

Evolución de elementos móviles en ortópteros



Universidad de Granada

Eugenia Elisabeth Montiel Jiménez

Tesis Doctoral

Marzo, 2012

Editor: Editorial de la Universidad de Granada
Autor: Eugenia Elisabeth Montiel Jiménez
D.L.: GR 2201-2012
ISBN: 978-84-9028-208-3

UNIVERSIDAD DE GRANADA

Facultad de Ciencias

Departamento de Genética



Evolución de elementos móviles en ortópteros

Memoria de Tesis Doctoral presentada por la Licenciada en Biología
Eugenia Elisabeth Montiel Jiménez para optar al grado de Doctor por la
Universidad de Granada.

Dirigida por las Doctoras:

Fdo. Dra. Josefa Cabrero Hurtado

Fdo. Dra. M^a Dolores López León

Granada, 2012

El presente trabajo se ha realizado en el Grupo de Genética Evolutiva del Departamento de Genética de la Universidad de Granada. La investigación ha sido financiada por el Ministerio de Ciencia e Innovación, a través de los proyectos CGL2006-06307 y CGL2009-11917 y del Plan Andaluz de Investigación (CVI-1664).

Durante la realización de este proyecto de Tesis Doctoral he disfrutado de una Beca Predoctoral incluida en el Proyecto de Excelencia con referencia P06-CVI-01664 de la Junta de Andalucía.

A los que creyeron en mí.
A mis padres y a mis hermanos.
A CQ

“If you know you are on the right track, if you have this inner knowledge, then nobody can turn you off... no matter what they say.”

“They thought I was crazy, absolutely mad.”

Barbara McClitock (Premio Nobel)

AGRADECIMIENTOS

La Tesis Doctoral, más que un mero proyecto científico, se convierte para los doctorandos en toda una experiencia vital que trasciende los límites del laboratorio. Con un espíritu inquieto como el mío, mi vida durante este periodo ha estado llena, no sólo de pipetas y electroforesis, sino también de retos, obstáculos, amistad y música. Muchas son las personas que de una forma u otra han contribuido a que me encuentre en este punto, y aunque no sean nombradas en estas líneas, quiero agradecerles, desde aquí, todo lo que me aportaron.

En primer lugar, me gustaría agradecer a mis directoras, las Dras. Josefa Cabrero y M^a Dolores López-León, todo el esfuerzo y tiempo que han dedicado a este proyecto. Ellas, con su sentido práctico, han conseguido en todo momento concretar las múltiples ideas que surgían en las reuniones de grupo, ahorrándome mucho tiempo con ello. Gracias Dra. Josefa Cabrero por enseñarme el gusto por el laboratorio y por mostrarme, que con tensión suficiente, es posible conseguir buenos resultados. Gracias Dra. M^a Dolores López-León, por haberme tutelado en todo momento, guiándome y acompañándome en este camino, obligándome a buscar la perfección del trabajo bien hecho.

Gracias también al Dr. Juan Pedro Martínez Camacho, que es director *de facto* de esta tesis. Gracias, porque ha sido para mí un ejemplo del científico que me gustaría ser, y porque sin su enorme creatividad y capacidad de relación no estarían aquí algunos de los experimentos, y muchos párrafos de discusión.

Gracias al Dr. Francisco Perfectti, que me descubrió lo dura de entender que llega a ser la genética, por confiar en mí desde el principio, brindándome su ayuda y amistad en todo momento.

Por otro lado, no puede faltar mi agradecimiento para todos los compañeros y amigos que han compartido mi día a día en el laboratorio. A mis compañeras, y ya doctoras, María Teruel e Inmaculada Manrique, que estuvieron en el comienzo de mi singladura y fueron, cada una a su manera, un ejemplo a tener en cuenta. Y a los que llegaron luego, a Mercy y María, por permitirme ayudarlas con el laboratorio y ayudarme ellas con los transposones; a Laura, que nos dejó por una vida mejor en el CSIC; a mi PacoR, por su franca y friki amistad; a Bea, porque se ha convertido en una gran amiga sin apenas notarlo; a Eva, ese torbellino que llegará tan lejos en la ciencia; a Leticia, con esa perseverancia y manejo del tiempo tan admirables; a Carmen y Xio, la siguiente generación y a las nuevas incorporaciones, el Dr. Mohammed Bakkali y Rubén. Gracias también a Ángel por su asesoramiento y risas, sobre y entre ordenadores; y a Ester y Geli, por compartir siempre un buen cotilleo. A los técnicos y becarios del departamento de

Ecología por los momentos de desconexión que me ofrecían, los chistes y los partidos de volley-playa.

A Tati, muchas gracias, porque durante mi última etapa, tu ayuda para mí ha sido fundamental, y no sólo en el laboratorio.

Y a mis dos leones, al malo y al bueno, Jesús y Abdel, que entre grillos, eminencias, cuchillos y cucharas me habéis hecho reír y llorar, siendo mi familia durante estos años. Seguiremos juntos, aunque estemos lejos.

Gracias a todo el Departamento de Genética de la Universidad de Granada, por ser mi lugar en el mundo durante estos años, haciendo mención especial, al grupo de Citogenética Molecular y Evolutiva en Mamíferos de este departamento por su apoyo en todos los problemas que ocurrieron mientras intentaba cuantificar mis transposones con el termociclador en tiempo-real.

Me gustaría también agradecer al Dr. Thomas H. Eickbush que me abriera las puertas de su laboratorio en Rochester (New York, USA) así como a todo su grupo que me hicieron sentir como una más durante el tiempo que permanecí con ellos.

Pero no todo fue trabajo de laboratorio, necesité muchas dosis de motivación extra que vinieron de mano de los amigos y la música, por ello quiero agradecer a mis amigos Rafa, Pe, Ignacio y M^a José las partidas de mus, las de Bang y esas maravillosas cenas y declaraciones de renta. A mis compañeros de la Coral Polifónica de la Facultad de Ciencias, por el hueco que he hicieron entre sus filas. Pero más aún, si cabe, quiero agradecer el apoyo que el Cuarteto Madrigal ha sido para mí. Esther, Bert y Antonio, gracias por arroparme siempre que lo he necesitado, por mostrarme ese cariño y amistad tan incondicionales, por hacerme sentir tan feliz cantando con vosotros, gracias.

Y sobre todo, quiero dar las gracias a los que siempre han creído en mí.

A ti, José María, gracias por quererme, gracias por levantarme cada vez que me caía, gracias por estar todos los días a mi vera.

A mi familia, a mis padres y hermanos. A Eva, Manuel y Julia, por aguantar a esta pesada hermana mayor. A mis padres, para los que soy doctora desde hace mucho tiempo, por el orgullo que sienten por mí, porque sin ellos hace tiempo que lo hubiera dejado, por vuestra confianza, por vuestro amor. A vosotros dedico esta Tesis Doctoral.

Resumen	21
Capítulo 1. Introducción general y Objetivos	27
Los elementos transponibles _____	29
Historia del descubrimiento de los elementos transponibles _____	30
<i>Los elementos controladores en el maíz</i> _____	30
<i>Elementos transponibles en bacterias</i> _____	31
<i>Elementos transponibles en animales: Elemento P y disgénesis híbrida</i> _____	33
Clasificación de los elementos transponibles _____	34
Abundancia en los genomas _____	39
Distribución cromosómica _____	41
Efectos de los elementos transponibles sobre el genoma hospedador _____	42
<i>Generación de diversidad alélica, transducción de genes, duplicaciones y barajamiento de exones</i> _____	42
<i>Efectos epigenéticos</i> _____	44
<i>Reordenaciones cromosómicas</i> _____	44
Evolución de los elementos transponibles y sus hospedadores _____	45
<i>Autoregulación de los TEs</i> _____	46
<i>Regulación mediada por el hospedador</i> _____	47
<i>Autónomos y no autónomos</i> _____	48
<i>Transmisión vertical y transferencia horizontal</i> _____	48
<i>Domesticación</i> _____	50
<i>Otras estrategias evolutivas</i> _____	52
Cromosomas B _____	53
Sistema de cromosomas B en <i>Eyprepocnemis plorans</i> _____	56
Sistema de cromosomas B en <i>Locusta migratoria</i> _____	57
Objetivos _____	59
Bibliografía _____	60

Capítulo 2. Gypsy, RTE and Mariner transposable elements populate *Eyrepocnemis plorans* and *Locusta migratoria* genomes. 75

Introduction	77
Materials and Methods	79
<i>Materials and DNA isolation</i>	79
<i>PCR amplification</i>	79
<i>DNA cloning and sequence analyses</i>	80
<i>RNA isolation and expression analysis</i>	81
<i>Copy number analysis</i>	81
<i>Fluorescence in situ hybridization (FISH)</i>	82
Results	83
<i>Presence of transposable elements: PCR amplification in E. plorans and DNA sequence analysis</i>	83
<i>Chromosome distribution of Transposable Elements</i>	86
<i>Expression of transposable elements</i>	86
<i>Copy number estimations</i>	87
Discussion	88
Acknowledgements	91
References	91

Capítulo 3. Preferential occupancy of R2 retroelements for B chromosomes in the grasshopper *Eyrepocnemis plorans* 99

Introduction	101
Materials and Methods	104
<i>Animals and cytological characterization of B chromosomes</i>	104
<i>DNA and RNA isolation</i>	104
<i>B chromosome microdissection</i>	104
<i>PCR amplification, DNA cloning and sequence analyses</i>	105
<i>Fluorescent in situ hybridization</i>	106
<i>R2 expression and 5' truncation profile</i>	107
<i>Copy number estimations</i>	108
Results	109
<i>Molecular characterization of R2 in E. plorans</i>	109
<i>Chromosome distribution of R2</i>	111
<i>Copy number quantification</i>	112
<i>R2 Expression and Activity</i>	114
Discussion	114
Acknowledgements	119
References	119

Capítulo 4. Conserved phylogenetic signal for two mobile elements in acridid grasshoppers	129
Introduction _____	131
Materials and methods _____	132
Results _____	136
Discussion _____	143
Acknowledgements _____	146
References _____	146
Capítulo 5. Analysis of intragenomic distribution of RTE retrotransposons through single-cell microdissection of individual bivalents	153
Introduction _____	155
Materials and methods _____	156
Results _____	157
Discussion _____	162
Acknowledgements _____	166
References _____	166
Capítulo 6. Discusión general	173
Bibliografía _____	189
Capítulo 7. Conclusiones	199
Capítulo 8. Perspectivas	205
Anexos	211
Anexo A. Tablas Suplementarias _____	213
Anexo B. Figuras Suplementarias _____	215
Anexo C. Símbolos y Abreviaturas _____	219
Anexo D. Alineamientos _____	221
<i>COI 5'</i> _____	221
<i>COI 3'</i> _____	222
<i>ND 5</i> _____	223
<i>Gypsy</i> _____	226
<i>RTE</i> _____	228

RESUMEN

La importancia de los elementos móviles o transponibles (TEs) como componentes del genoma y agentes de su evolución ha ido creciendo en las últimas décadas con el avance en su estudio. Su ubicuidad y la enorme fracción que representan en los genomas hasta ahora secuenciados, nos ha llevado a plantearnos el papel que juegan estos elementos en los genomas de ortópteros.

Para ello hemos analizado la presencia, en varias especies de ortópteros, de TEs que son abundantes en los genomas de otros insectos como ocurre con los elementos de Clase I, *Gypsy*, *RTE* y *R2*, y el elemento de Clase II, *Mariner*. La distribución y abundancia de los TEs difiere según el tipo de elemento y el organismo en el que se estudia. Así, los elementos móviles estudiados en el genoma de *Eyprepocnemis plorans* y *Locusta migratoria* han mostrado dos tendencias distintas. Por un lado, los elementos *Gypsy*, *RTE* y *Mariner* presentan una localización dispersa a lo largo de la eucromatina de todos los cromosomas A, pero con escasa presencia en la de los Bs, hallándose ausentes de las regiones heterocromáticas, los clústeres de ADN ribosómico y los de histonas. La otra tendencia es la encontrada en el elemento *R2*, cuya presencia está limitada al ADN ribosómico, debido a su mecanismo de retrotransposición, pero que muestra una acumulación en los clústeres ribosómicos del cromosoma B y una escasa, o incluso nula, presencia en los clústeres ribosómicos de los cromosomas A de la especie *E. plorans*. El elemento *R2* en *E. plorans*, *EploR2*, fue caracterizado completamente, observando que su estructura es similar a la encontrada en los *R2* de otras especies ya que posee una longitud de 3,5 kb y un único ORF que codifica tres dominios (unión al ADN, retrotranscriptasa y endonucleasa) compartiendo la diana de inserción en los genes 28S ribosómicos y careciendo de las delecciones que este gen presenta en otras especies

Por otra parte, la abundancia de estos elementos, estudiada por estimación de número de copias mediante PCR cuantitativa, muestra diferencias entre los tipos de TEs, siendo los elementos *Gypsy* los más abundantes en *E. plorans* (~7000 copias), seguidos por los elementos *RTE* (~4200), *Mariner* (~350) y *R2* (~ 100). En *L. migratoria*, fue el elemento *Mariner* el que mostraba una mayor presencia con valores similares de los de *E. plorans*, mientras que los elementos *Gypsy* y *RTE* presentaron un número de copias extremadamente escaso debido probablemente a la especificidad de los cebadores diseñados a partir de los elementos encontrados en *E. plorans*. Aún así, tanto la distribución observada mediante FISH como las copias estimadas, sugieren que los TEs suponen una proporción importante de estos genomas.

La acumulación del elemento R2 en la variante de cromosoma B, B₂₄, de *E. plorans*, nos llevó a preguntarnos si ocurriría igual en las diferentes variantes observadas de este cromosoma en otras poblaciones, por lo que estimamos el número de elementos R2 en individuos (0B y 1B) de distintas poblaciones, cuantificando además, el número de unidades de ADN ribosómico para analizar la dinámica de ocupación de este elemento. De media, los individuos portadores de cromosoma B poseían 10 veces más copias de R2 que los individuos sin este cromosoma. De igual forma, las unidades de ADN ribosómico presentes en el cromosoma B fueron tantas como las presentes en el complemento normal de esta especie, apareciendo diferencias entre las distintas variantes de cromosoma B. La ocupación por R2 del 2% de los clúster de ADN ribosómico, fue más baja que en otros insectos, mostrando también diferencias entre individuos con y sin cromosoma B y entre las variantes de Bs, siendo mayor en la variante de B más antigua, B₁, con un 8%, que en las más modernas, B₂ y B₂₄, con alrededor del 5%. A la luz de estos datos, se desprende que el elemento R2 parece estar fundamentalmente ligado al cromosoma B, donde encuentra un lugar de inserción paradisiaco debido a la gran cantidad de dianas disponibles, con alta tolerancia a las inserciones por su escasa repercusión en el hospedador, y debido a la ausencia de recombinación con los cromosomas A lo que dificulta la eliminación de copias.

Sin embargo, puesto que la actividad de R2 se encuentra ligada a la expresión del gen ribosómico 28S, la inserción en los cromosomas B (donde los clústeres ribosómicos están usualmente inactivos) no favorece su actividad. De hecho, no hemos conseguido amplificar transcritos de R2 en el ADN complementario de machos de Torrox portadores de B₂₄, a pesar de que en esta población se produce con mucha frecuencia la activación de la NOR del B₂₄. Este hecho contrasta fuertemente con las evidencias de actividad obtenidas mediante el análisis de las truncaciones 5' de este elemento, que muestra que las poblaciones con cromosomas B tienen mayor número de eventos de retrotransposición. Estos datos sugieren que o bien la expresión de R2 ocurre en un tejido o fase del desarrollo distinto al estudiado aquí, o bien la actividad observada se debe a un pasado reciente coincidente con el estrés celular provocado durante la fase parasítica del B. Por otro lado, el análisis de la expresión de los otros TEs estudiados sí dio resultados positivos, encontrando transcritos en ADN complementario de *E. plorans* y *L. migratoria* tanto de *Gypsy*, como *RTE* y *Mariner*, aunque no podemos asegurar que exista actividad transposicional porque aún podrían afectarles la regulación post-transcripcional que pueda ejercer el genoma hospedador o las propias copias defectivas de estos elementos.

La reconstrucción de la historia evolutiva de los TEs en ortóperos fue realizada estudiando la relación filogenética de los retroelementos, *Gypsy* y *RTE*, y tres marcadores del ADN mitocondrial en 31 especies de la familia Acridoidea. Ambos elementos móviles fueron amplificados en la mayoría de las especies y la topología de sus árboles filogenéticos mostraron una congruencia de entre el 67-72% con el árbol filogenético obtenido con el ADN mitocondrial, por lo que estos elementos parecen ser antiguos componentes de estos genomas que han evolucionado junto con sus hospedadores. Pero como muestran los porcentajes, esta congruencia no es completa, encontrándose casos que sugieren la existencia de eventos de transferencia horizontal de transposones (HTT), como la alta relación mostrada entre los elementos *Gypsy* de las especies *E. plorans* y *H. littoralis*, que es mayor que la presentada con las otras especies de su mismo género, o bien la relación entre los *RTE* de *Ch. vagans* y el género *Dociostaurus*. Un análisis de la diversidad de copias de *Gypsy* en el genoma de *E. plorans*, sin embargo, puso en duda esta HTT puesto que se encontraron copias similares tanto al elemento *Gypsy* de *E. unicolor* como al de *H. littoralis*, por lo podríamos encontrarnos ante un efecto de “lineage sorting” incompleto o incluso de muestreo incompleto.

Para conocer la dinámica intragenómica de estos elementos realizamos un estudio de la diversidad del elemento *RTE* dentro del complemento cromosómico de *E. plorans*, mediante la microdissección por separado de todos los bivalentes (más el cromosoma B) de una célula meiótica y la posterior amplificación, clonación y secuenciación de las copias de *RTE* en cada uno de los cromosomas. Este abordaje nos permitió obtener una perspectiva filogeográfica e intragenómica de la evolución de estos elementos, al tratar cada cromosoma como una población. De este modo, pudimos establecer la existencia de dos linajes de elementos *RTE* en *E. plorans*, uno presente en todos los cromosomas, incluido el B (linaje B), y otro presente en todos los A, excepto el cromosoma S_{11} y el B (linaje A), sugiriendo la existencia de dos oleadas de invasiones de estos elementos, una por cada linaje y anteriores al surgimiento del cromosoma B. Además, la alta relación entre las secuencias del cromosoma B y S_{11} sugiere el origen, a partir del S_{11} , de este cromosoma accesorio.

Finalmente, el análisis filogenético de las secuencias de estos linajes junto a las secuencias de *RTE* de otras especies, obtenidas en el análisis anterior, nos mostró la posibilidad de una HTT desde oedipodinos que originara este linaje B. Esta HTT que sólo podría ser corroborada con un estudio similar de diversidad en los otros eyprepocnemidinos que descartara la presencia del linaje B en ellos.

CAPÍTULO 1

Introducción general y Objetivos

En este capítulo vamos a realizar una revisión del conocimiento acumulado sobre un tipo de elemento genético egoísta: los elementos transponibles. Estos elementos fueron llamados en un principio con el apelativo de “genes saltadores” puesto que pueden cambiar de localización genómica gracias a su propia maquinaria enzimática o a maquinaria ajena, produciendo un acusado efecto en los genomas que pueblan. Durante mucho tiempo han sido considerados únicamente como egoístas, causantes de mutaciones con su movimiento, pero actualmente el hallazgo de numerosas pruebas de su domesticación y contribución a procesos celulares ha obligado a una revisión de su carácter parasítico.

En saltamontes, no existen muchos estudios que analicen la presencia de elementos transponibles pero, debido al intenso estudio citogenético realizado sobre este grupo, sí existe abundante literatura sobre otro elemento genético egoísta, los cromosomas B. En la composición molecular de estos cromosomas B, se ha descrito la presencia de elementos móviles, mostrando una interesante posibilidad para el estudio de la relación existente entre transposones y dos sistemas de cromosomas B muy estudiados como son los de los ortópteros, *Eyprepocnemis plorans* y *Locusta migratoria*. Así en este capítulo también realizaremos una breve descripción de estos cromosomas accesorios y de los sistemas de cromosomas B presentes ambas especies.

Los elementos transponibles

Los elementos transponibles se definen como secuencias de ADN capaces de cambiar de localización genómica, generando frecuentemente nuevas copias de si mismos, lo que les permite extenderse por los genomas. Fueron descubiertos por primera vez en plantas, pero hoy en día sabemos que son componentes genómicos de la mayoría de los eucariotas. Su propia capacidad para transponerse de forma replicativa, además de la posibilidad de incrementar su número de copias a través de mecanismos indirectos, dependientes del genoma hospedador, hace que sean considerados elementos egoístas (Doolittle & Sapienza 1980; Orgel & Crick 1980; Craig et al. 2002). De hecho, Richard Dawkins comentaba ya en su libro “El Gen Egoísta” (1989) que la última edición del Oxford English Dictionary introducía una nueva acepción del término “egoísta” relacionada con lo que se consideraba en ese momento “ADN egoísta” y que englobaba a los elementos transponibles. Actualmente, los elementos transponibles cuentan con entradas en 27 idiomas distintos en Wikipedia (<http://es.wikipedia.org/>) y el término “transposable” genera resultados que superan las miles de decenas en buscadores como Web of KnowledgeSM o Google Scholar (<http://scholar.google.com/>).

El interés científico sobre los elementos móviles o transponibles ha cambiado mucho a lo largo de los últimos cuarenta años, pasando del más profundo ostracismo a ser considerados piezas fundamentales en la estructura y evolución genómica. Los elementos móviles se encuentran implicados en las investigaciones más vanguardistas de la ciencia actual, ya sea desde una perspectiva puramente evolutiva, a través del análisis filogenético de sus secuencias y su interacción con los genomas hospedadores, como bioquímica, mediante el estudio de los mecanismos que posibilitan su transposición y la similitud de su estructura con virus o factores de transcripción, o incluso, desde una perspectiva práctica, debido a la dificultad de su anotación en la secuenciación de genomas o su uso como vectores en técnicas de ingeniería genética.

Historia del descubrimiento de los elementos transponibles

El protagonismo que los elementos transponibles tienen en las investigaciones actuales ha sido difícil de conseguir. Por ello, es interesante hacer un pequeño repaso de los hitos que han marcado el desarrollo de nuestro conocimiento de los transposones a través de la historia.

Los elementos controladores en el maíz

Barbara McClintock recibió en 1983 el premio Nobel de Medicina “*por su descubrimiento de los elementos genéticos móviles*”

McClintock descubrió los elementos transponibles (“genes saltadores”) mientras trabajaba con el maíz (*Zea mays*) en los años 40. Su trabajo se basaba principalmente en el análisis fenotípico de los granos, donde la variegación observada no podía explicarse por las leyes mendelianas de la herencia y se hallaba asociada a la rotura de cromosomas. Las observaciones de este trabajo indicaron que algunos eventos de rotura siempre se producían en la misma región cromosómica, por lo que McClintock asumió que esos eventos eran producidos por un elemento genético llamado *Ds* (*Dissociator*) que se movía. En investigaciones posteriores ella dedujo que la inestabilidad del elemento *Ds*, causante de roturas cromosómicas, era dependiente de la presencia de otro tipo de elemento denominado *Ac* (*Activator*). Estos elementos fueron llamados inicialmente “elementos controladores” puesto que alteraban la función de los genes vecinos. En una serie de avanzados experimentos realizados entre 1948 y 1951, McClintock mapeó diferentes

familias de estos elementos controladores que afectaban no sólo al patrón de pigmentación sino también a otras propiedades del grano (Fedoroff, 1993). Estos hallazgos la llevaron a sugerir que esos elementos genéticos móviles probablemente estaban presentes no sólo en plantas sino también en insectos y animales superiores.

El pionero trabajo de McClintock, que ocupa un puesto destacado en la historia de la genética, fue ignorado durante 25 años por la comunidad científica hasta que la existencia de los elementos transponibles fue confirmada por clonación molecular a mitad de los años 1980 (Fedoroff et al. 1983). Su trabajo simplemente no fue comprendido por sus contemporáneos cuando fue presentado por primera vez en el Cold Spring Harbor Symposium en 1951 (McClintock 1951), pareciendo extravagante e, incluso, herético. La visión del genoma en aquellos momentos consistía en la de una cadena estable de genes ordenados de un extremo al otro del cromosoma que evolucionaban lentamente por la acumulación de mutaciones aleatorias a frecuencias constantes (Miller & Capy 2004). Por lo que no es sorprendente que pocos genetistas estuvieran preparados para aceptar que los genes pudieran saltar de la forma irresponsable que Barbara McClintock proponía para los elementos controladores.

Las técnicas moleculares revelaron que el sistema *Ac-Ds* está compuesto por una copia autónoma (*Ac*) de unas 4,5 kb de longitud que contiene un gen de transposasa y cuyos extremos tienen repeticiones invertidas de 11 bp. El elemento no autónomo (*Ds*) presenta las mismas repeticiones invertidas pero con una delección interna que inactiva el gen de la transposasa (Fedoroff 1989). Aunque el trabajo genético de McClintock fue la primera prueba clara de que la existencia de elementos de ADN móviles servían como una fuente genética para la plasticidad de los genomas, se necesitaron más de 30 años para que su concepto de la dinámica del genoma fuera aceptada de forma general (Kidwell & Lisch 2001).

Elementos transponibles en bacterias

En procariotas, probablemente debido al pequeño tamaño de su genoma, el estudio de los elementos transponibles se realizó pronto, datando los primeros estudios de finales de los años 60. Su presencia varía entre un 0 y un 21%, encontrándose inicialmente dos tipos de elementos: las secuencias de inserción (*IS*) y los transposones (*Tn*). Las secuencias de inserción (*IS*) consistían en pequeñas secuencias de ADN (0,5-2,5 kb), fenotípicamente

crípticas y con genes relacionados con la inserción (Mahillon & Chandler 1998), mientras que los elementos *Tn* presentaban una mayor complejidad, siendo generalmente de mayor longitud y conteniendo a veces elementos *IS* así como otros genes no relacionados con el mecanismo de inserción.

Las secuencias *IS* fueron descubiertas como mutaciones altamente polares e identificadas como inserciones en el cromosoma bacteriano (Shapiro 1969) y en bacteriófagos (Brachet et al. 1970). Muchas de estas mutaciones mostraban ser inserciones de los mismos pequeños fragmentos de ADN pero en diferentes posiciones y orientaciones. Cuando se encontró que estas pequeñas secuencias eran residentes naturales del genoma de *Escherichia. coli*, se reconoció su parecido con los elementos transponibles del maíz descubiertos por Barbara McClintock.

Por otro lado, se encontró que el determinante de resistencia a antibióticos encontrado en este factor R se encontraba en un elemento transponible (*Tn*) que mostraba propiedades muy similares a las de las secuencias *IS* (Berg et al. 1975) con lo que los elementos *Tn* pasaron a ser de una elevada importancia médica. Por ejemplo, los miembros de la familia *Tn3* son usualmente encontrados en plásmidos de bacterias resistentes a antibióticos, pero que pueden transponerse a bacteriófagos y al cromosoma de *E. coli* y otras bacterias (Datta et al. 1971, Heffron et al. 1979, Arthur et al. 1984).

Las secuencias *IS* así como los *Tn* eran de tal complejidad que pronto fue necesario establecer un sistema de nomenclatura y una clasificación (Campbell et al. 1979; Mahillon & Chandler 1989).

Otro tipo particular de elementos móviles bacterianos descritos más recientemente son los intrones autocatalíticos del grupo I y II (Edgell et al 2000; Lambowitz & Zimmerly, 2004). La abundancia e importancia de estos elementos en procariontes contribuyó al cambio de concepción del genoma como una estructura fija a otra más plástica capaz de reaccionar ante los factores externos.

Elementos transponibles en animales: Elemento P y disgénesis híbrida

A pesar de los descubrimientos de Barbara McClintock y la demostración de la presencia de elementos móviles en bacterias, para la Ciencia de la época seguía siendo difícil de asumir la posibilidad de que los elementos transponibles pudieran influir en los polimorfismos genéticos, y en la diversidad genética. El pensamiento que persistía dictaba que aunque se diera el caso de que estos elementos se movieran, no lo harían en una frecuencia significativa como para contribuir a la diversidad genética del genoma.

Sin embargo, la década de los 70 presencié la emergencia del estudio en *Drosophila melanogaster* del fenómeno de la disgénesis híbrida (Kidwell et al. 1977), en virtud del cual los cruces entre líneas específicas de *D. melanogaster* producía cambios genéticos, que incluían esterilidad en uno o ambos sexos, e incremento de la proporción de mutaciones y reordenaciones cromosómicas. Años después, este fenómeno fue finalmente asociado a la movilización de unos elementos transponibles específicos: el elemento *P* (sistema *P/M*) y el elemento *I* (sistema *I/R*). El elemento *P* mostró ser un transposon de ADN (elementos que se transponen vía intermediario de ADN), y el elemento *I*, un retrotransposón sin largas repeticiones terminales (LTR) (elementos que se transcriben vía intermediario de ARN).

La disgénesis híbrida producida por el sistema *P-M* ocurría cuando machos de la línea *P* se cruzaban con hembras pertenecientes a la línea *M*, pero no se producía en la progenie del cruce recíproco. En 1979, Engel estableció el paradigma para los siguientes estudios sobre este fenómeno al proponer que un determinado factor genético interaccionaba con un determinado “citotipo” materno (Engel 1979). Este autor propuso que el citotipo de las hembras de la línea *M* era permisivo a la actividad del factor cromosómico *P*, que en las moscas disgénicas era heredado de los padres de la línea *P*. El citotipo *P*, en cambio, no era permisivo y de ahí se concluía la naturaleza no recíproca de esta disgénesis. Finalmente, Engel propuso que el factor *P* debía ser un elemento transponible. La caracterización molecular del elemento *P* a partir de las inserciones de secuencias repetidas en el gen *white* presente en moscas de la progenie de cruces *P-M* (Rubin et al. 1982), confirmó las afirmaciones sobre la naturaleza transponible de este elemento. Pero el corazón de la disgénesis híbrida realmente es la rotura del férreo control de la transposición que ejerce la célula hospedadora, como se demuestra con el descubrimiento de la represión producida por una forma truncada de la propia transposasa del elemento *P* (Misra & Rio 1990) así como de pequeñas secuencias de ARN *piwi* (ARN_{pi}) que pueden silenciar a los elementos (Vagin et al 2006; Jensen et al 2008), y que explicarían la no reciprocidad de los cruces en la disgénesis híbrida del sistema *P-M*.

Actualmente, la importancia de los elementos transponibles en la evolución de los genomas que quedado patente con la llegada de la era genómica y post-genómica. El desarrollo de programas de secuenciación así como el creciente interés por estos elementos ha permitido que en las últimas décadas se hayan descubierto y caracterizado numerosas familias de elementos móviles, revelando una gran diversidad natural en cuanto a su abundancia y distribución entre los genomas estudiados.

Además, el descubrimiento de la implicación de estos elementos en la formación de secuencias reguladoras y génicas, y su papel clave en los procesos celulares, la respuesta genómica ante el estrés o la transferencia lateral de genes, como hemos visto en procariontes, ha hecho que el concepto de elemento transponible se haya extendido para considerar un escenario posible continuo, desde el parasitismo hasta el mutualismo o la simbiosis, en las interacciones de los elementos móviles y las especies que habitan.

Clasificación de los elementos transponibles

Los estudios sobre los elementos transponibles no han hecho más que crecer desde su descubrimiento, aumentando a su vez la dimensión del problema que supone su clasificación en una estructura clara y jerarquizada que permita encuadrar la diversidad de estos elementos. A este problema ya se enfrentaron los autores dedicados al estudio de los elementos transponibles en procariontes (ver *Elementos transponibles en bacterias*) pero en eucariotes la clasificación tardó algo más en llegar. El conocimiento, cada vez más profundo, de la estructura molecular de los elementos móviles y de su forma de moverse en los genomas hizo finalmente que Finnegan en 1989 propusiera una primera clasificación de los elementos transponibles eucarióticos. Así, estos elementos se dividían en dos clases según su mecanismo de transposición: los elementos de Clase I que se transponen mediante la retrotranscripción de un intermediario de ARN y los elementos de Clase II que se transponen directamente de ADN a ADN. Además, Finnegan contemplaba dentro de la Clase I, dos subclases: los elementos estructuralmente similares a los retrovirus, con largas secuencias repetidas terminales, o elementos LTR y los que no poseían estas repeticiones sino una característica secuencia rica en adenina en posición 3' y que no eran semejantes a los retrovirus, a los que denominó elementos no-LTR. Esta primera dicotomía fue rápidamente aceptada por los investigadores, pero el número de elementos transponibles descritos creció de forma tan rápida que fue necesario hacer una redefinición de la misma.

Una clasificación más detallada y basada en la de Finnegan, fue propuesta por Capy et al. (1997) donde se hacía una síntesis de los términos usados por diferentes autores. En la Clase I se establecían dos subclases, la Subclase I que contenía a los elementos LTR o retrotransposones y la Subclase II con los elementos no-LTR o retroposones. A su vez, en la Subclase I se definían cuatro superfamilias (*Gypsy*, *Ty3*, *Tf1* y *Ty1*) en función de la presencia de características tales como el orden de los dominios en el gen *pol*, la existencia de un cebador de ARN y/o la existencia del gen *env*. La Subclase II se dividía en dos superfamilias dependiendo de la presencia o ausencia de un región codificante: la Superfamilia LINE (con región codificante y autónomos en su transposición) y la Superfamilia SINE (sin región codificante y no autónomos).

Por otro lado, una clasificación similar a la empleada por los virólogos surgió para clasificar a los elementos LTR (Hull 2001) mientras que una clasificación por clados fue sugerida para los elementos no-LTR mediante el análisis filogenético del dominio de retrotranscriptasa de estos elementos (Malik et al. 1999).

Finalmente, en 2006 se constituyó, durante la primera conferencia internacional sobre elementos transponibles, un Comité Internacional para la Clasificación de los Elementos Transponibles (Kapitonov & Jurka, 2008) encargado de establecer una clasificación actualizada (Tabla 1) pero que conservara la estructura principal de la propuesta por Finnegan. Según esta clasificación todos los elementos transponibles eucariotas se dividen en dos tipos (retrotransposones y transposones de ADN) que están compuestos por cinco clases principales: retrotransposones con largas repeticiones terminales (LTR), retrotransposones no-LTR, transposones de ADN con el mecanismo *escisión/reparación*, transposones de ADN con el mecanismo de *círculo rodante* (*Helitrons*) y transposones de ADN con auto-sintetizamiento (*Politons* o *Mavericks*). Además, cada clase está compuesta por un pequeño número de superfamilias, y éstas, a su vez, por numerosas familias de elementos transponibles.

Los elementos transponibles de la clase de los retrotransposones LTR constan de regiones parcialmente solapadas de tres diferentes genes: *gag*, proteasa *prt* y polimerasa *pol*; además en ocasiones existe un cuarto dominio *env* (Figura 1). El gen *pol* contiene dominios para una retrotranscriptasa (RT), una RNasaH, y una endonucleasa (EN). Las regiones LTR no codifican para ninguna proteína conocida pero contienen promotores y terminadores de la transcripción. Las proteínas traducidas son sintetizadas como una poliproteína que será cortada en péptidos funcionales por la proteasa. El gen *gag* facilita

Tabla 1. Clasificación universal de elementos transponibles implementada en RepBase y propuesta por el Comité Internacional de Clasificación de Elementos Transponibles (Kapitonov & Jurka 2008).

Tipo 1: Transposones de ADN	Tipo 2: Retrotransposones	
Superfamilia	Retrotransposones no-LTR Superfamilia (Clado)	Retrotransposones LTR Superfamilia (Clado)
Chapaev	CRE	Copia
En/Spm (CACTA)	NeSL	Gypsy
hAT	R2	Bel
Harbinger (Pif)	R4	ERV1
ISL2EU (IS4EU)	L1	ERV2
Kolobok	RTE	ERV3
Mariner	Jockey	
Merlin	CR1	DIRS
Mirage	Rex1	
MuDR (MULE)	I	
Novosib	RandI(Dualen)	
P	Tx1	
PiggyBac	SINE1	
Rehavkus	SINE2	
Transib	SINE3	
Helitron	Penelope	
Politons (Maverick)		

la maduración y empaquetamiento de los retrotransposones de RNA y las proteínas para integrarse en el genoma. La enzima retrotranscriptasa RNasaH es requerida para la replicación-transposición del transposón, mientras que EN codifica una endonucleasa que permite al DNA del retrotransposón insertarse en una nueva localización cromosómica.

Los retroposones que carecen de LTR tienen una estructura formada por una región 5' no traducida (UTR) que actúa como promotor, uno o dos marcos abiertos de lectura (ORF1 y ORF2) separados por una región espaciadora, y una región UTR 3' con una cola de poliA (pA). Dentro de la clase también podemos diferenciar superfamilias, también llamadas clados, agrupadas filogenéticamente según la secuencia de su retrotranscriptasa (Malik et al. 1999; Kapitonov et al. 2009) y que comparten, además, las mismas características estructurales como el tipo de endonucleasa (similar a enzimas de restricción, RLE o apurínica-apirimidínica, APE), la presencia de ribonucleasa H o diferentes motivos como dedos de zinc o de reconocimiento del ARN.

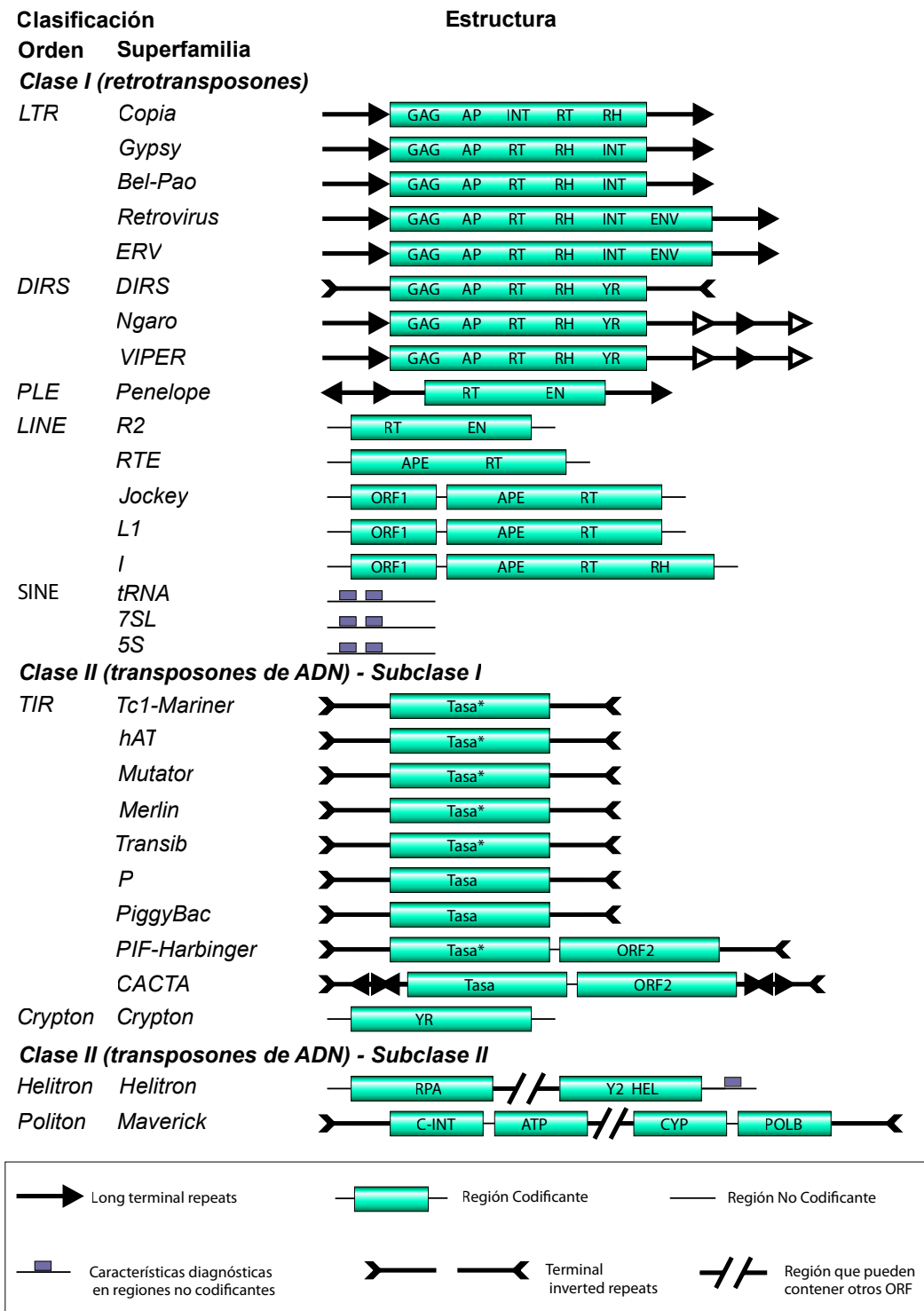


Figura 1. Sistema de clasificación propuesto por Wicker et al. (2007). AP, aspartil-proteasa; APE, endonucleasa apurínica; ATP, ATPasa; C-INT, C-integrasa; CYP, cisteína-proteasa; EN, endonucleasa; ENV, proteína env; GAG, proteína de la capsid; HEL, helicasa; INT, integrasa; ORF, ORF de función desconocida; POLB, ADN polimerasa B; RH, RNasa H; RPA, proteína de replicación A; RT, retrotranscriptasa; Tasa, transposasa (* con motivo DDE); YR tirosina-recombinasa; Y2, YR con motivo YY

Los elementos transponibles de ADN que se transponen mediante *escisión/reparación* están formados por un gran ORF en el que está incluido un dominio de reconocimiento y unión al DNA y otro dominio catalítico, todo ello flanqueado por repeticiones invertidas terminales cortas (TIR) (Figura 1). Una de las principales familias de esta clase es la familia *Tc1-mariner*, presente en multitud de órdenes de insectos (Robertson 1993; Robertson & Lampe 1995), otros invertebrados (Robertson 1995), plantas (Jarvi & Lark 1998) y mamíferos (Oosumi et al. 1995). Los elementos móviles mediante *círculo rodante* (*Helitrons*) poseen varios ORFs que incluyen una helicasa, proteínas iniciadoras de replicación y proteínas de unión a DNA de cadena simple, todo ello flanqueado por un dinucleótido TC 5' y un bucle más un dinucleótido CT 3' (Figura 1) (Kapitonov & Jurka 2007). Por último, los elementos transponibles llamados *Politons* son los elementos más complejos encontrados con una longitud que varía entre las 15-25 kb que codifica más de 10 proteínas diferentes que incluyen una ADN polimerasa B (POLB), un integrasa retroviral, una proteasa adenoviral (PRO) y una posible ATPasa (ATP) (Kapitonov & Jurka 2006).

Si bien es cierto que esta clasificación es la más aceptada actualmente, en 2007 se publicó una clasificación que permitía a los no expertos una fácil aplicación teniendo en cuenta la ingente cantidad de datos obtenidos en los diferentes proyectos de secuenciación de genomas (Wicker et al. 2007). Esta publicación fue rápidamente respondida por Kapitonov & Jurka (2008) comentando que en la base de datos RepBase (Jurka et al. 2005) ya estaba implementada una clasificación consensuada por la comunidad científica dedicada al estudio de los elementos transponibles. Sin embargo, la nueva clasificación propuesta por Wicker et al. (2007), aunque no contempla toda la diversidad presente en estos elementos y no refleja en algunos puntos las relaciones filogenéticas existentes, si es una buena herramienta para una primera aproximación al estudio de los elementos transponibles por la claridad de su estructura y es la que emplearemos en esta introducción (Figura 1).

Otra clasificación sería posible si atendiéramos a la capacidad de movimiento de los elementos móviles, de forma que pueden considerarse elementos autónomos, si codifican las enzimas necesarias para su propio movimiento y elementos no autónomos, si dependen de los primeros, ya sean de su misma familia o no, para moverse (Kidwell & Lisch 2001). Los elementos *MITE* son un ejemplo de este segundo grupo siendo elementos que tiene un tamaño pequeño (100-600 bp) y que están flanqueados por repeticiones

terminales inversas (TIRs), siendo encontrados frecuentemente cercanos a genes. En muchas especies, son los elementos *Mariner* los responsables del origen y movimiento de grandes poblaciones de estos elementos no autónomos, como los elementos *MITE Stowaway* presentes en arroz (Bureau & Wessler 1994).

Otro tipo diferente de elementos móviles que aparecen fundamentalmente en eucariotas inferiores y en orgánulos citoplasmáticos que son los intrones móviles de los grupos I, II y III, presentes en el ADN ribosómico, ADN de transferencia y en el de algunos otros genes (Davies et al 1982; Abelson et al. 1998; Christopher & Allic 1989)

Abundancia en los genomas

Es usual decir que los elementos transponibles se encuentran en todos los organismos vivientes. Pero lo cierto es que al menos dos excepciones han sido descritas. La primera corresponde al genoma de una línea de *Bacillus subtilis* en la que no se ha encontrado ninguna secuencia *IS* (Kunst et al. 1997). El otro ejemplo es el de *Plasmodium falciparum* cuyo genoma se encuentra libre de cualquier elemento transponible (Gardner et al. 2002). El resto de organismos conocidos presentan elementos transponibles de alguna clase y su abundancia puede ser cuantificada según los siguientes criterios: por el número de familias que presentan (diversidad), por el número de copias de una familia dada o por el porcentaje del genoma que ocupan estos elementos. En base a este último criterio, se observa una gran variación dependiendo de la especie en consideración (Tabla 2), existiendo una clara correlación entre el tamaño del genoma y la abundancia de elementos transponibles, (Kidwell 2002; Hua-Van et al. 2005). Si nos referimos a la proporción relativa de cada clase de elementos móviles, son los de la Clase I, retrotransposones LTR y no-LTR, los que generalmente se presentan en mayor proporción que los de la Clase II o transposones de ADN, siendo los retrotransposones los máximos responsables tanto del aumento como de la disminución del tamaño de los genomas (Boulesteix et al. 2006; Vitte & Bennetzen 2006; Sen et al. 2006). La distribución de las familias es muy variable entre los genomas estudiados mostrándose predominante en una especie una determinada familia, como posible consecuencia de una reciente amplificación, mientras que en otras especies la misma familia de elementos puede encontrarse ausente. En el caso de la especie *Saccharomyces cerevisiae*, los elementos *LINE* se encuentran ausentes mientras que en el genoma humano son la principal familia. Entorno al 40-45% del ADN humano corresponde a elementos transponibles, la mayoría de este porcentaje corresponde a retroelementos no-LTR representados por pocas familias, siendo dos de ellas las más representativas por su elevado número de copias: la familia *L1 (LINE)* y la

familia *Alu* (*SINE*). En el genoma del maíz, también la Clase I es predominante sobre la Clase II, consistiendo los primeros en casi el 50% del genoma con más de 20 familias de retrotransposones LTR (SanMiguel et al. 1996). También en insectos la Clase I es más abundante. En el genoma del mosquito *Anopheles gambiae* los elementos más importantes en número de copias son los elementos LTR y *SINE* (Holt et al. 2002). En *Drosophila*, los elementos transponibles comprenden del 4-5% del genoma eucromático, siendo destacable la ausencia de elementos no autónomos como los *SINE* y *MITE*, y la presencia de elementos *Foldback* (*FB*), pero siendo los retroelementos LTR los más abundantes, tanto en heterocromatina como en eucromatina, en términos de porcentaje (Kaminker et al. 2002). Pero no en todos los casos los retrotransposones son predominantes. Hay casos de genomas que han sido invadidos por elementos *MITE*, una forma no autónoma de algunas familias de transposones de ADN (Feschotte et al. 2002), como son los genomas de *Oriza sativa*, *Arabidopsis thaliana* (Le et al. 2000) y *Caenorhabditis elegans* (Duret et al. 2000).

Una imagen completamente distinta a la anterior es la encontrada en eucariotas unicelulares como *S. cerevisiae* donde no se han encontrado transposones de ADN y el número de familias de retrotransposones es extremadamente bajo, con solo cinco familias de elementos *Ty* con escaso número de copias (Kim et al. 1998).

Tabla 2. Proporción de TE y tamaños de genomas de 13 especies (Kidwell, 2002; Hua-Van et al. 2005).

Especies	Genoma (Mb)	% TE	Referencias
<i>S. cerevisiae</i>	12	3	(Kim et al., 1998)
<i>D. discoideum</i>	34	10	(Glockner et al., 2001)
<i>C. elegans</i>	100	6	(Duret et al. 2000; The <i>C. elegans</i> Sequencing Consortium 1998)
<i>A. thaliana</i>	125	14	(The Arabidopsis Genome Initiative, 2000; Kidwell 2002)
<i>D. melanogaster</i>	137	15	(Kaminker et al. 2002)
<i>A. quadrimaculatus</i>	245	16	(Rai & Black, 1999)
<i>A. gambiae</i>	278	16	(Holt et al. 2002)
<i>F. rubripes</i>	400	2	(Elgar et al., 1999)
<i>O. sativa</i>	430	14	(Turcotte et al., 2001)
<i>Z. mays</i>	2500	60	(SanMiguel et al., 1996)
<i>M. musculus</i>	2700	48	(Waterston et al. 2002)
<i>H. sapiens</i>	2900	45	(International Human Genome Sequencing Consortium, 2001).
<i>H. vulgare</i>	5000	55	(Kumar & Bennetzen, 1999; Vicient et al., 1999)

A la luz de estos datos es interesante recordar la discusiones que existieron en la comunidad científica sobre la posibilidad de la presencia o influencia de los elementos transponibles en organismos superiores a bacterias o *Drosophila*, como los humanos o las plantas, y que no fue hasta la secuenciación completa de estos genomas cuando su presencia en ellos fue completamente admitida (Biémont 2010).

Distribución cromosómica

La abundancia de los elementos transponibles puede variar a lo largo del complemento de una especie, observándose de forma usual que la densidad de estos elementos difiere entre cromosomas, e incluso entre los dos brazos cromosómicos. Esta variación se ha observado en el análisis cromosoma a cromosoma realizado en el arroz (Feng et al. 2002; Sasaki et al. 2002). En humanos son los cromosomas 21 y 22 los que portan el 40% de los elementos transponibles (Grover et al. 2003) mientras que en *Drosophila*, es el cromosoma 4 el que tiene una mayor cantidad de elementos, transposones de ADN y retroelementos no-LTR (Bartolomé et al. 2002; Kaminker et al. 2002). Pero es el cromosoma Y el que presenta en más especies (humano, ratón y *Drosophila*) una mayor riqueza de elementos móviles, explicada probablemente por su baja recombinación, su escasa densidad génica y su gran contenido de heterocromatina (Junakovic et al. 1998; Boissinot et al. 2001; Waterston et al. 2002). También el cromosoma X presenta un acúmulo de retrotransposones en humanos y ratón, lo que podría estar relacionado con la inactivación del cromosoma X (Bailey et al. 2000). La acumulación en ambos cromosomas sexuales también se ha descrito para el elemento L1 en las especies del grupo Arvicolinae (Acosta et al. 2008). En mosquito, el cromosoma X es el cromosoma con mayor densidad de elementos por Mb (Holt et al. 2002).

En cuanto a la localización de los transposones dentro de los cromosomas, se ha apreciado en gran número de genomas donde éstos parecen ser más abundantes en las regiones heterocromáticas (Dimitri et al. 2005). En Dípteros, estas secuencias se acumulan cerca de centrómeros y telómeros, representando en *Drosophila* el 50% de la heterocromatina (Hoskins et al. 2002). Mientras que en mosquito, este porcentaje llega hasta el 60% (siendo fundamentalmente heterocromatina pericentromérica) (Holt et al. 2002). En hormigas, Palomeque et al. (2006) han encontrado la presencia de MITEs y elementos Mariner en regiones de ADNsat, sugiriéndose la implicación del elemento Mboumar en el desarrollo y mantenimiento de este ADN repetido (Muñoz-López et al. 2008). En plantas como *A. thaliana*, aparece una curiosa compartimentalización, los retroelementos se agrupan en los centrómeros y los transposones de ADN en regiones pericentroméricas, pero algunas

familias de transposones de ADN, así como los no autónomos MITE y SINE, se muestran distribuidos a lo largo de todos los cromosomas (The Arabidopsis Genome Initiative, 2000). En maíz, grandes bloques de elementos LTR se disponen entre los genes pero en estado metilado e inactivo (SanMiguel et al 1998) al contrario que los transposones de ADN que se hallan en regiones eucromáticas y genéticamente activas (Bennetzen, 2000).

Efectos de los elementos transponibles sobre el genoma hospedador.

El tremendo potencial de los elementos transponibles para generar diversidad genética ha sido reconocido a lo largo del tiempo, pudiendo ocasionar mutaciones espontáneas, recombinación, reordenaciones cromosómicas y disgénesis híbrida (Engel 1979; Shiroishi et al. 1993; Mathiopoulus et al. 1998), por lo que los elementos transponibles han sido clasificados generalmente como ADN “egoísta” desde los años 80 (Doolittle & Sapienza 1980; Orgel & Crick 1980) como término opuesto a la hipótesis original de “elementos controladores” propuesta por McClintock, 1956. Este cambio de denominación fue debido a que los elementos transponibles se reconocieron como entidades independientes del genoma cuya habilidad replicativa le permitía expandirse independientemente del beneficio del hospedador. La cuestión era entonces si este ADN egoísta era sólo ADN “basura” para el hospedador o si por el contrario jugaba un importante papel en la evolución del organismo. Actualmente la consideración de los elementos transponibles se han quedado en un punto medio entre la hipótesis de ADN “basura” y “elementos controladores”. El genoma del hospedador se considera como una comunidad ecológica donde se producen complejas interacciones hospedador-elementos transponibles y entre los propios elementos (Kidwell & Lisch 2001; Brookfield 2005).

Generación de diversidad alélica, transducción de genes, duplicaciones y barajamiento de exones

Los elementos transponibles, al igual que otros mutágenos, pueden producir una gran variedad de efectos en rasgos fenotípicos así como cambios silenciosos que sólo son detectables a nivel de secuencia de ADN. Estos elementos, no sólo producen su efecto mutagénico con su inserción, sino que también producen mutaciones cuando se escinden de forma imprecisa dejando tras de sí fragmentos de su secuencia o “huellas” de su presencia. La escisión de un transposón de ADN puede, por lo tanto, generar rápidamente

diversidad alélica, por lo que ciertas modificaciones sutiles de genes o de promotores pueden derivarse de este fenómeno de inserción/escisión. Pero a menos que el transposón sea capturado en el acto, estas suposiciones son muy difíciles de demostrar (Girard & Freeling 1999; Brookfield 2004; Koga et al. 2006).

Otro interesante efecto del papel mutagénico de los elementos transponibles ha sido observado en la evolución de plantas. En estos genomas se produce una relativa inactividad de los elementos móviles durante un desarrollo normal y una activación de éstos, en situaciones de estrés (Wessler 1996). Un incremento en la tasa de transposición sería un mecanismo para asegurarse de que toda la descendencia del individuo estresado porta al menos una copia del elemento, pero al mismo tiempo, proporciona un incremento en la tasa de mutación aumentando la posibilidad de que alguna de esas mutaciones sea ventajosa para la descendencia. Por otro lado, la inserción de un elemento transponible puede interrumpir o alterar la función de un gen, pero si esta inserción se produce en un intrón, es más fácilmente tolerable y pueden sobrevivir al proceso de selección. Un elevado número de inserciones de elementos móviles se han fijado en algunas especies de mamíferos, dejando abierta la puerta a la posibilidad de que hayan sido fruto de selección positiva.

Otros efectos provocados por los elementos transponibles son la movilización de otras secuencias génicas y la formación de pseudogenes. Esnault et al. (2000) ha mostrado como las *LINE* humanas son capaces de movilizar transcritos de ADN ajenos al propio elemento, con lo que esta formación de pseudogenes procesados mediante la retrotransposición de *LINE* puede ser un importante mecanismo de duplicación de genes.

Algunos elementos pueden servir además como vehículos para movilizar regiones flanqueantes del genoma. Estas secuencias, ya sean genes, exones o promotores, pueden ser transducida o insertada en un nuevo gen proporcionando un mecanismo de evolución de genes. Un ejemplo es el barajamiento de exones producido por el elemento *LI* en cultivos celulares (Moran et al. 1999). Recientes estudios han mostrado que los transposones de ADN han transducido cientos de fragmentos de genes en el genoma de plantas. Los elementos *MULE* han sido durante mucho tiempo sospechosos de este fenómeno (Lisch 2002) y con la disponibilidad del genoma del arroz, se ha podido apreciar cuantitativamente su extensión. Jiang et al. (2004) identificaron más de 3000 elementos llamados *PACK-MULE* que portaban fragmentos de genes celulares. Este estudio establece el enorme potencial de este elemento para la duplicación génica y el barajamiento de genes. Incluso los elementos *Helitrons*, con su particular forma de amplificación, también han mostrado su capacidad para barajar y duplicar secuencias (Morgante 2006).

Efectos epigenéticos

Se han encontrado claras evidencias de que los transposones representan dianas para diversos procesos de silenciamiento de genes que implican a ARNi y modificaciones epigenéticas (Sijen & Plasterk 2003; Soltkin 2007). Así, la estructura de los transposones, con sus repeticiones terminales, y la propensión a movimientos locales predisponen contra ellos los mecanismos de silenciamiento basado en ARNi y de formación de islas de heterocromatina (Sijen & Plasterk 2003; Slotkin et al. 2005; Grewal & Jia 2007) con implícitas consecuencias para los genes que se hallan en sus proximidades. Igualmente, se ha propuesto que la función principal de la metilación de secuencias de ADN en mamíferos es regular la actividad transposicional de los elementos móviles (Yoder et al. 1997).

La evidencia más directa de que los elementos transponibles son un tipo de secuencias sobre las que actúa la maquinaria que forma y mantiene la heterocromatina, se ha obtenido de un estudio que compara dos regiones duplicadas en el cromosoma 4 de *Arabidopsis* (Lippman & Martienssen 2004) que muestran distinta conformación de la cromatina en función de la presencia (heterocromática) o ausencia (eucromática) de secuencias repetidas, incluyendo transposones de tipo *MULE*. Los transposones, junto con los elementos *insulator* y *boundary* derivados de repeticiones (Ohzeki et al. 2002; Noma et al. 2006), participan activamente en la compartimentalización del genoma en dominios cromosómicos con distintas marcas epigenéticas y actividad transcripcional. Estas marcas, heredables y normalmente estables, puede estar sujetas a cambios en respuesta a las condiciones ambientales o estrés genético como hibridaciones o poliploidizaciones (Kashkush et al. 2002; Adams & Wendel 2005; Noor & Chang 2006).

Reordenaciones cromosómicas

Los elementos transponibles no sólo son capaces de producir mutaciones sobre pequeños fragmentos de ADN sino que también son responsables de cambios estructurales en los genomas, incluyendo deleciones, duplicaciones, inversiones y traslocaciones. Los dos mecanismos implicados en estas reestructuraciones de los genomas son la recombinación ectópica producida por la recombinación de copias de elementos distribuidas por el genoma, y la inducción de reordenaciones cromosómicas mediante una transposición alternativa en la que los extremos de transposones diferentes se unen y se mueven a una nueva localización mediante *corte-escisión* (Gray 2000).

Muchas especies de *Drosophila* aportan buenos ejemplos de los cambios estructurales producidos por elementos transponibles. Durante la evolución del género un elevado número de reordenaciones cromosómicas se han fijado repetidamente. El mejor ejemplo es el estudio realizado en los puntos de rotura de las inversiones encontradas en las poblaciones naturales de *Drosophila buzzatii* donde se demostraba que la recombinación ectópica entre transposones *Foldback* orientados de forma opuesta generaba dos inversiones cromosómicas independientes (Cáceres et al. 1999). En cada caso, los puntos de rotura de la inversión parecen ser puntos calientes genómicos con una secuencia y estructura muy variable entre poblaciones (Cáceres et al. 2001; Casals et al. 2003) y caracterizados por un complejo anidamiento de transposones de ADN (la mayoría elementos FB y hAT).

También en el genoma humano se encuentran señales de la actividad de los elementos móviles. En el 27% de los extremos de los segmentos duplicados que presenta el genoma humano aparecen secuencias *Alu* y, aunque muchos mecanismos pueden estar implicados en la aparición de estas duplicaciones, la recombinación entre elementos *Alu* parece ser la explicación más plausible (Bailey et al. 2003).

Evolución de los elementos transponibles y sus hospedadores

La dinámica evolutiva de los elementos transponibles es compleja, en parte debido a su replicación y las interacciones que provoca con el genoma hospedador. Esta intrincada dinámica se complica, aún más, por la habilidad de algunos elementos para cruzar las barreras entre especies y dispersarse por nuevos genomas mediante transferencia horizontal. La transferencia horizontal se convierte en una parte importante del ciclo de vida de algunos elementos transponibles y contribuye a su continuado éxito en su evolución (Figura 2) (Silva et al. 2004; Schaack et al. 2010). La presencia de elementos transponibles de Clase I y II en todos los grupos eucariotas evidencia su éxito a largo plazo, adoptando para ello diversas estrategias para mantenerse.

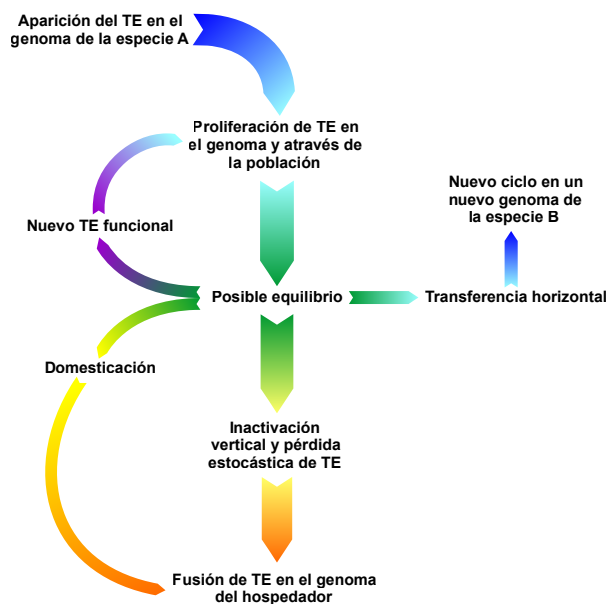


Figura 2. Modelo de dinámica evolutiva de los TE en los genomas eucariotas. Este hipotético modelo incorpora los datos de diversos trabajos (Hartl et al. 1999; Lampe et al. 2001; Silva et al. 2004).

Los elementos transponibles pueden determinar la historia evolutiva de sus genomas hospedadores a través tres tipos de potenciales efectos deletéreos: por mutagénesis insercional, puesto que pueden interrumpir la función o regulación génica; por coste transcripcional/traducciona de producción de los transcritos y proteínas que codifican y por recombinación ectópica entre copias homólogas del elemento situadas en diferentes localizaciones cromosómicas lo que resulta en duplicaciones, deleciones y nuevas relaciones entre genes (Kidwell & Lisch 2001; Bartolomé et al. 2002; Petrov et al. 2003; Feschotte & Pritham, 2007). La aportación de secuencias codificantes y no codificantes por parte del transposon también proporciona una oportunidad de cambio genómico al hospedador, a través de la creación de nuevas funciones génicas. El coste producido por los elementos transponibles puede incluir además el asociado con la replicación de ADN cuando estos son numerosos en el genoma. En esta sección discutiremos la dinámica intragenómica de las interacciones elemento transponible-hospedador, así como las diferentes estrategias seguidas para su mantenimiento.

Autoregulación de los TEs

La autorregulación ha sido observada en el elemento *Mariner* y *P* de *Drosophila* (Hartl et al. 1997; Kidwell & Lisch 2001). En el caso del elemento *P*, la autorregulación se realiza mediante la actividad de, al menos, dos tipos de represores codificados por el propio elemento. En el caso de *Mariner*, están involucrados muchos mecanismos como la inhibición por sobreproducción (un incremento de la cantidad de transposasa conduce a un decrecimiento neto en la transposición), el efecto de las mutaciones sin sentido (interferencia de la transposición por transposasas defectivas codificadas por copias defectivas del elemento), así como el secuestro de la transposasa por las copias defectivas.

Muchas familias de transposones, por otro lado, producen transcritos antisentido que pueden regular la transposición mediante interacciones ARN-ARN. En maíz se han encontrado dos transcritos antisentido que corresponden con los transcritos del elemento *MuDR* (Hershberger et al. 1995).

Por otro lado, las recombinaciones ectópicas producidas por regiones homólogas de elementos situados en diferentes lugares cromosómicos también podrían considerarse como un fenómeno de autorregulación involuntaria. Estos eventos producen duplicaciones, deleciones y nuevas uniones que a menudo reducen la fitness del hospedador y controlan el incremento del número de copias de los elementos. Un buen ejemplo de la importancia de la recombinación ectópica es el elemento *BARE-1* en cebada (Kalendar et al. 2000). En este genoma, además de elementos *BARE-1* completos se encuentran multitud de secuencias

huérfanas de LTR que corresponden con los restos producidos por recombinaciones entre LTR de elementos completos. La alta proporción de estas secuencias LTR, en relación a los elementos completos, es consistente con la idea de que una alta tasa de recombinación entre estos elementos ha reducido la población de copias intactas de *BARE-1*.

Regulación mediada por el hospedador

El ARN de interferencia (ARNi), que es un mecanismo basado en la supresión post-transcripcional en base a la homología de pequeños fragmentos de ARN de doble cadena, supone en el nematodo *Caenorhabditis elegans* un importante mecanismo de defensa frente a un amplio espectro de elementos transponibles (Tabara et al. 1999). En un reciente estudio en *Drosophila* se ha descubierto una compleja y diversa red de pequeñas secuencias de ARN que controlan la actividad de los elementos transponibles tanto en células de la línea germinal como somáticas (Malone & Hannon, 2009). El uso de estos tipos de ARN en el control de la actividad de los elementos transponibles en *Drosophila* fue inicialmente propuesta como mecanismo en la co-supresión del elemento *I* (Jensen et al. 2002). Los ARNpi, llamados así por ser ARNi que interaccionaban con proteínas del tipo Piwi (Vagin et al. 2006), se asocian a diferentes grupos de transposones, incluidos los elementos *I*, *roo* y *Gypsy* (Malone & Hannon 2009). Estos ARNpi difieren de los ARNsi en que tienen mayor longitud (24-29 nucleótidos) y se encuentran formando cluster derivados de elementos transponibles inactivos que pueblan el genoma de *Drosophila*. Este sistema de control puede heredarse vía materna a la descendencia confiriéndole la supresión epigenética de los elementos y explicando así, el fenómeno de la disgénesis híbrida comentado en el apartado *Elemento P y disgénesis híbrida*.

Existe además una relación entre el silenciamiento de los transposones por ARNpi y la heterocromatina. Recientemente se ha encontrado que la familia de proteínas HP1 (*heterochromatin protein 1*) es requerida para el silenciamiento de elementos transponibles, y la amplificación y producción de ARNpi (Klattenhoff et al. 2009). Así como que alteraciones en la maquinaria de la proteína chaperona Hsp90 afecta a la ruta de los ARNpi y permite la activación de los transposones (Specchia et al. 2010) con lo que también podrían actuar como un sistema de control de la actividad de los elementos transponibles.

Autónomos y no autónomos

Además de las interacciones genómicas descritas anteriormente, la mayoría de los elementos transponibles tienen que convivir con numerosas copias defectivas generadas durante o después de la transposición. Este proceso podría contribuir a su autorregulación, como ya comentamos, y puede conducir además a la total inactivación y extinción de los elementos si la población de elementos inactivos superara a la de las copias activas (Feschotte & Pritham 2007). Aún así, algunos elementos no autónomos han conseguido tener mucho éxito replicativo sin contar con mecanismo de replicación propio. Es el caso de las miles de copias de *SINE* encontradas en genomas de insectos como *Aedes aegypti* o *Bombyx moris* (Adams et al. 1986; Tu, 2004).

De forma similar, muchos *MITE* también han sido encontrados en alto número de copias en genomas de plantas (Le et al. 2000). El pequeño tamaño de los *MITE* y las *SINE* pueden producir un menor efecto deletéreo al hospedador, y por lo tanto más tolerable, puesto que su pequeño tamaño no es un sustrato suficiente para las recombinaciones (Petrov et al. 2003) o porque su impacto en los genes circundantes es menos severo. Así, tanto su reducida presión selectiva como otras propiedades inherentes a estos transposones no autónomos, han contribuido a su enorme éxito en algunos genomas. La pregunta que subyace es como estos *SINE* y *MITE* afectan a la evolución de los elementos autónomos que los movilizan.

Transmisión vertical y transferencia horizontal

Los principales mecanismos responsables de la amplificación y dispersión de los elementos transponibles, en y entre los genomas y poblaciones, son la transposición y la reproducción sexual, haciendo que los transposones se transmitan verticalmente de un organismo a otro en una población y especies. Pero con el tiempo, los mecanismos de control del hospedador así como las propias copias defectivas, o fenómenos como la selección purificadora o las recombinaciones ectópicas, llevan a la inactivación vertical de estos elementos (Lohe et al. 1995).

La habilidad para escapar a la inactivación vertical mediante la invasión de nuevos genomas mediante transferencia horizontal supondría un enorme logro para el éxito evolutivo de un elemento transponible, aunque no todos los transposones han adoptado el ciclo de vida de invasión, amplificación, senescencia y nueva invasión propuesto por muchos autores (Figura 2) (Hartl et al. 1997; Schaack et al. 2010). Esta transferencia horizontal suele estar apoyada por tres tipos de evidencias (Schaack et al. 2010). La primera

evidencia sería la detección de una alta similaridad en la secuencia de dos elementos de taxones distantes, aunque debería considerarse que evolucionen con diferente tasa de cambio. Una segunda evidencia sería la detección de incongruencias filogenéticas entre los elementos transponibles y sus hospedadores. Y finalmente, la tercera prueba que evidenciaría transferencia horizontal, sería una distribución parcheada de los elementos móviles entre taxones relacionados.

En la literatura se han encontrado 218 casos convincentes de transferencia horizontal de transposones, donde se hallan implicados elementos de casi todos los tipos existentes (Schaack et al. 2010). El primer caso de transferencia horizontal en eucariotas fue reportado en *Drosophila*, donde el elemento *P* había invadido el genoma de *D. melanogaster* desde el grupo de *D. willistoni* durante el pasado siglo (revisión en Kidwell 1992). Este caso de transferencia horizontal cuenta con los tres tipos de evidencias antes comentados y es ampliamente aceptado (Silva et al. 2004). Análisis posteriores de un elevado número de elementos *P* en diferentes especies de *Drosophila* ha mostrado que la transferencia horizontal ha sido recurrente en el grupo (Silva & Kidwell 2000). Otro caso espectacular de transferencia horizontal es el presentado por el elemento *Mariner*. Esta familia de transposones de ADN ha estado implicada en cientos de casos de transferencia horizontal entre un amplio rango de animales (Robertson 1993).

También se han encontrado casos de transferencia horizontal de elementos LTR, como es el caso del elemento *Copia* de *D. melanogaster* a *D. willistoni* y del elemento *Gypsy* dentro del grupo de los Drosophilidos (Terzian et al. 2000; Vazquez-Manrique et al. 2000). Sin embargo, son menos los eventos de transferencia horizontal de los elementos no-LTR. El caso más evidente es el producido por un elemento *RTE*, *Bov-B*, entre animales y reptiles (Kordis & Gubensek 1998) e ilustra la dificultad de afirmar la existencia de transferencia horizontal de un elemento transponible. Este evento fue primero apoyado por la similitud encontrada en una región de 550 bp del extremo 3' del elemento de reptiles y rumiantes así como por la distribución parcheada que éste presentaba entre los tetrápodos (Kordis & Gubensek 1998), pero fue refutado, años después, por el análisis de la secuencia de un fragmento mayor de la región de la retrotranscriptasa (Malik et al. 1999). Se necesitaron estudios más profundos de las secuencias de los elementos de más taxones (Zupunski et al. 2001, Gogolevsky et al. 2008) así como el descubrimiento de un fragmento de la secuencia de *Bov-B* de reptiles integrado en un poxvirus que infecta a mamíferos para la aceptación sin reservas de este caso de transferencia horizontal (Piskurek et al. 2007).

Dado que los proyectos de secuenciación de genomas se están extendiendo, el número de casos de transferencia horizontal puede incrementarse pronto, descubriendo que existen más casos de transferencia entre especies que comparten ecosistemas y dilucidando cuáles son los mecanismos que hacen posible esta transferencia lateral.

Domesticación

Existen muchos ejemplos de genes de los hospedadores identificados en diferentes linaje eucarióticos que derivan de transposasas (Smit, 1999). Quizás el mejor ejemplo estudiado de la domesticación de una transposasa, es la proteína activadora de la recombinación Rag1. Esta proteína inicia junto con Rag2 la recombinación V(D)J, una recombinación sitio-específica necesaria para el ensamblaje de la región variable de las inmunoglobulinas y el receptor de inmunoglobulinas de las células B y T. Este fenómeno genera una alta diversidad de inmunoglobulinas, y los receptores de los linfocitos T son capaces de reconocer un amplio rango de antígenos. El complejo Rag1/2 funciona como una endonucleasa produciendo roturas de doble cadena cerca de las secuencias señal de recombinación. Se ha podido demostrar que Rag1/2 puede trabajar como una transposasa y catalizar una transposición in vivo (Agrawal et al. 1998; Hiom et al. 1998). Además, muestra una clara relación entre el mecanismo de transposición del transposon *hAT* y la recombinación V(D)J. Estas observaciones suponen un fuerte apoyo para la hipótesis de que Rag1/2 deriva de una antigua transposasa de hace 500 millones de años (Roth & Craig 1998). Por otro lado, existen evidencias que sugieren que tanto Rag1 como las secuencias señal de recombinación son derivadas del transposón de ADN *Transib* que está presente en hongos, insectos, gusanos, hydras y erizos de mar (Kapitonov & Jurka 2005).

Otro gen posiblemente domesticado en eucariotas es el de la proteína CENP-B. Esta proteína asociada al centrómero se une específicamente a la caja CENP-B y se considera que deriva de la transposasa del elemento *Pogo* (Smit & Riggs 1996). Esta proteína está muy conservada en mamíferos y presenta homólogos en plantas y hongos sugiriendo un evento de domesticación muy antiguo. En mamíferos, muchos genes se consideran derivados de una transposasa *Pogo* (Smit & Riggs 1996), como el gen *jerky* que codifica una proteína de unión al ARNm y que es específica de cerebro con una alta expresión en neuronas. Esta proteína está relacionada con síndromes de epilepsia en ratones y humanos (Toth et al. 1995).

Otro ejemplo de domesticación en humanos es el gen para la proteína Metnase (también llamada Setmar) que está constituido por una transposasa de tipo *Mariner* fusionada con una región que codifica un dominio Set con actividad metiltransferasa de histonas (Robertson & Zumpano 1997; Lee et al. 2005). Setmar incrementa la resistencia a radiaciones ionizantes y la reparación de doble rotura de ADN, así como promueve la integración de ADN exógeno en las células hospedadoras (Lee et al. 2005).

En organismos tan diversos como animales, hongos, plantas o protozoos, los telómeros son replicados eficazmente por una enzima telomerasa, una ribonucleoproteína que contiene una subunidad de ARN que sirve de molde para la replicación del telómero. Esta telomerasa es una ADN polimerasa dependiente de ARN y pertenece a la familia de las retrotranscriptasas al igual que las de los retrotransposones (Nakamura et al. 1997; Lingner et al. 1997). Esta telomerasa usa para la replicación del telómero el extremo 3' OH que recuerda a la transcripción de los retroelementos no-LTR, con los que está filogenéticamente relacionada (Eickbush 1997). Retroelementos no-LTR son usados en algunas especies, como *Drosophila melanogaster*, en lugar de la telomerasa, para el mantenimiento de los telómeros (Pardue & DeBaryshe 2003).

Un último ejemplo de la domesticación de elementos transponibles sería las evidencias que existen sobre el papel que juega un gen retroviral *env* en la formación de la placenta. La proteína derivada de este gen *env* sería la *Sincitina-1* que es codificada en humanos por el provirus defectivo HERV-W (Mi et al. 2000). Otros genes de este endovirus han sido inactivados por mutaciones pero no es el caso de la *Sincitina-1* que es expresada principalmente en el sincitiotrofoblasto, una estructura que facilita el intercambio entre la madre y el feto. El papel de la *Sincitina-1* es provocar la fusión celular y la formación del sincitio. Esta actividad es inhibida en la línea celular trofoblástica pero es conservada en chimpancés, gorilas, orangután y gibones (Mi et al. 2000; Mallet et al. 2004).

Los múltiples ejemplos de genes que han sido derivados de elementos transponibles en los genomas eucariotas suponen una importante prueba del papel del ADN móvil como reservorio de nuevas funciones y regulaciones celulares, y por lo tanto en la evolución en eucariotas.

Otras estrategias evolutivas

Junto con la transferencia horizontal como vía de escape a la extinción vertical, recientes estudios sugieren una vía alternativa que pueden adoptar los elementos transponibles (Lampe et al. 2001). En base a la pérdida de interacción entre los transposones *Mariner* y unas TIR ligeramente diferentes, se ha propuesto que la diversificación intra-específica o intra-genómica de los elementos *Mariner* permitiría la aparición de un nuevo elemento que comenzara una nueva línea. Este proceso requeriría la coevolución de la transposasa con las repeticiones terminales, pero permitiría al elemento escapar de la inactivación vertical. La secuenciación de genomas ha proporcionado una excelente oportunidad para investigar la diversidad de las familias de elementos transponibles. Un reciente análisis ha encontrado en *A. gambiae* un enorme número de linajes de elementos no-LTR de los clados *CRI* y *Jockey* (Biedler & Tu 2003) que podrían derivar de la competición existente entre diferentes familias de elemento no-LTR o de sucesivos intentos de evadir los mecanismo de supresión producidos por el hospedador.

Por otro lado, algunos elementos transponibles pueden ser rescatados para cumplir funciones dentro del genoma del hospedador, llegando a ser “domesticados” (Lampe et al 2001) y dejando de ser, por lo tanto, elementos transponibles. Este positivo impacto de los elementos transponibles en el hospedador será discutido en el siguiente apartado.

Cromosomas B

El descubrimiento de los cromosomas B se debe a los estudios de Wilson (1907) en el insecto *Metapodius* (ahora llamado *Acanthocephal*) a principios del siglo XX, pero no fue hasta 1928 cuando Randolph utilizó por primera vez el término de “cromosomas B” para diferenciarlos de los otros cromosomas del complemento básico o “cromosomas A”. Estos cromosomas se han definido como “cromosomas supernumerarios e innecesarios que no recombinan con los cromosomas A y poseen una herencia irregular y no mendeliana” (Jones 1995). Estos cromosomas se encuentran presentes en más de 1300 especies de plantas y en casi 500 especies de animales, habiéndose descrito también en 10 especies de hongos, llegándose a estimar que un 15% de especies presentan cromosomas B (Jones y Rees 1982; Jones y Puertas 1993; Jones 1995). El número máximo de cromosomas B observado en un individuo ha sido de 50 cromosomas en la planta *Pachyphytum fitksui* (Camacho 2005).

La naturaleza de su interacción con el hospedador ha sido entendida de diferente forma a lo largo del tiempo, siendo considerados por algunos autores como parásitos en una constante “carrera armamentística” coevolutiva con el hospedador por algunos autores (Östergren 1945), o como beneficiosos al considerar otros investigadores un modelo *heterótico* (White 1973). La mayor parte de los cromosomas B se ajustan al modelo egoísta-parasítico, pero existen algunos cromosomas B, sin mecanismos de acumulación, que podrían considerarse casi-neutrales (Camacho et al. 1997).

Los principales constituyentes de los cromosomas B son las secuencias repetidas, tales como ADN satélite, genes ribosómicos 45S (ADNr) o elementos transponibles (Camacho et al. 2000; Camacho 2005). El ADN satélite que enriquece los cromosomas B es en algunos casos específico de éste, como es el ADN satélite del cromosoma PSR (Paternal Sex Ratio) de *Nasonia vitripennis* (Nur et al. 1988), pero en la mayoría de los casos es compartido con los cromosomas A (Jamilena et al. 1994; López-León et al. 1994). La presencia de genes ribosómicos en los cromosomas B se ha observado en multitud de especies, encontrándose generalmente inactivos por lo que son escasos los ejemplos de actividad nucleolar asociada a estos cromosomas extras (Schmid et al. 2002; Teruel et al. 2007, 2009b).

Estas secuencias repetidas contribuyen a la naturaleza heterocromática que presentan los cromosomas B (Camacho 2005), que junto con la ausencia de recombinación con el resto del genoma, lo convierten en un compartimiento genómico propicio para la acumulación de elementos transponibles. La presencia de elementos móviles se ha detectado en numerosos cromosomas B siendo el caso más llamativo el del cromosoma PSR de *N. vitripennis*, que se encuentra prácticamente invadido por el retrotransposón *NATE* (Nasonia Transposable Element) (McAllister 1995; McAllister & Werren 1997). Por otro lado, en maíz se ha documentado la formación de una secuencia específica del cromosoma B a partir de un elemento móvil, el elemento *StarkB* (Lamb et al. 2007), que muestra actividad transcripcional y está formado por secuencias derivadas de un elemento LTR presente en los cromosomas A y de secuencias específicas del cromosoma B. En el centeno, también se ha encontrado una secuencia altamente repetida y específica del cromosoma B que deriva de retrotransposones *Ty3-gypsy*, y que está implicada en la no-disyunción del cromosoma B (Langdon et al. 2000).

La heterocromatinización es una fuerza inactivadora y silenciadora de genes, entre ellos los elementos transponibles (Weiler & Wakimoto 1995), por lo que esta naturaleza heterocromática del B explicaría que sean escasos los cromosomas B que presentan efectos fenotípicos (Camacho 2005). Sin embargo, sí se han encontrado numerosas evidencias del efecto de los cromosomas B sobre caracteres asociados con el vigor, fertilidad y fecundidad, siendo este efecto negativo sobre la eficacia biológica de los individuos portadores, por lo que los cromosomas B podrían ser considerados parásitos genómicos (Camacho et al. 2000). La presencia de cromosomas B puede producir efectos sobre procesos celulares como la alteración en los niveles de expresión de proteínas, la expresión de las NORs o la frecuencia de quiasmas. En el caso del saltamontes *Eyprepocnemis plorans*, los cromosomas B influyen en el número de NORs activas aumentándolas en los individuos portadores de estos cromosomas extras (Teruel et al. 2007, 2009). También en esta especie se ha demostrado que los cromosomas B aumentan la frecuencia de quiasmas en el genoma hospedador cuando el cromosoma B es más parasítico, mostrando así una relación entre el efecto sobre los quiasmas y el estado evolutivo en que se encuentra el cromosoma supernumerario (Camacho et al. 2002).

El mantenimiento de los cromosomas B en las poblaciones, a pesar de ser prescindibles y de sus efectos deletéreos, se debe una tasa de transmisión más elevada que la de los cromosomas A (véase revisión de Camacho 2005; Jones et al. 2008). La existencia de un mecanismo de acumulación en la línea germinal, asegura a los cromosomas B la invasión de las poblaciones naturales. Es tal la rapidez que se espera de esta invasión

(Camacho et al. 1997) que está justificado el escaso número de casos reportados (Zurita et al. 1998; Cavallaro et al. 2000; Araujo et al. 2001; Riera et al. 2004). Si un cromosoma B inflige una carga a los individuos que lo portan, se genera un conflicto intragenómico que conduce a la selección, en los cromosomas A, de genes que se contrapongan a la acumulación del B y sus efectos, como ha ocurrido con el cromosoma B₂₄ en el saltamontes *Eyprepocnemis plorans* (Perfectti et al. 2004). Los cromosomas B neutralizados, es decir, aquellos que carecen de mecanismos de acumulación, pueden persistir durante largos periodos de tiempo en las poblaciones naturales antes de su total extinción, pudiendo ser regenerados por mutación (Camacho et al. 1997; Zurita et al. 1998). Este ciclo, de invasión, neutralización y extinción/regeneración ha sido propuesto por Camacho (2005) como un modelo general de la dinámica evolutiva de los cromosomas B, que integra los modelos propuestos para explicar el mantenimiento de los cromosomas B en las poblaciones (Figura 3).

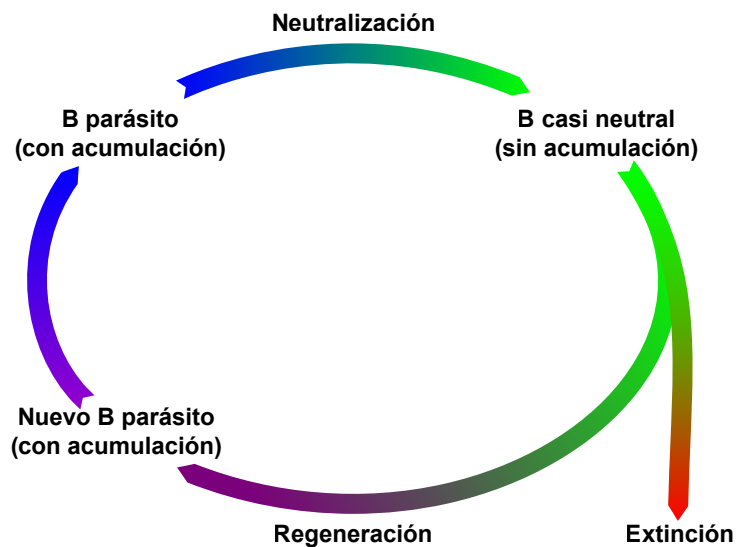


Figura 3. Modelo general de la dinámica evolutiva de un cromosoma B parásito.

A pesar de ser los primeros elementos genéticos egoístas descritos (Wilson 1907), en la mayor parte de los casos el origen de los cromosomas B es desconocido. Dos hipótesis se han propuesto para explicar el origen de estos cromosomas. Battaglia (1964) fue el primero en proponer que los cromosomas B podrían surgir durante los procesos de hibridación entre especies (origen interespecífico). Este es el caso del cromosoma PSR de la avispa *Nasonia*, donde un análisis filogenético de NATE muestra que las secuencias de este transposon, presentes en el cromosoma PSR, son más parecidas a las copias presentes en especies relacionadas del género *Trichomalopsis* que a las copias existentes en el propio

genoma de *Nasonia* (McAllister & Werren 1997), demostrándose empíricamente al observarse la formación de novo de un cromosoma B a través de cruzamientos controlados entre especies emparentadas (Perfectti & Werren 2001). Pero los cromosomas B también pueden aparecer con un origen intraespecífico como un producto de la evolución de los cromosomas A (Camacho 2005) lo que sería más probable. El origen intraespecífico implica que el cromosoma B debería poseer fundamentalmente secuencias de ADN de la misma clase que las presentes en los cromosomas A, como se ha descrito en el cromosoma B de *Crepis capillaris*, donde las secuencias aisladas por microdissección del B se encuentran también en los demás cromosomas del genoma (Jamilena et al. 1994, 1995).

Sistema de cromosomas B en *Eyprepocnemis plorans*

El sistema de cromosomas B de *E. plorans* ha sido intensamente estudiado y caracterizado desde su descripción (véase revisión de Camacho et al. 2003), aunque su presencia parece estar circunscrita solamente a la subespecie *E. plorans plorans* (López-León et al. 2008) que habita en la costa mediterránea, el Cáucaso, Turquía, Turkmenistán, Irán y el sudoeste de la península arábiga (Dirsh 1958). Los cromosomas B de este sistema muestran una extensa variación, con más de 50 tipos citogenéticos (López-León et al. 1993), algunos de ellos encontrados en la misma población. En *E. plorans*, los cromosomas B son mitóticamente estables, lo que implica que el mismo número de cromosomas B está presente en todas las células del individuo, no ocurriendo lo mismo con la frecuencia con la que son encontrados entre poblaciones que varía ampliamente (Camacho et al. 2003), pudiendo encontrar individuos que no portan este polimorfismo o individuos con hasta cuatro cromosomas accesorios.

La mayoría de las variantes de cromosomas B presentes en *E. plorans* carecen de mecanismos de acumulación como resultado de la neutralización de sus mecanismos de acumulación por parte del hospedador (Camacho et al. 1997). Siendo posible, además, que una variante de cromosoma B presente acumulación en una población mientras que en otra no (Bakkali et al. 2002). Las variantes de cromosomas B más extendidas (B_1 , B_2 y B_5) carecen tanto de mecanismos de acumulación como de efectos deletéreos para la fitness del hospedador (López-León et al. 1992; Herrera et al. 1996; Muñoz et al. 1998). La variante neutral B_2 ha sido reemplazada recientemente por la variante parasítica (con acumulación) B_{24} en la población de Torrox. Esta nueva variante se relaciona con un significativo descenso en la fertilidad de los huevos pero es transmitido a un 70% de la descendencia a través de las hembras (Zurita et al. 1998).

En *E. plorans*, como ocurre en la mayoría de los cromosomas B descritos, todos los tipos de cromosomas B son heterocromáticos y están enriquecidos en dos clases de secuencias de ADN repetido (Cabrero et al. 1999) y que además son compartidas con los cromosomas A: un ADN satélite de 180 bp y ADN ribosómico (López-León et al. 1994). La presencia de estas dos clases de secuencias en el cromosoma X, organizadas en el mismo orden (centromero-ADNsat-ADNr) que en el que presenta el cromosoma B₂, permitió postular a López-León et al. (1994) que el cromosoma B₂ derivaba del cromosoma X. Por otro lado, esta organización de las secuencias repetidas en el cromosoma está conservada en la mayoría de los cromosomas B de España y Marruecos (Bakkali et al. 1999; Cabrero et al. 1999) sugiriendo un origen común. Sin embargo, las poblaciones del Mediterráneo oriental, incluyendo poblaciones de Grecia, Turquía y Armenia, muestran diferentes tipos de cromosomas B que los de las poblaciones del Mediterráneo occidental, con mayores cantidades de ADNr y menores de ADNsat que los cromosomas B de las poblaciones españolas y marroquíes (Abdelaziz et al. 2007; López-León et al. 2008). Además, los cromosomas B orientales parecen contener ADN ribosómico 5S, que no es detectado en los cromosomas B españoles ni marroquíes (Cabrero et al. 2003). Esta diferente composición de secuencias entre los cromosomas B de las poblaciones orientales y occidentales de esta especie ha permitido proponer un origen multirregional para el cromosoma B, pudiendo derivar cada tipo de B de distintos cromosomas A. Sin embargo, en un reciente estudio, Muñoz-Pajares et al. (2011) han conseguido aislar un marcador SCAR específico del cromosoma B que proporciona una nueva perspectiva acerca del origen del B al encontrar una alta similitud entre las secuencias de este marcador (1510 bp) provenientes de poblaciones muy distantes (Marruecos y Armenia) lo que sugiere con fuerza un origen único y reciente para el cromosoma B.

Sistema de cromosomas B en *Locusta migratoria*

La especie *Locusta migratoria* es portadora de uno de los sistemas de cromosomas B más extendidos de todos los estudiados en plantas y animales, puesto que se ha descrito la presencia de cromosomas B en poblaciones naturales de Japón (Nur 1969; Kayano 1971), China (Hsiang 1958), Mali (Dearn 1974), Australia (King & John 1980) y España (Cabrero et al. 1984), en definitiva, en todos los continentes donde esta especie habita.

Los cromosomas B de *L. migratoria*, al contrario de lo comentado para *E. plorans*, son mitóticamente inestables. Esta inestabilidad es debida a la no disyunción mitótica de estos cromosomas durante el desarrollo embrionario que constituye un mecanismo de acumulación premeiótica en los machos de la especie (Nur 1969; Cabrero et al.

1984; Viseras et al. 1990). Por esta razón, existe una gran variación intraindividual en el número de cromosomas B que portan las células de un individuo. La inestabilidad de los cromosomas B es ya patente en los embriones de 3 días, alcanzando su máximo al quinto día de desarrollo (Pardo et al. 1995). En cruces controlados, sin embargo, se muestra que los cromosomas B también se acumulan por vía materna mediante la migración preferencial del cromosoma B hacia el oocito secundario durante la primera división meiótica (Pardo et al. 1994).

El análisis citogenético de estos cromosomas B durante la espermatogénesis muestra que su mitad proximal tiene un carácter eucromático mientras que la mitad distal es heterocromática (Cabrero et al. 1984). La técnica del bandeo C corrobora esta naturaleza heterocromática del extremo distal (Cabrero et al. 1984), lo que es consistente con el patrón observado en las poblaciones españolas (Santos 1980). El bandeo con fluorescencia también muestra que la zona distal de este cromosoma B es una región DAPI+, lo que es usual en zonas heterocromáticas ricas en A+T, mientras que la región proximal es CMA3+ que muestra que esta región es rica en ADN G+C (Camacho et al. 1991).

Los primeros datos sobre la composición molecular del cromosoma B de *L. migratoria* se obtuvieron mediante la microdissección de este cromosoma realizada por Teruel et al. (2009) encontrando que los cromosomas B comparten al menos dos tipos de secuencias repetidas con los cromosomas A. Una de estas secuencias repetidas se localizan en la región distal y heterocromática del B y en las regiones pericentroméricas de la mitad de los cromosomas A, incluyendo el cromosoma X. El otro tipo de ADN repetido se localiza distribuido a lo largo de las regiones eucromáticas de todos los cromosomas A y en la zona eucromática proximal del cromosoma B, lo que sugiere que estas secuencias puedan corresponder con elementos móviles. Estos resultados apoyaban con fuerza el posible origen intraespecífico del cromosoma B de *L. migratoria*, pero no delimitaban cual de los cromosomas A sería el precursor del B. En un trabajo posterior, Teruel et al. (2010) profundizaron en el conocimiento molecular de este cromosoma B al identificar genes para las histonas H3 y H4 tanto en el cromosoma B como en el autosoma 8, lo que apunta al cromosoma 8 como originario del B. El análisis de la similaridad entre las secuencias de las histonas H3 y H4 de los dos cromosomas corrobora este origen, así como sugieren la inactividad de las copias del cromosoma B. Las estimas del tiempo de divergencia entre estos genes de histonas de los cromosomas A y B indican, además, que la presencia de cromosomas B en *L. migratoria* constituyen un polimorfismo antiguo (más de 750000 años) que ilustra la capacidad de los cromosomas B para mantenerse en las poblaciones naturales durante largos periodos de tiempo.

Objetivos

El objetivo general planteado en esta Tesis Doctoral es conocer el papel que juegan los elementos móviles como componentes de los genomas de ortópteros y la dinámica evolutiva e intragenómica que desarrollan. Para abordar este objetivo general nos planteamos los siguientes objetivos específicos:

1. Analizar la presencia de los elementos transponibles, *Gypsy*, *RTE*, *Mariner* y *R2*, en los genomas de *Eyprepocnemis plorans* y *Locusta migratoria* y el aislamiento de los mismos.
2. Establecer la localización cromosómica de estos elementos en el complemento cromosómico de ambas especies incluyendo los cromosomas B.
3. Comparar la cantidad relativa de elementos transponibles presentes en diferentes poblaciones, así como en individuos portadores y no portadores de cromosomas B.
4. Comprobar la actividad transcripcional de estos elementos en los saltamontes *E. plorans* y *L. migratoria*.
5. Reconstruir la historia evolutiva de los elementos transponibles *Gypsy* y *RTE* en la familia de ortópteros Acridoidea.
6. Analizar la dinámica intragenómica del elemento *RTE* en el genoma de *E. plorans*.

Bibliografía

- Abdelaziz M, Teruel M, Chobanov D, Camacho JPM & Cabrero J (2007). Physical mapping of rDNA and satDNA in A and B chromosomes of the grasshopper *Eyprepocnemis plorans* from a Greek population. *Cytogenet Genome Res* 119: 143-146.
- Abelson J, Trotta CR & Li H (1998). tRNA splicing. *J Biol Chem* 273: 12685-12688.
- Acosta MJ, Marchal JA, Fernández-Espartero CH, Bullejos M & Sánchez A (2008). Retroelements (LINEs and SINEs) in vole genomes: differential distribution in the constitutive heterochromatin. *Chromosome Res* 16: 949-959.
- Adams DS, Eickbush TH, Herrera RJ & Lizardi PM (1986). A highly reiterated family of transcribed oligo (A)-terminated, interspersed DNA elements in the genome of *Bombyx mori*. *J Mol Biol* 187: 465-478.
- Agrawal A, Eastman QM & Schatz DG (1998). Implications of transposition mediated by V(D)J-recombination proteins RAG1 and RAG2 for origins of antigen-specific immunity. *Nature* 394: 744-751.
- Araújo SM, Pompolo SG, Perfectti F & Camacho JPM (2001). Integration of a B chromosome into the A genome of a wasp. *Proc R Soc Lond B* 268: 1127-31.
- Arthur A, Nimmo E & Hettle S (1984). Transposition and transposition immunity of transposon Tn3 derivatives having different ends. *EMBO J* 3: 1723-1729.
- Bailey JA, Carrel L, Chakravarti A & Eichler EE (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc Natl Acad Sci USA* 97: 6634-9.
- Bailey JA, Liu G & Eichler EE (2003) An Alu transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* 73: 823-34.
- Bartolomé C, Maside X & Charlesworth B (2002). On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol Biol Evol* 19: 926-937.
- Battaglia, E., 1964 Cytogenetics of B chromosomes. *Caryologia* 17: 245-299.
- Bennetzen J L (2000). Transposable element contributions to plant gene and genome evolution. *Plant Mol Biol* 42: 251-69.

- Biedler J & Tu Z (2003). Non-LTR retrotransposons in the African malaria mosquito, *Anopheles gambiae*: unprecedented diversity and evidence of recent activity. *Mol Biol Evol* 20: 1811-25.
- Biémont C (2010). A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* 186: 1085-93.
- Boissinot S, Entezam A & Furano AV (2001). Selection against deleterious LINE-1-containing loci in the human lineage. *Mol Biol Evol* 18: 926-35.
- Boulesteix M, Weiss M & Biémont C (2006). Differences in genome size between closely related species: the *Drosophila melanogaster* species subgroup. *Mol Biol Evol* 23: 162-7.
- Burt A. & Trivers R. (1998). Selfish DNA and breeding system in flowering plants. *Proc R Soc Lond B* 265: 141-146.
- Cabrero J, Bugrov A, Warchałowska-Sliwa E, et al. (2003). Comparative FISH analysis in five species of Eyprepocnemidine grasshoppers. *Heredity* 90: 377-81.
- Cabrero J, Viseras E & Camacho JPPM (1984). The B-chromosomes of *Locusta migratoria* I. Detection of negative correlation between mean chiasma frequency and the rate of accumulation of the B's; a reanalysis of the available data about the transmission of these B-chromosomes. *Genetica* 64: 155-164.
- Cabrero J, López-León MD, Bakkali M & Camacho JPM (1999). Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* 83: 435-439.
- Camacho JPM, Bakkali M, Corral JM, et al. (2002). Host recombination is dependent on the degree of parasitism. *Proc R Soc Lond B* 269: 2173-2177.
- Camacho JPM, Cabrero J, Viseras E, et al. (1991). G banding in two species of grasshopper and its relationship to C, N, and fluorescence banding techniques. *Genome* 34: 638-643.
- Camacho JPM, Sharbel TF & Beukeboom LW (2000). B-chromosome evolution. *Phil Trans R Soc Lond B* 355: 163-78.
- Camacho JPM, Shaw MW, López-León MD, Pardo MC & Cabrero J (1997). Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper *Eyprepocnemis plorans*. *Am Nat* 149: 1030-1050.

- Camacho JPM (2005) B Chromosomes, in E Inc. (ed), *The Evolution of the Genome*, Gregory, T. R., San Diego, pages. 223-286.
- Camacho JPM, Cabrero J, López-León MD, Bakkali M & Perfectti F (2003). The B chromosomes of the grasshopper *Eyprepocnemis plorans* and the intragenomic conflict. *Genetica* 117: 77-84.
- Campbell A, Berg D, Botstein D & Lederberg E (1979). Nomenclature of transposable elements in prokaryotes. *Gene* 5: 197-206.
- Capy Pierre, Langin T & Higuete D (eds.) (1997) *Evolution and impact of transposable elements*, Kluwer Academic, Dordrecht.
- Casals F, Cáceres M, Manfrin MH, González J & Ruiz A (2005). Molecular characterization and chromosomal distribution of Galileo, Kepler and Newton, three foldback transposable elements of the *Drosophila buzzatii* species complex. *Genetics* 169: 2047-2050.
- Cavallaro ZI, Bertollo L a, Perfectti F & Camacho JP (2000). Frequency increase and mitotic stabilization of a B chromosome in the fish *Prochilodus lineatus*. *Chromosome Res* 8: 627-34.
- Christopher DA & Hallick RB (1989). *Euglena gracilis* chloroplast ribosomal protein operon: a new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. *Nucleic Acids Res* 17: 7591-7608.
- Cáceres M, Puig M & Ruiz A (2001). Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. *Genome Res* 11: 1353-64.
- Davies RW, Waring RB, Ray JA, Brown TA & Scazzocchio C (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature* 300: 719-724.
- Dawkins R (1993). *El gen egoísta* Salvat Editores Sa (ed). *Salud P* 50 :page. 289.
- Dearn JM (1974). Phase transformation and chiasma frequency variation in locusts. *Chromosoma* 45: 321–338.
- Delprat A, Negre B, Puig Marta & Ruiz A (2009) .The transposon Galileo generates natural chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS one* 4: page. e7883.

- Dimitri P, Corradini N, Rossi F, et al. (2005). Transposable elements as artisans of the heterochromatic genome in *Drosophila melanogaster*. *Cytogenet Genome Res* 110: 165-72.
- Dirsh VM (1956). The phallic complex in Acridoidea (Orthoptera) in relation to taxonomy. *Transactions of the Royal Entomological Society of London* 108: 223-270.
- Doolittle WF & Sapienza C (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284: 601-603.
- Duret L, Marais G & Biémont C (2000). Transposons but not retrotransposons are located preferentially in regions of high recombination rate in *Caenorhabditis elegans*. *Genetics* 156: 1661-9.
- Edgell D & Belfort M (2000). Barriers to intron promiscuity in bacteria. *J Bacteriol* 182: 5281-5289.
- Eickbush TH (1997). Telomerase and retrotransposons: which came first? *Science* 277: 911-912.
- Engels WR (1979). Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genetical Res* 33: 219-236.
- Esnault C, Maestre J & Heidmann T (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 24: 363-367.
- Fedoroff N (1989). Maize transposable elements, in Douglas Berg & M Howe (eds), *Mobile DNA*, American Society for Microbiology, pages. 375-411.
- Fedoroff N, Wessler S & Shure M (1983) Isolation of the transposable maize controlling elements Ac and Ds. *Cell* 35: 235-242.
- Feng Q, Zhang Y, Hao P, et al. (2002) Sequence and analysis of rice chromosome 4. *Nature* 420: 316–320.
- Feschotte C & Zhang X (2002). Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons, in N Craig, R Craigie, M Gellert, & A Lambowitz (eds), *Mobile DNA II*, American Society for Microbiology Press, Washington DC, pages. 1147–1158.
- Finnegan DJ (1989). Eukaryotic transposable elements and genome evolution. *Trends Genet* 5: 103–107.

- Gardner MJ, Hall N, Fung E, et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498-511.
- Gogolevsky KP, Vassetzky NS & Kramerov DA (2008). Bov-B-mobilized SINEs in vertebrate genomes. *Gene* 407: 75-85.
- Grewal SIS & Jia S (2007). Heterochromatin revisited. *Nat Rev Genet* 8: 35-46.
- Grover D, Majumder PP, B Rao C, Brahmachari SK & Mukerji M (2003). Nonrandom distribution of alu elements in genes of various functional categories: insight from analysis of human chromosomes 21 and 22. *Mol Biol Evol* 20: 1420-1424.
- Hartl DL, Lohe AR & Lozovskaya ER (1997). Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31: 337-358.
- Heffron F, McCarthy BJ, Ohtsubo H & Ohtsubo E (1979). DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3. *Cell* 18: 1153-1163.
- Herrera J, López-León M, Cabrero J, Shaw M & Camacho JPM (1996). Evidence for B chromosome drive suppression in the grasshopper *Eyprepocnemis plorans*. *Heredity* 76: 633-639.
- Hewitt GM (1979) Grasshopper and crickets. In: John B (ed) *Animal cytogenetics*, vol. 3: insecta 1 Orthoptera. Gebruder Borntraeger, Berlin
- Hiom K, Melek M & Gellert M (1998). DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* 94: 463-70.
- Holt RA, Subramanian GM, Halpern A, et al. (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129-149.
- Hoskins RA, Smith CD, Carlson JW, et al. (2002). Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome biology* 3: 0085.1-0085.16.
- HsiangW(1958) Cytological studies on migratory locust hybrid, *Locusta migratoria migratoria* L. X *Locusta migratoria manilensis* Meyen. *Acta Zoologica Sinica* 10: 53-59
- Hua-Van A, Rouzic A, Maisonhaute C & Capy P (2005). Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet Genome Res* 110: 426-440.

- Hull R (2001). Classifying reverse transcribing elements: a proposal and a challenge to the ICTV. *Arch Virol* 146: 2255-2261.
- Initiative AG (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796-815.
- Jamilena M & Garrido-Ramos M (1994). Molecular relationship between the A and B chromosomes of *Crepis capillaris*. *Heredity* 73: 527-531.
- Jamilena M, Garrido-Ramos M, Ruiz Rejón M, Ruiz Rejón C & Parker JS (1995). Characterisation of repeated sequences from microdissected B chromosomes of *Crepis capillaris*. *Chromosoma* 104: 113-120.
- Jarvik T & Lark KG (1998). Characterization of Soymar1, a mariner element in soybean. *Genetics* 149: 1569-1574.
- Jensen PA, Stuart JR, Goodpaster MP, Goodman JW & Simmons MJ (2008) Cytotype regulation of P transposable elements in *Drosophila melanogaster*: repressor polypeptides or piRNAs? *Genetics* 179: 1785-1793.
- Jiang N, Bao Z, Zhang X., Eddy SR & Wessler SR (2004). Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431: 569–573.
- Jones RN. 1995. Tansley review no. 85: B chromosomes in plants. *New Phytol* 131: 411–434.
- Jones, R. N. & Puertas, M. J. 1993 The B-chromosomes of rye (*Secale cereale* L.). In *Frontiers in plant science research* (ed. K. K.Dhir & T. S. Sareen), pp. 81-112. Delhi: BhagwatiEnterprises.
- Jones RN, Rees H. 1982. B Chromosomes. New York: Academic Press.
- Jones RN, Viegas W & Houben A (2008). A century of B chromosomes in plants: so what? *Ann Bot* 101: 767-75.
- Junakovic N, Terrinoni A, Franco C Di, Vieira C & Loevenbruck C (1998). Accumulation of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. *J Mol Evol* 46: 661-668.
- Jurka J, Kapitonov VV, Pavlicek A, et al. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110: 462-467.

- Kaminker JS, Bergman CM, Kronmiller B, et al. (2002). The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol* 3 : 0084.1-0084.19.
- Kapitonov VV & Jurka J (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet* 9: 411-2; author reply 414.
- Kapitonov VV & Jurka J (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet* 23: 521-529.
- Kapitonov VV & Jurka J (2005). RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS biology* 3: e181.
- Kapitonov VV & Jurka J (2006). Self-synthesizing DNA transposons in eukaryotes. *Proc Natl Acad Sci USA* 103: 4540-4545.
- Kashkush K, Feldman M & Levy AA (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet* 33: 102-106.
- Kayano H (1971). Accumulation of B chromosomes in the germ line of *Locusta migratoria*. *Heredity* 27: 119-123.
- Kidwell MG (1992). Horizontal transfer of P elements and other short inverted repeat transposons. *Genetica* 86: 275-286.
- Kidwell MG (2002). Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49-63.
- Kidwell MG & Lisch DR (2001). Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution* 55: 1-24.
- Kidwell MG, Kidwell JF & Sved JA (1977). Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* 86: 813.
- Kim JM, Vanguri S, Boeke JD, Gabriel A & Voytas DF (1998). Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res* 8: 464-478.
- King M & John B (1980). Regularities and restrictions governing C-band variation in acridoid grasshoppers. *Chromosoma* 76: 123–150.

- Kordis D & Gubensek F (1998). Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *Proc Natl Acad Sci USA* 95: 10704.
- Lamb JC, Riddle NC, Cheng YM, Theuri J & Birchler JA. (2007). Localization and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Res.* 15: 383-398.
- Lampe DJ, Walden KK & Robertson HM (2001). Loss of transposase-DNA interaction may underlie the divergence of mariner family transposable elements and the ability of more than one mariner to occupy the same genome. *Mol Biol Evol* 18: 954-961.
- Langdon T, Seago C, Jones RN, et al. (2000). De novo evolution of satellite DNA on the rye B chromosome. *Genetics* 154: 869-884.
- Le QH, Wright S, Yu Z & Bureau T (2000). Transposon diversity in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 97: 7376-81.
- Lee S-H, Oshige M, Durant ST, et al. (2005). The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci USA* 102: 18075-18080.
- Levin DA, Palestis BG, Jones RN & Trivers R (2005) Phyletic hot spots for B chromosomes in angiosperms. *Evolution* 59: 962-969.
- Lingner J, Hughes TR, Shevchenko A, et al. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276: 561-567.
- Lippman Z & Martienssen R (2004). The role of RNA interference in heterochromatic silencing. *Nature* 431: 364-370.
- Lisch D (2002). Mutator transposons. *Trends Plant Sci* 7: 498-504.
- Lohe AR, Moriyama EN, Lidholm DA & Hartl DL (1995). Vertical inactivation and stochastic loss of mariner-like transposable elements. *Mol Biol Evol* 12:62-72.
- López-León MD, Cabrero J., Camacho JPM, Cano MI & Santos JL (1992). A widespread B chromosome polymorphism maintained without apparent drive. *Evolution* 46: 529–539.

- López-León MD, Cabrero J, Dzyubenko VV, et al. (2008). Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of the grasshopper *Eyprepocnemis plorans plorans*. *Cytogenet Genome Res* 121: 260-265.
- López-León MD, Cabrero J, Pardo M C, et al. (1993). Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity* 71: 352-362.
- López-León MD, Neves N, Schwarzacher T, et al. (1994). Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Res* 2: 87-92.
- Mahillon J & Chandler M (1998). Insertion sequences. *Microbiol Mol Biol Rev* 62: 725–774.
- Mallet F, Bouton O, Prudhomme S, et al. (2004). The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc Natl Acad Sci USA* 101: 1731-1736.
- McAllister BF (1995). Isolation and characterization of a retroelement from B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Insect Mol Biol* 4: 253-262.
- McAllister BF & Werren J.H. (1997). Phylogenetic analysis of a retrotransposon with implications for strong evolutionary constraints on reverse transcriptase. *Mol Biol Evol* 14: 69-80.
- McClintock B (1951). Chromosome organization and genic expression. *Cold Spring Harbor Symposia on Quantitative Biology* 16: 13-47.
- Mi S, Lee X, Li X, et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 403: 785–789.
- Miller WJ & Capy P (2004). Mobile genetic elements as natural tools for genome evolution, in *Mobile Genetic Elements*, Humana Press, pages. 1-20.
- Misra S & Rio DC (1990). Cytotype control of *Drosophila* P element transposition: the 66 kd protein is a repressor of transposase activity. *Cell* 62: 269-84.
- Moran JV, DeBerardinis RJ & Kazazian HH (1999). Exon shuffling by L1 retrotransposition. *Science* 283: 1530-1534.
- Morgante M (2006). Plant genome organisation and diversity: the year of the junk! *Curr Opin Biotechnol* 17: 168-173.

- Muñoz-López M, Siddique A, Bischerour J, et al. (2008) Transposition of Mboumar-9: identification of a new naturally active mariner-family transposon. *J Mol Biol* 382: 567-572.
- Muñoz-Pajares AJ, Martínez-Rodríguez L, Teruel M, et al. (2011). A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* 187: 853-863.
- Nakamura TM, Morin GB, Chapman KB, et al. (1997). Telomerase catalytic subunit homologs from fission yeast and human. *Science* 277: 955-959.
- Noma K, Cam HP, Maraia RJ & Grewal SIS (2006). A role for TFIIC transcription factor complex in genome organization. *Cell* 125: 859-872.
- Nur U (1969). Mitotic instability leading to an accumulation of B-chromosomes in grasshoppers. *Chromosoma* 27: 1-19.
- Nur U, Werren JH, Eickbush D, Burke W & Eickbush TH (1988). A “selfish” B chromosome that enhances its transmission by eliminating the paternal genome. *Science* 240: 512-514.
- Ohzeki J, Nakano M, Okada T & Masumoto H (2002). CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J Cell Biol* 159: 765-775.
- Oosumi T, Garlick B & Belknap WR (1995). Identification and characterization of putative transposable DNA elements in solanaceous plants and *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 92: 8886-8890.
- Orgel LE & Crick FH (1980). Selfish DNA: the ultimate parasite. *Nature* 284: 604-607.
- Östergren G (1945) Parasitic nature of extra fragment chromosomes. *Botaniska Notiser* 2: 157–163.
- Palestis BG, Trivers R, Burt A & Jones RN (2004). The distribution of B chromosomes across species. *Cytogenet Genome Res* 106: 151-158.
- Palestis BG, Cabrero J, Trivers R & Camacho JPM (2010). Prevalence of B chromosomes in Orthoptera is associated with shape and number of A chromosomes. *Genetica* 138: 1181-1189.

- Palestis BG, Burt A, Jones RN & Trivers R (2004). B chromosomes are more frequent in mammals with acrocentric karyotypes: support for the theory of centromeric drive. *Proc R Soc Lond B* 271 Suppl: S22-24.
- Palomeque T, Carrillo AJ, Muñoz-López M & Lorite P (2006) Detection of a mariner-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution. *Gene* 371: 194-205.
- Pardo MC, Viseras E, Cabrero J & Camacho JPM (1995). Mitotic instability of B chromosomes during embryo development in *Locusta migratoria*. *Heredity* 74: 164-169.
- Pardo MC, López-León MD, Cabrero J & Camacho JPM (1994). Transmission analysis of mitotically unstable B chromosomes in *Locusta migratoria*. *Genome* 37: 1027-1034.
- Pardue ML & DeBaryshe PG (2003). Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Ann Rev Genet* 37: 485-511.
- Perfectti F & Werren JH (2001). The interspecific origin of B chromosomes: experimental evidence. *Evolution* 55: 1069-1073.
- Perfectti F, Corral JM, Mesa JA, et al. (2004). Rapid suppression of drive for a parasitic B chromosome. *Cytogenet Genome Res* 106: 338-43.
- Petrov DA, Aminetzach YT, Davis JC, Bensasson D & Hirsh AE (2003). Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol Biol Evol* 20: 880-892.
- Piskurek O & Okada N (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proc Natl Acad Sci USA* 104: 12046-12051.
- Randolph L (1928). Types of supernumerary chromosomes in maize. *Anat Rec* 41: 102.
- Riera L, Petitpierre E, Juan C, Cabrero J & Camacho JPM (2004). Evolutionary dynamics of a B chromosome invasion in island populations of the grasshopper *Eyprepocnemis plorans*. *J Evolution Biol* 17: 716-719.

- Robertson HM & Lampe DJ (1995). Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. *Mol Biol Evol* 12: 850-62.
- Robertson H.M. (1995). The Tc1-mariner superfamily of transposons in animals. *J Insect Physiol* 41: 99–105.
- Robertson HM (1993). The mariner transposable element is widespread in insects. *Nature* 362: 241-245.
- Robertson HM & Zumpano KL (1997). Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene* 205: 203-217.
- Roth DB & Craig NL (1998). VDJ recombination: a transposase goes to work. *Cell* 94: 411-414.
- Rubin GM, Kidwell MG & Bingham PM (1982). The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. *Cell* 29: 987-994.
- SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y & Bennetzen JL (1998). The paleontology of intergene retrotransposons of maize. *Nat Genet* 20: 43-45.
- SanMiguel P., Tikhonov A, Jin YK, et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765-768.
- Santos JL (1980) Variación de la heterocromatina constitutiva en el cariotipo de los Acridoidea y su efecto en el comportamiento cromosómico en meiosis. Ph.D. thesis, Universidad Complutense de Madrid, Spain
- Sasaki T, Matsumoto T, Yamamoto K, et al. (2002). The genome sequence and structure of rice chromosome 1. *Nature* 420: 312–316.
- Schaack S, Gilbert C & Feschotte C (2010). Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* 25: 537-546.
- Schmid M, Ziegler CG, Steinlein C, Nanda I & Haaf T (2002). Chromosome banding in Amphibia. XXIV. The B chromosomes of *Gastrotheca espeletia* (Anura, Hylidae). *Cytogenet Genome Res* 97: 205-218.
- Sen SK, Han K, Wang J, et al. (2006). Human genomic deletions mediated by recombination between Alu elements. *Am J Hum Genet* 79: 41–53.
- Sijen T & Plasterk RHA (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426: 310–314.

- Silva JC & Kidwell MG (2000). Horizontal transfer and selection in the evolution of P elements. *Mol Biol Evol* 17: 1542-57.
- Silva JC, Loreto EL & Clark JB (2004). Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol* 6: 57-71.
- Slotkin RK, Freeling M & Lisch D (2005). Heritable transposon silencing initiated by a naturally occurring transposon inverted duplication. *Nat Genet* 37: 641-644.
- Slotkin RK & Martienssen R (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8: 272–285.
- Smit A. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Develop* 9: 657–663.
- Smit AF & Riggs AD (1996). Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci USA* 93: 1443.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2007). Nucleolus size variation during meiosis and NOR activity of a B chromosome in the grasshopper *Eyprepocnemis plorans*. *Chromosome Res* 15: 755–765.
- Teruel M, Cabrero J, Perfectti F, Acosta MJ, et al. (2009a). Microdissection and chromosome painting of X and B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenet Genome Res* 125: 286-291.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2009b). Quantitative analysis of NOR expression in a B chromosome of the grasshopper *Eyprepocnemis plorans*. *Chromosoma* 118: 291-301.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* 119: 217–225.
- Terzian C, Ferraz C, Demaille J & Bucheton A (2000). Evolution of the Gypsy endogenous retrovirus in the *Drosophila melanogaster* subgroup. *Mol Biol Evol* 17: 908-14.
- Toth M, Grimsby J, Buzsaki G & Donovan GP (1995). Epileptic seizures caused by inactivation of a novel gene, jerky, related to centromere binding protein-B in transgenic mice. *Nat Genet* 11: 71–75.
- Trivers R, Burt A & Palestis BG (2004). B chromosomes and genome size in flowering plants. *J Ecol* 8: 1-8.

- Vagin VV, Sigova A, Li C, et al. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313: 320-324.
- Viseras E, Camacho JPM, Cano MI & Santos JL (1990). Relationship between mitotic instability and accumulation of B chromosomes in males and females of *Locusta migratoria*. *Genome* 33: 23-29.
- Vitte C & Bennetzen JL (2006). Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. *Proc Natl Acad Sci USA* 103: 17638-43.
- Vázquez-Manrique RP, Hernández M, Martínez-Sebastián MJ & Frutos R de (2000). Evolution of gypsy endogenous retrovirus in the *Drosophila obscura* species group. *Mol Biol Evol* 17: 1185-1193.
- Waterston RH, Lindblad-Toh K, Birney E, et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-562.
- Weiler KS & Wakimoto BT (1995). Heterochromatin and gene expression in *Drosophila*. *Annu Rev Genet* 29: 577-605.
- White WJD. 1973. *Animal Cytology and Evolution*, 3rd ed. London: Cambridge University Press.
- Wicker T, Sabot F, Hua-Van A, et al. (2007). A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8: 973-982.
- Wilson E (1907). The supernumerary chromosomes of Hemiptera. *Science* 26: 870-871.
- Yoder JA, Walsh CP & Bestor TH (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 13: 335-340.
- Zurita S, Cabrero J, López-León MD & Camacho JPM (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution* 52: 274-277.

CAPÍTULO 2

Gypsy, RTE and Mariner transposable elements populate *Eyprepocnemis plorans* and *Locusta migratoria* genomes.

Eugenia E. Montiel, Josefa Cabrero, Juan Pedro M. Camacho, M^a Dolores López-León

Introduction

Genetic redundancy is a general property of eukaryotic genomes. The remarkable presence of transposable elements (TEs) found in most species studied, points to the idea that TE expansion across genomes strongly contributes to that redundancy. The abundance of TEs has been estimated in a number of species of plants and animals (Kidwell & Lisch 1997; Kidwell 2005). In humans, particularly, TE elements represent nearly 45% of the genome (Lander et al. 2001) and, in maize, its presence is even higher than 50% (SanMiguel et al. 1996). This high prevalence is not ubiquitous since, for instance, TEs constitute only 15% of the *Drosophila melanogaster* genome (Vieira et al. 1999).

Mobile elements are usually classified into Class I (retrotransposons), which have an RNA intermediate, and Class II (transposons sensu stricto), which directly move as DNA (Wicker et al. 2007). These two types are themselves classified into Orders including different Superfamilies, on the basis of replication strategy. Retrotransposons include elements with long terminal repeats (LTR) and others lacking them (non-LTR), also known as LINES (long interspersed nuclear elements), which are considered the most widespread eukaryotic transposable elements. Other poorly characterized DNA transposons with unknown transpositional mechanism are grouped into the called *MITE* elements or miniature inverted repeat transposable elements (Wessler et al. 1995).

Replicative TE behavior allows them to accumulate in the genomes. This fact is on the basis of their consideration as selfish DNA or genomic parasites, at least at their initial descriptions. But, their contribution to genome size, their role in the evolution of other repetitive sequences and their genetic and chromosomal mutagenic capacity, have turned them into an important evolutionary force contributing to building and reorganizing eukaryotic genomes (Kidwell 2002; Feschotte & Pritham 2007). Although it is believed that most transposable elements must be inactive due to their deleterious effects, evidences for transposon transcription have been frequently obtained (e.g. Grandbastien 1998; Muñoz-López et al. 2008).

TE frequency and long-term population dynamics could be explained by an equilibrium between the self-spreading ability of these parasitic elements and copy elimination due to deleterious effects on genes where they are inserted, as well as the negative consequences of the ectopic recombination and chromosomal rearrangements they promote (Charlesworth et al. 1994; Kidwell & Lish 1997). However, the biological significance of TEs continues under discussion mainly due to increased evidence for a beneficial role to the host genome.

For instance, variability of the human immune system and telomere length maintenance in *Drosophila* depend on “domesticated” mobile elements (Pardue 1995; Kapitonov & Jurka 2005). In *Bombyx mori*, the implication of *TRAS* and *SART* non-LTR retrotransposons in shaping telomere structure has also been suggested (Fujiwara et al. 2005). Furthermore, their presence as normal constituents of regulatory or coding regions of a number of heterochromatic and euchromatic genes (Nekrutenko & Li 2001; Dimitri et al. 2003) has suggested an adaptive role for them.

Transposable elements mostly target heterochromatic regions (Dimitri et al. 2003) where the existence of low gene density and suppressed or low recombination soften their deleterious effects. However, euchromatic regions also harbor these intriguing elements, some of them being preferentially located in genetically active chromosome domains (Rebuzzini et al. 2009).

The two grasshopper species, *Eyprepocnemis plorans* and *Locusta migratoria* have large genomes (Ruíz-Ruano et al. 2011) which, in some individuals, include dispensable chromosomes called B chromosomes. These extra elements have been reported in about 1540 species of plants and animals, and are mainly composed of different types of repetitive DNA (for review, see Camacho 2005). Due to their large genome size and the heterochromatic nature of B chromosomes, both species are good candidates to harbour TEs. *E. plorans* shows a complex B chromosome system with many different B variants mainly composed of a 180bp satellite DNA and ribosomal DNA (López-León 1994; Cabrero et al. 1999). The molecular composition of B chromosomes in *L. migratoria* remains nearly unknown, excepting the presence of genes for H3 and H4 histone proteins (Teruel et al. 2010). Here we try to go forward in the knowledge of the molecular nature of *E. plorans* and *L. migratoria* genomes, and the B chromosomes they carry, by analyzing the abundance and chromosome distribution of *Gypsy* (LTR retrotransposon), *RTE* (non-LTR retrotransposon) and *Mariner* (transposon) elements. Our research thus represents an opportunity to explore the relationship between two kinds of selfish genetics elements, TEs and B chromosomes, in these two species.

Materials and Methods

Materials and DNA isolation

Adult specimens of *Eyprepocnemis plorans* were collected in Torrox (Málaga, Spain) and Salobreña (Granada), and those of *Locusta migratoria* were caught in Padul (Granada). Adults were bred in the laboratory in order to obtain embryos. Testes were fixed in 1:3 acetic acid-ethanol. Females were injected with 0.1 ml of 0.05% colchicine in insect saline solution, 6 h prior to dissection and fixation of ovarioles in 1:3 acetic acid-ethanol. The number of B chromosomes was analysed in squash preparations of two testis tubules or one ovariole in acetic orcein. Egg pods were incubated at 28° C for twelve (*E. plorans*) or six (*L. migratoria*) days, after which embryos were fixed in 3:1 ethanol-acetic acid, for cytogenetic studies, as described in Camacho et al. (1991). Adult bodies were frozen at -80° C for DNA and RNA isolation. Genomic DNA from both species was extracted using the Gen Elute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich).

PCR amplification

In order to amplify members of *Gypsy* and *Mariner* TE families we used primers designed for different species by several authors (PRO / HA pair for *Gypsy* and R2-1 / 28S-R2 pair for *Mariner*). To amplify *RTE* elements, however, we designed primers (RTE-F and RTE-R) for the reverse transcriptase gene using the CODEHOP software (Rose 2003). Table 3 shows all primers used.

Table 3. Primers used for amplification and real-time qPCR assays of different types of transposable elements

Primer name	Sequence 5'->3'	Reference
PRO Gypsy	GTKTTIKTIGAYACIGGIKC	(Miller et al. 1999)
HA Gypsy	GCGTTIKTIAGICCRAAIGGCAT	(Miller et al. 1999)
RTE-F RTE	CTGGACGCAGAGAARGCVTTYGAC	-
RTE-R RTE	CGGGCTCAGNGGWCANCCYTGA	-
R2-1 Mariner	CGCGCATGAATGGATTAACG	(Burke et al. 1993)
28S-R2 Mariner	AAGAGCCGACATCGAAGGATC	(Burke et al. 1993)
qPCR_GypF	GTTCGTATAGGATGTCGCCAG	-
qPCR_GypR	TTAGCACCCCTTCAAGCAGTC	-
Gyp_Int900F	GATACGGGCGCAACTGTTCTG	-
Gyp_Int1300R	TTAGGGATAAGTTGAGACTGCTC	-
qPCR_RTEF	AGATTGGGAAACGAGGCACTG	-
RTE_IntR	CATCCATACAAGGCAACACTC	-
qPCR_MarF	TATACCGTTAGCAGAGCCAGTTG	-
qPCR_MarR	TACGAAAGAAGCGGCGACAC	-

PCR experiments were performed in a 50 µl reaction containing 50 ng of genomic DNA, 0.2 µM of each primer, 0.2 mM dNTPs and 1 unit of Taq polymerase (MBL). PCR conditions for *Gypsy* elements were: initial denaturation for 4 min at 80°C, 30 cycles of 30 s at 45°C, 1 min at 74°C, and 30 s at 94°C, followed by 3 min at 45°C and a final extension for 10 min at 74°C. For *Mariner* and *RTE*, PCR conditions were: initial denaturation 5 min at 94°C, 30 cycles of 30 s at 94°C, 30 s at 45°C, and 2 min at 72°C, and a final elongation 7 min at 72°C. PCR products were visualized in a 1% agarose gel.

DNA cloning and sequence analyses

PCR products obtained from *E. plorans* genomic DNA were cloned. For *Gypsy* elements, we excised two bands (1,000 and 1,300 bp) from the gel, and purified them with the GenElute Gel Extraction Kit (Sigma-Aldrich). For *RTE* and *Mariner* elements, we cleaned the PCR product with the GenElute PCR Clean-up kit (Sigma-Aldrich). All fragments were cloned into pCR2.1-TOPO vector (Invitrogen). Plasmid DNA was obtained with the Gen Elute Mammalian Genomic DNA Miniprep kit (Sigma). Clones obtained were sequenced at the DNA Sequencing Service of the Department of Genetics (University of Granada) and DNA sequences were analysed with BioEdit (Hall et al. 1999) and deposited in the GeneBank database (accession numbers: JN604089 for EploGypI, JN604088 for EploGypII, JN604091 for EploRTE5, JN604090 for EploRTE20 and JN604094 for EploMar20). Clone similarity to TEs was tested subjecting DNA sequences to BLAST and BLASTX tools (Altschul et al. 1997) at the NCBI site. To confirm and refine the identity of the elements found within the vast families of LTR and non-LTR retrotransposons, we tested cloned DNA sequences with the CENSOR online software (Jurka et al. 2005).

Table 4. Description of clones used in this work.

Accessions Number	Clone Name	Fragment Length bp	Primers
JN604089	EploGypI	1107	PRO/HA
JN604088	EploGypII	1302	PRO/HA
JN604091	EploRTE5	1242	RTE-F/RTE-R
JN604090	EploRTE20	932	RTE-F/RTE-R
JN604094	EploMar20	1112	R2-1/28S-R2
JN604093	EploGyp_1	432	Gyp_Int900F/Gyp_Int1300R
JN604092	EploRTE_1	218	qPCR_LineF/LINE_IntR

RNA isolation and expression analysis

Total RNA was isolated from 30 mg of frozen body using the Real total RNA spin plus kit (Durviz), following manufacturer's protocol including DNAase I treatment. RNA was later submitted to a second treatment with 20 units of DNase I to ensure total elimination of genomic DNA. The total amount and quality of RNA obtained was checked in a denaturing agarose gel. cDNA was obtained with SuperScript III First-Strand Síntesis SuperMix (Invitrogen), according to manufacturer's recommendations, using random hexamers.

TE expression was investigated, in both species, by PCR amplification on the obtained cDNA, using the primer pairs Gyp_Int900F / Gyp_Int1300R for *Gypsy*, qPCR_LinrF / RTE_IntR for *RTE*, and qPCR_MarF / qPCR_MarR for *Mariner* elements (see Table 1 for DNA sequences). PCR reactions were performed with 10 ng of cDNA, 0.25 μ M of each primer, 0.2 mM dNTPs, 1,25 mM MgCl₂ and 1U of Taq (MBL), using a thermocycler program with an initial denaturation step at 94°C for 5 min. and 30 cycles of 30 s at 94°C, 30 s at 55°C, and 30 s at 72°C with a final step at 72°C for 8 min. All amplifications performed included a control reaction with RNA to rule out DNA contamination of original RNA samples.

Copy number analysis

Copy number estimation for transposable elements was performed by real-time quantitative PCR (qPCR). DNA templates were obtained in *E. plorans* from three B-carrying and three B-lacking males and, in *L. migratoria*, from two B-carrying females. In *L. migratoria*, we opted for using females because B chromosomes show lower mitotic instability (and thus lower intraindividual variation in B number) in them (Viseras et al. 1990). DNA was quantified using PicoGreen® dsDNA Quantitation Reagent (Molecular Probes) in a fluorometer TBS-380, according to manufacturer's recommendations.

Real-time quantitative PCR assays contained 10 μ l of 2X SensiMixPlus SYBR (Quantace), 0.25 μ M of each primer and 10 ng of DNA sample in a 20 μ l final volume. Primers employed for *Gypsy* quantification were Gyp_Int900F and Gyp_Int1300R that yield a 476 bp amplicon. *RTE* and *Mariner* elements amplified 217 bp and 284 bp amplicons using qPCR_RTEF / RTE_IntR and qPCR_MarF / qPCR_MarR primers, respectively.

PCR programs employed consisted of a denaturation step at 95°C for 10 min, and 40 cycles of 15 s at 95°C, 15 s at 60°C (*Gypsy*, *RTE*) or 60.4°C (*Mariner*), 72°C for 30 s and a final step at 72°C for 5 min. Real-time qPCR assays were run using a Chromo4 (Biorad) thermocycler using the Opticon Monitor v3.1 software to export raw data and the LingRegPCR 11.0 program to analyse them. Each sample was amplified in triplicate and product specificity was verified through the generation of a melting curve following Real-time PCR.

Copy number estimations were performed through analysis with the LingRegPCR software, following the method described by Feng et al. (2008), which implies obtaining a calibration factor (CF) by amplification of reference samples with known copy number and the subsequent conversion of initial fluorescence values (F_0) to initial copy numbers (N_0). The CF was obtained with a serial dilution of clones which were inserted with the region to amplify: EploGyp_1, EploRTE_1 and EploMar20. Table 5 shows the obtained CF values.

We obtained the copy number of *Gypsy*, *RTE* and *Mariner* elements in both species on the basis of C values in *E. plorans* (10160 Mb) and *L. migratoria* (5850 Mb) and bearing in mind recent estimates of the size of X and B chromosomes (Ruíz-Ruano et al. 2011). Since both species show XX/X0 sex chromosome determinism and individuals could carry B chromosomes, and in order that these results were comparable among the different individuals (males and females with or without Bs), we divided the estimated copy number per nanogram between the 2C value calculated for each individual according to the number of X and B chromosomes that it carried.

Table 5. CFLinReg and Coefficient of Variation (CV) for TEs in the two grasshopper species.

TE	<i>E. plorans</i>		<i>L. migratoria</i>	
	CFLinReg	CV	CFLinReg	CV
Gypsy	2,13E-03	21,36%	2,99E-03	16,72%
RTE	2,06E-03	16,59%	2,43E-03	10,44%
Mariner	2,40E-02	16,87%	1,41E-02	17,99%

Fluorescence in situ hybridization (FISH)

Chromosome preparations were performed from embryos and spermatocytes according to Camacho et al. (1991) and Cabrero et al. (1999), respectively. Slides were dehydrated in a series of 70%, 90% and absolute ethanol and then incubated in an oven at 60° C overnight. We used the insert from the following clones as probes: EploGypI for *Gypsy*; EploRTE5 for *RTE* and EploMar20 for *Mariner* elements. All these probes were labeled

by nick translation using tetramethylrhodamine-11-dUTP (Roche) and hybridization signals were directly detected. Chromosomes were counterstained with DAPI (4', 6 diamidino-2-phenylindole). About 250 ng of DNA probe was used in each FISH experiment, following the technique described in Cabrero et al. (2003a). The analysis of chromosome preparations was made using a BX41 Olympus fluorescence microscope, and photographs were taken with an associated DP70 cooled camera. Images were afterwards optimized with THE GIMP freeware.

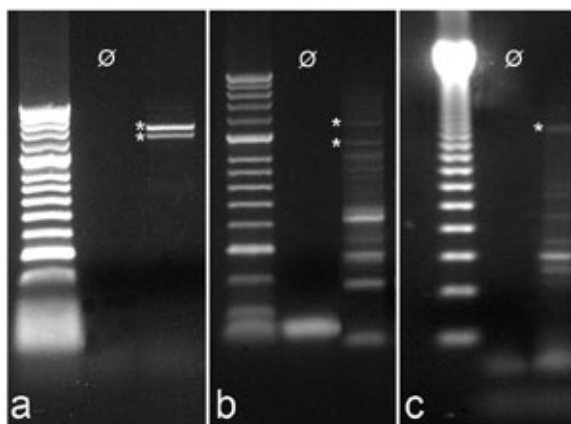


Figure 3. Electrophoresis of PCR products after amplification of Gypsy (a), RTE (b) and Mariner (c) elements with degenerate primers in *E. plorans*. Asterisks indicate the bands extracted for cloning.

Results

Presence of transposable elements: PCR amplification in E. plorans and DNA sequence analysis

For *Gypsy*-like elements, the PRO/HA primers pair, targeting conserved motifs of protease and retrotranscriptase genes, successfully amplified two DNA fragments of 1000 (EplOGypI) and 1300 bp (EplOGypII) (Figure 3a). The sequence analysis performed to these two DNA sequences revealed homology at protein level with Ty3/*gypsy*-like elements from *Drosophila melanogaster*, *Strongylocentrotus purpuratus* and *Danio rerio*, which was confirmed by the CENSOR software. PCR amplification of LINE-like elements yielded two fragments of 1180 (EplORTE5) and 900 bp (EplORTE20) (Figure 3b) which, after cloning and sequencing, turned out to be highly related to non-LTR retrotransposons present in the fish *Oryzias latipes*, *Mus musculus* and *Homo sapiens*. The analysis with CENSOR software indicated that these clones specifically contained members of the *RTE* family. In the following analyses we will focus on EplOGypI and EplORTE5 clones (Table 4).

Mariner elements were identified in the EplOMar20 clone (1100bp) (Figure 3c) obtained by amplification with the R2-1/28S-R2 primers (Table 3). This DNA sequence showed similarity with *Mariner* elements from several insect species.

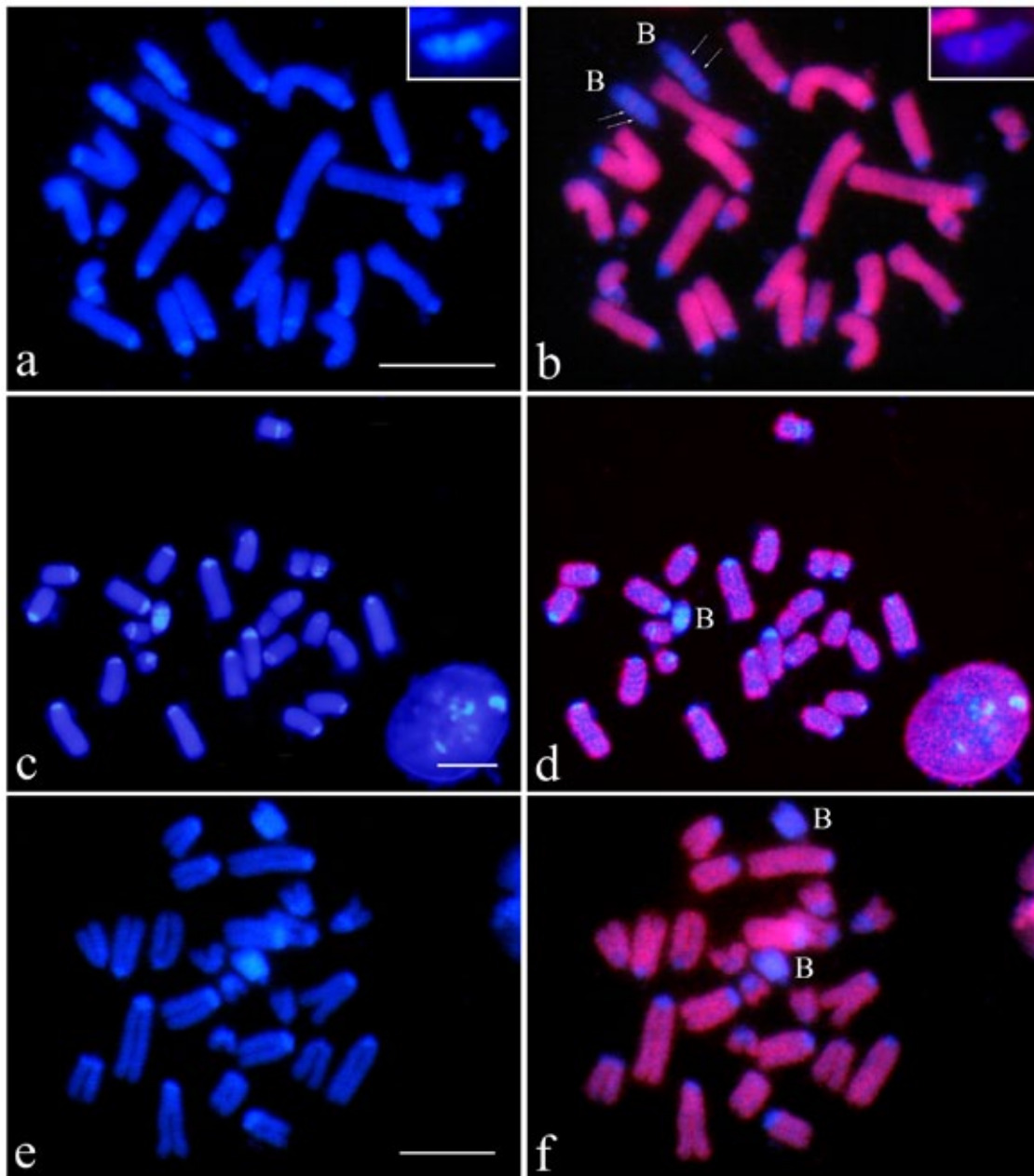


Figure 4. FISH on embryo mitotic metaphase chromosomes of *E. plorans*. Left column shows DAPI staining (a, c, e). Right column shows DAPI+FISH with the EploGypI (b), EploRTE5 (d) and EploMar20 (f) probes. Note in (a) that the B₂₄ chromosome shows three DAPI+ bands and, in (b), that it carries two interstitial RTE-carrying bands (arrows). Inset in (a) and (b) show a B₂ chromosome showing two DAPI+ bands and a single RTE-carrying interband. Note in (b), (d) and (f) the absence of FISH signals in the pericentromeric regions of all A chromosomes. Bars= 5 μ m

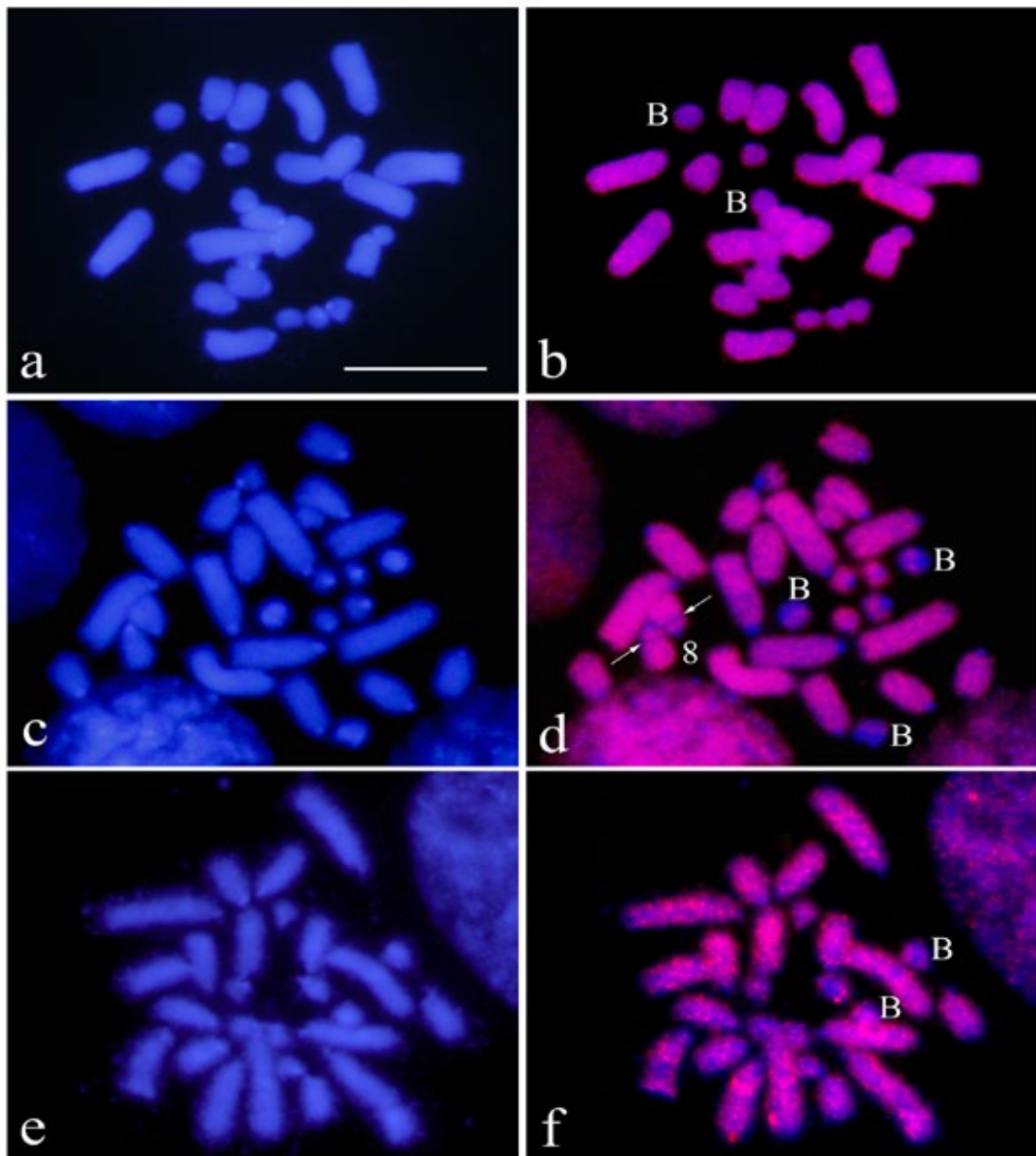


Figure 5. FISH on embryo mitotic metaphase chromosomes of *L. migratoria*. Left column shows DAPI staining (a, c, e). Right column shows DAPI+FISH with the EplGypI (b), EplORTE5 (d) and EplMar20 (f) probes. Note in (d) the absence of RTE elements from an interstitial band in chromosome 8 (arrows), where the histone gene cluster is located. Bar= 5 μ m

Chromosome distribution of Transposable Elements

FISH experiments showed high abundance of the three TEs analyzed in both *E. plorans* and *L. migratoria* chromosomes. Chromosome distribution patterns for all three elements were about similar in both species, with hybridization signals being dispersed over the whole euchromatic regions of all standard (A) chromosomes. In general, no small clustering yielding a band-like pattern was observed for these three TEs (Figure 4 and Figure 5).

The retrotransposon *Gypsy* showed strong signals covering the euchromatic regions of all chromosomes in *E. plorans*, but it was absent from pericentromeric heterochromatin in all chromosomes, including the nucleolar organizing regions (NOR) located on S₉, S₁₀, S₁₁ and X chromosomes. In the B₂₄ two interstitial signals were detected between the three DAPI+ heterochromatic blocks (Figure 4a,b). The B₂ chromosome variant, from which B₂₄ arose (Zurita et al. 1998), however, shows only two DAPI+ bands and a single interstitial region of *Gypsy* TEs (see inset in Figure 4a,b). *RTE* and *Mariner* elements were also abundant in euchromatic regions of *E. plorans* chromosomes (Figure 4c-f). For these TE elements, there are no remarkable differences in signal intensity on A chromosomes and, as for the *Gypsy* element, they were absent from heterochromatin and NOR regions. They are also poorly represented on B₂₄ and, again, they were confined to narrow regions between the DAPI+ blocks (Figure 4c-f).

In *L. migratoria*, *Gypsy*, *RTE* and *Mariner* elements showed hybridization signals distributed along the entire A chromosomes except for the small pericentromeric heterochromatin regions (Figure 5). In chromosome 8, which interstitially carries the histone gene cluster (Teruel et al. 2010), it was also apparent the absence of *RTE* elements in this region (Figure 5d). The higher heterochromatin content of B chromosomes was associated with lower hybridization for the three TEs, especially for *RTE* (Figure 5c,d).

Expression of transposable elements

We investigated the transcriptional activity of transposable elements, *Gypsy*, *RTE* and *Mariner*, searching for the presence of specific transcripts in the total RNA isolated from both species. cDNA obtained from RNA retrotranscription was submitted to PCR amplification with primers specific to *Gypsy* and *RTE* retrotranscriptases and *Mariner* transposase. Isolated RNA was used as a PCR negative control to discard DNA contamination and genomic DNA as a positive control for the suitability of the primers. The three mobile elements were successfully amplified in cDNA samples from both

species, *Gypsy* showing the most intense band and *Mariner* the least intense one (Figure 6). DNA sequencing of these bands confirmed their identity as the expected TE elements (accession number: JN850529 for LmigGyp, JN850531 for LmigRTE, JN850530 for LmigMar, JN850532 for EploGyp, JN850533 for EploRTE and JN850534 for EploMar).

Copy number estimations

The FISH technique suggested possible differences in the abundance of the different TE elements, but it is not a quantitative assay. Therefore, we performed an estimation of TE copy number by means of qPCR.

We observed remarkable differences in the number of copies for these elements (Fig. 5). In *E. plorans*, the *Gypsy* element is the most frequent (7001.6 ± 681.4 copies), followed by the *RTE* element showing about half copy number (4199.5 ± 231.1) and *Mariner* with a much smaller amount (344.0 ± 44.0). By contrast, in *L. migratoria*, *Mariner* is the most frequent element (223.3 ± 2.8), followed by *Gypsy* (37.5 ± 5.9) and *RTE* (3.5 ± 0.2). *Mariner* was the only element showing about similar copy number values in both species, whereas both *Gypsy* as *RTE* showed at least 100 times fewer copies in *L. migratoria* than *E. plorans*. This is a logical result since the primers used for this experiment in *L. migratoria* had been designed on the basis of *E. plorans* DNA sequence. Therefore, the elements detected in *L. migratoria* are presumably a subset conserving homology with primer DNA sequence.

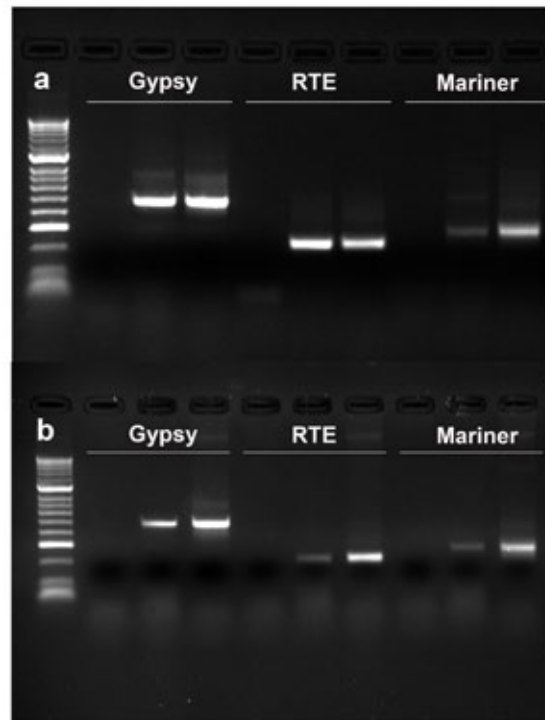


Figure 6. TE transcription in *E. plorans* (a) and *L. migratoria* (b). Three lanes are shown for each element, corresponding to PCR amplification on RNA (left), cDNA (center) and genomic DNA (right).

Discussion

We have shown that three TEs (Class I *Gypsy* and *RTE*, and Class II *Mariner*) are abundant in the genomes of the grasshoppers *E. plorans* and *L. migratoria*. Electrophoresis in agarose gel of the PCR products yielded several bands, especially for *RTE* and *Mariner* elements (see Figure 3), which may reflect the presence of degenerate copies with deletions and/or insertions in both species, in consistency with the well known fact that genomes are plenty of truncated or degenerated inactive TEs (Kazazian 1999).

The dynamics of TE expansion across genomes varies depending on the species and element type considered. Hence, *Gypsy* elements have been observed dispersed (Cermak et al. 2008) or clustered (Belyayev et al. 2001). Likewise LINEs have been found dispersed over all chromosomes (Rebuzzini et al 2009) or in restricted regions, such as G bands in mammals (Baker and Kass, 1994). *Mariner* may also be concentrated on specific regions, such as centromeres or distal regions (Mukabayire and Besansky, 1996; Jacobs et al. 2004) or scattered over chromosomes including or not the heterochromatic regions (Torti et al. 2000). All three TEs analysed in *E. plorans* and *L. migratoria* showed a dispersed but non-random chromosomal distribution, since FISH signals were observed scattered over euchromatic chromosome regions, but were absent from heterochromatin, including NORs (in *E. plorans*) and histone gene clusters (in *L. migratoria*). B chromosomes in both species showed hybridization signals but to a much lesser extent, especially in *E. plorans*, which could be explained by their mostly heterochromatic nature. We cannot rule out a low, undetectable by FISH, presence of some copies of these TEs in heterochromatic regions.

FISH signals suggested similar abundance in the euchromatin from different A chromosomes, contrary to the scenario described in other species where differential accumulation on sex chromosomes has been reported (Cermak et al. 2008). In *Drosophila* (Pardue 1995) and *Bombyx mori* (Fujiwara et al. 2005) chromosome ends are enriched in or associated with non-LTR elements. Likewise, the centromeric retrotransposon family (CR) is overrepresented in the centromere of many cereal species (Gaio et al. 2009). The ubiquitous euchromatic TE distribution, observed in *E. plorans* and *L. migratoria*, has also been reported in other species (Rebuzzini et al. 2009), and is quite interesting considering that TEs can be deleterious by affecting host gene expression and causing gene and chromosomal mutations (Kidwell and Lish 1997, 2000, Feschotte and Pritham 2007). However, perhaps euchromatin in giant genomes is plenty of potentially tolerable insertion sites for TEs. The genomes of *L. migratoria* and *E. plorans* are about two and

three times, respectively, larger than the human genome, and TEs could find abundance of intergenic spacers and pseudogenes dispersed over euchromatin. In *Drosophila melanogaster*, most TEs located in euchromatin are not inserted into genes, yielding a very low proportion of chimera TE-gene transcripts (Lipatov et al. 2005). It has been shown that ancient TE elements tend to accumulate in euchromatic regions associated with genes, whereas new and more deleterious elements tend to colonize the heterochromatin, thus avoiding the insertional mutagenesis effects (Santiago et al. 2002). On this basis, the presence of TEs in the euchromatin of *E. plorans* and *L. migratoria* could be an ancient event, presumably preceding these grasshoppers diversification. It is also conceivable that TEs play a beneficial role to the host genomes, by being part of the genes and their regulatory elements, such as it occurs, for instance, for MITEs being regulatory sequences with a binding-site function (Bureau and Wesler, 1994). Furthermore, TEs may generate lineage-specific regulatory sequences (Mariño-Ramirez et al. 2005) and even many genes seem to have derived from TEs such as *Mariner*, hAT or P elements (for review, see Feschotte and Pritham 2007).

The contribution of TEs to increasing genome size is evidenced by a direct, but not linear, relationship between the proportion of TEs in a genome and its C-value. For instance, barley (*Hordeum vulgare*) has a genome size of 5000 Mb and more than 50% of TEs, whereas *Drosophila melanogaster* has a genome with only 180 Mb and 15 % TEs (for review see Kidwell 2002). Quantitative PCR experiments in *E. plorans* revealed that *Gypsy* elements are the most frequent, followed by *RTE* and *Mariner*. Similar analyses in *L. migratoria* indicated, however, that *Mariner* is more frequent in this species, and copy number estimations were much lower than those in *E. plorans*, especially for *Gypsy* and *RTE*. This is a logical consequence of using primers that had been designed on the basis of *E. plorans* DNA sequence. These primers are expected to amplify only those elements conserving the same DNA sequence as in *E. plorans* at primer regions. On this basis, it was remarkable to find about similar copy numbers for *Mariner* in both species. This could be due to a different degree of conservatism for DNA sequences in Class I and Class II elements. Class I elements (e.g. *Gypsy* and *RTE*) are highly variable due to its amplification mechanism mediated by a reverse transcription step with mutation rates similar to those in virus (Drake et al 1998). The appearance of new variants would favor the divergence of families across species (Volf et al. 2000). This implies that the number of TE elements detected by qPCR is a subset of those detected by FISH since the latter technique also detects those variants showing partial homology. In *L. migratoria*, this subset is particularly small since it is restricted to those elements conserving, at primer regions, a DNA sequence homologous to that in *E. plorans*. In Class II elements (e.g.

Mariner), however, the dispersion using “cut and paste” as well as potential events of horizontal transmission contribute to sequence conservation among species (Hartl et al. 1997). Therefore, the higher conservatism expected for *Mariner* explains the almost similar copy number estimated in *L. migratoria* and *E. plorans*. In addition, this conservatism might be due to possible functional constraints operating on these elements, because these species belong to different grasshopper subfamilies which are separately evolving for the last 50 million years (Sharov 1968).

The abundance of TEs in the euchromatin of *E. plorans* and *L. migratoria* contrasts with their paucity in the heterochromatin, suggesting that these kinds of mobile elements have not played an important role in building the heterochromatin in both species, in contrast to evidence in other species (Hsieh and Fire 2000). A possible explanation is that TE spread in these grasshopper species preceded heterochromatin expansion. In *E. plorans*, pericentromeric heterochromatin is mainly made up of two repetitive DNAs: i) a 180 bp satDNA which is absent in *E. unicolor* and other Eypreocnemidinae species (Cabrero et al. 2003a), and ii) rDNA, which is restricted to chromosomes S₉ and S₁₁ in specimens from Dagestan (Caucasus) but has expanded to pericentromeric heterochromatin regions of most A chromosomes in Spanish populations (Cabrero et al. 2003b). It is thus likely that TE dispersion in the *E. plorans* genome preceded the expansion of these two repetitive DNAs, for which reason the pericentromeric heterochromatin regions are mostly empty of TEs (see Fig. 2). Alternatively, the absence of appropriate targets for TE integration in these repetitive DNAs is also conceivable. In *L. migratoria*, the precise molecular nature of pericentromeric chromatin is so far unknown, although the presence of DNA sequences belonging to the Cot1 fraction suggests that these chromosome domains are also made up of repetitive DNA sequences (López-León et al. unpublished).

The presence of TEs has been reported on B chromosomes, and TEs have been suggested as possible factors explaining B chromosome mutability (Camacho 2005, Lamb et al. 2007). *E. plorans* B chromosomes show high mutability (Henriques-Gil and Arana 1990; López-León et al. 1993; Bakkali and Camacho 2004) with new B variants appearing at rates as high as 0,001 per generation. The fact that the B₂₄ chromosome carries three DAPI+ bands lacking TEs and two interstitial bands carrying them, whereas the B₂ chromosome (from which B₂₄ arose) carries only two DAPI+ bands and one transposon band (see Figure 4), suggests that the mutation becoming B₂ into B₂₄ involved the duplication of one of the DAPI+ bands plus the transposon band. This is at least suggestive of a possible role of TEs in B chromosome instability.

The non-random distribution of TEs and their presence on potentially active euchromatic regions could suggest functionality. In fact, we have identified specific transcripts for *Gypsy*, *RTE* and *Mariner* elements in RNA isolated from both species, *Gypsy* yielding the most conspicuous band in the gel, in consistency with its higher copy number. Vicient (2010) found in maize that *Gypsy*-like retrotransposons exhibited the highest transcriptional activity among 56 families of TEs analysed. Transcription rate probably depends on multiple factors such as transpositional efficiency, time since genome invasion or the existence of molecular domestication. Although probably most TEs harboured by eukaryote genomes are deleterious or inactive due to epigenetic silencing mechanisms developed by host genomes, transcriptional activity has been described for many of them in several species. In some cases the chromosome from which TEs were transcribed has been shown. For instance, Lamb et al. (2007) reported transcriptional activity for a retrotransposon-derived element (*StarkB*) specific to maize B chromosomes. Despite having found transcripts for *Gypsy*, *RTE* and *Mariner* in RNA from *E. plorans* and *L. migratoria*, this does not guarantee their transpositional activity because they could be silenced, for instance, at postranscriptional level through small interfering RNA (Kasschau et al. 2007). This opens new pathways for future research.

Acknowledgements

This study was supported by grants from the Spanish Ministerio de Ciencia e Innovación (CGL2009-11917) and Junta de Andalucía (CVI-6649), and was partially performed by FEDER funds. E.E. Montiel was supported by a Junta de Andalucía fellowship. We thank David Martinez for English corrections.

References

- Altschul SF, Madden TL, Schäffer A et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402
- Baker RJ, Kass DH (1994) Comparison of chromosomal distribution of a retrotransposon (LINE) and a retrovirus-like element *mys* in *Peromyscus maniculatus* and *P. leucopus*. *Chromosome Res* 2:185-189
- Bakkali M, Camacho JPM (2004) The B chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa: III Mutation rate of B chromosomes. *Heredity* 92:428-433

- Belyayev A, Raskina O, Nevo E (2001) Chromosomal distribution of reverse transcriptase-containing retroelements in two Triticeae species. *Chromosome Res* 9:129-136
- Bureau TE, Wessler SR (1994) *Stowaway*: a new family of inverted repeat elements associated to genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* 6:907-916
- Burke WD, Eickbush DG, Xiong Y, Jakubczak J, Eickbush TH (1993) Sequence relationship of retrotransposable elements R1 and R2 within and between divergent insect species. *Mol Biol Evol* 10:163-85
- Cabrero J, López-León MD, Bakkali M, Camacho JPM (1999) Common origin of B chromosomes variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* 83:435-439
- Cabrero J, Bugrov A, Warchałowska-Sliwa E, López-León MD, Perfectti F, Camacho JPM (2003a) Comparative FISH analysis in five species of Eyprepocnemidine grasshoppers. *Heredity* 90:377-381
- Cabrero J, Perfectti F, Gómez R, Camacho JPM, López-León MD (2003b) Population variation in the A chromosome distribution of satellite DNA and ribosomal DNA in the grasshopper *Eyprepocnemis plorans*. *Chromosome Res* 11:375–381
- Camacho JPM, Cabrero J, Viseras E, López-León ND, Navas-Castillo J, Alché JD (1991) G banding in two species of grasshopper and its relationships to C, N and fluorescent banding techniques. *Genome* 34:638-643
- Camacho JPM (2005) B chromosomes. In: Gregory TR (ed) *The evolution of the genome*. San Diego: Elsevier, pp 223-286
- Cermak T, Kubat Z, Hobza R et al (2008) Survey of repetitive sequences in *Selene latifolia* with respect to their distribution on sex chromosomes. *Chromosome Res* 16:961-976
- Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215-220
- Dimitri P, Jukanovic N, Arcà B (2003) Colonization of heterochromatic genes by transposable elements in *Drosophila*. *Mol Biol Evol* 20:503-512
- Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. *Genetics* 148:1667-1686

- Feng J, Zeng, R, Chen, J (2008) Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. *Biotechniques* 44:901-912
- Feschotte C, Pritham EJ (2007) DNA Transposons and the evolution of eukaryotic Genomes. *Annu Rev Genet* 41:331-368
- Fujiwara H, Osanai M, Matsumoto T, Kojima KK (2005) Telomere-specific non-LTR retrotransposon and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res* 13:455-467
- Gaio D, Gill N, Kim HR et al (2009) A lineage-specific centromere retrotransposon in *Oryza brachyantha*. *Plant J* 60:820-831
- Grandbastien MA (1998) Activation of plant retrotransposons under stress conditions. *Trends Plant Sci* 3:181-187
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41:95-98
- Hartl DL, Lohe AR, Lozovskaya ER (1997) Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31:337-358
- Henriques-Gil N, Arana P (1990) Origin and substitution of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Evolution* 44:747-753
- Hsieh J, Fire A (2000) Recognition and silencing of repeated DNA. *Annu Rev Genet* 34:187-204
- Jacobs G, Dechryeva D, Menzel G, Dombrowski, Schmidt T (2004) Molecular characterization of Vulmar 1, a complete *mariner* transposon of sugar beet and diversity of *mariner* and En/Spm like sequences in the genus Beta. *Genome* 47:1192-1201
- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110:462-467
- Kapitonov VV, Jurka J (2005) RAGI core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biology* 3:e181
- Kasschau KD, Fahlgren N, Chapman EJ et al (2007) Genome-wide profiling and analysis of Arabidopsis siRNA. *PLOS Biol* 5:e57

- Kazazian HH (1999) An estimated frequency of endogenous insertional mutations in humans. *Nat Genet* 22:130
- Kidwell MG (2002) Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115:49-63
- Kidwell MG (2005) Transposable elements In: Gregory TR (ed) *The evolution of the genome*. San Diego: Elsevier, pp 165-221
- Kidwell MG, Lish D (1997) Transposable elements as sources of variation animals and plants. *Proc Natl Acad Sci USA* 94:7704-7711
- Kidwell MG, Lish D (2000) Transposable elements and genome evolution. *TREE* 15:95-99
- Lamb JC, Riddle NC, Cheng YM, Theuri J, Birchler JA (2007) Location and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Res* 15:383-398
- Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921
- Lipatov M, Lenko K, Petrov DA, Bergman CM (2005) Paucity of chimeric gene-transposable element transcripts in the *Drosophila melanogaster* genome. *BMC Biology* 3:24
- López-León MD, Cabrero J, Pardo MC, Viseras E, Camacho JPM, Santos JL (1993) Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity*: 71:352-362
- López-León MD, Neves N, Schwarzacher T, Heslop-Harrison JS, Hewitt GM, Camacho JPM (1994) Possible origin of B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Res* 2:87-92
- Mariño-Ramírez L, Lewis KC, Landsman D, Jordan IK (2005) Transposable elements donate lineage-specific regulatory sequences to host genomes. *Cytogenet Genome Res* 110:333-341
- Miller K, Lynch C, Martin J, Herniou E, Tristem M (1999) Identification of multiple *Gypsy* LTR-retrotransposon lineages in vertebrate genomes. *J Mol Evol* 49:358-66

- Mukabayire O, Besansky NJ (1996) Distribution of T1, Q, Pegasus and *mariner* transposable elements on the polytene chromosomes of PEST, a standard strain of *Anopheles gambiae*. *Chromosoma* 104:585-595
- Muñoz-López M, Siddique A, Bischerour J, et al (2008) Transposition of Mboumar-9: identification of a new naturally active *mariner*-family transposon. *J Mol Biol* 382:567-72
- Nekrutenko A, Li WH (2001) Transposable elements are found in a large number of human protein-coding genes. *Trends Genet* 17:619-621
- Pardue ML (1995) Drosophila telomeres: another way to end it all. In: Blackburn EH, Graider CW (eds) *Telomeres*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp 339-370
- Rebuzzini P, Castiglia R, Nergadze G et al (2009) Quantitative variation of LINE-1 sequences in five species and three subspecies of the subgenus *Mus* and in five Robertsonian races of *Mus musculus domesticus*. *Chromosome Res* 17:65-76
- Rose T (2003) CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res* 31:3763-3766
- Ruiz-Ruano F J, Ruiz-Estévez M, Rodríguez-Pérez J, López-Pino JL, Cabrero J, Camacho JPM (2011) DNA Amount of X and B Chromosomes in the Grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria* *Cytogenet Genome Res* 134:120-126
- SanMiguel P, Tikhonov A, Jin YK et al (1996) Nested retrotransposons in the intergenic regions of the maize genomes. *Science* 274:756-768
- Santiago M, Herraiz C, Goñi JR, Messeguer X, Casacuberta JM (2002) Genome wide analysis of the *Emigrant* family of MITEs of *Arabidopsis thaliana* *Mol Biol Evol* 19:2285-2293
- Sharov AG (1968) Phylogeny of Orthopteroid insects. *Akad Sci USSR (Moskow) Proc Paleont Inst* 18:1-216
- Teruel M, Cabrero J, Perfectti F, JPM Camacho (2010) B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* 119:217–225

- Torti C, Ludvik M, Gomulski M, Maralli D, Raimondi E, Robertson HM (2000) Evolution of different subfamilies of *mariner* elements within mdfly genome inferred from abundance and chromosomal distribution. *Chromosoma* 108:523-532
- Vicient C (2010) Transcriptional activity of transposable elements in maize. *BMC Genomics* 11:601-610
- Vieira C, Lepetit D, Dumont S, Biemont C (1999). Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol. Biol. Evol* 16:1251-1255
- Viseras E, Camacho JPM, Cano MI, Santos JL (1990) Relationship between mitotic instability and accumulation of B chromosomes in males and females of *Locusta migratoria*. *Genome* 33:23-29
- Volff JN, Körting C, Schartl M (2000) Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. *Mol Biol Evol* 17:1673-1684
- Wessler S, Bureau T, White SE (1995) LTR-retrotransposons and MITE: important players in the evolution of plant genomes. *Curr Opin Gent Dev* 5:814-821
- Wicker T, Sabot F, Hua-Van A et al (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973-982
- Zurita S, Cabrero J, López-León MD, Camacho JPM (1998) Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution* 52:274-277

CAPÍTULO 3

Preferential occupancy of R2 retroelements for B chromosomes in the grasshopper *Eyprepocnemis plorans*

Eugenia E. Montiel, Josefa Cabrero, William D. Burke, Thomas H. Eickbush,
Juan Pedro M. Camacho & María Dolores López-León

Introduction

Transposable elements (TEs) appeared to be an oddity when they were first uncovered by B. McClintock in the 1950s, but now they are known in essentially every sequenced genome, e.g. human beings, plants, yeast or fruit flies, with TEs constituting a major part of most eukaryotic genomes. In fact, they have been found in virtually all species where they have been searched for, except *Plasmodium falciparum* (Garner et al. 2002). Due to the huge variety of mobile elements – 22586 entries in RepBase Update (Jurka et al. 2005) – it was necessary to classify them on the basis of transposition mechanism and molecular structure (see RepBase Update). Within the retrotransposons (Class I), the R superfamily of non-LTR elements includes several members (e.g. R2 elements) which specifically insert into the 28S gene of the ribosomal DNA (rDNA) cluster in animals thus rendering it non-functional (Jacubzack et al. 1991). The R2 retrotransposon was first described in the fly *Drosophila melanogaster* (Rohia et al. 1981), but it has also been found in other arthropod species, such as *Bombix mori*, *Drosophila simulans*, *Apis mellifera*, *Rhyncosciara americana* and *Triops cancriformis* (Eickbush & Robins 1985; Zhang & Eickbush 2005; Gillespie et al. 2006; Rezende-Teixeira et al. 2009; Mingazzini et al. 2010), and also in other animal phyla such as Chordata, Echinodermata and Phatyhelminthes (Kojima & Fujiwara 2004, 2005).

R2 retrotransposable elements encode three protein domains needed for transpositional activity: a reverse transcriptase (RT) occupying a central position, an endonuclease (EN) at the C terminus and a DNA binding domain at the N terminus. The protein encoding domains are flanked by two untranslated sequences (5'UTR and 3'UTR). Unlike most other retroelements, the 5' UTR region in R2 does not contain promoter sequences. R2, however, encodes a self-cleaving ribozyme and allows R2 elements to be co-transcribed with the rDNA unit (Eickbush & Eickbush 2010). In the 3' UTR region, several families of R2 retrotransposons show poly(A) tails similar to several non-LTR elements (Burke et al. 1999).

The integration site for R2 elements (Figure 7) is shared by all arthropods (Burke et al. 1999) and the insertion mechanism has been extensively studied in *Bombyx mori* and the Drosophilidae genera (Luan et al 1993; Bibillo et al. 2002; Christensen et al 2005; Jamburuthugoda & Eickbush 2011). R2 inserts into rDNA units through cleavage of the 28S target site by an R2 encoded endonuclease, subsequent target primed reverse transcription (TPRT) reaction performed by its encoded reverse transcriptase, followed by second strand DNA cleavage, second strand DNA synthesis and finally DNA repair.

This reverse transcription frequently produces variable length deletions at 5' end of the R2 element, since reverse transcriptase is not always able to reach the end of the RNA template (George et al. 1996). These 5'-truncated copies serve as footprints of insertion events because recombination within the locus does not appear to produce such deletions (Lathe et al 1995; Lathe & Eickbush 1997). In *D. melanogaster* and all other organisms, rDNA units undergo concerted evolution, with recombination (through unequal crossovers and gene conversion) maintaining rDNA sequence homogeneity (Eickbush & Eickbush 2007). The escape of R2 elements from homogenization to R2-free rDNA sequences is explained by their high insertion rate (Lathe et al. 1995) and the existence of more rDNA units in the host genome than needed for organism survival (Zhang et al. 2008). A recent hypothesis has suggested that recombination in the rDNA locus constantly redistributes R2-inserted units. When large blocks of functional R2-free rDNA units are produced, R2-containing units are outside the transcribed area of the locus and prevent from transcription (Zhou & Eickbush 2009). This continuous reconfiguration of the rDNA cluster makes highly variable the number of R2 copies per genome -from tens to hundreds- depending on the individuals or species analyzed, but typically about 10-30% of the rDNA units are inserted (Burke et al. 1987; Jacukczab et al. 1992; Rezenda-Teixeira et al. 2009). The expression of R2 elements has been found in adults of *D. simulans* (Eickbush & Eickbush 2007) but the highest activity takes place during early embryo development and oogenesis (Eickbush & Eickbush 2011).

B chromosomes were discovered at the dawn of cytogenetics (Wilson 1907; Randolph 1928), i.e. long before TEs, and have been reported in about 15% of eukaryotic organisms. B chromosomes are dispensable, show a variable number among individuals, do not occur in pairs, in contrast to standard (A) chromosomes, and thus do not follow Mendelian inheritance. In addition, they are frequently heterochromatic, do not recombine with A chromosomes and may be deleterious for the organism carrying them. B chromosomes are most likely derived from A chromosomes, either intra- or interspecifically, but follow their own evolutionary pathway (for a recent review, see Camacho 2005). The evolutionary dynamics of B chromosomes usually follows a parasitic pathway, as was first claimed by Östergren (1945) and later recognized by several authors (Nur, 1966, 1977; Jones 1985; Shaw and Hewitt 1990). Alternatively, Darlington (1958) and White (1973) suggested a heterotic nature for B chromosomes arguing that they could be beneficial at low numbers. However, such a beneficial effect of B chromosomes has only been shown in a few cases (e.g. Plowman and Bougourd 1994; Miao et al 1991).

In the grasshopper *Eyprepocnemis plorans*, B chromosomes begin as parasitic elements in the newly invaded populations, and they rapidly increase in frequency because of meiotic drive in females (Camacho et al. 1997; Zurita et al. 1998; Bakkali et al. 2002). Bs are rapidly neutralized through the action of drive suppressor genes (Herrera et al. 1996; Camacho et al. 1997; Perfectti et al. 2004) and, during the long near-neutral pathway, new parasitic B variants (recovering drive) may invade the population replacing the original B (Zurita et al. 1998). B chromosomes in this species show a very high mutation rate and thus a very high number of B variants have been found (López-León et al. 1993; Bakkali and Camacho 2004), mainly differing in the amount and distribution of their two major repeated DNA sequences constituents: a 180 bp tandem repeat (satDNA) and ribosomal DNA (rDNA) (Cabrero et al. 1999). Recently, the activity of the rDNA in the B₂₄ chromosome from the Torrox population has been inferred from the presence of nucleoli attached to the distal region of B chromosomes during meiotic first prophase (Teruel et al. 2007, 2009) and of rRNA transcripts showing sequence specificity to the rDNA in the B chromosome (Ruiz et al. in preparation). Other nucleolar organizing regions (NORs) containing rDNA are located on X chromosome and S₉, S₁₀ and S₁₁ autosomes.

B chromosomes have been considered to be havens for TEs because B dispensability permits TEs escaping from the action of natural selection (Camacho et al. 2000). This may explain why TE accumulation has been reported in B chromosomes, for example, fish (Ziegler et al. 2003) and maize (Lamb et al. 2007). However, In *E. plorans*, *Gypsy* and *RTE* retrotransposons, and *Mariner* transposons appear to be most frequent in the euchromatin but are rarely found in the heterochromatin, including B chromosomes (Montiel et al. submitted). The fact that B chromosomes in this species show the largest amount of rDNA in the genome, makes it feasible that conditions will be different for an rDNA specific TE like R2. Here we characterize the complete DNA sequence of R2 in *E. plorans*, and analyze its abundance in the genome of B-carrying and B-lacking individuals in order to increase our understanding of the relationship between mobile elements and the *E. plorans* genome, in general, and B chromosomes, in particular. We found that B chromosomes are enriched in R2 elements, that these elements are not transcribed in adult males but that many truncations can be detected in all natural populations analysed, thus showing signs of past transposition activity.

Materials and Methods

Animals and cytological characterization of B chromosomes

Adult specimens of *Eyprepocnemis plorans* were collected at Spanish populations in Torrox (Málaga), Salobreña (Granada), Calasparra (Murcia), Mundo (Albacete) and Socovos (Albacete). Adults were bred in the laboratory in order to obtain embryos. Egg pods were incubated at 28°C for twelve days, after which embryos were fixed in 3:1 ethanol-acetic acid for cytogenetic studies, as described in (Camacho et al. 1991). Males were anaesthetized and testes were dissected out and fixed in freshly prepared 1:3 acetic acid-ethanol. Females were injected with 0.1 ml of 0.05% colchicine in insect saline solution 6 h prior to anesthesia, dissection and fixation of ovarioles in 1:3 acetic acid-ethanol. The number of B chromosomes in each individual was determined in 2% lactopropionic orcein squash preparations of two testis tubules or one ovariole. Adult bodies were frozen at -80°C for DNA and RNA isolation.

DNA and RNA isolation

Genomic DNA from adult individuals was extracted using the Gen Elute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich) following manufacturer's recommendations.

Total RNA was isolated from 30 mg of individual frozen bodies using the Real total RNA spin plus kit (Durviz), following manufacturer's protocol including DNase I treatment. RNA was later submitted to a second treatment with 20 units of DNase I to ensure complete removal of contaminating genomic DNA. The total amount and quality of the RNA obtained was checked in a denaturing agarose gel.

B chromosome microdissection

Fifteen B chromosomes were microdissected from spermatocytes in one B₂₄-carrying individual from the Torrox populations. Procedures for chromosome preparations and microdissection assay were described in Teruel et al. (2009). DNA of microdissected B chromosomes was amplified with the GenomePlex Single Cell Whole Genome Amplification Kit (WGA4-Sigma) following manufacturer's recommendations.

PCR amplification, DNA cloning and sequence analyses

To amplify R2 in *E. plorans* we designed degenerate primers, R2-F and R2-R, anchored on conservative parts of the reverse transcriptase domain, using the CODEHOP software (Rose 2003). Information about all primers is shown in Table 6.

Table 6. List of primer with their sequence and the method where they have been used. C: Initial cloning of element; CN: Copy number estimation; E: Expression analysis; TP: 5' Truncation profile.

Primer name	Sequence (5' -> 3')	Method
R2-F	AGCGCCACAGGATTTYCGNCCNAT	C
R2-R	AGGGCGAACATCCGCTGYTGBGGT	C
EploUp70-F	TGCCCAAGTGCTCTGAATTGTC	C, TP
EploDown70-R	AGATAGGGACAGTGGGAATC	C
EploR2_2800-R	CGAGATGGTAGAGCACTAATC	TP, CN, E
EploR2_2500-F	CAAGTCCCTCCATGCTCGCCA	TP, CN, E
Eplo R2_820-R	GGTGCCCGCTTAATGATGTCC	C, TP
Eplo R2_720-R	GGAGTGGGCCATCGCCAGATC	TP
Eplo R2_1200-R	GTAATCGGTGCCACAGATC	TP
EploR2_670-F	AAGGTCGACAACAGTGTCATC	C
Eplo R2_600-R	AATAGAGTCCAACGTCCGGTC	TP
ITSFISH4	ATATGCTTAAATTCAGCGGG	CN
ITS3	GCATCGATGAAGAACGCAGC	CN

PCR experiments were performed in a 50 µl reaction mixture containing 50 ng of genomic DNA, 0.2 µM of each primer, 0.2 mM dNTPs, 1.25 mM MgCl₂ and 1 unit of Taq polymerase (MBL). PCR conditions were as follows: initial denaturation 5 min at 94°C, 30 cycles of 1 min at 94°C, 1 min at 50°C, and 2 min at 72°C, and a final elongation 10 min at 72°C. PCR products were visualized in a 1% agarose gel and were cleaned with the GenElute PCR Clean-up kit (Sigma-Aldrich). All amplified fragments were cloned into pCR2.1-TOPO vector (Invitrogen). Plasmid DNA was obtained with the Gen Elute Mammalian Genomic DNA Miniprep kit (Sigma) and clones were sequenced at Macrogen Inc. (Seoul, Korea). DNA sequences were analyzed with BioEdit (Hall 1999) and Geneious Pro 4.8.5 (Biomatters Ltd.).

In order to characterize the complete sequence of R2 in *E. plorans*, we designed divergent primers, two of which being anchored in the *E. plorans* reverse transcriptase partial DNA sequence obtained, the EploR2_670-F and EploR2_820-R primers, and the other being anchored in the 28S rRNA gene described in Gomphocerine grasshoppers (EploUp70-F and EploDown70-R) (see positions in Figure 7). Thereby, we expected to amplify the 5' region with the EploUp70-F / EploR2_820-R primers pair and the 3' region

with EplOR2_670-F / EplODown70-R primers pair. PCR reactions were prepared as described above and were performed as follows: 2 cycles of 1 min at 97°C, 2 min at 55°C, and 3 min at 72°C, following for 28 cycles of 1 min at 94°C, 1 min at 60°C, and 3 min at 72°C, and final cycle of 1 min at 94°C, 1 min at 60°C and 10 min at 72°C. Amplifications using the EplUp70-F/EplOR2_820-R and EplOR2_670-F/EplODown70-R primers pairs, yielded a pool of fragments. In both cases we isolated the largest amplicons (about 2kb length) obtained from the gel and purified them with the GenElute Gel Extraction Kit (Sigma-Aldrich). Those fragments were cloned and sequenced. Clone similarity to TEs was tested using BLAST and BLASTX tools (Altschul et al. 1997) at the NCBI site. To confirm and refine the identity of the elements found within the vast families of non-LTR retrotransposons, we tested cloned DNA sequences with the CENSOR online software (Jurka et al. 2005).

To determine the molecular structure of this element we conducted an alignment of its predicted protein sequence with R2 elements from other species. From the Repbase Update database (<http://www.girinst.org/repbase/update/index.html>), we downloaded the aminoacid sequence of R2 elements from *Trichinella spiralis* (R2-1_TSP; nematodes), *Schmidtea mediterranea* (R2-1_SM; freshwater planarian), *Petromyzon marinus* (R2-1_PM; lamprey), *Danio rerio* (R2Dr; zebrafish), *Taeniopygia guttata* (R2-1_TG; zebra finch), *Limulus polyphemus* (R2_LP; horseshoe crab), *Anurida maritima* (R2_AM; springtails), *Apis mellifera* (R2Amel; honey bee), *Nasonia giraulti* (R2C_NGi), *Nasonia vitripennis* (R2B_NVi; jewel wasp), *Drosophila melanogaster* (R2_DM; fruit fly), *Drosophila simulans* (R2_DSi), *Bombyx mori* (R2_BM; silkworm) and *Forficula auricularia* (R2_FA; earwig). The alignment was done with the alignment tool implemented in Geneious Pro 4.8.5 (Biomatters Ltd.).

Fluorescent in situ hybridization

Chromosome preparations from embryos were performed according to Camacho et al. (1991). Slides were dehydrated in a series of 70%, 90% and absolute ethanol and then incubated in an oven at 60°C overnight. Insert from EplOR2_clon3 (Fig. 1) was used as probe, previously labeled by nick translation with tetramethylrhodamine-11-dUTP (Roche). Fluorescent hybridization signals were directly detected. About 250 ng of DNA probe was used in each FISH experiment following the technique described in (Cabrero et al. 2003). Chromosomes were counterstained with DAPI (4', 6 diamidino-2-phenylindole) and the analysis of chromosome preparations was made using an Olympus BX41 microscope for fluorescence and photographs were taken with an associated DP70 cooled camera. Images were optimized for bright and contrast with THE GIMP freeware.

R2 expression and 5' truncation profile

cDNA was obtained with SuperScript III First-Strand Synthesis SuperMix kit (Invitrogen), according to manufacturer's recommendations, using random hexamers. R2 retrotransposon expression was investigated by PCR amplification of the cDNA obtained from the isolated RNA of males from Torrox population, using the primers pair EploR2_2500-F / EploR2_2800-R. PCR reactions were performed with 10 ng of cDNA, 0.25 μ M of each primer, 0.2 mM dNTPs, 1,25 mM MgCl₂ and 1U of Taq (MBL), using a thermocycler program with an initial denaturation step at 94°C for 5 min. and 30 cycles of 30 s at 94°C, 30 s at 55°C, and 30 s at 72°C with a final step at 72°C for 8 min. All amplifications performed included a negative control without RNA to eliminate genomic DNA contamination of the original RNA samples, and a positive control with genomic DNA to test the suitability of the primers.

The analysis of R2 activity based on 5' truncation pattern was performed separately in individuals with and without B chromosomes (0B) in four populations: two from the Southeast of the Iberian Peninsula (Sovocos and Calasparra) where B chromosomes show low frequency, and two from the South (Salobreña and Torrox) where B chromosomes are frequent. We employed a total of 88 males: 14 from Socovos (all 0B) and 14 from Calasparra (all 0B). In the two populations showing high B-frequency, we analysed 28 males from Salobreña (14 lacking Bs, 9 with 1B and 5 with 2B) and 32 males from Torrox (17 lacking Bs, 7 with 1B, 6 with 2B, 1 with 3B and 1 with 4B). The 5' truncation patterns were obtained by PCR amplification using the EploUp70-F primer, annealing upstream of the R2 element insertion site, coupled with various R2-specific primers: EploR2_600-R, EploR2_1200-R, EploR2_720-R, EploR2_820-R, EploR2_2800-R. These primers anneal at 561, 1356, 1551, 2200 and 2747 positions downstream of the insertion site, respectively (Fig. 5). Amplifications were performed with 50 ng of genomic DNA and PCR reaction was as described above. PCR conditions were: 2 cycles of 1 min at 97°C, 2 min at 55°C, and 3 min at 72°C, following for 28 cycles of 1 min at 94°C, 1 min at 60°C, and 3 min at 72°C, and final cycle of 1 min at 94°C, 1 min at 60°C and 10 min at 72°C. PCR products were visualized in a 2% agarose gel and the pictures were taken with a Gel Doc XR System (BioRad Laboratories, Inc). Pictures were analyzed with Quantity One v4.6.3 (BioRad Laboratories, Inc).

Copy number estimations

The analysis of copy number for the R2 element in the *E. plorans* genome was performed by real-time quantitative PCR (qPCR). This approach was also employed to estimate the number of rDNA units, in order to calculate R2 occupancy, i.e. the proportion of rDNA loci carrying an inserted R2 element. Copy number quantifications per diploid genome were performed for R2 elements and rDNA units in five populations differing in B chromosome frequency and type: Socovos, with no B chromosomes, Calasparra, with low frequency of the B₁ variant, and Mundo, Salobreña and Torrox showing high frequency for B₁, B₂ and B₂₄ variants, respectively. DNA templates were obtained from three B-carrying and three B-lacking males from Torrox, Salobreña and Mundo, and three 0B from Calasparra and Socovos. DNA was quantified using PicoGreen® dsDNA Quantitation Reagent (Molecular Probes) in a fluorometer TBS-380, according to manufacturer's recommendations.

qPCR assays contained 10 µl of 2X SensiMixPlus SYBR (Quantace), 0.25 µM of each primer and 10 ng of genomic DNA sample in a 20 µl final volume. The quantification of R2 copy number was done with the EplOR2_2500F and EplOR2_2800R primers which yield a 295 bp amplicon, and rDNA copy number was analysed with the ITS3 / ITSFISH4 primers which amplified a 376 bp fragment.

Both sequences, R2 elements and rDNA units, were amplified with the same PCR programs, only differing in annealing temperature. PCR program consisted in a denaturation step at 95°C for 10 min, and 40 cycles of 15 s at 95°C, 15 s at 58.2°C (R2), 61°C (rDNA), 72°C for 30 s and a final step at 72°C for 5 min. Real-time qPCR assays were run in a Chromo4 (Biorad) thermocycler using Opticon Monitor v3.1 software to export raw data and LingRegPCR 11.0 programme to analyse them. Each sample was amplified in triplicate and product specificity was verified through generation of a melting curve following Real-time PCR assay. The calibration factor was obtained with a serial dilution of two clones containing the region to be amplified, EplOR2_2000.1 and EplOITS_1.6.

Copy number estimations were performed through analysis with LingRegPCR software, following the method described by Feng et al. (2008). We estimated copy number of R2 and rDNA sequences on the basis of C values in *E. plorans* (10160 Mb) and bearing in mind recent estimates of the size of X and B chromosomes (Ruíz-Ruano et al. 2011). Therefore, in order that these results were comparable among the different individuals (males with or without Bs), we expressed the estimated copy number (per nanogram) in terms of the 2C value, modified according to the number of X or B chromosomes present in each individual. Therefore, 2C value in B-lacking males was calculated as twice the C

value reported in Ruiz-Ruano et al. (2011) minus the DNA content of an X chromosome, since sex chromosome determinism is XO in *E. plorans* males. In B-carrying males, we estimated 2C value by also adding the DNA content of each B chromosome variant reported in Ruíz-Ruano et al. (2011).

Results

Molecular characterization of R2 in E. plorans

The general strategy to characterize R2 by using primers already described in literature (Jakubczak et al. 1991) was not suitable for *E. plorans*. Therefore, we designed new degenerate primers, R2-F and R2-R, to amplify R2 in this species. These primers were designed on the basis of an R2 protein alignment from several organisms and produced a pool of fragments ranging from 200 bp to 2 kb in length. We analyzed 15 clones from the PCR product of B-carrying males, and five of them contained an 859 bp insert showing sequence similarity with R2 elements in BLASTX and REPBASE. To complete the characterization of the R2 element in *E. plorans*, we employed EploUp70-F/EploR2_820-R and EploR2_670-F / EploDown70-R primer pairs. From both PCR reactions we obtained fragments of the expected size (2kb) that were subsequently cloned. Three of twelve clones analyzed from the 5' end corresponded to the 5' R2 region, and three of six clones from the 3' end corresponded to the R2 element. Combining the sequences in these clones allowed us to elucidate the complete sequence of the R2 element in *E. plorans* (Figure 7).

The full-length R2 element in *E. plorans* (EploR2) is 3566 bp long (GenBank acc. no. JQ082370). As a non-LTR family member, EploR2 lacks long terminal repeats and contains a 206 bp 5'UTR region and a 234bp 3' UTR region. The 3' UTR contains a canonical polyadenylation site (AATAAA) is present at nucleotide 3550, but does not contain a terminal poly(A) tail, as found in some species. The sequence contains a single open reading frame encoding a polypeptide consisting of 1041 amino acids ending with a termination codon (TGA) at position 3332.

To assess the similarity between EploR2 to other elements from the R2 family, we aligned the EploR2 protein sequence with R2 proteins present in Replibase from other species (Figure 7). The encoded protein contains the typical domains for an R2 element which are common to other species, with a DNA-binding domain in N-terminal region, a reverse transcriptase domain in the middle and an endonuclease domain in the C-terminal

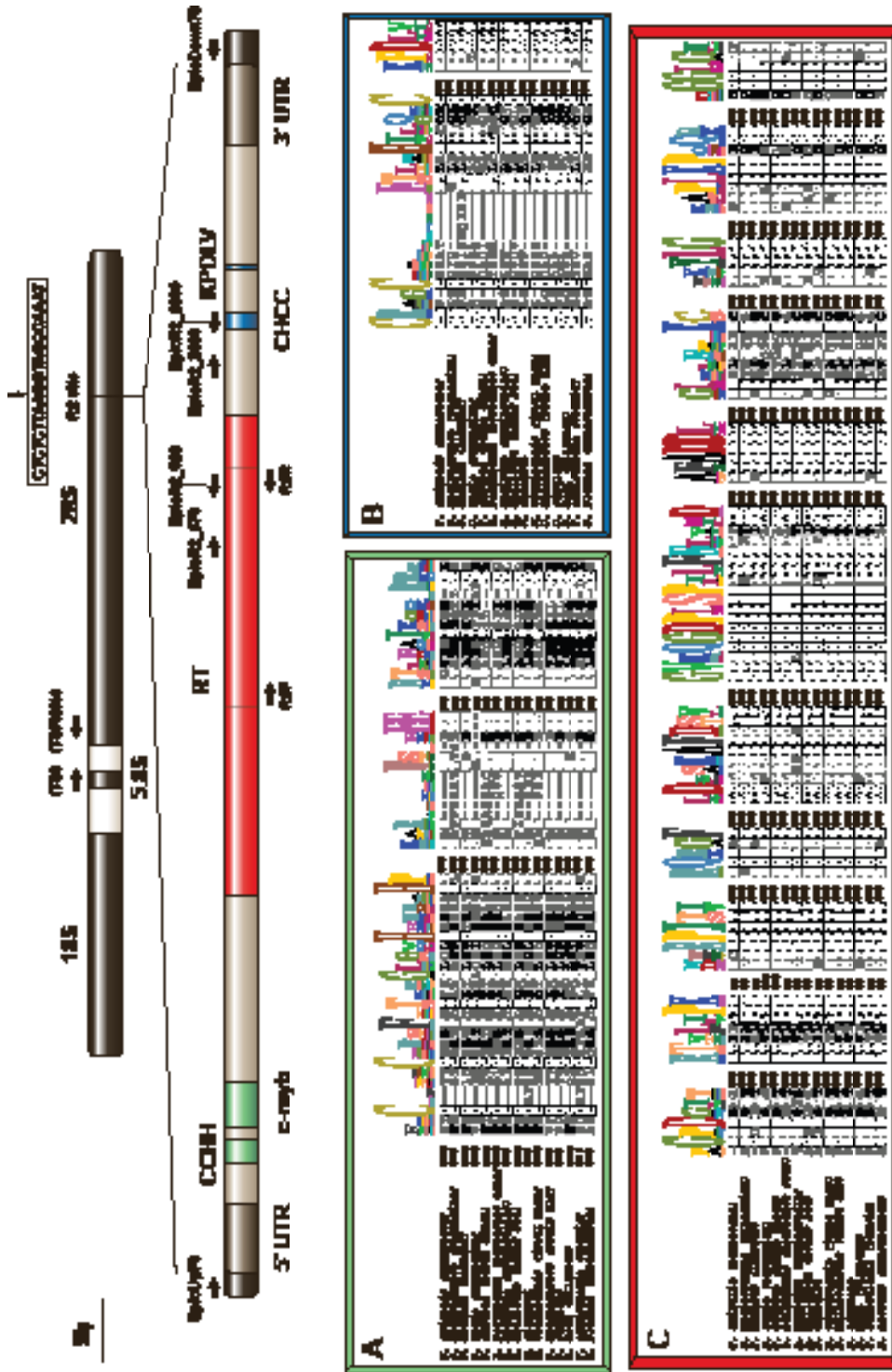


Figure 7. Diagram of Eplor2 insertion site, molecular structure and alignment with other R2 elements. rDNA locus is represented above, indicating the R2 site (blue mark) and the sequence of insertion target (box) with a vertical arrow in the cleavage site. The striped area indicates Eplor2_clon3 position, which was the first sequence obtained with the R2-F and R2-R degenerate primers. Arrows show position and orientation of primers used to characterize and copy number estimation assay. The boxes show the alignment of main domain with other R2 elements. A box contains CCHH and c-myb motifs of DNA-binding domain. B box represents CHCC motif and KPDLV of endonuclease domain. And finally, C box shows the 9 motifs of reverse transcriptase domain.

region. In the amino-terminal region, we found DNA-binding motifs, CCHH zinc finger and c-myb, involved in the recognition of DNA binding target (Christensen & T.H. Eickbush 2005). In the carboxyl-terminal region, the domain with endonuclease activity included Cys-His (CCHC) and KPDLV motifs.

Chromosome distribution of R2

Physical mapping of R2 elements was performed on embryos from the Torrox population, by means of fluorescence in situ hybridization (FISH) experiments. R2 elements were found on the distal DAPI- region of the B₂₄ chromosome, coinciding with location of the rDNA cluster in the B chromosome. Other less intense hybridization signals were also clustered on S₁₀ and S₁₁ autosomes, i.e. at the interstitial positions where they carry rDNA (Figure 8). No signals were observed on X and S₉ chromosomes, which are other chromosomes carrying rDNA. The B chromosome clearly showed the highest hybridization signal, in correspondence with its larger amount of rDNA compared to A chromosomes.

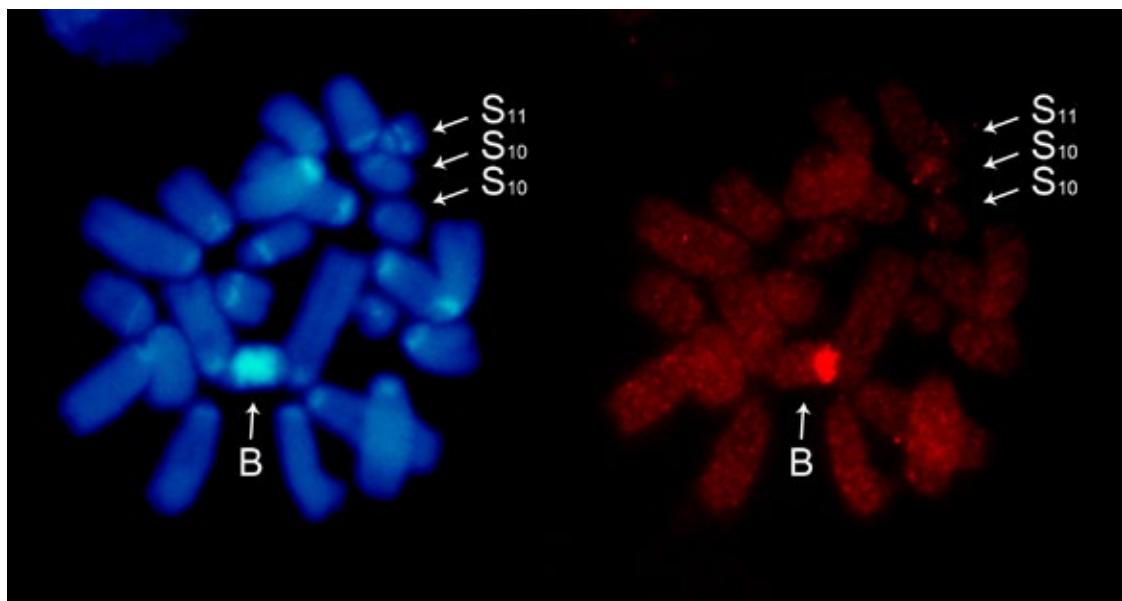


Figure 8. FISH analysis of mitotic metaphases from *E. plorans* embryos. The arrows indicate the observed fluorescent hybridization signals on S₁₁ and S₁₀ autosomes as well as on B₂₄ chromosome, which show the largest R2 cluster. S₁₁ autosome is polymorphic for the presence of R2 cluster. Indicate what the blue and red colors represent.

Copy number quantification

The presence of rDNA in the B chromosomes of *E. plorans* raised the possibility for differential R2 occupancy on A and B chromosomes. The results indicated that copy number for R2 and rDNA showed variation among populations (see Table 7), with 459 ± 586 (mean \pm standard deviation) R2 copies and $17,700 \pm 9,310$ rDNA units for all populations. In B-lacking individuals from the five populations analysed, the copy number of R2 elements was more similar, with a mean of 97 copies. In contrast, B-carrying males showed highly variable numbers of R2 copies, with a mean of 967 copies, i.e. about ten times larger than B-lacking ones (Figure 9b). This difference was most apparent in the Salobreña population, where B-carrying individuals carried 1225 copies but B-lacking ones carried only 87 (see also Table 7).

Table 7. R2 elements, and rDNA units copy number and percentage of rDNA units occupied by R2 elements. – B: individuals without B chromosomes; + B: individuals with B chromosomes. The number of R2 copies on B chromosomes was inferred from the difference between the mean numbers in +B and -B individuals.

Population	B chrom.	R2		rDNA		% R2-inserted rDNA	
		Mean	SD	Mean	SD	Mean	SD
Socovos	- B	111.01	25.81	11.91E+3	1.12E+3	0.94	0.21
Calasparra	- B	85.72	5.41	10.69E+3	2.79E+3	0.84	0.17
Mundo	- B	132.78	19.15	6.82E+3	0.54E+3	1.94	0.13
	+B ₁	1323.19	288.98	21.61E+3	1.39E+3	6.16	1.49
Salobreña	- B	87.76	42.19	13.12E+3	6.46E+3	0.70	0.30
	+B ₂	1225.58	677.65	36.25E+3	10.41E+3	3.17	1.13
Torrox	- B	80.06	13.98	16.80E+3	2.02E+3	0.47	0.05
	+B ₂₄	234.78	65.48	19.97E+3	0.36E+3	1.18	0.35
Total		459.84	586.83	17.76E+3	9.31E+3	2.10	2.13
	- B	97.30	28.14	12.23E+3	4.31E+3	0.91	0.50
	+B	967.38	623.07	25.51E+3	8.96E+3	3.77	2.44
B variants	B ₁	1190.41		14.79E+3		8.05	
	B ₂	1137.82		23.13E+3		4.92	
	B ₂₄	154.71		3.17E+3		4.88	

The number of rDNA units in B-lacking individuals from different populations also showed extensive variation, ranging from 6820 units in Mundo to 16800 units in Torrox (Figure 9c). In all B-carrying populations, the number of rDNA copies was higher in B-carrying than B-lacking individuals, because of the presence of rDNA on the B chromosomes. The highest difference between these two types of individuals was found in Mundo, as expected because the B₁ chromosome in this population shows the largest rDNA cluster, as deduced from FISH analysis (Cabrero et al. 1999). However, the

highest number of rDNA copies was found in B-carrying individuals from the Salobreña population, even though they carry B_2 which harbours lower amount of rDNA than B_1 . This was most likely due to the higher amount of rDNA in the A chromosomes in Salobreña than Mundo (compare 0B individuals between these two populations in Table 7). On average, B-carrying individuals carried about twice the number of rDNA copies found in B-lacking individuals, which suggests that B chromosomes carry as much rDNA as all A chromosomes together.

The number of R2 elements and rDNA units showed a positive and significant correlation (Pearson $r = 0.86$, $p = 0.000001$). Figure 9d also shows that rDNA occupancy by R2 in B-lacking individuals was appreciably lower than that in B-carrying individuals in all populations. This difference was lowest in Torrox where the B chromosome variant (B_{24}) carried the lowest level of rDNA. Therefore, R2 occupancy appears to be significantly higher on the B chromosomes (Figure 9), i.e. about 8% in B_1 and 5% in B_2 and B_{24} , compared to 1-2% for the autosomal rDNA.

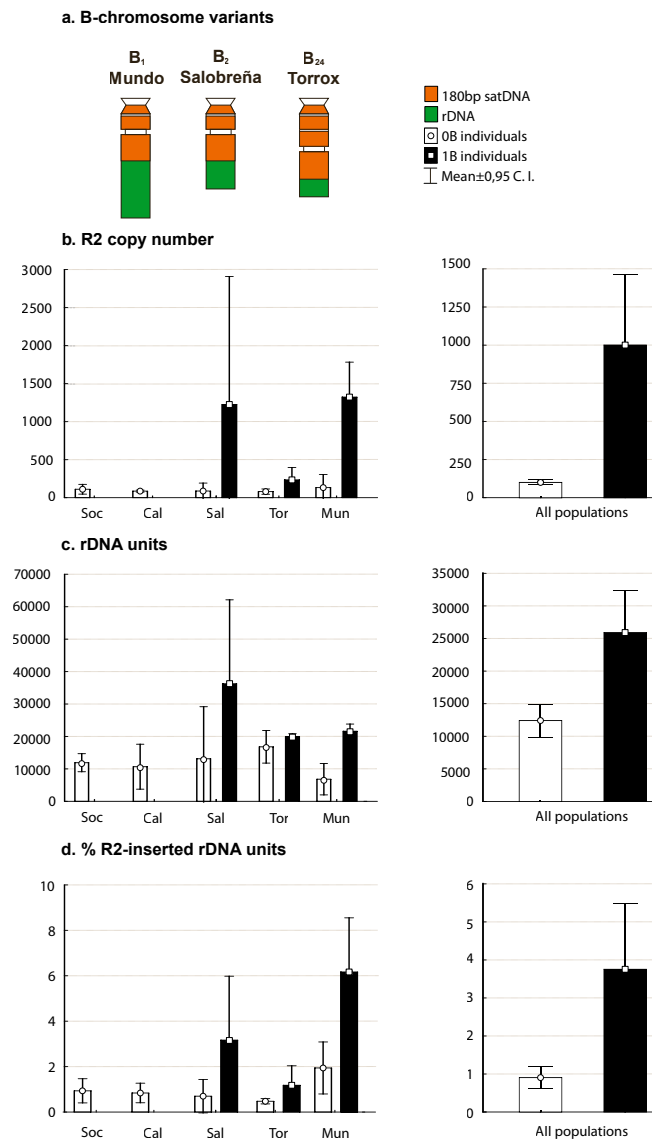


Figure 9. Molecular composition in the different B variants (the orange box indicates the satDNA and the green box, rDNA). (a.), and copy number estimation of R2 elements (b.), rDNA units (c.) and percentage of rDNA units inserted with EplOR2 (d.). The copy number of B-lacking and B-carrying individuals are showed separately for each population on the left and combined for all populations on the right. Soc: Socovos; Cal: Calasparra; Sal: Salobreña; Tor: Torrox; and Mun: Mundo.

R2 Expression and Activity

PCR amplification of R2 element sequences from total cDNA generated from RNA of different males in the Torrox population did not yield products, demonstrating the absence of R2 transcripts for these elements in most tissues of *E. plorans*.

To determine if R2 elements have recently been active in the populations the approach used to monitor R2 retrotransposition activity in *D. simulans* was followed (Zhang and Eickbush, 2005). PCR amplification of genomic DNA with different primer pairs to R2 sequences (Figure 10) allowed us to detect the presence of 5' truncated R2 copies in each specimen. The presence of these different 5' is correlated with their insertional activity. The 5' truncated profile in the Socovos population did not show any recent movement of R2, because all truncated copies were shared by several individuals and there were not any newly truncated element. Calasparra showed similar profile finding with only one new truncated copy. It is remarkable that these two populations from the Southeast region of the Iberian Peninsula shared four truncated copies. In the two populations from the South (Salobreña and Torrox) we detected high numbers of newly truncated elements. In Salobreña, these new truncations appeared most frequently in B-carrying individuals. In Torrox, however, this frequency was similar between individuals with and without B chromosomes. In this last population we observed some B-specific truncations that only appeared in B-carrying individuals and in the 5' truncated pattern yielded by DNA obtained through B chromosome microdissection. The analysis of PCR products also demonstrated the presence of R2 full-length elements in all individuals.

Discussion

R2 is a non-LTR retrotransposon deriving its mobility from the activity of its reverse transcriptase protein (Malik et al. 1999), a feature allowing rapid mobility, a high error rate and the production of 5' truncated copies (Luan et al. 1993). The rapid evolution of their nucleotide sequence explains why our initial attempts of using primers designed for *D. melanogaster* did not yield positive results. However, the amino acid sequence of the reverse transcriptase domain shows highly conserved motifs in R2 elements (Burke et al. 1999) and, on this basis, we designed degenerate primers which efficiently amplified R2 in *E. plorans*. In addition to the central reverse transcriptase domain, two other domains within R2 also have conserved motifs: the CCHC zinc-finger and KPDLV motifs within the C-terminal domain, which are essential for DNA cleavage, and multiple CCHH zinc-fingers and one c-myb motif, which contribute to DNA-binding function,

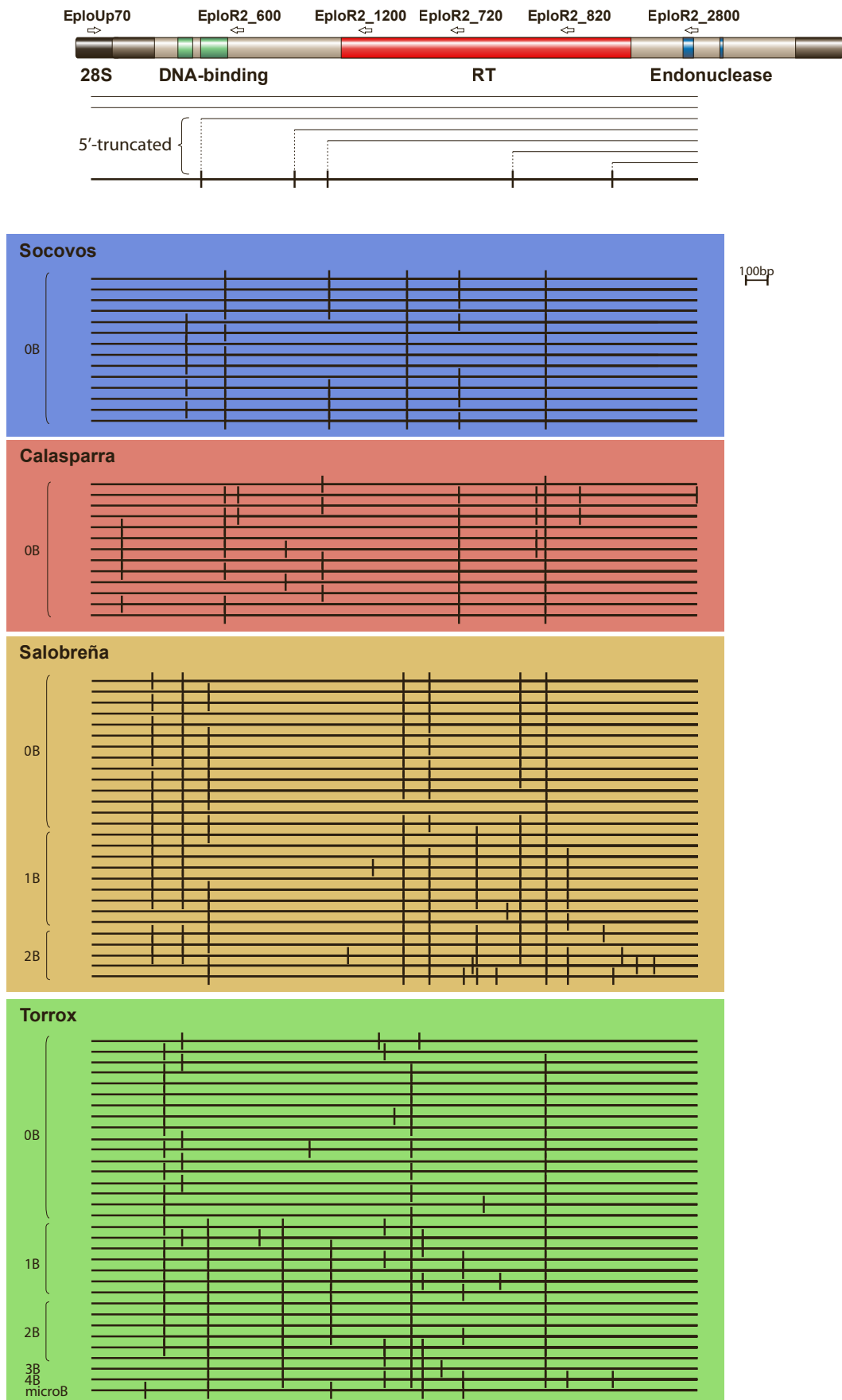


Figure 10. 5' truncation profile of R2 elements in different populations of *E. plorans*.

in the N-terminal domain (Yang et al. 1999). This structure is strongly shared by all R2 elements hitherto described, except for the number of CCHH zinc-finger repeats in the DNA-binding domain (Kojima et al. 2005). The R2 element in *E. plorans* (EplOR2) shows the three mentioned domains and includes a single CCHH motif, thus belonging to the R2-D clade, together with most other arthropod R2 elements, including those in *D. melanogaster* (Eickbush et al. 1997) and *T. cancriformis* (Mingazzini et al. 2011).

The insertion site within 28S rDNA is conserved in EplOR2, as has been reported for all R2 elements. We did not detect deletions of the adjacent 28S gene sequences as found in some other species (Burke et al. 1999; Kojima et al. 2005; Rezende-Terexeida et al. 2009; Ghesini et al. 2011; Mingazzini et al. 2011). However, in contrast to other species where R2 has been studied (Pérez-González & Eickbush 2001), *E. plorans* has rDNA loci on most chromosomes, with the largest clusters present on B chromosomes. Our FISH results showed higher levels of R2 elements on the rDNA of B chromosomes than on the rDNA of S₁₀ and S₁₁ autosomes, and no R2 on the other A chromosomes carrying rDNA such as the X and S₉ chromosomes. This in situ data was consistent with our estimations based on qPCR. On average, B chromosomes carried ten times more R2 copies than A chromosomes (see Table 7). An estimation of copy number on every B chromosome variant can be deduced from the difference between B-carrying and B-lacking individuals from a same population.

The B chromosome are important in explaining the population dynamics of EplOR2 elements. The B variants differ in length and in satDNA and rDNA proportion, with B₁ the highest amount of rDNA (as deduced from FISH analysis), and B₂ and B₂₄ carrying smaller rDNA clusters (Cabrero et al. 2003). Differences between the in situ data and the qPCR data (Table 7) might be explain by a different molecular organization of rDNA units in different B variants, i) the differences could be due to the proportion of degenerate rDNA copies which may not be detect by the PCR analysis, or ii) differences in the amount of interspersed sequences between ribosomal cistrons, with would reduce the intensity of the in situ signal. Cytogenetic and molecular analyses have shown that *Gypsy* and *RTE* retrotransposons and *Mariner* transposons are very abundant in the *E. plorans* genome, although more on the A than B chromosomes since they are absent from regions where rDNA is located (Montiel et al. submitted). Interestingly, some authors have shown an accumulation of mobile DNA on B chromosomes (Theuri et al. 2005; McAllister & Werren 1997) or other chromosomes that have no (few) essential genes, e.g. Y chromosome of *D. miranda* (Bachtrog 2003).

Some of the more than 50 different B chromosome variants reported in Spanish populations of *E. plorans* presumably arose through chromosome rearrangements (e.g. inversions, deletions and translocations) involving breakpoints within the lightly C-banded distal region (López-León et al. 1993), which corresponds with the rDNA harbouring region of B chromosomes in *E. plorans* (López-León et al. 1994, Cabrero et al. 1999). In Moroccan populations one B variant is also known that appears to have been derived from rDNA deletion (Bakkali & Camacho 2004). Therefore, at least part of B chromosome mutations in *E. plorans* involving breakpoints into rDNA could be caused by the R2 mobile element. While there is no direct evidence of chromosome rearrangements induction by R2 elements, like that reported for the FoldBack transposon in *D. buzzatti* (Cáceres et al. 1999), R2 can produce large deletions in the rDNA locus by recombination of two elements in the same rDNA locus (Zhang et al. 2008). However, the most frequent variation in *E. plorans* B chromosomes of Spanish and Moroccan populations (López-León et al. 1993; Bakkali & Camacho 2004) is centromere misdivision, so that it is likely that other mobile elements also provoke chromosome rearrangements in the B chromosomes.

The number of rDNA units estimated in B-lacking individuals (ca. 11000) is larger than that in other orthopterans like *Ch. parallelus* (ca. 1000) (Parkin & Butlin 2004) and *L. migratoria* (ca. 3300) (Schafer & Kunz 1987). The fact that *Ch. parallelus* carries rDNA clusters on only two (or three) chromosome pairs, *L. migratoria* on three pairs and *E. plorans* on most A chromosomes (Cabrero & Camacho 2008) could explain these differences. In addition, rDNA represents about 1.3% of the *E. plorans* genome, which is not much different from those in *L. migratoria* (1%) or *D. melanogaster* (0.9%) (Schafer & Kunz 1987). However, the proportion of R2-inserted rDNA units in *E. plorans* is lower (2%) than that in *D. melanogaster* (10-30%), *B. mori* (10-30%) or *L. migratoria* (20%) (Jakubczak et al. 1991, 1992). In *Triops cancriformis* only 0.5-5% of rDNA units is occupied by R2 elements (Mingazinni et al. 2011) similar to that in *E. plorans*. Selection against R2-inserted rDNA units has been the most used explanation for low occupation rates. In *E. plorans*, rDNA units from 0B individuals show three times lower R2 occupancy than those in B-carrying individuals consistent with a deleterious role of R2 elements in the A chromosome rDNA. The dispensability of B chromosomes makes them excellent places for R2 landing since there are extra rDNA units to colonize. In addition rDNA in B chromosomes is usually inactive because of DNA methylation and histone hypoacetylation (López-León et al. 1991; Cabrero et al. 2007). This suggests the interesting possibility that B chromosomes would act as sinks of no return for R2 elements coming from the A chromosomes. Interestingly, the B₁ variant, which is considered the oldest variant in

the Iberian Peninsula (Henriques-Gil & Arana 1990), showed the highest proportion of rDNA clusters being occupied by R2, whereas the B₂₄ variant, which arose recently from the B₂ variant in the Torrox population (Henriques-Gil & Arana 1990; Zurita et al. 1998), showed about the same R2 occupancy than its ancestor variant, presumably because the short time for R2 accumulation. It is also remarkable that B₂₄ shows NOR expression in the Torrox population (Teruel et al. 2007; 2009), whereas other B variants (e.g. B₁ and B₂) usually show inactive rDNA (Cabrero et al. 1987; López-León et al. 1995). It is conceivable that the higher abundance of R2 copies in B chromosomes compared to A chromosomes might be why the rDNA units of older chromosomes are no longer active.

As suggested by Zhou & Eickbush (2009), a high rate of R2 occupation does not necessarily imply higher expression levels. These authors observed in *D. simulans* that for a similar number of R2 insertions between lines, those with high expression of R2 showed the R2-inserted rDNA units scattered over the whole rRNA locus, whereas lines with low or no R2 expression showed large rDNA regions free of R2 insertions. In this model those regions with no R2 elements are preferentially transcribed. Thus it is R2 localization within the ribosomal locus which determines R2 expression. The search for R2 transcripts has been unsuccessful in *E. plorans*, but we have obtained evidence for past retrotransposition events through the 5' truncation profile. Since R2 activity is unlikely without transcription (Zhou & Eickbush 2009) two possibilities can explain the situation in *E. plorans*. First, transcription events might take place in a tissue, or development time different to those analyzed here, i.e. adult males from the Torrox population. We chose this population because rDNA in the B₂₄ chromosome is active in many individuals from Torrox (Teruel et al. 2009) and thus it was most likely that R2 elements were also transcriptionally active. However, in *D. simulans*, it has been shown that early embryo stages and the female germline exhibit the highest transcriptional activity for R2 elements (Eickbush & Eickbush 2011). A second possibility is that R2 was more active in the past than now, as suggested by the low proportion of newly truncated elements in Torrox and other populations. Parasitic B chromosomes cause stress to host cells, and the cells may respond by increasing the number of chiasmata (Camacho et al. 2002). During its parasitic stage, males from Torrox carrying the B₂₄ chromosome showed higher increase in chiasma frequency than after its neutralization (Teruel et al. 2009). Higher numbers of chiasmata means increased unequal crossovers, thus rearranging the R2 inserted and R2-free rDNA regions thus explaining an increase of R2 activity and the 5' truncated profile from Torrox. Another consequence of R2 movement could be the generation of large deletions into the rDNA locus which would explain the low rDNA copy number estimated in the Torrox population.

This is consistent with the idea that mobile elements are activated in stress conditions (Kidwell & Lisch 1997; Capy et al. 2000) and with a higher frequency of retrotransposition in B-carrying individuals, as deduced from 5' truncation profiles. In *D. simulans*, high variation in the 5' truncation profile in a population suggests R2 activity (Perez-Gonzalez & Eickbush 2001; Zhang & Eickbush 2005). *E. plorans* populations lacking B chromosomes consistently showed low frequency of R2 5' truncations and complete absence of new truncations (see Figure 10) whereas those showing high B frequency (Salobreña and Torrox) showed the higher number of new truncations, most of them probably occurring in the B chromosome rDNA, as suggested by the new truncations observed in the rDNA microdissected from the B₂₄ chromosome.

The R2 elements in *E. plorans* have shown an accumulation in rDNA loci of B chromosomes whereas they are poorly represented in the rDNA loci from A chromosomes. This enrichment in transposable elements appears to be common in B chromosomes, and it can be explained by the absence of recombination between B and A chromosomes. Recombination is believed to eliminate transposable elements from chromosomes. B chromosomes seem to be new to the *E. plorans* genome (Muñoz-Pajares et al. 2011). In its evolutionary pathway, B chromosomes have established a special relationship with the R2 retroelement by serving as 'safe haven'. However, the fact that rDNA in the B chromosomes frequently become inactive (López-León et al. 1995) possibly by the insertion of the elements themselves is a clear obstacle for further R2 spread (George & Eickbush 1999). Some unanswered questions are: does R2 show transcription in females or embryos, or in other populations of *E. plorans*? Does R2 show similar R2 enrichment in rDNA-carrying B chromosomes in other species?

Acknowledgements

This study was supported by grants from the Spanish Ministerio de Ciencia y Tecnología (CGL2009-11917) and Plan Andaluz de Investigación (CVI-6649), and was partially performed by FEDER funds, and a grant from the National Institutes of Health (GM42790). We thank Danna Eickbush for helpful discussions.

References

- Altschul SF, Madden TL, Schäffer AA, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids res* 25: 3389-402.

- Bachtrog D (2003) Accumulation of Spock and Worf, Two Novel Non-LTR Retrotransposons, on the Neo-Y Chromosome of *Drosophila miranda*. *Mol Biol Evol* 20: 173-181.
- Bakkali M & Camacho JPM (2004) The B chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa: III. mutation rate of B chromosomes. *Heredity* 92: 428-33.
- Bibiłło A & Eickbush TH (2002) The reverse transcriptase of the R2 non-LTR retrotransposon: continuous synthesis of cDNA on non-continuous RNA templates. *J Mol Biol* 316: 459-73.
- Burke WD, Calalang CC & Eickbush TH (1987) The site-specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. *Mol Cell Biol* 7: 2221-30.
- Burke WD, Malik HS, Jones JP & Eickbush TH (1999) The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol Biol Evol* 16: 502-11.
- Cabrero J, Perfectti F, Gómez R, Camacho JPM & López-León MD (2003) Population variation in the A chromosome distribution of satellite DNA and ribosomal DNA in the grasshopper *Eyprepocnemis plorans*. *Chromosome Res* 11: 375-81.
- Cabrero J, Alché JD & Camacho JPM (1987) Effects of B chromosomes of the grasshopper *Eyprepocnemis plorans* on nucleolar organiser regions activity. Activation of a latent NOR on a B chromosome fused to an autosome. *Genome* 29: 116-121.
- Cabrero J, López-León MD, Bakkali M & Camacho JPM (1999) Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* 83: 435-9.
- Cabrero J, Teruel M, Carmona FD, Jiménez R & Camacho JPM (2007) Histone H3 lysine 9 acetylation pattern suggests that X and B chromosomes are silenced during entire male meiosis in a grasshopper. *Cytogenet Genome Res* 119: 135-142.
- Cabrero J & Camacho JPM (2008). Location and expression of ribosomal RNA genes in grasshoppers: abundance of silent and cryptic loci. *Chromosome* 16: 595-607.
- Camacho JPM, Bakkali M, Corral JM, et al. (2002) Host recombination is dependent on the degree of parasitism. *P Roy Soc B-Biol Sci* 269 (1505): 2173-2177.

- Camacho JPM, Cabrero J, Viseras E, et al. (1991) G banding in two species of grasshopper and its relationship to C, N, and fluorescence banding techniques. *Genome* 34: 638-643.
- Camacho JP, Sharbel TF & Beukeboom LW (2000) B-chromosome evolution. *Philos T Roy Soc B* 355: 163-78.
- Camacho JPM, Shaw MW, López-León MD, Pardo MC & Cabrero J (1997) Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper *Eyprepocnemis plorans*. *The American naturalist* 149: 1030-50.
- Camacho JPM (2005) B Chromosomes, in E Inc. (ed), *The Evolution of the Genome*, Gregory, T. R., San Diego, pages. 223-286.
- Capy P, Gasperi G, Biéumont C & Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85: 101-6.
- Christensen SM & Eickbush TH (2005) R2 Target-primed reverse transcription: ordered cleavage and polymerization steps by protein subunits asymmetrically bound to the target DNA. *Mol Cell Biol* 25: 6617-6628.
- Darlington CD. 1958. *Evolution of Genetic Systems*. Edinburgh and London: Oliver and Boyd.
- Eickbush MT & Eickbush TH (2011) Retrotransposition of R2 elements in somatic nuclei during the early development of *Drosophila*. *Mobile DNA* 2: 11.
- Eickbush TH & Robins B (1985) *Bombyx mori* 28S ribosomal genes contain insertion elements similar to the Type I and II elements of *Drosophila melanogaster*. *EMBO Journal* 4 (9): 2281-2285.
- Eickbush TH & Eickbush DG (2007) Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 175: 477-85.
- Eickbush TH, Burke WD, Eickbush DG & Lathe WC (1997) Evolution of R1 and R2 in the rDNA units of the genus *Drosophila*. *Genetica* 100: 49-61.
- Feng J, Zeng R & Chen J (2008) Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. *BioTechniques* 44: 901-12.
- Gardner MJ, Hall N, Fung E, et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498-511.

- George JA & Eickbush TH (1999) Conserved features at the 5' end of *Drosophila* R2 retrotransposable elements: implications for transcription and translation. *Insect Mol Biol* 8: 3-10.
- George JA, Burke WD & Eickbush TH (1996) Analysis of the 5' junctions of R2 insertions with the 28S gene: implications for non-LTR retrotransposition. *Genetics* 142: 853-863.
- Ghesini S, Luchetti A, Marini M & Mantovani B (2011) The non-LTR retrotransposon R2 in termites (Insecta, Isoptera): characterization and dynamics. *J Mol Evo* 72: 296-305.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, in *Nucleic acids symposium series*, pages. 95–98.
- Henriques-Gil N & Arana P (1990) Origin and substitution of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Evolution* 44: 747–753.
- Jakubczak JL, Burke WD & Eickbush TH (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *P Natl Acad Sci USA* 88: 3295-9.
- Jakubczak JL, Zenni MK, Woodruff RC & Eickbush TH (1992) Turnover of R1 (Type I) and R2 (Type II) retrotransposable elements in the ribosomal DNA of *Drosophila melanogaster*. *Genetics* 131 (1): 883-894.
- Jamburuthugoda VK & Eickbush TH (2011) The reverse transcriptase encoded by the non-LTR retrotransposon R2 is as error-prone as that encoded by HIV-1. *J Mol Biol* 407: 661-72.
- Jones RN (1985) Are B chromosomes selfish?, in T Cavalier-Smith (ed), *The Evolution of Genome Size*, Wiley, pages. 397-425.
- Jurka J, Kapitonov VV, Pavlicek A, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110: 462-467.
- Kidwell MG & Lisch D (1997) Transposable elements as sources of variation in animals and plants. *P Natl Acad Sci USA* 94: 7704-11.
- Kojima KK & Fujiwara H (2004) Cross-genome screening of novel sequence-specific non-LTR retrotransposons: various multicopy RNA genes and microsatellites are selected as targets. *Mol Biol Evol* 21: 207-17.

- Kojima KK & Fujiwara H (2005) Long-term inheritance of the 28S rDNA-specific retrotransposon R2. *Mol Biol Evol* 22: 2157-65.
- Lamb JC, Meyer JM, Corcoran B, et al. (2007) Distinct chromosomal distributions of highly repetitive sequences in maize. *Chromosome Res* 15: 33-49.
- Lathe W, Burke WD, Eickbush DG & Eickbush TH (1995) Evolutionary stability of the R1 retrotransposable element in the genus *Drosophila*. *Mol Biol Evol* 12 (6): 1094.
- Lathe WC & Eickbush TH (1997) A single lineage of R2 retrotransposable elements is an active, evolutionarily stable component of the *Drosophila* rDNA locus. *Mol Biol Evol* 14: 1232-41.
- Luan DD, Korman MH, Jakubczak JL & Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 72: 595-605.
- López-León MD, Cabrero J & Camacho JPM (1991) A nucleolus organizer region in a B chromosome inactivated by DNA methylation. *Chromosoma* 100: 134-138.
- López-León MD, Neves N, Schwarzacher T, et al. (1994) Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Res* 2: 87-92.
- López-León MD, Vázquez P, Hewitt GM & Camacho JPM (1995) Cloning and sequence analysis of an extremely homogeneous tandemly repeated DNA in the grasshopper *Eyprepocnemis plorans*. *Heredity* 75: 370-375.
- López-León MD, Cabrero J, Pardo MC, et al. (1993) Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity* 71: 352-362.
- Malik HS, Burke WD & Eickbush TH (1999) The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16: 793-805.
- McAllister BF & Werren JH (1997) Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma* 106: 243-253.
- Miao VP, Covert SF & VanEtten HD (1991) A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. *Science* 254: 1773-1776.
- Mingazzini V, Luchetti A & Mantovani B (2011) R2 dynamics in *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca): turnover rate and 28S concerted evolution. *Heredity* 106: 567-575.

- Muñoz-Pajares AJ, Martínez-Rodríguez L, Teruel M, et al. (2011) A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* 187: 853-63.
- Nur U (1977) Maintenance of a “Parasitic” B Chromosome in the Grasshopper *Melanopus femur-rubrum*. *Genetics* 87: 499-512.
- Nur U (1966) The effect of supernumerary chromosomes on the development of mealy bugs. *Genetics* 54: 1239-1249.
- Östergren G (1945) Parasitic nature of extra fragment chromosomes. *Botaniska Notiser* 2: 157–163.
- Parkin EJ & Butlin RK (2004). Within- and between-individual sequence variation among ITS1 copies in the meadow grasshopper *Chorthippus parallelus* indicates frequent intrachromosomal gene conversion. *Mol Biol Evol* 21: 1595-1601.
- Plowman AB & Bougourd SM (1994) Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. *Heredity* 72: 587-593.
- Pérez-González CE & Eickbush TH (2001) Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics* 158: 1557-67.
- Randolph L (1928) Types of supernumerary chromosomes in maize. *Anat Rec*: 102.
- Rezende-Teixeira P, Siviero F, Costa Rosa M da & Machado-Santelli GM (2009) The R2 mobile element of *Rhynchosciara americana*: molecular, cytological and dynamic aspects. *Chromosome Res* 17: 455-67.
- Roiha H, Miller JR, Woods LC & Glover DM (1981) Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in *D. melanogaster*. *Nature* 290: 749-754.
- Rose T (2003) CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic acids res* 31: 3763-3766.
- Ruíz-Ruano FJ, Ruiz-Estévez M, Rodríguez-Pérez J, et al. (2011) DNA Amount of X and B Chromosomes in the Grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenet Genome Res* 134: 120-6.
- Schafer M & Kunz W (1987). Ribosomal gene amplification does not occur in the oocytes of *Locusta migratoria*. *Developmental Biology* 120: 418–424.

- Shaw MW & Hewitt GM (1990) B chromosomes, selfish DNA and theoretical models: where next? *Oxford Surveys in Evolutionary Biology* 7: 197–223.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2007) Nucleolus size variation during meiosis and NOR activity of a B chromosome in the grasshopper *Eyprepocnemis plorans*. *Chromosome Res* 15 (6): 755–765.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2009) Quantitative analysis of NOR expression in a B chromosome of the grasshopper *Eyprepocnemis plorans*. *Chromosoma* 118: 291-301.
- Teruel M, Cabrero J, Perfectti F, Acosta MJ, et al. (2009) Microdissection and chromosome painting of X and B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenet Genome Res* 125: 286-291.
- Theuri J, Phelps-Durr T, Mathews S & Birchler J (2005) A comparative study of retrotransposons in the centromeric regions of A and B chromosomes of maize. *Cytogenet Genome Res* 110: 203-8.
- Weinstock GM, Robinson GE, Gibbs RA, et al. (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443: 931-949.
- White WJD. 1973. *Animal Cytology and Evolution*, 3rd ed. London: Cambridge University Press.
- Wilson E (1907) The supernumerary chromosomes of Hemiptera. *Science* 26: 870-871.
- Yang J, Malik HS & Eickbush TH (1999) Identification of the endonuclease domain encoded by R2 and other site-specific, non-long terminal repeat retrotransposable elements. *P Natl Acad Sci USA* 96: 7847-52.
- Zhang X & Eickbush TH (2005) Characterization of active R2 retrotransposition in the rDNA locus of *Drosophila simulans*. *Genetics* 170: 195-205.
- Zhang X, Zhou J & Eickbush TH (2008) Rapid R2 retrotransposition leads to the loss of previously inserted copies via large deletions of the rDNA locus. *Mol Biol Evol* 25: 229-37.
- Zhou J & Eickbush TH (2009) The pattern of R2 retrotransposon activity in natural populations of *Drosophila simulans* reflects the dynamic nature of the rDNA locus. *PLoS genetics* 5: e1000386.

Ziegler CG, Lamatsch DK, Steinlein C, et al. (2003) The giant B chromosome of the cyprinid fish *Alburnus alburnus* harbours a retrotransposon-derived repetitive DNA sequence. *Chromosome Res* 11: 23-35.

Zurita S, Cabrero J, López-León MD & Camacho JPM (1998) Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution* 52: 274-277.

CAPÍTULO 4

Conserved phylogenetic signal for two mobile elements in acridid grasshoppers

Eugenia E. Montiel, A. Jesús Muñoz-Pajares, Josefa Cabrero, Juan Pedro M. Camacho,
Francisco Perfectti, M^a Dolores López-León

Introduction

Transposable elements (TEs) have revealed as a main actor in genome evolution. Recent advance in genome sequencing projects has shown ubiquitous presence of TEs in eukaryote genomes shaping a high proportion of their nucleotide composition (e.g. 45% in human and 85% in maize) (Lander et al. 2001; Schnable et al. 2009). The giant grasshopper genomes (2-5 folds the human genome) might also harbour huge amounts of TEs, but no genome has yet been sequenced in these organisms.

About 9600 species have been described within the superfamily group Acridomorpha (see <http://orthoptera.speciesfile.org>), 6500 of which belong to the family Acrididae, the subfamily Gomphocerinae showing the highest number of species. In spite of the large number of species and a huge amount of studies performed on them, TEs have scarcely been studied in these organisms. The first analyses focused on the R1 and R2 retrotransposons in insects, including Orthoptera (Jakubczak et al. 1991), and on the *mariner* transposon in several species of grasshopper (Robertson et al. 1993). Only recently, the role of mobile elements in chromosomal rearrangements in the oedipodine grasshopper *Trimerotropis pallidipennis* (Flores et al. 2007) or in neuronal response during phase change in the locust *Locusta migratoria* (Guo et al. 2010) has been investigated.

Here we analyze the evolutionary history of two class-I (RNA-mediated) TEs, *Gypsy* and *RTE*, in 31 grasshopper species belonging to the family Acrididae. An analysis of the distribution and abundance of these elements in the species *Eyprepocnemis plorans* and *Locusta migratoria* has shown that they are abundant in these genomes, with high preference for euchromatic regions (Montiel et al. submitted). The *RTE* element is non-LTR and presumably quite old since it is broadly spread in animals. It was first described in *Caenorhabditis elegans* (Youngman et al. 1996) but it has also been found in mammals, reptiles, amphibia, insects, fish and marine invertebrates, showing, in all of them, a very conserved molecular structure (Zupunski et al. 2001). The *Gypsy* element, however, is an LTR retrotransposon thus showing a molecular structure similar to retroviruses and, in some cases, it can harbour the baculoviral gene *env* coding for capsid proteins (Malik et al. 2000). *Gypsy* was, in fact, the first retrovirus found in invertebrates (Kim et al. 1994).

TEs are genomic parasites theoretically showing obligate vertical (parent to offspring) transmission, on which basis they should show molecular phylogenies being congruent with those of the host species. But their mobile nature, with presence of multiple paralogous copies within a same genome, makes it difficult to trace back their evolutionary

history, and thus they frequently show phylogenies being inconsistent with the phylogeny of the host species. Some of these inconsistencies have been explained by horizontal transposon transfer (HTT), a process having profusely been documented in the main TE groups (for review, see Schaack et al. 2010). For instance, the evolutionary history of the *Gypsy* element in the Drosophilids appears to have involved HTT events (Héredia et al. 2004; Setta et al. 2009), likewise some non-LTR elements among reptiles and ruminants (Piskurek et al. 2007), the best studied TEs being *Mariner* (Hartl et al. 1997) and the *P* element (Houck et al. 1991; Clark & Kidwell 1997; Kidwell 2000).

Molecular phylogenetics has been very useful in acridid grasshoppers, and Orthoptera in general, to elucidate some phylogenetic relationships that had remained obscure with morphological analysis (Flook & Rowell 1997, 1999; Rowell & Flook 1998; Chapco et al. 1999; Amédegnato et al. 2003; Litzenberger & Chapco 2003; Zhang et al. 2005; Contreras & Chapco 2006; Bugrov et al. 2006; Fries et al. 2007; López et al. 2007; Yin et al. 2008; Fenn et al. 2008). In most cases, mitochondrial DNA markers were used.

We analyse here the evolutionary pattern of two TEs in 31 acridid species, by means of phylogenetic analysis of DNA sequences for mtDNA (as a representative compartment of the host genome) and the *RTE* and *Gypsy* retrotransposons.

Materials and methods

Specimens belonging to 31 grasshopper species belonging to the family Acrididae, and one species belonging to the family Pamphagidae (*Eumigus monticolus*) to be used as outgroup, were collected in Spain and Russia (Table 8). DNA was extracted with the Gen Elute Mammalian Genomic DNA Miniprep kit (Sigma-Aldrich) following manufacturer's instructions. This DNA was used to PCR amplify three fragments of mtDNA and one fragment of the *Gypsy* and *RTE* retrotransposons. In the mtDNA, we amplified two fragments of the cytochrome oxidase I gene, one at 5' end (COI5') and another at the 3' end (COI3'), as well as a fragment of the NADH dehydrogenase 5 (ND5) gene. In mobile elements, we amplified conserved fragments from the *pol* domain in *Gypsy* and ORF2 in *RTE* (see Table 9 for a description of used primers). PCR reaction was similar for all sequences: 10 ng DNA, 0.25 µM each primer, 0.2 mM dNTPs, 1.25 mM MgCl₂ and 1U Taq polimerasa (MBL) in 50 µl. We used a standard PCR program consisting in a denaturation step at 94°C for 5', followed by 30 cycles of 30 s at 94°C, 30 s at 55°C and 30 s at 72°C, and a final step at 72°C for 8 min, with minor modifications for each amplicon. Namely, to amplify the ND5 fragment, we had to increase each cycle time to

Table 8. Species analysed and geographical origin..

Subfamily/Tribe	Species	Localization
Calliptaminae		
/Calliptamini	<i>Calliptamus barbarus</i>	Sierra Nevada, Granada
Cantantopinae		
/Pezotettigini	<i>Pezotettix giornae</i>	Sierra Nevada, Granada
Cyrtacanthacridinae		
/Cyrtacanthacridini	<i>Anacridium aegyptium</i>	Granada, Granada
	<i>Schistocerca gregaria</i>	Laboratory culture
Eyprepocnemidinae		
/Eyprepocnemidini	<i>Eyprepocnemis plorans</i>	Torrox, Málaga
	<i>Eyprepocnemis unicolor</i>	Tajikistán, Rusia
	<i>Heteracris adspersa</i>	Dhagestán, Rusia
	<i>Heteracris littoralis</i>	Torrox, Málaga
Gomphocerinae		
/Gomphocerini	<i>Chorthippus binotatus</i>	Sierra Nevada, Granada
	<i>Chorthippus jacobsi</i>	Sierra Nevada, Granada
	<i>Chorthippus jucundus</i>	Capileira, Granada
	<i>Chorthippus nevadensis</i>	Sierra Nevada, Granada
	<i>Chorthippus vagans</i>	Sierra Nevada, Granada
	<i>Stauroderus scalaris</i>	Trevenque, Granada
/Dociostaurini	<i>Dociostaurus genei</i>	Salobreña, Granada
	<i>Dociostaurus jagoi</i>	Monachil, Granada
	<i>Dociostaurus marrocanus</i>	El Purche, Granada
/Stenobothrini	<i>Omocestus bolivari</i>	Sierra Nevada, Granada
	<i>Omocestus minutissimus</i>	Sierra Nevada, Granada
	<i>Omocestus panteli</i>	Granada, Granada
	<i>Omocestus raymondi</i>	Monachil, Granada
	<i>Stenobothrus festivus</i>	Sierra Nevada, Granada
Oedipodinae		
/Acrotylini	<i>Acrotylus fischeri</i>	El Purche, Granada
	<i>Acrotylus patruelis</i>	Algarrobo, Málaga
/Aiolopini	<i>Aiolopus thalassimus</i>	Torrox, Málaga
/Locustini	<i>Locusta migratoria</i>	El Padul, Granada
	<i>Oedaleus decorus</i>	Capileira, Granada
/Oedipodini	<i>Oedipoda caerulescens</i>	El Purche, Granada
	<i>Oedipoda fuscocincta</i>	Sierra Nevada, Granada
/Sphingonotini	<i>Sphingonotus azureus</i>	Nerja, Málaga
	<i>Sphingonotus caerulans</i>	Sierra Nevada, Granada
Outgroups		
	<i>Gryllotalpa orientalis</i>	Genbank
	<i>Ixodes scapularis</i>	RepBase
	<i>Eumigus monticolus</i>	Sierra Nevada, Granada

1 min (instead of 30 s), whereas to amplify TE fragments we had to decrease annealing temperature till 50°C, all remaining parameters being the same. Products were visualized in a 1.5% agarose electrophoresis gel. Fragments of the expected size were cut out from the gel, separated from the agarose with the GenElute Gel Extraction Kit (Sigma-Aldrich) and sequenced in both directions by Macrogen Inc. Sequence identity was checked using Blast against GenBank sequences. Outgroup species for each molecular marker were obtained from the GenBank and RepBase (<http://www.girinst.org/repbase/update/index.html>). To select the most appropriate outgroup for retroelements, we employed the on-line CENSOR tool (Jurka et al. 2005).

Table 9. Primer description and length of the amplified fragment.

Primer	Sequence 5'->3'	Target
UEA3*	TATAGCATTCCCACGAATAAATAA	Cytochrome oxidase I gene
UEA4*	AATTTTCGGTCAGTTAATAATATAG	Cytochrome oxidase I gene
UEA9*	GTAAACCTAACATTTTTTCCTCAACA	Cytochrome oxidase I gene
UEA10*	TCCAATGCACTAATCTGCCATATTA	Cytochrome oxidase I gene
N5-J**	ACTCACCTCAACCAGAATCAA	NADH deshydrogenase subunits 5
N5-N**	ACTCATGCTTTATTTAAGGCTTTA	NADH deshydrogenase subunits 5
Gyp_Int900F***	GATACGGGCGCAACTGTTCTG	Gypsy reverse polymerase
Gyp_Int1300R***	TTAGGGATAAGTTGAGACTGCTC	Gypsy reverse polymerase
qPCR_RTEF***	AGATTGGGAAACGAGGCACTG	RTE ORF2
RTE_IntR***	CATCCATACAAGGCAACACTC	RTE ORF2

Amplicon	Primers Pair	Fragment Length
COI 5'	UEA3 / UEA 4	371
COI 3'	UEA9 / UEA 10	307
ND5	N5-J / N5-N	688
Gypsy	Gyp_Int900F / Gyp_Int1300R	432
RTE	qPCR_RTEF / RTE_IntR	218

* Lunt et al. 1996.

** Bensasson et al. 2000

*** Montiel et al. Unpublished

DNA sequences were aligned with ClustalW (Thompson et al. 1994) and visually inspected using Bioedit software (Hall 1999). In the case of TEs, the electropherogram of every DNA sequence obtained from PCR product was checked for the presence of more than one peak, in which case we chose the nucleotide showing the highest peak in both readings. This thus resulted in a consensus DNA sequence per species. In mtDNA, however, all DNA sequences showed single peaks. Prior to phylogenetic analysis, a saturation analysis for each marker was performed with the software DAMBE (Xia &

Xie 2001), following the procedure described in Xia et al. (2003). For this purpose, we compared the I_{ss} (Index of Substitution Saturation) of the sequences with the $I_{ss,c}$ (critical level at which the sequences become uninformative due to saturation). In addition, we used the AIC criterion in the jModelTest program (Posada 2008) to determine the best fitted evolutionary model for each marker. Phylogenetic trees were obtained using maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) criteria. MP analyses were performed in PAUP 4.0b8 (Swofford 2003) with 10,000 bootstrapping replicates. ML analyses were done in PhyML-aLRT (Anisimova & Gascuel 2006), with branch support calculated by SH-like approximate likelihood ratio tests (aLRT). The BI analyses were performed with MrBayes (Ronquist & Huelsenbeck 2003) using two independent MCMC with 2,000,000 generations and discarding (burnin) the first 25% of trees in each chain to obtain a 50% majority rule consensus tree with a posterior probability (PP) for each node (Ronquist & Huelsenbeck 2003). As outgroups for TEs, we tried to use *Eumigus monticolus* but only the *RTE* primers amplified in this species. In the case of *Gypsy*, we used as outgroup that element which in the RepBase showed the highest similarity with grasshopper sequences, i.e. the *Gypsy-23_IS* element from the tick *Ixodes scapularis*.

We performed a second congruence analysis, following Campbell et al. (2009), by means of the CADM (Congruence Among Distance Matrices) test, where the null hypothesis is that matrices are incongruent. This test was performed in the CADM.global and CADM.post instructions of the APE package (Paradis et al. 2004) implemented in R (R Development Core Team 2010). We used matrices of patristic distances obtained from the different phylogenetic trees to compare topology and branch length. In the case of mitochondrial markers, we used matrices of distances calculated with the K80 model of substitution. The significance of matrices congruence was obtained by 9999 permutations.

The congruence among trees obtained for different molecular markers was analyzed using a SH-test (Shimodaira & Hasegawa 1999) implemented in PAUP. In this test, which compares multiples topologies, the null hypothesis is that all trees explain the data equally well, so that the trees are congruent. In this SH-test we compared the topology obtained from each marker with the mtDNA data. The significance was obtained by 1000 bootstrap replicates.

To get deeper insight on *Gypsy* intragenomic diversity, we ligated the fragments obtained from *E. plorans* into TOPO-TA vector (Invitrogen) and transformed One Shot TOP10 cell (Invitrogen). We chose this species because it had shown presumptive HTT. In this case, all DNA sequences obtained after cloning showed single peaks for all nucleotide positions. Alignments and trees including these sequences were performed as described above.

Results

In all 31 species analyzed, PCR amplification with the five primer pairs yielded a single fragment in every case, which was sliced out from the gel and sequenced in both directions. Sequences were blasted and the homology with the expected sequences was confirmed. Alignments were obtained without difficulty due to the absence of indels. Stop codons were not observed.

In order to build a mtDNA tree, we first performed a matrix congruence analysis for the three mitochondrial fragments (COI5', COI3' and ND5) (Table S1). The CADM test rejected the null hypothesis of incongruence ($P < 0.001$) and showed a congruence close to 80% (Kendal's $W = 0.79$). In addition, post-hoc tests comparing the matrices by pairs showed that all matrix pairs were significantly congruent. These results allowed concatenation of the three mtDNA fragments in a single DNA sequence.

A total of 1098 positions were available for the mtDNA phylogenetic analysis, 625 of which were variable and 409 were parsimoniously informative sites. In the case of *Gypsy*, we analyzed 353 nucleotides, 243 of which were variable and 61 were parsimoniously informative. Finally, out of the 176 nucleotides analyzed for the *RTE* element, 113 were variable and 57 were parsimoniously informative. A saturation analysis (see methods) showed no saturation for retroelements sequences in any of the three codon positions, but it showed saturation in the third codon position in the case of the COI and ND5 mtDNA sequences (Table S2). However, the trees obtained including or eliminating the third codon position were roughly similar (data not shown), for which reason we included them in the subsequent analyses. The evolutionary model for the concatenated mtDNA sequences was GTR+I+G, whereas *Gypsy* and *RTE* fitted to the HKY+G model (Table 10). These models were employed in the three methods for phylogenetic analysis.

Table 10. Evolutionary models best fitted for each marker under the AIC criterion in the jModelTest program (Posada 2008).

	Mitochondrial combined	Gypsy	RTE
Model	GTR+I+G	HKY+G	HKY+G
-lnL	9777.41	1639.84	926.80
Base freq.			
A	0.3060	0.2828	0.3306
C	0.1198	0.2436	0.1436
G	0.1267	0.1999	0.2329
T	0.4474	0.2737	0.2929
Rate matrix			
A-C	1.7160		
A-G	9.6816		
A-T	1.8769		
C-G	0.9760		
C-T	7.4818		
G-T	1.0000		
Gamma shape	0.6960	1.6180	0.3770
Pinv	0.2690		
Kappa		4.5741	6.2854
Ti/Tv		2.2569	3.0425

The mtDNA trees obtained by MP and ML showed similar topology, but that built by BI differed in the branches being closer to the outgroup (*Grylotalpa orientalis*). Nine clades with high branch support, coinciding with different tribes, were observed in the three trees (Figure 11, Figure S1). Clade A included four *Omocestus* species belonging to the tribe Stenobothrini, but another member of this tribe, i.e. *Stenobothrus festivus*, was included into clade B along with species of the genus *Chorthippus* (Tribe Gomphocerini). Clade C included the three *Dociostaurus* species (tribe Dociostaurini). These three clades show common descent and all species are included in the subfamily Gomphocerinae. The tribu Cyrtacanthacridinae is coincident with clade D with large branch support in the ML and MP tree (1 and 93, respectively). Clade E corresponds with the tribe Eyprepocnemidini and, although branch support is less than 50 in the MP tree, it is larger for the ML (0.88) and BI (0.93) trees. Likewise, clades F, G, H and I included species belonging to tribes Oedipodini, Sphingonotini, Locustini and Acrotylini, respectively. These four clades thus include species belonging to the subfamily Oedipodinae. But only the BI tree suggested the monophyletic origin of this subfamily, the ML and MP indicating that they could possibly be paraphyletic.

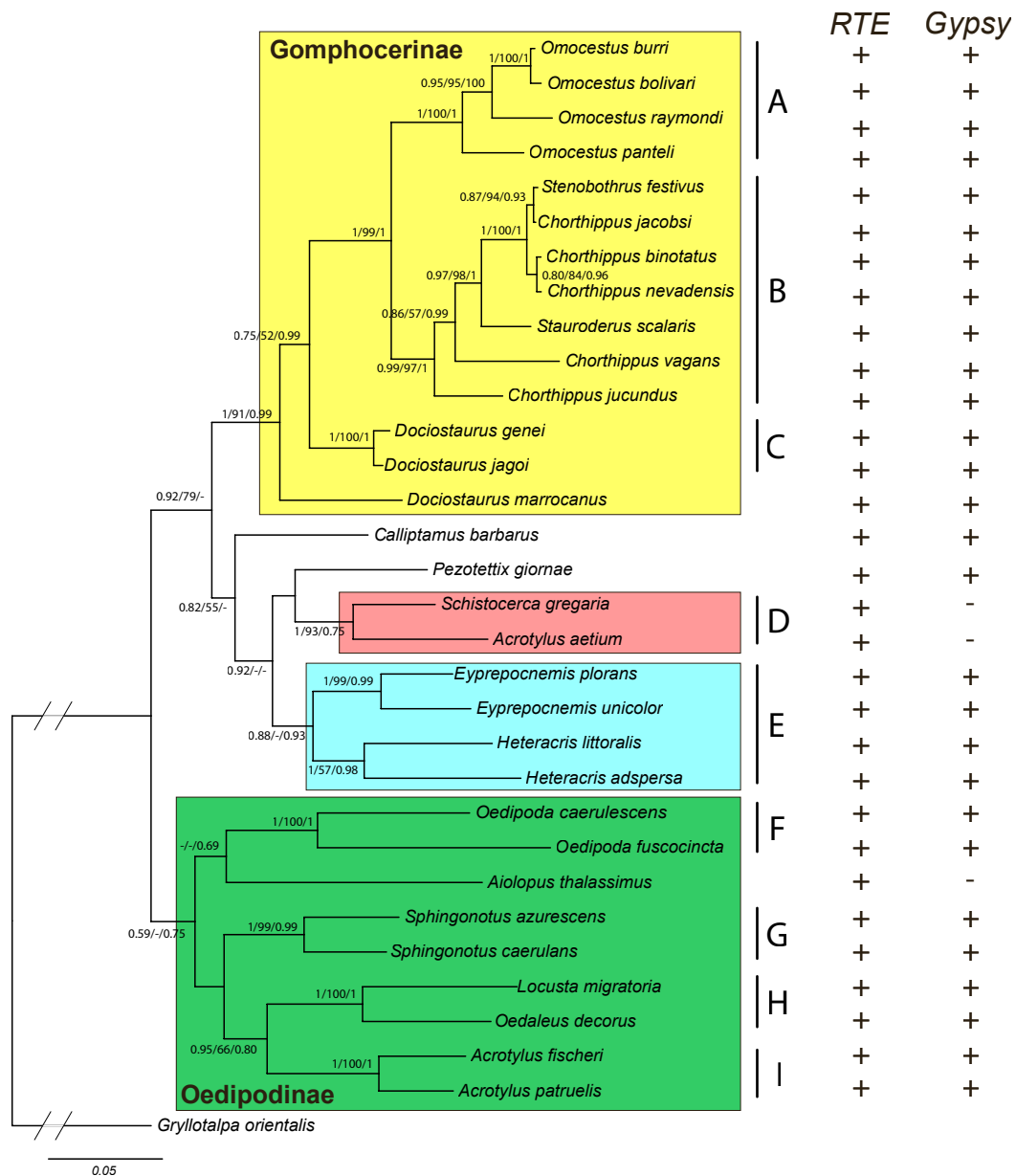


Figure 11. Phylogenetic tree obtained, by maximum likelihood, with the three mtDNA markers combined. Only branch supports higher than 50% (maximum parsimony, MP) or 0.5 (maximum likelihood, ML, and bayesian inference, BI) are shown in the order ML/MP/BI. Capital letters correspond with clades coinciding with different tribes: A. Stenobothrini; B. Gomphocerini; C. Dociostaurini; D. Eyprepocnemidini; E. Oedipodini; F. Sphingonotini; G. Locustini and H. Acrotylini. Clades A, B and C contain species belonging to the Subfamily Gomphocerinae and clades E, F, G and H, include species from the Subfamily Oedipodinae. The two columns on the right indicate the result of PCR amplification for RTE and Gypsy.

The trees built for the *Gypsy* and *RTE* retroelements showed lower resolution than the mtDNA markers, yielding a high number of polytomies. However, some of the patterns observed for the mtDNA tree are conserved in them. In the case of *RTE*, the ML tree (see it in Figure 12, and MP and BI trees in Figure S2) showed significant branch supports for five of the nine tribe level clades supported by mtDNA (clades C, E, G, H and I). *RTE* elements appeared to be almost monophyletic in the subfamily Gomphocerinae,

with the single exception of *Ch. nevadensis* which seems to carry *RTE* elements being rather different from those in the remaining gomphocerine species analysed here, for which reason this species is grouped with the Eyprepocnemidinae in the MP and BI trees (in this latter also including *C. barbarus* and *P. giornae*). The *RTE* elements in the four Eyprepocnemidinae species are monophyletic in the three trees, with significant branch support. In the subfamily Oedipodinae, however, *RTE* elements in the two *Sphingonotus* species are clearly different from those in the remaining oedipodine species, and this difference is clear in the three trees. A last interesting result is the similarity of the *RTE* elements in *Ch. vagans* with those in the *Dociostaurus* species, with remarkably high significance (0.98 in ML, 90 in MP and 1 in BI). This high similarity points to a possible case of recent HTT.

The trees built for *Gypsy* (Figure 13, Figure S3) did only maintain significance in clade G (tribe Locustini), and all elements at subfamily level were paraphyletic. The species from the subfamily Oedipodinae are a good example (clades F-I). In the subfamily Gomphocerinae, the *Gypsy* elements in the *Dociostaurus* genus are different from those in the remaining gomphocerine species, and this is apparent in the three trees. In addition, the *Gypsy* elements in the subfamily Eyprepocnemidinae share a highly supported clade (1 in ML, 84 in MP and 0.86 in BI) with the Catantopinae *P. giornae* and the Calliptaminae *C. barbarus*. And, most remarkably, the *Gypsy* elements in *E. plorans* and *H. littoralis* are grouped together with very high branch support (0.95 in ML, 91 in MP and 0.96 in BI). In fact, these elements show 97.77% similarity between *E. plorans* and *H. littoralis*, which is higher than those shown with their respective species of the same genus (94.98% in *Eyprepocnemis* and 96.94% in *Heteracris*). This also points to possible HTT.

In order to know the level of agreement between the retrotransposon and mtDNA phylogenies we conducted two congruence tests with opposite null hypothesis, thus supplying complementary information. On the other hand, the CADM test showed evidence for some congruence between TE and mtDNA trees, being it higher for *RTE* ($W = 0.72$, $\text{Chi}^2 = 669.22$, $p = 0.0001$) than for *Gypsy* ($W = 0.67$, $\text{Chi}^2 = 504.47$, $p = 0.0013$). The SH test showed that retrotransposon phylogenies were incongruent with the topology of grasshopper mtDNA phylogeny (*RTE*: $-2\Delta\ln L = 465.33$ $p < 0.001$; *Gypsy*: $-2\Delta\ln L = 637.06$ $p < 0.001$). The joint conclusion from both tests is that TEs and mtDNA trees are roughly congruent, but there exists some significant incongruences, some of them pointing to possible HTTs (see above).

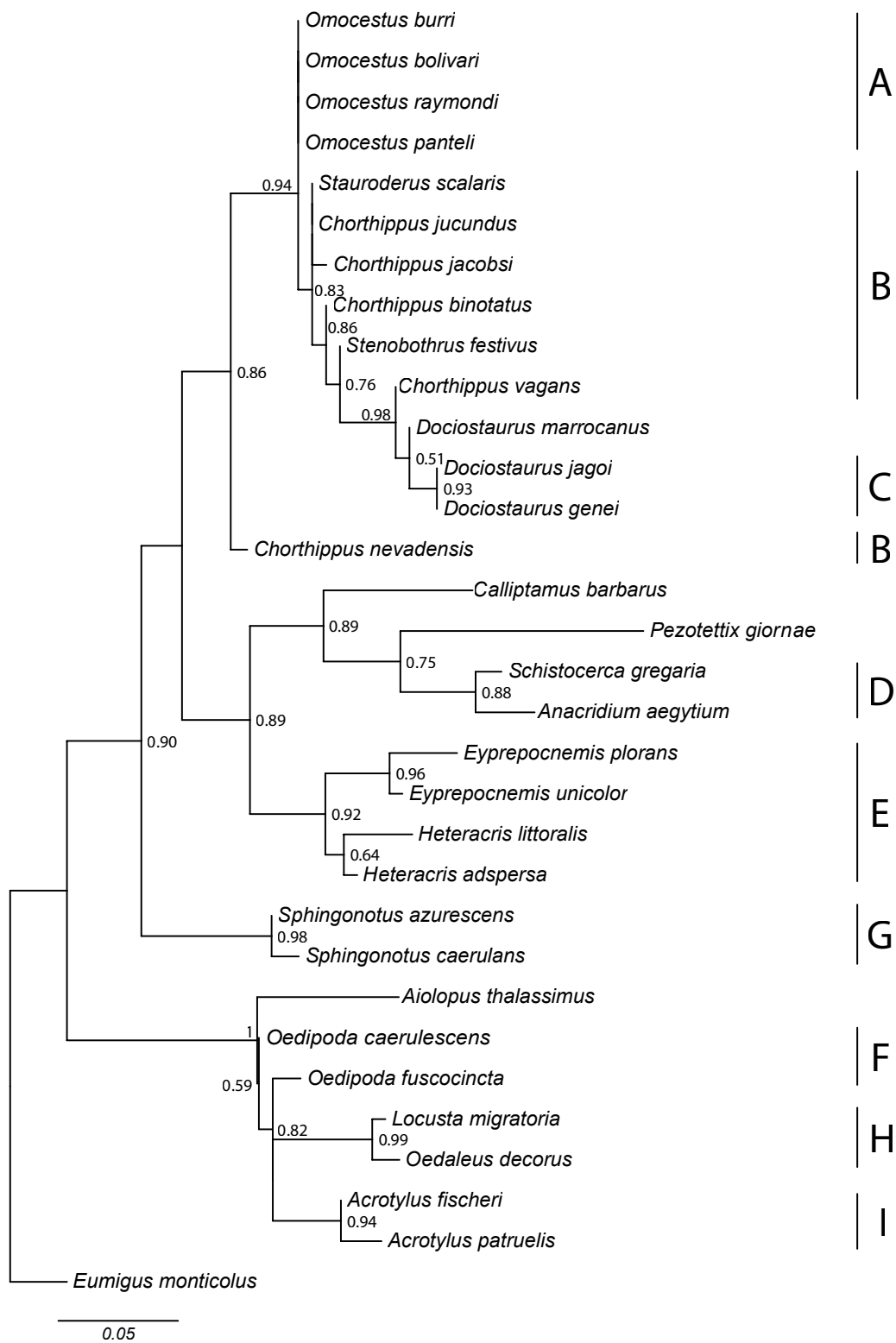


Figure 12. Phylogenetic tree for RTE elements obtained by ML. Only branch supports higher than 0.5 are shown. Clades defined by mtDNA (see Figure 11) are depicted on the right.

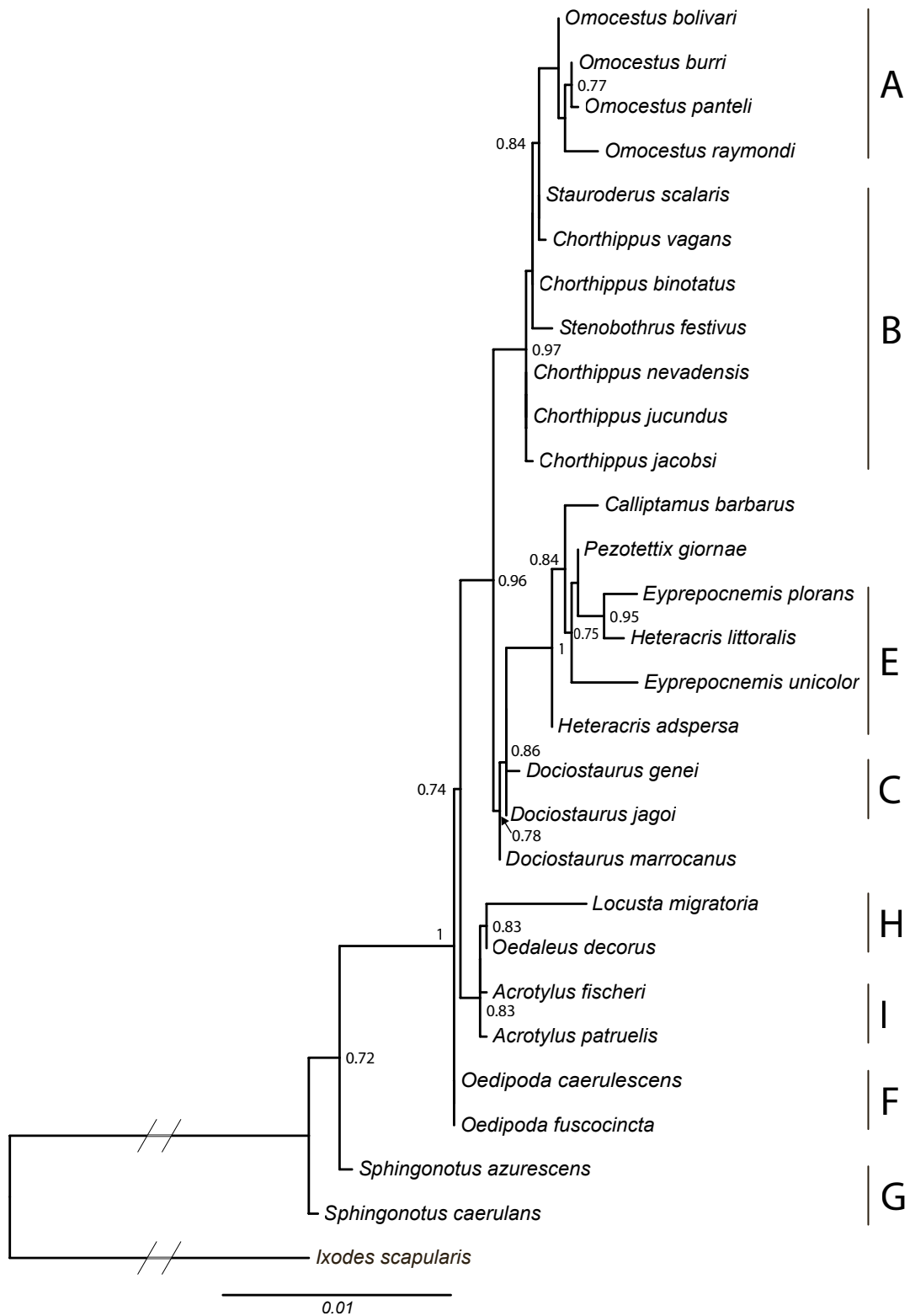


Figure 13. Phylogenetic tree for Gypsy element obtained by ML. Only branch supports higher than 0.5 are shown. Clades defined by mtDNA (see Figure 11) are depicted on the right.

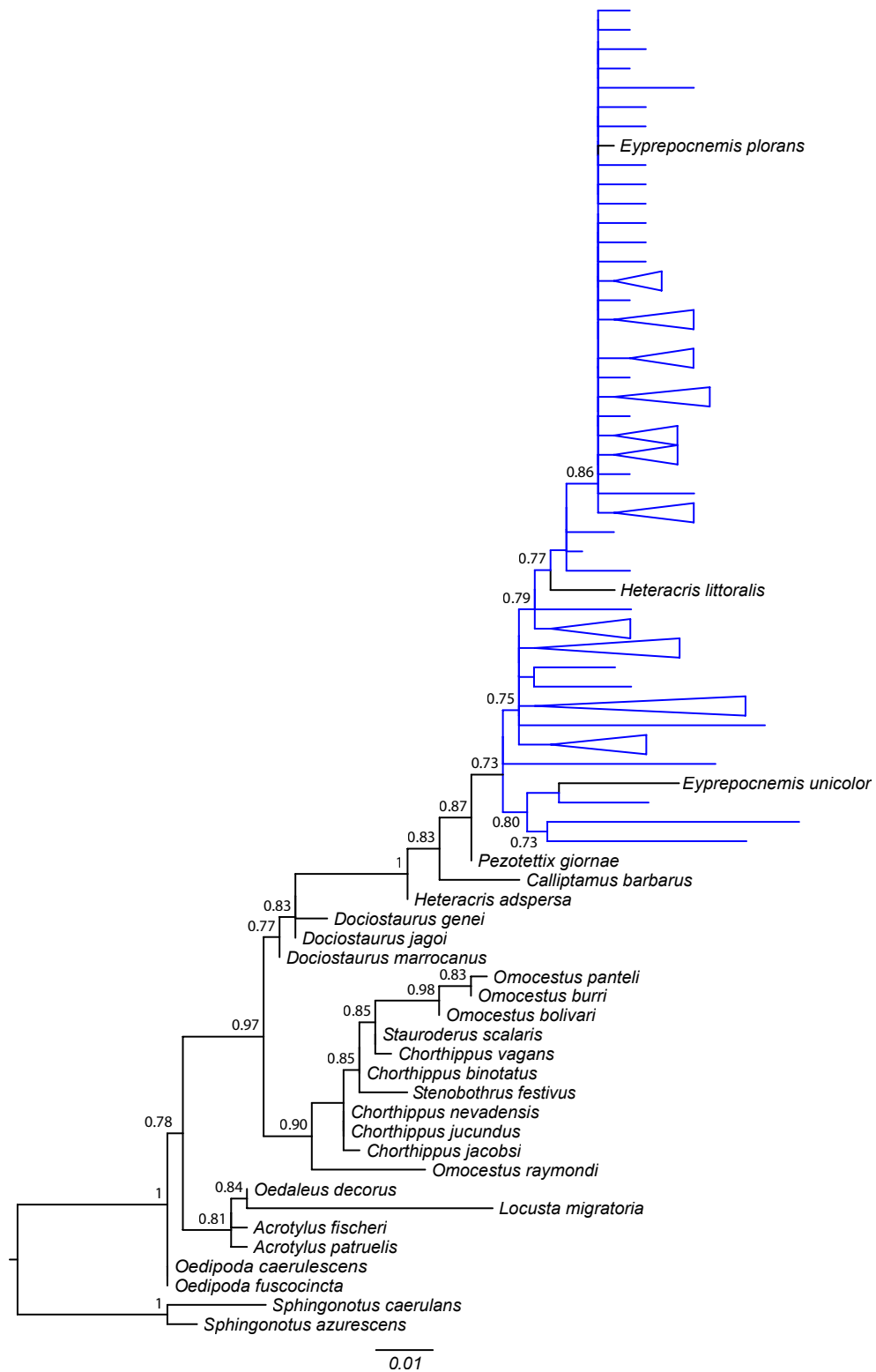


Figure 14. Phylogenetic tree (obtained by ML) for Gypsy elements in 31 species of acridid grasshoppers, including the 84 elements obtained in *E. plorans* (in blue). Only branch supports higher than 0.5 are shown.

To thoroughly analyze the diversity for *Gypsy* elements in one of the species involved in a possible HTT, we cloned the PCR product from *E. plorans*. A total of 120 clones were sequenced and 84 of them showed a non-defective sequence, without gap and stop codons. These sequences of *E. plorans* were aligned with the 31 sequences of *Gypsy* elements obtained from different grasshopper species and an ML tree was built (Figure 14). This tree was very similar to that obtained previously (see Figure 13), and showed all *E. plorans* sequences grouped in a clade along with *Gypsy* elements in *E. unicolor*, in a basal position, and *H. littoralis*, in an internal position. At first sight, this tree would appear to be consistent with the possibility of HTT between *E. plorans* and *H. littoralis*, but the presence of some *E. plorans* *Gypsy* elements grouped with *E. unicolor* in the basal clade call attention on the possibility of ILS for *Gypsy* elements during species differentiation in the Eypreocnemidinae subfamily.

Discussion

TEs have been scarcely studied in Orthoptera (but see Robertson et al. 1993; Jakubczak et al. 1991; Florencia et al. 2007; Guo et al. 2010). We analyse here the retrotransposons *Gypsy* and *RTE* in 31 species belonging to the family Acrididae, which is, to our knowledge, the broader screening of TE presence in these insects. Our analysis of mtDNA, in addition to TEs, has illuminated some interesting pathways of the evolutionary history of these parasitic elements.

The recent proliferation of genome sequencing projects has shown that TEs are a majoritary component in genomes of most kinds of eukaryotes (Kidwell & Lisch 2000, Biemont 2010) representing the 45% in *H. sapiens* (Lander et al. 2001), the 39% in *M. musculus* (Waterston et al. 2002) or 85% in B73 maize genome (Schnable et al. 2009). TEs have thus remarkably contributed, through successive replication waves, to genome size increase (Kidwell 2002). No orthopteran genome has yet been completely sequenced, and nothing can be said about which proportion of their giant genomes (up to 16.93 pg DNA in the grasshopper *Podisma pedestris*) is constituted by TEs, but cytological observation of several TEs, e.g. *Gypsy*, *RTE* and *Mariner* indicate that they are very abundant in euchromatic regions, which was also suggested by qPCR estimations of copy number (Montiel et al. submitted). The phylogenies shown here also suggest that TEs appear to be an old component of grasshopper genomes, since they appeared in all groups analysed and they roughly resemble grasshopper taxonomy. However, we want to

call attention to the fact that our analysis is based on PCR amplification and this creates an artificial filter by which we can only analyse those TE elements conserving DNA sequence complementarily with the primers employed, whereas many defective elements showing mutations in this region could not be recovered.

The mtDNA phylogeny for the 31 acridid species analyzed here is consistent with the current taxonomy in the Orthoptera Species File online database (<http://orthoptera.speciesfile.org>), and also with previous molecular trees built for these organisms (Contreras & Chapco 2006; Bugrov et al. 2006), which suggests that this gene tree most likely represents the true species tree. The tree provided good resolution at genus and tribe levels, but somewhat poorer resolution at subfamily level. It is remarkable the basal position of the three *Dociostaurus* species, within the subfamily Gomphocerinae, in coincidence with their ancestral $2n\♂=23$ chromosome number, compared to that in the derived A and B clades including the tribes Stenobothrini and Gomphocerini showing $2n\♂=17$ chromosomes as a result of three fixed centric fusions (Hewitt 1979). The only inconsistency with taxonomy was the inclusion of a *Stenobothrus* species (*S. festivus*) within the Gomphocerini (clade B) instead of within the Stenobothrini (clade A), suggesting that the taxonomy of the genera *Stenobothrus*, *Omocestus* and *Chortippus* probably needs some revision.

The molecular phylogenies of the two TEs analysed resemble that of mtDNA, with high values of congruence provided by the CADM test: 72% between *RTE* and mtDNA and 67% between *Gypsy* and mtDNA. However, there are also some remarkable incongruences (see below), as shown by the SH test. *RTE* phylogeny shows a high general consistence with mtDNA phylogeny, with the subfamily Oedipodinae as the most basal group, and the subfamily Gomphocerinae as the most recent one. This coincidence with the mtDNA phylogeny is not unexpected since it is known that the *RTE* element is one of the most widespread, and allegedly very old, TEs in eukaryotes (Malik & Eickbush 1998; Malik et al. 1999).

An interesting feature of *RTE* elements is the high conservatism of their aminoacid sequence, even among distantly related groups, in comparison with other sequences in the host genome (Zupunski et al. 2001). This can explain the high degree of polytomy found in the *RTE* grasshopper phylogeny shown here, which is most remarkable in the Gomphocerinae with $2n=17$, which appear to be the youngest clade in the mtDNA phylogeny. Remarkably, one of the few well resolved clades, with 0.98 branch support, within this subfamily includes *Ch. vagans* ($2n=17$) with the three *Dociostaurus* species ($2n=23$), which is suggestive of possible HTT between these species.

The *Gypsy* phylogeny also shows polytomies within the Gomphocerinae and Oedipodinae subfamilies, which suggests a scarce degree of differentiation of this element within these groups. But the most remarkable feature in this phylogeny is the grouping of *E. plorans* and *H. littoralis* in a common clade with 0.95 support also suggesting HTT between these species. HTT has been well documented for *Gypsy* in drosophilids (Herédia et al. 2004) being TE inclusion within a virus capsid is a mechanism suggested for HTT (Friesen & Nissen 1990; Piskurek et al. 2007). In addition, if the TE itself would acquire the ability to generate a capsid, it would be able to perform the HTT. The viral nature of *Gypsy* has been previously reported, and indeed this TE was eventually considered the first virus described for invertebrates (Kim et al. 1994; Pelisson et al. 1997), although a later analysis suggested that the *Gypsy* TE became a virus by acquiring the *env* genes from a baculovirus (Malik et al. 2000). Hence, the cuasi-viral nature of *Gypsy* supports the possible HTT between *E. plorans* and *H. littoralis* suggested by the phylogeny built here. This is conceivable because these two species usually share the same habitats in coastal zones and are closely associated, frequently being found in the same places (Barranco & Pascual 1993; Olmo-Vidal 2002). Likewise, *Ch. vagans* is frequently found in the same habitats and localities than one or more *Dociostaurus* species (see Badih et al. 2003). This might have provided ample opportunity for common infection by mites or retroviruses acting as vehicles for HTT (Schaack et al. 2010).

However, other processes, such as genetic drift or a different evolutionary dynamics for TEs among species, being dependent on transposon activity level (for review, see Capy et al. 1994), could also explain the inconsistencies between TE and host phylogenies. In fact, the analysis of 84 paralogous copies of *Gypsy* within a same *E. plorans* individual showed the presence in this genome of some copies being similar to the *Gypsy* element found in *H. littoralis* (as observed previously), but there were other copies being similar to those in *E. unicolor*, a congeneric species. This throws serious doubts on HTT as a possible explanation of the phylogenetic incongruence observed, and points to possible incomplete lineage sorting for a high variety of elements that were already present in the common ancestor's genome of these two genera.

Given the paralogous nature of the multiple intragenomic copies of TEs, the likelihood of observing phylogenetic incongruences with host phylogeny is high. In fact, the high congruence between TE and mtDNA phylogenies observed here is surprising bearing also in mind that, in all species but *E. plorans*, we sampled a single TE sequence. Anyway, phylogenetic incongruences do not grant HTT occurrence since many of them

could simply be the product of incomplete lineage sorting (Schaack et al. 2010) or even incomplete sampling. Ascertaining whether this caveat applies to other presumptive cases of HTT, including the *Dociostaurus-Ch. vagans* for RTE suggested by our present results (see above), needs further research analyzing many TE DNA sequences in each species.

Acknowledgements

This study was supported by grants from the Spanish Ministerio de Ciencia y Tecnología (CGL2009-11917) and Junta de Andalucía (CVI-6649), and was partially supported by FEDER funds. We thank Mohammed Bakkali for providing *Schistocerca gregaria* material, Alexander Bugrov for sending us Russian material, and Tatiana López for technical assistance.

References

- Amédégno C, Chapco W & Litzemberger G (2003). Out of South America? Additional evidence for a southern origin of melanopline grasshoppers. *Mol Phylogenet Evol* 29: 115-119.
- Anisimova M & Gascuel O (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* 55: 539-552.
- Badih A, Fortis PJ, Pascual F (2003). Diversidad faunística y ecología de los ortópteros de la comarca sierra sur de Jaén (Insecta: Orthoptera). In: "In memoriam" al profesor Dr. Isidoro Ruiz Martínez, coord. JM Pérez Jiménez, págs. 295-328, ISBN 84-8439-195-7.
- Barranco P, Pascual F (1993). Estudio ecológico de los ortópteros de la vega del río Andarax (Insecta: Orthoptera). Distribución en relación a la vegetación. *Boln. Asoc. Esp. Ent.* 17: 285-297.
- Bensasson D, Zhang D X & Hewitt GM (2000). Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Mol Biol Evol* 17: 406-15.
- Biémont C (2010). A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* 186 (4): 1085-93.
- Bugrov A, Novikova O, Mayorov V, Adkison L & Blinov A (2006). Molecular phylogeny of Palaearctic genera of Gomphocerinae grasshoppers (Orthoptera, Acrididae). *Syst Entomol* 31: 362-368.

- Campbell V, Legendre P & Lapointe FJ (2009). Assessing congruence among ultrametric distance matrices. *J Classif* 26: 103-117.
- Capy P, Anxolabéhère D & Langin T (1994). The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet* 10: 7-12.
- Chapco W., Kuperus WR & Litzenberger G (1999). Molecular Phylogeny of Melanopline Grasshoppers (Orthoptera: Acrididae): The Genus *Melanoplus*. *Ann Entomol Soc Am* 92: 7.
- Clark JB & Kidwell M G (1997). A phylogenetic perspective on P transposable element evolution in *Drosophila*. *P Natl Acad Sci USA* 94: 11428-11433.
- Contreras D & Chapco W (2006). Molecular phylogenetic evidence for multiple dispersal events in gomphocerine grasshoppers. *Journal of Orthoptera Research* 15: 91-98.
- Fenn JD, Song H, Cameron SL & Whiting MF (2008). A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. *Mol Phylogenet Evol* 49: 59-68.
- Flook PK & Rowell CH (1997a). The phylogeny of the Caelifera (Insecta, Orthoptera) as deduced from mtrRNA gene sequences. *Mol Phylogenet Evol* 8: 89-103.
- Flook PK & Rowell CH (1997b). The effectiveness of mitochondrial rRNA gene sequences for the reconstruction of the phylogeny of an insect order (Orthoptera). *Mol Phylogenet Evol* 8: 177-192.
- Flook PK, Klee S & Rowell CH (1999). Combined molecular phylogenetic analysis of the Orthoptera (Arthropoda, Insecta) and implications for their higher systematics. *Syst Biol* 48: 233-53.
- Florencia T, Guzman N, Gonzalez G, et al. (2007). Mobile elements and inverted rearrangements in *Trimerotropis pallidipennis* (Orthoptera: Acrididae). *Biosystems* 60: 212-221.
- Fries M, Chapco W & Contreras D (2007). A molecular phylogenetic analysis of the Oedipodinae and their intercontinental relationships. *Journal of Orthoptera Research* 16: 115-125.

- Friesen PD & Nissent MS (1990). Gene organization and transcription of TED , a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol Cell Biol* 10: 3067-3077
- Guo W, Wang XH, Zhao DJ, Yang PC & Kang L (2010). Molecular cloning and temporal-spatial expression of I element in gregarious and solitary locusts. *J Insect Physiol* 56: 943-8.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, in *Nucleic acids symposium series*, pages. 95–98.
- Hartl DL, Lohe AR & Lozovskaya ER (1997). Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31: 337-58.
- Herédia F, Loreto ELS & Valente VLS (2004). Complex evolution of gypsy in *Drosophilid* species. *Mol Biol Evol* 21: 1831-42.
- Hewitt GM. 1979.. Grasshopper and crickets. In: John B ed. *Animal Cytogenetics*, vol. 3: Insecta 1 Orthoptera. Berlin, Stuttgart: Gebruder Borntraeger.
- Houck MA, Clark JB, Peterson KR & Kidwell MG (1991). Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* 253: 1125-8.
- Jakubczak JL, Burke WD & Eickbush TH (1991). Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *P Natl Acad Sci USA* 88: 3295-9.
- Jurka J, Kapitonov VV, Pavlicek A, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110: 462-467.
- Kidwell MG (2002). Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49-63.
- Kidwell MG & Lisch D (2000). Transposable elements and host genome evolution. *Trends Ecol Evol* 15: 95–99.
- Kim A, Terzian C, Santamaria P, et al. (1994). Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *P Natl Acad Sci USA* 91: 1285-1289.
- Lander ES, Linton LM, Birren B, et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.

- Litzenberger G & Chapco W (2003). The North American Melanoplinae (Orthoptera: Acrididae): a molecular phylogenetic study of their origins and taxonomic relationships. *Ann Entomol Soc Am* 96: 491-497.
- López H, Contreas-Díaz HG, Oromí P, Juan C (2007). Delimiting species boundaries for endangered Canary Island grasshoppers based on DNA sequence data. *Conserv Genet* 8: 587-598.
- Malik HS. (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 10: 1307-1318.
- Malik HS & Eickbush TH (1998). The RTE class of non-LTR retrotransposons is widely distributed in animals and is the origin of many SINEs. *Mol Biol Evol* 15: 1123-1134.
- Malik HS, Burke WD & Eickbush TH (1999). The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16: 793-805.
- Olmo-Vidal JM (2002) Atlas de los Ortópteros de Cataluña. Atlas de Biodiversidad no. 1: 209-336. Generalitat de Catalunya. Departament de Medioambient.
- Paradis E, Claude J & Strimmer K (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20: 289-290.
- Piskurek O & Okada N (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *P Natl Acad Sci USA* 104: 12046-12051.
- Posada D (2008). jModelTest: Phylogenetic averaging. *Mol Biol Evol* 25: 1253-1256.
- Pélisson A, Teyssset L, Chalvet F, et al. (1997). About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the *Drosophila flamenco* host gene. *Genetica* 100: 29-37.
- Robertson HM (1993). The mariner transposable element is widespread in insects. *Nature* 362: 241-245.
- Ronquist F & Huelsenbeck JP (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
- Rowell CHF & Flook PK (1998). Phylogeny of the Caelifera and the Orthoptera as derived from ribosomal gene sequences. *Journal of Orthoptera Research* 7: 147-156 .

- Schaack S, Gilbert C & Feschotte C (2010). Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* 25: 537-546.
- Schnable PS, Ware D, Fulton RS, et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-5.
- Setta N de, Sluys M-A Van, Capy P & Carareto CMA (2009). Multiple invasions of Gypsy and Microplasmid retroelements in genus *Zaprionus* and melanogaster subgroup of the genus *Drosophila*. *BMC Evol Biol* 9: 279.
- Shimodaira H & Hasegawa M (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16: 1114–1116.
- Swofford DL (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. DL Swofford (ed). Sinauer Associates Massachusetts.
- Thompson JD, Higgins DG & Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- Waterston RH, Lindblad-Toh K, Birney E, et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-62.
- Xia X. & Xie Z. (2001). DAMBE: software package for data analysis in molecular biology and evolution. *Hered J* 92: 371–373.
- Xia X, Xie Z, Salemi M, Chen L & Wang Y (2003). An index of substitution saturation and its application. *Mol Phylogenet Evol* 26: 1-7.
- Yin X, Li X, Wang W, et al. (2008). Phylogenetic analyses of some genera in Oedipodidae (Orthoptera : Acridoidea) based on 16S mitochondrial partial gene sequences. *Insect Sci*: 471-476.
- Youngman S, Luenen HG van & Plasterk RH (1996). *Rte-1*, a retrotransposon-like element in *Caenorhabditis elegans*. *FEBS letters* 380 (1-2): 1-7.
- Zhang D, Li X, Wang W, et al. (2005). Molecular phylogeny of some genera of Pamphagidae (Acridoidea, Orthoptera) from China based on mitochondrial 16S rDNA sequences. *Zootaxa* 1103: 41–49.
- Zupunski V, Gubensek F & Kordis D (2001). Evolutionary dynamics and evolutionary history in the *RTE* clade of non-LTR retrotransposons. *Mol Biol Evol* 18: 1849-63.

CAPÍTULO 5

Analysis of intragenomic distribution of *RTE* retrotransposons through single-cell microdissection of individual bivalents

Eugenia E. Montiel, Josefa Cabrero, Juan Alberto Marchal, Antonio Sánchez,
Francisco Perfectti, María Dolores López-León & Juan Pedro M. Camacho

Introduction

Whole genome sequencing has emerged as a breakthrough key to study the evolutionary dynamics of DNA sequences at intragenomic level. Complete sequencing of relevant organisms such as human beings (Lander et al. 2001), *Drosophila melanogaster* (Adams et al. 2000), mice (Consortium et al. 2002), *Caenorhabditis elegans* (*C. elegans* Sequencing Consortium, 1998) or *Arabidopsis thaliana* (Initiative 2000) has enabled countless research about genome composition and DNA sequences distribution. However, this kind of information is absent for most organisms, although the advent of next generation sequencing (NGS) will undoubtedly fill this gap in a few years. All DNA sequencing projects, however, face the problem of assembling chromosome regions being rich in repetitive DNA, such as centromere and telomere regions, and this problem is no less for paralogous gene copies located in several or most chromosomes.

Chromosome microdissection was developed by Scalenghe et al. (1981) in polytene chromosomes of *Drosophila*, and short after it was applied to human chromosomes (Bates et al. 1986; Kao 1987). From then on, many different applications have emerged (see Cannizzaro 1996; Zhou and Hu 2007 for review). It is currently very used in both plants and animals to characterize chromosome domains or chromosome regions (Miklos et al. 1988; Jung et al. 1992), chromosome arms (Liu et al. 1997) or whole chromosomes (Zhou et al. 1999), autosomes or sexual chromosomes (Zimmer et al. 1997). One of the main current uses of chromosome microdissection is for chromosome painting (Guan et al. 1994; Houben et al. 2001; Marchal et al. 2004; Teruel et al. 2009). But it has rarely been used for the analysis of specific DNA sequences from specific chromosomes (Cheng & Lin 2003; Lamb et al. 2005; Teruel et al. 2010), even though this latter approach may be very useful for a broad range of genomic studies. For instance, clinical genetics has taken particular advantage of microdissection facilitating genetic disease analysis such as the identification of the affected genes in Prader-Willi syndrome (Buiting et al. 1990) or characterization of marker chromosomes in prenatal diagnosis (Mahjoubi et al. 2005).

Separate chromosome microdissection of every chromosome in a single cell can be useful to reveal the intragenomic distribution of paralogous gene copies, by simply amplifying them separately in the DNA extracted from each chromosome. This may also be useful for gene mapping purposes, the analysis of genomic structure and even to

unveil- the evolutionary history of paralogous genes. One of the most abundant classes of this kind of genes are transposable elements. Their mobility and ubiquity within most eukaryotic genomes make them suitable elements to perform intragenomic analysis by comparing sequence variation within and among chromosomes.

Eyprepocnemis plorans is a grasshopper species with a giant genome (10.39 pg; see Ruíz-Ruano et al. 2011) being rich in mobile elements, such as *Gypsy* and *RTE* retrotransposons and *Mariner* transposons (Montiel et al., submitted). They are widely represented in euchromatic regions of all standard (A) chromosomes and, in a lesser extent, on B chromosomes (a kind of dispensable heterochromatic chromosomes; see Camacho 2005). In this work we microdissected all 11 autosome bivalents, and the X and B chromosome univalents from a single first meiotic metaphase cell, in order to analyse the intragenomic evolutionary dynamics of the *RTE* elements in *E. plorans*.

Materials and methods

We microdissected all chromosomes (11 bivalents and X and B univalents) in a single spermatocyte at first meiotic metaphase, from a B-carrying male captured in Torrox (Málaga, Spain), and stored them into separate 0.2 ml eppendorf tubes with 1X Buffer Taq polymerase (MBL) and 2 mg/ml proteinase K (Sigma), to be subsequently denatured at 50°C for 1 hour and at 99°C for 4 minutes. Chromosomes were named in order of decreasing area, excepting X and B univalents which were easily identified. Chromosome area was measured (in arbitrary units) using the ImageJ program, version 1.44p [<http://imagej.nih.gov/ij>]. Measures were repeated three times, in different days, to corroborate the results. PCR reactions were performed, on each chromosome, to amplify the ORF2 region of the *RTE* element, with the primers qPCR_RTEf (5'-AGA TTG GGAAAC GAG GCA CTG-3') and RTE_IntR (5'-CAT CCA TAC AAG GCA ACA CTC-3') which yield a 218 bp fragment of the *RTE* retrotransposon. For this purpose, we added 2mM MgCl₂, 0.16 mM dNTP, 4% DMSO, 0.4 uM of each primer and 1U Taq polymerase (MBL). The reaction conditions consisted in an initial denaturation step at 95°C for 5 min followed by 30 cycles of 20 s at 94°C, 30 s at 55°C and 45 s at 72°C, and a final extension for 8 min at 72°C. PCR products were visualized in a 1% agarose gel, and the fragment obtained from each chromosome was extracted from the gel with the Gel Elute Kit (Sigma), cloned into TOPO-TA vector (Invitrogen) and transformed into One Shot TOP10 cell (Invitrogen).

The insert of 17-28 clones per chromosome was sequenced in both directions (Macrogen) using the standard M13 primers. *RTE* elements were also obtained by PCR amplification with the same primers in 31 grasshopper species, and PCR products were sequenced in both directions.

DNA sequences were aligned with the program BioEdit (Hall 1999) using ClustalW (Thompson et al. 1994). The analysis of nucleotide polymorphism was performed with DnaSP v.5 (Librado & Rozas 2009). We obtained the following parameters: number of segregating sites (S), number of haplotypes per chromosome (H), average number of pairwise differences per site (π), and heterozygosity per site expected in a population at mutation-drift equilibrium given the observed number of mutations value (Θ or Watterson's estimator; Watterson 1975). The MEGA 4.0 program (Tamura et al. 2007) was used to test for positive and purifying selection comparing the number of synonymous and nonsynonymous substitutions per site. Codon-based tests of neutrality were conducted using the Nei-Gojobori method (Nei and Gojobori 1986) and all positions containing gaps and missing data were eliminated from the dataset. Arlequin 3.0 software (Excoffier et al. 2005) was used to test the genomic structure of these sequences. Saturation tests, performed following the procedure described in Xia et al. (2003), indicated the absence of saturation for *RTE* sequences. In addition, DNA sequences were cleaned of poorly aligned and uninformative positions through the software GBlocks (Castresana 2000), and thereafter we built phylogenetic trees of these mobile elements through the maximum likelihood method with the online version of PHYML (Dereeper et al. 2008; Stéphane Guindon & Olivier Gascuel 2003), using a GTR evolutionary model, and employing the aLTR method (Anisimova and Gascuel 2006) to calculate branch support and the Splitstree4 program (Huson & Bryant 2006) to process the obtained tree. Statistical analyses were performed with the Statistica 8.0 software.

Results

Once discarded primer regions, a total of 178 bp from the ORF2 region of the *RTE* element were obtained in 289 clones, 74.4% of which lacked indels or stop codons, suggesting that they were non-defective, and thus they could be potentially active. The proportion of non-defective elements did not show significant differences among chromosomes (contingency $\chi^2 = 8.03$, d.f.= 12, $p = 0.78$).

Non-defective sequences showed similar values for nucleotide diversity (π) and heterozygosity per site (Θ) among chromosomes, excepting chromosomes 11 and B which showed remarkably lower values (Table 11). Nucleotide diversity observed in chromosomes 11 and B (≈ 0.05) was about one third that observed in the remaining A chromosomes (0.13-0.16).

Table 11. Summary of diversity analysis to *RTE*. N: number of sequences; H: number of haplotypes; s: number of sites; S: number of segregating sites; π : nucleotide diversity (per site) and Θ : heterozygosity (per site) from number of mutations.

Chromosome	N	H	s	S	π	Θ
1	20	18	177	86	0.14618	0.13695
2	19	19	166	107	0.16050	0.20420
3	24	24	178	92	0.14167	0.13841
4	21	19	171	92	0.13876	0.14954
5	21	20	173	96	0.14776	0.15424
6	25	22	177	84	0.12992	0.12568
7	26	25	171	97	0.14456	0.14865
8	23	17	174	88	0.10979	0.13703
9	18	17	178	86	0.14555	0.14047
10	21	18	176	91	0.14194	0.14371
11	17	14	161	50	0.05595	0.09186
B	28	20	172	60	0.04802	0.08964
X	26	17	174	87	0.13675	0.13103
Non-Defective	227	186	178	168	0.13245	0.15730
Defective	62	56	124	120	0.18991	0.20607
All	289	223	124	124	0.14364	0.16021

We also compared synonymous and non-synonymous substitutions in non-defective *RTE* sequences to test whether selection was operating on them using a codon-based tests of neutrality. It showed that purifying selection was acting for these retroelements in most chromosomes excepting, again, chromosomes 11 and B, where neutrality could not be rejected (Table 12).

Considering that our collection of *RTE* elements came from specific locations in the *E. plorans* genome (i.e. each of the 12 standard chromosomes and the B chromosome), we analysed the genetic structure of these 13 population samples of DNA sequences by means of an AMOVA with all the haplotypes obtained with DnaSP v.5 (including defective and non-defective elements and considering gap). The results indicated that most genetic variation was found within chromosomes (85.53%), although there was also significant variation among chromosomes (14.47%) thus resulting in an F_{ST} value (0.145) being

Table 12. Codon-based test of neutrality, positive selection and purifying selection for analysis averaging over all sequence pairs within each chromosome. The probability of rejecting the null hypothesis of strict-neutrality ($dN = dS$), in favor of the alternative hypothesis, positive selection, ($dN > dS$) and purifying selection, ($dN < dS$) is shown. Values of P lower than 0.05 are highlighted in bold-type letter. dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively. The variance of the difference was computed using the bootstrap method (500 replicates). Analyses were conducted using the Nei-Gojobori method in MEGA4. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were a total of 57 codons RTE in the final dataset.

Chromosome	Neutrality ($dN = dS$)		Positive Selection ($dN > dS$)		Purifying Selection ($dN < dS$)	
	Prob.	Stat.	Prob.	Stat.	Prob.	Stat.
1	0.000	-6.749	1.000	-6.572	0.000	6.886
2	0.000	-4.480	1.000	-4.412	0.000	4.471
3	0.000	-5.545	1.000	-5.462	0.000	5.500
4	0.000	-4.626	1.000	-4.626	0.000	4.824
5	0.000	-4.472	1.000	-4.429	0.000	4.520
6	0.000	-3.657	1.000	-3.579	0.000	3.598
7	0.000	-4.753	1.000	-4.653	0.000	4.892
8	0.006	-2.818	1.000	-2.795	0.002	2.937
9	0.000	-5.128	1.000	-4.930	0.000	5.158
10	0.000	-4.599	1.000	-4.869	0.000	4.942
11	0.365	-0.910	1.000	-0.928	0.173	0.945
X	0.000	-3.893	1.000	-3.914	0.000	4.004
B	0.156	-1.427	1.000	-1.551	0.068	1.498

significantly higher than zero (Table 13). Pairwise F_{ST} values between the *RTE* elements in the B chromosome were lower when compared with *RTE* elements in chromosome 11 than when compared with all other A chromosomes (Table 14), suggesting the highest similarity among elements in chromosomes 11 and B.

Table 13. . AMOVA with the 244 haplotypes found.

Source of variation	df	Sum of squares	Variance components	Percentage of variation
Among chromosomes	12	655.28	1.94	14.47
Within chromosomes	276	3171.94	11.49	85.53
Total	288	3827.22	13.43	

$$F_{ST} = 0.145$$

Significance tests (1023 permutations)

$$V_a \text{ and } F_{ST} : P(\text{rand. value} > \text{obs. value}) < 0.00001$$

$$P(\text{rand. value} = \text{obs. value}) < 0.00001$$

$$P\text{-value} < 0.00001 \pm 0.00001$$

Table 14. Pairwise F_{ST} values between RTE elements in B chromosomes and those in every A chromosome. Note in bold-type letter the lowest F_{ST} values between B and S₁₁ RTE elements.

B chromosome vs. A chromosome no.	Pairwise F_{ST} for RTE elements		
	Non-defective	Defective	Total
1	0.43124	0.44898	0.44617
2	0.08155	0.09118	0.10515
3	0.29541	0.23117	0.32225
4	0.14183	0.10294	0.13024
5	0.19843	0.20161	0.21746
6	0.28161	0.11268	0.26112
7	0.4594	0.38909	0.4373
8	0.69384	0.66912	0.60566
9	0.44122	0.25	0.40362
10	0.42066	0.29752	0.35711
11	0.01653	0.04878	0.00221
X	0.33341	0.36066	0.30157

All RTE sequences, defective and non-defective, were aligned, and poorly aligned positions were removed with the Gblocks software. The resulting phylogenetic tree showed the existence of two kinds of RTE elements, one of them (lineage A) being present in all chromosomes excepting chromosomes 11 and B, and the other (lineage B) being present in all chromosomes, including A and B chromosomes (Figure 15). Both kinds of elements included defective and non-defective copies, with defective ones showing longer branches (Mann-Whitney test, $P < 0.01$). Nucleotide diversity (π) and heterozygosity per site (Θ) of lineage A elements (0.077 and 0.164, respectively) were higher than those of lineage B elements (0.053 and 0.148, respectively). Likewise, codon-based tests of neutrality, applied separately to both kinds of RTE elements, revealed purifying selection for lineage A elements in most chromosomes where they were present (with the only exception of chromosome 2) but neutrality for lineage B elements in most chromosomes, including chromosomes 11, B and five other chromosomes (Table S3).

A detailed analysis of the DNA sequence of both RTE lineages found in *E. plorans* showed the existence of 32 nucleotide differences, 23 of which were synonymous changes whereas 9 of them implied aminoacid replacements (Figure 16). This indicates that both RTE lineages are rather divergent. A phylogenetic tree of RTE elements in *E. plorans* and those found in 31 species of grasshoppers also belonging to the family Acrididae (Figure 17) suggested that lineage A elements are similar to RTE elements found in close

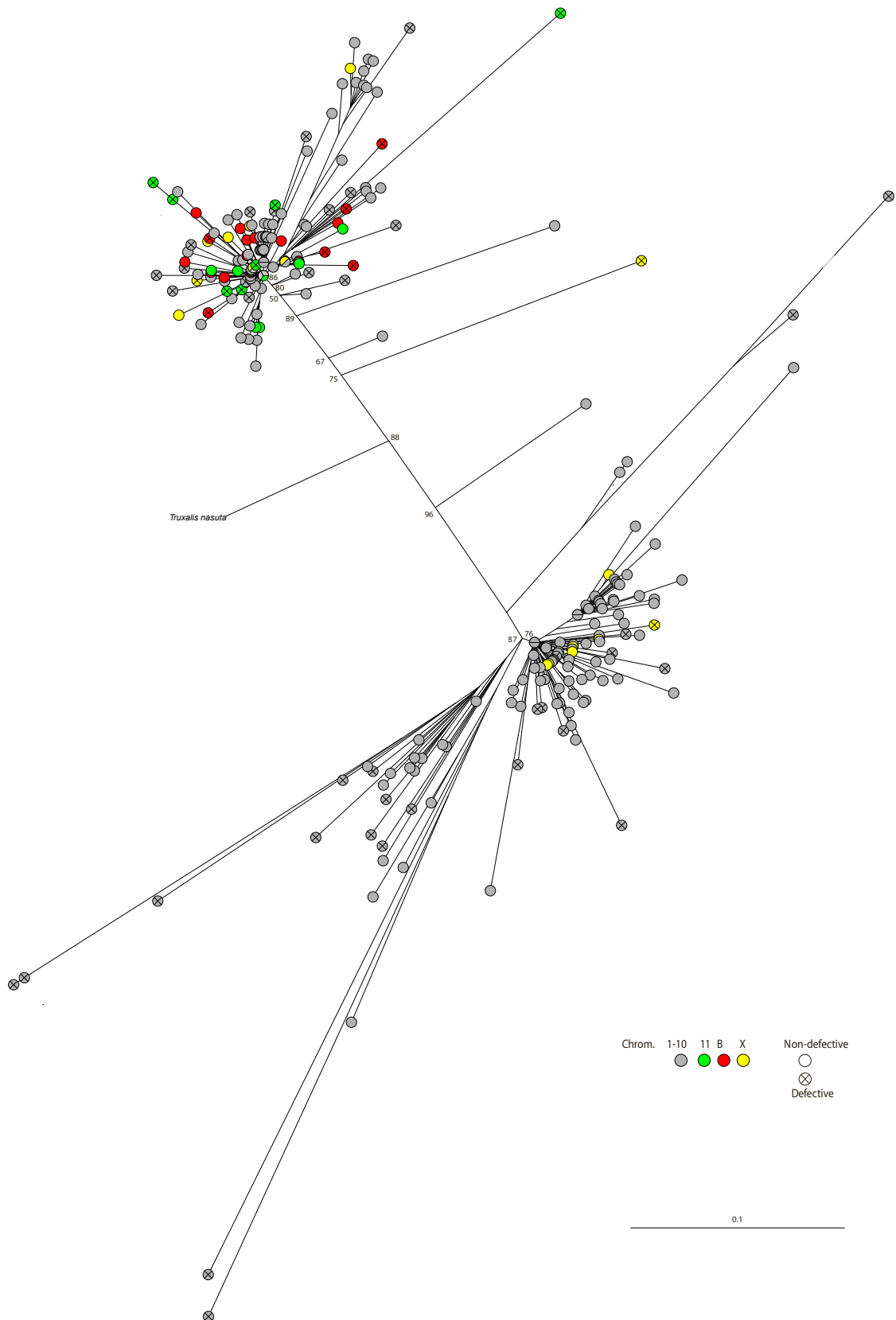


Figure 15. Phylogenetic tree of *RTE* element fragments amplified from every chromosome in *E. plorans*. Chromosomes are differentiated by color: grey= autosomes 1-10, yellow= X chromosome, green= autosome 11, red= B chromosome. Defective copies are marked with an x. Note the presence of B and 11 chromosomes only in the upper cluster.

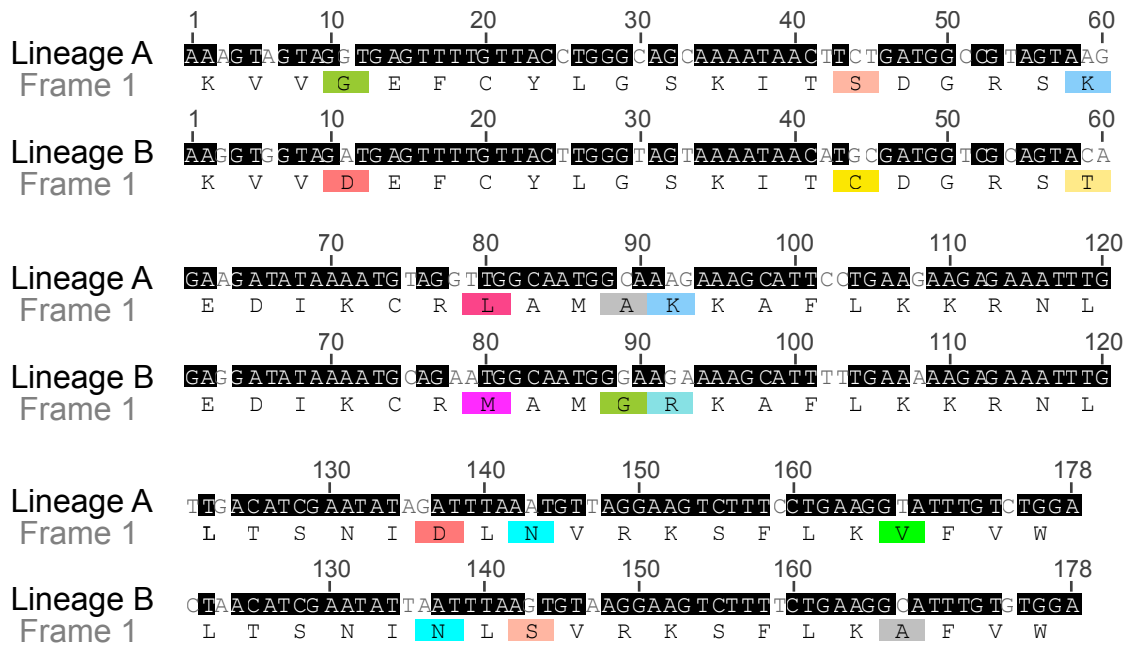


Figure 16. Alignment of consensus sequences for the two types of RTE elements found in *E. plorans* after analysing a 178 bp fragment of the ORF2 region. Note the presence of 32 variable positions, 23 of which correspond with synonymous changes whereas 9 of them provoke aminoacid replacements (marked in different colors).

relative species within the subfamily Eyprepocnemidinae (e.g. *Eyprepocnemis unicolor*, *Heteracris littoralis* and *H. adspersa*) whereas lineage B elements (i.e. those being present in all chromosomes, including the B chromosome) are most similar to RTE elements found in species belonging to the subfamily Oedipodinae.

Discussion

Our analysis of part of the retrotranscriptase gene from RTE elements in the grasshopper *E. plorans*, by means of microdissection of individual chromosomes, PCR amplification, cloning and sequencing, has provided information on their intragenomic structure. Most of the analysed DNA sequences lacked stop codons or indels, and we considered them as non-defective. However, since the analysed sequence may represent less than 10% of total DNA sequence in full-length elements (*RTE-1* is 4kb length in *Danio rerio*, and 2.8kb in *N. vitripennis* (Jurka et al. 2005)) it is conceivable that the proportion of defective RTE elements in *E. plorans* is surely much higher, bearing in mind that many of the elements considered here as non-defective might carry loss of function mutations in other parts of the unanalysed DNA. In any case, more than 25% of RTE elements in the *E. plorans* genome are defective thus forming a part of the presumable huge amounts of junk DNA in this giant genome containing more than 10^{10} bp (Ruiz-Ruano et al. 2011).

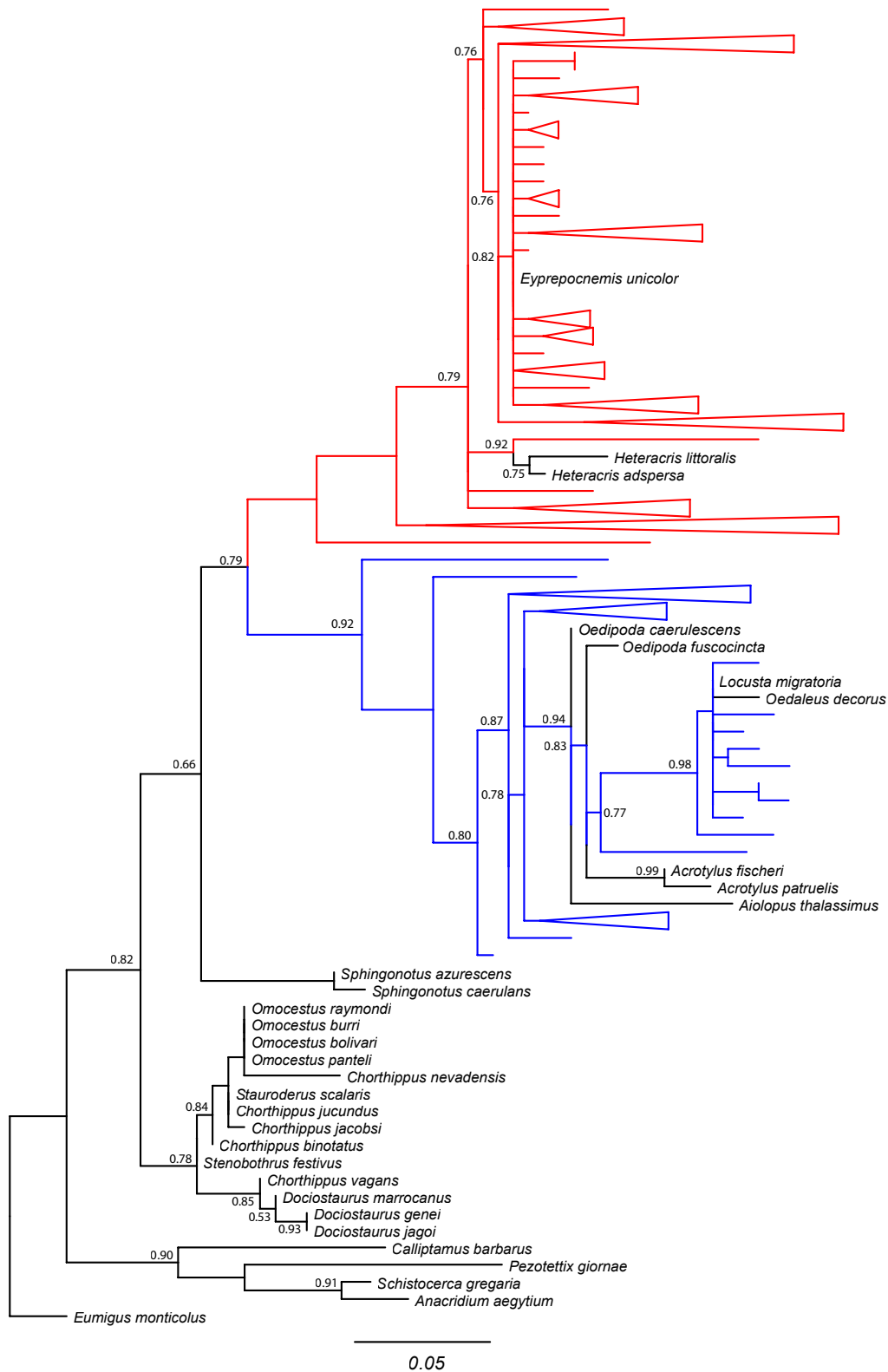


Figure 17. Phylogenetic tree, built by the ML method, of non-defective RTE elements in *E. plorans* and RTE elements in 31 species of grasshoppers belonging to the family Acrididae. Lineage A RTE elements in *E. plorans* are represented in red, and lineage B elements in blue. *Eumigus monticolus* (family Pamphagidae) was used as outgroup. Note how lineage A elements in *E. plorans* are similar to RTE elements in other species belonging, likewise *E. plorans*, to the subfamily Eyprepocnemidinae, whereas lineage B elements are most similar to RTE elements in species belonging to the subfamily Oedipodinae.

With the above caution in mind, the proportion of defective and non-defective elements was similar among chromosomes, including B chromosomes. This is not consistent with the expectation that B chromosomes are havens for mobile elements (Camacho, Sharbel & Beukeboom 2000), under which these chromosomes should show a higher proportion of non-defective (putatively active) elements. But two facts may explain why B chromosomes behave like A chromosomes in this specie. First, *E. plorans* B chromosomes appear to be rather young (Muñoz-Pajares et al. 2011) and thus they have not had much time to accumulate mobile elements. Second, *RTE* elements are preferentially located at euchromatic regions in *E. plorans* (Montiel et al., submitted) and B chromosomes in this species are mostly heterochromatic. These two facts also explain the lower nucleotide diversity found in *RTE* elements from chromosomes 11 and B, this also being indicative of a smaller population size for *RTE* elements in these two chromosomes.

The degree of nucleotide diversity found here for *RTE* elements (0.05-0.16, see Table 11) is larger than those observed for other DNA sequences in *E. plorans*, such as, for instance, the ITS-1 region in the rDNA ($\pi = 0.014$) (Teruel 2009), a fact being expected for elements using an RNA molecule as transposition intermediate. These diversity values are one order of magnitude larger than those reported for active copies of LTR and non-LTR elements in *Drosophila melanogaster* (Sánchez-Gracia et al. 2005), but are similar to those reported for *TOPI* elements in *Anopheles gambiae* ($\pi = 0.051$) (Subramanian et al. 2008).

Although the diversity of retro- and transposon families has been widely analysed (e.g., Bartolomé et al. 2002, Maside and Charlesworth 2002; Roy et al. 2000; Jin-Shan et al. 2005; Gomulski et al. 2004; Marin and Llorens 2000; Hollister and Gaut 2007), only a few studies have connected diversity with chromosome location. Good examples of this kind of works are found in Hood et al.(2005), where they analyze the genetic structure of *Copia* and *Helitron* families in electrophoretically separated chromosomes from *Microbotryum violaceum*. Likewise, Kejnovsky et al. (2007) analysed diversity for *Retand*, a *Gypsy*-like retrotransposon, in microdissected sex chromosomes and autosomes in *Silene latifolia*. Our present study, however, is the first trying to ascertain the intragenomic distribution of *RTE* retroelements including all standard chromosomes and also B chromosomes. This has provided valuable information on the evolutionary history of the *RTE* elements in the *E. plorans* genome, and also suggests a possible origin for the B chromosome in this species.

Our phylogenetic analysis revealed the existence of two kinds of *RTE* retroelements in the *E. plorans* genome, only one of them (lineage B) being present in chromosomes 11 and B, and the remaining chromosomes carrying both kinds. This explains the lower nucleotide diversity for *RTE* elements in these two chromosomes, since two kinds of rather divergent elements were found in all chromosomes but 11 and B. The existence of two different lineages for *RTE* in the *E. plorans* genome points to the occurrence of two different invasions by *RTE* elements in the past history of its genome, both events predating B chromosome origin. One of them (lineage A) reached all A chromosomes except chromosome 11 and B chromosome whereas the other (lineage B) reached all chromosomes, including 11 and B. Since lineage A *RTE* elements are also present in close relatives within the subfamily Eyprepocnemidinae, they were probably present in the common ancestor of these species and thus could simply have derived from vertical transmission. However, lineage B elements in *E. plorans* are most similar to the *RTE* elements found in several species of the subfamily Oedipodine, a rather phylogenetically distant group. Therefore, there exists the possibility that these elements recently invaded the *E. plorans* genome by means of horizontal transfer from an oedipodine species. Such a conclusion would need the demonstration that lineage B elements are absent from other Eyprepocnemidinae species, but this could only be tested by analysing a very high number of paralogous copies in these genomes, since lineage B-like elements could be minority in them. An intragenomic phylogeographic approach similar to that performed here for *E. plorans* would be useful to detect possible types of elements being minority represented in some chromosomes.

It is also remarkable the suggestion that the B chromosome might have arisen from chromosome 11, for which reason both B and 11 chromosomes carry lineage B elements but not lineage A ones. The lowest pairwise F_{ST} value between *RTE* elements in the B chromosome and those in chromosome 11 (see Table 14) supports the conclusion that the B derived from chromosome 11. This is reinforced by our recent finding that the ITS regions of the ribosomal DNA contained in the B chromosome are more similar to those in chromosome 11 than to those in other A chromosomes (Teruel et al., in preparation).

To conclude, the methodology reported here to obtain DNA sequences from each chromosome, i.e. microdissection and PCR amplification with appropriate primers for a specific DNA sequence (*RTE* elements in *E. plorans*, in this case), can be useful for many other paralogous sequences and organisms. Availability of full genome sequences is nowadays restricted to model or commercially interesting organisms. *E. plorans* is a grasshopper species with a giant genome (three times larger than human genome), which

makes it necessary approaching different strategies to address genomic studies in this species. Individual chromosome microdissection provides a collection of paralogous DNA sequences from each chromosome which may be treated as different population samples to apply population genetics tools. The same as these latter provide valuable information on population structure, and demographic and historical events, in an evolutionary context, the sampling of a collection of copies for a paralogous gene coming from the different chromosomes in a genome can give valuable spatial and historical information about the genome. This provides answers to some interesting questions on the physical arrangement of the analysed DNA sequences within the genome, and may give interesting insights on the past evolutionary history of the genome. For instance, we have shown here the existence of two waves of *RTE* invasion separated in time revealing the conceivable origin of a B chromosome from the smallest autosome. The combination of these two kinds of information (space and time) is analogous to that obtained in phylogeographic analysis, for which reason, our present approach could be named "intragenomic phylogeography".

Acknowledgements

This study was supported by grants from the Spanish Ministerio de Ciencia y Tecnología (CGL2009-11917) and Plan Andaluz de Investigacion (CVI-6649), and was partially performed by FEDER funds.

References

- Adams MD, Celniker SE, Holt RA, et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185-2195.
- Anisimova M & Gascuel O (2006) Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* 55: 539-552.
- Bates GP, Wainwright BJ, Williamson R & Brown SD (1986) Microdissection of and microcloning from the short arm of human chromosome 2. *Mol Cell Biol* 6: 3826-3830.
- Buiting K, Neumann M, Lüdecke HJ, et al. (1990) Microdissection of the Prader-Willi syndrome chromosome region and identification of potential gene sequences. *Genomics* 6: 521-527.
- Camacho JP, Sharbel TF & Beukeboom LW (2000) B-chromosome evolution. *Phil Trans R Soc Lond B* 355: 163-78.

- Camacho JPM (2005) B Chromosomes, in E Inc. (ed), *The Evolution of the Genome*, Gregory, T. R., San Diego, pages. 223-286.
- Cannizzaro LA (1996) Chromosome microdissection: a brief overview. *Cytogenetics Cell Genet* 74: 157-160.
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17: 540-552.
- Cheng YM & Lin BY (2003) Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* 164: 299-310.
- Consortium MGS, Waterston RH, Lindblad-Toh K, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-562.
- Dereeper A, Guignon V, Blanc G, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36 (Web Server issue): W465-W469.
- Excoffier L, Laval G & Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.
- Guan XY, Meltzer PS & Trent JM (1994) Rapid generation of whole chromosome painting probes (WCPs) by chromosome microdissection. *Genomics* 22: 101-107.
- Guindon S & Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696-704.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, in *Nucleic acids symposium series*, pages. 95-98.
- Hood ME, Katawczik M & Giraud T (2005) Repeat-induced point mutation and the population structure of transposable elements in *Microbotryum violaceum*. *Genetics* 170: 1081-1089.
- Houben A, Field BL & Saunders VA (2001) Microdissection and chromosome painting of plant B chromosomes. *Methods in cell science : an official journal of the Society for In Vitro Biology* 23: 115-24.
- Huson DH & Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23: 254-267.

- Initiative AG (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796-815.
- Jung C, Claussen U, Horsthemke B, Fischer F & Herrmann RG (1992) A DNA library from an individual *Beta patellaris* chromosome conferring nematode resistance obtained by microdissection of meiotic metaphase chromosomes. *Plant Mol Biol* 20: 503-11.
- Jurka J, Kapitonov VV, Pavlicek A, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic Genome Res* 110: 462-7.
- Kao FT (1987) Chromosome microdissection and microcloning in human molecular genetics. *Somatic Cell Mol Genet* 13: 375-380.
- Kejnovsky E, Hobza R, Kubat Z, et al. (2007) High intrachromosomal similarity of retrotransposon long terminal repeats: evidence for homogenization by gene conversion on plant sex chromosomes? *Gene* 390: 92-97.
- Lamb JC, Kato A & Birchler JA (2005) Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. *Chromosoma* 113: 337-349.
- Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
- Librado P & Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Liu B, Segal G, Vega JM, Feldman M & Abbo S (1997) Isolation and characterization of chromosome-specific DNA sequences from a chromosome arm genomic library of common wheat. *Plant J* 11: 959-965.
- Mahjoubi F, Peters GB, Malafiej P, et al. (2005) An alphoid marker chromosome inv dup(15)(q26.1qter), detected during prenatal diagnosis and characterized via chromosome microdissection. *Cytogenet Genome Res* 109: 485-490.
- Marchal JA, Acosta MJ, Bullejos M, Guardia RD la de & Sánchez A (2004) A repeat DNA sequence from the Y chromosome in species of the genus *Microtus*. *Chromosome Res* 12: 757-765.

- Miklos GL, Yamamoto MT, Davies J & Pirrotta V (1988) Microcloning reveals a high frequency of repetitive sequences characteristic of chromosome 4 and the beta-heterochromatin of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 85: 2051-2055.
- Muñoz-Pajares AJ, Martínez-Rodríguez L, Teruel M, et al. (2011) A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* 187: 853-63.
- Nei M & Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418-426.
- Scalenghe F, Turco E, Edstrom JE, Pirrotta V & Melli M (1981) Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. *Chromosoma* 82: 205-216.
- Sequencing Consortium *C. elegans* (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282: 2012-2018.
- Subramanian RA, Akala OO, Adejinmi JO & O'Brochta DA (2008) *Topi*, an IS630/Tc1/mariner-type transposable element in the African malaria mosquito, *Anopheles gambiae*. *Gene* 423: 63–71.
- Sánchez-Gracia A, Maside X & Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 21: 200-203.
- Tamura K, Dudley J, Nei Masatoshi & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.
- Teruel M (2009). Origen, expresión y efectos fenotípicos de un parásito genómico. Ph. D. thesis, Universidad de Granada, Spain.
- Teruel M, Cabrero J, Perfectti F, et al. (2009) Microdissection and chromosome painting of X and B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenet Genome Res* 125: 286-291.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2010) B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma* 119: 217–225.

- Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
- Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 7: 256-276.
- Zhou RN & Hu ZM (2007) The development of chromosome microdissection and microcloning technique and its applications in genomic research. *Current Genomics* 8: 67-72.
- Zhou Y, Hu Z, Dang B, et al. (1999) Microdissection and microcloning of rye (*Secale cereale* L.) chromosome 1R. *Chromosoma* 108: 250-5.
- Zimmer R, Haberfeld A & Gibbins AM (1997) Microisolation of the chicken Z chromosome and construction off microclone libraries. *Genome* 40: 865-72.

CAPÍTULO 6

Discusión general

El estudio de los elementos transponibles (TEs) en ortópteros ha sido muy escaso, y en la mayoría de los casos, no ha estado focalizado en las especies de saltamontes. En esta Tesis Doctoral hemos realizado un análisis de estos elementos móviles en hasta 31 especies de saltamontes, centrándonos, para un estudio de mayor profundidad, en las especies *Eyprepocnemis plorans* y *Locusta migratoria* que presentan sistemas de cromosomas B muy estudiados.

La reciente proliferación de proyectos de secuenciación de genomas ha demostrado que los TEs constituyen un componente importante en los genomas de la mayoría de los eucariotas (Kidwell & Lisch 2000; Biemont 2010), representando el 45% en *Homo sapiens* (Lander et al. 2001), el 39% en *Mus musculus* (Waterston et al. 2002) o el 85% en el genoma del maíz B73 (Schnable et al. 2009). Además, los TEs han contribuido, a través de sucesivas oleadas de replicación, al aumento de tamaño de algunos genomas (Kidwell 2002) mostrándose una relación directa entre la proporción de TEs y el tamaño de los genomas, como se aprecia en el hecho de que la cebada (*Hordeum vulgare*) con un tamaño de genoma de 5000 Mb contenga una proporción mayor del 50% de TEs, mientras que *Drosophila melanogaster* con un genoma de tan sólo 180 Mb presente un 15% de TEs. Ningún genoma de ortópteros ha sido aún completamente secuenciado por lo que no se conoce la proporción que los TEs ocupan en estos genomas gigantes (hasta 16.93 pg de ADN en el saltamontes *Podisma pedestris*), pero según la tendencia observada, un genoma como el de *E. plorans* con 10100 Mb debe estar constituido por una elevada cantidad de TEs.

En esta Tesis Doctoral hemos analizado la presencia, localización y abundancia de cuatro tipos de TEs: el elemento LTR *Gypsy* y los elementos no LTR, *RTE* y *R2*, todos ellos pertenecientes a la Clase I, además del elemento *Mariner* perteneciente a la Clase II. Mientras que *Gypsy*, *RTE* y *Mariner* muestran una presencia abundante y una distribución muy extendida en los genomas de *E. plorans* y *L. migratoria*, el elemento *R2* (el único elemento de este estudio con una diana de inserción específica localizada en los genes ribosómicos 28S) mostró una localización preferencial en el cromosoma B y muy escasa en los cromosomas A de *E. plorans*.

La dinámica de la expansión de los TEs a través de los genomas varía dependiendo de las especies y los elementos estudiados. Así, el elemento *Gypsy* se ha observado tanto disperso (Cermak et al, 2008) como en clústeres (Belyayev et al. 2001). De igual forma, las *LINE* han sido encontradas dispersas sobre todos los cromosomas (Rebuzzini et al. 2009) o en regiones restringidas, como las bandas G en mamíferos (Baker & Kass 1994) o los transposones *Mariner* que pueden concentrarse en regiones específicas, como los

centrómeros o las regiones distales (Mukabayire & Besansky 1996; Jacobs et al. 2004) o esparcidos por todos los cromosomas, incluyendo las regiones heterocromáticas (Torti et al. 2000). Estos tres elementos, analizados en *E. plorans* y *L. migratoria*, mostraron una distribución dispersa sobre todos los cromosomas pero no aleatoria, puesto que la señal de FISH fue observada sobre las regiones eucromáticas pero ausentes de la heterocromatina, incluida las NORs (en *E. plorans*) y los clústeres de genes de histonas (en *L. migratoria*). Los cromosomas B en ambas especies mostraron señal de hibridación pero mucho menos extensa, especialmente en *E. plorans*, que podría explicarse por su naturaleza eminentemente heterocromática. Aun así, no podemos descartar la presencia de algunas copias en las regiones heterocromáticas que escapen a la resolución de un experimento de FISH.

Las señales de FISH sugieren una abundancia similar en la eucromatina de todos los cromosomas A, lo que no ocurre en otras especies donde se ha encontrado una acumulación de TEs en los cromosomas sexuales (Marchal et al. 2006; Cermak et al. 2008). En *Drosophila* (Pardue 1995) y en *Bombyx mori* (Fujiwara et al. 2005) los extremos cromosómicos están asociados con elementos no LTR. Del mismo modo, la familia de retrotransposones centroméricos (CR) está sobrerrepresentada en el centrómero de muchas especies de cereal (Gaio et al. 2009). No obstante, la ubicua distribución de los TE en la eucromatina, observada en estas especies de saltamontes, también ha sido encontrada en otras especies (Acosta et al. 2008; Rebuzzini et al. 2009), a pesar de los posibles efectos deletéreos que pueden producir estos TEs al afectar a la expresión génica del hospedador o al causar alteraciones tanto en genes como en cromosomas por mutagénesis insercional (Kidwell & Lisch 2001, Feschotte & Pritham 2007). Sin embargo, no podemos descartar que la eucromatina de estos genomas gigantes de saltamontes contenga un mayor número de sitios tolerables de inserción de TEs. De hecho, el genoma de *L. migratoria* y de *E. plorans* supone de dos a tres veces el genoma humano, y los TEs podrían encontrar en ellos numerosos espaciadores intergénicos y pseudogenes sobre los que insertarse a lo largo de la eucromatina. Mientras que los elementos antiguos tienden a acumularse en regiones eucromáticas asociados con genes, se ha observado que los elementos más recientes colonizan la heterocromatina, evitando así el efecto de mutagénesis insertacional (Santiago et al. 2002). En base a esto, la presencia de TEs en la eucromatina de *E. plorans* y *L. migratoria* podría ser un evento antiguo, presumiblemente precedente a la diversificación de los saltamontes. Por otro lado, esta situación también es coherente con un posible papel beneficioso de los TEs en el genoma hospedador, al

formar parte de genes o regiones reguladoras, como ocurre con los *MITEs* que forman parte de las regiones reguladoras como sitio de unión al ADN (Bureau & Wessler 1994) o la producción de novedades génicas derivadas de TEs como *Mariner*, *hAT* o el elemento *P* (véase para revisión Feschotte & Pritham 2007).

La abundancia de TEs en la eucromatina de *E. plorans* y *L. migratoria* contrasta con su escasez en la heterocromatina, lo que sugiere que estos tipos de elementos móviles no han jugado un papel importante en la generación de la heterocromatina en ambas especies, en contraste con lo observado en otros casos (Hsieh & Fire 2000). Una posible explicación es que la dispersión de los TEs en estas especies de saltamontes fue anterior a la expansión de la heterocromatina. En *E. plorans*, la heterocromatina pericentromérica está principalmente formada por dos tipos de ADN repetitivo: un ADNsat de 180 bp que está ausente en *E. unicolor* y otras especies de Eypreopcnemidinae (Cabrero et al. 2003a) y ADNr que está restringido a los cromosomas S₉ y S₁₁ en muestras de Dhagestan (Cáucaso) pero que se ha ampliado a las regiones de heterocromatina pericentromérica de los cromosomas de la mayoría de las poblaciones españolas (Cabrero et al. 2003b). Por lo tanto, es probable que la dispersión de TEs en el genoma de *E. plorans* precediera a la expansión de estos dos ADN repetidos, razón por la cual las regiones de heterocromatina pericentromérica se encuentran en su mayoría sin estos TEs (véase Figure 4). Otra posible explicación podría ser la ausencia de dianas adecuadas para la integración de TE en estas secuencias de ADN repetitivo. En *L. migratoria*, la naturaleza molecular de la heterocromatina pericentromérica se desconoce hasta el momento, aunque la presencia de secuencias de ADN pertenecientes a la fracción Cot1 sugiere que estos dominios cromosómicos también se componen de secuencias de ADN repetitivo (León-López et al. sin publicar).

El elemento R2, como comentábamos anteriormente, presenta una diana de inserción dentro de los genes ribosómicos 28S y en *E. plorans* hemos realizado la caracterización completa de este elemento. El elemento EploR2 ha mostrado una estructura similar a la presente en los elementos R2 descritos hasta el momento, con un ORF que codifica tres dominios (unión al ADN, retrotranscriptasa y endonucleasa), que incluye un único motivo CCHH, por lo que se encuentra en el clado R2-D (Kojima et al. 2005), junto con otros elementos R2 de artrópodos, como los de *D. melanogaster* (Eickbush et al. 1997) o *Triops cancriformis* (Mingazzini et al. 2011) que también son similares a EploR2 en longitud (unos 3,5 kb). Este elemento comparte la diana de inserción en los genes 28S ribosómicos con todos los elementos de R2 y también, hemos encontrado intactas las regiones adyacentes en la secuencia de los genes 28S, sin las deleciones que presentan

en otras especies (Burke et al. 1999; Kojima et al. 2005; Rezende-Terexeida et al. 2009; Ghesini et al. 2011; Mingazzini et al. 2011). Sin embargo, a diferencia de otras especies donde R2 ha sido previamente estudiado (Pérez-González & Eickbush 2001), *E. plorans* muestra ADNr en la mayoría de los cromosomas, siendo el mayor cluster el presente en los cromosomas B, seguido de los correspondientes a las NORs primarias de los cromosomas S₉, S₁₀, S₁₁ y X. Nuestros resultados de FISH muestran una gran cantidad de elementos R2 en el ADNr de los cromosomas B y, en menor medida, en el ADNr de los autosomas S₁₀ y S₁₁, sin mostrarse señal alguna en los otros cromosomas que portan ADNr como los cromosomas X y S₉.

Como comentamos en la introducción general, más de 50 variantes diferentes de cromosomas B han sido descritas en las poblaciones españolas de *E. plorans*, surgidas presumiblemente a través de reorganizaciones cromosómicas (inversiones, deleciones y translocaciones) donde los puntos de rotura coinciden con la región distal (López-León et al. 1993), que se corresponde con el ADNr de los cromosomas B de *E. plorans* (López-León et al. 1994; Cabrero et al. 1999). En las poblaciones de Marruecos, sin embargo, sólo una variante de B apareció derivada de la eliminación del ADNr (Bakkali & Camacho 2004). Por lo tanto, al menos una parte de las mutaciones cromosómicas del B de *E. plorans*, aquellas que implican los puntos de rotura en ADNr, podrían ser causadas por el elemento móvil R2. Aunque no hay evidencias directas de la inducción de reordenaciones cromosómicas por elementos R2, sí existen evidencias para otros transposones como el elemento *FoldBack* en *D. buzzatti* (Cáceres et al. 1999). Por otro lado, R2 sí puede producir grandes deleciones en el locus ADNr por recombinación de dos elementos en el mismo cromosoma (Zhang et al. 2008). Sin embargo, la mutación más frecuente en los cromosomas B de *E. plorans*, en las poblaciones españolas y marroquíes (López-León et al. 1993; Bakkali & Camacho 2004) es la misdivisión centromérica, por lo que es probable que otros elementos móviles distintos de R2 puedan provocar reordenaciones cromosómicas en los cromosomas B de esta especie.

Una vez conocida su localización cromosómica, nos propusimos conocer la abundancia de estos TEs, por lo que realizamos una cuantificación del número de copias mediante experimentos de PCR cuantitativa. Estos experimentos en *E. plorans* revelan que el elemento *Gypsy* es el más frecuente seguido por *RTE* y *Mariner*. Un análisis similar en *L. migratoria* indica que *Mariner* es más frecuente en esta especie y que la estimación del número de copias es mucho menor que en *E. plorans*, especialmente para *Gypsy* y *RTE*, siendo una consecuencia lógica del uso de cebadores que han sido diseñados en base a la secuencia de los elementos encontrados en *E. plorans*, por lo que es destacable encontrar

un número similar de copias para *Mariner* en las dos especies. Estas diferencias entre la estima del número de copias entre ambas especie puede deberse a diferentes grados de conservación para las secuencias de ADN de los elementos de clase I y II. Los elementos de clase I (*Gypsy* y *RTE*) son altamente variables puesto que su amplificación se hace mediante un paso de transcripción reversa y con una tasa de mutación similar a la de virus (Drake et al. 1998). La aparición de nuevas variantes favorecería la divergencia de las familias a través de la especie (Volff et al. 2000). Esto implica que el número de TEs detectado por PCR cuantitativa es un subconjunto de los detectados por FISH, puesto que esta técnica también detecta aquellas variantes que muestra homología parcial. En *L. migratoria*, este subconjunto es particularmente pequeño ya que además está restringido a aquellos elementos que conservan la región de anclaje de los cebadores similar a la de *E. plorans*. En los elementos de clase II (ej. *Mariner*), sin embargo, la dispersión mediante *corte/escisión*, así como los posibles eventos de transferencia horizontal contribuyen a la conservación de secuencia entre especies (Hartl et al. 1997). Por lo tanto, la mayor conservación esperada para *Mariner* explica el similar número de copias estimado en *L. migratoria* y *E. plorans*.

El retroelemento EploR2 se mostró menos abundante que *Gypsy* o *RTE*, pero si comparamos el número de copias de R2 de *E. plorans* con el de otros insectos como *Drosophila*, éste es notablemente superior (Zhou & Eickbush 2009). Puesto que R2 se localizaba preferentemente en el cromosoma B, para este elemento también realizamos una estima del número de copias de R2 entre individuos portadores y no portadores de cromosomas B, así como entre poblaciones con las variantes B₁, B₂ y B₂₄. Los resultados obtenidos indican que los cromosomas B presentan diez veces más copias de R2 que los cromosomas A, como se deduce de las comparaciones entre individuos portadores y no portadores de cromosoma B. De este modo, la débil o quizás inadvertida señal de hibridación de R2 en los clústeres de ADN_r de los cromosomas A, fue por lo tanto debida al escaso 10% de copias de R2 presentes en ellos, ya que el 90% de las copias de este elemento estaban presentes en los cromosomas B (véase Table 7). Una estimación del número de copias en cada variante de cromosoma B puede deducirse de la diferencia entre los individuos portadores y no portadores de este cromosoma en una misma población. Las poblaciones con distintas variantes de cromosoma B mostraron diferencias en el número de copias de R2, sugiriendo que el cromosoma B es muy importante a la hora de explicar la dinámica de la población de elementos EploR2. Las variantes de B, tal como mencionamos en la introducción, difieren en el tamaño y en la proporción de ADN_{sat} y ADN_r, siendo el cromosoma B₁ la variante con la mayor cantidad de ADN_r

(como es deducido del análisis de FISH), mientras que las variantes B_2 y B_{24} son las que portan aparentemente clústeres de ADN_r menores (Cabrero et al. 2003). Esta diferente proporción ADN_r está en consonancia con el número de copias de R2 encontrados en los individuos portadores de diferentes variantes del cromosoma B (Table 7).

Junto con la estima del número de copias de R2 también se realizó la cuantificación de unidades de ADN_r para conocer la ocupación del locus ribosómico por este elemento. El número de unidades de ADN_r estimada en individuos no portadores de B (aprox. 11000) es mayor que el presente en otros ortópteros como *Ch. parallelus* (aprox. 1000) (Parkin & Butlin 2004) y *L. migratoria* (aprox. 3300) (Schafer & Kunz 1987). Sin embargo, el hecho de que *Ch. parallelus* presente clústeres de ADN_r, estudiados mediante FISH, en sólo dos (o tres) pares de cromosomas, y *L. migratoria* los presente en tres pares cromosómicos mientras que *E. plorans* muestre ADN_r en la mayoría de los cromosomas A (Cabrero & Camacho 2008) podría explicar estas diferencias. Además, el ADN_r representa aproximadamente el 1.3% del genoma de *E. plorans*, que no es muy diferente a lo observado en *L. migratoria* (1%) o *Apis mellifera* (0.9%) (Schafer & Kunz 1987). Sin embargo, las estimas de ADN_r de las diferentes variantes de cromosomas B no concuerdan completamente con la relación esperada según los datos de FISH, puesto que el cromosoma B_2 parece contener más unidades de ADN_r que el cromosoma B_1 . Esto podría sugerir una organización molecular diferente de las unidades de ADN_r en distintas variantes de cromosoma B. Así, estas diferencias se deberían, por una parte, a la proporción de unidades de ADN_r degeneradas que pueden ser más difíciles de detectar por estimación de número de copias mediante PCR y, por otro lado, a las diferencias en la cantidad de secuencias dispersas entre los cistrones de ADN ribosómico, presentando el cromosoma B_2 una mayor densidad de genes ADN_r en su tercio distal, que es donde la FISH revela la presencia de ADN_r.

La proporción de unidades ADN_r con inserciones de R2 es del 2%. Este valor es menor que el encontrado en *D. melanogaster* (30%), *B. mori* (10-30%) o *L. migratoria* (20%) (Jakubczak et al. 1991, 1992). En otras especies, por ejemplo, *T. cancriformis*, sólo el 0,5-5% de las unidades de ADN_r están ocupadas por elementos R2 (Mingazinni et al. 2011), una cifra que es muy similar a la de *E. plorans*. La selección contra unidades de ADN_r insertadas con R2 ha sido la explicación más utilizada para explicar las bajas tasas de ocupación. En *E. plorans*, las unidades de ADN_r de individuos 0B muestran tres veces menos ocupación por R2 que los individuos portadores de B. Esto es consistente con un efecto negativo de los elementos R2 en el ADN_r de los cromosomas A, y con la prescindibilidad de los cromosomas B, que les convierte en excelentes destinos para la

inserción de R2 puesto que portan numerosas unidades de ADNr para colonizar. Pero el ADNr en los cromosomas B de esta especie es usualmente inactivo debido a la metilación de su secuencia de ADN y la hipoacetilación de las histonas (López-León et al. 1991; Cabrero et al. 2007) que evitaría la transcripción de R2. Esto sugiere la interesante posibilidad de que los cromosomas B actúen como sumideros de no retorno para los elementos R2 procedentes de los cromosomas A, explicando así el porqué son tan escasos en el genoma estándar de esta especie. Curiosamente, la variante B₁, que es considerada la variante más antigua en la Península Ibérica (Henriques-Gil & Arana 1990), mostró la mayor proporción de clústeres de ADNr ocupados por R2, mientras que la variante B₂₄, que surgió recientemente de la variante de B₂ en la población de Torrox (Henriques-Gil & Arana 1990; Zurita et al. 1998), demostró una ocupación similar a la de su variante antecesora, presumiblemente por el escaso tiempo disponible para la acumulación de R2.

La presencia de algunos elementos en las regiones eucromáticas potencialmente activas sugiere la funcionalidad de estos TEs. De hecho, hemos identificado transcritos específicos de *Gypsy*, *RTE* y *Mariner* en el ARN de *E. plorans* y *L. migratoria*, correspondiendo la banda más notoria en el gel al transcrito del elemento *Gypsy*, en consonancia con su mayor número de copias. Vicent (2010) encontró en el maíz que *Gypsy* exhibía la mayor actividad transcripcional de entre las 56 familias de TEs analizados. La tasa de transcripción probablemente depende de múltiples factores tales como la eficiencia en la transposición, el tiempo transcurrido desde la invasión del genoma o la existencia de domesticación. La actividad transcripcional ha sido descrita para muchos TEs en varias especies. Lamb et al. (2007) reportó actividad transcripcional de un elemento derivado de retrotransposones (*StarkB*) específico del cromosoma B de maíz, así como Muñoz-López et al. (2008) describió la funcionalidad de la transposasa del elemento Mbourmar en la hormiga *Messor bouvieri*. A pesar de haber encontrado transcritos de *Gypsy*, *RTE* y *Mariner* en el ARN de *E. plorans* y *L. migratoria*, esto no garantiza su actividad transposicional debido a podrían ser silenciados, por ejemplo, a nivel postranscripcional a través de ARN de interferencia (Kasschau et al. 2007).

Para el elemento R2 de *E. plorans* no sólo testamos su transcripción sino que también estudiamos su actividad mediante el análisis de elementos truncados en el extremo 5'. Según sugieren Zhou & Eickbush (2009), una alta tasa de ocupación por R2 no necesariamente implica un alto nivel de expresión. Estos autores observaron en *D. simulans* que en líneas con un número similar de inserciones de R2, aquellas que mostraban una alta expresión de R2 presentaban las unidades de ADNr ocupadas por R2 diseminadas en todo el locus de ADNr, mientras que las líneas con baja o nula expresión mostraban grandes regiones

de ADNr libre de inserciones de R2, siendo estas regiones transcritas preferentemente. Estos autores sugieren que es la localización de R2 dentro del locus ribosómico lo que determina su expresión. La búsqueda de transcritos de R2 no tuvo éxito en *E. plorans*, pero sí hemos obtenido evidencia de hechos pasados de retrotransposición mediante el perfil de truncamiento del extremo 5'. La actividad de R2 no es concebible sin su transcripción (Zhou & Eickbush 2009) por lo que dos explicaciones son posibles a la situación encontrada en *E. plorans*. En primer lugar, que los eventos de transcripción podrían ocurrir en un tejido, tiempo de desarrollo o población diferente a los analizados aquí, es decir, los machos de la población de Torrox. La elección de esta población se realizó porque el ADNr en el cromosoma B₂₄ está activo en muchos individuos de Torrox (Teruel et al. 2009) y así era más probable que los elementos R2 también fueran transcripcionalmente activos. Sin embargo, en *D. simulans*, se ha demostrado que las primeras etapas del desarrollo embrionario y la línea germinal femenina exhiben la mayor actividad transcripcional para elementos R2 (Eickbush & Eickbush 2011). Una situación posible es que R2 fuera más activo en el pasado que en la actualidad, como sugiere la baja proporción encontrada de nuevos elementos truncados en Torrox y otras poblaciones. De hecho, los cromosomas B causan estrés a las células del hospedador, y estos últimos pueden responder aumentando el número de quiasmas (Camacho et al. 2002). Durante la etapa parasítica del cromosoma B₂₄, los machos de Torrox portadores mostraron una mayor frecuencia de quiasmas que después de su neutralización (Teruel et al. 2009). El mayor número de quiasmas puede aumentar la probabilidad de entrecruzamientos desiguales, movilizandando así elementos R2 hacia regiones de ADNr libre de R2 lo que explicaría un aumento de la actividad de R2 y el perfil de truncaciones en 5' de Torrox. Pero una consecuencia del aumento del movimiento y copias de R2 es la generación de grandes deleciones en el locus de ADNr por entrecruzamiento desigual entre copias de R2, disminuyendo así su número, lo que contribuiría a explicar el bajo número de copias de ADNr estimada en la población de Torrox.

Esto es consistente con la idea de que los elementos móviles son activados en condiciones de estrés (Kidwell & Lisch 1997; Capy et al. 2000) y con una mayor frecuencia de retrotransposición en individuos portadores de B, como se deduce de los perfiles de truncaciones 5'. En *D. simulans*, la alta variación en el perfil de truncaciones 5' en una población sugiere actividad del elemento R2 (Pérez-González & Eickbush 2001, Zhang & Eickbush 2005). En *E. plorans*, las poblaciones carentes de cromosomas B han mostrado baja frecuencia de truncaciones de R2 o ausencia total de nuevas truncaciones (véase

Figure 10) mientras que aquéllas que muestran alta frecuencia de B (Salobreña y Torrox) mostraban el mayor número de nuevas truncaciones, la mayoría de ellas probablemente producidas en el ADNr del cromosoma B, como sugiere las truncaciones observadas en el ADNr del cromosoma B₂₄ microdisecionado.

Los elementos R2 han mostrado en *E. plorans* acumulación en los loci de ADNr de los cromosomas B mientras que están insuficientemente representadas en los loci ADNr de los cromosomas A. Este enriquecimiento en elementos transponibles parece ser común en los cromosomas B, tal y como se ha comentado con anterioridad, y puede explicarse por la ausencia de recombinación entre los B y los cromosomas A que impide la eliminación de estas secuencias de ADN. Los cromosomas B parecen ser un miembro joven del genoma de *E. plorans* (Muñoz-Pajares et al. 2011) que en su camino evolutivo han establecido una relación especial con el retroelemento R2 al servirles de refugio. Sin embargo, el hecho de que el ADNr en los cromosomas B esté frecuentemente inactivo (López León et al. 1995) es un claro obstáculo para la expansión de R2, ya que su actividad depende de la expresión de ADNr de forma crucial (George & Eickbush 1999).

Una vez conocida la localización, abundancia y posible actividad de los TEs estudiados en los saltamontes *E. plorans* y *L. migratoria*, decidimos estudiar la historia evolutiva de dos de los elementos, *Gypsy* y *RTE*, dentro de la familia Acridoidea. Para esto realizamos una reconstrucción filogenética con las secuencias de estos elementos encontradas en 31 especies de este grupo y las comparamos con la filogenia producida por el ADN mitocondrial de estas mismas especies. Estas filogenias sugieren que los TEs parecen ser antiguos componentes de los genomas de saltamontes, ya que se han encontrado en todos los grupos analizados (al menos el elemento *RTE*) y su topología se asemeja a la taxonomía de saltamontes. Sin embargo, es necesario tener en cuenta que este análisis se basa en la amplificación por PCR y esto crea un filtro artificial que sólo nos permite analizar aquellos elementos TEs que han conservado la secuencia de ADN complementaria con los cebadores empleados, mientras que se escapan del análisis muchos elementos que muestren mutaciones en esta región.

La filogenia del ADN mitocondrial de las 31 especies de acrididos analizada es consistente con la taxonomía actual presente en la base de datos *Orthoptera Species File* y también con los árboles moleculares construidos anteriormente para estos organismos (Contreras & Chapco 2006; Bugrov et al. 2006), lo que sugiere que este árbol probablemente representa el árbol real de la especie. El árbol proporciona una buena resolución a nivel de tribu y de género, pero ésta disminuye a nivel de subfamilia. Es notable la posición basal de las tres especies de *Dociostaurus*, dentro de la subfamilia Gomphocerinae, en

coincidencia con su número ancestral de cromosomas $2n^{\sigma}=23$, comparado con el de los clados derivados A y B, incluyendo las tribus Stenobothrini y Gomphocerini que muestran $2n^{\sigma}=17$ cromosomas como resultado de la fijación de tres fusiones céntricas (Hewitt 1979). La única incoherencia con la taxonomía fue la inclusión de una especie del género Stenobothrus (*S. festivus*) dentro de la tribu Gomphocerini (clado B) en lugar de dentro de la Stenobothrini (clado A), sugiriendo que la taxonomía de los géneros Stenobothrus, Omocestus y Chorthippus probablemente necesita alguna revisión.

Las filogenias moleculares de los dos TEs analizados se asemejan al del ADN mitocondrial, con altos valores de congruencia proporcionados por el test CADM: 72% entre *RTE* y el ADN mitocondrial, y 67% entre *Gypsy* y el ADN mitocondrial. Sin embargo, también hay notables incongruencias, como se muestra en el test de SH. La filogenia de *RTE* muestra una alta consistencia con la filogenia ADN mitocondrial, con la subfamilia Oedipodinae como el grupo más basal y la subfamilia Gomphocerinae como la más reciente. Esta coincidencia con la filogenia de ADN mitocondrial no es inesperada, ya que se sabe que el elemento de *RTE* es uno de los más extendidos y más antiguos TEs en eucariotas (Malik & Eickbush 1998, Malik et al. 1999).

Una característica interesante de los elementos *TE* es la alta conservación de su secuencia de aminoácidos, incluso entre grupos lejanamente relacionados, en comparación con otras secuencias del genoma hospedador (Zupunski et al. 2001). Esto puede explicar el alto grado de politomías que se encuentra en la filogenia de los *RTE* de saltamontes que se muestra aquí, que es más notable en la subfamilia Gomphocerinae con $2n=17$, que parecen ser el clado más joven en la filogenia de ADN mitocondrial. Sorprendentemente, uno de los pocos clados bien resueltos, con un soporte de rama de 0,98, está dentro de esta subfamilia e incluye a *Ch. vagans* ($2n=17$) con las tres especies de *Dociostaurus* ($2n=23$) y que podría ser indicio de una posible transferencia horizontal de transposón (HTT) entre estas especies.

La filogenia de *Gypsy* también muestra politomías dentro de las subfamilias Gomphocerinae y Oedipodinae, lo que sugiere un escaso grado de diferenciación de este elemento dentro de estos grupos. Pero la característica más notable en esta filogenia es la agrupación de las especies *E. plorans* y *H. littoralis* en un clado común con un 0,95 de soporte de rama, sugiriendo también una posible HTT entre estas especies. La HTT ha sido bien documentada para el elemento *Gypsy* en drosofilidos (Herédia et al. 2004) siendo la inclusión del elemento en una cápsida vírica el mecanismo propuesto para esta HTT (Friesen & Nissen 1990; Piskurek et al. 2007). Además, si el TE por sí mismo pudiera adquirir la capacidad de generar una cápsida, podría ser capaz de realizar la HTT.

En la literatura se encuentran referencias a la naturaleza viral de *Gypsy* y, de hecho, este TE fue considerado el primer virus descrito en invertebrados (Kim et al. 1994; Pelisson et al. 1997), aunque un análisis posterior sugirió que el elemento *Gypsy* se convirtió en virus al adquirir los genes *env* desde un baculovirus (Malik et al. 2000). Por lo tanto, el carácter cuasi viral de *Gypsy* admite la posible HTT entre *E. plorans* y *H. littoralis* sugerida por la filogenia obtenida en este estudio. Esto es concebible porque estas dos especies normalmente comparten los mismos hábitats en las zonas costeras y están estrechamente relacionadas, siendo encontradas frecuentemente en los mismos lugares (Barranco & Pascual 1993; Olmo-Vidal 2002). Asimismo, *Ch. vagans* frecuentemente se encuentra en el mismo hábitat y localidades que una o más especies de *Dociostaurus* (véase Badih et al. 2003) lo que podría haber proporcionado una amplia oportunidad para infección común por ácaros o retrovirus, actuando como vehículos para la HTT (Schaack et al. 2010).

Sin embargo, otros procesos, como la deriva genética o una diferente dinámica evolutiva de los TEs entre especies que depende del nivel de actividad del transposón (véase para revisión Capy et al. 1994) también podría explicar las contradicciones entre las filogenias del hospedador y el TE. De hecho, el análisis de 84 copias parálogas de *Gypsy* dentro de un mismo individuo de *E. plorans*, mostró la presencia en este genoma de algunas copias de *Gypsy* similares al elemento *Gypsy* en *H. littoralis*, pero hubo otras copias similares al elemento de *E. unicolor*. Esto arroja serias dudas sobre la HTT como una posible explicación de la incongruencia filogenética observada y sugiere la existencia un “lineage sorting” incompleto a partir de una gran variedad de elementos presentes en el genoma de antepasados comunes de estos dos géneros.

Dada la naturaleza paróloga de las múltiples copias intragenómicas de los TEs, la probabilidad de observar incongruencias filogenéticas con la filogenia de hospedador es alta. De hecho, la alta congruencia entre filogenias de TEs y ADN mitocondrial observada aquí es sorprendente teniendo también en cuenta que, en todas las especies, excepto *E. plorans*, hemos empleado una única secuencia de TE. De todos modos, las incongruencias filogenéticas no suponen que haya ocurrido una HTT, ya que muchas de ellas podrían ser simplemente el producto de “lineage sorting” incompleto (Schaack et al. 2010) o de un muestreo incompleto. Por lo tanto, habría que determinar si esta advertencia se aplica a otros casos de supuesta HTT, incluyendo el de *Ch. vagans*-*Dociostaurus* para el elemento *RTE* sugerido por nuestros resultados, siendo necesaria una mayor investigación analizando la secuencia de un mayor número de copias de los TEs en cada especie.

Así, procedimos a un análisis en profundidad de las copias de *RTE* presentes en *E. plorans* mediante la microdissección de cada bivalente de una célula meiótica, incluyendo también el cromosoma X y el B. Este análisis de parte del gen de la retrotranscriptasa del elemento *RTE* presente en los diferentes cromosomas ha proporcionado una importante información sobre su estructura de intragenómica. La mayoría de las secuencias de ADN analizadas carecía de codones de parada o indels, y se les ha considerado como no defectivos. Sin embargo, dado que la secuencia analizada puede representar menos de 10% de la secuencia de ADN total del elemento completo (*RTE-1* tiene una longitud de 4 kb en *Danio rerio* y de 2,8 kb en *N. vitripennis*, Jurka et al. 2005) es probable que la proporción de elementos *RTE* defectivos en *E. plorans* sea mucho mayor, puesto que muchos de éstos no defectivos pueden llevar mutaciones que supongan una pérdida de la función en otras partes del elemento no analizadas. En cualquier caso, más del 25% de los elementos *RTE* en el genoma de *E. plorans* son defectivos formando parte de las supuestamente enormes cantidades de “ADN basura” presentes en este gigantesco genoma.

Con la anterior advertencia en mente, la proporción de elementos defectivos y no defectivos fue similar en los diferentes cromosomas, incluido el cromosoma B. Esto contradice la idea de que los cromosomas B son paraísos para los elementos móviles (Camacho, Sharbel & Beukeboom 2000), ya que deberían entonces mostrar una mayor proporción de elementos no defectivos (supuestamente activos). Dos hechos pueden explicar por qué los cromosomas B se comportan como los cromosomas A en esta especie. En primer lugar, los cromosomas B de *E. plorans* parecen ser bastante jóvenes (Muñoz-Pajares et al. 2011) y por lo tanto no han tenido mucho tiempo para acumular elementos móviles. En segundo lugar, los elementos *RTE* se encuentran preferentemente insertados en las regiones eucromáticas de *E. plorans* como hemos discutido con anterioridad mientras que los cromosomas B en esta especie son principalmente heterocromáticos. Estos dos hechos también explican la menor diversidad nucleotídica encontrada entre los elementos *RTE* presentes en los cromosomas S_{11} y B, que también puede ser indicativo de un menor tamaño de población de este elemento en estos dos cromosomas.

El grado de diversidad nucleotídica encontrada para los elementos *RTE* (0.05-0.16, véase Table 11) es mayor que la observada para otras secuencias de ADN en *E. plorans*, tales como, por ejemplo, la región ITS-1 del ADN_r ($\pi = 0.014$) (Teruel 2009) un hecho esperable para elementos cuya transposición se realiza mediante una molécula de ARN.

Estos valores de diversidad son un orden de magnitud superiores a los encontrados para copias activas de elementos LTR y no LTR en *D. melanogaster* (Sánchez-Gracia et al. 2005), pero son similares a los encontrados para los elementos *TOPI* en *Anopheles gambiae* ($\pi = 0.051$) (Subramanian et al. 2008).

Aunque la diversidad de familias de retrotransposones y transposones ha sido ampliamente analizada (Bartolomé et al. 2002, Roy et al. 2000; Jin-Shan et al. 2005; Gomulski et al. 2004; Marin & Llorens 2000, Hollister & Gaut 2007), sólo unos pocos estudios han relacionado la diversidad con la ubicación en los cromosomas. Un buen ejemplo de este tipo de trabajos se encuentran en Hood et al. (2005), donde analizan la estructura genética de las familias *Copia* y *Helitron* en cromosomas separados electroforéticamente de *Microbotryum violaceum*. Asimismo, Kejnovsky et al. (2007) analizaron la diversidad de *Retand*, un retrotransposón como *Gypsy*, en cromosomas sexuales y autosomas microdisecionados de *Silene latifolia*. Nuestro actual estudio, sin embargo, es el primero que intenta determinar la distribución de intragenómica de los retroelementos *RTE*, incluyendo todos los cromosomas A y también el cromosoma B. Esto ha proporcionado información valiosa sobre la historia evolutiva de los elementos *RTE* en el genoma de *E. plorans* sugiriendo, además, un posible origen del cromosoma B en esta especie.

Nuestro análisis filogenético intragenómico revela la existencia de dos tipos de retroelementos *RTE* en el genoma de *E. plorans*, pero sólo uno de ellos (linaje B) está presente en los cromosomas S_{11} y B, mientras que los restantes cromosomas portan ambos tipos. Esto explica la menor diversidad de nucleótidos entre los *RTE* en estos dos cromosomas. La existencia de dos linajes diferentes de *RTE* en el genoma de *E. plorans* apunta hacia dos posibles invasiones por parte de los elementos *RTE* en el pasado histórico de este genoma, siendo ambos eventos anteriores al origen del cromosoma B. Una de las invasiones (la del linaje A) alcanzó a todos los cromosomas A excepto al cromosoma S_{11} , mientras que la otra (la del linaje B) alcanzó a todos los cromosomas. Dado que los elementos *RTE* del linaje A también están presentes en parientes cercanos dentro de la subfamilia Eyprepocnemidinae, estaban probablemente presentes en el ancestro común de estas especies y así simplemente pudieron haber derivado por transmisión vertical. Sin embargo, los elementos de linaje B en *E. plorans* son más similares a los elementos *RTE* de varias especies de la subfamilia Oedipodinae, un grupo filogenéticamente distante. Por lo tanto, existe la posibilidad de que estos elementos invadieran recientemente el genoma de *E. plorans* por medio de una transferencia horizontal desde una especie de oedipodinos. Para probar dicha conclusión sería necesaria la demostración de que los elementos del linaje

B están ausentes en las otras especies de la subfamilia Eyprepocnemidinae analizando un número muy alto de copias parálogas de estos genomas, puesto que los elementos del linaje B podrían ser minoritarios en ellas. Una aproximación similar a la filogeografía intragenómica realizada aquí en *E. plorans*, sería útil para detectar estos posibles tipos minoritarios de elementos transponibles, presentes sólo en algunos cromosomas.

También es destacable la suposición de que el cromosoma B podría haber surgido del cromosoma S_{11} , razón por la cual ambos B y S_{11} portan elementos del linaje B pero no del linaje A. Los menores valores F_{ST} se encontraron entre elementos *RTE* del cromosoma B y el cromosoma S_{11} (Table 14) apoyando la conclusión de que el B deriva del cromosoma S_{11} . A su vez, esto se ve reforzado por nuestro hallazgo reciente de que las regiones ITS del ADN ribosómico contenidas en el cromosoma B son más similares a los del cromosoma S_{11} que a los de otros cromosomas A (Teruel et al. en preparación).

Para concluir, la metodología aquí expuesta para obtener las secuencias de ADN de cada cromosoma, es decir, la microdissección y la amplificación por PCR con cebadores apropiados para una secuencia específica de ADN (elementos *RTE* en *E. plorans*, en este caso), puede ser útil para muchas otras secuencias parálogas y organismos. La disponibilidad de secuencias del genoma completo en la actualidad está restringida a organismos modelo o comercialmente interesantes. *E. plorans* es una especie de saltamontes con un genoma gigante (tres veces más grande que el genoma humano), lo que hace necesario abordar diferentes estrategias para estudios genómicos en esta especie. La microdissección individual de cada cromosoma proporciona una colección de secuencias parálogas de cada cromosoma que pueden ser tratadas como muestras de diferentes poblaciones al aplicar herramientas de genética de poblaciones. Así también puede proporcionar una valiosa información sobre la estructura poblacional y los eventos históricos y demográficos, en un contexto evolutivo. El muestreo de una colección de copias parálogas de un gen procedentes de diferentes cromosomas puede dar una importante información histórica y espacial sobre el genoma. Esto proporciona respuestas a algunas preguntas interesantes sobre la disposición física de las secuencias de ADN analizadas dentro del genoma y puede dar ideas interesantes sobre la historia evolutiva del mismo. Por ejemplo, hemos demostrado aquí la existencia de dos oleadas de invasión de *RTE* separadas en el tiempo, revelando como concebible el origen del cromosoma B desde el más pequeño de los autosomas. La combinación de estos dos tipos de información (espacio y tiempo) es similar a la obtenida en análisis filogeográficos, razón por la cual, nuestro enfoque actual podría llamarse filogeografía intragenómica.

En esta Tesis Doctoral hemos comenzado a responder al interrogante que planteaban los TEs en los genomas de ortópteros, demostrando que estos elementos son antiguos compañeros de viajes de los saltamontes con una colonización amplia de los genomas y un origen que puede situarse con el del grupo. Esto sugiere que son anteriores a la aparición de los cromosomas B, al menos en *E. plorans*, que han demostrado influir en la dinámica del retrotransposón R2. La relación de estos dos parásitos genómicos, TEs y cromosomas B, dentro de unos genomas tan masivos como los de saltamontes ha quedado patente en esta Tesis Doctoral, pero la diversidad de TEs que todavía debe subyacer en estos genomas pueden guardarnos muchas más sorpresas.

Bibliografía

- Acosta MJ, Marchal JA, Fernández-Espartero CH, Bullejos M & Sánchez A (2008). Retroelements (LINEs and SINEs) in vole genomes: differential distribution in the constitutive heterochromatin. *Chromosome Res* 16: 949-959.
- Badih A, Fortis PJ, Pascual F (2003). Diversidad faunística y ecología de los ortópteros de la comarca sierra sur de Jaén (Insecta: Orthoptera). In: "In memoriam" al profesor Dr. Isidoro Ruiz Martínez, coord. JM Pérez Jiménez, págs. 295-328, ISBN 84-8439-195-7.
- Baker RJ, Kass DH (1994). Comparison of chromosomal distribution of a retrotransposon (LINE) and a retrovirus-like element mys in *Peromyscus maniculatus* and *P. leucopus*. *Chromosome Res* 2:185-189.
- Bakkali M & Camacho JPM (2004). The B chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa: III Mutation rate of B chromosomes. *Heredity* 92:428-433.
- Barranco P, Pascual F (1993). Estudio ecológico de los ortópteros de la vega del río Andarax (Insecta: Orthoptera). Distribución en relación a la vegetación. *Boln. Asoc. Esp. Ent.* 17: 285-297.
- Bartolomé C, Maside X & Charlesworth B (2002). On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol Biol Evol* 19 (6) :pages. 926-937.
- Belyayev A, Raskina O & Nevo E (2001). Chromosomal distribution of reverse transcriptase-containing retroelements in two Triticeae species. *Chromosome Res* 9:129-136.

- Biémont C (2010). A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* 186: 1085-1093.
- Bugrov A, Novikova O, Mayorov V, Adkison L & Blinov A (2006). Molecular phylogeny of Palaearctic genera of Gomphocerinae grasshoppers (Orthoptera, Acrididae). *Syst Entomol* 31: 362-368.
- Bureau TE, Wessler SR (1994). Stowaway: a new family of inverted repeat elements associated to genes of both monocotyledonous and dicotyledonous plants. *Plant Cell* 6:907-916.
- Burke WD, Malik HS, Jones JP & Eickbush TH (1999). The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol Biol Evol* 16: 502-511.
- Cabrero J, López-León MD, Bakkali M, Camacho JPM (1999). Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity* 83:435-439.
- Cabrero J, Bugrov A, Warchałowska-Sliwa E, López-León MD, Perfectti F, Camacho JPM (2003a). Comparative FISH analysis in five species of Eyprepocnemidine grasshoppers. *Heredity* 90:377-381.
- Cabrero J, Perfectti F, Gómez R, Camacho JPM, López-León MD (2003b). Population variation in the A chromosome distribution of satellite DNA and ribosomal DNA in the grasshopper *Eyprepocnemis plorans*. *Chromosome Res* 11:375-381.
- Cabrero J, Teruel M, Carmona FD, Jiménez R & Camacho JPM (2007). Histone H3 lysine 9 acetylation pattern suggests that X and B chromosomes are silenced during entire male meiosis in a grasshopper. *Cytogenet Genome Res* 119: 135-142.
- Cabrero J & Camacho JPM (2008). Location and expression of ribosomal RNA genes in grasshoppers: abundance of silent and cryptic loci. *Chromosome* 16: 595-607.
- Camacho JP, Sharbel TF & Beukeboom LW (2000). B-chromosome evolution. *Phil Trans R Soc Lond B* 355: 163-78.
- Camacho JPM, Bakkali M, Corral JM, et al. (2002). Host recombination is dependent on the degree of parasitism. *P Roy Soc B-Biol Sci* 269 (1505): 2173-2177.
- Capy P, Anxolabéhère D & Langin T (1994). The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet* 10: 7-12.

- Capy P, Gasperi G, Biémont C & Bazin C (2000). Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85: 101-106.
- Cermak T, Kubat Z, Hobza R, et al. (2008). Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosome* 16: 961-976.
- Cáceres M, Ranz JM, Barbadilla A, Long M & Ruiz A (1999). Generation of a widespread *Drosophila* inversion by a transposable element. *Science* 285: 415-418.
- Cermak T, Kubat Z, Hobza R et al (2008) Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosome Res* 16:961-976
- Contreras D & Chapco W (2006). Molecular phylogenetic evidence for multiple dispersal events in gomphocerine grasshoppers. *Journal of Orthoptera Research* 15: 91-98.
- Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998). Rates of spontaneous mutation. *Genetics* 148:1667-1686.
- Eickbush TH, Burke WD, Eickbush DG & Lathe WC (1997). Evolution of R1 and R2 in the rDNA units of the genus *Drosophila*. *Genetica* 100: 49-61.
- Eickbush MT & Eickbush TH (2011). Retrotransposition of R2 elements in somatic nuclei during the early development of *Drosophila*. *Mobile DNA* 2: 11.
- Feschotte C, Pritham EJ (2007). DNA Transposons and the evolution of eukaryotic Genomes. *Annu Rev Genet* 41:331-368.
- Friesen PD & Nissent MS (1990). Gene organization and transcription of TED , a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol Cell Biol* 10: 3067-3077.
- Fujiwara H, Osanai M, Matsumoto T, Kojima KK (2005). Telomere-specific non-LTR retrotransposon and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res* 13:455-467.
- Gaio D, Gill N, Kim HR et al (2009). A lineage-specific centromere retrotransposon in *Oryza brachyantha*. *Plant J* 60:820-831.

- George JA & Eickbush TH (1999). Conserved features at the 5' end of *Drosophila* R2 retrotransposable elements: implications for transcription and translation. *Insect Mol Biol* 8: 3-10.
- Ghesini S, Luchetti A, Marini M & Mantovani B (2011). The non-LTR retrotransposon R2 in termites (Insecta, Isoptera): characterization and dynamics. *J Mol Evol* 72: 296-305.
- Gomulski LM, Torti C, Murelli V, et al. (2004). Medfly transposable elements: diversity, evolution, genomic impact and possible applications. *Insect Biochem Mol Biol* 34: 139–148.
- Hartl DL, Lohe AR, Lozovskaya ER (1997). Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31:337-358.
- Henriques-Gil N, Arana P (1990). Origin and substitution of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Evolution* 44:747-753.
- Herédia F, Loreto ELS & Valente VLS (2004). Complex evolution of gypsy in *Drosophilid* species. *Mol Biol Evol* 21: 1831-42.
- Hewitt GM. 1979. Grasshopper and crickets. In: John B ed. *Animal Cytogenetics*, vol. 3: Insecta 1 Orthoptera. Berlin, Stuttgart: Gebruder Borntraeger.
- Hollister JD & Gaut BS (2007). Population and evolutionary dynamics of Helitron transposable elements in *Arabidopsis thaliana*. *Mol Biol Evol* 24: 2515-2524.
- Hood ME, Katawczik M & Giraud T (2005). Repeat-induced point mutation and the population structure of transposable elements in *Microbotryum violaceum*. *Genetics* 170: 1081-1089.
- Hsieh J, Fire A (2000). Recognition and silencing of repeated DNA. *Annu Rev Genet* 34:187-204.
- Jacobs G, Dechyeva D, Menzel G, Dombrowski, Schmidt T (2004). Molecular characterization of Vulmar 1, a complete mariner transposon of sugar beet and diversity of mariner and En/Spm like sequences in the genus Beta. *Genome* 47:1192-1201.
- Jakubczak JL, Burke WD & Eickbush TH (1991). Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *P Natl Acad Sci USA* 88: 3295-3299.

- Jakubczak JL, Zenni MK, Woodruff RC & Eickbush TH (1992). Turnover of R1 (Type I) and R2 (Type II) retrotransposable elements in the ribosomal DNA of *Drosophila melanogaster*. *Genetics* 131 (1): 883-894.
- Jin-Shan X, Qing-You X, Jun L, Guo-Qing P & Ze-Yang Z (2005). Survey of long terminal repeat retrotransposons of domesticated silkworm (*Bombyx mori*). *Insect Biochem Mol Biol* 35: 921–929.
- Jurka J, Kapitonov VV, Pavlicek A, et al. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110: 462-467.
- Kasschau KD, Fahlgren N, Chapman EJ et al (2007). Genome-wide profiling and analysis of Arabidopsis siRNA. *PLOS Biol* 5:e57.
- Kejnovsky E, Hobza R, Kubat Z, et al. (2007). High intrachromosomal similarity of retrotransposon long terminal repeats: evidence for homogenization by gene conversion on plant sex chromosomes? *Gene* 390: 92-97.
- Kidwell MG, Lish D (1997). Transposable elements as sources of variation animals and plants. *Proc Natl Acad Sci USA* 94:7704-7711.
- Kidwell MG & Lisch DR (2001). Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution* 55: 1-24.
- Kidwell MG (2002). Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49-63.
- Kim A, Terzian C, Santamaria P, et al. (1994). Retroviruses in invertebrates: the *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *P Natl Acad Sci USA* 91: 1285-1289.
- Kojima KK & Fujiwara H (2005). Long-term inheritance of the 28S rDNA-specific retrotransposon R2. *Mol Biol Evol* 22: 2157-2165.
- Lamb JC, Riddle NC, Cheng YM, Theuri J, Birchler JA (2007). Location and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Res* 15:383-398.
- Lander ES, Linton LM, Birren B et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
- López-León MD, Cabrero J & Camacho JPM (1991). A nucleolus organizer region in a B chromosome inactivated by DNA methylation. *Chromosoma* 100: 134-138.

- López-León MD, Cabrero J, Pardo MC, Viseras E, Camacho JPM, Santos JL (1993). Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity*: 71:352-362.
- López-León MD, Neves N, Schwarzacher T, Heslop-Harrison JS, Hewitt GM, Camacho JPM (1994). Possible origin of B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Res* 2:87-92.
- López-León MD, Vázquez P, Hewitt GM & Camacho JPM (1995). Cloning and sequence analysis of an extremely homogeneous tandemly repeated DNA in the grasshopper *Eyprepocnemis plorans*. *Heredity* 75: 370-375.
- Malik HS & Eickbush TH (1998). The RTE class of non-LTR retrotransposons is widely distributed in animals and is the origin of many SINEs. *Mol Biol Evol* 15: 1123-1134.
- Malik HS, Burke WD & Eickbush TH (1999). The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16: 793-805.
- Malik HS. (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 10: 1307-1318.
- Marchal JA, Acosta Manuel J, Bullejos Mónica, et al. (2006). Distribution of L1-retroposons on the giant sex chromosomes of *Microtus cabreræ* (Arvicolidae, Rodentia): functional and evolutionary implications. *Chromosome Res* 14: 177-186.
- Marin I & Llorens C (2000). Ty3/Gypsy retrotransposons: description of new *Arabidopsis thaliana* elements and evolutionary perspectives derived from comparative genomic data. *Mol Biol Evol* 17: 1040-1049.
- Mingazzini V, Luchetti A & Mantovani B (2011). R2 dynamics in *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca): turnover rate and 28S concerted evolution. *Heredity* 106: 567-575.
- Mukabayire O, Besansky NJ (1996). Distribution of T1, Q, Pegasus and mariner transposable elements on the polytene chromosomes of PEST, a standard strain of *Anopheles gambiae*. *Chromosoma* 104:585-595.
- Muñoz-López M, Siddique A, Bischerour J, et al (2008). Transposition of Mboumar-9: identification of a new naturally active mariner-family transposon. *J Mol Biol* 382:567-72.

- Muñoz-Pajares AJ, Martínez-Rodríguez L, Teruel M, et al. (2011). A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* 187: 853-63.
- Olmo-Vidal JM (2002) Atlas de los Ortópteros de Cataluña. Atlas de Biodiversidad no. 1: 209-336. Generalitat de Catalunya. Departament de Medioambient.
- Pardue ML (1995). Drosophila telomeres: another way to end it all. In: Blackburn EH, Graider CW (eds) *Telomeres*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp 339-370.
- Parkin EJ & Butlin RK (2004). Within- and between-individual sequence variation among ITS1 copies in the meadow grasshopper *Chorthippus parallelus* indicates frequent intrachromosomal gene conversion. *Mol Biol Evol* 21: 1595-1601.
- Pélisson A, Teyssset L, Chalvet F, et al. (1997). About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the *Drosophila flamenco* host gene. *Genetica* 100: 29-37.
- Pérez-González CE & Eickbush TH (2001). Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics* 158: 1557-1567.
- Piskurek O & Okada N (2007). Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *P Natl Acad Sci USA* 104: 12046-12051.
- Rebuzzini P, Castiglia R, Nergadze G et al (2009). Quantitative variation of LINE-1 sequences in five species and three subspecies of the subgenus *Mus* and in five Robertsonian races of *Mus musculus domesticus*. *Chromosome Res* 17:65-76.
- Rezende-Teixeira P, Siviero F, Costa RM da & Machado-Santelli GM (2009). The R2 mobile element of *Rhynchosciara americana*: molecular, cytological and dynamic aspects. *Chromosome Res* 17: 455-467.
- Roy AM, Carroll ML, Nguyen SV, et al. (2000). Potential gene conversion and source genes for recently integrated Alu elements. *Genome Res* 10: 1485-1495.
- Sánchez-Gracia A, Maside X & Charlesworth B (2005). High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 21: 200-203.
- Santiago M, Herraiz C, Goñi JR, Messeguer X, Casacuberta JM (2002). Genome wide analysis of the Emigrant family of MITEs of *Arabidopsis thaliana* *Mol Biol Evol* 19:2285-2293.

- Schaack S, Gilbert C & Feschotte C (2010). Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* 25: 537-546.
- Schafer M & Kunz W (1987). Ribosomal gene amplification does not occur in the oocytes of *Locusta migratoria*. *Developmental Biology* 120: 418–424.
- Schnable PS, Ware D, Fulton RS, et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-1115.
- Subramanian RA, Akala OO, Adejinmi JO & O’Brochta DA (2008). Topi, an IS630/Tc1/mariner-type transposable element in the African malaria mosquito, *Anopheles gambiae*. *Gene* 423: 63–71.
- Teruel M (2009). Origen, expresión y efectos fenotípicos de un parásito genómico. Ph. D. thesis, Universidad de Granada, Spain.
- Teruel M, Cabrero J, Perfectti F & Camacho JPM (2009). Quantitative analysis of NOR expression in a B chromosome of the grasshopper *Eyprepocnemis plorans*. *Chromosoma* 118: 291-301.
- Torti C, Ludvik M, Gomulski M, Maralli D, Raimondi E, Robertson HM (2000). Evolution of different subfamilies of mariner elements within mdmly genome inferred from abundance and chromosomal distribution. *Chromosoma* 108:523-532.
- Vicient C (2010). Transcriptional activity of transposable elements in maize. *BMC Genomics* 11:601-610.
- Volff JN, Körting C, Schartl M (2000). Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. *Mol Biol Evol* 17:1673-1684.
- Waterston RH, Lindblad-Toh K, Birney E, et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-562.
- Zhang X & Eickbush TH (2005). Characterization of active R2 retrotransposition in the rDNA locus of *Drosophila simulans*. *Genetics* 170: 195-205.
- Zhang X, Eickbush MT & Eickbush TH (2008). Role of recombination in the long-term retention of transposable elements in rRNA gene loci. *Genetics* 180: 1617-1626.

- Zhou J & Eickbush TH (2009). The pattern of R2 retrotransposon activity in natural populations of *Drosophila simulans* reflects the dynamic nature of the rDNA locus. *PLoS genetics* 5: e1000386.
- Zupunski V, Gubensek F & Kordis D (2001). Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons. *Mol Biol Evol* 18: 1849-63.
- Zurita S, Cabrero J, López-León MD, Camacho JPM (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution* 52:274-277.

CAPÍTULO 7

Conclusiones

1. Los elementos *Gypsy*, *RTE* y *Mariner* son muy abundantes en los genomas de *E. plorans* y *L. migratoria*, mostrando una localización cromosómica no aleatoria ya que se distribuyen uniformemente por la eucromatina, encontrándose ausentes en la heterocromatina, en las NORs (en *E. plorans*) y clúster de histonas (en *L. migratoria*) y son menos abundantes en los cromosomas B. El elemento R2 se localiza mayoritariamente en el clúster de ADN ribosómico de los cromosomas B de *E. plorans* y en menor medida en los clústeres ribosómicos de los cromosomas S_{10} y S_{11} .

2. Las estimas del número de copias muestran que *Gypsy* es el elemento mayoritario en *E. plorans*, seguido por *RTE* y *Mariner*, siendo el elemento *Mariner* el que muestra un número similar de copias en *E. plorans* y *L. migratoria*, mientras que *Gypsy* y *RTE* presentan valores muy bajos en *L. migratoria*.

3. El elemento R2 de *E. plorans* muestra la estructura típica para este elemento y presenta menor número de copias que los elementos *Gypsy* y *RTE*, con un 90% de sus copias en el clúster de ADN ribosómico del cromosoma B y un 10% en los de los A. Esto apunta a los cromosomas B como un excelente destino cromosómico para la inserción de R2.

4. La estima de unidades de ADNr en *E. plorans* es muy superior a la de otros insectos pero la proporción que representan del genoma (1,3%) es muy similar. La ocupación de estas unidades por R2 es del 2%, siendo mayor en la variante de B más antigua, B_1 , que en las más recientes B_2 y B_{24} , presumiblemente debido al menor tiempo disponible para la acumulación de R2 de éstas últimas.

5. Los elementos *Gypsy*, *RTE* y *Mariner* son transcripcionalmente activos en los genomas de *E. plorans* y *L. migratoria*. Sin embargo, no hemos encontrado expresión del elemento R2 aunque sí hemos detectado muestras de su actividad en los perfiles de truncaciones del extremo 5' lo que sugiere un pasado reciente activo.

6. Los elementos *Gypsy* y *RTE* están presentes en la familia Acridoidea, habiéndose encontrado el elemento *RTE* en 32 especies de este grupo mientras que *Gypsy* fue aislado en 28 especies.

7. La filogenia obtenida a partir de los elementos *Gypsy* y *RTE* se ajusta a la filogenia de ADN mitocondrial y la taxonomía del grupo de los acrididos aunque se presentan ciertas incongruencias que apuntan a posibles casos de HTT.

8. El estudio de la diversidad de copias del elemento *Gypsy* en *E. plorans* mostró la posibilidad de un “lineage sorting” incompleto en lugar de una HTT entre *E. plorans* y *H. littoralis*.

9. La microdissección de todos los cromosomas de una célula por separado y la amplificación por PCR y secuenciación de una determinada secuencia en cada uno de ellos es una aproximación directa al estudio intragenómico de copias parálogas, como las de los elementos *RTE*, que permite realizar una “filogeografía intragenómica”.

10. El análisis filogenético intragenómico de las copias de *RTE* muestra la existencia de dos linajes en *E. plorans*: Un linaje A que está presente en todos los cromosomas excepto el cromosoma S_{11} y B, y un linaje B que está presente en todos. Estos linajes sugieren la posibilidad de dos oleadas de invasión de *RTE* anteriores a la aparición del cromosoma B.

11. El bajo valor de F_{ST} entre los cromosoma S_{11} y B, así como que compartan un único tipo de *RTE* apoyan el origen del cromosoma B a partir de este autosoma.

12. El linaje A de *RTE* debe ser el transmitido de forma vertical en el grupo de los eyprepocnemidinos mientras que el linaje B parece ser fruto de un evento de HTT desde los oedipodinos como sugiere el análisis filogenético de estos linajes junto con 31 especies de acrididos. Aunque se necesita un análisis similar al realizado en *E. plorans* para poder corroborar la HTT.

CAPÍTULO 8

Perspectivas

Las investigaciones desarrolladas en esta Tesis Doctoral han permitido conocer muchos e interesantes aspectos del papel de los elementos transponibles en el genoma de ortópteros, especialmente de las especies *Eyprepocnemis plorans* y *Locusta migratoria*, lo que ha permitido plantear, a su vez, nuevos interrogantes que serán muy interesantes de abordar en investigaciones futuras. En este apartado de perspectivas nos proponemos describir algunas líneas de trabajo que pretendemos desarrollar en nuestro grupo de investigación para profundizar en el conocimiento de los elementos transponibles que componen los genomas de ortópteros.

La enorme diversidad de TEs observada en eucariotas, especialmente entre los artrópodos, sugiere que deben existir una mayor variedad de tipos de TEs poblando los genomas de saltamontes. Así, el aislamiento de nuevos TEs mediante la amplificación con cebadores degenerados basados en zonas conservadas de la transposasa o retrotranscriptasa, tal y como se ha realizado en este proyecto, podría llevarse cabo en *E. plorans* y *L. migratoria* y ser extendido a otras especies de ortópteros. Con un análisis más intenso de tipos de TE y número de secuencias amplificadas aumentaríamos las posibilidades de cubrir toda la diversidad de TEs existente en estos saltamontes, pudiendo también conocer aquellos elementos que colonizan la heterocromatina.

Su posible papel dentro de la eucromatina podrá ser abordado conociendo las secuencias que flanquean dichos elementos, lo que aclarará su función en los genomas pudiendo aparecer como ADN basura que forma parte de los espaciadores intergénicos, intrones o pseudogenes o formando parte de secuencias funcionales como las reguladoras.

Del mismo modo, la caracterización completa de los elementos encontrados en estas especies, *Gypsy*, *RTE* y *Mariner*, nos permitiría diseñar nuevos cebadores para ahondar en la actividad y evolución de los mismos.

El cromosoma B, en contra de lo que establece la teoría, no ha mostrado una acumulación de este tipo de elementos (salvo el elemento R2). Sin embargo, la asequibilidad que los proyectos de secuenciación masiva están alcanzando, permitiría la secuenciación de este cromosoma y el análisis posterior de las secuencias obtenidas podría desvelar la presencia de nuevos TEs que ayudaran a explicar la dinámica evolutiva de este cromosoma. Una vez conocida la presencia de nuevos tipos de TE en el cromosoma B, su estudio podría ampliarse a las diferentes variantes. Además, el establecimiento de nuevas filogeografías intragenómicas podrían proporcionarnos una mejor visión del origen de las variantes de Bs y de la influencia de la población considerada en la dinámica de dichos elementos.

Hasta el momento, del único elemento del que tenemos constancia de acumulación en el cromosoma B es el elemento R2. El estudio de la expresión de este elemento ha quedado incompleto, necesitando la extensión del mismo tanto a embriones como a hembras, donde se ha descrito su expresión en otras especies. Asimismo, podríamos aumentar el estudio poblacional de este elemento en *E. plorans* incluyendo a poblaciones de Europa oriental con distintas variantes de cromosomas B que nos permitiría una visión más global de la relación R2-B en esta especie. Además, la posible implicación de R2 en los niveles de actividad de la NOR del B podrá ser aclarada estudiando su presencia y abundancia en los cromosomas Bs cuya NOR se ha encontrado activa.

Por último, los distintos eventos de HTT sugeridos en las filogenias de *RTE* y *Gypsy* requerirían ser tratados en profundidad, analizando la diversidad existentes de estos elementos en las especies implicadas y en otras del mismo género, como sugerimos en la discusión general, para poder corroborar si estamos ante un “lineage sorting” incompleto o realmente ante una HTT.

ANEXOS

Anexo A. Tablas Suplementarias

Table S1. CADM test of pairwise mitochondrial DNA genetic distance for predicting congruence. Holm adjustment for multiple tests was performed.

Pairwise genetic distance			
Global congruence	H_0 : matrices are incongruent		
Kendal's W	0.7977		
Friedman's X	902.23***		
Posterior pairwise congruence	H_0 : matrix is incongruent with remaining matrices H_1 : matrix is congruente with at least one matrix		
	P-value	Adjusted P-value	
COI (5'end)	***	***	
COI (3'end)	***	***	
ND5	***	***	
One-tailed Mantel test	H_0 : $r=0$; H_1 : $r>0$		
	COI (5'end)	COI (3'end)	ND5
COI (5'end)	1.0000	0.6327***	0.7464***
COI (3'end)		1.0000	0.7106***
ND5			1.0000

CADM, congruence among distance matrices; COI, cytochrome oxidase subunit I gene; ND5: NADH subunit 5 gene. * P = 0.05; ** P = 0.01; *** P = 0.001; ns: not significant.

Table S2. Results of the substitution saturation test following the procedure described in Xia et al. (2003). The test compares the estimated index of substitution saturation (ISS) with the critical value of the index of substitution saturation, ISS.C Sym, when the true tree is symmetrical, or ISS.C Asym, if the true tree is asymmetrical. A significant value implies no substitution saturation.

Gene	Codon Position	I_{SS}	$I_{SS.C}$ Sym.	P	$I_{SS.C}$ Asym.	P
COI (5' end)	1 & 2	0.1005	0.6055	***	0.3582	***
	3	0.5985	0.5244	ns	0.4302	***
COI (3' end)	1 & 2	0.0793	0.5886	***	0.3617	***
	3	0.4591	0.5076	ns	0.4691	ns
ND5	1 & 2	0.2307	0.6204	***	0.3589	***
	3	0.4740	0.6210	***	0.3590	***
Gypsy	1 & 2	0.0853	0.6302	***	0.3710	***
	3	0.1172	0.5133	***	0.3533	***
RTE	1 & 2	0.0431	0.5589	***	0.3808	***
	3	0.3069	0.4784	**	0.5735	***

COI: cytochrome oxidase subunit I gene; ND5: NADH subunit 5 gene.

* P = 0.05; ** P = 0.01; *** P = 0.001; ns: not significant.

Table S3. Codon-based test of neutrality, positive selection and purifying selection for each lineage of RTE element (lineage A and lineage B) within each chromosome. The probability of rejecting the null hypothesis of strict-neutrality ($dN = dS$), in favor of the alternative hypothesis, positive selection, ($dN > dS$) and purifying selection, ($dN < dS$) is shown. Values of P lower than 0.05 are highlighted in bold-type letter. dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively. The variance of the difference was computed using the bootstrap method (500 replicates). Analyses were conducted using the Nei-Gojobori method in MEGA4. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). There were 57 codons for lineage A and 58 for lineage B.

<i>RTE</i> lineage A in chromosome no.	N	Neutrality ($dN = dS$)		Positive Selection ($dN > dS$)		Purifying Selection ($dN < dS$)	
		P	Stat	P	Stat	P	Stat
1	11	0.002	-3.165	1.000	-3.125	0.001	3.249
2	2	0.266	-1.117	1.000	-1.090	0.133	1.116
3	10	0.001	-3.250	1.000	-2.965	0.002	2.974
4	6	0.052	-1.965	1.000	-1.885	0.035	1.829
5	6	0.025	-2.268	1.000	-2.373	0.011	2.337
6	9	0.004	-2.904	1.000	-2.844	0.002	2.895
7	15	0.001	-3.466	1.000	-3.304	0.000	3.564
8	17	0.034	-2.141	1.000	-1.924	0.018	2.128
9	11	0.004	-2.970	1.000	-3.141	0.003	2.828
10	10	0.028	-2.222	1.000	-1.991	0.017	2.136
X	12	0.001	-3.423	1.000	-3.297	0.000	3.393
Total	109	0.000	-5.490	1.000	-5.605	0.000	5.742
<i>RTE</i> lineage B in chromosome no.							
1	5	0.011	-2.576	1.000	-2.644	0.005	2.606
2	13	0.005	-2.888	1.000	-3.118	0.001	3.184
3	10	0.004	-2.897	1.000	-2.926	0.003	2.809
4	13	0.194	-1.307	1.000	-1.415	0.103	1.270
5	10	0.169	-1.383	1.000	-1.442	0.078	1.425
6	11	0.121	-1.564	1.000	-1.627	0.052	1.643
7	5	0.248	-1.161	1.000	-1.156	0.130	1.132
8	1	-	-	-	-	-	-
9	5	0.026	-2.261	1.000	-2.238	0.011	2.334
10	4	0.191	-1.316	1.000	-1.391	0.109	1.241
11	9	0.238	-1.185	1.000	-1.199	0.096	1.310
B	22	0.134	-1.510	1.000	-1.534	0.068	1.503
X	10	0.011	-2.567	1.000	-2.544	0.009	2.414
Total	118	0.001	-3.282	1.000	-3.505	0.000	3.579

Anexo B. Figuras Suplementarias

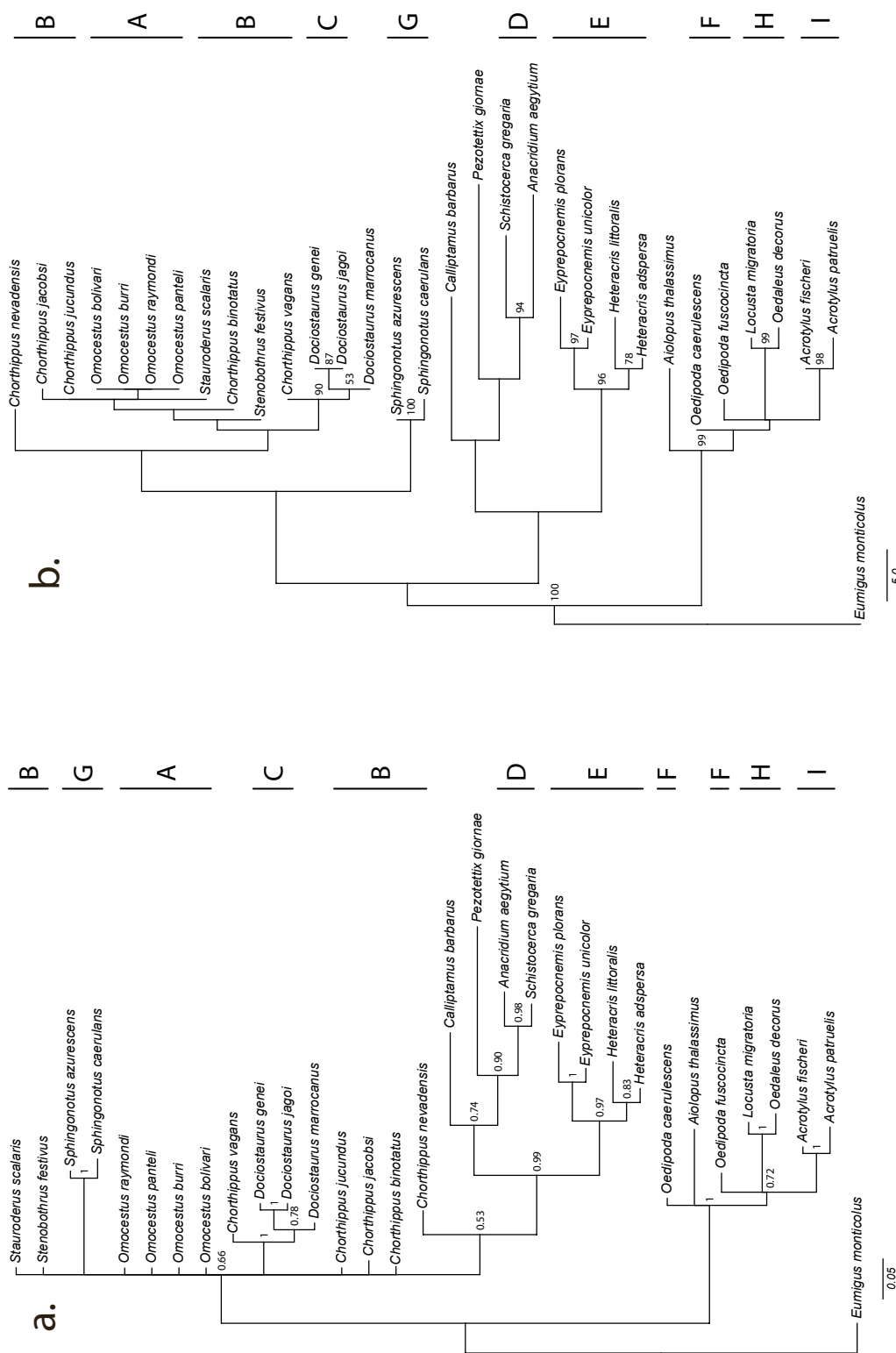


Figure S1. Phylogenetic trees for mtDNA obtained by Bayesian inference (a) and maximum parsimony (b). Only branch supports higher than 50% or 0.5 are shown.

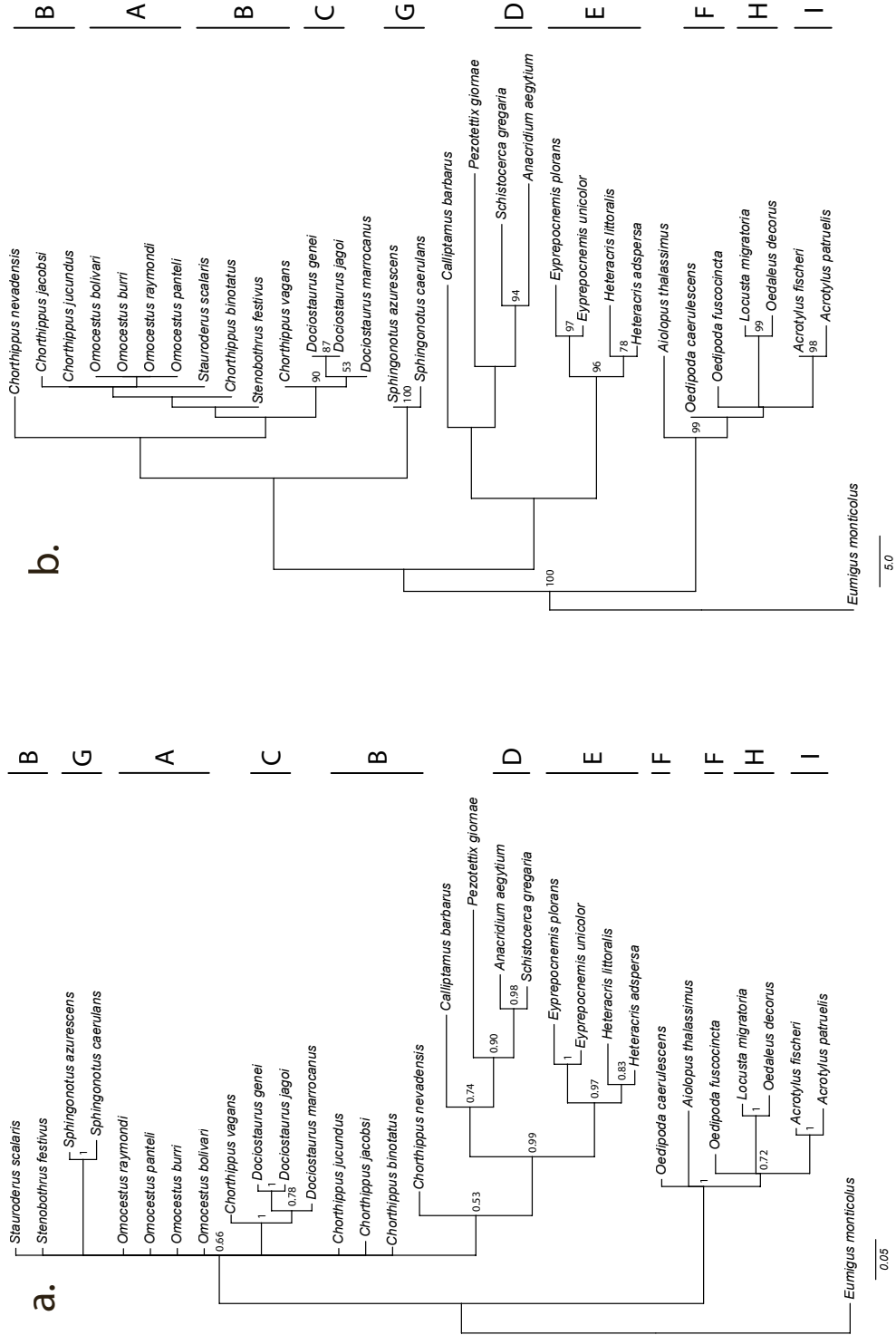


Figure S2. Phylogenetic trees for *RTE* obtained by Bayesian inference (a) and maximum parsimony (b). Only branch supports higher than 50% or 0.5 are shown.

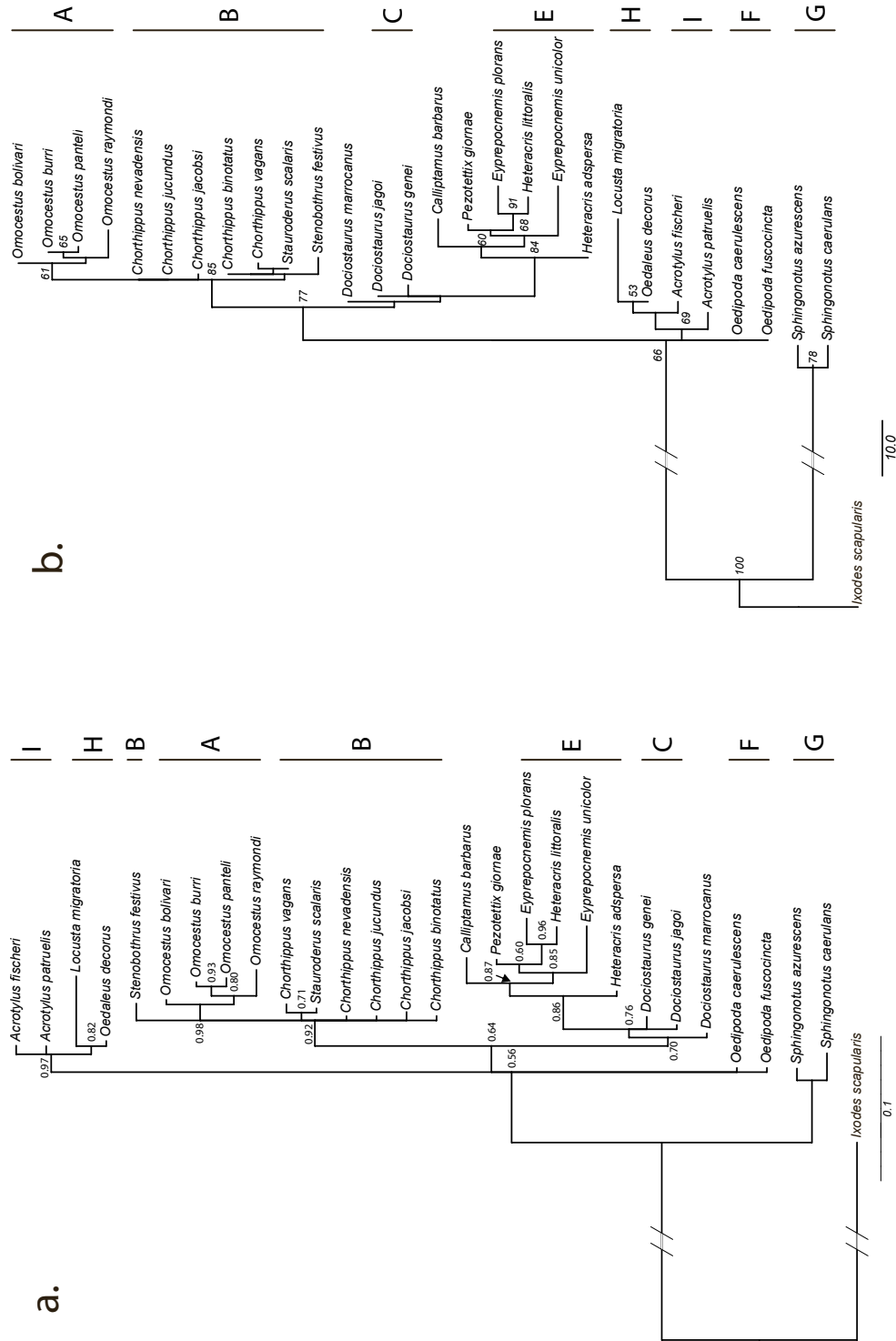


Figure S3. Phylogenetic trees for *Gypsy* obtained by Bayesian inference (a) and maximum parsimony (b). Only branch supports higher than 50% or 0.5 are shown.

Anexo C. Símbolos y Abreviaturas

μg	microgramo	IS	insertion sequence
μl	microlitro	ISS	index of substitution saturation
μM	micromolar	ISS.C	critical value of the index of substitution saturation
A	adenina	Kb	kilobase
ADN / DNA	ácido desoxirribonucleico / deoxyribonucleic acid	LINE	long interspersed nuclear element
ADN c / cDNA	ADN complementario / complementary DNA	LTR	long terminal repeat
ADNr / rDNA	ADN ribosómico / ribosomal DNA	Mb	megabase
ADNsat / satDNA	ADN satélite / satellite DNA	mg	miligramo
AMOVA	Analysis of Molecular Variance	min	minuto
ARN / RNA	ácido ribonucleico / ribonucleic acid	MITE	miniature inverted repeat transposable element
ARNi	ARN de interferencia	ML	maximum likelihood
ARNr / rRNA	ARN ribosómico / ribosomal RNA	ml	mililitro
BI	Bayesian inference	mM	milimolar
BLAST	Basic Local Alignment Search Tool	MP	maximum parsymonia
bp	base pair	mtDNA	mitochondrial DNA
C	citocina	MULE	Mutator-like element
CADM	congruence among distance matrices	N	número de secuencias
CMA3	cromomicina A3	NATE	Nasonia trasposable element
COI	cytochrome oxidase subunit I gene	NCBI	National Center for Biotechnology Information
COI3'	cytochrome oxidase subunit I gene (3' end)	ND5	NADH subunit 5 gene
COI5'	cytochrome oxidase subunit I gene (5' end)	ng	nanogramo
DAPI	4', 6' diamidino-2-fenilindol	NOR	nucleolar organizing region
dN	número de sustituciones no sinónimas por sitio	ORF	open reading frame
dS	número de sustituciones sinónimas por sitio	PCR	polymerase chain reaction
EN	endonucleasa	pg	picogramo
FISH	fluorescence in situ hybridization	PSR	paternal sex ratio
F _{ST}	Fixation index	qPCR	quantitative polymerase chain reaction
G	guanina	RT	retrotranscriptasa
h	hora	RTE	retrotransposable element
H	número de haplotipos	s	segundo
HTT	horizontal transposon transference	S	número de sitios segregantes

SINE	short interspersed nuclear element	Tn	Transposon
T	timina	UTR	untranslated region
TE	transposable element	⊖	heterozigosidad (por sitio)
TIR	terminal inverted repeat	π	diversidad nucleotídica (por sitio)

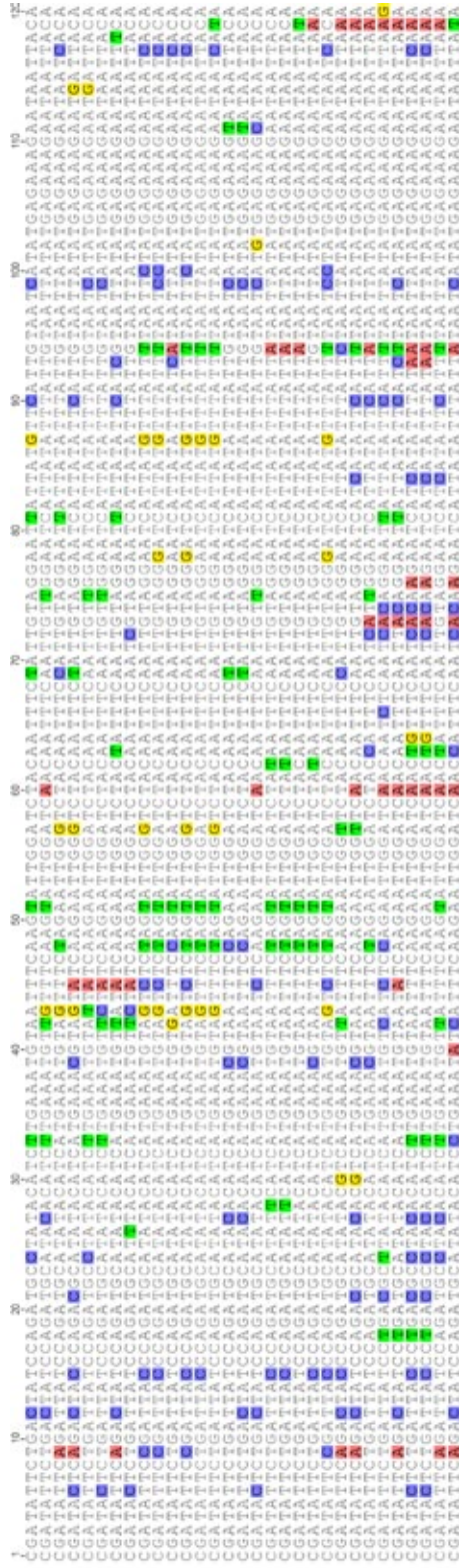
Anexo D. Alineamientos

COI 5'

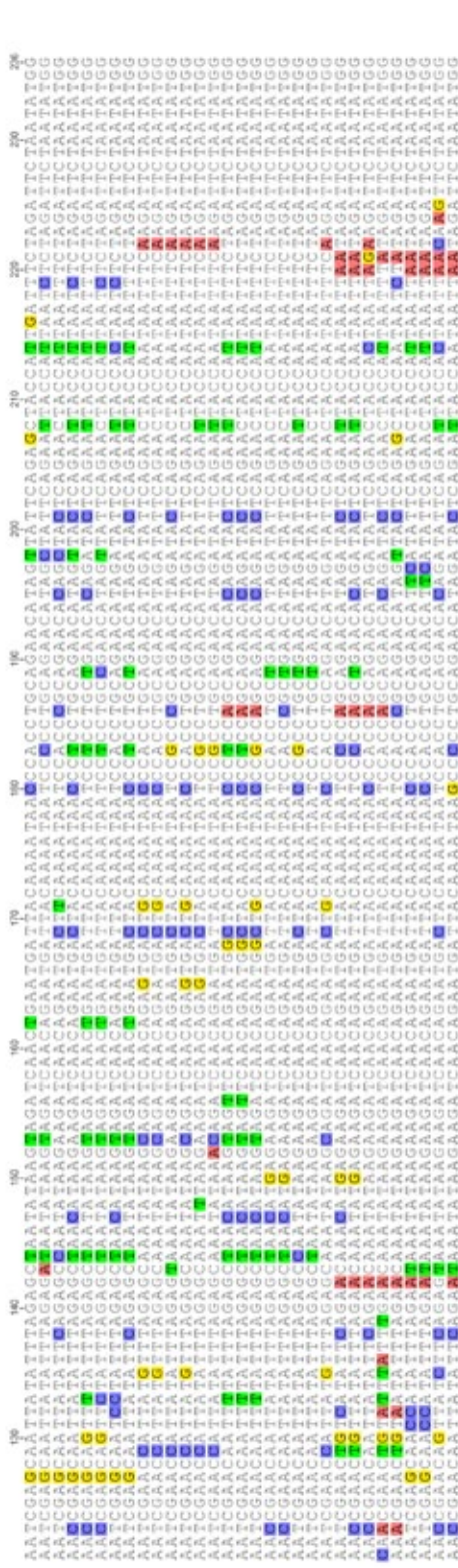
Callitamus barbarus
Pezomachus gramineus
Anacridium aegyptium
Schistocerca gregaria
Eyprocneis plirans
Eyprocneis umicolor
Heteracris adspersa
Heteracris illiriana
Chorthippus binotatus
Chorthippus lucidulus
Chorthippus luteiventris
Chorthippus viridis
Chorthippus scaber
Decostaurus legallii
Decostaurus macromacrus
Omocestus bellii
Omocestus parvulus
Omocestus raymondii
Stenobothrus festinus
Acrotylus fischeri
Acrotylus palustris
Acrotylus thalassinus
Locusta migratoria
Deutellus decoratus
Oedipoda caerulescens
Oedipoda fuscescens
Sphingonotus azureus
Sphingonotus caeruleus
Eumigulus monticola

Callitamus barbarus
Pezomachus gramineus
Anacridium aegyptium
Schistocerca gregaria
Eyprocneis plirans
Eyprocneis umicolor
Heteracris adspersa
Heteracris illiriana
Chorthippus binotatus
Chorthippus lucidulus
Chorthippus luteiventris
Chorthippus viridis
Chorthippus scaber
Decostaurus legallii
Decostaurus macromacrus
Omocestus bellii
Omocestus parvulus
Omocestus raymondii
Stenobothrus festinus
Acrotylus fischeri
Acrotylus palustris
Acrotylus thalassinus
Locusta migratoria
Deutellus decoratus
Oedipoda caerulescens
Oedipoda fuscescens
Sphingonotus azureus
Sphingonotus caeruleus
Eumigulus monticola

COI 3'



- Calligramma barbatum*
- Pezomachus barbatum*
- Anacardium acuminatum*
- Schistocerca gregaria*
- Epinipponema plerumque*
- Epinipponema unicolor*
- Heteracris adspersa*
- Heteracris illivensis*
- Chromyppus binotatus*
- Chromyppus jacoboi*
- Chromyppus nigritarsis*
- Chromyppus reuteri*
- Chromyppus scalaris*
- Stenobothrus scallaris*
- Dociostaurus gemini*
- Dociostaurus jingui*
- Dociostaurus mairianus*
- Oncocetia burri*
- Oncocetia parviflora*
- Oncocetia raymondii*
- Oncocetia scalaris*
- Stenobothrus fastuosus*
- Arotopylus fischeri*
- Arotopylus pituitus*
- Leucilla migratoria*
- Oedaleus decorus*
- Oedaleus decorus*
- Oedaleus decorus*
- Sahlgreniella azurescens*
- Sahlgreniella azurescens*
- Sahlgreniella azurescens*
- Elumyga monticola*



- Calligramma barbatum*
- Pezomachus barbatum*
- Anacardium acuminatum*
- Schistocerca gregaria*
- Epinipponema plerumque*
- Epinipponema unicolor*
- Heteracris adspersa*
- Heteracris illivensis*
- Chromyppus binotatus*
- Chromyppus jacoboi*
- Chromyppus nigritarsis*
- Chromyppus reuteri*
- Chromyppus scalaris*
- Stenobothrus scallaris*
- Dociostaurus gemini*
- Dociostaurus jingui*
- Dociostaurus mairianus*
- Oncocetia burri*
- Oncocetia parviflora*
- Oncocetia raymondii*
- Oncocetia scalaris*
- Stenobothrus fastuosus*
- Arotopylus fischeri*
- Arotopylus pituitus*
- Leucilla migratoria*
- Oedaleus decorus*
- Oedaleus decorus*
- Oedaleus decorus*
- Sahlgreniella azurescens*
- Sahlgreniella azurescens*
- Sahlgreniella azurescens*
- Elumyga monticola*

ND 5

Calliptenopus barbarus
Pezolelethys gnomae
Anacardium aegoptilum
Schistocerca gregaria
Eyprepocnemis plorans
Eyprepocnemis unicolor
Heteracris adspersa
Heteracris fibrillata
Chorthippus boidulatus
Chorthippus jacoboli
Chorthippus jumcunus
Chorthippus nevadensis
Chorthippus vespignani
Saurodonnina scabaris
Decostaurus gemeli
Decostaurus legai
Omocestus nabhari
Omocestus burni
Omocestus parvelli
Omocestus rymondii
Sternobothrus festinus
Acrotylus fischeri
Acrotylus palustris
Alolopus thalassinus
Locusta migratoria
Oedipoda decorata
Oedipoda caeruleascens
Oedipoda fuscoconcha
Schlingonotus azurescens
Schlingonotus citerulens
Eumrigus monticolus

Calliptenopus barbarus
Pezolelethys gnomae
Anacardium aegoptilum
Schistocerca gregaria
Eyprepocnemis plorans
Eyprepocnemis unicolor
Heteracris adspersa
Heteracris fibrillata
Chorthippus boidulatus
Chorthippus jacoboli
Chorthippus jumcunus
Chorthippus nevadensis
Chorthippus vespignani
Saurodonnina scabaris
Decostaurus gemeli
Decostaurus legai
Omocestus nabhari
Omocestus burni
Omocestus parvelli
Omocestus rymondii
Sternobothrus festinus
Acrotylus fischeri
Acrotylus palustris
Alolopus thalassinus
Locusta migratoria
Oedipoda decorata
Oedipoda caeruleascens
Oedipoda fuscoconcha
Schlingonotus azurescens
Schlingonotus citerulens
Eumrigus monticolus

ND 5 continuación

Calliptamus barbarus
 Pezomachus gnomae
 Anacardium acytophllum
 Schistoscoera grogna
 Eypriocnemis plarans
 Eypriocnemis unicolor
 Heteractis adspersa
 Heteractis fibrilis
 Chorthippus binoctatus
 Chorthippus jacobae
 Chorthippus junceus
 Chorthippus nevadensis
 Chorthippus vagans
 Sauridromus scalaris
 Sauridromus gmelii
 Doceataurus mirrocanus
 Doceataurus bolivari
 Omocestus burri
 Omocestus parvelli
 Omocestus raymondii
 Stenobothrus festinus
 Stenobothrus fischeri
 Acrotylus fhalassinus
 Acrotylus migratorius
 Oedaeus decorus
 Oedipoda caerulescens
 Oedipoda fuscoconcha
 Sahingonotus azurescens
 Sahingonotus caerulens
 Eurymys monticola

Calliptamus barbarus
 Pezomachus gnomae
 Anacardium acytophllum
 Schistoscoera grogna
 Eypriocnemis plarans
 Eypriocnemis unicolor
 Heteractis adspersa
 Heteractis fibrilis
 Chorthippus binoctatus
 Chorthippus jacobae
 Chorthippus junceus
 Chorthippus nevadensis
 Chorthippus vagans
 Sauridromus scalaris
 Doceataurus gmelii
 Doceataurus mirrocanus
 Doceataurus bolivari
 Omocestus burri
 Omocestus parvelli
 Omocestus raymondii
 Stenobothrus festinus
 Stenobothrus fischeri
 Acrotylus fhalassinus
 Acrotylus migratorius
 Oedaeus decorus
 Oedipoda caerulescens
 Oedipoda fuscoconcha
 Sahingonotus azurescens
 Sahingonotus caerulens
 Eurymys monticola

Gypsy continuación

Callipterus barbarus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Pezetolittix ghomae	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Eyrepocnemis plorans	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Eyrepocnemis unicolor	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Heteractis adspersa	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Heteractis iloraltis	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Chorthippus binotatus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Chorthippus leucobeli	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Chorthippus leucundus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Chorthippus nevadensis	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Chorthippus vulgans	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Stauroderus scotaris	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Dociostaurus genei	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Dociostaurus jippei	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Dociostaurus marrocanus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Omocestus bolivari	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Omocestus burri	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Omocestus pantei	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Omocestus raymondii	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Stenobothrus festinus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Acrotylus fischeri	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Acrotylus patruelis	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Locusta migratoria	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Oedipoda caerulescens	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Oedipoda fuscoocincta	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Sphingonotus azureus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354
Sphingonotus caenulatus	AGMAGGTTCCAGTCCACCAACACAGATGACCGTTAAACCTGTGAAGTCTGTGCGATAGTAAAGTTTTCAGGCTCCAAAAGAAAGTATCATTAAGCAATTAC	354

