudiados, permitió reconstruir con una muy buena cobertura los transcriptomas embrionarios secuenciados.
- El enfoque cualitativo del análisis de expresión génica, realizado en este trabajo, permitió corroborar la existencia de un amplio repertorio de genes con expresión activa durante el desarrollo embrionario temprano.
- Las ambigüedades encontradas en la identificación de transcritos, provenientes del mapeo contra el genoma anotado de *R. prolixus*, exponen las limitaciones que presenta dicha predicción *ab initio*, y las necesidades de ser mejorados para su uso futuro.
- La identificación de genes embrionarios conservados evolutivamente, sugiere que los eventos que atraviesan las diferentes especies durante su desarrollo embrionario, ineludibles para construir su plan corporal final, se encuentran regulados por los mismos actores genéticos.
- La identificación de una expresión génica generalizada de todos los genes estudiados a lo largo de los tiempos del desarrollo analizados, sugiere la existencia de una transcripción, de origen materna, extensiva.
- La utilización de tinciones con fluorescencia permitió identificar a los distintos tipos celulares que componen a la ovariola, así como también seguir los cambios que experimentan durante el proceso de oogénesis.
- Durante el proceso de vitelogénesis, se observó un aumento de la presencia de vitelina en el tejido ovárico, más precisamente en los cuerpos vitelínicos, como consecuencia de la captación mediada por endocitosis de la vitelogenina sintetizada por el cuerpo graso.

- La identificación de la localización de vitelina en el epitelio folicular del estadio previtelogénico se explicaría a través de la síntesis de proteínas del vitelo por el epitelio folicular.
- *Bicaudal C*, en *R. prolixus*, actúa manteniendo la estructura del epitelio folicular de manera organizada durante todo el proceso de oogénesis, dando lugar a una acumulación controlada de vitelo y a un correcto establecimiento del corion.
- *Bicaudal D* es necesario para la producción de huevos embrionados en *R. prolixus*. Su silenciamiento lleva a la oviposición de huevos en los cuales no es posible identificar ninguna estructura característica que permita inferir los ejes embrionarios corporales.
- *cornichon* no desempeña ninguna función característica durante la embriogénesis de *R. prolixus*, aunque si presenta una activa expresión génica durante los diferentes estadios del desarrollo estudiados (0-12-24-48 horas post oviposición y huevos no fertilizados).
- *Bicaudal D* y *Bicaudal C* presentan un patrón de expresión materno y folicular en el ovario. Por otro lado, a nivel embrionario *Bicaudal D* no se expresa en el estadio correspondiente al estado de cigoto (0 horas post oviposición).

# Suplementario

## 1. Suplementario 1

*Script* codificado en lenguaje *Python* utilizado para la remoción de secuencias no deseadas en el set de datos crudos obtenidos por la plataforma de secuenciación. A modo de ejemplo se escribe, en rojo, uno de los archivos que se utilizó durante este trabajo.

```
#!/usr/bin/python

from Bio import SeqIO

fasta_file= "mpg_L11463-1_DmSF-RNA_S4_R1_001.fasta" # Input
fast[a|q]file
wanted_file= "hit_Univec.out" # Input interesting sequence IDs, one
per line
result_file  = "mpg_L11463-1_DmSF-RNA_R1_SUCIAS.fasta" # Output
fast[a|q] file. Sucias
result_file_2 = "mpg_L11463-1_DmSF-RNA_R1_LIMPIAS.fasta" # Output
fast[a|q] file. limpias

wanted = set()
with open(wanted_file) as f:
    for line in f:
        line = line.strip()
        if line != "":
            wanted.add(line)

fasta_sequences = SeqIO.parse(open(fasta_file),"fasta")
file = open(result_file_2, "w")

with open(result_file, "w") as f:
    for seq in fasta_sequences:
        if seq.id in wanted:
            SeqIO.write([seq], f, "fasta")
        else:
            SeqIO.write([seq], file, "fasta")

file.close()
```

## 2. Suplementario 2

Corresponde a un archivo en formato Excel donde se listan los transcriptos reconstruidos, para cada uno de los transcriptomas secuenciados, junto al correspondiente valor de normalización FPKM. Cada transcriptoma reconstruido se lo halla en las distintas hojas del archivo, en donde, cada tiempo es representado con un color específico.

Referencias: (SE) *single-end*, (PE) *paired-end*, (Bo) *Bowtie*, (Top) *Tophat*. *Ohpo* (verde), *12hpo* (rosa claro), *24hpo* (violeta), *48hpo* (amarillo), *ovario* (gris) y *no fertilizados* (naranja). Debido al extenso tamaño del archivo el mismo fue adjuntado como archivo extra al presente trabajo.

### 3. Suplementario 3

Alineamiento múltiple generado por *ClustalW* para el gen *BicC*. Además de las predicciones proteicas correspondientes a las dos unidades transcripcionales identificadas como *BicC* en *R. prolixus* (RPRC001613 + RPRC001612) se incluyeron las secuencias de *BicC* en: *Harpegnathos saltator* (gi|749730745 - gi|749730739-X1), *Bombus terrestris* (gi|808147069), *Apis florea* (gi|820865347), *Megachile rotundata* (gi|805824678), *Acromyrmex echinatior* (gi|332022439), *Tribolium castaneum* (gi|91089717), *Pediculus humanus* (gi|242009357), *Drosophila melanogaster* (gi|24584541), *Mus musculus* (AAK27347.1), *Nilaparvata lugens* (Zhang et al., 2015).

La conservación de aminoácidos se visualiza con bloques negros, la conservación a nivel de grupo aminoacídico con bloques en color gris. Los guiones muestran la ausencia de secuencia alineada a lo largo del alineamiento.

10 20 30 40 50 60

RPRC001613  
RPRC001612  
*Harpegnathos saltator* -----MRPDEVVQSERRAAETMSETSEGTATTSVSGKSGFESREEM  
*Harpegnathos saltator* X1 -----MRPDEVVQSERRAAETMSETSEGTATTSVSGKSGFESREEM  
*Bombus terrestris* -----MRPDEVIQNERKATETMSETSEGTMTTSVSGKSGFESRQEL  
*Apis florea* -----  
*Megachile rotundata* -----  
*Acromyrmex echinaior* -----M  
*Tribolium castaneum* -----MSCNCATPRGIEVKNSDTQSEVSE-SGTATSGFNSNWSREEL  
*Pediculus humanus* MVECYVTIECTKLTFAQTIIIIIIIMRPTYASTDNLSEISESGTGRTGSTSGRSGYSDNR  
*Drosophila melanogaster* -----  
*Mus musculus* -----MASQSEPGYLAQAQSDPGSNSESTDSVPVAGSE  
*Nilaparvata lugens* -----MAKAKGSDTMSEVSESGTSACTFNRSR-----REEL

70 80 90 100 110 120

RPRC001613  
RPRC001612  
*Harpegnathos saltator* -----MQETKTHIAWP  
*Harpegnathos saltator* X1 -----MQETKTHIAWP  
*Bombus terrestris* -----MQETKTHIAWP  
*Apis florea* -----MTIMEETNTEVTWP  
*Megachile rotundata* -----MEETNTEVTWP  
*Acromyrmex echinaior* GYLKRIKLNH-----YLMKYSGLTELNYDNQS---TVIFWHMVMEERTNTEVTWP  
*Tribolium castaneum* RDFVAKHGLKDISDLYQDRFRVDRKKLEQMLLGEVGP---ADTFFHNVMEETNTEVTWP  
*Pediculus humanus* DELLSILGLTEDDEIFTDRFRIDRKKLEDMIMSDNDAL-DATQAFFORIMEGSGTHIAWP  
*Drosophila melanogaster* -----MNTITDTYVSWP  
*Mus musculus* DDLVAAAPLLHSPWSEERFRVDRKKLEAMLQAAAEGKGRSGEDFFOKIMEETNTQIAWP  
*Nilaparvata lugens* QELARTLGLSSVDLLQQDRFRVDRKKLEAMLLGSSDID-ISANDFFOKIMTETNTEVLP

130 140 150 160 170 180

RPRC001613  
RPRC001612  
*Harpegnathos saltator* SRLKLGAKSKKDPHIVIGR-AEDVKAARERIVAAALDTK-SNRVNMKIDVSYTDHSHIIG  
*Harpegnathos saltator* X1 SRLKIGAKSKKDPHIKVAGR-LDDVRSAKEKIMEILDTRQSNRVTMKLDVSYTDHSHIIG  
*Bombus terrestris* SRLKIGAKSKKDPHIKVAGR-PDDVRAAKEKIMEILDTRQSNRVTMKLDVSYTDHSHIIG  
*Apis florea* SRLKIGAKSKKDPHIKVAGR-PDDVRAAKEKIMEILDTRQSNRVTMKLDVSYTDHSHIIG  
*Megachile rotundata* SRLKIGAKSKKDPHIKVAGR-PDDVRAAKEKIMEILDTR-SNRVTKMLEVSYTDHSHIIG  
*Acromyrmex echinaior* SRLKIGAKSKKDPHIKVAGR-LDDVRSAKEKIMEILDTRQSNRVTMKLDVSYTDHSHIIG  
*Tribolium castaneum* NKLGKLGAKSKKDPHVRVAGK-EEDVKAADRIMTVLYTR-CNRVTKMMDISYTDHSHIIG  
*Pediculus humanus* IRLKIGAKSKKDPHVRVAGRSLQDIQTAKDTILSALGIR-NKQVTMKNVSYTHSHIIG  
*Drosophila melanogaster* CRLKIGAKSKKDPHVRIVGK-VDQVQRAKERILSSLDSR-GTRVIMKMDVSYTDHSHIIG  
*Mus musculus* SRLKIGAKSKKDPHIKVSQK-KEDVKEAKEMIMSVLDTK-SNRVTKMDVSETEHSHVIG  
*Nilaparvata lugens* SRLKIGSKSHKDPVIRVIGR-LDDIKIAKKNKIMSVLDTR-GSRVIMKLEVSYTDHSHIIG

190 200 210 220 230 240

RPRC001613 **KGGLTIKRVMEETGCHIHFPDSNRNNPTEKSNQVSIAGEVWGVEKARARVRELTPLI**FCF

RPRC001612

*Harpegnathos saltator* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Harpegnathos saltator X1* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Bombus terrestris* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Apis florea* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Megachile rotundata* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Acromyrmex echinatio* **KGGLTIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Tribolium castaneum* **RGGLSIKRVMEETGCHIHFPDSNRSNHODKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Pediculus humanus* **EGGQTITNVRKQTGCHIHFPDSNRSGTETKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Drosophila melanogaster* **RGGNLIKRVMEETGCHIHFPDSNRNNOAEKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Mus musculus* **KGGNLIKRVMEETGCHIHFPDSNRNNOAEKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

*Nilaparvata lugens* **KGGLVIKRVMEETGCHIHFPDSNRNNOAEKSNQVSIAGEVMEGVERARARVRNLTPLIFSF**

250 260 270 280 290 300

RPRC001613 **ELPILGALSLASDCYSAYIKSLQEKYNVQVMFTRPKLNS TLVVVKGSEWEVARVKDATQ**

RPRC001612

*Harpegnathos saltator* **ELPIMGASQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Harpegnathos saltator X1* **ELPIMGASQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Bombus terrestris* **ELPIMSASQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Apis florea* **ELPIMSSTQALPDSTIPYLVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Megachile rotundata* **ELPIMGSSQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Acromyrmex echinatio* **ELPIMGASQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Tribolium castaneum* **ELPIMAON--IDNTCPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Pediculus humanus* **ELPVLGSIQSTPDCSTSPYVVKIQEKYNVQVMFTRPKLHATLVVVKGCEWEVSQVKEATV**

*Drosophila melanogaster* **ELPVMGPNKQPDHETPYIKMIETKENVQVIFSTRPKLHATLVVVKGCEWEVSQVKEATV**

*Mus musculus* **ELPIAGILQPVDPNTPSIQHSQTYSVSVSKQRSRMYGATVIVRGSQNTNAVKEGTA**

*Nilaparvata lugens* **QLPVMGTVHSASELASSYITGLQERYNVHVKFCQSKLYSTLVVVKGCEWEVSQVKEATV**

310 320 330 340 350 360

RPRC001613  
RPRC001612  
*Harpegnathos saltator*  
*Harpegnathos saltator* X1  
*Bombus terrestris*  
*Apis florea*  
*Megachile rotundata*  
*Acromyrmex echinaior*  
*Tribolium castaneum*  
*Pediculus humanus*  
*Drosophila melanogaster*  
*Mus musculus*  
*Nilaparvata lugens*

370 380 390 400 410 420

RPRC001613  
RPRC001612  
*Harpegnathos saltator*  
*Harpegnathos saltator* X1  
*Bombus terrestris*  
*Apis florea*  
*Megachile rotundata*  
*Acromyrmex echinaior*  
*Tribolium castaneum*  
*Pediculus humanus*  
*Drosophila melanogaster*  
*Mus musculus*  
*Nilaparvata lugens*

430 440 450 460 470 480

RPRC001613  
RPRC001612  
*Harpegnathos saltator*  
*Harpegnathos saltator* X1  
*Bombus terrestris*  
*Apis florea*  
*Megachile rotundata*  
*Acromyrmex echinaior*  
*Tribolium castaneum*  
*Pediculus humanus*  
*Drosophila melanogaster*  
*Mus musculus*  
*Nilaparvata lugens*

490 500 510 520 530 540

RPRC001613 ---NAGPLFKTIWN---FP-----VYPELP-----SSPLMLGPIVQQTSPGE---N

RPRC001612

*Harpegnathos saltator* ---NAGSIFQGNSNNNDVR-----

*Harpegnathos saltator X1* ---NAGSIFQGNSNNNDVSSSNLMSSVPDNFSNMLTINTQ-----NSPYCVSPIPH

*Bombus terrestris* ---NAGNIFQGNSTGNGSSPNSLVSSLSNDLSSILTINTQ-----NSPYCVSPVPH

*Apis florea* ---NVGNIFQGNSSNGSSPPSSLVSSLSNDLSSILTINTQ-----TSPYCMSPISH

*Megachile rotundata* ---NAENVFQNGS-NGSATNNNLVGNLSDLSNILTINTQ-----NPPYCVSPMSH

*Acromyrmex echinator* ---NAGNIFQ-SSSNNDSNNNLI SGVQENFSSILNVNTQ-----NSSYCVSPISH

*Tribolium castaneum* -----

*Pediculus humanus* ---NPPKIFDVNEINDINSTRALPSILTLDNGSSFPQLSLS-----DHSPLLMPSL

*Drosophila melanogaster* KDLNLAYRTQLTALLAGYVDSPKTPSLLPPSLAGQLTPYAN-----NNHLLLNGANG

*Mus musculus* DILASAGLGLTGLGLLGPPTLSLNTSATPNSLLNALNTSVSPLQSSSSGTPSPTLWAPPI

*Nilaparvata lugens* -----

550 560 570 580 590 600

RPRC001613 VWPSL--PHLMAAQHVP--VLTSLSQPTHRLH-----QHLLQQQR-----HHL-

RPRC001612

*Harpegnathos saltator* -----THHPYPYT-HLNHLLTIQHMLHNNMMPH---PHGM

*Harpegnathos saltator X1* SPNPMGLSPHWG-LPHLPSVFSPLGTHHPYPYT-HLNHLLTIQHMLHNNMMPH---PHGM

*Bombus terrestris* SPNPMALSPHWGHLPLPSIFSPPLPHHSYPYP-HLNHLLTIQHVMHNNAMPP--HSHGL

*Apis florea* SPSPLGLSPHWG-LPPIPSMFSPPLPHHTYTYQQHLNHLLTIQHVMHNNAI PP--HTHAQ

*Megachile rotundata* SPNPMGLSPHWG-LSTIPSMFSSIPLHHAYPYP-HLNHLLTIQHVLNNAMPT--HSHGS

*Acromyrmex echinator* SPNSLGLSPHWG-LTHHYVS-----SLCPYPYS-HLNHLLTIQHMLHNNIMGTSHGPHGL

*Tribolium castaneum* -----STSPNNDLTIPTNLNPSPLTPIFTWHFSS

*Pediculus humanus* YNPIWGSQLPPGLMSIAWGQMMLPQQQQQQQQQQQQQHQQQQQTHQQQQPMIMS--RPRP

*Drosophila melanogaster* LATPTGVCAPTQKYMQLHNSFQQAQNRSMVAGGQSNNGNYLQVPGAVAPPLKPTVSPRN

*Mus musculus* ANTASATGFSTIPHLMLPSTAQATLTNILLSGVPTYGHTAPSPPPGLTPVDVHINSMQTE

*Nilaparvata lugens* -----NGPTPMKPKVKNPPMGCQNIPLPLG

610 620 630 640 650 660

RPRC001613 -DHLGH-----FSYSSLSLPELPVERF

RPRC001612

*Harpegnathos saltator* SNHVMS-----GLGPFHHNASTNFPGIHGLNTNSMTDSKEGSSAYSSLSVSS-----

*Harpegnathos saltator X1* SNHVMS-----GLGPFHHNASTNFPGIHGLNTNSMTDSKEGSSAYSSLSVSS-----

*Bombus terrestris* QNHAF-----GIGQLHSNGPVGPHSTHGLNGNSLVESKEG-SAYSSLSVSS-----

*Apis florea* QNHGFS-----GIAQLHSNVPPGFHSIHGLNGNSLTDSEK-SAYSSLSMSS-----

*Megachile rotundata* TNHGFT-----GMGQLRSNSSAGFHGIHGPNSSLPDSKEG-STYSSLSVSS-----

*Acromyrmex echinator* SNHVMPGP---PGIGAFHHNNVPTNFSGIHGLNANSMLDMSKDGGSAYSSLSVSS-----

*Tribolium castaneum* DQTTP-----FEIVQKPQMLYNNNNNKMQQSSNS-----

*Pediculus humanus* TSTVTGR-----DISVGSYGYSQTLNQFGSMSTSPISDTNGLSAISSFSNTS---

*Drosophila melanogaster* SCSQNTSGYQSFSSSTTSLEQSYPPYAQLPGTVSSTSSSTAGSQNRAHYSPPDSTYGSEG-

*Mus musculus* GKNISASINGHVQAPANMKYGPLSTSSSLGEKVLSSNHGDP SMQTAGPEQASPKNSVVEGCN

*Nilaparvata lugens* NLPEP-----DLNQLFYQISQFFSQHQQQGSLNENALGQVSSLSLDPD-----

```

          670          680          690          700          710          720
  . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
RPRC001613  G-----
RPRC001612  -----
Harpegnathos saltator  -----SLSSPAISPRNISPVNPTDTSPNLADLSSML-----
Harpegnathos saltator X1  -----SLSSPAISPRNISPVNPTDTSPNLADLSSML-----
Bombus terrestris  -----SLSSPAISPRNISPVNPIESSHNIDLSSML-----
Apis florea  -----SLSSPAISPRNISPVNPTENNPNIDLSSML-----
Megachile rotundata  -----SLSSPAISPRNVSPVNPVETSPNLDLCSML-----
Acromyrmex echinaior  -----SLSSPAISPRNISPVNPTDTNTNLADLSSML-----
Tribolium castaneum  -----GFQSFEQVKADTSLSSSISSNSSNYVSSP-----
Pediculus humanus  -----SLSTPGVSPRNQSPTMQMDHQLSLGKTKEMKRQR
Drosophila melanogaster  -----GGVGGGGGGGARLGRRLSDGVLLGLSNNGGGGNSGG---
Mus musculus  DAFVEVGMPRSPSHSGNAGDLKQMLGASKVSCAKRQTVELLOQTKNSHLHGTDRLLS---
Nilaparvata lugens  -----IFKETHSSONSRCSSMLSPNSKSAHGSGQ-----

```

```

          730          740          750          760          770          780
  . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
RPRC001613  -----
RPRC001612  -----MAISDSTSRFSD-QV-ERKAPGFDRPMLT-----PQD
Harpegnathos saltator  -----SDLSVTDRRRAPGCEKKILEMAAQQNLT-----PFD
Harpegnathos saltator X1  -----SDLSVTDRRRAPGCEKKILEMAAQQNLT-----PFD
Bombus terrestris  -----SDLTVPDRRRAPGCEKKSLEMAVQQNLV-----PFD
Apis florea  -----SDLTVSDRRRAPGCEKKSLEMAVQQNLV-----PFD
Megachile rotundata  -----SDLSVADRRRAPGCEKKSLEMAVQQNFA-----PFD
Acromyrmex echinaior  -----SDLPVTDRRRAPGCEKKSLEMAAQQNLT-----PFD
Tribolium castaneum  -----PTVFSQSDNNIPTTSDYQSMNEKKSA-----SFD
Pediculus humanus  TIDSAENLLLKNVKEDLSNMMKVTPDWSDQRNLGFDSTPIQTLT-----PID
Drosophila melanogaster  -----AHLLPGSAESYRSLHYDLGGNKHSGHR-----AFD
Mus musculus  -----DPELSATESPLADKKAPGSERAAERAAAQQKSERARLASQPTYVHMQAFD
Nilaparvata lugens  -----IPNISSDLSLSANSNDWRAPGFGRKS-----HIE

```

```

          790          800          810          820          830          840
  . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
RPRC001613  -----
RPRC001612  Y-QRKKIHAHQALQ-KPVSCREIRVPNSSWEGYGESQSSPTAALREQ-KIIDEKRAVQEW
Harpegnathos saltator  YEQKKLLA-AKAMQSKP-SG-DYRVPTSAWSGYGLSQSIPPIITTSDLSKE-----
Harpegnathos saltator X1  YEQKKLLA-AKAMQSKP-SG-DYRVPTSAWSGYGLSQSIPPIITTSDLSKE-----
Bombus terrestris  YEQKKLLA-AKVIQCKPSTN-DYRVPTSAWSGYGLSQSIPPIINTGDIKKE-----
Apis florea  YEQKKLLA-TKAIQIKPATN-DYRVPTSAWSGYGLSQTIPPMSATDLNKEL-----
Megachile rotundata  YEQKKLLA-AKAIQSKPSFN-EYRVPTSAWSGYGLSQSIPPIGTTDLNKEL-----
Acromyrmex echinaior  YEQKKLLA-TKAMQNKSNTN-DFRVPTSAWSGYGLSQSIPPIITTSDLSKE-----
Tribolium castaneum  YDCSKRLAGFKAMQGRPNPD-VYRVPTAMWAGYGISHTSPAG-----
Pediculus humanus  YEQKKFLA-VKAIQNMPNSN-EYRVPTSAWSGYGESLTSPSSLFSDOKEFMNS-----
Drosophila melanogaster  FDMKRALG-YKAMERTPVAG-ELRIPTIAWYGMGLSSTSPAPAPLENGENG-----
Mus musculus  YEQKKLLA-TKAMLKKPVVT-EVRIPTNTWSGLGESSKSMPAETIKELRRANHVSYKPTMT
Nilaparvata lugens  Y-QEMKQQALIALQ-KPHPD-DPRVPTSIWAGYGLSRSCPAVVMNELQNS-----

```

850 860 870 880 890 900

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RPRC001613  
RPRC001612 LGN-----GGVENDYSSTQNNFDIVSSRTNVSD---DQ  
*Harpegnathos saltator* -----LTSHPSDLWKEPTTP-VFSTEMDFG-MISSKDRIGQ---I  
*Harpegnathos saltator X1* -----LTSHPSDLWKEPTTP-VFSTEMDFG-MISSKDRIGQ---I  
*Bombus terrestris* -----STTSHPSDLWKEPTTP-VFSREIDFG-IGSEKDRVGQ---I  
*Apis florea* -----TTTSHPSDLWKEPTAS-VFSRDIDFG-IGSEKDRVGQ---I  
*Megachile rotundata* -----TTASHPSDPWKEPTTP-VFGKEVDFG-IGSEKDRVGQ---I  
*Acromyrmex echinatio* -----LTSHPSDLWKEPTTPV-VFSTEMDFGSIISSKDRVGQ---I  
*Tribolium castaneum* -----MLKKEDDIWKP-----AEVFPSTSG---L  
*Pediculus humanus* -----HNDLFDGDPWKNSSDLPPFSN-----LTERDQTSS---F  
*Drosophila melanogaster* -----AAGGGASSGWRLP-----PGLGSPYG---L  
*Mus musculus* TAYEGSSLSLSRSSSREHLASGSESDNWRDRNGIGPMGHSEFSAPIGSEPKRKQNKREHY  
*Nilaparvata lugens* -----SEVDGAAWGG-----ASQGN----QYGNISNILEAVS

910 920 930 940 950 960

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RPRC001613  
RPRC001612 PSVSNYLEAVSQAILSSITNLDLSDQ---DLYTILATVGLEKYSOLFKKHEIDTSTFRILT  
*Harpegnathos saltator* GMASNYMDHTPTSHLN---KITS-HRYNDLASMLTSLGLEKYIRLFTSHEVDMATFPSLT  
*Harpegnathos saltator X1* GMASNYMDHTPTSHLN---KITS-HRYNDLASMLTSLGLEKYIRLFTSHEVDMATFPSLT  
*Bombus terrestris* GLSSNYMEHTPTSHLN---KITS-HRYDDLISMLVSVGLEKYIRLFTSHEVDMATFPSLT  
*Apis florea* GINSNYLEHTPSAHLK---NITS-NRYDDLISMLTSLVGLEKYIRLFTSHEVDMATFPSLT  
*Megachile rotundata* GVTISNYVEHTPTSHLN---RITS-HRYDDLISMLTSLVGLEKYIRLFTSHEVDMATFPSLT  
*Acromyrmex echinatio* GMASNYMDHTPTSHLN---KITSSHRFNDLGSMLTSLGLEKYIRLFTSHEVDMATFPSLT  
*Tribolium castaneum* HNTISNILDHTPTHFIN---RIAN-TTWSDLPSLLSALGLDRYISLFIHFVDFLPTFTTSL  
*Pediculus humanus* MSNSNAIESVPSGMLK---RIQN---CDTDLPTLLASLGLEKYIOLFKSQEIIDMSSFLTFN  
*Drosophila melanogaster* SATTGLLDATPVNRRM---QLAK---HK---DIQTLTSLGLEHYIKIFVLNEIDLEVFITLT  
*Mus musculus* LSSSNYMDCISSLTGSNGCNLNSCFKGSDLPELFSKLGKGYTDVFQQQEIIDLQTFLLT  
*Nilaparvata lugens* SQHME---VITS-SESYTIPSLLASLNLEMYIPLFHMHHIDLPTFRITLTEODLKETGVSL

970 980 990 1000 1010 1020

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RPRC001613  
RPRC001612 ENDLKEIGISMVGP RRKIML-----CIAGSPVFTIYSAKISPWIVGIPMQVCRE  
*Harpegnathos saltator* EKDLCEIGITAWGARRKIMLLISEMNKRIVSPFSGSAA--PGAERKASTSSASTSSLGKC  
*Harpegnathos saltator X1* EKDLCEIGITAWGARRKIMLLISEMNKRIVSPFSGSAA--PGAERKASTSSASTSSLGKC  
*Bombus terrestris* EKDLCEIGITAWGARRKIMLLIAEMNKR-TSPFCGSAA--PGAERKTTIS-STSSMEKC  
*Apis florea* EKDLCEIGITAWGARRKIMLLIAEMNKR-TSPFCGSAA--PGAERKTTIS-TTSSMEKC  
*Megachile rotundata* EKDLCEIGISAWGARRKIMLLIAEMNKR-TSPFCGSAA--PGAERKISTAGIASAPMEKC  
*Acromyrmex echinatio* EKDLCEIGICAWGARRKIMLLISEMNKR-TSPFSGSAA--PGAERKASSSAPTSAIGKC  
*Tribolium castaneum* DKDLMTIGVTAFGSRRKMLLAISELNKR--SCTFSAA--PGAERKSSSSNSPPSHDNSP  
*Pediculus humanus* DVNLKEIGVLAAGAREKMLLAITELKKR--QSSLISAA--PGAERKRSHGWPKPEQ---  
*Drosophila melanogaster* EENLMEIGIAAGGARKLLTAHTLLANEAACSTMPSS--SSSQNSSSPRFSGSAAAPGAE  
*Mus musculus* DQDLKELGITTFGARRKMLLAISELSKNRRKIFEPNASCSTSFLEGGASGRLPRQYHSDI  
*Nilaparvata lugens* VGARRRILASIAELNRS--QSNLAPGL--WDYMERSSNSNASSNGLE-----

1030 1040 1050 1060 1070 1080

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RPRC001613  
RPRC001612  
*Harpegnathos saltator*  
*Harpegnathos saltator X1*  
*Bombus terrestris*  
*Apis florea*  
*Megachile rotundata*  
*Acromyrmex echinatio*  
*Tribolium castaneum*  
*Pediculus humanus*  
*Drosophila melanogaster*  
*Mus musculus*  
*Nilaparvata lugens*

**KVAVHHYQ**AMDYRIEIDKDISMYEWTEIQVALYFERMAEVVGKMNHLVICFSVELVWMLD

N-LDTKW  
N-LDTKW  
N-LDTKW  
N-LDTKW  
N-LDSRW  
N-LDTKW  
R---QSW  
  
RRPSNQW  
ASVSGRW

1090 1100 1110

.....|.....|.....|

RPRC001613  
RPRC001612  
*Harpegnathos saltator*  
*Harpegnathos saltator X1*  
*Bombus terrestris*  
*Apis florea*  
*Megachile rotundata*  
*Acromyrmex echinatio*  
*Tribolium castaneum*  
*Pediculus humanus*  
*Drosophila melanogaster*  
*Mus musculus*  
*Nilaparvata lugens*

YIWTQNITSSAKAKGKERFLIYSIE**YVNKRYVKYF**

# **Bibliografía**

- Aanes, H., Collas, P. and Alestrom, P., 2014. Transcriptome dynamics and diversity in the early zebrafish embryo. *Brief Funct Genomics*. 13, 95-105.
- Agarwal, A., Koppstein, D., Rozowsky, J., Sboner, A., Habegger, L., Hillier, L.W., Sasidharan, R., Reinke, V., Waterston, R.H. and Gerstein, M., 2010. Comparison and calibration of transcriptome data from RNA-Seq and tiling arrays. *BMC Genomics*. 11, 383.
- Ali-Murthy, Z. and Kornberg, T.B., 2016. Bicoid gradient formation and function in the *Drosophila* pre-syncytial blastoderm. *Elife*. 5.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990. Basic local alignment search tool. *J Mol Biol*. 215, 403-10.
- Anders, S. and Huber, W., 2010. Differential expression analysis for sequence count data. *Genome Biol*. 11, R106.
- Anderson, E., 1974. Comparative aspects of the ultrastructure of the female gamete. *Int Rev Cytol. Suppl* 4, 1-70.
- Araujo, R.N., Santosa, A., Pintoa, F.S., Gontijoa, N.F., Lehaneb, M.J. and Pereira, M.H., 2007. RNA Interference of the Salivary Gland Nitrophorin 2 in the Triatomine Bug *Rhodnius Prolixus* (Hemiptera : Reduviidae) by ARNdc Ingestion or Injection. *Insect Biochem*. 36(9), 683–693.
- Arnosti, D.N., Gray, S., Barolo, S., Zhou, J. and Levine, M., 1996. The gap protein knirps mediates both quenching and direct repression in the *Drosophila* embryo. *EMBO J*. 15, 3659-66.
- Atella, G.C., Gondim, K.C., Machado, E.A., Medeiros, M.N., Silva-Neto, M.A. and Masuda, H., 2005. Oogenesis and egg development in triatomines: a biochemical approach. *An Acad Bras Cienc*. 77, 405-30.
- Au, K.F., Jiang, H., Lin, L., Xing, Y. and Wong, W.H., 2010. Detection of splice junctions from paired-end RNA-seq data by SpliceMap. *Nucleic Acids Res*. 38, 4570-8.
- Beament, J.W., 1946. The formation and structure of the chorion of the egg in an hemipteran, *Rhodnius prolixus*. *Q J Microsc Sci*. 87, 393-439.
- Behura, S.K., Haugen, M., Flannery, E., Sarro, J., Tessier, C.R., Severson, D.W. and Duman-Scheel, M., 2011. Comparative genomic analysis of *Drosophila melanogaster* and vector mosquito developmental genes. *PLoS One*. 6, e21504.
- Bellés, X., 1998. . In: Coast, G.M., Webster, S.G. (Eds.), *Recent Advances in Arthropod Endocrinology*. Cambridge University Press, Cambridge, pp. 71–90, 1998. Endocrine effectors in insect vitellogenesis, in: Coast, G.M. and Webster, S.G. Eds.), *Recent Advances in Arthropod Endocrinology*. Cambridge, Cambridge University Press, pp. 71-90.
- Benoit, B., He, C.H., Zhang, F., Votruba, S.M., Tadros, W., Westwood, J.T., Smibert, C.A., Lipshitz, H.D. and Theurkauf, W.E., 2009. An essential role for the RNA-binding protein Smaug during the *Drosophila* maternal-to-zygotic transition. *Development*. 136, 923-32.
- Benton, M., 2018. A revised understanding of *Tribolium* morphogenesis further reconciles short and long germ development. *PLOS Biology*.
- Berenbaum, M.R., 1996. Bugs in the System, in: Wesley., N.Y.A. (Ed.).
- Berger, M.F., Levin, J.Z., Vijayendran, K., Sivachenko, A., Adiconis, X., Maguire, J., Johnson, L.A., Robinson, J., Verhaak, R.G., Sougnez, C., Onofrio, R.C., Ziaugra, L., Cibulskis, K., Laine, E., Barretina, J., Winckler, W., Fisher, D.E., Getz, G., Meyerson, M., Jaffe, D.B., Gabriel, S.B., Lander, E.S., Dummer, R., Gnirke, A., Nusbaum, C. and Garraway, L.A., 2010. Integrative analysis of the melanoma transcriptome. *Genome Res*. 20, 413-27.
- Berni, M., Ribeiro, J.M., Tobias-Santos, V., Caceres-Rodrigues, A., Borges Muty, F., Vionette-Amaral, R.J., Masuda, H., Sorgine, M.H., Nunes da Fonseca, R. and Araujo, H., 2014. Toll signals regulate dorsal–ventral patterning and anterior–posterior placement of the embryo in the hemipteran *Rhodnius prolixus*. *EvoDevo*. 5, 38.

- Blariza, M.J., Leyria, J., Canavoso, L.E., Soria, N.W. and Garcia, B.A., 2016. Dynamics of expression of two vitellogenin genes in the Chagas' disease vector *Triatoma infestans*: Analysis throughout pre-vitellogenesis and vitellogenesis. *Acta Trop.* 156, 100-7.
- Blekhman, R., Marioni, J.C., Zumbo, P., Stephens, M. and Gilad, Y., 2010. Sex-specific and lineage-specific alternative splicing in primates. *Genome Res.* 20, 180-9.
- Bökel, C., Dass, S., Wilsch-Bräuninger, M. and Roth, S., 2005. *Drosophila* Cornichon acts as cargo receptor for ER export of the TGF-like growth factor Gurken. *Development* 133, 459-470.
- Bolognesi, R., Farzana, L., Fischer, T.D. and Brown, S.J., 2008. Multiple Wnt genes are required for segmentation in the short-germ embryo of *Tribolium castaneum*. *Curr Biol.* 18, 1624-9.
- Bradford, J.R., Hey, Y., Yates, T., Li, Y., Pepper, S.D. and Miller, C.J., 2010. A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. *BMC Genomics.* 11, 282.
- Braz, G.R., Moreira, M.F., Masuda, H. and Oliveira, P.L., 2002. *Rhodnius* heme-binding protein (RHBP) is a heme source for embryonic development in the blood-sucking bug *Rhodnius prolixus* (Hemiptera, Reduviidae). *Insect Biochem Mol Biol.* 32, 361-7.
- Brent, A.E., Yucel, G., Small, S. and Desplan, C., 2007. Permissive and instructive anterior patterning rely on mRNA localization in the wasp embryo. *Science.* 315, 1841-3.
- Brody, T., 1999. The Interactive Fly: gene networks, development and the Internet. *Trends Genet.* 15, 333-4.
- Bucher, G., Farzana, L., Brown, S.J. and Klingler, M., 2005. Anterior localization of maternal mRNAs in a short germ insect lacking bicoid. *Evol Dev.* 7, 142-9.
- Bucher, G., Scholten, J. and Klingler, M., 2002. Parental RNAi in *Tribolium* (Coleoptera). *Curr Biol.* 12, R85-6.
- Buxton, P.A., 1930. The biology of a blood-sucking bug, *Rhodnius prolixus*. *Transactions of the Entomological Society London.* 78, 227-236.
- Clark, E., 2017. Dynamic patterning by the *Drosophila* pair-rule network reconciles long-germ and short-germ segmentation. *PLoS Biol.* 15, e2002439.
- Clark, E. and Peel, A., 2017. Evidence for the temporal regulation of insect segmentation by a conserved set of developmental transcription factors. *bioRxiv.*
- Claussen, M. and Suter, B., 2005. BicD-dependent localization processes: from *Drosophila* development to human cell biology. *Ann Anat.* 187, 539-53.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X. and Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17, 13.
- Copf, T., Schroder, R. and Averof, M., 2004. Ancestral role of caudal genes in axis elongation and segmentation. *Proc Natl Acad Sci U S A.* 101, 17711-5.
- Coura JR and B.-P.J., 2012. Chagas disease. What is known and what should be improved: a systemic review. *Rev Soc Bras Med Trop.* 45, 286-296.
- Coutelis, J.B., Petzoldt, A.G., Speder, P., Suzanne, M. and Noselli, S., 2008. Left-right asymmetry in *Drosophila*. *Semin Cell Dev Biol.* 19, 252-62.
- Cummings, M.R., 1972. Formation of the vitelline membrane and chorion in developing oocytes of *Ephestia kuhniella*. *Z Zellforsch Mikrosk Anat.* 127, 175-88.
- Cummings, M.R., Brown, N.M. and King, R.C., 1971. The cytology of the vitellogenic stages of oogenesis in *Drosophila melanogaster*. 3. Formation of the vitelline membrane. *Z Zellforsch Mikrosk Anat.* 118, 482-92.
- Chagas, C.R.J., 1909. Nova tripanosomíase humana. Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen. n. esp., agente da nova entidade mórbida do homem. . *Mem Inst Oswaldo Cruz.* 1, 159-218.
- Chen, G., Wang, C. and Shi, T., 2011. Overview of available methods for diverse RNA-Seq data analyses. *Sci China Life Sci.* 54, 1121-8.

- Damen, W.G., Janssen, R. and Prpic, N.M., 2005. Pair rule gene orthologs in spider segmentation. *Evol Dev.* 7, 618-28.
- Davey, K., 2007. The interaction of feeding and mating in the hormonal control of egg production in *Rhodnius prolixus*. *J Insect Physiol.* 53, 208-15.
- Davey, K.G., 1987. Inputs to the hormonal control of egg development in *Rhodnius prolixus*. *Mem Inst Oswaldo Cruz.* 82 Suppl 3, 103-8.
- Davey, K.G. and Huebner, E., 1974. The response of the follicle cells of *Rhodnius prolixus* to juvenile hormone and antigonadotropin in vitro. *Can J Zool.* 52, 1407-12.
- Davis, G.K. and Patel, N.H., 2002. Short, long, and beyond: molecular and embryological approaches to insect segmentation. *Annu Rev Entomol.* 47, 669-99.
- De Bona, F., Ossowski, S., Schneeberger, K. and Ratsch, G., 2008. Optimal spliced alignments of short sequence reads. *Bioinformatics.* 24, i174-80.
- de Cuevas, M., 2005. *Drosophila* Oogenesis, *ENCYCLOPEDIA OF LIFE SCIENCES*.
- Dearden, P.K., Wilson, M.J., Sablan, L., Osborne, P.W., Havler, M., McNaughton, E., Kimura, K., Milshina, N.V., Hasselmann, M., Gempe, T., Schioett, M., Brown, S.J., Elsik, C.G., Holland, P.W., Kadowaki, T. and Beye, M., 2006. Patterns of conservation and change in honey bee developmental genes. *Genome Res.* 16, 1376-84.
- Denoeud, F., Aury, J.M., Da Silva, C., Noel, B., Rogier, O., Delledonne, M., Morgante, M., Valle, G., Wincker, P., Scarpelli, C., Jaillon, O. and Artiguenave, F., 2008. Annotating genomes with massive-scale RNA sequencing. *Genome Biol.* 9, R175.
- Dias, J.C., Prata, A. and Schofield, C.J., 2002a. Chagas' disease in the Amazon: an overview of the current situation and perspectives for prevention. *Rev Soc Bras Med Trop.* 35, 669-678.
- Dias, J.C., Silveira, A.C. and Schofield, C.J., 2002b. The impact of Chagas disease control in Latin America: a review. *Mem Inst Oswaldo Cruz.* 97, 603-612.
- Driever, W. and Nusslein-Volhard, C., 1988a. The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell.* 54, 95-104.
- Driever, W. and Nusslein-Volhard, C., 1988b. A gradient of bicoid protein in *Drosophila* embryos. *Cell.* 54, 83-93.
- Edgar, B.A. and O'Farrell, P.H., 1990. The three postblastoderm cell cycles of *Drosophila* embryogenesis are regulated in G2 by string. *Cell.* 62, 469-80.
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. and Turner, S., 2009. Real-time DNA sequencing from single polymerase molecules. *Science.* 323, 133-8.
- Esponda-Behrens, N., 2014. Estudios funcionales de la evolución de la segmentación en insectos Modelo de estudio: *Rhodnius prolixus* (Stähl, 1859) (Hemiptera, Reduviidae), Facultad de Ciencias Naturales y Museo. Universidad Nacional de La Plata.
- Ewen-Campen, B., Shaner, N., Panfilio, K.A., Suzuki, Y., Roth, S. and Extavour, C.G., 2011. The maternal and early embryonic transcriptome of the milkweed bug *Oncopeltus fasciatus*. *BMC Genomics.* 12, 61.
- Farzana, L. and Brown, S.J., 2008. Hedgehog signaling pathway function conserved in *Tribolium* segmentation. *Dev Genes Evol.* 218, 181-92.
- Fedurco, M., Romieu, A., Williams, S., Lawrence, I. and Turcatti, G., 2006. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res.* 34, e22.

- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 391, 806-11.
- Foe, V.E., 1989. Mitotic domains reveal early commitment of cells in *Drosophila* embryos. *Development*. 107, 1-22.
- Frasch, M., Hoey, T., Rushlow, C., Doyle, H. and Levine, M., 1987. Characterization and localization of the even-skipped protein of *Drosophila*. *EMBO J.* 6, 749-59.
- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M., 1986. Structure of the segmentation gene paired and the *Drosophila* PRD gene set as part of a gene network. *Cell*. 47, 735-746.
- Garber, M., Grabherr, M.G., Guttman, M. and Trapnell, C., 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods*. 8, 469-77.
- Garcia-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F. and Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*. 28, 2678-9.
- Garcia, E.S. and Azambuja, P., 2004. Lignoids in insects: chemical probes for the study of ecdysis, excretion and *Trypanosoma cruzi*-triatomine interactions. *Toxicol.* 44, 431-440.
- Gaul, U. and Jackle, H., 1987. Pole region-dependent repression of the *Drosophila* gap gene *Kruppel* by maternal gene products. *Cell*. 51, 549-55.
- Gaul, U., Seifert, E., Schuh, R. and Jackle, H., 1987. Analysis of *Kruppel* protein distribution during early *Drosophila* development reveals posttranscriptional regulation. *Cell*. 50, 639-47.
- Giorgi, F., Snigirevskaya, E.S. and Raikel, A.S.I.R., A.S. (Ed.), *Reproductive Biology of Invertebrates. Progress in Vitellogenesis*, vol. XII. Part B. Science Publishers Inc., Enfield, USA, pp. 33-68., 2005. The cell biology of yolk protein precursor synthesis and secretion, in: Raikel, A.S. (Ed.), *Reproductive Biology of Invertebrates. Progress in vitellogenesis*. Science Publishers Inc., Enfield, USA, pp. 33-68.
- Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., Rinn, J.L., Lander, E.S. and Regev, A., 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol.* 28, 503-10.
- Haas, B.J. and Zody, M.C., 2010. Advancing RNA-Seq analysis. *Nat Biotechnol.* 28, 421-3.
- Hardison, R.C., 2003. Comparative genomics. *PLoS Biol.* 1, E58.
- Harker, B.W., Behura, S.K., deBruyn, B.S., Lovin, D.D., Mori, A., Romero-Severson, J. and Severson, D.W., 2013. Stage-specific transcription during development of *Aedes aegypti*. *BMC Dev Biol.* 13, 29.
- Harrison, R.E. and Huebner, E., 1997. Unipolar microtubule array is directly involved in nurse cell-oocyte transport. *Cell Motil Cytoskeleton*. 36, 355-62.
- Homer, N., Merriman, B. and Nelson, S.F., 2009. BFAST: an alignment tool for large scale genome resequencing. *PLoS One*. 4, e7767.
- Hopkins, C.R. and King, P.E., 1966. An electron-microscopic and histochemical study of the oocyte periphery in *Bombus terrestris* during vitellogenesis. *J. Cell Sei.* 1, 201-216.
- Huebner, E., 1981. Oocyte-follicle cell interaction during normal oogenesis and atresia in an insect. *J Ultrastruct Res.* 74, 95-104.
- Huebner, E. and Anderson, E., 1970. The effects of vinblastine sulfate on the microtubular organization of the ovary of *Rhodnius prolixus*. *J Cell Biol.* 46, 191-8.
- Huebner, E. and Anderson, E., 1972a. A cytological study of the ovary of *Rhodnius prolixus*. Cytoarchitecture and development of the trophic chamber. *J Morphol.* 138, 1-40.
- Huebner, E. and Anderson, E., 1972b. A cytological study of the ovary of *Rhodnius prolixus*. I. The ontogeny of the follicular epithelium. *J Morphol.* 136, 459-93.
- Huebner, E. and Anderson, E., 1972c. A cytological study of the ovary of *Rhodnius prolixus*. II. Oocyte differentiation. *J Morphol.* 137, 385-415.

- Hülskamp, M. and Tautz, D., 1991. Gap genes and gradients – the logic behind the gaps. *BioEssays*. 13, 261-268.
- Hunter, C.P., 1999. Genetics: a touch of elegance with RNAi. *Curr Biol*. 9, R440-442.
- Huynh, J.R. and St Johnston, D., 2000. The role of BicD, Egl, Orb and the microtubules in the restriction of meiosis to the *Drosophila* oocyte. *Development*. 127, 2785-94.
- Jacquier, A., 2009. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet*. 10, 833-44.
- Janssen, R. and Budd, G.E., 2013. Deciphering the onychophoran 'segmentation gene cascade': Gene expression reveals limited involvement of pair rule gene orthologs in segmentation, but a highly conserved segment polarity gene network. *Dev Biol*. 382, 224-34.
- Jiang, Y., Gao, X., Liu, S., Zhang, Y., Liu, H., Sun, F., Bao, L., Waldbieser, G. and Liu, Z., 2013. Whole genome comparative analysis of channel catfish (*Ictalurus punctatus*) with four model fish species. *BMC Genomics*. 14, 780.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R. and Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 30, 1236-40.
- Kelly, G.M. and Huebner, E., 1989. Embryonic development of the hemipteran insect *Rhodnius prolixus*. *J Morphol*. 199, 175-196.
- Kent, W.J., 2002. BLAT--the BLAST-like alignment tool. *Genome Res*. 12, 656-64.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 14, R36.
- Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D. and Ren, B., 2005. A high-resolution map of active promoters in the human genome. *Nature*. 436, 876-80.
- King, R.C. and Koch, E.A., 1963. Studies on ovarian follicle cells of *Drosophila*. . *Quart. J. micr. Sci.* . 104, 297-320
- Kollien, A.H. and Schaub, G.A., 2000. The development of *Trypanosoma cruzi* in triatominae. . *Parasitol Today*. 16, 381-387.
- Konrad, K.D., Wang, D. and Marsh, J.L., 1993. Vitelline membrane biogenesis in *Drosophila* requires the activity of the alpha-methyl dopa hypersensitive gene (*l(2)amd*) in both the germline and follicle cells. *Insect Mol Biol*. 1, 179-87.
- Krause, G., 1939. Die Eitypen der Insekten. *Biol Zentralbl*. 59.
- Kunkel, J.G. and Nordin, J.H., 1985. Yolk protein, in: Kerkut, G.A. and Gilbert, L.I. Eds.), *Comprehensive Insect Physiology, Biochemistry and Pharmacology*. . PergamonPress, Oxford, pp. 83-111.
- Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10, R25.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23, 2947-8.
- Lavore, A., Pagola, L., Esponda-Behrens, N. and Rivera-Pomar, R., 2012. The gap gene giant of *Rhodnius prolixus* is maternally expressed and required for proper head and abdomen formation. *Dev Biol*. 361, 147-55.
- Lavore, A., Pascual, A., Salinas, F.M., Esponda-Behrens, N., Martinez-Barnette, J., Rodriguez, M. and Rivera-Pomar, R., 2015. Comparative analysis of zygotic developmental genes in *Rhodnius prolixus* genome shows conserved features on the tracheal developmental pathway. *Insect Biochem Mol Biol*. 64, 32-43.

- Lawrence, P.A., 1992. *The Making of a Fly: The Genetics of Animal Design*, Blackwell Scientific Publications, OXFORD.
- Levine, M. and Harding, K.W., 1989. *Drosophila: The zygotic contribution*. Genes and Embryos. IRL, 39-94.
- Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25, 1754-60.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25, 2078-9.
- Li, H., Ruan, J. and Durbin, R., 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 18, 1851-8.
- Liu, H., Wang, T., Wang, J., Quan, F. and Zhang, Y., 2013. Characterization of Liaoning cashmere goat transcriptome: sequencing, de novo assembly, functional annotation and comparative analysis. *PLoS One*. 8, e77062.
- Liu, P.Z. and Kaufman, T.C., 2005. Short and long germ segmentation: unanswered questions in the evolution of a developmental mode. *Evol Dev*. 7, 629-46.
- Lu, X., Li, J.M., Elemento, O., Tavazoie, S. and Wieschaus, E.F., 2009. Coupling of zygotic transcription to mitotic control at the *Drosophila* midblastula transition. *Development* 136, 2101-2110.
- Lunter, G. and Goodson, M., 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res*. 21, 936-9.
- Lutz, D.A. and Huebner, E., 1980. Development and cellular differentiation of an insect telotrophic ovary (*Rhodnius prolixus*). *Tissue Cell*. 12, 773-94.
- Lutz, D.A. and Huebner, E., 1981. Development of nurse cell-oocyte interactions in the insect telotrophic ovary (*Rhodnius prolixus*). *Tissue Cell*. 13, 321-35.
- Lynch, J.A., Brent, A.E., Leaf, D.S., Pultz, M.A. and Desplan, C., 2006. Localized maternal orthodenticle patterns anterior and posterior in the long germ wasp *Nasonia*. *Nature*. 439, 728-32.
- Lynch, J.A., Ozuak, O., Khila, A., Abouheif, E., Desplan, C. and Roth, S., 2011. The phylogenetic origin of oskar coincided with the origin of maternally provisioned germ plasm and pole cells at the base of the Holometabola. *PLoS Genet*. 7, e1002029.
- Lynch, J.A., Peel, A.D., Drechsler, A., Averof, M. and Roth, S., 2010. EGF signaling and the origin of axial polarity among the insects. *Curr Biol*. 20, 1042-7.
- Lynch, J.A. and Roth, S., 2011. The evolution of dorsal-ventral patterning mechanisms in insects. *Genes Dev*. 25, 107-18.
- Macdonald, P.M., Ingham, P. and Struhl, G., 1986. Isolation, structure, and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell*. 47, 721-34.
- Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N. and Chinnaiyan, A.M., 2009. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 458, 97-101.
- Mahone, M., Saffman, E.E. and Lasko, P.F., 1995. Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. *EMBO J*. 14, 2043-55.
- Margaritis, L.H., Kafatos, F.C. and Petri, W.H., 1980. The eggshell of *Drosophila melanogaster*. I. Fine structure of the layers and regions of the wild-type eggshell. *J Cell Sci*. 43, 1-35.
- Margulies, E.H., Program, N.C.S., Maduro, V.V., Thomas, P.J., Tomkins, J.P., Amemiya, C.T., Luo, M. and Green, E.D., 2005. Comparative sequencing provides insights about the structure and conservation of marsupial and monotreme genomes. *Proc Natl Acad Sci U S A*. 102, 3354-9.
- Martínez-Arias, A. and Lawrence, P.A., 1985. Parasegments and compartments in the *Drosophila* embryo. *Nature*. 639-642.
- Martinez, J. and Tuschl, T., 2004. RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev*. 18, 975-980.

- McClure, R., Balasubramanian, D., Sun, Y., Bobrovskyy, M., Sumbly, P., Genco, C.A., Vanderpool, C.K. and Tjaden, B., 2013. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res.* 41, e140.
- McGregor, A.P., Pechmann, M., Schwager, E.E., Feitosa, N.M., Kruck, S., Aranda, M. and Damen, W.G., 2008. Wnt8 is required for growth-zone establishment and development of opisthosomal segments in a spider. *Curr Biol.* 18, 1619-23.
- McIntyre, L.M., Lopiano, K.K., Morse, A.M., Amin, V., Oberg, A.L., Young, L.J. and Nuzhdin, S.V., 2011. RNA-seq: technical variability and sampling. *BMC Genomics.* 12, 293.
- McLaughlin, J.M. and Bratu, D.P., 2015. *Drosophila melanogaster* Oogenesis: An Overview. *Methods Mol Biol.* 1328, 1-20.
- Medeiros, M.N., Logullo, R., Ramos, I.B., Sorgine, M.H., Paiva-Silva, G.O., Mesquita, R.D., Machado, E.A., Coutinho, M.A., Masuda, H., Capurro, M.L., Ribeiro, J.M., Cardoso Braz, G.R. and Oliveira, P.L., 2011. Transcriptome and gene expression profile of ovarian follicle tissue of the triatomine bug *Rhodnius prolixus*. *Insect Biochem Mol Biol.* 41, 823-31.
- Melo, A.C., Valle, D., Machado, E.A., Salerno, A.P., Paiva-Silva, G.O., Cunha, E.S.N.L., de Souza, W. and Masuda, H., 2000. Synthesis of vitellogenin by the follicle cells of *Rhodnius prolixus*. *Insect Biochem Mol Biol.* 30, 549-57.
- Mellanby, H., 1935. The early embryonic development of *Rhodnius prolixus* (Hemiptera, Heteroptera). *Quarterly Journal of Microscopic Science.*
- Mellanby, H., 1936. The latter embryology of *Rhodnius prolixus*. *Quarterly Journal of Microscopic Science.* 1-40.
- Mesquita, R.D., Vionette-Amaral, R.J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F.A., Minx, P., Spieth, J., Carvalho, A.B., Panzera, F., Lawson, D., Torres, A.Q., Ribeiro, J.M., Sorgine, M.H., Waterhouse, R.M., Montague, M.J., Abad-Franch, F., Alves-Bezerra, M., Amaral, L.R., Araujo, H.M., Araujo, R.N., Aravind, L., Atella, G.C., Azambuja, P., Berni, M., Bittencourt-Cunha, P.R., Braz, G.R., Calderon-Fernandez, G., Carareto, C.M., Christensen, M.B., Costa, I.R., Costa, S.G., Dansa, M., Dumas-Filho, C.R., De-Paula, I.F., Dias, F.A., Dimopoulos, G., Emrich, S.J., Esponda-Behrens, N., Fampa, P., Fernandez-Medina, R.D., da Fonseca, R.N., Fontenele, M., Fronick, C., Fulton, L.A., Gandara, A.C., Garcia, E.S., Genta, F.A., Giraldo-Calderon, G.I., Gomes, B., Gondim, K.C., Granzotto, A., Guarneri, A.A., Guigo, R., Harry, M., Hughes, D.S., Jablonka, W., Jacquín-Joly, E., Juárez, M.P., Koerich, L.B., Lange, A.B., Latorre-Estivalis, J.M., Lavore, A., Lawrence, G.G., Lazoski, C., Lazzari, C.R., Lopes, R.R., Lorenzo, M.G., Lugon, M.D., Majerowicz, D., Marcet, P.L., Mariotti, M., Masuda, H., Megy, K., Melo, A.C., Missirlis, F., Mota, T., Noriega, F.G., Nouzova, M., Nunes, R.D., Oliveira, R.L., Oliveira-Silveira, G., Ons, S., Orchard, I., Pagola, L., Paiva-Silva, G.O., Pascual, A., Pavan, M.G., Pedrini, N., Peixoto, A.A., Pereira, M.H., Pike, A., Polycarpo, C., Prosdocimi, F., Ribeiro-Rodrigues, R., Robertson, H.M., Salerno, A.P., Salmon, D., Santesmasses, D., Schama, R., Seabra-Junior, E.S. et al., 2015. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc Natl Acad Sci U S A.* 112, 14936-41.
- Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nat Rev Genet.* 11, 31-46.
- Mohler, J., Eldon, E.D. and Pirrotta, V., 1989. A novel spatial transcription pattern associated with the segmentation gene, *giant*, of *Drosophila*. *EMBO J.* 8, 1539-48.
- Mortazavi, A., Schwarz, E.M., Williams, B., Schaeffer, L., Antoshechkin, I., Wold, B.J. and Sternberg, P.W., 2010. Scaffolding a *Caenorhabditis* nematode genome with RNA-seq. *Genome Res.* 20, 1740-7.
- Mullis, K.B., 1990. Target amplification for DNA analysis by the polymerase chain reaction. *Ann Biol Clin (Paris).* 48, 579-82.
- Mundall, E. and Engelmann, F., 1977. Endocrine control of vitellogenin synthesis and vitellogenesis in *Triatoma protracta*. *J Insect Physiol.* 23, 825-36.

- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M., 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*. 320, 1344-9.
- Noriega, F.G., 1992. Autogeny in three species of Triatominae: *Rhodnius prolixus*, *Triatoma rubrovaria*, and *Triatoma infestans* (Hemiptera: Reduviidae). *J Med Entomol*. 29, 273-7.
- Nunes-da-Fonseca, R., Berni, M., Tobias-Santos, V., Pane, A. and Araujo, H.M., 2017. *Rhodnius prolixus*: From classical physiology to modern developmental biology. *Genesis*. 55.
- Nunes da Fonseca, R., Lynch, J. and Roth, S., 2009. Evolution of axis formation: mRNA localization, regulatory circuits and posterior specification in non-model arthropods. *Current Opinion in Genetics & Development* 19, 404-411.
- Nusslein-Volhard, C., Frohnhofer, H.G. and Lehmann, R., 1987. Determination of anteroposterior polarity in *Drosophila*. *Science*. 238, 1675-81.
- Nusslein-Volhard, C., Lohs-Schardin, M., Sander, K. and Cremer, C., 1980. A dorso-ventral shift of embryonic primordia in a new maternal-effect mutant of *Drosophila*. *Nature*. 283, 474-6.
- Nusslein-Volhard, C. and Roth, S., 1989. Axis determination in insect embryos. *Ciba Found Symp*. 144, 37-55; discussion 55-64, 92-8.
- Nusslein-Volhard, C. and Wieschaus, E., 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature*. 287, 795-801.
- Nusslein-Volhard, C., Wieschaus, E. and Kluding, H., 1984. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* : I. Zygotic loci on the second chromosome. *Wilehm Roux Arch Dev Biol*. 193, 267-282.
- Olesnick, E.C., Brent, A.E., Tonnes, L., Walker, M., Pultz, M.A., Leaf, D. and Desplan, C., 2006. A caudal mRNA gradient controls posterior development in the wasp *Nasonia*. *Development*. 133, 3973-82.
- Oshlack, A., Robinson, M.D. and Young, M.D., 2010. From RNA-seq reads to differential expression results. *Genome Biol*. 11, 220.
- Ozsolak, F. and Milos, P.M., 2011. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet*. 12, 87-98.
- Pagola, L., 2012. Análisis genómico y molecular de la embriogénesis de *Rhodnius prolixus* (Stähl, 1859) (Hemíptera, Reduviidae): implicancias morfológico-evolutivas en insectos
- Facultad de Ciencias Naturales y Museo. Universidad Nacional de La Plata.
- Pankratz, M.J., Hoch, M., Seifert, E. and Jackle, H., 1989. Kruppel requirement for knirps enhancement reflects overlapping gap gene activities in the *Drosophila* embryo. *Nature*. 341, 337-40.
- Papantonis, A., Swevers, L. and Iatrou, K., 2015. Chorion genes: a landscape of their evolution, structure, and regulation. *Annu Rev Entomol*. 60, 177-94.
- Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D. and Yeates, T.O., 1999. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A*. 96, 4285-8.
- Raikhel, A.S., 2005. Vitellogenesis of disease vectors, from physiology to genes, in: Marquardt, W. (Ed.), *Biology of Disease Vectors*. Elsevier Academic Press, London, UK.
- Ribeiro, J.M., Genta, F.A., Sorgine, M.H., Logullo, R., Mesquita, R.D., Paiva-Silva, G.O., Majerowicz, D., Medeiros, M., Koerich, L., Terra, W.R., Ferreira, C., Pimentel, A.C., Bisch, P.M., Leite, D.C., Diniz, M.M., da, S.G.V.J.J.L., Da Silva, M.L., Araujo, R.N., Gandara, A.C., Brosson, S., Salmon, D., Bousbata, S., Gonzalez-Caballero, N., Silber, A.M., Alves-Bezerra, M., Gondim, K.C., Silva-Neto, M.A., Atella, G.C., Araujo, H., Dias, F.A., Polycarpo, C., Vionette-Amaral, R.J., Fampa, P., Melo, A.C., Tanaka, A.S., Balczun, C., Oliveira, J.H., Goncalves, R.L., Lazoski, C., Rivera-Pomar, R., Diambra, L., Schaub, G.A., Garcia, E.S., Azambuja, P., Braz, G.R. and Oliveira, P.L., 2014. An

- insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *PLoS Negl Trop Dis.* 8, e2594.
- Rivera-Pomar, R. and Jackle, H., 1996. From gradients to stripes in *Drosophila* embryogenesis: filling in the gaps. *Trends Genet.* 12, 478-83.
- Roth, S., 2003. The origin of dorsoventral polarity in *Drosophila*. *Philos Trans R Soc Lond B Biol Sci.* 358, 1317-29; discussion 1329.
- Roth, S., Neuman-Silberberg, F.S., Barcelo, G. and Schupbach, T., 1995. *cornichon* and the EGF receptor signaling process are necessary for both anterior-posterior and dorsal-ventral pattern formation in *Drosophila*. *Cell.* 81, 967-78.
- Rozen, S. and Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.* 132, 365-86.
- Sander, K., 1981. Pattern generation and pattern conservation in insect ontogenesis—problems, data, and models. *Progress in Developmental Biology.* 101-19.
- Sanson, B., 2001. Generating patterns from fields of cells. Examples from *Drosophila* segmentation. *EMBO Rep.* 2, 1083-8.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 9, 676-82.
- Schupbach, T. and Roth, S., 1994. Dorsoventral patterning in *Drosophila* oogenesis. *Curr Opin Genet Dev.* 4, 502-7.
- Shendure, J. and Ji, H., 2008. Next-generation DNA sequencing. *Nat Biotechnol.* 26, 1135-45.
- Shigenobu, S., Bickel, R.D., Brisson, J.A., Butts, T., Chang, C.C., Christiaens, O., Davis, G.K., Duncan, E.J., Ferrier, D.E., Iga, M., Janssen, R., Lin, G.W., Lu, H.L., McGregor, A.P., Miura, T., Smagghe, G., Smith, J.M., van der Zee, M., Velarde, R.A., Wilson, M.J., Dearden, P.K. and Stern, D.L., 2010. Comprehensive survey of developmental genes in the pea aphid, *Acyrtosiphon pisum*: frequent lineage-specific duplications and losses of developmental genes. *Insect Mol Biol.* 19 Suppl 2, 47-62.
- Small, S., Blair, A. and Levine, M., 1992. Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *EMBO J.* 11, 4047-57.
- Smith, A.D., Xuan, Z. and Zhang, M.Q., 2008. Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics.* 9, 128.
- Soneson, C. and Delorenzi, M., 2013. A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics.* 14, 91.
- Sorrivas de Lozano, V., Morales, A. and Yañez, M.J., 2014. Principios y práctica de la Microscopía Electrónica, in: Sorrivas de Lozano, V. (Ed.). UAT- CONICET, Bahía Blanca.
- Speder, P., Adam, G. and Noselli, S., 2006. Type ID unconventional myosin controls left-right asymmetry in *Drosophila*. *Nature.* 440, 803-7.
- Spradling, A.C., 1993. Germline cysts: communes that work. *Cell.* 72, 649-51.
- St Johnston, D. and Nusslein-Volhard, C., 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell.* 68, 201-19.
- Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z. and Brow, D.A., 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell.* 24, 735-746.
- Stoka, A.M., Salomón, O.D. and Noriega, F.G., 1978. Physiology of Triatominae's reproduction, in: Brenner, R.R. and Stoka, A.M. Eds.), *Chagas' Disease Vectors*. CRC Press, Florida, pp. 109-129.
- Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques.* 28, 1102, 1104.
- t Hoen, P.A., Ariyurek, Y., Thygesen, H.H., Vreugdenhil, E., Vossen, R.H., de Menezes, R.X., Boer, J.M., van Ommen, G.J. and den Dunnen, J.T., 2008. Deep sequencing-based expression analysis

- shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res.* 36, e141.
- Tatusov, R.L., Koonin, E.V. and Lipman, D.J., 1997. A genomic perspective on protein families. *Science.* 278, 631-7.
- Tautz, D., Friedrich, M. and Schröder, R., 1994. Insect embryogenesis - what is ancestral and what is derived? *Development Supplement.* 193-199.
- Telfer, W.H., 1975. Development and physiology of the oocyte-nurse cells syncytium. *Advances in Insect Physiology.* 11, 223-319.
- Tisseur, M., Kwapisz, M. and Morillon, A., 2011. Pervasive transcription - Lessons from yeast. *Biochimie.* 93, 1889-96.
- Trapnell, C., Pachter, L. and Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 25, 1105-11.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 7, 562-78.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 28, 511-5.
- Tucker, T., Marra, M. and Friedman, J.M., 2009. Massively parallel sequencing: the next big thing in genetic medicine. *Am J Hum Genet.* 85, 142-54.
- Tufail, M., Nagaba, Y., Elgendy, A.M. and Takeda, M., 2014. Regulation of vitellogenin genes in insects. *Entomological Science*
- Tufail, M. and Takeda, M., 2008. Molecular characteristics of insect vitellogenins. *J Insect Physiol.* 54, 1447-58.
- Valdimarsson, G. and Huebner, E., 1989. The development of microtubular arrays in the germ tissue of an insect telotrophic ovary. *Tissue Cell.* 21, 123-38.
- Valle, D., Lima Gomes, J., Goldenberg, S. and Garcia, E., 1987. *Rhodnius prolixus* vitellogenesis: Dependence upon a blood source. *Journal of Insect Physiology.* 33, 249-254.
- van Eeden, F. and St Johnston, D., 1999. The polarisation of the anterior–posterior and dorsal–ventral axes during *Drosophila* oogenesis. *Curr. Opin. Genet. Dev.* 9, 396–404.
- Van Verk, M.C., Hickman, R., Pieterse, C.M. and Van Wees, S.C., 2013. RNA-Seq: revelation of the messengers. *Trends Plant Sci.* 18, 175-9.
- Vazquez-Pianzola, P., Schaller, B., Colombo, M., Beuchle, D., Neuenschwander, S., Marcil, A., Bruggman, R. and Suter, B., 2017. The mRNA transportome of the BicD/Egl transport machinery. *RNA BIOLOGY.* 14, 73-89.
- Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., Hieter, P., Vogelstein, B. and Kinzler, K.W., 1997. Characterization of the yeast transcriptome. *Cell.* 88, 243-51.
- Walser, C.B. and Lipshitz, H.D., 2011. Transcript clearance during the maternal-to-zygotic transition. *Curr Opin Genet Dev.* 21, 431-43.
- Wanderley-Teixeira, V., Teixeira, A.A., Cunha, F.M., Costa, M.K., Veiga, A.F. and Oliveira, J.V., 2006. Histological description of the midgut and the pyloric valve of *Tropidacris collaris* (Stoll, 1813) (Orthoptera: Romaleidae). *Braz J Biol.* 66, 1045-9.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature.* 456, 470-6.
- Wang, K., Singh, D., Zeng, Z., Coleman, S.J., Huang, Y., Savich, G.L., He, X., Mieczkowski, P., Grimm, S.A., Perou, C.M., MacLeod, J.N., Chiang, D.Y., Prins, J.F. and Liu, J., 2010. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* 38, e178.

- Wang, Y., Ghaffari, N., Johnson, C.D., Braga-Neto, U.M., Wang, H., Chen, R. and Zhou, H., 2011. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics*. 12 Suppl 10, S5.
- Wang, Z., Gerstein, M. and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 10, 57-63.
- Wanninger, A., 2015. Evolutionary Developmental Biology of Invertebrates 5, in: Wanninger, A. (Ed.), *Ecdysozoa III: Hexapoda*. Springer-Verlag Wien, pp. VII, 219.
- Wesolowski, S., Birtwistle, M.R. and Rempala, G.A., 2013. A Comparison of Methods for RNA-Seq Differential Expression Analysis and a New Empirical Bayes Approach. *Biosensors (Basel)*. 3, 238-58.
- Wigglesworth, V.B., 1934. The physiology of ecdysis in *Rhodnius prolixus* (Hemiptera). II. Factors controlling moulting and metamorphosis. *Quarterly Journal of Microscopic Science*. 77, 191-222.
- Wigglesworth, V.B., 1936. The function of corpus allatum in the growth and reproduction of *R. prolixus* (Hemiptera). *Quarterly Journal of Microscopic Science*. 79, 91-121.
- Wigglesworth, V.B., 1943. The fate of the hemoglobin in *R. prolixus* and other blood-sucking arthropods. *Proceedings of the Royal Society of London Series B*. 131, 313-339.
- Wilhelm, B.T., Briau, M., Austin, P., Faubert, A., Boucher, G., Chagnon, P., Hope, K., Girard, S., Mayotte, N., Landry, J.-R., Hébert, J. and Sauvageau, G., 2011. RNA-seq analysis of 2 closely related leukemia clones that differ in their self-renewal capacity. *Blood*. 117, e27-e38.
- Wu, T.D. and Nacu, S., 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*. 26, 873-81.
- Yassour, M., Kaplan, T., Fraser, H.B., Levin, J.Z., Pfiffner, J., Adiconis, X., Schroth, G., Luo, S., Khrebtukova, I., Gnirke, A., Nusbaum, C., Thompson, D.A., Friedman, N. and Regev, A., 2009. Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing. *Proc Natl Acad Sci U S A*. 106, 3264-9.
- Zhang, B.X., Huang, H.J., Yu, B., Lou, Y.H., Fan, H.W. and Zhang, C.X., 2015. Bicaudal-C plays a vital role in oogenesis in *Nilaparvata lugens* (Hemiptera: Delphacidae). *J Insect Physiol*. 79, 19-26.