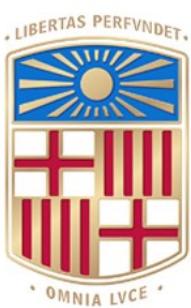


# **Identificación y caracterización de isomiRNA involucrados en el desarrollo de insectos.**



**Universitat  
Oberta  
de Catalunya**



**UNIVERSITAT DE  
BARCELONA**

**Jesús Alarcón Gil**

MU Bioinf. i ABioest.  
Área de trabajo final

**Nombre Tutor/a de TF**  
Guillem Ylla Bou

**Profesor/a responsable de  
la asignatura**  
Guillem Ylla Bou

**Fecha Entrega**  
20 / 06 / 2023



Esta obra está sujeta a una licencia de  
Reconocimiento-NoComercial-  
SinObraDerivada [3.0 España de Creative  
Commons](#)

**FICHA DEL TRABAJO FINAL**

<b>Título del trabajo:</b>	<i>Identificación y caracterización de isomiRNA involucrados en el desarrollo de insectos.</i>
<b>Nombre del autor:</b>	<i>Jesús Alarcón Gil</i>
<b>Nombre del consultor/a:</b>	<i>Guillem Ylla Bou</i>
<b>Nombre del PRA:</b>	<i>Guillem Ylla Bou</i>
<b>Fecha de entrega (mm/aaaa):</b>	<i>06 / 2023</i>
<b>Titulación o programa:</b>	Máster en Bioinformática y Bioestadística
<b>Área del Trabajo Final:</b>	<i>Análisis de datos ómicos</i>
<b>Idioma del trabajo:</b>	<i>Castellano</i>
<b>Palabras clave</b>	<i>Insectos, isomiRNA, desarrollo.</i>

**Resumen del Trabajo**

El descubrimiento de los isomicroRNA (isomiRNA) hace más de una década abrió un gran campo de estudio en la regulación génica. Estos isomiRNA son RNA pequeños no codificantes que difieren sutilmente en secuencia, tamaño o ambas respecto a los microRNA (miRNA) canónicos, pero que pueden variar considerablemente en su función. Por ello, elucidar completamente la importancia y los mecanismos de acción de los isomiRNA es fundamental para comprender todos los mecanismos de regulación génica en los seres vivos. Por ello, este trabajo tiene como objetivo identificar y caracterizar los isomiRNA que participan en la regulación génica del desarrollo del insecto *Blattella germanica*. Para ello, se disponen de datos transcriptómicos de distintas etapas del desarrollo de este insecto, de los que se anotaron los miRNA e isomiRNA, así como sus cuentas. Tras realizar el análisis visual y de enriquecimiento de estos datos, se predijeron sus dianas y las diferencias en las dianas entre los isomiRNA y sus miRNA canónicos. Finalmente, se realizó un modelo de las consecuencias funcionales de los miRNA y isomiRNA a lo largo del desarrollo del insecto. Así, se han obtenido las diferencias de enriquecimiento de los miRNA, isomiRNA y sus tipos en las distintas etapas del desarrollo; las características de las diferencias en las dianas entre los isomiRNA y miRNA; y la participación de ambos en la regulación funcional del desarrollo del insecto. De esta forma, estos resultados son sumamente relevantes para comprender la biología de *Blattella germanica*, el desarrollo y los mecanismos de regulación génica.

## **Abstract**

The discovery of isomicroRNA (isomiRNA) more than a decade ago opened up a large field of study in gene regulation. These isomiRNAs are small noncoding RNAs that differ subtly in sequence, size, or both from canonical microRNAs (miRNAs) but can vary considerably in function. Therefore, fully elucidating the importance and mechanisms of action of isomiRNAs is essential to understand all mechanisms of gene regulation in living beings. Therefore, this work aims to identify and characterize the isomiRNAs that participate in the gene regulation of the development of the insect *Blattella germanica*. To do this, we have transcriptomic data from different stages of the development of this insect, of which the miRNA and isomiRNA were noted, as well as their counts. After performing visual and enrichment analysis of these data, we predict their targets and the differences in the targets between the isomiRNAs and their canonical miRNAs. Finally, we obtain a model of the functional consequences of miRNA and isomiRNA throughout the development of the insect. Thus, we obtained the enrichment differences of miRNA, isomiRNA, and their types in the different stages of development, the characteristics of the differences in the targets between the isomiRNA and miRNA, and the participation of both in the functional regulation of insect development. Thus, these results are highly relevant for understanding *Blattella germanica* biology, development, and gene regulation mechanisms.

# Índice

<b>1. Introducción .....</b>	<b>6</b>
1.1. Contexto y justificación del trabajo .....	6
1.2. Objetivos del trabajo.....	7
1.3. Impacto en sostenibilidad, ético-social y de diversidad .....	7
1.4. Enfoque y método seguido.....	8
1.5. Planificación del trabajo .....	9
1.6. Breve sumario de productos obtenidos .....	10
1.7. Breve descripción de los capítulos de la memoria .....	10
<b>2. Estado del arte .....</b>	<b>11</b>
2.1 miRNA.....	13
2.2 isomiRNA .....	15
2.3 <i>Blattella germanica</i> .....	16
2.4 Justificación y objetivos del trabajo .....	18
<b>3. Materiales y métodos.....</b>	<b>19</b>
3.1. Entornos y datos de entrada. ....	19
3.2. Detección de iso/miRNA. ....	20
3.3. Análisis gráfico y de enriquecimiento de iso/miRNA. ....	21
3.4. Predicciones de los “targets” de los iso/miRNA en los mRNA. ....	22
3.5. Sumarización de los “targets” en los mRNA.....	23
3.6. Análisis gráfico, de enriquecimiento y “GO term” de los mRNA. ....	25
<b>4. Resultados.....</b>	<b>27</b>
4.1. Identificación y cuantificación de los isomiRNA y miRNA presentes en las distintas etapas del desarrollo de <i>Blattella germanica</i> . ....	27
4.2. Análisis de las diferencias en los isomiRNA y miRNA entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	30
4.3. Análisis de las diferencias en las categorías de isomiRNA entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	37
4.4. Predicción de los “targets” en los mRNA de los miRNA e isomiRNA..	39
4.5. Análisis de las diferencias en los “targets” de los isomiRNA respecto a sus miRNA canónicos. ....	40
4.6. Análisis gráfico y de enriquecimiento de la extrapolación de las cuentas de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	43

4.7. Análisis de clusteres funcionales de la extrapolación de las cuentas de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	48
4.8. Interpretación de los resultados funcionales de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	53
<b>5. Discusión</b> .....	55
<b>6. Conclusiones</b> .....	59
<b>7. Bibliografía</b> .....	60
<b>8. Anexos</b> .....	64
8.1 Anexo I .....	64
8.2 Anexo II .....	65
8.3 Anexo III .....	65
8.4 Anexo IV.....	66
8.5 Anexo V.....	67
8.6 Anexo VI.....	68
8.7 Anexo VII.....	69
8.8 Anexo VIII.....	70
8.9 Anexo IX.....	71

# **Lista de figuras**

Figura 1. Representación en detalle del dogma central de la Biología.....	11
Figura 2. Clasificación de los distintos tipos de RNA. ....	12
Figura 3. Estructura y procesamiento del mRNA. ....	13
Figura 4. Mecanismo de biogénesis canónica de los miRNA.....	14
Figura 5. Clasificación de los isomiRNA.....	15
Figura 6. Distribución de la regulación génica a lo largo del desarrollo de <i>Blattella germanica</i> .....	17
Figura 7. Diagrama de flujo desarrollado del pipeline seguido para ejecutar este trabajo. ....	26
Figura 8. Análisis de calidad de 2 de las muestras de RNA-seq no codificante del desarrollo de <i>Blattella germanica</i> . ....	28
Figura 9. Representación de la alineación de los datos brutos de una de las muestras con el genoma de <i>Blattella germanica</i> mediante el software IGV....	29
Figura 10. Gráficas de análisis de componentes principales de los miRNA (A) y los isomiRNA (B) entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	31
Figura 11. Heatmap de distancias de las anotaciones de los miRNA en las muestras de los estadios del desarrollo de <i>Blattella germanica</i> . ....	33
Figura 12. Heatmap de distancias de las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de <i>Blattella germanica</i> . ....	33
Figura 13. Heatmap clusterizado de las anotaciones de los miRNA en las muestras de los estadios del desarrollo de <i>Blattella germanica</i> .....	34
Figura 14. Heatmap clusterizado de las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de <i>Blattella germanica</i> . ....	35
Tabla 1. Correlación entre los grupos observados en los heatmaps de distancias, los heatmaps clusterizados y las etapas del desarrollo para los miRNA y los isomiRNA.....	35
Tabla 2. Número de miRNA con un enriquecimiento significativo (p-valor menor de 0.001) entre cada una de las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	36
Tabla 3. Número de isomiRNA con un enriquecimiento significativo (p-valor menor de 0.001) entre cada una de las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	36
Figura 15. Gráfico de análisis de componentes principales de los isomiRNA según su categoría entre las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	37
Figura 16. Heatmap de distancias de las categorías de isomiRNA según las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de <i>Blattella germanica</i> .....	38

Tabla 4. Número de categorías de isomiRNA enriquecidas significativamente (p-valor < 0.001) entre cada una de las distintas etapas del desarrollo de <i>Blattella germanica</i> .....	39
Figura 17. Ejemplo gráfico de la predicción de los targets de los miRNA / isomiRNA .....	40
Figura 18. Heatmap del número de las diferencias en los targets en el conjunto de los isomiRNA respecto a su miRNA canónico para cada uno de los mRNA. ....	41
Figura 19. Heatmap del número de las diferencias en los targets de los isomiRNA respecto a su miRNA canónico para cada uno de los mRNA y agrupados por la categoría de los isomiRNA. ....	42
Figura 20. Heatmap de distancias de las “cuentas de inhibición” por parte de los miRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> ..	44
Figura 21. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los miRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> .....	45
Figura 22. Heatmap de distancias de las “cuentas de inhibición” por parte de los isomiRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> . ....	46
Figura 23. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> .....	46
Figura 24. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> .....	47
Figura 25. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA sobre los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> .....	48
Figura 26. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa adulta. ....	50
Figura 27. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa ED0. ....	51
Figura 28. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa NFE.....	52
Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa NFE. ....	66
Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa ED0.....	67

Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa NFE .....	68
Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa adulta.....	69
Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa ED0. ....	70
Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de <i>Blattella germanica</i> respecto a la etapa adulta. ....	71

# 1. Introducción

## 1.1. Contexto y justificación del trabajo

La regulación génica es el proceso que controla la correcta expresión de los distintos genes en las dimensiones espaciales y temporales. De esta manera, éste es un proceso dinámico en el que participan numerosos mecanismos para lograr el correcto funcionamiento y adaptación de los organismos al entorno mediante la correcta expresión de los distintos genes. Exactamente, uno de los mecanismos incluidos dentro de esos mecanismos de regulación génica (epigenéticos, transcripcionales, traduccionales, etc.) es el mecanismo de regulación post-transcripcional realizado por los micro-RNA (miRNA) (1, 2).

Los miRNA son moléculas pequeñas de RNA no codificante, exactamente de entre 19 y 24 nucleótidos, pertenecientes al grupo de RNA de interferencia. Por ello, la principal función de los miRNA es la regulación de la expresión génica mediante la represión postranscripcional de determinados RNA mensajeros (mRNA) (2-4). Sin embargo, según se avanza en el estudio de los mecanismos de acción de los miRNA, se va descubriendo una mayor diversidad de mecanismos de acción de estos miRNA. Estos mecanismos de acción van desde miRNA no canónicos (isomiRNA, miRtrons, etc.) hasta la activación o silenciamiento transcripcional directamente en el núcleo (3, 4).

Exactamente, durante las últimas décadas se descubrieron los isomiRNA, pequeñas modificaciones en longitud y/o secuencia de los miRNA canónicos originalmente anotados que, en un principio, se identificaron como errores técnicos, pero que, posteriormente, se han demostrado reales y funcionales. La importancia de estos isomiRNA, aparte de la evidencia existente de su abundancia en los organismos, reside en que pueden provocar cambios funcionales respecto a sus miRNA canónicos. En concreto, los isomiRNA pueden provocar cambios en la secuencia que reconoce los mRNA y, por tanto, cambiar la fortaleza de la unión a los mRNA diana o, directamente, los mRNA diana sobre los que actúan respecto a sus miRNA canónicos. Además, también pueden cambiar su propia estabilidad e incluso provocar su degradación respecto a sus miRNA canónicos, por lo que también alteran el alcance de sus efectos. Sin embargo, los cambios funcionales que provocan los isomiRNA respecto a sus miRNA canónicos están siendo ampliamente investigados actualmente por la necesidad de ampliar el conocimiento de éstos (3, 5, 6).

Estos estudios para conocer el papel e importancia de los isomiRNA en la regulación génica se basan en técnicas ómicas, concretamente, epitranscriptómicas, las cuales necesitan un gran desarrollo y recursos bioinformáticos. Por ello, en los últimos años se han desarrollado varias herramientas orientadas a la identificación y caracterización de estos isomiRNA (6). Sin embargo, actualmente es necesario obtener y aplicar conclusiones mediante estas herramientas bioinformáticas sobre el papel de los isomiRNA en distintos procesos biológicos, como pueden ser procesos oncológicos, envejecimiento o el desarrollo. Es por ello por lo que este trabajo propone identificar y caracterizar el papel de los isomiRNA en el desarrollo de los insectos, concretamente en la cucaracha *Blattella germanica*, mediante la aplicación de las herramientas bioinformáticas disponibles sobre datos transcriptómicos de distintas fases del desarrollo de este organismo.

## **1.2. Objetivos del trabajo**

### **2.1 Objetivo general**

El objetivo general propuesto en este trabajo es el siguiente:

- 1.2.1.1 Caracterizar los isomiRNA involucrados en el desarrollo de *Blattella germanica* y sus funciones.

### **2.2 Objetivos específicos**

Los objetivos específicos planteados en este trabajo son:

- 1.2.2.1. Hallar y clasificar los isomiRNA presentes en cada etapa del desarrollo de *Blattella germanica*.
- 1.2.2.2. Hallar los isomiRNA significativamente enriquecidos en cada etapa del desarrollo de *Blattella germanica*.
- 1.2.2.3. Hallar los mRNA sobre los que ejercen su función los isomiRNA hallados previamente en cada etapa del desarrollo de *Blattella germanica*.
- 1.2.2.4. Hallar las diferencias de regulación funcionales de familias génicas (mediante “Gene Ontology”) por parte de los isomiRNA en cada una de las etapas del desarrollo.
- 1.2.2.5. Relacionar las diferencias funcionales observadas por parte de los isomiRNA en las etapas del desarrollo con los procesos morfológicos y fisiológicos del desarrollo.

## **1.3. Impacto en sostenibilidad, ético-social y de diversidad**

Este trabajo respecta, comparte y aplica las “Competencias de Compromiso Ético y Global” (CCEG) en sus tres grandes dimensiones: “sostenibilidad”, “comportamiento ético y responsabilidad social (RS)” y “diversidad (género, entre otros) y derechos humanos”. Exactamente, este trabajo se alinea principalmente con el tercer “Objetivo de Desarrollo Sostenible 2030” (ODS) de la ONU “Good Health and Well-Being”. Asimismo, también se aplican los siguientes ODS: sexto ODS (“Clean water and sanitation”), octavo ODS (“Decent work and economic growth”), noveno ODS (“Industry, innovation and infrastructure”), decimo ODS (“Reduce inequality within and among countries”), duodecimo ODS (“Responsible consumption and production”), decimocuarto ODS (“Life below water”), y el decimoquinto ODS (“Life on land”).

Concretamente, este trabajo persigue aumentar el conocimiento sobre la regulación génica, lo cual podría ayudar a desarrollar nuevas técnicas diagnósticas y terapéuticas en un futuro (ODS 3). De esta forma, el aumento del conocimiento científico aportado por este trabajo podría dar lugar a un desarrollo económico al aplicarse este conocimiento a la sociedad (ODS 8 y 9), y, al realizarse este trabajo en parte en instituciones públicas, la accesibilidad del conocimiento y su aplicabilidad derivada sería más accesible al conjunto de sociedades (ODS 10). Además, la aplicación de herramientas bioinformáticas permite el desarrollo de experimentos “*in silico*” que reducen y afinan el uso de experimentos “*in vitro*” e “*in vivo*”, los cuales conllevan un mayor gasto de recursos energéticos, agua, químicos, etc. De esta manera, con el plan de trabajo propuesto se reduce el malgasto de recursos y los contaminantes derivados de otros posibles planes de trabajo (ODS 6, 12, 14 y 15).

#### **1.4. Enfoque y método seguido**

Como se ha detallado previamente, el objetivo principal de este trabajo es obtener los datos de identificación y caracterización funcional de los isomiRNA entre las distintas etapas del insecto hemimetábolο *Blattella germanica*. Para ello se disponen de los datos transcriptómicos brutos de secuenciación masiva (obtenidos mediante la tecnología “Illumina NextSeq 500 platform”) previamente publicados (7) (GEO Database: GSE87031) y correspondientes a datos por duplicado de 11 etapas distintas del desarrollo de esta especie: huevos no fertilizados, 5 estadios de la etapa embrionaria, 4 estadios de la etapa de ninfa y un adulto hembra.

Así, la primera tarea de este trabajo fue evaluar las herramientas bioinformáticas disponibles actualmente y descritas en la bibliografía (8-11) para identificar los isomiRNA y los miRNA en los datos de transcriptómica mencionados previamente. De esta manera, tras la evaluación de varias herramientas bioinformáticas para la identificación de isomiRNA (12-15), se concluyó el uso de la herramienta miR&moRe2 (16) en Linux para llevar a cabo esta tarea (objetivo específico 1) por ser la más adecuada para el trabajo propuesto. Para ello, se obtuvo el genoma de *Blattella germanica* (“Bger\_2.0 Jan. 2018 Genome Browser - GCA\_000762945.2 assembly”) de la base de datos de UCSC (17) y la base de datos de miRNA de *Blattella germanica* de la base de datos mirGeneDB (18). Así, una vez identificados y categorizados los datos de los isomiRNA y los miRNA, así como su abundancia en cada una de las etapas del desarrollo de *Blattella germanica*, se realizó en análisis gráfico y de enriquecimiento en las etapas citadas mediante el paquete de R DESeq2 (19) en RStudio para los isomiRNA y los miRNA (objetivo específico 2).

Una vez obtenidos los isomiRNA y los miRNA enriquecidos significativamente entre al menos dos de las etapas del desarrollo de *Blattella germanica*, se procedió a evaluar las herramientas disponibles y descritas en la bibliografía (20) para predecir las posibles dianas en los mRNA de esos isomiRNA y miRNA. Exactamente, las secuencias de los mRNA de *Blattella germanica* se obtuvieron de la bibliografía existente (21) (ENA: Project PRJNA248247). Así, tras la evaluación de algunas de las distintas herramientas bioinformáticas disponibles (22-25), con distintos métodos para la predicción de las dianas en los mRNA, se concluyó por el uso de la herramienta scanMiR (26), un paquete de Bioconductor para R que permite realizar la predicción de los targets de los miRNA e isomiRNA en mRNA mediante “seed pairing”. De esta forma se obtuvieron las predicciones de las dianas de los isomiRNA y los miRNA en los mRNA.

Entonces, mediante un script desarrollado en R, se resumió la regulación génica que ejercen los distintos isomiRNA y miRNA sobre los distintos mRNA según las predicciones obtenidas previamente de las dianas de cada isomiRNA y cada miRNA. De esta forma, se obtuvieron los datos de regulación génica totales para cada mRNA por parte de los distintos isomiRNA y miRNA (separados y juntos) de *Blattella germanica*. Además, también mediante un script desarrollado en R, se cruzaron los datos de cuentas de los isomiRNA y los miRNA en cada etapa del desarrollo de *Blattella germanica* con los datos previo de “targets” para cada mRNA de *Blattella germanica* para obtener las “cuentas de inhibición” por parte del total de los isomiRNA, los miRNA y ambos sobre cada mRNA en cada etapa del desarrollo (objetivo específico 3).

A continuación, al no existir o no encontrar en la bibliografía existente los datos funcionales de los mRNA de *Blattella germanica*, se procedió a obtenerlos mediante la herramienta BLAST2GO en OmicsBox (27). Así, mediante esta herramienta, se anotaron las proteínas a las que con mayor probabilidad corresponderían algunos de los mRNA estudiados por homología con la base de datos swiss-prot de Uniprot (28) mediante blast. Una vez obtenidas las anotaciones de las proteínas posiblemente correspondientes a cada mRNA, también mediante BLAST2GO en OmicsBox, se anotaron los términos funcionales de “Gene Ontology” (29) (GO) de las proteínas anotadas, y, por tanto, de algunos de los mRNA estudiados.

Finalmente, mediante los datos obtenidos previamente de la regulación génica de los mRNA por parte de los isomiRNA y los miRNA (y la suma de ellos) en las distintas etapas del desarrollo de *Blattella germanica* (“cuentas de inhibición”), se realizó un análisis gráfico y de enriquecimiento de los mRNA a lo largo de las etapas del desarrollo de *Blattella germanica* utilizando las “cuentas de inhibición” de cada mRNA por parte de los isomiRNA, los miRNA y la suma de ellos. Además, también se realizó un análisis de clusteres funcionales (mediante los términos GO) con los datos citados previamente y la anotación funcional (GO ID) de los mRNA de este organismo obtenida previamente mediante el paquete de R clusterProfiler (30). De esta forma, como resultado final de este trabajo, se obtuvieron los clusteres funcionales significativamente distintos sobre los que actúa la regulación génica por parte de los isomiRNA, los miRNA y la suma de ellos a lo largo de las distintas etapas del desarrollo de *Blattella germanica*.

## 1.5. Planificación del trabajo

Para llevar a cabo el trabajo descrito previamente, se propusieron las tareas e hitos relacionados con los objetivos específicos del trabajo mostrados a continuación. Además, en el Anexo I, se muestra el Diagrama de Gantt de la distribución temporal de este plan de trabajo.

- A. **Objetivo específico 1:** hallar y clasificar los isomiRNA presentes en cada etapa del desarrollo de *Blattella germanica*.
  - i. **Hito 1:** elección de la herramienta bioinformática para realizar el análisis de isomiRNA.
    - a. **Tarea 1:** Evaluar las herramientas bioinformáticas disponibles para el análisis de isomiRNA.
- B. **Objetivo específico 2:** hallar los isomiRNA significativamente enriquecidos en cada etapa del desarrollo de *Blattella germanica*.
  - i. **Hito 2:** identificar y anotar las diferencias de los isomiRNA entre las etapas de desarrollo de insectos.
    - a. **Tarea 2:** Realizar el análisis bioinformático de isomiRNA de los datos de las distintas etapas de desarrollo de insectos.
    - b. **Tarea 3:** Observar y analizar las diferencias existentes en los isomiRNA entre las distintas etapas del desarrollo.

- C. **Objetivo específico 3:** hallar los mRNA sobre los que ejercen su función los isomiRNA hallados previamente en cada etapa del desarrollo de *Blattella germanica*.
- i. **Hito 3:** elección de la herramienta bioinformática para predecir la interacción entre los isomiRNA y los mRNA.
    - a. **Tarea 4:** Evaluar las herramientas bioinformáticas disponibles para predecir la interacción entre los isomiRNA y los mRNA.
- D. **Objetivo específico 4:** hallar las diferencias de regulación funcionales de familias génicas (“Gene Ontology”) por parte de los isomiRNA en cada una de las etapas del desarrollo de *Blattella germanica*.
- i. **Hito 4:** identificar y esclarecer las diferencias funcionales de los isomiRNA entre las etapas de desarrollo.
    - a. **Tarea 5:** Realizar el análisis bioinformático de la interacción de los isomiRNA y los mRNA para cada etapa del desarrollo.
    - b. **Tarea 6:** Identificar los efectos de la regulación génica de los isomiRNA en cada etapa del desarrollo.
- E. **Objetivo específico 5:** relacionar las diferencias funcionales observadas por parte de los isomiRNA en las etapas del desarrollo con los procesos morfológicos y fisiológicos de las etapas del desarrollo *Blattella germanica*.
- i. **Hito 5:** asociar las consecuencias funcionales de los isomiRNA con los procesos de cada etapa del desarrollo.
    - a. **Tarea 7:** Extraer conclusiones sobre la relación entre las consecuencias funcionales de los isomiRNA y los procesos asociados a cada etapa del desarrollo.

## 1.6. Breve sumario de productos obtenidos

Al final de este trabajo se obtendrán como producto los datos sobre las implicaciones funcionales de los isomiRNA, lo cual incluye tanto su papel en el desarrollo de *Blattella germanica* como las implicaciones biológicas que tienen los datos sobre las diferencias existentes entre los miRNA y los isomiRNA en la regulación génica. Además, el pipeline usado en este trabajo para comenzar con datos de transcriptómica hasta su efecto sobre la expresión génica puede volverse a utilizarse para otros datos de transcriptómica, incluyéndose los códigos desarrollados en: <https://github.com/jesusalarcongil/TFM.git>.

## 1.7. Breve descripción de los capítulos de la memoria

**Contexto:** se revisa la bibliografía existente acerca de los miRNA, los isomiRNA (incluyendo su biogénesis y mecanismo de acción), y el desarrollo de *Blattella germanica*.

**Métodos:** se describe en detalle el enfoque y método anteriormente descrito en el apartado 1.4.

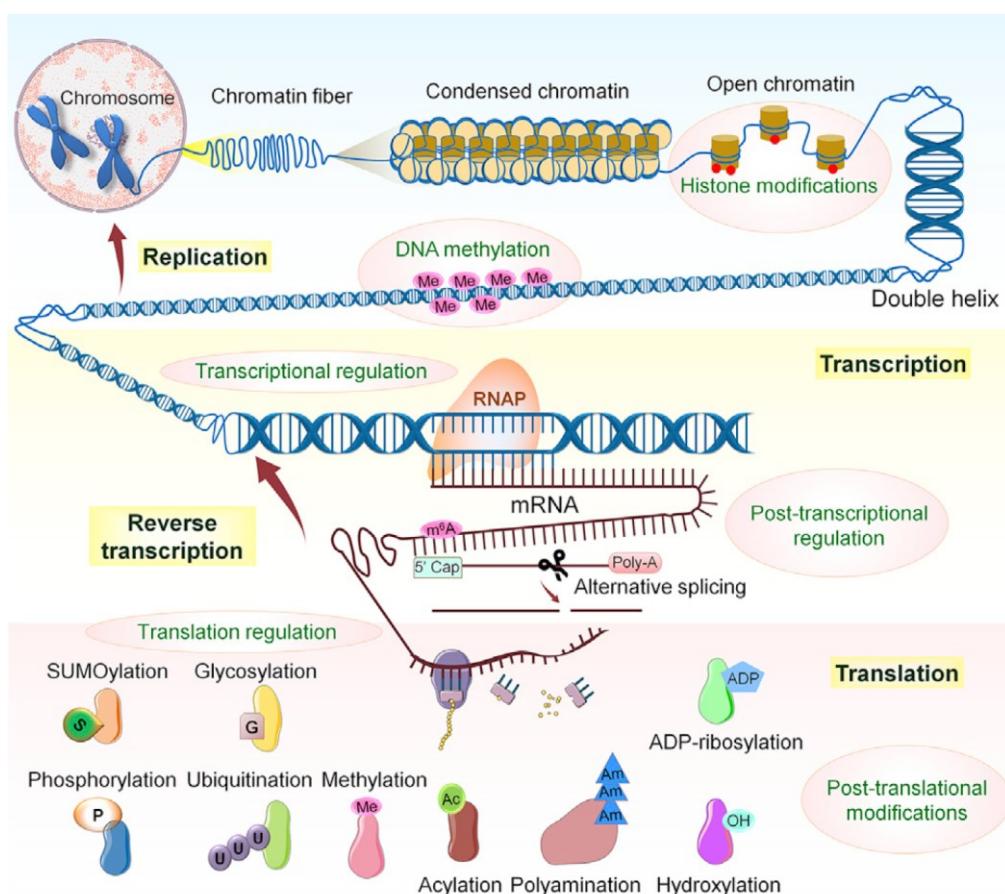
**Resultados:** se presentan los resultados obtenidos tras la consecución de los hitos propuestos.

**Conclusiones:** se resumen y discuten los hitos científicos logrados en este trabajo.

## 2. Estado del arte

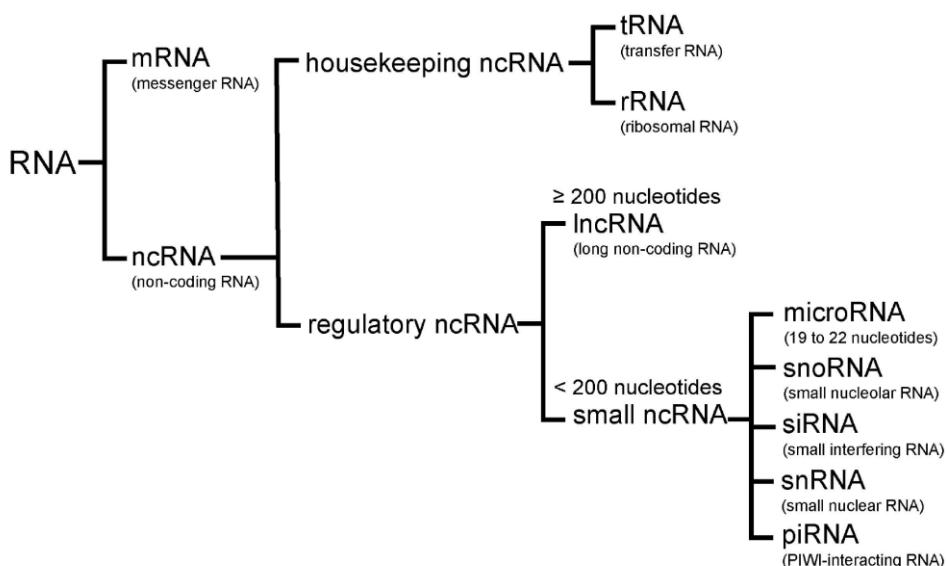
Todos los aspectos relacionados con el correcto desarrollo, funcionamiento y adaptación de un ser vivo en uno o varios entornos están relacionados con la correcta expresión de su material genético. Esta expresión del material genético se basa en el dogma central de la biología, por el cual se postula que la información contenida en el material genético en forma de ácido desoxirribonucleico (DNA) se transcribe en forma de ácido ribonucleico (ARN) para finalmente traducirse en una proteína, la cual es la molécula efectora de la información contenida en el DNA. De esta forma, la completa información disponible en el DNA de los organismos les permite desarrollarse, funcionar y adaptarse a distintos entornos mediante la correcta expresión de las distintas partes adecuadas del DNA en cada momento en forma de proteínas. Sin embargo, para lograr esta adecuada expresión del DNA de un organismo según sus circunstancias temporales y espaciales es necesario una exquisita regulación de su expresión, es decir, la expresión génica (1, 2, 31, 32).

La regulación de la expresión génica es un proceso sumamente complejo que incluye una enorme cantidad de procesos y variables, los cuales, a pesar de la cantidad de investigaciones realizadas, todavía no son del todo conocidos, cambiándose y actualizándose sus conocimientos cada día (2, 31-33).



**Figura 1. Representación en detalle del dogma central de la Biología.** Este esquema muestra todos los distintos procesos del dogma central de la Biología (subrayados de amarillo: replicación, transcripción, traducción y transcripción inversa), los elementos que participan del mismo y los procesos de regulación génica (en verde: modificación de las histonas, metilación del DNA, regulación transcripcional, regulación post-transcripcional, regulación de la traducción y modificaciones post-traduccionales) (33).

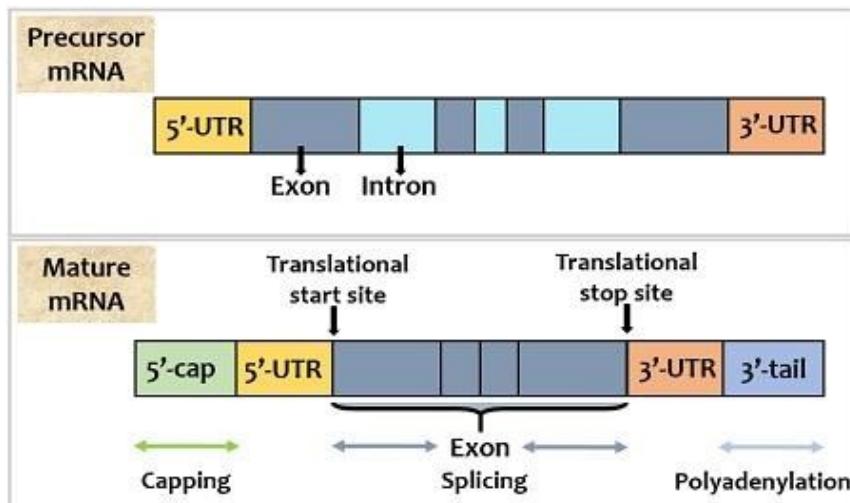
Resumiendo los procesos del dogma central de la Biología a nivel molecular en los organismos eucariotas (Figura 1), el DNA se encuentran en el núcleo de las células de forma condensada (heterocromatina) o accesible (eucromatina), siendo esta última la más permisiva para la transcripción de los genes. Asimismo, pueden tener lugar modificaciones sobre las histonas (moléculas que permiten la organización del DNA) y el propio DNA que provocan la regulación de la transcripción, conociéndose como epigenética. De esta forma, la transcripción de los genes se realiza mediante la RNA polimerasa, la cual lee una de las dos hebras del DNA en la dirección 3' a 5' de esta molécula y genera una molécula de RNA monocadena en dirección 5' a 3' complementaria de ese DNA. Sin embargo, este proceso de transcripción puede regularse por distintos mecanismos como factores de transcripción, miRNA o niveles de metabolitos (2, 32-34).



**Figura 2. Clasificación de los distintos tipos de RNA.** Este esquema muestra la clasificación de los RNA según sean: codificantes de una proteína (mRNA) o no (ncRNA), y se expresen de manera constitutiva (tRNA y rRNA) o regulen la expresión de los RNA codificantes (lncRNA y sncRNA) (35).

La molécula de RNA obtenida tras el proceso de transcripción puede tratarse de distintos tipos de RNA funcionales, las cuales se clasifican en un primer momento como RNA codificantes para proteínas (mRNA) o RNA no codificantes. A su vez, estos RNA no codificantes se clasifican en RNA constitutivos (como los RNA ribosomales (rRNA) que forman los ribosomas o los RNA de transferencia (tRNA) que adaptan los aminoácidos a la secuencia del mRNA) o RNA reguladores, clasificándose éstos últimos en largos o cortos (como los miRNA) según su longitud (31, 35) (Figura 2).

De todos estos tipos de RNA, cabe resaltar la estructura de los mRNA por su importancia (Figura 3). El mRNA se transcribe desde el DNA como precursor del mRNA, el cual se caracteriza por presentar dos regiones que no se traducen a ambos lados de la cadena (5'-UTR y 3'-UTR) y regiones intrónicas (no codificantes para proteínas) y exónicas (si codificantes) en el interior. De esta forma, el precursor del mRNA se procesa para dar lugar al mRNA maduro eliminando las secuencias intrónicas ("splicing") y añadiéndole dos estructuras protectoras a ambos lados del mRNA (5'-cap y 3'-poliA) (34).



**Figura 3. Estructura y procesamiento del mRNA.** En esta figura se muestra la estructura del precursor del mRNA, el resultado de la transcripción, y la estructura resultante del procesamiento del precursor del mRNA en el mRNA maduro (34).

En este proceso de maduración del mRNA se pueden dar procesos de regulación génica como puede ser el “splicing” alternativo (distinta eliminación y combinación de los intrones y los exones). Además, una vez obtenido el mRNA maduro, también se dan procesos de regulación génica como la regulación post-transcripcional por miRNA, la cual se explicará posteriormente (2, 34).

Posteriormente, el mRNA se traduce en los ribosomas (unos orgánulos formados por rRNA y proteínas ribosomales) gracias, entre otros participantes, a los tRNA, los cuales permiten introducir el aminoácido correcto en la nueva proteína codificada por el mRNA según los tripletes de nucleótidos (codones). Sin embargo, en este proceso también se dan procesos de regulación génica (regulación traduccional), como el bloqueo de la traducción por modificación del RNA o modificaciones del ribosoma que modifican la traducción. Además, una vez traducida la proteína codificada en el mRNA, también se pueden dar modificaciones de la información transmitida mediante modificaciones post-traduccionales, las cuales se basan en modificaciones de la proteína resultante que pueden provocar modificaciones funcionales de la proteína (31, 33).

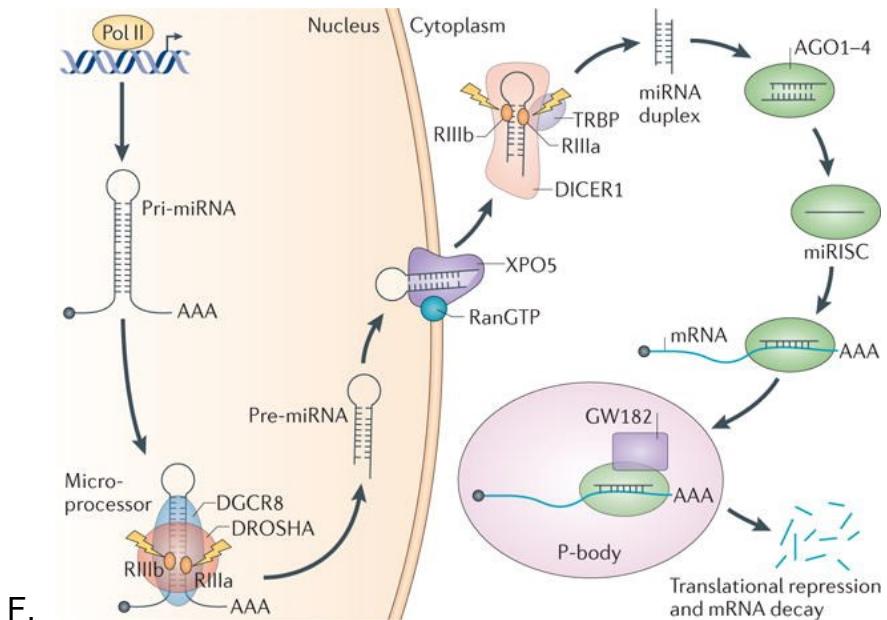
Así, una vez explicados los participantes, los procesos y la regulación del dogma central de la Biología, es necesario explicar en mayor detalle algunas de las partes de este, como los miRNA.

## 2.1 miRNA

Como se ha mencionado previamente, los miRNA son moléculas de RNA de una cadena no codificantes con un tamaño medio de 22 nucleótidos cuya función es regular la expresión génica.

En cuanto a la biogénesis canónica de estas moléculas (Figura 4), los miRNA se transcriben a partir de sus genes propios o, en ocasiones, de los intrones de genes codificantes. Así, la transcripción de éstos por parte de la RNA polimerasa II (y en algunos casos la III) da lugar a los pri-miRNA, los cuales presentan una estructura en forma de cerradura con dos colas y que, como los mRNA, presentan una 5'-cap y algunas veces una cola 3'-poliA. Estos pri-miRNA son procesados por el complejo microprocesador, el cual está conformado por la RNasa III DROSHA y la proteína de unión a RNA DGCR8.

Como resultado de este procesamiento se recortan las dos colas del pri-miRNA y se forma el pre-miRNA, el cual tiene estructura de “hairpin” y se exporta al citoplasma desde el núcleo mediante la exportina 5 (3, 4, 10, 36).



**Figura 4. Mecanismo de biogénesis canónica de los miRNA.** Este esquema representa el proceso de biogénesis de los miRNA desde su transcripción hasta su función (36).

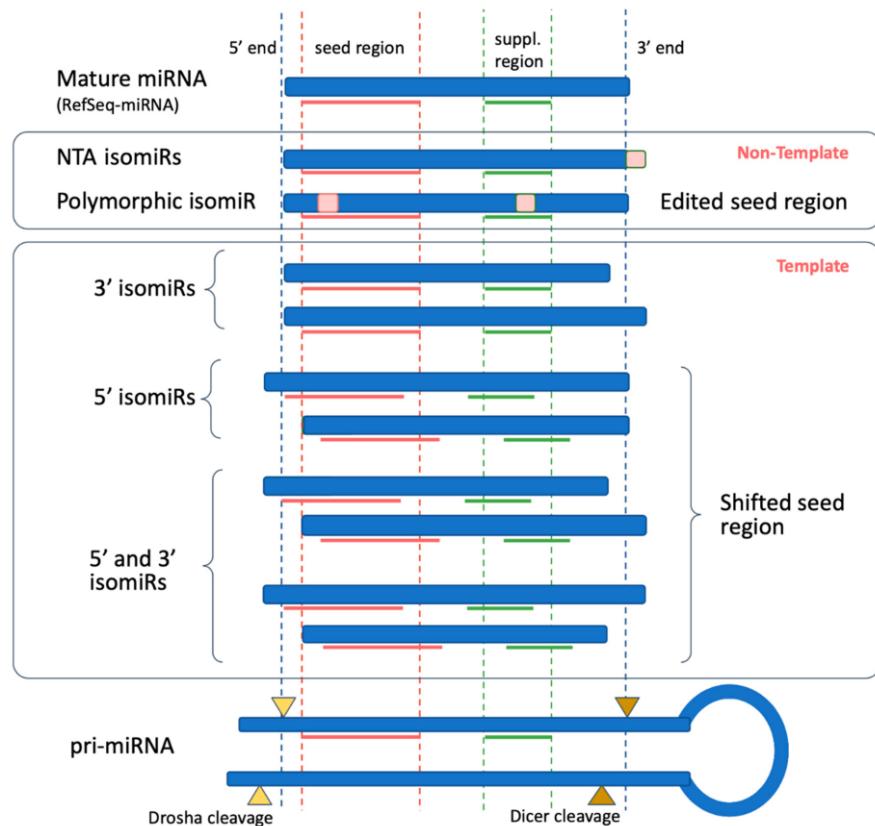
Una vez en el citoplasma, el pre-miRNA es procesado por la RNAsa III DICER y la proteína de unión a RNA TRPB, de modo que eliminan el loop (extremo unido) del pre-miRNA y da lugar a un dúplex de miRNA, el cual se integra con una de las proteínas Argonauta para pasar a formar parte del complejo silenciador inducido por miRNA miRISC. Entonces se desecha una de las hebras del dúplex de miRNA y se mantiene la otra, la cual será la responsable, de la regulación negativa de la expresión de un gen mediante la degradación de su mRNA, ya que por complementariedad de los nucleótidos 2 al 7 de su extremo 5' (“seed”) se une a la región 3'-UTR de ese mRNA (3, 4, 10, 36).

De esta manera, la regulación post-transcripcional por parte de los miRNA depende enormemente de la secuencia “seed” de estos, aunque no es el único factor que interviene en la identificación del mRNA. Además, de la complementariedad de la región “seed”, se han observado indicios de que, aunque la complementariedad de ésta con el mRNA sea muy alta, si los nucleótidos del centro no coinciden, se produce una degradación directa del miRNA. Por otro lado, también existen indicios acerca de la existencia de otro tipo de interacciones más complejas entre la secuencia completa del miRNA y el mRNA, desde la complementariedad de la región 3' del miRNA hasta la regulación de la transcripción (4).

Por todo ello, pequeños cambios en las secuencias del miRNA pueden implicar importantes diferencias funcionales en cuanto a los mRNA sobre los que actúa. Estas pequeñas diferencias son las que aparecen en los llamados isomiRNA, variaciones en longitud y/o nucleótidos de los miRNA que, al principio de su estudio, se creyeron que eran fallos técnicos en su detección, pero que posteriormente se demostraron su existencia y efecto. Por ello, los isomiRNA pueden provocar un gran cambio en la regulación génica, por lo que es fundamental estudiar el efecto de los isomiRNA sobre la misma (3).

## 2.2 isomiRNA

De esta forma, los isomiRNA son miRNA con diferencias en secuencia y/o longitud respecto a los miRNA canónicos. Según sean esas diferencias (Figura 5), los isomiRNA se clasifican respecto al miRNA de referencia en (3, 10, 37):



**Figura 5. Clasificación de los isomiRNA.** Los isomiRNA se pueden clasificar en distintos tipos según los cambios que sufren respecto a su miRNA canónico (10).

- IsomiRNA polimórficos, en los que se cambia al menos un nucleótido de la secuencia, lo cual puede provocar cambios en la secuencia “seed”. El origen de este tipo de isomiRNA se puede deber a polimorfismos de un solo nucleótido, aunque éstos son bastante menos comunes que en el resto del genoma. Además, este tipo de isomiRNA también pueden surgir por la modificación enzimática de nucleótidos, causando el cambio de secuencia.
- IsomiRNA de longitud, en los que cambia la longitud de estos, pudiendo ser por el extremo 3' (no cambia la secuencia “seed”, pero aumenta la estabilidad de la unión con el mRNA), por el extremo 5' (cambia la secuencia “seed”), o por ambos extremos. Los isomiRNA en los que cambia el extremo 3' suelen ser bastante más frecuentes que en los que cambia el 5'. En cuanto a su origen, podrían deberse a la acción de exoribonucleasas o a cortes imprecisos de las RNAsas DROSHA y/o DICER.
- IsomiRNA de tipo “no-template”, los cuales son también un cambio en la longitud del isomiRNA respecto al miRNA de referencia, pero que no coincide con los nucleótidos del pri-miRNA, siendo debido a la adición de nucleótidos posterior por nucleotidiltransferasas. Mayoritariamente, se suelen agregar adeninas (aumentan la estabilidad del microRNA) y uridinas (favorecen la degradación del microRNA).

- IsomiRNA de tipo mixto, los cuales combinan las características de los anteriores tipos (3, 31, 37).

A pesar de todos los tipos de isomiRNA y mecanismos para su generación observados en la bibliografía, todavía no se conocen bien los mecanismos de cómo surge toda esta diversidad de isomiRNA y sus implicaciones. Por ello, es de gran importancia continuar investigando este campo tanto para el conocimiento de la regulación génica en sí como para su posible aplicabilidad a la sociedad, ya que ya se ha comprobado la implicación de los isomiRs en procesos como el desarrollo o las diferencias entre tejidos y en patologías como el cáncer o las enfermedades neurodegenerativas (3, 10, 31).

### **2.3 *Blattella germanica***

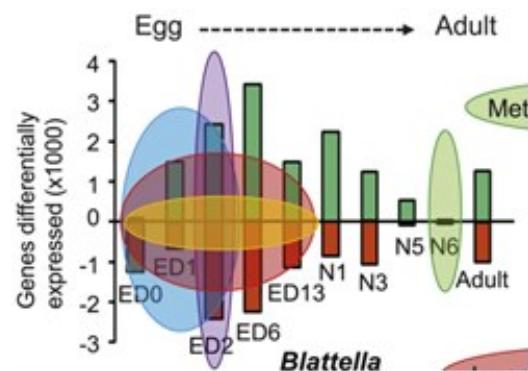
Como se ha indicado previamente, es sumamente necesario estudiar los efectos de los isomiRNA y, para ello, es necesario un modelo de laboratorio evolutivamente conservado, lo más fácil de manejar posible, con la posibilidad de obtener grandes muestras sin problemas éticos y con interés biomédico. Entre uno de esos modelos, se encuentra el organismo *Blattella germanica*.

*Blattella germanica*, o cucaracha rubia, es un organismo eucariota perteneciente al reino *Animalia*, filo *Arthropoda*, clase *Insecta* y del orden *Blattodea*. Este orden incluye al conjunto de las cucarachas, unos organismos que han mantenido su morfología durante cientos de millones de años, siendo uno de los insectos con alas más primitivos. La importancia de estos organismos reside en las plagas domésticas que provocan algunas de las especies incluidas en este orden, incluida *Blattella germanica*. Las plagas de esta especie son de gran relevancia, ya que distribuyen mecánicamente diversos y numerosos agentes patológicos (incluyendo bacterias con resistencia a antibióticos) y se han relacionado con un aumento en la incidencia de alergias respiratorias y de asma. Además, esta especie se ha adaptado al humano hasta depender en parte de él para su supervivencia, por lo que coloniza un gran número de ambientes humanos. Incluso, a pesar del aumento de la higiene y del mantenimiento de nuestros ambientes durante los últimos tiempos, estas plagas no solo no se han reducido, sino que han aumentado. Por ello, es sumamente necesario investigar cómo reducir la propagación de las plagas de *Blattella germanica* en los ambientes humanos para reducir la exposición a agentes patológicos y la incidencia de alteraciones respiratorias (21, 38, 39).

Además, el estudio de la biología de la cucaracha *Blattella germanica* no solo resulta interesante por la posibilidad de acabar con las plagas de este organismo, sino que, al tratarse de uno de los organismos más conservados evolutivamente con una metamorfosis hemimetábola (metamorfosis gradual sin fase de pupa, a diferencia de la metamorfosis holometábola y, por tanto, aunque alejado, más cercano evolutivamente al humano), es un modelo biológico muy interesante para el estudio de procesos con una gran regulación génica conservados evolutivamente, en concreto, su desarrollo. Por ello, y por otras razones de su biología, se explica que este organismo se utilice ampliamente como modelo biológico de laboratorio para el estudio de la reproducción y el desarrollo (incluyendo el proceso de metamorfosis), la etología, el metabolismo, la neurofisiología, la alergología y el desarrollo de nuevos mecanismos insecticidas y sus resistencias (21, 40).

En concreto, este organismo permite estudiar los procesos de exquisita regulación génica (incluyendo los miRNA y los isomiRNA) que ocurren durante su desarrollo, incluyendo su metamorfosis hemimetábola. Así, durante el desarrollo de *Blattella germanica*, el organismo pasa por las etapas de huevo no fecundado, huevo fecundado, etapas embrionarias (en las que ya se desarrolla la morfología básica del adulto), etapas de ninfa y adulto. Exactamente, durante estas etapas ocurren distintos procesos fundamentales en el completo desarrollo del organismo adulto (7, 41-43) (Figura 6).

- Etapas embrionarias (según el artículo publicado por Tanaka en 1976 (41), resumido por Bellas *et al.* en 2017 (7) y detallado celular y molecularmente por Belles *et al.* en 2018 y 2022 (42, 43)):
- Etapa embrionaria ED0 (8 horas después de la oviposición): solo se observan “yolk granules”. Entre esta y la siguiente etapa descrita se produce la primera oleada de expresión de genes relacionados con la diferenciación de los segmentos torácicos y otras zonas del embrión (incluyendo el sistema nervioso y las alas). Exactamente, esta oleada parece regular el comienzo de la diferenciación de distintas zonas.
- Etapa embrionaria ED1 (24 horas después de la oviposición): se observan “energids” distribuidos entre los “yolk granules” y corresponde con la primera etapa de Tanaka. Durante esta etapa aún tiene lugar la oleada de expresión que ocurre en la anterior etapa.
- Etapa embrionaria ED2 (48 horas después de la oviposición): se observan muchas más “energids”, además de empezar a esbozarse las bandas germinales bien definidas, y corresponderse con la segunda etapa de Tanaka. Hasta esta etapa también se expresan más genes asociados a la transición maternal-cigoto, y, en concreto en esta etapa, a la regulación de la metilación de DNA.
- Etapa embrionaria ED6 (144 horas después de la oviposición): se observan claramente pleuropodios, además de surgir el espacio caudal, comenzase a observar los ojos y la segmentación de las patas, pudiéndose incluso a percibir el latido del corazón. En este caso, corresponde con la octava etapa de Tanaka.
- Etapa embrionaria ED13 (312 horas después de la oviposición): se observan ya las antenas, los ojos coloreados y las patas más desarrolladas, correspondiéndose con la quinceava etapa de Tanaka. En esta etapa se produce la segunda, aunque más leve, oleada de



**Figura 6. Distribución de la regulación génica a lo largo del desarrollo de *Blattella germanica*.** En el gráfico se muestran los resultados obtenidos por Ylla *et al.* en 2018 (43) sobre la distribución de la regulación de distintos sets de genes a lo largo del desarrollo de *Blattella germanica* (43)

expresión de genes relacionadas con la diferenciación de los segmentos torácicos y otras zonas del embrión (incluyendo el sistema nervioso y las alas). Esta segunda oleada se cree que ya comienza el desarrollo de las zonas diferenciadas en la primera oleada, como el desarrollo de las alas. Así, alrededor de esta etapa comienza la organogénesis.

- Etapas de ninfa (según el artículo publicado por *Bellas et al.* en 2017): estas etapas incluyen los distintos estadios ninfales que tienen lugar después de la eclosión, pudiéndose resumir en: N1 (primer estadio ninfal, de 1 a 2 días), N3 (tercer estadio ninfal, de 2 a 4 días), N5 (quinto estadio ninfal, de 3 a 5 días) y N6 (sexto estadio ninfal, de 5 a 7 días). En todos estos estadios, de forma general, los procesos morfológicos y fisiológicos que se dan son bastante estables y parecidos entre todos los estadios. Exactamente, en ellas se da un aumento de la diferenciación de las estructuras ya definidas durante las etapas del embrión y un crecimiento de las mismas, como, por ejemplo: coloración, crecimiento del cuerpo, alas, antenas, patas, etc (7, 44, 45).

De esta manera, se ha comprobado que es en los estadios embrionarios del desarrollo de *Blattella germanica*, y más en concreto en los primeros de ellos, donde más procesos de regulación génica tienen lugar, en concreto, de regulación génica mediada por miRNA (42, 43).

## 2.4 Justificación y objetivos del trabajo

Con toda la información aportada previamente, se pueden deducir varios objetivos existentes y urgentes para la investigación contemporánea. Algunos de ellos son: ampliar y comprobar el conocimiento existente sobre la regulación génica (incluyendo los miRNA y los isomiRNA) y de ésta sobre procesos biológicos como el desarrollo; o conocer más ampliamente la Biología de *Blattella germanica*, incluyendo su desarrollo, por las implicaciones que tiene en la evolución y en el posible descubrimiento de formas de controlar las plagas que produce por los riesgos para la salud humana de esas plagas.

Así, teniendo en cuenta estos objetivos científicos, una posible estrategia de investigación que los incluiría sería mejorar el entendimiento de la regulación génica del desarrollo de *Blattella germanica* para aumentar el conocimiento de la regulación génica en general, de la evolución de *Blattella germanica* y la conservación de la regulación génica en la evolución, y encontrar posibles dianas biológicas para el control de las plagas de esa especie.

De esta forma, siguiendo la estrategia de investigación anteriormente propuesta, el objetivo principal de este trabajo es “caracterizar los isomiRNA involucrados en el desarrollo de *Blattella germanica* y sus funciones”. Además, para cumplir ese objetivo principal, se proponen los siguientes objetivos específicos en este trabajo: “Hallar y clasificar los isomiRNA presentes en cada una de las etapas del desarrollo de *Blattella germanica*”, “Hallar los isomiRNA significativamente enriquecidos en cada una de las etapas del desarrollo”, “Hallar los mRNA sobre los que ejercen su función los isomiRNA hallados previamente en cada una de las etapas del desarrollo”, “Hallar las diferencias de regulación funcional de familias génicas (mediante “Gene Ontology”) por parte de los isomiRNA en cada una de las etapas del desarrollo”, y “Relacionar las diferencias funcionales observadas por parte de los isomiRNA en las etapas del desarrollo con los procesos del desarrollo de *Blattella germanica*”.

### 3. Materiales y métodos

#### 3.1. Entornos y datos de entrada.

Para llevar a cabo este trabajo, se utilizaron dos entornos distintos. En primer lugar, un entorno con un sistema operativo Windows 11 de 64 bits montado en un equipo con una CPU y procesador “Intel<sup>(R)</sup> Core<sup>(TM)</sup> i7-1165G7” de onceava generación a 2.80 GHz, con 4 núcleos y 12 procesadores, y con 16 Gb de memoria RAM. En segundo lugar, un entorno con un sistema operativo Ubuntu-MATE de 64 bits montado mediante una imagen en el software Oracle VM VirtualBox instalado en el anterior sistema operativo, destinándole 6 procesadores y 11 Gb de memoria RAM del primer equipo.

De esta forma, en el entorno Ubuntu-Mate se trabaja directamente sobre el terminal para ejecutar las distintas herramientas; mientras que en el entorno Windows se ha trabajado con RStudio para el lenguaje R y con el software “Integrative Genomics Viewer” 2.16.1 (46) para la visualización de los datos.

Por otro lado, para llevar a cabo este trabajo, se han necesitado una serie de datos originales:

- **Datos de secuenciación transcriptómica de RNA no codificante** en las distintas etapas del desarrollo de *Blattella germanica* obtenidos por tecnología “Illumina NextSeq 500 platform”. Estos se obtuvieron gracias a Guillem Ylla en formato “.bam”, y también están disponibles en el repositorio “Gene Expression Omnibus” (GEO) del “National Center for Biotechnology Information” (NCBI) mediante el número de acceso GSE87031. Exactamente, se disponen de 22 archivos “.bam” correspondientes a los duplicados de las etapas del desarrollo de *Blattella germanica*: ED0, ED1, ED2, ED6, ED13, N1, N3, N5, N6, Adulto y huevo no fecundado (NFE).
- **El genoma de *Blattella germanica*** en formato fasta. Estos datos se obtienen de la base de datos del “University of California Santa Cruz (UCSC) Genome Browser” (47), exactamente del genoma: “Bger\_2.0 Jan. 2018 Genome Browser - GCA\_000762945.2 assembly”.
- **La base de datos de los miRNA de *Blattella germanica*** en formato “.gff3”. Estos datos se obtienen de la base de datos MirGeneDB (18) en formato “.bed” (<https://mirgenedb.org/static/data/bge/bge-all.bed>), y posteriormente se modifican al formato “.gff3” mediante un código creado en RStudio para este trabajo (Código I). Esto es debido a que la herramienta usada en este trabajo para identificar los isomiRNA necesita la base de datos de miRNA con la información que da la base de datos miRBase (48), pero los datos de *Blattella germanica* no existen en esta última base de datos. Por ello, se obtiene la información de la base de datos de MirGeneDB y se formatea mediante el código creado en R (Código I) para obtener la información necesaria según miRBase en formato “.gff3”. Así, el archivo “.gff3” obtenido con este código de R se corresponde con los datos obtenidos de mirGeneDB necesarios en miRBase, a excepción de los caracteres “=” y “%” que no se cambian correctamente y del orden de las entradas. Por ello, se corrigen estos caracteres en el archivo “.gff3” mediante el editor de texto y el orden mediante Microsoft Excel.

- **Las secuencias DNA de los mRNA de *Blattella germanica*** en formato “.fasta”, incluyendo las regiones 3'UTR y 5'UTR de los mRNA. Estas secuencias se obtienen del repositorio “European Nucleotide Archive” (ENA) del “European Molecular Biology Laboratory - European Bioinformatics” Institute (EMBL-EBI) mediante el número de acceso (GBID01000000) (21). En el estudio en el que se generaron estos datos (21), éstos se obtuvieron a partir de huevos maduros, ninfas de cuarto estadío, y adulto machos y hembras, considerándose estos últimos después del séptimo día después de la eclosión del huevo.

Además, indicar que todos los códigos desarrollados y necesarios para la correcta ejecución de este pipeline, a los cuales se hace referencia a continuación, se encuentran en: <https://github.com/jesusalarcongil/TFM.git>.

### **3.2. Detección de iso/miRNA.**

Para llevar a cabo la detección de isomiRNA y miRNA a partir de los datos de secuenciación transcriptómica de RNA no codificante, se intentaron poner a punto distintas herramientas y pipelines para el análisis de isomiRNA y miRNA. Para ello, se llevó a cabo una búsqueda bibliográfica sobre las herramientas y pipelines disponibles actualmente para el análisis de isomiRs. Concretamente, se buscaron artículos científicos y de revisión que describiesen y comparasen el estado actual de este tipo de herramientas, como la revisión de Schmartz *et al.* (8), de Desvignes *et al.* (11) o la publicación que describió la herramienta miRge3.0 (15). Así, se encontró un número considerable de herramientas desarrolladas para el análisis de isomiRNA y miRNA, de las cuales se descartaron aquellas herramientas especialmente desarrolladas para trabajar con datos procedentes de muestras humanas, de muestras de algunas especies concretas, de otras tecnologías diferentes a la usada en los datos de origen, o herramientas con fines específicos (encontrar variantes patológicas). De esta manera, se seleccionaron 5 herramientas para intentar realizar el análisis de isomiRNA y miRNA en el entorno Linux: SeqBuster (12), sRNAbench (13), Jasmine (14), miRge3.0 (15) y miR&moRe2 (16).

Tras intentar ejecutar la herramienta Jasmine en el entorno Linux, pero obtener un error en su ejecución, se intentaron ejecutar los paquetes de Bioconda “SeqBuster” e “isomirs”. Para ello, y para el resto de herramientas probadas posteriormente, hubo que convertir los archivos “.bam” de los datos de secuenciación transcriptómica de RNA no codificante en archivos “.fastq”, ya que este y los siguientes pipelines solo aceptan como input “.fastq”. Para ello se ejecutó la siguiente orden de la herramienta samtools (49) en el entorno Linux: “`samtools bam2fq *.bam > *.fastq`”. Sin embargo, tampoco se consiguieron ejecutar de forma correcta estos paquetes de Bioconda ni las herramientas sRNABench y miRge3.0. La última herramienta que se ejecutó para detectar isomiRNA y miRNA, en este caso eficazmente, fue miR&more2. Sin embargo, es necesario indicar que, durante su instalación, fue necesario sustituir manualmente el paquete ViennaRNA (50), aparentemente por un problema en las versiones del mismo.

La herramienta miR&moRe2 permite analizar datos de secuenciación transcriptómica procedentes de la tecnología Illumina en formato “.fastq” para obtener la anotación y el conteo de los isomiRNA y los miRNA en distintas muestras. Para ello, esta herramienta utiliza: el genoma del organismo en

formato “.fasta”; la anotación de los miRNA en formato “.gff3” procedente de miRBase (se utiliza el archivo obtenido mediante el código de R (Código I) previamente explicado); un archivo “.csv” con el nombre, la ruta y la condición experimental de los archivos de las muestras; y un archivo “.py” con los parámetros necesarios para ejecutar la herramienta. Brevemente, esta herramienta realiza los siguientes procesos: eliminado de los adaptadores de la secuenciación (no realizado en este caso porque ya estaban eliminados en los archivos aportados por Guillem Ylla), análisis de calidad (mediante la herramienta FastQC (51)), filtrado por tamaño y calidad, colapsados, filtrado por el número de cuentas (en este caso no, ya que se reduce el filtro a 1 cuenta), alineado con el genoma de referencia y anotación de miRNA (mediante Bowtie2 v1.1.2 (52)), detección de nuevos posibles precursores y miRNA (mediante MiRDeep2 v0.1.2 (53) y ViennaRNA v2.4.14 (50)), detección de los isomiRNA y miRNA, y resumen y generación de los datos.

De esta forma, tras instalar miR&moRe2 en el entorno Linux, se crea una carpeta para el trabajo y, en el terminal, en la ruta de la carpeta del trabajo, se ejecuta: “bash /home/jalarcongi/TFM/mirandmore2/mirandmore “-j 4””. Este comando indica la ruta de la instalación de mir&more2 y el comando “-j 4” indica el número de tareas a realizar en paralelo. Además, en la carpeta del trabajo debe haber dos archivos: el archivo “.csv” indicado previamente y el archivo “.py” con los parámetros a utilizar (detallado en el Anexo II y los parámetros usados en este trabajo en el Anexo III). Los valores de los parámetros remarcados en el Anexo II se han seleccionado siguiendo las indicaciones mostradas en el artículo de referencia de miR&moRe2 para conseguir la máxima sensibilidad en la ejecución de la herramienta. De este modo, la ejecución de miR&moRe2 con los 22 archivos “.fastq” de las muestras genera distintos resultados: una carpeta “mirandmore\_dbs” con las bases de datos correspondientes a las distintas secuencias utilizadas y generadas en el análisis; una carpeta “samples” con una subcarpeta por cada muestra en la que se generan los distintos archivos intermedios de la herramienta (secuencias filtradas, alineadas, anotaciones, etc.); y una carpeta “results” con el resumen de los resultados de los resultados de todas las muestras individualizados.

### **3.3. Análisis gráfico y de enriquecimiento de iso/miRNA.**

Para llevar a cabo el análisis gráfico y de enriquecimiento de los datos obtenidos en la anotación y conteo de isomiRNA y miRNA en las distintas etapas del desarrollo de *Blattella germanica* se usó el paquete DESeq2 (19) de R en RStudio en Windows. En general, este paquete permite: la importación de los datos de cuentas de las distintas anotaciones en las distintas muestras, el filtrado de éstas, la obtención del análisis de enriquecimiento de éstas, su transformación y su exploración visual.

Así, el código utilizado en R para aplicar el paquete DESeq2 (basado en la documentación de la web de DESeq2) es el mostrado en el Código II para los miRNA y en el Código III para los isomiRNA. Este código, elimina las anotaciones correspondientes a secuencias pri-miRNA, secuencias loop, nuevas predicciones de miRNA, y nuevas predicciones de MicroRNA-offset RNAs (moRNA) de los datos resultantes de miRNA de miR&moRe y los datos correspondientes a secuencias pri-moRNA, secuencias loop, isomiRNA exactos, nuevas predicciones de miRNA, y nuevas predicciones de moRNA de los datos resultantes isomiRNA de miR&moRe2.

Entonces, se incorporan esos datos resultantes de miR&moRe2 modificados como un objeto DESeq2. A continuación, este código filtró las anotaciones de isomiRNA y miRNA por sus contajes para obtener unos datos más precisos y sin ruido experimental. Esto es debido a que, durante la ejecución previa de miR&moRe2, no se aplicó ningún filtro al número de cuentas para aumentar la sensibilidad en la detección. Por ello, este código eliminó las anotaciones con menos de 220 cuentas entre todas las 22 muestras. Se aplicó este filtro total en lugar de aplicar uno de 10 cuentas por muestra por si existieran diferencias muy grandes entre las distintas muestras.

A continuación, ese código transforma el objeto DESeq2 para su representación gráfica, obteniendo las gráficas: heatmap por distancias, heatmap por distancias clusterizado y “principal component analysis” (PCA). Entonces, el código realiza el análisis de enriquecimiento de los datos DESeq2 y se generan unas tablas con los datos de significancia, con las anotaciones enriquecidas entre dos estadios del desarrollo con una significancia menor o igual entre 0.001, y con el número de anotaciones significativamente enriquecidas entre cada uno de los estadios de *Blattella germanica*. Además, se obtiene una lista con las anotaciones (isomiRNA o miRNA) significativamente enriquecidas entre al menos dos de los estadios.

Por otro lado, el análisis llevado a cabo con el paquete DESeq2 de R es más exhaustivo para los isomiRNA (Codigo III) que para los miRNA (Codigo II), ya que se analizan los isomiRNA también según su clasificación (Figura 5). De esta forma, se crea una tabla con la suma de las cuentas de cada categoría de isomiRNA a lo largo de las distintas etapas del desarrollo de *Blattella germanica* y, con esa tabla, se repite el análisis realizado previamente (creación de su objeto DESeq2, transformación para su representación gráfica, representación gráfica, análisis de enriquecimiento y summarización del mismo).

### **3.4. Predicciones de los “targets” de los iso/miRNA en los mRNA.**

A partir de las listas obtenidas previamente de anotaciones de isomiRNA y miRNA significativamente enriquecidas entre algunos de los estadios de *Blattella germanica* y el archivo “.fasta” con las secuencias de los mRNA de *Blattella germanica* (21), se realiza la predicción de las dianas o “targets” de los isomiRNA y miRNA en los distintos mRNA con R en el entorno Windows.

Para ello, se evaluaron varias herramientas de entre las disponibles en la bibliografía para realizar las predicciones de los targets de isomiRNA y miRNA mediante distintos métodos. En esa evaluación, se tuvo en cuenta para elegir la herramienta más adecuada, entre otros motivos, la posibilidad de adaptarlo al genoma de *Blattella germanica* y el método de predicción usado por la herramienta. De esta forma, finalmente se decidió utilizar el paquete de Bioconductor para R scanMiR (26), el cual permite la identificación de targets de isomiRNA y miRNA dadas las secuencias tanto de éstos como de los mRNA a analizar. Además, scanMiR acepta cualquier especie y aplica un modelo de “seed pairing”, es decir, la complementariedad de la región “seed” de los isomiRNA y miRNA en la región 3'-UTR de los mRNA. Así, se ha conseguido ejecutar correctamente la herramienta scanMir de R en el entorno Windows, aunque para ello es necesario modificar previamente una de sus funciones (“.checkSeedsInput”). Esta función posee unas restricciones de tamaño para los miRNA de 18 a 24 nucleótidos y, en este caso, se han detectado miRNA e

isomiRNA mayores. Por ello, la función “.checkSeedsInput” del paquete scanMir se cambia mediante el siguiente código en R para modificar los límites de longitud de los isomiRNA y miRNA de 14 a 32 nucleótidos: “library(scanMir); trace("scanMiR:::checkSeedsInput", edit=TRUE)”.

Además, para ejecutar la herramienta scanMiR en R, es necesario modificar, combinar y cargar: las listas de isomiRNA y miRNA significativamente enriquecidos entre al menos dos de los estadios del desarrollo de *Blattella germanica* obtenidas previamente, las secuencias DNA obtenidas por miR&moRe2 de los isomiRNA (incorporada en sus anotaciones) y de los miRNA (un archivo), y las secuencias DNA de los mRNA (21). Esto se lleva a cabo mediante un código desarrollado en R (Codigo IV y V) en el entorno Windows, mediante el cual primero se filtran las listas de isomiRNA y miRNA para eliminar las secuencias “loop”, las precursoras (“pri”) y los moR. Luego se convierten las secuencias DNA de los miRNA e isomiRNA a ARN mediante su complementariedad (incluyendo el cambio de timina por uracilo) y su cambio de dirección (darle la vuelta), y se combinan las secuencias de los isomiRNA y miRNA con la lista de los isomiRNA y miRNA enriquecidos significativamente. A continuación, se cargan las secuencias DNA de los mRNA en R como un objeto DNAString. Finalmente, se ejecuta la herramienta scanMiR, mediante su función “.findSeedMatches”, para cada una de las anotaciones de los isomiRNA y los miRNA. Para ello, se indican como argumentos para esta función: el objeto DNAString de las secuencias DNA de los mRNA, las secuencias RNA de los isomiRNA y los miRNA, la detección solamente de las posibles dianas canónicas en los mRNA y el formato en el que se obtienen los resultados (“dataframe” en este caso). Así, mediante el código desarrollado en R (Codigo IV y V) se almacenan los resultados de las predicciones en dos listas de “dataframes”, una para los isomiRNA y otra para los miRNA. Además, esta función permite cambiar ciertos parámetros del algoritmo que realiza la predicción de los “targets”, aunque, en este caso, se han dejado por defecto.

De esta manera, en cada uno de los “dataframes” (uno por cada anotación) de las listas de isomiRNA y miRNA obtenidas se detalla: el mRNA en el que se ha encontrado el “target”; si pudiera tratarse de un target de degradación o rotura; la posición de inicio del “target”; el tipo de “target” encontrado según la clasificación k-mer; el número de “mismatches” en la identificación; y un score del “target” predicho según el algoritmo. Así, se consiguen predecir los posibles “targets” para cada uno de los isomiRNA y miRNA significativamente enriquecidos entre las distintas etapas del desarrollo de *Blattella germanica*.

Además, el paquete scanMiR, mediante su función “viewTargetAlignment”, también permite observar gráficamente los resultados de los alineamientos realizados cuando se exportan en formato “GRanges”.

### **3.5. Sumarización de los “targets” en los mRNA.**

Posteriormente, para poder extraer conclusiones de los resultados de predicción de “targets” en mRNA por parte de los isomiRNA y los miRNA significativamente enriquecidos entre al menos dos etapas del desarrollo de *Blattella germanica*, se desarrolla un código en R (Codigo VI) que permite la sumarización de esos resultados y observar las diferencias existentes entre los “targets” en los mRNA de un miRNA y sus isomiRNA.

En primer lugar, se filtran las predicciones de los “targets” en los mRNA de los isomiRNA y los miRNA por la puntuación (“score”) que les asigna el paquete scanMiR (en función de la complementariedad de la región “seed” entre otros parámetros), desechando aquellas predicciones que presentan un “score” menor de 3, ya que son las menos probables de las predicciones detectadas.

Entonces se generan dos tablas de contingencia que resumen los “targets” predichos para cada mRNA con cada miRNA y con cada isomiRNA. Esto se realiza mediante la preparación de unas tablas generales con ceros, una con todos los mRNA y todos los miRNA y otra con los mRNA y todos los isomiRNA. Luego se cruzan los resultados y se sustituyen los 0 por la clasificación k-mer de la secuencia “seed” de los “targets” predichos para cada miRNA o isomiRNA y cada mRNA. Así, se obtienen dos tablas resumen para las predicciones de los isomiRNA y de los miRNA con la clasificación k-mer para cada mRNA.

Posteriormente, se procede a comprobar las diferencias existentes en los “targets” predichos de los isomiRNA respecto a su miRNA canónico para el conjunto de los mRNA. Para ello se comparan los “targets” predichos de cada uno de los isomiRNA con su miRNA canónico para el conjunto de los mRNA y se puntúa como 1 la existencia de una diferencia y como 0 la presencia de no diferencia. De esta forma, se genera una tabla de contingencia entre todos los isomiRNA y los mRNA con ceros indicando la existencia de no diferencia en la presencia de un “target” del isomiRNA con respecto a su miRNA canónico en ese mRNA y con unos indicando la existencia de una diferencia. Además, se colapsa esta tabla sumando todas las diferencias (unos) halladas para todos los isomiRNA que se referencian al mismo miRNA canónico (“familia de isomiRNA”), generando la misma tabla que se ha citado previamente, pero con los miRNA canónicos y la suma de las diferencias en los “targets” con sus isomiRNA para cada mRNA. Esta última tabla también se filtra para eliminar aquellos miRNA que no presentan ninguna diferencia en los “targets” con sus isomiRNA, y se representan estos últimos datos mediante un heatmap.

También desde la tabla de contingencia de las diferencias (unos) de los “targets” de cada isomiRNA con su miRNA canónico para el conjunto de los mRNA, se colapsa esta tabla en una tabla de contingencia, pero, en este caso, con la suma de las diferencias individuales según el tipo de isomiRNA. La tabla de contingencia, resultante de las diferencias acumuladas por tipos de isomiRNA en los “targets” de los isomiRNA respecto a su miRNA canónico en los mRNA, también se representa mediante un heatmap.

Por último, se crearon dos tablas de contingencia (una para los resultados de los isomiRNA y otra para los de los miRNA) que cruzan los mRNA analizados y las distintas muestras originales (22) de los estadios del desarrollo de *Blattella germanica* para obtener las “cuentas de inhibición”. Exactamente, como “cuentas de inhibición” se define a la suma total de las cuentas del conjunto de miRNA y/o isomiRNA que presentan a un mRNA como target, por lo que son las cuentas que inhiben la expresión de ese mRNA. Así, tomando como origen las tablas de contingencia de los mRNA con los isomiRNA o los miRNA que representan los “targets” predichos mediante la clasificación k-mer y las cuentas de los isomiRNA y miRNA obtenidas de miR&moRe2, se anotan las cuentas de cada isomiRNA o miRNA, en cada muestra de cada estadio del desarrollo, en los mRNA que se les predice un “target” (si hay varios isomiRNA o miRNA se van sumando progresivamente). De esta forma, se generan dos

tablas con las cuentas de la suma de los miRNA o isomiRNA que presentan un “target” para cada mRNA en cada muestra (“cuentas de inhibición”). Además, se genera otra tabla con la suma total de estas dos tablas.

### 3.6. Análisis gráfico, de enriquecimiento y “GO term” de los mRNA.

Una vez obtenidas las tablas de las supuestas cuentas de la regulación negativa (“cuentas de inhibición”) de los mRNA por parte de los isomiRNA, los miRNA y la suma de ambos según los “targets” predichos de éstos, se procedió a realizar el análisis de enriquecimiento de estos datos (procedentes de isomiRNA, miRNA y total) mediante el paquete DESeq2 de R en Windows. Para ello se generó el objeto DESeq2 para cada uno de estos tres sets de datos (isomiRNA, miRNA y total), se transformaron los datos para su representación y se generaron los heatmaps, en general y clusterizados.

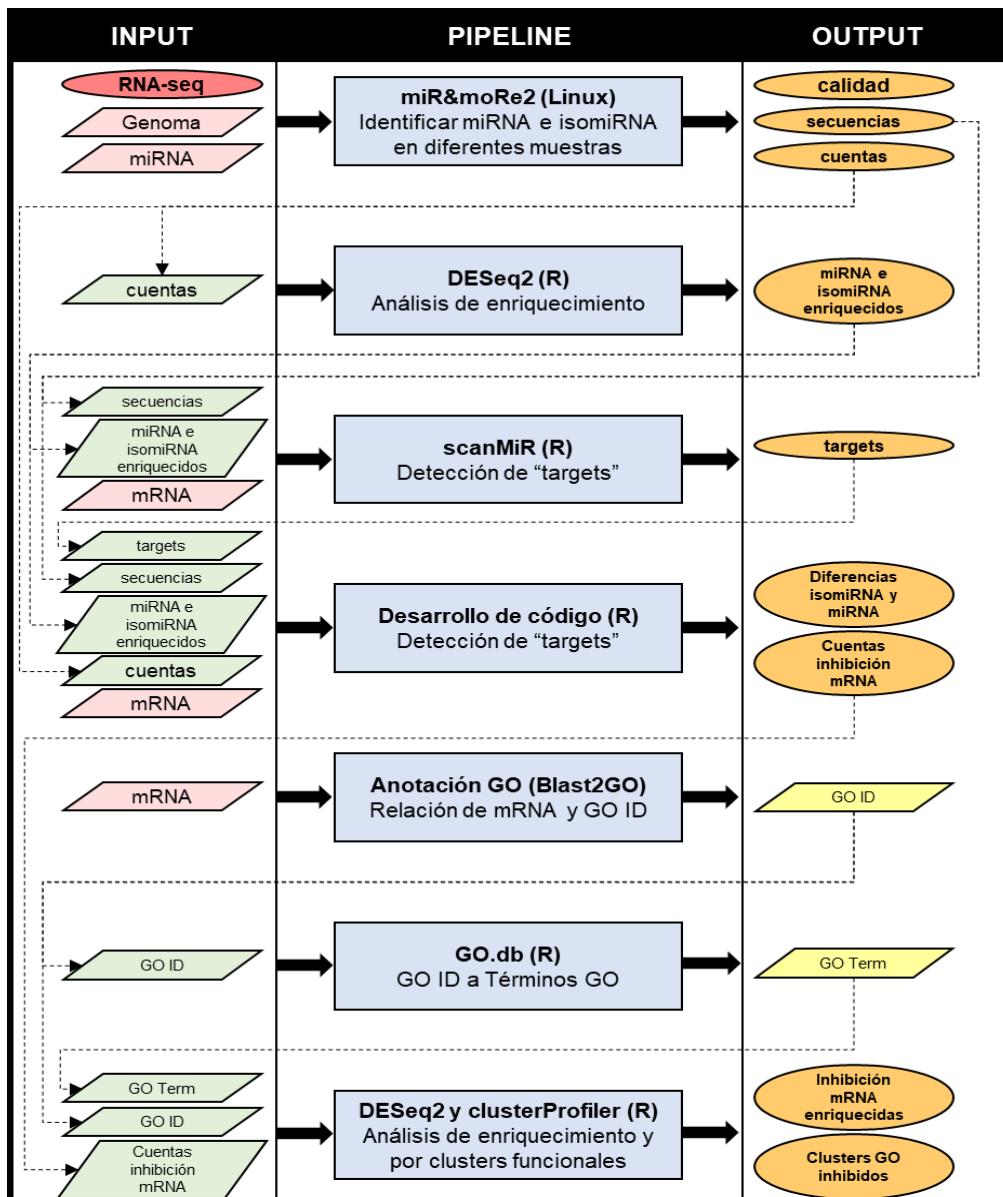
Posteriormente, se realizaron los análisis de enriquecimiento de los tres sets de datos de “cuentas de inhibición”, obteniendo los resultados del análisis para todas las comparaciones entre los 11 estadios de desarrollo de *Blattella germanica*. Entonces se generaron tres tablas (una para cada uno de los sets de datos) en las que se anotaron los datos de “log<sub>2</sub>FoldChange” (el logarítmico en base 2 de la medida del cambio de pliegue) de los análisis de enriquecimiento de los distintos estadios del desarrollo en comparación con los estadios NFE, ED0 o adulto del desarrollo de *Blattella germanica*.

Así, obtenidas las tablas de los datos de “log<sub>2</sub>FoldChange” anteriormente citadas, se procedió a realizar el análisis de enriquecimiento funcional mediante los términos GO (54) de los mRNA para cada uno de los tres sets de datos y en base a NFE, ED0 o adulto. Para realizar este análisis, en primer lugar, se obtuvieron los “GO ID” para los mRNA analizados de *Blattella germanica* mediante la herramienta BLAST2GO del software OmicsBox (27), la cual realizó un blast (identificación de secuencias mediante alineamiento por homología) contra la base de datos swiss-prot de Uniprot (28) en todos los organismos, devolviendo un único y mejor resultado por homología. Lamentablemente, mediante este método solo se consiguieron identificar más de la mitad de los mRNA de *Blattella germanica*. Esos mRNA restantes se intentaron identificar mediante blast contra la base de datos de Refseq “non-redundant proteins”, pero el excesivo tiempo computacional necesario para ello con los recursos disponibles imposibilitó su ejecución. Por ello, los resultados existentes del blast de cada mRNA, se mapearon y se obtuvieron los “GO ID” en el mismo software, los cuales se guardaron en un archivo “.annot”.

Después, ese archivo se cargó en R y, junto con los datos procedentes del paquete GO.db (55) de R, se obtuvieron las relaciones entre los mRNA de *Blattella germanica*, los “GO ID” y los términos GO. Sin embargo, se tuvieron que eliminar las anotaciones de los términos de la base de datos funcional Expasy-ENZYME por no encontrar la forma de analizarlos conjuntamente con los términos GO (56). A continuación, se agruparon los “log<sub>2</sub>FoldChange” en infraexpresado (menor de 0) y sobreexpresado (mayor de 1) para facilitar el análisis de clusteres funcionales mediante términos GO. Así, se realizaron estos análisis obteniendo la comparación de clústeres funcionales entre las distintas etapas del desarrollo de *Blattella germanica*, mediante las funciones “compareCluster” y “enricher” del paquete clusterProfiler de R (30). Estas funciones se ejecutaron según el pipeline mostrado en la documentación del

paquete y con sus parámetros por defecto, a excepción del p-valor, el cual se cambia a 0,1 para aumentar la sensibilidad del análisis.

Finalmente, los resultados de estas funciones para realizar las comparaciones de clústeres funcionales entre las distintas etapas del desarrollo de *Blattella germanica* se expresaron mediante gráficos dotplot para los sets de datos de isomiRNA, miRNA y totales en referencia al estadio NFE, ED0 o adulto. Además, es importante recalcar que el Enriquecimiento detectado en estos casos es un Enriquecimiento en su regulación por miRNA y/o isomiRNA, por lo que en realidad es un Enriquecimiento en la regulación negativa de esos clústeres funcionales.



**Figura 7. Diagrama de flujo desarrollado del pipeline seguido para ejecutar este trabajo.**

En este esquema se muestran los inputs realizados en los distintos procesos (rectángulos azules) del pipeline como: elipse en color rojo oscuro para los datos originales para el trabajo; trapezoides en color rojo claro para los datos obtenidos de la bibliografía existente; y trapezoides en verde para los datos obtenidos de otros procesos de este pipeline. Además, los outputs de este trabajo se muestran como: trapezoides en color amarillo para los resultados no fundamentales necesarios para otros procesos del pipeline; y elipses de color naranja para los resultados fundamentales del pipeline.

## 4. Resultados

### 4.1. Identificación y cuantificación de los isomiRNA y miRNA presentes en las distintas etapas del desarrollo de *Blattella germanica*.

La existencia de datos transcriptómicos de RNA no codificante de distintas etapas del desarrollo de *Blattella germanica* permite estudiar ampliamente la regulación génica que tiene lugar durante el desarrollo de este organismo. Ylla *et al.* (7, 43) han publicado estudios sobre la implicación de los miRNA en este proceso, pero no se han publicado estudios acerca de la implicación de los isomiRNA en el desarrollo de *Blattella germanica*.

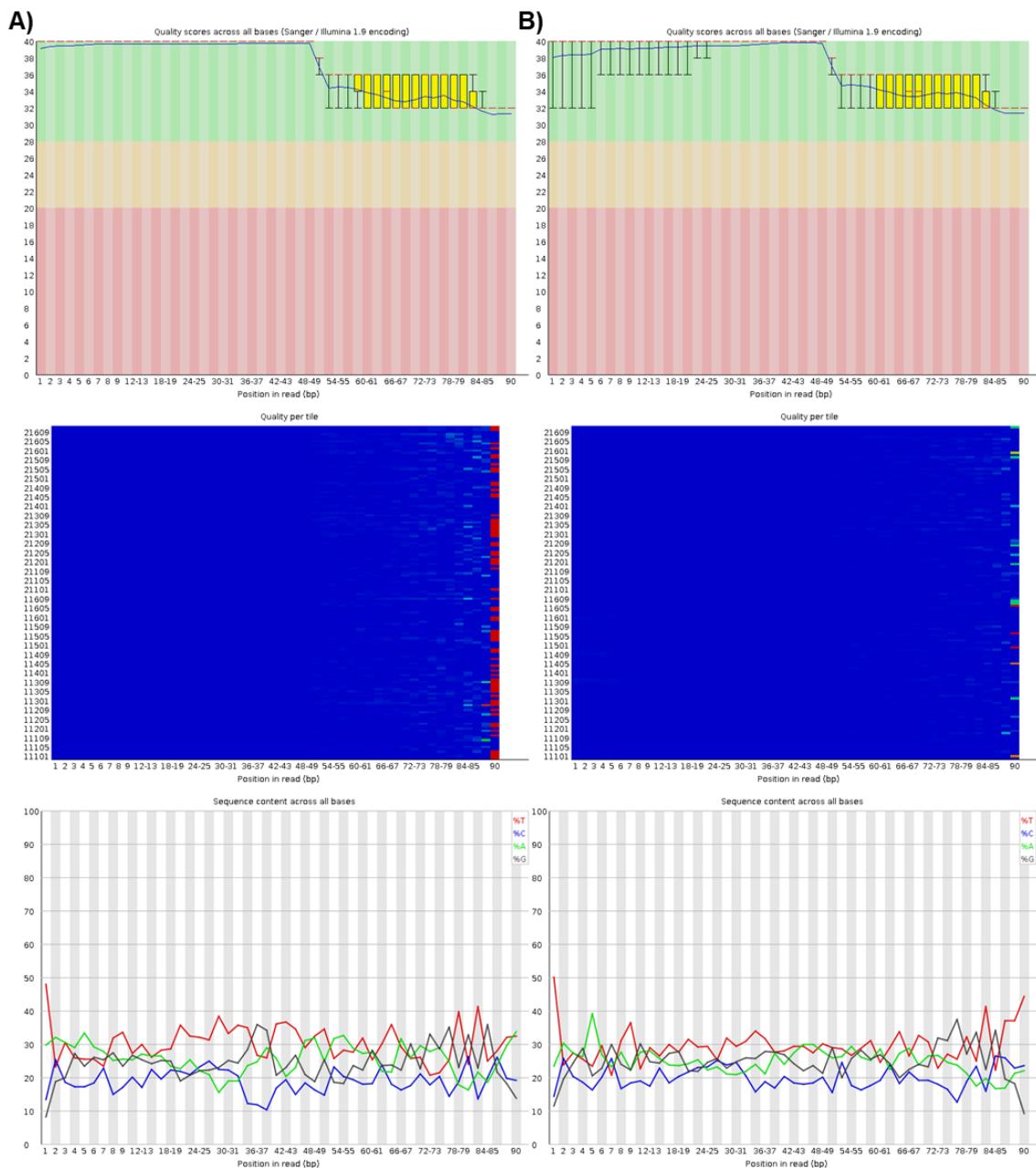
De este modo, para elegir la herramienta más acertada para la identificación y cuantificación de los isomiRNA en esos datos, se realizó una búsqueda bibliográfica acerca de las herramientas disponibles para ese objetivo. Por ello, se experimentó con la ejecución de las herramientas “Jasmine”, “SeqBuster” e “isomirs”, “sRNABench”, “miRGe3.0” y “miR&moRe2”. De entre todas ellas, se decidió el uso de la herramienta “miR&moRe2” para conseguir el objetivo propuesto, ya que permite la detección de isomiRNA, miRNA, sus secuencias y sus cuentas en distintas muestras. Además, la mayoría de las otras herramientas citadas no disponen de la configuración adecuada para poder analizar datos del organismo *Blattella germanica*, al ser éste un organismo menos estudiado actualmente en comparación con otros.

En primer lugar, la herramienta miR&moRe2 realiza un análisis de la calidad de las muestras mediante la herramienta FastQC (Figura 8). En general, en este análisis de la calidad de las muestras se puede observar que las muestras cumplen los estándares necesarios para continuar analizando y extrayendo conclusiones de estas. Por ejemplo, las gráficas “Per base sequence quality” mostradas en la Figura 8 (primera fila), muestran la calidad en todas las bases en cada una de las posiciones a lo largo de las secuencias, por lo que permite evaluar la calidad general de la secuenciación. En este caso, las 2 muestras presentan unos valores de calidad buenos a lo largo de las secuencias, descendiendo un poco en ambas muestras aproximadamente a partir de la posición 50. Sin embargo, a pesar de esta disminución, la calidad sigue siendo buena. Además, los miRNA e isomiRNA son RNA no codificantes pequeños, por lo que, en cualquier caso, no afectarían demasiado al análisis propuesto.

Por otro lado, en la segunda fila de la Figura 8 se muestran las gráficas “Per tile sequence quality”, las cuales muestran la calidad de cada posición de las secuencias en cada una de las unidades de secuenciación de la tecnología *Illumina*, por lo que permite observar si ha habido problemas técnicos en la secuenciación. De esta forma, como se observa en las gráficas de la segunda fila, no han existido problemas técnicos graves en la secuenciación, ya que se observarían filas con colores cálidos. Únicamente se observan colores cálidos en todas las muestras al final de las secuencias (alrededor del nucleótido 80), por lo que, en caso de haber ocurrido problemas técnicos en los últimos ciclos de secuenciación, éstos no afectarían al análisis de miRNA e isomiRNA por la longitud de sus secuencias.

También en la Figura 8, en la tercera fila, se pueden observar las gráficas “Per base sequence content”, las cuales muestran la proporción de cada una de las bases nitrogenadas a lo largo de las posiciones de las secuencias. En los

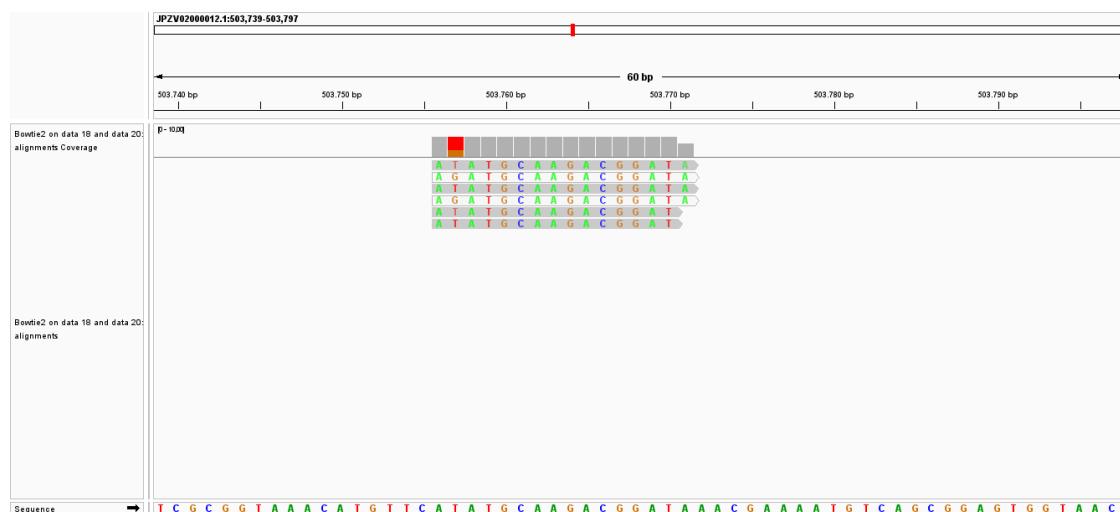
casos mostrados en la figura, a lo largo de las secuencias, en general, no se observan grandes diferencias en la proporción de las bases nitrogenadas a lo largo de las secuencias. Sin embargo, al principio de las secuencias (en la primera posición), sí que se observa en todas las muestras una clara elevación de la proporción de timina. Este hecho probablemente se deba a la forma en la que se construyó la librería para la secuenciación y, a pesar de representar un cierto sesgo técnico, no representa ningún problema para continuar analizando los datos, ya que se reduce a la primera posición de las muestras.



**Figura 8. Análisis de calidad de 2 de las muestras de RNA-seq no codificante del desarrollo de *Blattella germanica*.** Se muestran 3 gráficas del análisis de calidad (“Per base sequence quality” en la primera fila, “Per tile sequence quality” en la segunda, y “Per base sequence content” en la tercera) de 2 muestras seleccionadas al azar (la columna A es la muestra NFE\_1 y la B es la N5\_2) llevado a cabo con miR&moRe2 y la herramienta FastQC.

De esta forma, una vez comprobada la calidad de las muestras, se continuó explorando los resultados obtenidos en la ejecución de la herramienta miR&moRe2. En primer lugar, se observa como esta herramienta procesó 210570351 lecturas entre las 22 muestras, de las cuales: 177587896 lecturas se consideran adecuadas (84,3%); 29238584 lecturas (13,9%) se consideran adecuadas, aunque presentan más de 31 nucleótidos; y 3743871 lecturas (un 1,8 %) se descartaron por tener 2 o más bases con baja calidad. Además, también se observa cómo esta herramienta ha anotado 283 miRNA canónicos y 37977 isomiRNA entre las 22 muestras; además de moRNA, isomoRNA, secuencias loop, variantes de las secuencias loop y nuevas secuencias predichas de estas moléculas. Sin embargo, para los consiguientes análisis se tienen únicamente en cuenta las anotaciones de miRNA e isomiRNA con el objetivo de centralizar el estudio en estas moléculas y obtener los resultados más exactos y precisos posibles. Además, es necesario indicar que, para aumentar la sensibilidad de la detección, en la anotación de estos isomiRNA y miRNA no se aplicó ningún tipo de filtro por cuentas, por lo que también se anotó un miRNA o isomiRNA con una única lectura en una sola muestra.

Mediante la herramienta miR&moRe2 se obtienen las anotaciones y las cuentas de los isomiRNA y los miRNA en las distintas etapas del desarrollo de *Blattella germanica*, además de sus secuencias y, en el caso de los isomiRNA varios datos descriptivos de cada uno, como su tipo de isomiRNA. Como ilustración, en la Figura 9, se puede ver el alineamiento de los datos crudos con el genoma de *Blattella germanica* mediante el software IGV. Exactamente, se pueden observar un miRNA y sus isomiRNA (de longitud, polimórfico y mixto) alineados con el genoma del insecto.



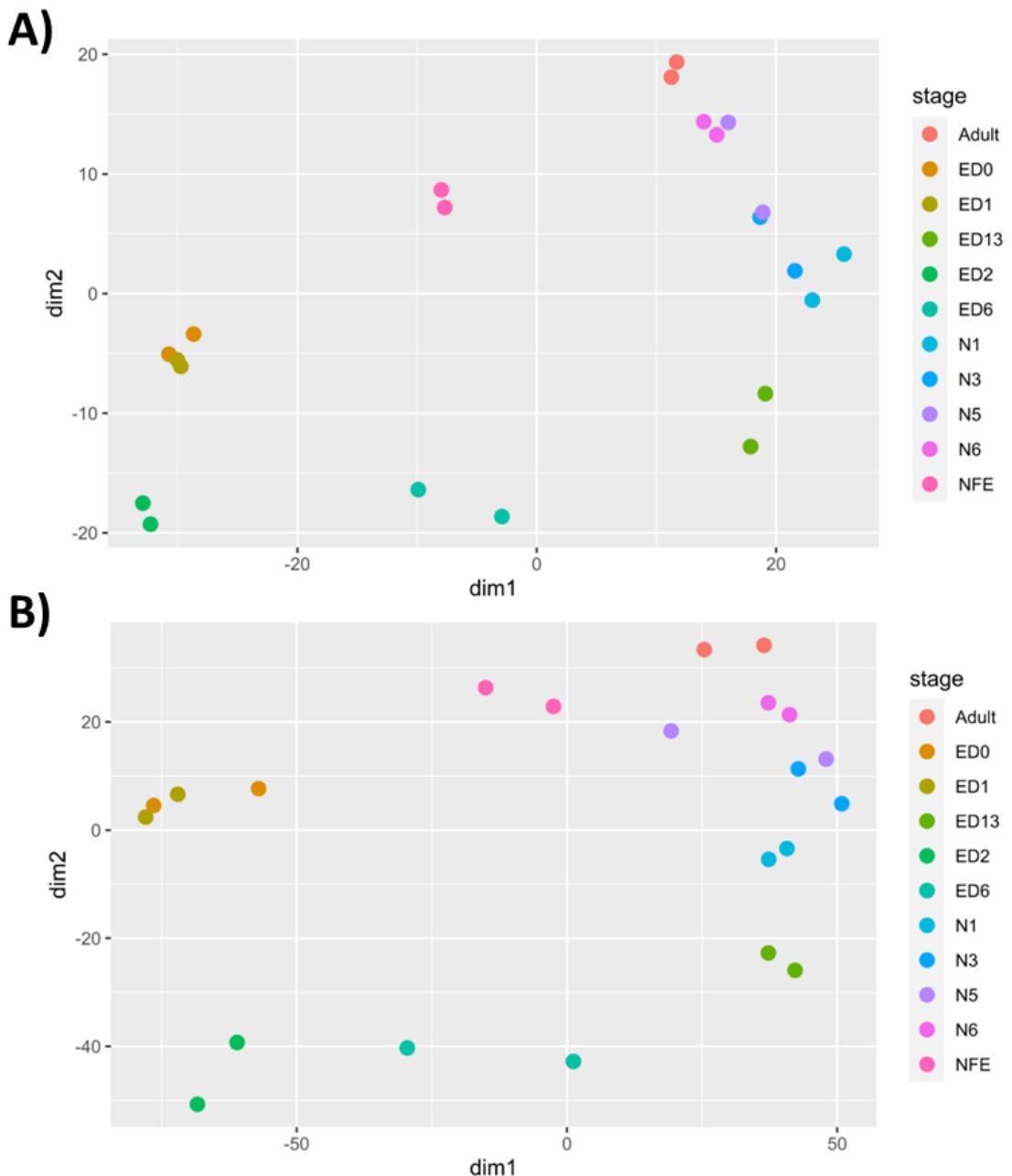
**Figura 9. Representación de la alineación de los datos brutos de una de las muestras con el genoma de *Blattella germanica* mediante el software IGV.** En esta figura se puede observar, en la parte superior la ubicación en el total del genoma de *Blattella germanica* (línea roja y el número de pares de bases), la cobertura de las lecturas analizadas (barras verticales grises), el número de lecturas alineadas en esa zona del genoma con las secuencias de cada una de esas lecturas y sus nucleótidos en distintos colores según cada base nitrogenada, y la secuencia de nucleótidos (también con un color para cada base nitrogenada) del genoma de *Blattella germanica* en esa zona concreta.

#### **4.2. Análisis de las diferencias en los isomiRNA y miRNA entre las distintas etapas del desarrollo de *Blattella germanica*.**

A partir de los datos obtenidos por miR&moRe2 de anotaciones y cuentas de los isomiRNA y los miRNA, se procedió a su análisis en R mediante el paquete DESeq2 para filtrar, observar y analizar el enriquecimiento de los isomiRNA y los miRNA en las distintas etapas del desarrollo de *Blattella germanica*.

Así, en primer lugar, se filtraron las anotaciones de isomiRNA y miRNA por el número de cuentas para reducir el ruido experimental, ya que, como se ha detallado previamente, en la anotación de éstos isomiRNA y miRNA por miR&moRe2 no se aplicó este filtro. Esto es debido a que algunos isomiRNA y miRNA pueden estar muy poco o nada presentes en algunas muestras y más presentes en otras muestras. Por ello, en este punto se realiza un filtro de cuentas por el que se eliminaron aquellas anotaciones con un número de cuentas menor a 220 cuentas entre todas las 22 muestras, es decir, una media de 10 cuentas por anotación. La diferencia entre este filtro por cuentas y aplicar un filtro por cuentas en cada muestra (como el que se podía haber aplicado en miR&moRe2), es que el aplicado en este caso es más sensible y respeta diferencias más grandes de cuentas entre distintas muestras. De esta forma, tras aplicar este filtro de cuentas, las anotaciones de miRNA pasan de 283 a 189 miRNA y las anotaciones de isomiRNA pasan de 37977 a 1688 isomiRNA. Así, se puede observar como el filtro de cuentas afecta más a las anotaciones de los isomiRNA (se reduce en un 95%) que a las de los miRNA (se reduce en un 33%), siendo este hecho evidente al ser los isomiRNA pequeñas variaciones de los miRNA canónicos, las cuales se pueden confundir con pequeños fallos en el proceso de secuenciación. Por ello, es fundamental aplicar un filtro de cuentas adecuado en la anotación de isomiRNA.

Una vez filtradas las anotaciones de isomiRNA y miRNA por el número de cuentas, se observaron gráficamente las diferencias existentes en las cuentas de las anotaciones de isomiRNA y miRNA entre las distintas etapas del desarrollo de *Blattella germanica* mediante gráficos de componentes principales (PCA), heatmaps de distancias y heatmaps de distancias clusterizados. En concreto, el PCA (Figura 10) permite mantener el número de variables (anotaciones), reduciendo su dimensionalidad en la representación gráfica (dos dimensiones para la representación), pudiéndose observar simplemente las diferencias existentes entre los grupos experimentales.



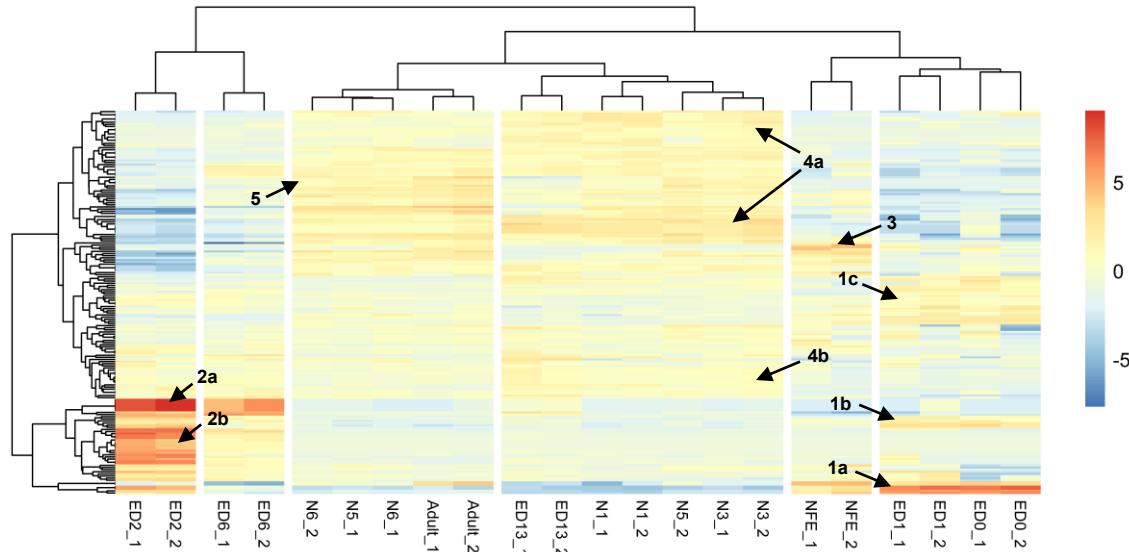
**Figura 10.** Gráficas de análisis de componentes principales de los miRNA (A) y los isomiRNA (B) entre las distintas etapas del desarrollo de *Blattella germanica*. Las leyendas situadas a la derecha de cada una de las gráficas muestran los colores con los que se representan cada uno de los estadios del desarrollo.

De este modo, en la Figura 10 se puede observar cómo el PCA de los isomiRNA (B) y el de los miRNA (A) se asemejan considerablemente, por lo que, es probable que las diferencias existentes en el conjunto de las anotaciones de los miRNA y los isomiRNA se asemejen bastante. Exactamente, en ambos casos, las distintas muestras de los estadios de *Blattella germanica* se distribuyen a lo largo del plano de dos dimensiones formando una elipse ordenada en la que el estadio de huevos no fecundados se sitúa en la parte superior de la misma. Siguiendo el perímetro de la supuesta elipse desde el estadio de huevos no fecundados hacia la izquierda y levemente hacia abajo se observan los estadios embrionarios ED0, ED1 y ED2, los cuales se encuentran considerablemente separados entre sí. Posteriormente hacia la derecha en el perímetro de la elipse, se encuentran los estadios embrionarios ED6 y ED13, también considerablemente separados entre sí. Y, desde ED13 hacia arriba en el perímetro de la elipse, se encuentran los estadios ninfales y el adulto considerablemente cerca los unos de los otros. Así, de estas gráficas se puede intuir que las anotaciones y cuentas de isomiRNA y miRNA detectadas representan el ciclo biológico de *Blattella germanica* y que no existen grandes diferencias a este nivel de análisis entre ambos (isomiRNA y miRNA), más que una leve mayor separación en los estados embrionarios en los isomiRNA respecto a los miRNA.

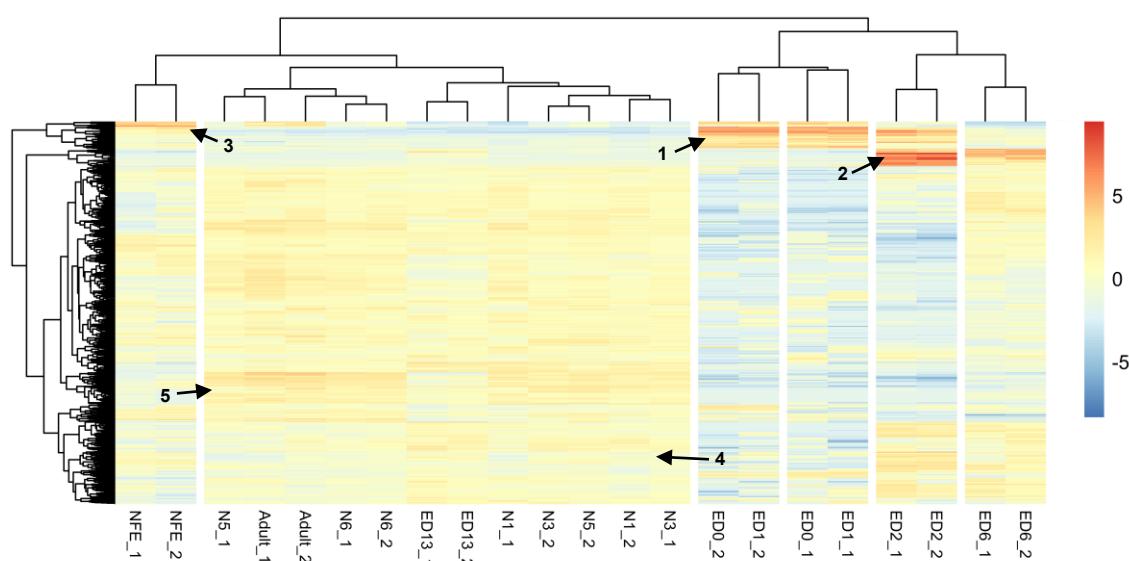
Tras observar la coherencia de las distancias existentes en el análisis de PCA entre los estadios del desarrollo de *Blattella germanica* en los datos de isomiRNA y miRNA, se realizó la representación de heatmaps para los grupos experimentales y el conjunto de anotaciones en los isomiRNA (Figura 12) y los miRNA (Figura 11). Estos gráficos permiten observar la agrupación entre los estadios del desarrollo y las anotaciones para observar sus diferencias.

Así, el heatmap de los miRNA (Figura 11) muestra diferencias considerables en la expresión de distintos miRNA en los distintos estadios del desarrollo de *Blattella germanica*. Exactamente, se observan tres grupos de miRNA que se expresan desde el huevo no fecundado hasta la etapa embrionaria ED2, presentando un pico en su expresión en la etapa ED0-ED1 y no expresándose prácticamente en el resto de las etapas (1a, 1b y 1c). Uno de esos grupos de miRNA presenta una mayor expresión (1a) y un mayor pico de expresión en la etapa ED1 que el otro grupo (1b) de miRNA, mientras que el tercer grupo (1c) presenta una expresión muy leve. Por otro lado, hay otros dos grupos de miRNA más grandes que se expresan mayoritariamente en el estadio embrionario ED2 (2a y 2b), aunque aún mantienen algo de su expresión durante la etapa ED6 y una expresión muy muy leve en la etapa ED13, no expresándose prácticamente nada en el resto de las etapas del desarrollo. En el caso de estos dos grupos de miRNA ocurre como con los anteriores, uno se expresa más (2a) y con un mayor pico de expresión que el otro (2b). También existe un grupo de miRNA (3) el cual presenta una mayor expresión en los huevos no fecundados y una expresión moderada en el individuo adulto, mientras que no se expresan en el resto de las etapas del desarrollo. Además, se pueden observar dos grupos de miRNA (4a y 4b) que caracterizan principalmente al conjunto de las etapas ninfales. Exactamente ambos grupos presenta una mayor expresión desde ED13 hasta la primera mitad de las etapas ninfales y una leve menor expresión desde ese punto hasta el adulto, pero uno de ellos no se expresa en el resto de las etapas (4a) mientras que el

otro (4b) sí que se expresa levemente también en las etapas ED2 y ED6. Y, finalmente en este heatmap, se puede observar otro grupo de miRNA (5) el cual se expresa comienza a expresar levemente en las etapas ninfales y presenta su mayor expresión en la etapa adulta. De este modo, el conjunto de grupos de miRNA que se observan en el heatmap de miRNA permiten intuir una caracterización de cada estadio del desarrollo de *Blattella germanica* por los gradientes existentes en esos grupos de miRNA, la cual se comprueba por la correcta agrupación que realiza el heatmap de los estadios del desarrollo



**Figura 11. Heatmap de distancias de las anotaciones de los miRNA en las muestras de los estadios del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de expresión estandarizados de los distintos miRNA (filas) en las distintas muestras, de modo que se organizan gráficamente según esa expresión tanto en las muestras (columnas) como los miRNA. Las flechas en el gráfico indican los distintos de miRNA que se citan en el texto.

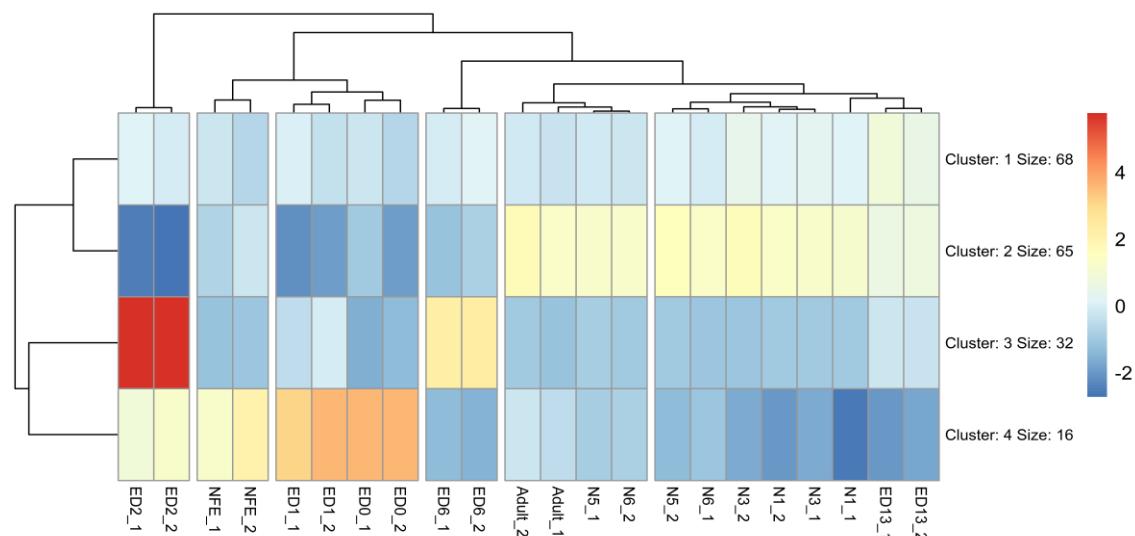


**Figura 12. Heatmap de distancias de las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de expresión estandarizados de los distintos isomiRNA (filas) en las distintas muestras, de modo que se organizan gráficamente según esa expresión tanto en las muestras (columnas) como los isomiRNA. Las flechas en el gráfico indican los grupos de isomiRNA citados.

Por otro lado, el heatmap de los isomiRNA (Figura 12) muestra unos resultados similares, aunque ligeramente diferentes a los resultados descritos en el heatmap de los miRNA (Figura 11) para la caracterización de los distintos estadios del desarrollo de *Blattella germanica*. En primer lugar, el grupo 1 del heatmap de los isomiRNA, al igual que el grupo 1 de los miRNA, caracteriza principalmente las etapas ED0-ED1 y se extiende, con una menor expresión, hasta la etapa ED2. Además, en el caso de los isomiRNA, el grupo 1 también se expande levemente a la etapa ED6. Por otro lado, los grupos 2, 3, 4 y 5 se asemejan mucho entre el heatmap de los miRNA y isomiRNA, caracterizando en ambos casos: el grupo 2 las etapas ED2, ED6 y levemente ED13; el grupo 3 las etapas de huevo y adultas; el grupo 4 las etapas ninfales tempranas y en menos las etapas tardías y adulto; y el grupo 5 al contrario que el grupo 4.

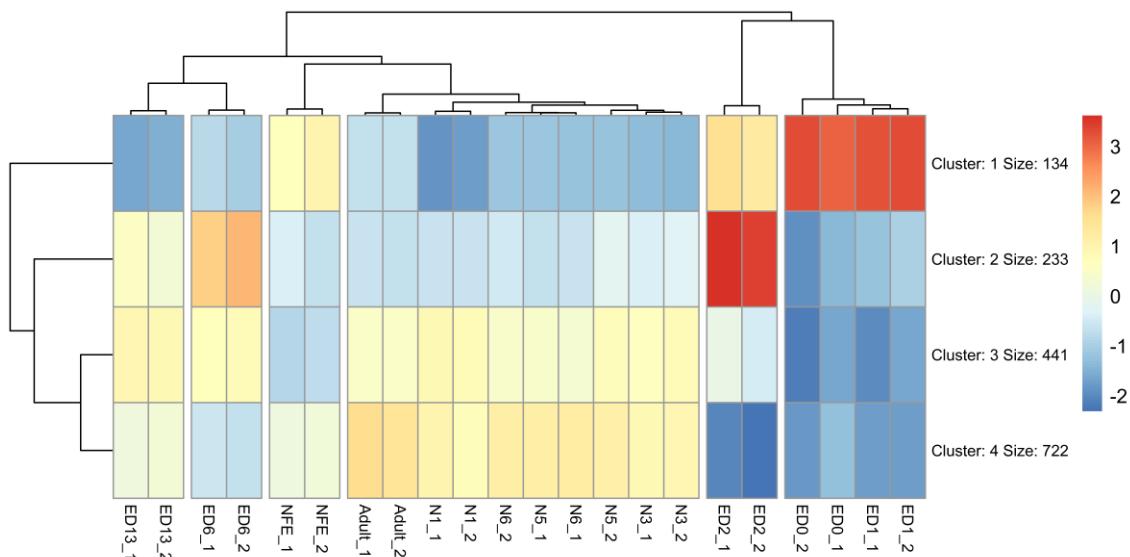
Es por todo ello, que las visualizaciones de los heatmaps refutan los resultados obtenidos en los anteriores análisis PCA, ya que se observa un ciclo, una relación y una diferenciación entre los distintos grupos de miRNA e isomiRNA en las diferentes etapas del desarrollo de *Blattella germanica*, apareciendo pequeñas únicamente diferencias en las etapas embrionarias entre los miRNA y los isomiRNA, en las cuales los isomiRNA presentan más detalles.

Para reforzar y comprobar aún más estos resultados, se realizaron representaciones de heatmaps por clusteres para los miRNA (Figura 13) y los isomiRNA (Figura 14), en los que se agruparon 4 clusteres distintos. De esta forma, en el de los miRNA (Figura 13), se puede observar como hay un cluster que se corresponde al anterior grupo 1 y 3 (cluster 4) del heatmap de los miRNA, ya que presenta un pico de su expresión ED0-ED1 empezando en NFE y terminando en ED2. Además, el cluster 3 se correspondería al grupo 2 del heatmap de los miRNA, ya que presenta su pico de expresión en la etapa ED2 y se alarga hasta ED6. Y, el cluster 1 y el cluster 2 se corresponderían a los grupos 4 y 5 respectivamente, ya que el primero caracteriza ED13 y las etapas ninfales tempranas y el segundo las etapas ninfales tardías y el adulto.



**Figura 13. Heatmap clusterizado de las anotaciones de los miRNA en las muestras de los estadios del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de expresión de los clusteres de miRNA (agrupados por la expresión de los miRNA en las distintas muestras), organizándose según la expresión de los clusteres (filas) en las muestras (columnas).

En el caso del heatmap clusterizado de los isomiRNA, éste se parece enormemente al de los miRNA y se corresponde también con los heatmaps de distancias previos. De esta forma, y al igual que en el caso anterior, en este caso el grupo 1 y 3 de los heatmaps de distancias se referenciaría al cluster 1 de este heatmap (NFE, ED0, ED1, ED2), el grupo 2 al cluster 2 (ED2, ED6, ED13), el grupo 4 al cluster 3 (ED6, ED13, primeras etapas ninfales), y el grupo 5 al cluster 4 (últimas etapas ninfales y etapa adulta)



**Figura 14. Heatmap clusterizado de las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de expresión de los clusters de isomiRNA (agrupados por la expresión de los isomiRNA en las muestras), organizándose según la expresión de los clusteres (filas) en las muestras (columnas).

De esta forma, los datos de los heatmaps clusterizados de los miRNA y los isomiRNA a lo largo de las etapas del desarrollo de *Blattella germanica* concuerdan con los datos e interpretaciones obtenidos previamente. Por ello, se resumen todas estas interpretaciones acerca de la correlación entre ciertos grupos de miRNA e isomiRNA con el ciclo biológico de *Blattella germanica* en la siguiente tabla (Tabla 1).

Grupos heatmaps	Cluster miRNA	Cluster isomiRNA	Etapa del desarrollo
Grupo 1	4	1	ED0 – ED1
Grupo 2	3	2	ED2 – ED6
Grupo 3	4	1	NFE
Grupo 4	1	3	ED13 – Ninfa (1º)
Grupo 5	2	4	Ninfa (2º) – Adulto

**Tabla 1. Correlación entre los grupos observados en los heatmaps de distancias, los heatmaps clusterizados y las etapas del desarrollo para los miRNA y los isomiRNA.** La primera columna de esta tabla representa los grupos mostrados en los heatmaps de distancias de miRNA e isomiRNA, la segunda los clusteres del heatmap clusterizado de miRNA, la tercera los clusteres del heatmap clusterizado de isomiRNA, y la cuarta columna las etapas del desarrollo a las que se asocian cada uno de estos grupos.

Como se ha observado, a pesar de las pequeñas diferencias observadas entre los isomiRNA y los miRNA en las etapas embrionarias tempranas, los miRNA y

los isomiRNA presentan agrupaciones comunes de los mismos que se corresponden con las etapas del desarrollo de *Blattella germanica*.

Una vez observadas gráficamente las diferencias de los miRNA y los isomiRNA en el desarrollo de *Blattella germanica*, se realizaron los análisis de enriquecimiento de los miRNA y los isomiRNA entre las distintas etapas para obtener resultados estadísticos acerca de las diferencias apreciadas anteriormente. En esos análisis de enriquecimiento se usó el p-valor de 0.001 como límite para que una diferencia se considere significativa. Así, tras realizar los análisis, se resumieron los resultados de las anotaciones significativamente enriquecidas entre al menos dos etapas en las siguientes tablas (Tablas 2 y 3):

	NFE	ED0	ED1	ED2	ED6	ED13	N1	N3	N5	N6	Adult
NFE	0	30	31	79	62	58	59	70	45	44	42
ED0	30	0	2	69	47	47	40	46	46	47	42
ED1	31	2	0	57	38	54	58	64	57	61	64
ED2	79	69	57	0	64	103	110	104	110	105	114
ED6	62	47	38	64	0	60	76	83	79	76	82
ED13	58	47	54	103	60	0	18	26	27	49	40
N1	59	40	58	110	76	18	0	4	13	25	27
N3	70	46	64	104	83	26	4	0	3	8	13
N5	45	46	57	110	79	27	13	3	0	0	6
N6	44	47	61	105	76	49	25	8	0	0	3
Adult	42	42	64	114	82	40	27	13	6	3	0

**Tabla 2. Número de miRNA con un enriquecimiento significativo (p-valor menor de 0.001) entre cada una de las distintas etapas del desarrollo de *Blattella germanica*.**

	NFE	ED0	ED1	ED2	ED6	ED13	N1	N3	N5	N6	Adult
NFE	0	71	95	300	187	177	205	146	169	138	166
ED0	71	0	0	104	135	126	138	116	127	127	138
ED1	95	0	0	95	136	141	151	137	142	130	159
ED2	300	104	95	0	151	268	410	354	391	417	460
ED6	187	135	136	151	0	107	206	173	161	189	216
ED13	177	126	141	268	107	0	56	59	87	96	182
N1	205	138	151	410	206	56	0	2	11	28	53
N3	146	116	137	354	173	59	2	0	2	5	33
N5	169	127	142	391	161	87	11	2	0	0	12
N6	138	127	130	417	189	96	28	5	0	0	10
Adult	166	138	159	460	216	182	53	33	12	10	0

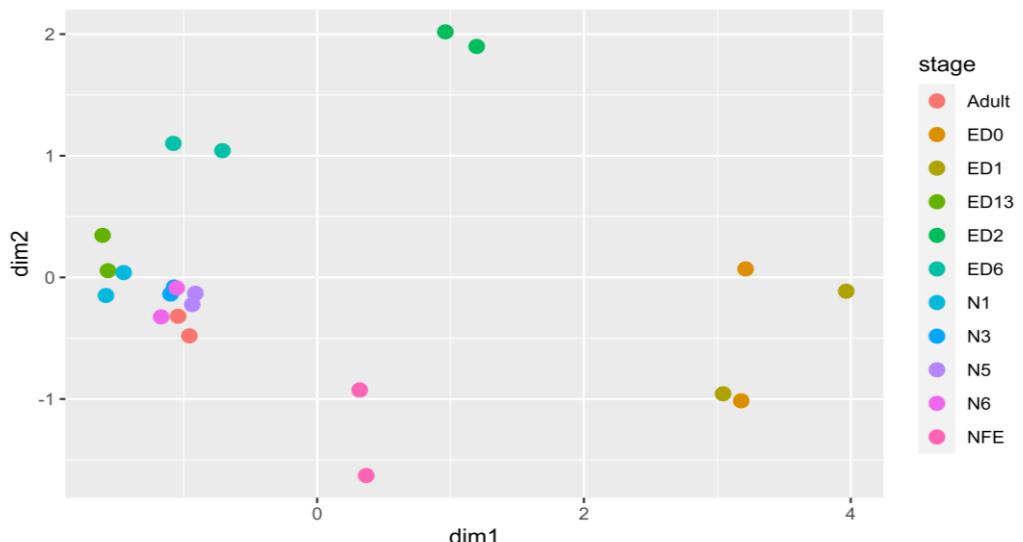
**Tabla 3. Número de isomiRNA con un enriquecimiento significativo (p-valor menor de 0.001) entre cada una de las distintas etapas del desarrollo de *Blattella germanica*.**

Al observar estas tablas sobre las anotaciones significativamente enriquecidas de miRNA y isomiRNA entre las etapas del desarrollo, se puede observar cómo, en ambos casos hay un mayor número de anotaciones significativas entre las distintas etapas del periodo embrionario y de huevo no fecundado (coloreadas de amarillo) que entre las distintas etapas del periodo de ninfa y adulto (coloreadas en azul), por lo que se puede concluir que tanto los miRNA como los isomiRNA actúan regulando la expresión génica mayoritariamente en las etapas embrionarias y de huevo no fecundado. Además, entre los miRNA y los isomiRNA, esta implicación predominante en la regulación génica de la etapa embrionaria y de huevo no fecundado es aún más predominante en el caso de los isomiRNA que en el caso de los miRNA.

#### 4.3. Análisis de las diferencias en las categorías de isomiRNA entre las distintas etapas del desarrollo de *Blattella germanica*.

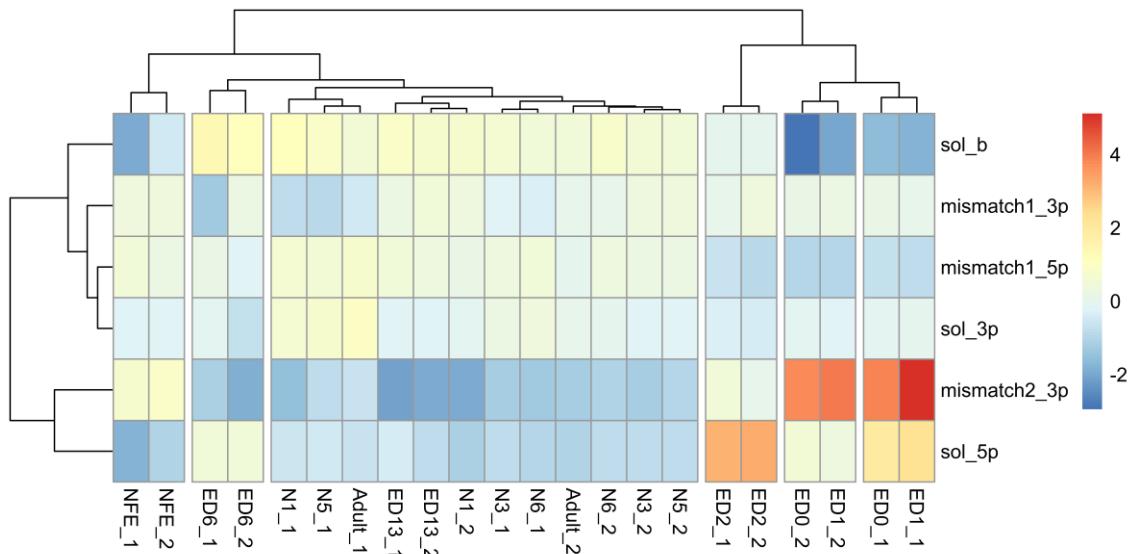
Dadas las diferencias observadas previamente sobre la mayor involucración de los isomiRNA respecto a los miRNA en la regulación génica de las etapas embrionarias del desarrollo de *Blattella germanica*, se procedió a estudiar si esas diferencias eran debidas a una categoría de isomiRNA concreta. Para ello, se modificaron los resultados de anotaciones y cuentas de isomiRNA en las distintas etapas del desarrollo de *Blattella germanica*, obtenidos de miR&moRe2. Exactamente, se sumaron las cuentas de las anotaciones de los isomiRNA (previamente filtradas por el número de cuentas) pertenecientes a una misma categoría en cada etapa del desarrollo. Con estos datos de cuentas de cada categoría de isomiRNA en las distintas etapas del desarrollo, se procedió a explorar visualmente esos datos y realizar un análisis de enriquecimiento como en el apartado anterior.

En primer lugar, se obtiene el gráfico del análisis de componentes principales de las cuentas de las categorías de los isomiRNA en las distintas etapas del desarrollo de *Blattella germanica*. Así, esta gráfica muestra una distribución muy parecida a la que se obtuvo en los análisis de componentes principales previos para los miRNA y los isomiRNA (Figura 10). Sin embargo, en este caso, el estado de huevo no fecundado y las etapas embrionarias (menos E13) del desarrollo se encuentran aún más alejadas entre sí y del resto de etapas, mientras que las etapas de ninfa (además de E13) y adulto se encuentran aún más cerca entre sí. Estos resultados indican que las diferencias de categorías de isomiRNA en las distintas etapas del desarrollo influyen mucho en las etapas del desarrollo embrionario de *Blattella germanica*.



**Figura 15. Gráfico de análisis de componentes principales de los isomiRNA según su categoría entre las distintas etapas del desarrollo de *Blattella germanica*.** Las leyendas situadas a la derecha de la gráfica muestran los colores con los que se representan cada uno de los estadios del desarrollo.

Por ello, a continuación, se obtuvo un heatmap de distancias con los datos de cuentas de las categorías de isomiRNA en las distintas etapas del desarrollo de *Blattella germanica* para comprobar si existen diferencias en el tipo concreto de categoría de isomiRNA en los distintos estados del desarrollo, ya que el gráfico anterior nos indica que sí que influyen las categorías de los isomiRNA.



**Figura 16. Heatmap de distancias de las categorías de isomiRNA según las anotaciones de los isomiRNA en las muestras de los estadios del desarrollo de *Blattella germanica*.** Esta gráfica muestra los niveles estandarizados de expresión de los isomiRNA según las categorías de isomiRNA (filas), organizándose por su expresión en las muestras (columnas).

Observando el heatmap de distancias citado (Figura 16), se puede observar como los isomiRNA de tipo **longitud-5'** (sol\_5p) son los que se expresan mayoritariamente en el estadio ED2, aunque también se expresan levemente en los estadios previos y en ED6, por lo que, aproximadamente, se asocia al grupo 2 de los heatmaps citado en la Tabla 1. En cuanto a los isomiRNA de tipo **polimórficos-3' con 2 nucleótidos diferentes** del miRNA canónico (mismatch2\_3p), éstos se expresan preferentemente en los estadios ED0-ED1 y, levemente, en los huevos no fecundados (NFE), por lo que se asocian con el grupo 1 y 3 de los heatmaps de la Tabla 1. Por otro lado, los isomiRNA de tipo **longitud tanto 5' como 3'** (sol\_b), parecen estar más enriquecidos desde ED6 hasta las primeras etapas ninfales, aunque no es un efecto tan consistente como los anteriores, por lo que, se podría llegar a encuadrar en el grupo 4 de los heatmaps de la Tabla 1. También en el grupo 4, y además en el grupo 5, se incluirían los isomiRNA de tipo **polimórficos-5' con un nucleótido diferente** (mismatch1\_5p), ya que se encuentran más expresados en las etapas ninfales y adulta, además de en los huevos no fecundados. Además, los isomiRNA de tipo **longitud-3'** (sol\_3p) parecen presentar una leve tendencia a encontrarse más enriquecidos en las etapas ninfales finales y el adulto, por lo que se incluirían en el grupo 5 de los heatmaps de la Tabla 1. En cambio, en los isomiRNA de tipo **polimórficos-3' con un nucleótido diferente** (mismatch1\_3p) no se observa una tendencia clara a estar enriquecidos en un estadio concreto del desarrollo de *Blattella germanica*.

De esta forma, se obtienen indicios claros de que, no solo los isomiRNA tienen un papel más importante en la regulación génica de los estados embrionarios de *Blattella germanica* que los miRNA, sino que parece que los distintos tipos de isomiRNA juegan papeles distintos en las distintas etapas del desarrollo de *Blattella germanica*. En concreto, los tipos de isomiRNA que destacan en la regulación de las etapas embrionarias serían: los isomiRNA de longitud-5', los isomiRNA polimórficos-3' con 2 nucleótidos distintos, y, en menor medida, los isomiRNA de tipo longitud tanto 3' como 5'.

A continuación, para comprobar el papel de los distintos tipos de isomiRNA en las distintas etapas del desarrollo embrionario de *Blattella germanica*, se realizó un análisis de enriquecimiento sobre los datos previamente analizados. De este modo, posteriormente se sumarizaron los resultados de los tipos de isomiRNA significativamente enriquecidos (con un p-valor menor de 0.001) entre algunas de las etapas del desarrollo, cuyos resultados se muestran en la Tabla 4.

	NFE	ED0	ED1	ED2	ED6	ED13	N1	N3	N5	N6	Adult
NFE	0	3	3	2	3	2	2	2	2	2	2
ED0	3	0	0	2	2	4	4	4	4	4	4
ED1	3	0	0	2	3	4	4	4	4	4	4
ED2	2	2	2	0	3	3	3	3	3	3	3
ED6	3	2	3	3	0	0	0	0	0	0	0
ED13	2	4	4	3	0	0	0	0	0	0	0
N1	2	4	4	3	0	0	0	0	0	0	0
N3	2	4	4	3	0	0	0	0	0	0	0
N5	2	4	4	3	0	0	0	0	0	0	0
N6	2	4	4	3	0	0	0	0	0	0	0
Adult	2	4	4	3	0	0	0	0	0	0	0

**Tabla 4. Número de categorías de isomiRNA enriquecidas significativamente (p-valor < 0.001) entre cada una de las distintas etapas del desarrollo de *Blattella germanica*.**

Para interpretar esta tabla, primero hay que ser conscientes de que se han estudiado 6 tipos distintos de isomiRNA, por lo que aquellos estadios que presentan un enriquecimiento significativo de 3 tipos de isomiRNA respecto a otro estadio, en realidad presentan un enriquecimiento significativo de la mitad de los tipos de isomiRNA estudiados. De esta manera, podemos observar en la tabla de tipos de isomiRNA significativamente enriquecidos (Tabla 4), que las etapas ninfales y adultas no presentan ningún tipo de isomiRNA enriquecido entre sí, mientras que la mayoría de las etapas embrionarias y huevos no fecundados (a excepción de ED0 con ED1 y de ED6 con ED13) presentan más de un tipo de isomiRNA enriquecidos entre sí. Es por ello por lo que, junto con los resultados previos observados, se puede concluir que los isomiRNA y concretamente, los distintos tipos de isomiRNA, juegan un papel predominante en la regulación génica del desarrollo embrionario.

#### 4.4. Predicción de los “targets” en los mRNA de los miRNA e isomiRNA.

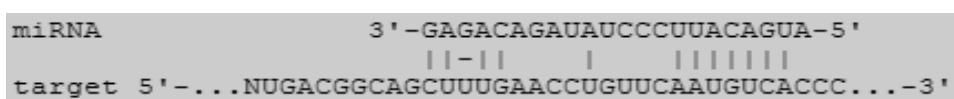
A través de los análisis de enriquecimiento de miRNA e isomiRNA en las distintas etapas del desarrollo de *Blattella germanica*, se obtuvieron dos listas de anotaciones, una con los miRNA y otra con los isomiRNA, enriquecidos entre, al menos, dos etapas del desarrollo. Mediante estas listas de anotaciones, más las secuencias de las anotaciones obtenidas por miR&moRe2 y las secuencias de los mRNA de *Blattella germanica* (21), se objetivizó realizar predicciones sobre los posibles “targets” de los miRNA y los isomiRNA en los mRNA de *Blattella germanica*.

Así, se realizó una revisión bibliográfica sobre las herramientas bioinformáticas disponibles y los métodos desarrollados actualmente para este objetivo (20). Algunos de los métodos desarrollados actualmente para predecir los “targets” de los miRNA se basan en: el “seed pairing” (la complementariedad de la secuencia “seed” del miRNA y la secuencia 3'-UTR del mRNA, además de

otros parámetros del resto de la secuencia), la estabilidad termodinámica (las propiedades termodinámicas de la unión del miRNA y el “target”); el mantenimiento evolutivo de los “targets”; la accesibilidad y número de “targets” en la región 3'-UTR del mRNA, o “Machine Learning”. Todos estos métodos pueden ejecutarse de forma complementaria y/o paralela para predecir los “targets” de un miRNA en un mRNA. Sin embargo, por la escasez actual de estudios previos sobre las interacciones de los miRNA y los mRNAs en *Blattella germanica* se desecharon los métodos dependientes de “Machine Learning”, la estructura de las regiones 3'-UTR, y el mantenimiento evolutivo.

Entonces, se buscaron las herramientas bioinformáticas disponibles para la detección de “targets” de miRNA mediante los métodos “seed pairing” y estabilidad termodinámica. Entre esas herramientas, se desecharon aquellas que no permitían realizar el análisis en cualquier organismo (por no encontrarse *Blattella germanica* entre las especies comúnmente estudiadas), como son los casos de TargetScan (22) o miRDB (23). Además, se intentaron ejecutar las herramientas miRanda (24) o MIRZA-G (25), pero no se obtuvieron resultados satisfactorios. Por ello, finalmente se escogió la herramienta scanMiR (26), un paquete de Bioconductor para R que predice los miRNA en mRNAs mediante “seed pairing”. Exactamente, esta herramienta aplica modelos lineares simples para hallar posibles targets mediante “seed pairing” y otras características del miRNA y el mRNA (complementariedad del resto de la secuencia y sustituciones T/G).

De este modo, se ejecutó la herramienta scanMiR sobre las secuencias de los isomiRNA y miRNA significativamente Enriquecidos entre algunas de las etapas del desarrollo de *Blattella germanica* y los mRNA de esa misma especie. Así, se obtuvieron los posibles “targets” (Figura 17) de cada uno de los miRNA e isomiRNA seleccionados previamente en los mRNA de *Blattella germanica* con la clasificación “k-mer” del tipo de “target” y una calificación de la predicción del “target”. Por ello, posteriormente se filtraron las predicciones de los “targets” para eliminar aquellos “targets” con una calificación baja y obtener unos resultados más precisos con los que continuar los análisis.



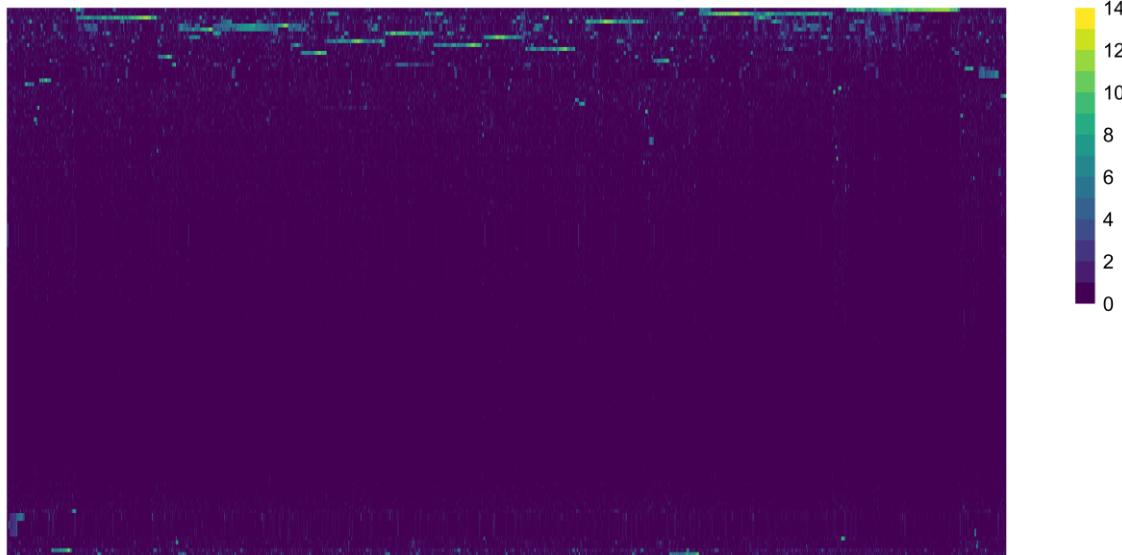
**Figura 17. Ejemplo gráfico de la predicción de los targets de los miRNA / isomiRNA.** La imagen muestra el resultado de la función “viewTargetAlignment” del paquete de R “scanMiR”, la cual representa una de las predicciones de “targets” realizadas para este trabajo.

#### **4.5. Análisis de las diferencias en los “targets” de los isomiRNA respecto a sus miRNA canónicos.**

Una vez obtenidos los “targets” de los isomiRNA y los miRNA, se pretendió comprobar si los cambios estructurales existentes en los isomiRNA provocan cambios funcionales en sus miRNA canónicos, así como caracterizar esos cambios funcionales por categorías de miRNA.

Para ello, en primer lugar, se sumarizaron las diferencias en los datos de los “targets” de los isomiRNA y sus miRNA canónicos, de manera que se compararon los “targets” de cada isomiRNA en un mRNA con los “targets” de su miRNA canónico en el mismo mRNA. Así, se contaron y sumaron el número

total de diferencias en cada mRNA en todos los isomiRNA de un mismo miRNA canónico. De este modo, se obtuvo una tabla de contingencia de los miRNA frente a los mRNA en la que las cuentas se corresponden al número de diferencias en los targets observadas en el conjunto de los isomiRNA de cada miRNA. A continuación, estos datos de diferencias de los “targets” se representaron en un heatmap (Figura 18) que enfrenta los mRNA y los miRNA, estando los datos escalados para cada uno de los mRNA.

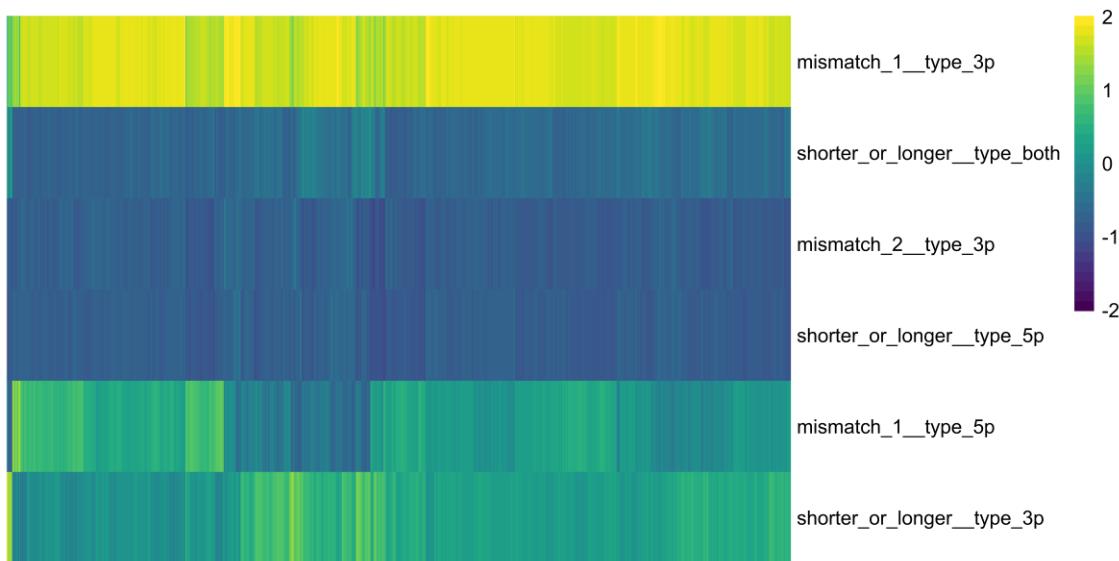


**Figura 18. Heatmap del número de las diferencias en los targets en el conjunto de los isomiRNA respecto a su miRNA canónico para cada uno de los mRNA.** La gráfica muestra el conteo de las diferencias de los targets del conjunto de los isomiRNA de un mismo miRNA canónico (filas) para cada mRNA (columnas), relativizados por cada mRNA.

Este heatmap muestra que la mayoría de los conjuntos de isomiRNA de cada miRNA canónico (“familias de isomiRNA”) no muestran unas exacerbadadas diferencias en sus “targets”, aunque prácticamente todas las “familias de isomiRNA” presentan algún cambio en los “targets” respecto a su miRNA canónico. Sin embargo, es cierto que hay algunas “familias de isomiRNA” que si que presentan grandes y consistentes cambios en los “targets” que presentan respecto a su miRNA canónico, las cuales se pueden observar en la parte superior del heatmap (Figura 18). Algunas de estas “familias de isomiRNA” con grandes cambios en sus “targets” son las asociadas a los miRNA canónicos: Mir-276-3p, Bantam-3p, Mir-1-3p, Mir-8-3p, Mir-10-P1-v1-5p, Mir-10-P1-v2-5p, Mir-9-o13-5p-5p, Mir-7\_5p, Mir-279-o36-3p, Mir-10-P2-5p, Mir-9-o12-5p, Mir-184-3p, Mir-750-3p, Mir-281-3p, y Mir-9-o14-5p. Así, con todos estos datos, se puede concluir que los cambios funcionales de los isomiRNA respecto a los miRNA dependen de cada miRNA concreto, habiendo desde “familias de isomiRNA” que provocan pocos cambios en los “targets” a “familias de isomiRNA” que provocan grandes cambios en los “targets”.

Posteriormente, se repitió este mismo análisis, pero, en este caso, agrupando las diferencias en los “targets” de los isomiRNA respecto a su miRNA en los mRNA a por el tipo de isomiRNA en lugar de por el miRNA canónico. Así, se obtuvieron los datos de las diferencias en los “targets” que provocan los isomiRNA agrupados por el tipo de isomiRNA. Estos datos se representaron mediante un heatmap en el que se enfrentaron los mRNA y los tipos de isomiRNA, escalando los datos para cada uno de los mRNA.

Así, en este heatmap de las diferencias de los “targets” según el tipo de isomiRNA se puede observar como las diferencias escaladas en los “targets” se acumulan principalmente en los isomiRNA de tipo polimórfico-3’ con un nucléotido diferente del miRNA canónico (mismatch\_1\_type\_3p). A continuación de este tipo de isomiRNA, los tipos de isomiRNA que presentan más cambios en los targets son los isomiRNA de longitud-3’ (shorter\_or\_longer\_type\_3p) y los isomiRNA polimórficos-5’ con un nucléotido diferente del miRNA canónico (mismatch\_1\_type\_5p). Por último, respecto a los otros tres tipos de isomiRNA estudiados, éstos son los que menos diferencias de “targets” presentan respecto a sus miRNA canónicos y respecto al resto de tipos de isomiRNA.



**Figura 19. Heatmap del número de las diferencias en los targets de los isomiRNA respecto a su miRNA canónico para cada uno de los mRNA y agrupados por la categoría de los isomiRNA.** La gráfica muestra el contejo de las diferencias de los targets del conjunto de los isomiRNA de un mismo miRNA canónico (filas) para cada mRNA (columnas), relativizados por cada mRNA.

Con todos estos datos, se puede llegar a la conclusión que las variaciones funcionales asociadas a los cambios de “targets” que provocan los isomiRNA respecto a los miRNA dependen enormemente tanto del miRNA en concreto como de los tipos de isomiRNA. Por ejemplo, los tipos de isomiRNA de longitud-5’, longitud 3’ y 5’, y polimórficos con dos nucleótidos diferentes respecto al miRNA canónico son los tipos de isomiRNA que menores diferencias de “targets” presentan respecto a sus miRNA canónicos (aunque también presentan diferencias), pero, a su vez, en resultados previos son los isomiRNA que más enriquecidos se encuentras en las etapas del desarrollo embrionario ED0, ED1 y ED2 de *Blattella germanica*. Esto puede ser debido a que los cambios de “targets” que provoquen estos tipos de isomiRNA sean más localizados y precisos respecto al resto de tipos de isomiRNA; o que no causen grandes cambios en los “targets” respecto al miRNA canónico, pero cambien la estabilidad del isomiRNA y su unión a los targets, como ya se ha visto en la bibliografía para algunos casos. Lo que es indudable es que la regulación génica de los miRNA y de los isomiRNA también se ve influenciada por las “familias de isomiRNA” y por el tipo de isomiRNA, estableciéndose correlaciones claras (como la anteriormente citada) con el desarrollo de *Blattella germanica*.

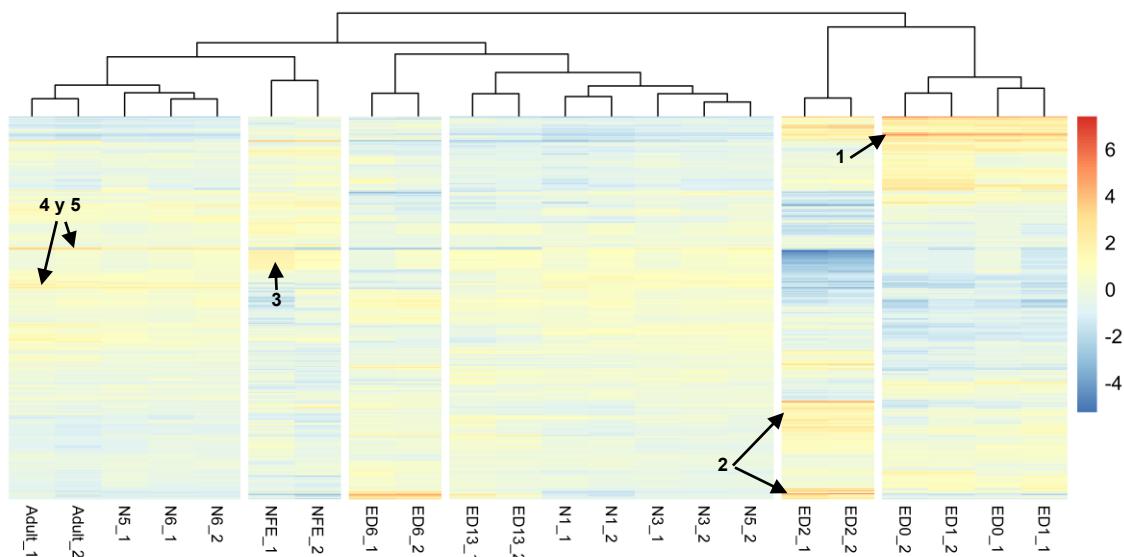
#### **4.6. Análisis gráfico y de enriquecimiento de la extrapolación de las cuentas de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de *Blattella germanica*.**

A pesar de ya haber observado en los resultados anteriores la importancia de algunos isomiRNA y algunos de los tipos de éstos en las diferencias de “targets” respecto a sus miRNA canónicos, todavía no queda claro qué efectos funcionales provocan sobre la regulación génica y fisiológica en las distintas etapas del desarrollo de *Blattella germanica*. Por ello, como parte final de este trabajo, se intentó realizar un análisis de enriquecimiento funcional mediante los “targets” predichos, los mRNA y las cuentas de los isomiRNA y miRNA de *Blattella germanica*. Para ello se crearon tablas de contingencia que enfrentaron las anotaciones de los mRNA y los distintos estadios del desarrollo con los datos de la suma de las cuentas de los isomiRNA, los miRNA o ambos, que poseen un “target” en cada mRNA (“cuentas de inhibición”).

Sin embargo, es necesario aclarar que la aproximación seguida en este análisis es solamente un modelaje aproximado, ya que se necesitaría de mucha más información biológica para poder desarrollar este modelaje de forma precisa. En primer lugar, esto es debido a que los “targets” de los isomiRNA y miRNA son predicciones bioinformáticas no corroboradas, por lo que sería necesaria comprobar cuáles de estos targets son reales y cuáles no. En segundo lugar, en el siguiente análisis se realiza una aproximación por la cual la suma de las cuentas de los miRNA y/o isomiRNA en una etapa del desarrollo concreta y que presentan un target para un mRNA concreto equivale a las “cuentas de inhibición” de ese mRNA por parte de los miRNA y/o isomiRNA en esta etapa concreta del desarrollo. Sin embargo, esto no es real, ya que, como se ha revisado en la introducción de este trabajo, los miRNA presentan una gran diversidad de mecanismos de acción, los cuales, muchos de ellos todavía no se conocen en profundidad. Tanto es así que algunos de los isomiRNA, aunque se les detecte bioinformáticamente cambios de “targets”, puede que estos no sean reales y en realidad el isomiRNA provoque una ganancia o pérdida de función. Además, la unión de los miRNA e isomiRNA al mRNA no solo depende de la secuencia “seed”, sino de muchas más características biológicas. Por ello sería necesario disponer de modelos bioinformáticos que predigan, no solo los “targets” de los miRNA e isomiRNA en los mRNA, sino los posibles efectos cuantitativos de los miRNA e isomiRNA sobre la expresión génica de los mRNA. Lamentablemente, no se disponen ni de estos modelos ni de estos datos en la bibliografía científica actual para el organismo *Blattella germanica*. Por ello, se realiza esta aproximación mediante “cuentas de inhibición”, las cuales, si se dispusiese de esos datos, luego habría que multiplicar por los efectos concretos de cada miRNA o isomiRNA sobre la expresión de un mRNA. A pesar de ello, se cree que este modelo aproximado puede aportar una idea generalizada de la función de los miRNA e isomiRNA, así como sus diferencias, sobre la regulación génica y el desarrollo de *Blattella germanica*.

Para llevar a cabo este modelo aproximado, a partir de la tabla de contingencia citada previamente de las anotaciones de mRNA contra las muestras de los estados del desarrollo y completada con la suma de las cuentas de los miRNA y/o isomiRNA que poseen un “target” para los mRNA (“cuentas de inhibición”), se obtuvieron heatmaps de distancias para los datos procedentes de los isomiRNA, miRNA y el total.

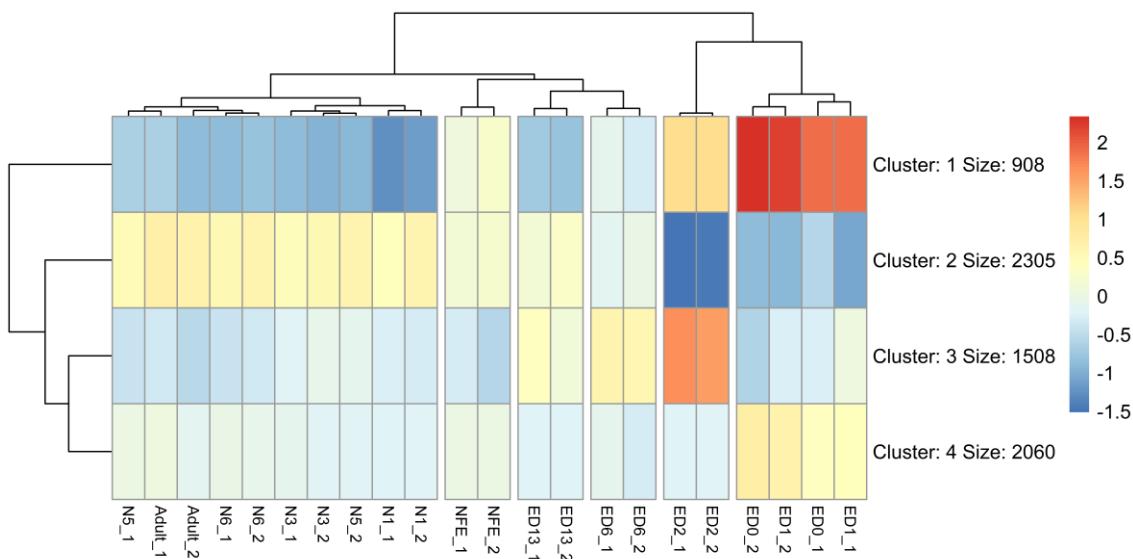
En primer lugar, el heatmap de las “cuentas de inhibición” de los miRNA sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica* (Figura 20) muestra un resultado semejante al heatmap de los miRNA en las distintas etapas del desarrollo (Figura 11). Exactamente, en la Figura 20 se observa un grupo de mRNA enriquecidos en “cuentas de inhibición” mayoritariamente en las etapas ED0 y ED1 del desarrollo y algo menos para la etapa ED2, al igual que el grupo 1 mostrado en la Figura 11. También se observa un grupo de mRNA enriquecidos en “cuentas de inhibición” principalmente para las etapas ED2 y ED6, asemejándose al grupo 2 mostrado en la Figura 11. Por otro lado, se observa también un grupo de mRNA levemente enriquecido en “cuentas de inhibición” en el huevo no fecundado (parecido al grupo 3 mostrado en la Figura 11) y un pequeño grupo de mRNA progresivamente enriquecido en “cuentas de inhibición” desde las etapas de ninfa hasta el adulto (como una combinación de los grupos 4 y 5 de la Figura 11). Con todo ello, aunque diferentes, este heatmap muestra una distribución similar de “cuentas de inhibición” por los miRNA en los mRNA a la distribución de cuentas de miRNA en las etapas del desarrollo, más en las embrionarias.



**Figura 20. Heatmap de distancias de las “cuentas de inhibición” por parte de los miRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por los miRNA en los distintos mRNA (filas) en las distintas muestras (columnas), organizándose gráficamente según esas “cuentas de inhibición” en las muestras (columnas). Las flechas en el gráfico indican los grupos en el texto.

Al observarse la presencia de diferencias correlacionadas con las etapas del desarrollo en el heatmap previo de “cuentas de inhibición” por parte de los miRNA, se procedió a realizar un heatmap clusterizado de esos datos para comprobar las observaciones realizadas (Figura 20). Ese heatmap clusterizado muestra un cluster de mRNA enriquecido en “cuentas de inhibición” en las etapas ED0 y ED1 (cluster 4); y otro en ED0 y ED1 mayoritariamente y en menor medida en las etapas ED2 y huevo no fecundado (cluster 1), lo cual se asemeja al grupo 1 de la Figura 11. Además, se observa un cluster de mRNA principalmente enriquecido en “cuentas de inhibición” en la etapa ED2 que desciende progresivamente en las etapas ED6 y ED13 (cluster 2), al igual que el grupo 2 de la Figura 11. Finalmente, se puede observar también un cluster de mRNA enriquecido en “cuentas de inhibición” en las etapas de ninfa y adulto

(cluster 2), al igual que los grupos 4 y 5 de la Figura 11. Así, se confirma que la regulación negativa que ejercen los miRNA sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica* también presenta una correlación.

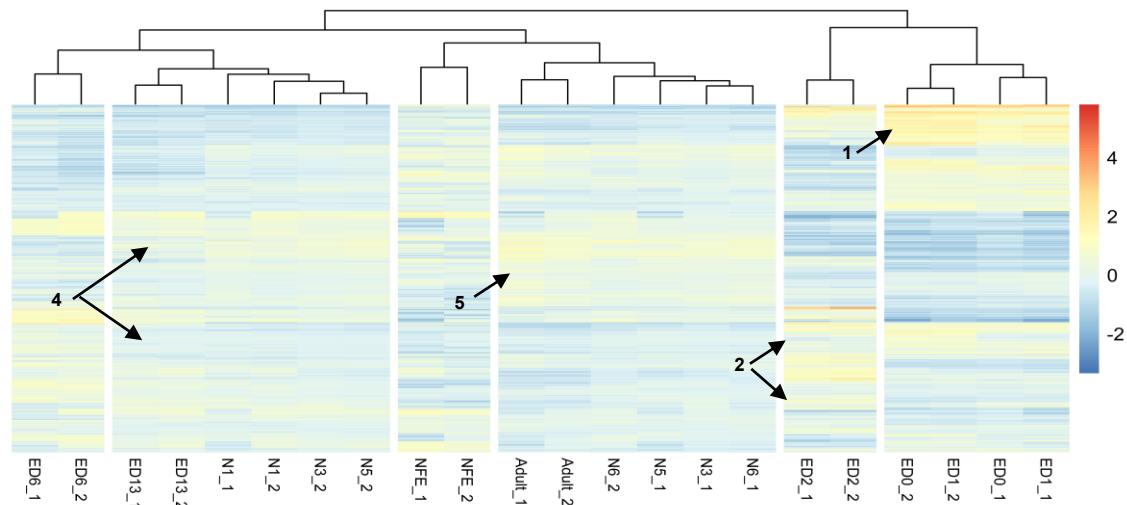


**Figura 21. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los miRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por parte de los miRNA en los distintos mRNA clusterizados (filas) en las distintas muestras (columnas), de modo que se organizan gráficamente según esas “cuentas de inhibición” en las muestras (columnas) y en los clusteres de mRNA (filas).

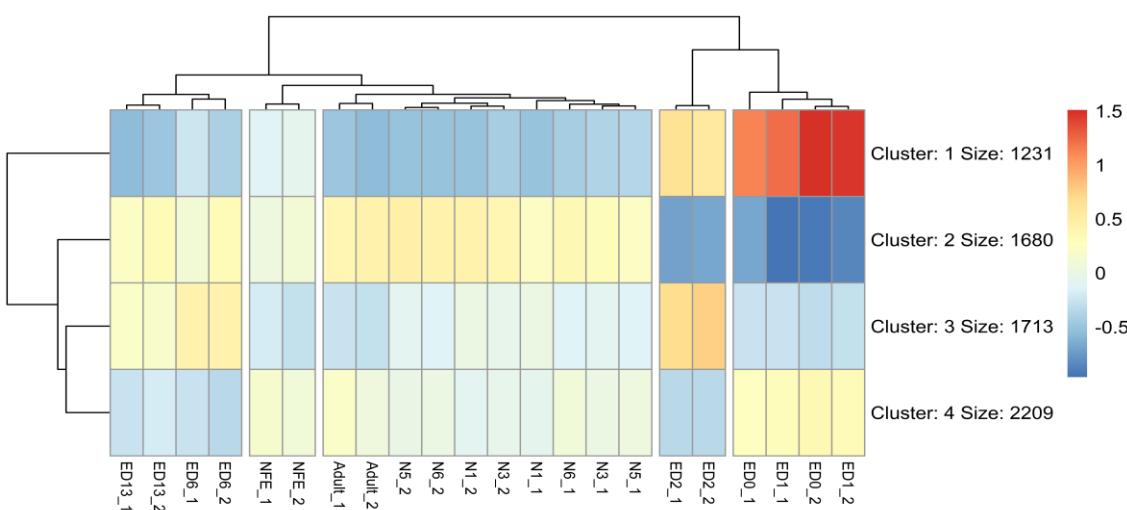
Una vez comprobada la relación entre la regulación negativa de los miRNA sobre los mRNA y las etapas del desarrollo de *Blattella germanica*, se continuó con la comprobación de esta misma relación con la regulación negativa de los isomiRNA. Así, el heatmap de distancias de las “cuentas de inhibición” de los isomiRNA sobre los mRNA en las etapas del desarrollo (Figura 22) muestra un grupo de mRNA enriquecidos negativamente en las etapas ED0, ED1 y levemente en ED2 (grupo 1) que se parece al grupo 1 de la Figura 12. Además, se observa un grupo de mRNA enriquecidos en “cuentas de inhibición” en la etapa ED2 y también, pero algo menos en la etapa ED6 (grupo 2), con lo que se asemeja al grupo 2 de la Figura 11. También se observa un grupo de mRNA enriquecidos en “cuentas de inhibición” en las etapas finales de ninfa y en el adulto (grupo 5), lo cual sería similar al grupo 5 de la Figura 11. Por último, se observa un grupo de mRNA enriquecidos en “cuentas de inhibición” en ED13 y las primeras etapas de ninfa (grupo 4), pareciéndose al grupo 4 de la Figura 11. De esta manera, este gráfico también parece indicar que la regulación negativa de los isomiRNA sobre los mRNA se correlaciona con las etapas del desarrollo de *Blattella germanica* y, más aún, en las etapas embrionarias.

Para comprobar las observaciones realizadas en el anterior heatmap, y como en el caso anterior de los miRNA, se obtiene un heatmap clusterizado de la regulación negativa de los isomiRNA sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica* (Figura 23). En ese heatmap clusterizado se puede observar un cluster de mRNA enriquecido en “cuentas de inhibición” en las etapas del desarrollo ED0 y ED1, y algo menos en la etapa ED2 (cluster 1), por lo que se parece al grupo 1 de la Figura 11. También se observa un cluster de mRNA enriquecido en “cuentas de inhibición” en las etapas ED0 y

ED1 mayoritariamente, pero que luego aparece progresivamente desde las etapas de ninfa hasta el huevo no fecundado (cluster 4), pasando por el adulto. Aparte de ello, también se observa un cluster de mRNA enriquecido en “cuentas de inhibición” en la etapa ED2 y un poco menos en las etapas ED6 y ED13 (cluster 3), por lo que se parece al grupo 2 de la Figura 11. Además, también se observa un cluster de mRNA enriquecido en “cuentas de inhibición” desde la etapa ED6 hasta el adulto y levemente en el huevo no fecundado, el cual se asemeja a los grupos 3 y 5 de la Figura 11. De esta forma, mediante este heatmap se confirma que la regulación negativa que ejercen los isomiRNA sobre los mRNA se correlaciona con las etapas del desarrollo de *Blattella germanica*, relacionándose aún más con las etapas embrionarias del mismo.



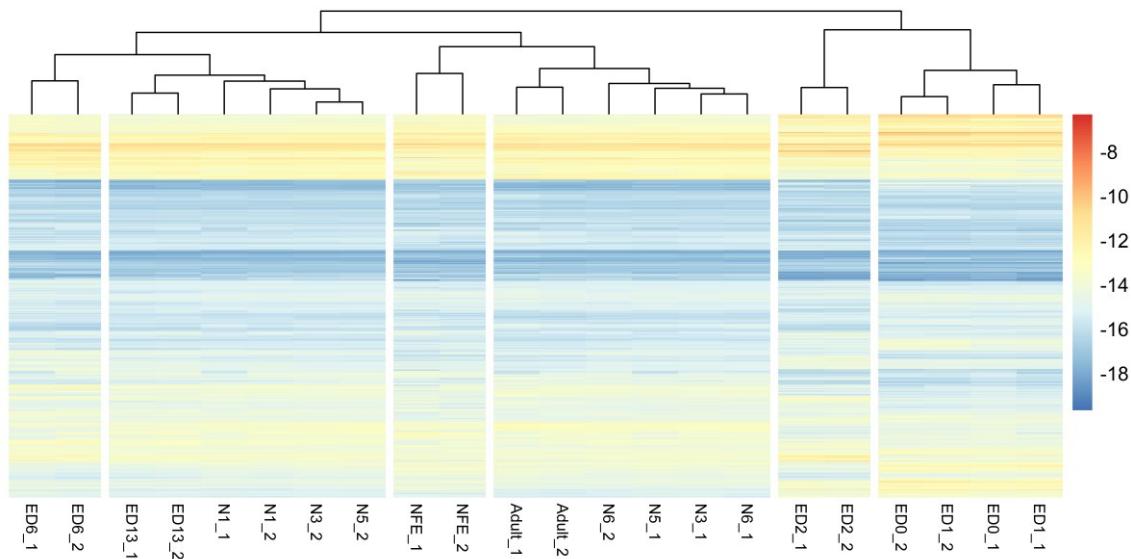
**Figura 22. Heatmap de distancias de las “cuentas de inhibición” por parte de los isomiRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por parte de los isomiRNA en los distintos mRNA (filas) en las distintas muestras (columnas), de modo que se organizan gráficamente según esas “cuentas de inhibición” en las muestras (columnas). Las flechas en el gráfico indican los grupos citados en el texto.



**Figura 23. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por parte de los isomiRNA en los distintos mRNA clusterizados (filas) en las distintas muestras (columnas), de modo que se organizan gráficamente según esas “cuentas de inhibición” en las muestras (columnas) y en los clusteres de mRNA (filas).

Para completar el análisis de la regulación génica de los isomiRNA y miRNA sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica*, se volvió a realizar el análisis previo, pero con las “cuentas de inhibición” de los isomiRNA y los miRNA sobre los mRNA en el desarrollo de *Blattella germanica*.

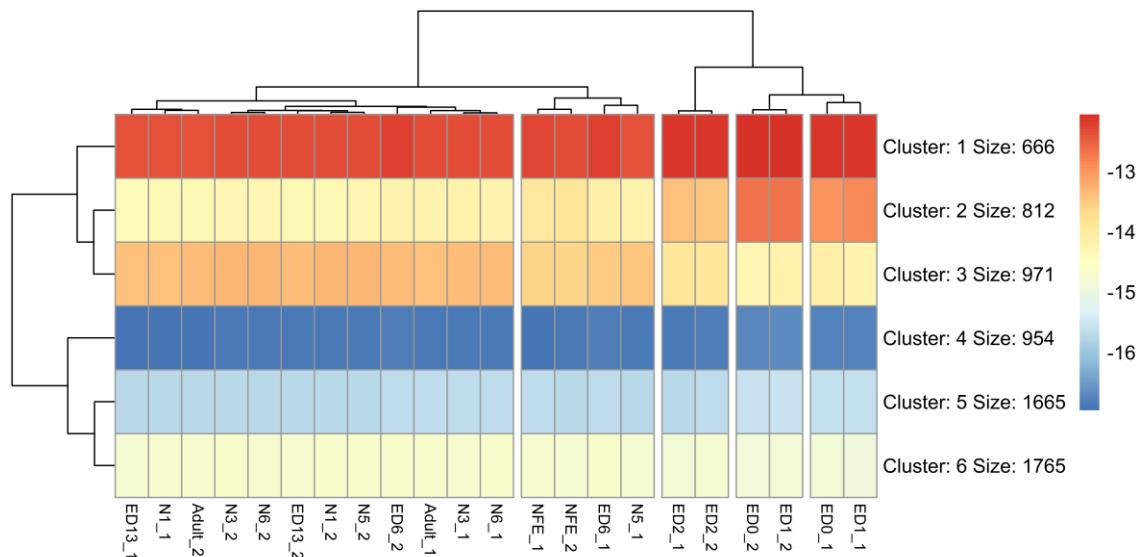
De este modo, el heatmap de distancias de las “cuentas de inhibición” por los isomiRNA y los miRNA sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica* muestra una distribución muy diferente respecto a los heatmaps de distancias anteriores. Exactamente, se observan varios grupos de mRNA igualmente enriquecidos o no en “cuentas de inhibición” por parte de los isomiRNA y los miRNA, por lo que no varían entre las distintas etapas del desarrollo. Sin embargo, sí que se observan ciertos grupos de mRNA enriquecidos en “cuentas de inhibición” por los isomiRNA y miRNA y que se encuentran asociados a las etapas de huevo no fecundado, ED0, ED1 y ED2 o a las etapas posteriores hasta el adulto. De este modo, este heatmap podría indicar un tipo de regulación génica por parte de los isomiRNA y miRNA desde el huevo no fecundado hasta la etapa embrionaria ED2, y otro tipo de regulación génica desde la etapa ED6 hasta la etapa adulta. Sin embargo, también podría indicar una regulación más precisa por parte de los miRNA e isomiRNA que no es observada correctamente en este heatmap.



**Figura 24. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los distintos mRNA clusterizados (filas) en las distintas muestras (columnas), de modo que se organizan gráficamente según esas “cuentas de inhibición” en las muestras (columnas) y en los clusteres de mRNA (filas).

Así, a continuación, tras no obtener unos datos concluyentes en el heatmap de distancias anterior, se comprobó el efecto de regulación negativa del conjunto de isomiRNA y miRNA sobre los mRNA en las distintas etapas de *Blattella germanica* mediante un heatmap clusterizado, el cual, en este caso, también muestra un resultado muy diferente a los mostrados previamente. Exactamente, en este heatmap clusterizado se pueden observar dos clusteres contrarios, un cluster de mRNA enriquecido en “cuentas de inhibición” por isomiRNA y miRNA que se presenta principalmente en las etapas de ED0 a ED2 (cluster 2) y otro que se presenta principalmente en el resto de las etapas

del desarrollo (cluster 3). Además, aparecen otro dos clusteres parecidos a los dos anteriores, pero con menores diferencias en el enriquecimiento en “cuentas de inhibición” entre los distintos grupos de etapas del desarrollo: el cluster 1 que sería el homólogo del cluster 2 y el cluster 4 que sería el homólogo del cluster 3. En cambio, los otros dos clusteres que aparecen en este heatmap clusterizado no presentan cambios aparentes entre las distintas etapas del desarrollo. De esta manera, todos estos datos parecen indicar que la suma de los efectos de regulación negativa sobre los mRNA por parte de los isomiRNA y los miRNA provoca la aparición de un set de regulación génica para los estadios del desarrollo embrionario de ED0 a ED2 y otro set de regulación génica para el resto de los estadios del desarrollo de *Blattella germanica* (desde ED6 hasta adulto pasando por ninfa y huevo).



**Figura 25. Heatmap de distancias clusterizado de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA sobre los mRNA en las etapas del desarrollo de *Blattella germanica*.** En esta gráfica se observan los niveles de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los distintos mRNA clusterizados (filas) en las distintas muestras (columnas), de modo que se organizan gráficamente según esas “cuentas de inhibición” en las muestras (columnas) y en los clusteres de mRNA (filas).

Con todos estos datos, se puede concluir que la regulación génica sobre los mRNA en el desarrollo de *Blattella germanica* por parte de la suma de los isomiRNA y los miRNA despliega un programa de regulación génica para la etapa embrionaria y otro para la etapa de ninfa y adulta. Sin embargo, cuando se desengranan estos sets de regulación génica en la regulación individual de los isomiRNA y los miRNA, se comprueba que cada uno de ellos regulan prácticamente todos los distintos estados del desarrollo. En concreto, tanto los isomiRNA como los miRNA regulan mayoritariamente los distintos estados del desarrollo embrionario, interviniendo más en ello los isomiRNA que los miRNA.

#### 4.7. Análisis de clusteres funcionales de la extrapolación de las cuentas de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de *Blattella germanica*.

Tras observar las diferencias estructurales en la regulación génica inhibitoria por parte de los isomiRNA y los miRNA en los mRNA entre las distintas etapas del desarrollo de *Blattella germanica*, se procedió a intentar relacionar esas diferencias con sus correspondientes diferencias funcionales. Para ello se

realizó un análisis de clusteres funcionales mediante los GO ID de los mRNA de *Blattella germanica*, los términos GO asociados a esos GO ID y los datos previos sobre las “cuentas inhibitorias” de los isomiRNA, miRNA y ambos sobre los mRNA en las distintas etapas del desarrollo de *Blattella germanica*.

De esta manera, este análisis de clusteres funcionales se realizó tomando como referencia los estados del desarrollo huevo no fecundado, ED0 y adulto para los isomiRNA, los miRNA y ambos juntos. Como resultado de ellos se obtuvieron los términos GO enriquecidos (es decir, más inhibidos al ser las “cuentas inhibitorias” lo que se encuentra enriquecido) o empobrecidos (menos inhibidos por el razonamiento previo) en cada estado del desarrollo en comparación con el estado del desarrollo de referencia (Código VII).

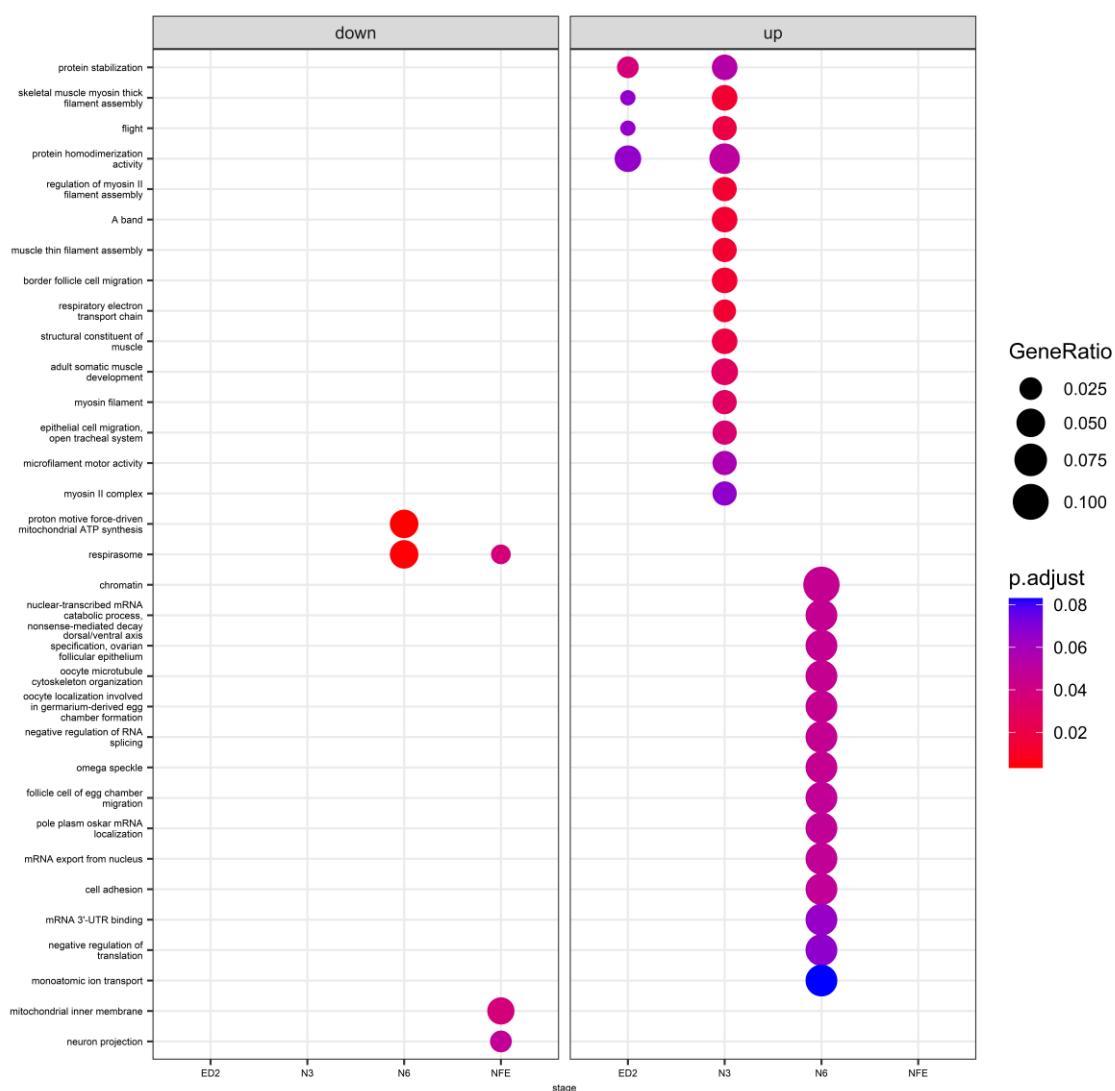
Así, resumiendo los dotplots sobre las comparaciones más relevantes de los clústeres funcionales de las “cuentas inhibitorias” por parte de los miRNA en los mRNA de las distintas etapas del desarrollo de *Blattella germanica* respecto a las etapas de huevo (Anexo IV), ED0 (Anexo V) y adulto (Figura 26), los resultados obtenidos indican:

- ED0: presenta una mayor inhibición del ribosoma y una menor inhibición de las uniones enzimáticas respecto al huevo no fecundado. Además, también presenta una mayor inhibición de las uniones enzimáticas respecto a otros estados (ED6, ED13 y N3).
- Adulto: presenta una mayor inhibición de la proyección neural respecto al huevo no fecundado y de la mitocondria respecto al huevo no fecundado y a N6. También presenta una menor inhibición de la biogénesis de proteínas, la función muscular, el sistema reproductivo, el vuelo y la mitocondria respecto a estados del desarrollo como ED2 y N3, además de la adhesión celular, la regulación de la transcripción y el sistema reproductivo respecto a N6.
- NFE (huevo no fecundado): presenta una menor inhibición de la proyección neuronal respecto a N5, N6 y el adulto. Además, también presenta una menor inhibición de la mitocondria respecto al adulto, del remodelaje del citoplasma y el citoesqueleto respecto a N6 y de la regulación de procesos biológicos respecto a N5.

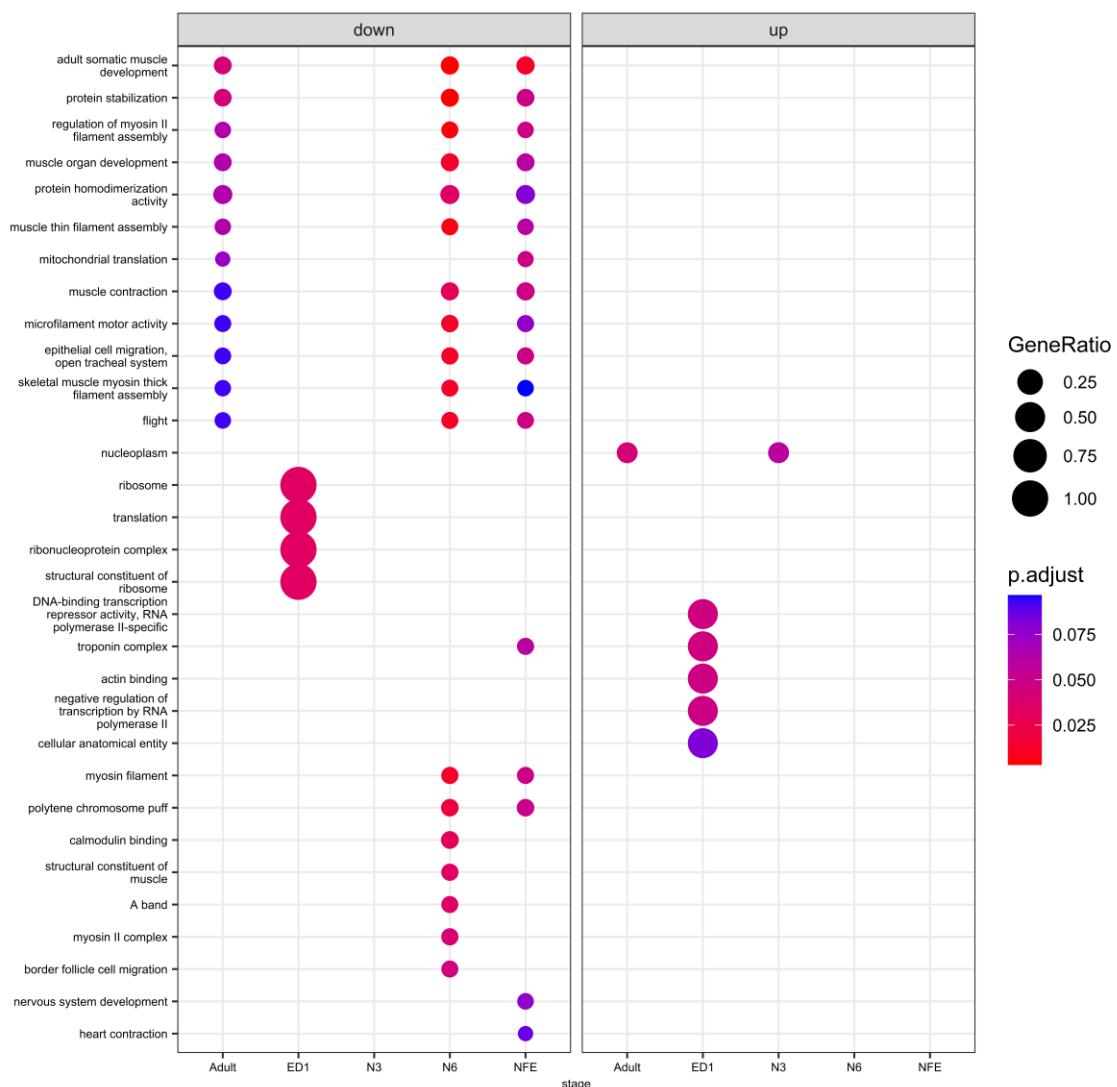
Por otro lado, también se resumieron los dotplots que comparan los clusteres funcionales de las “cuentas inhibitorias” por los isomiRNA en los mRNA durante el desarrollo de *Blattella germanica*, también respecto a las etapas de huevo (Anexo VI), ED0 (Figura 27) y adulto (Anexo VII). Así, el resumen de los resultados más relevantes muestra:

- ED0: aparece una mayor inhibición de la biogénesis de proteínas, el desarrollo y la función muscular, y el vuelo respecto a los estados de huevo no fecundado, N6 y adulto. Además, también aparece una mayor inhibición del desarrollo del sistema nervioso frente al huevo no fecundado; el ribosoma y la traducción frente a ED1; la reproducción frente a N6, y la traducción mitocondrial frente al adulto. Por otro lado, presentan una menor inhibición de la identidad celular anatómica y la regulación de la transcripción respecto a ED1; y del nucleoplasma respecto a N3 y la etapa adulta.

- Adulto: presenta una mayor inhibición del metabolismo respecto a la etapa de huevo y ED1; de la traducción respecto a ED2 y ED6; de la mitocondria respecto a ED1; y de la actividad antioxidante respecto a N5. Por otro lado, presenta una menor inhibición del desarrollo muscular y la biogénesis de proteínas respecto a ED1 y ED2. Además, también presenta una menor inhibición del tejido muscular, del vuelo, del proceso de biosíntesis de la hormona juvenil y de la reproducción respecto a ED2.
- NFE: presenta una menor inhibición del metabolismo catabólico respecto a las etapas ED2, N3, N5, N6; del desarrollo del sistema nervioso respecto a ED0; de la biogénesis de proteínas y el desarrollo muscular respecto a ED0 y ED1



**Figura 26. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa adulta.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa adulta. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-value de la comparación.

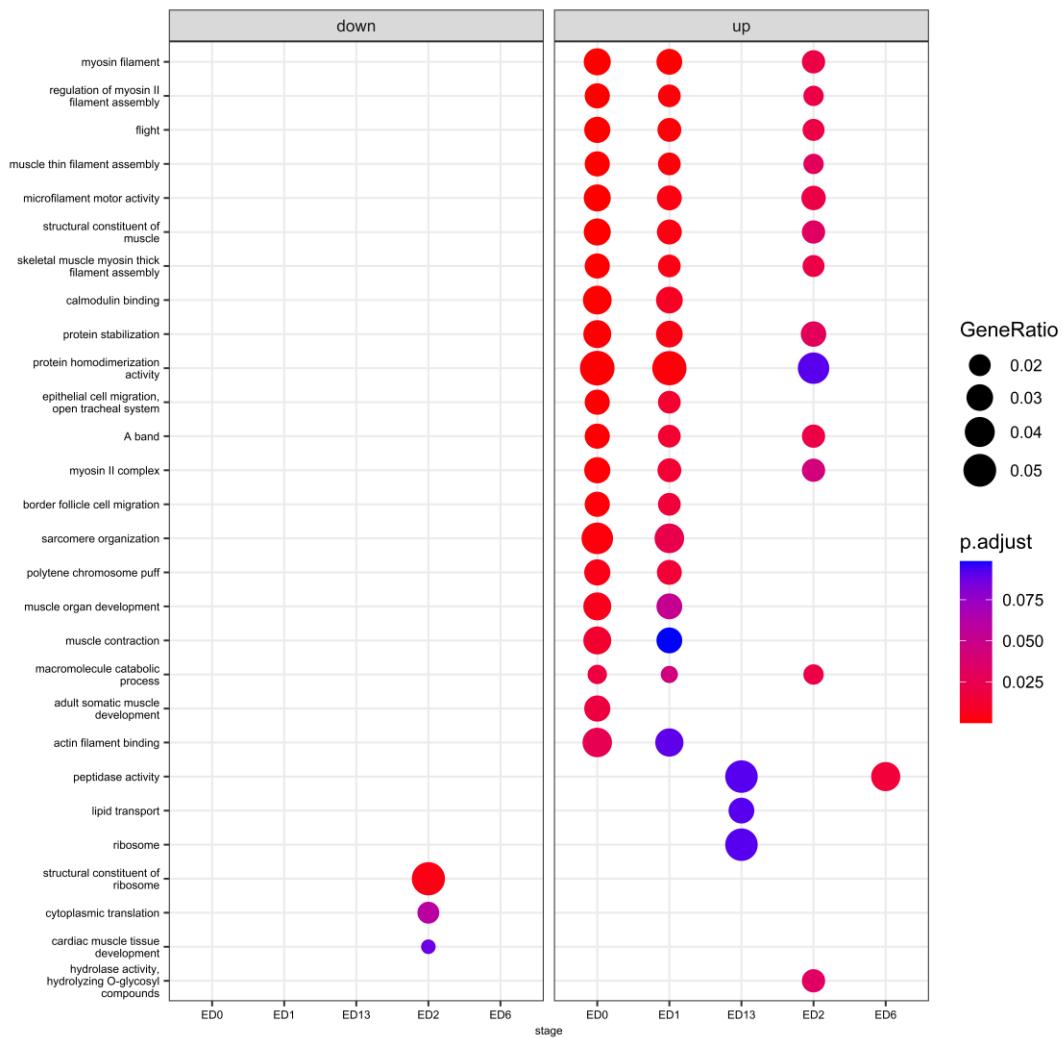


**Figura 27. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa ED0.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa ED0. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

Finalmente, se sumarizaron los dotplots sobre las comparaciones de los clústeres funcionales de las “cuentas inhibitorias” tanto procedentes de los isomiRNA como de los miRNA durante el desarrollo de *Blattella germanica*, igualmente respecto a las etapas de huevo (Figura 28), ED0 (Anexo VIII) y adulto (Anexo IX). De esta manera, los resultados más relevantes obtenidos son:

- ED0: aparece una mayor inhibición del metabolismo catabólico, de la biogénesis de proteínas, del tejido muscular y del vuelo respecto a las etapas del desarrollo NFE, ED6, N1, N3, N5, N6 y adulto. Además, también presenta una mayor inhibición de la reproducción, respecto a NFE, N6 y adulto.

- Adulto: presenta una mayor inhibición del ribosoma respecto a ED2 y ED6; de las dinámicas citoplasmáticas respecto a ED2; y del metabolismo catabólico respecto a ED13. Además, también presenta una menor inhibición de la biogénesis de proteínas, el tejido muscular, el vuelo y la reproducción respecto a las etapas ED0, ED1 Y ED2.
- NFE: aparece una menor inhibición del vuelo, el tejido muscular, la biogénesis de proteínas, la reproducción y el desarrollo muscular respecto a ED0 Y ED1; del ribosoma, transporte lipídico y peptidasas respecto a ED13.



**Figura 28. Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa NFE.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa NFE. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

#### **4.8. Interpretación de los resultados funcionales de los isomiRNA y miRNA con sus “targets” en los mRNA entre las distintas etapas del desarrollo de *Blattella germanica*.**

Así, una vez resumidas las comparaciones de clusteres funcionales obtenidas en los dotplots de miRNA, isomiRNA y ambos respecto a los estadios del desarrollo huevo no fecundado, ED0 y adulto, se integraron estas diferencias en el desarrollo de *Blattella germanica*.

En cuanto a los datos de las comparaciones de clusteres funcionales de los miRNA, estos indican que los huevos no fecundados presentan una baja inhibición de la proyección neural, de la mitocondria, del remodelaje del citoplasma y el citoesqueleto, y de la regulación de procesos biológicos. De todo ello se extrae que las diferencias en los clusteres funcionales de los miRNA detectados en este estadio (NFE) son fundamentales para el comienzo del desarrollo por tratarse de procesos celulares básicos necesarios para ello. En cambio, los datos de isomiRNA en los huevos no fecundados indican una menor inhibición del metabolismo catabólico y del desarrollo del sistema nervioso y muscular, por lo que las diferencias en los clusteres funcionales de los isomiRNA en el huevo no fecundado también parecen estar indicados al comienzo del desarrollo del organismo, pero, en este caso, a preparar la organización a un nivel más tisular que los miRNA. Por último, en las comparaciones de los clusteres funcionales obtenidos por la suma de los isomiRNA y miRNA en el estadio NFE se encuentran menos inhibidos las funciones asociadas al tejido muscular, el vuelo, la biogénesis de proteínas, la reproducción, el desarrollo muscular, al ribosoma, al transporte lipídico y la actividad peptidasa. En este último caso, las funciones obtenidas menos inhibidas podrían asociarse a un resto del organismo previo (la hembra), ya que, a excepción del ribosoma, el transporte lipídico y la actividad peptidasa, solo son comparaciones significativas con ED0 y ED1. Por todo ello, se intuye que los miRNA podrían permitir la mayor expresión de procesos celulares y los isomiRNA la preparación de la organización espacial en el huevo no fecundado para poder comenzar el desarrollo en cuanto se dé la fecundación, en comparación con el resto de etapas del desarrollo del insecto. Además, es necesario resaltar que el nivel de inhibición por parte de los miRNA en el huevo no fecundado es considerablemente menor que en el resto de las etapas del desarrollo de *Blattella germanica*.

En cuanto a las comparaciones de clusteres funcionales de los miRNA en la etapa ED0, estos muestran una mayor inhibición del ribosoma respecto al estado NFE y de las uniones enzimáticas en numerosos estadios distintos del desarrollo. De esta forma, los miRNA también actuarían a un nivel más infracelular en el estadio ED0. Por otro lado, las comparaciones de clusteres funcionales de los isomiRNA en ED0 muestran una mayor inhibición de la biogénesis de proteínas, el desarrollo y función muscular, el desarrollo nervioso, la reproducción, y el vuelo frente a varios estadios; y también muestran una menor inhibición de la identidad celular anatómica y de la regulación génica respecto a varios estadios del desarrollo. Además, las comparaciones de clusteres funcionales totales de isomiRNA y miRNA muestran resultados muy aproximados al de los isomiRNA. Así, en general, los isomiRNA regulan procesos celulares y fisiológicos, por lo que, al igual que en el estadio NFE, los isomiRNA regulan procesos de un orden superior a los

miRNA en ED0. Aparte de ello, hay que recalcar que, contrariamente a lo observado en el estado de huevo no fecundado, en este estadio (ED0) el nivel de inhibición de los isomiRNA es superior a la mayoría del resto de estadios del desarrollo de *Blattella germanica* y a la de miRNA en este estadio.

Finalmente, se exploraron las comparaciones de clusteres funcionales de los miRNA en la etapa adulta, los cuales muestran una mayor inhibición de la proyección neural y la mitocondria respecto a otros estados; y una menor inhibición de la biogénesis de proteínas, la función muscular, el sistema reproductivo, y el vuelo respecto a otros estados. De esta manera, en el adulto, los miRNA parecen presentar una mayor inhibición de procesos celulares de desarrollo y una menor inhibición de procesos fisiológicos de mantenimiento. Por otro lado, en las comparaciones de clusteres funcionales de los isomiRNA en la etapa adulta, los isomiRNA presentan una mayor inhibición del metabolismo, de la traducción, de la mitocondria y de la actividad antioxidante respecto a algunos estadios del desarrollo. Además, los isomiRNA también presentan una menor inhibición del desarrollo muscular, la biogénesis de proteínas, el tejido muscular, el vuelo, la biosíntesis de la hormono juvenil (implicada en la metamorfosis) y la reproducción. Así, en el adulto, los isomiRNA se comportan, al igual que los miRNA, con una mayor inhibición de procesos celulares de desarrollo y mantenimiento y una menor inhibición de procesos fisiológicos de mantenimiento. Posteriormente, se procedió a explorar los resultados de las comparaciones de clusteres funcionales del total los isomiRNA y los miRNA en la etapa adulta, pudiéndose observar como se presenta una mayor inhibición del ribosoma, las dinámicas citoplasmáticas, y el metabolismo catabólico; y una menor inhibición de la biogénesis de proteínas, del tejido muscular, el vuelo y la reproducción respecto a algunas etapas del desarrollo. Con estos resultados, se puede comprobar cómo, en este caso, independientemente de si se trata de miRNA, isomiRNA o ambos, las mayores inhibiciones se producen sobre procesos celulares de mantenimiento y las menores inhibiciones se refieren a procesos fisiológicos de mantenimiento. Además, es necesario recalcar que en esta etapa adulta también (al igual que en NFE) se da una menor regulación génica mediante el aumento de la inhibición respecto a otras etapas tanto por miRNA como por isomiRNA.

Con todos estos resultados, se pueden obtener varias conclusiones del análisis funcional asociado a las “cuentas de inhibición” durante el desarrollo. En primer lugar, los isomiRNA provocan un aumento de la regulación génica funcional en las etapas embrionarias respecto al resto de etapas del desarrollo. En cambio, los miRNA no muestran una distribución funcional clara a lo largo del desarrollo. Sin embargo, la suma de ambos muestra un claro incremento de la regulación génica funcional mediante miRNA e isomiRNA en las etapas embrionarias del desarrollo embrionario. Por otro lado, este mayor aumento de la represión génica durante las fases embrionarias por parte, sobre todo de los isomiRNA, se debe sobre todo a la represión de mRNA asociados a clusters funcionales fisiológicos (vuelo, función muscular o reproducción) y a la permisividad de su expresión durante el resto de las etapas del desarrollo. Sin embargo, durante algunas de las etapas de ninfa, también son los miRNA los que reprimen estos clusters funcionales fisiológicos y morfológicos. En cambio, los clústeres funcionales infracelulares (proyección neuronal, mitocondria, o

actividades enzimáticas) son regulados tanto por miRNA como por isomiRNA, aumentando su represión génica durante las etapas no embrionarias.

En resumen, funcionalmente se produce un aumento de la represión génica provocada por los miRNA e isomiRNA (sobre todo estos últimos) de los clusters funcionales fisiológicos y morfológicos en las etapas embrionarias, la cual se reduce en las etapas no embrionarias. Además, se produce un efecto contrario con los clusters funcionales infracelulares (más reprimidos en las etapas no embrionarias), en este caso, tanto por los miRNA como por los isomiRNA.

## 5. Discusión

El principal objetivo con el que partía este estudio era identificar y caracterizar los isomiRNA involucrados en el desarrollo de insectos, exactamente en *Blattella germanica* y, con los resultados expuestos previamente, este objetivo se considera cumplido.

En primer lugar, se anotaron eficazmente los miRNA e isomiRNA presentes en cada una de las etapas del desarrollo de *Blattella germanica*, obteniendo también las secuencia y cuentas de estos. Posteriormente, se realizó una exploración gráfica de las diferencias en las cuentas de las anotaciones de los isomiRNA y los miRNA presentes en cada etapa del desarrollo de *Blattella germanica*. Esta exploración permitió obtener indicios acerca del enriquecimiento de ciertos clusteres de miRNA e isomiRNA en ciertas etapas del desarrollo. Aproximadamente, se observaron clusteres de miRNA e isomiRNA enriquecidos en las etapas: NFE; de ED0 a ED2 más NFE; de ED2 a ED13; de ED13 a la primera mitad de las etapas de ninfa; y de la segunda mitad de las etapas de ninfa hasta la etapa adulta. Para estudiar cuantitativamente estas diferencias se realizaron análisis de enriquecimiento sobre esos datos de isomiRNA y miRNA, los cuales demostraron que tanto en isomiRNA como en miRNA, existen un mayor número de anotaciones significativas entre las etapas embrionarias y NFE respecto a las etapas de ninfa y adulta. De esta manera, se demostró que tanto los isomiRNA como los miRNA ejercen su regulación génica preferentemente entre las etapas embrionarias del desarrollo de *Blattella germanica*. Además, esos resultados también revelaron que la regulación génica ejercida preferentemente sobre las etapas embrionarias respecto al resto de etapas del desarrollo es mayor a través de los isomiRNA que a través de los miRNA. Estos resultados concuerdan con los publicados por Ylla *et al.* (7), en los cuales demostraron la existencia de ciertas olas de miRNA que regulaban las etapas del desarrollo de *Blattella germanica*, sobre todo las etapas embrionarias. Sin embargo, este es el primer indicio acerca del mismo mecanismo de acción por parte de los isomiRNA, aunque levemente diferente al estar aún más enfocado en las etapas embrionarias.

Debido a este papel preferente en la regulación génica de las etapas embrionarias de *Blattella germanica* por parte de los isomiRNA, se estudiaron las diferencias existentes entre los distintos tipos de isomiRNA en las etapas del desarrollo del insecto. Así, la exploración gráfica de esta aproximación demostró: que los isomiRNA polimórficos 3' con dos nucleótidos distintos están enriquecidos en las etapas NFE, ED0 y ED1 del desarrollo; que los isomiRNA de tipo longitud-5' están enriquecidos en las etapas ED0, ED1 y ED2 del

desarrollo; que los isomiRNA de tipo longitud 3' y 5' están enriquecidos en las etapas desde la etapa ED6 hasta las primeras etapas ninfales del desarrollo, y que el resto de tipos de isomiRNA detectados no presentan un patrón claro correlacionado con las etapas de desarrollo de *Blattella germanica*. Para comprobar cuantitativamente estas observaciones se hizo un análisis de enriquecimiento de las distintas categorías de isomiRNA en las distintas etapas del desarrollo del insecto. Así, mediante este análisis, se comprobó que los distintos tipos de isomiRNA varían significativamente entre las distintas etapas embrionarias y NFE, pero no varían entre las etapas de ninfa y adulta. Con ello, este resultado aporta más seguridad a la tesis de que los isomiRNA y sus distintos tipos están más involucrados en la regulación génica de las etapas embrionarias de *Blattella germanica* que los miRNA.

A continuación, para estudiar en más profundidad las diferencias entre los isomiRNA y su miRNA canónico, se predijeron posibles "targets" en los mRNA de *Blattella germanica* para cada una de las anotaciones. Una vez obtenidas y filtradas las predicciones de los "targets" se contabilizaron las diferencias existentes entre los "targets" del conjunto de los isomiRNA asociados a un mismo miRNA canónico y los "targets" de ese miRNA canónico, a lo cual se ha denominado "familia de isomiRNA". Con estos datos se realiza una representación gráfica escalada por cada gen que permite observar la heterogeneidad de las "familias de isomiRNA" en cuanto a sus diferencias en sus "targets" con los "targets" de su miRNA canónico, habiendo "familias" con un gran número de diferencias y otras "familias" con muy pocas diferencias. Para ahondar más en el origen de la heterogeneidad de los "targets" se realiza otra representación de estos datos escalados por cada gen, pero, en este caso, agrupando las diferencias según el tipo de isomiRNA. De este modo se puede observar que los tipos de isomiRNA polimórficos-3' y polimórficos-5' con un nucleótido distinto y de longitud 3' presentan cambios considerables en sus targets respecto a sus miRNA canónicos, pero los tipos de isomiRNA polimórficos-3' con dos nucleótidos distintos y longitud 5' y longitud 3' y 5' no presentan muchos cambios en comparación con los otros tipos de miRNA.

Curiosamente (o no) los tipos de isomiRNA que menos cambios presentan en sus targets respecto a sus miRNA canónicos son los mismos que se asociaban preferentemente con las etapas embrionarias del desarrollo de *Blattella germanica*. Este hecho puede implicar que estos tipos de isomiRNA sean más precisos en los cambios que provocan en la regulación génica o que sufren algún cambio que altere su funcionalidad (estabilización o degradación del miRNA, mayor estabilidad termodinámica, regulación génica transcripcional, etc.). Para comprobar cuáles son exactamente los efectos que provocan estos tipos de isomiRNA sería necesario realizar más análisis bioinformáticos sobre cada uno de estos isomiRNA y realizar experimentos para comprobar las predicciones bioinformáticas. Además, la obtención de datos biológicos precisos sobre las interacciones y funciones de estos miRNA e isomiRNA permitiría realizar posteriormente análisis bioinformáticos biológicamente más avanzados mediante el desarrollo de modelos cuantificables acerca de la regulación génica de estos miRNA e isomiRNA sobre los mRNA. Lamentablemente, al no tratarse *Blattella germanica* de una especie bien estudiada en la bibliografía existente, no se dispone de estos datos.

Por ello, el siguiente paso de este trabajo fue intentar realizar un modelo aproximado del efecto sobre la regulación génica de las cuentas de los isomiRNA y miRNA en el desarrollo de *Blattella germanica* mediante los "targets" anteriormente predichos. De este modo, se multiplicó cada "target" de un miRNA o isomiRNA en un mRNA por el número de cuentas de ese miRNA o isomiRNA en un estado del desarrollo, por lo que la suma de las cuentas de todos los miRNA y/o isomiRNA que presentan como target un mismo mRNA en una etapa del desarrollo, representan las "cuentas de inhibición" de un mRNA en una etapa concreta del desarrollo por parte de los miRNA y/o isomiRNA.

Mediante este método, se obtuvieron gráficas acerca del enriquecimiento en "cuentas de inhibición" por parte de los miRNA, isomiRNA y ambos juntos en cada mRNA durante el desarrollo de *Blattella germanica*. Estas gráficas obtenidas acerca del enriquecimiento en "cuentas inhibitorias" ejercido por los isomiRNA y los miRNA por separado presentaron bastantes homologías con las gráficas obtenidas previamente sobre la distribución de las cuentas de esos isomiRNA y miRNA en las etapas del desarrollo. A pesar de ello, estas nuevas gráficas presentaban sus diferencias temporales y de número de mRNA afectados de forma más precisa, restringiéndose aún más esas diferencias provocadas por los isomiRNA y miRNA a las etapas embrionarias del desarrollo de *Blattella germanica*. Este hecho significaría que la distribución cualitativa, cuantitativa y temporal de los isomiRNA y miRNA se mantendría desde su presencia hasta la realización de su función en la regulación génica, pero de forma más precisa. Sin embargo, las diferencias observadas en la gráfica del enriquecimiento en "cuentas inhibitorias" provocado por los isomiRNA y los miRNA de forma conjunta sobre los mRNA no presentaban la misma distribución. En este caso, solo existen dos clusteres que presentan diferencias en su enriquecimiento en "cuentas inhibitorias" a lo largo del desarrollo del insecto. Exactamente, uno de esos clusteres se ve enriquecido inhibitoriamente en las etapas embrionarias y el otro clusteres en el resto de las etapas del desarrollo. Aparte de estos, también se observan otros clusteres de mRNA más o menos enriquecidos inhibitoriamente entre ellos, pero que no cambian a lo largo del desarrollo de *Blattella germanica*. De este modo, estos resultados nos aportan indicios acerca del posible efecto más preciso de los miRNA e isomiRNA en la regulación génica de ciertas etapas del desarrollo y a ciertos mRNA, pero que convergen en la existencia, en líneas generales, de dos sets de mRNA regulados por ellos y orientados a las etapas embrionarias y no embrionarias del desarrollo de *Blattella germanica*.

Finalmente, para observar las implicaciones funcionales de estas diferencias en el enriquecimiento en "cuentas inhibitorias" por parte de los isomiRNA y miRNA a lo largo del desarrollo, se realizó una comparación de clusteres funcionales mediante términos GO entre las distintas etapas del desarrollo del insecto. Mediante este análisis, se pudieron obtener indicios acerca de la tendencia general de los miRNA e isomiRNA a regular algunos procesos moleculares y celulares asociados al desarrollo mediante su permisividad en las etapas embrionarias y su represión en las etapas no embrionarias. También se obtuvieron indicios acerca de la tendencia, en este caso principalmente llevada a cabo por parte los isomiRNA, a regular procesos supracelulares fisiológicos y morfológicos mediante su represión durante las fases embrionarias y su permisividad durante las fases no embrionarias. Además, se observó que tanto

los miRNA como los isomiRs, ejercen una menor regulación génica (permisividad de la expresión) en el huevo no fecundado, una mayor regulación génica en las etapas embrionarias (represión de la expresión), y mixta en el adulto, aunque menor que en las etapas embrionarias. Así, estos resultados parecen indicar que, en cada etapa del desarrollo de *Blattella germanica*, existe una mayor o menor permisividad a la regulación génica por parte de los isomiRNA y los miRNA según los clusters funcionales que se necesitan regular en cada caso. Sin embargo, para poder comprobar estos indicios sería necesario realizar más análisis bioinformáticos y experimentos, no sólo en *Blattella germanica*, sino en más insectos y organismos, incluyendo insectos holometábolos para comprobar, por ejemplo, el posible cambio de sets de regulación génica en los mismos al realizar éstos una metamorfosis completa.

En resumen, las conclusiones científicas de este estudio es que se ha comprobado la asociación de ciertos isomiRNA concretos y tipos de isomiRNA a ciertas etapas del desarrollo del insecto *Blattella germanica*, predominando la asociación con las etapas embrionarias de su desarrollo. También se ha observado la heterogeneidad en las diferencias de "targets" en los isomiRNA respecto a sus miRNA canónicos, pero a su vez la preferencia de ciertos tipos de isomiRNA respecto a otros por presentar diferencias en sus "targets" en *Blattella germanica*. Y, además, se ha observado la posible existencia de una relación entre los isomiRNA y la regulación funcional de procesos fisiológicos y morfológicos mediante su represión en las etapas embrionarias. Sin embargo, estos resultados no deberían ser suficientes a largo plazo debido al futuro potencial de transferencia a la sociedad de este campo. Esto es debido a que el estudio en profundidad de la regulación génica en un organismo tan evolutivamente conservado como *Blattella germanica* puede permitir encontrar y desarrollar herramientas diagnósticas y terapeúticas para enfermedades humanas sumamente prevalentes. Además, este insecto es un modelo de laboratorio muy ventajoso, por el argumento anteriormente dado, y por el ahorro económico y de recursos naturales que supone su uso. Pero incluso, además, un mayor conocimiento de este insecto podría ayudar a desarrollar herramientas para controlar las plagas que provoca (y sus consecuencias en la salud humana), así como extrapolar ese conocimiento a otros insectos y desarrollar herramientas para la preservación de especies y ecosistemas. Así, por todas estas razones, y por los resultados científicos aportados, es fundamental seguir ampliando el conocimiento que aporta este trabajo.

## 6. Conclusiones

Las principales conclusiones del presente trabajo son:

- Las distintas anotaciones de isomiRNA caracterizan distintas etapas del desarrollo de *Blattella germanica*, sobre todo, y en comparación con los miRNA, las etapas embrionarias.
- Las diferencias entre las distintas categorías de isomiRNA caracterizan principalmente las etapas embrionarias del desarrollo *Blattella germanica*.
- Las diferencias de targets entre los isomiRNA y sus miRNA canónicos es muy heterogénea, aunque se agrupan según el tipo de isomiRNA en *Blattella germanica*.
- Los isomiRNA parecen regular más los clusters funcionales asociados a procesos fisiológicos y morfológicos respecto a los miRNA, regulando ambos (tanto miRNA como isomiRNA) igualmente los procesos celulares.
- Los isomiRNA y sus implicaciones funcionales, en comparación con los miRNA, se asocian principalmente con la regulación de las etapas embrionarias del desarrollo de *Blattella germanica*.

## 7. Bibliografía

1. Segre J. Gene regulation 2023 [Available from: <https://www.genome.gov/genetics-glossary/Gene-Regulation>].
2. Behzadi P, Issakhanian L. Introductory chapter: gene regulation, an RNA network-dependent architecture. *Gene regulation*: IntechOpen; 2019.
3. Tomasello L, Distefano R, Nigita G, Croce CM. The MicroRNA Family Gets Wider: The IsomiRs Classification and Role. *Front Cell Dev Biol*. 2021;9:668648.
4. Stavast CJ, Erkeland SJ. The Non-Canonical Aspects of MicroRNAs: Many Roads to Gene Regulation. *Cells*. 2019;8(11).
5. Bofill-De Ros X, Yang A, Gu S. IsomiRs: Expanding the miRNA repression toolbox beyond the seed. *Biochim Biophys Acta Gene Regul Mech*. 2020;1863(4):194373.
6. Glogovitis I, Yahubyan G, Wurdinger T, Koppers-Lalic D, Baev V. IsomiRs-Hidden Soldiers in the miRNA Regulatory Army, and How to Find Them? *Biomolecules*. 2020;11(1).
7. Ylla G, Piulachs MD, Belles X. Comparative analysis of miRNA expression during the development of insects of different metamorphosis modes and germ-band types. *BMC Genomics*. 2017;18(1):774.
8. Schmartz GP, Kern F, Fehlmann T, Wagner V, Fromm B, Keller A. Encyclopedia of tools for the analysis of miRNA isoforms. *Brief Bioinform*. 2021;22(4).
9. Vasconcelos AM, Carmo MB, Ferreira B, Viegas I, Gama-Carvalho M, Ferreira A, et al. IsomiR\_Window: a system for analyzing small-RNA-seq data in an integrative and user-friendly manner. *BMC Bioinformatics*. 2021;22(1):37.
10. Glogovitis I, Yahubyan G, Wurdinger T, Koppers-Lalic D, Baev V. miRGalaxy: Galaxy-Based Framework for Interactive Analysis of microRNA and isomiR Sequencing Data. *Cancers (Basel)*. 2021;13(22).
11. Desvignes T, Loher P, Eilbeck K, Ma J, Urgese G, Fromm B, et al. Unification of miRNA and isomiR research: the mirGFF3 format and the mirtop API. *Bioinformatics*. 2020;36(3):698-703.
12. Pantano L, Estivill X, Martí E. SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. *Nucleic Acids Res*. 2010;38(5):e34.
13. Aparicio-Puerta E, Gomez-Martin C, Giannoukakos S, Medina JM, Scheepbouwer C, Garcia-Moreno A, et al. sRNAbench and sRNAtoolbox 2022 update: accurate miRNA and sncRNA profiling for model and non-model organisms. *Nucleic Acids Res*. 2022;50(W1):W710-W7.
14. Zhong X, Pla A, Rayner S. Jasmine: a Java pipeline for isomiR characterization in miRNA-Seq Data. *Bioinformatics*. 2019;36(6):1933-6.
15. Patil AH, Halushka MK. miRge3.0: a comprehensive microRNA and tRF sequencing analysis pipeline. *NAR Genom Bioinform*. 2021;3(3):lqab068.

16. Gaffo E, Bortolomeazzi M, Bisognin A, Di Battista P, Lovisa F, Mussolin L, et al. MiR&moRe2: A Bioinformatics Tool to Characterize microRNAs and microRNA-Offset RNAs from Small RNA-Seq Data. *Int J Mol Sci.* 2020;21(5).
17. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res.* 2002;12(6):996-1006.
18. Fromm B, Hoye E, Domanska D, Zhong X, Aparicio-Puerta E, Ovchinnikov V, et al. MirGeneDB 2.1: toward a complete sampling of all major animal phyla. *Nucleic Acids Res.* 2022;50(D1):D204-D10.
19. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
20. Riolo G, Cantara S, Marzocchi C, Ricci C. miRNA Targets: From Prediction Tools to Experimental Validation. *Methods Protoc.* 2020;4(1).
21. Zhou X, Qian K, Tong Y, Zhu JJ, Qiu X, Zeng X. De novo transcriptome of the Hemimetabolous German cockroach (*Blattella germanica*). *PLoS One.* 2014;9(9):e106932.
22. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife.* 2015;4.
23. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* 2020;48(D1):D127-D31.
24. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome Biol.* 2003;5(1):R1.
25. Gumienny R, Zavolan M. Accurate transcriptome-wide prediction of microRNA targets and small interfering RNA off-targets with MIRZA-G. *Nucleic Acids Res.* 2015;43(3):1380-91.
26. Soutschek M, Gross F, Schrott G, Germain PL. scanMiR: a biochemically based toolkit for versatile and efficient microRNA target prediction. *Bioinformatics.* 2022;38(9):2466-73.
27. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 2008;36(10):3420-35.
28. UniProt C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* 2023;51(D1):D523-D31.
29. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25-9.
30. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb).* 2021;2(3):100141.
31. Adamopoulos PG, Athanasopoulou K, Daneva GN, Scorilas A. The Repertoire of RNA Modifications Orchestrates a Plethora of Cellular Responses. *Int J Mol Sci.* 2023;24(3).

32. Costa Dos Santos G, Jr., Renovato-Martins M, de Brito NM. The remodel of the "central dogma": a metabolomics interaction perspective. *Metabolomics*. 2021;17(5):48.
33. Pramanik D, Shelake RM, Kim MJ, Kim JY. CRISPR-Mediated Engineering across the Central Dogma in Plant Biology for Basic Research and Crop Improvement. *Mol Plant*. 2021;14(1):127-50.
34. Supriya N. Messenger RNA 2023 [Available from: <https://biologyreader.com/messenger-rna.html>].
35. Inamura K. Major Tumor Suppressor and Oncogenic Non-Coding RNAs: Clinical Relevance in Lung Cancer. *Cells*. 2017;6(2).
36. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer*. 2015;15(6):321-33.
37. Arao Y, Nakayama M, Tsuji Y, Hamano Y, Otsuka C, Vecchione A, et al. EpisomiR, a New Family of miRNAs, and Its Possible Roles in Human Diseases. *Biomedicines*. 2022;10(6).
38. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)*. 2020;2020.
39. Nasirian H. Infestation of cockroaches (Insecta: Blattaria) in the human dwelling environments: A systematic review and meta-analysis. *Acta Trop*. 2017;167:86-98.
40. Truman JW, Riddiford LM. The evolution of insect metamorphosis: a developmental and endocrine view. *Philos Trans R Soc Lond B Biol Sci*. 2019;374(1783):20190070.
41. Tanaka A. Stages in the embryonic development of the German cockroach, *Blattella germanica* Linné (Blattaria, Blattellidae). 1976;44(4):512-25.
42. Fernandez-Nicolas A, Ventos-Alfonso A, Kamsoi O, Clark-Hachtel C, Tomoyasu Y, Belles X. Broad complex and wing development in cockroaches. *Insect Biochem Mol Biol*. 2022;147:103798.
43. Ylla G, Piulachs MD, Belles X. Comparative Transcriptomics in Two Extreme Neopterans Reveals General Trends in the Evolution of Modern Insects. *iScience*. 2018;4:164-79.
44. Fujita M, Machida RJAs, phylogeny. Reproductive biology and postembryonic development of a polyphagid cockroach *Eucorydia yasumatsui* Asahina (Blattodea: Polyphagidae). 2014;72(2):193-211.
45. Tanaka A, Ogawa E, Ohtake M. Relative growth of leg segments of the German cockroach, *Blattella germanica*. *Growth*. 1986;50(2):273-86.
46. Robinson JT, Thorvaldsdottir H, Turner D, Mesirov JP. igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics*. 2023;39(1).
47. Navarro Gonzalez J, Zweig AS, Speir ML, Schmelter D, Rosenbloom KR, Raney BJ, et al. The UCSC Genome Browser database: 2021 update. *Nucleic Acids Res*. 2021;49(D1):D1046-D57.

48. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155-D62.
49. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2).
50. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011;6:26.
51. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online; 2010.
52. Langmead B, Wilks C, Antonescu V, Charles R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics*. 2019;35(3):421-32.
53. Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 2012;40(1):37-52.
54. Gene Ontology C, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. *Genetics*. 2023;224(1).
55. M C. GO.db: A set of annotation maps describing the entire Gene Ontology. R package version 3.8.2. ed: Bioconductor; 2019.
56. Bairoch A. The ENZYME database in 2000. *Nucleic Acids Res.* 2000;28(1):304-5.

## 8. Anexos

### 8.1 Anexo I

Semana \ Tarea	T1	T2	T3	T4	T5	T6	T7	E	D
S01 20/03 – 26/03									
S02 27/03 – 02/04									
S03 03/04 – 09/04									
S04 10/04 – 16/04									
S05 17/04 – 23/04									
S06 24/04 – 30/04									
S07 01/05 – 07/05									
S08 08/05 – 14/05									
S09 15/05 – 21/05									
S10 22/05 – 28/05									
S11 29/05 – 04/06									
S12 05/06 – 11/06									
S13 12/06 – 18/06									
S14 19/06 – 25/06									
S15 26/06 – 02/07									

E: Escritura del Trabajo Fin de Máster // D: Defensa del Trabajo Fin de Máster

## 8.2 Anexo II

CPUS: el número de CPUS a usar.

GFF\_FILE: la ruta al archivo de anotación de miRs en formato “.gff3”.

GENOME: la ruta al archivo del genoma del organismo en formato fasta.

META: el archivo “.csv” con los datos de las muestras a analizar.

NOAPADAPTER: realizar la eliminación de los adaptadores o no. En este caso se indica que no se realice, ya que ya estaban eliminados en los datos de origen.

SPECIES: la especie a analizar.

MIN\_COUNT: el filtro del mínimo número de cuentas.

MIN\_MORNA\_LEN: un filtro para el tamaño mínimo de los moRs.

MAX\_LEN\_FILTER: un filtro para el tamaño máximo de una lectura.

MEAN\_QUAL\_FILTER: un filtro de la calidad media total de una lectura.

MORFILTER: en este caso permite contabilizar las secuencias que se alinean también con los miRs para la expresión de moRs.

ALLOWED\_OVERHANG: un filtro para el tamaño que se permite extender un isomiR respecto al miR.

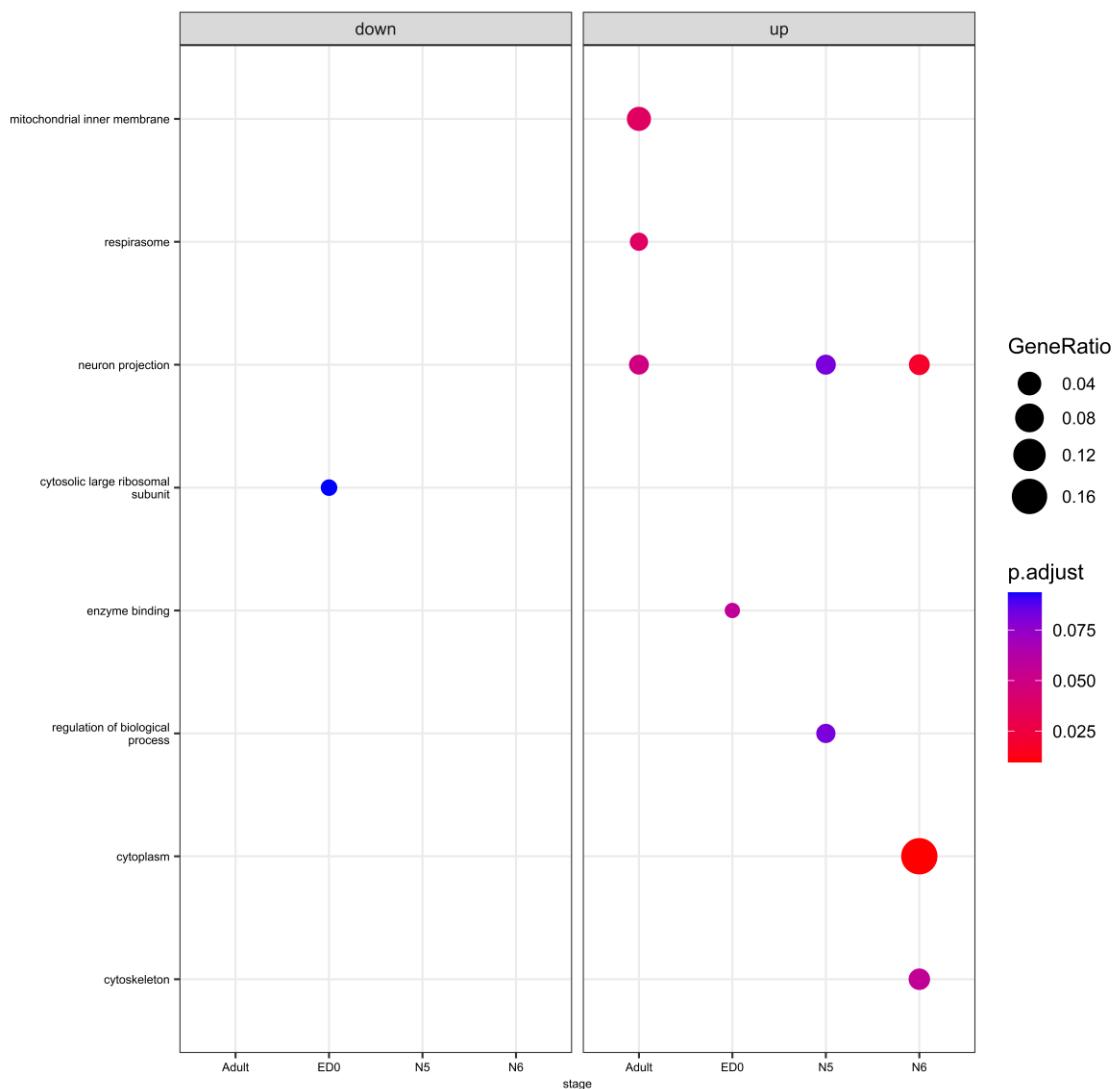
PREDICT\_NOVEL\_PREMIRS: activar la predicción de nuevos miRs.

NON\_MIRNAS: activar la caracterización de secuencias no alineadas con los miRs anotados.

## 8.3 Anexo III

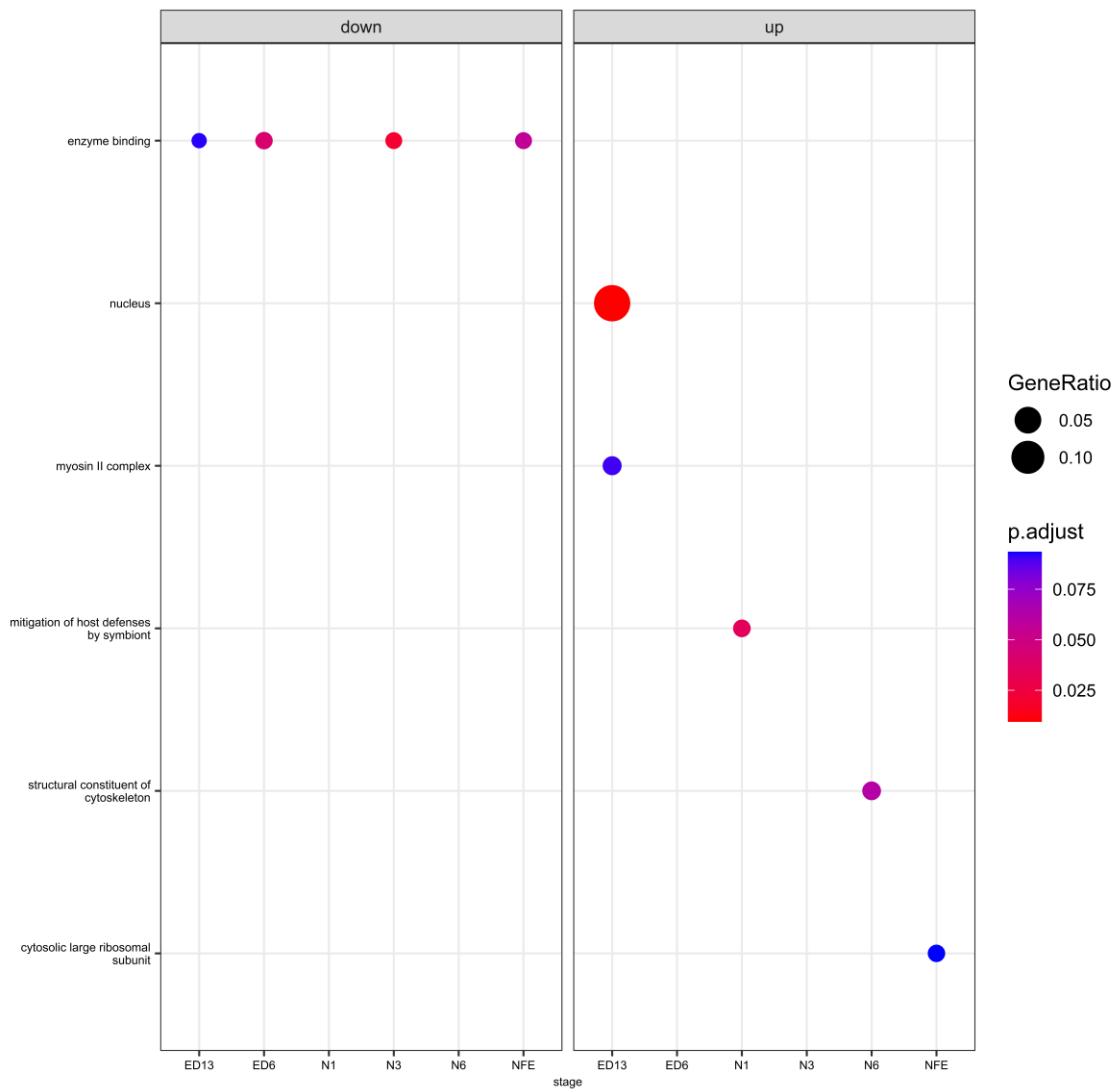
```
CPUS = '2'  
GFF_FILE = '/home/jalarcongi/TFM/bge_mirbase.gff3'  
GENOME = '/home/jalarcongi/TFM/GCA_000762945.2.fa'  
META = 'samples.csv'  
NOADAPTER = 'True'  
SPECIES = 'GCA'  
MIN_COUNT = '1'  
MIN_MORNA_LEN = '15'  
MAX_LEN_FILTER = '31'  
MEAN_QUAL_FILTER = '26'  
MORFILTER = 'permissive'  
ALLOWED_OVERHANG = '4'  
PREDICT_NOVEL_PREMIRS = 'True'  
NON_MIRNAS = 'True'
```

## 8.4 Anexo IV



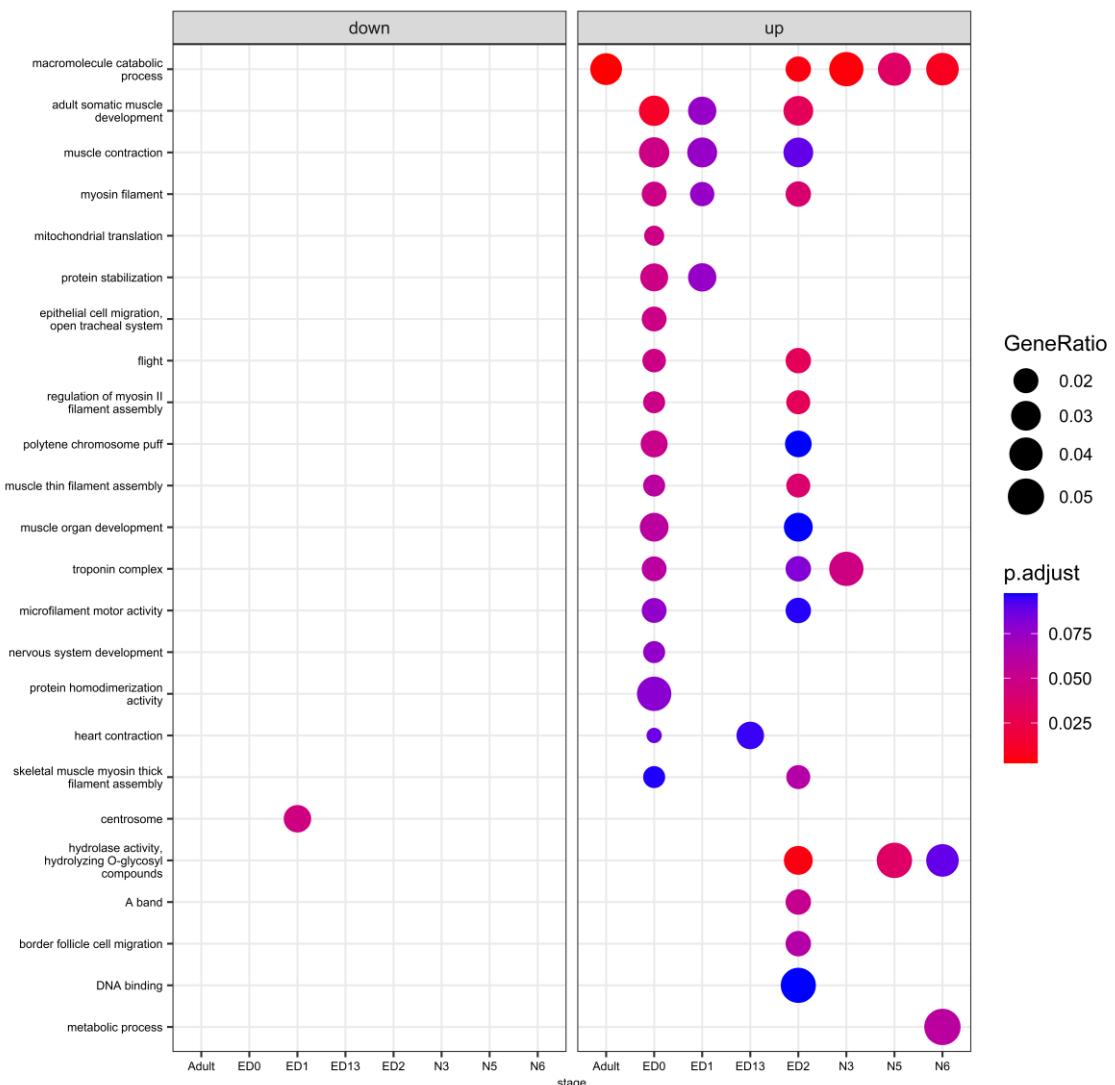
**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa NFE.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa NFE. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

## 8.5 Anexo V



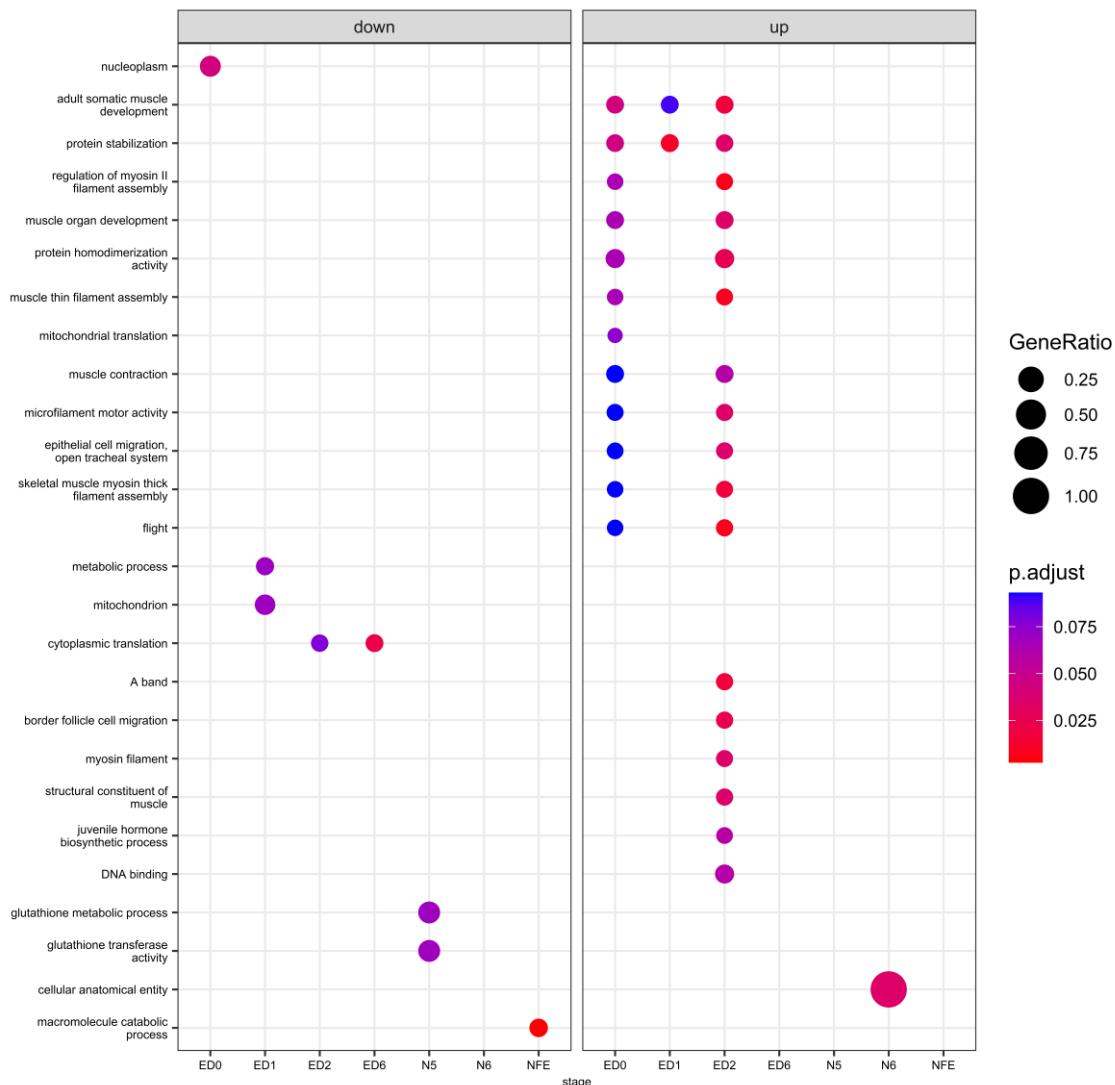
**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa ED0.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa ED0. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

## 8.6 Anexo VI



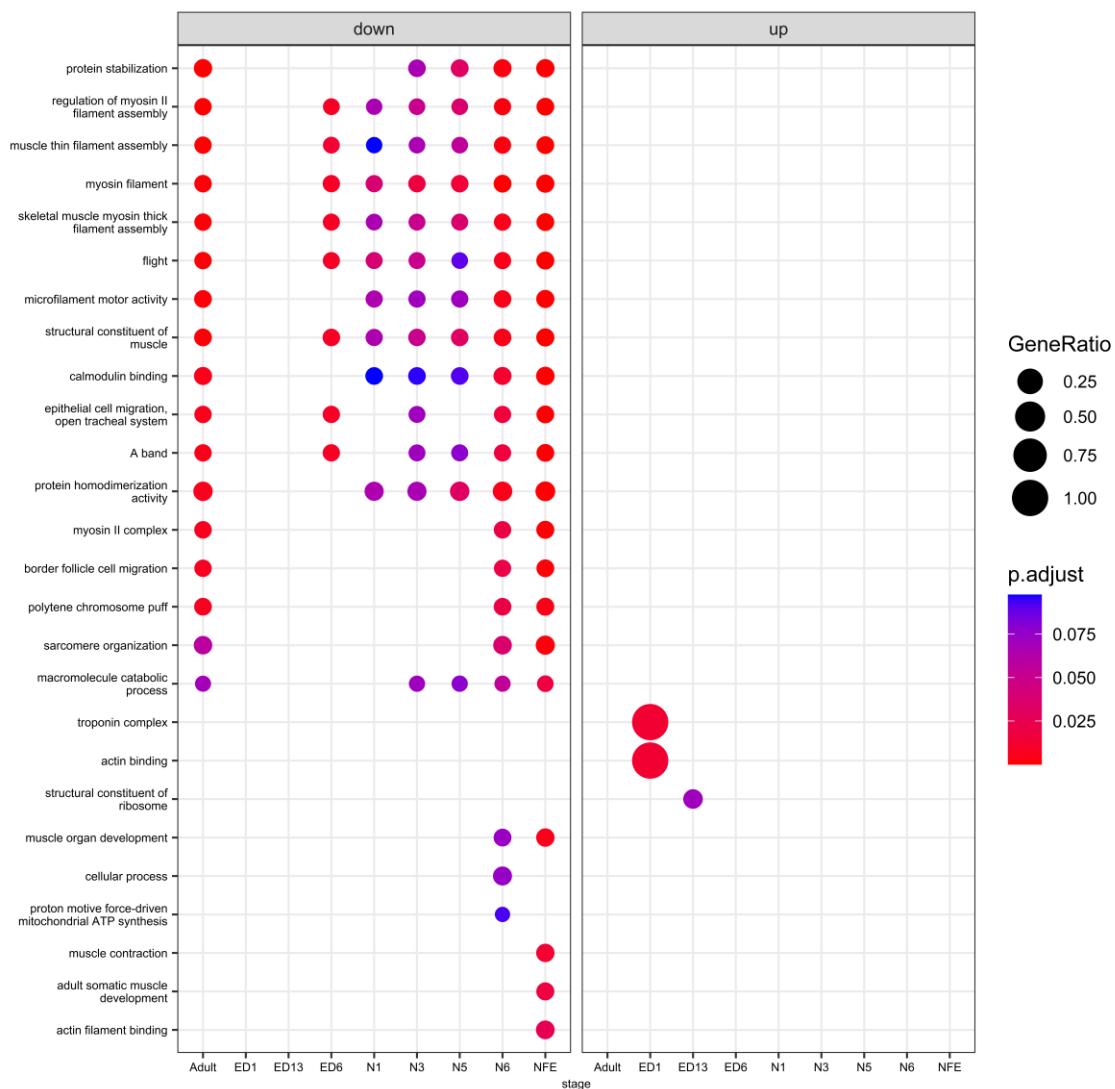
**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa NFE.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa NFE. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

## 8.7 Anexo VII



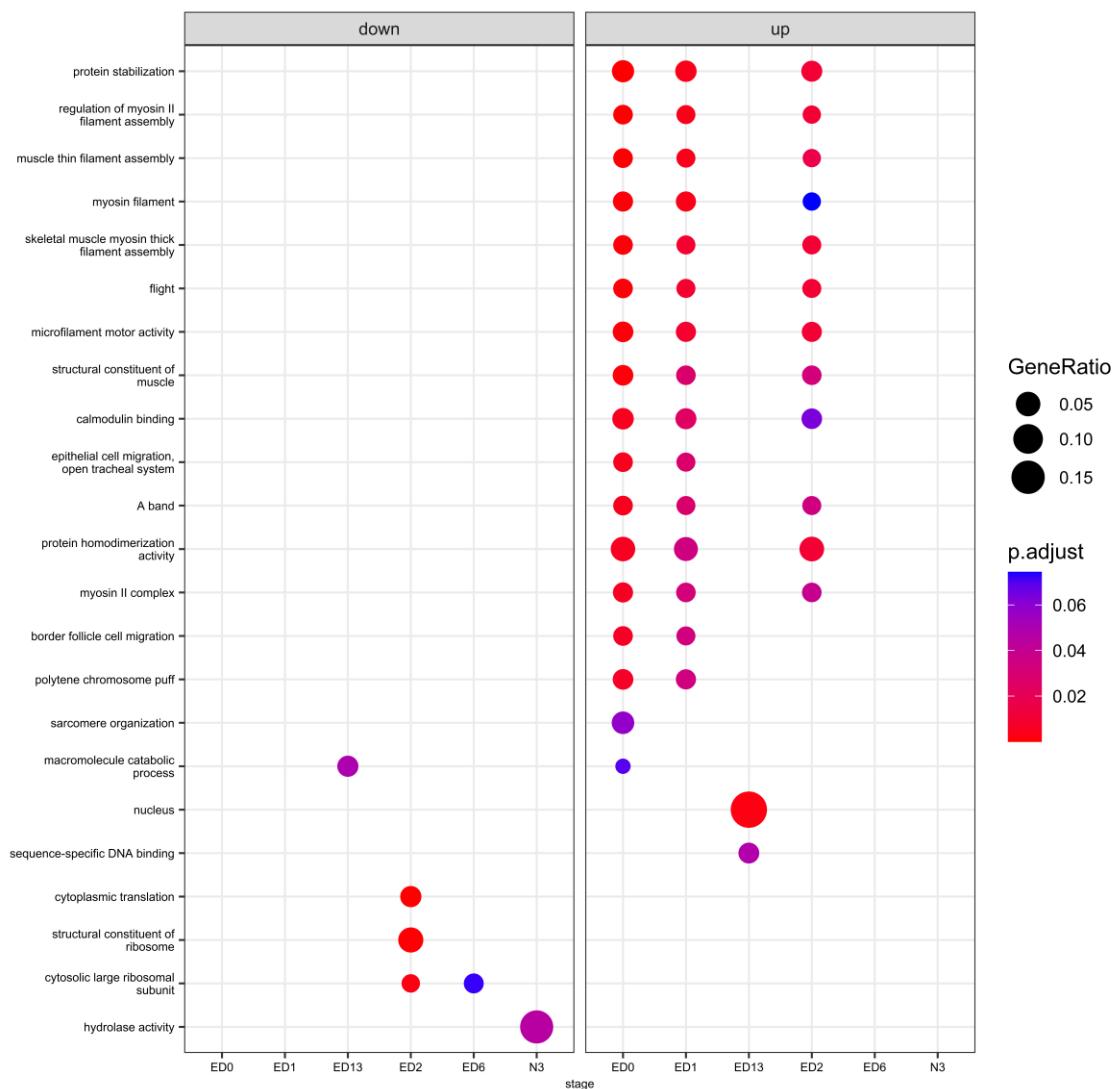
**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa adulta.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa adulta. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

## 8.8 Anexo VIII



**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa ED0.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa ED0. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.

## 8.9 Anexo IX



**Dotplot de la comparación de clústeres funcionales a partir de las “cuentas de inhibición” por parte de los isomiRNA y los miRNA en los mRNA en las etapas del desarrollo de *Blattella germanica* respecto a la etapa adulta.** En esta gráfica se muestra la comparación de clústeres funcionales (puntos) de distintas etapas del desarrollo (eje X) para distintos clústeres funcionales (eje Y) en comparación con la etapa adulta. Además, el tamaño de los puntos denota el “gene ratio” de una comparación para un clúster y el color de los puntos indica el p-valor de la comparación.