

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

MAESTRÍA EN BIOLOGÍA COMPUTACIONAL

**Identificación de péptidos antimicrobianos y antifúngicos involucradas en la
respuesta inmune en *Drosophila mesophragmatica***

Disertación previa a la obtención del título de Magister en Biología Computacional

Franco Stelios Cabrera Ruiz

Quito, 2022

INTRODUCCIÓN

Sistema inmune en insectos

Los insectos son el grupo más diverso de los artrópodos ocupando al menos el 75% del total de especies descritas (Behura, 2015), incluyen organismos carnívoros, detritívoros, hervívoros y parásitos; los cuales pueden ocupar diversos ecosistemas (Thomas *et al.*, 2020). Los insectos están expuestos a varios organismos como hongos, bacterias y virus que pueden infectar su cavidad interna u homocelo (Siva, Morett y Rolff, 2005) (Vega y Kaya, 2012), por lo que han desarrollado estrategias para defenderse de factores patógenos (Hanson, Hamilton y Perlman, 2016). El primer mecanismo de defensa se basa en la conducta, evitando acercarse a poblaciones infectadas, sin embargo, este mecanismo no es totalmente efectivo y necesitan de una protección externa. La cutícula es una estructura quitinosa que forma parte del exoesqueleto de los insectos y actúa como segundo filtro para evitar la infección de patógenos (Siva, Morett y Rolff, 2005). No obstante, los patógenos pueden ingresar a través de heridas o el tracto digestivo al homocelo, causando una respuesta inmune de las células inmunitarias que en el caso de los insectos son conocidas como hemocitos, estas células forman parte del último mecanismo y más eficaz contra infecciones de todo tipo (Hillyer, 2016).

Los hemocitos brindan al organismo una protección inmunitaria innata, la cual actúa de manera rápida e inespecífica contra infecciones (Netea *et al.*, 2019) desencadenando procesos celulares como la fagocitosis, lisis, melanización y respuestas humorales en la que actúan moléculas efectoras solubles como los péptidos antimicrobianos (Strand, 2008). Los hemocitos pueden clasificarse como sésiles y móviles, en donde los hemocitos sésiles son aquellos que se fijan a tejidos mientras que los hemocitos móviles se encuentran circulando por la hemolinfa (Hillyer y Stran, 2014). Para que se genere la respuesta inmunitaria en el insecto debe ocurrir un reconocimiento de la entidad invasora; este proceso se produce por la interacción de los receptores de reconocimiento (PRR) y los patrones moleculares asociados a patógenos (PAMP). Los PRR involucran una familia diversa de proteínas implicadas en procesos como la fagocitosis, melanización y activación de vías de señalización intercelular (Hillyer, 2016).

Existen tres vías de señalización principales en insectos, la vía de señalización de sensores de reconocimiento de membrana (Toll) , la vía de señalización de deficiencia

inmune (Imd) y la vía de señalización de tiroquinasas de la familia Janus con su proteína activadora de transcripción Stat (Jack Stat) (Hillyer, 2016) (Larsen, Reynaldi y Novia, 2019) . La vía de señalización Toll se encuentra altamente conservada y está relacionada con el desarrollo e inmunidad (Hegde *et al.*, 2022). Cuando se detecta una entidad invasora los PRRs interactúan con la citoquina Spätzle localizada en el exterior de la célula provocando la activación de la vía Toll intracelularmente y generando un gradiente molecular, en donde, el factor de transcripción NF-κB es activado, empezando la producción de péptidos antimicrobianos y la transcripción de genes relacionados con el sistema inmune. En el caso de la vía Imd, el factor de transcripción NF-κB, los diferentes péptidos antimicrobianos y la transcripción de genes inmunológicos son activados cuando existe la interacción entre PAMP y el receptor extracelular PGRP-LC. Por último, la vía Jack Stat es activada mediante dos procesos, el primero involucra la interacción del receptor Domeless y la citoquina extracelular Unpaired mientras que el segundo involucra la atracción de la proteína Stat, ambos procesos promueven la transcripción de genes inmunitarios (Hillyer, 2016).

El correcto funcionamiento del sistema inmune además de prevenir infecciones mantiene estables los niveles de homeostasis del organismo (Sheehan *et al.*, 2018). Los insectos han logrado desarrollar un sistema inmunológico innato exitoso que les ha permitido ocupar varios nichos ecológicos. Este hito evolutivo en los insectos está asociado a una preparación inmunológica generada en la primera exposición con un patógeno, brindando al organismo una memoria inmune que incluso puede ser heredada a las siguientes generaciones (Sulek, Kordakzuk y Wojda, 2021). Sin embargo, hay que tener en consideración que los niveles de estrés altos pueden generar un déficit inmunológico (Adamo, 2017).

Sistema inmune en *Drosophila*

Los estudios inmunológicos en *Drosophila* han permitido comprender varios rasgos referentes a la inmunidad innata (Hoffmann y Reichhart, 2002) (Igaki *et al.*, 2002) (Jang *et al.*, 2006), por ejemplo, Yu *et al.* (2022) clasifican la inmunidad innata de *Drosophila* en inmunidad humoral y celular, en donde la inmunidad humoral se encuentra relacionada con el cuerpo graso y la hemolinfa, mientras que la inmunidad celular se relaciona con los hemocitos. A pesar que esta clasificación puede facilitar la comprensión de los componentes inmunológicos innatos, esta sugerencia puede ser relativa debido a que ambos componentes presentan una estrecha relación funcional (Hillyer y Stran, 2014).

Casi en su totalidad, los hemocitos están involucrados en la inmunidad celular. En *Drosophila* se han identificado tres tipos de hemocitos: plasmatocitos, células cristalinas y lamelocitos (Yu *et al.*, 2022). Los plasmatocitos representan alrededor del 95% de la cantidad total de hemocitos y su función principal es intervenir en el proceso de fagocitosis, además en el estado embrionario; los plasmatocitos ayudan en la secreción de proteínas MEC encargadas de la morfogénesis y viabilidad de los embriones (Benerjee *et al.*, 2019). Por otra parte, las células cristalinas intervienen en la melanización (Kurucz *et al.*, 2007) a través de la producción de profeniloxidasas; estas células corresponden al porcentaje restante del total de hemocitos (Dudzic *et al.*, 2015). Finalmente, los lamelocitos son células diferenciadas a partir de plasmatocitos cuando existe parasitismo por parte de avispas (Yu *et al.*, 2022).

Como se mencionó anteriormente, las respuestas humorales involucran péptidos antimicrobianos o AMPs, estas moléculas se introducen en la zona hidrofóbica de las membranas lipídicas de los microorganismos con el fin de alterar su composición y causar la muerte celular (Hanson y Lamaitre, 2020); actualmente existen siete AMPs con respuesta a diferentes patógenos, la drosomicina y metchnikowin se encargan de combatir levaduras, la attacina, cecropina, drosocina y diptericina generan respuestas contra bacterias gram negativas y la defensina combate bacterias gram positivas (Yu *et al.*, 2022). Existe una estrecha relación entre la producción de AMPs y las vías de señalización. Por ejemplo, la vía de señalización Toll interviene en la producción de AMPs en presencia de hongos y bacterias gram positivas (Hanson y Lamaitre, 2020) (Yu *et al.*, 2022), así como en el proceso melanización (Ligoxygakis *et al.*, 2002). En segundo lugar, la vía de señalización Imd participa en la producción de AMPs (Hanson y Lamaitre, 2020) en presencia de bacterias gram negativas (Gendrin *et al.*, 2009) y la vía de señalización Jack Stat está involucrada en el proceso de diferenciación celular de hemocitos en lamelocitos en respuesta a la parasitación de avispas (Benerjee *et al.*, 2019), a su vez, puede ser activada por estrés como radiación ultravioleta, exposición a metales pesados, choque térmico, entre otros (Ekengren y Hultmark, 2001).

A pesar de la especificidad de los AMPs la mayoría de estudios se basan en la especie *Drosophila melanogaster* y actualmente se conoce que las secuencias de estas proteínas evolucionan frente a cambios ecológicos y patógenos (Hanson y Lamaitre, 2020), por lo tanto, ampliar los estudios en el género *Drosophila* permitirá comprender la evolución de estos péptidos antimicrobianos.

Grupo *mesophragmatica* y rasgos ecológicos

El grupo *mesophragmatica* (Brncic y Koref Santibañez, 1957) pertenece al subgénero *Drosophila*, radiación *virilis-repleta* (Mota *et al.*, 2008). Actualmente, está conformado por alrededor de 19 especies (Duda, 1927) (Brncici, 1957) (Jaeger y Salzano, 1953) (Brncic y Koref Santibañez, 1957) (Mota *et al.*, 2008) (Cespédes y Rafaél, 2012) (Guillín y Rafael, 2017) distribuidas en países andinos y ocupando rangos altitudinales entre los 2500 a 4000 msnm (Figuro, Rafael y Cespédes, 2012). Estas especies presentan una evolución alotrópica (Mota *et al.*, 2008) debido a la formación de la cordillera de los Andes (Cespédes y Rafaél, 2012), en donde, este rasgo ecológico puede causar variaciones genéticas en estas especies (Kellermann *et al.*, 2009) fundamentalmente cuando se encuentran en zonas de variabilidad climática como trópicos y neotrópicos (Kellermann *et al.*, 2012).

En general, los estudios inmunológicos no incluyen especies pertenecientes a otros grupos y el grupo *mesophragmatica* es de importancia Andina debido a la diversidad ecológica que presenta, por lo tanto, se espera que con este y próximos estudios se puedan incluir más especies en este tipo de análisis.

Justificación y Objetivos

Los estudios inmunológicos en especies andinas del género *Drosophila* son escasos, por lo que en este proyecto se decidió identificar los péptidos antimicrobianos y antifúngicos de la especie *Drosophila mesophragmatica*, en primer lugar, debido a que es la especie más representativa del grupo con una distribución amplia en relación a sus especies hermanas. Y en segundo lugar, porque actualmente se conoce que los péptidos antimicrobianos pueden variar frente a cambios ecológicos. Por lo tanto, la identificación de estas secuencias peptídicas permitirán contribuir en trabajos relacionados con la inmunidad en respuesta a cambios ecológicos en zonas andinas, promoviendo el uso de especies del grupo *mesophragmatica* cuya diversidad y ecología es amplia en los Andes y en el Ecuador.

El objetivo fundamental de este proyecto es identificar los péptidos antimicrobianos y antifúngicos en la especie *D. mesophragmatica* con el fin de: verificar la presencia o ausencia de las secuencias peptídicas, comparar la variabilidad genética con la especie modelo *D. melanogaster* y comprobar las diferentes vías de señalización involucradas.

Metodología

Datos crudos

Los datos fueron proporcionados por el laboratorio de genética evolutiva de la Pontificia Universidad Católica del Ecuador. Se facilitó el transcriptoma no ensamblado de *D. mesophragmatica*, secuenciado a través de la plataforma Illumina, en donde se tiene un total de 8 lecturas paired-end correspondientes a 2 machos y 2 hembras.

Calidad y limpieza de secuencias

Para verificar la calidad de las secuencias crudas se usó el programa fastqc (Babraham Bioinformatics, 2022), en donde, los 8 archivos de lecturas fueron sometidos al control de calidad. Para la limpieza de las lecturas se usó el programa fastp (Chen *et al.*, 2018) a través de los siguientes parámetros “- - detect adapter for pe - - trim front1 20 - - trim front2 20” para eliminar la contaminación de las lecturas con el adaptador Illumina y a su vez recortar las bases de baja calidad. Finalmente, se volvió a usar fastqc para verificar la calidad de las lecturas.

Ensamblaje del transcriptoma

Las lecturas procesadas se usaron para generar el ensamblaje de machos y hembras usando Trinity v2.1.1 (Grabherr *et al.*, 2011) con los parámetros predefinidos por el sistema y modificando solo la cantidad de CPU y memoria “- - max memory 16 - - CPU 2”. La calidad y la integridad del ensamblaje se evaluaron usando ortólogos universales de copia de única mediante la herramienta BUSCO v5.4.3 (Waterhouse *et al.*, 2018) con una búsqueda automática en eucariotas “- -auto-lineage-euk” y diptera “- diptera_odb10”.

Identificación de proteínas

Las secuencias de proteínas atacina, cecropina, defensina y drosomicina fueron obtenidas de la base de datos del NCBI de la especie *Drosophila melanogaster*. La búsqueda de las secuencias de proteína en el transcriptoma de *D. mesophragmatica* se realizó mediante el programa tblastn (Camacho *et al.*, 2009), en donde, mediante una base de datos local con los transcriptomas de machos y hembras de *D. mesophragmatica* se llevó a cabo una búsqueda de similitud de cada una de las proteínas.

Resultados

Calidad y limpieza de las lecturas

En general los 8 archivos de lecturas crudas presentan una calidad por base superior a 30, sin embargo, se muestran fallos con respecto al contenido de secuencia por base, en donde, se encontró una composición de secuencia sesgada al inicio de todas las lecturas entre 1-20 pares de bases, también se presentó sesgo en los niveles de duplicación de las secuencias los cuales superan el 50% y por último, el contenido de GC en las lecturas presentan un pico entre 31-35 %. Al realizar la limpieza de las lecturas crudas se obtuvo una mejora con respecto al contenido de secuencia por base, sin embargo, los dos parámetros restantes no presentaron mejoras significativas. El recorte en la longitud de las lecturas crudas para eliminar adaptadores, secuencias ambiguas o de baja calidad dió como resultado una pérdida del 0.1 % en cada uno de los archivos de lectura (Tabla 1).

Tabla 1. Longitud de secuencias de datos crudos y curados.

| Nombre de la secuencia | Total de lecturas datos crudos | Total de lecturas datos procesados | % de lecturas eliminadas |
|------------------------|--------------------------------|------------------------------------|--------------------------|
| DmesoRNA_fem4_1.fastq | 55860874 | 55355642 | 0.1 |
| DmesoRNA_fem4_2.fastq | 55860874 | 55355642 | 0.1 |
| DmesoRNA_fem8_1.fastq | 47488549 | 47083796 | 0.1 |
| DmesoRNA_fem8_2.fastq | 47488549 | 47083796 | 0.1 |
| DmesoRNA_male8_1.fastq | 51163443 | 50639923 | 0.1 |
| DmesoRNA_male8_2.fastq | 51163443 | 50639923 | 0.1 |
| DmesoRNA_male9_1.fastq | 48984537 | 48616901 | 0.1 |
| DmesoRNA_male9_2.fastq | 48984537 | 48616901 | 0.1 |

Ensamblaje del transcriptoma

El ensamblaje final de machos presentó un total de 73489 transcritos, agrupados en 42819 genes. El total de bases ensambladas fue de 107037115 y presentó un contenido de GC de 42.22%, en donde, el valor N50 para todos los contigs transcritos es de 2780 (Tabla 2). En el análisis de calidad para eucariota se obtuvo un total de 98.8% de genes completos, en donde, el 38% son de copia única y el 60.8% duplicados. Adicionalmente, se obtuvo un 1.2% de genes fragmentados y 0% de genes faltantes. Con respecto al análisis más específico de dípteros, se obtuvo un 94.6% de genes completos; en donde, el 32.8% corresponde a genes de copia única y el 61.8% a

genes duplicados. Adicionalmente, se obtuvo un 2.1% de genes fragmentados y un 3.3% de genes faltantes (Tabla 3).

En el caso del ensamblaje final de hembras, el total de transcritos fue de 67229, agrupados en 39819 genes; en donde, el total de bases ensambladas fue de 103030921 y presentó un contenido de GC de 45.17% con un valor N50 para todos los contigs transcritos de 3030 (Tabla 2). En el análisis de calidad para eucariota se obtuvo un total del 100% de genes completos, en donde, el 47.5% corresponde a secuencias de copia única y el 52.5% a secuencias duplicadas. Con respecto al análisis en dípteros, se obtuvo un 93.8% de genes completos, en donde, el 34.9% corresponde a secuencias de una solo copia y el 58.9% a secuencias duplicadas. Adicionalmente, se obtuvo un 2.3% de genes fragmentados y un 3.9% de genes faltantes (Tabla 4).

Tabla 2. Estadístico de ensamblaje para machos y hembras

| Ensamblaje | Total de bases ensambladas | Total de transcritos | Total de genes | Contenido de GC (%) | N50 |
|------------|----------------------------|----------------------|----------------|---------------------|------|
| Machos | 107037115 | 73489 | 42819 | 42.22 | 2780 |
| Hembras | 103030921 | 67229 | 39819 | 45.17 | 3030 |

Tabla 3. Calidad de ensamblaje para machos. A representa la búsqueda en eucariota y B representa la búsqueda en diptera

| Genes completos (%) | Genes de una copia (%) | Genes duplicados (%) | Genes fragmetados (%) | Genes faltantes (%) |
|---------------------|------------------------|----------------------|-----------------------|---------------------|
| A 98.8 | 30 | 60.8 | 1.2 | 0 |
| B 94.6 | 32.8 | 61.8 | 2.1 | 3.3 |

Tabla 4. Calidad de ensamblaje para hembras A representa la búsqueda en eucariota y B representa la búsqueda en diptera

| Genes completos (%) | Genes de una copia (%) | Genes duplicados (%) | Genes fragmetados (%) | Genes faltantes (%) |
|---------------------|------------------------|----------------------|-----------------------|---------------------|
| A 100 | 47.5 | 52.5 | 0 | 0 |
| B 93.8 | 34.9 | 58.9 | 2.3 | 3.9 |

Identificación de proteínas

El alineamiento de Blast permitió obtener varias secuencias con similitudes significativas para cada una de las proteínas en machos y hembras. En el caso de

machos, al realizar al alineamiento con respecto a la proteína atacina la secuencia más significativa presenta un e-value de $6e-104$, un porcentaje de identidad del 80%, un porcentaje de sustituciones conservativas del 88% y 2% de gaps. La secuencia más significativa para la proteína cecropina presenta un e-value de $5e-28$, un porcentaje de identidad del 90%, un porcentaje de sustituciones conservativas del 98% y 0% de gaps. Para la proteína defensina la secuencia más significativa presenta un e-value de $2e-22$, un porcentaje de identidad del 55%, un porcentaje de sustituciones conservativas del 67% y 0% de gaps. Finalmente, para la proteína drosomicina no se encontró ninguna secuencia significativa.

En el caso de hembras, al realizar al alineamiento con respecto a la proteína atacina la secuencia más significativa presenta un e-value de $1e-107$, un porcentaje de identidad del 80%, un porcentaje de sustituciones conservativas del 88% y 2% de gaps. La secuencia más significativa para la proteína cecropina presenta un e-value de $2e-28$, un porcentaje de identidad del 90%, un porcentaje de sustituciones conservativas del 98% y 0% de gaps. Para la proteína defensina la secuencia más significativa presenta un e-value de $1e-22$, un porcentaje de identidad del 55%, un porcentaje de sustituciones conservativas del 67% y 0% de gaps. Finalmente, para la proteína drosomicina no se encontró ninguna secuencia significativa.

Discusión

Los análisis de calidad mostraron fallos en el contenido de secuencia por base, niveles de duplicación y contenido de GC, siendo este último la advertencia más preocupante del análisis. Con respecto al contenido de secuencia por base las variaciones significativas de las 4 bases al inicio 5' de las secuencias son normales en datos Illumina y pueden deberse principalmente a un fallo en la producción de la biblioteca (Dudic *et al.*, 2021), razón por la cual al realizar la limpieza de la secuencia las primeras bases fueron cortadas.

Por otro lado, el sesgo producido en los niveles de duplicación no afectaron nuestro análisis, debido a que este valor no produce cambios significativos en la calidad general de los datos y se considera normal en secuencias de RNA debido a la presencia de lecturas largas, productos de corte y empalme, y secuencias repetidas (Ji y Sadreyev, 2018, Dudic *et al.*, 2021).

Finalmente en el contenido de GC, el pico obtenido puede relacionarse con contaminación por adaptadores u otros organismos, lecturas duplicadas y transposones

ricos en GC (Dudic *et al.*, 2021). La contaminación por adaptadores o por otros organismos no se consideran razones de explicación para esta anomalía, debido a que en la verificación de la calidad de los datos crudos y procesados no existió la presencia de adaptadores, y en el caso de existir contaminación por otros organismos se habría esperado una calidad menor en todas las secuencias, por lo tanto, se considera que esta variación puede deberse a la gran cantidad de elementos transponibles que en el género *Drosophila* se aproximan a 150 (Malone y Hannon, 2009), a su vez se debe tener en consideración la evolución del contenido de GC entre las diferentes especies que comprenden el género *Drosophila* (Rodríguez, Tarrío y Ayala, 2000) cuyos estudios son inexistentes en *D. mesophragmatica* y pueden ser motivo para investigaciones futuras.

Debido al alta calidad de las secuencias y teniendo en consideración los puntos antes mencionados las lecturas procesadas pasaron al filtro de ensamblaje. El ensamblaje de lecturas para machos y hembras se realizó teniendo en consideración el 42% de variabilidad del transcriptoma entre ambos sexos (Huang *et al.*, 2015), por lo tanto, la existencia de variaciones en la detección de proteínas. Para la evaluación de calidad el valor N50 no se tomó en consideración debido a la imprecisión de su resultado; razón por la cual se obtuvo los valores de coincidencia de ortólogos (Jauhal y Newcomb, 2021), en donde, ambos ensamblajes presentaron puntajes de identidad altos superiores al 90%. Por lo tanto, se considera a los ensamblajes con una buena calidad; con respecto al valor alto de secuencias duplicadas no se consideran significantes o que representen una baja calidad, esto se debe a que los transcritos pueden presentar duplicaciones o isoformas producidas por el empalme (Raghavan *et al.*, 2022).

Finalmente en la identificación de proteínas en el transcriptoma de *D. mesophragmatica* de machos y hembras se pudo encontrar la presencia de las proteínas atacina, cecropina y defensina mientras que la proteína drosomicina se encontraba ausente. Cabe recalcar que la estructura general y funcional de las proteínas presentan un alto grado de conservación, sin embargo, la variabilidad genética aumenta entre especies estrechamente relacionadas (Lazzaro y Clark, 2001) y puede ser explicada por cambios postraduccionales o ecológicos (Rabel *et al.*, 2004) (Hanson y Lamaitre, 2020). Por lo tanto, la ausencia de la proteína drosomicina indica un cambio evolutivo importante en la especie *D. mesophragmatica*.

En el caso de la proteína atacina el grado de conservación es relativamente alto debido al 80% de similitud existente, sin embargo el 88% de sustituciones conservativas

pueden ser explicadas por los altos niveles de recombinación meiótica de esta proteína lo cual ha generado una gran cantidad de haplotipos que a su vez son conservados entre taxones cercanos y distantes. La función principal de esta proteína es provocar la alteración de la membrana externa de bacterias gram negativas con el fin de aumentar su permeabilidad e inhibir la producción de proteínas (Lazzaro y Clark, 2001).

Para la proteína cecropina el porcentaje de similitud y de sustituciones conservativas es mayor en comparación con la proteína atacina con valores de 90% y 98% respectivamente. En este caso la proteína cecropina pertenece a una familia multigénica que incluye tanto genes funcionales y pseudogenes (Kylsten, Samakovlis y Hultmark, 1990) motivo por el cual se puede explicar su alto porcentaje de sustituciones conservativas. La función de esta proteína es contrarrestar la infección por bacterias gram negativas al igual que la proteína atacina, sin embargo, la cecropina actúa más al interior de la célula (Lazzaro y Clark, 2001).

Las proteínas defensina y drosomicina son las menos conservadas en *D. mesophragmatica*, en el caso de la defensina el porcentaje de similitud es del 55% mientras que la drosomicina se encuentra ausente. La variación en la presencia parcial y ausencia de estas proteínas está relacionada con patrones evolutivos, en donde, la variabilidad de los AMPs es afectada en su mayoría por eventos de duplicación los cuales pueden provocar ganancia, pérdida o cambios significativos en sus secuencias (Hanson, Lemaitre y Unckless, 2019). A su vez los diferentes grados de exposición a microorganismos pueden provocar una mayor o menor resistencia (Seto y Tamura, 2013), por ejemplo la ausencia de drosomicina en *D. mesophragmatica* puede relacionarse indirectamente con un hábitat más estéril. Sin embargo, existe evidencia de la presencia de esta proteína solo en especies del grupo *melanogaster* (Seto y Tamura, 2013).

Conclusiones y recomendaciones

Las especies del género *Drosophila* están expuestas a una gran cantidad de organismos patógenos, por lo tanto, su sistema inmunitario debe asegurar la supervivencia de estos organismos en diferentes hábitats. La identificación de proteínas antimicrobianas en *D. mesophragmatica* ha revelado la variabilidad existente en las proteínas atacina, cecropina, defensina y drosomicina de especies de un mismo taxón, las cuales pueden producirse por eventos ecológicos o evolutivos. Por lo tanto, la principal recomendación es utilizar un mayor número de especies del grupo

mesophragmatica y grupos externos en futuros análisis con el fin de obtener evidencia evolutiva robusta. A su vez, se determinó que las tres vías de señalización son activadas, sin embargo, la vía de señalización Toll no es activada ante la presencia de levaduras.

Referencias

Adamo S. A. (2017). The stress response and immune system share, borrow, and reconfigure their physiological network elements: Evidence from the insects. *Hormones and behavior*, 88, 25–30. <https://doi.org/10.1016/j.yhbeh.2016.10.003>

Behura S. K. (2015). Insect phylogenomics. *Insect molecular biology*, 24(4), 403–411. <https://doi.org/10.1111/imb.12174>

Banerjee, U., Girard, J. R., Goins, L. M., & Spratford, C. M. (2019). *Drosophila* as a Genetic Model for Hematopoiesis. *Genetics*, 211(2), 367–417. <https://doi.org/10.1534/genetics.118.300223>

Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. (s. f.). Recuperado 26 de septiembre de 2022, de <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Brcic, D. y Koref-Santibañez, S. 1957. The mesophragmatica group of species of *Drosophila*. *Evolution*. 11. 300–310

Brcic. (1957). *Drosophila camaronensis* Brcic, 1957.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC bioinformatics*, 10, 421. <https://doi.org/10.1186/1471-2105-10-421>

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)*, 34(17), i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>

Duda O. (1927). Die Südamerikanischen Drosophilidae unter Berücksichtigung auch der anderen Neotropischen sowie Nearktischen Arten. *Naturgesch.*

Dudić, D., Banović Deri, B., Pajić, V y Pavlović-Lažetić, G. (2021). Demystification of RNAseq Quality Control. *Journal of Information Technology and Applications*. 11. 73-86.

Dudzic, J. P., Kondo, S., Ueda, R., Bergman, C. M., & Lemaitre, B. (2015). *Drosophila* innate immunity: regional and functional specialization of prophenoloxidases. *BMC biology*, 13, 81. <https://doi.org/10.1186/s12915-015-0193-6>

Ekengren, S., & Hultmark, D. (2001). A family of Turandot-related genes in the humoral stress response of *Drosophila*. *Biochemical and biophysical research communications*, 284(4), 998–1003. <https://doi.org/10.1006/bbrc.2001.5067>

Figuero, L., Rafael, V., Cespédes, D. (2012). Grupo *Drosophila* asiri (Diptera, Drosophilidae), un nuevo grupo de especies andinas con la descripción de dos nuevas especies y la redescrición de *Drosophila* asiri. *Scielo*, 102(1), 33-42. <https://doi.org/10.1590/S0073-47212012000100005>

Gendrin, M., Welchman, D. P., Poidevin, M., Hervé, M., & Lemaitre, B. (2009). Long-range activation of systemic immunity through peptidoglycan diffusion in *Drosophila*. *PLoS pathogens*, 5(12), e1000694. <https://doi.org/10.1371/journal.ppat.1000694>

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644–652. <https://doi.org/10.1038/nbt.1883>

Hanson, M. A., Lemaitre, B., & Unckless, R. L. (2019). Dynamic Evolution of Antimicrobial Peptides Underscores Trade-Offs Between Immunity and Ecological Fitness. *Frontiers in immunology*, 10, 2620. <https://doi.org/10.3389/fimmu.2019.02620>

Hanson, M. A., Hamilton, P. T., & Perlman, S. J. (2016). Immune genes and divergent antimicrobial peptides in flies of the subgenus *Drosophila*. *BMC evolutionary biology*, 16(1), 228. <https://doi.org/10.1186/s12862-016-0805-y>

Hanson, M. A., & Lemaitre, B. (2020). New insights on *Drosophila* antimicrobial peptide function in host defense and beyond. *Current opinion in immunology*, 62, 22–30. <https://doi.org/10.1016/j.coi.2019.11.008>

Hegde, S., Sreejan, A., Gadgil, C. J., & Ratnaparkhi, G. S. (2022). SUMOylation of Dorsal attenuates Toll/NF- κ B signaling. *Genetics*, 221(3), iyac081. <https://doi.org/10.1093/genetics/iyac081>

- Hillyer, J. F., & Strand, M. R. (2014). Mosquito hemocyte-mediated immune responses. *Current opinion in insect science*, 3, 14–21. <https://doi.org/10.1016/j.cois.2014.07.002>
- Hillyer J. F. (2016). Insect immunology and hematopoiesis. *Developmental and comparative immunology*, 58, 102–118. <https://doi.org/10.1016/j.dci.2015.12.006>
- Hoffmann, J. A., & Reichhart, J. M. (2002). Drosophila innate immunity: an evolutionary perspective. *Nature immunology*, 3(2), 121–126. <https://doi.org/10.1038/ni0202-121>
- Huang, W., Carbone, M. A., Magwire, M. M., Peiffer, J. A., Lyman, R. F., Stone, E. A., Anholt, R. R., & Mackay, T. F. (2015). Genetic basis of transcriptome diversity in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 112(44), E6010–E6019. <https://doi.org/10.1073/pnas.1519159112>
- Igaki, T., Kanda, H., Yamamoto-Goto, Y., Kanuka, H., Kuranaga, E., Aigaki, T., & Miura, M. (2002). Eiger, a TNF superfamily ligand that triggers the *Drosophila* JNK pathway. *The EMBO journal*, 21(12), 3009–3018. <https://doi.org/10.1093/emboj/cdf306>
- Jaeger, C. Salzano F. (1953). *Drosophila gaucha*, a new species from Brazil. *Brazil. Biol.*
- Jang, I. H., Chosa, N., Kim, S. H., Nam, H. J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., Brey, P. T., & Lee, W. J. (2006). A Spätzle-processing enzyme required for toll signaling activation in *Drosophila* innate immunity. *Developmental cell*, 10(1), 45–55. <https://doi.org/10.1016/j.devcel.2005.11.013>
- Jauhal, A. A., & Newcomb, R. D. (2021). Assessing genome assembly quality prior to downstream analysis: N50 versus BUSCO. *Molecular ecology resources*, 21(5), 1416–1421. <https://doi.org/10.1111/1755-0998.13364>
- Ji, F., & Sadreyev, R. I. (2018). RNA-seq: Basic Bioinformatics Analysis. *Current protocols in molecular biology*, 124(1), e68. <https://doi.org/10.1002/cpmb.68>
- Kellermann, V., van Heerwaarden, B., Sgrò, C. M., & Hoffmann, A. A. (2009). Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science (New York, N.Y.)*, 325(5945), 1244–1246. <https://doi.org/10.1126/science.1175443>
- Kellermann, V., Overgaard, J., Hoffmann, A. A., Fløjgaard, C., Svenning, J. C., & Loeschcke, V. (2012). Upper thermal limits of *Drosophila* are linked to species

distributions and strongly constrained phylogenetically. *Proceedings of the National Academy of Sciences of the United States of America*, 109(40), 16228–16233. <https://doi.org/10.1073/pnas.1207553109>

Kurucz, E., Váczi, B., Márkus, R., Laurinyecz, B., Vilmos, P., Zsámboki, J., Csorba, K., Gateff, E., Hultmark, D., & Andó, I. (2007). Definition of *Drosophila* hemocyte subsets by cell-type specific antigens. *Acta biologica Hungarica*, 58 Suppl, 95–111. <https://doi.org/10.1556/ABiol.58.2007.Suppl.8>

Kylsten, P., Samakovlis, C., & Hultmark, D. (1990). The cecropin locus in *Drosophila*; a compact gene cluster involved in the response to infection. *The EMBO journal*, 9(1), 217–224. <https://doi.org/10.1002/j.1460-2075.1990.tb08098.x>

Larsen, A., Reynaldi, F., Novoa, E. (2019). Bases del sistema inmune de la abeja melífera (*Apis mellifera*). *SCIELO*, 10 (3). <https://www.scielo.org.mx/pdf/rmcp/v10n3/2448-6698-rmcp-10-03-705.pdf>

Lazzaro, B. P., & Clark, A. G. (2001). Evidence for recurrent paralogous gene conversion and exceptional allelic divergence in the Attacin genes of *Drosophila melanogaster*. *Genetics*, 159(2), 659–671. <https://doi.org/10.1093/genetics/159.2.659>

Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J. A., & Reichhart, J. M. (2002). A serpin mutant links Toll activation to melanization in the host defence of *Drosophila*. *The EMBO journal*, 21(23), 6330–6337. <https://doi.org/10.1093/emboj/cdf661>

Malone, C. D., & Hannon, G. J. (2009). Molecular evolution of piRNA and transposon control pathways in *Drosophila*. *Cold Spring Harbor symposia on quantitative biology*, 74, 225–234. <https://doi.org/10.1101/sqb.2009.74.052>

Mota, N. R., Robe, L. J., Valente, V. L., Budnik, M., & Loreto, E. L. (2008). Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. *Zoological science*, 25(5), 526–532. <https://doi.org/10.2108/zsj.25.526>

Rabel, D., Charlet, M., Ehret-Sabatier, L., Cavicchioli, L., Cudic, M., Otvos, L., Jr, & Bulet, P. (2004). Primary structure and in vitro antibacterial properties of the *Drosophila melanogaster* attacin C Pro-domain. *The Journal of biological chemistry*, 279(15), 14853–14859. <https://doi.org/10.1074/jbc.M313608200>

- Raghavan, V., Kraft, L., Mesny, F., & Rigerte, L. (2022). A simple guide to de novo transcriptome assembly and annotation. *Briefings in bioinformatics*, 23(2), bbab563. <https://doi.org/10.1093/bib/bbab563>
- Rodríguez-Trelles, F., Tarrío, R., & Ayala, F. J. (2000). Evidence for a high ancestral GC content in *Drosophila*. *Molecular biology and evolution*, 17(11), 1710–1717. <https://doi.org/10.1093/oxfordjournals.molbev.a026269>
- Thomas, G., Dohmen, E., Hughes, D., Murali, S. C., Poelchau, M., Glastad, K., Anstead, C. A., Ayoub, N. A., Batterham, P., Bellair, M., Binford, G. J., Chao, H., Chen, Y. H., Childers, C., Dinh, H., Doddapaneni, H. V., Duan, J. J., Dugan, S., Esposito, L. A., Friedrich, M., ... Richards, S. (2020). Gene content evolution in the arthropods. *Genome biology*, 21(1), 15. <https://doi.org/10.1186/s13059-019-1925-7>
- Netea, M. G., Schlitzer, A., Placek, K., Joosten, L., & Schultze, J. L. (2019). Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens. *Cell host & microbe*, 25(1), 13–26. <https://doi.org/10.1016/j.chom.2018.12.006>
- Seto, Y., & Tamura, K. (2013). Extensive differences in antifungal immune response in two *Drosophila* species revealed by comparative transcriptome analysis. *International journal of genomics*, 2013, 542139. <https://doi.org/10.1155/2013/542139>
- Sheehan, G., Garvey, A., Croke, M., & Kavanagh, K. (2018). Innate humoral immune defences in mammals and insects: The same, with differences ?. *Virulence*, 9(1), 1625–1639. <https://doi.org/10.1080/21505594.2018.1526531>
- Siva, M., Moret, Y., Rolff, Y. (2005). Insect Immunity: An Evolutionary Ecology Perspective. *Advances in Insect Physiology*, 32. https://biogeosciences.u-bourgogne.fr/documents/articles_pdf/2005_siva-jothy_AdvInsectPhysiol.pdf
- Strand, M. R. (2008). *The insect cellular immune response*. *Insect Science*, 15(1), 1–14. <https://doi.org/10.1111/j.1744-7917.2008.00183.x>
- Sulek, M., Kordaczuk, J., & Wojda, I. (2021). Current understanding of immune priming phenomena in insects. *Journal of invertebrate pathology*, 185, 107656. <https://doi.org/10.1016/j.jip.2021.107656>
- Vega, F., Kaya, H. (2012). *Insects pathology*. ELSEVIER. <https://books.google.com.ec/books?hl=es&lr=&id=gg0sDWP->

[hhkC&oi=fnd&pg=PP1&ots=RakwAq1Zv8&sig=I6CDS8wDQSo1sNrQAUs2DVfQVNE&redir_esc=y#v=onepage&q&f=false](https://doi.org/10.1093/molbev/msx319)

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V., & Zdobnov, E. M. (2018). BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular biology and evolution*, 35(3), 543–548. <https://doi.org/10.1093/molbev/msx319>

Yu, S., Luo, F., Xu, Y., Zhang, Y., & Jin, L. H. (2022). *Drosophila* Innate Immunity Involves Multiple Signaling Pathways and Coordinated Communication Between Different Tissues. *Frontiers in immunology*, 13, 905370. <https://doi.org/10.3389/fimmu.2022.905370>