

Análisis de la variabilidad nucleotídica en poblaciones naturales de *Drosophila melanogaster* en una región del cromosoma X influenciada por el gen *Megalin*.

Lourdes González Pizarro

Máster en Bioinformática y Bioestadística

Área 3 – Evolución molecular

Nombre Consultor/a:

Dorcas Orengo Ferriz

Nombre Profesor/a responsable de la asignatura:

Laura Calvet Liñan

06/2021



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	Análisis de la variabilidad nucleotídica en poblaciones naturales de <i>Drosophila melanogaster</i> en una región del cromosoma X influenciada por el gen <i>Megalin</i>
Nombre del autor:	Lourdes González Pizarro
Nombre del consultor/a:	Dorcas Orengo Ferriz
Nombre del PRA:	Laura Calvet Liñan
Fecha de entrega (mm/aaaa):	06/2021
Titulación:	Máster en Bioinformática y Bioestadística
Área del Trabajo Final:	Área 3 – Evolución molecular
Idioma del trabajo:	Castellano
Número de créditos:	15
Palabras clave	Variabilidad nucleotídica, selección natural, <i>Drosophila melanogaster</i>
Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i>	
<p><i>Drosophila melanogaster</i> es una especie cosmopolita que se encuentra en todos los continentes. Originaria del África Subsahariana, la expansión a territorios fuera de África es relativamente reciente en el tiempo, por lo que la variabilidad nucleotídica que se observa en su genoma todavía podría verse afectada por los efectos de su expansión y adaptación a nuevos hábitats.</p> <p>En un estudio de una población catalana se detectó una región de aproximadamente 67 kb que podría ser candidata de haber experimentado los efectos de la selección positiva. En este trabajo se ha estudiado la variabilidad de esta región en tres poblaciones de África, dos de Europa y una de EE. UU., utilizando diversos programas bioinformáticos, además de analizar la diferenciación genética entre ellas.</p> <p>La secuencia se ha analizado en ventanas de 2.000 nucleótidos con un desplazamiento de 1.000 nucleótidos. El patrón de variabilidad nucleotídica observado en las poblaciones derivadas es muy similar al observado anteriormente en la población catalana y está diferenciado del que se observa en las poblaciones africanas. Se detectan zonas en la secuencia en las que se rechaza la neutralidad tanto en las poblaciones africanas como en las derivadas, destacando una zona común a todas las poblaciones de unos 4 Kb. Además, se comprueba que las seis poblaciones se diferencian genéticamente entre ellas, aunque al hacer comparaciones dos a dos, algunos pares no muestran diferenciación entre ellos.</p>	

Abstract (in English, 250 words or less):

Drosophila melanogaster is a cosmopolitan species found on all continents. Originally from sub-Saharan Africa, the expansion to territories outside Africa is relatively recent in time, so the nucleotide variability observed in its genome could still be affected by the effects of its expansion and adaptation to new habitats.

In a study of a Catalan population, a region of approximately 67 kb was detected that could be a candidate for having experienced the effects of positive selection. In this work, we have studied the variability of this region in three populations from Africa, two from Europe and one from the USA, using different bioinformatics programs, as well as analyzing the genetic differentiation between them.

The sequence has been analyzed in 2,000 nucleotide windows with a 1,000 nucleotide displacement. The pattern of nucleotide variability observed in the derived populations is very similar to that previously observed in the Catalan population and is different from that observed in the African populations. We detected regions in the sequence in which neutrality is rejected both in the African and in the derived populations, highlighting a region of about 4 Kb common to all populations. In addition, it is found that the six populations differ genetically between them, although when making pairwise comparisons, some pairs do not show differentiation between them.

Índice

1	Resumen.....	1
2	Introducción	2
2.1	Contexto y justificación del Trabajo	2
2.2	Objetivos del Trabajo	3
2.3	Enfoque y método seguido	4
2.4	Planificación del Trabajo.....	5
2.5	Breve resumen de contribuciones y productos obtenidos	8
2.6	Breve descripción de los otros capítulos de la memoria	8
3	Estado del arte.....	10
4	Metodología	11
5	Resultados	14
5.1	Análisis de los resultados de la diferenciación genética entre poblaciones.....	14
5.2	Análisis de los indicadores de polimorfismo y divergencia genética.	17
6	Discusión	29
6.1	Análisis de los resultados de la diferenciación genética entre poblaciones.....	29
6.2	Análisis de los indicadores de polimorfismo y divergencia genética.	29
7	Conclusiones	31
7.1	Conclusiones	31
7.2	Líneas de futuro	32
7.3	Seguimiento de la planificación	32
8	Glosario.....	33
9	Bibliografía	35
10	Anexos.....	37

Lista de figuras

Figura 1: Diagrama de Gantt con la planificación del proyecto.....	7
Figura 2: Correlación entre la diferenciación genética y la distancia geográfica	15
Figura 3: Correlación entre diferenciación genética y la distancia geográfica excluyendo la población americana.	15
Figura 4: Correlación entre los estadísticos <i>Snn</i> y <i>Fst</i>	16
Figura 5: Porcentaje de variación explicado con los tres primeros ejes para el estadístico <i>Snn</i>	16
Figura 6: Porcentaje de variación explicado con los tres primeros ejes para el estadístico <i>Fst</i>	17
Figura 7: Distribución de la diversidad nucleotídica (π).....	18
Figura 8: Distribución de la divergencia nucleotídica (<i>K</i>).....	19
Figura 9: Distribución de la relación π /K para las poblaciones de Francia, Países Bajos y EE. UU.....	20
Figura 10: Análisis de la variación de la región del estudio realizado en la población catalana.	21
Figura 11: Localización de las posiciones 48.000 a 51.000 de la secuencia analizada.....	22
Figura 12: Localización de las posiciones 20.000 a 25.000 de la secuencia analizada.....	23
Figura 13: Localización de las posiciones 60.000 a 66.000 de la secuencia analizada.....	23
Figura 14: Localización de las posiciones 14.000 a 19.000 de la secuencia analizada.....	24
Figura 15: Localización de las posiciones 67.001 a 69.194 de la secuencia analizada.....	24
Figura 16: Distribución de los estadísticos de neutralidad <i>D</i> de Tajima y <i>H</i> de Fay y Wu en las poblaciones africanas	27
Figura 17: Distribución de los estadísticos de neutralidad <i>D</i> de Tajima y <i>H</i> de Fay y Wu en las poblaciones derivadas	28

Lista de tablas

Tabla 1: Planificación de las tareas	6
Tabla 2: Resultados estadísticos S_{nn} y F_{st}	14
Tabla 3: Valores medios de los indicadores de polimorfismo y diferenciación genética	17
Tabla 4: Fragmentos significativos para D en las tres poblaciones africanas.	22
Tabla 5: Fragmentos significativos para D en las tres poblaciones derivadas	24
Tabla 6: Número de fragmentos para los que se obtiene un valor de H significativo. .	25
Tabla 7: Fragmentos significativos para los estadísticos D y H	26

1 Resumen

Drosophila melanogaster es una especie cosmopolita que se encuentra en todos los continentes. Originaria del África Subsahariana, la expansión a territorios fuera de África es relativamente reciente en el tiempo, por lo que la variabilidad nucleotídica que se observa en su genoma todavía podría verse afectada por los efectos de su expansión y adaptación a nuevos hábitats.

En un estudio de una población catalana se detectó una región de aproximadamente 67 kb que podría ser candidata de haber experimentado los efectos de la selección positiva. En este trabajo se ha estudiado la variabilidad de esta región en tres poblaciones de África, dos de Europa y una de EE. UU., utilizando diversos programas bioinformáticos además de analizar la diferenciación genética entre ellas.

La secuencia se ha analizado en ventanas de 2.000 nucleótidos con un desplazamiento de 1.000 nucleótidos. El patrón de variabilidad nucleotídica observado en las poblaciones derivadas es muy similar al observado anteriormente en la población catalana y está diferenciado del que se observa en las poblaciones africanas. Se detectan zonas en la secuencia en las que se rechaza la neutralidad tanto en las poblaciones africanas como en las derivadas, destacando una zona común a todas las poblaciones de unos 4 Kb. Además, se comprueba que las seis poblaciones se diferencian genéticamente entre ellas, aunque al hacer comparaciones dos a dos, algunos pares no muestran diferenciación entre ellos.

2 Introducción

2.1 Contexto y justificación del Trabajo

Una de las bases de la genética de poblaciones es el estudio de la variación nucleotídica que se observa entre los individuos de una población. Esta variación, originada por las mutaciones en el ADN, es necesaria para que los procesos de evolución puedan tener lugar. La variabilidad que se observa puede estar influenciada por diversos factores como la deriva genética, los procesos de migración que dan lugar al efecto fundador, cuellos de botella que reducen drásticamente el número de individuos de la población y, por último, procesos de selección natural.

Según la teoría neutralista de la evolución¹ la mayor parte de la variación que se observa en una población tiene un carácter neutro, por tanto, no se encuentra sujeta a procesos de selección natural. Es decir, según esta teoría, la causa principal de los procesos evolutivos en la población es la deriva genética.

El estudio de la variación genética que observamos en los individuos de distintas poblaciones nos proporciona información de sus orígenes y los eventos de adaptación que han experimentado a lo largo del tiempo.

Una disminución en la variación nucleotídica dentro de una región genómica que se aleje de la teoría neutralista de Kimura puede ser el resultado de dos tipos de procesos que no hay que confundir. Por un lado, están los procesos de selección natural que tienen lugar en la región y, por otro lado, los procesos demográficos. Los procesos demográficos en los que la población ha experimentado una fuerte disminución del tamaño, ya sea por la existencia de cuellos de botella o por un efecto fundador al colonizar nuevos territorios, también reducen la variabilidad. Por este motivo, hay que poder distinguir si la disminución en la variabilidad se ha originado por la reciente expansión en nuevos territorios, o si ha sido originada por procesos relacionados con eventos de selección. Para poder diferenciar entre estos dos procesos se deben realizar estudios a lo largo del genoma, ya sea haciendo estudios multilocus o utilizando secuencias completas del genoma. Mientras que se espera que el efecto demográfico sea homogéneo a lo largo del genoma, el efecto de la selección natural estará más localizado en la región seleccionada (efecto *hitchhiking* o barrido selectivo). Esto último se debe a que, a medida que la región analizada se aleja de la región seleccionada, se pierde la huella que deja la selección que va siendo erosionada por el efecto de la recombinación.

Uno de los organismos más utilizados por parte de la comunidad científica en los estudios genéticos es *Drosophila melanogaster*. Originaria del África Subsahariana, posiblemente en África central y meridional², actualmente se ha convertido en una especie cosmopolita repartida por prácticamente todo el mundo^{3, 4}. Además, su expansión fuera del continente africano ha tenido lugar en épocas relativamente recientes. Todo esto, hace que sea un organismo muy utilizado, entre otros, para el estudio de los modelos de adaptación local.

En un estudio de una población catalana sobre la variación nucleotídica de 15 regiones del cromosoma X de *D. melanogaster* seleccionadas al azar y que evolucionan de manera independiente, se detectó una región de aproximadamente 67 kb que podría ser candidata de haber experimentado los efectos de la selección positiva, ya que la variabilidad detectada era menor que la esperada⁵. La región identificada se localizaba dentro del gen *Megalin* (*mgl*). Se ha descrito que este gen está involucrado en procesos biológicos relacionados con el desarrollo de la cutícula

y la regulación de sustancias por medio de endocitosis. Una de estas sustancias reguladas por endocitosis es la de la proteína del gen *yellow*, necesaria para la formación de melanina negra⁶. Además, dentro de esta región también se encuentra anotado el gen *CG44815*, para el que actualmente se desconoce su funcionalidad.

Como se ha comentado anteriormente, la expansión de *Drosophila melanogaster* a territorios fuera de África, es relativamente reciente en el tiempo. Por ello la variabilidad nucleotídica que se observa en las poblaciones derivadas todavía podría verse afectada por los efectos de cuellos de botellas relacionados con su expansión.

Para comprobar si el patrón que se observó en la región de influencia del gen *mgl* en una población catalana también se repite en otras áreas geográficas, se estudiará esta región genómica de interés en distintas poblaciones naturales de *D. melanogaster* de distintas áreas geográficas (África, Europa y América). De este modo, se comprobará si el efecto de la selección que se observó en la población catalana es algo local o, por el contrario, se trata de algo más general. Al seleccionar poblaciones de África, de donde es originaria la especie, y poblaciones no africanas se puede comprobar también si es un efecto exclusivo de las poblaciones derivadas, y, por tanto, relacionado con el proceso colonizador o, en cambio, también se observa en las africanas lo que podría indicar que la selección experimentada no respondió a las nuevas condiciones ambientales a que fue expuesta la especie durante la colonización.

2.2 Objetivos del Trabajo

Para este trabajo se plantean los objetivos generales y específicos que se detallan a continuación:

2.2.1. Objetivos generales

El estudio que se llevará a cabo engloba tres objetivos generales:

1. Estudiar la variabilidad nucleotídica en poblaciones naturales de *D. melanogaster* de distintas áreas geográficas (África, Europa y América) en una región candidata de 67 kb del cromosoma X.
2. Analizar la diferenciación genética entre poblaciones.
3. Determinar si el patrón de variación nucleotídica observado en una población catalana se repite en otras poblaciones naturales.

2.2.2. Objetivos específicos

Cada objetivo general se desglosa en los siguientes objetivos específicos:

1. Estudiar la variabilidad nucleotídica en poblaciones naturales de *D. melanogaster* de distintas áreas geográficas (África, Europa y América) en una región candidata de 67 kb del cromosoma X.
 - 1.1. Identificar las coordenadas de la secuencia de 67kb en el genoma de referencia
 - 1.2. Obtener la secuencia de la región genómica de interés para los individuos de poblaciones naturales en la base de datos pública PopFly⁷.

- 1.3. Seleccionar 15 individuos de cada población de estudio.
- 1.4. Alinear las secuencias de los individuos con la de *Drosophila simulans* para poder calcular los indicadores y estadísticos que necesitan de la secuencia de un outgroup.
- 1.5. Calcular los indicadores del polimorfismo S , K y π .
- 1.6. Calcular los estadísticos de neutralidad D de Tajima⁸ y H de Fay y Wu⁹.
- 1.7. Obtener la significación estadística de los estadísticos mediante simulaciones.
- 1.8. Analizar los resultados de la variabilidad en las distintas poblaciones.
2. Analizar la diferenciación genética entre poblaciones.
 - 2.1. Calcular los estadísticos de diferenciación genética Snn de Hudson¹⁰ y Fst ¹¹.
 - 2.2. Discutir las posibles causas del patrón de diferenciación genética observado entre las poblaciones.
 - 2.3. Comparar los estadísticos Snn y Fst entre ellos.
3. Determinar si el patrón de variación nucleotídica observado en una población catalana se repite en otras poblaciones naturales.
 - 3.1. Comparar los resultados con los obtenidos anteriormente para la población catalana.

2.3 Enfoque y método seguido

El estudio de la variabilidad y la diferenciación genética se ha realizado mediante el cálculo de indicadores y test estadísticos que permitan asignar un valor de significación entre los valores observados y los esperados según las teorías neutralistas.

Para poder llevar a cabo estos cálculos se seleccionaron muestras de 15 individuos de poblaciones de distintas áreas geográficas. Se han elegido tres poblaciones de diferentes regiones de África (Zambia, Ruanda y Etiopía) de dónde es originaria la especie y tres poblaciones no africanas, dos europeas (Francia y Países Bajos) y una americana (EE. UU.). De esta manera, se ha podido evaluar si el efecto observado en el estudio de la población catalana es algo que también se observa en otras poblaciones o es algo local. Además, al incluir poblaciones africanas, también se ha podido evaluar si es un efecto que solo se da en las poblaciones derivadas como posible efecto de la adaptación a climas menos cálidos o, también existe en las poblaciones africanas.

Dado que *D. melanogaster* es un organismo modelo ampliamente utilizado por la comunidad científica se dispone de gran información, de acceso público, tanto sobre su genoma de referencia como del genoma particular de numerosos individuos de distintas poblaciones naturales. Por este motivo, para realizar el estudio propuesto se obtuvo la información de los individuos seleccionados y de la secuencia a estudiar de bases de datos públicas como PopFly⁷ y FlyBase¹².

Además de las secuencias de *D. melanogaster*, era necesario tener una secuencia *outgroup*. Por esta razón, se obtuvo la secuencia de la región para *D. simulans*. Se eligió esta especie porque es una de las especies más cercanas a *D. melanogaster* de la que se dispone la secuencia completa de su genoma. La secuencia también se extrajo de la base de datos de acceso público Ensembl¹³. Para alinear esta secuencia *outgroup* con las secuencias de *D. melanogaster* se utilizó la versión web del programa MAFFT¹⁴.

Una vez seleccionados los individuos de las diferentes poblaciones, obtenidas sus secuencias y alineadas con la especie *outgroup*, se procedió al cálculo de distintos indicadores de polimorfismo y estadísticos de neutralidad como los D de Tajima y H de Fay y Wu. También se calcularon los estadísticos de diferenciación genética Snn y Fst y se realizaron simulaciones para poder asignar un valor de significación estadística entre los valores observados y los esperados según las teorías neutralistas. Para poder realizar todos estos cálculos se utilizaron distintos programas bioinformáticos.

Una vez finalizados los cálculos y las simulaciones se analizaron los resultados obtenidos y se compararon con los del estudio de la población catalana.

2.4 Planificación del Trabajo

Para poder realizar la planificación del trabajo se definen las tareas a realizar en base a los objetivos definidos y se planifica su inicio y fin. Además, se evalúan los posibles riesgos que se pueden presentar a lo largo del trabajo y las posibles acciones para mitigarlos

2.4.1. Tareas

En función de los objetivos, tanto generales como específicos, se identifican las siguientes tareas a realizar:

1. Identificación de las coordenadas de la secuencia de 67 Kb en el genoma de referencia.
2. Obtención de la secuencia correspondiente a la región genómica para las distintas poblaciones.
3. Selección de 15 individuos de cada una de las poblaciones de estudio.
4. Alineamiento de las secuencias.
5. Cálculo de los indicadores del polimorfismo y estadísticos de neutralidad.
6. Cálculo de los estadísticos de diferenciación genética Snn y Fst .
7. Realización de simulaciones para la obtención de la significación estadística.
8. Análisis de los resultados de la variabilidad y de la diferenciación genética.
9. Comparación de los resultados obtenidos con los del estudio de la población catalana.
10. Ajustes y corrección de errores.

11. Redacción de la memoria.

Aunque se propone una revisión constante de las tareas a realizar con el objetivo de la detección y corrección de posibles errores lo más tempranamente posible, también se propone una tarea previa a la redacción de la memoria para ajuste y corrección de potenciales errores no detectados.

2.4.2. Calendario

Para la realización del calendario, se ha tenido en cuenta una dedicación aproximada de 24 h semanales incluyendo fines de semana y festivos. En función de esto, las distintas tareas se planifican según se muestra en la siguiente tabla:

Nombre de la tarea	F. Inicio	F. Fin
PEC0 - Definición de los contenidos	17/02/2021	01/03/2021
PEC1 - Plan de Trabajo	02/03/2021	16/03/2021
PEC2 - Desarrollo del trabajo Fase 1	17/03/2021	19/04/2021
Identificación de las coordenadas de la secuencia de 67 Kb	17/03/2021	17/03/2021
Obtención de la secuencia correspondiente a la región genómica para las distintas poblaciones	17/03/2021	18/03/2021
Selección de 15 individuos de cada una de las poblaciones de estudio	19/03/2021	20/03/2021
Alineamiento de las secuencias	21/03/2021	23/03/2021
Cálculo de los indicadores del polimorfismo y estadísticos de neutralidad	24/03/2021	02/04/2021
Cálculo de los estadísticos S_{nn} y F_{st}	03/04/2021	12/04/2021
Realización de simulaciones para la obtención de la significación estadística	13/04/2021	19/04/2021
PEC3 - Desarrollo del trabajo Fase 2	20/04/2021	17/05/2021
Realización de simulaciones para la obtención de la significación estadística (cont.)	20/04/2021	26/04/2021
Análisis de los resultados de la variabilidad y de la diferenciación genética	27/04/2021	06/05/2021
Comparación de los resultados obtenidos con los del estudio de la población catalana	07/05/2021	12/05/2021
Ajustes y corrección de errores	13/05/2021	17/05/2021
PEC4 -Cierre de la memoria	18/05/2021	08/06/2021
Redacción de la memoria	18/05/2021	02/06/2021
Ajustes y corrección de errores	03/06/2021	08/06/2021
PEC5a - Elaboración de la presentación	09/06/2021	13/06/2021

Tabla 1: Planificación de las tareas

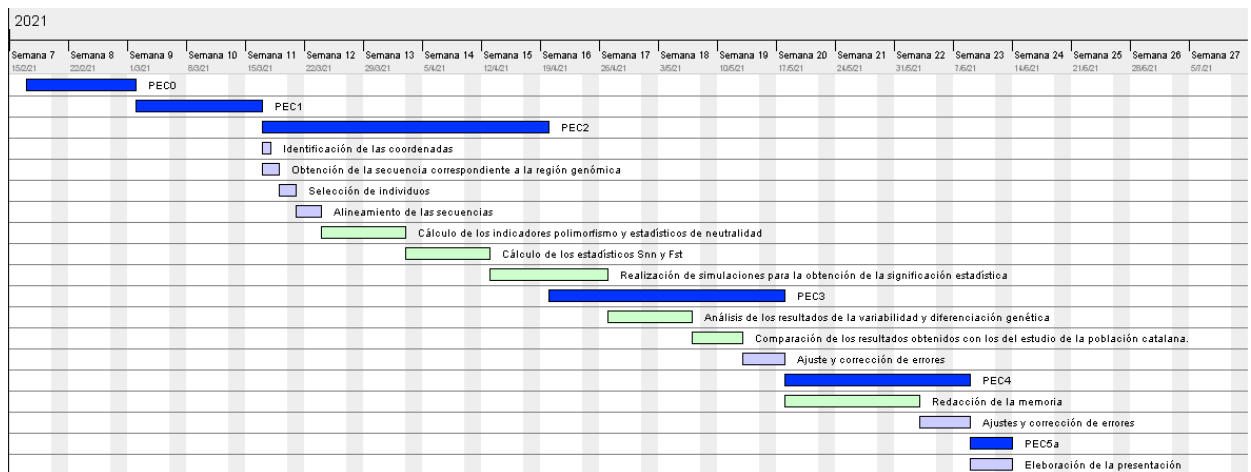


Figura 1: Diagrama de Gantt con la planificación del proyecto

2.4.3. Hitos

De acuerdo con la planificación realizada, las tareas que se establecen como hitos y en las que un posible retraso en el cumplimiento de las fechas propuestas podría poner en riesgo la consecución de los objetivos planteados, son las siguientes:

1. Cálculo de los indicadores del polimorfismo y estadísticos de neutralidad.
2. Cálculo de los estadísticos Snn y Fst .
3. Realización de simulaciones para la obtención de la significación estadística.
4. Análisis de los resultados de la variabilidad y de la diferenciación genética.
5. Comparación de los resultados obtenidos con los del estudio de la población catalana.
6. Redacción de la memoria.

Estas tareas, se han marcado en verde en el diagrama de Gantt.

2.4.4. Análisis de riesgos

Se identifican una serie de factores de riesgos que podrían tener lugar durante el desarrollo del trabajo y que podrían poner en riesgo el cumplimiento de los objetivos planteados. Estos factores son:

- Inexperiencia en el uso de los programas informáticos que se utilizarán. A pesar de que se ha realizado un trabajo previo para intentar familiarizarse con los programas informáticos que se utilizarán mediante los tutoriales existentes, la falta de experiencia en su uso podría suponer retrasos en las fechas. Este factor se ha tenido en cuenta a la hora de realizar la planificación de las tareas para intentar disminuir su impacto.
- Baja calidad de las secuencias de los individuos de las muestras, con secuencias que contengan un porcentaje elevado sin identificar y que dificulten el análisis. Para minimizar este impacto, para cada población se seleccionarán

los individuos, en los que, para la región de estudio tengan las secuencias más completas.

- Posibles problemas de rendimiento de los distintos programas bioinformáticos debido al tamaño de la secuencia a analizar (67Kb). Para intentar minimizar este riesgo se intentará optimizar el tiempo, ejecutando los procesos que puedan ser más pesados como, por ejemplo, las simulaciones en horario no productivo (dejarlos ejecutando por la noche o durante el horario laboral). Si, aun así, se detectaran problemas que pudieran suponer un riesgo, se podría plantear fragmentar la secuencia a analizar.
- Detección tardía de posibles problemas. Para evitar este punto, se proporcionará feedback de manera regular, de modo que se puedan identificar errores o problemas con la mayor antelación posible.
- Mala gestión del tiempo debido a imprevistos tanto de tipo profesional como personal o la aparición de problemas técnicos. En el caso de que surgiera algún imprevisto no contemplado a la hora de realizar la planificación de las tareas, ese tiempo se intentaría recuperar realizando algún sobreesfuerzo si fuera necesario.

2.5 Breve resumen de contribuciones y productos obtenidos

Una vez finalizado el proyecto, se habrán generado los siguientes entregables:

- Plan de Trabajo. Se generará un documento en el que se determinan los objetivos y el alcance del proyecto. Además de establecer el calendario y planificación de las distintas tareas junto con el análisis de riesgos.
- Memoria. Se generará un documento dónde se informará del trabajo realizado durante el proyecto en el que se detallará el contexto y la justificación del trabajo, los métodos utilizados para llevar a cabo el proyecto, los resultados obtenidos y las conclusiones derivadas del análisis de los resultados.
- Presentación virtual. Se realizará una presentación en la que se resumirá el desarrollo del trabajo y los resultados obtenidos recopilando los aspectos más significativos del proyecto.
- Informe de autoevaluación del proyecto.

2.6 Breve descripción de los otros capítulos de la memoria

La memoria, además de la introducción, se estructura en los siguientes capítulos:

- Estado del arte: Se explica la situación actual de los estudios de la variación nucleotídica que se observa en el genoma y la problemática de discernir si una disminución en la variabilidad se ha originado por procesos demográficos o por eventos de selección.
- Metodología: Se detalla la metodología seguida para llevar a cabo el análisis de la variabilidad nucleotídica y la diferenciación genética de la secuencia objeto del trabajo y los programas bioinformáticos utilizados para ello.
- Resultados: Se detallan los resultados obtenidos tras realizar el análisis de la secuencia en las distintas poblaciones seleccionadas.
- Discusión: Se discuten los resultados obtenidos respecto a la variabilidad

observada y la diferenciación genética entre las poblaciones estudiada.

- Conclusiones: Se incluyen las conclusiones que se extraen del trabajo realizado, se valora el seguimiento de la planificación establecida al inicio del semestre y se plantean posibles líneas de trabajo futuro.
- Glosario: Donde se definen los términos y acrónimos utilizados en la redacción de la memoria.
- Bibliografía: Dónde se listan las referencias bibliográficas utilizadas como fuentes de información para la redacción de la memoria.
- Anexos: Se incluye un apartado de anexos que contiene distintos gráficos y la tabla con la información de los valores de los indicadores de polimorfismo y estadísticos de neutralidad.

3 Estado del arte

La expansión de *Drosophila melanogaster* desde el África Subsahariana, de dónde se cree que es originaria la especie, a otros territorios es relativamente reciente. La expansión a territorios del África occidental se data entre ~ 70.000-72.000 años coincidiendo con cambios climáticos⁴, mientras que la expansión a territorios fuera de África se cree que tuvo lugar hace ~10.000-15.000 años.

Se espera que durante el transcurso de esta expansión hayan tenido lugar procesos demográficos pero que también hayan tenido lugar eventos de selección debidos a la adaptación a nuevos hábitats. Para poder diferenciar los procesos demográficos de los eventos de selección se deben realizar estudios a lo largo del genoma, ya sea haciendo estudios multilocus o utilizando secuencias completas del genoma. El efecto demográfico se espera que se presente de manera homogénea a lo largo del genoma mientras que los efectos de selección estarán más localizados en la zona seleccionada (efecto *hitchhiking* o barrido selectivo). Esto último se debe a que, a medida que la región analizada se aleja de la región seleccionada, se pierde la huella que deja la selección que va siendo erosionada por el efecto de la recombinación

En diversos estudios multilocus realizados en distintas poblaciones naturales de *Drosophila melanogaster*, dónde se han analizado diferentes regiones del cromosoma X, se ha observado que los efectos demográficos no son suficientes para explicar por completo los patrones de variabilidad observada.

Glinka y colaboradores¹⁵ llevaron a cabo un estudio multilocus en una población africana y otra europea de *D. melanogaster*. Analizaron 105 fragmentos del cromosoma X y obtuvieron resultados para diversos loci en la población europea compatibles con eventos de selección natural. Resultados similares se obtuvieron en un estudio de Orengo y Aguadé¹⁶ dónde se analizaban 109 fragmentos del cromosoma X de una población catalana. Kapun y colaboradores¹⁷, en un estudio utilizando pool-seq para secuenciar 48 muestras de poblaciones de 32 localizaciones también encuentra indicios de diferentes genes que han experimentado selección en poblaciones europeas. Estos son solo algunos ejemplos en los que se ha encontrado indicios de eventos de selección natural en poblaciones derivadas de *D. melanogaster*.

Se han encontrado pruebas de barrido selectivo en diversos loci del cromosoma X de *D. melanogaster*, entre otros, en un fragmento de 348 pb dentro del quinto intrón del gen *wap1*⁸, en una región que incluye el gen *phantom*¹⁹ y en una región de aproximadamente 67 kb dentro del gen *Megalín*. Esta última región es el objeto de estudio del presente trabajo.

Por otra parte, al disponer de muchos datos genómicos poblacionales de acceso público en bases de datos como PopFly, es posible evaluar si lo observado en una población concreta es algo aislado o el patrón se repite en otras pudiendo así comprobar si la variación observada es un caso especial de la población estudiada o si es algo más general.

4 Metodología

Para empezar, se localizaron las secuencias de la población catalana de *D. melanogaster* en la que se había detectado una región de aproximadamente 67 Kb que podría ser candidata de haber experimentado los efectos de la selección positiva accediendo a la base de datos públicos de GenBank ²⁰.

Una vez localizadas las secuencias de los fragmentos de la población catalana, se identificaron las coordenadas de la secuencia completa en el genoma de referencia utilizando la base de datos FlyBase¹² y seleccionando la opción que accede al programa BLAST de alineamiento de secuencias. La localización de las secuencias se realizó utilizando la *release* 5.57 de FlyBase. Se utilizó esta *release* porque es la versión del genoma de referencia de *D. melanogaster* utilizada para el alineamiento de las secuencias de los individuos de las distintas poblaciones que se encuentran en PopFly⁷.

Como resultado, se identificó que la región de estudio se encontraba comprendida entre las coordenadas 8.856.806 y 8.923.083 del cromosoma X de *D. melanogaster*.

Una vez obtenidas las coordenadas que comprendían la secuencia a analizar se descargaron, para esta región, las secuencias de los individuos de las poblaciones de *D. melanogaster* de Zambia, Ruanda, Etiopia, Francia, Países Bajos y EE. UU. a través del browser de PopFly.

También se obtuvo la secuencia de la especie *D. simulans* que se utilizó como *outgroup* para el cálculo de los estadísticos que lo necesitaban. Esta secuencia de referencia se consiguió accediendo al buscador de Ensembl¹³.

Para cada una de las poblaciones, se seleccionaron 15 individuos teniendo en cuenta los siguientes criterios:

- Que la secuencia se hubiera obtenido del mismo modo para todos los individuos seleccionados dentro de la misma población. Para las dos poblaciones europeas, la población de EE. UU. y la población de Etiopia se seleccionaron individuos en los que la secuencia se había obtenido mediante líneas endogámicas. Para las poblaciones de Zambia y Ruanda se seleccionaron individuos en los que la secuencia se había obtenido a partir de embriones haploides.
- Que los individuos no fuesen portadores de inversiones en el cromosoma X.
- Que la secuencia presentara el menor % de nucleótidos sin determinar.

La información relativa a los dos primeros puntos se obtuvo a partir de la documentación sobre los individuos de las poblaciones de la que dispone PopFly. Para la obtención de los individuos con menor porcentaje de secuencia sin identificar se creó un *script* para calcular el porcentaje de secuencia sin determinar para todos los individuos de una población y seleccionar los 15 mejor identificados.

El alineamiento de las secuencias se realizó accediendo a la web del programa MAFFT¹⁴, recomendado para alineamiento de secuencias grandes. Una vez obtenido el alineamiento automático de las secuencias, se realizó una revisión de la calidad

realizando diversos ajustes manuales en algunas regiones conflictivas para mejorar la calidad obtenida.

El cálculo de los estadísticos de polimorfismo y divergencia S , π y K y los estadísticos de neutralidad D de Tajima y H de Fay y Wu se han realizado utilizando el programa bioinformático DnaSP6²¹.

Las secuencias se analizaron en ventanas de 2.000 nucleótidos con un desplazamiento de 1.000 nucleótidos entre ellas. Además, como las posiciones sin identificar de las secuencias variaban bastante entre las distintas poblaciones de *D. melanogaster* se determinó que, al definir la longitud de las ventanas, se tuvieran en cuenta las posiciones que presentaban *gaps*. De esta manera se ha mantenido el tamaño de las ventanas y sus coordenadas en todas las poblaciones estudiadas, pero ha variado el número de nucleótidos analizados en cada ventana.

Para comprobar la diferenciación genética entre las poblaciones se han calculado los estadísticos Snn y Fst utilizando dos programas bioinformáticos diferentes. El estadístico Snn se calculó usando el programa DnaSP6 mientras que para el cálculo del estadístico Fst se utilizó el programa mstastpop²².

Se han realizado comparaciones dos a dos entre las distintas poblaciones y una comparación global entre las seis poblaciones de estudio. También se ha realizado la comparación considerando como poblaciones únicas los individuos de las poblaciones de África (Etiopia, Ruanda y Zambia), los individuos de las poblaciones de Europa (Francia y Países Bajos) y como otra población los individuos de América (EE. UU.).

Para poder comprobar la significación estadística de los estadísticos D de Tajima y H de Fay y Wu se realizaron simulaciones de cada una de las poblaciones de *D. melanogaster* de 1.000 iteraciones, considerando cada ventana definida en el cálculo de los distintos indicadores y estadísticos como loci independientes.

Las simulaciones se han realizado utilizando el programa mlcoalsim²³. Para poder generar los ficheros de entrada al programa, ha sido necesario calcular el número de nucleótidos por cada loci (cada ventana de la secuencia analizada) y la recombinación asociada a cada uno de ellos.

El número de nucleótidos de cada ventana se ha obtenido mediante la creación de un *script* que "eliminara" todas las posiciones que contenían *gaps* o algún nucleótido sin identificar en la ventana para alguna de las secuencias de los individuos de la población. Los resultados se han validado realizando algunas pruebas al azar con el programa DnaSP en distintos rangos de ventanas.

El cálculo del parámetro de recombinación de la población ($R= 4Nr$) para esta región se ha calculado considerando una tasa de recombinación de 4,32 cM/Mbp obtenida a partir de la estima de Hey y Kliman²⁴. Este dato es facilitado por la directora del trabajo.

Una vez obtenidas todas las simulaciones y las probabilidades, se identifican los fragmentos que son significativos con un nivel de significación α igual a 0,025 al estar evaluando una distribución de dos colas donde podemos tener valores observados muy pequeños (cola de la izquierda de la distribución) o muy elevados (cola de la derecha de la distribución).

Los resultados de los estadísticos Snn y Fst se analizan realizando tests de Mantel y análisis de PCoA con ayuda del complemento de Excel GenALEx²⁵.

5 Resultados

5.1 Análisis de los resultados de la diferenciación genética entre poblaciones

Los resultados obtenidos en la comparación de las seis poblaciones entre ellas, tanto para los estadísticos *Snn* como para *Fst*, indican que las seis poblaciones son estadísticamente diferentes con una $p < 0,05$.

Con los resultados de las comparaciones dos a dos de las poblaciones, en cambio, observamos diferencias entre los dos estadísticos. Mientras que con el estadístico *Snn* observamos diferenciación significativa entre la población de Etiopía y el resto, entre la población de Ruanda y las poblaciones derivadas y también entre la de Países Bajos y Estados Unidos; con el estadístico *Fst*, detectamos diferenciación genética significativa entre cualquier par de poblaciones excepto para las de Francia y Países Bajos y Francia y EE. UU.

Como al realizar las comparaciones por pares, estamos realizando comparaciones múltiples que aumentan la probabilidad de obtener falsos positivos, se aplica la corrección de Bonferroni secuencial²⁶ a los resultados obtenidos para las comparaciones dos a dos. Se ha decidido aplicar esta corrección ya que no es tan estricta como la de Bonferroni clásica. Aplicando la corrección, para el estadístico *Snn* aceptaríamos que entre las poblaciones de Francia y Ruanda por un lado y entre Países Bajos y EE. UU. por otro no se aprecian diferencias. Para el estadístico *Fst* también aceptaríamos que entre las poblaciones de Países Bajos y EE. UU. y entre Zambia y Ruanda no se observan diferencias.

Poblaciones	Snn	P-Value	Ajuste Bonferroni	Fst	P-Value	Ajuste Bonferroni
EF-RG-ZI-FR-NTH-RAL	0,511	0	N/A	0,239	0	N/A
EF-RG	0,717	0,004		0,074	0	*
EF-ZI	0,750	0,001	*	0,109	0	*
EF-FR	1	0	*	0,409	0	*
EF-NTH	1	0	*	0,411	0	*
EF-RAL	1	0	*	0,406	0	*
FR-NTH	0,627	0,092		0,019	0,118	
FR-RAL	0,567	0,242		0,011	0,183	
FR-RG	0,700	0,007		0,302	0	*
FR-ZI	0,633	0,067		0,269	0	*
NTH-RAL	0,767	0,009		0,034	0,031	
NTH-RG	0,817	0	*	0,306	0	*
NTH-ZI	0,606	0,121		0,271	0	*
RAL-RG	0,867	0	*	0,298	0	*
RAL-ZI	0,633	0,116		0,267	0	*
RG-ZI	0,633	0,082		0,014	0,026	
FR-NTH-RAL	0,600	0,001	N/A	0,027	0,012	N/A
EF-RG-ZI	0,556	0,002	N/A	0,067	0	N/A
(EF-RG-ZI) / (FR-NTH) / RAL	0,661	0	N/A	0,264	0	N/A

Tabla 2: Resultados estadísticos *Snn* y *Fst*

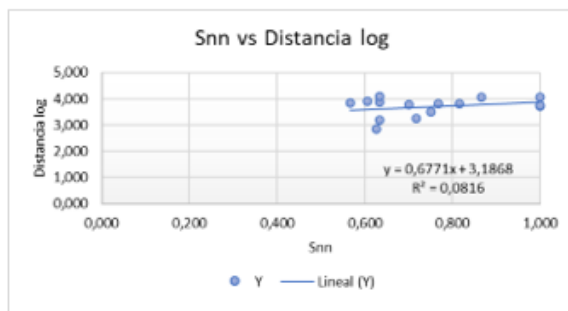
Se ha marcado con un * las comparaciones que mantienen la significación estadística tras el ajuste de Bonferroni.

Al comparar entre ellas, por un lado, las tres poblaciones africanas y por otro las tres no africanas, los resultados que se obtienen también son que los dos grupos de poblaciones son significativamente diferentes entre ellas para los dos estadísticos. Lo mismo sucede si comparamos las poblaciones “agrupándolas” por continente (tabla2).

Al evaluar las correlaciones entre el estadístico *Snn* y la distancia geográfica y entre el estadístico *Fst* y la distancia geográfica (tanto en Km como en escala logarítmica) con el test de Mantel, comprobamos que no existe correlación ($p > 0,05$) en ninguno de los dos casos, a pesar de que el coeficiente de correlación para el caso del estadístico *Fst* cuando se analiza la correlación con la distancia geográfica en escala logarítmica es elevado (0,622).

A.

SSx	SSy	SPxy	Rxy	P(rxy-rand >= rxy-data)
0,320	1,796	0,216	0,286	0,149



B.

SSx	SSy	SPxy	Rxy	P(rxy-rand >= rxy-data)
0,328	1,796	0,477	0,622	0,053

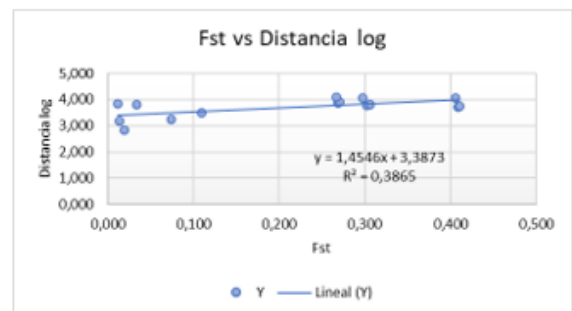


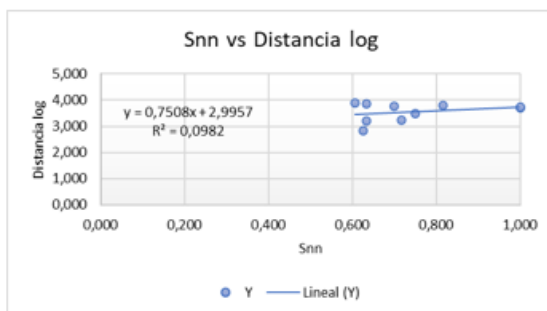
Figura 2: Correlación entre la diferenciación genética y la distancia geográfica

A: Correlación de *Snn* con la distancia geográfica en escala logarítmica, **B:** Correlación de *Fst* con la distancia geográfica en escala logarítmica

Si se excluye la población americana, ya que entre esta población y el resto existe la barrera del Atlántico que puede distorsionar los resultados, únicamente se obtiene una correlación significativa entre el estadístico *Fst* y la distancia geográfica en escala logarítmica ($p < 0,05$).

A.

SSx	SSy	SPxy	Rxy	P(rxy-rand >= rxy-data)
0,196	1,127	0,147	0,313	0,187



B.

SSx	SSy	SPxy	Rxy	P(rxy-rand >= rxy-data)
0,208	1,127	0,408	0,844	0,039

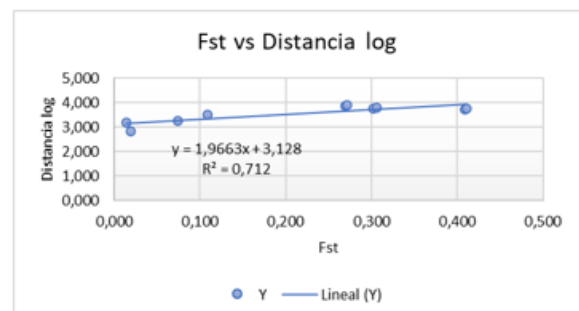


Figura 3: Correlación entre diferenciación genética y la distancia geográfica excluyendo la población americana.

A: Correlación de *Snn* y la distancia geográfica en escala logarítmica, **B:** Correlación de *Fst* y la distancia geográfica en escala logarítmica.

Los resultados del test de Mantel para los estadísticos *Snn* y *Fst* indican que los dos estadísticos presentan una alta correlación entre ellos ($p < 0,05$).

SSx	SSy	SPxy	Rxy	P(rxy-rand \geq rxy-data)
0,320	0,328	0,223	0,688	0,002

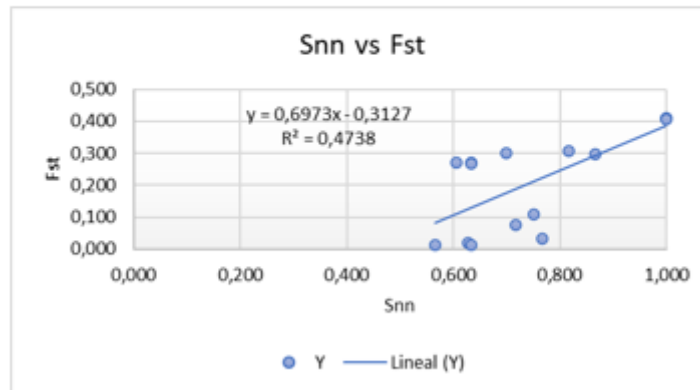


Figura 4: Correlación entre los estadísticos *Snn* y *Fst*

Para completar el análisis de los estadísticos, se realizó también un análisis de coordenadas principales (PCoA). Con este análisis, para *Snn* se observa que se diferencian claramente las poblaciones y con los dos primeros ejes ya se explica un 53,77 % de la variabilidad (figura 5). El PCoA para *Fst* explica el 95,09 % de la variabilidad observada con los dos primeros ejes y aunque, en la representación gráfica de la distribución de los ejes 1 y 2 las poblaciones derivadas aparentemente no se diferencian tan bien como para *Snn*, cuando se representa la distribución respecto al eje 3, sí que se observa una diferenciación clara entre todas las poblaciones (figura 6).

Axis	1	2	3
%	33,14	20,63	18,91
Cum %	33,14	53,77	72,68

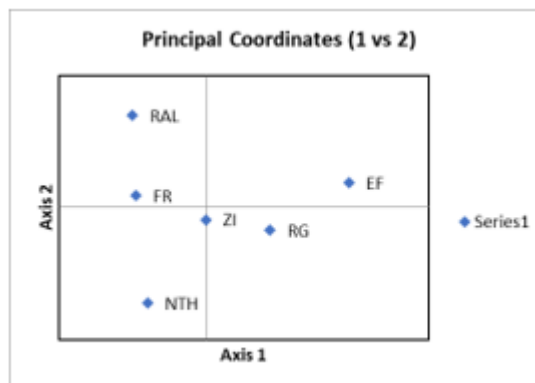


Figura 5: Porcentaje de variación explicado con los tres primeros ejes para el estadístico *Snn*.

Axis	1	2	3
%	84,96	10,13	3,26
Cum %	84,96	95,09	98,35



Figura 6: Porcentaje de variación explicado con los tres primeros ejes para el estadístico F_{st} .

5.2 Análisis de los indicadores de polimorfismo y divergencia genética.

Conforme a los resultados obtenidos para los indicadores de polimorfismo S y π podemos observar dos patrones diferentes. Por un lado, un patrón similar para las poblaciones africanas dónde el valor medio de π varía entre 0,00578 y 0,00829 y por otro, un patrón diferente para las poblaciones derivadas dónde los valores de los indicadores S y π son menores (con valores medios de π entre 0,00177 y 0,00211). En cambio, para el indicador de divergencia genética K , las seis poblaciones tienen valores similares y también siguen un patrón de variación muy parecido (figura 7 y anexos 1 y 2).

Población	Valor medio S	Valor medio π	Valor medio K
Zambia (ZI)	57,89855	0,00829	0,04832
Ruanda (RG)	54,00000	0,00764	0,04851
Etiopia (EF)	34,92754	0,00578	0,04785
Francia (FR)	9,43478	0,00177	0,04986
Países Bajos (NTH)	9,37681	0,00194	0,05069
EE. UU. (RAL)	11,44928	0,00211	0,05163

Tabla 3: Valores medios de los indicadores de polimorfismo y diferenciación genética

En las poblaciones derivadas, tanto para S como para π , se observa una zona comprendida aproximadamente entre las posiciones 25.000 y 50.000 de la secuencia analizada dónde los valores observados se sitúan por debajo de la media, aunque en la población francesa hay un pico aproximadamente sobre la posición 34.000-35.000 en el que estos valores son superiores a la media.

Además de éste, también hay otro punto localizado aproximadamente entre las posiciones 12.000 y 13.000 dónde los valores descienden, siendo más marcado este descenso para la población de Países Bajos.

Los datos que se observan para estos dos indicadores en las poblaciones africanas también son bastantes similares entre ellos. Hay dos zonas en las que se observa un descenso pronunciado de los valores que se encuentran por debajo de la media. Estas zonas se encuentran aproximadamente en las posiciones 27.000-28.000 y 37.000-38.000 de la secuencia. Además, destaca que, en la población de Etiopia, los valores de los indicadores S y π son menores en comparación con las otras dos poblaciones africanas, aunque no son tan bajos como los de las poblaciones derivadas.

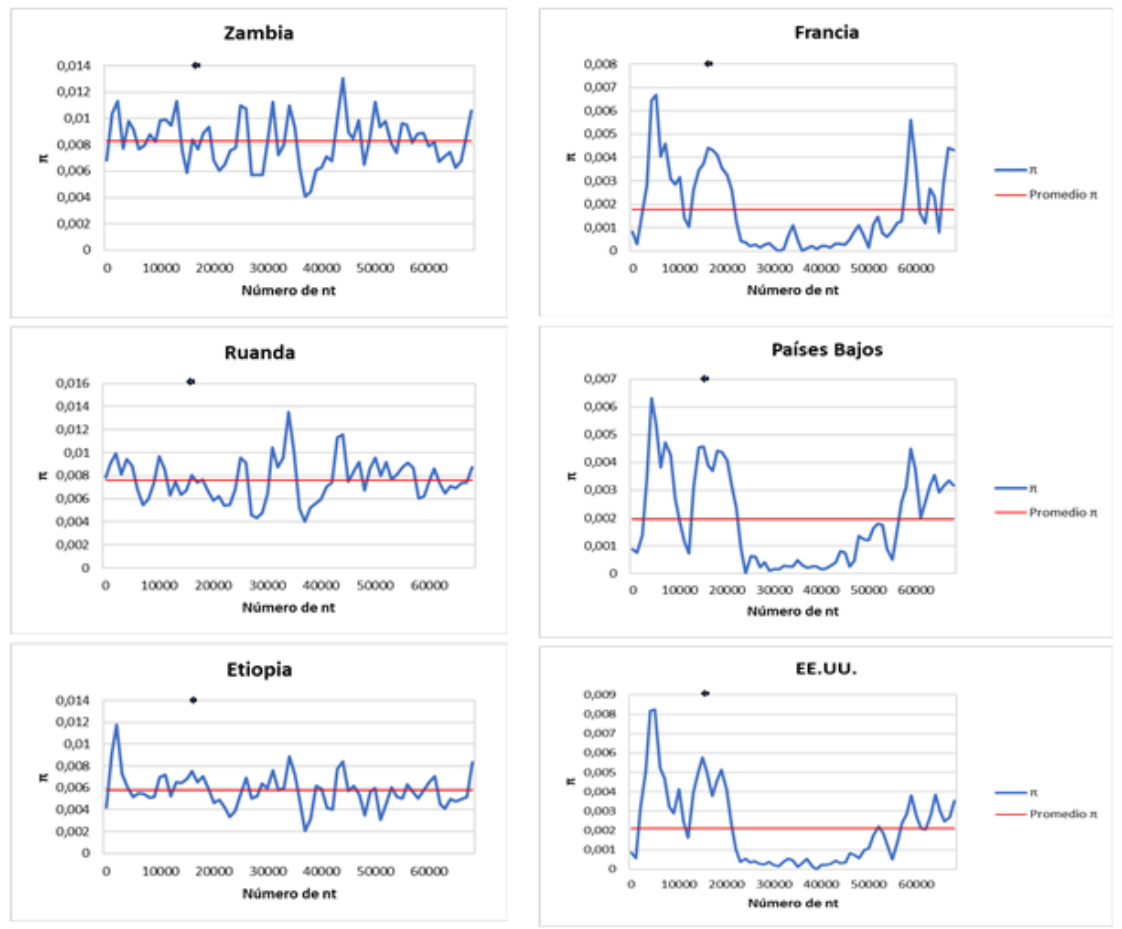


Figura 7: Distribución de la diversidad nucleotídica (π)
 La flecha de la parte superior indica la posición aproximada el gen CG44815 que se encuentra anotado dentro de la región analizada.

Respecto a los resultados obtenidos para el indicador de divergencia nucleotídica K , destaca una zona muy por encima de la media en todas las poblaciones localizada aproximadamente en las posiciones 34.000-36.000 de la secuencia (figura 8).

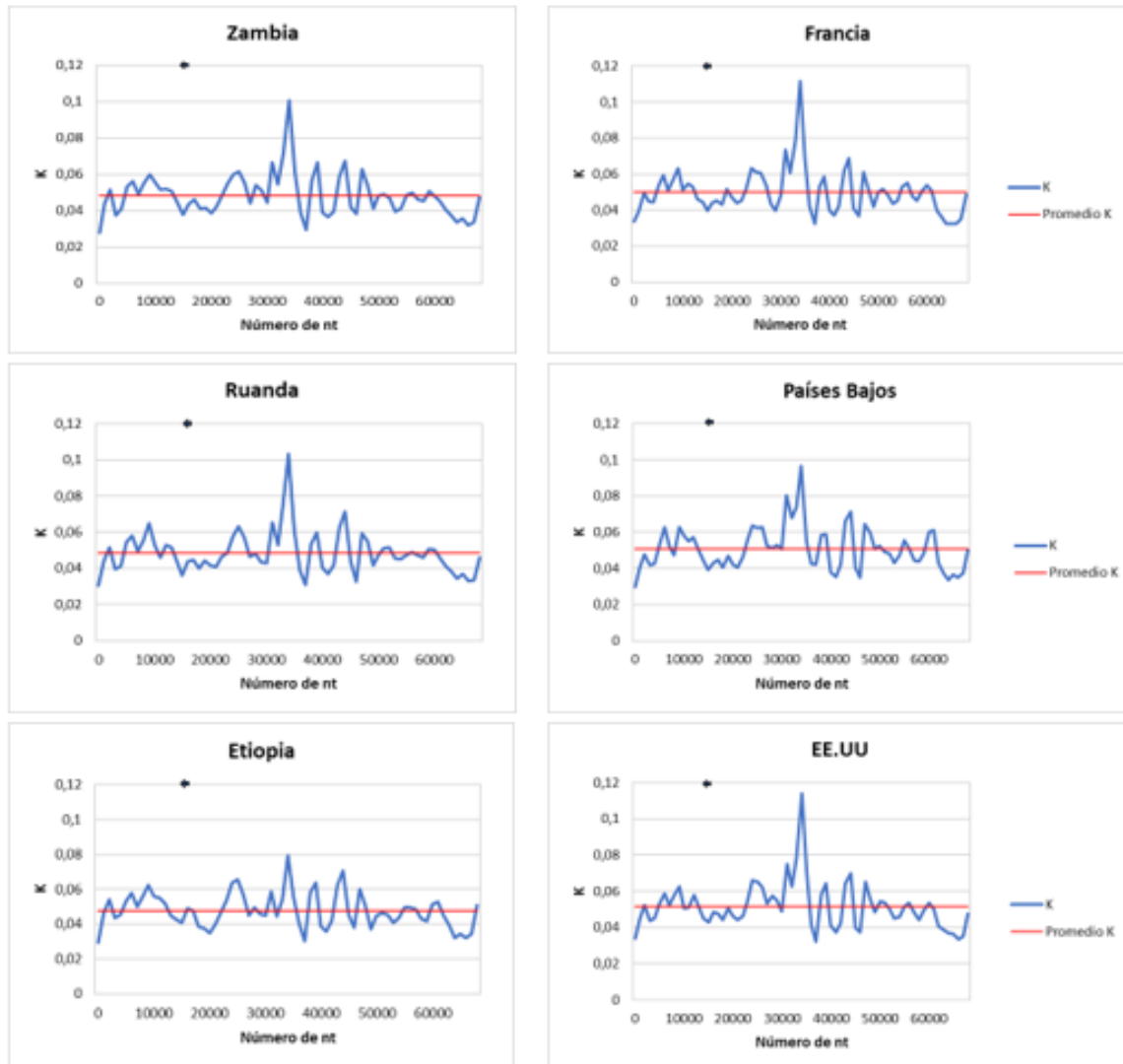


Figura 8: Distribución de la divergencia nucleotídica (K)
 La flecha de la parte superior indica la posición aproximada el gen CG44815 que se encuentra anotado dentro de la región analizada.

Los resultados obtenidos para los indicadores S , π y K en las poblaciones derivadas, son similares a los resultados del estudio que se llevó a cabo en la población catalana⁵, donde también se observa un descenso pronunciado del polimorfismo en la región localizada aproximadamente en las posiciones 22.000 – 50.000 de la secuencia. Esto nos lleva a pensar que el efecto observado en la población catalana es común a las poblaciones derivadas y por lo tanto relacionado con el efecto colonizador.

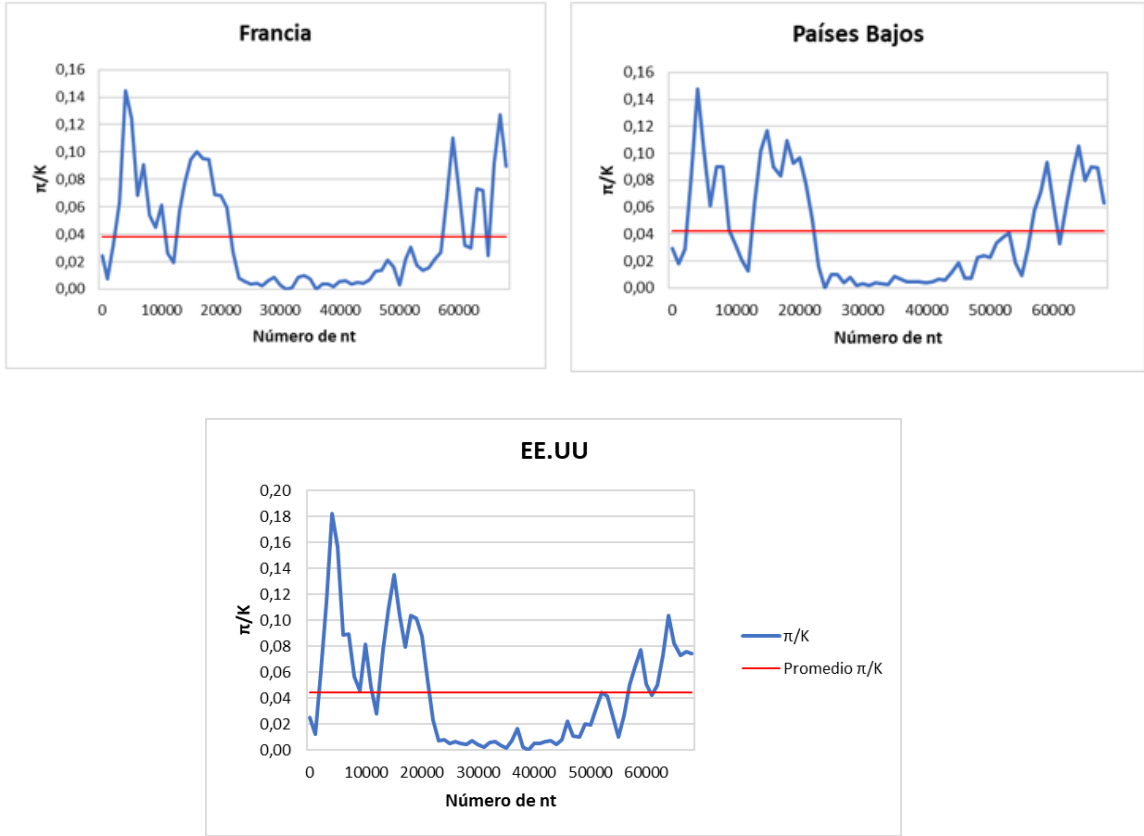


Figura 9: Distribución de la relación π/K para las poblaciones de Francia, Países Bajos y EE. UU.

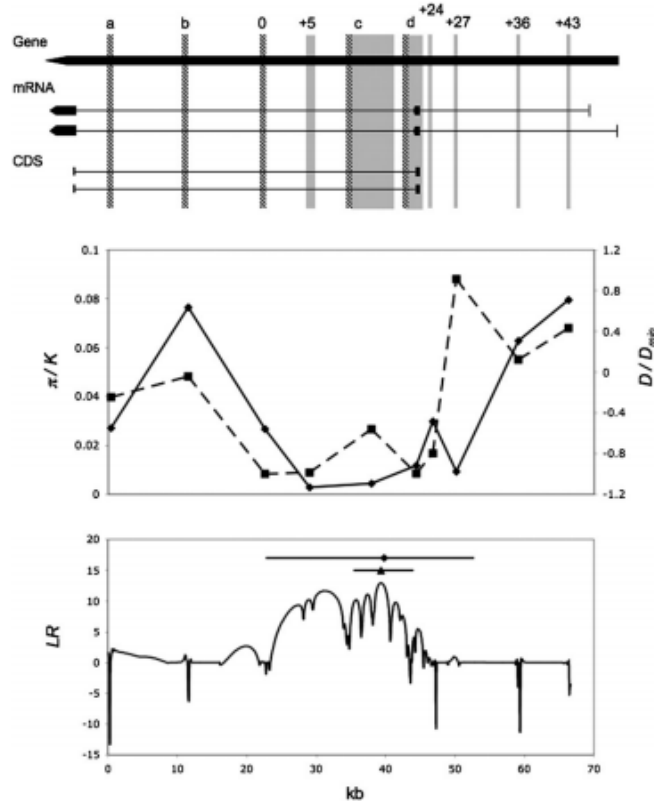


Figura 10: Análisis de la variación de la región del estudio realizado en la población catalana. En el gráfico central, la línea continua indica la relación π/K . Imagen extraída directamente del artículo publicado por Orengo, & Aguadé en 2010 sobre un estudio de la región en una población catalana⁵

A pesar de realizar comparaciones múltiples, en el análisis de significación estadística de los estadísticos de neutralidad no se han realizado correcciones FDR o de Bonferroni porque la secuencia se ha analizado en fragmentos solapados y que, por tanto, no son independientes entre ellos.

Por otra parte, como se trata de una distribución de dos colas en las que podemos tener valores observados muy pequeños (cola de la izquierda) o muy elevados (cola de la derecha), para obtener un nivel de confianza al 95 % se rechaza la hipótesis de la existencia de neutralidad si el nivel de significación α es $\leq 0,025$ al considerar cada cola independientemente.

Para el test D de Tajima, las poblaciones en las que se obtienen mayor número de fragmentos estadísticamente significativos son las poblaciones de Ruanda y Zambia. Aproximadamente el 83% de los fragmentos es significativo para la población de Ruanda y el 74 % para la población de Zambia. También hay que destacar que para estas poblaciones los valores de D que se observan son valores negativos lo que indica un exceso de polimorfismos a baja frecuencia.

Hay dos fragmentos solapados que tienen valores de D de Tajima estadísticamente significativos para las seis poblaciones. Estos fragmentos son los que se corresponden con las posiciones 48.001 - 50.000 y 49.001 - 51.000, que equivalen a las coordenadas 9.475.650 - 9.478.649 según la versión FB2021_01 de FlyBase (figura 11). Se trata de una región próxima al exón 2 del gen *mgl* que se localiza aproximadamente a 3 kb y que forma parte de una región en la que hay anotado un origen de replicación. En todos los casos el valor de D es negativo y los valores de H para esta zona no son significativos

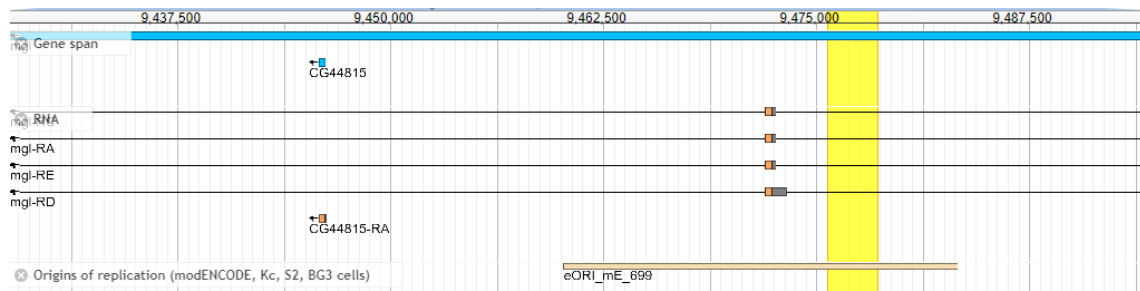


Figura 11: Localización de las posiciones 48.000 a 51.000 de la secuencia analizada. La región se ha marcado en amarillo.

Dentro de la secuencia de 67 Kb analizada en el trabajo, se encuentra anotado el gen *CG44815*. Este gen se encuentra aproximadamente a unas 29 kb de esta zona.

A partir de los resultados obtenidos para las poblaciones africanas observamos que:

- Para las poblaciones de Ruanda y Zambia el 100 % de los fragmentos significativos tienen valor de D negativo mientras que para la población de Etiopia este porcentaje baja al 95 %, aunque sigue siendo un porcentaje elevado.
- Hay una serie de fragmentos con valores de D de Tajima que son significativos para las 3 poblaciones. Alguno de ellos también es significativo en alguna de las poblaciones derivadas. Estos fragmentos, excluyendo los dos que son significativos para todas las poblaciones comentado anteriormente, son los que se muestran en la tabla 4.

Fragmento	Valor D	Observaciones
1-2000	Negativo	
12001-14000	Negativo	Significativo también en la población de Francia
20001-22000	Negativo	
21001-23000	Negativo	Significativo también en la población de EE. UU.
22001-24000	Negativo	Significativo también en la población de EE. UU.
23001-25000	Negativo	Significativo también en la población de Francia
28001-30000	Negativo	
35001-37000	Negativo	Significativo también en la población de Francia
46001-48000	Negativo	Significativo también en la población de Francia y EE. UU.
60001-62000	Negativo	
61001-63000	Negativo	
64001-66000	Negativo	Significativo también en la población de EE. UU. pero con valor de D positivo.

Tabla 4: Fragmentos significativos para D en las tres poblaciones africanas.

La zona comprendida entre las posiciones 20.000 - 25.000, que se corresponde con las coordenadas 9.447.649 - 9.452.649 y que se sitúa aproximadamente a 1,5 kb del gen *CG44815* (figura 12). En esta zona hay anotada una región reguladora de la transcripción.

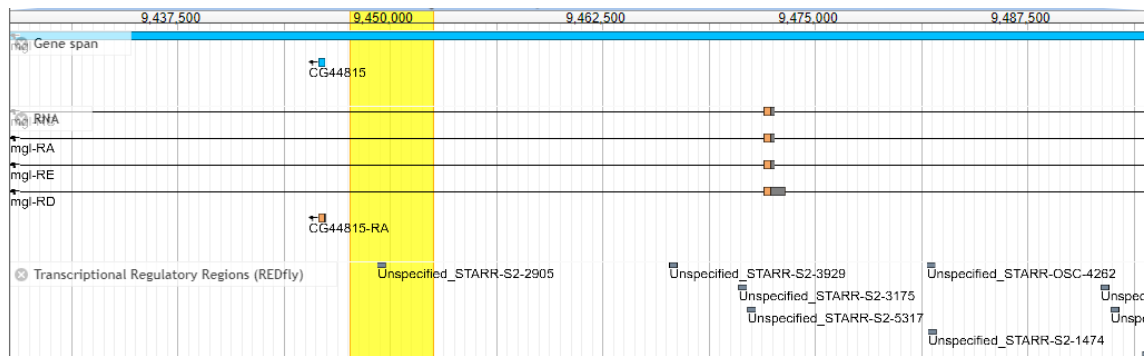


Figura 12: Localización de las posiciones 20.000 a 25.000 de la secuencia analizada. La región se ha marcado en amarillo.

La zona comprendida entre las posiciones 60.000 - 66.000 se corresponde con las coordenadas 9.487.649 – 9.493.649 (figura 13). En esta región también se encuentran anotadas regiones reguladoras de la transcripción

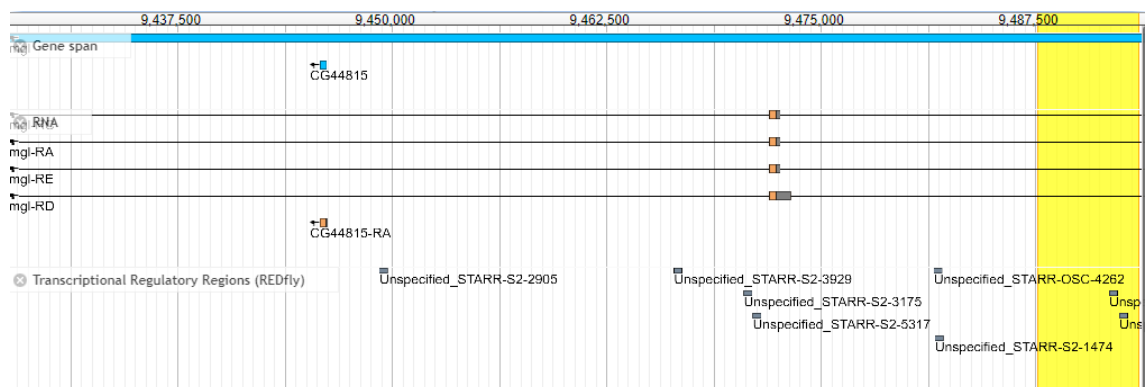


Figura 13: Localización de las posiciones 60.000 a 66.000 de la secuencia analizada. La región se ha marcado en amarillo.

En el caso de las poblaciones derivadas, observamos que:

- Se obtienen fragmentos significativos tanto para valores de D positivos como para valores de D negativos. El porcentaje de fragmentos con valores D negativo para las poblaciones varía entre el 55 % de la población de Francia y el 41 % de la de Países Bajos. El porcentaje de la población de EE. UU. está más cercano al valor de la población de Países Bajos siendo de un 44 %.
- Una serie de fragmentos son significativos para las tres poblaciones. En el caso de estos fragmentos, encontramos que también son significativos en alguna de las poblaciones africanas. En estos casos, en las poblaciones derivadas el valor de D es positivo por lo que encontramos un déficit de polimorfismos a baja frecuencia, mientras que las poblaciones africanas el valor de D es negativo. Estos fragmentos, excluyendo los dos en los que es significativo para las seis poblaciones y que ya se ha comentado con anterioridad son los que se muestran en la tabla 5.

Fragmento	Valor D	Observaciones
4001-6000	Positivo	Significativo también para las poblaciones de Ruanda y Zambia con valor de D negativo.
5001-7000	Positivo	Significativo también para las poblaciones de Ruanda y Zambia con valor de D negativo.
14001-16000	Positivo	Significativo también para las poblaciones de Ruanda y Zambia con valor de D negativo.
17001-19000	Positivo	Significativo también para las poblaciones de Ruanda y Zambia con valor de D negativo.
67001-69000	Positivo	Significativo también para la población de Zambia con valor de D negativo.
68001-69194	Positivo	Significativo también para la población de Zambia con valor de D negativo.

Tabla 5: Fragmentos significativos para D en las tres poblaciones derivadas

La zona comprendida entre las posiciones 14.000-19.000 se corresponde con las coordenadas 9.441.649 - 9.446.649. En esta región se encuentra anotado el gen **CG44815**, además de un sitio de inicio de la transcripción.

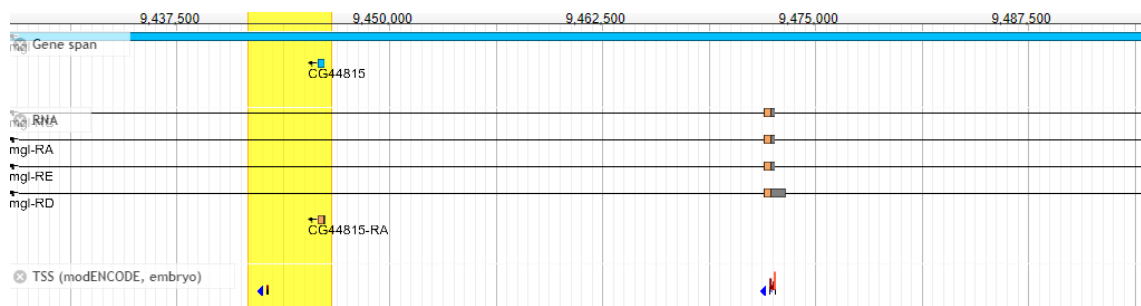


Figura 14: Localización de las posiciones 14.000 a 19.000 de la secuencia analizada. La región se ha marcado en amarillo.

La zona comprendida entre las posiciones 67.001-69.194 se corresponde con las coordenadas 9.494.650 – 9.496.843. En esta región se encuentra anotado el exón 1 del gen **mgl** para los transcritos **mgl-RE** y **mgl-RD**



Figura 15: Localización de las posiciones 67.001 a 69.194 de la secuencia analizada. La región se ha marcado en amarillo.

Para la zona comprendida entre las posiciones 4.001 -7.000, que se corresponde con las coordenadas 9.431.650 - 9.434.649, no se observa ninguna anotación especial en FlyBase.

Además de lo comentado, encontramos también un fragmento más en el que el valor de D obtenido es significativo en cinco de las poblaciones estudiadas. Este fragmento, que

se corresponde con las posiciones 46.001 - 48.000 de la secuencia y se localiza en las coordenadas 9.473.650-9.475.649, es significativo para todas las poblaciones excepto para la población de Países Bajos. Los valores de D que se obtienen son negativos mientras que los valores de H , aunque no significativos, son negativos para las poblaciones derivadas y positivos para las poblaciones africanas. Se trata de una región muy próxima al exón 2 del gen *mgI* que se sitúa aproximadamente a 400 pb.

Para el estadístico H de Fay y Wu se observa que el número de fragmentos que son estadísticamente significativos es mucho menor que los obtenidos para la D de Tajima.

La población para la que se ha obtenido más fragmentos significativos es la de Etiopia, en contraste con las otras dos poblaciones africanas que son para las que menos fragmentos significativos se han obtenido. La población de Etiopia tiene aproximadamente tres veces más de fragmentos significativos que los obtenidos para Ruanda y Zambia.

Los valores de H para los que se ha obtenido significación estadística son para casi todas las poblaciones, mayoritariamente negativos, lo que significa un exceso de variantes derivadas a frecuencias elevada. A excepción de la población de Zambia, donde el 100 % de los fragmentos significativos tienen valor de H positivo lo que significa un déficit de variantes derivadas a frecuencias elevadas. Para la población de Ruanda el porcentaje de valores positivos y negativos es del 50 %. En la otra población africana, la de Etiopia, el porcentaje de H negativo es del 88 %, un porcentaje elevado en comparación con las otras dos poblaciones africanas.

Para poblaciones derivadas, los valores de H negativos se encuentran entre el 100 % de la población de EE. UU. y el 78 % de la población de Francia.

Población	N.º total	N.º H positivo	N.º H negativo
Zambia (ZI)	3	3	0
Ruanda (RG)	4	2	2
Etiopia (EF)	11	9	2
Francia (FR)	9	7	2
Países Bajos (NTH)	8	7	1
EE. UU. (RAL)	6	0	6

Tabla 6: Número de fragmentos para los que se obtiene un valor de H significativo.

No hay ningún fragmento que sea significativo en las seis poblaciones analizadas. Hay un único fragmento significativo en las tres poblaciones africanas que es el fragmento localizado en las posiciones 65.001-67.000 de la secuencia que se corresponde con las coordenadas 9.492.650 - 9.494.649 y para el que se obtienen valores de H positivos. En valor de D en estos casos es negativo y significativo para las poblaciones de Ruanda y Zambia por lo que en base a los resultados de los dos indicadores esta zona podría estar sufriendo eventos de selección purificadora.

También hay un único fragmento significativo para las tres poblaciones derivadas localizado en las posiciones 63.001-65.000 que se corresponden con las coordenadas 9.490.650 - 9.492.649 y para el que se obtienen valores de H negativos. En este caso, la D no es significativa y salvo para la población de Francia tiene valores positivos.

De los 41 fragmentos para los que se ha obtenido significación para H en alguna de las poblaciones, hay 19 que son significativos para los dos estadísticos y de estos, en cinco

el valor es de signo contrario (tabla 7). En cuatro de ellos, para D se obtienen valores negativos y para H se obtienen valores positivos. Estos cuatro fragmentos son de las poblaciones de Ruanda y Zambia. En el quinto fragmento, que pertenece a la población de EE. UU., en cambio se obtiene un valor de D positivo y un valor de H negativo

Población	Ventana	D_obs	D_P	H_obs	H_P
Etiopia (EF)	35001-37000	-0,8488	0,012	-11,9333	0,001
Etiopia (EF)	46001-48000	-0,7703	0,025	-11,4857	0
Etiopia (EF)	47001-49000	-1,045	0,004	-10,6381	0
Francia (FR)	11001-13000	-1,3469	0,01	-8,9429	0
Francia (FR)	12001-14000	-1,2725	0,018	-6,6286	0
Francia (FR)	35001-37000	-2,04	0	-2,9714	0,019
Francia (FR)	44001-46000	-1,8159	0,008	-3,3429	0,003
Francia (FR)	62001-64000	-1,865	0	-3,7048	0,016
P. Bajos (NTH)	10001-12000	-1,483	0,004	-4,8571	0,008
P. Bajos (NTH)	11001-13000	-1,3074	0,014	-6,6762	0,001
P. Bajos (NTH)	55001-57000	-1,4512	0,016	-3,219	0,006
P. Bajos (NTH)	56001-58000	-1,2067	0,019	-5,1333	0,004
EE. UU. (RAL)	21001-23000	-0,9278	0,024	-4,8286	0,022
EE. UU. (RAL)	64001-66000	1,4206	0,002	-5,2	0,013
Ruanda (RG)	9001-11000	-1,1528	0	-13,0571	0,001
Ruanda (RG)	38001-40000	-1,332	0	4,5524	0,023
Ruanda (RG)	65001-67000	-1,3965	0	7,9619	0,002
Zambia (ZI)	64001-66000	-1,2387	0	8,7714	0,001
Zambia (ZI)	65001-67000	-1,1258	0	6,7429	0,007

Tabla 7: Fragmentos significativos para los estadísticos D y H .

Al representar gráficamente la distribución de los estadísticos D y H se observa que los valores positivos de H en las poblaciones derivadas son valores muy cercanos a cero y que en las zonas donde los valores de H son más negativos también se observan valores de D negativo (figura 17).

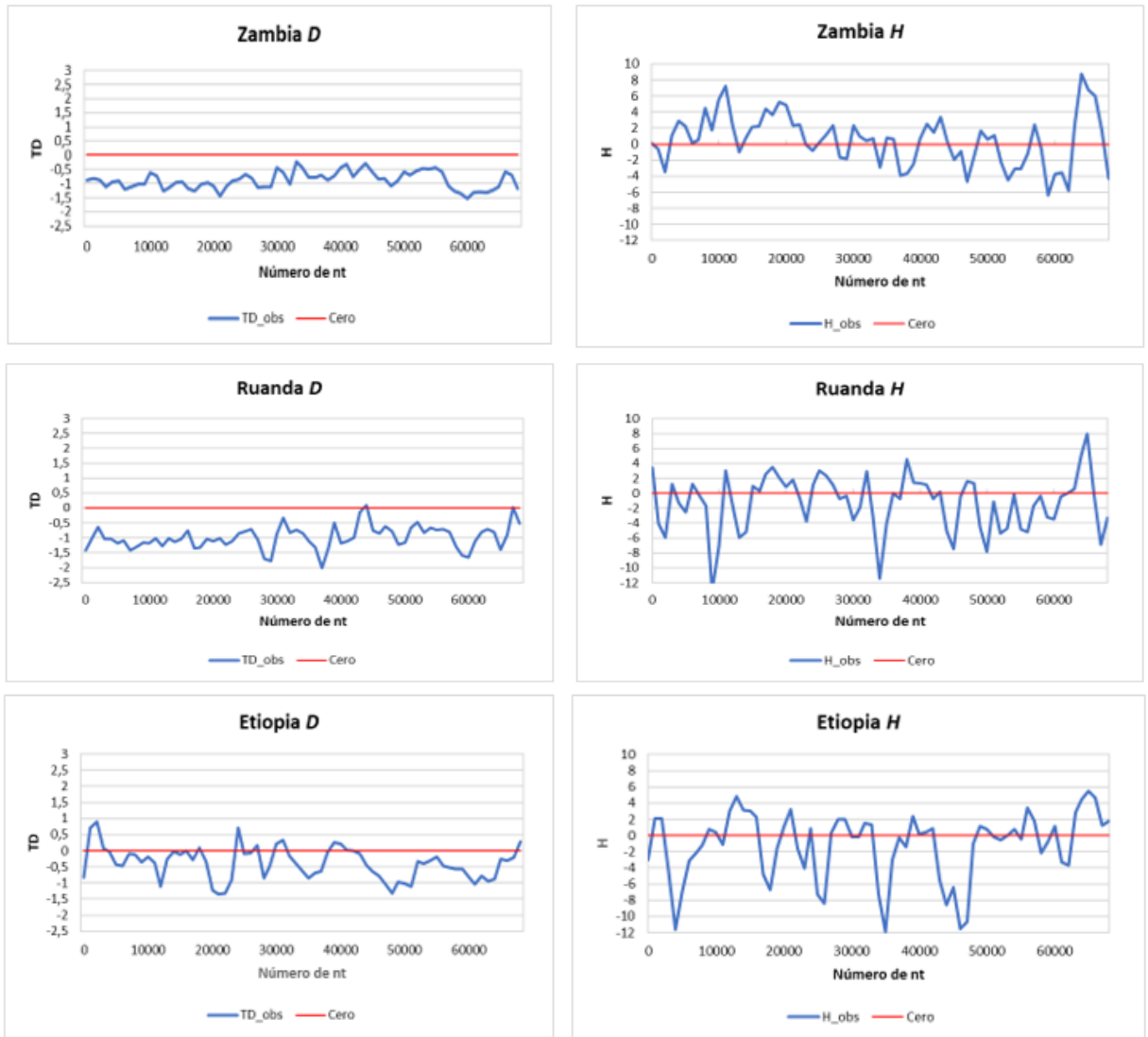


Figura 16: Distribución de los estadísticos de neutralidad D de Tajima y H de Fay y Wu en las poblaciones africanas

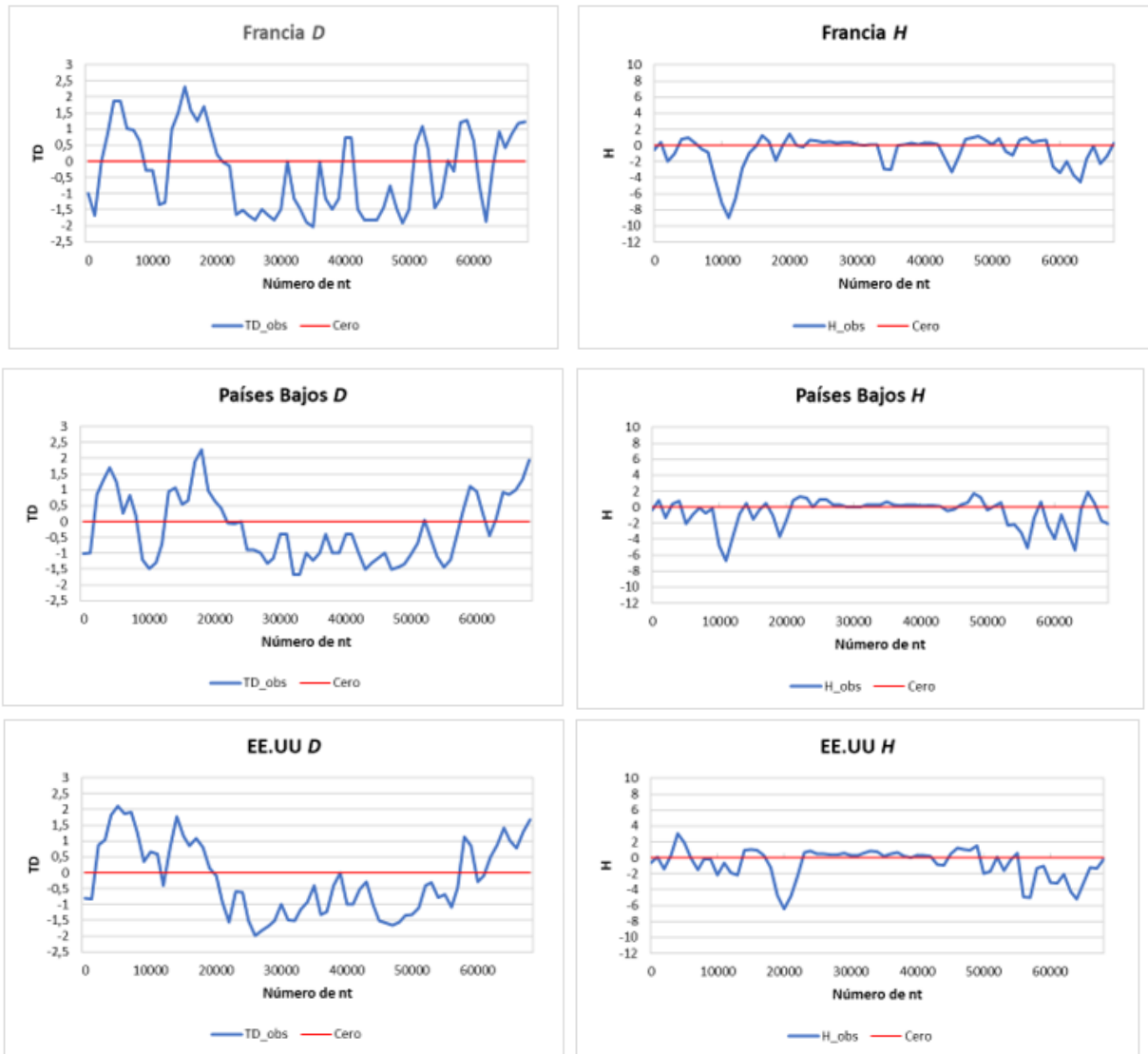


Figura 17: Distribución de los estadísticos de neutralidad D de Tajima y H de Fay y Wu en las poblaciones derivadas

6 Discusión

6.1 Análisis de los resultados de la diferenciación genética entre poblaciones.

Los resultados obtenidos en el análisis de la diferenciación genética nos indica que las seis poblaciones estudiadas son diferentes entre ellas.

Llama la atención los resultados obtenidos para la población de Etiopia en las comparaciones dos a dos que indican que es una población diferenciada de las demás. Los individuos de esta población etíope pertenecen a la localidad de Fiche. Esta región se encuentra aproximadamente a unos 2.700 metros sobre el mar en el macizo etíope en una zona de clima más templado. Estas dos características pueden haber influido en la diferenciación observada, la primera porque puede haber supuesto una barrera geográfica y la segunda por procesos adaptativos a climas más templados.

Para la población de EE. UU. no se ha obtenido diferenciación con respecto a las poblaciones europeas (estadístico *Fst*). La expansión de *Drosophila melanogaster* a América es relativamente reciente, tan solo de unos cientos de años y estos resultados puede que nos estén indicando que la expansión a este territorio tuviese su origen en poblaciones europeas.

Los dos estadísticos utilizados para el análisis de la diferenciación genética se encuentran muy correlacionados entre ellos con un coeficiente de correlación de 0,688 aunque parece ser más eficiente *el estadístico Fst*. Con el estadístico *Snn*, tras la corrección de Bonferroni secuencial, no se mantiene la significación de la diferenciación entre las poblaciones de Francia y Ruanda, Francia y Zambia, Países Bajos y EE. UU. mientras que el estadístico *Fst* sí que es capaz de mantener esta significación en la diferenciación entre poblaciones africanas y derivadas. Por otro lado, en el análisis de PCoA, los tres primeros ejes, para el estadístico *Fst* explican un 98,35 % de la variabilidad frente al 72,68 % de *Snn*.

6.2 Análisis de los indicadores de polimorfismo y divergencia genética.

Los resultados obtenidos para los indicadores de polimorfismo en la población etíope pueden ser compatibles con la existencia de un cuello de botella durante la colonización de Etiopia. La diversidad observada en la población de Zambia es aproximadamente un 70 % mayor que en la población etíope lo que concuerda con los resultados obtenidos por Pool y colaboradores en estudios sobre la diversidad africana de *D. melanogaster*² en los que se observó, para dos muestra etíopes, una variabilidad menor respecto a otras poblaciones africanas y que esta disminución llegaba a ser tres cuartas partes menor que la variabilidad observada en individuos de una población de Zambia.

En las poblaciones derivadas, al representar gráficamente la distribución de los estadísticos *D* y *H* observamos que los valores positivos de *H* son valores muy cercanos a cero y que, en esas zonas, los valores de *D* son negativos lo que podría indicar que la variación del polimorfismo puede ser debida a la existencia de un efecto de cuello de botella durante la colonización (figura 17).

Las regiones para las que se ha obtenido significación en los estadísticos de neutralidad son regiones que se encuentran cerca de regiones codificadoras o en regiones reguladoras de la transcripción lo que puede hacer que estas regiones se encuentren sometidas a procesos de selección purificadora.

Los resultados obtenidos para los indicadores S , π y K en las poblaciones derivadas, son similares a los resultados del estudio que se llevó a cabo en la población catalana⁵, donde también se observa un descenso pronunciado del polimorfismo en la región localizada aproximadamente en las posiciones 22.000 – 50.000 de la secuencia. Esto nos lleva a pensar que el efecto observado en la población catalana es común a las poblaciones derivadas y por lo tanto relacionado con el efecto colonizador.

Como se puede comprobar a partir de la relación entre π/K que se observa para las poblaciones derivadas (figura 9), se repite el patrón observado en la población catalana a pesar de que en el análisis de esta población se analizó una secuencia de 10 fragmentos más o menos equidistantes y en este trabajo se ha utilizado la secuencia continua analizando 69 ventanas solapadas. Esto nos hace pensar que todo el efecto que pueda tener esta región ya se recogía en el estudio previo, reforzando la posición de la diana de selección predicha por el estadístico LR de Kim y Stephan²⁷ que se describió muy cerca del segundo exón del gen *mgl* (figura 10).

La región analizada es una región que se encuentra dentro del gen *mgl*. Se ha descrito que este gen interviene en los procesos de pigmentación de la cutícula, regulando mediante endocitosis los niveles de la proteína Yellow⁶. Una disminución en los niveles de *mgl* hace que la pigmentación de la cutícula sea más oscura. Una teoría, totalmente especulativa, nos podría llevar a pensar que, al colonizar nuevas zonas dónde la exposición a la radiación solar es distinta que en los lugares de origen, alguna variante de un gen relacionado con la coloración pueda estar seleccionada.

7 Conclusiones

7.1 Conclusiones

En base a los resultados obtenidos durante la realización del trabajo se puede llegar a las siguientes conclusiones:

- Las seis poblaciones del estudio están genéticamente diferenciadas entre ellas, aunque cuando se realizan comparaciones dos a dos no se observa diferenciación entre algunas de las poblaciones. Las poblaciones que son significativamente diferentes varían en función del estadístico evaluado. Mientras que con el estadístico *Snn* observamos diferenciación significativa entre la población de Etiopía y el resto, entre la población de Ruanda y las poblaciones derivadas y también entre la de Países Bajos y Estados Unidos; con el estadístico *Fst*, detectamos diferenciación genética significativa entre cualquier par de poblaciones excepto para las de Francia y Países Bajos y Francia y EE. UU.
- La diferenciación genética entre las poblaciones, en base a los resultados obtenidos al realizar el test de Mantel, no está correlacionada con la distancia geográfica. Solo se aceptaría cierta correlación entre los resultados del estadístico *Fst* y la distancia geográfica en escala logarítmica excluyendo la población de EE. UU.
- Al comparar los dos estadísticos utilizados para estudiar la diferenciación genética, parece que *Fst* es más potente que *Snn*. Por un lado, se ha observado que el estadístico *Snn* debe saturarse antes que el *Fst* puesto que los *P-value* obtenidos por test de permutaciones son mayores, especialmente entre pares de poblaciones ancestrales y derivadas. Por otro lado, en el análisis de PCoA encontramos que el estadístico *Fst* explica mayor porcentaje de variabilidad en los dos primeros ejes.
- La diversidad nucleotídica observada en las poblaciones derivadas es menor que la observada en las poblaciones africanas. Además, podemos diferenciar dos patrones de variación diferentes, por un lado, el patrón seguido por las poblaciones derivadas y por otro, el patrón seguido por las poblaciones africanas.

El patrón observado para la población catalana es un patrón similar al observado en las poblaciones derivadas por lo que el efecto observado en la región parece que no es algo local, si no que es algo más general de las poblaciones derivadas.

- Las poblaciones de Zambia y Ruanda son las que presentan un mayor número de fragmentos para los que se rechazaría la neutralidad en función del test *D* de Tajima y con valor de *D* negativo. En cambio, para las poblaciones derivadas, los fragmentos para los que se rechazaría la neutralidad tienen un valor de *D* positivo.

Para el estadístico *H* de Fay y Wu, el número de fragmentos para los que rechazaría la hipótesis de neutralidad es bastante menor, siendo las poblaciones de Zambia y Ruanda las que menos tienen. Para la población de Zambia, todos los fragmentos significativos tienen un valor de *H* positivo. En cambio, para las poblaciones derivadas el valor de *H* para los fragmentos significativos tiene mayoritariamente valor negativo con un 89%.

En la secuencia analizada se detectan zonas en las que podría estar actuando los procesos de selección tanto para las poblaciones africanas como en las poblaciones derivadas.

7.2 Líneas de futuro

Basándonos en los resultados obtenidos tras el análisis de los dos estadísticos de neutralidad, se identifican una serie de fragmentos que podrían ser interesantes de cara a profundizar más en su estudio ya que en base a los resultados obtenidos parece que están sujetos a procesos de selección purificadora, al menos en las poblaciones africanas. Estos fragmentos son los siguientes:

- La región comprendida entre las posiciones 48.000 - 51.000 en la que la D de Tajima es significativa para todas las poblaciones y que quizá se podía ampliar a las posiciones 46.000 - 51.000 ya que el fragmento 46.000 - 48.000 es significativo para todas las poblaciones menos para la población de Países Bajos.
- La región comprendida entre las posiciones 14.000 - 19.000 donde se obtiene un valor de D positivo para las poblaciones derivadas y negativo para Ruanda y Zambia. Este fragmento no es significativo para Etiopia.
- La región comprendida entre las posiciones 20.000 - 25.000 en la que la D es significativa para todas las poblaciones africanas y algún fragmento de esta región también lo es para alguna de las poblaciones derivadas.
- La región comprendida entre las posiciones 60.000 - 66.000 en la que la D es significativa para las poblaciones africanas y que se podría extender hasta la posición 67.000. Dentro de esta zona encontramos el fragmento 65.000 - 67.000 que es significativo para los dos estadísticos.
- La región comprendida entre las posiciones 35.000 - 37.000 que es significativa para los dos estadísticos para las poblaciones de Etiopia y Francia pero que también es significativa para D en las poblaciones Ruanda y Zambia.

7.3 Seguimiento de la planificación

La planificación y la metodología establecida al inicio del semestre se ha podido cumplir sin incidentes destacables que hayan puesto en peligro la consecución de los objetivos planteados, que en líneas generales se han conseguido cumplir.

El único problema destacable se presentó al llevar a cabo la tarea de alineación. Inicialmente se había planteado realizar la alineación de las secuencias utilizando el programa bioinformático MUSCLE²⁸ pero debido al gran tamaño de la secuencia, el programa fallaba por requerimiento de memoria por lo que finalmente el alineamiento de las secuencias se realizó con el programa MAFFT recomendado para alineamiento de secuencias grandes a través de la opción web del mismo

8 Glosario

- *D*: Estadístico *D* de Tajima. Compara los sitios segregantes con la media de la variación nucleotídica en comparaciones dos a dos. Se utiliza para evaluar hipótesis de neutralidad.
- DnaSP: DNA Sequence Polymorphism. Programa bioinformático para realizar análisis de polimorfismos en secuencias de ADN²¹.
- *Fst*: Estadístico para determinar diferenciación genética entre poblaciones a partir de las frecuencias en los sitios polimórficos, tratando cada sitio como un locus separado¹¹.
- GenALEx: Complemento de Excel para el análisis genético de poblaciones²⁵.
- *H*: Estadístico *H* de Fay y Wu. Test de neutralidad que utiliza los datos de polimorfismo de la población y también datos de una especie *outgroup*. Mide la desviación de la neutralidad por alelos derivados de alta frecuencia. Sensible a los barridos selectivos.
- *K*: Divergencia observada. Número medio de diferencias nucleotídicas observadas al comparar secuencias dos a dos.
- MAFFT: Programa bioinformático para realizar alineamientos múltiples de secuencias de nucleótidos o aminoácidos¹⁴.
- *Megalin (mgl)*: Gen que codifica una proteína relacionada con el receptor de lipoproteínas de baja densidad (LRP). Se expresa en los tejidos epiteliales e interviene en la endocitosis de una amplia gama de sustancias, una de ellas es la proteína del gen *yellow*⁶.
- mlcoalsim: Multilocus coalescent simulations. Programa bioinformático para realizar simulaciones y calcular análisis multilocus²³.
- mstatspop: Programa bioinformático para realizar análisis de polimorfismos de secuencias de ADN²².
- π : Diversidad nucleotídica. Número medio de diferencias que se espera observar por nucleótido.
- PCoA: Análisis de coordenadas principales representamos los valores de la distancia entre entidades en distintas dimensiones (coordenadas). Método para explorar y visualizar las similitudes o disimilitudes de los datos a partir de una matriz de distancias.
- *S*: Número de posiciones segregantes.
- *Snn* (Statistic nearest-neighbor) Estadístico para determinar diferenciaciones genéticas. Es una medida de la frecuencia con la que los "vecinos más cercanos" (en el espacio de la secuencia) de las secuencias son de la misma localidad en el espacio geográfico¹⁰.
- Test de Mantel: Test estadístico para medir la correlación entre dos matrices del mismo rango.

9 Bibliografía

1. KIMURA, M. Evolutionary Rate at the Molecular Level. *Nature* **217**, 624–626 (1968).
2. Pool, J. E. *et al.* Population Genomics of Sub-Saharan *Drosophila melanogaster*: African Diversity and Non-African Admixture. *PLoS Genet.* **8**, (2012).
3. Sprengelmeyer, Q. D. *et al.* Recurrent Collection of *Drosophila melanogaster* from Wild African Environments and Genomic Insights into Species History. *Mol. Biol. Evol.* **37**, 627–638 (2020).
4. Kapopoulou, A., Pfeifer, S. P., Jensen, J. D. & Laurent, S. The demographic history of african *Drosophila melanogaster*. *Genome Biol. Evol.* **10**, 2338–2342 (2018).
5. Orengo, D. J. & Aguadé, M. Uncovering the footprint of positive selection on the X chromosome of *Drosophila melanogaster*. *Mol. Biol. Evol.* **27**, 153–160 (2010).
6. Riedel, F., Vorkel, D. & Eaton, S. Megalin-dependent Yellow endocytosis restricts melanization in the *Drosophila* cuticle. *Development* **138**, 149–158 (2011).
7. Hervas, S., Sanz, E., Casillas, S., Pool, J. E. & Barbadilla, A. PopFly: the *Drosophila* population genomics browser. *Bioinformatics* **33**, 2779–2780 (2017).
8. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
9. Fay, J. C. & Wu, C. I. Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405–1413 (2000).
10. Hudson, R. R. A new statistic for detecting genetic differentiation. *Genetics* **155**, 2011–2014 (2000).
11. Hudson, R. R., Slatkin, M. & Maddison, W. P. Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**, 583–589 (1992).
12. Larkin, A. *et al.* FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res.* **49**, D899–D907 (2021).
13. Howe, K. L. *et al.* Ensembl Genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Res.* **48**, D689–D695 (2020).
14. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2018).
15. Glinka, S., Ometto, L., Mousset, S., Stephan, W. & De Lorenzo, D. Demography and Natural Selection Have Shaped Genetic Variation in *Drosophila melanogaster*: A Multi-locus Approach. *Genetics* **165**, 1269–1278 (2003).
16. Orengo, D. J. & Aguadé, M. Detecting the footprint of positive selection in a European population of *Drosophila melanogaster*: Multilocus pattern of variation and distance to coding regions. *Genetics* **167**, 1759–1766 (2004).
17. Kapun, M. *et al.* Genomic analysis of european *Drosophila melanogaster* populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Mol. Biol. Evol.* **37**, 2661–2678 (2020).
18. Beisswanger, S., Stephan, W. & De Lorenzo, D. Evidence for a selective sweep in the wapl region of *Drosophila melanogaster*. *Genetics* **172**, 265–274 (2006).
19. Orengo, D. J. & Aguadé, M. Genome scans of variation and adaptive change:

- Extended analysis of a candidate locus close to the phantom gene region in *Drosophila melanogaster*. *Mol. Biol. Evol.* **24**, 1122–1129 (2007).
20. <https://www.ncbi.nlm.nih.gov/genbank/>.
 21. Rozas, J. *et al.* DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302 (2017).
 22. <https://github.com/cragenomica/mstatspop>.
 23. Ramos-Onsins, S. E. & Mitchell-Olds, T. Mlcoalsim: Multilocus coalescent simulations. *Evol. Bioinforma.* **3**, 41–44 (2007).
 24. Hey, J. & Kliman, R. M. Interactions between natural selection, recombination and gene density in the genes of *Drosophila*. *Genetics* **160**, 595–608 (2002).
 25. Peakall, R. & Smouse, P. E. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539 (2012).
 26. Rice, W. R. Analyzing Tables of Statistical Tests. *Evolution* vol. 43 223 (1989).
 27. Kim, Y. & Stephan, W. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* **160**, 765–777 (2002).
 28. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

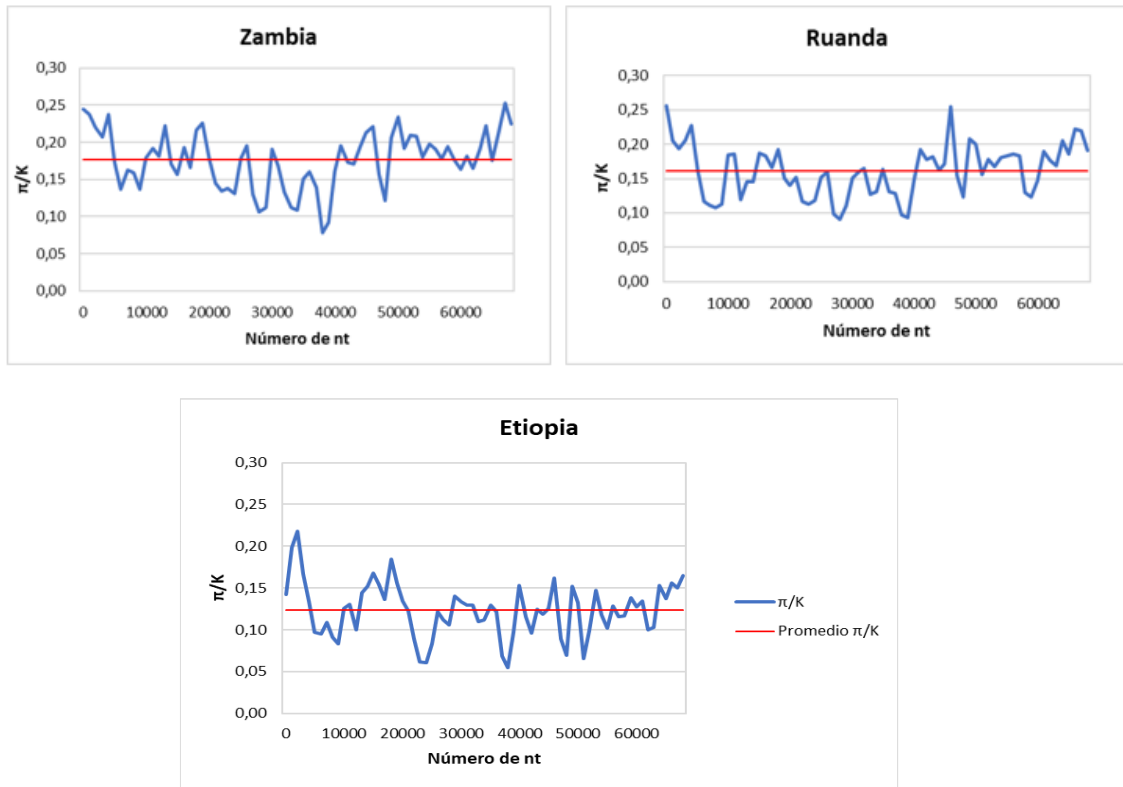
10 Anexos

Anexo 1: Distribución del número de posiciones segregantes (S)



La flecha de la parte superior indica la posición aproximada en la que se encuentra el gen *CG44815*.

Anexo 2: Distribución de la relación π/K para las poblaciones africanas



Anexo 3: Calculo de indicadores de polimorfismos y estadísticos de neutralidad.

Anexo 3: Indicador	Ventana	π	S	K	π / K	D	P-D	H	P-H
EF	1-2000	0,0042	31	0,02945	0,14261	-0,8299	0,022	-3,0476	0,123
EF	1001-3000	0,00918	45	0,04633	0,19814	0,699	0,04	2,1238	0,272
EF	2001-4000	0,01178	57	0,05407	0,21787	0,8959	0,004	2,1143	0,289
EF	3001-5000	0,00729	41	0,04368	0,16690	0,0678	0,423	-4,419	0,091
EF	4001-6000	0,00607	36	0,04519	0,13432	-0,0275	0,451	-11,6095	0,001
EF	5001-7000	0,00516	35	0,05338	0,09667	-0,4363	0,131	-6,9143	0,012
EF	6001-8000	0,00553	34	0,05787	0,09556	-0,4705	0,114	-3,1048	0,135
EF	7001-9000	0,00544	31	0,05001	0,10878	-0,0887	0,412	-2,1714	0,202
EF	8001-10000	0,0051	32	0,056	0,09107	-0,1281	0,385	-1,2476	0,306
EF	9001-11000	0,00518	30	0,06247	0,08292	-0,3509	0,205	0,7524	0,416
EF	10001-12000	0,00701	38	0,0558	0,12563	-0,1774	0,341	0,419	0,472
EF	11001-13000	0,00718	46	0,05493	0,13071	-0,3683	0,143	-1,1429	0,333
EF	12001-14000	0,00521	41	0,05203	0,10013	-1,1036	0	3,0095	0,151
EF	13001-15000	0,00653	40	0,04517	0,14456	-0,2666	0,231	4,8381	0,033
EF	14001-16000	0,00648	38	0,04253	0,15236	-0,0351	0,458	3,0952	0,144
EF	15001-17000	0,00688	39	0,04104	0,16764	-0,1266	0,378	3,0476	0,143
EF	16001-18000	0,00753	40	0,04903	0,15358	0,001	0,488	2,2762	0,225
EF	17001-19000	0,00653	40	0,04775	0,13675	-0,2765	0,246	-4,7905	0,073
EF	18001-20000	0,00708	41	0,03843	0,18423	0,0969	0,396	-6,6857	0,034
EF	19001-21000	0,0059	38	0,0377	0,15650	-0,3231	0,217	-1,5714	0,269
EF	20001-22000	0,00465	38	0,03469	0,13404	-1,2185	0	1,0476	0,381
EF	21001-23000	0,00491	41	0,03992	0,12300	-1,3425	0	3,1905	0,133
EF	22001-24000	0,00418	36	0,047	0,08894	-1,3247	0	-1,619	0,306
EF	23001-25000	0,00332	25	0,05354	0,06201	-0,9039	0,016	-4,0952	0,049
EF	24001-26000	0,0039	19	0,06391	0,06102	0,7077	0,073	0,8667	0,386
EF	25001-27000	0,00545	33	0,06539	0,08335	-0,0859	0,427	-7,2571	0,011
EF	26001-28000	0,00695	43	0,0571	0,12172	-0,0696	0,438	-8,4381	0,013
EF	27001-29000	0,00503	27	0,04507	0,11160	0,154	0,348	0,3333	0,475
EF	28001-30000	0,00525	34	0,04963	0,10578	-0,8563	0,018	2,0095	0,241
EF	29001-31000	0,00641	39	0,04565	0,14042	-0,4719	0,114	2,0095	0,265
EF	30001-32000	0,00598	32	0,0449	0,13318	0,2152	0,306	-0,1619	0,416
EF	31001-33000	0,0076	33	0,05862	0,12965	0,3229	0,212	-0,2	0,414
EF	32001-34000	0,00576	30	0,04433	0,12993	-0,1727	0,362	1,5619	0,277
EF	33001-35000	0,00597	34	0,05421	0,11013	-0,3856	0,181	1,3143	0,366
EF	34001-36000	0,00885	45	0,07926	0,11166	-0,6173	0,046	-7,2762	0,016
EF	35001-37000	0,00719	48	0,05558	0,12936	-0,8488	0,012	-11,9333	0,001
EF	36001-38000	0,0049	28	0,04023	0,12180	-0,6963	0,066	-2,9905	0,114
EF	37001-39000	0,00209	10	0,03043	0,06868	-0,6473	0,164	-0,3048	0,349
EF	38001-40000	0,00317	17	0,05822	0,05445	-0,0366	0,48	-1,4476	0,209

EF	39001-41000	0,00618	34	0,0636	0,09717	0,2626	0,259	2,3714	0,181
EF	40001-42000	0,00593	33	0,0388	0,15284	0,2158	0,312	0,181	0,492
EF	41001-43000	0,00416	25	0,03585	0,11604	0,0139	0,511	0,4286	0,481
EF	42001-44000	0,00405	23	0,042	0,09643	0,0071	0,505	0,8476	0,393
EF	43001-45000	0,00776	45	0,0623	0,12456	-0,0949	0,375	-5,4667	0,046
EF	44001-46000	0,00842	57	0,07082	0,11889	-0,4549	0,084	-8,5524	0,022
EF	45001-47000	0,0057	41	0,04531	0,12580	-0,6389	0,051	-6,4381	0,038
EF	46001-48000	0,00616	42	0,03802	0,16202	-0,7703	0,025	-11,4857	0
EF	47001-49000	0,00537	38	0,06018	0,08923	-1,045	0,004	-10,6381	0
EF	48001-50000	0,00349	28	0,0505	0,06911	-1,3175	0,004	-1	0,339
EF	49001-51000	0,00567	39	0,03719	0,15246	-0,9662	0,005	1,1048	0,407
EF	50001-52000	0,00595	41	0,04482	0,13275	-1,0068	0,004	0,7524	0,435
EF	51001-53000	0,00305	23	0,04632	0,06585	-1,1074	0,013	-0,1333	0,45
EF	52001-54000	0,00449	28	0,04553	0,09862	-0,3346	0,233	-0,5429	0,382
EF	53001-55000	0,00602	38	0,04101	0,14679	-0,3995	0,153	0,0286	0,53
EF	54001-56000	0,00514	31	0,04362	0,11784	-0,3077	0,264	0,7429	0,401
EF	55001-57000	0,00505	28	0,04944	0,10214	-0,1956	0,315	-0,5048	0,364
EF	56001-58000	0,00635	39	0,04945	0,12841	-0,4787	0,093	3,3714	0,115
EF	57001-59000	0,00566	36	0,04871	0,11620	-0,5281	0,095	1,781	0,277
EF	58001-60000	0,00504	28	0,04325	0,11653	-0,5664	0,087	-2,2	0,193
EF	59001-61000	0,00579	32	0,04174	0,13872	-0,5777	0,086	-0,7048	0,36
EF	60001-62000	0,00652	43	0,0512	0,12734	-0,8249	0,012	1,1905	0,381
EF	61001-63000	0,00708	47	0,05262	0,13455	-1,0324	0,002	-3,2762	0,147
EF	62001-64000	0,00448	26	0,04494	0,09969	-0,7833	0,03	-3,6571	0,085
EF	63001-65000	0,00407	28	0,03964	0,10267	-0,9512	0,009	2,8476	0,119
EF	64001-66000	0,00494	35	0,03225	0,15318	-0,8754	0,007	4,4762	0,029
EF	65001-67000	0,00477	27	0,03461	0,13782	-0,2681	0,306	5,5048	0
EF	66001-68000	0,00499	29	0,03191	0,15638	-0,3026	0,232	4,6571	0,012
EF	67001-69000	0,00519	32	0,03454	0,15026	-0,2058	0,303	1,2667	0,362
EF	68001-69194	0,00837	29	0,05086	0,16457	0,276	0,278	1,8381	0,26
FR	1-2000	0,00081	7	0,03365	0,02407	-0,9885	0,077	-0,5905	0,282
FR	1001-3000	0,0003	3	0,03911	0,00767	-1,685	0,044	0,3714	0,524
FR	2001-4000	0,00164	6	0,04943	0,03318	0,0045	0,485	-2,019	0,05
FR	3001-5000	0,00283	12	0,04473	0,06327	0,8897	0,08	-1,0762	0,212
FR	4001-6000	0,00644	26	0,04449	0,14475	1,8609	0	0,7619	0,392
FR	5001-7000	0,00668	28	0,05365	0,12451	1,8722	0	0,981	0,404
FR	6001-8000	0,00406	18	0,05952	0,06821	1,0111	0,021	0,3238	0,476
FR	7001-9000	0,0046	21	0,05059	0,09093	0,9693	0,016	-0,4667	0,373
FR	8001-10000	0,00308	16	0,05732	0,05373	0,6324	0,111	-0,819	0,314
FR	9001-11000	0,00285	15	0,06315	0,04513	-0,2834	0,311	-4,1429	0,028

FR	10001-12000	0,00314	17	0,05122	0,06130	-0,2868	0,289	-7,2	0
FR	11001-13000	0,00142	13	0,05453	0,02604	-1,3469	0,01	-8,9429	0
FR	12001-14000	0,00103	9	0,05303	0,01942	-1,2725	0,018	-6,6286	0
FR	13001-15000	0,00264	12	0,04601	0,05738	0,9902	0,054	-2,8095	0,037
FR	14001-16000	0,00345	15	0,04458	0,07739	1,5134	0,001	-0,8952	0,28
FR	15001-17000	0,00375	14	0,03977	0,09429	2,314	0	-0,1238	0,412
FR	16001-18000	0,0044	18	0,04387	0,10030	1,584	0,001	1,2	0,313
FR	17001-19000	0,00431	20	0,0452	0,09535	1,2441	0,01	0,5048	0,446
FR	18001-20000	0,0041	18	0,0433	0,09469	1,7097	0	-1,9048	0,157
FR	19001-21000	0,00355	17	0,05163	0,06876	0,9496	0,032	0,0095	0,544
FR	20001-22000	0,00324	18	0,04727	0,06854	0,2148	0,312	1,4095	0,244
FR	21001-23000	0,00261	16	0,04382	0,05956	-0,0286	0,488	0,0381	0,566
FR	22001-24000	0,00125	8	0,04509	0,02772	-0,1331	0,45	-0,1333	0,407
FR	23001-25000	0,00041	5	0,0516	0,00795	-1,6601	0,014	0,7048	0,323
FR	24001-26000	0,00035	4	0,06339	0,00552	-1,5181	0,039	0,581	0,336
FR	25001-27000	0,00021	3	0,06124	0,00343	-1,685	0,039	0,3714	0,491
FR	26001-28000	0,00027	4	0,06053	0,00446	-1,8159	0,011	0,4952	0,4
FR	27001-29000	0,00014	2	0,05394	0,00260	-1,4905	0,117	0,2476	0,611
FR	28001-30000	0,00026	3	0,04353	0,00597	-1,685	0,037	0,3714	0,489
FR	29001-31000	0,00034	4	0,03983	0,00854	-1,8159	0,01	0,3714	0,519
FR	30001-32000	0,00014	2	0,04835	0,00290	-1,4905	0,108	0,1238	0,656
FR	31001-33000	0	0	0,07336	0,00000	0	n.a	0	1
FR	32001-34000	0,00008	1	0,06064	0,00132	-1,1594	0,335	0,1238	0,73
FR	33001-35000	0,00067	7	0,07886	0,00850	-1,4664	0,016	0,1429	0,528
FR	34001-36000	0,0011	13	0,11157	0,00986	-1,8902	0	-2,9524	0,054
FR	35001-37000	0,0005	7	0,06919	0,00723	-2,04	0	-2,9714	0,019
FR	36001-38000	0	0	0,04219	0,00000	0	n.a	0	1
FR	37001-39000	0,00011	1	0,03219	0,00342	-1,1594	0,348	0,1238	0,758
FR	38001-40000	0,0002	2	0,05312	0,00377	-1,4905	0,118	0,2476	0,621
FR	39001-41000	0,00009	1	0,05832	0,00154	-1,1594	0,329	0,1238	0,779
FR	40001-42000	0,00022	1	0,03955	0,00556	0,7421	0,385	0,2667	0
FR	41001-43000	0,00022	1	0,03689	0,00596	0,7421	0,356	0,2667	0
FR	42001-44000	0,00015	2	0,04202	0,00357	-1,4905	0,116	0,1238	0,66
FR	43001-45000	0,0003	4	0,06175	0,00486	-1,8159	0,011	-1,6095	0,052
FR	44001-46000	0,00029	4	0,06894	0,00421	-1,8159	0,008	-3,3429	0,003
FR	45001-47000	0,00028	4	0,04077	0,00687	-1,8159	0,016	-1,3619	0,077
FR	46001-48000	0,00048	5	0,03681	0,01304	-1,4512	0,025	0,7524	0,297
FR	47001-49000	0,00082	6	0,06108	0,01343	-0,7544	0,152	0,9238	0,239
FR	48001-50000	0,00109	10	0,05191	0,02100	-1,4724	0,007	1,1238	0,26
FR	49001-51000	0,00069	8	0,04199	0,01643	-1,9149	0	0,7048	0,321

FR	50001-52000	0,00015	2	0,04991	0,00301	-1,4905	0,11	0,1238	0,646
FR	51001-53000	0,00113	6	0,05147	0,02195	0,5104	0,267	0,8476	0,274
FR	52001-54000	0,00148	7	0,0487	0,03039	1,0826	0,066	-0,781	0,26
FR	53001-55000	0,00075	4	0,04365	0,01718	0,3681	0,307	-1,2381	0,113
FR	54001-56000	0,00061	6	0,04531	0,01346	-1,441	0,021	0,6381	0,379
FR	55001-57000	0,00082	7	0,05274	0,01555	-1,1159	0,059	0,9905	0,245
FR	56001-58000	0,00119	7	0,05491	0,02167	0,0311	0,51	0,3905	0,446
FR	57001-59000	0,00129	8	0,04795	0,02690	-0,3184	0,337	0,5429	0,386
FR	58001-60000	0,00305	12	0,04533	0,06728	1,2114	0,018	0,6476	0,394
FR	59001-61000	0,00562	22	0,05096	0,11028	1,2752	0,001	-2,6286	0,101
FR	60001-62000	0,0039	19	0,05365	0,07269	0,6279	0,103	-3,4	0,04
FR	61001-63000	0,00161	11	0,05069	0,03176	-0,7727	0,098	-2	0,087
FR	62001-64000	0,0012	12	0,03965	0,03026	-1,865	0	-3,7048	0,016
FR	63001-65000	0,00266	16	0,03627	0,07334	-0,1919	0,356	-4,5333	0,018
FR	64001-66000	0,00234	11	0,03249	0,07202	0,92	0,072	-1,6667	0,137
FR	65001-67000	0,00079	4	0,03232	0,02444	0,4177	0,316	-0,0762	0,346
FR	66001-68000	0,00295	14	0,03222	0,09156	0,8409	0,071	-2,2762	0,098
FR	67001-69000	0,00442	20	0,0348	0,12701	1,1806	0,009	-1,2952	0,239
FR	68001-69194	0,00433	12	0,04849	0,08930	1,2315	0,013	0,3048	0,518
NTH	1-2000	0,00088	7	0,02974	0,02959	-1,0203	0,072	-0,3429	0,32
NTH	1001-3000	0,00074	5	0,04048	0,01828	-0,9916	0,09	0,8857	0,197
NTH	2001-4000	0,00138	5	0,04764	0,02897	0,847	0,119	-1,2952	0,094
NTH	3001-5000	0,00346	13	0,0416	0,08317	1,3132	0,01	0,381	0,452
NTH	4001-6000	0,00632	24	0,04274	0,14787	1,7139	0	0,8	0,403
NTH	5001-7000	0,00538	23	0,05257	0,10234	1,2499	0,006	-2,0667	0,167
NTH	6001-8000	0,00382	19	0,06282	0,06081	0,2556	0,34	-0,981	0,282
NTH	7001-9000	0,00472	21	0,05232	0,09021	0,8238	0,056	-0,1048	0,435
NTH	8001-10000	0,00426	19	0,04723	0,09020	0,1891	0,364	-0,7333	0,31
NTH	9001-11000	0,00271	12	0,06283	0,04313	-1,2015	0,023	-0,219	0,393
NTH	10001-12000	0,00185	12	0,05798	0,03191	-1,483	0,004	-4,8571	0,008
NTH	11001-13000	0,00115	10	0,05486	0,02096	-1,3074	0,014	-6,6762	0,001
NTH	12001-14000	0,00072	5	0,05734	0,01256	-0,6991	0,179	-3,7048	0,002
NTH	13001-15000	0,0031	13	0,04985	0,06219	0,9386	0,029	-0,8381	0,286
NTH	14001-16000	0,00454	19	0,04435	0,10237	1,0667	0,019	0,4571	0,479
NTH	15001-17000	0,00457	22	0,03915	0,11673	0,555	0,119	-1,4571	0,235
NTH	16001-18000	0,0039	18	0,04326	0,09015	0,6618	0,083	-0,3905	0,385
NTH	17001-19000	0,00371	14	0,04467	0,08305	1,8931	0	0,4381	0,449
NTH	18001-20000	0,00443	16	0,04044	0,10955	2,2808	0	-1,0476	0,257
NTH	19001-21000	0,00436	20	0,04696	0,09284	0,9777	0,015	-3,7143	0,035
NTH	20001-22000	0,00408	20	0,04235	0,09634	0,6352	0,104	-1,5524	0,208

NTH	21001-23000	0,00314	16	0,0407	0,07715	0,4302	0,192	0,8952	0,355
NTH	22001-24000	0,00236	13	0,04643	0,05083	-0,0356	0,481	1,3524	0,213
NTH	23001-25000	0,0009	5	0,05608	0,01605	-0,0723	0,493	1,1524	0,01
NTH	24001-26000	0	0	0,06352	0,00000	0	n.a	0	1
NTH	25001-27000	0,00063	5	0,06215	0,01014	-0,908	0,12	0,9619	0,117
NTH	26001-28000	0,00061	5	0,06297	0,00969	-0,908	0,125	0,9619	0,109
NTH	27001-29000	0,00022	2	0,05253	0,00419	-1,0016	0,21	0,3333	0,443
NTH	28001-30000	0,0004	3	0,05131	0,00780	-1,3165	0,101	0,3333	0,52
NTH	29001-31000	0,00011	1	0,05298	0,00208	-1,1594	0,331	0	0,773
NTH	30001-32000	0,00015	1	0,05061	0,00296	-0,3988	0,484	0	0,785
NTH	31001-33000	0,00016	1	0,08028	0,00199	-0,3988	0,508	0	0,826
NTH	32001-34000	0,00027	3	0,06808	0,00397	-1,685	0,039	0,2476	0,584
NTH	33001-35000	0,00024	3	0,07372	0,00326	-1,685	0,036	0,2476	0,558
NTH	34001-36000	0,00025	2	0,09677	0,00258	-1,0016	0,217	0,3333	0,456
NTH	35001-37000	0,00048	4	0,05463	0,00879	-1,2203	0,075	0,6667	0,243
NTH	36001-38000	0,00029	2	0,04242	0,00684	-1,0016	0,244	0,3333	0,454
NTH	37001-39000	0,00019	1	0,04213	0,00451	-0,3988	0,508	0,2095	0,42
NTH	38001-40000	0,00025	2	0,05844	0,00428	-1,0016	0,236	0,3333	0,422
NTH	39001-41000	0,00026	2	0,05892	0,00441	-1,0016	0,197	0,3333	0,457
NTH	40001-42000	0,00015	1	0,03802	0,00395	-0,3988	0,506	0,2095	0,396
NTH	41001-43000	0,00017	1	0,03535	0,00481	-0,3988	0,489	0,2095	0,415
NTH	42001-44000	0,00028	2	0,04221	0,00663	-1,0016	0,233	0,2095	0,599
NTH	43001-45000	0,00041	4	0,06631	0,00618	-1,5181	0,042	0,1238	0,595
NTH	44001-46000	0,00079	7	0,07143	0,01106	-1,3071	0,03	-0,4476	0,312
NTH	45001-47000	0,00075	6	0,04063	0,01846	-1,1519	0,05	-0,2381	0,388
NTH	46001-48000	0,00026	2	0,03493	0,00744	-1,0016	0,248	0,3333	0,454
NTH	47001-49000	0,00045	4	0,06455	0,00697	-1,5181	0,043	0,581	0,309
NTH	48001-50000	0,00136	10	0,05991	0,02270	-1,4488	0,009	1,6667	0,085
NTH	49001-51000	0,00123	9	0,05061	0,02430	-1,3499	0,017	1,2095	0,212
NTH	50001-52000	0,00119	9	0,05257	0,02264	-1,0403	0,046	-0,4	0,326
NTH	51001-53000	0,00162	11	0,04885	0,03316	-0,6859	0,112	0,1429	0,514
NTH	52001-54000	0,00181	10	0,04822	0,03754	0,0363	0,48	0,5714	0,409
NTH	53001-55000	0,00175	11	0,04301	0,04069	-0,4906	0,214	-2,2381	0,085
NTH	54001-56000	0,00088	7	0,04721	0,01864	-1,1159	0,055	-2,1429	0,054
NTH	55001-57000	0,00051	5	0,05558	0,00918	-1,4512	0,016	-3,219	0,006
NTH	56001-58000	0,00153	11	0,05122	0,02987	-1,2067	0,019	-5,1333	0,004
NTH	57001-59000	0,00257	12	0,04453	0,05771	-0,4776	0,204	-1,3048	0,179
NTH	58001-60000	0,00312	12	0,04383	0,07118	0,3669	0,275	0,6952	0,344
NTH	59001-61000	0,0045	17	0,04816	0,09344	1,1115	0,012	-2,3619	0,114
NTH	60001-62000	0,00375	13	0,06018	0,06231	0,9573	0,04	-3,9429	0,02

NTH	61001-63000	0,00201	6	0,06109	0,03290	0,2574	0,366	-0,8952	0,19
NTH	62001-64000	0,00258	11	0,04295	0,06007	-0,4472	0,249	-3,0857	0,036
NTH	63001-65000	0,00316	16	0,03695	0,08552	0,1036	0,423	-5,4	0,006
NTH	64001-66000	0,00354	16	0,03365	0,10520	0,9123	0,046	-0,3048	0,394
NTH	65001-67000	0,00292	13	0,03668	0,07961	0,8449	0,078	1,9048	0,115
NTH	66001-68000	0,00315	13	0,03498	0,09005	0,9948	0,043	0,4476	0,431
NTH	67001-69000	0,00334	13	0,0375	0,08907	1,332	0,008	-1,7238	0,147
NTH	68001-69194	0,00317	7	0,05054	0,06272	1,9429	0,004	-2,0857	0,058
RAL	1-2000	0,00086	7	0,03384	0,02541	-0,7973	0,139	-0,6286	0,289
RAL	1001-3000	0,00056	4	0,0448	0,01250	-0,8232	0,149	0,1143	0,561
RAL	2001-4000	0,00316	13	0,05221	0,06052	0,8636	0,065	-1,3905	0,178
RAL	3001-5000	0,00504	23	0,04397	0,11462	1,0437	0,008	0,2857	0,469
RAL	4001-6000	0,00817	34	0,04494	0,18180	1,8137	0	3,0762	0,137
RAL	5001-7000	0,00823	33	0,05259	0,15649	2,1131	0	1,8667	0,233
RAL	6001-8000	0,00522	21	0,05879	0,08879	1,8666	0,001	-0,0667	0,424
RAL	7001-9000	0,00468	19	0,05218	0,08969	1,9044	0	-1,5238	0,198
RAL	8001-10000	0,00324	15	0,05735	0,05650	1,2661	0,009	-0,1905	0,437
RAL	9001-11000	0,00287	16	0,06258	0,04586	0,3524	0,247	-0,181	0,409
RAL	10001-12000	0,00413	20	0,05048	0,08181	0,6606	0,08	-2,1238	0,143
RAL	11001-13000	0,00247	12	0,05101	0,04842	0,5881	0,176	-0,6286	0,315
RAL	12001-14000	0,00164	11	0,05801	0,02827	-0,4038	0,241	-1,9143	0,09
RAL	13001-15000	0,00398	20	0,05148	0,07731	0,7747	0,046	-2,1238	0,14
RAL	14001-16000	0,00495	21	0,04514	0,10966	1,7696	0	0,9143	0,362
RAL	15001-17000	0,00578	27	0,04281	0,13502	1,1615	0,006	1,0857	0,391
RAL	16001-18000	0,005	25	0,04828	0,10356	0,8595	0,018	0,9143	0,378
RAL	17001-19000	0,00379	19	0,04763	0,07957	1,0933	0,011	0,2476	0,513
RAL	18001-20000	0,00456	24	0,04409	0,10342	0,8087	0,035	-1,1905	0,27
RAL	19001-21000	0,00513	31	0,05049	0,10160	0,1472	0,342	-4,6857	0,053
RAL	20001-22000	0,0041	26	0,04646	0,08825	-0,0974	0,396	-6,4	0,012
RAL	21001-23000	0,00239	19	0,04399	0,05433	-0,9278	0,024	-4,8286	0,022
RAL	22001-24000	0,00106	11	0,04639	0,02285	-1,554	0,003	-2,2476	0,085
RAL	23001-25000	0,00039	3	0,05489	0,00711	-0,5796	0,265	0,6286	0,104
RAL	24001-26000	0,00052	4	0,06605	0,00787	-0,6246	0,242	0,8381	0,077
RAL	25001-27000	0,00034	4	0,06527	0,00521	-1,5181	0,041	0,4571	0,409
RAL	26001-28000	0,00041	6	0,06229	0,00658	-1,9831	0	0,4952	0,42
RAL	27001-29000	0,00028	4	0,05329	0,00525	-1,8159	0,011	0,3714	0,497
RAL	28001-30000	0,00024	3	0,05743	0,00418	-1,685	0,038	0,3714	0,473
RAL	29001-31000	0,00038	4	0,05478	0,00694	-1,5181	0,043	0,581	0,36
RAL	30001-32000	0,0002	2	0,04876	0,00410	-1,0016	0,212	0,3333	0,457
RAL	31001-33000	0,00015	2	0,07506	0,00200	-1,4905	0,12	0,2476	0,584

RAL	32001-34000	0,00037	4	0,06273	0,00590	-1,5181	0,044	0,581	0,357
RAL	33001-35000	0,00052	5	0,07824	0,00665	-1,1587	0,062	0,8762	0,199
RAL	34001-36000	0,00046	4	0,11397	0,00404	-0,9225	0,124	0,7524	0,159
RAL	35001-37000	0,00013	1	0,07176	0,00181	-0,3988	0,496	0,2095	0,385
RAL	36001-38000	0,00031	3	0,04178	0,00742	-1,3165	0,097	0,4571	0,407
RAL	37001-39000	0,00053	4	0,03237	0,01637	-1,2203	0,078	0,6667	0,267
RAL	38001-40000	0,00015	1	0,05853	0,00256	-0,3988	0,494	0,2095	0,403
RAL	39001-41000	0	0	0,06451	0,00000	0	n.a	0	1
RAL	40001-42000	0,0002	2	0,04104	0,00487	-1,0016	0,23	0,3333	0,43
RAL	41001-43000	0,0002	2	0,03724	0,00537	-1,0016	0,213	0,3333	0,454
RAL	42001-44000	0,00027	2	0,04267	0,00633	-0,5127	0,343	0,2095	0,608
RAL	43001-45000	0,00045	3	0,06418	0,00701	-0,2726	0,365	-0,819	0,158
RAL	44001-46000	0,00032	3	0,06998	0,00457	-1,0095	0,094	-0,9048	0,136
RAL	45001-47000	0,00033	4	0,03994	0,00826	-1,5181	0,038	0,4571	0,45
RAL	46001-48000	0,00083	8	0,03729	0,02226	-1,5728	0,007	1,2476	0,173
RAL	47001-49000	0,00072	7	0,06507	0,01107	-1,6576	0,007	1,0381	0,241
RAL	48001-50000	0,00058	6	0,05568	0,01042	-1,5494	0,013	0,9143	0,226
RAL	49001-51000	0,00098	9	0,04848	0,02021	-1,3499	0,016	1,5429	0,087
RAL	50001-52000	0,00107	10	0,05454	0,01962	-1,331	0,018	-2	0,097
RAL	51001-53000	0,00175	15	0,05381	0,03252	-1,1076	0,011	-1,7143	0,158
RAL	52001-54000	0,00222	15	0,0502	0,04422	-0,3988	0,206	0,0667	0,549
RAL	53001-55000	0,00186	12	0,04488	0,04144	-0,2966	0,314	-1,6095	0,158
RAL	54001-56000	0,00118	9	0,04591	0,02570	-0,7822	0,117	-0,3048	0,361
RAL	55001-57000	0,00051	4	0,05128	0,00995	-0,6743	0,22	0,5429	0,349
RAL	56001-58000	0,00142	12	0,0538	0,02639	-1,0808	0,027	-4,9238	0,008
RAL	57001-59000	0,0024	16	0,04796	0,05004	-0,4718	0,189	-5	0,008
RAL	58001-60000	0,00284	12	0,04436	0,06402	1,131	0,033	-1,3429	0,191
RAL	59001-61000	0,00381	17	0,04927	0,07733	0,8466	0,045	-1,0762	0,257
RAL	60001-62000	0,00273	17	0,05359	0,05094	-0,2868	0,29	-3,1429	0,065
RAL	61001-63000	0,00211	12	0,05027	0,04197	-0,0956	0,43	-3,2286	0,032
RAL	62001-64000	0,00204	10	0,04095	0,04982	0,5077	0,198	-2,0286	0,09
RAL	63001-65000	0,00282	14	0,03878	0,07272	0,876	0,063	-4,2667	0,02
RAL	64001-66000	0,00383	17	0,03688	0,10385	1,4206	0,002	-5,2	0,013
RAL	65001-67000	0,00301	14	0,03668	0,08206	1,0163	0,029	-3,2381	0,036
RAL	66001-68000	0,00245	12	0,0335	0,07313	0,7892	0,085	-1,2	0,201
RAL	67001-69000	0,00268	12	0,0353	0,07592	1,2918	0,011	-1,3143	0,189
RAL	68001-69194	0,00353	9	0,04752	0,07428	1,6691	0,001	-0,1429	0,408
RG	1-2000	0,00784	68	0,03055	0,25663	-1,4247	0	3,4095	0,217
RG	1001-3000	0,00908	65	0,04413	0,20576	-1,0029	0,001	-4,0571	0,177
RG	2001-4000	0,00995	63	0,05149	0,19324	-0,6361	0,028	-5,9143	0,073

RG	3001-5000	0,00809	56	0,03934	0,20564	-1,0452	0,002	1,219	0,395
RG	4001-6000	0,00945	66	0,04144	0,22804	-1,0417	0,002	-1,3143	0,338
RG	5001-7000	0,00887	69	0,05445	0,16290	-1,1925	0	-2,5619	0,268
RG	6001-8000	0,00678	52	0,05794	0,11702	-1,0917	0,001	1,2476	0,381
RG	7001-9000	0,00549	49	0,04921	0,11156	-1,4311	0	-0,3143	0,429
RG	8001-10000	0,00596	51	0,05554	0,10731	-1,3052	0	-1,7333	0,289
RG	9001-11000	0,00733	56	0,06472	0,11326	-1,1528	0	-13,0571	0,001
RG	10001-12000	0,00971	73	0,05274	0,18411	-1,1844	0	-7,1143	0,058
RG	11001-13000	0,00856	62	0,04618	0,18536	-1,02	0,002	3,0095	0,235
RG	12001-14000	0,00631	50	0,05269	0,11976	-1,2682	0	-1,4762	0,321
RG	13001-15000	0,00748	55	0,05142	0,14547	-1,007	0,002	-5,9238	0,063
RG	14001-16000	0,00635	44	0,04355	0,14581	-1,1365	0	-5,1905	0,069
RG	15001-17000	0,00671	44	0,03585	0,18717	-1,0461	0,005	0,9714	0,426
RG	16001-18000	0,00806	55	0,04394	0,18343	-0,7589	0,009	0,2476	0,499
RG	17001-19000	0,00745	64	0,04463	0,16693	-1,3535	0	2,5714	0,256
RG	18001-20000	0,00769	66	0,04001	0,19220	-1,331	0	3,5048	0,192
RG	19001-21000	0,00667	52	0,04437	0,15033	-1,0454	0,001	2,0952	0,281
RG	20001-22000	0,00583	46	0,04151	0,14045	-1,1051	0	0,9048	0,418
RG	21001-23000	0,00624	48	0,04096	0,15234	-1,0235	0,003	1,8381	0,327
RG	22001-24000	0,0054	45	0,04625	0,11676	-1,243	0	-0,6381	0,375
RG	23001-25000	0,0055	40	0,04859	0,11319	-1,1159	0	-3,7333	0,115
RG	24001-26000	0,00676	44	0,05739	0,11779	-0,847	0,01	1,1238	0,422
RG	25001-27000	0,00956	70	0,06305	0,15163	-0,7936	0,005	3,0286	0,236
RG	26001-28000	0,0091	66	0,05696	0,15976	-0,7157	0,012	2,4	0,292
RG	27001-29000	0,00457	34	0,04657	0,09813	-1,0647	0,001	1,1048	0,386
RG	28001-30000	0,00435	36	0,04816	0,09032	-1,7048	0	-0,7048	0,369
RG	29001-31000	0,00476	42	0,04321	0,11016	-1,7858	0	-0,3714	0,433
RG	30001-32000	0,00646	46	0,0429	0,15058	-0,8653	0,005	-3,581	0,137
RG	31001-33000	0,01045	56	0,0655	0,15954	-0,3402	0,163	-1,8381	0,29
RG	32001-34000	0,00874	57	0,05301	0,16487	-0,8238	0,009	2,9429	0,228
RG	33001-35000	0,00953	70	0,0748	0,12741	-0,7378	0,011	-3,0571	0,227
RG	34001-36000	0,01353	90	0,10345	0,13079	-0,8538	0,005	-11,4095	0,026
RG	35001-37000	0,01004	74	0,06138	0,16357	-1,1183	0,001	-4,1048	0,175
RG	36001-38000	0,00513	38	0,03914	0,13107	-1,3157	0,003	-0,0286	0,476
RG	37001-39000	0,004	35	0,03104	0,12887	-2,0053	0	-0,7333	0,375
RG	38001-40000	0,00523	36	0,0535	0,09776	-1,332	0	4,5524	0,023
RG	39001-41000	0,00562	31	0,05982	0,09395	-0,5098	0,133	1,4667	0,335
RG	40001-42000	0,00595	46	0,04024	0,14786	-1,1831	0,001	1,3714	0,365
RG	41001-43000	0,00707	54	0,03682	0,19202	-1,1134	0,001	1,1048	0,41
RG	42001-44000	0,00744	54	0,04169	0,17846	-0,9896	0,002	-0,7429	0,392

RG	43001-45000	0,01133	68	0,0624	0,18157	-0,1343	0,324	0,1619	0,495
RG	44001-46000	0,01158	68	0,07127	0,16248	0,1032	0,381	-5,1238	0,108
RG	45001-47000	0,00749	53	0,04373	0,17128	-0,7498	0,013	-7,5048	0,028
RG	46001-48000	0,00837	55	0,03277	0,25542	-0,8489	0,008	-0,5048	0,411
RG	47001-49000	0,00921	55	0,05944	0,15495	-0,6057	0,037	1,6286	0,348
RG	48001-50000	0,00675	45	0,05494	0,12286	-0,7855	0,023	1,3238	0,34
RG	49001-51000	0,00864	64	0,04153	0,20804	-1,2254	0	-4,5333	0,144
RG	50001-52000	0,00955	67	0,04756	0,20080	-1,1557	0,001	-7,8095	0,05
RG	51001-53000	0,00798	51	0,05129	0,15559	-0,6616	0,03	-1,1048	0,331
RG	52001-54000	0,00921	58	0,05177	0,17790	-0,478	0,08	-5,3714	0,084
RG	53001-55000	0,00767	52	0,04564	0,16805	-0,8272	0,008	-4,7048	0,114
RG	54001-56000	0,00813	55	0,04509	0,18031	-0,6592	0,032	-0,1905	0,427
RG	55001-57000	0,00872	61	0,04745	0,18377	-0,7431	0,012	-4,8476	0,128
RG	56001-58000	0,00909	61	0,04883	0,18616	-0,7211	0,015	-5,1714	0,1
RG	57001-59000	0,0087	60	0,04742	0,18347	-0,8036	0,004	-1,6762	0,301
RG	58001-60000	0,00602	44	0,04617	0,13039	-1,3054	0	-0,4095	0,405
RG	59001-61000	0,00624	48	0,05054	0,12347	-1,5921	0	-3,2095	0,167
RG	60001-62000	0,0074	63	0,05028	0,14718	-1,6679	0	-3,4857	0,214
RG	61001-63000	0,00861	58	0,04535	0,18986	-1,1156	0,001	-0,5048	0,412
RG	62001-64000	0,00733	45	0,04136	0,17722	-0,8121	0,011	-0,0095	0,48
RG	63001-65000	0,00649	43	0,03831	0,16941	-0,7016	0,038	0,5333	0,444
RG	64001-66000	0,00709	49	0,03451	0,20545	-0,803	0,007	5,0571	0,039
RG	65001-67000	0,00689	56	0,03695	0,18647	-1,3965	0	7,9619	0,002
RG	66001-68000	0,00735	51	0,03301	0,22266	-0,9049	0,006	-0,1524	0,466
RG	67001-69000	0,00739	42	0,03356	0,22020	0,0338	0,435	-6,8857	0,024
RG	68001-69194	0,00874	36	0,04577	0,19095	-0,5208	0,114	-3,3048	0,124
ZI	1-2000	0,00683	49	0,02785	0,24524	-0,8818	0,003	0,1238	0,542
ZI	1001-3000	0,01039	69	0,04375	0,23749	-0,8042	0,004	-0,6286	0,421
ZI	2001-4000	0,01132	77	0,05134	0,22049	-0,8809	0,005	-3,4476	0,219
ZI	3001-5000	0,00775	56	0,03747	0,20683	-1,1073	0,001	1,0476	0,419
ZI	4001-6000	0,00978	70	0,04116	0,23761	-0,9398	0,004	2,9333	0,29
ZI	5001-7000	0,00921	68	0,05317	0,17322	-0,8963	0,001	2,1905	0,297
ZI	6001-8000	0,00767	60	0,05602	0,13692	-1,2059	0,001	0,1238	0,524
ZI	7001-9000	0,00793	62	0,04889	0,16220	-1,1175	0	0,4857	0,46
ZI	8001-10000	0,0088	68	0,05529	0,15916	-1,013	0	4,5333	0,14
ZI	9001-11000	0,00823	59	0,05993	0,13733	-1,0252	0	1,7429	0,322
ZI	10001-12000	0,00984	63	0,05511	0,17855	-0,6105	0,037	5,5619	0,072
ZI	11001-13000	0,00992	70	0,05163	0,19214	-0,7281	0,007	7,2476	0,026
ZI	12001-14000	0,00946	76	0,0521	0,18157	-1,2558	0	2,381	0,298
ZI	13001-15000	0,01131	86	0,05083	0,22251	-1,141	0	-1	0,395

ZI	14001-16000	0,00755	57	0,04429	0,17047	-0,9717	0,001	0,8286	0,445
ZI	15001-17000	0,00589	43	0,03762	0,15657	-0,9452	0,004	2,1429	0,256
ZI	16001-18000	0,0084	65	0,04362	0,19257	-1,1808	0	2,2	0,302
ZI	17001-19000	0,00767	63	0,04621	0,16598	-1,2671	0	4,3905	0,134
ZI	18001-20000	0,00889	68	0,04102	0,21672	-1,017	0	3,6571	0,183
ZI	19001-21000	0,00937	72	0,04143	0,22616	-0,9697	0	5,2476	0,098
ZI	20001-22000	0,00681	54	0,03861	0,17638	-1,0961	0	4,8571	0,05
ZI	21001-23000	0,00606	54	0,04175	0,14515	-1,455	0	2,3524	0,28
ZI	22001-24000	0,0065	52	0,0483	0,13458	-1,0891	0	2,419	0,258
ZI	23001-25000	0,00758	56	0,05477	0,13840	-0,897	0,003	-0,0571	0,449
ZI	24001-26000	0,00778	54	0,05966	0,13041	-0,8609	0,008	-0,8	0,379
ZI	25001-27000	0,01097	75	0,0616	0,17808	-0,6654	0,024	0,2095	0,502
ZI	26001-28000	0,01076	80	0,05521	0,19489	-0,8164	0,007	1,1524	0,408
ZI	27001-29000	0,00573	44	0,04409	0,12996	-1,1365	0,001	2,3238	0,245
ZI	28001-30000	0,00573	39	0,05404	0,10603	-1,1152	0,003	-1,6095	0,248
ZI	29001-31000	0,00572	40	0,05092	0,11233	-1,1291	0,001	-1,8	0,23
ZI	30001-32000	0,00849	53	0,0446	0,19036	-0,4398	0,122	2,3429	0,265
ZI	31001-33000	0,01125	65	0,06657	0,16900	-0,6162	0,035	0,9524	0,429
ZI	32001-34000	0,00722	50	0,05428	0,13301	-1,0122	0,002	0,4381	0,474
ZI	33001-35000	0,00803	50	0,07175	0,11192	-0,2389	0,249	0,6762	0,458
ZI	34001-36000	0,01097	64	0,10074	0,10889	-0,4318	0,096	-2,8762	0,181
ZI	35001-37000	0,00932	62	0,06159	0,15132	-0,7775	0,006	0,7714	0,471
ZI	36001-38000	0,00625	40	0,03912	0,15976	-0,7854	0,021	0,581	0,444
ZI	37001-39000	0,00406	24	0,02929	0,13861	-0,691	0,051	-3,8857	0,05
ZI	38001-40000	0,00443	32	0,05665	0,07820	-0,8883	0,018	-3,5905	0,086
ZI	39001-41000	0,0061	40	0,06628	0,09203	-0,7226	0,022	-2,4571	0,219
ZI	40001-42000	0,00626	38	0,03906	0,16027	-0,4238	0,13	0,7524	0,411
ZI	41001-43000	0,0071	46	0,0363	0,19559	-0,322	0,19	2,4952	0,245
ZI	42001-44000	0,00677	46	0,03919	0,17275	-0,7728	0,016	1,4762	0,35
ZI	43001-45000	0,01001	62	0,0584	0,17140	-0,4809	0,07	3,3619	0,209
ZI	44001-46000	0,01307	83	0,06751	0,19360	-0,2741	0,207	0,2857	0,528
ZI	45001-47000	0,00896	62	0,04197	0,21349	-0,5762	0,036	-1,9619	0,282
ZI	46001-48000	0,00844	59	0,03808	0,22164	-0,8548	0,009	-0,8952	0,384
ZI	47001-49000	0,00988	66	0,06285	0,15720	-0,8115	0,009	-4,6952	0,132
ZI	48001-50000	0,00649	48	0,05347	0,12138	-1,0873	0	-1,5524	0,299
ZI	49001-51000	0,00844	58	0,04092	0,20626	-0,9054	0,002	1,7048	0,338
ZI	50001-52000	0,01128	69	0,04822	0,23393	-0,5915	0,041	0,5905	0,481
ZI	51001-53000	0,00938	61	0,04893	0,19170	-0,7145	0,021	1,0476	0,424
ZI	52001-54000	0,00981	64	0,04675	0,20984	-0,5578	0,049	-2,3524	0,248
ZI	53001-55000	0,00817	51	0,03931	0,20784	-0,4549	0,103	-4,5048	0,103

ZI	54001-56000	0,00737	47	0,04096	0,17993	-0,4832	0,103	-3,1048	0,182
ZI	55001-57000	0,00965	61	0,04875	0,19795	-0,4241	0,121	-3,0571	0,214
ZI	56001-58000	0,00952	63	0,04978	0,19124	-0,5935	0,03	-1,1619	0,374
ZI	57001-59000	0,00816	61	0,04597	0,17751	-1,084	0	2,419	0,262
ZI	58001-60000	0,00884	63	0,04539	0,19476	-1,2714	0	-0,6	0,391
ZI	59001-61000	0,00888	66	0,05057	0,17560	-1,3596	0	-6,3905	0,065
ZI	60001-62000	0,00788	67	0,04827	0,16325	-1,5412	0	-3,7429	0,175
ZI	61001-63000	0,00819	59	0,04506	0,18176	-1,3114	0	-3,5524	0,196
ZI	62001-64000	0,00672	48	0,04082	0,16463	-1,3037	0	-5,8476	0,045
ZI	63001-65000	0,00709	58	0,03714	0,19090	-1,3327	0	2,3048	0,284
ZI	64001-66000	0,00743	59	0,03338	0,22259	-1,2387	0	8,7714	0,001
ZI	65001-67000	0,00627	47	0,03574	0,17543	-1,1258	0	6,7429	0,007
ZI	66001-68000	0,00676	43	0,032	0,21125	-0,5752	0,071	6,0286	0,009
ZI	67001-69000	0,00846	58	0,03346	0,25284	-0,7067	0,014	1,6952	0,362
ZI	68001-69194	0,01061	53	0,04711	0,22522	-1,1657	0	-4,3238	0,126