

Análisis comparativo de genes productores de piRNA en mamíferos

Autor: José Carlos Montañés Domínguez

Máster de bioinformática y estadística

Área 2

Nombre Consultor/a: Tanya Vavouri

Nombre Profesor/a responsable de la asignatura: Carles Ventura Royo

05/06/2019



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

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Análisis comparativo de genes productores de piRNA en mamíferos.</i>
Nombre del autor:	<i>José Carlos Montañés Domínguez</i>
Nombre del consultor/a:	<i>Tanya Vavouri</i>
Nombre del PRA:	<i>Carles Ventura Royo</i>
Fecha de entrega (mm/aaaa):	06/2019
Titulación::	<i>Máster de Bionformatica y bioestadística</i>
Área del Trabajo Final:	<i>Bioinformática y Bioestadística Area 2</i>
Idioma del trabajo:	<i>Español</i>
Palabras clave	<i>PIWI-interacting RNAs, comparative genomics, Data mining</i>
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i></p>	
<p>Los piRNAs son secuencias de RNA de 21-35 nucleótidos. Estas secuencias provienen de locus específicos conocidos con piRNA clústers (piCs). Los piRNAs son las encargadas de guiar a las proteínas de la familia PIWI a reprimir la expresión de los transposones (represión transcripcional) o bien de eliminar las secuencias nacientes de estos (represión post-transcripcional). Pese a que se ha comprobado la conservación de las proteínas de la familia PIWI, se sabe muy poco sobre la conservación tanto de los piRNAs como de los piCs. En este trabajo hemos estudiado la conservación de los piRNA producidos en genes codificantes y el contenido de regiones repetitivas en una serie de piC predichos en 10 mamíferos distintos. En primer lugar, hemos encontrado que la mayoría de piRNAs génicos se producen en intrones y exones. Seguidamente analizamos la conservación de los genes que producen piRNAs. Hemos obtenido que, a pesar de que la mayoría de los genes se conservan entre los distintos clades observados sólo una pequeña fracción de genes productores de piRNA se encuentran conservados. Finalmente hemos analizado como la densidad de repeticiones LTR en los piCs es mayor respecto a los genes y a la región solapada entre piCs y genes ortólogos productores de piRNA. Como conclusión del trabajo obtenemos que existen un número limitado de genes productores de piRNA que conserven esta función entre los diversos mamíferos estudiados. Sin embargo, estos genes conservados mantienen características similares entre ellos sugiriendo un mecanismo de creación distinto al resto de piCs.</p>	

Abstract (in English, 250 words or less):

PIWI-interacting RNAs (piRNAs) are 21-35 nucleotide long noncoding RNAs. This type of small RNAs are created in specific loci called piRNA clusters (piCs). The function of piRNAs is to repress the propagation of transposons over the genome. To defend the cell against transposons piRNAs guide PIWI proteins to transposons causing their repression (transcriptional repression) or cleaving the transcripts (post-transcriptional repression). Although PIWI proteins are conserved among different species piRNAs are not. In this project we studied the conservation of piRNA produced in protein-coding genes and their repeat content in a set of previously predicted piRNA clusters from 10 different mammals. First, we found that in all mammals, protein-coding genes produce piRNAs from both exons and introns. Second, we analyzed the conservation of those genes that generate piRNAs. Although most mammalian genes are conserved, only a small fraction produce piRNAs in more than one species. Nevertheless, we found a small number of deeply conserved piRNA producing protein coding genes. Last, we analysed repeats in piCs, genic piCs and protein-coding genes. We found that there is higher density of LTR repeats in piC when compared to other protein-coding genes and the intersection of piCs and orthologues piRNA generating genes. We conclude that production of piRNAs from protein-coding genes evolves fast, with notable exceptions, and that the difference in repeat content between genic and intergenic piCs suggests different functions or modes of evolution.

Índice

Introducción.....	1
1.1 Contexto y justificación del Trabajo.....	1
piRNAs, pequeñas moléculas con gran importancia.....	1
Biogénesis y función.....	1
Elementos transponibles.....	3
Evolución de los piRNAs y piRNA clústers.....	3
1.2 Objetivos del Trabajo.....	4
1.3 Enfoque y método seguido.....	5
1.4 Planificación del Trabajo.....	7
1.5 Breve resumen de productos obtenidos.....	8
1.6 Breve descripción de los otros capítulos de la memoria.....	8
Materiales y métodos.....	9
Identificación y anotación de los piC génicos.....	9
Identificación de genes ortólogos.....	10
Identificación de regiones repetidas.....	11
Resultados y discusión.....	11
Clasificación de los piCs.....	11
Genes ortólogos productores de piRNAs.....	13
Localización de los piCs sobre los genes.....	16
Regiones repetitivas en genes productores de piRNAs.....	16
3. Conclusiones.....	20
4. Glosario.....	21
5. Bibliografía.....	22
6. Anexos.....	24

Lista de figuras

Figura 1	2
Figura 2	6
Figura 3	8
Figura 4	10
Figura 5	12
Figura 6	14
Figura 7	14
Figura 8	15
Figura 9	15
Figura 10	17
Figura 11	18
Figura 12	19
Tabla 1.....	8

Introducción

1.1 Contexto y justificación del Trabajo

piRNAs, pequeñas moléculas con gran importancia

Los piRNAs primero fueron definidos en *Drosophila* relacionando su función con el silenciamiento del gen *Stellate*. Cuando *Stellate* no está regulado correctamente tiende a formar agregados proteicos produciendo infertilidad en los testículos. Aravin et al., descubrieron que *Stellate* era reprimido por los transcritos producidos por el locus de *Suppressor of Stellate*¹. Estos transcritos serían conocidos posteriormente como *PIWI-interacting RNAs* (piRNAs)². Los piRNAs son secuencias nucleotídicas no codificantes de entre 21 y 35 nucleótidos. Este tipo de RNA se distinguen de otros RNAs no codificantes por una fuerte tendencia a tener un uracilo en el extremo 5' así como por la presencia de una modificación 2'O-metil en su extremo 3'. Todos los RNAs pequeños no codificantes actúan mediante la unión con la familia de proteínas Argonauta. Los piRNAs se unen específicamente a una sub-familia de las proteínas Argonauta llamadas PIWI. Las proteínas PIWI son mayoritariamente expresados en las células de las gónadas. Cabe mencionar que a pesar de que la mayoría de las proteínas relacionadas con los piRNAs están conservadas las secuencias de los piRNA no lo están³.

Biogénesis y función

La biogénesis de los piRNA está relacionada con su función. La mayoría de los piRNA son transcritos en regiones cromosómicas conocidas como piRNA clústers (piCs). Los piCs son regiones cromosómicas que contienen transposones en su secuencia y que, además generan piRNAs a partir de largas secuencias no codificantes de RNA. Los piRNAs son producidos por la escisión de transcritos largos monocatenario de RNA no codificante producidos en los piCs⁴. Cuando estos transcritos son producidos se exportan al citoplasma. En el citoplasma las secuencias largas precursoras de los piRNA serán escindidas y utilizadas para bloquear sus secuencias objetivo o entrarán en la vía de amplificación ping-pong.

Drosophila es uno de los organismos donde más se ha estudiado la función de los piRNAs. En *Drosophila*, la vía ping-pong de amplificación empieza con la proteína *Aubergine* (Aub) uniéndose a una secuencia de piRNA. Una vez unido el piRNA a Aub la proteína será guiada a la secuencia complementaria al piRNA el cual será reconocido por Aub produciendo la escisión de la secuencia diana. El fragmento de secuencia obtenido después del corte de Aub será recogido por la proteína Ago3. La secuencia de pre-piRNA será recortado por una 3' exonucleasa llamada Zucchini. Posteriormente la secuencia aún inmadura de piRNA será metilada por el extremo 3'. Esta secuencia metilada será la forma ya madura del piRNA que será utilizada para la obtención de otra secuencia de piRNA que es idéntica a la primera secuencia producida por Aub³. El esquema de la vía ping-pong se encuentra en la figura 1.

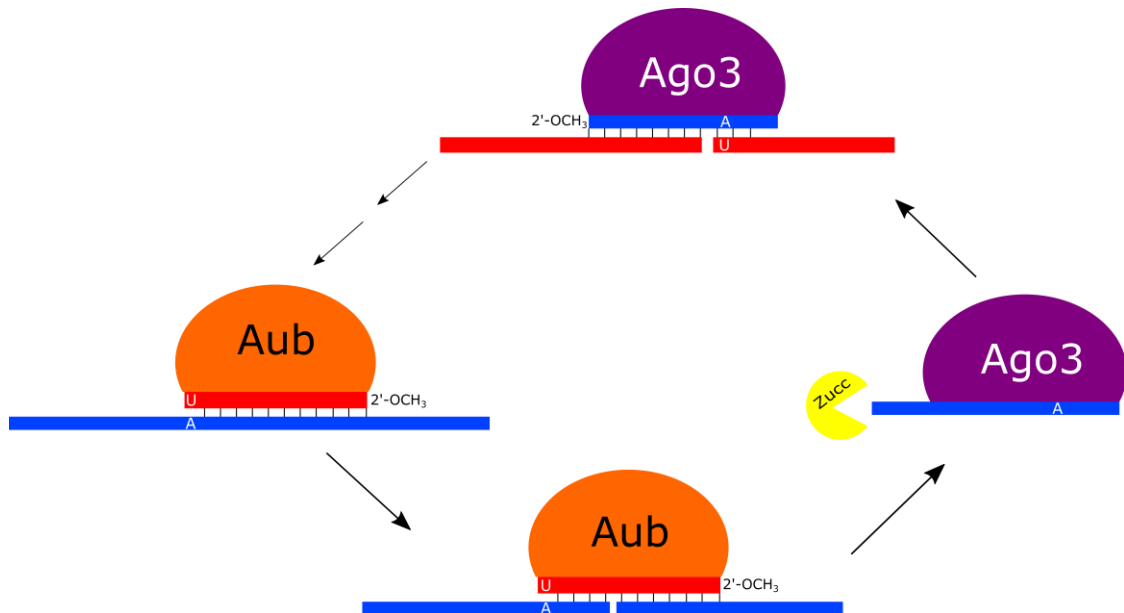


Figura 1. Esquema de la vía ping-pong en *Drosophila*. La vía comienza con la unión de una secuencia de piRNA a la proteína *Aubergine* (Aub). Después de la unión de Aub con el piRNA se procederá al reconocimiento y unión con la secuencia complementaria al piRNA. Al producirse la unión Aub cortará la secuencia objetivo 10 nucleótidos desde el inicio de la unión de las secuencias. Esta nueva secuencia será recogida por Ago3 y recortada por Zucchini para volver a ser utilizada para reconocer la secuencia inicial volviendo a iniciar el ciclo.

En ratón existen tres proteínas de la familia PIWI: MILI, MIWI y MIWI2. MILI es la única proteína de la familia PIWI que se expresa en oocitos. En la línea germinal masculina las tres proteínas son expresadas, pero en distintos estadios del desarrollo⁵. Los piRNAs expresados durante el estadio pre-paquiteno de la espermatogénesis son producidos principalmente a partir de mRNA codificantes y muchos son complementarios a secuencias de elementos transponibles (TEs). MILI y MIWI2 son las dos proteínas expresadas durante este estadio. Por otro lado, existen los piRNAs del estadio paquiteno. Estos piRNAs no son complementarios a TEs y se unen a las proteínas MILI y MIWI⁵. Los piRNAs del estadio paquiteno son los mayoritarios en los testículos de ratón y son producidos a partir de largas secuencias de RNA no codificante. El factor de transcripción A-MYB es el encargado de coordinar la transcripción de los piRNAs paquitenos así como las proteínas relacionadas con estas secuencias como MIWI y MILI. Algunos de los aspectos sobre el mecanismo de producción de piRNAs a día de hoy todavía son desconocidos como porque algunos mRNAs producen piRNAs y porque el factor de transcripción A-MYB puede regular al mismo tiempo la transcripción de mRNAs y piRNAs³.

La principal función descrita para los piRNAs es silenciar transposones. Esta descrito que los piRNAs actúan como elemento guía para las proteínas de la familia PIWI para poder reprimir los transposones. Cuando se muta proteínas de la familia PIWI se ha comprobado un aumento en la expresión de los transposones⁶. El complejo piRNA-PIWI es capaz de reprimir tanto transcripcional como post-transcripcionalmente sus secuencias objetivo. A nivel transcripcional de inhibición las proteínas PIWI son capaces de unirse al transcrito objetivo naciente y promover la metilación H3K9 haciendo que se

condense la región cromosómica y evitando la transcripción de la secuencia objetivo. A nivel post-transcripcional se ha comprobado en *Drosophila* que PIWI tiene actividad hidrolasa capaz de escindir sus secuencias objetivo³.

Elementos transponibles

Los elementos transponibles (TEs) son inserciones de DNA en los cromosomas que se pueden expandir dentro del genoma del huésped infectando otras regiones genómicas. Los TEs son clasificados en función de los mecanismos que utilizan para propagarse por el genoma del huésped en transposones de DNA o retrotransposones⁷. Los TEs de DNA se insertan en los distintos locos mediante un proceso de corte y empalme. Las secuencias de DNA son extraídas de un locus genómico y posteriormente se insertan en otra región. Para hacer esta función los DNA transposones codifican proteínas transposasas encargadas de realizar la función de corte y empalme. El otro tipo de transposones son los retrotransposones. Los retrotransposones se expanden por el genoma por un proceso de copia y pega. Los retrotransposones son transcritos y mediante una transcriptasa reversa que codifica de RNA a DNA se insertarán en otro lugar del genoma. Los retrotransposones son subdivididos según su estructura y la forma en que se replican en: repeticiones terminales largas (LTRs), elementos largos intercalados (LINEs) y elementos cortos intercalados (SINEs). Los LTR codifican la enzima transcriptasa reversa y están flanqueados por secuencias repetitivas en ambos extremos. Los LINEs son transposones autónomos que codifican transcriptasa reversa y endonucleasas y los SINEs son transposones no autónomos que necesitan las enzimas producidos por los LINEs para poder expandirse por el genoma⁸.

Los TEs pueden producir muchos efectos perjudiciales en su huésped. Esto es debido a que los transposones pueden insertarse en regiones específicas dentro de genes o promotores provocando mutaciones en proteínas o represión de genes. Sin embargo, dependiendo del lugar de inserción de los transposones no tiene por qué tener un efecto perjudicial para el huésped. También se ha comprobado que las proteínas codificadas por los TEs como las transposasas pueden producir daños en la célula huésped⁷.

Evolución de los piRNAs y piRNA clústers

Los piRNAs y los TEs están constantemente evolucionando. La supervivencia de los transposones consiste en evitar el sistema de defensa del huésped basado en piRNAs. Por otro lado, los piRNAs deben evolucionar para detener la propagación de los transposones por el genoma del huésped. Este proceso lleva a una carrera armamentística entre los piRNAs y los transposones para intentar sobrepasar al rival. Esta carrera armamentística es una de las hipótesis actuales para entender el porqué de la redundancia de secuencias de piRNA en los genomas de distintos organismos⁸. Esta redundancia es positiva para los piRNAs ya que en caso de que se inhiba su expresión, por ejemplo, por mutación de su región promotora, se seguirían teniendo copias de esas secuencias expresadas en otro lugar del genoma. Actualmente se cree que esta rápida evolución de los piRNAs para proteger el genoma de la célula

huésped es lo que provoca que las secuencias no se encuentren conservadas incluso en especies evolutivamente cercanas³. Sin embargo, recientemente se ha observado que existen piCs en regiones sinténicas en especies evolutivamente cercanas como los mamíferos⁹.

Actualmente el modelo más aceptado para la generación de nuevos piCs es el modelo de trampa. Cuando un TE invade un nuevo genoma se replica hasta insertarse en un piC. La inserción del transposón en el piC produce un incremento del tamaño del clúster así como nuevos piRNAs complementarios a la secuencia insertada que serán capaces de bloquear el movimiento del TE en cualquier ubicación del genoma^{7,10}.

Los piRNA clústers son clasificados en función de sus posiciones en el genoma en génicas e intergénicas. Los clústers intergénicos normalmente son más grandes que los génicos (de 10 a 100 kb) y algunos de ellos tienen secuencias complementarias a los transposones. Los clústers génicos son localizados normalmente en las regiones 3'UTR de genes codificantes para proteínas. Un ejemplo en *Drosophila* de piC génico es *Traffic jam* (tj) el cual se ha visto que está relacionado con el desarrollo celular¹¹. Sin embargo todavía se desconoce mucho sobre todas las funciones y mecanismos de producción de los piCs.

A pesar de que se piensa que la principal función de los piCs es silenciar transposones se ha comprobado que muchos apenas tienen secuencias de transposones en su secuencia⁹. Porque muchos piCs en mamíferos no son complementarios a secuencias de transposones y como genes que codifican proteínas son redirigidos a la vía PIWI para crear piRNAs son preguntas aún en el aire. Además, la amplia mayoría de estudios en mamíferos sobre los piRNAs se han realiza en ratón. Por ello el objetivo de este trabajo es entender mejor la evolución y la biogénesis de piCs génicos y de los genes que contienen los piCs (genes productores de piRNAs).

1.2 Objetivos del Trabajo

Los tres objetivos generales de este trabajo son:

- A. Entender qué tipo de genes en mamíferos producen piRNAs y la localización de los piCs en estos genes.
 - A1. Identificar aquellos genes que se superponen con los piCs en 10 mamíferos distintos.
 - A2. Describir la ubicación de la superposición de los piCs dentro de los genes que codifican proteínas.
- B. Probar que la producción de piRNAs de genes ortólogos está conservada entre los mamíferos estudiados.
 - B1. Identificar los genes productores de piRNA, así como aquellos que sean ortólogos entre las distintas especies.

B2. Comprobar que la probabilidad de producción de piRNA en genes productores de piRNA ortólogos es mayor que la esperada por azar.

C. Comprobar la existencia de una asociación entre transposones y las regiones solapantes entre piCs y genes codificantes.

C1. Describir los transposones asociados a los distintos genes y piCs, así como aquellas regiones solapantes entre ambas partes.

C2. Comprobar si existe alguna asociación entre algún tipo específico de transposón y los piCs génicos.

1.3 Enfoque y método seguido

Para conseguir los objetivos anteriormente mencionados hemos utilizado los siguientes programas:

R¹²: Programa estadístico que puede ser utilizado para manipular las anotaciones genómicas (con el paquete GRanges¹³) y descargar las anotaciones genómicas de Ensembl¹⁴ (con el paquete biomaRT¹⁵).

Ensembl Compara Perl API¹⁴: Interfaz de programación de aplicaciones que hemos utilizado para descargar los identificadores de familias génicas (Gene Trees IDs, GTIDs). Los GTIDs se utilizan para identificar genes que son ortólogos en distintas especies. Hemos utilizado este identificador porque Ensembl tiene uno de los protocolos más robustos actualmente para la identificación de genes ortólogos¹⁶.

Repeat masker¹⁷: Programa de detección de regiones repetitivas. Repeat masker es el programa con el que se elaboraron las tablas de regiones repetitivas que descargamos de UCSC¹⁸.

La estrategia que se ha utilizado para la comparación de los diversos genes productores de piRNAs se encuentra en la figura 2.

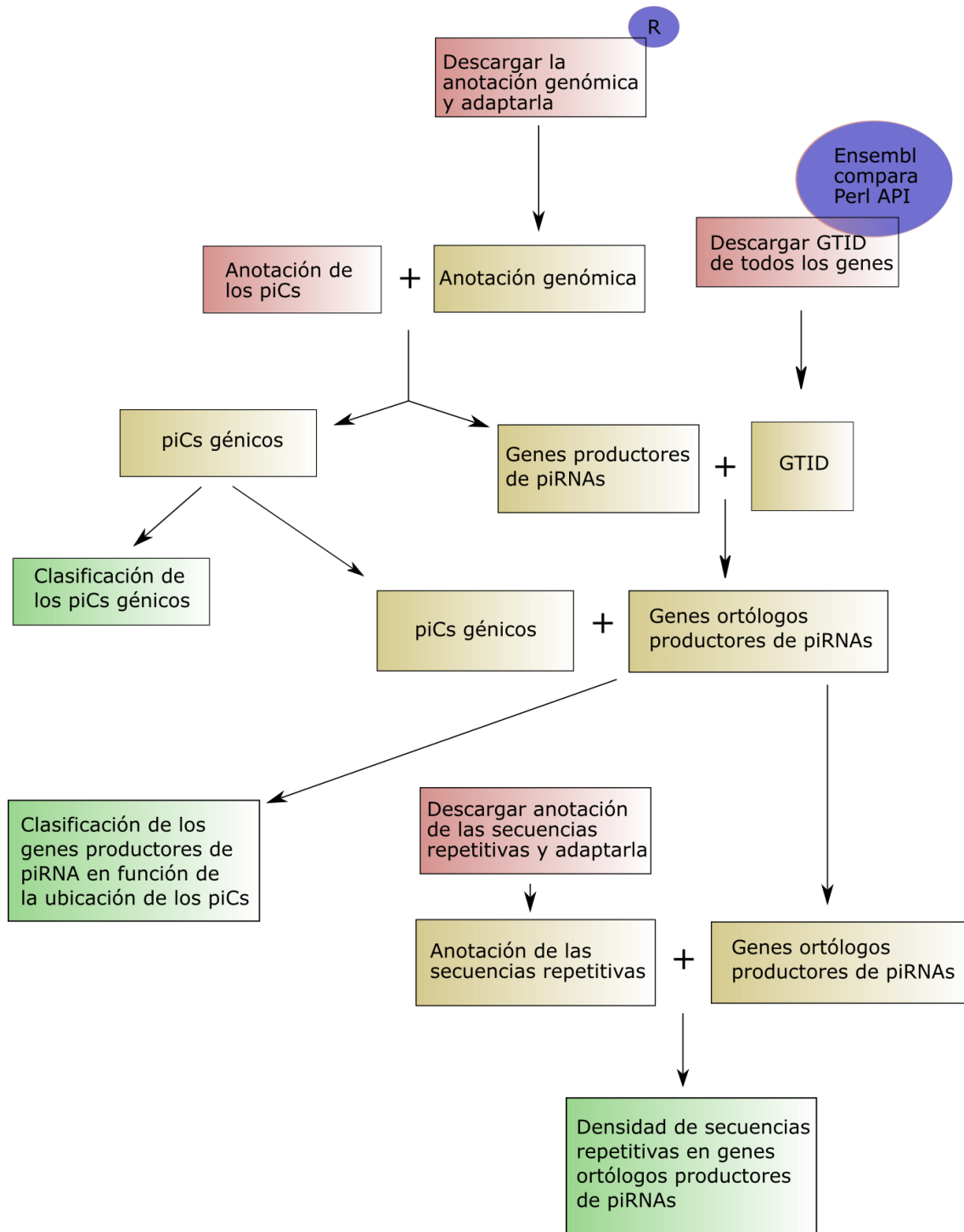


Figura 2. Esquema de la hoja de ruta utilizada para la realización de este trabajo. En rojo se indican inputs descargados o proporcionados por la tutora. En el caso que se haya necesitado software adicional aparte del navegador para descargar algún tipo de archivo se ha indicado con un círculo morado con el nombre del programa utilizado. Los cuadrados amarillos son inputs de un proceso y outputs de otro. Los cuadrados verdes son outputs definitivos del cual se han obtenido las figuras que se muestran en trabajo. El software utilizado para el análisis de los datos ha sido R.

1.4 Planificación del Trabajo

Para la realización de este trabajo se han utilizado programas que se han descrito en el apartado anterior. Las tareas que se han realizado se muestran a continuación de forma secuencial:

1. Descargar la anotación genética de los mamíferos utilizados en este trabajo.
2. Adaptar las coordenadas genéticas para que sean comparables con la anotación de los piCs.
3. Superponer las anotaciones genómicas de los genes y los piCs.
4. Calcular, para cada especie, la proporción de piCs que se superponen en ubicaciones génicas.
5. Analizar las regiones superpuestas de los piCs génicos sobre los genes, es decir, conocer la ubicación de los piCs sobre los genes.
6. Comparar las ubicaciones de los piCs entre los distintos animales estudiados.
7. Descargar las GTIDs de todos los genes de los diferentes mamíferos estudiados.
8. Encontrar genes ortólogos entre los diferentes mamíferos que produzcan piRNAs.
9. Comprobar que la frecuencia de encontrar genes ortólogos que produzcan piRNAs es mayor que la esperada por azar.
10. Descargar la anotación de las regiones repetidas en el genoma de los diversos mamíferos. Para ello utilizaremos las tablas de anotación de la UCSC¹⁸.
11. Comprobar si hay algún tipo de repetición asociada con piCs génicos.
12. Comprobar si algún tipo de repetición conservada entre los distintos piCs génicos de los animales estudiados.

La planificación se ha esquematizado en un diagrama de Gantt en la figura 3.

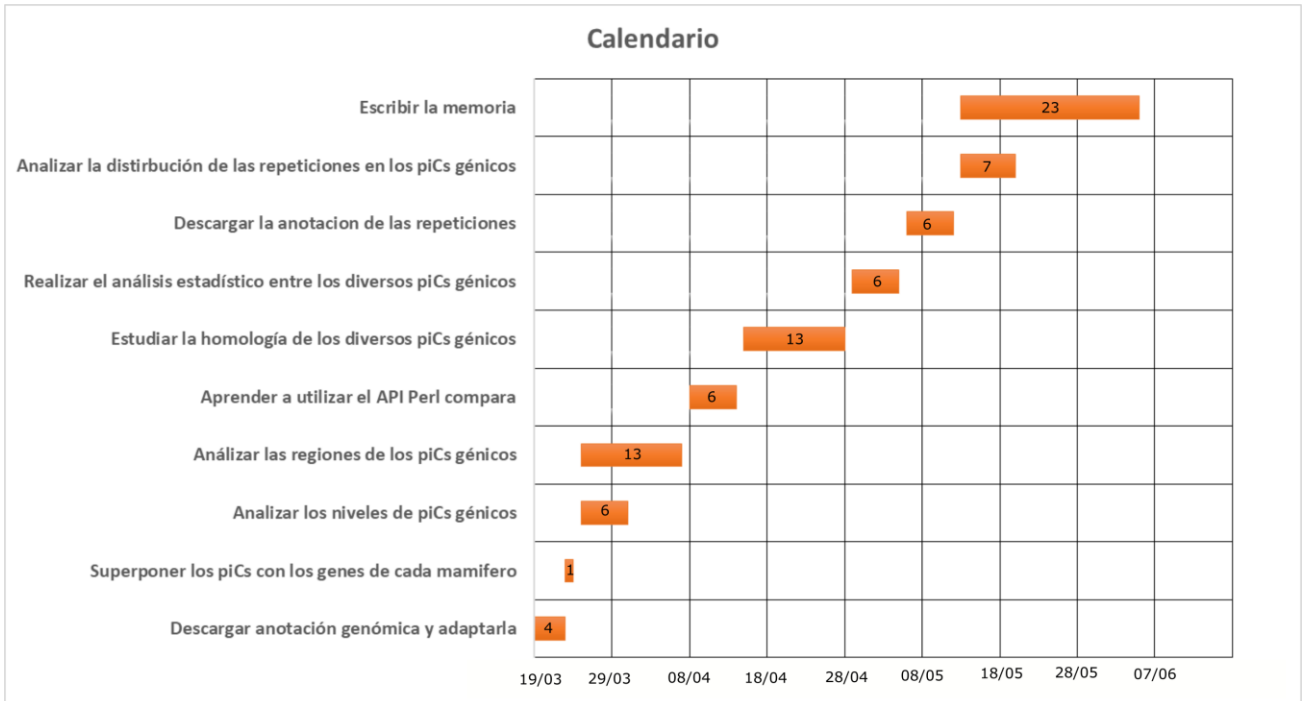


Figura 3. Diagrama de Gantt de la organización realizada para este trabajo.

1.5 Breve resumen de productos obtenidos

Los productos obtenidos de este trabajo son la memoria del proyecto y la presentación virtual del proyecto. Así mismo también se obtendrán archivos propios de la elaboración del proyecto como los archivos de anotación de piRNAs génicos en los diez animales estudiados y las GTIDs correspondientes a los identificadores de Ensembl de todos los genes codificantes de los 10 mamíferos.

1.6 Breve descripción de los otros capítulos de la memoria

Este trabajo va a seguir la distribución de un artículo científico, por ello se divide en 4 apartados: introducción, materiales y métodos, resultados y discusión y conclusión.

Introducción: Se expone el tema de estudio y se realiza aporta la información suficiente como para que el lector comprenda el motivo del trabajo.

Materiales y métodos: Se definen las herramientas y métodos utilizados a lo largo del trabajo con todo detalle para que el trabajo pueda ser reproducible.

Resultados y discusión: Se exponen los resultados obtenidos a lo largo del trabajo y se valoran.

Conclusión: Se realiza una valoración global del trabajo incidiendo en el porqué de los resultados y se proponen acciones para futuros proyectos.

Materiales y métodos

Identificación y anotación de los piC génicos

La anotación de los piCs predichos en los 10 mamíferos utilizados en este estudio (humano, macaco, tití, ratón, rata, conejo, perro, caballo, cerdo y zarigüeya) fue provista por la tutora de este proyecto Tanya Vavouri. La anotación de los piCs fue estimada con el programa proTRAC¹⁹ utilizando librerías de RNA pequeño de testículos de los diversos mamíferos. Los parámetros utilizados en proTRAC fueron los mismos en las 10 especies para que las anotaciones pudieran ser comparadas. Las versiones utilizadas de cada genoma para poder crear la anotación de los piCs se encuentran en la tabla 1.

Tabla 1. Tabla donde se muestra las versiones del genoma utilizadas para cada mamífero usado en este estudio

Mamífero	Versión genómica
Humano	hg38
Macaco	rheMac8.0.1
Tití	C_jacchus3.2.1
Ratón	GRCm38.p2
Rata	Rnor_6.0
Conejo	OryCun2.0
Cerdo	Sscrofa10.2
Caballo	Equ Cab 2
Perro	CanFam3.1
Zarigüeya	monDom5

Una vez obtenida la anotación de los genes de los diversos animales utilizando un script propio de R y el paquete biomaRT se procedió a superponer piCs y genes^{12,14,15}. Al conocer aquellos piCs génicos lo catalogamos en función de su ubicación en el gen en el que se encontraban (Figura 4). En este trabajo hemos considerado un total de 6 tipos distintos de piCs. Los piCs exónicos se sitúan enteramente dentro de una región exónica o en la región exónica y fuera del gen. Los piCs intrónicos se sitúan enteramente dentro de intrones. Los piCs exón-intrón se superponen a la vez en regiones exónicas e intrónicas. Los piCs intergénicos se sitúan fuera de las regiones génicas. Los piCs 3'UTR son piCs exónicos que además se superponen en regiones 3'UTR. Finalmente, los piCs multigénicos se encuentran superpuestos en dos genes a la vez.

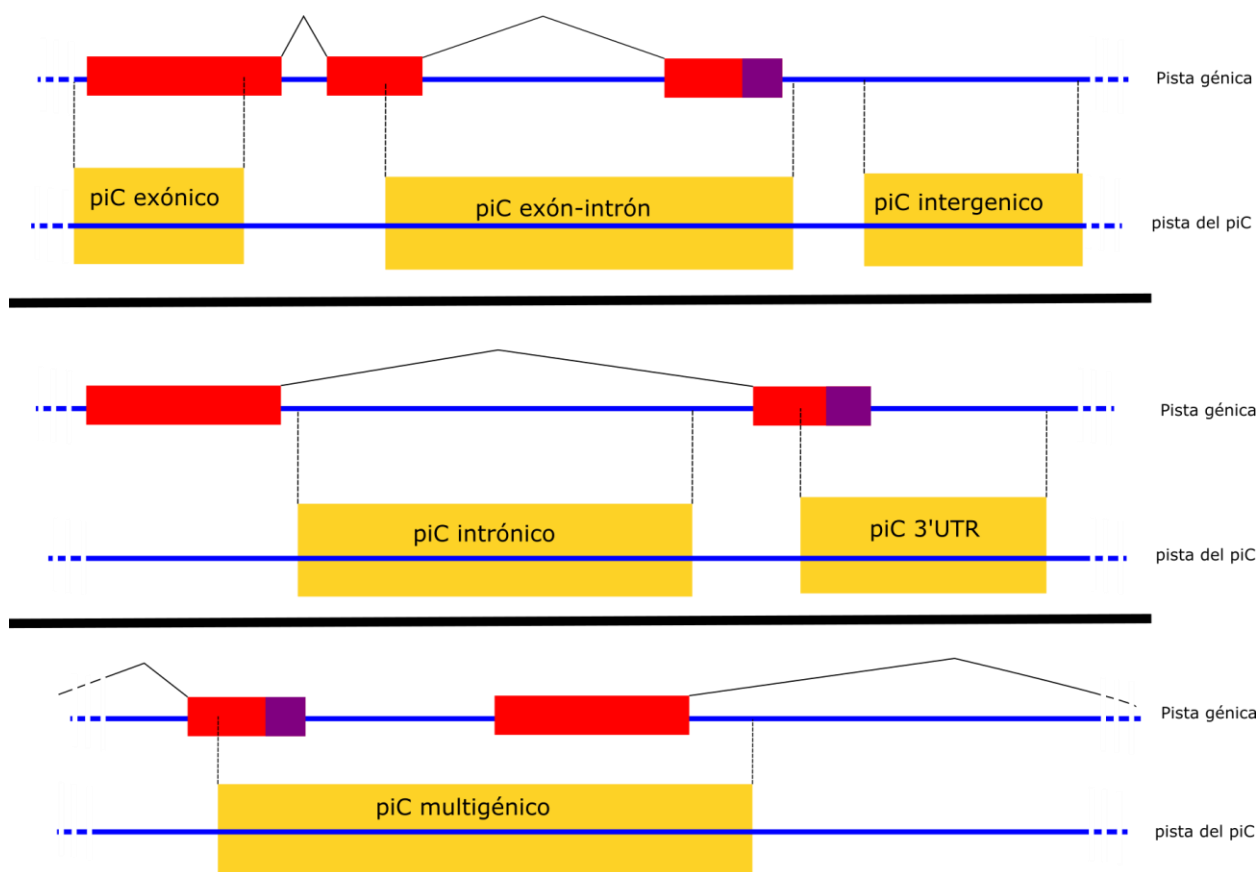


Figura 4. Esquema de la clasificación de los piRNA clústers (piCs) en función de su ubicación sobre los genes. En el esquema se han separado las pistas de piCs y de los genes para facilitar su visualización. Las regiones rojas son exones, las púrpuras regiones 3'UTR, las azules con líneas delgadas negras encima son intrones y las regiones amarillas son los piCs. En la figura se muestran los 6 tipos de piCs que hemos estudiado.

Identificación de genes ortólogos

Para definir aquellos genes productores de piRNAs que son ortólogos entre las distintas especies hemos utilizado las familias génicas de Ensembl Compara²⁰. Hemos identificado mediante el Ensembl Compara Perl API los identificadores de familia génica (GTIDs) de cada uno de los genes de todas las especies utilizando su identificador Ensembl único²¹. Ensembl se encarga de agrupar los genes de los diversos organismos que contienen en sus bases de datos en función de su región codificante (CDS). Una vez ha agrupado los diversos genes en familias les otorga un identificador estable de familia génica a los genes agrupados. Aquellos genes con el mismo GTID son considerados ortólogos²².

Los nombres de los genes que aparecen en las figuras de este documento han sido extraídos de la base de datos de Ensembl partir de su identificador Ensembl.

Para comprobar si la probabilidad de producción de piRNA en genes productores de piRNA ortólogos es mayor que la esperada por azar se realizó

el Test exacto de Fisher. Además, el p-valor de los distintos test estadísticos se ha modificado por la corrección de Bonferroni para evitar errores de tipo I.

Identificación de regiones repetidas

La anotación de las repeticiones en los distintos genomas fue descargada de la página UCSC¹⁸. En este trabajo únicamente nos hemos basado en las repeticiones de los transposones conocidos como DNA, LTR, SINE y LINE. Conociendo la ubicación de las repeticiones en los distintos genomas calculamos la densidad de repeticiones en 3 regiones distintas: piCs, genes y la región solapada entre piCs y genes (intersección piC-gen). Para calcular la densidad de las repeticiones se dividió el número total de pares de bases (bp) de repeticiones dentro de la región elegida dividido entre el total de bp de la región escogida. Para hacer este proceso utilizamos R y el paquete GRanges¹³.

Resultados y discusión

Clasificación de los piCs

Previamente en trabajos anteriores se ha descrito que los piRNAs producidos por genes son principalmente procesados desde la región 3'UTR en *Drosophila*, ratón y ranas²³. En este trabajo nos hemos preguntado si esto también sucede en mamíferos. Para conocer esto en este trabajo hemos predicho con los mismos parámetros los piCs en 10 especies que van desde la zarigüeya hasta los humanos. Posteriormente hemos superpuesto las posiciones de los genes y los piCs para conocer la posición de los piCs dentro de los genes. Hemos separado los piCs en 6 categorías distintas en función de su ubicación en el genoma: piCs exónicos, intrónicos, exón-intrón, 3'UTR, multigénicos e intergénicos. Sorprendentemente realizando este análisis hemos encontrado que la mayoría de piCs génicos son intrón-exón (Figura 5A). La proporción de 3'UTR piCs, entre un 1-19%, es más baja en todos los animales respecto los piCs exón-intrón. Además, en conejo la proporción de piCs intergénicos es más elevada que el resto. Debido a estos resultados inesperados tenemos varias hipótesis al respecto. En primer lugar, debemos tener en cuenta que, como hemos hecho la clasificación aquellos piCs que se superpongan no solo sobre la región 3'UTR si no también en intrones serán clasificados como piCs exón-intrón. Esto puede producir un menor porcentaje de piCs 3'UTR. Otro motivo para explicar esta baja proporción de piCs 3'UTR es la anotación genómica. Debido a que la mayoría de análisis genómicos se realizan en humanos y ratones y que existen muchos más datos de expresión génica estas anotaciones han sido corregidas manualmente más veces y por ello son más exactas. Esto podría explicar el mayor porcentaje de piCs 3'UTR en humano y ratón. Para comprobar esto decidimos analizar el total de kilobases (kb) de 3'UTR respecto al total de kb en exones en todos los mamíferos (Figura 6). Observando la proporción 3'UTR es claro que este factor también ha influido en la clasificación de los piCs 3'UTR. Humano, macaco y ratón tienen una mayor densidad de regiones 3'UTR respecto a los exones anotados y caballo presenta una cantidad mucho menor de 3'UTR anotadas respecto al resto de animales.

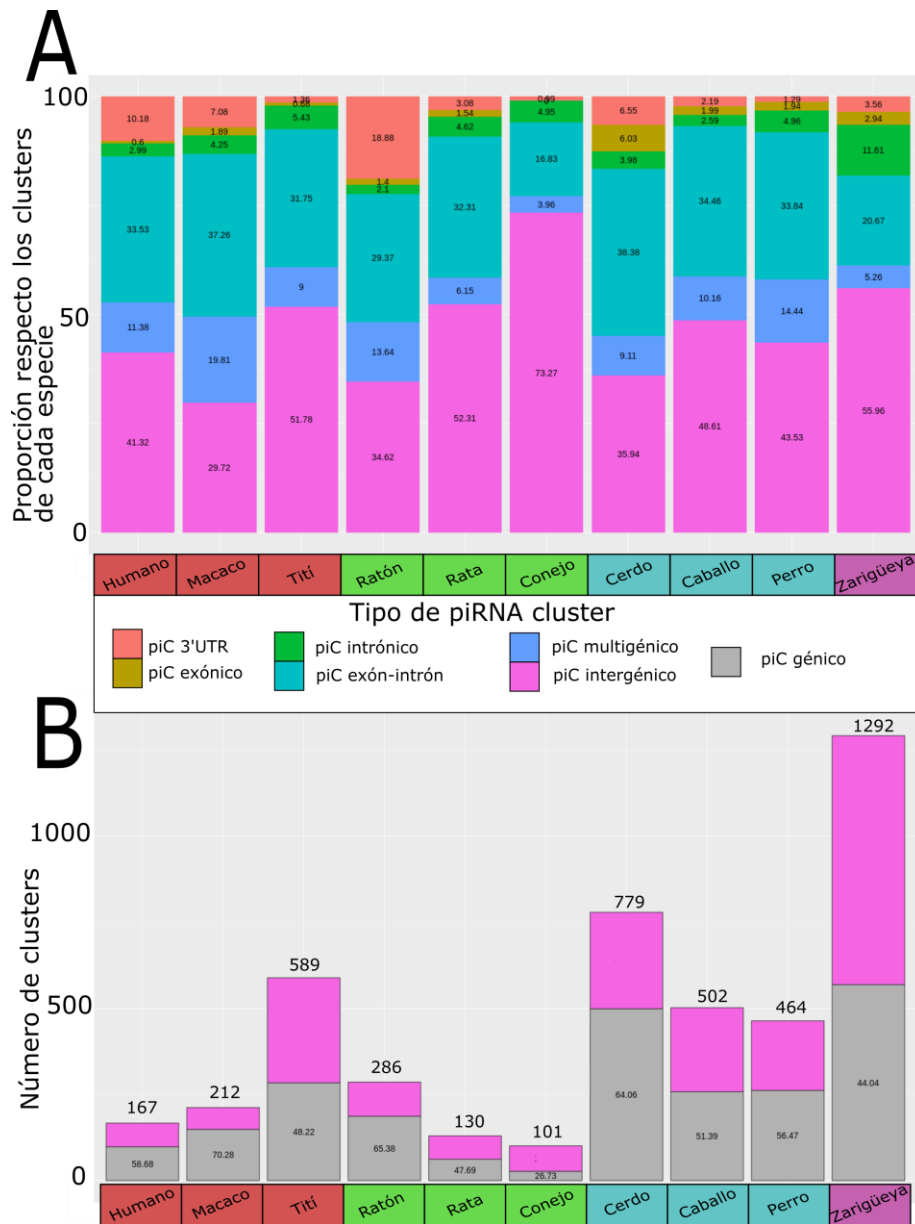


Figura 5. Clasificación de los tipos de piCs en 10 especies mamíferas distintas. En esta figura hemos definido un total de 7 tipos de piCs: aquellos que solapan regiones 3'UTR (rojo), únicamente exones (dorado), únicamente intrones (verde), exones e intrones al mismo tiempo (aguamarina), aquellos que solapan más de un gen a la vez (azul marino), los que no solapan ningún gen (lila). La proporción de cada tipo de piC está definido dentro de cada barra. La sección A muestra la proporción de cada tipo de piC dentro del total de piCs de cada mamífero estudiado. La sección B muestra el total de piCs en cada mamífero, así como el porcentaje de piCs génicos (gris).

En este trabajo también decidimos estudiar el número total de piCs en cada especie (Figura 5B). Los laurasiaterios (Cerdo, caballo y perro) y la zarigüeya son las especies con mayor número de piCs y piCs génicos. Vemos que, a pesar de que los valores de piCs en tití son más elevados, en primates (Humano, macaco y tití) y glires (Ratón, rata y conejo) los niveles de piCs y piCs génicos son parecidos y más bajos que los laurasiaterios. Esta diferencia entre estos clades puede ser debido a que los laurasiaterios se separaron de glires y primates hace más de 80 millones de años⁹. Recientemente una nueva teoría de Kofler, R. Describió como una selección negativa en los piCs puede

llevar a un estado de equilibrio entre TEs y piCs ya que estos últimos tienden a retener más tipos de transposones evitando que sea necesario el mantenimiento de una cantidad muy elevada de piCs⁷.

Genes ortólogos productores de piRNAs

Después de haber analizado los piCs génicos y su ubicación en los genes que solapan decidimos estudiar los genes que contienen los piCs. En este trabajo llamaremos genes productores de piRNAs a aquellos genes que contienen piCs. Como se esperaba los niveles de genes productores de piRNAs son similares a los niveles de piCs génicos (Figura 7). También podemos apreciar, como era esperable por los resultados de la distribución de los piCs, que el conejo es el animal con menor número de genes productores de piRNAs.

Para comprender mejor la conservación de genes productores de piRNAs decidimos hacer dos tipos de clasificaciones por GTID (Figura 8). En primer lugar, comparamos los GTIDs de todas las proteínas codificantes de los 3 clades que disponemos; primates, glires y Laurasiaterios (Figura 8A). Como era esperable, en esta primera comparación podemos ver como la gran mayoría de GTIDs se encuentran en los 3 animales de cada clade. Sin embargo, cuando comparamos el los GTIDs de aquellos genes productores de piRNAs la proporción de GTIDs desciende notablemente llegando hasta 0 en los glires (Figura 8B). Esto quiere decir que la mayoría de genes productores de piRNAs se conservan muy poco entre los diversos clades. Este resultado sugiere una rápida evolución de los genes productores de piRNAs.

Analizando los GTIDs compartidos descubrimos un total de 374 genes productores de piRNAs ortólogos al menos entre dos especies (Figura 8B y apéndice 1). Además, comprobamos que entre parejas de especies el número de genes productores de piRNA ortólogos podía variar entre 2 y 44 genes. Los genes que se conservan a lo largo de más especies son CBL, KIF1B y CBX5. CBL es una proteína con actividad ubiquitin ligasa. Se ha comprobado que mutaciones en el gen CBL están relacionadas la aparición de diversos tipos de cancer²⁴. Curiosamente también se ha observado que en moscas CBL es reprimido mediante piRNAs²⁵. KIF1B es un gen relacionado con la apoptosis de las neuronas²⁶. CBX5 es una proteína relacionada con la heterocromatina y el silenciamiento génico.

Después de conocer aquellos genes productores de piRNA conservados en las diversas especies decidimos comprobar si la conservación de estos genes había sido producida por azar o si ciertamente se conservan evolutivamente. Para ello realizamos una tabla de contingencia de 2x2 comparando cada pareja de animales (Figura 9 y apéndice 2). Observamos que, excepto en conejo y zarigüeya, los mamíferos estudiados conservaban entre ellos un mayor número de genes productores de piRNA que los que se esperaban por azar. Estos resultados dan a entender que a pesar del bajo número de genes productores de piRNA conservados entre los diversos estos son más de los que se generarían por azar lo que indica que son genes conservados entre las distintas especies estudiadas.

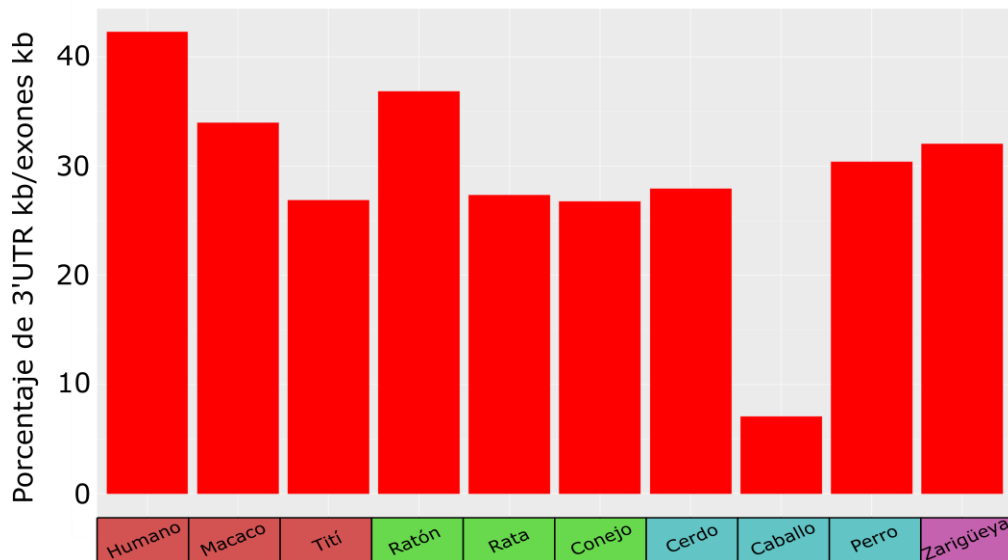


Figura 6. Diagrama de barras que muestra el porcentaje de pares de bases de las regiones 3'UTR respecto al número total de pares de bases de los exones en 10 mamíferos distintos.

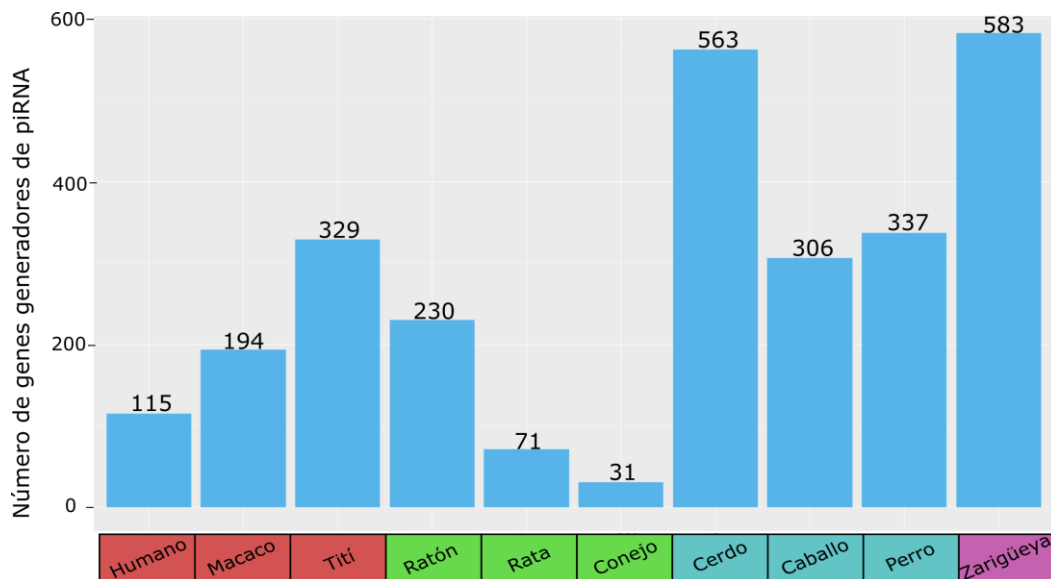


Figura 7. Diagrama de barras que muestra el número total de genes productores de piRNAs en 10 mamíferos distintos. En la parte superior de cada barra se muestra el total de genes productores de piRNAs de cada animal.

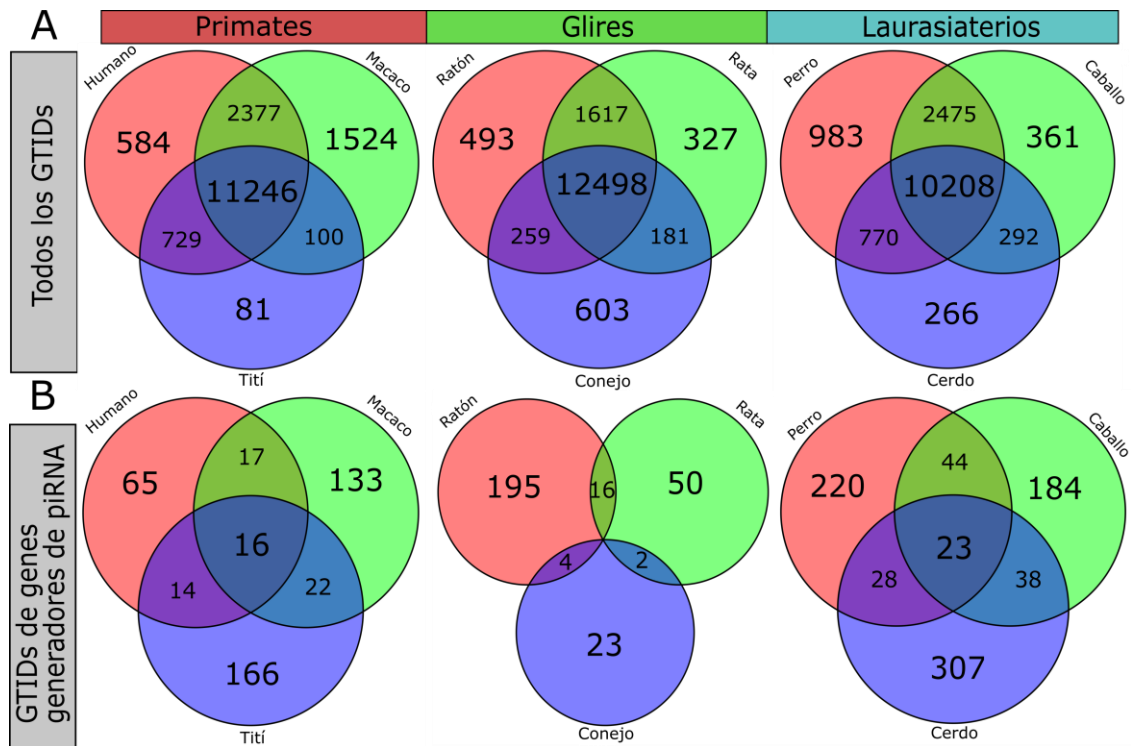


Figura 8. Diagramas de Venn de los GTIDs compartidos en los tres principales clades de este estudio; primates, glires y laurasiaterios. (A) Comparación de todos los GTID de cada uno de los mamíferos. (B) Comparación de los GTID de aquellos genes productores de piRNAs de cada mamífero.

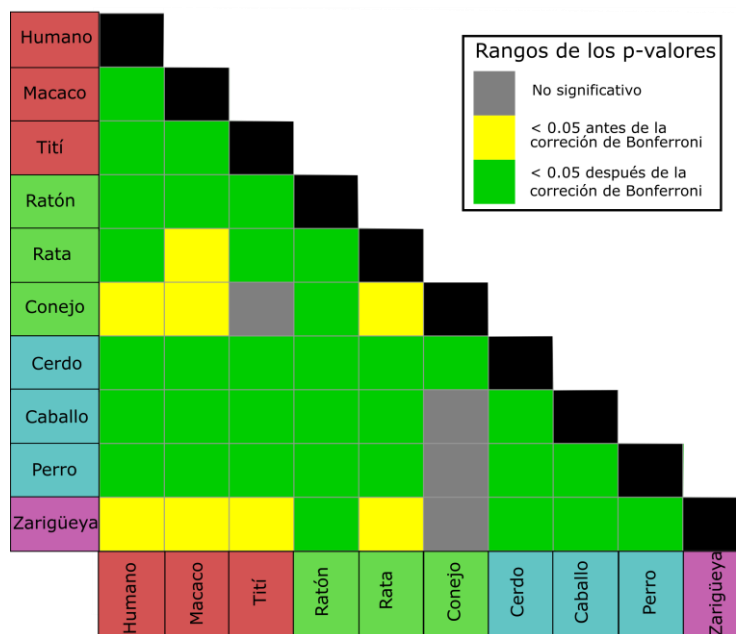


Figura 9. Representación gráfica de los p-valores obtenidos de la comparación 2 a 2 del número de genes ortólogos productores de piRNAs entre cada pareja de mamíferos. La comparación se ha realizado entre cada animal en la columna de la izquierda y los de la fila inferior. Los valores en gris muestran p-valores superiores a 0.05 incluso sin la aplicación de la corrección de Bonferroni. Los valores en amarillo muestran un p-valor inferior a 0.05 antes de la corrección de Bonferroni. Los valores en verde tienen un p-valor inferior a 0.05 después de la corrección de Bonferroni.

Localización de los piCs sobre los genes

Una vez conocemos los genes ortólogos productores de piRNAs decidimos estudiar la localización del piC en ellos (Figura 10). Hemos podido comprobar que la mayoría de piCs están ubicados en regiones exón-intrón en todos los genes ortólogos generadores de piCs. Este resultado es esperable ya que la mayoría de piCs génicos que hemos estudiados son exón-intrón (Figura 5). Otro hecho que hemos podido comprobar es que la segunda región más habitual para encontrar los piCs es 3'UTR en primates y en la comparación ratón-rata. Sin embargo, en Laurasiaterios el segundo tipo de piC más abundante son los exónicos. Finalmente podemos comprobar como los genes ortólogos que generan piRNAs minoritarios son aquellos que contienen más de un piC y aquellos alojados en intrones. En este análisis también quisimos comprobar como de conservadas estaban las ubicaciones en los genes productores de piRNAs más conservados en las distintas especies. Vemos que mayoritariamente están ubicados en regiones exón-intrón siendo una proporción de 6/8 en CBL, 6/6 en KIF1B y 4/5 en CBX5.

Regiones repetitivas en genes productores de piRNAs

Por último, decidimos investigar las secuencias repetitivas de los genes, piCs, piCs génicos y la intersección de los genes productores de piRNAs (Figuras 11 y 12). En primer lugar, observamos las densidades de distintas repeticiones y de todas juntas (Figura 11). Pudimos observar notables diferencias entre la densidad de repeticiones LTR y SINE entre los grupos. Las repeticiones LTR son mucho más comunes en piCs y piCs génicos respecto a los genes codificantes. Respecto a las densidades de los SINEs hemos visto que la densidad de repeticiones SINEs es mayor en genes de primates respecto al resto de mamíferos.

Teniendo en cuenta los resultados anteriores decidimos investigar las regiones solapadas entre piCs y genes que se encuentran en aquellos genes ortólogos más conservados (Figura 12). Observando las densidades de cada tipo de repetición podemos observar que los niveles de repeticiones LTR son ínfimos. Estos niveles de densidad de LTR indican que las regiones solapadas entre piCs y genes son más parecidas a genes que a piCs (Figura 11).

Estos resultados nos llevan a pensar que las repeticiones LTR tienen relación con la generación de piCs, tanto los génicos como los intergénicos. Hemos visto que esto no sucede en genes productores de piRNAs ortólogos ya que tienen una concentración mucho baja de repeticiones LTR, hecho que también sucede en genes. Esto nos lleva a pensar que es probable que los genes productores de piRNAs ortólogos se generasen por un mecanismo alternativo al resto de piCs que ha permitido su conservación a lo largo del tiempo.

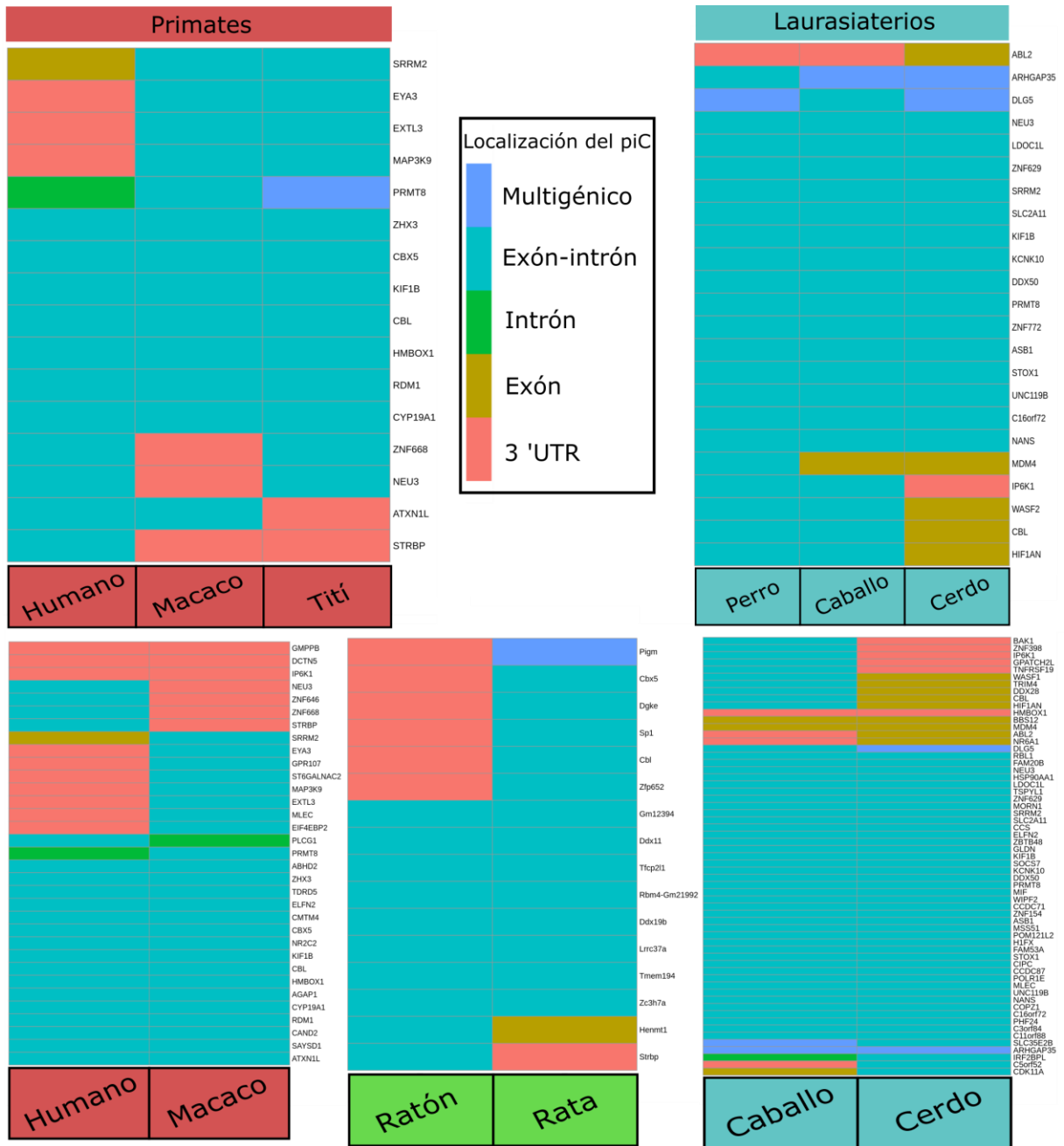


Figura 10. Posición de los piCs dentro de los genes ortólogos productores de piRNAs en diferentes especies de mamíferos indicado en la parte inferior de cada *heatmap*. Hemos dividido la figura en dos partes. La parte superior incluye los clades primates y laurasiaterios y la parte inferior agrupa aquellos genes ortólogos productores de piRNAs de especies más cercanas entre sí; humano-macaco, ratón-rata y caballo-cerdo. En cada *heatmap* hemos indicado, en la parte derecha, los símbolos de los genes ortólogos. Los piCs se han catalogado en función de su ubicación en el gen como: 3'UTR (rojo), exón (dorado), intrón (verde), exón-intrón (aguamarina) y multigénicos (azul marino).

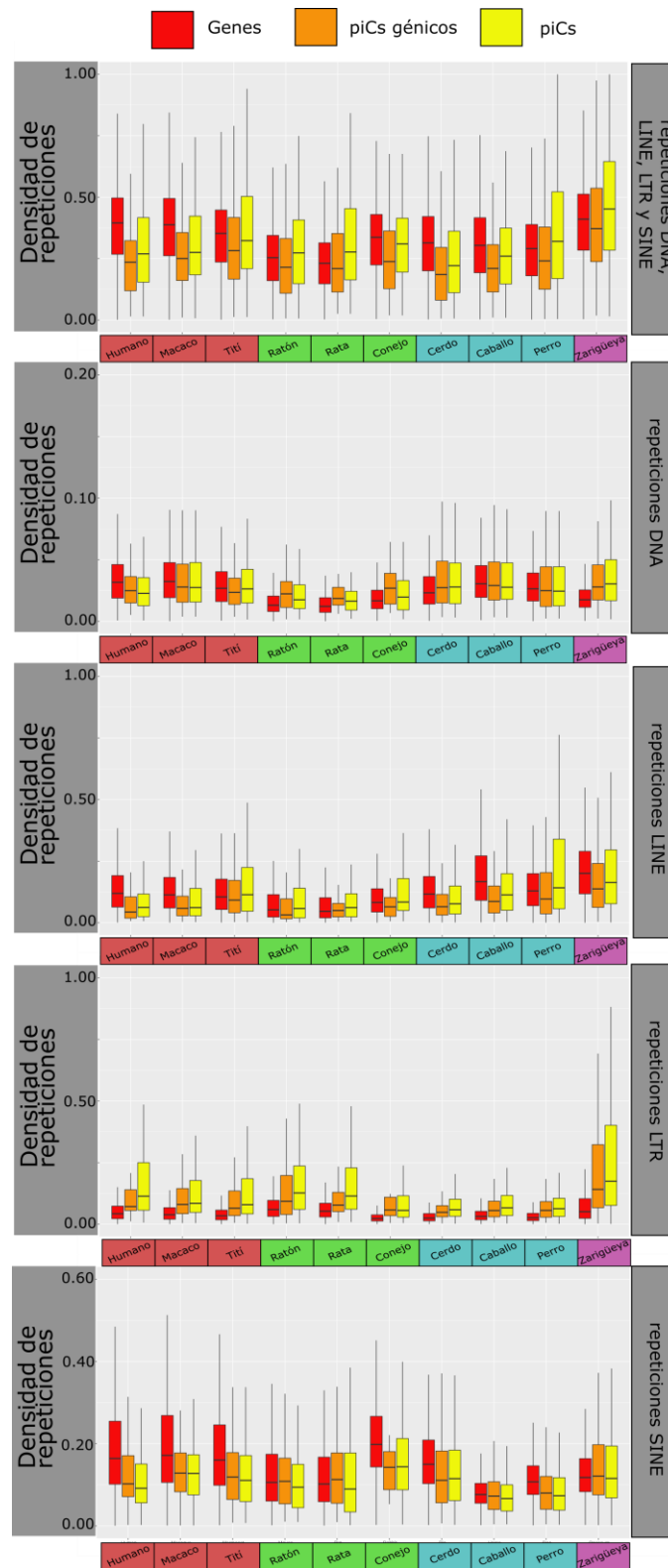


Figura 11. Diagramas de cajas de las densidades en 10 mamíferos de los distintos tipos de repeticiones en 3 regiones genómicas distintas; genes (rojo), piCs (amarillo) y piCs génicos (naranja). De arriba hacia abajo se muestran las distribuciones de la densidad de; todas las repeticiones estudiadas juntas, repeticiones de virus DNA, repeticiones LINE, repeticiones LTR y repeticiones SINE. Para mejorar la visualización de los datos se ha omitido la representación de los valores atípicos.

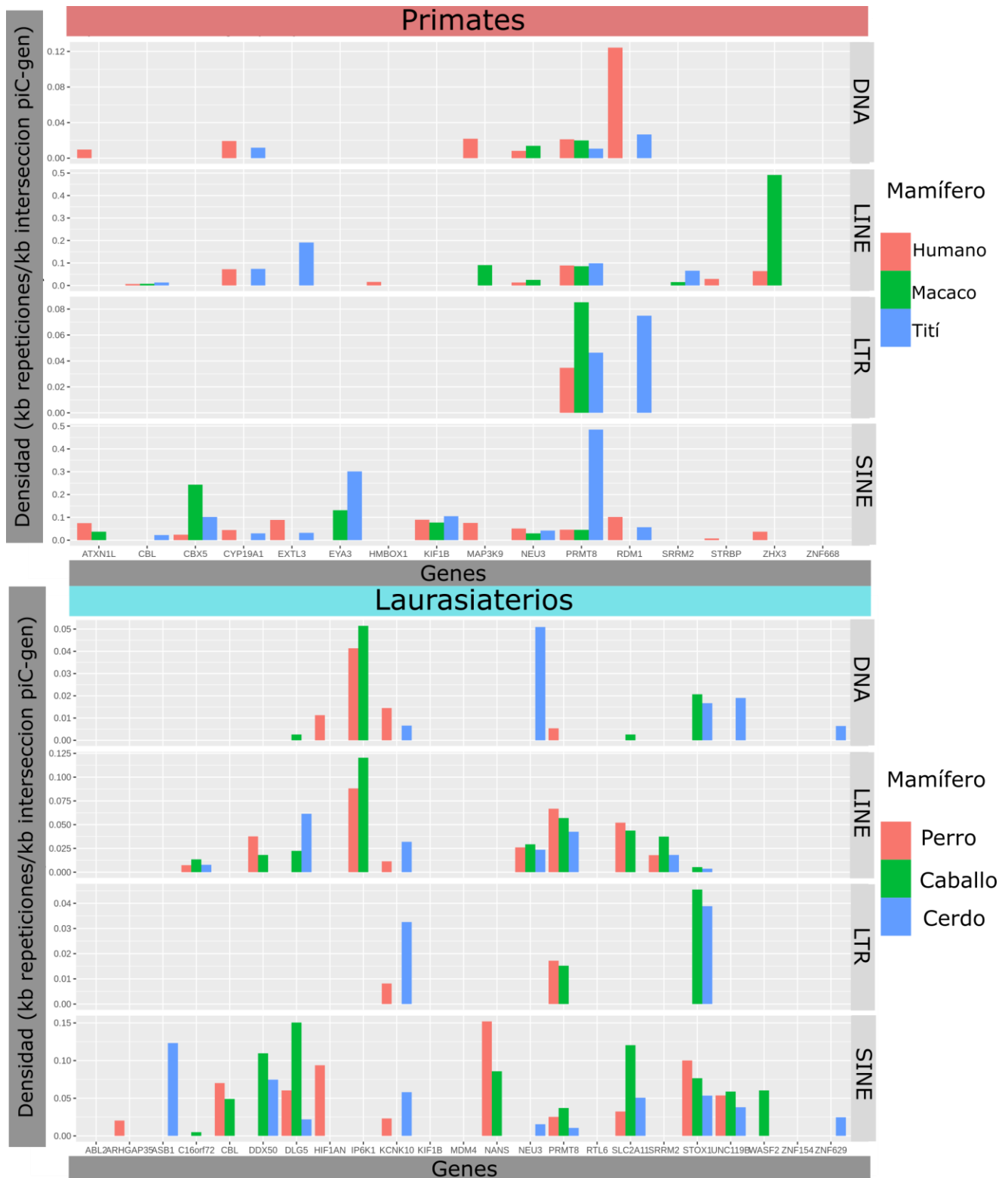


Figura 12. Diagramas de barras que muestra la densidad de 4 tipos de repeticiones en primates (superior) y laurasiaterios (inferior). La región que se ha analizado la densidad es la región solapada entre piC y gen. En la parte inferior de cada bloque se indica el nombre del gen que se ha estudiado.

3. Conclusiones

Esta descrito en la bibliografía que tanto piRNAs como piCs evolucionan rápidamente entre las distintas especies debido a una selección positiva. Esta selección positiva favorece un mayor número de piCs para que puedan detener la propagación de un mayor número de transposones²⁷. Sin embargo, Chirn, et al., describieron que algunos piCs están conservados entre distintas especies de mamíferos. La función de estos piCs conservados se ha hipotetizado que tiene relación con la correcta formación de la espermatogénesis y el correcto desarrollo del embrión ya que mutaciones en los genes que producen piRNAs también producen una baja en los niveles de piRNAs que protegen a los oocitos y a los espermatozoides frente a los transposones⁹. En este estudio nos hemos basado en estudiar los genes que generan piRNAs en lugar de estudiar los piRNAs en sí mismos. Hemos encontrado que alrededor del 50% de los piCs son génicos. También hemos visto que un total de 374 genes que codifican proteínas también generan piRNAs en al menos dos de todos los animales que hemos estudiado (Apéndice 1).

En este trabajo hemos visto como la mayoría de los genes ortólogos productores de piRNAs contenían su piC en regiones situadas entre exones e intrones. Normalmente las regiones exón-intrón solapan parcialmente regiones CDS en el gen. Nuestra hipótesis es que estos piCs tienden a conservarse más que el resto por la importancia del gen en el que se encuentran. Mutaciones en los piCs situados en CDS podrían significar producir mutaciones en el gen en el que se encuentran lo que podría provocar alteraciones celulares importantes dependiendo del gen mutado.

Otra hipótesis que proponemos al hecho de que los piCs conservados estén situados en regiones exón-intrón sería que la generación de piRNAs se produce por un fallo en el proceso de *splicing* del mRNA. Estos fallos en el *splicing* producirían que en lugar de degradarse el mRNA entrara en la vía de los piRNAs. Esta hipótesis también se respaldaría en el hecho de que los piCs conservados presentan niveles de repeticiones LTR más parecidas a genes que a piCs.

A lo largo del trabajo hemos seguido la planificación que realizamos en los primeros estadios. Sin embargo, ha habido una leve limitación en la elaboración del trabajo. En un principio teníamos la anotación de piCs de un total de 11 animales. Sin embargo, tuvimos que descartar uno de ellos (ornitorrinco) por la imposibilidad de encontrar la anotación genómica en las dos bases de datos que hemos utilizado en este trabajo, Ensembl y UCSC. Finalmente, ante la imposibilidad de añadir más animales al estudio, se ha decidido que la utilización de 10 animales y 3 clades era suficiente para obtener resultados significativos de la evolución de los piRNAs.

Entre las posibles líneas de trabajo futuras destacamos las siguientes:

- Análisis de secuencias y motivos conservados de genes productores de piRNAs.
- Análisis de las regiones circundantes a los genes productores de piRNAs para encontrar motivos o repeticiones conservadas entre diversos mamíferos.
- Estudios funcionales de los genes productores de piRNAs que se encuentran más conservados para encontrar relación entre ellos.

4. Glosario

piC (piRNA clúster): Región genómica que contiene una cobertura muy abundante de piRNAs.

Gen generador de piRNAs: gen que contiene uno más piC en su loci.

3'UTR: Región transcrita de un gen pero que no es codificante. Esta región se sitúa en el extremo 3' del gen.

Sintenia: Bloques de genes u otros marcadores que conservan evolutivamente el orden²⁸.

Ortología: en genética hace referencia la término de genes o proteínas que son homologas, es decir, comparten un alto grado de similitud y que, además provienen del mismo gen o proteína común²⁹.

GTID (Gene Tree IDs): Identificadores otorgados por Ensembl. Estos identificadores sirven para indicar aquellos genes que se han conservado entre distintas especies y que, por lo tanto, son ortólogos.

5. Bibliografía

1. Aravin, A. A. *et al.* Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11**, 1017–27 (2001).
2. Kim, V. N. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev.* **20**, 1993–1997 (2006).
3. Ozata, D. M., Gainetdinov, I., Zoch, A., O’Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2019).
4. Stuwe, E., Tóth, K. F. & Aravin, A. A. Small but sturdy: small RNAs in cellular memory and epigenetics. *Genes Dev.* **28**, 423–31 (2014).
5. Iwasaki, Y. W., Siomi, M. C. & Siomi, H. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu. Rev. Biochem.* **84**, 405–433 (2015).
6. Sienski, G., Dönertas, D. & Brennecke, J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**, 964–80 (2012).
7. Kofler, R. Dynamics of Transposable Element Invasions with piRNA Clusters. *Mol. Biol. Evol.* (2019). doi:10.1093/molbev/msz079
8. Parhad, S. S. & Theurkauf, W. E. Rapid evolution and conserved function of the piRNA pathway. *R. Soc. Open Biol.* (2019). doi:10.1098/RSOB.180181
9. Chirn, G. *et al.* Conserved piRNA Expression from a Distinct Set of piRNA Cluster Loci in Eutherian Mammals. *PLOS Genet.* **11**, e1005652 (2015).
10. Coline, G., ThÄ©ron, E., Brassat, E. & Vaury, C. History of the discovery of a master locus producing piRNAs: the flamenco/COM locus in *Drosophila melanogaster*. *Front. Genet.* **5**, 257 (2014).
11. Saito, K. *et al.* A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* **461**, 1296–1299 (2009).
12. R Core Team (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing* (2018).
13. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
14. Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761 (2018).
15. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
16. Herrero, J. *et al.* Ensembl comparative genomics resources. *Database (Oxford)*. **2016**, (2016).
17. Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. (2015).
18. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, 493D – 496 (2004).
19. Rosenkranz, D. & Zischler, H. proTRAC - a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **13**, 5 (2012).
20. Vilella, A. J. *et al.* EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res.* **19**, 327–35 (2009).

21. Yates, A. *et al.* The Ensembl REST API: Ensembl Data for Any Language. *Bioinformatics* **31**, 143–145 (2015).
22. Ensembl genome browser 96. Available at: <https://www.ensembl.org/index.html>. (Accessed: 18th May 2019)
23. Robine, N. *et al.* A Broadly Conserved Pathway Generates 3'UTR-Directed Primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).
24. Naramura, M. *et al.* Mutant Cbl proteins as oncogenic drivers in myeloproliferative disorders. *Oncotarget* **2**, 245 (2011).
25. Rojas-Ríos, P., Chartier, A., Pierson, S. & Simonelig, M. Aubergine and piRNAs promote germline stem cell self-renewal by repressing the proto-oncogene Cbl. *EMBO J.* **36**, 3194–3211 (2017).
26. Schlisio, S. *et al.* The kinesin KIF1B acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. *Genes Dev.* **22**, 884–893 (2008).
27. Assis, R. & Kondrashov, A. S. Rapid repetitive element-mediated expansion of piRNA clusters in mammalian evolution. *Proc. Natl. Acad. Sci.* **106**, 7079–7082 (2009).
28. Sinha, A. U. & Meller, J. Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. *BMC Bioinformatics* **8**, 82 (2007).
29. Koonin, E. V. Orthologs, Paralogs, and Evolutionary Genomics. *Annu. Rev. Genet.* **39**, 309–338 (2005).

6. Anexos

Apéndice 1. Lista de genes productores de piRNA que están conservados entre 10 especies distintas. En la columna GTID se indica la familia génica a la que pertenecen. En la columna número de animales compartidos se indica el número de animales que comparten dicho gen generador de piRNAs. En la columna animales conservados se indican los animales que conservan el gen generador de piRNAs. En la columna genes se indica el nombre del gen.

GTID	Número de animales compartidos	Animales conservados	Genes
ENSGT00940000155772	9	Perro Caballo Humano Macaco Tití Ratón Zarigüeya Cerdo Rata	CBL
ENSGT00940000157445	7	Perro Caballo Humano Macaco Tití Ratón Cerdo	KIF1B
ENSGT00940000158801	7	Perro Caballo Humano Macaco Tití Ratón Rata	CBX5
ENSGT00520000055589	6	Perro Caballo Humano Tití Ratón Cerdo	STOX1
ENSGT00940000159082	6	Perro Caballo Humano Tití Ratón Zarigüeya	PDPR
ENSGT00940000161127	6	Perro Caballo Humano Macaco Tití Cerdo	SRRM2
ENSGT00940000156147	6	Perro Caballo Macaco Tití Ratón Cerdo	Kcnk10
ENSGT00950000182944	6	Perro Caballo Humano Macaco Tití Cerdo	NEU3
ENSGT00940000153969	6	Perro Caballo Humano Tití Ratón Cerdo	ASB1
ENSGT00940000155867	6	Perro Caballo Humano Macaco Tití Cerdo	PRMT8
ENSGT00940000157409	6	Perro Caballo Macaco Ratón Zarigüeya Cerdo	Hif1an
ENSGT00940000154928	6	Caballo Humano Macaco Tití Ratón Cerdo	HMBOX1
ENSGT00950000182648	5	Perro Caballo Macaco Tití Ratón	Kpna4
ENSGT00950000182978	5	Perro Humano Macaco Tití Cerdo	EYA3
ENSGT00940000155342	5	Perro Humano Macaco Ratón Zarigüeya	EIF4EBP2
ENSGT00940000155303	5	Perro Caballo Macaco Tití Cerdo	DLG5
ENSGT00390000002762	5	Perro Caballo Humano Tití Ratón	SHMT1
ENSGT00390000010052	5	Perro Caballo Humano	SMCR8

		Tití Ratón	
ENSGT0039000005939	5	Perro Humano Macaco Tití Ratón	ATXN1L
ENSGT00390000013127	5	Perro Caballo Tití Zarigüeya Rata	MIEF1
ENSGT00390000011081	5	Perro Caballo Macaco Tití Cerdo	NANS
ENSGT00940000157802	5	Perro Caballo Humano Macaco Cerdo	IP6K1
ENSGT00840000129915	5	Perro Caballo Humano Macaco Tití	CYP19A1
ENSGT00530000063539	5	Perro Caballo Ratón Zarigüeya Cerdo	Mdm4
ENSGT00390000016504	5	Caballo Humano Macaco Ratón Cerdo	MLEC
ENSGT00940000159737	5	Caballo Humano Macaco Ratón Cerdo	ELFN2
ENSGT00940000158393	5	Humano Macaco Ratón Cerdo Conejo	NR2C2
ENSGT00940000154687	5	Humano Macaco Tití Ratón Rata	STRBP
ENSGT00940000158303	4	Perro Caballo Macaco Zarigüeya	TXLNA
ENSGT00950000182962	4	Perro Caballo Humano Cerdo	WASF2
ENSGT00940000156701	4	Perro Caballo Humano Tití	TOP3A
ENSGT00390000013836	4	Perro Tití Ratón Cerdo	Tbl2
ENSGT00940000161588	4	Perro Caballo Humano Macaco	ZNF646
ENSGT00390000015360	4	Perro Caballo Humano Macaco	DCTN5
ENSGT00390000002886	4	Perro Caballo Macaco Cerdo	C16orf72
ENSGT00940000153838	4	Perro Caballo Ratón Cerdo	Abl2
ENSGT00530000063282	4	Perro Ratón Cerdo Rata	Lrrc37a
ENSGT00940000161010	4	Perro Caballo Macaco Tití	TRIM14
ENSGT00940000161262	4	Perro Macaco Tití Ratón	Chst1
ENSGT00940000158909	4	Perro Caballo Humano Macaco	GMPPB
ENSGT00940000161576	4	Perro Caballo Macaco Tití	ZKSCAN8
ENSGT00940000159351	4	Caballo Tití Zarigüeya Cerdo	SLC35E2B
ENSGT00390000005140	4	Caballo Humano Tití Zarigüeya	FAM168A
ENSGT00940000155557	4	Caballo Macaco Tití	WIPF2

		Cerdo	
ENSGT00920000149141	4	Caballo Humano Macaco Tití	ZNF668
ENSGT00950000183043	4	Caballo Macaco Ratón Rata	Gm12394
ENSGT00940000158243	4	Humano Macaco Tití Cerdo	MAP3K9
ENSGT00940000159573	4	Humano Macaco Ratón Zarigüeya	CMTM4
ENSGT00940000155271	4	Humano Tití Ratón Cerdo	GNAL
ENSGT00950000182893	4	Humano Macaco Tití Ratón	ZHX3
ENSGT00940000158877	4	Macaco Tití Ratón Cerdo	Cds2
ENSGT00940000154734	3	Perro Caballo Cerdo	ZNF772
ENSGT00950000182819	3	Perro Caballo Cerdo	ARHGAP35
ENSGT00940000161660	3	Perro Tití Zarigüeya	AEN
ENSGT00390000009416	3	Perro Macaco Tití	WDR1
ENSGT00940000155901	3	Perro Caballo Cerdo	DDX50
ENSGT00940000159421	3	Perro Caballo Tití	ZNF821
ENSGT00940000153104	3	Perro Zarigüeya Cerdo	ZNF3
ENSGT00940000161909	3	Perro Caballo Cerdo	ZNF629
ENSGT00710000106769	3	Perro Caballo Tití	TRIP11
ENSGT00400000022095	3	Perro Caballo Humano	STRIP1
ENSGT00530000063651	3	Perro Caballo Ratón	Tmed8
ENSGT00950000182753	3	Perro Tití Cerdo	GK2
ENSGT00950000182629	3	Perro Zarigüeya Rata	CPT1A
ENSGT00940000162853	3	Perro Caballo Cerdo	LDOC1L
ENSGT00940000157896	3	Perro Macaco Zarigüeya	TCF20
ENSGT00950000182657	3	Perro Caballo Tití	RFX4
ENSGT00940000160246	3	Perro Humano Zarigüeya	ZBTB5
ENSGT00940000157048	3	Perro Caballo Zarigüeya	POLH
ENSGT00940000157288	3	Perro Zarigüeya Cerdo	PLEKHA8
ENSGT00390000004945	3	Perro Humano Tití	POMK
ENSGT00940000153123	3	Perro Tití Rata	ALMS1
ENSGT00950000182641	3	Perro Humano Cerdo	HIST2H2BF
ENSGT00950000182715	3	Perro Cerdo Rata	H3F3A
ENSGT00950000183212	3	Perro Macaco Ratón	Snx27
ENSGT00940000155031	3	Perro Cerdo Rata	MARK2
ENSGT00940000155864	3	Perro Caballo Ratón	Rab7
ENSGT00940000153680	3	Perro Tití Ratón	Rnf169
ENSGT00940000161238	3	Perro Caballo Conejo	ENSSSCG00000036465
ENSGT00940000157210	3	Perro Caballo Ratón	Slc9a8
ENSGT00940000153661	3	Perro Zarigüeya Cerdo	Ankrd26
ENSGT00390000015354	3	Perro Caballo Zarigüeya	ORAI2
ENSGT00390000014595	3	Perro Caballo Cerdo	UNC119B
ENSGT00940000161061	3	Perro Caballo Cerdo	SLC2A11
ENSGT00950000182686	3	Perro Zarigüeya Cerdo	LRRC14

ENSGT00390000018606	3	Perro Tití Ratón	Gnb1l
ENSGT00390000015552	3	Perro Tití Ratón	Bcl2l13
ENSGT00940000163685	3	Perro Caballo Ratón	1700006A11Rik
ENSGT00530000063187	3	Caballo Ratón Zarigüeya	Zyg11b
ENSGT00940000158568	3	Caballo Humano Zarigüeya	AGO1
ENSGT00940000158981	3	Caballo Macaco Cerdo	ZBTB48
ENSGT00940000158459	3	Caballo Zarigüeya Cerdo	CDK11A/B
ENSGT00940000157422	3	Caballo Tití Ratón	Slc4a8
ENSGT00390000004405	3	Caballo Macaco Cerdo	COPZ1
ENSGT00950000182674	3	Caballo Humano Tití	CTDSP2
ENSGT00390000002677	3	Caballo Zarigüeya Cerdo	C11orf88
ENSGT00940000155608	3	Caballo Ratón Cerdo	Mif
ENSGT00940000161453	3	Caballo Macaco Zarigüeya	YPEL1
ENSGT00390000006939	3	Caballo Macaco Tití	DCAF7
ENSGT00940000156314	3	Caballo Humano Cerdo	SOCS7
ENSGT00940000160156	3	Caballo Ratón Zarigüeya	Cyb5d2
ENSGT00940000157436	3	Caballo Ratón Zarigüeya	Rbm14
ENSGT00940000161327	3	Caballo Humano Ratón	KCTD7
ENSGT00530000063371	3	Caballo Ratón Cerdo	Fam53b
ENSGT00730000111536	3	Caballo Ratón Cerdo	H1fx
ENSGT00940000155349	3	Caballo Macaco Zarigüeya	SLC7A1
ENSGT00410000025698	3	Caballo Macaco Cerdo	GPATCH2L
ENSGT00390000002865	3	Caballo Zarigüeya Cerdo	PHF24
ENSGT00940000160451	3	Caballo Humano Macaco	GPR107
ENSGT00940000153253	3	Caballo Cerdo Conejo	POM121L12
ENSGT00390000018542	3	Caballo Ratón Rata	Zc3h7a
ENSGT00940000159158	3	Humano Ratón Rata	TFCP2L1
ENSGT00940000154793	3	Humano Macaco Zarigüeya	AGAP1
ENSGT00940000156692	3	Humano Macaco Tití	EXTL3
ENSGT00950000182902	3	Humano Macaco Cerdo	ABHD2
ENSGT00730000111392	3	Humano Tití Ratón	FAM174B
ENSGT00390000018397	3	Humano Macaco Tití	RDM1
ENSGT00490000043329	3	Humano Ratón Cerdo	MAPRE1
ENSGT00940000158901	3	Humano Macaco Ratón	PLCG1
ENSGT00440000039706	3	Macaco Tití Cerdo	FBXO42
ENSGT00640000091558	3	Macaco Ratón Conejo	Mrps25
ENSGT00530000063320	3	Macaco Ratón Zarigüeya	Mfsd4
ENSGT00940000155027	3	Macaco Zarigüeya Cerdo	NSD1
ENSGT00390000014208	3	Macaco Ratón Conejo	Gm4884
ENSGT00940000161379	3	Macaco Zarigüeya Cerdo	TESMIN
ENSGT00660000095569	3	Macaco Tití Cerdo	MED1
ENSGT00940000153434	3	Tití Cerdo Rata	HNRNPK

ENSGT00940000158604	3	Tití Ratón Rata	Dgke
ENSGT00940000157416	3	Tití Ratón Rata	Zfp652
ENSGT00950000183176	3	Ratón Zarigüeya Cerdo	Cbfa2t2
ENSGT00390000003084	3	Ratón Zarigüeya Cerdo	AU022252
ENSGT00940000154417	3	Ratón Zarigüeya Rata	Ddx19b
ENSGT00390000002174	3	Ratón Zarigüeya Rata	Tmem194
ENSGT00390000016694	3	Ratón Zarigüeya Cerdo	Mettl16
ENSGT00940000163387	2	Perro Caballo	ZNF132
ENSGT00950000182806	2	Perro Zarigüeya	ZNF850
ENSGT00940000161799	2	Perro Zarigüeya	ZNF574
ENSGT00940000161781	2	Perro Cerdo	ZNF585A
ENSGT00940000161777	2	Perro Zarigüeya	4930430A15Rik
ENSGT00390000000002	2	Perro Macaco	APBB2
ENSGT00950000182827	2	Perro Cerdo	SLC35A4
ENSGT00940000153181	2	Perro Caballo	RGD1559804
ENSGT00940000159268	2	Perro Tití	SYNM
ENSGT00950000182810	2	Perro Zarigüeya	PRELID2
ENSGT00390000011424	2	Perro Cerdo	DPAGT1
ENSGT00940000155542	2	Perro Zarigüeya	PGM1
ENSGT00950000183108	2	Perro Cerdo	CALB2
ENSGT00650000093476	2	Perro Caballo	CHTF8
ENSGT00390000010337	2	Perro Cerdo	HAS1
ENSGT00940000162625	2	Perro Caballo	SRRM3
ENSGT00940000157752	2	Perro Macaco	PLK1
ENSGT00940000162456	2	Perro Cerdo	ZNF75A
ENSGT00950000182888	2	Perro Zarigüeya	TOR1B
ENSGT00940000162691	2	Perro Conejo	CATSPERE
ENSGT00940000160472	2	Perro Cerdo	DENND4B
ENSGT00390000001278	2	Perro Zarigüeya	ITPK1
ENSGT00550000075074	2	Perro Cerdo	IPO4
ENSGT00940000158667	2	Perro Cerdo	TM9SF1
ENSGT00940000158098	2	Perro Tití	ZNF410
ENSGT00940000159998	2	Perro Zarigüeya	FAM161B
ENSGT00940000158336	2	Perro Humano	KCTD2
ENSGT00940000156241	2	Perro Zarigüeya	AARSD1
ENSGT00940000161498	2	Perro Caballo	SP6
ENSGT00940000160918	2	Perro Cerdo	SETX
ENSGT00940000154208	2	Perro Caballo	Tob1
ENSGT00390000005650	2	Perro Caballo	ANKRD53
ENSGT00390000010888	2	Perro Caballo	TEX261
ENSGT00390000017030	2	Perro Caballo	Cuta
ENSGT00950000183180	2	Perro Caballo	PHF19
ENSGT00390000007975	2	Perro Humano	KCTD20
ENSGT00940000155636	2	Perro Caballo	GTPBP2
ENSGT00390000014963	2	Perro Cerdo	BEND4
ENSGT00940000154572	2	Perro Zarigüeya	ANKS1B
ENSGT00940000162604	2	Perro Caballo	ZNF786
ENSGT00940000155320	2	Perro Cerdo	SLC4A1AP

ENSGT00940000158503	2	Perro Zarigüeya	THNSL2
ENSGT00940000157203	2	Perro Caballo	RNF115
ENSGT00940000156302	2	Perro Rata	H2A2A
ENSGT00950000182940	2	Perro Humano	SV2B
ENSGT00940000158791	2	Perro Tití	LRIG2
ENSGT00940000157402	2	Perro Caballo	TCP11L1
ENSGT00390000005232	2	Perro Ratón	Bbs1
ENSGT00390000018439	2	Perro Cerdo	HMCES
ENSGT00940000159273	2	Perro Caballo	GLT8D1
ENSGT00940000158320	2	Perro Caballo	GNL3
ENSGT00940000160561	2	Perro Cerdo	RBM15B
ENSGT00440000038032	2	Perro Zarigüeya	GTF2F1
ENSGT00390000016048	2	Perro Macaco	XRRA1
ENSGT00940000159443	2	Perro Caballo	RNF24
ENSGT00940000157626	2	Perro Rata	PANK2
ENSGT00720000108866	2	Perro Ratón	Actr5
ENSGT00940000155407	2	Perro Caballo	TTPAL
ENSGT00950000182793	2	Perro Cerdo	SERINC4
ENSGT00940000159787	2	Perro Zarigüeya	STK4
ENSGT00950000182724	2	Perro Zarigüeya	SLC13A5
ENSGT00940000160868	2	Perro Zarigüeya	TRIM35
ENSGT00940000158642	2	Perro Zarigüeya	FARP2
ENSGT00550000074985	2	Perro Tití	DHX37
ENSGT00940000155165	2	Perro Cerdo	VPS33A
ENSGT00950000183200	2	Perro Cerdo	PIWIL3
ENSGT00940000159296	2	Perro Caballo	PATZ1
ENSGT00940000155261	2	Perro Tití	ATF7
ENSGT00390000009802	2	Perro Tití	FAM105A
ENSGT00940000153783	2	Perro Rata	EIF4A2
ENSGT00940000158071	2	Perro Zarigüeya	FARS2
ENSGT00940000153092	2	Perro Humano	HIST1H2AE
ENSGT00950000182855	2	Perro Tití	CENPA
ENSGT00950000182994	2	Perro Zarigüeya	GDI2
ENSGT00940000155501	2	Perro Humano	HIST1H1E
ENSGT00940000154173	2	Perro Humano	SPDYE1
ENSGT00390000013399	2	Caballo Rata	2010208K18Rik
ENSGT00940000158020	2	Caballo Cerdo	GLDN
ENSGT00940000157291	2	Caballo Macaco	MTF1
ENSGT00940000161806	2	Caballo Cerdo	PEX10
ENSGT00390000008984	2	Caballo Cerdo	BBS12
ENSGT00940000155883	2	Caballo Zarigüeya	GNAO1
ENSGT00940000161738	2	Caballo Cerdo	DDX28
ENSGT00940000161603	2	Caballo Cerdo	ZNF398
ENSGT00940000158900	2	Caballo Ratón	EIk4
ENSGT00950000182951	2	Caballo Cerdo	FAM20B
ENSGT00940000157300	2	Caballo Zarigüeya	VTCN1
ENSGT00940000160669	2	Caballo Humano	ALX3
ENSGT00940000155665	2	Caballo Tití	SLC35D1

ENSGT00940000160722	2	Caballo Macaco	ZBTB39
ENSGT00390000011578	2	Caballo Cerdo	C5orf52
ENSGT00940000160366	2	Caballo Macaco	PRX
ENSGT00390000007946	2	Caballo Rata	Dhdh
ENSGT00950000182651	2	Caballo Cerdo	BAK1
ENSGT00940000162821	2	Caballo Cerdo	TSPYL1
ENSGT00940000157430	2	Caballo Macaco	SSH2
ENSGT00930000151062	2	Caballo Zarigüeya	ZNF408
ENSGT00390000012935	2	Caballo Ratón	1110051M20Rik
ENSGT00390000018647	2	Caballo Cerdo	CCDC87
ENSGT00940000159785	2	Caballo Cerdo	CCS
ENSGT00390000013104	2	Caballo Ratón	Tollip
ENSGT00940000163330	2	Caballo Cerdo	TRIM4
ENSGT00940000155267	2	Caballo Tití	PDPK1
ENSGT00940000157857	2	Caballo Macaco	NDST1
ENSGT00940000153259	2	Caballo Cerdo	TNFRSF19
ENSGT00940000153820	2	Caballo Cerdo	MSS51
ENSGT00940000155781	2	Caballo Macaco	RNF123
ENSGT00390000002177	2	Caballo Cerdo	C3orf84
ENSGT00940000155306	2	Caballo Cerdo	CCDC71
ENSGT00940000158548	2	Caballo Macaco	NKTR
ENSGT00390000012929	2	Caballo Zarigüeya	CCDC115
ENSGT00940000156332	2	Caballo Ratón	Epb4.1l5
ENSGT00390000011294	2	Caballo Humano	CCDC93
ENSGT00940000156417	2	Caballo Macaco	LRFN2
ENSGT00950000183202	2	Caballo Cerdo	RBL1
ENSGT00940000160596	2	Caballo Zarigüeya	MROH8
ENSGT00940000155597	2	Caballo Tití	DNAJC5
ENSGT00940000162596	2	Caballo Cerdo	IRF2BPL
ENSGT00510000048522	2	Caballo Cerdo	CIPC
ENSGT00950000182747	2	Caballo Cerdo	TRAP1
ENSGT00940000159636	2	Caballo Ratón	Pax5
ENSGT00390000018004	2	Caballo Cerdo	POLR1E
ENSGT00940000157936	2	Caballo Cerdo	NR6A1
ENSGT00950000182836	2	Caballo Ratón	Agpat3
ENSGT00940000159167	2	Caballo Rata	CHST11
ENSGT00940000156578	2	Caballo Macaco	TEF
ENSGT00390000016392	2	Humano Conejo	OTUD3
ENSGT00940000159902	2	Humano Macaco	TDRD5
ENSGT00390000017740	2	Humano Macaco	CAND2
ENSGT00940000157882	2	Humano Zarigüeya	NUDT3
ENSGT00390000004313	2	Humano Macaco	SAYSD1
ENSGT00390000011873	2	Humano Cerdo	GTF3C4
ENSGT00940000153385	2	Humano Cerdo	BRD3
ENSGT00950000183206	2	Humano Cerdo	HSPA8
ENSGT00940000155499	2	Humano Tití	B3GALTL
ENSGT00940000160367	2	Humano Ratón	TTBK2
ENSGT00390000003337	2	Humano Tití	CRAMP1L

ENSGT00390000007453	2	Humano Cerdo	IST1
ENSGT00940000156623	2	Humano Tití	USP22
ENSGT00440000033966	2	Humano Rata	CNTD1
ENSGT00940000160433	2	Humano Macaco	ST6GALNAC2
ENSGT00940000157831	2	Macaco Ratón	Pou2f1
ENSGT00940000156669	2	Macaco Cerdo	IL17RD
ENSGT00530000063965	2	Macaco Cerdo	WDR82
ENSGT00940000160677	2	Macaco Tití	EXOG
ENSGT00510000046470	2	Macaco Tití	TRAM2
ENSGT00950000182682	2	Macaco Ratón	Zfp41
ENSGT00940000153189	2	Macaco Cerdo	PPIL2
ENSGT00940000154184	2	Macaco Cerdo	KIAA1671
ENSGT00940000160998	2	Macaco Cerdo	FAM83F
ENSGT00940000158866	2	Macaco Tití	ACADS
ENSGT00390000017613	2	Macaco Cerdo	RETSAT
ENSGT00530000064712	2	Macaco Tití	TGOLN2
ENSGT00940000160208	2	Macaco Tití	BET1L
ENSGT00390000005534	2	Macaco Cerdo	RPL27A
ENSGT00390000006189	2	Macaco Cerdo	RRP8
ENSGT00940000160158	2	Macaco Tití	AMOTL1
ENSGT00500000044966	2	Macaco Ratón	Ccdc15
ENSGT00940000155147	2	Macaco Cerdo	SLC31A1
ENSGT00940000158571	2	Macaco Zarigüeya	GRIN3A
ENSGT00950000182721	2	Macaco Tití	FAM57A
ENSGT00390000018418	2	Macaco Ratón	Lsmd1
ENSGT00440000037582	2	Macaco Ratón	Cyb5d1
ENSGT00940000164244	2	Macaco Ratón	Gm20422
ENSGT00940000157319	2	Macaco Tití	Vps4a
ENSGT00390000008931	2	Macaco Zarigüeya	RFWD3
ENSGT00940000157543	2	Macaco Tití	KLHL36
ENSGT00940000157216	2	Tití Cerdo	RABGAP1
ENSGT00940000158365	2	Tití Ratón	Cbx7
ENSGT00950000183104	2	Tití Zarigüeya	CDC20B
ENSGT00940000159660	2	Tití Cerdo	LEF1
ENSGT00390000000489	2	Tití Cerdo	MLN
ENSGT00940000155913	2	Tití Conejo	TULP4
ENSGT00940000154206	2	Tití Zarigüeya	CHGB
ENSGT00950000182764	2	Tití Cerdo	STX16
ENSGT00940000159165	2	Tití Zarigüeya	ZNF512B
ENSGT00940000156179	2	Tití Ratón	Ankfy1
ENSGT00940000158444	2	Tití Zarigüeya	NUDCD3
ENSGT00940000158039	2	Tití Cerdo	KDM7A
ENSGT00950000182981	2	Tití Cerdo	AHCYL2
ENSGT00940000155449	2	Tití Ratón	Sbno1
ENSGT00390000000860	2	Tití Zarigüeya	FAM219B
ENSGT00940000155838	2	Tití Zarigüeya	MCAM
ENSGT00940000156393	2	Tití Cerdo	ZFP91
ENSGT00940000158103	2	Tití Cerdo	SLC3A1

ENSGT00390000015114	2	Tití Zarigüeya	LARS2
ENSGT00940000160590	2	Tití Ratón	Efcab2
ENSGT00510000048991	2	Tití Cerdo	IRGQ
ENSGT00950000182890	2	Tití Zarigüeya	ZNF883
ENSGT00390000004368	2	Tití Rata	PLD6
ENSGT00530000063801	2	Tití Cerdo	ATF4
ENSGT00390000017728	2	Ratón Rata	Cerdom
ENSGT00940000157226	2	Ratón Zarigüeya	ALKBH3
ENSGT00390000015634	2	Ratón Cerdo	Pofut1
ENSGT00530000063956	2	Ratón Zarigüeya	Spata2
ENSGT00730000111202	2	Ratón Zarigüeya	Zmat3
ENSGT00390000004798	2	Ratón Rata	Henmt1
ENSGT00910000144273	2	Ratón Zarigüeya	Wdr31
ENSGT00940000161580	2	Ratón Cerdo	Zscan20
ENSGT00390000001901	2	Ratón Zarigüeya	0610030E20Rik
ENSGT00950000182755	2	Ratón Cerdo	Zfp866
ENSGT00940000155273	2	Ratón Zarigüeya	Gan
ENSGT00950000182974	2	Ratón Cerdo	Aff4
ENSGT00390000005712	2	Ratón Cerdo	1200014J11Rik
ENSGT00440000033742	2	Ratón Cerdo	Rab11fip4
ENSGT00390000006544	2	Ratón Zarigüeya	Cep112
ENSGT00390000018085	2	Ratón Conejo	Rpl36al
ENSGT00400000022103	2	Ratón Cerdo	Ssr1
ENSGT00940000160106	2	Ratón Zarigüeya	Pou6f1
ENSGT00940000157804	2	Ratón Rata	Sp1
ENSGT00950000183107	2	Ratón Zarigüeya	Brwd1
ENSGT00950000182970	2	Ratón Rata	Ddx11
ENSGT00940000154421	2	Ratón Rata	RBM4B
ENSGT00940000162361	2	Ratón Zarigüeya	Zdhhc24
ENSGT00940000159574	2	Zarigüeya Rata	GMPR2
ENSGT00400000022326	2	Zarigüeya Rata	TINF2
ENSGT00940000160318	2	Zarigüeya Conejo	MYH15
ENSGT00940000158161	2	Zarigüeya Cerdo	RBM15
ENSGT00940000154951	2	Zarigüeya Rata	SENP2
ENSGT00940000159560	2	Zarigüeya Cerdo	SIN3B
ENSGT00890000139406	2	Zarigüeya Cerdo	WSB2
ENSGT00940000159279	2	Zarigüeya Cerdo	PLA1A
ENSGT00950000183031	2	Zarigüeya Cerdo	RRAGA
ENSGT00940000161946	2	Zarigüeya Cerdo	DCDC2B
ENSGT00390000003692	2	Zarigüeya Cerdo	C20orf194
ENSGT00940000159240	2	Zarigüeya Cerdo	TRIM66
ENSGT00940000156806	2	Zarigüeya Cerdo	Aptx
ENSGT00940000153578	2	Zarigüeya Cerdo	PPL
ENSGT00940000154389	2	Zarigüeya Cerdo	RANBP2
ENSGT00440000035348	2	Zarigüeya Cerdo	NUP50
ENSGT00940000155977	2	Zarigüeya Cerdo	KLHDC10
ENSGT00950000183040	2	Zarigüeya Cerdo	ARFIP2
ENSGT00390000000510	2	Cerdo Rata	GTF3C6

ENSGT00730000111224	2	Cerdo Rata	ELAC1
ENSGT00940000155689	2	Cerdo Conejo	NFYC
ENSGT00940000160086	2	Cerdo Rata	SETD2
ENSGT00530000063983	2	Cerdo Conejo	ZCCHC3
ENSGT00390000006954	2	Cerdo Conejo	DCTN4
ENSGT00950000182735	2	Conejo Rata	LPCAT3
ENSGT00390000000305	2	Conejo Rata	EMG1

Apéndice 2. Resultados de lo p-valores obtenido con el test exacto de Fisher de las comparaciones 2 a 2 y ajustado por Bonferroni.

	Humano	Macaco	Tití	Ratón	Rata	Conejo	Cerdo	Caballo	Perro	Zarigüeya
Humano	-									
Macaco	1.9117E-36	-								
Tití	4.15833E-28	6.52234E-32	-							
Ratón	4.09872E-20	2.48166E-21	3.71374E-18	-						
Rata	0.007139574	0.345222813	1.39514E-05	1.00643E-13	-					
Conejo	0.784119137	0.207472553	16.42442366	0.032976214	0.35736035	-				
Cerdo	2.16722E-15	4.76812E-22	2.90674E-14	2.99761E-12	0.010506636	0.03062598	-			
Caballo	8.23474E-26	3.15945E-33	2.30409E-19	7.40065E-22	0.001393431	1	4.74387E-38	-		
Perro	1.35756E-25	4.23367E-18	3.14082E-26	8.11523E-14	3.43463E-06	1	3.35963E-25	4.77203E-49	-	
Zarigüeya	0.841400072	0.228752324	1	1	0.120491362	1	0.001070352	0.004238675	2.25685E-08	-