

PONTIFICIA UNIVERSIDAD CATOLICA DEL ECUADOR
FACULTAD DE HÁBITAT, INFRAESTRUCTURA Y CREATIVIDAD
MAESTRÍA EN BIOLOGIA COMPUTACIONAL

Anotación funcional y comparación genómica de *D. quitensis* (Díptera, Drosophilidae)
perteneiente al grupo *guarani*.

Trabajo previo a la obtención del título de
Magister en Biología Computacional

WILSON FERNANDO PAREDES CEDEÑO

Quito, 2025

Índice de contenidos

Resumen	6
Capítulo I: Introducción	8
Planteamiento del problema	8
Justificación	9
Objetivos	10
Capítulo II : Fundamentación Teórica y referencial	10
Grupo guarani	11
Drosophila quitensis	12
Estudio de elementos móviles en Drosophila	13
Capítulo III: Metodología	15
Evaluación del ensamblaje (QUAST)	15
Evaluación funcional del ensamblaje (BUSCO)	15
Alineamiento de genomas (NUCmer y MuMmer)	16
Anotación estructural de genes (AUGUSTUS)	16
Anotación funcional (InterProScan)	16
Comparación genómica (Orthofinder)	17
Identificación y caracterización de elementos repetitivos (RepeatModeler y RepeatMasker).	17
Capítulo IV: Resultados	17
Calidad de Ensamblaje con QUAST Y BUSCO	17
Alineamiento genómico (MuMmer)	20
Anotación estructural	23
Anotación Funcional	23

Elementos repetitivos	24
Capítulo V: Discusión	25
Capítulo VI: Conclusiones	28
Referencias	30

Índice de tablas

Tabla 1. Pipeline del análisis de datos de la secuencia ensamblada <i>D. quitensis</i>	15
Tabla 2. Reporte principal de QCAST para el ensamblaje de <i>D. quitensis</i>	17
Tabla 3. Resultados de evaluación de ensamblaje en BUSCO	19
Tabla 4. Resumen de anotación estructural por AUGUSTUS	23
Tabla 5. Resumen de anotación funcional por Interproscan	23
Tabla 6. Comparación genómica entre <i>D. quitensis</i> , <i>D. melanogaster</i> y <i>D. mojavensis</i>	24
Tabla 7 Resumen de elementos repetitivos en Repeatmasker	24

Índice de Figuras

Figura 1 <i>Comando de ejecución BUSCO</i>	15
Figura 2 <i>Comando de ejecución AUGUSTUS</i>	16
Figura 3 <i>Comando de ejecución Interproscan</i>	16
Figura 4 <i>Distribución del contenido GC de D.quitensis</i>	18
Figura 5 <i>Porcentaje de BUSCO</i>	19
Figura 6 <i>Alineamiento entre D.mojavensis vs D. quitensis</i>	20
Figura 7 <i>Porcentaje de identidad y número de genes entre D.mojavensis vs D.quitensis</i>	20
Figura 8 <i>Genes divergentes D.mojavensis vs D. quitensis</i>	21
Figura 9 <i>Genes conservados entre D.mojavensis vs D. quitensis</i>	21
Figura 10 <i>Alineamiento entre D. melanogaster vs D. quitensis</i>	21
Figura 11 <i>Porcentaje de identidad y número de genes entre D.melanogaster vs D.quitensis</i>	21
Figura 12 <i>Genes conservados entre D.melanogaster vs D.quitensis</i>	22
Figura 13 <i>Genes más divergentes entre D.melanogaster vs D.quitensis</i>	22
Figura 14 <i>Distribución de longitudes de genes predichos en D. quitensis</i>	23

Resumen

El género *Drosophila*, ha sido históricamente un modelo fundamental en biología, parte de la diversidad del género permanece poco explorada, especialmente aquellas especies que habitan en regiones neotropicales y andinas. Una de ellas es *Drosophila quitensis*, descubierta en los bosques montanos del Ecuador y perteneciente al grupo guaraní, un linaje con gran variabilidad y escasa información genómica. Su presencia en ecosistemas de altura, sometidos a condiciones ambientales exigentes, convierte a esta especie en un organismo clave para comprender su adaptación a escenarios extremos de temperatura, oxigenación y estacionalidad.

En este trabajo se presenta el estudio de anotación funcional y caracterización del genoma de *D. quitensis*. Se partió desde el ensamblaje el cual alcanzó un tamaño cercano a los 323 Mb y mostró un nivel de completitud superior al 90 %. La anotación estructural y funcional permitió la identificación de cerca de 18.800 genes codificantes. La mayoría de estos genes se asociaron a funciones básicas de la célula, procesos metabólicos y respuestas al estrés.

El análisis comparativo con otras especies del género reveló regiones muy conservadas, pero también áreas divergentes que reflejan la particularidad evolutiva del grupo *guaraní*. Además, se observó una fracción considerable de elementos repetitivos, que abarcan casi la mitad del genoma. Entre ellos destacan los transposones Rolling-circles, cuya abundancia es inusualmente alta y sugiere un papel importante en la reorganización del material genético y en la generación de innovaciones adaptativas.

Palabras clave: Anotación, Transposones, *Drosophila quitensis*.

Abstract

The genus *Drosophila* has historically been a fundamental model in biology, yet much of its diversity remains poorly explored, particularly species inhabiting neotropical and Andean regions. One of them is *Drosophila quitensis*, discovered in the montane forests of Ecuador and belonging to the guaraní group, a lineage characterized by high variability but limited genomic information. Its presence in high-altitude ecosystems, exposed to demanding environmental conditions, makes this species a key organism for understanding adaptation to extreme scenarios of temperature, oxygen availability, and seasonality.

This study presents the functional annotation and genomic characterization of *D. quitensis*. The assembly reached a total size of approximately 323 Mb and showed a completeness level above 90%. Structural and functional annotation allowed the identification of nearly 18,800 protein-coding genes. Most of these genes were associated with fundamental cellular functions, metabolic processes, and stress responses.

Comparative analysis with other *Drosophila* species revealed highly conserved regions, as well as divergent areas that reflect the evolutionary uniqueness of the guaraní group. A substantial fraction of repetitive elements was also observed, accounting for nearly half of the genome. Among these, rolling-circle transposons stood out for their unusually high abundance, suggesting a major role in genome reorganization and the generation of adaptive innovations.

Keywords: Annotation, Transposons, *Drosophila quitensis*.

Capítulo I: Introducción

Planteamiento del problema

El género *Drosophila* ha sido ampliamente utilizado como modelo biológico para estudios genéticos, evolutivos y de biodiversidad (Jennings, 2011). Dentro de este género, el grupo *guarani* se distingue por su diversidad y adaptación en ambientes neotropicales, especialmente en ecosistemas montañosos (Thompson et al., 2021). Comprende diversas especies de moscas distribuidas principalmente en la región neotropical, especialmente en regiones andinas (Peñañiel-Vinueza & Rafael, 2018). A pesar de su relevancia ecológica y evolutiva, el conocimiento genómico de estas especies es limitado, la falta de secuencias completas y anotaciones precisas dificulta la comprensión que explican su adaptación y evolución.

Entre las especies que integran este grupo, *Drosophila quitensis* (*D. quitensis*) fue descrita en el año 2004 como una especie endémica de los Andes ecuatorianos, específicamente del volcán Pasochoa (Vela & Rafael, 2004). A pesar de su relevancia biogeográfica y su potencial importancia evolutiva dentro del grupo *guarani*, *D. quitensis* ha sido escasamente estudiada a nivel molecular y genómico limitando la comprensión sobre su biología (Vela & Rafael, 2004). Adams y colaboradores (2000), indican que la comparación con especies modelo como *D. melanogaster*, cuyo genoma ha sido secuenciado y anotado exhaustivamente podría revelar diferencias y similitudes clave en elementos genéticos móviles y estructuras génicas, aportando información valiosa sobre la evolución genómica en el género *Drosophila*.

Justificación

El grupo *guarani* del género *Drosophila* constituye una línea evolutiva de gran interés dentro de los drosophilidos neotropicales, especialmente por su diversificación en ecosistemas complejos y ambientes neotropicales (Figuro & Rafael, 2013). A diferencia de especies modelo como *Drosophila melanogaster* (*D. melanogaster*), cuyo genoma y funciones génicas han sido ampliamente caracterizados (Adams et al., 2000), las especies del grupo *guarani* permanecen escasamente exploradas desde una perspectiva genómica, lo que limita la comprensión de sus procesos evolutivos, ecológicos y adaptativos (Vela & Rafael, 2004).

Para ello la anotación funcional y comparación genómica de *D. quitensis* a partir de una secuencia ya conocida, permitirá, describir la organización de su genoma, identificar sus genes codificantes, proteínas, regiones reguladoras y elementos móviles, aportando información esencial sobre su regulación génica y su adaptación a los ecosistemas (Clark et al., 2007).

La comparación con otras especies de *Drosophila* pertenecientes a distintos grupos neotropicales servirá para detectar patrones de conservación y divergencia que ayudarán a reconstruir los procesos evolutivos que han dado lugar a la diversidad actual del género en la región (Clark et al., 2007).

Los aportes de la investigación permitirán fortalecer el conocimiento sobre la biodiversidad genética de *D. quitensis* generando recursos genómicos que beneficiarán al grupo guaraní del género *Drosophila*.

Objetivos

Objetivo General

- Describir la estructura genómica de la especie *D. quitensis* y compararla con especies de otros grupos de *Drosophila*.

Objetivos Específicos

- Anotar el genoma de *D. quitensis* identificando genes, proteínas y regiones reguladoras mediante herramientas bioinformáticas.
- Caracterizar la estructura y distribución de los elementos móviles, analizando su impacto en la organización genómica y evolución de la especie.
- Comparar la organización genómica de *D. quitensis* (grupo *guarani*) con *Drosophila mojavensis* (grupo *repleta*) y *D. melanogaster* (grupo *melannogaster*), identificando similitudes, diferencias y su relevancia en la evolución del género *Drosophila* en la región neotropical.

Capítulo II : Fundamentación Teórica y referencial

La biodiversidad andina se caracteriza por su alto grado de endemismo y complejidad ecológica, resultado de procesos geológicos, climáticos y evolutivos únicos que han favorecido la diversificación de múltiples organismos, incluyendo los insectos del género *Drosophila* (Sprengelmeyer et al., 2022). Dentro de este grupo, destacan especies adaptadas a los ecosistemas altoandinos, como *D. quitensis*, perteneciente al grupo *guarani* (Vela & Rafael, 2004).

Si bien *D. melanogaster* ha sido históricamente el referente central en estudios genómicos, la exploración de especies no modelo, como las del grupo *guarani*, ha sido escasa, especialmente a nivel molecular y funcional (Vela & Rafael, 2004). sobre herencia, mutación, especiación y evolución molecular (Thompson et al., 2021).

Hoy en día se han secuenciado múltiples especies del género, lo que ha permitido estudios comparativos que han revelado patrones de conservación y divergencia en partes del ADN que contiene genes esenciales y regiones reguladoras (Singh et al., 2009). Esto ha ampliado la comprensión de los mecanismos evolutivos dentro del linaje, identificando adaptaciones moleculares específicas a diversos nichos ecológicos, desde zonas áridas hasta ambientes tropicales (Sprengelmeyer et al., 2022).

La importancia del estudio de *D. quitensis*, como especie endémica de los Andes ecuatorianos, representa un modelo ideal para explorar adaptaciones como condiciones ambientales extremas que incluyen bajas temperaturas, alta radiación ultravioleta, baja presión de oxígeno y una marcada variabilidad térmica diaria, lo cual se vería influenciado a nivel genómico (Feng et al., 2024; Flatt, 2020). Su anotación funcional y comparación genómica permitirá, describir la organización de su genoma, identificar sus genes codificantes, proteínas, regiones reguladoras y elementos móviles, aportando información esencial sobre su regulación génica y su adaptación a los ecosistemas (Clark et al., 2007).

Grupo guarani

El grupo *guarani* del género *Drosophila* fue propuesto originalmente por Dobzhansky y Pavan en 1943 como un conjunto de especies neotropicales con distribución exclusiva en América del Sur. Está compuesto por especies neotropicales restringidas a América Central y del

Sur, muchas de ellas adaptadas a hábitats específicos como bosques nublados andinos (Dobzhansky & Pavan, 1943; Peñafiel-Vinueza & Rafael, 2018).

A lo largo de los años, este grupo ha ido creciendo hasta incluir al menos 24 especies reconocidas, entre ellas *D. quitensis*, descrita por Vela y Rafael en 2004 a partir de ejemplares recolectados en el bosque andino de Pasochoa en Ecuador (Vela & Rafael, 2004; Peñafiel-Vinueza & Rafael, 2018). Están distribuidas principalmente en regiones tropicales y subtropicales de Sudamérica, con presencia destacada en Brasil, Colombia, Perú, Bolivia y Ecuador (Llangarí-Arizo & Rafael, 2018).

Este grupo ofrece un modelo ideal para estudiar procesos de especiación, adaptación altitudinal y evolución de cariotipos en ambientes andinos permitiendo investigar eventos de diversificación reciente en los Andes del norte, donde varias especies coexisten en rangos geográficos relativamente pequeños, pero con marcada diferenciación morfológica y cromosómica (Peñafiel-Vinueza & Rafael, 2018).

Estudios recientes han incrementado notablemente la diversidad conocida del grupo en los Andes del sur, con la descripción de cinco nuevas especies: *D. zamorana*, *D. quinarensis*, *D. sachapuyu*, *D. caxarumi* y *D. misi* (Peñafiel-Vinueza & Rafael, 2018). Estas especies, además de compartir características morfológicas típicas del grupo, presentan una notable diferenciación genética y patrones de pigmentación, reflejando procesos de diversificación adaptativa en hábitats de bosque nublado (Peñafiel-Vinueza & Rafael, 2018).

Drosophila quitensis

Fue descrita en el año 2004 por Vela & Rafael, a partir de especímenes recolectados en el volcán Pasochoa, en la provincia de Pichincha, Ecuador. Se distingue por características morfológicas particulares en la terminalia del macho y patrones de pigmentación abdominal,

epandrio, placa anal y surestilo. A nivel taxonómico, *D. quitensis* pertenece al grupo *guarani*, pero hasta ahora no se había explorado a fondo desde el punto de vista molecular ni genómico. Dejando vacíos importantes en el entendimiento de su historia evolutiva, de cómo se diferencia de especies cercanas y de los mecanismos que le han permitido adaptarse a los ecosistemas andinos (Vela & Rafael, 2004). Hoy en día, gracias a las herramientas bioinformáticas, es posible responder a estas preguntas utilizando datos generados en especies que ya cuentan con secuencias genómicas disponibles, empleando herramientas bioinformáticas como QUAST y BUSCO, AUGUSTUS para predecir genes y evaluar calidad de secuencias, InterProScan para su caracterización funcional, y RepeatMasker para identificar los elementos repetitivos. MUMmer y OrthoFinder para comparar el genoma con especies modelo, haciendo posible la caracterización genómica y aportando información novedosa que ayuda a comprender mejor la diversidad y evolución.

Estudio de elementos móviles en Drosophila

El estudio del género *Drosophila* ha sido fundamental para comprender los mecanismos de transposición y su regulación, ofreciendo un modelo clave para el análisis evolutivo y epigenético de los elementos transponibles (TE) (Kapitonov & Jurka, 2003). Los TEs son componentes dinámicos del genoma que pueden movilizarse e insertarse en nuevas posiciones, influyendo en la organización y evolución del material genético (Mérel et al., 2020). En especies del género *Drosophila*, estos elementos representan entre el 15% y 25% del genoma, dependiendo de la especie, siendo responsables de una fracción significativa de la variabilidad estructural (Kapitonov & Jurka, 2003).

Los TE en *Drosophila* se clasifican en dos grandes clases: retrotransposones, que se movilizan mediante un intermediario de ARN (incluyen los elementos LINE, SINE y LTR), y transposones

de ADN, que se movilizan directamente mediante una enzima transposasa (Kaminker et al., 2002). Participan en la creación de nuevos elementos reguladores cis, modifican la expresión de genes vecinos e incluso generan innovaciones genéticas mediante la exaptación de sus secuencias se ha demostrado que algunos LTRs actúan como promotores alternativos y que ciertos TEs pueden responder a señales ambientales, favoreciendo la adaptación fenotípica rápida (Gebrie, 2023). En especies como *D. mojavensis*, del grupo *repleta*, se ha detectado una alta tasa de inserción de TEs, lo que sugiere que estos elementos podrían haber desempeñado un papel clave en su diversificación ecológica y genómica (Bergman & Bensasson, 2007). Sin embargo, en el caso de las especies ecuatorianas del grupo *guarani*, como *D. quitensis*, el conocimiento sobre la diversidad, abundancia y distribución de los TEs sigue siendo limitado (Clark et al., 2007). Se ha propuesto que los TEs pueden contribuir a procesos de especiación mediante la generación de incompatibilidades genómicas y cambios en la arquitectura de los cromosomas sexuales (Kofler, 2019; Obbard et al., 2008).

Para lograr comprender todo este proceso, el uso de herramientas como RepeatModeler y RepeatMasker permite identificar y clasificar familias de TEs en genomas ensamblados (Stanke & Morgenstern, 2005). Esta aproximación integrativa es crucial para caracterizar el impacto funcional de los TEs y su posible asociación con rasgos adaptativos en especies del grupo *guarani*.

Capítulo III: Metodología

Los análisis se realizaron combinando el uso de comandos en Ubuntu y la plataforma Galaxy, integrando ambas herramientas para optimizar los procesos bioinformáticos

Tabla 1. Pipeline del análisis de datos de la secuencia ensamblada *D. quitensis*.

Paso	Herramienta	Archivo de Entrada	Archivo de Salida
1. Evaluación del Ensamblaje	QUAST, BUSCO	Secuencia ensamblada (FASTA)	Estadísticas del ensamblaje (TXT, CSV)
2. Alineamiento	MUMmer3 (nucmer)	Secuencia ensamblada (FASTA), Genomas de referencia (FASTA)	Archivos de alineación (Delta, MAF, Dot Plot)
3. Identificación de Elementos Repetidos	RepeatModeler, RepeatMasker	Secuencias ensambladas (FASTA)	Base de datos de repeticiones (FASTA, GFF, TXT)
4. Anotación Estructural de Genes	AUGUSTUS.	Secuencia ensamblada (FASTA)	GFF con predicciones de genes
5. Anotación Funcional	InterProScan	GFF con genes anotados	(TSV, CSV)
6. Comparación Genómica con <i>D. melanogaster</i>	Orthofinder	Genes anotados (FASTA, GFF)	Listas de genes ortólogos y divergentes (CSV, TSV)
7. Análisis de Elementos Móviles	RepeatMasker, Dfam	Secuencia ensamblada (FASTA)	Distribución de elementos móviles (GFF, TXT)
8. Visualización de Resultados	IGV, R	Datos anotados y alineados	Diagramas de distribución cromosómica, dot plots

Evaluación del ensamblaje (QUAST)

La calidad estructural del ensamblaje genómico de *D. quitensis* fue evaluada mediante la herramienta QUAST (Quality Assessment Tool for Genome Assemblies).

Evaluación funcional del ensamblaje (BUSCO)

Para estimar la completitud funcional del ensamblaje genómico, se utilizó la herramienta BUSCO (Benchmarking Universal Single-Copy Orthologs).

Figura 1 Comando de ejecución BUSCO

```
47 2025-08-03 20:54:04 busco -i masuca_Dquitensis_final.genome.scf
.fasta -o busco_dquitensis -l /home/wilson/Documentos/busco_downloads
/lineages/diptera_odb10 -m genome --cpu 4 -f
```

Alineamiento de genomas (NUCmer y MuMmer)

Se empleó el programa NUCmer, incluido en la herramienta MUMmer4, se designó un genoma como referencia y el otro como query, y se ejecutó NUCmer/ MuMmer para obtener los alineamientos globales. obtener tanto tablas de coordenadas de alineamientos como representaciones gráficas

Anotación estructural de genes (AUGUSTUS)

La predicción de genes en el genoma de *D. quitensis* se realizó mediante la herramienta AUGUSTUS. Se utilizó el parámetro `--species=fly`, correspondiente al modelo entrenado en *D. melanogaster*. Permitiendo una predicción *ab initio*, siendo archivo de entrada el ensamblaje final.

Figura 2 Comando de ejecución AUGUSTUS

```
48 2025-08-03 20:54:04 augustus --help
49 2025-08-03 20:54:04 augustus --species=fly
50 2025-08-03 20:54:04 top
```

Anotación funcional (InterProScan)

Para la anotación funcional, se utilizó la herramienta (InterProScan). El análisis se realizó sobre las secuencias proteicas generadas a partir de la predicción estructural con AUGUSTUS. Se ejecutó sin pre-cálculos y con los modos de búsqueda completos activados.

Figura 3 Comando de ejecución Interproscan

```
695 2025-08-03 21:12:30 du -h ~ --max-depth=1 | sort -hr | head -n
10
696 2025-08-03 21:14:56 tar -czvf interproscan_respaldo.tgz ~/inte
rproscan-5.65-97.0
697 2025-08-04 23:23:12 HISTTIMEFORMAT="%F %T " History
```

Comparación genómica (Orthofinder)

En la comparación genómica, se utilizó la herramienta Orthofinder la misma que nos permite agrupar genes en ortogrupos, es decir, conjuntos de genes derivados de un mismo gen ancestral, y reconstruir relaciones filogenéticas tanto a nivel génico como de especie.

Identificación y caracterización de elementos repetitivos (RepeatModeler y RepeatMasker).

La identificación y caracterización de elementos repetitivos en el genoma de *D. quitensis* se realizó mediante un enfoque de novo utilizando las herramientas RepeatModeler y RepeatMasker.

RepeatModeler

Se ejecutó la herramienta RepeatModeler sobre el ensamblaje genómico, que permitió generar una biblioteca específica de secuencias repetitivas basadas en modelos de consenso y clasificadas por tipo y familia.

RepeatMasker

Para la anotación de elementos repetidos, se utilizó RepeatMasker empleando como biblioteca de entrada el conjunto de repeticiones generado previamente por RepeatModeler.

Capítulo IV: Resultados

Calidad de Ensamblaje con QUAST Y BUSCO

Tabla 2. Reporte principal de QUAST para el ensamblaje de *D. quitensis*

Métrica	Valor
# contigs (≥ 0 bp)	70751
# contigs (≥ 1000 bp)	37067
# contigs (≥ 5000 bp)	14979
# contigs (≥ 10000 bp)	9375
# contigs (≥ 25000 bp)	2952
# contigs (≥ 50000 bp)	580
Total length (≥ 0 bp)	323432929

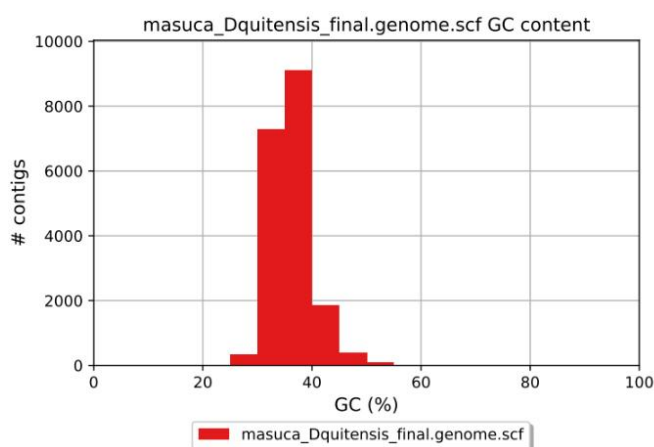
Total length (\geq 1000 bp)	306121781
Total length (\geq 5000 bp)	258872463
Total length (\geq 10000 bp)	220091179
Total length (\geq 25000 bp)	118710122
Total length (\geq 50000 bp)	34991282
# contigs	14913
Largest contig	1427780
Total length	275781862
GC (%)	36,09
N50	21572
N75	11671
L50	3820
L75	8152
# N's per 100 kbp	287,69

Nota: Todas las estadísticas se basan en contigs de tamaño \geq 3000 pb, a menos que se indique lo contrario (por ejemplo, '# contigs (\geq 0 pb)' y 'Longitud total (\geq 0 pb)' incluyen todos los contigs).

En la evaluación de la secuencia ensamblada de *D. quitensis*, se obtuvo una longitud total ensamblada de 275.8 Mb distribuida en 70,751 contigs, de los cuales 37,067 superan los 1 kb y 528 superan los 50 kb. El valor de N50 fue de 21,532 pb, indicando un ensamblaje moderadamente fragmentado, con un L50 de 3,820 contigs. La secuencia más larga alcanzó 142,780 pb y el contenido de GC fue del 36.09%.

Figura 4.

Figura 4 Distribución del contenido GC de *D. quitensis*

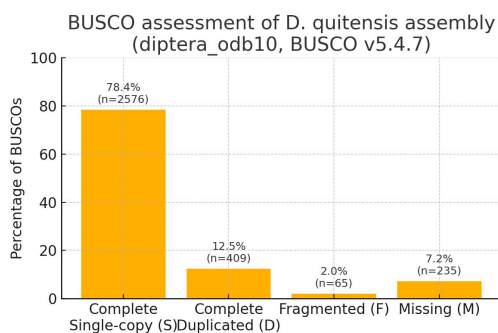


La distribución del contenido GC en el ensamblaje de *D. quitensis* muestra un rango concentrado entre 30% y 40%, con un valor promedio de aproximadamente, consistente con otros genomas del género *Drosophila*.

Tabla 3. Resultados de evaluación de ensamblaje en BUSCO

Categoría	Valor	Porcentaje
Complete BUSCOs (C)	2985	90.9
└ Single-copy (S)	2576	78.4
└ Duplicated (D)	409	12.5
Fragmented BUSCOs (F)	65	2.0
Missing BUSCOs (M)	235	7.1
Total BUSCO groups	3285	100.0
Number of scaffolds	70751	
Number of contigs	87988	
Total length (bp)	323324929	
Percent gaps	0.287%	
Scaffold N50 (bp)	17000	
Contigs N50 (bp)	10000	
Dependency: HMMER	3.1	
Dependency: BLAST+ (tools)	39.06	
Dependency: Metaeuk	7.bba0d80	
Dependency: BUSCO version	5.4.7	

Figura 5 Porcentaje de BUSCO



La evaluación de la completitud con BUSCO indicó que el ensamblaje presenta un alto nivel de integridad genómica, con un 90,9 % de genes completos, de los cuales la mayoría corresponden a copias únicas y un menor porcentaje a genes duplicados. Solo un 2 % se encontró fragmentado y un 7,1 % estuvo ausente, valores aceptables para un ensamblaje de novo.

En cuanto a las estadísticas generales, el genoma mostró una longitud total aproximada de 323 Mb, distribuida en 87.988 contigs agrupados en 70.751 scaffolds, lo que se encuentra dentro del rango esperado para especies del género *Drosophila*. Los valores de N50 (107 kb para scaffolds y 10 kb para contigs) confirman que se trata de un ensamblaje aún fragmentado; sin embargo, la completitud génica superior al 90 % respalda su calidad y lo hace apto para la anotación funcional y los análisis comparativos posteriores.

Alineamiento genómico (MuMmer)

Figura 6 Alineamiento entre *D.mojavensis* vs *D. quitensis*

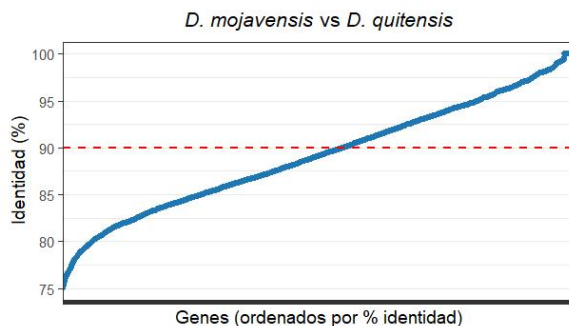


Figura 7 Porcentaje de identidad y número de genes entre *D.mojavensis* vs *D. quitensis*

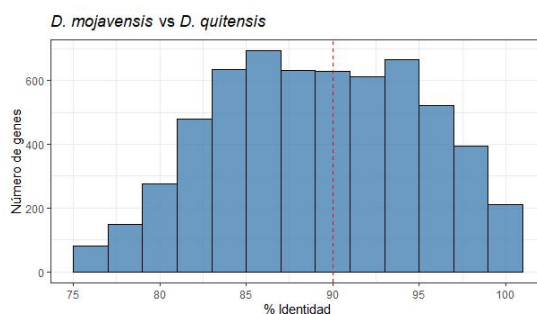


Figura 8 Genes divergentes *D.mojavensis* vs *D. quitensis*

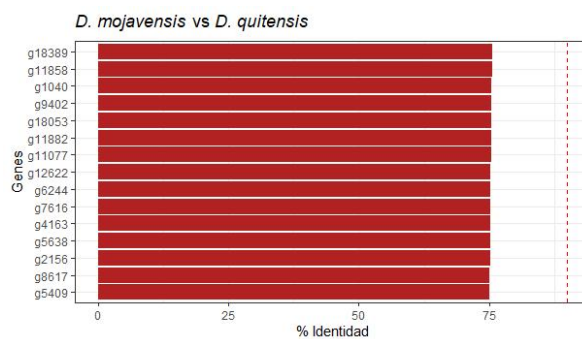


Figura 9 Genes conservados entre *D.mojavensis* vs *D. quitensis*

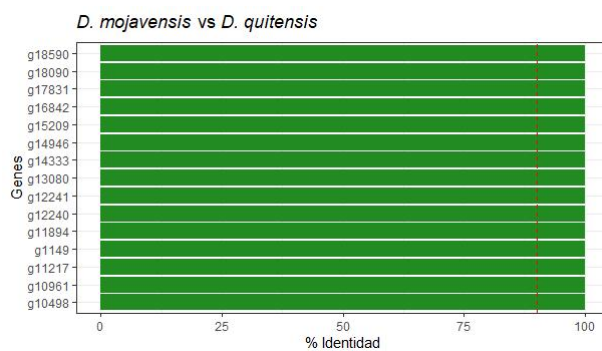


Figura 10 Alineamiento entre *D. melanogaster* vs *D. quitensis*

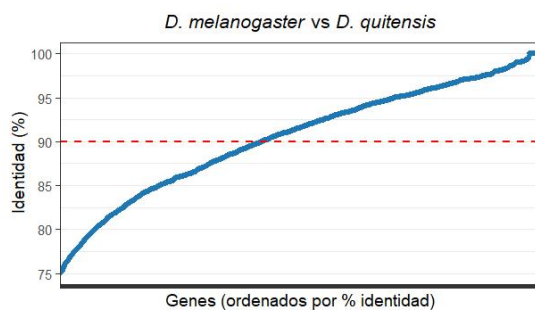


Figura 11 Porcentaje de identidad y número de genes entre *D.melanogaster* vs *D.quitensis*

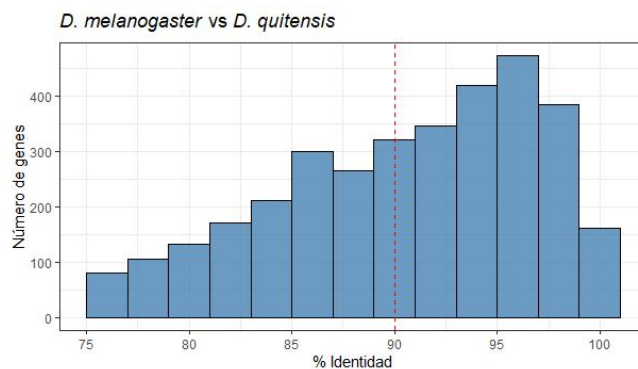


Figura 12 Genes conservados entre *D.melanogaster* vs *D.quitensis*

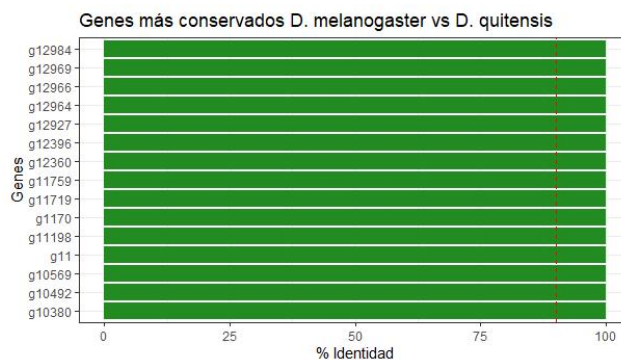
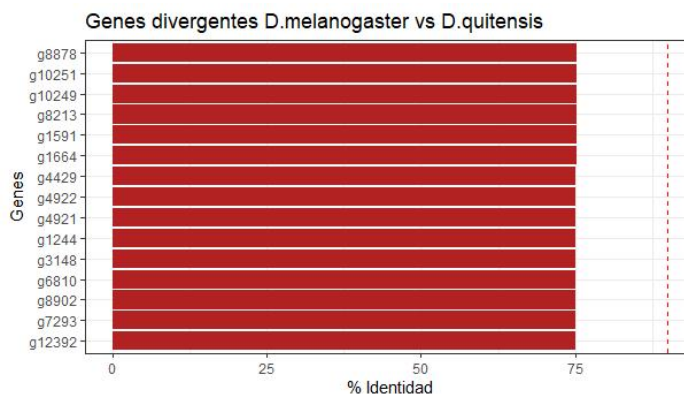


Figura 13 Genes más divergentes entre *D.melanogaster* vs *D.quitensis*



En los alineamientos genómicos realizados entre *D. mojavensis* y *D. quitensis* (Figuras 7–10), así como entre *D. melanogaster* y *D. quitensis* (Figuras 11–14), se observó una tendencia general a identidades de secuencia moderadas, con valores que en la mayoría de los genes superan el 75% de identidad. Las distribuciones muestran que una gran proporción de genes se agrupan alrededor de valores intermedios de identidad, mientras que los más conservados alcanzan identidades cercanas al 100%. En contraste, también se identificaron genes altamente divergentes, con identidades apenas superiores al umbral de filtrado (75%), lo que refleja posibles procesos de diferenciación adaptativa en *D. quitensis*. Los gráficos de barras permiten visualizar de forma clara los genes más conservados y los más divergentes para cada comparación, destacando que, aunque existe un núcleo de genes altamente preservados, la presencia de un número considerable

de genes divergentes evidencia la singularidad genómica de *D. quitensis* frente a especies modelo ampliamente estudiadas como *D. melanogaster* y *D. mojavensis*.

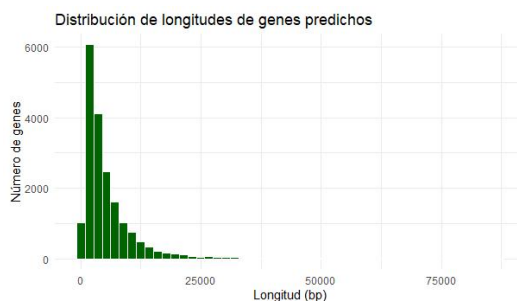
Anotación estructural

Tabla 4. Resumen de anotación estructural por AUGUSTUS

Métrica	Valor
Número de genes	18873
Número de transcritos	18873
Número de exones	79575
Número de intrones	49766
Longitud media de genes (bp)	5346
Longitud media de exones (bp)	586
Longitud media de intrones (bp)	762

Se predijo un total de 18873 genes codificantes y el mismo número de transcritos. En promedio, las longitudes medias de los exones e intrones son de 586 bp y 762 bp, respectivamente la longitud media de los genes fue de 5,346 bp.

Figura 14 Distribución de longitudes de genes predichos en *D. quitensis*



Anotación Funcional

Tabla 5. Resumen de anotación funcional por Interproscan

Métrica	Valor
Genes totales analizados	13781
Genes con términos GO	8355
Genes con anotación InterPro	13527
Genes con dominios Pfam	13781

Tabla 6. Comparación genómica entre *D. quitensis*, *D. melanogaster* y *D. mojavensis*.

Métrica	<i>D. quitensis</i>	<i>D. melanogaster</i>	<i>D. mojavensis</i>
Número de genes	18873	30802	23305
Número de genes en ortogrupos	16937	28949	22196
Número de genes no asignados	1936	1853	1109
Porcentaje de genes en ortogrupos	89,7	94	95,2
Porcentaje de genes no asignados	10,3	6	4,8
Número de ortogrupos que contienen la especie	10518	12754	12514
Porcentaje de ortogrupos que contienen la especie	74,9	96,8	89,1

Elementos repetitivos

El análisis mostró que, cerca del 45,1 % del genoma de *D. quitensis* está compuesto por secuencias repetitivas, lo que pone en evidencia una notable abundancia de elementos móviles dentro de su arquitectura genómica. Entre los grupos más representativos destacan los retroelements 9,97 %, especialmente los LINEs 9,87 %, donde sobresalen las familias L2/CR1/Rex 3,85 % y Jockey 3,22 %. También se identificaron elementos LTR del tipo Gypsy/DIRS1 2,6 %.

Por otro lado, los transposones de ADN constituyen un 2,06 % del total, con predominio de las familias hobo-Activator y Tc1-IS630-Pogo. Asimismo, los Rolling-circles asociados con mecanismos de transposición por círculo rodante, representan un porcentaje considerable 7,83 %, lo que sugiere que podrían desempeñar un papel en la captura y reorganización de fragmentos génicos, además de secuencias que permanecen sin clasificar 21,3 %.

Tabla 7 Resumen de elementos repetitivos en Repeatmasker

Category	Number of elements	Length (bp)	Percentage of sequence
Retroelements	71301	32142355	9.97 %
SINEs	2793	228808	0.07 %
Penelope	0	0	0.00 %
LINEs	39859	31913547	9.87 %
CRE/SLACS	0	0	0.00 %
L2/CR1/Rex	12343	12424969	3.85 %

R1/LOA/Jockey	19460	10420025	3.22 %
R2/R4/NeSL	0	0	0.00 %
RTE/Bov-B	244	42910	0.01 %
L1/CIN4	0	0	0.00 %
LTR elements	29494	12224905	3.79 %
BEL/Pao	259	112220	0.03 %
Ty1/Copia	578	692993	0.20 %
Gypsy/DIRS1	17576	8422503	2.60 %
Retroviral	1081	263989	0.08 %
DNA transposons	33653	6628093	2.06 %
hobo-Activator	7847	1188388	0.37 %
Tc1-IS630-Pogo	14083	3221199	1.00 %
En-Spm	0	0	0.00 %
MULE-MuDR	0	0	0.00 %
PiggyBac	284	87782	0.03 %
Tourist/Harbinger	689	60441	0.02 %
Other (Mirage, P-element, Transib)	259	80194	0.03 %
Rolling-circles	89294	25254751	7.83 %
Unclassified	330538	68655547	21.30 %
Total interspersed repeats:	107552503	107552503	33.36 %
Small RNA:	3172	408102	0.13 %
Satellites:	0	0	0.00 %
Simple repeats:	245415	10740608	3.33 %
Low complexity:	35159	1774823	0.55 %

Capítulo V: Discusión

El genoma ensamblado de *Drosophila quitensis* alcanzó un tamaño total de 275,8 Mb, distribuido en 70.751 contigs, con un valor de N50 de 21.532 pb. Aunque la fragmentación es evidente, el tamaño global se encuentra dentro del rango esperado para el género *Drosophila*: supera al de *D. melanogaster* (Adams et al., 2000; Ranz et al., 2007) y se aproxima al de *D. mojavensis* con 200 Mb aproximadamente (Richards et al., 2005) y *D. virilis* con 328 Mb (Clark et al., 2007). Estas diferencias entre especies reflejan los procesos agrandamiento o disminución de familias de elementos transponibles, ya que el número de genes codificantes suele permanecer

relativamente estable (Clark et al., 2007). El contenido de GC del ensamblaje (36,09 %) se ubica dentro del rango característico de las drosófilas (35–40 %), lo que descarta la presencia de contaminantes y complementa la calidad del ensamblaje (Seetharam & Stuart, 2012).

Si bien el bajo valor de N50 indica que sería deseable mejorar la continuidad con tecnologías de lectura larga, tal y como ya se ha demostrado en *D. suzukii* y *D. pseudoobscura* (Zhang et al., 2020), el ensamblaje obtenido para *D. quitensis* fue suficientemente robusto y útil para avanzar con análisis genómicos y comparativos. La complementación se la hizo con la evaluación en BUSCO la misma que mostró que el ensamblaje recupera un 90,9 % de genes completos, de los cuales un 78,4 % corresponden a copias únicas y un 12,5 % a duplicadas. Estos valores son comparables a los de genomas de referencia como *D. pseudoobscura* y *D. persimilis* (Richards et al., 2005), y apenas inferiores a *D. melanogaster*, donde la completitud supera generalmente el 95 % (Ranz et al., 2007). El porcentaje de duplicados probablemente se relacione con la fragmentación del ensamblaje, aunque no se descarta que refleje expansiones génicas genuinas, fenómeno descrito en otras especies neotropicales de *Drosophila* (Clark et al., 2007). Además que la baja proporción de genes ausentes respalda la fiabilidad del ensamblaje para estudios funcionales y evolutivos.

La predicción con AUGUSTUS identificó 18.873 genes codificantes, cifra superior a la reportada para *D. melanogaster* (Adams et al., 2000; Ranz et al., 2007) y *D. mojavensis* (Richards et al., 2005). Parte de esta diferencia podría explicarse por redundancias derivadas de la fragmentación, aunque también puede deberse a eventos de duplicación génica. Las longitudes promedio de genes (5.346 pb), exones (586 pb) e intrones (762 pb) resultaron consistentes con lo observado en otros drosófilos (Clark et al., 2007), respaldando así la validez biológica de la anotación.

En cuanto a la anotación funcional, los análisis con InterProScan permitieron vincular la mayoría de las proteínas a familias y dominios conservados, además de identificar su participación en rutas metabólicas clave. El valor de los términos GO más representativos estuvieron relacionados se los puede considerar con procesos celulares básicos, respuesta al estrés y metabolismo, en concordancia con lo observado en estudios de transcriptómica comparativa en diversas especies de *Drosophila* (Ma et al., 2018).

El alineamiento frente a *D. melanogaster* y *D. mojavensis* mostró regiones altamente conservadas con identidades superiores al 90 %, así como segmentos más divergentes. Estos patrones reflejan la distancia evolutiva esperada: *D. quitensis*, como parte del grupo guaraní, se encuentra más cercano a especies neotropicales como *D. mojavensis* y más alejado del subgrupo *melanogaster* (Seetharam & Stuart, 2012).

Los genes conservados evidencian la estabilidad funcional necesaria, mientras que los más divergentes sugieren adaptaciones específicas, destacándose aquellos vinculados a la respuesta inmune pudiendo destacarse la adaptación ambiental, en línea con la evolución dinámica de los sistemas de defensa en drosófilos (Obbard et al., 2009).

El análisis con RepeatMasker reveló que un gran porcentaje (33,36%) del genoma de *D. quitensis* corresponde a elementos repetitivos, un valor significativamente mayor al de *D. melanogaster* de 20–25 % (Kaminker et al., 2002; Bergman et al., 2006) y más cercano al de especies con genomas expandidos como *D. virilis* con 40 % (Clark et al., 2007) y *D. mojavensis* 35 % (Richards et al., 2005).

Dentro de este repertorio, los LINEs 9,9 % y los LTRs del tipo Gypsy/DIRS1 2,6 % fueron los más representativos, siguiendo un patrón observado en otras drosófilas neotropicales (Seetharam & Stuart, 2012). Los transposones de ADN alcanzaron el 2 %, con predominio de las

familias hobo-Activator y Tc1-IS630-Pogo, mientras que los Rolling-circles mostraron una abundancia particularmente alta en comparación con *D. melanogaster* (Mérel et al., 2020).

Los Rolling-circles conocidos por su capacidad de capturar fragmentos génicos y modificar la regulación transcripcional (Kapitonov & Jurka, 2001; Wicker et al., 2007), podrían ser los principales causantes de contribuir a procesos de innovación adaptativa en ambientes de estrés. Además existe 21,3 % de repeticiones no clasificadas siendo este un valor elevado que sugiere la existencia de familias aún desconocidas o exclusivas del grupo guaraní, lo que convierte a *D. quitensis* en un modelo importante de estudio en cuanto a la evolución de elementos transponibles en drosófilos.

Capítulo VI: Conclusiones

La anotación estructural y funcional del genoma de *Drosophila quitensis* permitió identificar un conjunto de genes de acuerdo a lo reportado en otras especies del género, esto genera confianza en la calidad del ensamblaje y ofrece una base sólida para explorar procesos biológicos clave como el metabolismo, la respuesta al estrés y la adaptación ambiental.

La caracterización de genes y elementos repetitivos generó datos particulares de expansión de Rolling-circles y otros transposones, sugiriendo que los elementos móviles han desempeñado un papel importante en la evolución del genoma de esta especie y generar la posibilidad de estar enlazados con los procesos de adaptación de *D. quitensis*.

El análisis comparativo con especies modelo como *D. melanogaster* y *D. mojavensis* ayudó a identificar regiones altamente conservadas, además de zonas divergentes que reflejan la historia evolutiva del grupo guaraní. Estos hallazgos resaltan la relevancia de *D. quitensis* como un modelo para estudiar los procesos de diversificación y adaptación en drosófilos neotropicales.

Referencias

- Acurio, A., Rafael, V., Céspedes, D., & Ruiz, A. (2013). Description of a New Spotted-Thorax *Drosophila* (Diptera: Drosophilidae) Species and its Evolutionary Relationships Inferred by a Cladistic Analysis of Morphological Traits. *Annals of the Entomological Society of America*, 106(6), 695-705. <https://doi.org/10.1603/AN13028>
- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., ... Venter, J. C. (2000). The genome sequence of *D. melanogaster*. *Science* (New York, N.Y.), 287(5461), 2185-2195. <https://doi.org/10.1126/science.287.5461.2185>
- Adrion, J. R., Song, M. J., Schrider, D. R., Hahn, M. W., & Schaack, S. (2017). Genome-Wide Estimates of Transposable Element Insertion and Deletion Rates in *D. melanogaster*. *Genome Biology and Evolution*, 9(5), 1329-1340. <https://doi.org/10.1093/gbe/evx050>
- Ayhan, N., Güler, P., & Onder, B. S. (2016). Variación altitudinal de la longevidad de las poblaciones de *D. melanogaster* del valle de Firtina, noreste de Turquía. *Journal of Thermal Biology*, 61, 91-97. <https://doi.org/10.1016/j.jtherbio.2016.09.002>
- Behura, S. K., Haugen, M., Flannery, E., Sarro, J., Tessier, C. R., Severson, D. W., & Duman-Scheel, M. (2011). Comparative Genomic Analysis of *D. melanogaster* and Vector Mosquito Developmental Genes. *PLOS ONE*, 6(7), e21504. <https://doi.org/10.1371/journal.pone.0021504>
- Bergman, C. M., & Bensasson, D. (2007). Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *D. melanogaster*. *Proceedings of the*

- National Academy of Sciences of the United States of America, 104(27), 11340-11345.
<https://doi.org/10.1073/pnas.0702552104>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.
<https://doi.org/10.1093/bioinformatics/btu170>
- Clark, A. G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., Kaufman, T. C., Kellis, M., Gelbart, W., Iyer, V. N., Pollard, D. A., Sackton, T. B., Larracunte, A. M., Singh, N. D., Abad, J. P., Abt, D. N., Adryan, B., Aguade, M., Akashi, H., ... *Broad Institute Genome Sequencing Platform. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, 450(7167), 203-218. <https://doi.org/10.1038/nature06341>
- Dobzhansky, T., & Pavan, C. (1943). Chromosome Complements of Some South-Brazilian Species of *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 29(11), 368-375. <https://doi.org/10.1073/pnas.29.11.368>
- Diversity and geographical distribution of *Drosophila* (Diptera, Drosophilidae) in Ecuador. (s. f.). ResearchGate. Recuperado 2 de marzo de 2025, de https://www.researchgate.net/publication/258879757_Diversity_and_geographical_distribution_of_Drosophila_Diptera_Drosophilidae_in_Ecuador
- Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1), 51-63. <https://doi.org/10.1016/j.tree.2013.09.008>
- Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1), 238. <https://doi.org/10.1186/s13059-019-1832-y>

- Feng, S., DeGrey, S. P., Guédot, C., Schoville, S. D., & Pool, J. E. (2024). Genomic Diversity Illuminates the Environmental Adaptation of *Drosophila suzukii*. *Genome Biology and Evolution*, 16(9), evae195. <https://doi.org/10.1093/gbe/evae195>
- Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nature Reviews Genetics*, 9(5), 397-405. <https://doi.org/10.1038/nrg2337>
- Gebrie, A. (2023). Transposable elements as essential elements in the control of gene expression. *Mobile DNA*, 14(1), 9. <https://doi.org/10.1186/s13100-023-00297-3>
- Figuro, M. L., & Rafael, V. (2013). Diversidad del género *Drosophila* (Diptera, Drosophilidae) en el páramo de Papallacta, Pichincha, Ecuador. *Revista Ecuatoriana de Medicina y Ciencias Biológicas*, 34(1-2), Article 1-2. <https://doi.org/10.26807/remcb.v34i1-2.241>
- Flatt, T. (2020). Life-History Evolution and the Genetics of Fitness Components in *D. melanogaster*. *Genetics*, 214(1), 3-48. <https://doi.org/10.1534/genetics.119.300160>
- Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, 36(8), 2628-2629. <https://doi.org/10.1093/bioinformatics/btz931>
- Jennings, B. H. (2011). *Drosophila*: Un modelo versátil en biología y medicina. *Materials Today*, 14(5), 190-195. [https://doi.org/10.1016/S1369-7021\(11\)70113-4](https://doi.org/10.1016/S1369-7021(11)70113-4)
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Leopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics (Oxford, England)*, 30(9), 1236-1240. <https://doi.org/10.1093/bioinformatics/btu031>

- Kaminker, J. S., Bergman, C. M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D. A., Lewis, S. E., Rubin, G. M., Ashburner, M., & Celniker, S. E. (2002). The transposable elements of the *D. melanogaster* euchromatin: A genomics perspective. *Genome Biology*, 3(12), research0084.1. <https://doi.org/10.1186/gb-2002-3-12-research0084>
- Kapitonov, V. V., & Jurka, J. (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nature Reviews. Genetics*, 9(5), 411-412; author reply 414. <https://doi.org/10.1038/nrg2165-c1>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780. <https://doi.org/10.1093/molbev/mst010>
- khuram.shahzad@yahoo.com. (2024, enero 6). The genetic diversity of *Drosophila* flies based on cytochrome-c oxidase subunit 1 gene from North Sulawesi [Text]. *Advancements in Life Sciences*. <https://www.als-journal.com/1049-23/>
- Kofler, R. (2019). Dynamics of Transposable Element Invasions with piRNA Clusters. *Molecular Biology and Evolution*, 36(7), 1457-1472. <https://doi.org/10.1093/molbev/msz079>
- Koonin, E. V. (2012). *The logic of chance: The nature and origin of biological evolution*. Pearson Education.
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, 5(2), R12. <https://doi.org/10.1186/gb-2004-5-2-r12>

- Llangarí-Arizo, L. M., & Rafael, V. (2018). Cuatro especies nuevas del género *Drosophila* (Diptera, Drosophilidae) en las provincias de Pichincha, Napo y Santo Domingo de los Tsáchilas, Ecuador. *Iheringia. Série Zoologia*, 108, e2018040.
<https://doi.org/10.1590/1678-4766e2018040>
- Ma, S., Avanesov, A. S., Porter, E., Lee, B. C., Mariotti, M., Zenskaya, N., Guigo, R., Moskalev, A. A., & Gladyshev, V. N. (2018). Comparative transcriptomics across 14 *Drosophila* species reveals signatures of longevity. *Aging Cell*, 17(4), e12740.
<https://doi.org/10.1111/accel.12740>
- Mérel, V., Boulesteix, M., Fablet, M., & Vieira, C. (2020). Transposable elements in *Drosophila*. *Mobile DNA*, 11(1), 23. <https://doi.org/10.1186/s13100-020-00213-z>
- Obbard, D. J., Gordon, K. H. J., Buck, A. H., & Jiggins, F. M. (2009). The evolution of RNAi as a defence against viruses and transposable elements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1513), 99-115.
<https://doi.org/10.1098/rstb.2008.0168>
- Oti, M., Pane, A., & Sammeth, M. (2018). Comparative Genomics in *Drosophila*. En J. C. Setubal, J. Stoye, & P. F. Stadler (Eds.), *Comparative Genomics: Methods and Protocols* (pp. 433-450). Springer. https://doi.org/10.1007/978-1-4939-7463-4_17
- (PDF) Karyotype Description of Two Andean Species of the guarani Group of *Drosophila* (Diptera: Drosophilidae) and Cytological Notes. (2024). ResearchGate.
<https://doi.org/10.1093/jisesa/icab032>
- Peñafiel-Vinueza, A. D., & Rafael, V. (2018). Five new species of *Drosophila* guarani group from the Andes of southern Ecuador (Diptera, Drosophilidae). *ZooKeys*, 781, 141-163.
<https://doi.org/10.3897/zookeys.781.22841>

- Ranz, J. M., Maurin, D., Chan, Y. S., Grotthuss, M. von, Hillier, L. W., Roote, J., Ashburner, M., & Bergman, C. M. (2007). Principles of Genome Evolution in the *D. melanogaster* Species Group. *PLOS Biology*, 5(6), e152. <https://doi.org/10.1371/journal.pbio.0050152>
- Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Meisel, R. P., Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussemaker, H. J., van Batenburg, M. F., Howells, S. L., Scherer, S. E., ... Gibbs, R. A. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Research*, 15(1), 1-18. <https://doi.org/10.1101/gr.3059305>
- Seetharam, A., & Stuart, G. W. (2012). Whole genome phylogenies for multiple *Drosophila* species. *BMC Research Notes*, 5(1), 670. <https://doi.org/10.1186/1756-0500-5-670>
- Singh, N. D., Larracunte, A. M., Sackton, T. B., & Clark, A. G. (2009). Comparative Genomics on the *Drosophila* Phylogenetic Tree. *Annual Review of Ecology, Evolution, and Systematics*, 40(Volume 40, 2009), 459-480. <https://doi.org/10.1146/annurev.ecolsys.110308.120214>
- Sprenkelmeyer, Q. D., Lack, J. B., Braun, D. T., Monette, M. J., & Pool, J. E. (2022). The evolution of larger size in high-altitude *D. melanogaster* has a variable genetic architecture. *G3: Genes|Genomes|Genetics*, 12(3), jkab454. <https://doi.org/10.1093/g3journal/jkab454>
- Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, 33(suppl_2), W465-W467. <https://doi.org/10.1093/nar/gki458>

- Thompson, J., Ross, D., & Hellack, J. (2021). Dros_Inf_Serv_104_2021.pdf.
https://www.ou.edu/journals/dis/DIS104/Dros_Inf_Serv_104_2021.pdf
- Vela, D., & Rafael, V. (2004). Tres nuevas especies del grupo *guarani*, genero *Drosophila* (Diptera, Drosophilidae) en el Bosque Pasochoa, Provincia de Pichincha. *Revista Ecuatoriana de Medicina y Ciencias Biológicas: REMCB*, 26(1-2), 14-21.
- Vela, D., & Villavicencio, E. (2021). Karyotype Description of Two Andean Species of the *guarani* Group of *Drosophila* (Diptera: Drosophilidae) and Cytological Notes. *Journal of Insect Science (Online)*, 21(3), 6. <https://doi.org/10.1093/jisesa/ieab032>
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, 8(12), 973-982. <https://doi.org/10.1038/nrg2165>
- Young, I., & Coleman, A. W. (2004). Las ventajas de la región ITS2 del cistron nuclear del ADNr para el análisis de las relaciones filogenéticas de los insectos: Un ejemplo de *Drosophila*. *Molecular Phylogenetics and Evolution*, 30(1), 236-242.
[https://doi.org/10.1016/S1055-7903\(03\)00178-7](https://doi.org/10.1016/S1055-7903(03)00178-7)
- Zhang, H., Mitchell-Olds, T., Mujacic, I., & Song, B.-H. (2020). De novo Genome Assembly, Annotation, and SNP Identification of an Endangered Rockcress, *Boechera fecunda*. *Frontiers in Ecology and Evolution*, 8. <https://doi.org/10.3389/fevo.2020.550936>

