



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
DOCTORADO EN CIENCIAS BIOMÉDICAS  
INSTITUTO DE BIOTECNOLOGÍA

**Transcriptómica comparativa de la embriogénesis en *Anastrepha ludens***

TESIS  
QUE PARA OPTAR POR EL GRADO DE  
DOCTOR EN CIENCIAS

PRESENTA  
XIMENA GUTIÉRREZ RAMOS

DIRECTOR DE TESIS  
DR. MARIO ENRIQUE ZURITA ORTEGA  
INSTITUTO DE BIOTECNOLOGÍA  
COMITÉ TUTOR  
DRA. ROSA ESTELA NAVARRO GONZALEZ  
INSTITUTO DE FISIOLÓGÍA CELULAR  
DR. FÉLIX RECILLAS TARGA  
INSTITUTO DE FISIOLÓGÍA CELULAR

CUERNAVACA, MORELOS. ENERO DE 2021



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hay tres temas: el amor, la muerte y las moscas. Desde que el hombre existe, ese sentimiento, ese temor, esas presencias lo han acompañado siempre. Traten otros los dos primeros. Yo me ocupo de las moscas...

*Las moscas, Augusto Monterroso*

Este trabajo se presenta como tesis doctoral dentro del Programa de Doctorado en Ciencias Biomédicas de la Universidad Nacional Autónoma de México, el cual se llevó a cabo en el Laboratorio de Transcripción, Epigenética y Desarrollo del Instituto de Biotecnología bajo la dirección del Dr. Mario Zurita Ortega en colaboración con el Dr. Mario Arteaga Vázquez del Instituto de Biotecnología y Ecología Aplicada de la Universidad Veracruzana. Y con apoyo del Consejo Nacional de Ciencia y Tecnología.

## **Agradecimientos**

Este proyecto empezó con la propuesta por parte del Dr. Mario Arteaga-Vázquez, en donde fue fundamental el apoyo del Dr. Francisco Díaz Fleischer y mis compañeros de laboratorio Luis Manzanares, Dulce Flores, Cruz Perea y Evaristo Calihua en el Instituto de Biotecnología y Ecología Aplicada de la Universidad Veracruzana.

Posteriormente la tutoría pasó a manos del Dr. Mario Zurita, con apoyo y asesoría de la Dra. Martha Vázquez, los cuales fueron un dúo fundamental para mi crecimiento académico y personal durante mi doctorado. Así como las enseñanzas y ayuda de mis compañeros de laboratorio Alyeri, Samantha, Maritere, Eduardo y Adriana, ya que su compañía hizo más agradable el trabajo de laboratorio. Y sin olvidar a nuestra laborista, Carmen Muñoz por su ayuda con la preparación de comida para las moscas; así como a Jerome Jean Verleyen por su ayuda en el uso del clúster de la Unidad de Bioinformática del Instituto de Biotecnología.

A mi comité tutorial, Dra. Rosa Navarro González y Dr. Félix Recillas Targa, que durante mis estudios doctorales me han brindado una guía para desarrollar mi proyecto.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca de doctorado otorgada a Ximena Gutiérrez Ramos no. 448454; así como el financiamiento por parte de PAPPIT/UNAM IN200315 y CONACYT 25088 a Mario Zurita, y Fundación Agropolis 1502-307 y Royal Society Newton Advanced Fellowship NA150181 a Mario Arteaga Vázquez.

A mi comité sinodal, Dra. Rosa María Gutiérrez Ríos, Dra. Ingrid Fetter Pruneda, Dr. Diego Cortez Quezada y Dr. Eria Rebolgar Caudillo, que realizaron comentarios y sugerencias que enriquecieron mi tesis.

A mi familia que estuvo a mi lado.

# Índice

**Índice de figuras ... 9**

**Índice de tablas ... 11**

**Resumen ... 12**

**Abstract ... 13**

**Capítulo 1. Introducción ... 14**

1.1 *Anastrepha ludens*, mosca mexicana de la fruta ... 14

1.2 Importancia del estudio de *A. ludens* ... 17

**Capítulo 2. Antecedentes ... 19**

2.1. Impacto del estudio de la embriogénesis en los dípteros ... 19

2.2 Embriogénesis en *Drosophila melanogaster* ... 20

2.2.1 Estudios morfológicos de la embriogénesis de *D. melanogaster* ... 20

2.2.2 Mecanismos de regulación del eje antero – posterior del embrión de *D. melanogaster* ... 21

2.3 Embriogénesis en los tefrítidos ... 24

2.3.1 Embriogénesis en el género *Anastrepha* ... 25

2.4 Embriogénesis de *A. ludens* ... 25

**Capítulo 3. Objetivos ... 30**

3.1 Objetivo general ... 30

3.2 Objetivos particulares ... 30

**Capítulo 4. Metodología ... 31**

4.1 Cultivo de moscas *A. ludens* ... 31

4.2 Colecta y crecimiento de embriones de *A. ludens* ... 31

4.3 Extracción de RNA total de embriones de <i>A. ludens</i> ...	32
4.4 Construcción de librerías y secuenciación de muestras de RNA de embriones de <i>A. ludens</i> ...	33
4.5 Análisis bioinformáticos de RNA-seq de embriones de <i>A. ludens</i> ...	33
4.5.1 Análisis de calidad de secuencias crudas y ensamble <i>de novo</i> ...	33
4.5.2 Anotación de ensamblajes <i>de novo</i> ...	34
4.5.3 Análisis de expresión diferencial ...	35
4.5.4 Comparación a lo largo de varias especies de moscas ...	35
4.5.5 Anotación de proteínas de <i>A. ludens</i> ...	36
4.6 Hibridación <i>in situ</i> de embriones de <i>A. ludens</i> ...	37
4.6.1. Generación de sonda de RNA ...	37
4.6.2. Comprobación de generación y marcaje de sonda de RNA ...	37
4.6.2.1 Transferencia de sonda a partir de gel de agarosa ...	37
4.6.2.2 Dot blot ...	38
4.6.3. Hibridación <i>in situ</i> de embriones de <i>A. ludens</i> ...	38
4.7. Toma y análisis de imágenes ...	40
<b>Capítulo 5. Resultados</b> ...	<b>41</b>
5.1. Selección y generación de datos transcriptómicos de <i>A. ludens</i> ...	41
5.1.1. Selección de etapas de embriogénesis de <i>A. ludens</i> ...	41
5.1.2. Estandarización del protocolo para la extracción de RNA total de embriones de <i>A. ludens</i> ...	43
5.2 Generación y análisis de los datos transcriptómicos de <i>A. ludens</i> ...	45
5.2.1 Generación de librerías y secuencias de muestras de <i>A. ludens</i> ...	45
5.2.2 Análisis de calidad de secuencias ...	47

5.2.3	Ensamble <i>de novo</i> de los datos transcriptómicos ...	49
5.2.4.	Anotación de transcriptomas ...	52
5.3	Dinámicas de expresión entre las etapas de desarrollo embrionario temprano de <i>A. ludens</i> ...	55
5.3.1	Análisis de expresión diferencial entre las etapas de desarrollo embrionario de <i>A. ludens</i> ...	57
5.3.2	Identificación de transcritos importantes en la determinación sexual en <i>A. ludens</i> ...	58
5.4	Comparación de expresión transcripcional en la embriogénesis temprana entre diferentes moscas modelo y <i>A. ludens</i> ...	60
5.5	Expresión de genes con un posible papel en el desarrollo embrionario temprano de <i>A. ludens</i> ...	63
5.5.1	Genes que participan en el desarrollo temprano de <i>D. melanogaster</i> identificados en <i>A. ludens</i> ...	63
5.5.2	Comparación de genes que participan en el desarrollo temprano de <i>D. melanogaster</i> identificados en <i>A. ludens</i> , <i>B. dorsalis</i> y <i>C. capitata</i> ...	64
5.6	Patrones de expresión de genes del desarrollo embrionario temprano de <i>A. ludens</i> ...	66
5.6.1	Gen <i>exuperantia</i> de <i>A. ludens</i> ...	66
5.6.2	Gen <i>hunchback</i> de <i>A. ludens</i> ...	67
5.6.3	Gen <i>even skipped</i> de <i>A. ludens</i> ...	68
5.6.4	Gen <i>engrailed</i> de <i>A. ludens</i> ...	69
5.6.5	Comparación de patrones de expresión de genes del desarrollo embrionario temprano entre <i>A. ludens</i> y <i>D. melanogaster</i> ...	70
<b>Capítulo 6. Discusión ...</b>		<b>72</b>
<b>Capítulo 7. Conclusión y perspectivas ...</b>		<b>76</b>



**Capítulo 8. Referencias ... 77**

**Anexo 1.** Soluciones y métodos complementarios ... 90

**Anexo 2.** Esquema del análisis bioinformático de datos de secuenciación de *A. ludens* ...  
94

**Anexo 3.** Oligonucleótidos ... 95

**Anexo 4.** Secuencias de los genes anotados de *A. ludens* ... 96

**Anexo 5.** Secuencias repetidas identificadas ... 99

**Anexo 6.** Secuencias de pre - miRNAs ... 103

**Anexo 7.** Secuencias de PFAM ... 105

**Anexo 8.** Hibridaciones *in situ* de embriones de *D. melanogaster* con sondas a partir de los transcritos de *A. ludens* ... 106

**Anexo 9.** Artículo publicado ... 107

## Índice figuras

- Figura 1.1.** Características de *A. ludens* ... 16
- Figura 2.1.** Genes regulatorios que participan en la segmentación antero - posterior del embrión de *D. melanogaster* ... 23
- Figura 2.2.** Descripción del huevo de *A. ludens* ... 26
- Figura 2.3.** Principales etapas de desarrollo embrionario de *A. ludens* ... 29
- Figura 4.1.** Cronograma de las colectas de los embriones ... 32
- Figura 5.1.** Características de ellos embriones tempranos de *A. ludens* ... 43
- Figura 5.2.** Esquema para la colecta de embriones para la generación de los datos de RNA-seq ... 44
- Figura 5.3.** Diagramas del contenido de GC por secuencia ... 49
- Figura 5.4.** Correlación de similitud entre transcritos presentes en las réplicas de cada etapa de desarrollo ... 53
- Figura 5.5.** Porcentaje de transcritos anotados de las etapas de desarrollo embrionario de *A. ludens* ... 54
- Figura 5.6.** Anotación de la ontología de genes de cada transcriptoma ... 56
- Figura 5.7.** Dinámicas de expresión en la embriogénesis temprana de *A. ludens* ... 58
- Figura 5.8.** Temporalidad de la embriogénesis entre *A. ludens* con otras moscas ... 61
- Figura 5.9.** Transcritos compartidos entre los embriones de *A. ludens* con *B. dorsalis*, *C. capitata* y *D. melanogaster* ... 62
- Figura 5.10.** Genes del desarrollo temprano de *D. melanogaster* identificados en *A. ludens* ... 64
- Figura 5.11.** Genes maternos y cigóticos presentes en *A. ludens*, *C. capitata*, *B. dorsalis* y *D. melanogaster* ... 65
- Figura 5.12.** Gen *exuperantia* de *A. ludens* ... 67

**Figura 5.13.** Gen *hunchback* de *A. ludens* ... 68

**Figura 5.14.** Gen *even skipped* de *A. ludens* ... 69

**Figura 5.15.** Gen *engrailed* de *A. ludens* ... 70

**Figura 5.16.** Localización de genes involucrados en el patrón antero - posterior del embrión de *A. ludens* ... 71

## Índice de tablas

**Tabla 5.1.** Descripción de las características morfológicas de las etapas de sincicio, celularización y gastrulación en embriones de *A. ludens* ... 42

**Tabla 5.2.** Relación de colectas entre las diferentes etapas de desarrollo temprano de *A. ludens* ... 45

**Tabla 5.3.** Calidad de las muestras de RNA utilizadas para la construcción de librerías ... 46

**Tabla 5.4.** Resultados globales de cada librería secuenciada ... 47

**Tabla 5.5.** Resumen del análisis de calidad de cada librería generada ... 48

**Tabla 5.6.** Métricas estadísticas de los transcriptomas ensamblados *de novo* ... 51

**Tabla 5.7.** Resumen del análisis comparativo de los transcritos evolutivamente conservados en Artrópodos ... 52

**Tabla 5.8.** Presencia de los genes de determinación sexual en *A. ludens*, *B. dorsalis* y *C. capitata* ... 59

## Resumen

*Anastrepha ludens*, comúnmente conocida como *mosca mexicana de la fruta*, es un insecto plaga que tiene un impacto económico muy importante; las frutas a las cuales infesta son principalmente cítricos y mangos en México y regiones de América Central. A pesar de su importancia económica, el conocimiento del desarrollo embrionario de *A. ludens* es escaso; sin embargo, en años recientes se han realizado esfuerzos para contar con una descripción detallada de sus eventos morfológicos embrionarios. A partir de descripciones embrionarias se realizó el primer panorama transcripcional de las etapas embrionarias de sincicio (0 a 9 h), celularización (9 a 18 h) y gastrulación (18 a 30 h). La caracterización de las dinámicas transcripcionales entre sincicio, celularización y gastrulación presentó un aproximado de 9 000 transcritos diferentes en cada etapa de desarrollo; los cuales están involucrados en procesos de segmentación, genes de determinación sexual, metabolismo, entre otros. Asimismo, se realizó un análisis transcripcional comparativo con moscas modelo (*Drosophila melanogaster*, *Bactrocera dorsalis* y *Ceratitidis capitata*), a partir de estos datos se identificó una gran cantidad de transcritos conservados entre las cuatro moscas. No obstante, se encontró un número de transcritos no presentes en *D. melanogaster*, pero sí en *A. ludens*, *B. dorsalis* y *C. capitata*, por lo que podrían ser transcritos específicos de tefritidos; así como la ausencia de transcritos descritos en el desarrollo embrionario de *D. melanogaster*, como el caso de *bicoid*. La localización de los transcritos es importante para determinar el plan corporal de *D. melanogaster*, razón por la cual se realizaron hibridaciones *in situ* en *A. ludens*. Los transcritos seleccionados son importantes en la segmentación antero - posterior del embrión (*exu*, *hb*, *eve* y *en*); en *A. ludens* se observó que los patrones de localización son similares a los reportados para *D. melanogaster* con excepción de *exu*. En general observamos características únicas en la embriogénesis de *A. ludens*, que podrían representar las bases para comprender mejor los mecanismos regulatorios, así como para el diseño de estrategias para el control de la plaga.

## Abstract

*Anastrepha ludens*, commonly known as the Mexican fruit fly, is an important pest insect that primarily infests citrus and mangoes trees in Mexico and regions of Central America. Despite its economic importance, knowledge of *A. ludens* embryonic development is scarce; however, in recent years efforts have been made to describe its embryonic morphological events in detail. From embryonic descriptions, the first transcriptional landscape of the embryonic stages of syncytium (0 to 9 h), cellularization (9 to 18 h) and gastrulation (18 to 30 h) was evaluated. The characterization of the transcriptional dynamics between syncytium, cellularization and gastrulation presented approximately 9 000 different transcripts at each stage; which are mostly involved in segmentation processes, sexual determination, and metabolism, among others. Furthermore, a comparative transcriptional analysis was performed with model flies (*Drosophila melanogaster*, *Bactrocera dorsalis* and *Ceratitis capitata*), based on these data, a large number of conserved transcripts were identified among the four flies. However, a number of transcripts were not found in *D. melanogaster*, but were found in *A. ludens*, *B. dorsalis* and *C. capitata*, which could be tephritid specific transcripts; and some transcripts described in the embryonic development of *D. melanogaster*, as in the case of *bicoid*. The location of the transcripts is important to determine the *Drosophila* body plan, which is why *in situ* hybridizations were performed in *A. ludens*. The selected transcripts are important in the anteroposterior segmentation of the embryo (*exu*, *hb*, *eve* and *en*); in *A. ludens* it was observed that the localization patterns are similar to those reported for *D. melanogaster* with the exception of *exu*, but unique characteristics were also observed in *A. ludens* embryogenesis. These results could help to generate a platform for understanding the regulatory mechanisms of tephritid development, as well as aid in the design of strategies for pest control.

## Capítulo 1. Introducción

Los dípteros son, sin lugar a duda, de los insectos más importantes en la vida del ser humano y los animales (Wiegmann *et al.*, 2003), por su gran variedad de hábitats y papeles ecológicos. Los dípteros ocupan muchos nichos ecológicos como degradadores, saprófagos, depredadores o parasitoides; así como vectores de enfermedades y especies plaga (Wiegmann *et al.*, 2003; Wiegmann and Yeates, 2017). Una familia de dípteros de gran relevancia, por sus características de insecto plaga, es la Tephritidae (Hernandez-Ortiz, Guillén-Aguilar and Lopez, 2010).

Las moscas de la familia Tephritidae son conocidas como “las verdaderas moscas de la fruta” debido a que la larva crece y se alimenta dentro de frutos, muchos de ellos de importancia económica (Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944; Hernandez-Ortiz, Guillén-Aguilar and Lopez, 2010). Los tefrítidos cuentan con 5000 especies descritas hasta el momento, su distribución es cosmopolita y los géneros de mayor importancia económica son: *Anastrepha*, *Bactrocera*, *Ceratitis*, *Dacus* y *Rhagoletis* (Aluja and Norrbom, 1999). Dentro del género *Anastrepha*, el cual se encuentra distribuido a lo largo del continente americano, encontramos siete especies particularmente importantes, debido a su amplia distribución y el rango de hospederos de los que se alimentan. Estas son: *Anastrepha ludens*, *A. obliqua*, *A. fraterculus*, *A. suspensa*, *A. serpentina*, *A. striata* y *A. grandis* (Aluja and Norrbom, 1999; (Conabio), 2011).

### 1.1 *Anastrepha ludens*, la mosca mexicana de la fruta

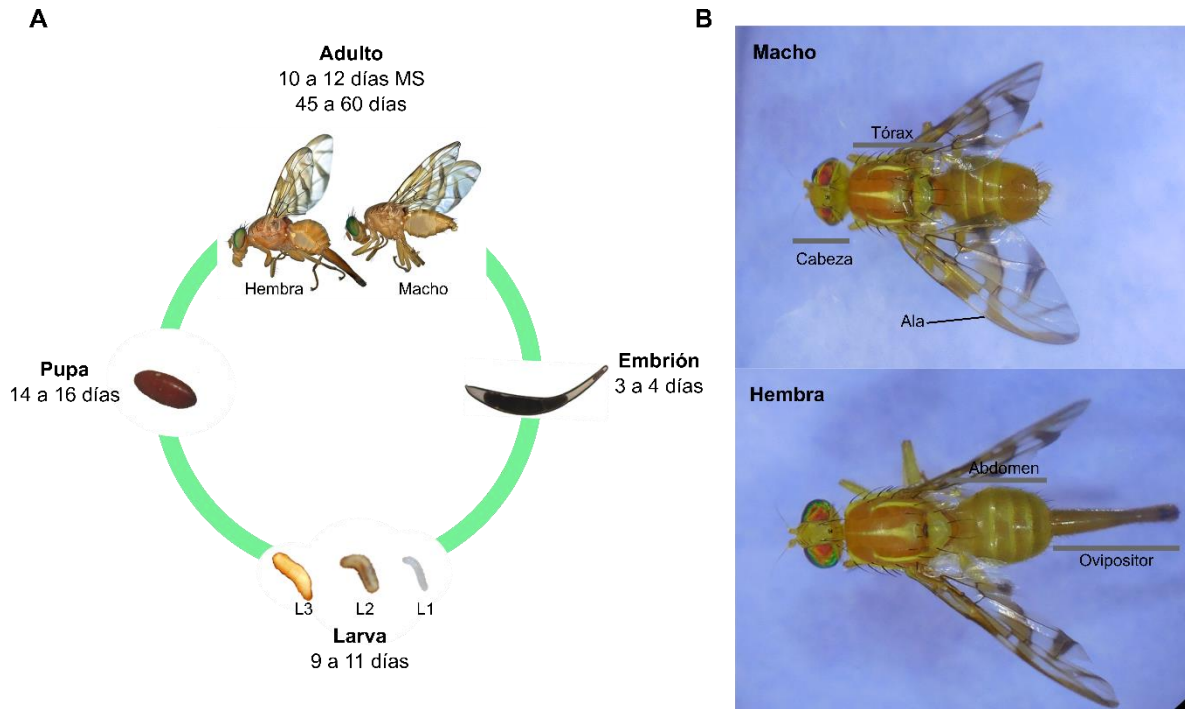
A mediados del siglo XIX, el entomólogo F. Hermann Loew realizó observaciones de especies de dípteros en América del Norte. Entre 1862 y 1873 publicó una serie de libros en donde describe una gran cantidad de dípteros endémicos de esta región. En su libro publicado en 1873 realizó la primera descripción de *Trypetina ludens*, ahora conocida como *Anastrepha ludens* (Loew, 1873). A inicios del siglo XX, empezó a observar el poder destructor que tenía *A. ludens* en los cultivos de naranja en el estado de Morelos, México. Debido a la temperatura tropical, y la cantidad de hospederos disponibles en los cultivos de naranjas del estado, la región es adecuada para albergar a esta mosca, lo cual produce un efecto negativo en la producción y exportación de la fruta (Herrera *et al.*, 1903). Debido a la

importancia de las pérdidas económicas por insectos plaga en los cultivos frutales, el gobierno de México, en cooperación con el gobierno Norteamericano, comenzaron a realizar estudios para conocer las características de *A. ludens*, sus hospederos, su ciclo de vida, así como plantear estrategias de acción para su control (Herrera *et al.*, 1903; McPhail and Bliss, 1933).

*A. ludens* es conocida como “mosca mexicana de la fruta” debido a que es endémica de México; su distribución abarca desde el nordeste de México y hasta América Central (Stone, 1942). Las hembras de *A. ludens* causan el daño, ya que ellas puncionan la piel del fruto con su ovipositor para depositar los huevos. En el mesocarpio del fruto se desarrollan los huevos y, después de un tiempo eclosionan las larvas, que se alimentan de éste, arruinando el fruto (Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944). Los principales frutos que infecta pertenecen a la familia de las rutáceas, como los cítricos y especies endémicas de México como el chicozapote, chapote amarillo y el zapote blanco; también a especies como durazno, membrillo y mango (Stone, 1942; Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944).

El ciclo de vida de *A. ludens* se describió gracias a las observaciones de campo (Figura 1.1.A): la etapa embrionaria se desarrolla entre 3 a 4 días (McPhail and Bliss, 1933; Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944); después el estado larvario, que cuenta con 3 etapas dura de 9 a 11 días. Las etapas del estado larvario son: L1 donde la larva recién eclosionada presenta un color blanquecino; L2 donde la larva se alimenta del fruto tomando una pigmentación amarilla y se observa su intestino, el cual se distingue por su color marrón y L3, donde ya no se observa el intestino y la larva cuenta con un color amarillo claro uniforme. En la etapa de pupa, la larva sale del fruto y comienza a pupar en el suelo, periodo que dura de 14 a 16 días (Greene, 1929). Posteriormente, eclosiona la mosca adulta, en donde se aprecian características fenotípicas propias de *A. ludens* las cuales son: en el tórax se observan franjas en ambos lados a lo largo del escudo de color amarillo pálido que se juntan en el escutelo; sus alas tienen bandas de color marrón amarillento en forma de S; y en las hembras se observa el ovipositor (Figura 1.1.B) (Loew, 1873; Aluja and Norrbom, 1999).





**Figura 1.1.** Características de *A. ludens*. En A, esquema del ciclo de vida de *A. ludens* creciendo en condiciones de vida silvestre, en donde llega a vivir hasta 100 días. En B, moscas adultas de machos y hembras, algunas de sus partes corporales se encuentran señaladas.

El tiempo en el cual se llevan a cabo las diferentes etapas del ciclo de vida de *A. ludens* está relacionado con las condiciones ambientales y del tipo de hospedero (Messenger and Flitters, 1957); considerando el tamaño y madurez del fruto (Leyva, Browning and Gilstrap, 1991), contando con una adaptación específica para cada hospedero (Celedonio-Hurtado *et al.*, 1988). La nutrición de las moscas adultas es importante para su desarrollo y tiempo de vida, ya que una vez que emerge el adulto cuenta con un periodo de madurez sexual (MS) que dura de 10 a 12 días, por lo que requiere de una fuente de aminoácidos, vitaminas, minerales, carbohidratos y agua. Estos nutrientes los consigue de heces de aves, gotas de lluvia y líquidos de frutos dañados. Las moscas adultas requieren de alimentación balanceada para desarrollarse de forma óptima y poder reproducirse, es decir, los machos desarrollan sus glándulas salivales, importantes para la producción de feromonas, y las hembras el desarrollo de sus ovarios (Dickens, J. C., Solis, E., Hart, 1982; Aluja, 1994). La longevidad promedio del adulto es de 45 a 60 días; sin embargo, se sabe que con una alimentación fuerte en carbohidratos las moscas llegan a

vivir hasta 11 meses si son hembras y hasta 16 meses si son machos (McPhail and Bliss, 1933; Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944).

## **1.2 Importancia del estudio de *A. ludens***

La mayor parte del conocimiento del género *Anastrepha* se obtuvo entre 1900 y 1944; posteriormente, dependiendo de la región, las investigaciones se enfocaron en *Anastrepha* endémicas. Los principales campos de conocimiento desarrollado son: taxonomía y filogenia, distribución, relación planta-hospedero, demografía, genética de poblaciones, ecología del comportamiento y los métodos para el manejo de la plaga (Aluja, 1994; Aluja and Norrbom, 1999). A pesar de contar con información restringida a pocas áreas, ésta ha sido de vital importancia para el manejo y control de la plaga. En México, como anteriormente se mencionó, desde inicios del siglo XX se conoce el papel de *A. ludens* como insecto plaga en frutos como mango y naranja, razón por la cual, en 1900 se fundó la comisión de Parasitología para desarrollar herramientas e investigar a las moscas de la fruta. Conforme avanzó la investigación se crearon los lineamientos para contar con un control integrado de plagas para el saneamiento y producción frutícola del país. Estos acuerdos se vieron registrados en 1992 en la Campaña Nacional contra Moscas de la Fruta (CNCMF), cuyo objetivo es la administración de los recursos para controlar, suprimir y erradicar las moscas de la fruta (Gutiérrez *et al.*, 2013).

Actualmente, el Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria (SENASICA), de la Secretaría de Agricultura y Desarrollo Rural, es el encargado de dirigir la CNCMF, con el manejo integrado de plagas (control mecánico, químico, autocida y biológico) y acciones de trampeo y muestreo de frutos (SENASICA, 2020). Como consecuencia de las acciones de la Comisión, los estados de Baja California, Baja California Sur, Coahuila, Sonora y el norte de Sinaloa están libres de la plaga y los estados restantes se consideran como zonas de baja prevalencia o zonas bajo control fitosanitario (DGSV, 2020; SENASICA, 2020; SIRVEF, 2020).

No obstante, a pesar de contar con más de un siglo de conocimiento sobre *A. ludens*, la información generada deja a un lado estudios sobre la biología de su desarrollo; los cuales podrían proveer las bases para el desarrollo de nuevas estrategias para combatir la plaga (Gutiérrez-Ramos *et al.*, 2020). Ejemplos del uso de nuevas herramientas

tecnológicas y de investigación se pueden observar en los tefrítidos *Ceratitis capitata* y *Bactrocera dorsalis*. Ambas moscas cuentan con sus genomas de referencia (USDA-ARS, 2015; Papanicolaou *et al.*, 2016), cuentan con estudios centrados en genes de determinación sexual con intención de realizar mutaciones dirigidas (Pane *et al.*, 2002; Chen *et al.*, 2008; Gomulski *et al.*, 2012), identificar genes de resistencia a insecticidas (Hsu *et al.*, 2012), o el uso de herramientas y técnicas de vanguardia que permitan conocerlas mejor, como el uso de CRISPR-Cas 9 (Meccariello *et al.*, 2017; Bai *et al.*, 2019); por mencionar un ejemplo.

## Capítulo 2. Antecedentes

En el siglo XIX se observó que existían similitudes en el periodo del desarrollo embrionario entre diferentes especies, por lo que surgió la duda de cómo es que se presentan las novedades morfológicas y los planes corporales entre los diferentes organismos (Kalinka and Tomancak, 2012), (Angelini and Kaufman, 2005).

Razón por lo cual, el estudio del desarrollo y la biología evolutiva son importantes para explorar los cambios de los organismos a lo largo del tiempo (Raff, 2000). El estudio del desarrollo embrionario, sus características morfológicas y mecanismos de regulación, se ha centrado en seis modelos de estudio principalmente: *Gallus gallus*, *Caenorhabditis elegans*, *Mus musculus*, *Xenopus laevis*, *Danio rerio* y *Drosophila melanogaster*; lo cual limita las características que se comparten, dejando de lado la gran biodiversidad que existe (Jenner and Wills, 2007). Estudiar la conservación de los patrones de desarrollo embrionario de forma comparativa en organismos diversos nos puede ayudar a tener un mejor entendimiento de la relación entre el desarrollo, la ecología y los patrones corporales que generan biodiversidad (Kalinka and Tomancak, 2012).

### 2.1 Impacto del estudio de la embriogénesis en los dípteros

Los insectos, debido a su gran variedad de roles ecológicos, hábitats, plasticidad y simbiosis con otros organismos, además de su facilidad del manejo en el laboratorio son excelentes para estudiar las novedades morfológicas (Angelini and Kaufman, 2005; Peel, 2008). Los insectos se clasifican según su tipo de desarrollo en hemimetábolos y holometábolos. Los insectos holometábolos, los cuales sufren una metamorfosis completa de larva a adulto vía pupa, son muy importantes para los seres humanos, principalmente las ordenes siguientes: Díptera (moscas de dos alas), Lepidóptera (mariposas y palomillas), Coleóptera (escarabajos) e Himenóptera (abejas, avispas y hormigas) (Peel, 2008).

Uno de los grandes organismos modelo es el díptero *D. melanogaster*, del que ya se conocen sus redes de regulación genética embrionaria (Peel, 2008). Las redes de regulación genética describen y explican las complejas dependencias entre factores de transcripción, claves en el desarrollo y sus genes blanco (Levine and Davidson, 2005;

Wilczynski and Furlong, 2010). De los primeros genes que se identificaron con un comportamiento conexo entre la embriogénesis y el adulto, mostrando un impacto en el plan corporal, son los genes *Hox* en *D. melanogaster*. Posteriormente, los genes *Hox* se identificaron en la mayoría de los metazoarios, desde escarabajos y gusanos hasta el ser humano; siendo genes fundamentales en la determinación de la identidad de los segmentos del cuerpo (Carroll, 2008; Heffer and Pick, 2013).

Para identificar las redes de regulación genética, es importante considerar estos factores para generar una red precisa: I) patrón espacio - temporal de los factores transcripcionales; II) cuál módulo *cis*-regulatorio se une a quién; y III) una conexión causal entre la actividad de los factores transcripcionales con la expresión de los genes blanco (Levine and Davidson, 2005; Wilczynski and Furlong, 2010). Debido a las características de *D. melanogaster* (Jenner and Wills, 2007) se ha convertido en un modelo clave para identificar redes de regulación genética, por ejemplo: la segmentación a lo largo del embrión en el eje antero-posterior y en el eje dorso-ventral (Wilczynski and Furlong, 2010). Esto se ha logrado gracias a la información generada con el estudio de *D. melanogaster*, desde la perspectiva morfológica hasta las aproximaciones ómicas como la genómica, la transcriptómica y la proteómica, por nombrar algunas.

## **2.2 Embriogénesis en *Drosophila melanogaster***

A inicios del siglo XX se crío por primera vez *D. melanogaster*, probablemente por su corto tiempo de generación y su facilidad de cultivo. Fue con la investigación de Thomas H. Morgan que *D. melanogaster* se convirtió en organismo modelo, debido a la identificación de los genes de *D. melanogaster* y el establecer que éstos se encuentran en los cromosomas (Hales *et al.*, 2015; Markow, 2015). El ciclo de vida de *D. melanogaster* es corto, el cual ocurre de 9 a 10 días a 25 °C, considerando que su etapa embrionaria es menos de 24 h; lo cual es un gran atributo para los estudios del desarrollo embrionario (Hales *et al.*, 2015).

### **2.2.1 Estudios morfológicos de la embriogénesis de *D. melanogaster***

Uno de los primeros trabajos enfocados en ordenar y complementar estudios de la embriogénesis de *D. melanogaster* fue el de Morris Rabinowitz (Campos-Ortega and

Hartenstein, 1985). Su trabajo se centró en caracterizar la ovogénesis en *D. melanogaster*, realizando una descripción detallada de, entre otras cosas, el proceso de formación del blastema, la migración y desaparición de dos de los tres núcleos polares del oocito (Rabinowitz, 1941). En 1975 se publicó un estudio fotográfico cronológico del desarrollo embrionario de *D. melanogaster*, realizado mediante observaciones de embriones vivos por microscopía de campo claro. En este trabajo se lograron describir los principales cambios morfológicos durante las primeras 22 h del desarrollo embrionario, se clasificaron en catorce etapas. Los principales cambios que caracterizan cada etapa son: I) etapa de multiplicación; II) acortamiento del embrión; III) migración de los núcleos a la superficie del huevo; IV) blastodermo sincicial; V) extensión de las membranas celulares; VI) blastodermo celular; VII) extensión de los rudimentos del intestino medio; VIII) extensión de la banda germinal; IX) plegamiento de los rudimentos delanteros y posteriores; X) acortamiento de la banda germinal; XI) cierre dorsal e involución de la cabeza; XII) unión de parte cefálica con los segmentos torácicos; XIII) digestión de la yema y enroscamiento del intestino; y XIV) absorción total de la yema y eclosión de larva (Bownes, 1975).

Posteriormente, utilizando tinciones nucleares, se realizaron estudios más detallados de las primeras etapas de la embriogénesis de *D. melanogaster*. Un estudio clave es el de Foe y Alberts *et al.*, en donde se observaron los primeros ciclos mitóticos en el desarrollo del embrión de *D. melanogaster* (Foe and Alberts, 1983). Conforme se incrementaba la cantidad de información de la embriogénesis de *D. melanogaster*, en 1985 se realizó una compilación de la información sobre la embriogénesis donde se propuso que el desarrollo del embrión se puede clasificar en diecisiete etapas, las cuales se resumen en siete eventos principales de desarrollo: escisión, blastodermo, gastrulación, elongación de la banda germinal, retracción de la banda germinal, involución de la cabeza y cierre dorsal, y diferenciación (Campos-Ortega and Hartenstein, 1985).

### 2.2.2 Mecanismos de regulación del eje antero - posterior del embrión de *D. melanogaster*

Una de las tantas ventajas de utilizar insectos para estudiar los procesos de organización durante la embriogénesis es que tienen una estructura segmentada, es decir, tienen una ubicación espacial para la cabeza, el tórax y el abdomen; y se pueden realizar manipulaciones experimentales simples como trasplantes de células de un lugar a otro o realizar ligaciones en distintas partes del embrión afectando el plan corporal del insecto

(Meinhardt, 1977; Russell, 1985). Una pregunta interesante era saber cómo las células sabían en qué momento debían desarrollarse. Se llegó a creer que era entre el periodo de la oviposición y la formación del blastodermo (Meinhardt, 1977). Experimentos realizados con el embrión de *D. melanogaster*, consistentes en la inyección de células del plasma polar (células germinales ubicadas en la parte posterior) a un sitio anterior del embrión, dieron como resultado que las células inyectadas no lograron transformarse en células germinales. Gracias a estos resultados se propuso la existencia de un sistema de gradiente que limita la formación de los segmentos a lo largo del eje antero - posterior, el cual estaría presente desde el ovocito (Simcox and Sang, 1983; Russell, 1985; French, 1988).

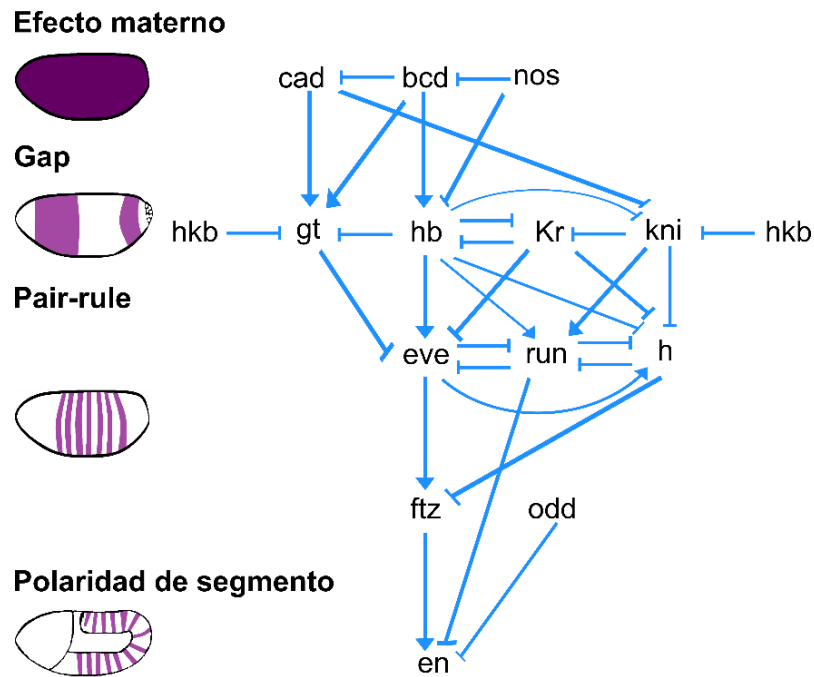
Estos determinantes encontrados tendrían la función de estructurar el eje antero - posterior en el embrión de *D. melanogaster*. Posteriormente, se encontró que estos determinantes son genes que se identificaron por medio de mutaciones y selección de fenotipos que cambiaban la forma de la larva o la mosca adulta (French, 1988; Gilbert, 2014). Los pioneros en identificar y jerarquizar los genes que establecen la polaridad antero - posterior, y la división del embrión en segmentos específicos, fueron Christiane Nüsslein-Volhard y Eric Wieschaus (Gilbert, 2014). La jerarquía de los genes involucrados en establecer los gradientes, se clasifica en genes producidos durante la ovogénesis como genes de efecto materno, y los que están expresados dentro del embrión como genes cigóticos (French, 1988; Gilbert, 2014).

El primer gen identificado como morfógeno fue *bicoid* (*bcd*), ya que se observó que su proteína se encontraba en forma de gradiente a lo largo del embrión. En experimentos con embriones carentes de la proteína Bcd se observó que no formaban la cabeza, no obstante, en experimentos de rescate, es decir, mediante la inyección de ARN mensajero de *bcd* sí formaban la cabeza (Kimelman and Martin, 2012). Posteriormente, se sabría cuáles son los genes que participan en la determinación antero - posterior del embrión en *D. melanogaster*, se podrían considerar como una caja de herramientas génica, debido a que son importantes para definir la localización, la formación de segmentos, la formación de tipos celulares y la especificación de los planes corporales (Carroll, Grenier and Weatherbee, 2005; Gilbert, 2014).

Los genes que pertenecen a esta caja de herramientas se pueden clasificar en: genes de efecto materno; genes gap, los cuales regulan la formación de segmentos continuos; genes pair-rule, que dividen al embrión en unidades periódicas; genes de

polaridad de segmento, que dividen al embrión en 14 segmentos; y los genes de selección homeótica, los cuales determinan el destino del desarrollo de cada segmento (Figura 2.1) (Carroll, Grenier and Weatherbee, 2005; Gilbert, 2014).

El descubrimiento de estos genes se realizó a partir de experimentos de *forward genetics*, en donde, después de realizar una mutación, se identifica el conjunto de genes o el gen que da origen a ese fenotipo particular (Gilbert, 2014). El uso de aproximaciones ómicas ha sido importante para conocer a detalle los genes que regulan este proceso de desarrollo. Poco después de la secuenciación del genoma de *D. melanogaster*, se llevaron a cabo estudios de su transcriptoma a lo largo de su ciclo de vida. Debido a su corto tiempo embrionario, se tomaron muestras cada hora durante 24 h, así como entre horas para contar con posibles transcritos que se perdieran entre colectas. Los resultados mostraron que el 88% de los transcritos se expresan en las primeras 20 h del desarrollo de la mosca (Arbeitman *et al.*, 2002).



**Figura 2.1.** Genes regulatorios que participan en la segmentación antero - posterior del embrión de *D. melanogaster*. Los patrones de expresión de los genes que participan en la segmentación antero - posterior del embrión de *D. melanogaster* son representados según su orden de función. Los genes maternos son *caudal (cad)*, *bicoid (bcd)* y *nanos (nos)*; los genes gap son *huckebein (hkb)*, *giant (gt)*, *hunchback (hb)*, *Kruppel (Kr)* y *knirps (kni)*; los genes pair rule son *even skipped (eve)*, *runt (run)*,



*hairy (h)*, *fushi tarazu (ftz)* y *odd skipped (odd)* y el gen de polaridad de segmento es *engrailed (en)*.  
Fuente: Carroll, Grenier and Weatherbee, 2005.

Conforme mejoraron las técnicas de secuenciación, se realizaron más ensayos del tipo transcriptómico en embriones de *D. melanogaster*, por ejemplo, en algunos casos estos ensayos se complementaron con estudios de hibridaciones *in situ* para conocer la localización de los transcritos encontrados (Tomancak *et al.*, 2007); en otros se logró identificar nuevas isoformas de genes previamente descritos (Graveley *et al.*, 2011); otros se han encargado de estudiar el desarrollo a nivel de embrión único (Lott *et al.*, 2011) y, algunos más actuales han estudiado los patrones de accesibilidad de la cromatina en genes involucrados en la segmentación del embrión (Haines and Eisen, 2018). Todo esto nos brinda información y herramientas que permiten trasladar este conocimiento a otros insectos que son importantes para el humano, como son los insectos plaga.

### **2.3 Embriogénesis de los tefrítidos**

El estudio del ciclo de vida de los tefrítidos es muy importante, pero se ha centrado en aquellas especies que tienen importancia económica (Aluja and Norrbom, 1999). Dichos estudios se centran en el desarrollo de las pupas y los insectos adultos, debido a que su estudio es sencillo desde el punto de vista de su manejo experimental. En el caso del estadio de embrión o huevo las observaciones llevadas a cabo son principalmente de la forma externa (Greene, 1929), no obstante, hay excepciones.

El primer miembro de la familia de los tefrítidos del que se conoció su desarrollo embrionario fue *Bactrocera tryoni*, conocida como la Mosca de la fruta de Queensland. Sus descripciones incluyeron los cambios morfológicos desde la fecundación hasta la organogénesis, realizando observaciones de embriones vivos a campo claro y embriones teñidos con hematoxilina-eosina, realizando un cronograma muy similar al descrito para *D. melanogaster* (Anderson, 1962), lo que permitió conocer los fenómenos clave de la embriogénesis de esta mosca. Posteriormente, mediante el uso de herramientas de biología comparativa y genómicas, se comenzó a realizar estudios de embriogénesis en tefrítidos con una aproximación de genómica funcional (Scolari *et al.*, 2014). Una mosca muy utilizada es la mosca de la fruta del mediterráneo, *C. capitata*, que tiene desde la secuenciación de

su genoma (Papanicolaou *et al.*, 2016) hasta estudios transcriptómicos en embriones (Salvemini *et al.*, 2014).

### 2.3.1 Embriogénesis en el género *Anastrepha*

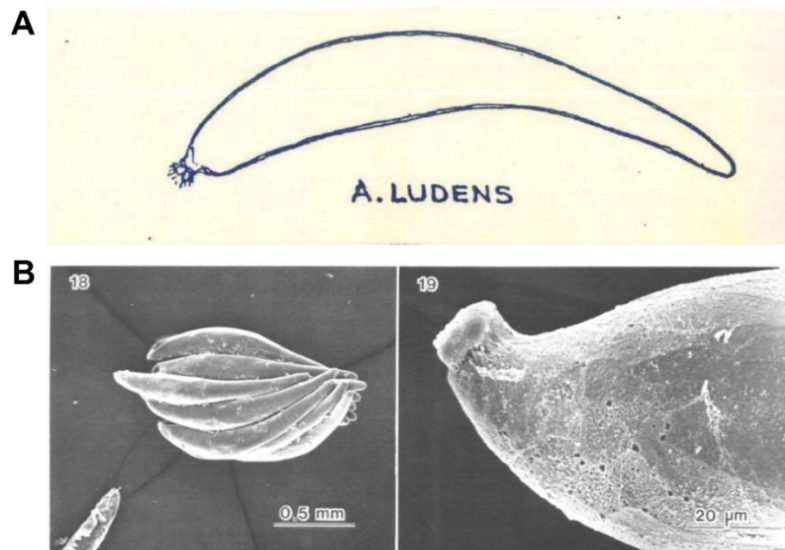
Las primeras etapas de desarrollo embrionario se han descrito en muy pocas moscas de la fruta, ya que crecen dentro del fruto (Stone, 1942). Las descripciones realizadas en el estadio de huevo son raras y llegan, en la mayoría de los casos, a ser muy generales (Dutra *et al.*, 2013). Las especies que cuentan con una descripción en su etapa de huevo que toman en cuenta la curvatura del huevo, el ahusamiento extremo anterior y posterior, localización del micropilo, etcétera (Figueiredo *et al.*, 2013), son las siguientes: *Anastrepha alveata*, *A. consobrina*, *A. amita*, *A. pickeli*, *A. bristrigata*, *A. grandis*, *A. obliqua*, *A. pseudoparallela*, *A. striata*, *A. suspensa*, *A. zenildae* (Figueiredo *et al.*, 2013), *A. curtis*, *A. leptozona* (Dutra *et al.*, 2013), *A. sororcula*, *A. manihoti*, *A. monteji*, *A. atrox*, *A. cordata*, *A. nigrifascia*, *A. pallens*, *A. pittieri*, *A. suspensa* (Aluja and Norrbom, 1999), *A. serpentina*, *A. striata*, *A. fraterculus* y *A. ludens* (Emmart, 1933). A pesar de contar con descripciones morfológicas del huevo, son pocas las especies que tienen una descripción más centrada en los eventos morfológicos del embrión como son: *A. fraterculus* (Selivon, Morgante and Perondini, 1997), *A. obliqua*, *A. sororcula*, *A. serpentina* (Perondini, Selivon and Morgante, 1998), y de forma más reciente *A. ludens* (Gutiérrez-Ramos, 2016; Gutiérrez-Ramos *et al.*, 2020). No obstante, a pesar de contar con pocas descripciones detalladas de la embriogénesis temprana de moscas *Anastrepha*, se ha llegado a responder algunas preguntas relacionadas con la embriogénesis, desde un punto de vista de biología comparativa y biología funcional, como es el caso de *A. suspensa* (Nirmala *et al.*, 2013), *A. fraterculus* (Rezende *et al.*, 2016) y *A. obliqua* (Sarno *et al.*, 2010; Gonçalves *et al.*, 2013).

### 2.4 Embriogénesis de *A. ludens*

Las primeras observaciones de los huevos de *A. ludens* se realizaron como parte de la descripción de su ciclo de vida. En campo, normalmente los huevos se mantienen justo debajo de la piel de los frutos que infesta, por lo que el periodo de tiempo de su desarrollo varía dependiendo de la temperatura del ambiente (McPhail and Bliss, 1933). Los datos con los que se cuenta sobre la embriogénesis de *A. ludens* se pueden agrupar en: morfología

del huevo; tiempo de desarrollo en campo; tamaño del *clutch* (lote); relación embrión - fruto hospedero; y, de forma más reciente una descripción morfológica del embrión a lo largo de toda su embriogénesis; los cuales se describen enseguida.

**Morfología del huevo.** El huevo es color blanco crema, elongado y con bordes redondeados, tiene forma de huso, es decir, es más ancho en el polo anterior y va disminuyendo conforme va al polo posterior. El micropilo se encuentra en el polo anterior, el cual cuenta con dos estructuras en forma de hoja que protegen el orificio y están rodeadas de un grupo de cerdas finas (Figura 2.2.A) (Emmart, 1933). Cerca del micropilo se observa una reticulación en el corión en forma de pentágonos y hexágonos irregulares que se desvanecen en la parte posterior del embrión. El tamaño del huevo puede medir entre 1.37 y 1.60 mm (desviación estándar de 0.010, n de 30) (Figura 2.2.B) (Carroll and Wharton, 1989).



**Figura 2.2.** Morfología externa del huevo de *A. ludens*. En A, huevo en una posición antero- posterior, donde se observa su forma de huso y las cerdas del micropilo en la parte anterior. Fuente: Emmart, 1933. En B, huevos observados por microscopía electrónica de barrido, a la izquierda se observa el clutch y a la derecha el micropilo y la reticulación del corión en la parte anterior. Fuente: Carroll and Wharton, 1989.

**Tiempo de desarrollo en campo.** En experimentos en campo se observó que el periodo embrionario a los 19.6 °C es de 9 días y conforme aumenta la temperatura el periodo se acorta hasta los 6 días a 24 °C (McPhail and Bliss, 1933). Posteriormente, los estudios de laboratorio se centraron en detectar posibles cambios en las tasas de crecimiento de los huevos, en los diferentes frutos o dietas artificiales en las cuales se crecieron. El desarrollo embrionario llega a durar 3.27 días en dieta artificial, presentando arriba de un 75% de eclosión de larvas (Celedonio-Hurtado *et al.*, 1988).

**Tamaño del clutch (lote).** *A. ludens* deposita sus huevos en lotes (clutch), es decir, un grupo de huevos fertilizados al mismo tiempo. En experimentos con diferentes tamaños de frutos se observó que en frutos pequeños el lote puede contener hasta 9 huevos; mientras que en frutos de tamaño medio a grande se ovipositan de 15 a 19 huevos (Leyva, Browning and Gilstrap, 1991). En otros estudios se determinó que la mosca llega a poner de 1 a 28 huevos por clutch, donde el clutch más pequeño llega a ser de 3.5 huevos/clutch (desviación estándar = 3.47, n=3377) y hasta un máximo de 10.8 huevos/clutch (desviación estándar = 5.50, n=165). Esto nos permite considerar que existe una fuerte correlación entre el tamaño del fruto y la cantidad de huevos puestos por clutch. También se observó que el color del fruto y la densidad de hospederos disponibles no afecta el tamaño del clutch (Berrigan *et al.*, 1988).

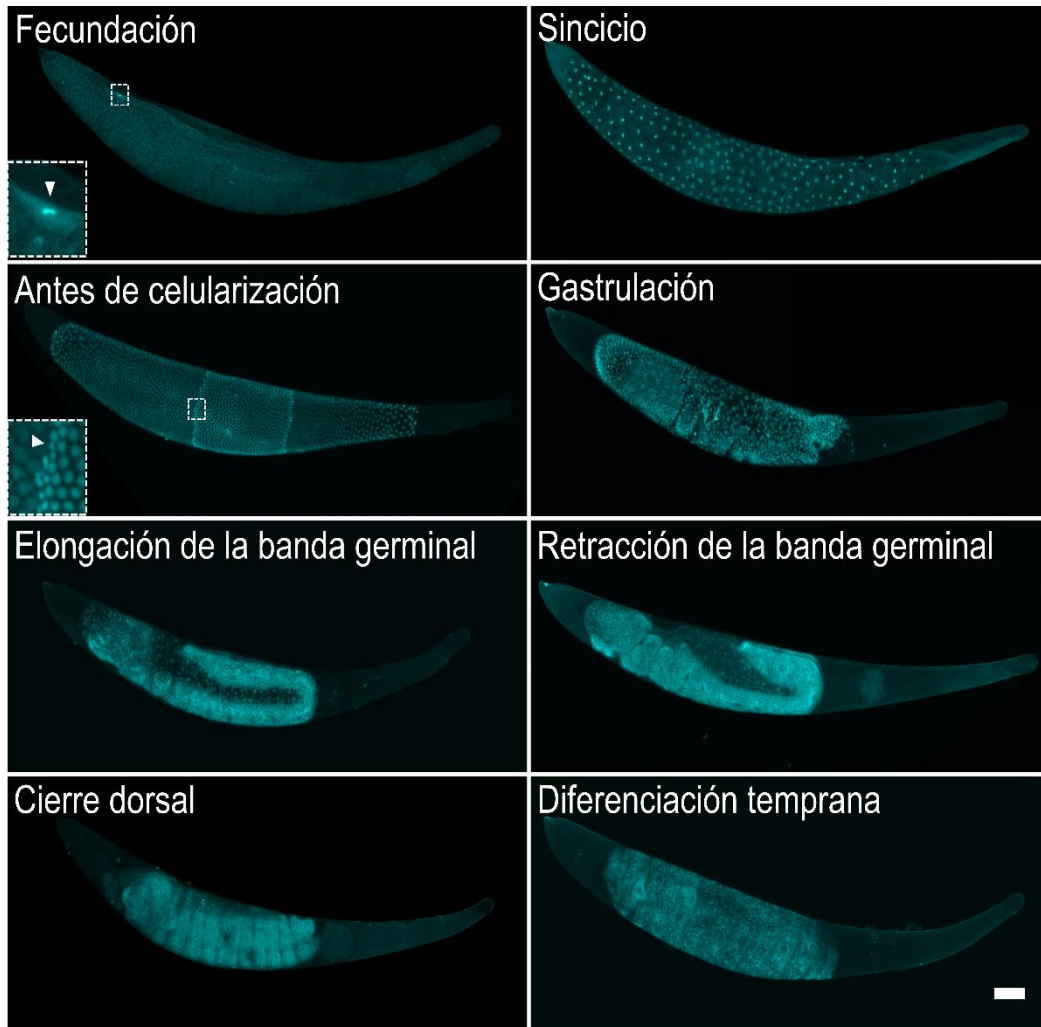
**Relación embrión – fruto hospedero.** En estudios en distintas especies hospederas de *A. ludens* la incubación del huevo puede ser corta de 3.8 días en toronja y larga de 4.6 días en mango (Leyva, Browning and Gilstrap, 1991); sin embargo, estas diferencias se consideran dentro de los rangos previamente reportados en campo y laboratorio (Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944; Berrigan *et al.*, 1988). Otra característica estudiada en los frutos hospederos es el nivel de madurez, el cual relacionan con la firmeza y color del fruto. En estudios con mangos inmaduros (verdes y firmes) y maduros (amarillos y suaves) y en bolas de gelatina que simulan los mangos inmaduros y maduros; se observó una preferencia de las hembras a ovipositar en mangos inmaduros. En el caso de las bolas de gelatina, las cuales contaban con el mismo nivel de firmeza, pero distinto color las hembras preferían las de color verde (Díaz-Fleischer and Aluja, 2003). Las hembras de *A. ludens* prefieren ovipositar en hospederos inmaduros, sin importar su tamaño (McPhail and Bliss, 1933; Díaz-Fleischer and Aluja, 2003).

**Morfología del embrión.** En observaciones realizadas en embriones vivos por microscopía de campo claro y en embriones fijados y teñidos por microscopía de fluorescencia, se logró identificar y cronometrar los eventos morfológicos a lo largo de toda la embriogénesis de *A. ludens*. Los estudios se realizaron a 25 °C bajo condiciones de crecimiento controlado, con una duración de desarrollo embrionario de 120 h, en donde se pudieron observar los principales eventos de desarrollo, previamente descritos en *D. melanogaster* (Figura 2.3) (Gutiérrez-Ramos, 2016).

El embrión de *A. ludens* presenta los principales eventos embrionarios observados en insectos (Figura 2.3). No obstante, cuenta con características que hacen único su desarrollo embrionario, desde su temporalidad de 120 horas hasta las extrusiones que presenta la yema (Gutiérrez-Ramos, 2016). Las extrusiones de la yema son una característica previamente observada en otras *Anastrepha* (Selivon *et al.*, 1996; Selivon, Morgante and Perondini, 1997; Perondini, Selivon and Morgante, 1998; Ruiz *et al.*, 2007), pero en el caso de *A. ludens* las extrusiones parecen marcar el desarrollo del embrión, es decir, si éstas no se presentan a su debido tiempo o etapa de desarrollo, el embrión muere. Esto parece ser un evento muy importante para la progresión del desarrollo, así como un factor que determina el tamaño de la larva (Gutiérrez-Ramos *et al.*, 2020).

## 2. 5 Justificación

En trabajos anteriores se cuenta con una descripción morfológica en donde se aprecian los cambios a un nivel de forma externa e interna, celular y nuclear, en la embriogénesis de *A. ludens*. Así como protocolos para coleccionar embriones y tinción de los tejidos importantes para aumentar la cantidad de herramientas que se pueden utilizar para estudiar la mosca mexicana de la fruta. Por lo que conocer a detalle la biología del desarrollo a un nivel transcripcional de *A. ludens* es muy importante para entender su regulación, las dinámicas de expresión de genes involucrados en la formación del patrón corporal y, a la larga, tener más opciones para el control de la plaga y el estudio de cambios fenotípicos por modificaciones genéticas de genes clave en el desarrollo embrionario.



**Figura 2.3.** Principales etapas de desarrollo embrionario de *A. ludens*. Embriones teñidos con un marcador de ADN fluorescente (DAPI) en diferentes etapas de desarrollo embrionario observados por microscopía confocal. En la fecundación observamos la primera división celular (en la magnificación se observa un núcleo en división). Continúa con la etapa de sincicio en donde los núcleos se encuentran distribuidos a lo largo del huevo; antes de llegar a la celularización se observan los núcleos en el proceso de citocinesis (en la magnificación se observan las diferentes formas de los núcleos). En la gastrulación se observan las invaginaciones a lo largo del eje dorsal; posteriormente, se observa la elongación de la banda germinal con el engrosamiento de las células a lo largo del eje dorsal – ventral y finaliza con la retracción de la banda germinal. En el cierre dorsal se observan con mayor detalle los segmentos a lo largo del eje dorsal y la parte cefálica del embrión. Para terminar el desarrollo embrionario se continúa con la etapa de diferenciación del embrión en donde la cabeza involuciona y se aprecian las estructuras de la probóscide de la larva. La barra de escala representa 100  $\mu\text{m}$ .

## Capítulo 3. Objetivos

### 3.1 Objetivo general

Caracterizar la dinámica transcripcional de las etapas de sincicio, celularización y gastrulación en la embriogénesis de la mosca *Anastrepha ludens*.

### 3.2 Objetivos particulares

- I. Generar un compendio de los transcritos de las etapas de sincicio, celularización y gastrulación de la embriogénesis de la *A. ludens*.
- II. Establecer un protocolo que permita determinar los patrones de localización de transcritos involucrados en las etapas de sincicio, celularización y gastrulación de la embriogénesis de *A. ludens*.
- III. Caracterizar la dinámica transcripcional de la embriogénesis temprana de *A. ludens*.

## Capítulo 4. Metodología

### 4.1 Cultivo de moscas *A. ludens*

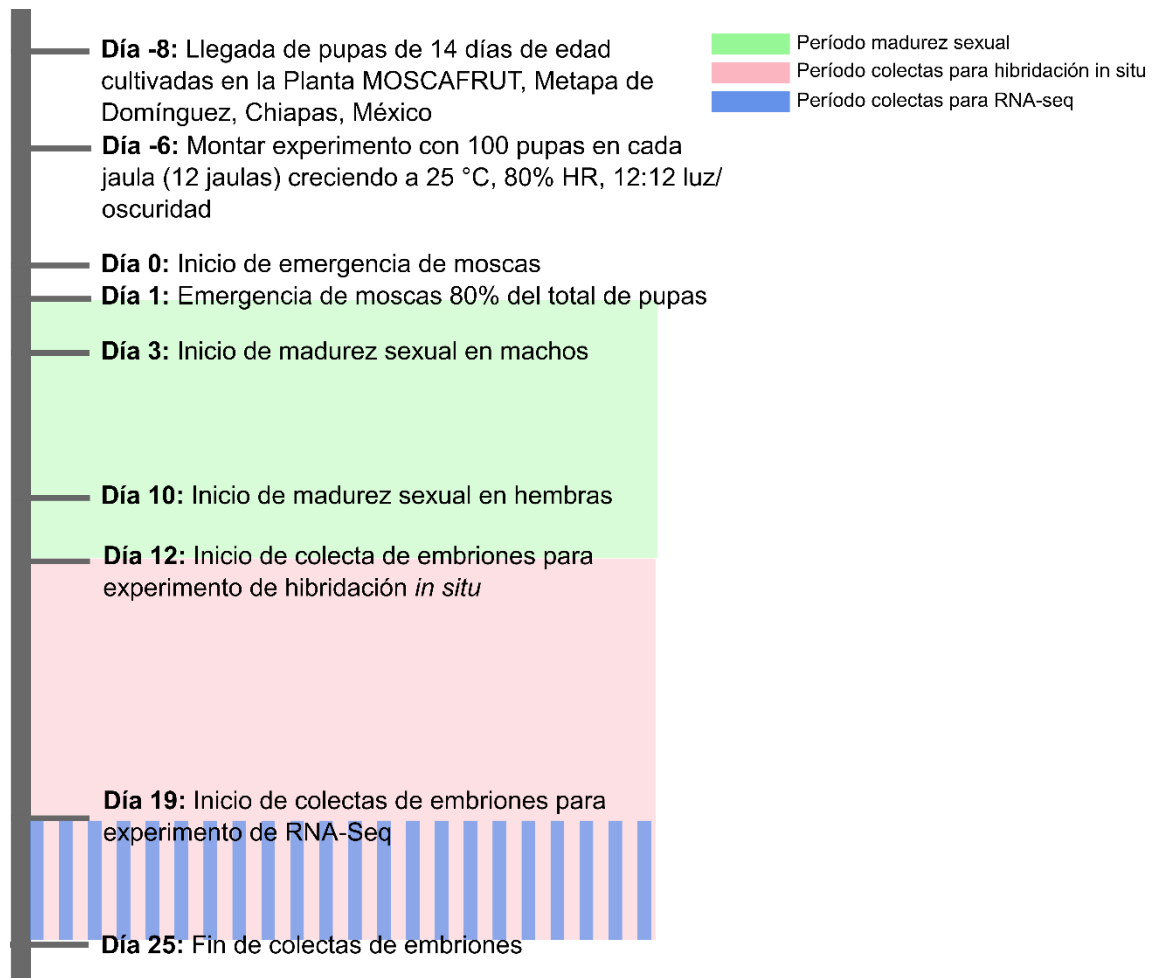
Las moscas *A. ludens* se obtuvieron de pupas cultivadas de la cepa de cría masiva producida en la Planta MOSCAFRUT (Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria - Secretaría de Agricultura y Desarrollo Rural (SENASICA-SADER)), localizada en Metapa de Domínguez, Chiapas, México. Las moscas adulto se criaron con dieta artificial (3:1 mezcla de azúcar: levadura liofilizada (MO Biomedicals, LLC), a 25 °C con 80% de humedad relativa y con un fotoperiodo de 12:12 h de luz/oscuridad.

### 4.2. Colecta y crecimiento de embriones de *A. ludens*

La colecta de los embriones se realizó en moscas adultas sexualmente maduras, entre los 12 y 25 días de adulto (Figura 4.1) (Dickens, J. C., Solis, E., Hart, 1982). Las moscas ovipositan en dispositivos llenos de agua limpia, los cuales son cajas de Petri con el fondo retirado y cubierto con tela satín negra con aplicación de silicón transparente (Sista, Henkel Mexicana S.A., Huixquilucan, Estado de México, México) colocados encima de una jaula de madera. Las colectas de embriones se realizan de la tercera a la séptima hora del ciclo de luz del fotoperiodo, con intervalos de 30 minutos.

Los embriones colectados se crecieron en cajas de Petri de vidrio con agua limpia, se observaron bajo un microscopio de disección para clasificarlos en las etapas de interés, que son, sincicio (0 a 9 h), celularización (9 a 18 h) y gastrulación (18 a 30 h).





**Figura 4.1.** Cronograma de las colectas de los embriones. La colecta de los embriones empieza con pupas de 14 días de edad, las cuales se seleccionan en grupos de 100 y se crecen en jaulas con alimento y agua a 25 °C y 80% de humedad relativa (HR). Una vez que emergen las moscas adultas, inicia su periodo de madurez sexual hasta las primeras oviposiciones. A partir del día 12 inician las colectas de embriones para las hibridaciones *in situ*; a partir del día 19 al 25 se realizan las colectas para los experimentos de RNA-Seq.

### 4.3 Extracción de RNA total de embriones de *A. ludens*

Una vez que se contó con los embriones de la etapa de desarrollo deseada, se depositaron en un tubo Eppendorf de 1.5 mL, en una cantidad aproximada de 500  $\mu$ L, y se congelaron en nitrógeno líquido. Se continuó con la molienda de los embriones congelados, haciendo uso de perlas de tungsteno en un pulverizador de tejidos (MM300 Grinder Shaker Pulverizer Tissuelyser, RETSCH) bajo condiciones de congelación. Una vez con el tejido pulverizado

se adicionaron 500 mL de TRI Reagent (SIGMA, T9424) para continuar con la extracción de ácidos ribonucleicos (RNA) totales de las muestras según el protocolo estándar. La integridad del RNA se analizó por medio de un gel desnaturizante de agarosa (ver Anexo 1) y la cuantificación de la absorbancia 260/280 por celda de cuarzo se realizó en un espectrofotómetro (ver Anexo 1).

#### **4.4 Construcción de librerías y secuenciación de muestras de RNA de embriones de *A. ludens***

Para cada etapa de desarrollo los experimentos biológicos, mencionados anteriormente, se realizaron por triplicado. La construcción de las librerías y la secuenciación de las muestras de RNA total de las etapas de desarrollo embrionarias se realizó bajo el servicio de Beijing Genomic Institute (BGI) Genomic Services. Las librerías fueron pareadas y la secuenciación obtenida fue de entre 106-58 millones de lecturas pareadas con 100 nucleótidos de longitud utilizando la plataforma de Illumina HiSeq 4000.

#### **4.5 Análisis bioinformáticos de RNA-seq de embriones de *A. ludens***

Los análisis bioinformáticos se llevaron a cabo en el clúster del Instituto de Biotecnología de la Universidad Nacional Autónoma de México. Una vez que se contó con los datos de la secuenciación de las muestras se siguió con el flujo de trabajo esquematizado en el Anexo 2.

##### **4.5.1 Análisis de calidad de secuencias crudas y ensamble *de novo***

Una vez que se contó con las lecturas crudas, el servicio de BGI Genomic Services realizó la remoción de lecturas de baja calidad y de adaptadores para la secuenciación. De este paso se obtuvieron 566,454,896 millones de lecturas pareadas, a las cuales se les realizó el análisis de calidad de secuenciación utilizando el programa de FastQC v0.11.7 (Andrews, 2010), así como un análisis de calidad de las réplicas biológicas por el programa de MutiQC v1.5.dev0 (Ewels *et al.*, 2016). Para conocer el paisaje transcriptómico de cada etapa de desarrollo, se realizaron tres estrategias de ensamble *de novo*: 1) a nivel de etapa de desarrollo, 2) a nivel de etapa de desarrollo concatenada y sin transcritos redundantes y 3)

un ensamble global con todas las muestras. Ambos tipos de estrategias se llevaron a cabo utilizando el programa Trinity v2.8.4 (Grabherr *et al.*, 2011). Posteriormente se continuó con un análisis de calidad del ensamble *de novo*, a nivel estadístico y a nivel de similitud con genes conservados en el grupo de los Artrópodos utilizando el programa BUSCO v4.0.5 (Simão *et al.*, 2015).

#### 4.5.2 Anotación de ensamblados *de novo*

Los transcriptomas que se trabajaron fueron los ensamblados a nivel de las etapas de desarrollo, de los cuales para cada ensamble *de novo* se identificaron los marcos de lectura abiertos (Open Reading Frames, ORFs) utilizando el programa TransDecoder v5.3.0 (Brian and Papanicolaou, 2019). Posteriormente se realizó la búsqueda de similitud a nivel de proteína-proteína y transcrito-proteína utilizando BLASTp y BLASTx, respectivamente, v2.7.1 contra el proteoma de *D. melanogaster* descargado de FlyBase vFB2018-06 y con la base de datos no redundante de proteínas de BLAST v2.7.1.

Para realizar la selección de transcritos similares a proteínas ya reportadas se realizó la estimación de abundancia de los transcritos en cada transcriptoma utilizando el programa de kallisto v0.44.0, el cual es un método para cuantificar la abundancia de los transcritos por medio de pseudoalineamientos de las lecturas o los fragmentos, concentrándose en identificar los transcritos de donde las lecturas se originaron de forma rápida (Bray *et al.*, 2016), con un bootstrap de 100. La anotación de los transcritos similares con los reportados en *D. melanogaster* y en la base de datos de BLAST se llevó a cabo por transcritos que mostraron una identidad igual o mayor de 25% con una proteína anotada, y que tuvieran similar número de FPKMs (fragments per kilobase per million reads mapped) en sus tres replicas biológicas (Log2 de sus FPKMs de cada replica  $\pm 1$  de desviación estándar). Los transcritos de las proteínas que mostraron un porcentaje de identidad igual o mayor de 25% de *A. ludens* en comparación con los de *D. melanogaster* fueron tomados en cuenta con un complemento de los *e-values* (menores de  $10^{-20}$ ) ya que es más significativo y razonable como una distancia evolutiva entre dos secuencias (Pearson, 2013). De dichos transcritos que cumplían con este parámetro de identidad se realizó un análisis de anotación funcional y un análisis de ontología de genes utilizando el programa DAVID v6.8 (Huang, Sherman and Lempicki, 2009a, 2009b).

Para los transcritos que no contaron con un valor significativo de similitud con proteínas anotadas de *D. melanogaster*, y las ya reportadas en BLAST, se realizaron los siguientes análisis: i) Identificación de secuencias repetidas previamente publicadas y un modelado *de novo* de familias de secuencias repetidas utilizando el programa RepeatModeler v1.0.11, que consiste en los programas RECON v1.08, RepeatScout v1.0.5, and Tandem Repeats Finder TRF v4.09 (Smit, Hubley and Green, 2019b, 2019a); ii) La determinación de precursores de miRNAs basados en comparaciones de secuencia con los reportados en la base de datos de miRBase v22.1 (Kozomara and Griffiths-Jones, 2011) y los reportados en InsectBase consultado en febrero del 2020 (Yin *et al.*, 2016) y iii) La anotación de dominios de proteínas utilizando el programa HMMER/PFAM v3.2.1.

#### 4.5.3 Análisis de expresión diferencial

El análisis de expresión diferencial a través de las tres etapas de desarrollo se realizó a partir de un concatenado de los transcriptomas de dichas etapas, posteriormente se realizó la remoción de transcritos redundantes utilizando el programa cd-hit v4.8.1 (Li and Godzik, 2006; Fu *et al.*, 2012) con un tamaño de palabras entre 10-11 con una similitud de 95% en ambas cadenas. La estimación de las abundancias de lecturas se realizó para las tres réplicas de cada etapa de desarrollo utilizando el programa kallisto v0.44.0 (Bray *et al.*, 2016) y el análisis de expresión diferencial se realizó utilizando el paquete de edgeR (Robinson, McCarthy and Smyth, 2009; McCarthy, Chen and Smyth, 2012). Para realizar los agrupamientos de los genes que cuentan con abundancias y dinámicas de expresión similares se realizó un agrupamiento por *soft clustering*, el cual es un agrupamiento que toma en cuenta que cada transcrito (n) podría pertenecer a un grupo (k) en el que cada transcrito pertenece al grupo cuyo valor medio es el más cercano. Tomando en cuenta estas características de agrupamiento se utilizó el paquete de Mfuzz con el agrupamiento *soft* (Kumar and Futschik, 2007).

#### 4.5.4 Comparación a lo largo de varias especies de moscas

Para la comparación entre diferentes especies de moscas se utilizaron los datos de transcriptomas de embriones de *B. dorsalis* de 0 a 24 h (SRA SRP007559) y de 0 a 8 h (GEO GSE118472) (Peng *et al.*, 2020), *C. capitata* de 8 a 10 h (SRA SRR1380982)

(Salvemini *et al.*, 2014), y los valores de FPKM reportados para *D. melanogaster* en FlyBase versión FB2018-06. Los datos de las moscas de *B. dorsalis* y *C. capitata* se analizaron como se describe a continuación: el análisis de las lecturas crudas se realizó con el programa FastQC v0.11.7 (Andrews, 2010) y la remoción de secuencias de adaptadores se realizó con el programa Trimmomatic v0.38 (Bolger, Lohse and Usadel, 2014). Los ensamblajes *de novo* se realizaron utilizando el programa Trinity v2.8.4 (Grabherr *et al.*, 2011) con los parámetros predeterminados para cada set de datos de cada mosca; después se continuó con la anotación, primero identificando los ORFs con el programa TransDecoder v5.3.0 (Brian and Papanicolaou, 2019) y se realizó una búsqueda por similitud a nivel proteína-proteína y transcrito-proteína utilizando BLASTp y BLASTx respectivamente, v2.7.1 contra el proteoma de *D. melanogaster* descargado de FlyBase vFB2018-06 y con las proteínas no redundantes reportadas en BLAST v2.7.1. Se realizó la estimación de abundancia de los transcritos en cada transcriptoma utilizando el programa de kallisto v0.44.0 (Bray *et al.*, 2016) con un bootstrap de 100. La anotación de los transcritos similares con los reportados en *D. melanogaster* y en la base de datos de BLAST se llevó a cabo por transcritos que mostraban una identidad igual o mayor de 25% con una proteína anotada, y que tuvieran similar número de FPKMs (fragments per kilobase per million reads mapped) en sus tres replicas biológicas (Log2 de sus FPKMs de cada replica  $\pm 1$  de desviación estándar).

#### 4.5.5 Anotación de proteínas de *A. ludens*

Una vez identificados todos los transcritos que son de un mismo gen, se seleccionaron todas las secuencias de cada transcriptoma de la etapa de desarrollo. Una vez que se tuvieron todas las secuencias se realizó un alineamiento múltiple con el programa Clustal Omega (Sievers *et al.*, 2011) versión enero 2020; teniendo las secuencias alineadas se visualizaron las secuencias con el programa SeaView v4.7 (Galtier, Gouy and Gautier, 1996) para identificar la secuencia consenso. Una vez identificada, se realizó la búsqueda de ORFs utilizando el programa Serial Cloner v2.6 ('Serial Cloner', 2020), tomando como transcrito la secuencia más larga y completa. A partir de esta secuencia se tradujo a aminoácidos utilizando Serial Cloner y se continuó con la búsqueda de dominios proteicos utilizando el programa PFAM v (Finn *et al.*, 2014).

## 4.6 Hibridación *in situ* de embriones de *A. ludens*

### 4.6.1 Generación de sonda de RNA

Las sondas de RNA específicas para *A. ludens* se realizaron a partir del producto de la amplificación del gen blanco (Hodar and Cambiazo, 2018), a la cual por medio de oligonucleótidos se le integra la secuencia del promotor de la RNA polimerasa T7 (5'-GAAATAATACGACTCACTATAGGGAGA-3') (ver Anexo 3). Los transcritos seleccionados fueron *exuperantia*, *hunchback*, *even-skipped* y *engrailed*; todos basados en secuencias de *A. ludens* para la construcción de las sondas (revisar Anexo 4 para conocer la secuencia de cada gen, así como la región en donde se hibrida la sonda de cada gen). Una vez que se contó con el producto de amplificación del gen de interés, se realizó la purificación y se procedió con la transcripción *in vitro* del transcrito, el cual se marcó con UTPs marcadas con digoxigenina (Anexo 1).

### 4.6.2 Comprobación de generación y marcaje de sonda de RNA

Para comprobar la generación de la sonda se realizó la transferencia de la sonda marcada y posteriormente se realizó la purificación de la sonda por el kit NucAway Spin Columns (Invitrogen, AM10070) y se cuantificó por nanodrop. Para conocer qué tan marcada se encuentra la sonda se realizó un dot blot a diferentes concentraciones de la sonda.

#### 4.6.2.1 Transferencia de sonda a partir de gel de agarosa

A partir del producto de la transcripción *in vitro* se reservó 1  $\mu$ L, al que se adicionaron 5  $\mu$ L de búfer de carga para RNA y se incubó a 65 °C por 10 minutos. De manera simultánea, se realizó un gel de agarosa al 2% en búfer MOPS 1x, sin formaldehído. Después del tiempo de incubación, la muestra se pasó a hielo y se adicionó 1  $\mu$ L de bromuro de etidio 5 mg/mL (dilución 1:100); el gel se corre por 45 min a 70 V en búfer MOPS 1x.

Para la transferencia del gel se utilizó una membrana para ácidos nucleicos (Anexo 1). La solución utilizada para la transferencia es solución SSC 20x (100 mL) para incubar toda la noche (aproximadamente 16 h). Una vez pasada la incubación, se retiró la membrana y se fijó por crosslinking UV, en el cual se utiliza un pulso de autocrosslinking, por ambos lados de la membrana. Una vez terminado el crosslinking se realizó una serie

de lavados en agitación: lavado DIG1 por 15 min, lavado de DIG1 + 0.5% de leche en polvo por 30 min, lavado de DIG1 más anticuerpo anti-digoxigenina (Roche, 11093274910) dilución 1:5000 por 30 min, lavado de DIG1 por 10 min, lavado de DIG1 por 10 min y lavado de DIG3 por 5 min. La membrana se reveló en obscuridad con la solución NBT/BCIP por el tiempo necesario.

#### 4.6.2.2 Dot blot

Las concentraciones utilizadas para el Dot blot fueron las siguientes: 1 (1 ug/ $\mu$ L), 1:10 (100 pg/ $\mu$ L), 1:100 (10 pg/ $\mu$ L), 1:1 000 (1 pg/ $\mu$ L), 1:10 000 (0.1 pg/ $\mu$ L) y 1:100 000 (0.01 pg/ $\mu$ L). Las muestras se prepararon con 1  $\mu$ L de cada dilución con 3 volúmenes de solución de incubación de RNA (Anexo 1), incubadas a 65 °C por 5 min. Posteriormente se adicionó 1 volumen de SSC 20X.

La membrana para ácidos nucleicos se lavó con agua destilada estéril por 5 minutos, después se lavó con SSC 10x por 10 min y se dejó secar sobre un papel filtro limpio. Luego se cargó la membrana con 2  $\mu$ L de cada dilución dejando secar entre muestras, se fijó por crosslinking UV, en la cual se utilizó un pulso de autocrosslinking, por ambos lados de la membrana. Una vez realizado el crosslinking se realizaron una serie de lavados en agitación: lavado DIG1 por 15 min, lavado de DIG1 + 0.5% de leche en polvo por 30 min, lavado de DIG1 más anticuerpo anti-digoxigenina dilución 1:5000 por 30 min, lavado de DIG1 por 10 min, lavado de DIG1 por 10 min y lavado de DIG3 por 5 min. La membrana se reveló en obscuridad con la solución NBT/BCIP por el tiempo necesario.

#### 4.6.3. Hibridación *in situ* de embriones de *A. ludens*

Para realizar la hibridación *in situ* las colectas de *Anastrepha* se llevaron a cabo como se menciona en el punto 4.2, los embriones de *D. melanogaster* se colectaron en un periodo de 0 a 2 h, 2 a 4 h y 0 a 17 h (Rothwell and Sullivan, 2000). Los embriones de *D. melanogaster* se utilizaron para tener un control positivo de las sondas que a nivel de secuencia son muy similares, es decir, las sondas de los genes *hunchback* y *engrailed*; así como tener un control positivo de las soluciones, la incubación con proteínasa K y el revelado de los embriones ya que se encuentra muy establecido el protocolo de *whole mount* hibridación *in situ* en embriones de *D. melanogaster* (Anexo 8).

Una vez que se colectaron los embriones de *A. ludens* se descorionizaron con 6.15% de cloro por 3 min en agitación, después se fijaron con 4% formaldehído por 20 min en agitación y se realizaron tres lavados con 2 mL de metanol en agitación. Los siguientes pasos son modificaciones del protocolo reportado por Tautz *et al.*, (Tautz and Pfeifle, 1989), como sigue: se continuó con una segunda fijación con 1 mL de 4% de paraformaldehído en PBS y 3 mL de heptano por 20 min en agitación, luego se realizaron dos lavados con 3 mL de metanol y un lavado con 3 mL de etanol. Se realizó un clareo del tejido con xileno, luego un lavado con etanol:xileno 1:1 v/v por 10 min, un lavado con xileno por 3 h, un lavado con etanol:xileno 1:1 v/v por 10 min, un lavado con etanol:xileno 2:1 v/v por 10 min, un lavado con etanol:xileno 3:1 v/v por 10 min y tres lavados con etanol por 10 min.

Los siguientes pasos se realizaron en tubos de 1.5 mL en agitación en un roscicero. Se continuó con un lavado en metanol por 10 min, un lavado con metanol:PBT 1:1 v/v por 10 min y, una tercera fijación con 4% paraformaldehído en PBT 1x por 20 min. Después se realizaron 3 lavados con PBT 1x por 10 min, posteriormente se incubó con proteinasa K 50 µg/mL en PBT 1x a 55 °C en agitación por 3 horas; pasado el tiempo de incubación se realizaron dos lavados con glicina 2 mg/mL en PBT por 5 min y con dos lavados con PBT 1x por 10 min. Siguió un lavado con PBT 1x por 10 min, una fijación con 4% paraformaldehído en PBT 1x por 20 min y cuatro lavados con PBT 1x por 10 min. Se comenzó con la prehibridación con un lavado con PBT:búfer ISH 1:1 v/v por 5 min, un lavado con búfer ISH por 5 min y una incubación a 45 °C con un mínimo de 1 hora en búfer ISH en agitación leve. La hibridación se realizó sin agitación a 55 °C con búfer ISH y con una concentración de 100 µg de sonda a utilizar por 14 a 15 h.

Después de la incubación con la sonda, se realizaron lavados en un roscicero a la misma temperatura de hibridación, comenzando con un lavado con búfer de ISH por 20 min, un lavado con búfer de ISH:PBT 4:1 v/v por 20 min, un lavado con búfer de ISH:PBT 3:2 v/v por 20 min, un lavado con búfer de ISH:PBT 2:3 v/v por 20 min y un lavado con búfer de ISH:PBT 1:4 v/v por 20 min. Los siguientes lavados se realizaron a temperatura ambiente, dos lavados con PBT por 20 min, una incubación con 1:2000 de anticuerpo anti-digoxigenina en PBT por 1 h, seguido de cuatro lavados con PBT 1x por 20 min, 3 lavados con la solución de detección por 5 min e incubar con la solución de NBT/BCIP, la cual puede ser desde minutos hasta toda la noche, sin agitación y a temperatura ambiente. Para mejores resultados del revelado de la sonda se recomienda realizar el revelado a 4 °C.



#### **4.7. Toma y análisis de imágenes**

La toma de imágenes de las hibridaciones *in situ* se llevaron a cabo en el microscopio Nikon ECLIPSE E600 y el análisis se realizó utilizando el programa de Fiji v1.52.p (Schindelin *et al.*, 2012).

## Capítulo 5. Resultados

### 5.1 Selección y generación de datos transcriptómicos de *A. ludens*

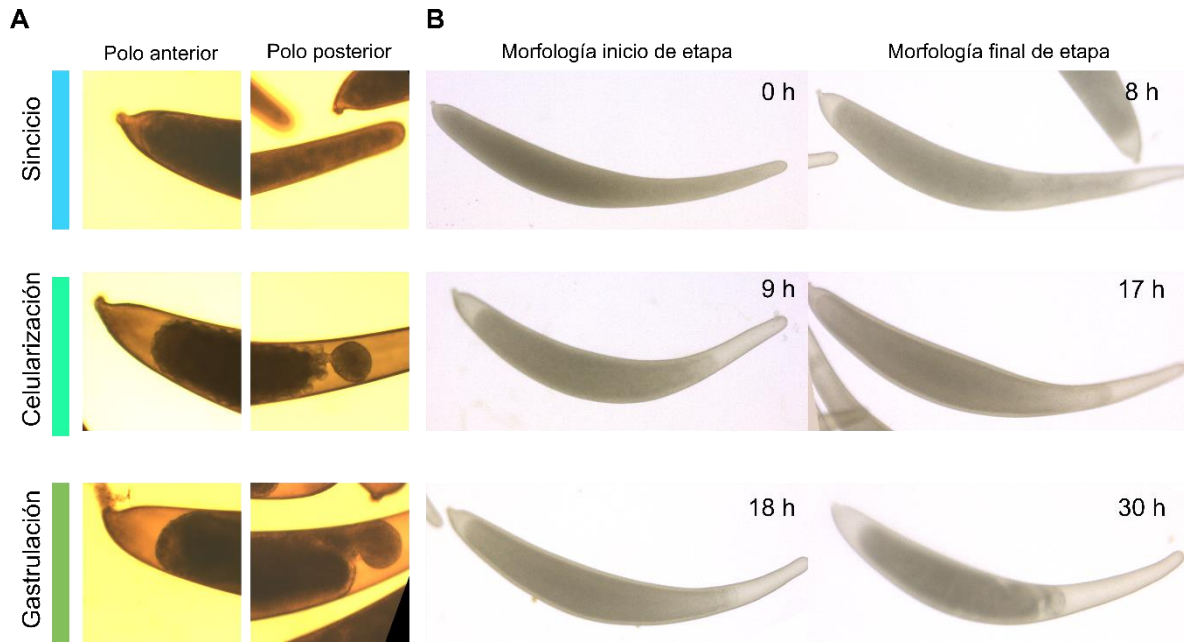
En esta sección se describe la selección de las etapas de desarrollo utilizadas en los experimentos de RNA-seq, así como la estandarización de la extracción de RNA total de embriones.

#### 5.1.1 Selección de etapas de embriogénesis de *A. ludens*

En observaciones previas se observó que el desarrollo embrionario de *A. ludens* transcurre a lo largo de 120 h a 25 °C, periodo que se puede dividir en desarrollo temprano y el desarrollo tardío, que abarca de las 0 a 30 h y 30 a 120 h respectivamente. Mediante las observaciones realizadas con un microscopio de disección, se identificaron las etapas del desarrollo embrionario temprano, que son: sincicio, celularización y gastrulación (Gutiérrez-Ramos, 2016). Las observaciones de embriones se llevaron a cabo a nivel individual en función de las características morfológicas descritas en la Tabla 5.1 y Figura 5.1.

**Tabla 5.1.** Descripción de las características morfológicas de las etapas de sincicio, celularización y gastrulación en embriones de *A. ludens*.

<b>Etapas de desarrollo</b>	<b>Características morfológicas</b>
<p>Sincicio</p> <p>Tiempo: 0 a 9 h</p>	<p>Una vez ovipositado el huevo fecundado, la yema se encuentra distribuida a lo largo de todo el huevo. Conforme pasan las primeras 3 h de desarrollo se observa una retracción de la yema en la parte anterior, permitiendo distinguir la yema del corión. Después de pasadas las 6 h la retracción de la yema se observa en ambos polos, anterior y posterior. Los cambios observados entre las 6-9 h de desarrollo son principalmente en la parte posterior. Una vez que se retrae la yema de la parte posterior se comienza a ver en el último cuarto de la periferia del embrión pequeñas invaginaciones.</p>
<p>Celularización</p> <p>Tiempo: 9 a 18 h</p>	<p>El cambio más importante en esta etapa es que la punta posterior sufre una extrusión, es decir, se separa la punta de la yema del embrión; las invaginaciones se observan más marcadas en la periferia de la parte posterior del embrión. Conforme avanza el desarrollo del embrión, a las 12 h, las invaginaciones se comienzan a observar en toda la periferia del embrión. Este proceso de invaginación a lo largo del embrión permite observar un halo de tejido menos denso que la parte media del embrión, siendo éste un anillo de células en citocinesis.</p>
<p>Gastrulación</p> <p>Tiempo: 18 a 30 h</p>	<p>Pasadas las 18 h de desarrollo se observa el anillo de células en toda la periferia del embrión, así como el tejido residual de la extrusión. A las 20 h, en la parte posterior del embrión, se observa un grupo de células menos densas, las cuales conforme pasan las horas comienzan a migrar a la parte ventral del embrión. También se visualizan los primeros surcos de escisión en la parte dorsal del embrión, uno muy marcado se encuentra en la parte dorsal-anterior el cual es una invaginación que distingue la parte cefálica del embrión. En esta etapa ocurre la segunda extrusión polar, que se distingue del grupo de células que van a migrar por su apariencia más blanquecina y menos densa.</p>



**Figura 5.1.** Características de los embriones tempranos de *A. ludens*. En A, se observan las características de los embriones en la parte posterior y anterior del huevo. En B, se observa el huevo en la primera y la última hora de cada etapa de desarrollo seleccionada.

Mediante nuevas observaciones se confirmó que las etapas que se clasifican en embriogénesis temprana ocurren en un lapso de 0 a 30 horas, desde la fecundación hasta la gastrulación, en donde el sincicio es de las 0 a 9 h, celularización de 9 a 18 h y gastrulación de 18 a 30 h. Los embriones clasificados según estas características morfológicas se utilizaron para los experimentos de RNA-seq y las hibridaciones *in situ*.

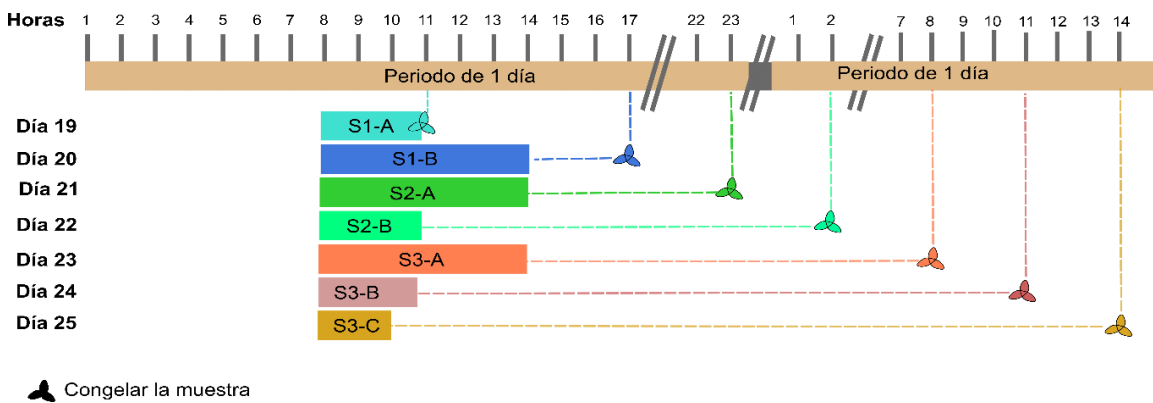
### 5.1.2 Estandarización del protocolo para la extracción de RNA total de embriones de *A. ludens*

Debido a que *A. ludens* no es considerado un insecto modelo de estudio, fuera de estudios ecológicos y manejo de plaga, la estandarización de los protocolos para poder estudiarla es muy importante. Teniendo en cuenta esto, se realizó la estandarización de la extracción de ARN total de embriones tomando como referencia los protocolos descritos para *D. melanogaster* (Rothwell and Sullivan, 2000). El molido de los embriones para realizar la extracción de RNA total es un paso crítico para obtener una cantidad con integridad adecuada de RNA. El protocolo utilizado para la extracción de RNA total es el protocolo

estándar de TRI Reagent (Sigma, 93289). El primer paso consiste en realizar pruebas con embriones en diferentes condiciones: embriones con corión, embriones descoronizados, embriones molidos con pistilo de plástico y embriones congelados molidos con pistilo de plástico; en estas cuatro pruebas el resultado predominante es que no se logra realizar un molido completo de los embriones, ya que la mayoría se encuentra íntegra y no logra romperse, por lo que la recuperación de RNA es poca o nula.

Considerando estos resultados se realizó una prueba con embriones congelados molidos con perlas de tungsteno manteniendo las condiciones de congelación, una forma que se utiliza para pulverizar tejidos difíciles de moler, como es el caso de semillas de plantas, dientes y huesos. La prueba con las perlas de tungsteno resultó la más eficaz para moler el tejido, ya que no es necesario realizar algún tratamiento a los embriones, es decir descoronizarlos, por lo que se ahorra tiempo. Una vez establecido el método de molido del tejido se realizó el protocolo descrito de TRI Reagent, en que se obtuvo un promedio de 26.53 µg de RNA total de 500 embriones. Con esta relación de cantidad de embriones y RNA recuperado podemos realizar estimaciones de cuánto tejido coleccionar para realizar la secuenciación de las muestras.

Una de las características del periodo de desarrollo temprano de *A. ludens* es que comprende muchas horas, por lo que se realizaron colectas a lo largo de 3 h para contar con una representatividad de todas las horas de desarrollo de cada una de las etapas, como se muestra en la Tabla 5.2. Se realizaron 2 colectas para la etapa de sincicio, 2 colectas para la etapa de celularización y 3 colectas para la etapa de gastrulación (Figura 5.2).



**Figura 5.2.** Esquema para la colecta de embriones para la generación de datos de RNA-seq. Entre el día 19 y 25 de las colectas para los experimentos de RNA-seq (ver Figura 4.1) las colectas de los

embriones se realizaron en un periodo de las 8 a las 14 horas; para después dejar crecer los embriones en agua hasta el periodo de sincicio, celularización y gastrulación. Cada día se realiza la colecta de los embriones tomando en consideración una subetapa de cada etapa de desarrollo y se dejan crecer hasta que cumpla las horas que se describen en la Tabla 5.2.

**Tabla 5.2.** Relación de colectas entre las diferentes etapas de desarrollo temprano de *A. ludens*. De las etapas seleccionadas se realizaron subcolectas a tiempos establecidos para contar con una representación total de cada etapa de desarrollo.

Etapa	Subetapa	Horas de colecta	Número de muestras	ARN total ( $\mu\text{g}$ )
Sincicio	S1-A	0 h a 3 h	10	22.368
	S1-B	3 h a 9 h	10	20.964
Celularización	S2-A	9 h a 15 h	14	29.178
	S2-B	15 h a 18 h	10	22.743
Gastrulación	S3-A	18 h a 24 h	11	20.363
	S3-B	24 h a 27 h	10	22.046
	S3-C	27 h a 30 h	10	24.621

## 5.2 Generación y análisis de los datos transcriptómicos de *A. ludens*

Una vez que se contó con el RNA de cada subetapa de desarrollo embrionario de *A. ludens*, se realizó una mezcla para tener 10 mg de RNA total de cada etapa. Es decir, de cada subetapa se mezclan cantidades equimolares del RNA total para cada etapa de desarrollo, sincicio, celularización y gastrulación. La generación de las librerías y la secuenciación se realizó a partir de 10  $\mu\text{g}$  de RNA total de cada etapa de desarrollo, con sus respectivas réplicas. El procesamiento de las muestras se realizó bajo el servicio de Beijing Genomic Institute (BGI) Genomic Services.

### 5.2.1 Generación de librerías y secuenciación de muestras de *A. ludens*

Los experimentos se llevaron a cabo por triplicado de réplica biológica de cada etapa de desarrollo. Antes de proceder a la construcción de las librerías se realizó una prueba de

calidad del RNA de cada etapa de desarrollo, como se observa en la Tabla 5.3, en la cual se muestra la medición de la integridad del RNA, por sus siglas en inglés RIN (RNA Integrity Number), mediante un algoritmo que se aplica a las medidas electroforéticas del RNA que muestra qué tan intacta está la muestra, mostrando un parámetro de contaminación de DNA (Mueller, Lightfoot and Schroeder, 2004).

**Tabla 5.3.** Calidad de las muestras de RNA utilizadas para la construcción de librerías. Rep: representa réplicas de cada etapa de desarrollo.

Muestra	Nombre de muestra	Concentración (µg/uL)	Masa total (µg)	RIN	Nivel de calidad
1	S1-Rep1	240	7.20	6.7	A
2	S2-Rep1	170	5.10	6.7	A
3	S3-Rep1	304	9.12	6.7	A
4	S1-Rep2	284	8.52	6.8	A
5	S2-Rep2	284	8.52	6.7	A
6	S3-Rep2	194	5.82	6.7	A
7	S1-Rep3	228	6.84	6.3	A
8	S2-Rep3	322	9.66	6.9	A
9	S3-Rep3	342	10.26	7.1	A

El valor de integridad de RNA, RIN, de todas las muestras utilizadas corresponde al nivel A, es decir, el valor de contaminación de la muestra es mínimo y la cantidad es suficiente para realizar la construcción de dos librerías o más. Una vez obtenidos estos resultados, se continuó con la generación de una librería de ARN poliadenilado y la secuenciación de las muestras, utilizando la plataforma de Illumina HiSeq 4000 pareada de 2x100. El resultado de la secuenciación es de 58 a 106 millones de lecturas, teniendo un contenido de GC uniforme entre las distintas etapas de desarrollo y un valor de Q20 muy alto (Tabla 5.4).

**Tabla 5.4.** Resultados globales de cada librería secuenciada.

<b>Muestra</b>	<b>Lecturas limpias</b>	<b>Bases limpias</b>	<b>Longitud (pb)</b>	<b>Q20 (%)</b>	<b>Contenido GC (%)</b>
S1-Rep1	63,785,092	6,378,509,200	100	98.98	41.26
S2-Rep1	61,207,028	6,120,702,800	100	98.98	41.31
S3-Rep1	106,630,524	10,663,052,400	100	98.99	41.36
S1-Rep2	68,550,378	6,855,037,800	100	98.88	40.94
S2-Rep2	71,129,892	7,112,989,200	100	98.99	41.12
S3-Rep2	61,171,722	6,117,172,200	100	98.97	41.25
S1-Rep3	74,316,760	7,431,676,000	100	99.02	41.10
S2-Rep3	58,873,552	5,887,355,200	100	98.82	41.40
S3-Rep3	78,994,832	7,899,483,200	100	98.97	41.34

### 5.2.2 Análisis de calidad de secuencias

Una vez observado el análisis preliminar de la secuenciación, que se muestra en la Tabla 5.4, se continuó analizando la calidad de las secuencias; primero a nivel de secuencia generada y después el análisis por réplicas de cada etapa de desarrollo. Para el análisis de cada secuencia generada se contó con 18 archivos utilizando el programa de FastQC (Andrews, 2010); que genera una serie de gráficos de la distribución de cada nucleótido de las lecturas obtenidas, el porcentaje de GC, el contenido de secuencias adaptadoras para la secuenciación, entre otros datos. En la Tabla 5.5 se muestra un resumen de los datos.

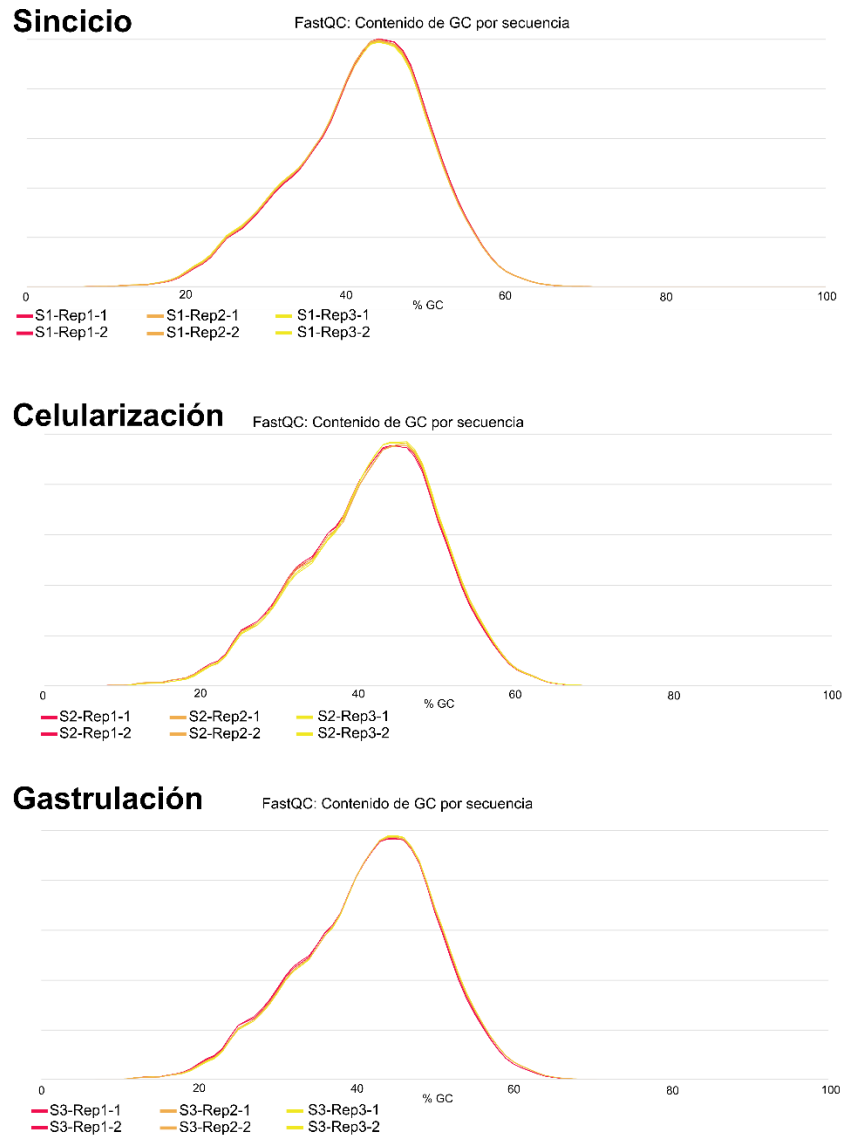
Como todas las lecturas de los 18 archivos generados tienen valores de calidad buenos, se continuó con la revisión de la calidad de las muestras que corresponden a cada etapa de sincicio, celularización y gastrulación. Utilizando el programa MultiQC (Ewels *et al.*, 2016) se analizó la calidad de las secuencias que conforman cada etapa de desarrollo como se observa en la Figura 5.2. En dicha figura se observa que el comportamiento de las tres réplicas generadas para cada etapa de desarrollo se distribuye de manera muy similar a la posición de sus bases; se observa también que el porcentaje de GC entre muestras por etapa se empalman por lo que se puede considerar como réplicas de cada etapa. Este análisis, posterior a la generación de las secuencias, es muy importante; ya que nos permite conocer que nuestras librerías son de alta calidad, las réplicas son similares en cada etapa



y que cada etapa cuenta con un comportamiento único, en el caso de su distribución de las bases.

**Tabla 5.5.** Resumen del análisis de calidad de cada librería generada. Los datos mostrados fueron generados por el programa de FastQC.

<b>Muestra</b>		<b>Secuencias totales</b>	<b>Contenido de GC (%)</b>	<b>Contenido de adaptadores</b>
S1-Rep1	Forward	31,892,546	41	No
	Reverse	31,892,546	41	No
S2-Rep1	Forward	30,603,514	41	No
	Reverse	30,603,514	41	No
S3-Rep1	Forward	53,315,262	41	No
	Reverse	53,315,262	41	No
S1-Rep2	Forward	34,275,189	40	No
	Reverse	34,275,189	40	No
S2-Rep2	Forward	35,564,946	41	No
	Reverse	35,564,946	41	No
S3-Rep2	Forward	30,585,861	41	No
	Reverse	30,585,861	41	No
S1-Rep3	Forward	37,158,380	41	No
	Reverse	37,158,380	41	No
S2-Rep3	Forward	29,436,776	41	No
	Reverse	29,436,776	41	No
S3-Rep3	Forward	39,497,416	41	No
	Reverse	39,497,416	41	No



**Figura 5.3.** Diagramas del contenido de GC por secuencia. Gráfica de la distribución del contenido de GC de las lecturas generadas de cada librería. Las tendencias observadas nos permiten identificar el comportamiento entre las réplicas, si es similar se concluye que las réplicas biológicas entre cada etapa son robustas.

### 5.2.3 Ensamble *de novo* de los datos transcriptómicos

Existen diversas estrategias para realizar el ensamble *de novo* en organismos no modelo, por lo que la pregunta biológica y la experiencia del investigador son importantes para elegir cuál es la más adecuada para el estudio que se está realizando. Las estrategias propuestas son: i) mezcla de todos los datos, a pesar de que provengan de diferentes etapas de

desarrollo; ii) ensamble individual de cada etapa de desarrollo; iii) utilizar un genoma de referencia de la especie modelo más cercana y mapear los transcritos y iv) ensamble global de todas las etapas de desarrollo y después mapearlo en el genoma de la especie modelo más cercana (Ockendon *et al.*, 2016; Ungaro *et al.*, 2017). En el presente caso se utilizaron tres estrategias, que se describen a continuación:

- I. **Ensamble *de novo* por etapa de desarrollo:** se realizó un ensamble por cada etapa de desarrollo, utilizando las tres réplicas de cada etapa. El resultado obtenido fueron tres transcriptomas: transcriptoma de sincicio, transcriptoma de celularización y transcriptoma de gastrulación.
- II. **Ensamble *de novo* global sin redundancias:** a partir de los tres transcriptomas obtenidos (transcriptoma de sincicio, transcriptoma de celularización y transcriptoma de gastrulación) se realizó un concatenado de todos los datos y, posteriormente, una remoción de los transcritos redundantes presentes en los tres transcriptomas; quedando solo transcritos comunes en los tres transcriptomas y los transcritos únicos de cada etapa de desarrollo.
- III. **Ensamble *de novo* global de todas las etapas:** se realizó un ensamble con la unión de las 9 librerías, es decir, todas las etapas con sus réplicas, esto permitió contar con una gran cantidad de datos para su ensamblaje.

Cada estrategia utilizada tiene sus ventajas y desventajas, en el caso de I) la estrategia por etapa de desarrollo, la ventaja que tenemos es que podríamos encontrar transcritos únicos de cada etapa de desarrollo, la desventaja es que la cantidad de datos es mayor ya que se cuenta con tres transcriptomas; II) la ventaja es que ahora solo tienes un conjunto de datos transcripcionales que son transcritos robustos ya que quitas la redundancia de que exista el mismo transcrito en todas las etapas de desarrollo, pero la desventaja es que se pierden aquellos transcritos que cuentan con isoformas únicas de cada etapa de desarrollo y II) la ventaja es que cuentas con un universo de datos mayor por lo que es una estrategia que podría servir para contar con un transcriptoma de referencia, ya que tendría mejores valores de calidad (como se puede observar en la Tabla 5.6 y 5.7).

Para obtener información de cada transcriptoma se realizaron dos análisis de calidad del ensamble; que son medidas estadísticas de identificación de genes clave evolutivamente conservados. En el caso de las medidas estadísticas de los transcriptomas se hizo la cuantificación de los transcritos, la media y mediana del tamaño de los transcritos

y la N50, que es una medida que denota la longitud del transcrito que cubre un 50% del ensamble. En los ensambles por etapa se observó que más del 95% de los transcritos ensamblados tiene más de 300 pb (pares de bases) de longitud; sin embargo, cuando se realizó el ensamble de forma global dicho porcentaje disminuyó a un 60%. A pesar de estos detalles, los datos cuentan con un valor de N50 adecuado para el tamaño total del transcriptoma, como se puede observar en la Tabla 5.6.

**Tabla 5.6.** Métricas estadísticas de los transcriptomas ensamblados *de novo*. En donde Trans es transcritos, pb es pares de bases.

Muestras	Ensamble de novo						
	# de trans	Mediana del tamaño (pb)	Media del tamaño (pb)	Transcritos ≥ 300 pb	Transcritos ≥1000 pb	N50	Tamaño total del transcriptoma (pb)
Sincicio	102,771	538	1,397.5	74,439	102,763	3,148	143,622,249
Celularización	111,012	500	1,344.7	79,220	111,003	3,129	149,208,196
Gastrulación	102,451	562	1,436.5	74,849	102,442	3,214	147,174,131
Global no redundantes	130,752	358	906.4	80,788	27,152	2,270	118,509,361
Global todas etapas	181,888	437	1,182.1	125,003	54,348	2,887	215,007,531

Para la identificación de transcritos evolutivamente conservados en los transcriptomas ensamblados, se utilizó una base de datos de transcritos esenciales conservados en el filo de los Artrópodos (1,013 genes). En esta identificación se encontró que en el caso de los transcritos ensamblados por etapas de desarrollo los transcritos completos son más de un 90% del total; presentando en la etapa de gastrulación hasta un 94%. En el caso del ensamble global el porcentaje de transcritos completos es de 95.3%, siendo el ensamble donde hay menos transcritos que no llegan a encontrarse (no reportados) (Tabla 5.7).

Se puede concluir que las tres estrategias utilizadas son buenas, ya que se obtuvieron valores estadísticos de alta calidad en el ensamblado con nivel funcional y se logró detectar más del 90% de transcritos evolutivamente conservados en el filo de los

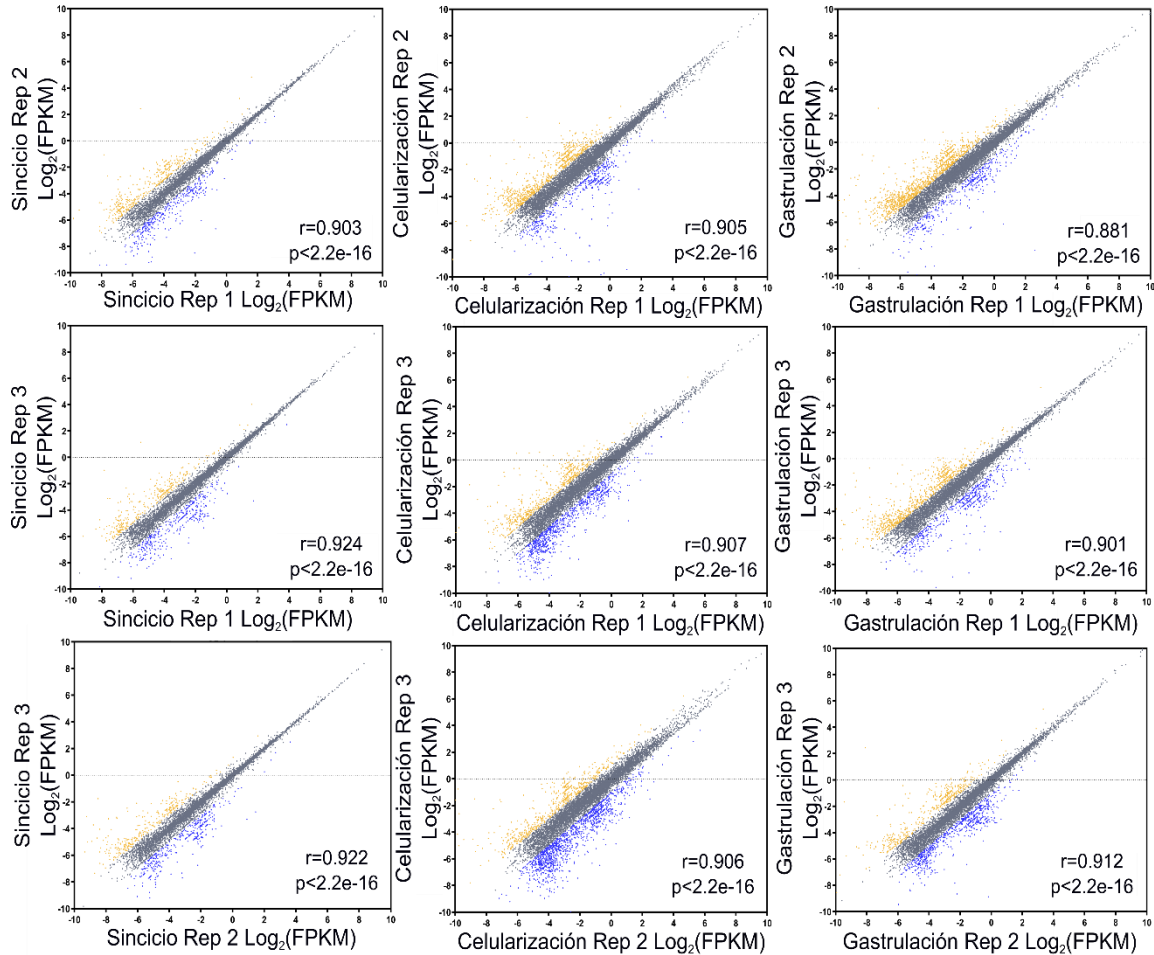
Artrópodos. De acuerdo con estos resultados, el uso de cualquiera de los ensambles es adecuado; sin embargo, se decidió utilizar los ensambles por etapas de desarrollo ya que con la calidad encontrada los datos se pueden considerar altamente representativos, esperando encontrar los transcritos que son únicos para cada etapa de desarrollo.

**Tabla 5.7.** Resumen del análisis comparativo de los transcritos evolutivamente conservados en Artrópodos.

Muestras	Transcritos evolutivamente conservados en Artrópodos				
	Completo sencillos	Completo duplicados	Fragmentados	Perdidos	Total
Sincicio	4.7% (48)	88.7% (899)	0.3% (3)	6.3% (63)	100% (1013)
Celularización	5.2% (53)	88.7% (899)	0.8% (8)	5.3% (53)	100% (1013)
Gastrulación	4.9% (50)	89.0% (902)	0.4% (4)	5.7% (57)	100% (1013)
Global no redundantes	45.1% (457)	43.2% (438)	0.6% (6)	11.1% (112)	100% (1013)
Global todas etapas	5.3% (54)	90.4% (916)	0.3% (3)	4.0% (40)	100% (1013)

#### 5.2.4 Anotación de transcriptomas

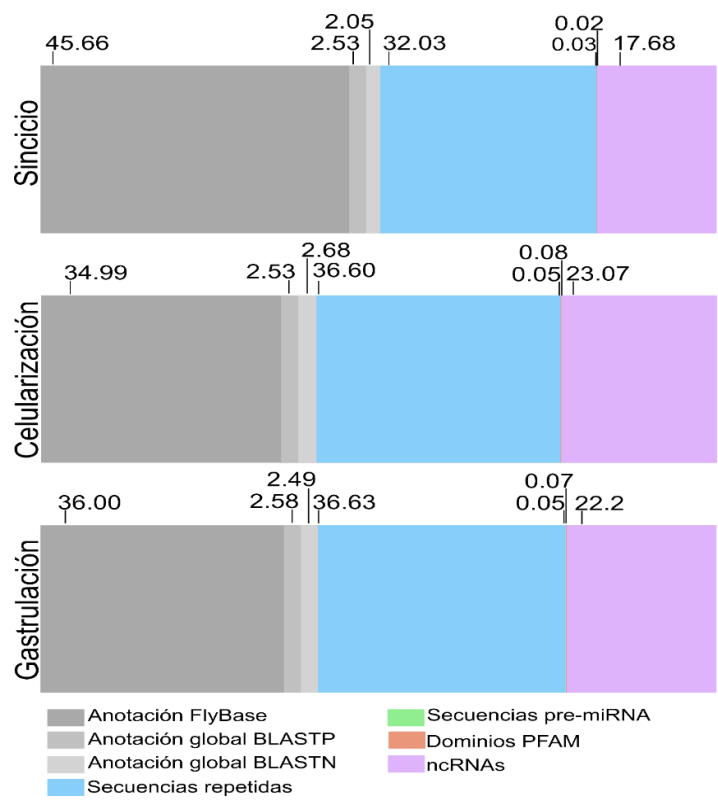
Una vez seleccionados los transcriptomas con los que se va a trabajar, las dos aproximaciones utilizadas para la anotación de cada transcriptoma por etapa son los datos de los transcritos y proteínas de *D. melanogaster* y la base de datos de NCBI. La selección de los transcritos robustos no anotados, más los anotados se realizó considerando la presencia de cada transcrito en por lo menos dos de las tres réplicas de cada etapa. La correlación de Pearson entre las etapas de desarrollo fue muy buena, considerando a 1 como similitud del 100% entre dos réplicas, ya que todas las etapas cuentan con correlaciones entre 0.924 y 0.88. En las etapas de sincicio y celularización todas las réplicas correlacionan arriba de un 0.9, sin embargo, la réplica 1 de la etapa de gastrulación cuenta con una mayor dispersión ( $r^2=0.8$ ) (Figura 5.3).



**Figura 5.4.** Correlación de similitud entre transcritos presentes en las réplicas de cada etapa de desarrollo. Para conocer la similitud entre las réplicas de cada etapa de desarrollo se grafican los FPKMs Log<sub>2</sub> de cada transcrito anotado, los transcritos representados en color gris son los que tienen una desviación estándar entre  $\pm 1$ , en amarillo los  $> 1$  y en azul  $< 1$ . El valor de la correlación de Pearson se muestra como r.

Los transcritos anotados por estas dos bases de datos fueron hasta un 40% en celularización y gastrulación y hasta un 50% en sincicio, es decir, que por lo menos la mitad de cada transcriptoma se anota, como se aprecia en la Figura 5.4. Parte de los análisis para identificar secuencias repetidas, precursores de miRNAs y dominios de proteínas fueron para contar con la mayor cantidad de transcritos identificados. En los tres transcriptomas se lograron identificar secuencias de repetidos anteriormente reportadas (Anexo 5) y secuencias repetidas modeladas *de novo* (Figura 5.4, color azul) encontrándose hasta un

tercio de su total de transcritos. En el caso de las secuencias de precursores de miRNAs identificadas (Anexo 6) se logró identificar la secuencia madura de cada miRNA, y se encontró que casi todos cuentan con una función previamente reportada, como en el caso del miR-9/79 que participa en la regulación de la neurogénesis (Coolen, Katz and Bally-Cuif, 2013), o como el caso de mir-14 que participa en el metabolismo de *D. melanogaster* (Xu *et al.*, 2003). En el caso de los dominios que se encontraron similares a la base de datos de PFAM se observa que en sincicio se encuentran, en celularización y en gastrulación (Anexo 7). Estos dominios que son muy abundantes no muestran una relevancia biológica para los fines de nuestros resultados. Sin embargo, entre 17-21% de secuencias no se lograron identificar y se denominaron ncRNAs, los cuales corresponden a secuencias que pueden ser errores de ensamblado, secuencias con características de RNAs largos no codificantes (lncRNAs), entre otros.



**Figura 5.5.** Esquema del porcentaje de transcritos anotados de las etapas de desarrollo embrionario de *A. ludens*.

De los transcritos anotados con las bases de datos de *D. melanogaster* y NCBI se realizaron análisis de ontología de genes, para conocer la clasificación de dichos transcritos en relación con procesos biológicos, participación en componentes celulares y función molecular (Figura 5.5). Los procesos biológicos más representados fueron neurogénesis, transcripción, reparación de DNA y respuesta celular a inanición y, en el caso de las funciones moleculares, fueron unión a proteínas, unión a ATP y unión a RNA. Todos estos procesos y funciones moleculares son esperadas debido a la serie de procesos llevados a cabo en la embriogénesis temprana.

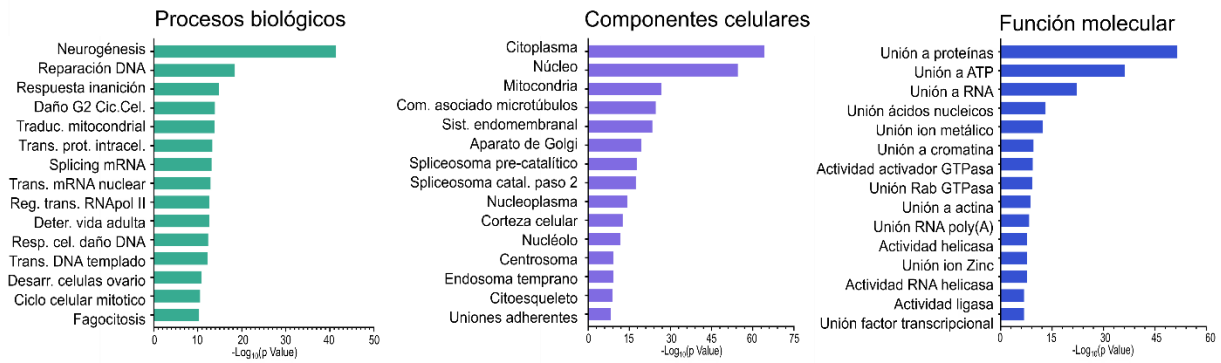
Adicionalmente, se identificaron transcritos que codifican 488 genes de los que no está reportada su expresión en el embrión temprano de *D. melanogaster*. La mayoría de estos transcritos codifican proteínas que cuentan con motivos o dominios que son similares a otras reportadas en animales, como humano y ratón. Las funciones descritas en los otros organismos son en vías de transducción de señales, metabolismo y desarrollo. Estos resultados de los transcriptomas anotados por etapa significan un robusto análisis global de expresión de la embriogénesis temprana de *A. ludens*.

### **5.3 Dinámicas de expresión entre las etapas de desarrollo embrionario temprano de *A. ludens***

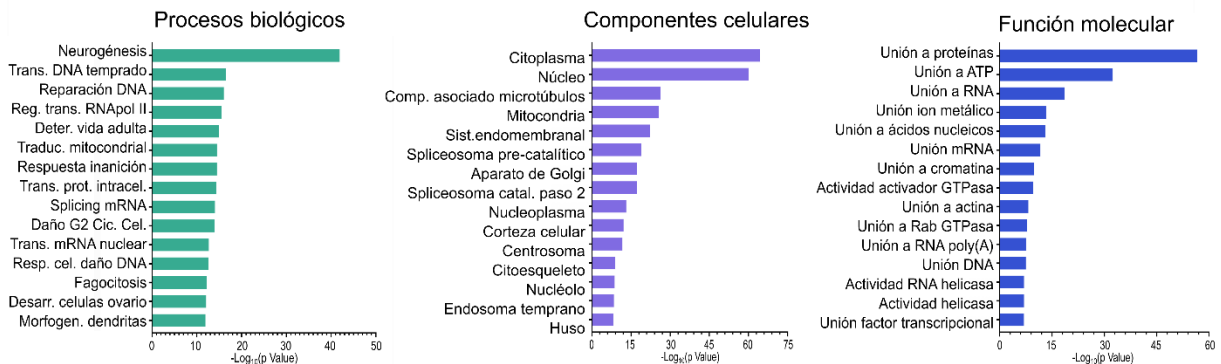
Una vez que se contó con los transcriptomas anotados se realizaron las comparaciones de las dinámicas de expresión entre las etapas de desarrollo embrionario temprano de *A. ludens*. La razón principal para la selección de estas etapas es que en el blastodermo sincicial los transcritos depositados por la madre participan en los primeros procesos embrionarios; la activación del genoma del cigoto ocurre antes de la celularización y el establecimiento de las capas germinales ocurre durante la gastrulación (Lott *et al.*, 2011).



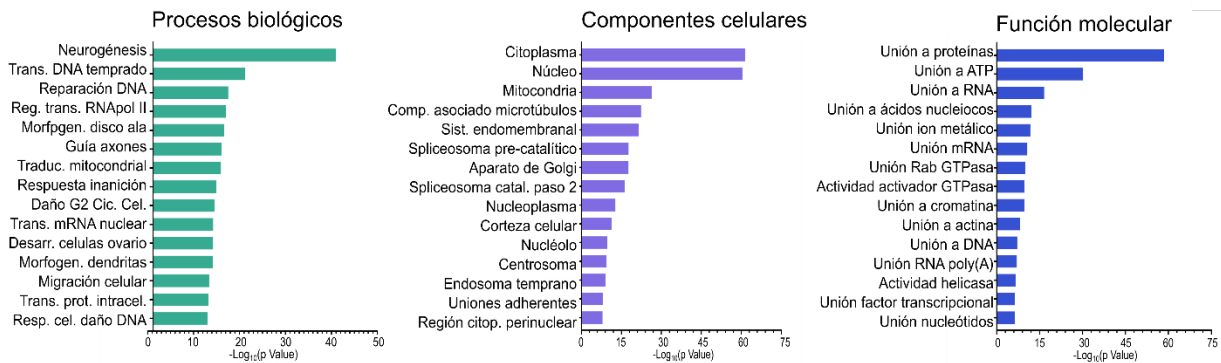
## Sincicio



## Celularización



## Gastrulación

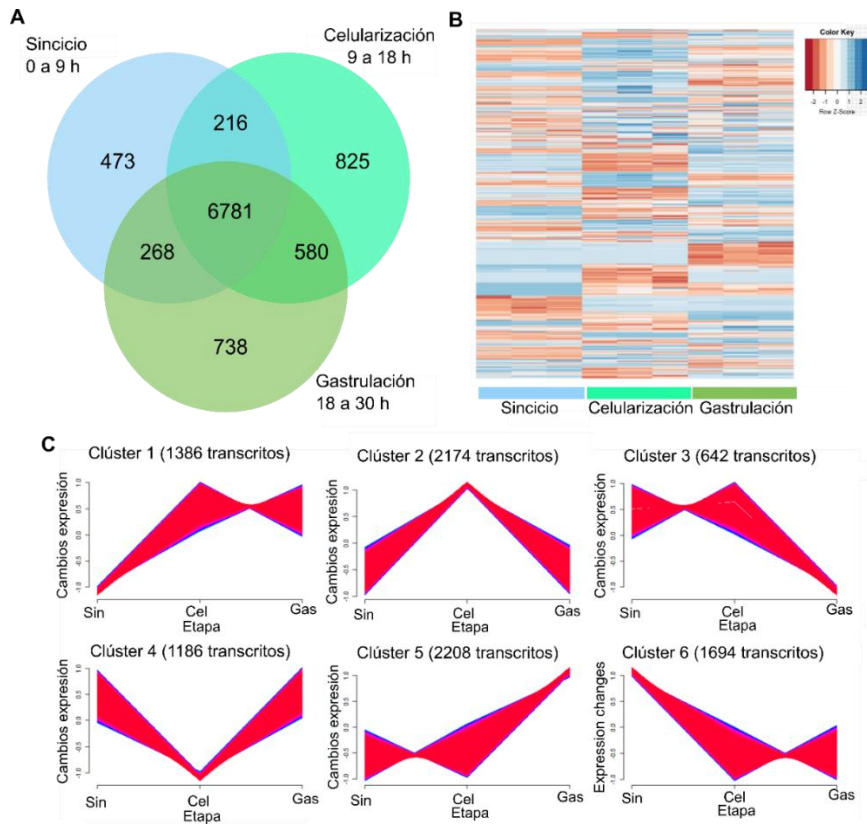


**Figura 5.6.** Anotación de la ontología de genes de cada transcriptoma. En el histograma se observan los primeros 15 grupos mayormente representados por los transcritos anotados. Las abreviaturas usadas son: Trans: transcripción; Reg: regulación; Morfogen: morfogénesis; Traduc: traducción; Cic. Cel: ciclo celular; Desarr. desarrollo; prot: proteínas; intracel: intracelular; Resp: respuesta; cel; celuar; Deter: determinación; Sist: sistema; catal: catalítico y citop: citoplasmática.

### 5.3.1 Análisis de expresión diferencial entre las etapas de desarrollo embrionario de *A. ludens*

En la etapa de sincicio se identificaron 21 000 transcritos, en celularización 26 000 transcritos y en gastrulación 32 000 transcritos; posteriormente se realizó una correlación de Pearson comparando las etapas de desarrollo, en donde se observó una correlación de 0.56 entre sincicio vs. celularización, en celularización vs. gastrulación una correlación de 0.63 y entre sincicio vs. gastrulación la correlación fue de 0.48. La tendencia de estas correlaciones era esperada, observando una mayor correlación entre las etapas de celularización y gastrulación, las cuales serían mayormente similares debido a sus eventos de desarrollo.

Se encontraron aproximadamente 9 000 unigenes expresados en cada etapa del desarrollo, de los cuales 6,781 se comparten entre las tres etapas de desarrollo en cantidades similares a las que se presentan en *D. melanogaster* y menores que en *C. capitata* (Lott *et al.*, 2011; Salvemini *et al.*, 2014). De estos unigenes se encontraron 473 únicos para sincicio, 825 únicos para celularización y 738 únicos para gastrulación; algunos otros se encuentran expresados en la transición entre etapas (Figura 5.6.A). Cuando realizamos una comparación en el nivel de abundancia de estos transcritos (Figura 5.6.B), se observa que hay transcritos que muestran patrones de abundancia diferencial entre las tres etapas de desarrollo. Para conocer si estos transcritos encontrados en las etapas de desarrollo se pueden clasificar en grupos, se realizó un análisis de soft clustering el cual permite agrupar transcritos con dinámicas similares de comportamiento entre las tres etapas de desarrollo embrionario (Peters *et al.*, 2013) (Figura 5.6.C); este análisis nos ayudó a agrupar la dinámica de los transcritos que encontramos, es decir, se agruparon dependiendo a la similitud de su nivel de abundancia en las tres etapas de desarrollo. Se encontraron seis grandes grupos que presentan transcritos con tendencias similares entre las etapas de desarrollo; por ejemplo, el clúster 3 en donde los transcritos son abundantes en sincicio y celularización disminuyendo su densidad en gastrulación, un comportamiento de simetría especular que muestra con el clúster 5.



**Figura 5.7.** Dinámicas de expresión en la embriogénesis temprana de *A. ludens*. En A, se observa la presencia de proteínas únicas y compartidas entre las tres etapas de desarrollo. En B, un mapa de calor en donde se observa el comportamiento general de los transcritos, es decir, en color rojo se representan transcritos poco abundantes, mientras que en color azul son transcritos muy abundantes. En C, las dinámicas de la abundancia de los transcritos en las tres etapas de desarrollo.

### 5.3.2 Identificación de transcritos importantes en la determinación sexual en *A. ludens*

Un tema de gran interés en la familia de tefritidos es conocer los genes que participan en la determinación sexual, no sólo por su importancia biológica si no por su potencial para ser manipulados como una herramienta para el control de poblaciones (Saccone, Pane and Polito, 2002). En *D. melanogaster* se conocen muy bien los genes que participan en la determinación sexual, algunos están expresados desde el embrión temprano hasta su metamorfosis; situación que no se presenta en *C. capitata* (Saccone, Pane and Polito, 2002; Salvemini *et al.*, 2009).

**Tabla 5.8.** Presencia de los genes de determinación sexual identificados en *A. ludens*, *B. dorsalis* y *C. capitata*. Los genes presentes son Si, los genes ausentes son No, los genes no encontrados son NE. Alu es *A. ludens*, Bdo es *B. dorsalis* y Cca es *C. capitata*.

Nombre	Símbolo	Alu			Bdo			Cca
		0-9 h	9-18 h	18-30 h	0-1 h	2-4 h	5-8 h	8-10 h
<i>fruitless</i>	Fru	Si	No	Si	No	No	No	No
<i>Scute</i>	sc (sisB)	Si	Si	Si	No	No	Si	Si
<i>deadpan</i>	Dnp	Si	Si	Si	No	Si	Si	Si
<i>Runt</i>	Run	Si	Si	Si	No	Si	Si	Si
<i>Doublesex</i>	Dsx	Si	Si	Si	Si	Si	Si	No
<i>Male-specific-lethal 2</i>	m <sub>sl</sub> -2	Si	Si	Si	Si	Si	Si	No
<i>Virilizer</i>	Vir	Si	Si	Si	Si	Si	Si	Si
<i>transformer-2</i>	tra2	Si	Si	Si	Si	Si	Si	Si
<i>daughterless</i>	Da	Si	Si	Si	Si	Si	Si	Si
<i>sex-lethal</i>	Sxl	Si	No	Si	No	Si	Si	Si
<i>Male-specific-lethal 3</i>	m <sub>sl</sub> -3	Si	Si	Si	Si	Si	Si	Si
<i>Male-specific-lethal 1</i>	m <sub>sl</sub> -1	Si	Si	Si	Si	Si	Si	Si
<i>intersex</i>	ix	Si	Si	Si	Si	Si	Si	Si
<i>hopscotch</i>	hop	Si	Si	Si	Si	Si	Si	Si
<i>extra macrochaetae</i>	emc	Si	Si	Si	No	No	No	Si
<i>sex-lethal</i>	Sxl	Si	Si	Si	Si	Si	Si	Si
<i>sisterless A</i>	sisA	Si	Si	Si	Si	Si	Si	Si
<i>degringolade</i>	dgrn	Si	Si	Si	Si	Si	Si	Si
<i>modifier of mdg4</i>	mod(mdg4)	Si	Si	Si	Si	Si	Si	No
<i>dissatisfaction</i>	dsf	NE	NE	NE	NE	NE	NE	NE
<i>female lethal d</i>	fl(2)d	Si	Si	Si	Si	Si	Si	Si

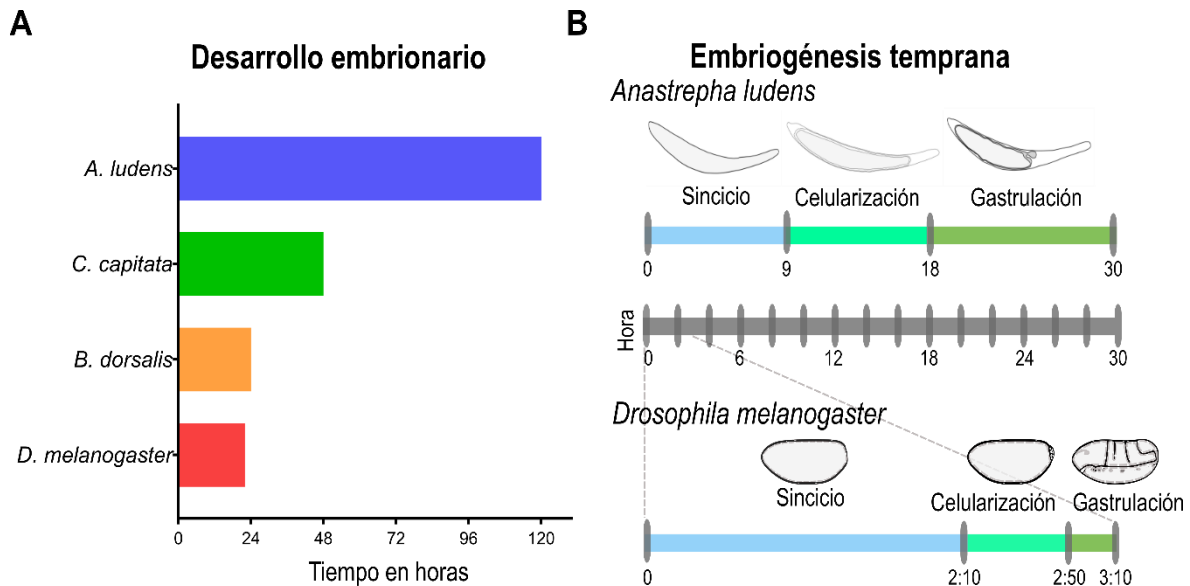
<i>hermaphrodite</i>	her	NE	NE	NE	NE	NE	NE	NE
<i>outstretched</i>	os (sisC)	NE	NE	NE	NE	NE	NE	Si
<i>Signal-transducer and activator of transcription protein at 92E</i>	Stat92E	Si	No	No	Si	Si	Si	Si
<i>transformer</i>	tra	Si	Si	Si	Si	Si	Si	Si
<i>groucho</i>	gro	Si	Si	Si	Si	Si	Si	Si
<i>Mes-4</i>	Mes-4	Si	Si	Si	Si	Si	Si	No

Se realizó una búsqueda de los principales genes que participan en la determinación sexual en *D. melanogaster*, encontrando 27 genes que participan en este proceso. En el caso de *A. ludens* se lograron identificar 24 genes en los tres transcriptomas (Tabla 5.8), algunos de ellos muy importantes en la determinación sexual de *D. melanogaster*, como el caso de *transformer-2* o *double-sex*. Los tres genes que no se lograron identificar son *dissatisfaction (dsf)*, *hermaphrodite (her)* y *outstretched (Os)*. Utilizando los datos de *C. capitata* y *B. dorsalis*, se logró identificar la presencia de genes de determinación sexual en los transcriptomas reportados de embriogénesis temprana, como se muestra en la Tabla 5.8. Los genes no encontrados en *A. ludens* están ausentes también en *B. dorsalis* y en *C. capitata*, con excepción de *Os* en *C. capitata*.

#### **5.4 Comparación de expresión transcripcional en la embriogénesis temprana entre diferentes moscas modelo y *A. ludens***

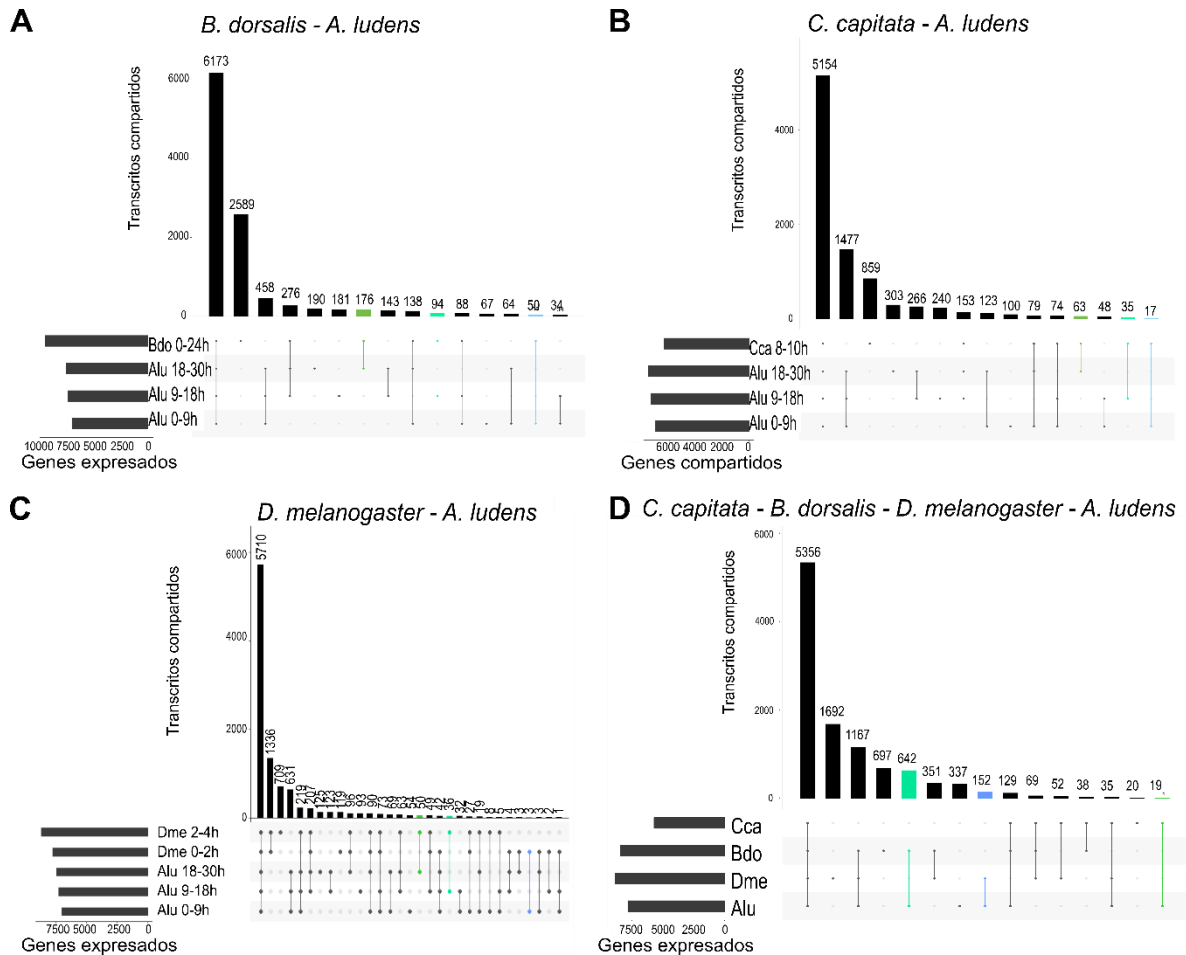
Para contar con una mejor percepción de las características únicas y compartidas entre el desarrollo embrionario de los tefrítidos, se realizó la comparación entre los datos de expresión de *A. ludens* con los datos reportados de *B. dorsalis* de embriones de 0 - 1 h, 2 - 4 h, 5 - 8 h y 0 - 24 h (Shen *et al.*, 2011), así como los reportados de *C. capitata* en etapa de gastrulación de 8 a 10 h (Salvemini *et al.*, 2014). Complementario a estos datos se agregaron los reportados de 0 - 2 h y 2 - 4 h de *D. melanogaster*. Un punto importante en la embriogénesis de estas moscas es su diversa temporalidad (Figura 5.7.A), en *A. ludens*

dura 120 h, en *C. capitata* dura 48 h y en *B. dorsalis* y *D. melanogaster* de 24 a 22 h, respectivamente. Debido a estas grandes diferencias temporales, se realizó la comparación de su embriogénesis temprana, en *A. ludens* ésta transcurre en las primeras 30 h en comparación con *D. melanogaster* que sólo lleva 3 h (Figura 5.7.B), es decir, en *A. ludens* transcurre en 10 veces más tiempo.



**Figura 5.8.** Temporalidad de la embriogénesis entre *A. ludens* con otras moscas. En A, se observa la temporalidad de la embriogénesis entre *A. ludens*, *C. capitata*, *B. dorsalis* y *D. melanogaster* bajo condiciones de crecimiento en laboratorio a 25 °C. En B, se observa la comparación entre las etapas de la embriogénesis temprana de *A. ludens* y *D. melanogaster*.

Posteriormente se continuo con la comparación a nivel de presencia de transcritos reportados en transcriptomas de *B. dorsalis* y *C. capitata*. Entre *B. dorsalis* (0 a 24 h) y *A. ludens* se comparten 6 173 transcritos en todas las etapas de desarrollo (Figura 5.8.A); cuando se realizan comparaciones entre transcritos conservados entre las etapas hay 176 en gastrulación, 94 en celularización y 50 transcritos en sincicio. En el caso de los transcritos compartidos entre *C. capitata* (8 a 10 h) y *A. ludens*, estos son 5 154 en todas las etapas de desarrollo; en gastrulación se comparten 63 transcritos, 35 en celularización y en sincicio 17. Entre los comportamientos observados se aprecia una similitud mayor entre *B. dorsalis* y *A. ludens*, observando menor similitud con *C. capitata*; esto podría ser debido a que los datos de *B. dorsalis* son de toda su embriogénesis (Shen *et al.*, 2011) y de *C. capitata* sólo son de la etapa de gastrulación (Salvemini *et al.*, 2014).



**Figura 5.9.** Transcritos compartidos entre los embriones de *A. ludens* con *B. dorsalis*, *C. capitata* y *D. melanogaster*. En A, la comparación entre los datos transcriptómicos de *B. dorsalis* (Bdo). En B, la comparación de los datos transcriptómicos entre *C. capitata* (Cca). En C, la comparación entre los datos transcriptómicos de *D. melanogaster* (Dme). En D, la comparación de los datos transcriptómicos globales, es decir, el conjunto de transcritos presentes entre todas sus etapas de desarrollo de las moscas estudiadas. Los colores que se observan representan las etapas de desarrollo estudiadas en *A. ludens*, en color verde claro es celularización, en color azul es la etapa de sincicio, en color verde claro es la etapa de celularización y en color verde oscuro es la etapa de gastrulación.

En el caso de las tendencias observadas en *D. melanogaster* se aprecian 5,710 transcritos compartidos con *A. ludens* en todas las etapas de desarrollo; en la etapa de sincicio son 3 transcritos, en celularización 36 y en gastrulación 50 (Figura 5.8.C). Estas tendencias son un tanto particulares, ya que se espera que se conserve gran parte de los

transcritos en las etapas de desarrollo; tomando en cuenta que los grupos de genes importantes en la determinación de estas etapas de desarrollo temprano son similares en dípteros. Este tema se retomará más adelante.

Continuando con este análisis de transcritos compartidos entre las moscas utilizadas, se realizó un análisis global con todos los datos utilizados de cada mosca, es decir, *B. dorsalis* (0 a 8 h), *C. capitata* (8 a 10 h), *D. melanogaster* (0 a 4 h) y *A. ludens* (0 a 30 h) (Figura 5.8.D). Se encontraron 5 356 transcritos conservados entre las cuatro moscas; 129 transcritos entre los tefrítidos (*B. dorsalis*, *C. capitata* y *A. ludens*), 642 transcritos entre *B. dorsalis* y *A. ludens*, 19 transcritos entre *C. capitata* y *A. ludens* y 152 transcritos entre *D. melanogaster* y *A. ludens*. Estas relaciones de transcritos compartidos muestran que gran parte de éstos están presentes en las cuatro moscas; observando también una mayor similitud entre *B. dorsalis* y *A. ludens* a diferencia de los compartidos con *C. capitata*.

## **5.5 Expresión de genes con un posible papel en el desarrollo embrionario temprano de *A. ludens***

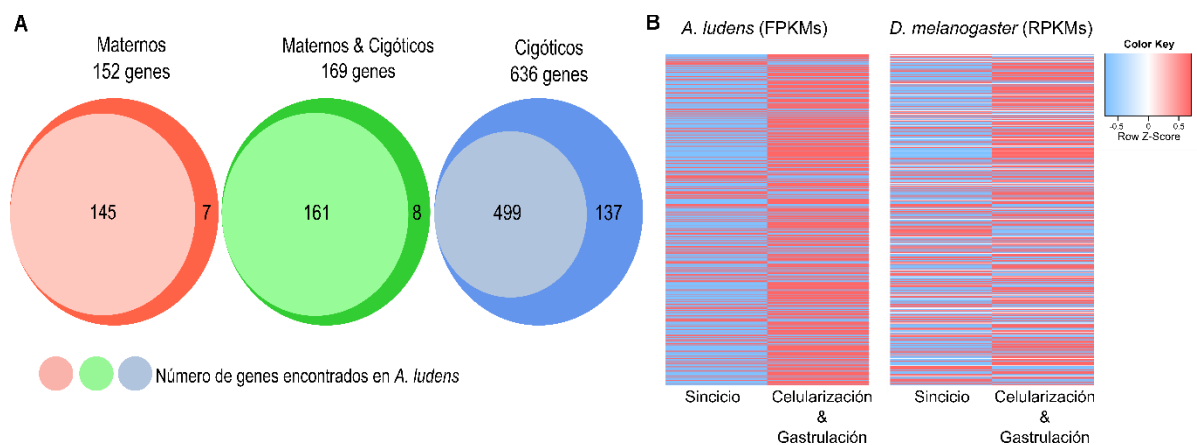
*D. melanogaster* es un buen modelo para estudiar el desarrollo embrionario de los animales, ya que los múltiples genes que participan en el establecimiento del plan corporal del embrión son bien conocidos (Gilbert, 2014). Por esta razón se decidió realizar una búsqueda de cuáles de los genes expresados en *D. melanogaster* se encuentran en el embrión temprano de *A. ludens*. Para este análisis se realizó una clasificación en tres categorías a dichos genes: maternos, los cuales son depositados en el embrión durante la ovogénesis y tienen un papel en el desarrollo temprano; maternos y cigóticos, genes que son expresados de forma materna pero también durante la activación del genoma del cigoto (ZGA) y cigóticos que sólo se expresan durante la activación del genoma del cigoto (Society for Developmental Biology, 2020a, 2020b).

### **5.5.1 Genes que participan en el desarrollo temprano de *D. melanogaster* identificados en *A. ludens***

Para realizar la búsqueda de la presencia de los genes que participan en el desarrollo temprano de *D. melanogaster* en *A. ludens*, se utilizaron solamente los genes que tienen una función reportada en *D. melanogaster* (Society for Developmental Biology, 2020a,



2020b), los cuales son 152 con expresión materna, 636 con expresión cigótica y 169 de expresión materna y cigótica. En *A. ludens* se encuentran 145 de expresión materna, 499 de expresión cigótica y 161 de expresión materna y cigótica (Figura 5.9.A). La gran mayoría de genes que participan en el desarrollo temprano de *D. melanogaster* fueron detectados en *A. ludens*, no obstante, no se encontraron todos los genes cigóticos. Al realizar un análisis de agrupamiento de estos genes se observó que el comportamiento de su expresión es similar entre ambas moscas (Figura 5.9.B), a pesar de que su temporalidad es muy distinta.

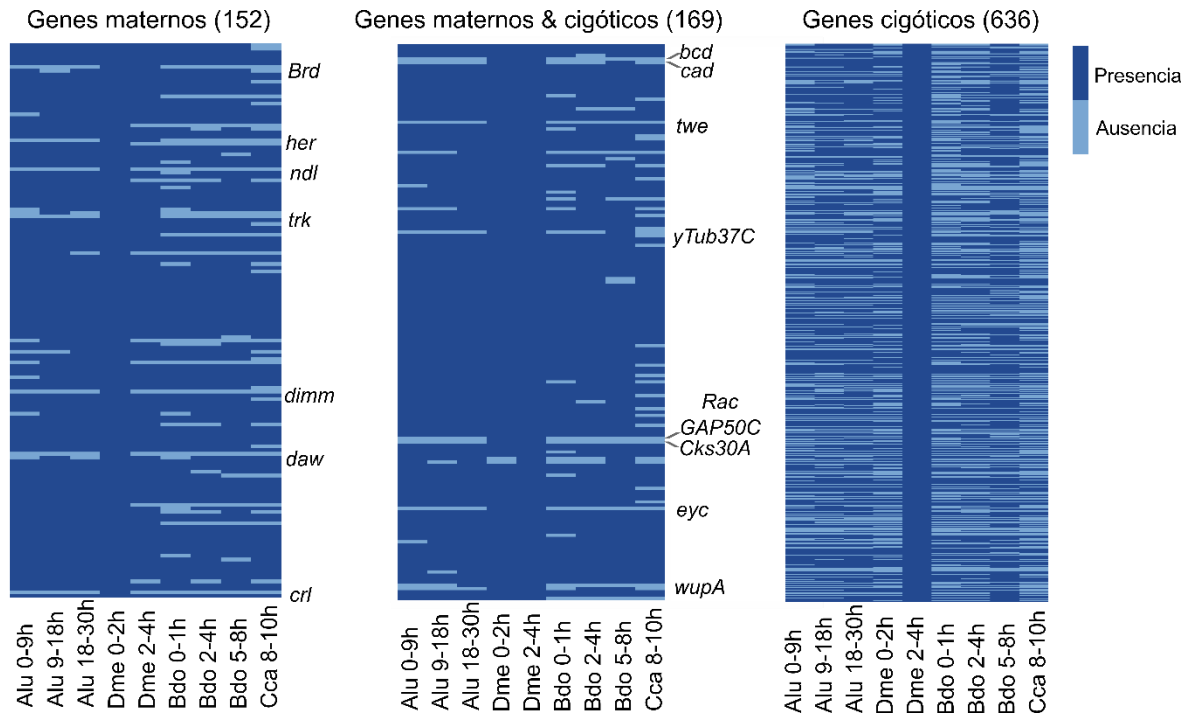


**Figura 5.10.** Genes del desarrollo embrionario temprano de *D. melanogaster* identificados en *A. ludens*. En A, se observan los genes presentes en *A. ludens* en comparación con los descritos en *D. melanogaster*. En B, un mapa de calor donde se observa el comportamiento general de los transcritos, el color azul representa transcritos poco abundantes y el color rojo transcritos muy abundantes.

### 5.5.2 Comparación de genes que participan en el desarrollo temprano de *D. melanogaster* identificados en *A. ludens*, *B. dorsalis* y *C. capitata*

Después se identificó la presencia de los genes de desarrollo temprano de *D. melanogaster* en los datos transcriptómicos de embriones tempranos de *B. dorsalis*, *C. capitata* y *A. ludens*. En la Figura 5.10 se muestra la presencia o ausencia de dichos genes en las cuatro moscas; a grandes rasgos su comportamiento es similar, destacando el de *C. capitata* ya que sólo se contó con un conjunto de datos y no se pudieron encontrar varios de los genes maternos. Se destaca que un grupo de genes maternos muy importantes en el desarrollo

de *D. melanogaster* no fueron encontrados en *A. ludens*, *B. dorsalis* y *C. capitata*, como los genes *bicoid* (*bcd*), *caudal* (*cad*) y *trunk* (*trk*), que en *D. melanogaster* son fundamentales para la determinación antero-posterior del embrión (Gilbert, 2014; Society for Developmental Biology, 2020a).



**Figura 5.11.** Genes maternos y cigóticos presentes en *A. ludens*, *C. capitata*, *B. dorsalis* y *D. melanogaster*. Presencia (azul fuerte) o ausencia (azul claro) de los 152 genes maternos, 169 genes maternos y cigóticos, y los 636 genes cigóticos en las moscas estudiadas.

Tampoco se identificaron genes que participan en el establecimiento de la capa germinal, como es *twine* (*twe*) y *nudel* (*ndl*). No obstante, la ausencia del gen *bcd* no es sorpresa, ya que en reportes previos se muestra que es un gen específico para moscas de la familia Drosophilidae (MacDonald, 1990; Gregor, McGregor and Wieschaus, 2008) y que no está presente en moscas *Anastrepha* (de Oliveira *et al.*, 2017). Con esta información, los análisis indican que la gran mayoría de genes necesarios para el desarrollo embrionario de *D. melanogaster* se encuentran presentes en *A. ludens*; sin embargo, parece que algunos genes importantes en el desarrollo de *D. melanogaster* no se encuentran en tefrítidos.

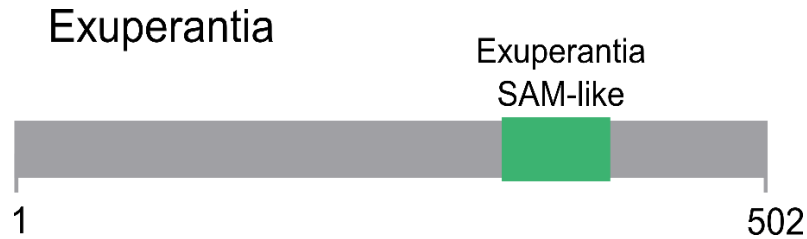
## 5.6 Patrones de expresión de genes del desarrollo embrionario temprano de *A. ludens*

Los patrones de transcritos que modulan la expresión del embrión temprano de *D. melanogaster* son herramientas poderosas para entender el papel de los genes en la especificación de los ejes antero - posterior y dorso - ventral, así como en la determinación de los segmentos del organismo. Debido a estas características se analizaron e identificaron las secuencias de genes que participan en estos eventos de especificación del plan corporal, y a partir de estas secuencias se desarrollaron sondas de RNA mensajero para observar su localización por medio de hibridaciones *in situ*.

### 5.6.1 Gen *exuperantia* de *A. ludens*

En *D. melanogaster* se conoce que el transcrito del gen *exuperantia* (*exu*) es importante para la localización del transcrito del gen *bcd* y la regulación del plasma polar en conjunto con el gen *oskar* (*osk*) en el embrión, así como en la línea germinal (Hazelrigg *et al.*, 1990). La proteína Exu tiene una composición alta en residuos de serina, los cuales llegan a estar presentes en una unidad de repetidas de Ser-X, en donde X es usualmente glicina, con una longitud de 532 aminoácidos (Hazelrigg *et al.*, 1990; Macdonald, Luk and Kilpatrick, 1992). En datos de expresión de RNA-Seq su expresión es principalmente de las 0 a las 2 h de vida del embrión (FlyBase, 2020c).

En estudios recientes en *A. fraterculus* se identificó que se ha perdido el gen *bcd* en estas moscas, por lo que se hipotetiza que *exu* podría tener la función similar de *bcd* en *D. melanogaster*, esto en el suborden Cyclorrhapha (de Oliveira *et al.*, 2017). En el análisis llevado a cabo en *A. ludens* se encontró al transcrito que codifica a la proteína Exu de 502 aminoácidos, donde se identificó el dominio Exuperantia SAM-like (Figura 5.11) y, en comparación con la identidad de la proteína de *D. melanogaster* existe una similitud de 59.79% con *A. ludens*. Utilizando el transcrito de *exuperantia* de *A. ludens* se construyó la sonda de RNA mensajero para la hibridación *in situ* con una longitud de 117 nucleótidos (Anexo 4).

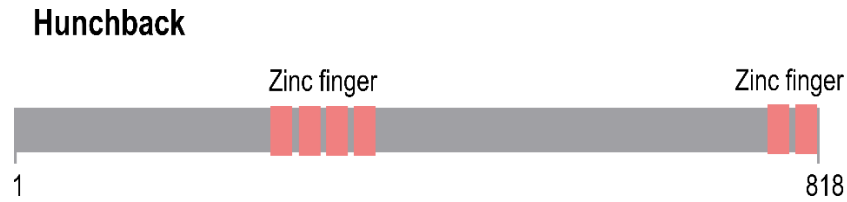


**Figura 5.12.** Gen *exuperantia* de *A. ludens*. Se observa un esquema de la proteína Exu donde se señala su dominio más importante en color verde.

### 5.6.2 Gen *hunchback* de *A. ludens*

Existen tres clases de genes cigóticos de segmentación los cuales se relacionan, dependiendo de los diferentes niveles de organización, durante la segmentación del embrión. Los genes *gap* constituyen una clase de genes con información materna y cigótica, los cuales incluyen a *hunchback* (*hb*), *Krüppel* (*Kr*) y *knirps* (*kni*); que son importantes para la localización del eje antero - posterior. *Hunchback* cuenta con por lo menos dos transcritos importantes, el materno y el cigótico, en donde ambos son transcritos que provienen del mismo RNA pero sus regiones 5' son diferentes (Tautz *et al.*, 1987; Struhl, Johnston and Lawrence, 1992). La proteína Hb cuenta con dos dominios diferentes, compuestos por 4 dedos de Zinc y dos dedos de Zinc en la parte carboxilo, con una longitud de 758 aminoácidos (Tautz *et al.*, 1987). En datos de expresión de RNA-Seq su expresión es principalmente de las 0 a las 4 h del desarrollo del embrión, disminuyendo su expresión de forma progresiva (FlyBase, 2020d).

En el análisis llevado a cabo en *A. ludens* se encontró al transcrito que codifica la proteína Hb de 818 aminoácidos, donde se identificaron seis dominios dedos de Zinc (Figura 5.12) y, en comparación con la de identidad de la proteína de *D. melanogaster*, su similitud es de 55.97% con *A. ludens*. Utilizando el transcrito de *hunchback* de *A. ludens* se construyó la sonda de ARN mensajero para la hibridación *in situ* con una longitud de 406 nucleótidos (Anexo 4). El diseño de la sonda permitió su uso en embriones de *D. melanogaster* (Anexo 8), donde se observó el patrón previamente reportado.

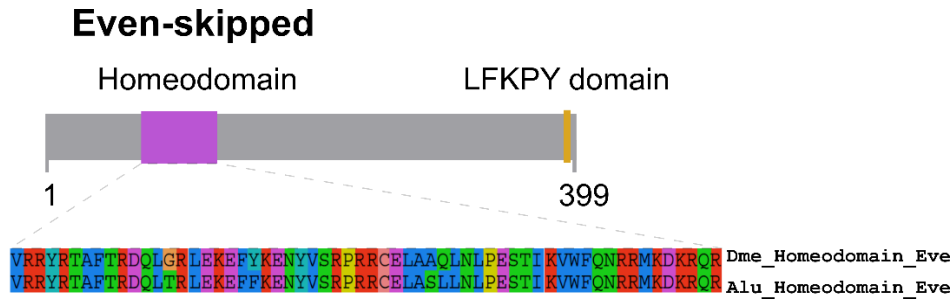


**Figura 5.13.** Gen *hunchback* de *A. ludens*. Se observa un esquema de la proteína Hb donde se señala sus dominios dedos de Zinc identificados en color anaranjado.

### 5.6.3 Gen *even skipped* de *A. ludens*

La cascada de genes de segmentación en *D. melanogaster* ha sido muy estudiada, razón por lo cual cada gen de segmentación tiene una jerarquía bien definida. Los genes *pair-rule* primarios se encuentran en series de rayas repetidas de forma transversal a lo largo del embrión. Uno de estos genes *pair-rule* primarios es *even skipped* (*eve*) (Small, Blair and Levine, 1992), cuenta con un dominio homeobox y con un patrón de expresión temprana de “cebra” (7 bandas a lo largo del eje antero - posterior) (Macdonald, Ingham and Struhl, 1986). La proteína Eve es de 376 aminoácidos y cuenta con un dominio homeobox y un dominio LFKPY que se utiliza para la interacción con la proteína Groucho (Gro) (Frasch *et al.*, 1987; Fujioka *et al.*, 2002). En datos de expresión de RNA-Seq el transcrito se encuentra de las 2 a las 4 h en el desarrollo del embrión (FlyBase, 2020b).

En el análisis llevado a cabo en *A. ludens* se encontró el transcrito que codifica la proteína Eve de 399 aminoácidos y que tiene una similitud de 64.16% con la proteína Eve de *D. melanogaster*. En la proteína de *A. ludens* se identificó el homeodominio (Figura 5.13.), donde se presenta una similitud de 96.67% con el reportado en *D. melanogaster*; y también se identificó el dominio de LFKPY en la parte carboxilo terminal de la proteína. Utilizando el transcrito de *even skipped* de *A. ludens* se construyó la sonda de RNA mensajero para la hibridación *in situ* con una longitud de 422 nucleótidos (Anexo 4). El diseño de la sonda permitió su uso en embriones de *D. melanogaster* (Anexo 8), donde se observó el patrón previamente reportado.

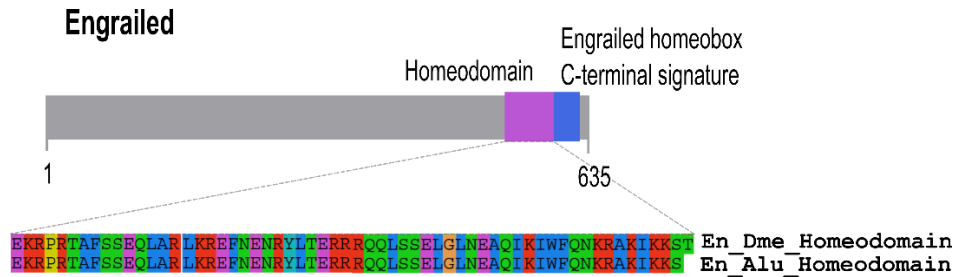


**Figura 5.14.** Gen *even skipped* de *A. ludens*. Se observa un esquema de la proteína Eve donde se señala su dominio homeobox, así como la secuencia de aminoácidos entre *A. ludens* y *D. melanogaster*, y su dominio LFKPY de color amarillo.

#### 5.6.4 Gen *engrailed* de *A. ludens*

*Engrailed* es un gen que participa en la formación del patrón corporal en el embrión y el adulto de *D. melanogaster*. A las 3.5 h del desarrollo del embrión, el transcrito del gen se acumula en bandas estrechas a lo largo del eje antero – posterior y su proteína se encuentra alternada entre los segmentos ya que se sabe que está regulada por los productos de genes *pair-rule* (DiNardo *et al.*, 1985). La proteína En cuenta con un dominio homeobox de 60 aminoácidos, que le permite unirse a secuencias de DNA; así como un dominio adicional cerca del carboxilo terminal de 39 aminoácidos, dejando una proteína total de 552 aminoácidos (Desplan, Theis and O’Farrell, 1985). El embrión celularizado presenta un patrón de 14 bandas y conforme avanza su desarrollo estas son más notorias (Desplan, Theis and O’Farrell, 1985; DiNardo *et al.*, 1985). Por estudios de RNA-seq se sabe que el transcrito es muy abundante de las 4 a las 14 h de la embriogénesis en *D. melanogaster* (FlyBase, 2020a).

En el análisis llevado a cabo en *A. ludens* se encontró al transcrito que codifica la proteína En de 635 aminoácidos, que presenta una similitud de 68.96% con la proteína En de *D. melanogaster*. En la proteína de *A. ludens* se identificó al homeodominio (Figura 5.13.A), donde se presenta una similitud de 98.34% con el reportado en *D. melanogaster*, y también se identificó el dominio de Engrailed homeobox C-terminal en la parte carboxilo de la proteína. Utilizando el transcrito de *engrailed* de *A. ludens* se construyó la sonda de RNA mensajero para la hibridación *in situ* con una longitud de 332 nucleótidos (Figura 5.14.B). El diseño de la sonda permitió utilizarla en embriones de *D. melanogaster* (Anexo 8), donde se observó el patrón previamente reportado.



**Figura 5.15.** Gen *engrailed* de *A. ludens*. Se observa un esquema de la proteína donde se señala su dominio homeobox, así como la secuencia de aminoácidos entre *A. ludens* y *D. melanogaster*, y su dominio Engrailed Homeobox C-terminal.

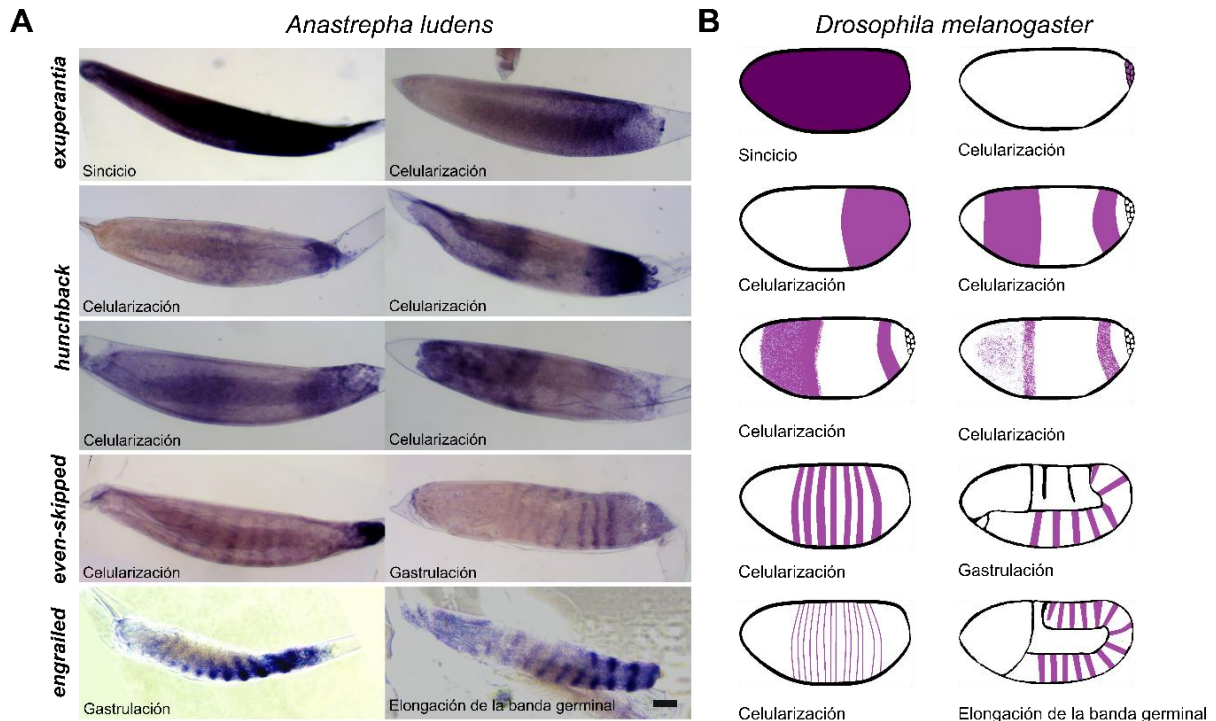
### 5.6.5 Comparación de patrones de expresión de genes del desarrollo embrionario temprano entre *A. ludens* y *D. melanogaster*

Los patrones de expresión génica modulan el desarrollo temprano del embrión de *D. melanogaster*, por lo que se desarrolló la técnica de hibridación *in situ* para los embriones de *A. ludens*. Fue interesante observar los patrones de expresión relevantes para la formación del embrión, por lo que se seleccionaron los cuatro genes previamente descritos, *exu*, *hb*, *eve* y *en* de *A. ludens*. Una vez que se contó con la anotación manual de los genes se construyeron las sondas y las hibridaciones en las etapas correspondientes a sincicio, celularización y gastrulación, así como en la etapa de elongación de la banda germinal. A continuación, se presenta una descripción de los patrones observados en los embriones de *A. ludens* (Figura 5.15.A):

- En el caso de *exu* se observó una señal ubicua en embriones en sincicio, la cual desaparece de forma gradual en la etapa de celularización. Cuando se observaron embriones en gastrulación no se encontró señal de la sonda de *exu* en *A. ludens* (n=25 de embriones observados).
- En el caso de *hb* se observó la señal en la etapa de celularización, donde se encontraron los diferentes patrones (Fig. 5.15.A) previamente reportados en *D. melanogaster*. En los embriones gastrulados de *A. ludens* se observó la señal de la sonda en lo que podría ser el ectodermo del embrión (n=35 de embriones observados).
- En el caso de *eve* se observaron en el embrión celularizado de *A. ludens* las siete bandas que corresponden al transcrito, las cuales se continúan observando en la

gastrulación temprana, como se aprecia en la figura 5.15.A (n=12 de embriones observados).

- En el caso de *en* se observó la señal desde el embrión gastrulado hasta el embrión en la retracción de la banda germinal. En la Fig. 5.15.A se observan las 14 bandas finas a lo largo del embrión en gastrulación temprana y en la banda germinal (n=12 de embriones observados).



**Figura 5.16.** Localización de genes involucrados en el patrón antero - posterior del embrión de *A. ludens*. En A, los patrones de localización de genes expresados en la embriogénesis de *A. ludens* utilizando hibridaciones *in situ* (el color morado indica la localización de la sonda). La barra representa 100  $\mu$ m. En B, esquema de referencia de la localización de estos genes en *D. melanogaster* en la etapa de desarrollo indicada. Fuente: BDGP, 2020.

Estos resultados muestran una gran similitud con lo reportado en *D. melanogaster* (Fig. 5.15.B) por lo menos en cuatro de los genes analizados. En *exu* se aprecia una diferencia en la localización del transcrito, ya que en *D. melanogaster* éste, en la etapa de celularización, se encuentra localizado en las células polares mientras que *exu*, en *A. ludens*, se observa con un gradiente en la parte posterior del embrión.



## Capítulo 6. Discusión

Desde inicios de 1900, poco después de la primera descripción de *A. ludens*, se observó que las larvas emergían de frutos como naranjas y mangos (Herrera *et al.*, 1903); debido a esto, tomó importancia el estudio de esta mosca. Posteriormente, en la primera serie de estudios que describen su comportamiento en relación con la fruta, se clasificó como mosca plaga, principalmente de frutos cítricos, mangos y especies endémicas de México; como el chapote y el zapote (Greene, 1929; McPhail and Bliss, 1933; Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, 1944). A pesar de su importancia, los estudios relacionados con su estado embrionario se centraron en su morfología externa (Emmart, 1933; Carroll and Wharton, 1989), la cual es importante para la identificación en campo de la plaga; sin embargo, contando con una descripción detallada de sus etapas de desarrollo embrionario (Gutiérrez-Ramos, 2016), los enfoques genómico y molecular son de gran relevancia para conocer los mecanismos y procesos de su desarrollo.

En la primera aproximación en condiciones de laboratorio de la embriogénesis de *A. ludens* se notó que su desarrollo es de 120 h, es decir, 5 veces más largo que el periodo de desarrollo embrionario de *D. melanogaster* (Campos-Ortega and Hartenstein, 1985). En cambio, la diferencia con otros tefrítidos no es tan importante, como se observa en la Figura 5.7. El tiempo de desarrollo del embrión de *A. ludens* tiene una correlación con el tamaño de su genoma, el cual está reportado como de 1555 Mb (De Teodoro Pardo, 2011), en comparación con las 139 Mb de *D. melanogaster*, 322 Mb de *B. dorsalis* y 591 Mb de *C. capitata* (Tsoumani and Mathiopoulos, 2012).

Con esto en mente, esta extensión de 120 h de tiempo embrionario de *A. ludens* fue de gran interés debido a la temporalidad necesaria para realizar procesos regulatorios de desarrollo a un nivel transcriptómico; por ejemplo, conocer los transcritos de herencia materna en las primeras horas del desarrollo embrionario para continuar con el inicio de la transcripción del genoma del cigoto. Las etapas que más interesaron fueron las de sincicio, celularización y gastrulación, donde se encontró una gran cantidad de transcritos presentes en las tres etapas de desarrollo. Dichos transcritos participan en procesos metabólicos, regulación de moléculas de olor, eventos de desarrollo como en neurogénesis, etc (observado en la Figura 5.5). La gran mayoría de los transcritos que resultaron únicos por etapa de desarrollo no se anotaron, debido a que no se encontraron similitudes con otros

previamente reportados, por lo que éstos podrían ser transcritos únicos de *A. ludens*. De los transcritos identificados con función en la formación del plan corporal en *D. melanogaster* se logró identificar a la mayoría de dichos genes (observado en la Figura 5.9); no obstante, se notó un comportamiento diferente al reportado en *D. melanogaster*. Las tendencias del comportamiento de dichos transcritos apuntan a que una gran mayoría de éstos se encuentran desde la etapa de sincicio (con muy pocas lecturas), así como en la etapa de celularización y gastrulación. Esto resulta importante ya que podrían llegar a conservar la función en *A. ludens* pero, a un nivel de regulación transcripcional, resulta diferente en comparación con lo reportado para *D. melanogaster*. También se cree que debido a que la temporalidad de las etapas de hasta 9 h de sincicio y celularización, y 12 h de gastrulación, se podrían encontrar subetapas de transcripción; es decir, que se encuentre un empalme de los transcritos. Esto no se ha observado en *D. melanogaster* ya que se encuentra de manera muy restringida la participación de los genes en cada etapa de desarrollo; así como mecanismos que permitan evitar la degradación de dichos transcritos durante tanto tiempo.

Una observación importante es que a pesar de que *A. ludens* cuenta con una cantidad de transcritos que podrían contar con una (o varias) funciones, se encontraron hasta un 40% de secuencias repetidas en los tres transcriptomas lo que es muy interesante ya que se cuenta con poca evidencia de la función de dichos elementos del genoma (Palomeque and Lorite, 2008; Ding *et al.*, 2017; Petersen *et al.*, 2019). Asimismo, se realizó una búsqueda de precursores de miRNAs, los cuales cuentan con funciones muy importantes en el desarrollo de los animales (Asgari, 2013). Algunos de los pre-miRNAs identificados en *A. ludens* cuentan con funciones en el desarrollo de *D. melanogaster* y otras moscas (observado en Anexo 6), por lo que se podría hipotetizar que tienen una función similar en *Anastrepha*.

En las comparaciones realizadas entre *D. melanogaster* y *A. ludens* se encontró una tendencia de expresión similar entre las etapas de celularización y gastrulación. Cuando se realizaron las comparaciones de los datos transcriptómicos entre los tefrítidos, *A. ludens*, *B. dorsalis* y *C. capitata*, el comportamiento observado es más similar entre *Bactrocera* y *Anastrepha*, lo cual podría ser consecuencia de que se contó con mayor número de datos transcriptómicos en comparación con los de *Ceratitis* de la que solo se contó con los datos de gastrulación. En el caso de los genes de determinación sexual, muy importantes para aplicaciones de control de plaga, se encontraron los genes de expresión temprana

esenciales para la determinación sexual de *D. melanogaster* (observado en Tabla 5.8). No se encontraron los genes *dsf*, *her* y *Os* en el análisis de *A. ludens*; un caso particular es *sexlethal (sxl)* el cual no está relacionado con la determinación sexual en todas las especies de dípteros (Saccone, Pane and Polito, 2002).

La comparación realizada de los transcritos expresados en la embriogénesis temprana de *A. ludens* con las otras moscas demostró que un gran porcentaje de genes se encuentran conservados entre ellas (observado en Figura 5.6), y sólo unos cuantos transcritos parecen ser únicos de *A. ludens*. Debido a las observaciones morfológicas previamente reportadas del embrión temprano de *A. ludens* (Gutiérrez-Ramos, 2016), nos interesó conocer la localización de dichos transcritos, ya que a pesar de que estén conservados éstos podrían contar con algún cambio en su patrón de localización. Considerando que los eventos de extrusión de la yema se presentan a lo largo de su embriogénesis, éstos se consideraron marcadores morfológicos del desarrollo del embrión de *A. ludens* ya que permiten determinar la transición del desarrollo, es decir, de sincicio a celularización, de celularización a gastrulación y de retracción de la banda germinal a la involución de la cabeza.

Los genes utilizados para conocer los patrones de localización del embrión de *Anastrepha* son muy importantes para la determinación del patrón corporal en *D. melanogaster*, los cuales son *exu*, *hb*, *eve* y *en* (observado en Figura 5.15). En el caso de *hb*, *eve* y *en* se observó el mismo patrón que el reportado en *D. melanogaster*, sin embargo, en el caso de *exu* se observó una localización ubicua a lo largo del embrión sincicial para continuar con una degradación del transcrito a lo largo de la celularización y la gastrulación; un comportamiento también observado en los datos transcriptómicos. Este comportamiento difiere del reportado en *D. melanogaster*, el cual en la etapa de sincicio se encuentra ubicuo y, conforme avanza a la celularización, el transcrito se va concentrando en las células polares y disminuyendo su presencia. La localización del transcrito *exu* de *A. ludens* en el sincicio es ubicuo, a pesar de sufrir las extrusiones, conforme avanza la celularización y gastrulación se comienza a observar una localización posterior. Este comportamiento de localización es interesante ya que podríamos hipotetizar que por lo menos *exu* no es un transcrito que defina la cantidad de yema que se debe extruir, a pesar de que se reporta en de Oliveira *et al* que *exu* podría suplir la función del gen *bcd* (de Oliveira *et al.*, 2017). En su localización en *A. ludens* el gen *exu* no cuenta con una distribución similar a la reportada en *D. melanogaster*.

Para concluir, también se observó que los principales genes que participan en las vías de la determinación del plan corporal en *D. melanogaster* se encuentran en *A. ludens* y, por el momento los tres estudiados (*hb*, *eve* y *en*) cuentan con una localización igual a la reportada. Se podría hipotetizar que los genes de herencia materna son los más divergentes entre *A. ludens* y *D. melanogaster*, los cuales determinarían las diferencias morfológicas y regulatorias entre ambas moscas. Valdría la pena observar la localización y función de un grupo de genes maternos en ambas moscas para conocer las diferencias evolutivas entre los mecanismos regulatorios embrionarios.

## Capítulo 7. Conclusión y perspectivas

El estudio de *A. ludens* se encuentra en sus primeras etapas, en comparación con *D. melanogaster* que se convirtió en un modelo de estudio desde inicios del siglo XX. La ventaja de tener la embriogénesis de *A. ludens* en etapas y homologada a la descrita en *D. melanogaster*, es que permite tener preguntas más específicas, observar diferencias y comenzar a estudiar eventos que pueden ser relevantes para el estudio de esta mosca. Este trabajo muestra la importancia de contar con información de las etapas morfológicas del embrión de *A. ludens* y la descripción de su embriogénesis temprana a un nivel transcriptómico, considerando las etapas de sincicio, celularización y gastrulación; así como con la información de las similitudes y transcritos únicos identificados y comparados con tefrítidos como es el caso de *B. dorsalis* y *C. capitata*. La estandarización de la técnica de hibridación *in situ* es importante para implantar técnicas más a la vanguardia y poder responder preguntas más puntuales del desarrollo de *A. ludens*.

El estudio de *A. ludens* es valioso debido a su importancia económica ya que los resultados pueden generar una plataforma para comprender mejor sus mecanismos regulatorios, así como diseñar estrategias para el control de la plaga. De manera similar se puede generar nuevo conocimiento para tener una visión más amplia sobre los diferentes planes corporales en animales, considerando las diferencias de localización y función de genes, como se observa en *exuperantia* o la ausencia de *bicoid* observada en tefrítidos.

Las preguntas que surgen a partir de este trabajo se pueden abordar desde diferentes áreas de estudio, por ejemplo: cuáles son los transcritos que determinan el eje antero - posterior de efecto materno; cómo se determina la yema que debe extruirse en el embrión; determinar si el papel de *exu* suple la función de *bcd* en la determinación anterior en el embrión; conocer si hay subetapas de transcritos maternos y transcritos cigóticos. La lista de cuestionamientos es muy larga, sin embargo, lo más intrincado es comenzar.

## Capítulo 8. Referencias

(Conabio), C. N. para el C. y uso de la B. (2011) *La biodiversidad en Veracruz: estudio de estado*. Edited by C. N. para el C. y uso de la B. (Conabio) et al. Xalapa.

Aluja, M. (1994) 'Bionomics and Management of *Anastrepha*', *Annual Review of Entomology*, 39(1), pp. 155–178.

Aluja, M. and Norrbom, A. L. (eds) (1999) *Fruit flies (Tephritidae): phylogeny and evolution of behavior*. Boca Raton, Florida: CRC Press.

Anderson, D. T. (1962) 'The Embryology of *Dacus tryoni* ( Frogg.) [Diptera, Trypetidae (=Tephritidae)], the Queensland Fruit-Fly', *J. Embryol. exp. Morph*, 10(September), pp. 248–292.

Andrews, S. (2010) 'FastQC'. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Angelini, D. R. and Kaufman, T. C. (2005) 'Comparative Developmental Genetics and the Evolution of Arthropod Body Plans', *Annual Review of Genetics*, 39(1), pp. 95–119. doi: 10.1146/annurev.genet.39.073003.112310.

Arbeitman, M. N. *et al.* (2002) 'Gene expression during the life cycle of *Drosophila melanogaster*', *Science*, 297(5590), pp. 2270–2275. doi: 10.1126/science.1072152.

Asgari, S. (2013) 'MicroRNA functions in insects', *Insect Biochemistry and Molecular Biology*. Elsevier Ltd, 43(4), pp. 388–397. doi: 10.1016/j.ibmb.2012.10.005.

Bai, X. *et al.* (2019) 'CRISPR/Cas9-mediated knockout of the eye pigmentation gene *white* leads to alterations in colour of head spots in the oriental fruit fly, *Bactrocera dorsalis*', *Insect Molecular Biology*, 28(6), pp. 837–849. doi: 10.1111/imb.12592.

Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, M. (1944) *A review of studies on the mexican fruitfly and related mexican species*, *Miscellaneous Publication*. Washington D.C.: United States Department of Agriculture.

Berrigan, D. A. *et al.* (1988) 'Age and host effects on clutch size in the Mexican fruit fly , *Anastrepha ludens*', *Entomol. exp. appl.*, 47, pp. 73–80.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: A flexible trimmer for Illumina

sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. doi: 10.1093/bioinformatics/btu170.

Bownes, M. (1975) 'A photographic study of development in the living embryo of *Drosophila melanogaster*', *Journal of Embryology and Experimental Morphology*, 33(3), pp. 789–801.

Bray, N. L. *et al.* (2016) 'Near-optimal probabilistic RNA-seq quantification', *Nature Biotechnology*, 34(5), pp. 525–527. doi: 10.1038/nbt.3519.

Brian, H. and Papanicolaou, A. (2019) 'Transdecoder (Finding Coding Regions Within Transcripts)'. Available at: <https://github.com/TransDecoder/TransDecoder/wiki>.

Campos-Ortega, J. A. and Hartenstein, V. (1985) *The Embryonic Development of Drosophila melanogaster*, Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-662-02454-6.

Carroll, L. E. and Wharton, R. A. (1989) 'Morphology of the Immature Stages of *Anastrepha ludens* (Diptera: Tephritidae)', *Entomological Society of America*, 89, pp. 201–214.

Carroll, S. B. (2008) 'Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution', *Cell*, 134(1), pp. 25–36. doi: 10.1016/j.cell.2008.06.030.

Carroll, S. B., Grenier, J. K. and Weatherbee, S. D. (2005) *From DNA to Diversity: molecular genetics and the evolution of animal design*. 2nd edn, *Journal of Paleontology*. 2nd edn. Malden, MA: Blackwell Publishing. doi: 10.1666/0022-3360(2003)077<0598:r>2.0.co;2.

Celedonio-Hurtado, H. *et al.* (1988) 'Demography of *Anastrepha ludens*, *A. obliqua* and *A. serpentina* (Diptera: Tephritidae) in Mexico', *Florida Entomologist*, 71(2), pp. 111–120.

Chen, S. L. *et al.* (2008) 'Female-specific *doublesex* dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, *Bactrocera dorsalis* (Hendel)', *Insect Biochemistry and Molecular Biology*, 38(2), pp. 155–165. doi: 10.1016/j.ibmb.2007.10.003.

Coolen, M., Katz, S. and Bally-Cuif, L. (2013) 'miR-9: A versatile regulator of neurogenesis', *Frontiers in Cellular Neuroscience*, 7(Nov), pp. 1–11. doi: 10.3389/fncel.2013.00220.

Desplan, C., Theis, J. and O'Farrell, P. H. (1985) 'The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity', *Nature*, 318(6047), pp. 630–635. doi: 10.1038/318630a0.

DGSV, D. G. de S. V. (2020) 'Campaña Moscas de la Fruta, quinto informe mensual'. Ciudad de México: SENASICA.

Díaz-Fleischer, F. and Aluja, M. (2003) 'Clutch size in frugivorous insects as a function of host firmness : the case of the tephritid fly *Anastrepha ludens*', *Ecological Entomology*, 28, pp. 268–277.

Dickens, J. C., Solis, E., Hart, W. G. (1982) 'Sexual development and mating behavior of the Mexican fruit fly, *Anastrepha ludens* (Loew)', *The Southwestern Entomologist*, 7(1), pp. 9–15.

DiNardo, S. *et al.* (1985) 'Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear engrailed protein', *Cell*, 43(1), pp. 59–69. doi: 10.1016/0092-8674(85)90012-1.

Ding, S. *et al.* (2017) 'Large-scale analysis reveals that the genome features of simple sequence repeats are generally conserved at the family level in insects', *BMC Genomics*. BMC Genomics, 18(1), pp. 1–10. doi: 10.1186/s12864-017-4234-0.

Dutra, V. S. *et al.* (2013) 'Description of Eggs of *Anastrepha curitis* and *Anastrepha leptozona* (Diptera: Tephritidae) Using SEM', *Annals of the Entomological Society of America*, 106(1), pp. 13–17. doi: 10.1603/an12054.

Emmart, E. W. (1933) 'The eggs of four species of fruit flies of the genus *Anastrepha*', *Proc. Ent. Soc. Wash.*, 35(8), pp. 184–191.

Ewels, P. *et al.* (2016) 'MultiQC: Summarize analysis results for multiple tools and samples in a single report', *Bioinformatics*, 32(19), pp. 3047–3048. doi: 10.1093/bioinformatics/btw354.

Figueiredo, J. V. A. *et al.* (2013) 'External eggshell morphology of *Anastrepha* fruit flies (Diptera: Tephritidae)', *Acta Zoologica*, 94(2), pp. 125–133. doi: 10.1111/j.1463-6395.2011.00533.x.

Finn, R. D. *et al.* (2014) 'Pfam: The protein families database', *Nucleic Acids Research*, 42(D1), pp. 222–230. doi: 10.1093/nar/gkt1223.

FlyBase (2020a) '*Engrailed*'. Available at: <http://flybase.org/reports/FBgn0000577>.

FlyBase (2020b) '*Even skipped*'. Available at: <http://flybase.org/reports/FBgn0000606>.

FlyBase (2020c) '*Exuperantia*'. Available at: <http://flybase.org/reports/FBgn0000615>.

FlyBase (2020d) '*Hunchback*'. Available at: <http://flybase.org/reports/FBgn0001180>.



- Foe, V. E. and Alberts, B. M. (1983) 'Studies of nuclear and cytoplasmic behavior during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis', *Journal of Cell Science*, Vol. 61, pp. 31–70.
- Frasch, M. *et al.* (1987) 'Characterization and localization of the even-skipped protein of *Drosophila*.', *The EMBO Journal*, 6(3), pp. 749–759. doi: 10.1002/j.1460-2075.1987.tb04817.x.
- French, V. (1988) 'Gradients and insect segmentation', *Development*, 104(SUPPL.), pp. 3–16.
- Fu, L. *et al.* (2012) 'CD-HIT: Accelerated for clustering the next-generation sequencing data', *Bioinformatics*, 28(23), pp. 3150–3152. doi: 10.1093/bioinformatics/bts565.
- Fujioka, M. *et al.* (2002) 'The repressor activity of *Even-skipped* is highly conserved, and is sufficient to activate engrailed and to regulate both the spacing and stability of parasegment boundaries', *Development*, 129(19), pp. 4411–4421. doi: 10.1038/jid.2014.371.
- Galtier, N., Gouy, M. and Gautier, C. (1996) 'Seaview and phylo\_win: Two graphic tools for sequence alignment and molecular phylogeny', *Bioinformatics*, 12(6), pp. 543–548. doi: 10.1093/bioinformatics/12.6.543.
- Gilbert, S. F. (2014) 'The genetics of axis specification in *Drosophila*', in Aquadro, A. (ed.) *Developmental Biology*. Tenth. Sunderland, MA: Sinauer, Andrew D, pp. 179–215.
- Gomulski, L. M. *et al.* (2012) 'Transcriptome profiling of sexual maturation and mating in the mediterranean fruit fly, *Ceratitis capitata*', *PLoS ONE*, 7(1). doi: 10.1371/journal.pone.0030857.
- Gonçalves, V. R. *et al.* (2013) 'Transcriptome analysis of female reproductive tissues of *Anastrepha obliqua* and molecular evolution of eggshell proteins in the fraterculus group', *Insect Molecular Biology*, 22(5), pp. 551–561. doi: 10.1111/imb.12045.
- Grabherr, M. G. *et al.* (2011) 'Full-length transcriptome assembly from RNA-Seq data without a reference genome', *Nature Biotechnology*, 29(7), pp. 644–652. doi: 10.1038/nbt.1883.
- Graveley, B. R. *et al.* (2011) 'The developmental transcriptome of *Drosophila melanogaster*', *Nature*, 471(7339), pp. 473–479. doi: 10.1038/nature09715.

Greene, C. T. (1929) 'Characters of the larvae and pupae of certain fruit flies', *J.Agric.Res*, 38(9), pp. 489–504. Available at: <file:///Y:/14293.pdf>.

Gregor, T., McGregor, A. P. and Wieschaus, E. F. (2008) 'Shape and function of the *Bicoid* morphogen gradient in dipteran species with different sized embryos', *Developmental Biology*, 316(2), pp. 350–358. doi: 10.1016/j.ydbio.2008.01.039.

Gutiérrez-Ramos, X. (2016) *Análisis citológico de la embriogénesis de Anastrepha ludens*. Universidad Autónoma de Querétaro.

Gutiérrez-Ramos, X. *et al.* (2020) 'Novel tephritid-specific features revealed from cytological and transcriptomic analysis of *Anastrepha ludens* embryonic development', *Insect Biochemistry and Molecular Biology*. Elsevier, 122(March), p. 103412. doi: 10.1016/j.ibmb.2020.103412.

Gutiérrez, J. M. *et al.* (2013) *Moscafrut, 'Los programas de moscas de la fruta en México Su historia reciente'*.

Haines, J. E. and Eisen, M. B. (2018) 'Patterns of chromatin accessibility along the anterior-posterior axis in the early *Drosophila* embryo', *PLoS Genetics*, 14(5). doi: 10.1371/journal.pgen.1007367.

Hales, K. G. *et al.* (2015) 'Genetics on the fly: A primer on the *Drosophila* model system', *Genetics*, 201(3), pp. 815–842. doi: 10.1534/genetics.115.183392.

Hazelrigg, T. *et al.* (1990) 'The *exuperantia* gene is required for *Drosophila* spermatogenesis as well as anteroposterior polarity of the developing oocyte, and encodes overlapping sex-specific transcripts', *Genetics*, 126(3), pp. 607–617.

Heffer, A. and Pick, L. (2013) 'Conservation and Variation in Hox Genes: How Insect Models Pioneered the Evo-Devo Field', *Annual Review of Entomology*, 58(1), pp. 161–179. doi: 10.1146/annurev-ento-120811-153601.

Hernandez-Ortiz, V., Guillén-Aguilar, J. and Lopez, L. (2010) 'Taxonomía e identificación de moscas de la fruta de importancia económica en América', in *Moscas de la fruta: Fundamentos y Procedimientos para su manejo*, pp. 49–80. Available at: [http://www.programamoscamed.mx/EIS/biblioteca/libros/articulos/HernandezOrtiz\\_etal\\_2010\\_2 .pdf](http://www.programamoscamed.mx/EIS/biblioteca/libros/articulos/HernandezOrtiz_etal_2010_2.pdf).

Herrera, A. L. *et al.* (1903) *Boletín de la Comisión de Parasitología Agrícola*. México: Ministerio de Fomento.

Hodar, C. and Cambiazo, V. (2018) 'The dorsoventral patterning of *Musca domestica* embryos: Insights into BMP/Dpp evolution from the base of the lower cyclorhaphan flies', *EvoDevo*. BioMed Central, 9(1), pp. 1–12. doi: 10.1186/s13227-018-0102-5.

Hsu, J. C. *et al.* (2012) 'Discovery of genes related to insecticide resistance in *Bactrocera dorsalis* by functional genomic analysis of a *De novo* assembled transcriptome', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0040950.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009a) 'Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists', *Nucleic Acids Research*, 37(1), pp. 1–13. doi: 10.1093/nar/gkn923.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009b) 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources', *Nature Protocols*, 4(1), pp. 44–57. doi: 10.1038/nprot.2008.211.

Jenner, R. A. and Wills, M. A. (2007) 'The choice of model organisms in evo–devo', *Nature Reviews Genetics*, 8(April), pp. 311–319.

Kalinka, A. T. and Tomancak, P. (2012) 'The evolution of early animal embryos: Conservation or divergence?', *Trends in Ecology and Evolution*. Elsevier Ltd, 27(7), pp. 385–393. doi: 10.1016/j.tree.2012.03.007.

Kimelman, D. and Martin, B. L. (2012) 'Anterior-Posterior Patterning in Early Development: Three Strategies', *Wiley Interdiscip Rev Dev Biol*, 1(2), pp. 253–266. doi: 10.1016/j.physbeh.2017.03.040.

Kozomara, A. and Griffiths-Jones, S. (2011) 'miRBase'. Available at: <http://www.mirbase.org/>.

Kumar, L. and Futschik, M. E. (2007) 'Mfuzz: A software package for soft clustering of microarray data', *Bioinformatics*, 2(1), pp. 5–7. doi: 10.6026/97320630002005.

Levine, M. and Davidson, E. H. (2005) 'Gene regulatory networks for development', *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp. 4936–4942. doi: 10.1073/pnas.0408031102.

- Leyva, J. L., Browning, H. W. and Gilstrap, F. E. (1991) 'Development of *Anastrepha ludens* (Diptera : Tephritidae) in Several Host Fruit', *Entomological Society of America*, 20(4), pp. 1160–1165.
- Li, W. and Godzik, A. (2006) 'Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences', *Bioinformatics*, 22(13), pp. 1658–1659. doi: 10.1093/bioinformatics/btl158.
- Loew, F. H. (1873) *Monographs of the Diptera of North America. Part III*. Edited by S. Institution. Washington D.C.: Smithsonian Institution.
- Lott, S. E. *et al.* (2011) 'Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-Seq', *PLoS Biology*, 9(2). doi: 10.1371/journal.pbio.1000590.
- MacDonald, P. M. (1990) '*bicoid* mRNA localization signal: Phylogenetic conservation of function and RNA secondary structure', *Development*, 110(1), pp. 161–171.
- 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(5), pp. 721–734. doi: 10.1016/0092-8674(86)90515-5.
- Macdonald, P. M., Luk, S. K. S. and Kilpatrick, M. (1992) 'Protein encoded by the *exuperantia* gene is concentrated at site of *bicoid* mRNA accumulation in *Drosophila* nurse cells but not in oocytes or embryos', *Genes and Development*, 6(4), pp. 2455–2466.
- Markow, T. A. (2015) 'The secret lives of *Drosophila* flies', *eLife*, 4(JUNE), pp. 1–9. doi: 10.7554/eLife.06793.
- McCarthy, D. J., Chen, Y. and Smyth, G. K. (2012) 'Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation', *Nucleic Acids Research*, 40(10), pp. 4288–4297. doi: 10.1093/nar/gks042.
- McPhail, M. and Bliss, C. I. (1933) 'Observations on the mexican fruit fly and some related species in Cuernavaca, Mexico, in 1928 and 1929.'
- Meccariello, A. *et al.* (2017) 'Highly efficient DNA-free gene disruption in the agricultural pest *Ceratitidis capitata* by CRISPR-Cas9 ribonucleoprotein complexes', *Scientific Reports*. Springer US, 7(1), pp. 1–11. doi: 10.1038/s41598-017-10347-5.

- Meinhardt, H. (1977) 'A model of pattern formation in insect embryogenesis', *Journal of Cell Science*, Vol. 23, pp. 117–139.
- Messenger, P. S. and Flitters, N. E. (1957) 'Bioclimatic studies of the mexican fruit fly', *California Avocado Society*, Yearbook 4, pp. 119–127.
- Mueller, O., Lightfoot, S. and Schroeder, A. (2004) 'RNA Integrity Number ( RIN ) - Standardization of RNA Quality Control', *Agilent Technologies*. Waldbronn, Germany, pp. 1–17. doi: 10.1101/gr.189621.115.7.
- Nirmala, X. *et al.* (2013) 'An EST database of the Caribbean fruit fly, *Anastrepha suspensa* (Diptera: Tephritidae)', *Gene*. Elsevier B.V., 517(2), pp. 212–217. doi: 10.1016/j.gene.2012.12.012.
- Ockendon, N. F. *et al.* (2016) 'Optimization of next-generation sequencing transcriptome annotation for species lacking sequenced genomes', *Molecular Ecology Resources*, 16(2), pp. 446–458. doi: 10.1111/1755-0998.12465.
- de Oliveira, J. L. *et al.* (2017) 'Evolutionary coincidence of adaptive changes in exuperantia and the emergence of *bicoid* in Cyclorrhapha (Diptera)', *Development Genes and Evolution*. *Development Genes and Evolution*, 227(5), pp. 355–365. doi: 10.1007/s00427-017-0594-3.
- Palomeque, T. and Lorite, P. (2008) 'Satellite DNA in insects: A review', *Heredity*, 100(6), pp. 564–573. doi: 10.1038/hdy.2008.24.
- Pane, A. *et al.* (2002) 'The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate', *Development*, 129(15), pp. 3715–3725.
- Papanicolaou, A. *et al.* (2016) 'The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species', *Genome Biology*. *Genome Biology*, 17(1), pp. 1–31. doi: 10.1186/s13059-016-1049-2.
- Pearson, W. R. (2013) 'An Introduction to Sequence Similarity ("Homology") Searching', *Curr Protoc Bioinformatics*, 1(03), pp. 1286–1292. doi: 10.1002/0471250953.bi0301s42.
- Peel, A. D. (2008) 'The evolution of developmental gene networks: Lessons from comparative studies on holometabolous insects', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1496), pp. 1539–1547. doi: 10.1098/rstb.2007.2244.

- Peng, W. *et al.* (2020) 'Transcriptome analysis of the oriental fruit fly *Bactrocera dorsalis* early embryos', *Insects*, 11(5), pp. 1–16. doi: 10.3390/insects11050323.
- Perondini, A. L. P., Selivon, D. and Morgante, J. S. (1998) 'Facultative polar extrusion of yolk masses and of hatching at the posterior egg pole in tephritid fruit flies (Diptera)', *International Journal of Insect Morphology and Embryology*, 27(3), pp. 249–254. doi: 10.1016/S0020-7322(98)00017-8.
- Peters, G. *et al.* (2013) 'Soft clustering - Fuzzy and rough approaches and their extensions and derivatives', *International Journal of Approximate Reasoning*. Elsevier Inc., 54(2), pp. 307–322. doi: 10.1016/j.ijar.2012.10.003.
- Petersen, M. *et al.* (2019) 'Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects', *BMC Evolutionary Biology*. BMC Evolutionary Biology, 19(1), pp. 1–15. doi: 10.1186/s12862-018-1324-9.
- Rabinowitz, M. (1941) 'Studies on the cytology and early embryology of the egg of *Drosophila melanogaster*', *Journal of Morphology*, 69(1), pp. 1–49.
- Raff, R. A. (2000) 'Evo-devo: the evolution of a new discipline', *Nat*, 1(October), pp. 302–307.
- Rezende, V. B. *et al.* (2016) 'Head transcriptomes of two closely related species of fruit flies of the *Anastrepha fraterculus* group reveals divergent genes in species with extensive gene flow', *G3: Genes, Genomes, Genetics*, 6(10), pp. 3283–3295. doi: 10.1534/g3.116.030486.
- Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2009) 'edgeR: A Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics*, 26(1), pp. 139–140. doi: 10.1093/bioinformatics/btp616.
- Rothwell, W. F. and Sullivan, W. A. (2000) 'Chapter 9: Fluorescent Analysis of *Drosophila*', in *Drosophila Protocols*. Cold Spring, New York: Cold Spring Harbor Laboratory Press, p. xiv + 697. doi: 10.1007/978-1-61779-034-8\_15.
- Ruiz, M. F. *et al.* (2007) 'The gene *doublesex* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects', *Development Genes and Evolution*, 217(10), pp. 725–731. doi: 10.1007/s00427-007-0178-8.
- Russell, M. A. (1985) 'Positional information in insect segments', *Developmental Biology*,

108(2), pp. 269–283. doi: 10.1016/0012-1606(85)90031-4.

Saccone, G., Pane, A. and Polito, L. C. (2002) 'Sex determination in flies, fruitflies and butterflies', *Genetica*, 116(1), pp. 15–23. doi: 10.1023/A:1020903523907.

Salvemini, M. *et al.* (2009) '*Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination', *International Journal of Developmental Biology*, 53(1), pp. 109–120. doi: 10.1387/ijdb.082681ms.

Salvemini, M. *et al.* (2014) '*De Novo* assembly and transcriptome analysis of the mediterranean fruit fly *Ceratitis capitata* early embryos', *PLoS ONE*, 9(12), pp. 1–27. doi: 10.1371/journal.pone.0114191.

Sarno, F. *et al.* (2010) 'The gene *transformer-2* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects', *BMC Evolutionary Biology*, 10(1). doi: 10.1186/1471-2148-10-140.

Schindelin, J. *et al.* (2012) 'Fiji: An open-source platform for biological-image analysis', *Nature Methods*, 9(7), pp. 676–682. doi: 10.1038/nmeth.2019.

Scolari, F. *et al.* (2014) 'How functional genomics will impact fruit fly pest control: The example of the Mediterranean fruit fly, *Ceratitis capitata*', *BMC Genetics*, 15(Suppl 2). doi: 10.1186/1471-2156-15-S2-S11.

Selivon, D. *et al.* (1996) 'Extrusion of masses of yolk during embryony development of the fruit fly *Anastrepha fraterculus*', *Invertebrate Reproduction and Development*, 29(1), pp. 1–7. doi: 10.1080/07924259.1996.9672489.

Selivon, D., Morgante, J. S. and Perondini, A. L. P. (1997) 'Egg size, yolk mass extrusion and hatching behavior in two cryptic species of *Anastrepha fraterculus* (Wiedemann) (Diptera, Tephritidae)', *Brazilian Journal of Genetics*, 20(4), pp. 587–594. doi: 10.1590/s0100-84551997000400006.

SENASICA (2020) 'Moscas Nativas de la Fruta'. Available at: <https://www.gob.mx/senasica/documentos/moscas-nativas-de-la-fruta-110869>.

'Serial Cloner' (2020). Francia. Available at: [http://serialbasics.free.fr/Serial\\_Cloner.html](http://serialbasics.free.fr/Serial_Cloner.html).

Shen, G.-M. *et al.* (2011) 'Transcriptome analysis of the oriental fruit fly (*Bactrocera*

*dorsalis*’, *PLoS ONE*, 6(12). doi: 10.1371/journal.pone.0029127.

Sievers, F. *et al.* (2011) ‘Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega’, *Molecular Systems Biology*, 7(539). doi: 10.1038/msb.2011.75.

Simão, F. A. *et al.* (2015) ‘BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs’, *Bioinformatics*, 31(19), pp. 3210–3212. doi: 10.1093/bioinformatics/btv351.

Simcox, A. A. and Sang, J. H. (1983) ‘When does determination occur in *Drosophila* embryos?’, *Developmental Biology*, 97(1), pp. 212–221. doi: 10.1016/0012-1606(83)90078-7.

SIRVEF (2020) ‘Mapa Dinámico Fitosanitario’. Available at: <http://sinavef.senasica.gob.mx/MDF/>.

Small, S., Blair, A. and Levine, M. (1992) ‘Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo.’, *The EMBO Journal*, 11(11), pp. 4047–4057. doi: 10.1002/j.1460-2075.1992.tb05498.x.

Smit, A. F., Hubley, R. and Green, P. (2019a) ‘RepeatMasker open-3.0’.

Smit, A. F., Hubley, R. and Green, P. (2019b) ‘RepeatModeler open-1.0’.

Society for Developmental Biology (2020a) *The interactive Fly. Maternally transcribed genes*. Available at: <https://www.sdbonline.org/sites/fly/aimain/4matrnl.htm>.

Society for Developmental Biology (2020b) *The interactive Fly. Zygotically transcribed genes*. Available at: <https://www.sdbonline.org/sites/fly/aimain/5zygotic.htm>.

Stone, A. (1942) ‘The fruitflies of the genus *Anastrepha*’, *U.S. Department of Agriculture. Miscellaneous publication*, 439, pp. 6–8. doi: ia:fruitfliesofgenu439ston.

Struhl, G., Johnston, P. and Lawrence, P. A. (1992) ‘Control of *Drosophila* body pattern by the *hunchback* morphogen gradient’, *Cell*, 69(2), pp. 237–249. doi: 10.1016/0092-8674(92)90405-2.

Tautz, D. *et al.* (1987) ‘Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes’, *Nature*, 327(6121), pp. 383–389. doi: 10.1038/327383a0.



- Tautz, D. and Pfeifle, C. (1989) 'A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*', *Chromosoma*, 98(2), pp. 81–85. doi: 10.1007/BF00291041.
- De Teodoro Pardo, C. V. (2011) *Evolución de genética de cuatro especies de Anastrepha spp. (Diptera: Tephritidae)*. Colegio de Postgraduados.
- Tomancak, P. *et al.* (2007) 'Global analysis of patterns of gene expression during *Drosophila* embryogenesis', *Genome Biology*, 8(7), pp. 1–24. doi: 10.1186/gb-2007-8-7-r145.
- Tsoumani, K. T. and Mathiopoulos, K. D. (2012) 'Genome size estimation with quantitative real-time PCR in two Tephritidae species: *Ceratitis capitata* and *Bactrocera oleae*', *Journal of Applied Entomology*, 136(8), pp. 626–631. doi: 10.1111/j.1439-0418.2011.01684.x.
- Ungaro, A. *et al.* (2017) 'Challenges and advances for transcriptome assembly in non-model species', *PLoS ONE*, 12(9), pp. 1–21. doi: 10.1371/JOURNAL.PONE.0185020.
- USDA-ARS (2015) '*Bactrocera dorsalis* Genome'. Available at: [https://www.ncbi.nlm.nih.gov/genome/annotation\\_euk/Bactrocera\\_dorsalis/100/](https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bactrocera_dorsalis/100/).
- Wiegmann, B. M. *et al.* (2003) 'Time Flies , a New Molecular Time-Scale for Brachyceran Fly Evolution Without a Clock', *Systematic Biology*, 52(6), pp. 745–756. doi: 10.1080/10635150390250965.
- Wiegmann, B. M. and Yeates, D. K. (2017) 'Phylogeny of Diptera', *Manual of Afrotropical Diptera*, 1(January), pp. 253–265.
- Wilczynski, B. and Furlong, E. E. M. (2010) 'Challenges for modeling global gene regulatory networks during development: Insights from *Drosophila*', *Developmental Biology*. Elsevier Inc., 340(2), pp. 161–169. doi: 10.1016/j.ydbio.2009.10.032.
- Xu, P. *et al.* (2003) 'The *Drosophila* microRNA mir-14 suppresses cell death and is required for normal fat metabolism', *Current Biology*, 13(2), pp. 790–795. doi: 10.1016/S0960-9822(03)00250-1.
- Yin, C. *et al.* (2016) 'InsectBase: A resource for insect genomes and transcriptomes', *Nucleic Acids Research*, 44(D1), pp. D801–D807. doi: 10.1093/nar/gkv1204.

## **Anexo 1. Soluciones y métodos complementarios**

### Medios y soluciones

#### Solución embryo wash 10x

Para 1 L utilizar 0.7% de NaCl y 1% de Triton x-100, no es necesario esterilizar y se guarda a temperatura ambiente.

#### Medio de piloncillo para colecta de embriones de *Drosophila*

Para 2 L, calentar 600 mL de agua hasta casi punto de ebullición. Disolver 200 g de levadura y 200 g de piloncillo, después agregar 400 mL de agua tibia. Agregar 18 g de agar, mezclar hasta homogenizar y autoclavar. Después de la esterilización, adicionar 28 g de grenetina previamente hidratada en 33 mL de agua tibia. Mezclar hasta homogenizar y esperar a que se atempere para adicionar 500 mL de jugo de manzana y 20 mL de ácido propiónico que funciona como antimicótico. Mezclar y vaciar en platos para realizar la colecta de embriones.

#### Búfer ISH para la hibridación *in situ*

Para realizar el búfer se necesita 50% de formamida, 5x de SSC, 15 mM de heparina, 0.1% de Tween-20 y 100 µg/mL de DNA de esperma de salmón. Proteger la solución de la luz y guardarla a -20 °C.

#### Búfer de fijación

Para el búfer se necesita 0.1 M Hepes pH 6.9, 2 mM de sulfato de magnesio y 1 mM de EGTA pH 8. Esterilizar por autoclave y guardar a temperatura ambiente.

#### Stock de bromuro de etidio (5 mg/ml)

Para la solución pesar 15 mg para diluirlo en 3 mL de agua destilada estéril, mezclar y guardar a temperatura ambiente protegido de la luz.

#### Stock de glicina 20 mg/mL

Para la solución pesar 60 mg para diluirlo en 3 mL de agua destilada estéril, mezclar y guardar a temperatura ambiente.

#### Solución de detección (para hibridaciones *in situ*)

Para la solución se necesita 100 mM de NaCl, 50 mM de MgCl<sub>2</sub>, 100 mM Tris pH 9.5, 1 mM de lavamisol y 0.1% de Tween-20. Hacerla al momento de uso para evitar la precipitación de las sales.

#### Solución PBT

Para la solución utilizar PBS (para hibridaciones *in situ*) a 1x con 1% de Tween 20, mezclar y guardar a temperatura ambiente.

#### Solución PBS 10x (para hibridaciones *in situ*)

Para 1 L utilizar 75.9 g de NaCl y 14.1 g de fosfato de sodio a un pH de 7. Esterilizar por autoclave y guardar a temperatura ambiente.

#### Solución SSC 20x

Para preparar 1 L utilizar 175.3 g de NaCl y 88.2 g de ácido cítrico. Esterilizar por autoclave y guardar a temperatura ambiente.

#### Membrana para ácidos nucleicos

La membrana utilizada es de las siguientes características: positively charged nylon transfer membrane, Hybond-N+, Amersham Biosciences, Cat# RPN303B, removal rating 0.45 µm.

#### Solución DIG1

Para la solución es 100 mM de Tris pH 7.5 y 150 mM de NaCl. Mezclar y almacenar a temperatura ambiente.

#### Solución DIG3

Para la solución es 100 mM de Tris pH 9.5, 100 mM NaCl y 50 mM MgCl<sub>2</sub>, la cual se realiza al momento de utilizarla para evitar que se precipiten las sales.

#### Solución NBT/BCIP

Para la solución se realiza con las pastillas de NBT/BCIP ready-to-use tablets (Roche, 1697471), diluyendo una pastilla en 10 mL de agua, guardar a 4 °C. Es viable la solución en un promedio de 2 meses.

#### Solución de incubación RNA

Para 1 mL se utiliza 657  $\mu\text{L}$  de formamida, 210  $\mu\text{L}$  de formaldehido 37% y 133  $\mu\text{L}$  de MOPS 10X. Esta solución se debe de preparar en el momento o se puede almacenar a  $-20\text{ }^{\circ}\text{C}$  de forma indefinida.

## Métodos complementarios

### Purificación de producto PCR por gel de agarosa de bajo punto de fusión

1. Cortar las bandas de los productos de PCR del gel de agarosa de bajo punto de fusión en fragmentos pequeños y ponerlos en tubos eppendorf 1.5 mL de tal manera que no ocupen más de 500  $\mu\text{L}$ .
2. Fundir por 10 min a  $65\text{ }^{\circ}\text{C}$ , asegurándose que la agarosa quede fundida y transparente. Adicionar 500  $\mu\text{L}$  de buffer, calentar por 10 min a  $65\text{ }^{\circ}\text{C}$ .
3. Adicionar 500  $\mu\text{L}$  de fenol. Vortex y centrifugar a 13 000 rpm por 10 min.
4. Recuperar la fase acuosa (aproximadamente 450  $\mu\text{L}$ ). Adicionar 500  $\mu\text{L}$  de fenol y 500  $\mu\text{L}$  de cloroformo. Vortex y centrifugar a 13 000 rpm por 10 min.
5. Recuperar la fase acuosa (aproximadamente 450  $\mu\text{L}$ ). Adicionar 500  $\mu\text{L}$  de cloroformo. Vortex y centrifugar a 13 000 rpm por 10 min.
6. Recuperar la fase acuosa y adicionar 1/10 de acetato de sodio 3 M pH 5.3 y 1  $\mu\text{L}$  de glucógeno. Vortex 5 s. Adicionar 2 volúmenes de etanol 100%, incubar ON a  $-20\text{ }^{\circ}\text{C}$ .
7. Decantar el etanol y lavar la pastilla con 500  $\mu\text{L}$  de etanol 70%. Mezclar y centrifugar a 12 000 rpm por 5 min.
8. Repetir paso 15 y 16.
9. Decantar el etanol y dejar secar la pastilla a TA hasta que se evapore el etanol.
10. Resuspender la pastilla en 20  $\mu\text{L}$  de agua.

### Transcripción *in vitro* de sonda de ARN

A partir del producto de PCR realizar la transcripción *in vitro* a partir de 500 a 1000 ng, la reacción es la siguiente:

- 500 - 1000 ng DNA
- 2  $\mu\text{L}$  Búfer de transcripción 10X (dejar descongelar en hielo ya que cuenta con ATP)
- 2  $\mu\text{L}$  DTT 0.1 M

- 1  $\mu$ L RNase Out
- 2  $\mu$ L RNA pol T7
- 2  $\mu$ L DIG RNA Labeling Mix, 10X (Roche, 11277073910) (dejar descongelar en hielo)

Para una reacción total de 20  $\mu$ L, la cual se incuba a 37 °C por aproximadamente 16 h. Una vez terminada la incubación, se toma 1  $\mu$ L de la reacción de transcripción *in vitro*, la cual se guarda en un tubo eppendorf a -20 °C. El producto restante (19  $\mu$ L) se guarda a -20 °C.

## Anexo 2. Esquema del análisis bioinformático de datos de secuenciación de *A. ludens*

### 1. Generación de datos

- Generación de librerías
- Secuenciación por Illumina 4000 paired-end
  - 3 etapas de desarrollo
  - 3 replicas biológicas por etapa
- Análisis de calidad de secuencias
  - Por muestra
  - Por replicas biológicas de cada etapa

### 2. Ensamble de los datos

- Ensamble *de novo*
  - Por etapa de desarrollo
  - Por etapa de desarrollo y después concatenada
  - Por todas las etapas juntas
- Análisis de calidad de ensamble *de novo*
  - Análisis estadísticos
  - Análisis de similitud con genes conservados en artrópodos

### 3. Anotación de los ensamblados

- Búsqueda de ORFs
- Anotación por similitud
  - Por proteínas de *D. melanogaster*
  - Por proteínas de base de datos de BLAST
- Anotación de secuencias repetidas
  - Por similitud de repetidas
  - Por modelado *de novo* de familias de repetidas
- Anotación de pre-miRNAs
- Anotación de dominios PFAM

### 4. Análisis de expresión diferencial

- Análisis de presencia o ausencia por valor de FPKM
- Análisis de expresión diferencial entre etapas de desarrollo

### 5. Comparación con datos de diferentes moscas

- Datos transcriptómicos de embriones
  - De *D. melanogaster*
  - De *C. capitata*
  - De *B. dorsalis*

### Anexo 3. Oligonucleótidos

En la siguiente tabla se muestran los oligonucleótidos utilizados en este trabajo. En color verde se muestra la secuencia del promotor de la ARN polimerasa de T7.

<b>Gen</b>	<b>Sentido</b>	<b>Secuencia (5' - 3')</b>
<i>exuperantia</i>	Forward	GAA ATA ATA CGA CTC ACT ATA GGG AGA CAT GAA GAA CGA CGG CAG CAA T
	Reverse	GAA ATA ATA CGA CTC ACT ATA GGG AGA GGC GGA GAT GGC GAC TTA ATT G
	Forward	CAT GAA GAA CGA CGG CAG CAA T
	Reverse	GGC GGA GAT GGC GAC TTA ATT G
<i>hunchback</i>	Forward	GAA ATA ATA CGA CTC ACT ATA GGG AGA CCG AAA TCG CCC AAA TCT AGG C
	Reverse	GAA ATA ATA CGA CTC ACT ATA GGG AGA GAG ACT TGC GGT GCG AAT TCA A
	Forward	CCG AAA TCG CCC AAA TCT AGG C
	Reverse	GAG ACT TGC GGT GCG AAT TCA A
<i>even-skipped</i>	Forward	GAA ATA ATA CGA CTC ACT ATA GGG AGA TTG CAG CGT TCA ACA TGC TAG G
	Reverse	GAA ATA ATA CGA CTC ACT ATA GGG AGA TGG CCT CAC TAC TTA ACC TGC C
	Forward	TTG CAG CGT TCA ACA TGC TAG G
	Reverse	TGG CCT CAC TAC TTA ACC TGC C
<i>engrailed</i>	Forward	GAA ATA ATA CGA CTC ACT ATA GGG AGA GAG GAT CCG CAC GCG TCG TCA CAT
	Reverse	GAA ATA ATA CGA CTC ACT ATA GGG AGA AGG AAT TCA GTT CGG GTC CAC GCT
	Forward	GAG GAT CCG CAC GCG TCG TCA CAT
	Reverse	AGG AAT TCA GTT CGG GTC CAC GCT
<i>histona H3</i>	Forward	ATC AGT TTT GAA ATC TTG AGC GAT
	Reverse	TAC TAA GCA GAC TGC TCG TAA ATC

#### Anexo 4. Secuencias de los genes anotados de *A. ludens*.

De las secuencias de los genes anotados de *A. ludens* se realizó el diseño de las sondas; cada gen muestra los dominios importantes en su proteína, así como la zona en donde hibrida la sonda de RNA.

#### Gen *exuperantia* de *A. ludens*

```
GTTCGTAGTTGGATTTTTTTTTATCGTTATTAAGCAAGCAATCAAACATGTGATAGACTACTCAGCTTGGAAAGTTTTCTTTTGAATTTGTTTGAATT-100
TCTATTCAAGTTTTATTTTTTAAATTTTAAGAAAAATCACTTGCTAAATTCAGTTTCTTTCTATATTCACAGGGATTAGATTGTCTTTAAGAAAAATT-200
TCGATTCAAGTACGTTAGAAAAATTCGTTTGTAGCAGAGAGAAAAGGAGTTCAGCTGATTGAGCGCATTTAACTAGTCAAATTAGTCCGATTTTCGCTTGG-300
ATTATTTTTATTGGAAATTTAGAAAATGTCGCTTCAACGAATGTAAGCCAAAACGGTGTAGTCAATCTCCAGTAGTAGGAACCTCTAATGGAGTAAC-400
AACCAATGGCACAATGAATGTGCTAGCGGACACAAGACTTTGGCCAAGGGCAGGTACACTTTAGTAGGCGTCGATATTGACACAACCTGGGCGTCGACTC-500
ATAGATGAAATCGTTCAGTTGGCCGCGTATTCGCCAAAGATCATTTTTGAACAATATATTTATGCCATACATGAACCTTGAATCTGTGCTGCGCAACGTC-600
ACCAAGTCCGTGTGATCTCTATTGGATTCTATCGTATGTTGAAATCAATGCAAACCTATAAGATTATCAAGTCTAAATCGGAGATCGCTGCCTTGAAAGA-700
TTTTCTGGATTGGCTGGAACGCGTTAAGGCCGAAGATACATCATCGAACGGTGTCAATCTAGTCTATCATGAGGAACGCAAATTTATTCATACATGATT-800
TTGGAATCGCTTAACAAATACAACCTGTTGGACAGGTTCCTTAAAACAGTCAAATCATTTGCCAATGGCCTCAATTTAGCCAAGCCAATGCCAACGAAT-900
CATTGCAACATTACAGTTGCGCAAACCTCCAAAAGTGTGAGCGCCAAAAGCACATGAAGTGATCAATAATCAAATGCAAGTGACGGTGAATCAAAAA-1000
CAATGTTAGCTCAACAACGGGCATAGCCGAGTGTGAGGCCCCAAAAGTCAATGCAAAAAGGGCGCTCTGGCACAGGACGCTGAGCTGTTGATGGTAAT-1100
GCTAGTGTGCGTGCCAAGTTGGCCTATGAGGTGGCACAACAGTTGAGCAATCAAGAACAACAGGTGCCTATCGAATCGTCGGCTGCATTAGAGAATATGA-1200
TTGCCGAGTGCAACCGTATGCACAGCCCATGGAATGCGATTGGAATGAACTACAAACTCAAACGAGAATTGGAACGTCAAAAATCTTTTCGGTCCGGT-1300
TTTCTGAATTACTTTAAAACCACTCTACCATCGAGTGCCTGCGGTCAAGTTTCGCATCATACTTGCAGAACAGGGTTA(GACCTGCAGACGCTAAAT-1400
GCAATTTGGAGTGAGAAACAAAAGGATGGATTGGCCGAAGTATTGCAAGTCAATTGAGGAGTTAAAACCTGAAGATAGGACGGAGTTGGCTGATTTATTG-1500
ACAGTTACTTCGATCCTGTAAAGACGACTATTAAGCCGATTGTCAAAAACAACAATCGCCGCCGAAATAGACGCGGAGTACGCAACATGAAGAACGACGG-1600
CAGCAATAGCGCTCGCACAAGTAGTGAGCAAGGAGCTGGCGGTGACAAGCAAGTCACTGGCGGTGAGCTGCCCGACACCACAATTAAGTGCCTATCTCCG-1700
CCAAAACGGCGGCAAGCCGAGCGTGGCAAGCGTAATCCGCGCTCACGCAATAACTCCTTAAAAACACAGCTGCCAAGCGCCGCGCTCG(CACGAGCAATA-1800
CAACGGCAACAGCAACCAATTAAAAACCTAAGAGTTTCATTCTAAGATTCAAGACAGGCGCTGTTGCAACATAAAAAACAAGTAATTTGTTTTATTTTT-1900
AAATTGTGGG
```

Regiones: 5'UTR-ATG-Región codificante-Exuperantia SAM-like domain-Secuencia de sonda-3'UTR



## Gen hunchback en *A. ludens*.

AGGCTGAATATAGAAATTA AAAAGCAAAC TGCAGTTGCGCTTTGATATTGTGGCGATAAATG GAAAAC TGGGACCATGCAACCATGCAAGCTGCATCTG-100  
GCAACTACGAACAACACAGCCCTTGGTACAACAATATGTTTCCGCCACAAATATCAAGAGCGAGCCTGGAAC TACCGCTGCCTCGCCGGAGCAGCAGAA-200  
GCCAATGGCGTCATCATCTAATGTGGATATATTGGCGCACTTCGAAAATTTGTAAAAACAACAGCACCAACAACAACAACAACAGCGCCCGCAAACG-300  
CCACGCCAGACATCGTTGATGGACACGCTAACAGCAATGACACCATCGCCAGCAACACGGGCAGCATGCAACAGTTGTTGCAGCAACAACAATACCAGC-400  
AGCAGTTTCAACAAGCCCAGCAAGCCGCTGCACAGCATCAGATCGCATCTTCGCATTTACCGCTCGGTTTTGATCCACTTACACGCCAGGCTTACCCAAA-500  
TGCCATTTTACCATCCATATCCCAATTTTATCACCATCAGTCGCACAACAATATGCCAATGGCCAACAACAACCCAGCCCCAAACCACCTCAAACGCAC-600  
ATAGATTCTAATGAGAACTGCGCACGGCTTTGACGCCACGCCCTTACTCCGCCAATGGATATCACGCCACCGAAATCGCCAAATCTAGGCCCTTAATGT-700  
CCACCAAAAAGATGGAAGTGCACGAGAAAGAACCGATTTTCAGTCTAGCAGTGAAGAAATGAAGAGTGTAAATGAGTCAGAGGACGACGAGTCGATTTCG-800  
GTTACCCATTTACAATTCGCATGGCAAAATGAAAACTCAAGTGCAGTCTGCAACATGGTTGGGATAACAAAAGTAAGCTTCTGGAACATGTACGC-900  
ACACATATGAAGCCGGAGAAGATTTACAATGCTCCAAATGCCGTTTGTACTGAAGTGAAGCACCCTCGAATACCACATACGCAAAACAAAAACA-1000  
TCAAGCCATTTCAAGTGCAGCAATGTTTCATACACTTGGCGTCAACAAGTCGATGTTGAATTGCGACCAGTCTCTCTTCCGTATACCCGTACCGTTG-1100  
CGCGGACTGTGATTATGTACCAAGTACTGTCTATTCAATTCGATCTACGCAAGTATGGCCAATCCGGGTGTGGTGTAGATGACGAAGGAATT-1200  
CCGAATCCGTCAGTGGTGTGACGTATATGGCACACGGCGGACCTAAAGCAAAGCCTTCTGCAACTAACAAACGTAGCAATAGCAAAAACAATGAAC-1300  
GCGGAAATGCAAAAACGCAAAACACCCAGAAAAGGCCAAAGACTTTAAATGTCCACATCCCAGCTTTCGGCAGCGCTACAATCTTTCAGTTGCCCTC-1400  
TCAACACCATCAGATGCCTATCTCACCAGCTCAGAGCCCATGCTCAATAGCCTGTGACCTGCCACGCAGTCCATTGATTTGTCTCAACAAC TAACCCAG-1500  
CAGCTGACGAAGGGTCTGAACAAACACAATCAACTTTTCAACAGAGATGGCAGTGAACAAAGCTTAAGGGTCAAAAATTCAAAAGCCTCAACCCGCA-1600  
GCACTACTATTCAGAAACGAAAGCGTCCATTTCAATCTTCTACCGCCACTAGCTACGATTTTCCAGCAGATGCCGGTATGGCGTTGTACCTTATTG-1700  
GAATTTGCACGTGCTGGCCGCAACAGCAAGCCCGCTTTTAGCCAAATACCAGCCAGTGCAGCGGAGAAAGCACTACGACAAC TGCAGAACTCCTCAA-1800  
TTAAATAAAGATGGTTATGATTCGTCACACGAAGATGTGATGAAGATGGCGGGAATGAGCTTGAACGGAGATCATCAGGCTCGCGGATGGACCTCTCAC-1900  
AAGGTTTACCTGTAAAAGATGAGGTCGAGCATTCTAAAACGCCCGCAATACAGGCTAATTTAAAAGCTCAAAATGAAACCAACGCCACGCCAATCATTAC-2000  
CGCAACGGCGCTCAAGACGAAAGGGTCTGTACTCAAATGAGGCCATACTGCAAGTGAAGTCTGCTTTGCAAGCAACAGAAAGCAAGCAACACCAA-2100  
GCTGTGAGTCTGTTAAAGAGCGATGGCGAAACAAGTATCCGACGCTTCATCGCGATTCCCTTCATTTGCTGCCAATTCGCTGACCAAGTTGAAGAAACT-2200  
CTGCTGAATGTGATGAACCTTATCAATAATATGACGCCTACTGCGAAAGATACCAACACTAACTCTGCAGCATCGTCATCCGGAAACAGTTCAAATAGTTG-2300  
CAGTACTTCTTCAGCTACGATATTAATGACCTGCTGCATTTGGCTCTATGATGAATGCAAAATTTGTGATATTTACTTCAAGGATGCTGTGCTCTAT-2400  
ACAATCCACATGGGCTATCACAGCTGCGACGACGCTTCAAGTGTAAATGTGCGGGCAGAAAGTGTGACAGCGCTGTGCGCCTTTTCGTTTCATATGGCGC-2500  
GAAATGCGCATTCATAAGTGGAGTTCTGCACCTTAAGCATTGGGTGCATGCGATCCGCTTTGCTGGAAAATGCGAGCATAGATTCAAATTTGTAATAT-2600  
ATGCGAGTATTAGTCGTAGTTCTAAGTTATTAAGGTGAAGCAATAAATTTATGCGCATTTTTCGTTATAAGAAATTTAGTTTCAAATCAAAATGAAT-2700  
CATGGATCTTAAGTAAAAAATGTGTAATGCGTTAAAAAATCAACAACGCAAAATAATTTTCGAAAAAATAAAAAATTTGATTGTAATATG-2800  
TGCTTAAGTGTAACTTGATTATAT

Regiones: 5' UTR-ATG-Región codificante-Dominio Zinc finger-Secuencia de sonda-3' UTR

## Gen even skipped de *A. ludens*.

ATGCCCGTTCGGAGGGACATGCAAATTAATGACGCCAAACAGGACACATCCGCTAAGCATTGAGGAAAAAGAGCCGTTTCGAAAACCAAGTCGAAAAAGT-100  
CTCTTAAATTCATAGCGAAAATATGCAAGGCTATAGACCTTACAATATGGACGGCACCATGCGTCACATCATCACGACACGCGGATCAAAGCCGTT-200  
AGTGGTGGATTTAATTGCCACGCAGTACGCGAAATCATGCACTCCGCCACCATCACCAAAATGAATGCCACTCCAGCCCCGATACGTCAGTACTGAGCAATGGT-300  
CGCACGAGTAGTGGCGAAATTCCTACAGATCCTTCGCTACGACGATACCGCACCGCTTTCACACGGGATCAACTTACACGCTTTGAAAAGGAGTTCCTTA-400  
AAGAGAATTATGTCTCCGTCACGTCGTTGCGAATGGCCCTACTGCTTAACTTCCCGAGTCAACAATCAAAGTTTGGTTTCAAATTCGACGCATGAA-500  
AGACAAGCGCAACGCATTCGCGTCGATGGCCCTACGCGCCGCTTACTCTGATCCCGCATTGCGCGCCTCGTACTACAAGCCGCCCAACAGTGT-600  
GGCATAACCTATCCGACCTATCCGCGCCAGGTGCGAATCCTATGCTAGCCGCATCTGTATGCCAATGGCAAGTCATTACGCGCCCTACCGCTATAACC-700  
CCTATCAATGGCCGTCGCTCCCGTCCGCCACCAATGCAACCACATCACCACGGCATAACCAGCACACGCTCCGCTCCAGCAGTTGCTGCCGTTATGCA-800  
TAATGGTCCCTTGGCTGACGCCACTGCCGTGGATACTTAGCATGTTGAACGCTGCAATGGCGCCACATCAGAGCTATGCCCATTTGGTGGCCATACCT-900  
TCAAAGCAACACACACCGCCATTTGATTTGCAATCGTCCGCTTCCGCACACTCCTCTACACTTTTCGCTGAGTCCCGACGGCTCGGACCACCAAAGTGT-1000  
TTGATCGCTCGCCACCAGCCCGTCGCAATTCAGCCACCAGCGGCAGCAGTGGTGTGAAATTTATGGCGTAGCACTTGAGAGCAACACCAGGACGGC-1100  
TAACACTTTGTACCGCAGCGAAAGCCCATGACAACGCCAACGCTGCCACCAGCAGCATCAATGAAATGAATCACCACCGCAGCCCCACACAGTCGGTT-1200  
GGTCTTTTGATGACGACGAATAATAATAGTCTTAAACGTTCCGCAAGTTCGCTCTCGCCACGCAAACCATTTATCGCAGAGCCTAAACCAAAGTTATTCA-1300  
AGCCCTACAAAACGGAACTTAA

Regiones: 5' UTR-ATG-Región codificante-Dominio Homeodomain-Secuencia de sonda-Dominio LFKPY-3' UTR

## Gen engrailed de *A. ludens*.

GTTCGCCGTGAACAGACGCTGTCGCTGCTAGTCTGTTTGGAGAGCATGTACAAAGAAGTAAAGAAGACGGAAGCAATACAAATAAGAAAAA-100  
TAAACGTGTTTTCTCTTTAGCCAACAAACAAGTGTTCGAGTAGAAGCAAATGAACTTCAATGAAAACGTTGAAACGCTACGGATTTGCGTGAAGCGAAA-200  
AAAAAGAAAACCAAAAACAAGTGTGAAAGTATAAGTGTGAAAGTAAAGCAGGGAATTAACAATTCATAAATAAAAATTAAGCCATTTTC-300  
TTAAGCAAATTTACAAATCGTTATCTTATGGCAGTGTGACAATTTTAAACTTAAACAATATGAAATTAACATTTGAAATTTCAAATAGTGT-400  
CGGTGAGAAAATTTTGAATTCGATTTGCGTCCGCGCCAGCTCAACGCAGCTGCATCATCAAATTTGCCACGGAACTACCAAGGCCCTAGAGGATCG-500  
TGCAAGTCCGCAATTCGCTCCAAGTCCGCTGGTATACCACAATCGCCGACCATCACCAGCCGAAACAGCAACCAACAACAATCACAACATTTACAAT-600  
ACACTCCTACTGCCAGCACAGTTCGGATATGAGCCTGCACCCACATCCGCCACCAGCCACTACCACATGTCATAGCCGCCCCGCAACGCTCCACAAT-700  
GCTGCCACACCCCGTGAATACGGTCCGCTACAAATGCAGTACCCTGCCACACTCTACCATCAACATCAGCAGCTCCAACTTTGCATCATTTGCAACAA-800  
CTGCAGCAGTGCACCAGCAACAATGGTGAATGCCGCGCCTTTCATCATCATCTGGTTTTGATGCTGCTGCAGCAACTGCTTTAGCCGACGCCAATC-900  
ATGCCGCTCATTTGCATCAGCGTTTGGCTGCCCAAGTGGCTCCTCTACACCCTCATCTACACTGCTCCACAATGACCATCAAGGAAGAGGAAAG-1000  
TGATTCGATTATAGAAGATAGCTTTCATGCGGATGCCGAAACGGCGACACTTAATAGCAAAGCTCAAACGCGCTACACTCCACAGAAGACGACGAGGCG-1100  
GAGGAGGATGAAGAGGACGATATCGATGTGGATGATGTGGACGATACTCCAAACGGCGCTTGTCAACGTACGCCACCACCCGCACATCAACAATCCGCTA-1200  
AGCCAAACATCGCCTTTTCCATTTCCAAATATACTCAGCGATCGCTTTGGCGGCMATGAGGCGAGTGGCCGCGGTGCCAATAGTGTGCTACCACAGCGGC-1300  
GAGCAAACAAACGCATTTGAATGCGAATAGTATTTCCGGCCCTTTGATGCCGACGCTTGCACACCCTCCGCTTTACGCGAGTGCATCTCCTGGAG-1400  
CTGAGTGCACAGCAGCAAGCCGCTGCGGCCGCGCCACTGCTATGATGCTGGAGCGGCCAACTTCTTGAAGTCTTCAATCCCGCAGCATAATCCGCTGA-1500  
TACATGAGGAGATTGTCAACAGTCCGATCGCTCGCAGTGGCGTCTACCAGCCGCTGCGGCAAAATGTACGAGCCCGGACTCATTAGTAACCGCAAT-1600  
CTCGCCAGCAGCCGTCGCGAGTGTAGGCGCATTTGGAGAAGTCCGCTTTGGGTTTCGCTTTGTAAGGCAGTCTCACAAATTTGGCCAGTCTGTAGCGGCACCA-1700  
CCATGTTCAAGCACTTCTCCGCGCAAGTGCATCAAGCATCGCCAGTCCACCACCCGCTCCAGTGCCTGCGGCGCATCTTCTTCGGCATCATCGTCCA-1800  
CATCGCGGGTGTTCATTAATTTCTCGCCACCAGCCGCTGCCAGTAAGAGTAGTAGTCCACAACCATAACCAGCCGCTCGGCCGTCAGCCGCTGA-1900  
CTCGGGCATGGAGTCTCCGATGATACAGCTCTGAAACCGGCTCTACGACAACGAAGTGGCAAAAATGAAGTGTGCCCGCGTGGGTGTATTGCAGC-2000  
CGCTACAGTATCGACCAAGTTCGGGTCCACGCTATCGCCGCCCAAGCAGCCGAAAGACAAAATTAATGACGAGAACGCGCCGCGCACCCGATTTTCCA-2100  
GTGAACAGTTGGCGGTTTGAAGCGCAATTTAACGAGAATCGTTATCTTACCAGCGCTCGCCAACTTGAATTCGGAGTTGGGTCTGAATGAGGC-2200  
GCAATCAAGATTTGGTTCAGAAATAAACGCGCTAAAATTAAGAAATCTTCAGGTTCTAAAACCCGCTCGCCTTGCATTTGATGGCACAGGCGCTGAC-2300  
AATCACACAACAGTCCGCTGACCAAGAGGAGGAGGAGCTGGAGATGCGTATGAGATGCTCTGGGTGAGCAAGAGTTGAGTTGAAGAGAAAGGCGGAA-2400  
GGAGAATGATTCGATGCTTTGAGCAGCGGATGTGACGACGCGTCAATAATAGTGGTGGGAGGGCAACTGACCCGCAATCAGGAAATTCAGTTGTGA-2500  
AGTCGTGAGGCGTTAATAGAAAATTAATGTAATAATTTCAAAAAATAATTTAAAATAAATAAGCAGCAAAGGAAAAGTTTGTATGTGATATAAATATAT-2600  
AAAGAGAAAACAAACCGCAATAAAGAAAGCTTAAAGAAATAATAAATAAATAAATAAATTAAGTACGAAATTAATAACAAAGCAAAATACAAC-2700  
AAACAAAAAATAAATAAATTTCTGACGAATTTGTAATAATTAGTTATAGATTCTTACTCTCAGCTTGGATATTAATAATATTTT

Regiones: 5' UTR-ATG-Región codificante-Secuencia de sonda-Dominio Homeodomain-Engrailed Homeobox C-terminal signature-3' UTR

## Anexo 5. Secuencias repetidas identificadas

### Etapa Sincicio

Elementos repetidos	Sincicio		
	Número de elementos	Longitud ocupada (pb)	Porcentaje de secuencia
Retroelementos	194	55,914	0.24%
SINEs:	0	0	0.00%
Penelope	0	0	0.00%
LINEs:	120	33,829	0.14%
CRE/SLACS	0	0	0.00%
L2/CR1/Rex	0	0	0.00%
R1/LOA/Jockey	120	33,829	0.14%
R2/R4/NeSL	0	0	0.00%
RTE/Bov-B	0	0	0.00%
L1/CIN4	0	0	0.00%
Elementos LTR:	74	22,085	0.09%
BEL/Pao	11	3,184	0.01%
Ty1/Copia	14	3,230	0.01%
Gypsy/DIRS1	49	15,671	0.07%
Retroviral	0	0	0.00%
Transposones DNA	251	77,776	0.33%
hobo-Activator	1	239	0.00%
Tc1-IS630-Pogo	240	74,484	0.31%
En-Spm	0	0	0.00%
MuDR-IS905	0	0	0.00%
PiggyBac	4	1,113	0.00%
Tourist/Harbinger	0	0	0.00%
Other (Mirage, P-element, Transib)	0	0	0.00%
Rolling cycles	0	0	0.00%
Sin clasificar	0	0	0.00%
Total de interspersed repeats		133,690	0.56%
RNAs pequeños	22	9,584	0.04%

Satélites	0	0	0.00%
Repetidos simples	19,385	815,684	3.43%
Baja complejidad	3,342	178,570	0.75%
Secuencias	42,389		
Longitud total	23,782,827 pb		
Nivel GC	34.24%		
Bases masked	1 137 136 pb (4.78%)		

### Etapa Celularización

Elementos repetidos	Celularización		
	Número de elementos	Longitud ocupada (pb)	Porcentaje de secuencia
Retroelementos	995	284,868	0.80%
SINEs:	0	0	0.00%
Penelope	0	0	0.00%
LINEs:	453	137,735	0.38%
CRE/SLACS	0	0	0.00%
L2/CR1/Rex	1	457	0.00%
R1/LOA/Jockey	452	137,278	0.38%
R2/R4/NeSL	0	0	0.00%
RTE/Bov-B	0	0	0.00%
L1/CIN4	0	0	0.00%
Elementos LTR:	542	147,133	0.41%
BEL/Pao	74	19,329	0.05%
Ty1/Copia	99	23,867	0.07%
Gypsy/DIRS1	369	103,937	0.29%
Retroviral	0	0	0.00%
Transposones DNA	423	131,668	0.37%
hobo-Activator	7	1,519	0.00%
Tc1-IS630-Pogo	365	114,228	0.32%
En-Spm	0	0	0.00%

MuDR-IS905	0	0	0.00%
PiggyBac	17	4,394	0.01%
Tourist/Harbinger	0	0	0.00%
Other (Mirage, P-element, Transib)	0	0	0.00%
Rolling cycles	0	0	0.00%
Sin clasificar	4	766	0.00%
Total interspersed repeats		417,302	1.17%
RNAs pequeños	47	24,686	0.07%
Satélites	0	0	0.00%
Repetidas simples	24,849	1,034,460	2.89%
Baja complejidad	4,292	219,930	0.61%
Secuencias		74,756	
Longitud total		35,811,346 pb	
Nivel GC		34.98%	
Bases masked		1,695,750 pb (4.74%)	

### Etapa Gastrulación

Elementos repetidos	Gastrulación		
	Número de elementos	Longitud ocupada (pb)	Porcentaje de secuencia
Retroelementos	1,001	281,676	0.77%
SINEs:	0	0	0.00%
Penelope	0	0	0.00%
LINEs:	445	129,972	0.35%
CRE/SLACS	0	0	0.00%
L2/CR1/Rex	1	340	0.00%
R1/LOA/Jockey	444	129,632	0.35%
R2/R4/NeSL	0	0	0.00%
RTE/Bov-B	0	0	0.00%
L1/CIN4	0	0	0.00%
Elementos LTR:	556	151,704	0.41%

BEL/Pao	83	21,165	0.06%
Ty1/Copia	81	19,646	0.05%
Gypsy/DIRS1	392	110,893	0.30%
Retroviral	0	0	0.00%
Transposones DNA	411	130,911	0.36%
hobo-Activator	5	1,075	0.00%
Tc1-IS630-Pogo	348	111,873	0.30%
En-Spm	0	0	0.00%
MuDR-IS905	0	0	0.00%
PiggyBac	13	2,622	0.01%
Tourist/Harbinger	0	0	0.00%
Other (Mirage, P-element, Transib)	0	0	0.00%
Rolling cycles	0	0	0.00%
Sin clasificar	10	2,255	0.01%
Total interspersed repeats		414,842	1.13%
RNAs pequeños	46	19,658	0.05%
Satélites	0	0	0.00%
Repetidas simples	26,282	1,102,692	3.00%
Baja complejidad	4,620	236,656	0.64%
Secuencias			74,436
Longitud total			36,811,633 pb
Nivel GC			34.78%
Bases masked			1,773,502 pb (4.82%)

## Anexo 6. Secuencias de pre - miRNAs

### Etapa Sincicio

ID <i>A. ludens</i>	Match	Organismo	% de ident	Valor esperado	miRNA maduro
DN11109_c1_g1_i1	Cca-mir-14	<i>Ceratitis capitata</i>	100	1.91E-09	TCA GTC TTT TTC TCT CTC CTA T

### Etapa celularización

ID <i>A. ludens</i>	Match	Organismo	% de ident	Valor esperado	miRNA maduro
DN10580_c0_g1_i1	Cca-mir-79	<i>Ceratitis capitata</i>	100	1.75E-32	GCT TTG GTA ATC TAG CTT TAT G
DN1779_c2_g1_i4	Cca-mir-79	<i>Ceratitis capitata</i>	100	2.60E-32	GCT TTG GCG CTT TAG CTG TAT G
DN1779_c2_g1_i5	Cca-mir-79	<i>Ceratitis capitata</i>	100	2.52E-32	GCT TTG GCG CTT TAG CTG TAT G
DN48332_c0_g1_i1	Cca-mir-9c	<i>Ceratitis capitata</i>	100	9.62E-12	TCT TTG GTA TTC TAG CTG TAG A
DN57733_c0_g1_i1	Cca-mir-2a-4	<i>Ceratitis capitata</i>	97.143	6.93E-11	TAT CAC AGC CAG CTT TGA GGA GC
DN68342_c0_g1_i1	Cca-mir-7	<i>Ceratitis capitata</i>	95.714	1.39E-26	CAA TAA ATC CCT TGT CTT CTT
DN72962_c0_g1_i1	Cca-mir-2a-2	<i>Ceratitis capitata</i>	94.545	3.96E-18	TAT CAC AGC CAG CTT TGA TGA GC
DN22854_c0_g1_i1	dsi-mir-193	<i>Drosophila simulans</i>	100	4.65E-09	TGG GAT TTT TTA GAT CAG CA
DN27968_c0_g1_i1	dvi-mir-iab-8	<i>Drosophila virilis</i>	97.647	7.85E-37	TAC GTA TAC TGA ATG TAT CCT G
DN55948_c0_g1_i1	dvi-mir-iab-8	<i>Drosophila virilis</i>	97.647	9.58E-36	TAC GTA TAC TGA AGG TAT ACC G
DN59729_c0_g1_i1	dvi-mir-994	<i>Drosophila virilis</i>	89.091	1.62E-32	TAA GGA AAT AGT AGC CGT GAT T
DN44797_c0_g1_i1	dwi-mir-9b-2	<i>Drosophila willistoni</i>	87.778	4.94E-24	TCA TTT GGT TTT TGG TAA TAA AGC TCT ATG

DN49737_c0_g1_i 1	sfr-mir- 10	<i>Spodoptera a frugiperda</i>	96	1.29E-27	TAC CCT GTA GAT CCG AAT TTG T
DN43777_c0_g1_i 1	xla-mir- 703	<i>Xenopus laevis</i>	86.885	1.96E-29	AAA CCT TCA GAA GGA AAG GA

### Etapa Gastrulación

ID <i>A. ludens</i>	Match	Organismo	% de identidad	Valor esperado	Mature miRNA
DN53499_c0_g1_i 1	Cca- mir-7	<i>Ceratitis capitata</i>	95.714	1.37E-26	CAA TAA ATC CCT TGT CTT CTT
DN6645_c2_g1_i2	Cca- mir-79	<i>Ceratitis capitata</i>	100	2.39E-32	GCT TTG GCG CTT TAG CTG TAT G
DN68048_c0_g1_i 1	Cca- mir- 278	<i>Ceratitis capitata</i>	93.243	2.95E-25	CCG GAT GAG GGT CCT CAA CGA CC
DN63475_c0_g1_i 1	cqu- mir-10	<i>Culex quinquefasciatus</i>	97.436	8.99E-12	CAA ATT CGG ATC TAC AGG GTA G
DN42940_c0_g1_i 1	dwi- mir- 9b-2	<i>Drosophila willistoni</i>	97.368	1.85E-12	TCA TTT GGT TTT TGG TAA TAA AGC TCT ATG
DN53068_c0_g1_i 1	sfr- mir-10	<i>Spodoptera frugiperda</i>	96	8.28E-28	TAC CCT GTA GAT CCG AAT TTG T

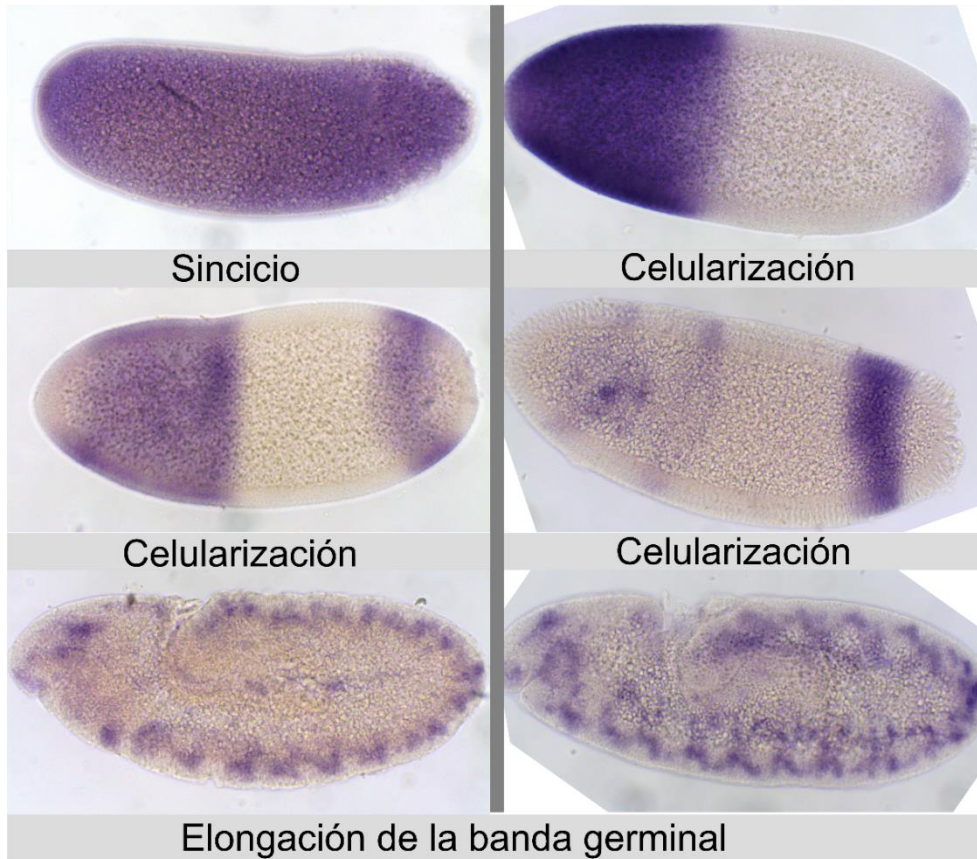


**Anexo 7.** Secuencias identificadas por PFAM.

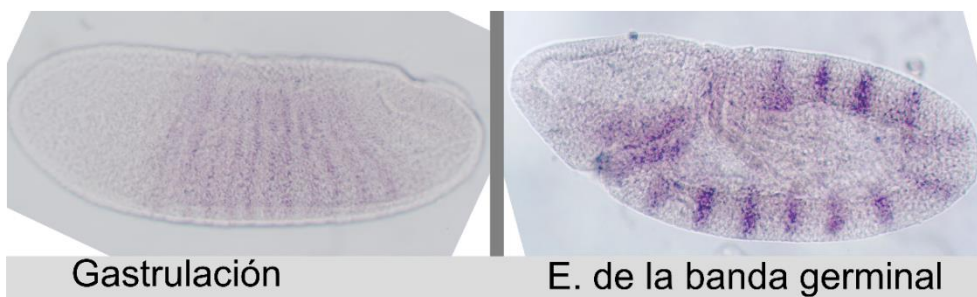
Etapa	ID transcrito	Nombre dominio	ID_PFAM	Secuencia blanco		
				E-value	score	bias
Sin	21242_c0_g1_i1	Collagen	PF01391.17	0.00072	19.2	4.4
Cel	30791_c0_g1_i1	Collagen	PF01391.17	0.00019	21.9	6.7
Cel	71000_c0_g1_i1	Collagen	PF01391.17	0.00028	21.4	10.1
Cel	47005_c0_g1_i1	Collagen	PF01391.17	0.00082	19.9	3.5
Cel	52835_c0_g1_i1	Collagen	PF01391.17	0.001	19.6	9.7
Gas	36890_c0_g1_i1	Collagen	PF01391.17	0.0002	21.9	60.3
Gas	48836_c0_g1_i1	Collagen	PF01391.17	9.30E-05	22.9	33.7

**Anexo 8.** Hibridaciones *in situ* de embriones de *D. melanogaster* con sondas a partir de los transcritos de *A. ludens*.

### Gen *hunchback*



### Gen *engrailed*



**Anexo 9.** Artículo publicado.



## Novel tephritid-specific features revealed from cytological and transcriptomic analysis of *Anastrepha ludens* embryonic development



Ximena Gutiérrez-Ramos<sup>a,b</sup>, Martha Vázquez<sup>a</sup>, Ana E. Dorantes-Acosta<sup>b</sup>,  
Francisco Díaz-Fleischer<sup>b</sup>, Carlos A. Peralta-Alvarez<sup>c</sup>, Hober N. Nuñez-Martínez<sup>c</sup>,  
Rodrigo G. Arzate-Mejía<sup>c</sup>, Félix Recillas-Targa<sup>c</sup>, Mario A. Arteaga-Vázquez<sup>b,\*\*</sup>, Mario Zurita<sup>a,\*</sup>

<sup>a</sup> Departamento de Genética Del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Mexico

<sup>b</sup> Group of Epigenetics and Developmental Biology, INBIOTECA, Instituto de Biotecnología y Ecología Aplicada, Universidad Veracruzana, Mexico

<sup>c</sup> Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Mexico

### ARTICLE INFO

#### Keywords:

*Anastrepha ludens*  
Embryo development  
Transcriptome  
Gene expression

### ABSTRACT

*Anastrepha ludens* is a major pest of fruits including citrus and mangoes in Mexico and Central America with major economic and social impacts. Despite its importance, our knowledge on its embryonic development is scarce. Here, we report the first cytological study of embryonic development in *A. ludens* and provide a transcriptional landscape during key embryonic stages. We established 17 stages of *A. ludens* embryogenesis that closely resemble the morphological events observed in *Drosophila*. In addition to the extended duration of embryonic development, we observed notable differences including yolk extrusion at both poles of the embryo, distinct nuclear division waves in the syncytial blastoderm and a heterochronic change during the involution of the head. Characterization of the transcriptional dynamics during syncytial blastoderm, cellular blastoderm and gastrulation, showed that approximately 9000 different transcripts are present at each stage. Even though we identified most of the transcripts with a role during embryonic development present in *Drosophila*, including sex determination genes, a number of transcripts were absent not only in *A. ludens* but in other tephritids such as *Ceratitis capitata* and *Bactrocera dorsalis*. Intriguingly, some *A. ludens* embryo transcripts encode proteins present in other organisms but not in other flies. Furthermore, we developed an RNA *in situ* hybridization protocol that allowed us to obtain the expression patterns of genes whose functions are important in establishing the embryonic body pattern. Our results revealed novel tephritid-specific features during *A. ludens* embryonic development and open new avenues for strategies aiming to control this important pest.

### 1. Introduction

*Anastrepha ludens*, commonly known as the Mexican fruit fly, is native to Northeastern Mexico (Baker et al., 1944) and belongs to the Tephritidae family, in which many flies of economic importance are found due to their characteristic oviposition behaviour within fruits (Hernández Livera, 2014). *Anastrepha ludens* is considered one of the main problems in fruit production in Mexico due to its high incidence and severity (Hernández Livera, 2014). Studies carried out in tephritids are generally based on basic biological features, which provide information on sexual development and reproductive abilities (Pérez-Staples et al., 2019). Studies in *Bactrocera dorsalis* and *Ceratitis capitata* have focused on the role of genes involved in sexual determination (Chen et al., 2008; Salvemini et al., 2009) but in recent years, genomic

and transcriptomic studies have been reported (Calla et al., 2014; Gomulski et al., 2012; Papanicolaou et al., 2016; Salvemini et al., 2014; Shen et al., 2011).

While studies on tephritids have focused mainly on larvae, pupae and adult stages, observations at the egg stage have been greatly restricted to the external features of the egg (Greene, 1929). The first tephritid whose embryonic development was studied was the Queensland fruit fly, *Bactrocera tryoni* and these observations were based on *Drosophila melanogaster* descriptions (Anderson, 1962). In *Anastrepha*, descriptions of the egg stage are rare, and those that are available are partial or incomplete (Dutra et al., 2013), even though the identification *Anastrepha* species is based on external egg characteristics (Figueiredo et al., 2013) and embryonic development has been described for only four: *A. fraterculus*, *A. obliqua*, *A. nigrifascia* and *A.*

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [maarteaga@uv.mx](mailto:maarteaga@uv.mx) (M.A. Arteaga-Vázquez), [marioz@ibt.unam.mx](mailto:marioz@ibt.unam.mx) (M. Zurita).

<https://doi.org/10.1016/j.ibmb.2020.103412>

Received 17 March 2020; Received in revised form 12 May 2020; Accepted 12 May 2020

Available online 15 May 2020

0965-1748/ © 2020 Elsevier Ltd. All rights reserved.

*pittieri*, but in those cases the descriptions were based on superficial characteristics only (do Nascimento and Kalisz de Oliveira, 1996; Genç, 2014).

Although *A. ludens* has been studied for more than a century, the information we have is derived from studies of its sexual behaviour, ecology and strategies for pest management (Aluja et al., 2000), leaving aside studies of its development, especially its embryonic development. Embryonic development is the stage at which the body plan is defined and is considered the preamble for the generation of the morphological diversity observed in adults (Kalinka and Tomancak, 2012; Peel, 2008). Information about *A. ludens* embryonic development is limited to its timing, which is 6 days at 24 °C and 9 days at 19.6 °C in wild conditions (McPhail and Bliss, 1933), or egg morphology (Carroll and Wharton, 1989; Emmart, 1933; Greene, 1929). These analyses are used as a tool for the identification of infested plantations (Pérez-Staples et al., 2019). Some investigations have indirectly described characteristics of *A. ludens* embryogenesis, including egg laying (4.24 eggs/day) and the mean age of reproduction (19.4 days) (Celedonio-Hurtado et al., 1988; Liedo et al., 1992). The characteristics of the clutch, or group of eggs that the females deposit in the host, have also been described: 4.4 eggs per clutch are loaded into a 2-cm host, and 12.7 eggs per clutch are loaded into an 11-cm host (Berrigan et al., 1988). The host characteristics, such as the firmness, ripeness, and use of an artificial host, have also been described (Díaz-Fleischer and Aluja, 2003). Oviposition-related descriptions are important for determining the process of egg growth in the fruit; however, at the level of developmental and regulatory events, these processes are not well understood.

Based on this information, we examined the embryonic development of *A. ludens* from morphological and gene expression perspectives. Detailed and systematic observations allowed us to describe for the first time the embryonic developmental stages of the *A. ludens* fly. Based on these observations, we carried out a transcriptomic study of three early embryo development stages: syncytium, cellularization and gastrulation. In addition, we reinforced the transcriptome data by establishing an *in situ* hybridization protocol to determine the expression patterns of some important genes involved in early embryonic development. This work provides the first description of the morphological characteristics and global gene expression levels of the early development of *A. ludens*, which provides a basis for developing new strategies to combat this economically and socially important pest.

## 2. Materials and methods

### 2.1. Insect strains

Wild type *Anastrepha ludens* flies were obtained as pupae from a mass-reared strain produced at the Moscafrut Plant (Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria - Secretaría de Agricultura y Desarrollo Rural (SENASICA-SADER)), located in Metapa de Domínguez, Chiapas, México. Adult flies were grown and maintained on an artificial diet (3:1 mixture of sugar:dry yeast (MO Biomedicals, LLC)), under laboratory conditions: 25 °C, at 80% RH with a photoperiod of 12:12 h of light:darkness.

### 2.2. Embryo collection and growth

*Anastrepha ludens* embryos were collected from sexually mature females of 12–25 days old (Dickens, J. C., Solis, E., Hart, 1982), in a growing chamber with 12 h light/darkness cycles. Flies oviposited on artificial water-filled oviposition devices made with a bottomless Petri dish covered with a black cloth whose surface was overlaid with a thin clear silicone layer (Sista, Henkel Mexicana S.A., Huixquilucan, Estado de Mexico, Mexico), and placed at the top of the wooden cage. Collections were done between the 3rd to 7th h of the light photoperiod. The embryos were collected in 30 min intervals and then they were grown in water until the desired stage. Embryos were grown at 25 °C in

a growth chamber which is a Petri dish with a piece of black satin fabric on top of a bed of wet cotton.

### 2.3. Embryo fixation and DAPI staining

The fixation protocol was based on *Drosophila*'s protocols (Sullivan et al., 2000) with modifications for *A. ludens*. The embryos were dechlorinated with 6.2% sodium hypochlorite for 3 min. Then they were fixed with 18.5% formaldehyde for 5 min, 3 washes of methanol, and a serial dehydration of 1:1 v/v, 1:2 v/v and 1:3 v/v of PBS:methanol. Rehydration was done with PBS:methanol 1:1 v/v and then embryos were washed with PBS-Tween 20 1% (PBST) necessary for vitelline membrane permeabilization. For nuclear staining, the embryos were incubated with a solution of 2.5 µg/ml DAPI (ThermoFisher Scientific, D1306) for 45 min at room temperature in darkness.

### 2.4. Embryo imaging

Microscopy imaging was performed at the Microscopy Unit of Instituto de Ecología A.C. (INECOL) - Cluster BioMimic. Living embryos were observed with a Leica S8 APO microscope. For fluorescent microscopy, we employed a Confocal Laser Scanning Microscope Leica TCS SP8 and an inverted microscope Leica DMI600B equipped with a Leica DFC450 camera. Imaging of *in situ* hybridization where done in a Nikon ECLIPSE E600 microscope at the Instituto de Biotecnología, Universidad Nacional Autónoma de México. Image analysis and processing were done with Leica Application Suite LAS EZ v2.0.0, Leica Application Suite LAS AF Lite v2.6.3 and Fiji v1.52.p (Schindelin et al., 2012) software.

### 2.5. Data collection and morphometric analysis

Embryo development was quantified with DAPI in a period of 1–96 h after-egg laying (AEL); from 96 to 120 h, staging was performed through microscopic observation of live embryos. Embryos were classified relative to their developmental stage as: i) expected growth embryos (EGE), ii) delayed growth embryos (DGE) and iii) unfertilized eggs. Eight-fold amplification images were taken every hour during 120 h of a collection of embryos arranged in an anteroposterior position. A total of 120 antero-posterior images, taken during embryonic development, were analysed for the position of 99 landmarks along the chorion and embryo periphery using the tpsDig232 software v2.30 (Rohlf, 2006). To determine the difference of embryo length, we compared the embryo occupied area with the one occupied by the chorion periphery using Procrustes superimposition to produce a canonical covariance matrix; a Canonical Variate Analysis was performed using the MorphoJ software v1.06d (Klingenberg, 2011).

### 2.6. Generation of RNA-seq data

RNA-Seq experiments were performed from early embryos laid by 19–25 days old females' flies (Fig. S1). Staging was done first by selecting them by hours post-fertilization and then individually staged by observation under a dissection microscope. The stages selected were syncytium (0–9 h), cellularization (9–18 h) and gastrulation (18–30 h). Total RNA from each pool, was extracted with TRI Reagent (Sigma Aldrich, 93289) according to manufacturer's protocol, experiments were performed by triplicate of biological replicas. Library construction (paired-end 2 × 100) and sequencing of ~106 - 58 million paired-end 2 × 100 reads (Table S1) per sample was performed at the Beijing Genomics Institute (BGI) in an Illumina HiSeq 4000 platform.

### 2.7. De novo transcriptome assembly and annotation

Bioinformatic analyses (Fig. S2A) were performed at the Instituto de Biotecnología, Universidad Nacional Autónoma de México. Quality

analysis of all reads from the experiment (566, 454, 896 million paired-end raw reads) was performed with FastQC v0.11.7 (Andrews, 2018) and a quality analysis for the biological replicates was performed by MutiQC v1.5.dev0 (Ewels et al., 2016) (Table S1). Low quality readings and sequence adapters removal was performed by BGI. For the comparison between the different fly species the data used were for *B. dorsalis* SRA SRP007559 (Shen et al., 2011) and GEO GSE118472, for *C. capitata* SRA SRRI380982 (Salvemini et al., 2014), and FPKM values reported for *D. melanogaster* in FlyBase FB2018-06. *B. dorsalis* and *C. capitata* data was performed as quality analysis of all reads with FastQC v0.11.7 (Andrews, 2018) and sequence adapters removal was performed by Trimmomatic v0.38 (Bolger et al., 2014). To preserve the unique transcriptomic landscape of each developmental event, as well as identified possible stage transcripts isoforms, we used Trinity v2.8.4 (Grabherr et al., 2011) with default parameters (Table S1) for *de novo* transcriptome assembly of each developmental stage of *A. ludens*, *B. dorsalis* and *C. capitata* data. However, we also performed an *A. ludens* global assembly (Table S1), but we work with the data generated from the assemblies by developmental stage. Annotation of each transcriptome was performed as it follows: for each assembly we identified open reading frames (ORFs) with TransDecoder v5.3.0 (Brian and Papanicolaou, 2019) and performed a protein-protein and transcript-protein search using BLASTp and BLASTx, respectively, v2.7.1 against the proteome of *D. melanogaster* downloaded from FlyBase vFB2018-06 and with the global non-redundant protein BLAST v2.7.1 database (Table S2). We estimated read abundance for all transcripts in each of the reconstructed transcriptomes using kallisto v0.44.0 (Bray et al., 2016) with a bootstrap of 100. We then selected for further analyses those transcripts showing both an identity equal or greater than 25% with an annotated protein in *Drosophila*, and with similar number of FPKMs in all three biological replicates (log2 across FPKMs of replicates  $\pm 1$  of standard deviation) (Fig. S2B). For transcripts in *A. ludens* showing identity equal or greater than 25% to *Drosophila* proteins, we performed functional annotation and gene ontology analysis using DAVID v6.8 (Huang, Sherman and Lempicki, 2009a, 2009b) (Fig. S2C).

For transcripts in *A. ludens* that did not score significant similarity to annotated proteins in *Drosophila* (Fig. S2D) we performed: i) the identification of previously published repeated sequences (Table S3) and *de novo* modelling of *A. ludens* unique repeat families using the RepeatModeler v1.0.11 suit consisting of the programs RECON v1.08, RepeatScout v1.0.5, and Tandem Repeats Finder TRF v4.09 (Smit, Hubley and Green, 2019b, 2019a), ii) detection of precursors of microRNAs based on sequence comparisons against the miRBase v22.1 database (Table S3) (Kozomara and Griffiths-Jones, 2011) and iii) annotation of protein domains using HMMER/PFAM v3.2.1. Identification of possible non-coding RNAs was considered, but due to limitations of lack of genomic data of *A. ludens* they were not performed.

## 2.8. Differential expression analysis of transcripts

Differential expression analyses across embryo stages were performed using the concatenate transcriptome of stages of development, subsequently debug of redundant transcript sequences in the transcriptome using cd-hit v4.8.1 (Fu et al., 2012; Li and Godzik, 2006) with a word size 10–11 for 0.95 threshold at both strands (Table S1). Estimated read abundance of the three replicates of each developmental stage was done using kallisto v0.44.0 (Bray et al., 2016) and differential expression analysis were performed using edgeR (McCarthy et al., 2012; Robinson et al., 2009) (Table S4).

## 2.9. RNA probes design and protocol for whole mount *in situ* hybridization

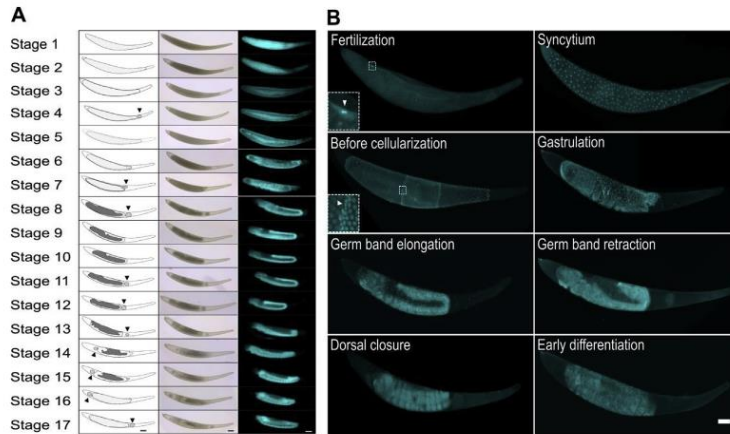
Probe design was performed as described in Hodar et al., (Hodar and Cambiasso, 2018). We used two pairs of primers for each gene (for more details see Table S5). For the sense probe (T7 promoter sequence + forward primer and reverse primer) and for the antisense

probe (forward primer and T7 promoter sequence + reverse primer) according to the sense of each gene. RT-PCR was performed from total embryo RNA and for each gene a pair of primers sets was used, followed by an *in vitro* transcription with digoxigenin-UTP labelling (Roche, 11277073910) and the digoxigenin-labelling probe following the manufacturer instructions was verified. The selected transcripts correspond to the *A. ludens* genes *exuperantia* (*exu*), *hunchback* (*hb*), *even skipped* (*eve*) and *engrailed* (*en*) (Table S5) found in our transcriptomic data. *In situ* hybridizations were performed in embryos from syncytium, cellularization, gastrulation, germ band elongation and germ band retraction stages. For the whole mount *in situ* hybridization the embryo collection was performed as described in “Embryo collection and growth”. Embryos from the desired developmental stage were dechorionized with 6.2% sodium hypochlorite for 3 min, fixed with 4% paraformaldehyde for 20 min and washed 2 times with 2 vol of methanol. The next steps are modifications of Tautz et al. (Tautz and Pfeifle, 1989), protocol as it follows: embryos were fixed with 4% paraformaldehyde for 20 min and washed 2 times with 2 vol of methanol followed by one xylene wash and 5 ethanol washes. Posteriorly, we fix the embryos again with 4% paraformaldehyde for 20 min, followed by 3 washings with PBST (1x PBS + 1% Tween 20) and then one incubation with 250 µg/mL proteinase K (Roche, 03115828001) in PBST at 55 °C for 3 h. Each new batch of proteinase K should be tested and adjusted for the desired balance of embryo permeabilization and integrity. Embryos then are washed with PBST and fixed with 4% paraformaldehyde for 20 min and washed again with PBST. Prehybridization treatment is realized in hybridization buffer (50% formamide, 5X SSC, 15 mM heparine, 1% Tween-20 and 100 µg/mL DNA salmon sperm) at least 1 h, to continue with an overnight hybridization in the same buffer with 100 ng of the biotin-labeled at 50 °C. The next day 2 washes are done with hybridization buffer and 4 washes others with PBST. Hybridized embryos are incubated 1 h with 1:2000 Anti-Digoxigenin-AP (Roche, 11093274910) and revealed with NBT/BCIP (Roche, 11697471001). This last step is performed overnight at 4 °C to get a better signal. *A. ludens* embryos cannot be devitelinized for this and other techniques. They lose their integrity even when we manually devitelinized them. Thus, precaution should be taken in setting the appropriate number of embryos in each sample to be hybridized, because the proper hybridization pattern is obtained following this protocol with an expected efficiency of 50% of embryos approximately.

## 3. Results

### 3.1. Overview of *A. ludens* embryogenesis

In order to characterize the embryogenesis of *A. ludens*, we performed direct phase contrast observations and established a protocol for tissue fixation and DAPI staining (see 2. Materials and Methods). Systematic observations were performed at 1 h intervals during a 120 h time window (from oviposition to hatching), under laboratory conditions at 25 °C in live individuals. Our description of *A. ludens* embryogenesis is based on classical observations in *Drosophila melanogaster*, which comprises 17 stages of morphogenetic embryo development (Campos-Ortega and Hartenstein, 1985). The stages and examples of key developmental events of *A. ludens* embryogenesis are depicted in Fig. 1A and B. Syncytial blastoderm covers stages 1–3 (0–3 h, stage 1), during which the egg is spindle-shaped (Carroll and Wharton, 1989), and once it has been laid, it turns uniformly white. The yolk shows a homogeneous distribution along the egg. The zygote fills the egg almost completely, except for the micropyle (Fig. 1A and B). Stage 2 (3–7 h): the retraction of the yolk initiates at the anterior pole, leaving an empty space from the vitelline envelope. The posterior retraction of the yolk begins 4 h after oviposition. Intriguingly, contrary to what happens in *Drosophila*, mitotic divisions during the syncytial blastoderm start in the middle of the embryo and expand to the poles (Fig. 1B). Stage 3 (7–9 h): pole bud formation takes place during this stage. The embryo acquires a



**Fig. 1.** Embryonic development of *A. ludens*. A) The 17 stages of the embryonic development of *A. ludens* are graphically represented. Extrusions of the yolk at different stages are indicated by an arrowhead (left). Representative pictures of embryos of the 17 stages described, observed in light field microscopy (center) or with DAPI-stained DNA to observe nuclei with fluorescence microscopy (right) (the bars represent 100  $\mu$ m). Embryos pictured at the center and right columns are not the same and each stage indicated follows a temporal sequence. B) Nuclei-stained embryos at key embryonic stages. At fertilization, the first embryo division is shown in the zoomed in panel; syncytium stage; before cellularization, nuclei cytokinesis is shown in the zoomed in panel, note that the nuclei morphology changes during mitosis; at gastrulation, invaginations along the ventral part, and at the posterior pole are noticed; in germ band elongation, thickening of the cells along the ventral to the dorsal region is

observed; in germ band retraction cells retract to the ventral region and the head structure is appreciated; in embryos at dorsal closure, there is segmentation and head involution; at differentiation a more defined segmentation and a larval shape is appreciated. Nuclei are stained with DAPI. The bar represents 100  $\mu$ m.

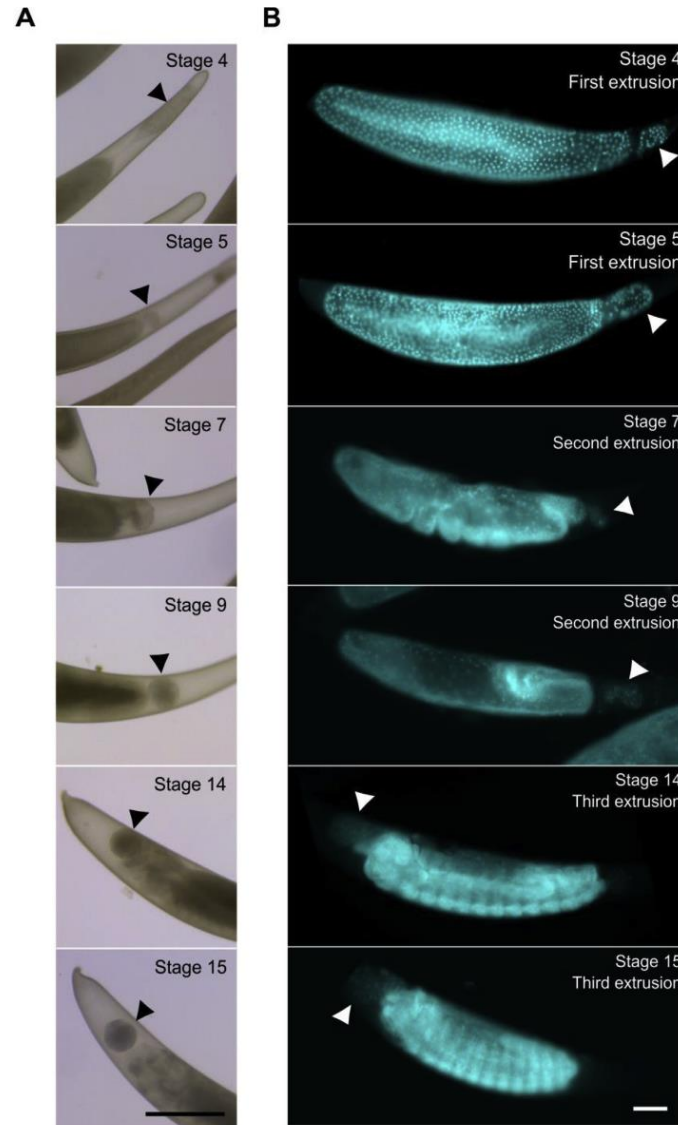
granulated appearance along the yolk, and the posterior part of the yolk has a less dense appearance. A cellularization event covers stage 4 (9–14 h): at the beginning, invaginations are formed on the less-dense part of the yolk, which later extend along the embryo periphery. On the posterior part of the embryo, the first yolk extrusions are observed. Once yolk extrusion occurs, the polar bud becomes visible on the embryo. Stage 5 (14–18 h): the invaginations along the embryo periphery become more visible. At the end of this stage, in *Drosophila*, the polar cells begin to migrate at the dorsal-posterior part of the embryo. Intriguingly, in *A. ludens* embryos, we cannot visualize the polar cells by phase contrast or DAPI staining at any developmental stage. The mid-ventral cells acquire an irregular, wavy appearance. A gastrulation event covers stage 6 (18–30 h): The second yolk extrusion occurs from the posterior part of the embryo. A germ band elongation event covers stage 7 (30–37 h): this event starts with the end of gastrulation, the posterior midgut and hindgut invaginate, and the dorsal folds can be observed. In addition to the cephalic furrow on the anterior part of the embryo, the amnioproctodeal invagination, the anterior transversal fold and the transversal posterior fold occur. Stage 8 (37–45 h): this stage is marked by the formation of an amnioproctodeal invagination and rapid germ band elongation. The second extrusion is completed at the end of this stage. Stage 9 (45–49 h): germ band elongation continues at the ventral part of the embryo. At the end of this stage, stomodeal invagination and extension of the procephalic lobe can be observed. In some cases, remnant tissues of the second yolk extrusion are still detectable on the embryos. Stage 10 (49–55 h): the beginning of this stage is characterized by the appearance of the stomodeum, which invaginates ventrally, and the germ band reaches its maximum size. Stage 11 (55–60 h): no major morphogenetic changes take place at this stage. Germ band retraction covers stage 12 (60–69 h): the shortening of the germ band takes place, and the yolk sac extends on the dorsal surface of the embryo as a consequence of germ band shortening (Fig. 1A). The acron and the embryonic larval segment are visible. Stage 13 (69–72 h): shortening of the germ band ceases. Once this occurs, the parasegments are observed along the dorsal side of the embryo. At the same time, head involution occurs, this is a major difference to that observed in *Drosophila*, in which the dorsal closure occurs before head involution. At the end of this stage, the segments of the embryo and the involuted head become visible. A dorsal closure event covers stage 14 (72–82 h): the dorsal epidermis grows at both sides of the amnioserosa and fuses at the midline (Fig. 1A). The third yolk extrusion begins on the anterior

part of the embryo. Differentiation events cover stage 15 (82–91 h): the dorsal epidermal segmentation is visible. At the ventral part of the embryo, the tissue becomes more condensed, and absorption of the amnioserosa begins (Fig. 1A). Stage 16 (91–108 h): the intersegmental grooves can be distinguished, and the absorption of the amnioserosa continues. At 99 h of development, parts of the mouth chitinize and can be observed from the ventral side. Stage 17 (108–120 h): this stage extends until hatching. The tracheal tubes are visible at the beginning of the stage. At 110 h of development, the third yolk extrusion is no longer distinguishable. In some cases, residual tissue is expelled from the micropyle. At the end of this stage, embryo head movements can be observed. Hatching begins at 114 h and can extend to 132 h of development, that commonly occurs through the micropyle or at the midline of the chorion. Together, these morphological observations show that *A. ludens* embryo development is similar, but not identical to the *Drosophila* descriptions. The major differences are the time required for each stage, which is extended in *A. ludens*, the anterior and posterior extrusions, mitotic divisions from the centre to the poles at the syncytial blastoderm and the occurrence of head involution before the dorsal closure.

Next, we determined the temporal accuracy of the different developmental stages by classifying the embryos into three categories (Table S6): i) expected growth embryos (EGE): embryos morphologically matching the phenotypic characteristics of a given stage, ii) delay growth embryos (DGE): embryos showing the phenotypic characteristics of a previous stage to that being analysed and iii) unfertilized eggs (UE): eggs that did not develop at all. The general distribution of the progress of embryonic development was 69.3% EGE, 16.3% DGE and 14.4% UE (total experimental embryos  $n = 14848$ ). On average, 60% of the embryos fell within the EGE category, and 14% of embryos across all stages fell within the DGE category. This allows us to conclude that approximately 75% of the oviposited embryos developed into larvae, and only a small fraction was not fertilized.

### 3.2. *A. ludens* embryonic yolk mass extrusions

Cellular extrusions are a phenomenon that has been reported in several animals; these events have not been reported to cause cellular or functional damage (Cameron, 1905; Clark, 1960; Thomson, 1919). Thus far, the information necessary to understand the function of these events is not available. In the *Anastrepha* genus, extrusions are observed



**Fig. 2.** Embryonic extrusion as markers of *A. ludens* development. A) Three yolk extrusions occur at different stages of development. The first four pictures show in detail the posterior pole of embryos, from where the first (between stages 4–5) and second extrusions (between stages 6–7) occur. The third extrusion that is anterior (at stage 14) (bottom picture). See Fig. 1A for stages. B) Yolk extrusions in nuclei-stained embryos. Embryos extruding yolk at different stages are shown. The extruded volume is approximately 20% of the embryo length in the first two extrusions. Nuclei are stained with DAPI. The bars represent 100  $\mu$ m.

as a part of embryogenesis (Perondini et al., 1998; Selivon et al., 1996; Selivon and Perondini, 1997). Our observations in *A. ludens* allowed us to identify three yolk mass extrusions (indicated with arrowheads in Fig. 2A): two occurred at the posterior part of the embryo during the transition of the syncytial to cellularized blastoderm and during gastrulation. The third occurred at the anterior part of the embryo during

the beginning of dorsal closure and the maturation of the embryo. To quantify the morphological changes occurring during embryogenesis, with particular attention to the morphological effects of the yolk mass extrusions, we employed MorphoJ geometric morphometrics software (Klingenberg, 2011) to analyse the changes in embryo length during embryogenesis. We analysed the shape of the egg chorion and the



**Table 1**  
Summary of the morphometric analysis of *A. ludens* embryogenesis.

Stage	Procrustes Analysis			Canonical Variate Analysis			
	Centroid Size		Size Difference	Change Percentage <sup>a</sup>	Eigenvalues	% Variance	Cumulative %
	Chorion	Yolk mass					
1	2807.87	2786.26	21.61	-0.77	15.5671	35.051	35.051
2	2614.68	2206.92	407.75	-15.59	8.3144	18.721	53.772
3	2558.22	2033.29	524.93	-20.52	2.935	6.609	60.381
4	2562.85	1936.71	626.15	-24.43	2.597	5.847	66.228
5	2561.64	1982.33	579.31	-22.61	2.521	5.676	71.905
6	2586.39	1816.22	770.17	-29.78	1.8945	4.266	76.171
7	2577.35	1556.35	1021	-39.61	1.7928	4.037	80.207
8	2594.37	1479.56	1114.81	-42.97	1.5089	3.398	83.605
9	2594.86	1420.41	1174.45	-45.26	1.3654	3.074	86.679
10	2603.68	1394.8	1208.88	-46.43	1.2137	2.733	89.412
11	2644.82	1441.84	1202.98	-45.48	1.0641	2.396	91.808
12	2636.62	1378.72	1257.9	-47.71	0.9367	2.109	93.917
13	2536.09	1348.29	1187.8	-46.84	0.9039	2.035	95.953
14	2626.15	1348.01	1278.14	-48.67	0.7317	1.648	97.6
15	2637.7	1329	1308.7	-49.62	0.5924	1.334	98.934
16	2629.96	1342.14	1287.82	-48.97	0.4734	1.066	100
17	2626.37	1433.35	1193.02	-45.42	0	0	100

embryo length by a Procrustes superposition (Table 1 and Fig. 2B), and then we compared the centroid size of the egg chorion and the embryo length to infer the size of the extruded yolk mass. With this comparison, we observed that the net loss of embryo length was approximately 48.97%, which was related to embryonic changes. These changes comprise yolk retraction (-15.6%) and yolk extrusions, mainly in the first (-24.4%) and the second (-42.9%) stage, respectively. To validate the variation in embryo sizes among the stages, we performed a canonical variate analysis (CVA). The classification criterion was the embryogenesis stage; the highest percentage of change between them was in the first three stages, resulting in 60.38% of the variance among them (Table 1). The changes resulted in a similar variation in the loss of yolk length previously observed in the centroid size measurement, so we can conclude that the loss of embryo length is an important feature for the development of *A. ludens* embryos.

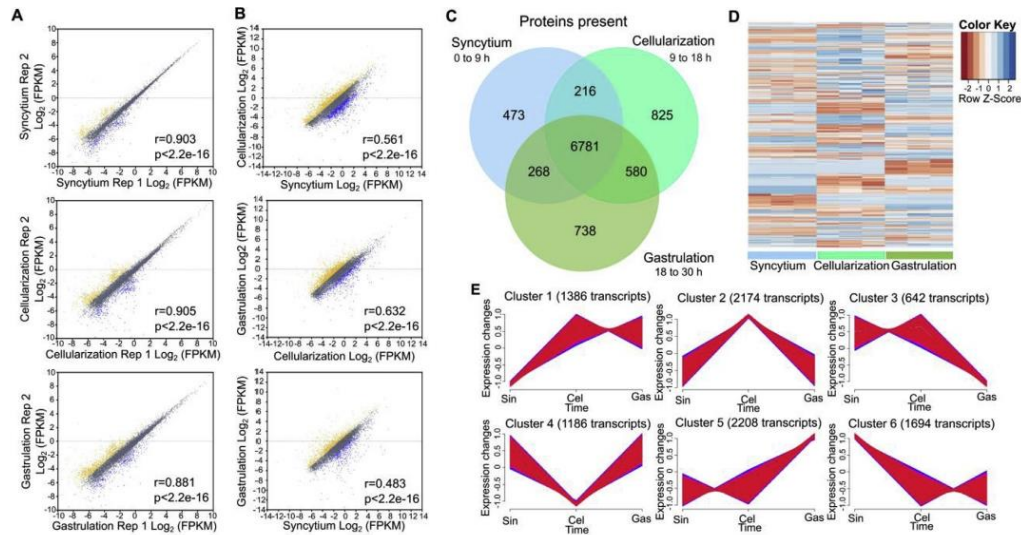
### 3.3. De novo assembly and transcriptome analysis of *A. ludens* early embryos

During early embryo development, the body plan of dipteran insects is established (Kalinka and Tomancak, 2012). Thus, to obtain an initial approximation of the genes that participate in the early embryo development of *A. ludens*, we performed an Illumina-based RNA-seq analysis from embryos at the syncytial blastoderm (0–9 h), cellular blastoderm (9–18 h) and the beginning of gastrulation (18–30 h) stages. The main reason for the selection of these developmental stages is that, in syncytial blastoderm, the transcripts deposited in the embryo by the mother participate in the first embryonic process, zygotic genome activation occurs just before cellularization, and the establishment of the germ layers occurs during gastrulation (Lott et al., 2011). Pearson correlation plots comparing the different replicates of each embryonic stage indicated high confidence (Fig. 3A; Fig. S2B); as did the correlation analysis between the different embryonic stages, which showed that genes were either down or up regulated (Fig. 3B). Approximately 9000 different unigenes were expressed in each developmental stage, of which 7000 corresponded to unique protein coding genes, which is similar than the number of genes reported for these developmental stages in *Drosophila* but less than that in *C. capitata* (Lott et al., 2011; Salvemini et al., 2014). From these, 7433 genes were expressed in the three analysed stages, 321 were shared only between the syncytial and cellularized embryos, 407 between syncytial and gastrulated embryos and 691 between cellularized and gastrulated embryos. In addition, 343 genes were unique to the syncytial stage, 573 to cellularization and 695

to gastrulation (Fig. 3C). This analysis indicated that most transcripts were present in the three developmental stages analysed and indeed, specific transcripts that are considered developmental stage markers by their expression pattern analysis in *Drosophila* were not as specific during the early embryogenesis stages in *A. ludens*. For instance, most of the *A. ludens* genes homologous to the *Drosophila* genes that participate in early embryo development, such as *hunchback* (*hb*), *exuperantia* (*exu*) and *engrailed* (*en*), were found to be present in the three analysed stages, although enriched in the corresponding developmental stage when *Drosophila* is taken as a reference (Table S2). Thus, the transcripts were classified into 6 clusters that had different dynamics during development (Fig. 3E). For instance, cluster 1 includes transcripts with low levels in the syncytial blastoderm that then increased in the next developmental stages (Fig. 3E), as well as transcripts with high levels in the syncytial blastoderm and low levels in cellularization and gastrulation (cluster 6). Other transcripts had a very dynamic behaviour, with expression oscillating up and down, depending on the developmental stage (Fig. 3E).

A major interest in the early development of Diptera comes from genes that participate in sex determination, not only for their biological importance but also for their potential to be manipulated to control pest populations. In *Drosophila*, the sex determination pathway is well known; the genes that participate in this process are expressed in the early embryo, and some related advances have also been made in *C. capitata* (Saccone et al., 2002; Salvemini et al., 2014). Therefore, we searched for the presence of transcripts that encode sex determination genes and found that among the 27 genes that participate in this process in *Drosophila*, 24 were identified in the *A. ludens* transcriptome, including the central participants for sex determination in *Drosophila* (Table 2). The missing genes were *dissatisfaction* (*dsf*), *hermaphrodite* (*her*) and *outstretched* (*Os*) (see 4. Discussion).

In addition, we identified transcripts that encode 488 genes that have been reported not to be expressed in early *Drosophila* embryo development or that have not been identified in *Drosophila*. (Table S2). Most of these genes encode proteins that have motifs or domains that are similar to other proteins in different organisms, suggesting a functional role of these genes in *A. ludens*. Many of the ORFs encode putative enzymes that may participate in different metabolic and general signal transduction pathways, but others may also have a regulatory role during embryo development (Table S2). Our analysis also identified transcripts from transposable elements (Table S3); we also identified precursor transcripts of microRNAs (Table S3), as well as different factors that participate in epigenetic mechanisms (Table S7) described

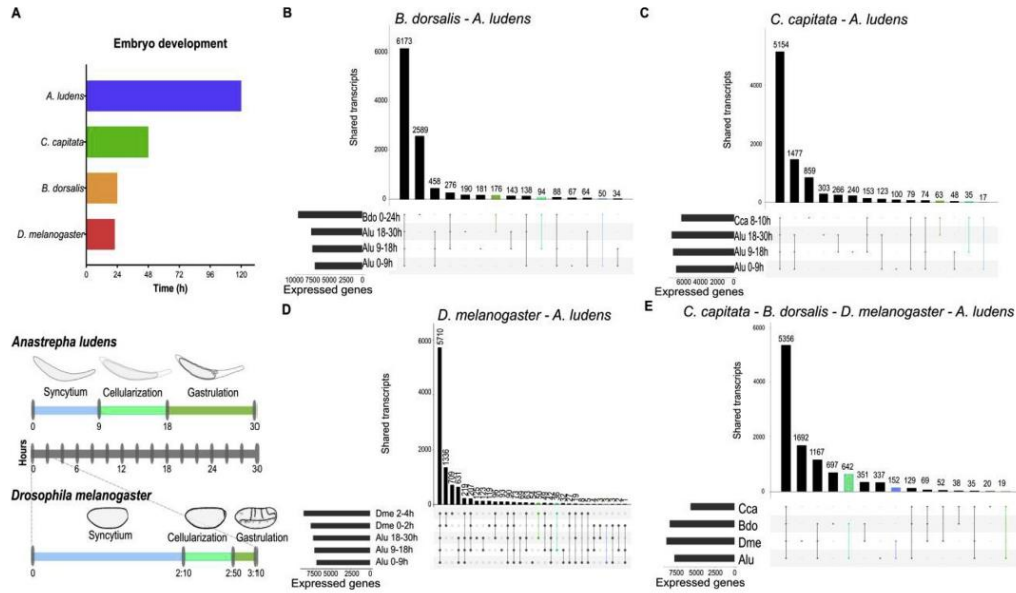


**Fig. 3.** Transcriptomic analysis of *A. ludens* early embryogenesis. A) Similarity of samples measured as Pearson's correlations between two of the three replicates of each stage of development, the correlation value,  $r$ , (lower right). B) Similarity of developmental events measured as Pearson's correlations between the mean of the three replicates of each stage of development; the correlation value,  $r$ , (lower right). C) Unique and shared proteins in FPKMs across the developmental stages depicted in Venn diagram. D) Heatmap depicting abundance dynamics of transcripts in the developmental stages analysed. E) Transcriptional dynamics during embryonic development of *A. ludens*, characterized using soft cluster analysis. The graphical depiction of each of the six clusters indicate the number of transcripts per clusters.

**Table 2**

*A. ludens* genes related to sex determination. Most of the genes important for sexual determination in flies can be identified in RNA-seq data from *A. ludens* (this study), *B. dorsalis* and *C. capitata* (Salvemini et al., 2014). We were not able to identify NF genes in neither of the tephritids analysed.

Name	Symbol	<i>A. ludens</i>			<i>B. dorsalis</i>			<i>C. capitata</i>
		0–9 h	9–18 h	18–30 h	0–1 h	2–4 h	5–8 h	8–10 h
fruitless	fru	Yes	No	Yes	No	No	No	No
scute	sc (sisB)	Yes	Yes	Yes	No	No	Yes	Yes
deadpan	dnp	Yes	Yes	Yes	No	Yes	Yes	Yes
runt	run	Yes	Yes	Yes	No	Yes	Yes	Yes
Doublesex	dsx	Yes	Yes	Yes	Yes	Yes	Yes	No
Male-specific-lethal 2	msl-2	Yes	Yes	Yes	Yes	Yes	Yes	No
virilizer	vir	Yes	Yes	Yes	Yes	Yes	Yes	Yes
transformer-2	tra2	Yes	Yes	Yes	Yes	Yes	Yes	Yes
daughterless	da	Yes	Yes	Yes	Yes	Yes	Yes	Yes
sex-lethal	Sxl	Yes	No	Yes	No	Yes	Yes	Yes
Male-specific-lethal 3	msl-3	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Male-specific-lethal 1	msl-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes
intersex	ix	Yes	Yes	Yes	Yes	Yes	Yes	Yes
hopscotch	hop	Yes	Yes	Yes	Yes	Yes	Yes	Yes
extra macrochaetae	emc	Yes	Yes	Yes	No	No	No	Yes
sex-lethal	Sxl	Yes	Yes	Yes	Yes	Yes	Yes	Yes
sisterless A	sisA	Yes	Yes	Yes	Yes	Yes	Yes	Yes
degringolade	dgrn	Yes	Yes	Yes	Yes	Yes	Yes	Yes
modifier of mdg4	mod (mdg4)	Yes	Yes	Yes	Yes	Yes	Yes	No
dissatisfaction	dsf	NF	NF	NF	NF	NF	NF	NF
female lethal d	fl (2)d	Yes	Yes	Yes	Yes	Yes	Yes	Yes
hermaphrodite	her	NF	NF	NF	NF	NF	NF	NF
outstretched	os (sisC)	NF	NF	NF	NF	NF	NF	Yes
Signal-transducer and activator of transcription protein at 92E	Stat92E	Yes	No	No	NF	NF	Yes	Yes
transformer	tra	Yes	Yes	Yes	Yes	Yes	Yes	Yes
groucho	gro	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mes-4	Mes-4	Yes	Yes	Yes	Yes	Yes	Yes	No



**Fig. 4.** Transcriptomic comparison of embryogenesis between *A. ludens* with *B. dorsalis*, *C. capitata* and *D. melanogaster*. A) Timing of *A. ludens*, *C. capitata*, *B. dorsalis* and *D. melanogaster* embryo development at laboratory conditions at 25 °C. Timing comparison between the *A. ludens* and *D. melanogaster* early embryogenesis at bottom. B) Embryo transcripts data comparison between FPKMs of *B. dorsalis* FPKMs data (0–24 h) and *A. ludens* early embryo transcripts data, 50 transcripts are shared in syncytium, 94 in cellularization and 176 in gastrulation. C) *C. capitata* (8–10 h) FPKMs gastrulation transcripts data comparison with *A. ludens* data, 17 transcripts are shared in syncytium, 35 in cellularization and 63 transcripts in gastrulation. D) Early embryo transcripts data comparison between FPKMs of *Drosophila* (0–2 h and 2–4 h) and *A. ludens* data, 3 transcripts are shared in syncytium, 36 in cellularization and 50 in gastrulation. E) Global comparison between *C. capitata* (8–10 h), *B. dorsalis* (0–8 h), *D. melanogaster* (0–4 h) and *A. ludens* (0–30 h) FPKMs. The number of transcripts shared among all the flies are 5356 transcripts, 152 in syncytium, 642 in cellularization, and 19 in gastrulation. *A. ludens*' stages in all the fly's transcriptomic data comparison are represented in colors, blue for syncytium, light green for cellularization and green for gastrulation embryo stages. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

in *Drosophila*. Together, the results presented in this section indicate that the transcriptomes that we examined represent, with high confidence, the global gene expression patterns of the first developmental stages of the *A. ludens* embryo.

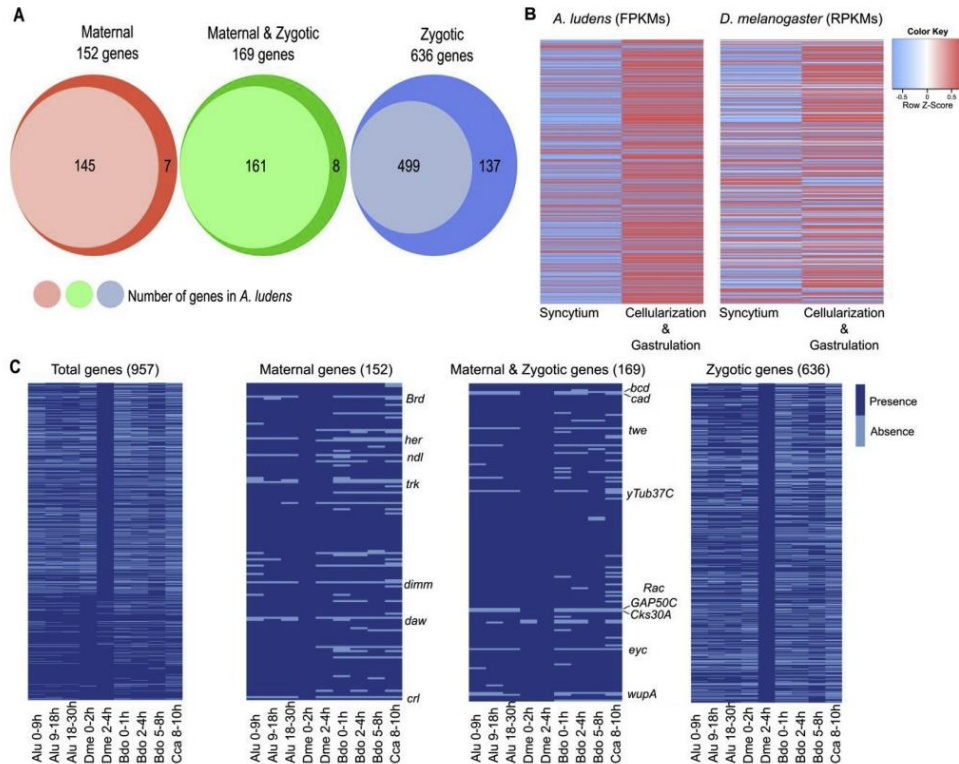
### 3.4. Comparison between the early embryo transcriptomes of *A. ludens* and *B. dorsalis*, *C. capitata* and *D. melanogaster*

In order to gain insight into shared and unique features of embryonic development in tephritids, we compared the expression data from the different embryogenesis stages of *A. ludens* with those reported for *B. dorsalis* from embryos at 0–1, 2–4, and 5–8 h of development and those reported for *C. capitata* from embryos at the gastrulation stage at 8–10 h (Salvemini et al., 2014). In addition, we included in this analysis the *D. melanogaster* transcriptome from embryos at 0–2 and 2–4 h that correspond to the same *A. ludens* developmental stages analysed here (Fig. 4A). An important point in embryo development between these flies is the difference in temporality (Fig. 4A), with the longest development time occurring in *A. ludens* and the stages that we analysed corresponding to 30 h of development, which in *D. melanogaster* only takes 3 h (Fig. 4A).

We determined the number of embryonic transcripts that are shared between *A. ludens* and the other flies and found that between 5000 and 6000 genes of the other three species were shared with *A. ludens* during the three embryonic stages analysed here (Fig. 4B–E). In addition, we identified transcripts that are stage specific in *A. ludens* that were also

identified in the other three flies (Fig. 4B–E). Intriguingly, an important number of transcripts identified in *A. ludens* early development were not reported in these developmental stages in the other three flies analysed. Some of these transcripts encode proteins that have been identified in other organisms but not in *D. melanogaster*, *C. capitata* and *B. dorsalis* (Fig. 4B–E; Table S8). These transcripts may encode products that are specific to the early development of *A. ludens*. Among these transcripts, we identified some gene products with high identity to the Zn-finger family present in humans, suggesting an evolutionarily conserved function for these proteins (Table S2).

Since *D. melanogaster* has been an excellent model to study embryo development in animals, multiple genes that participate in the establishment of embryo pattern formation in this species are well known. Thus, we decided to analyse which of these *Drosophila* genes are expressed in the *A. ludens* early embryo. For this analysis, we classified the developmentally relevant *Drosophila* genes into three categories: maternal, which are deposited in the embryo during oogenesis and have a relevant role in early development; zygotic and maternal, which are genes that are expressed maternally but also during zygote genome activation (ZGA); and zygotic genes that are only expressed during ZGA (Fig. 5A). From the 152 *Drosophila* genes expressed maternally that have a role in early development, 145 were present in the *A. ludens* embryo transcriptome (Fig. 5A). For the maternal and zygotic genes only 9 were not found in *A. ludens*, and 499 genes expressed in ZGA were present. Therefore, most of the *Drosophila* genes that have been linked with early embryo development were found in the *A. ludens*



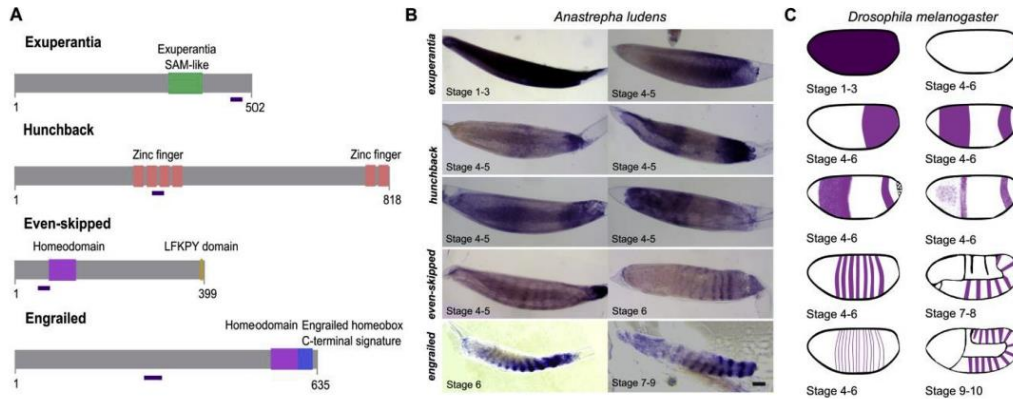
**Fig. 5.** Genes involved in maternal or zygotic developmental process in *A. ludens*. A) Genes involved in maternal or zygotic developmental programs originally described in *D. melanogaster* and identified in *A. ludens*. B) Comparison between the abundance of maternal or zygotic transcripts in syncytium relative to both cellularization & gastrulation events in *A. ludens* and *Drosophila*. C) Presence or absence of maternal, maternal & zygotic and zygotic transcripts (in FPKMs) reported in the embryos of *A. ludens* as Alu, *D. melanogaster* as Dme, *B. dorsalis* as Bdo, and *C. capitata* as Cca.

transcriptome. To know the expression behaviour of these developmentally relevant genes, we performed a cluster analysis that shows a similar behaviour between these two flies, even though the timing differences (Fig. 5B).

Next we compared these *Drosophila* homologues genes against *B. dorsalis*, *C. capitata* and *A. ludens* embryonic transcriptomes. The cluster analysis shows that *A. ludens* and *Bactrocera* early transcriptomes had higher similarity with *Drosophila* than with *Ceratitidis* (Fig. 5C); and *Ceratitidis* embryo transcriptome was the most different among the four fly species (Fig. 5C). Intriguingly, there is a significant number of gene transcripts that have been identified to be highly relevant for *Drosophila* early development that we did not find in *A. ludens*, *Bactrocera*, or *Ceratitidis*. For instance, *bicoid* (*bcd*), *caudal* (*cad*), and *trunk* (*trk*), which have a relevant role in the anterior-posterior determination, and genes that participate in germinal layer establishment, such as *twine* (*twe*) and *nudel* (*ndf*), were not detected (Table S8). Together, this analysis indicates that most of the important genes required for early embryo development in *Drosophila* are also present in *A. ludens*, however it seems that some important genes for early *Drosophila* development are not present in tephritids.

### 3.5. Expression patterns of genes required for early *A. ludens* embryo development

Gene expression patterns that modulate early embryo development in *Drosophila* have provided a powerful tool to understand the roles in the anterior-posterior and dorso-ventral specification as well as the determination of the segments in the organism. Thus, we developed a protocol for *in situ* hybridization optimized for *A. ludens*, to visualize the expression pattern of genes relevant to early embryo development (see 2. Materials and methods). We selected 4 genes that are well characterized in *Drosophila* embryos. These included *exuperantia* (*exu*), which is maternally derived and is initially ubiquitously distributed but preferentially located in the pole cells during cellularization (Hazelrigg et al., 1990); *hunchback* (*hb*), a GAP gene, which is initially maternally derived, and its transcript is ubiquitously distributed in the early embryo, but is also zygotically expressed, initially concentrating its transcript in the anterior region of the embryo and later in two broad bands in the middle and posterior part of the embryo (Tautz et al., 1987); *evenskipped* (*eve*) a pair rule gene, expressed in 7 strips in the trunk of the embryo (Macdonald et al., 1986); and *engrailed* (*en*) a segment polarity gene that it is expressed in 14 strips to determine the posterior part of each segment, also in the trunk of the embryo (DiNardo et al., 1985).



**Fig. 6.** Transcripts localization of genes involved in the anterior-posterior pattern of *A. ludens* embryo. A) Maternal and zygotic proteins involved in early embryogenesis whose transcripts are detected in B. Characteristic domains and the length for each protein are indicated. Exuperantia is a maternal protein involved in the determination of antero-posterior axis in some insects including *Drosophila*. Important for *bicoid* mRNA localization at the posterior pole of the embryo. Hunchback, Even-skipped and Engrailed proteins are zygotic segmentation proteins (gap, pair rule and segment-polarity, respectively). Hunchback also has a maternal contribution component. B) Expression patterns of genes expressed at early embryogenesis in *A. ludens*. Representative images of *A. ludens* embryos at different stages, where *in situ* localization of the mRNAs encoding the proteins indicated in A, (purple stripe represents the probe localization), was performed. Stages are indicated. Bar represents 100  $\mu\text{m}$ . C) A reference scheme of the expression pattern of the same genes in *Drosophila*. Stages are indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

For the set of selected sequences, protein homology between *D. melanogaster* and *A. ludens* was approximately 60% for *exu*, 56% for *hb*, 64% for *eve* and 68% for *en* (Fig. 6A). Based on this information, we generated *A. ludens* specific probes for *in situ* hybridization. In the case of *exu*, we only detected embryos in which the mRNA was in the whole embryo and others from the middle to the posterior part of embryo as expected ( $n = 25$ ), but not those with a specific localization in the pole cells in cellular blastoderm embryos as in *Drosophila* (Fig. 6C). In the case of *hb*, most of the cellularized embryos ( $n = 35$ ) showed an expression pattern of two broad bands in the middle and the posterior region, similar to the pattern observed in *Drosophila* embryos at the same stage (Fig. 6C). The *A. ludens eve* transcript was visualized as the expected 7 strips in the trunk of the embryos in cellularized and gastrulated embryos ( $n = 12$ ) (Fig. 6B). Finally, for *en*, as in *Drosophila*, the expression pattern of 14 strips in gastrulated embryos and during germ band elongation was also identified ( $n = 12$ ) (Fig. 6B). These results show that at least for the set of 4 genes analysed, their expression patterns in *A. ludens* early embryos was similar to that observed in *Drosophila*, and only *exu* showed a difference in the localization of its transcript at the pole cells during cellularization and gastrulation.

#### 4. Discussion

Since the beginning of the last century, when adult *A. ludens* flies were observed to emerge from infested oranges and mangos (Herrera et al., 1903), its study has gained great relevance. The initial studies addressed basic biological questions on their biology, such as their life cycle, their hosts, and their breeding rituals (Aluja et al., 2000). As the studies and technologies progressed, campaigns to eradicate the Mexican fruit fly were created, in which one of the main tools for pest management is the use of the sterile insect technique (SIT), an effective, environmentally friendly method to manage the fruit fly (Gutiérrez Ruelas et al., 2013). Despite their importance, studies of the embryonic stages of this pest have only focused on embryo shape (Carroll and Wharton, 1989; Emmart, 1933), which is used as a tool for the identification of flies in infested plantations. However, a detailed and chronological description of the embryogenesis of *A. ludens* is necessary

for our understanding of the biology of this socially and economically important pest species.

We performed observations in live and stained embryos to visualize the primordial stages of embryogenesis, based on those already established for *D. melanogaster* (Campos-Ortega and Hartenstein, 1985). This description covered the entire duration of the 120 h of *A. ludens* embryo development and additional processes specific to the fly. One of the most important distinctive characteristics of *A. ludens* embryogenesis is its timing; the temporality of *A. ludens* embryogenesis is a characteristic that is known to depend primarily on temperature, previously reported to take 9 days at 19.6 °C (McPhail and Bliss, 1933) or as few as 3.25 days at 28 °C (Celedonio-Hurtado et al., 1988). *A. ludens* embryogenesis takes 5.6 times longer than in model organisms such as *D. melanogaster*, which occurs over 22 h at 25 °C (Campos-Ortega and Hartenstein, 1985). Among tephritid flies, the difference is less dramatic see Fig. 4A; however, none of these time frames are especially long. The extremely long embryonic development of *A. ludens* could be related to the size of the egg, as well as the amount of time needed to carry out the cell divisions important for embryo development. Additionally, this time in *A. ludens* embryo development could be an important trait for the correct regulation of gene expression, as well as phenotypic embryo changes, such as yolk mass extrusions. Interestingly there is a correlation between the embryogenesis timing and the size of the genome since it is of 139 Mb for *D. melanogaster*, 322 Mb for *B. dorsalis*, 591 Mb for *C. capitata* and 1555 Mb in *A. ludens* (De Teodoro Pardo, 2011; Tsoumani and Mathiopoulos, 2012), suggesting that part of the long timing of the *A. ludens* embryo development, may be related to the replication of an extensive genome.

With this extended time of embryogenesis, we were interested in comparing the temporality of the stages reported in *Drosophila* embryo growth to understand the possible cellular and morphological regulatory mechanisms. Embryogenesis comparison with *D. melanogaster*, we observed a heterochronic phenomenon in the development of *A. ludens*. A heterochronic change in head involution occurs simultaneously with germ band shortening in *A. ludens* at 69–73 h; in *D. melanogaster*, head involution is observed after germ band shortening and is related to dorsal ventral closure at 10:20 to 11:20 h. These changes

relative to the timing of developmental events between *Drosophila* and *A. ludens*, such as head involution and embryogenesis timing, can serve as a breaking point to generate phenotypic diversity between both flies and changes in the onset, offset or rate of gene expression, as well as gene spatial expression patterns. The extended time of *A. ludens* embryogenesis could give us a better understanding of what happens in processes such as the maternal to zygotic transition, which in some species takes minutes, or provide an extensive time gap to study the processes of germinal line specification, which in *A. ludens* can last up to 12 h, among other events. However, these long embryonic stages may also allow non-synchronous zygotic genome activation in the early embryo (see ahead).

The phenomenon of extrusions has been reported in other flies of the genus *Anastrepha* (Selivon et al., 1996); (Perondini et al., 1998). However, in flies such as *C. capitata* and *B. tryoni*, there are no reports of extrusion events (Perondini et al., 1998; Selivon and Perondini, 1997). The moment of development at which these extrusions occur in *Anastrepha* flies corresponds to the morphogenetic events of gastrulation and head lobe formation (Selivon and Perondini, 1997), which is in agreement with the data that we observed in *A. ludens*, with the peculiarity that there is a first posterior extrusion before gastrulation. Therefore, this information could reinforce the idea that the extrusions could be a specific trait of flies of the genus *Anastrepha*, but in spite of its specificity, this may be different among the species of this genus. Because we observed that extrusions are a phenomenon that occurs normally in the embryonic development of *A. ludens*, the moments at which these events occur could be used as temporary markers of development. Intriguingly, we observed that the synchronized mitotic divisions that occur during the syncytial blastoderm start in the middle of the embryo and then extend to the poles. This is different from what has been observed in *Drosophila*, in which the mitotic waves start at the poles and progress toward the middle of the embryo. It is possible that this difference may be linked to the extrusion of the anterior and posterior regions, since nuclei are eliminated together with the yolk material.

The mechanisms underlying extrusions are still unknown; however, it has been suggested that they have a genetic component (Selivon and Perondini, 1997). We determined the loss of tissue due to extrusion by stages of development, that corresponds to 48.7% of embryo length of the total egg. We propose that the extrusions in *A. ludens* are a mechanism for the maintenance of adequate embryo size; in the spindle egg morphology at the beginning of embryogenesis, size needs to change according to the progress of development to maintain the appropriate embryo shape.

Once the morphological studies were completed, the description of the gene expression patterns in *A. ludens* embryo development was the next step to understand, in particular, its early stages, since the heterogeneity of different cell types is not very complex as it is in most advanced stages. The amount of different transcripts in the syncytial, cellularization and gastrulation stages was similar that reported for *Drosophila*. A similar situation has also been reported in *C. capitata* and *B. dorsalis* embryos (Salvemini et al., 2014). Intriguingly, we identified more than 400 gene transcripts that either have not been identified in *Drosophila* or that are expressed in other developmental stages. It is possible that in the case of *Drosophila*, several embryonic transcripts in the interval from syncytial blastoderm to gastrulation have not been identified, including non-coding RNAs as well as transcripts that encode for unknown ORFs. Additionally, it is possible that the extended developmental timing in the *A. ludens* embryos results in an overlap with transcripts that, in *Drosophila*, are restricted to specific developmental stages. Indeed, this is something that we observed in our analysis.

The sex determination genes are conserved in Diptera, although *sexlethal* (*sxl*) is not always related to sex determination in all species (Saccone et al., 2002). This pathway is relevant for the putative biological control of pest flies and mosquitoes. We found that most of the genes that participate in sex determination in *Drosophila* are also

expressed in the early embryo of *A. ludens*. Only *dsf*, *her* and *Os* were not found in our analysis. Intriguingly, *dsf* and *her* have also not been described in *C. capitata* or *B. dorsalis*, and therefore, it is possible that their role in sex determination has been replaced by other genes or through evolution, these genes were not needed for sex determination in tephritids.

The broad comparison of the *A. ludens* early embryo transcriptome with that of other flies revealed that most of these transcripts are also present in the other organism analysed here; however, some transcripts were only present in *A. ludens*. Intriguingly, some of these genes exist in other animals, including humans, suggesting that these missing genes are more similar to the human genes than to other fly genes or that they have not been annotated in these flies. When we compared our *A. ludens* data with the *Drosophila* genes that are known to have a direct role in early embryo development, we found that in the case of maternal genes and those that participate in the activation of the maternal to zygotic transition, only a few were not identified in *A. ludens*. For example, the absence of *bcd* is not surprising since this gene seem to be specific to *Drosophila* and is not present in the other flies analysed. In relation to this point, it should be noted that the transcriptome of *A. ludens* revealed fewer differences than those of *B. dorsalis* and *C. capitata* when compared with that of *Drosophila*. However, after ZGA in *Drosophila* (2–4 h of embryo development), we observed more differences in the presence or absence of genes identified between the four species. It is possible that this may be associated with a non-synchronized ZGA in tephritids and the extended timing of this developmental stage in *A. ludens*; and in future studies, it will be important to determine the timing of ZGA in these flies.

Among the genes that participate in the early development of *Drosophila* embryos and that were not identified in *A. ludens*, most are also not present in the transcriptomes of *C. capitata* and *B. dorsalis*. It is possible that the absence of these genes in our transcriptome analysis again may be related to the long embryo development of *A. ludens*; however, since these genes were also not found in the other flies, we hypothesize that these genes are not required in the early development in these organisms. Future transcript analysis of the next embryonic development stages or the complete genome sequence will be required to determine if these genes are present and expressed during *A. ludens* embryo development.

The standardization of a protocol for *in situ* hybridization optimized for *A. ludens* allowed us to analyse the expression patterns of classical *Drosophila* genes that have an important role in early development. In general, the expression patterns of the four genes we selected (*exu*, *hb*, *eve* and *en*) were similar to that in *Drosophila* embryos. The only difference corresponded to *exu*, as we did not detect a signal in polar cells, as in *Drosophila*. This may be a distinctive characteristic of *A. ludens*; however, polar cells in embryos of this fly are difficult to observe, even in embryos stained with DAPI. Further analysis of the expression patterns of additional genes will be important to obtain an expression atlas of early-stage *A. ludens* embryos.

In conclusion, in this work, we have generated the first detailed description of the embryogenesis of *A. ludens*. This is reinforced by the first transcriptome of the early embryonic stages of this organism, which has given us relevant information about the similarities and differences in genetic expression of these embryos compared to relative to other flies. We standardized a protocol for *in situ* hybridization to analyse the expression patterns of genes important for the development of this organism. In addition to its biological interest, *A. ludens* is of great economic and social interest, and we believe our studies represent a platform for a better understanding of the biological process and for future strategies to control this fruit pest. For example, the genetic manipulation by injecting embryos with DNA to include desirable traits in mass reared flies (Gabrieli and Scolari, 2016; Meza et al., 2014).

## 5. Funding

This work was supported by PAPPIT/UNAM grant No. IN200315 and CONACYT No. 25088 to MZ; and Agropolis Fondation No. 1502-307 and Royal Society Newton Advanced Fellowship NA150181 to MAAV.

## 6. Data availability

Accession number for the raw and processed RNA-seq data reported in this paper is GSE144185.

## 7. Author contributions

X.G.R. performed data collection and experiments, analysed data, verified the overall reproducibility of results, interpreted results and wrote the draft.

M.V. analysed data, interpreted results, verified the overall reproducibility of results, advised and reviewed the draft.

F.D.F. design methods, provided animals and laboratory facilities, interpreted results and advised.

C.A.P.A. Design and implementation of a De Novo transcriptome assembly annotation and differential expression analysis strategies.

H.N.N.M. Design and implementation of a De Novo transcriptome assembly annotation and differential expression analysis strategies.

R.G.A.M. Bioinformatic method adviser and analysis of De Novo transcriptome.

F.R.T. Method design adviser and draft reviewer.

A.E.D.A. analysed data, interpreted results, advised and oversaw the entire research endeavour and secured funds.

M.A.A.V. conceived the idea, designed methods and experiments, analysed data, interpreted results, advised and oversaw the entire research endeavour, secured funds and review the manuscript.

M.Z. designed methods and experiments, analysed data, interpreted results, advised and oversaw the entire research endeavour, secured funds and wrote the manuscript.

## Declaration of competing interest

No competing interests declared by the authors.

## Acknowledgments

X. G-R. is a doctoral student from the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), and received a fellowship (448454) from Consejo Nacional de Ciencia y Tecnología (CONACyT). We are grateful to Diego Cortez for bioinformatic approach suggestions; Evaristo Calihua, Cruz Perea Castellanos, José Luis Lorenzo Manzanarez and Dulce Olivia Flores Martínez for the technical assistance with flies; Luis Cruz Silva for technical microscopy assistance; and to José Salvador Meza Hernández of Programa Moscafrut. We thank the Instituto de Biotecnología-UNAM for giving us access to its computer cluster, especially to M.C. Jerome Jean Verleyen; and to Cuerpo Académico CA-UVER-234 of the Instituto de Biotecnología y Ecología Aplicada – Universidad Veracruzana.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ibmb.2020.103412>.

## References

- Aluja, M., Piñero, J., Jácome, I., Díaz-Fleischer, F., Sivinski, J., 2000. In: Aluja, M., Norrbom, A.L. (Eds.), Behavior of Flies of the Genus *Anastrepha*. Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior. CRC Press, DelRay Beach FL, pp. 375–408.

- Anderson, D.T., 1962. The embryology of *Dacus tryoni* (frogg.) [Diptera, tephritidae (= Tephritidae)], the Queensland fruit-fly. Development 10, 248–292.
- Andrews, S., 2018. FastQC.
- Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, M., 1944. A Review of Studies on the Mexican Fruitfly and Related Mexican Species. U. S. Department of Agriculture, Washington, D.C.
- Berrigan, D.A., Carey, J.R., Guillen, J., Celedonio, H., 1988. Age and host effects on clutch size in the Mexican fruit fly *Anastrepha ludens*. Entomol. Exp. Appl. 47, 73–80. <https://doi.org/10.1111/j.1570-7458.1988.tb02284.x>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34, 525–527. <https://doi.org/10.1038/nbt.3519>.
- Brian, H., Papanicolaou, A., 2019. Transdecoder (Finding Coding Regions within Transcripts).
- Calla, B., Hall, B., Hou, S., Geib, S.M., 2014. A genomic perspective to assessing quality of mass-reared SIT flies used in Mediterranean fruit fly (*Ceratitis capitata*) eradication in California. BMC Genom. 15, 1–19. <https://doi.org/10.1186/1471-2164-15-98>.
- Cameron, J., 1905. The development of the retina in Amphibia: an embryological and cytological study: Part I. J. Anat. Physiol. 39, 332–348.5.
- Campos-Ortega, J.A., Hartenstein, V., 1985. The Embryonic Development of *Drosophila melanogaster*, first ed. Springer-Verlag Berlin Heidelberg GmbH.
- Carroll, L.E., Wharton, R.A., 1989. Morphology of the immature stages of *Anastrepha ludens* (Diptera: Tephritidae). Ann. Entomol. Soc. Am. 82, 201–214. <https://doi.org/10.1093/aesa/82.2.201>.
- Celedonio-Hurtado, H., Liedo, P., Aluja, M., Guillen, J., Berrigan, D., Carey, J., 1988. Demography of *Anastrepha Ludens*, *A. Obligua* and *A. Serpentina* (Diptera: Tephritidae) in Mexico. Florida Entomol. <https://doi.org/10.2307/3495357>.
- Chen, S.L., Dai, S.M., Lu, K.H., Chang, C., 2008. Female-specific doublesex dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, *Bactrocera dorsalis* (Hendel). Insect Biochem. Mol. Biol. 38, 155–165. <https://doi.org/10.1016/j.ibmb.2007.10.003>.
- Clark, W.H., 1960. Electron microscope studies of nuclear extrusions in pancreatic acinar cells of the rat. J. Biophys. Biochem. Cytol. 7, 345–352. <https://doi.org/10.1083/jcb.7.2.345>.
- De Teodoro Pardo, C.V., 2011. Evolucion de genetica de cuatro especies de *Anastrepha* spp. (Diptera:Tephritidae). PhD tesis. Colegio de Postgraduados, Campus Montecillo.
- Díaz-Fleischer, F., Aluja, M., 2003. Clutch size in frugivorous insects as a function of host firmness: the case of the tephritid fly *Anastrepha ludens*. Ecol. Entomol. 28, 268–277. <https://doi.org/10.1046/j.1365-2311.2003.00517.x>.
- Dickens, J.C., Solis, E., Hart, W.G., 1982. Sexual Development and Mating Behavior of the Mexican Fruit Fly, *Anastrepha Ludens* (Loew). Entomol. Southwest.
- DiNardo, S., Kuner, J.M., Theis, J., O'Farrell, P.H., 1985. Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear engrailed protein. Cell 43, 59–69. [https://doi.org/10.1016/0092-8674\(85\)90012-1](https://doi.org/10.1016/0092-8674(85)90012-1).
- do Nascimento, J.C., Kalisz de Oliveira, A., 1996. Embryogenesis in *Anastrepha fraterculus* (Diptera: Tephritidae). Interciencia 21, 158–165.
- Dutra, V.S., Ronchi-Teles, B., Steck, G.J., Silva, J.G., 2013. Description of eggs of *Anastrepha curtis* and *Anastrepha leptozona* (Diptera: Tephritidae) using SEM. Ann. Entomol. Soc. Am. 106, 13–17. <https://doi.org/10.1163/an12054>.
- Emmert, E.W., 1933. The eggs of four species of fruit flies of the genus *Anastrepha*. Proc. Entomol. Soc. Wash. 35, 184–191.
- Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>.
- Figueiredo, J.V.A., Perondini, A.L.P., Ruggiro, E.M., Prezotto, L.F., Selivon, D., 2013. External eggshell morphology of *Anastrepha* fruit flies (Diptera: Tephritidae). Acta Zool. 94, 125–133. <https://doi.org/10.1111/j.1463-6395.2011.00533.x>.
- Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>.
- Gabrieli, P., Scolari, F., 2016. Delivery of nucleic acids through embryo microinjection in the worldwide agricultural pest insect, *Ceratitis capitata*. JoVE 116, e54528. <https://doi.org/10.3791/54528>.
- Genç, H., 2014. Embryonic development of the olive fruit fly, *Bactrocera oleae* Rossi (Diptera: Tephritidae), in vivo. Turk. J. Zool. 38, 598–602. <https://doi.org/10.3906/zoo-1305-19>.
- Gomulski, L.M., Dimopoulos, G., Xi, Z., Scolari, F., Gabrieli, P., Siciliano, P., Clarke, A.R., Malacrida, A.R., Gasperi, G., 2012. Transcriptome profiling of sexual maturation and mating in the mediterranean fruit fly, *Ceratitis capitata*. PLoS One 7. <https://doi.org/10.1371/journal.pone.0030857>.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. <https://doi.org/10.1038/nbt.1883>.
- Greene, C.T., 1929. Characters of the larvae and pupae of certain fruit flies. J. Agric. Res. 38, 489–504.
- Gutiérrez Ruelas, J.M., Santiago Martínez, G., Villaseñor Cortés, A., Enkerlin Hoefflich, W.R., Hernández López, F., 2013. Los programas de Moscas de la fruta en México, su historia reciente, "Los programas de moscas de la fruta en México Su historia reciente." Instituto Interamericano de Cooperación para la Agricultura (IICA), México, D.F.
- Hazelrigg, T., Watkins, W.S., Marcey, D., Tu, C., Karow, M., Lin, X., 1990. The *exuperantia*

- gene is required for *Drosophila* spermatogenesis as well as anteroposterior polarity of the developing oocyte, and encodes overlapping sex-specific transcripts. *Genetics* 126, 607–617.
- Hernández Livera, R.A., 2014. Manual técnico para la identificación de moscas de la fruta. Herrera, A.L., Rangel, A.F., De la Barreda, L., Bonansea, S., Tellez, O., 1903. Comisión de Parasitología agrícola, Boletín de la Comisión de Parasitología Agrícola. México.
- Hodar, C., Cambiazio, V., 2018. The dorsoventral patterning of *Musca domestica* embryos: insights into BMP/Dpp evolution from the base of the lower cycloraphan flies. *EvoDevo* 9, 1–12. <https://doi.org/10.1186/s13227-018-0102-5>.
- Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. <https://doi.org/10.1093/nar/gkn923>.
- Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. <https://doi.org/10.1038/nprot.2008.211>.
- Kalinka, A.T., Tomancak, P., 2012. The evolution of early animal embryos: conservation or divergence? *Trends Ecol. Evol.* 27, 385–393. <https://doi.org/10.1016/j.tree.2012.03.007>.
- Klingenberg, C.P., 2011. MorphoJ: an integrated software package for geometric morphometrics. *Mol. Ecol. Resour.* 11, 353–357. <https://doi.org/10.1111/j.1755-0998.2010.02924.x>.
- Kozomara, A., Griffiths-Jones, S., 2011. miRBase.
- Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
- Liedo, P., Carey, J.R., Celedonio, H., Guillen, J., 1992. Size specific demography of three species of *Anastrepha* fruit flies. *Entomol. Exp. Appl.* 63, 135–142.
- Lott, S.E., Villalta, J.E., Schroth, G.P., Luo, S., Tonkin, L.A., Eisen, M.B., 2011. Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-Seq. *PLoS Biol.* 9. <https://doi.org/10.1371/journal.pbio.1000590>.
- Macdonald, P.M., Ingham, P., Struhl, G., 1986. Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell* 47, 721–734. [https://doi.org/10.1016/0092-8674\(86\)90515-5](https://doi.org/10.1016/0092-8674(86)90515-5).
- McCarthy, D.J., Chen, Y., Smyth, G.K., 2012. Differential expression analysis of multi-factor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. <https://doi.org/10.1093/nar/gks042>.
- McPhail, M., Bliss, C.I., 1933. Observations on the Mexican Fruit Fly and Some Related Species in Cuernavaca, Mexico, in 1928. *Ann. Entomol. Soc. Am.* 26, 1–10.
- Meza, J.S., Diaz-Fleischer, F., Sánchez-Velázquez, L.R., Zepeda-Cisneros, C.C., Handler, A.M., Schettelig, M.F., 2014. Fitness cost implications of PhiC31-mediated site-specific integrations in target site strains of the Mexican Fruit Fly, *Anastrepha ludens* (Diptera: Tephritidae). *PLoS One* 9 (10), e109690.
- Papanicolaou, A., Schettelig, M.F., Arensburg, P., Atkinson, P.W., Benoit, J.B., Bourtzis, K., Castañera, P., Cavanaugh, J.P., Chao, H., Childers, C., Curral, I., Dinh, H., Doddapaneni, H.V., Dolan, A., Dugan, S., Friedrich, M., Gasperi, G., Geib, S., Georgakilas, G., Gibbs, R.A., Giers, S.D., Gomulski, L.M., González-Guzmán, M., Guillem-Amat, A., Han, Y., Hatzigeorgiou, A.G., Hernández-Crespo, P., Hughes, D.S.T., Jones, J.W., Karagkouni, D., Koskinioti, P., Lee, S.L., Malacrida, A.R., Manni, M., Mathiopoulos, K., Meccariello, A., Murali, S.C., Murphy, T.D., Muzny, D.M., Oberhofer, G., Ortego, F., Paraskevopoulou, M.D., Poelchau, M., Qu, J., Reczko, M., Robertson, H.M., Rosendale, A.J., Rosselot, A.E., Saccone, G., Salvemini, M., Savini, G., Schreiner, P., Scolari, F., Siciliano, P., Sim, S.B., Tsiamis, G., Ureña, E., Vlachos, I.S., Werren, J.H., Wimmer, E.A., Worley, K.C., Zacharopoulou, A., Richards, S., Handler, A.M., 2016. The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* 17, 1–31. <https://doi.org/10.1186/s13059-016-1049-2>.
- Peel, A.D., 2008. The evolution of developmental gene networks: lessons from comparative studies on holometabolous insects. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1539–1547. <https://doi.org/10.1098/rstb.2007.2244>.
- Pérez-Staples, D., Díaz-Fleischer, F., Montoya, P., Vera, T. (Eds.), 2019. Area-wide Management of Fruit Fly Pests. CRC Press, Boca Raton, FL, USA, pp. 414.
- Perondini, A.L.P., Selivon, D., Morgante, J.S., 1998. Facultative polar extrusion of yolk masses and of hatching at the posterior egg pole in tephritid fruit flies (Diptera). *Int. J. Insect Morphol. Embryol.* 27, 249–254. [https://doi.org/10.1016/S0020-7322\(98\)00017-8](https://doi.org/10.1016/S0020-7322(98)00017-8).
- Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2009. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
- Rohlf, F.J., 2006. tpsDig. Digitize Landmarks and Outlines.
- Saccone, G., Pane, A., Polito, L.C., 2002. Sex determination in flies, fruitflies and butterflies. *Genetica* 116, 15–23. <https://doi.org/10.1023/A:1020903523907>.
- Salvemini, M., Arunkumar, K.P., Nagaraju, J., Sanges, R., Petrella, V., Tomar, A., Zhang, H., Zheng, W., Saccone, G., 2014. De Novo assembly and transcriptome analysis of the mediterranean fruit fly *Ceratitis capitata* early embryos. *PLoS One* 9, 1–27. <https://doi.org/10.1371/journal.pone.0114191>.
- Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L.C., Saccone, G., 2009. *Ceratitis capitata transformer-2* gene is required to establish and maintain the auto-regulation of Cctra, the master gene for female sex determination. *Int. J. Dev. Biol.* 53, 109–120. <https://doi.org/10.1387/jidb.082681ms>.
- 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., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. <https://doi.org/10.1038/nmeth.2019>.
- Selivon, D., Morgante, J.S., Ribeiro, A.F., Perondini, A.L., 1996. Extrusion of masses of yolk during embryonic development of the fruit fly *Anastrepha fraterculus*. *Invertebr. Reprod. Dev.* 29, 1–7. <https://doi.org/10.1080/07924259.1996.9672489>.
- Selivon, D., Perondini, A.L.P., 1997. Extrusion of yolk masses by hybrid embryos of two cryptic species of *Anastrepha fraterculus* (Diptera, Tephritidae). *Braz. J. Genet.* 20, 253–255. <https://doi.org/10.1590/s0100-84551997000200015>.
- Shen, G.M., Dou, W., Niu, J.Z., Jiang, H.B., Yang, W.J., Jia, F.X., Hu, F., Cong, L., Wang, J.J., 2011. Transcriptome analysis of the oriental fruit fly (*Bactrocera dorsalis*). *PLoS One* 6. <https://doi.org/10.1371/journal.pone.0029127>.
- Smit, A.F., Hubley, R., Green, P., 2019a. RepeatModeler open-1.0.
- Smit, A.F., Hubley, R., Green, P., 2019b. RepeatMasker open-3.0.
- Sullivan, W.A., Ashburner, M., Hawley, R.S., 2000. *Drosophila* Protocols, *Drosophila* Protocols. [https://doi.org/10.1007/978-1-61779-034-8\\_15](https://doi.org/10.1007/978-1-61779-034-8_15).
- Tautz, D., Lehmann, R., Schürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., Jäckle, H., 1987. Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. *Nature* 327, 383–389. <https://doi.org/10.1038/327383a0>.
- Tautz, D., Pfeifle, C., 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* 98, 81–85. <https://doi.org/10.1007/BF00291041>.
- Thomson, A., 1919. The maturation of the human ovum. *J. Anat.* 53, 172–208.
- Tsoumani, K.T., Mathiopoulos, K.D., 2012. Genome size estimation with quantitative real-time PCR in two Tephritidae species: *Ceratitis capitata* and *Bactrocera oleae*. *J. Appl. Entomol.* 136, 626–631. <https://doi.org/10.1111/j.1439-0418.2011.01684.x>.