

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE CIENCIAS EXACTAS Y NATURALES

ESCUELA DE BIOLOGÍA

Ensamblaje del genoma mitocondrial de las especies andinas *Drosophila cashapamba*, *Drosophila machachensis*, *Drosophila mesophragmatica*, *Drosophila guayllabambae* y *Drosophila papallacta* (Diptera, Drosophilidae).

Disertación previa a la obtención del título de Biólogo

Vera Díaz Carlos Leonardo

Quito, 2023

Certifico que la Disertación de Licenciatura en Ciencias Biológicas del estudiante Carlos Leonardo Vera Díaz ha sido concluida de conformidad con las normas establecidas; por lo tanto, puede ser presentada para la calificación correspondiente.

Dra. Doris Jimena Vela Peralta

Directora de la Disertación

Quito, 12 de junio de 2023

Dedicatoria

La presente tesis la dedico en primer lugar a Dios, quién me ha guiado por el camino de la educación y de la bondad. Por haberme puesto en una familia alegre y buena. Por no dejarme caer en la mala vida y por darme fuerzas para ser una mejor persona cada día que pasa.

A mi familia, a la cual le debo todo lo que soy actualmente. A mis padres, por haber educado con amor y disciplina, por los buenos consejos y por el afecto constante que han demostrado. A mis hermanos, mis compañeros en crimen en las travesuras, por su compañía en los malos momentos y por la buena compañía que proveen.

A la Pontificia Universidad Católica del Ecuador, la institución que me abrió las puertas para poderme realizar como estudiante universitario y poder entablar mi futuro como investigador.

A mis profesores, quienes me impartieron su conocimiento, permitiéndome elevar mi nivel de conocimiento. Especialmente a la Dra. Vela Peralta Doris Jimena, quien me adoptó bajo su ala académica y me continúa llevando hacia la excelencia académica.

A mi pareja, Alcoser Pinto Samanta Mayte, que me inspira a salir adelante y ser un excelente profesional.

Agradecimiento

Quisiera agradecer a Dios, por todo lo que ha hecho por mí, darme unos buenos padres, una buena vida; por enseñarme el valor de cada persona que he conocido y por ser quien me ha enseñado a como ser mejor persona. Quisiera también agradecer a mis padres, que sin su apoyo incondicional no hubiera podido llegar tan lejos; los quiero con todo mi corazón y a sobremanera. Además, también a mis hermanos, quienes también me han apoyado y han sido lugares seguros de conversación. A mi tutora, Dr. Doris Vela, quien me ha guiado en la vida académica y que me ha proporcionado la oportunidad de poder observar el mundo de la investigación. Quisiera agradecer a mis compañeros, con los que he compartido la experiencia del aprendizaje a lo largo de estos años. Y, además, quisiera agradecer a mi pareja, Samanta Alcoser, que me ha enseñado a ser más maduro y tomar oportunidades cuando las tenga.

Muchas gracias con todos.

INDICE

RESUMEN	8
ABSTRACT	9
INTRODUCCIÓN	10
OBJETIVOS	14
OBJETIVO GENERAL	14
OBJETIVOS ESPECÍFICOS	14
MATERIALES Y MÉTODOS.....	14
Obtención de los genomas de referencia.....	14
Ensamblado mitocondrial y análisis	15
Resultados	15
Calidad de las lecturas.....	15
Figura de puntos de las dos referencias	16
<i>D. melanogaster</i> vs <i>D. mojavensis</i>	16
Figuras de puntos de <i>D. melanogaster</i> como referencia.....	16
<i>D. melanogaster</i> vs <i>D. cashapamba</i>	16
<i>D. melanogaster</i> vs <i>D. guayllabambae</i>	16
<i>D. melanogaster</i> vs <i>D. machachensis</i>	17
<i>D. melanogaster</i> vs <i>D. mesophragmatica</i>	17
<i>D. melanogaster</i> vs <i>D. papallacta</i>	17
Figuras de puntos de <i>D. mojavensis</i> como referencia	17
<i>D. mojavensis</i> vs <i>D. cashapamba</i>	17
<i>D. mojavensis</i> vs <i>D. guayllabambae</i>	17
<i>D. mojavensis</i> vs <i>D. machachensis</i>	18
<i>D. mojavensis</i> vs <i>D. mesophragmatica</i>	18
<i>D. mojavensis</i> vs <i>D. papallacta</i>	18
Figuras de puntos de los dos ensamblajes de especies	18
<i>D. cashapamba</i> (mojavensis) vs <i>D. cashapamba</i> (melanogaster).....	18
<i>D. gauayllabambae</i> (mojavensis) vs <i>D. guayllabambae</i> (melanogaster) ...	18
<i>D. machachensis</i> (mojavensis) vs <i>D. machachensis</i> (melanogaster)	19
<i>D. mesophragmatica</i> (mojavensis) vs <i>D. mesophragmatica</i> (melanogaster)	19
<i>D. papallacta</i> (mojavensis) vs <i>D. papallacta</i> (melanogaster).....	19
Análisis de secuencias por medio de Jalview	19
DISCUSIÓN	19

CONCLUSIONES.....	21
REFERENCIAS.....	21
ANEXOS	24
FIGURAS	24
Comparación entre referencias.....	24
Ensamblados de <i>D. melanogaster</i>	24
Ensamblados de <i>D. mojavensis</i>	26
Comparaciones entre ensamblajes	29
TABLAS	32

RESUMEN

El género *Drosophila* comprende muchas especies distribuidas a nivel global. Dentro de este género, se encuentran las especies andinas del Ecuador *D. cashapamba*, *D. guayllabambae*, *D. machachensis*, *D. papallacta* y *D. mesophragmatica* de las cuales no se posee información molecular suficiente. Este estudio busca comparar los ensamblajes mitocondriales de estas especies al utilizar diferentes referencias, *D. melanogaster* y *D. mojavensis*. La razón de utilizar estas dos especies como referencias radica en el origen geográfico de ellas; *D. melanogaster* es una especie africana, mientras que *D. mojavensis* es una especie norteamericana. Para realizar los ensamblajes se utilizó la herramienta bioinformática de NOVOPlasty, las referencias serán procuradas del Centro Nacional para Información Biotecnológica, NCBI. Posterior al ensamblaje de las especies andinas ecuatorianas, se realizó el análisis comparativo mediante herramientas bioinformáticas. Se obtuvieron 10 ensamblajes en total, dos por cada especie, y 16 figuras dot plot, una comparando las referencias, 10 comparando el ensamblaje de las especies versus las referencias, y 5 comparando los ensamblajes de una misma especie. Estos resultados demostraron que ambas referencias arrojaban resultados similares, con unas pequeñas diferencias, siendo fragmentos de la región A+T de *D. melanogaster*. Hubo un caso extraordinario de *D. papallacta* cuando se utilizó a *D. mojavensis* como referencias, obteniendo un ensamblaje de 727 pares de bases, que no se relaciona a la diferencia entre los genomas mitocondriales de referencias. Se recomienda más análisis para determinar la razón de este caso.

Palabras claves: Análisis, bioinformática, ensamblaje, mitocondrial, molecular.

ABSTRACT

The genus *Drosophila* comprises many species distributed globally. Within this genus, there are the Andean species of Ecuador *D. cashapamba*, *D. guayllabambae*, *D. machachensis*, *D. papallacta* and *D. mesophragmatica* for which there is insufficient molecular information. This study seeks to compare the mitochondrial assemblies of these species using different references, *D. melanogaster* and *D. mojavensis*. The reason for using these two species as references lies in their geographical origin; *D. melanogaster* is an African species, while *D. mojavensis* is a North American species. To carry out the assemblies, the NOVOPlasty bioinformatics tool was used, the references will be obtained from the National Center for Biotechnology Information, NCBI. After the assembly of the Ecuadorian Andean species, the comparative analysis was carried out using bioinformatics tools. A total of 10 assemblages were obtained, two for each species, and 16 dot plot figures, one comparing the references, 10 comparing the assemblage of the species versus the references, and 5 comparing the assemblages of the same species. These results showed that both references gave similar results, with a few small differences, being fragments of the A+T region of *D. melanogaster*. There was an extraordinary case of *D. papallacta* when *D. mojavensis* was used as references, obtaining an assembly of 727 base pairs, which is not related to the difference between the reference mitochondrial genomes. Further analysis is recommended to determine the reason for this case.

Keywords: Analysis, assembly, bioinformatics, mitochondrial, molecular.

INTRODUCCIÓN

El género *Drosophila* ha tenido un gran trayecto experimental desde los primeros experimentos con Castle y los experimentos genéticos de Morgan a inicios del siglo anterior (Yamaguchi y Yoshida, 2018). De este género se destaca principalmente *Drosophila melanogaster* (Meigen, 1830) debido a que sigue siendo una especie modelo usada comúnmente para estudios en las ciencias biomédicas gracias a su bajo costo de mantenimiento, rápida producción de generaciones y al hecho de que comparte alrededor del 60% de material genético con los humanos. Además, el estudio de embriones de *D. melanogaster* ha permitido avances significativos de investigación sobre genes relacionados con el desarrollo embrionario, que ahora sabemos son homólogos a los de los mamíferos (Tolwinski, 2017). Por otro lado, estudios genéticos sobre el microbioma intestinal de *Drosophila* han dado pautas importantes para manejar la salud y bienestar físico de los animales, a través de la regulación del microbiota utilizando recursos genéticos y genómicos (Douglas, 2018). Esta especie posee uno de los genomas mitocondriales más grandes del género *Drosophila*, teniendo 19 524 pb (Wan y Celniker, 2014).

Los estudios filogenéticos están relacionados con lo que se conoce como el árbol de la vida. Actualmente se lo conoce como el árbol temporal evolucionario de la vida y contiene la información de las relaciones filogenéticas entre especies y su correspondiente historia evolutiva. Este concepto está en constante ampliación, ya que van apareciendo nuevas técnicas de estudios filogenéticos, por ejemplo, los estudios genómicos mitocondriales, que ofrecen nuevas luces sobre las historias evolutivas de las especies (Hedges, et al. 2015). El acceso a información genómica, no solamente genes aislados, provee de una importante oportunidad para los estudios filogenéticos porque reduce el nivel de incongruencias resultantes de los estudios genéticos en los cuales solo se utilizaban unos cuantos genes. El beneficio principal de utilizar esta información es el análisis de procesos genéticos a lo largo de la historia (Liu, et al. 2015).

El desarrollo de tecnologías de secuenciación de siguiente generación ha provocado la producción de cantidades grandes de información genética y genómica; tanto nuclear como extra nuclear; por lo que la utilización de

herramientas bioinformáticas es necesaria para interpretar dicha cantidad de datos (Oliver, et al., 2015). Una de estas herramientas es NOVOPlasty, un algoritmo de ensamblaje *de novo* diseñado para el ensamblaje de genomas mitocondriales y de cloroplastos (Dierckxsens, et al., 2017) los cuales son de gran importancia filogenética y taxonómica (Brozynska, et al., 2014). Los genomas mitocondriales debido a su herencia materna guardan mucha información relevante, tanto para estudios filogenéticos, como para entender los procesos adaptativos que han permitido la alta radiación de *Drosophila* en la región Neotropical (Robe et al. 2005). Por tanto, los genomas mitocondriales ensamblados serían un aporte relevante para esclarecer las relaciones evolutivas de las especies andinas de *Drosophila* en Ecuador.

La especie *Drosophila mojavensis* (Patterson, 1940) es miembro del grupo repleta, cuyas especies se caracterizan por ser cactofílicas. *D. mojavensis* se distribuye principalmente en los desiertos del suroeste de Estados Unidos y en México. Además, es una de las 12 especies de *Drosophila* de las cuales está anotado su genoma secuenciado (2007). Esta especie ha sido de gran interés gracias a su capacidad de coevolución con las especies de cactus, denotando su rápida capacidad evolutiva (Benowitz, et al. 2020). Según los datos de GenBank, esta especie posee un tamaño de genoma mitocondrial de 14 904 pb (Drosophila 12 Genome Consortium, 2007).

En los animales, incluyendo insectos, y por ende al género *Drosophila*, la mitocondria cumple un papel de gran importancia, siendo este la respiración aerobia (Boore, 1999). Se pensaba que, al tener genes de importancia para la esta respiración, la mitocondria evolucionaba neutralmente, sino que tiene una tasa de mutación muy elevada por lo que depende de purificación de mutaciones (Montooth, et al, 2009). En *Drosophila*, la tasa de acumulación de mutaciones supera el 95% del genoma mitocondrial, pero la razón entre mutaciones no sinónimas y sinónimas es de 10 a 36, gracias a lo mencionado anteriormente. Además, es posible que algunas de estas mutaciones no sinónimas, o deletéreas, se fijen en diferentes organismos promoviendo la especiación (Montooth, et al, 2009). Dicha fijación de mutaciones deletéreas puede suceder

gracias al evento de cuello de botella que sufre el genoma mitocondrial durante la oogénesis (Montooth, et al, 2009). Además de participar en la respiración celular, la mitocondria también tiene papeles importantes dentro de las enfermedades; control del ciclo celular, o apoptosis; y envejecimiento (Boore, 1999).

El genoma mitocondrial animal tiene entre 15 000 y 20 000 pares de base de longitud, conteniendo 37 genes (Boore, 1999). Esta longitud es semi-conservada, ya que muchos organismos pueden variar este tamaño (Wolff, et al, 2016). Por ejemplo, en insectos, el tamaño mitocondrial es un poco más pequeño, de alrededor de 15 000 pares de base; característica que se mantiene en el género *Drosophila* (Kim, et al, 2016). Aunque, hay especies que superan este número gracias a aspectos específicos de las especies; por ejemplo *Drosophila melanogaster* y *Drosophila suzukii* tienen genomas más grandes, de 19 524 pb (Wan y Celniker, 2014) y 16 230 pb (Kim, et al, 2016) respectivamente. Esta diferencia en el tamaño del genoma mitocondrial reside en la región control, la cual es rica en secciones A+T (Wolff, et al, 2016). En *D. melanogaster* esta región de control ha sufrido varios eventos de replicación, lo que causó el crecimiento del genoma mitocondrial, de igual forma, la función biológica de esta región se desconoce (FlyBase, n.d).

Las especies andinas *Drosophila cashapamba* (Céspedes y Rafael, 2012), *Drosophila machachensis* (Vela y Rafael 2001), *Drosophila mesophragmatica* (Duda, 1927), *Drosophila guayllabambae* (Rafael y Arcos, 1988), y *Drosophila papallacta* (Figuro y Rafael, 2013) son especies distribuidas en la región andina de Ecuador y Perú. Estas especies pertenecen a diferentes grupos de especie de *Drosophila*. Existe un estudio realizado por Figuro, 2017, en la que se analiza la filogenia molecular de algunas de estas especies, excluyendo la especie de *D. machachensis*. En este estudio se ensamblaron los genomas de algunas especies con la metodología DeNOVO que ofrece Geneius, que a diferencia de este estudio, no utiliza un genoma mitocondrial de otra especie como referencia.

La herramienta NOVOplasty, utiliza un genoma referencia para ensamblar los nuevos genomas mitocondriales, por lo cual el genoma referencia podrían influir

en la calidad del genoma ensamblado. En este estudio se comparará la calidad del ensamblaje del genoma mitocondrial de cinco especies andinas de *Drosophila* al usar dos diferentes genomas referencia. Las especies cuyos genomas serán utilizados como referencia son *D. melanogaster* la cual es originaria de África (Lachaise et al. 1988) y *D. mojavensis* que es una especie Neártica (Desierto de Sonora) (Shaible & Matzkin 2021).

En esta investigación también se generan árboles filogenéticos basados en el genoma mitocondrial completo como una herramienta para un estudio más amplio y profundo de las relaciones filogenéticas, por lo tanto, tendría una ventaja frente a los estudios que emplean una secuencia individual o pocos genes para el análisis filogenético (Crampton-Platt, et al. 2016).

OBJETIVOS

OBJETIVO GENERAL

- Comparar la calidad del ensamblaje del genoma mitocondrial de cinco especies andinas de *Drosophila* en Ecuador. Se utilizarán como referencia los genomas de *Drosophila melanogaster* y *Drosophila mojavensis*.

OBJETIVOS ESPECÍFICOS

1. Conseguir los genomas de referencia de *Drosophila melanogaster* y *Drosophila mojavensis* del repositorio de GenBank NCBI.
2. Obtener el ensamblado de los genomas mitocondriales de especies andinas de *Drosophila* del Ecuador mediante herramientas bioinformáticas.
3. Analizar la calidad de los genomas ensamblados mediante herramientas bioinformáticas.

MATERIALES Y MÉTODOS

Obtención de los genomas de referencia

A través del repositorio en línea NCBI, se obtendrán los genomas mitocondriales de referencia pertenecientes a *Drosophila melanogaster* y *Drosophila mojavensis* que sean actuales y verificados por la comunidad científica.

Ensamblado mitocondrial y análisis

Se utilizará la información del secuenciamiento del genoma mitocondrial de las especies del género *Drosophila*: *D. cashapamba*, *D. machachensis*, *D. mesophragmatica*, *D. guallabambae* y *D. papallacta* proporcionadas por la Dra. Doris Vela.

Antes de comenzar el ensamblado, se realizará un control de calidad de las secuencias mediante la herramienta FastQC y se procederá con la limpieza pertinente utilizando la terminal de comandos de Linux, mediante el uso del programa FASTX-toolkit.

Con esta información, a través de NOVOPlasty, se ensamblarán por separado los genomas utilizando las referencias antes mencionadas. Utilizando la configuración por defecto del programa.

Una vez ensamblado, se utilizará el visualizador de secuencias Gepard, para comparar ambos grupos de ensamblajes.

La comparación se realizará por grupos. Primero, entre las referencias, para definir las diferencias iniciales. Segundo, entre cada genoma ensamblado de las especies de estudio con su referencia. Y tercero, entre los genomas ensamblado de ambas referencias.

Posteriormente, se utilizará Jalview para conocer la naturaleza de las diferencias entre los ensamblajes.

Resultados

Calidad de las lecturas

No se pudo realizar el análisis de calidad de las lecturas de *D. guayllabambae* porque, en repetidas ocasiones, ocurría un error en la extracción del archivo comprimido.

Las únicas lecturas que pasaron un proceso de filtración fueron las de *D. cashapamba*, de las cuales se eliminaron las 3 primeras bases para eliminar el ruido de los adaptadores.

Al resto de secuencias no se les realizó modificaciones puesto que tenían buena calidad.

Figura de puntos de las dos referencias

D. melanogaster vs D. mojavensis

Al comparar los genomas mitocondriales de las dos especies que serán utilizadas como referencias se puede observar una diferencia en el tamaño del genoma, que corresponden a secuencias repetitivas presentes en el genoma de *D. melanogaster* y que están ausentes en *D. mojavensis* no posee. *D. melanogaster* tiene un genoma mitocondrial de 19523 pb, mientras que la otra especie *D. mojavensis* tiene 14909 pb (Fig.1).

Figuras de puntos de *D. melanogaster* como referencia

D. melanogaster vs D. cashapamba

D. cashapamba no muestra la zona de secuencias repetitivas de *D. melanogaster* y en el ensamblaje se obtuvo un genoma mitocondrial de 15524 pb. Además que se puede observar que *D. cashapamba* tiene pares de bases antes de que comiencen las similitudes entre las dos secuencias (Fig.2).

D. melanogaster vs D. guayllabambae

D. guayllabambae mostró un resultado similar a *D. cashapamba*, región antes de que comiencen a tener similitud con *D. melanogaster*. Además, también posee una cantidad de pares de bases similar, teniendo 15567 pb, una diferencia de 43 pares de bases (Fig.3).

D. melanogaster* vs *D. machachensis

D. machachensis obtuvo un ensamblado mitocondrial muy menor al promedio de las otras especies vistas, de 5541 pb. Aunque el segmento inicial antes de la similitud con *D. melanogaster* es constante (Fig.4).

D. melanogaster* vs *D. mesophragmatica

D. mesophragmatica obtuvo un ensamblaje mitocondrial de 15150 y una región inicial más pequeña que el resto de las secuencias. Esta especie tampoco tiene el segmento repetitivo al final como *D. melanogaster* (Fig.5).

D. melanogaster* vs *D. papallacta

D. papallacta obtuvo un ensamblaje de 14893 pb, y a diferencia de las demás secuencias, esta no posee al inicio un segmento diferente al de *D. melanogaster*, en cambio, la situación es inversa. El genoma mitocondrial ensamblado comienza a coincidir con el de *D. melanogaster* después en ese genoma (Fig.6).

Figuras de puntos de *D. mojavensis* como referencia

D. mojavensis* vs *D. cashapamba

Al comparar, ahora, *D. cashapamba* con *D. mojavensis*, de 14983 se tiene un resultado similar al de *D. melanogaster*. Un ensamblado menor, de 15127 pb, y un segmento inicial no similar más pequeño. Además, este ensamblaje resultó más grande que el genoma de la referencia de *D. mojavensis* (Fig.7).

D. mojavensis* vs *D. guayllabambae

D. guayllabambae obtuvo resultados similares a *D. cashapamba*, un ensamblaje mitocondrial de 15161 pb (Fig.8).

D. mojavensis* vs *D. machachensis

Nuevamente *D. machachensis* obtuvo un ensamblaje sumamente pequeño, de apenas 1008 pb. Además, la pequeña secuencia no coincide mucho con la de *D. mojavensis* (Fig. 9).

D. mojavensis* vs *D. mesophragmatica

D. mesophragmatica tiene resultados similares con *D. cashapamba* y *D. guayllabambae* al utilizar de referencia a *D. mojavensis* en el ensamblaje. Esta secuencia tiene 15164 pb (Fig. 10).

D. mojavensis* vs *D. papallacta

D. Papallacta tiene resultados anormales, esta vez teniendo 727 pb (Fig. 11).

Figuras de puntos de los dos ensamblajes de especies

***D. cashapamba* (*D. mojavensis*) vs *D. cashapamba* (*D. melanogaster*)**

Las dos secuencias mitocondriales ensambladas de *D. cashapamba* son extremadamente similares, lo cual que se puede ver en la línea sólida en la figura. La única diferencia es en el tamaño del ensamblado, siendo el que utilizó de referencia a *D. mojavensis* el más pequeño de 15127 pb, mientras que el mayor pertenece al ensamblado de *D. melanogaster* con 15524 pb (Fig.12).

***D. gauayllabambae* (*D. mojavensis*) vs *D. guayllabambae* (*D. melanogaster*)**

Con *D. guayllabambae* sucede lo mismo que con en el ensamblaje de *D. cashapamba*. El ensamblaje con *D. mojavensis* tiene 15161 pb y con *D. melanogaster* tiene 15567 pb (Fig.13).

D. machachensis (D. mojavensis) vs D. machachensis (D. melanogaster)

En ambos ensamblajes de *D. machachensis* se obtuvieron valores anormales en el largo de las secuencias. En el ensamblaje con *D. melanogaster* se obtuvieron 5541 pb, mientras que con *D. mojavensis* se obtuvieron solamente 1008 pb (Fig.14).

D. mesophragmatica (D. mojavensis) vs D. mesophragmatica (D. melanogaster)

Los ensamblajes de *D. mesophragmatica* tuvieron longitudes similares, siendo la más grande la que fue realizada con *D. mojavensis* por solamente 14 pb (Fig.15).

D. papallacta (D. mojavensis) vs D. papallacta (D. melanogaster)

El ensamblaje de *D. papallacta* con *D. mojavensis* como referencia, arrojó una secuencia anormal, de solamente 727 pb. Mientras que con *D. melanogaster* como referencia, se obtuvo un ensamblaje con valores esperados (Fig.16).

Análisis de secuencias por medio de Jalview

Mediante Jalview, se observó que las diferencias entre los diferentes ensamblajes de *D. cashapamba*, *D. guayllabambae* y *D. mesophragmatica* pertenecen a regiones A+T.

DISCUSIÓN

D. melanogaster posee una sección en su genoma mitocondrial una región A+T. Esta región está principalmente compuesta por las bases de adenina y timina, que se encuentra desde la base 14917 hasta el final del genoma, la base 19523. No se conoce la función de esta región, ni en qué procesos biológicos se encuentren, pero se cree que una región de replicación (Flybase n.d). En general, este genoma tiene una extensión de 19523 pb (Código de acceso

NC_024511.2). Por otro lado, el otro genoma mitocondrial de referencia, *D. mojavensis*, no posee esta región, siendo su genoma de 14904 pb (Código de acceso BK006339.1).

Esta región repetitiva puede ser la razón de las diferencias entre los ensamblajes de las dos referencias. Estas diferencias son relativamente cortas, entre los ensamblajes mitocondriales de *D. cashapamba* solo hay una diferencia de 397 pb, 15524 pb para en ensamblaje de *D. melangaster* y 15127 pb para *D. mojavensis*. Esta región posee un alto contenido A+T que es similar a la región final del genoma mitocondrial de *D. melanogaster*. Excluyendo esta pequeña sección en el ensamblaje de *D. melanogaster* es idéntico al ensamblaje de *D. mojavensis*.

La comparación de los ensamblajes mitocondriales de *D. guayllabambae*, 15567 pb para *D. melanogaster* y 15161 pb para *D. mojavensis*. Esta diferencia de 406 pb también se puede atribuir a la región A+T de *D. melanogaster*, puesto que se trata de una región rica de A+T. Similarmente con *D. cashapamba*, excluyendo esta región, los ensamblajes mitocondriales son idénticos.

Los ensamblajes mitocondriales de *D. machachensis* arrojaron una cantidad de pares de base extrañas. 5541 pb para *D. melanogaster* y 1008 pb para *D. mojavensis*. Esto se le puede atribuir a una posible contaminación en las lecturas. Se puede observar un segundo pico de contenido %GC que puede indicar lo mencionado (Andrews, 2020).

En los ensamblajes mitocondriales de *D. mesophragmatica* se observó lo contrario a los de *D. cashapamba* y *D. guayllabambae* porque esta vez el ensamblado de *D. mojavensis* fue más grande que el de *D. melanogaster*, aunque por solo 14 pb; siendo 15164 y 15150 pb respectivamente. De la misma manera que *D. cashapamba* y *D. guayllabambae*, si se excluye esta región, los ensamblajes son idénticos.

En relación con los ensamblajes de *D. papallacta* se obtuvo un ensamblaje de extensión esperada por parte de *D. melanogaster* de 14893 pb, pero por parte de *D. mojavensis* se obtuvo un ensamblaje de solamente 727 pb. La calidad de estas secuencias fue buena, por lo que no se le puede atribuir ese error al número anormal de bases.

RECOMENDACIONES Y CONCLUSIONES

Se recomienda volver a realizar la secuenciación de *D. machachensis* ya que las lecturas se encontraban contaminadas. Además, se recomienda profundizar el análisis de ensamblaje de *D. papallacta* con *D. mojavensis* como referencia para poder explicar a futuro el resultado de 727 pb.

La comparación de los ensamblajes de las especies andinas del Ecuador de género *Drosophila*, utilizando dos referencias diferentes, *D. melanogaster* y *D. mojavensis* se concluye que la mayoría resultados son bastante similares, teniendo ligeras variaciones que se pueden otorgar a la región A+T que tiene *D. melanogaster* al final del genoma mitocondrial, segmento que *D. mojavensis* no posee. Por lo tanto, para el ensamblaje del genoma mitocondrial del género *Drosophila* se podría utilizar cualquiera de las opciones como referencia, habiendo la posibilidad de emplear una referencia en el caso de que una, ya sea utilizada y presente errores.

En general se puede observar que los ensamblajes de *D. cashapamba*, *D. guayllabamabe*, *D. mesophragmatica* obtuvieron un tamaño del genoma mitocondrial de alrededor de las 15 000 pb con ambas referencias, lo cual entra en el promedio de tamaño del genoma mitocondrial para los insectos.

REFERENCIAS

- Andrews, S. (2020). Per Sequence GC Content. Babraham.ac.uk. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/3%20Analysis%20Modules/5%20Per%20Sequence%20GC%20Content.html>
- Benowitz, K. M., Coleman, J. M., Allan, C. W., & Matzkin, L. M. (2020). Contributions of cis- and trans-Regulatory Evolution to Transcriptomic Divergence across Populations in the *Drosophila mojavensis* Larval Brain. *Genome Biology and Evolution*, 12(8), 1407–1418. <https://doi.org/10.1093/gbe/evaa145>
- Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8), 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>

- Brozynska, M., Furtado, A., & Henry, R. J. (2014). Direct chloroplast sequencing: comparison of sequencing platforms and analysis tools for whole chloroplast barcoding. *PLoS One*, 9(10), e110387. <https://doi.org/10.1371/journal.pone.0110387>
- Céspedes, D., y Rafael, V. (2012). Cuatro especies nuevas del grupo de especies *Drosophila mesophragmatica* (Diptera, Drosophilidae) de los Andes ecuatorianos. *Iheringia. Série Zoologia*, 102(1), 71–79. <https://doi.org/10.1590/S0073-47212012000100010>
- Crampton-Platt, A., Yu, D. W., Zhou, X., y Vogler, A. P. (2016). *Mitochondrial metagenomics: letting the genes out of the bottle*. *GigaScience*, 5(1), s13742-016. <https://doi.org/10.1186/s13742-016-0120-y>
- Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic acids research*, 45(4), e18-e18. doi: 10.1093/nar/gkw955
- Douglas, A. E. (2018). The *Drosophila* model for microbiome research. *Lab Animal*, 47(6), 157–164. <https://doi.org/10.1038/s41684-018-0065-0>
- Drosophila 12 Genomes Consortium. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, 450(7167), 203–218. <https://doi.org/10.1038/nature06341>
- Duda, O. (1927). Die Sudamerikanischen Drosophiliden unter Berücksichtigung auch der anderen Neotropischen sowie Nearktischen Arten. *Arch. Naturgesch.*, 91 A. 11-12: 1- 229.
- Evolution of genes and genomes on the *Drosophila* phylogeny. (2007). *Nature*, 450(7167), 203–218. <https://doi.org/10.1038/nature06341>
- Figero, M. (2017). Filogenia molecular de especies ecuatorianas del grupo *Drosophila mesophragmatica* (Diptera, Drosophilidae) [Pontificia Universidad Católica del Ecuador]. <http://repositorio.puce.edu.ec/handle/22000/13207>
- Figuero, M. L., & Rafael, V. (2013). Descripción de tres especies nuevas del género *Drosophila* (Diptera, Drosophilidae) en el Ecuador. *Iheringia. Série Zoologia*, 103(3), 246–254. <https://doi.org/10.1590/S0073-47212013000300006>
- FlyBase. (n.d.). *FlyBase gene report: Dmel\mt:Ori*. Flybase.org. Retrieved May 28, 2023, from <http://flybase.org/reports/FBgn0013687>
- Hedges, S. B., Marin, J., Suleski, M., Paymer, M., y Kumar, S. (2015). Tree of Life Reveals Clock-Like Speciation and Diversification. *Molecular Biology and Evolution*, 32(4), 835–845. <https://doi.org/10.1093/molbev/msv037>
- Kim, J. S., Kim, M. J., Choi, D.-S., & Kim, I. (2016). Complete mitochondrial genome of the spotted-wing drosophila, *Drosophila suzukii* (Diptera:

Drosophilidae). Mitochondrial DNA. Part B, Resources, 1(1), 222–223.
<https://doi.org/10.1080/23802359.2016.1155422>

Lachaise, D., Cariou, M. L., David, J. R., Lemeunier, F., Tsacas, L., y Ashburner, M. (1988). *Historical biogeography of the Drosophila melanogaster species subgroup*. In *Evolutionary biology* (pp. 159-225). Springer, Boston, MA. 10.1007/978-1-4613-0931-4_4

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic acids research*, 49(W1), W293-W296. <https://doi.org/10.1093/nar/gkab301>

Liu, L., Xi, Z., Wu, S., Davis, C. C., & Edwards, S. V. (2015). Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences*, 1360(1), 36–53. <https://doi.org/10.1111/nyas.12747>

Montooth, K. L., Abt, D. N., Hofmann, J. W., & Rand, D. M. (2009). Comparative genomics of *Drosophila* mtDNA: Novel features of conservation and change across functional domains and lineages. *Journal of Molecular Evolution*, 69(1), 94–114. <https://doi.org/10.1007/s00239-009-9255-0>

National center for biotechnology information. (n.d.). Nih.gov. Retrieved May 29, 2023, from <https://www.ncbi.nlm.nih.gov/>

Oliver, G. R., Hart, S. N., & Klee, E. W. (2015). Bioinformatics for clinical next generation sequencing. *Clinical chemistry*, 61(1), 124-135. <https://doi.org/10.1373/clinchem.2014.224360>

Rafael, V., Arcos, G., 1988. *Drosophila guayllabambae* n. sp., un nuevo miembro del grupo repleta, subgrupo hydei (Diptera, Drosophilidae). *Evolución Biológica* 2, 167–176.

Robe, L. J., Valente, V. L., Budnik, M., y Loreto, É. L. (2005). *Molecular phylogeny of the subgenus Drosophila (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach*. *Molecular Phylogenetics and Evolution*, 36(3), 623-640. <https://doi.org/10.1016/j.ympev.2005.05.005>

Shaible, T. M., y Matzkin, L. M. (2021). *Physiological and life history changes associated with seasonal adaptation in the cactophilic Drosophila mojavensis*. bioRxiv. <https://doi.org/10.1101/2021.11.09.447153>

Tolwinski, N. (2017). Introduction: *Drosophila*—A Model System for Developmental Biology. *Journal of Developmental Biology*, 5(3), 9. <https://doi.org/10.3390/jdb5030009>

Vela D. y Rafael V. (2001). Ocho nuevas especies del grupo tripunctata, genero *Drosophila* (Diptera, Drosophilidae), y el registro de *Drosophila paraguayensis* en el Bosque Protector Pasochoa, Pichincha – Ecuador. *Revista de la Pontificia Universidad Católica del Ecuador* 66:92 – 120.

Wan, K., & Celniker, S. (2014). *Drosophila melanogaster* mitochondrion, complete genome - Nucleotide - NCBI. Nih.gov.
https://www.ncbi.nlm.nih.gov/nucleotide/NC_024511.2

Wolff, J. N., Camus, M. F., Clancy, D. J., & Dowling, D. K. (2016). Complete mitochondrial genome sequences of thirteen globally sourced strains of fruit fly (*Drosophila melanogaster*) form a powerful model for mitochondrial research. *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis*, 27(6), 4672–4674.
<https://doi.org/10.3109/19401736.2015.1106496>

Yamaguchi, M. y Yoshida, H. (2018). *Drosophila as a Model Organism* (pp. 1–10). https://doi.org/10.1007/978-981-13-0529-0_1

ANEXOS

FIGURAS

Comparación entre referencias

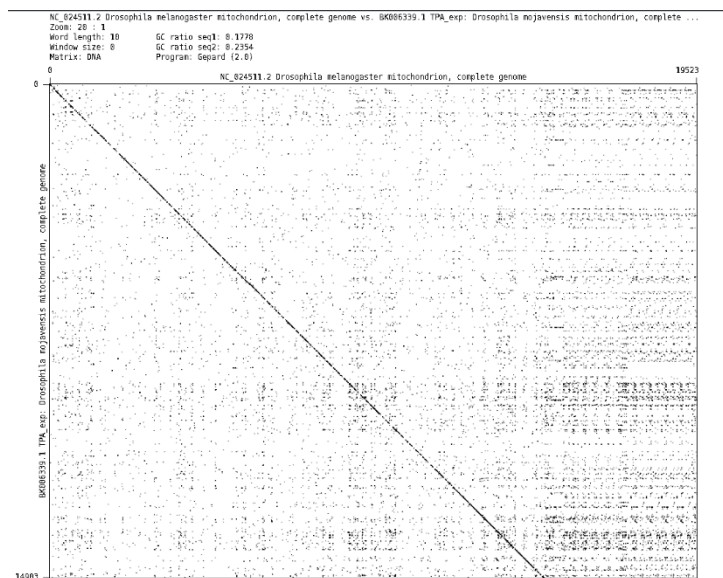


Fig.1.- Dot plot comparativo de los genomas mitocondriales de referencias *D. melanogaster* y *D. mojavensis*

Ensamblados mitocondriales con *D. melanogaster*

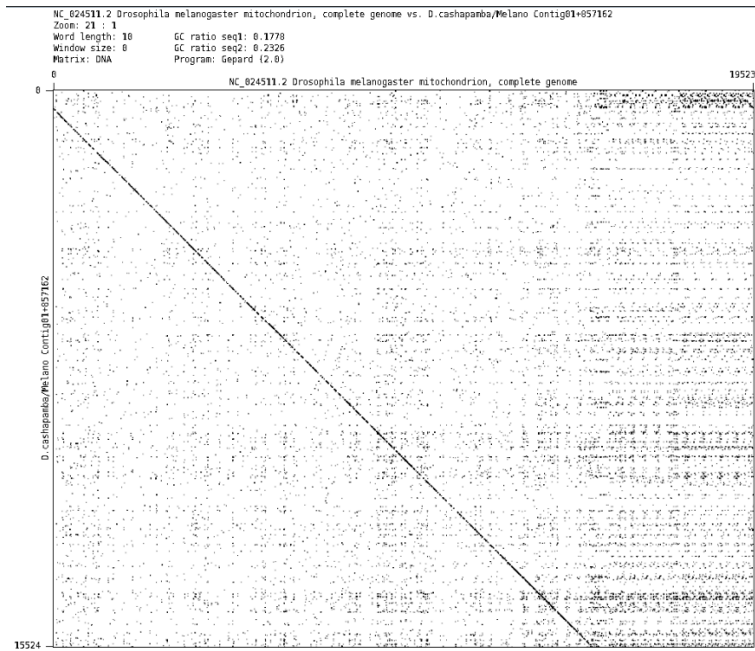


Fig.2.- Dot plot comparativo de los genomas mitocondriales de *D. melanogaster* y el ensamblaje de *D. cashapamba*

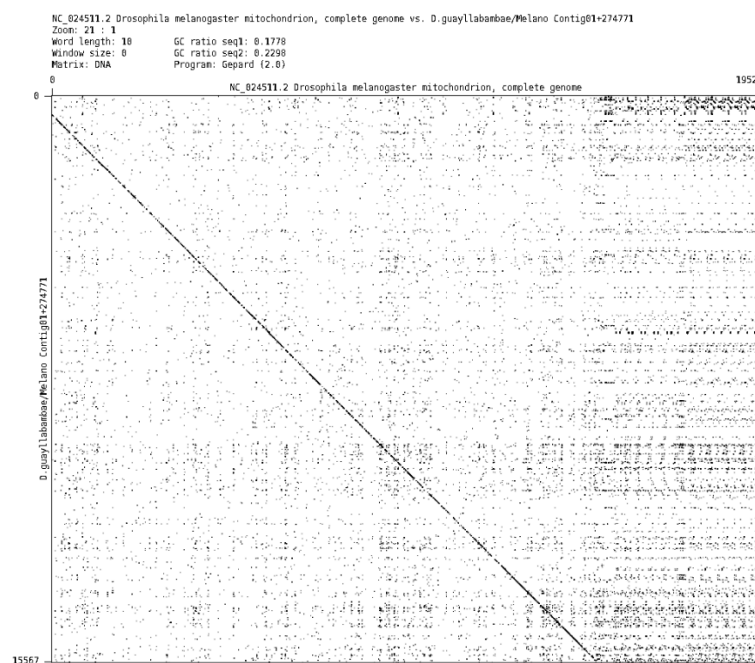


Fig.3.- Dot plot comparativo de los genomas mitocondriales de *D. melanogaster* y el ensamblaje de *D. guayllabambae*

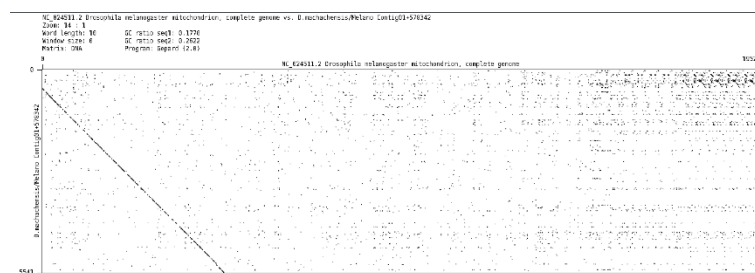


Fig.4.- Dot plot comparativo de los genomas mitocondriales de *D. melanogaster* y el ensamblaje de *D. machachensis*

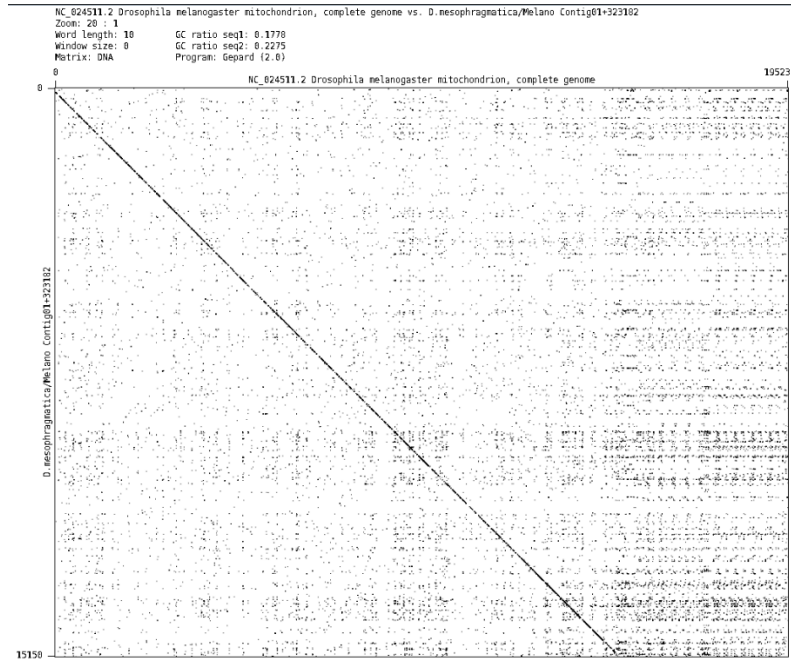


Fig.5.- Dot plot comparativo de los genomas mitocondriales de *D. melanogaster* y el ensamblaje de *D. mesophragmatica*

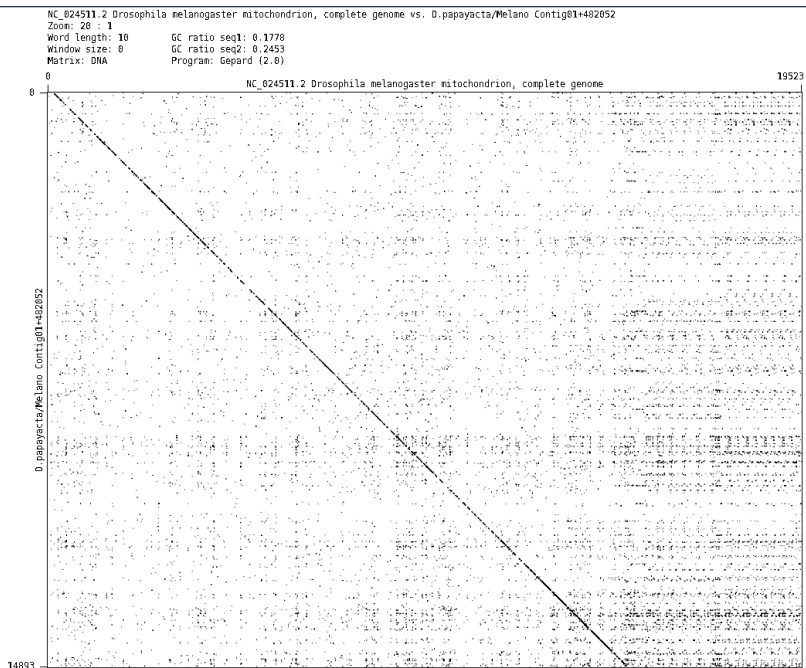


Fig.6.- Dot plot comparativo de los genomas mitocondriales de *D. melanogaster* y el ensamblaje de *D. papallacta*

Ensamblados mitocondriales con *D. mojavenis*

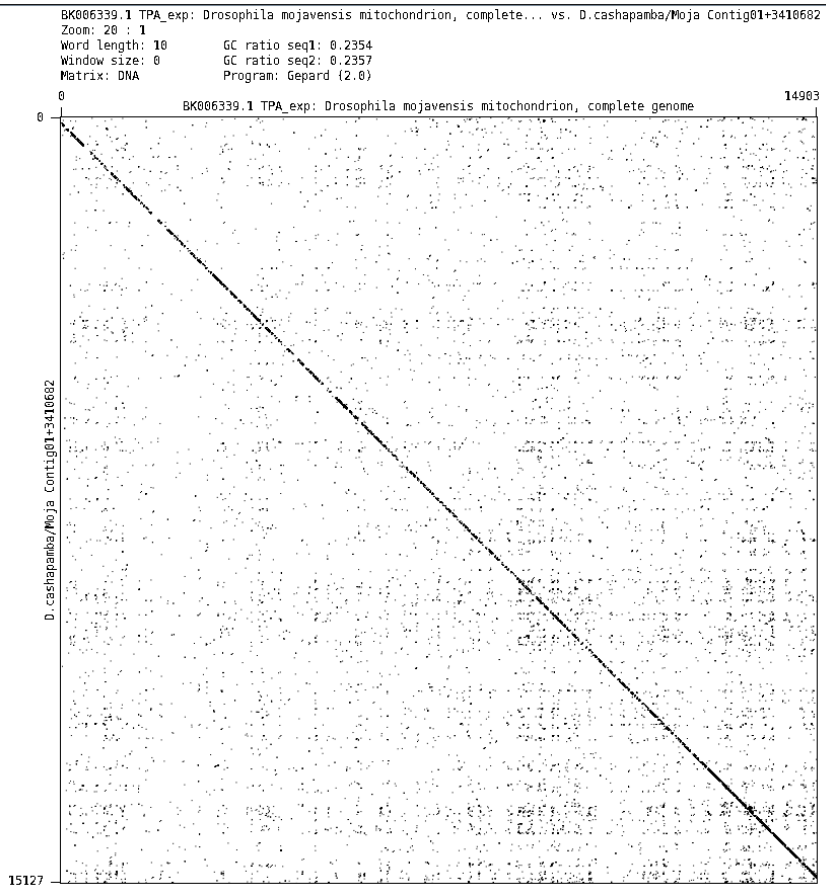


Fig.7.- Dot plot comparativo de los genomas mitocondriales de *D. mojavensis* y el ensamblaje de *D. cashapamba*

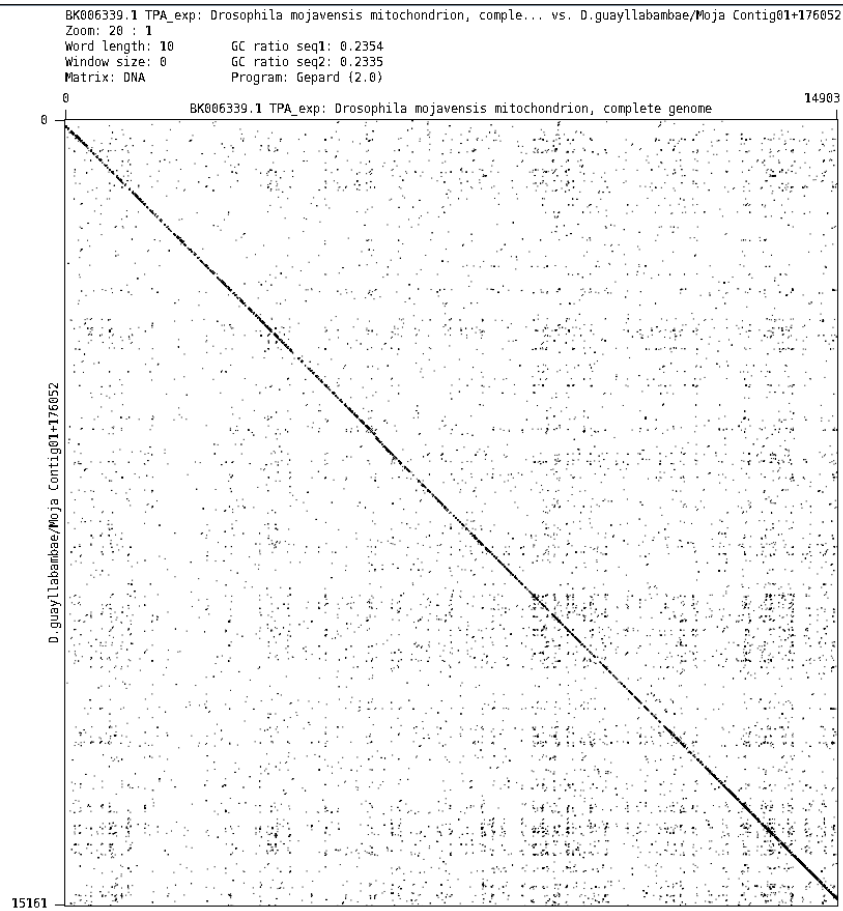


Fig.8.- Dot plot comparativo de los genomas mitocondriales de *D. mojavensis* y el ensamblaje de *D. guayllabambae*

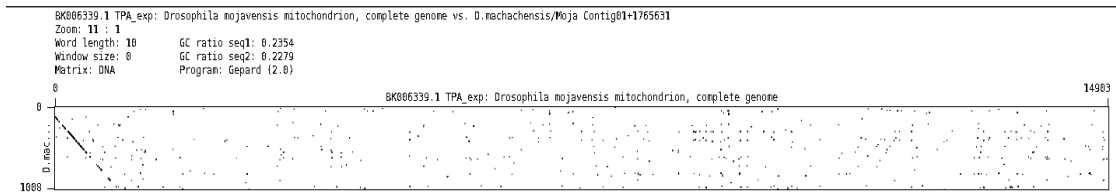


Fig.9.- Dot plot comparativo de los genomas mitocondriales de *D. mojavensis* y el ensamblaje de *D. machachensis*

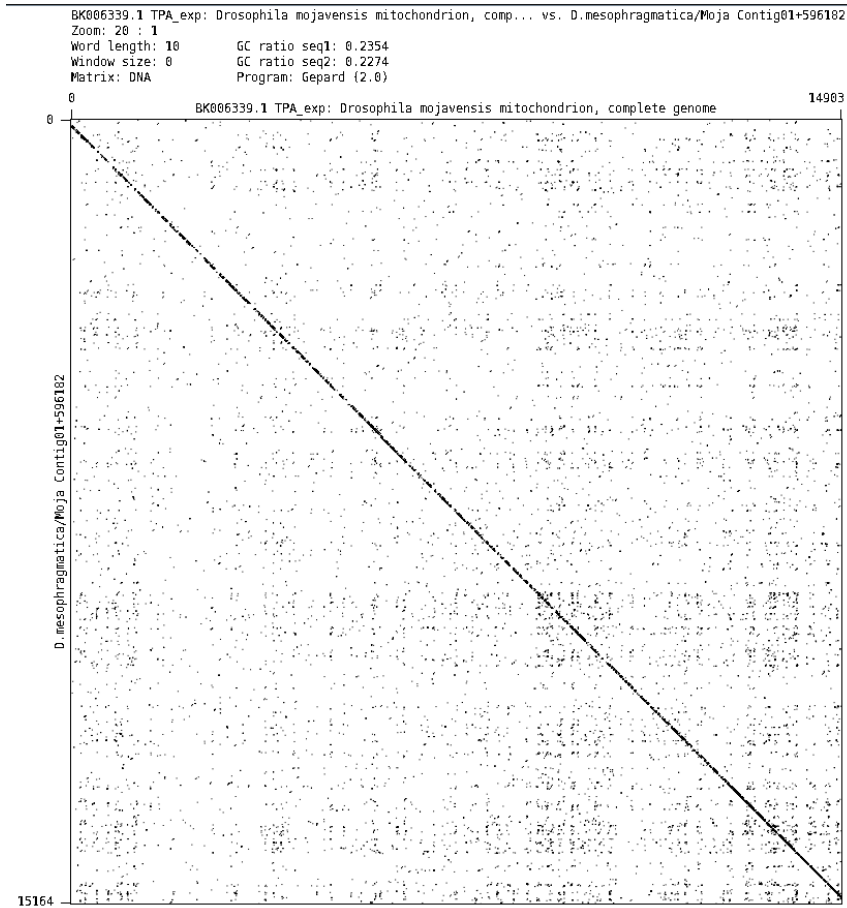


Fig.10.- Dot plot comparativo de los genomas mitocondriales de *D. mojavensis* y el ensamblaje de *D. mesophragmatica*

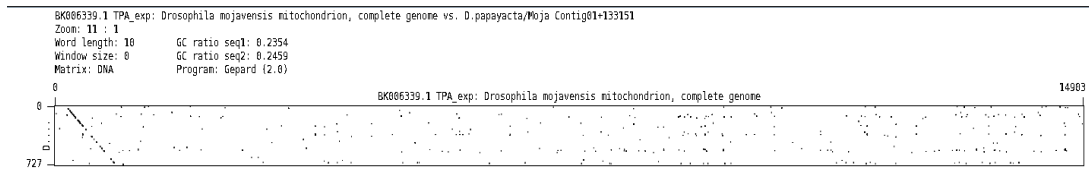


Fig.11.- Dot plot comparativo de los genomas mitocondriales de *D. mojavensis* y el ensamblaje de *D. papallacta*

Comparaciones entre ensamblajes mitocondriales

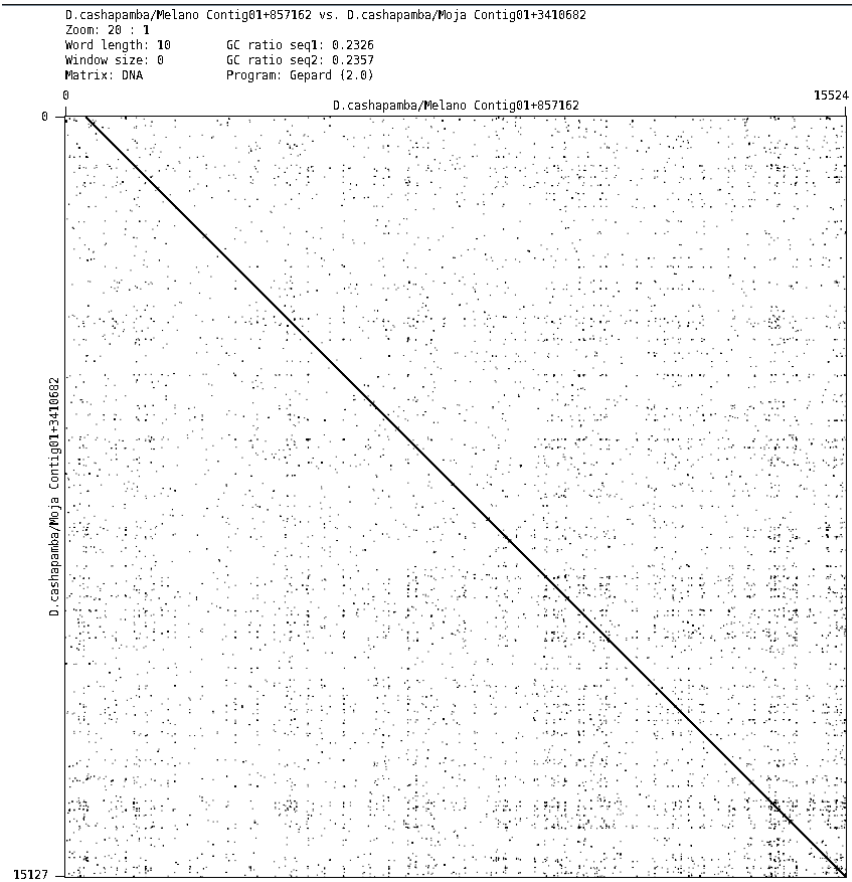


Fig.12.- Dot plot comparativo de los genomas mitocondriales de ensamblajes en *D. cashapamba*

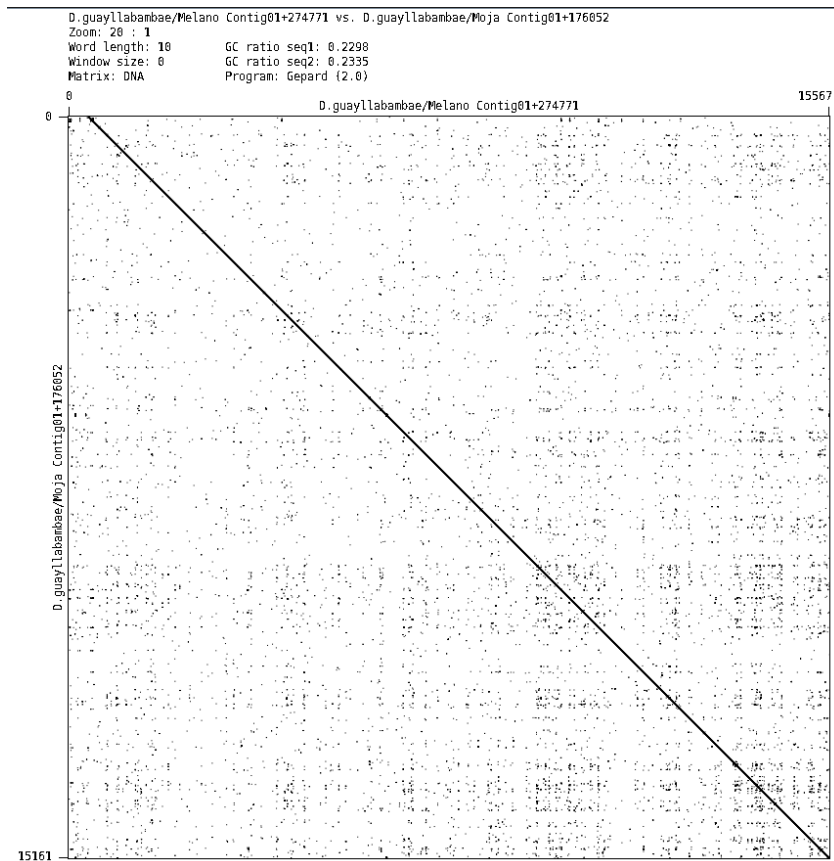


Fig.13.- Dot plot comparativo de los genomas mitocondriales de ensamblajes en *D. guayllabambae*

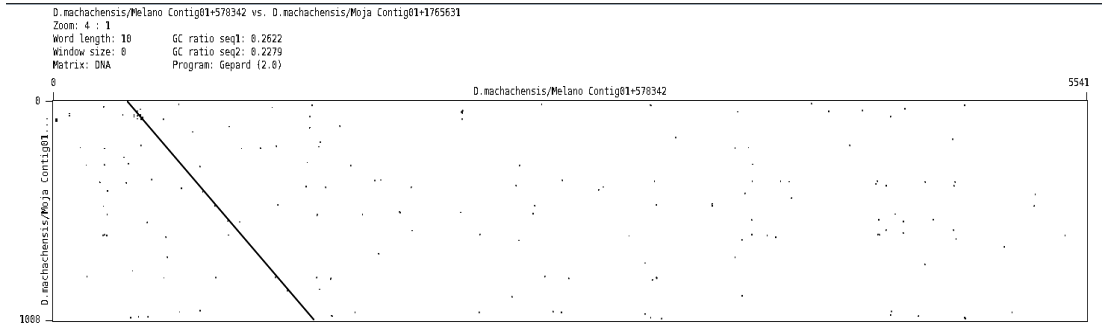


Fig.14.- Dot plot comparativo de los genomas mitocondriales de ensamblajes en *D. machachensis*

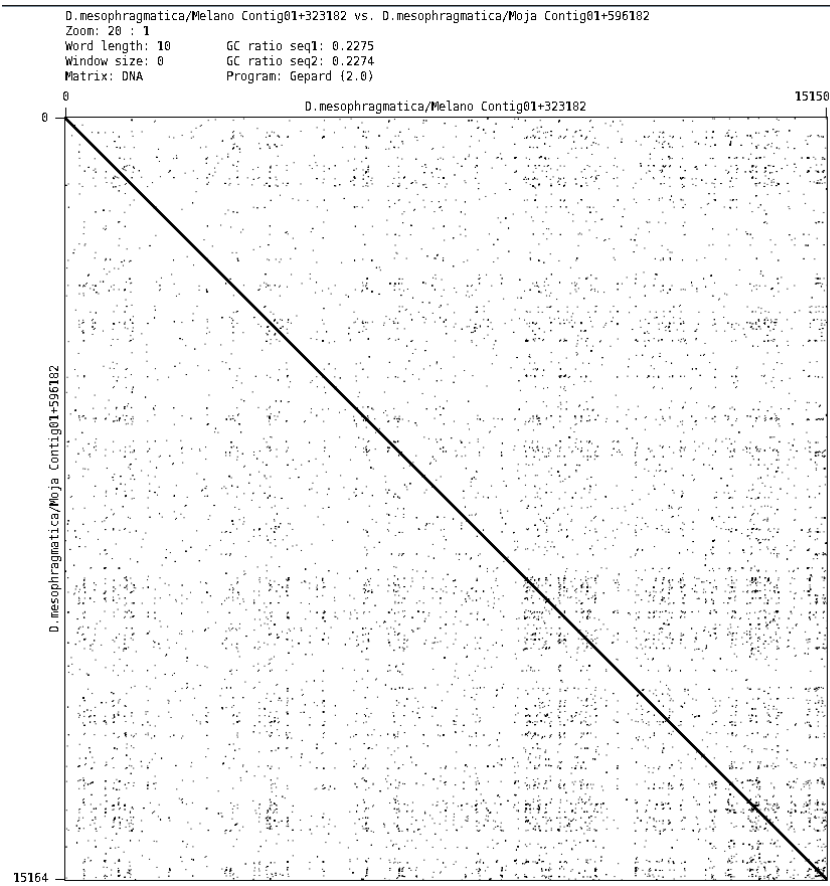


Fig.15.- Dot plot comparativo de los genomas mitocondriales de ensamblajes en *D. mesophragmatica*

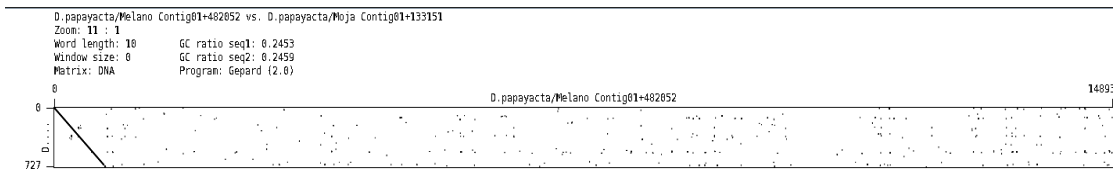


Fig.16.- Dot plot comparativo de los genomas mitocondriales de ensamblajes en *D. papallacta*

TABLAS

Tamaño de los genomas mitocondriales (pb)		
Ensamblajes	<i>D. melanogaster</i>	<i>D. mojavensis</i>
<i>D. cashapamba</i>	15524	15127
<i>D. guayllabambae</i>	15567	15161
<i>D. mesophragmatica</i>	15150	15164
<i>D. machachensis</i>	5541	1008
<i>D. papallacta</i>	14893	727

Referencias	
<i>D. melanogaster</i>	19523
<i>D. mojavensis</i>	14904

Tabla.1.- Tabla comparativa del tamaño del genoma mitocondrial de las referencias y de los ensamblados obtenidos.