UNIVERSIDAD DE GRANADA
DEPARTAMENTO DE GENÉTICA

Análisis de los cambios de expresión génica asociados a la presencia de cromosomas B en el saltamontes *Eyprepocnemis plorans*

Beatriz María Navarro Domínguez

Tesis Doctoral

Editor: Universidad de Granada. Tesis Doctorales
Autora: Beatriz María Navarro Domínguez
ISBN: 978-84-9125-997-8
URI: http://hdl.handle.net/10481/44096

Análisis de los cambios de expresión génica asociados a la presencia de cromosomas B en el saltamontes *Eyprepocnemis plorans*

Memoria de Tesis Doctoral presentada por la Licenciada Beatriz María Navarro Domínguez para optar al grado de Doctora por la Universidad de Granada

Dirigida por los Doctores:

Dra. Josefa Cabrero Hurtado

Dr. Juan Pedro Martínez Camacho

Dra. María Dolores López León

La doctoranda Beatriz María Navarro Domínguez y los directores de la tesis Juan Pedro Martínez Camacho, María Dolores López León y Josefa Cabrero Hurtado garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza en la realización del trabajo se han respetado los derechos de otros autores a ser citados cuando se han utilizado sus publicaciones

Granada, a 23 de Junio de 2016

Dra. Josefa Cabrero Hurtado	Dr. Juan Pedro Martínez Camacho	Dra. María Dolores López León

Beatriz María Navarro Domínguez

Este trabajo ha estado financiado por los proyectos CGL2009-11917 del Ministerio de Economía y Competitividad y P10-CVI-6649 de la Consejería de Economía, Innovación, Ciencia y Empleo de la Junta de Andalucía.

Durante el tiempo de realización de esta Tesis Doctoral he disfrutado de una Beca del Programa Nacional de Formación de Personal Investigador del Ministerio de Economía y Competitividad (BES-2010-030016), asociada al proyecto CGL2009-11917.

La investigación presentada en esta Tesis Doctoral se ha desarrollado casi en su totalidad en el Departamento de Genética de la Universidad de Granada. Se ha realizado una parte en el laboratorio del Dr. Timothy F. Sharbel en IPK Gatersleben (Alemania), durante dos estancias.

Cuando creíamos que teníamos todas las respuestas, de pronto, cambiaron todas las preguntas.

Mario Benedetti

Agradecimientos

Recordar: del latín re-cordis,
volver a pasar por el corazón

Eduardo Galeano

Seis largos años resumidos en más de trescientas páginas, que es sólo una pequeña parte de lo que ha sido para mi el doctorado. Han sido tantos buenos y malos momentos, he aprendido tantas cosas, hay tanta gente a la que recordar en el día en que lo único que me queda es agradecer a todos los que me han ayudado y han recorrido este camino junto a mi, ayudándome a levantarme una y otra vez. Muchísimas gracias por cada abrazo, cada palabra de ánimo, cada palmada en el hombro. Gracias por preocuparos por mi, por alegraros con cada pequeño éxito, por consolarme de cada pequeño fracaso. He sido incapaz de llevar la cuenta, pero seguro que suman más de trescientas las veces que habéis tenido palabras amables o un gesto de cariño y comprensión para mi.

Índice general

Resumen

Los cromosomas B y el conjunto de cromosomas estándar (A) experimentan una sucesión en el tiempo y en el espacio de diferentes relaciones coevolutivas, caracterizadas por la existencia de mecanismos de acumulación de los cromosomas B y la respuesta del hospedador para suprimir estos mecanismos y neutralizar los efectos deletéreos de los Bs.

El saltamontes *Eyprepocnemis plorans* presenta un sistema de cromosomas B muy polimórfico, con más de 50 variantes descritas. En la Península Ibérica, las principales variantes de cromosoma B (B1, B2 y B5) se encuentran en un estadio de evolución neutra o cercana a la neutralidad, sin mecanismos aparentes de acumulación, como resultado de esta coevolución con el hospedador. La aparición, en los años ochenta, de la variante B24 en la población española de Torrox (Málaga), derivada del B2, y que mostraba tasas de transmisión superiores a las mendelianas, evidenció la capacidad de regeneración de los mecanismos de acumulación de los cromosomas B.

De esta forma, las acciones antagonistas de los cromosomas parásitos y el genoma estándar, que se reflejan en esta dinámica evolutiva, sugieren la existencia de una interacción al nivel molecular, probablemente durante la expresión génica. Es aquí donde reside el principal objetivo de esta Tesis Doctoral, en la que trataremos de averiguar si los cromosomas B de esta especie contienen genes que codifican para proteínas, si están activos y si alteran los patrones de expresión génica del genoma hospedador, ya sea en respuesta a su mera presencia o a los productos génicos de los cromosomas B.

En primer lugar, hemos comprobado que la presencia de cromosomas B no provoca una mayor tasa de dobles roturas y reparación de ADN, pero hemos observado que la proteina Ku70, implicada en la reparación no homóloga del ADN, muestra un comportamiento único en *E. plorans* al localizarse en la región centromérica de todos los cromosomas mitóticos y meióticos (incluido el cromosoma B) durante metafase y anafase. Mediante inmunofluorescencia combinada con tratamientos con colchicina y ARN de interferencia, demostramos que la presencia de Ku70 en los centrómeros tiene relación con la función centromérica, ya que los *foci* centroméricos desaparecían con estos dos tratamientos y ello estaba asociado al aumento en la frecuencia de espermátidas poliploides, lo que sugiere que la citocinesis en esta especie requiere la presencia de Ku70 en los centrómeros. Esta nueva función para esta proteína parece ser una autapomorfía en *E. plorans*, puesto que no se observaron *foci* centroméricos en otras catorce especies de saltamontes ni en el ratón. No encontramos, sin embargo, ninguna evidencia de que esta función estuviera asociada con la presencia de cromosomas B en esta especie.

Entre los efectos de los cromosomas B, conocidos previamente en *E. plorans*, destaca el descenso de los niveles de proteína del estrés Hsp70 en gónadas de machos y hembras portadores de cromosomas B, que parecía tener alguna relación con el grado de parasitismo de los cromosomas B, ya que el efecto era dependiente de la dosis (es decir, del número de Bs) y del estado de neutralización de los mismos, ya que el descenso era más acusado para la variante B24 (en Torrox, Málaga) que para la variante B2 (en Salobreña, Granada), variantes que se encuentran en estado parasítico y neutralizado, respectivamente. Uno de los objetivos propuestos en esta Tesis Doctoral fue averiguar si este descenso en los niveles de Hsp70

se observa ya al nivel transcripcional o bien tiene su origen en una regulación post-transcripcional. Para abordarlo, analizamos, mediante PCR cuantitativa (qPCR), los niveles de transcripción del gen *Hsp70* en 80 machos y hembras de estas mismas dos poblaciones, tanto en gónadas como en el resto del cuerpo, en su conjunto. La ausencia de diferencias significativas entre individuos portadores y no portadores de cromosomas B, en casi todas las muestras, sugiere que el descenso en Hsp70, asociado a la presencia de cromosomas B, no se debe a cambios en la actividad transcripcional del gen sino a mecanismos de regulación post-transcripcionales, un hecho que parece ser frecuente en poblaciones expuestas a un estrés continuado.

El carácter predominantemente heterocromático de los cromosomas B, y la gran cantidad de ADN repetitivo que acumulan, han hecho creer durante años que los cromosomas B son elementos transcripcionalmente inertes. También se asumía que estos cromosomas no portaban secuencias génicas debido al característico proceso de degeneración que sufren por la ausencia de recombinación con los cromosomas A. Sin embargo, en los últimos años, las nuevas técnicas de secuenciación de alto rendimiento han permitido identificar la presencia de genes que codifican para proteínas, o versiones pseudogénicas de éstos, en los cromosomas B de algunas especies, revelando incluso su transcripción activa en algunos casos. Con el fin de obtener pruebas adicionales sobre este interesante aspecto de la investigación sobre los cromosomas B, hemos realizado una búsqueda de genes para proteínas en el cromosoma B24 de *E. plorans*, mediante el mapeo de secuencias Illumina de ADN genómico, procedentes de un macho 0B y otro 4B de Torrox, sobre las secuencias codificadoras extraídas de un transcriptoma ensamblado *de novo* con

lecturas de ARNseq realizado en esta misma población. El análisis comparativo de la cobertura diferencial en 0B y 4B reveló la presencia de, al menos, nueve genes para proteínas en el cromosoma B24 de *E. plorans*, que validamos posteriormente mediante qPCR al demostrar ésta que la abundancia genómica de estos genes aumentaba linealmente con el número de cromosomas B. Cuatro de estos genes (*CIP2A*, *GTPB6*, *KIF20A* y *MTG1*) mostraron su secuencia completa en el genoma 4B, mientras que los cinco restantes (*CKAP2*, *CND3*, *HYI*, *MYCB2* y *SLIT*) estaban truncados. El análisis de las lecturas de RNAseq de dos hembras de Torrox, una 0B y la otra 1B, y la subsiguiente confirmación por qPCR, mostró que cinco de ellos (dos completos y tres truncados) estaban sobre-expresados proporcionalmente al número de cromosomas B, tanto en gónadas como en el resto del cuerpo de machos y hembras, tal como se esperaría si el cromosoma B estuviera transcribiendo activamente estos genes. De hecho, los tres genes truncados (*CKAP2*, *CND3* y *MYCB2*) no mostraban esta sobrexpresión cuando la qPCR se realizaba con cebadores que anclaban en las regiones ausentes del cromosoma B. Por tanto, nuestros resultados demuestran que los cromosomas B de *E. plorans* no están tan silenciados como se creía hasta ahora. Cabe destacar que los cinco genes del cromosoma B que se transcriben, dos completos (*CIP2A* y *KIF20A*) y tres truncados (*CKAP2*, *CND3* y *MYCB2*) se han asociado con funciones potencialmente cruciales para la transmisión del cromosoma B, tales como la segregración y estabilidad cromosómicas, la citocinesis y el ensamblaje de los microtúbulos en la transición metafase-anafase. En conjunto, estos resultados permiten hipotetizar que el éxito evolutivo de los cromosomas B en esta especie se debe a su contenido génico.

Una vez demostrada la presencia de genes para proteínas en el cromosoma B, nos interesó averiguar la posible funcionalidad de los transcritos codificados por uno de esos genes, concretamente el gen que codifica para la subunidad *CAP-G* de la condensina I. La hibridación *in situ* fluorescente mediante amplificación de la señal con tiramidas (FISH-TSA), ha permitido visualizar la presencia de varias copias en tandem localizadas en la región distal del cromosoma B24. Hemos comprobado mediante qPCR, que hay también varias copias truncadas de este gen en el cromosoma B de la población de Salobreña, que es la variante B2. Además, en ambas poblaciones (Torrox y Salobreña), hemos observado la transcripción activa de la variante pseudogénica de *CAP-G* situada en el cromosoma B, pero esta sobre-expresión no afectaba a la expresión de los genes que codifican para las subunidades *CAP-D2* y *CAP-D3* de la condensina I y II, respectivamente. Esto indica que el exceso de transcritos de *CAP-G* procedente de los cromosomas B no está aumentando la actividad general de las condensinas. Por último, discutimos los posibles efectos, a varios niveles, del exceso de transcritos pseudogénicos.

Finalmente, nos interesaba averiguar cuál es la respuesta del genoma hospedador a la presencia de cromosomas B que transcriben al menos parte de su contenido génico. Mediante ARNseq y microarrays realizados en hembras con y sin cromosomas B de Torrox y Salobreña, respectivamente, hemos obtenido algunas respuestas a esta pregunta. En total, observamos 188 unigenes que mostraban cambios de expresión en el mismo sentido con ambas técnicas de análisis y, por tanto, en ambas poblaciones. Descartando los cambios que podían deberse a la expresión de los genes localizados en el cromosoma B, encontramos

46 genes que codifican para proteínas conocidas, cuyos cambios de expresión podrían estar relacionados con la adaptación del genoma hospedador a la presencia del cromosoma parásito. Estos cambios de expresión tenían relación con algunos de los efectos más conocidos de los cromosomas B, tales como los efectos nucleotípicos derivados de la presencia del ADN adicional que aportan los cromosomas B, la defensa química y la detoxificación, la modificación de proteínas, la respuesta al estrés, la función ovárica y la regulación de la expresión génica. Además, encontramos, asociado a la presencia de cromosomas B, un incremento en la expresión de muchos elementos transponibles, que era paralelo a la represión del gen *Dicer1*, y que podemos interpretar como un incremento en la actividad de éstos en presencia de cromosomas B, debido al debilitamiento de la defensa contra los elementos transponibles que proporciona *Dicer1* a través de ARN de interferencia.

En conjunto, los resultados de esta Tesis Doctoral han conseguido elucidar nuevos detalles de la relación entre los cromosomas A y B en *Eyprepocnemis plorans*, sugiriendo la interesante posibilidad de que el secreto del éxito de los cromosomas B resida en su contenido génico, su capacidad para transcribirlo y, por tanto, su habilidad para evitar el silenciamiento por parte de los cromosomas A. Esto convierte a los cromosomas B en verdaderos parásitos intragenómicos en íntima interacción transcripcional con el genoma hospedador, una visión completamente nueva para estos interesantes y enigmáticos elementos genómicos.

Summary

Supernumerary (B) and standard (A) chromosomes undergo a spatiotemporal succession of coevolutionary relationships characterized by the existence of accumulation mechanisms (drive) for B chromosomes and a variety of A chromosome responses directed toward drive suppression and diminish harmful effects.

The grasshopper *Eyprepocnemis plorans* shows a very polymorphic system of B chromosomes, as more than 50 variants have been hitherto described. As a result of their coevolutionary interaction with A chromosomes, the main variants of B chromosomes in the Iberian Peninsula (B1, B2 and B5) have been neutralized and thus lack apparent drive. However, a new variant (B24) arisen in the Spanish population of Torrox (Málaga), derived from B2, still showed drive when analyzed in the eighties. The appearance of this new variant proved the ability of these B chromosomes to recuperate drive mechanisms which had been lost in the ancestor neutralized B chromosome.

The antagonistic relationship between the parasitic chromosomes and the standard genome is revealed by their coevolutionary dynamics, and suggests the existence of an interaction at molecular level, probably through gene expression changes. At this point lies the main objective of the present PhD Thesis. Specifically, we will try to find out whether B chromosomes in this species harbor protein-coding genes, whether they are actively transcribed and whether they alter the expression of the genes located in the host genome, either in response to mere B presence or as consequence to B chromosome gene products.

First, we have observed that the presence of B chromosomes is not associated with higher levels of double strand breaks and repair, but we incidentally observed that the Ku70 protein, which is involved in the non-homologous end joining pathway of DNA repair, localizes in the centromeric region of all A and B chromosomes during mitosic and meiotic metaphase and anaphase, but this is a feature unique for *E. plorans*. By means of inmunofluorescence, we showed that Ku70 presence at centromeres is related with centromeric function, as the disappearance of the Ku70 *foci* from centromeres, after colchicine and interference RNA treatments, was associated with an increased frequency of polyploid spermatids, suggesting that Ku70 centromeric presence is required for cytokinesis in this species. This novel function for the Ku70 protein is an autapomorphy in *E. plorans*, as we did not observe centromeric *foci* in other fourteen grasshopper species or the mouse. Nevertheless, we did not find any evidence for the association between this function and the presence of B chromosomes in this species.

Among the previously known B chromosome effects in *E. plorans*, we can highlight the descent of cellular levels for the Hsp70 stress protein in gonads of B-carrying males and females. This effect appeared to be related with the parasitic degree of B chromosomes, as it depended on the number of B chromosomes and also on their evolutionary stage, as the descent was more marked for the parasitic B24 variant (from Torrox, Málaga) than for the neutralized B2 (from Salobreña, Granada). One of the objectives in this PhD Thesis was to find out whether this decrease in Hsp70 levels occurs at transcriptional level, or else is due to post-transcriptional regulation processes. To address this objective, we analyzed, by means of qPCR, the Hsp70 gene expression levels in 80

males and females from the same two populations mentioned above, in gonads and in the somatic body. The absence of significant differences between B-carrying and B-lacking individuals in almost all samples analyzed suggests that the descent of Hsp70 protein levels associated to B chromosome presence is not due to changes in the transcriptional activity of the gene, but to mechanisms of post-transcriptional regulation, a pattern that seems to be usual in populations submitted to a continuous source of stress.

The heterochromatic nature of B chromosomes and the high amounts of repetitive DNA they accummulate, have led researchers to believe that B chromosomes are transcriptionally inert. Furthermore, a typical feature of B chromosomes is that they do not recombine with A chromosomes, which leads them towards a process of chromosome degeneration. Therefore, it was assumed for long that B chromosomes do not harbor gene-derived sequences. However, during the last decade, the development of new high throughput sequencing techniques have allowed the identification of protein-coding genes, or pseudogenized copies, in the B chromosomes of some species, revealing even active transcription in some cases. Aiming to obtain additional evidence about this attractive facet of B chromosome research, we performed a search for protein-coding genes in the B24 chromosome of *E. plorans*, by means of Illumina sequencing of DNA and RNA. We thus sequenced genomic DNA from one 0B and one 4B males collected at Torrox, and obtained two libraries of Illumina reads which were mapped on a set of coding sequences, from a *de novo* transcriptome resulting from the assembling of RNAseq reads obtained from 0B and 1B females from this same population. The comparative analysis of differences in coverage for 0B

and 4B libraries revealed that at least nine protein-coding genes are located in the B24 chromosome of *E. plorans*. All of them were validated by means of qPCR, showing that the genomic abundance of these genes increased linearly with the number of B chromosomes. Four of these genes (*CIP2A*, *GTPB6*, *KIF20A* y *MTG1*) showed a complete coding sequence (CDS) in the 4B genome, whereas the five remaining (*CKAP2*, *CND3*, *HYI*, *MYCB2* and *SLIT*) were truncated. Mapping the RNA reads from the 0B and 1B females, on the same reference transcriptome, revealed that five of the former genes (two complete and three truncated) showed significant up-regulation which was proportional to the number of B chromosomes, in both gonads and bodies from males and females, as expected if that transcripts came directly from active expression of those genes. Indeed, in the case of truncated genes, we did not find differential expression when the qPCR was carried out with primers anchoring on the regions that were absent from the B chromosome. Therefore, our results demonstrate that B chromosomes in *E. plorans* are not so silenced as was hitherto believed. It is interesting to emphasize that the five genes that are actively transcribed, two complete (*CIP2A* and *KIF20A*) and three truncated (*CKAP2*, *CND3*, and *MYCB2*) are associated with functions being potentially crucial for B chromosome transmissions, such as chromosome segregation and stability, cytokinesis and microtubule assembly in metaphase-anaphase transition. As a whole, this result leads to hypothesize that the evolutionary success of B chromosomes in this species is a consequence of its genomic content and the ability to avoid silencing mechanisms of the host.

Once the presence of protein-coding genes in the B chromosome was demonstrated, we were interested in finding out the possible functionality

of the transcripts coded by one of these genes, specifically the gene coding for the condensin complex CAP-G subunit. We compared the aminoacid sequence of *CAP-G* from many organisms, and determined that the predicted protein coming from the transcripts yielded by the B chromosome would show at least two changes becoming it probably non-functional, and that the *CAP-G* gene copies in the B chromosome are pseudogenic. Fluorescence *in situ* hybridization with tyramide signal amplification (FISH-TSA) allowed to visualize the *B-CAP-G* pseudogene as several tandemly repeated copies located in the distal region of the B24 chromosome. By means of qPCR, we showed that, in two Spanish populations (Salobreña and Torrox, with B2 and B24 variants, respectively), the *B-CAP-G* pseudogene is actively transcribed, without apparent correlation with the expression of *CAP-D2* and *CAP-D3* subunits of condensin I and II, respectively. This indicates that the B chromosome transcripts are not enhancing the activity of the condensin complex. Possible effects of the *B-CAP-G* transcripts at several levels are discussed.

Finally, we were interested in finding out details on how the host genome responds to the presence of a B chromosome being able to transcribe part of its genetic content. By means of RNAseq and microarrays performed in females with and without B chromosomes from Torrox and Salobreña, respectively, we got some answers to this topic. We identified a total of 188 unigenes showing gene expression changes in the same direction for both techniques, and thus in both populations. Discarding changes that could be due to the expression of the genes located in the B chromosome, we found 46 known protein-coding genes, whose expression changes might be associated with the adaptation of the host genome to the

presence of a parasitic chromosome. These gene expression changes were related to some of the most known effects of B chromosomes, such as nucleotypic effects (i.e., effects derived from the increase in DNA amount due to the presence of the extra chromosomes), chemical defense and detoxification, protein modification and turnover, stress response, ovarian function and gene expression regulation. Moreover, we found, associated to the presence of B chromosomes, an increment in the expression of a high number of different transposable elements, which was parallel to the decrease on the expression of the *Dicer1* gene. We interpret this as a signal of increased transpositional activity in the presence of B chromosomes, derived from a weaker defense against transposable elements throught RNAi caused by Dicer1 down-regulation.

As a whole, the results of this PhD Thesis have untangled new details about the transcriptional relationship between B and A chromosomes in *Eyprepocnemis plorans*, suggesting the interesting possibility that the secret of the evolutionary success of B chromosomes may lie on their genetic content and their ability to avoid the mechanisms of silencing developed by the host, in order to transcribe their gene content. This results transform the current paradigm of B chromosomes and become them into true intragenomic parasites being in an intimate relationship at transcriptional level with the host genome, and providing new insights for future research on these interesting and enigmatic genomic elements.

Introducción general

Cualquier sistema altruista es,
inherentemente, inestable, ya que esta
sujeto al abuso por parte de individuos
egoístas, dispuestos a explotarlo.

Richard Dawkins

Los cromosomas B

El conflicto intragenómico

Los cromosomas B (también denominados accesorios o supernumerarios) son cromosomas adicionales cuya presencia en el genoma no es necesaria para desarrollar el ciclo biológico normal de los individuos portadores, y que no recombinan con los cromosomas del complemento estándard (también denominados cromosomas A) (Jones, 1995). Además, los cromosomas B se caracterizan por mostrar un comportamiento mitótico y/o meiótico irregular que les permite acumularse en la línea germinal y transmitirse a la descendencia a tasas que exceden la de los cromosomas A (Camacho et al., 2000).

Los cromosomas B constituyen el primer ejemplo descrito de los llamados elementos genéticos egoístas: elementos que obtienen una ventaja relativa en la transmisión respecto al resto del genoma, pero que son neutrales o perjudiciales para el organismo en su conjunto (Werren et al., 1988).

El descubrimiento de los cromosomas supernumerarios data de principios del siglo XX y fue debido a los estudios de Wilson (1907) en el hemíptero Metapodius (ahora llamado Acanthocepala). El término "cromosomas B" no fue utilizado hasta 1928 por Randolph, quien acuñó este término en contraposición con el de "cromosomas A" para referirse al complemento estándar (Randolph, 1928). Fue Östergren (1945), investigando la acumulación de cromosomas B en plantas, el primero en argumentar que a un elemento de estas características se le podía considerar parásito. Sin embargo, estas observaciones y sus implicaciones evolutivas pasaron bastante inadvertidas en su momento, principalmente porque tanto los cromosomas B como los elementos genéticos egoístas se percibían como peculiaridades más que como un fenómeno generalizado. Más tarde, Kimura y Kayano (1961) propusieron un elegante modelo para el mantenimiento de los cromosomas B en las poblaciones de *Lilium callosum*, basado en el impulso meiótico femenino.

La naturaleza egoísta de los cromosomas B no fue reconocida hasta cuarenta años después (Jones, 1985), cuando la existencia de elementos genéticos egoístas comenzó a considerarse más seriamente, principalmente por tres causas: la visión del gen como unidad de selección (Dawkins, 1976), el descubrimiento de la existencia en los genomas eucariotas de ingentes cantidades de ADN repetido sin una función clara (Doolittle y Sapienza, 1980) y un creciente número de estudios que, en diferentes organimos, describían elementos genéticos con comportamiento

egoísta, tales como los elementos transponibles, los distorsionadores de la segregación, factores citoplasmáticos, como microbios y orgánulos, que distorsionan la determinación sexual, y los propios cromosomas B (para revisión, ver Werren, 2011).

La invasión y propagación de los elementos genéticos egoístas, y la aparición de genes de resistencia que aparecen por selección en el genoma invadido, evidencian un conflicto de intereses entre componentes de un mismo genoma. Éste es el llamado conflicto intragenómico, una potente fuerza evolutiva cuyo estudio ha servido para explicar procesos clave en la historia de la biología, tales como el origen del sexo y los sistemas de determinación sexual, la diploidía y la herencia uniparental (Werren et al., 1988; Hurst, 1992; Burt y Trivers, 2006).

La mayoría de los cromosomas B se ajustan al modelo egoísta-parasítico, en el que el conflicto intragenómico se desarrolla como una constante "carrera de armamentos" coevolutiva entre el cromosoma parásito y el genoma hospedador (Östergren, 1945). Sin embargo, también se han descrito sistemas de cromosomas B donde los mecanismos de acumulación han sido neutralizados por el genoma hospedador (modelo casi-neutro, Camacho et al., 1997), e incluso cromosomas B que pueden ser ventajosos para el hospedador cuando se encuentran en número bajo, pero son perjudiciales en números elevados (modelo heterótico, White, 1973).

Tamaño, forma y frecuencia

No existe una morfología estándar para los cromosomas B. Suelen ser del mismo tamaño que los cromosomas del complemento, aunque

existen ejemplos de cromosomas B más grandes que los cromosomas A (Baverstock et al., 1982; Mestriner et al., 2000) o más pequeños (Peppers et al., 1997), incluso hay casos en que se reducen escasamente a un centrómero (Wolf et al., 1991). Hewitt (1979) observó que, generalmente, los cromosomas B de tamaño grande o mediano suelen ser mitóticamente estables, es decir, aparecen en igual número en todas las células del individuo portador, mientras que los cromosomas B pequeños suelen ser mitóticamente inestables y los individuos portadores suelen ser mosaicos numéricos. La posición del centrómero en los cromosomas supernumerarios también es variable, aunque parece tener cierta tendencia a ser la misma que la de los cromosomas A (John y Hewitt, 1965; López-León et al., 1993; Palestis et al., 2004).

Dependiendo de cada sistema, el grado de polimorfismo estructural de los cromosomas B puede ser bastante variable. Por ejemplo, en *Brachycome dichromosomatica* solamente se encuentran dos tipos (Houben et al., 1999), mientras que, en el otro extremo, existen sistemas de cromosomas B altamente polimórficos, como el de *Eyprepocnemis plorans*, donde se han descrito más de 50 variantes (López-León et al., 1993; Cabrero et al., 1997).

Los cromosomas B se han encontrado formando parte del genoma de algunos individuos en algunas poblaciones de unas 10 especies de hongos, 1300 de plantas y en más de 500 especies animales (Camacho, 2005). Se estima que un 15 % del total de especies presentan cromosomas B (Jones y Rees, 1982; Camacho et al., 2000). Su frecuencia en las poblaciones depende fundamentalmente de factores selectivos (e.g. selección negativa contra individuos con alto número de cromosomas B, desarrollo de mecanismos de resistencia en los hospedadores, tasa de acumulación del

cromosoma B o condiciones ambientales), y aleatorios (tasas de mutación y deriva genética), así como del estadio evolutivo en el que se encuentre el polimorfismo y de factores históricos como, por ejemplo, el tiempo transcurrido desde la llegada del cromosoma B a la población (Camacho et al., 2000).

El número de cromosomas B que porta cada individuo es variable atendiendo a diversos factores, tales como el tipo de cromosoma B, su virulencia, las características de su mecanismo de acumulación y la tolerancia del genoma hospedador. Salvo excepciones, en las poblaciones naturales no se encuentran individuos con más de tres o cuatro cromosomas B. La existencia de un límite superior para el número de cromosomas B sugiere que un número elevado de estos cromosomas puede ser perjudicial para los individuos portadores (Camacho et al., 2000).

Origen y evolución molecular

En la mayoría de los casos, se desconocen los mecanismos por los que puede surgir un nuevo cromosoma extra (llamémosle "neo-B") que adquiera las características propias de los cromosomas B. Existen dos modelos generales para explicar el origen de los cromosomas B: el modelo del origen intraespecífico, que postula que los cromosomas B derivan de los autosomas o de los cromosomas sexuales de la misma especie donde se encuentran (Jones y Rees, 1982), y el modelo del origen interespecífico, que propone que los cromosomas B surgen durante hibridaciones interespecíficas, por lo que derivarían de cromosomas de

una especie afín (Battaglia, 1964).

Ambos modelos coinciden en que para que se origine un cromosoma B es necesario un sustrato, que sería un cromosoma A en condición polisómica procedente de la misma especie, o de una especie diferente cuando es por hibridación, o bien un fragmento céntrico resultante de una fragmentación cromosómica. El cromosoma extra degenerará por su aislamiento meiótico respecto a los cromosomas A (es decir, por no recombinar con ellos), acumulando secuencias repetidas y transposones (Camacho et al., 2000).

Los neo-B de origen intraespecífico pueden emerger espontáneamente en muchas especies eucariotas como un subproducto de la evolución normal del cariotipo (Camacho et al., 2000). Sin embargo, para su incremento en frecuencia y mantenimiento, ese neo-B debería mostrar, desde el primer momento, algún tipo de mecanismo de acumulación porque, de lo contrario, sería rápidamente eliminado (Camacho et al., 1997). Como evidencia de su procedencia intraespecífica, algunos cromosomas B contienen secuencias de ADN similares a las de los cromosomas A de la misma especie, tal como se ha descrito en *Locusta migratoria*, donde la localización y la secuencia de los genes de las histonas H3 y H4 indican que el cromosoma B probablemente derivó del cromosoma 8 (Teruel et al., 2010). Otros ejemplos de cromosomas B de origen intraespecífico se han descrito en la planta *Crepis capillaris* (Jamilena et al., 1994) y la mosca *Drosophila subsilvestris* (Gutknecht et al., 1995).

Otro posible modo de origen de los cromosomas B es por amplificación de regiones pericentroméricas tras la fragmentación de un cromosoma A que proporcione un centrómero, como se ha propuesto en la planta

Plantago lagopus cuyo cromosoma B está compuesto mayoritariamente por repeticiones de ADNr 5S (Dhar et al., 2002).

Cuando el cromosoma B se origina como resultado de un cruzamiento entre especies afines (origen interespecífico), el cromosoma extra introgresado no recombina con los cromosomas hospedadores y degenera por su aislamiento genético. Una posible evidencia en favor de este modelo sería encontrar secuencias específicas del cromosoma B de una especie, que están ausentes en el genoma hospedador, y que guardan homología con el genoma de alguna especie emparentada. En el caso del cromosoma B de *Nasonia vitripennis*, denominado PSR ("Paternal Sex Ratio") por sus efectos sobre la proporción de sexos, un análisis filogenético reveló que las secuencias del transposón NATE del cromosoma PSR son más similares a las copias presentes en especies relacionadas del género Trichomalopsis que a las copias existentes en el propio genoma de *N. vitripennis* (McAllister y Werren, 1997). El origen interspecífico de cromosomas neo-B se ha demostrado empíricamente en Nasonia, durante experimentos de introgresión interespecífica en el laboratorio (Perfectti y Werren, 2001).

El proceso de degeneración y diferenciación de los cromosomas supernumerarios, tras su origen, se produce mediante procesos similares a los implicados en la degeneración evolutiva de los cromosomas sexuales heteromórficos. Por tanto, puede explicarse por el trinquete de Muller, que predice que, en ausencia de recombinación, se produce una acumulación irreversible de mutaciones que lleva a la pérdida de loci funcionales y de la homología con las secuencias del cromosoma progenitor, así como a la ganancia de heterocromatina (Green, 1990). En el caso de los cromosomas B generados a partir de un cromosoma

de la misma especie, este proceso debe ser bastante rápido, ya que en un principio el neo-B podría tener suficiente homología para recombinar con sus progenitores, por lo que la degeneración por aislamiento no se produciría. En los casos en los que se origina a partir de un fragmento cromosómico de otra especie, éste estaría directamente aislado (Camacho et al., 2000).

Estudios recientes en centeno (Martis et al., 2012) han propuesto el origen del B de esta especie como un subproducto de duplicaciones y translocaciones entre autosomas, concretamente el 3R y 7R, seguidos por inserciones de secuencias procedentes del genoma A y de los orgánulos, así como de la amplificación de repeticiones específicas en el cromosoma B. Los genes presentes en el proto-B serían en un principio silenciados por un mecanismo de compensación de dosis, y degenerarían rápidamente por la acumulación de mutaciones e inserciones de esas secuencias procedentes del genoma A y de los orgánulos, excepto la región donde se encuentran los factores desconocidos (codificantes o no) que facilitaron la invasión y el mantenimiento de este cromosoma B en las poblaciones naturales.

Composición molecular

Los cromosomas B son típicamente heterocromáticos, debido a que están compuestos en su mayoría por ADN repetido, principalmente ADN satélite y elementos transponibles (Camacho, 2005). Por lo general, su contenido heterocromático es similar al de los cromosomas A (Jamilena et al., 1994; López-León et al., 1994; Jones y Houben, 2003), pero se han descrito ADNs satélites específicos de los cromosomas B en

Nasonia vitripennis (Nur et al., 1988) y el centeno (Sandery et al., 1990; Martis et al., 2012), y una modificación del IGS del ADN ribosómico en *Eyprepocnemis plorans* (Muñoz-Pajares et al., 2011).

Al ser dispensables y no recombinar con el resto del genoma, los cromosomas B se convierten en un lugar susceptible de acumulación de elementos móviles (Camacho et al., 2000). Además, hay estudios que apuntan a que los transposones muestran cierta tendencia a insertarse directamente en la heterocromatina (Dimitri et al., 2003). Se ha descrito también la presencia de ADN telomérico en cromosomas B, en posición intersticial en algunos casos, como los de *Nyctereutes procyonides* (Wurster-Hill et al., 1988) y *Nectomys squamipes* (Silva y Yonenaga-Yassuda, 1998), sugiriendo que estos cromosomas se han originado mediante reordenaciones cromosómicas (Camacho, 2005).

Asimismo, los cromosomas B frecuentemente contienen ADN ribosómico 45S que suele estar inactivo (Cabrero et al., 1987). Sin embargo, la presencia de genes ADNr 5S sólo se ha detectado en algunos cromosomas B como, por ejemplo, los de *Plantago lagopus* (Dhar et al., 2002) y *Eyprepocnemis plorans* en poblaciones del Cáucaso (Cabrero et al., 2003).

Durante mucho tiempo se ha afirmado que los cromosomas B no contienen genes (Camacho et al., 2000; Jones y Houben, 2003; Burt y Trivers, 2006). Esto se debe, probablemente, a la dificultad de detectarlos de forma eficiente, mediante técnicas citogenéticas, y a la gran abundancia de elementos repetidos que puede ocultar cualquier señal de genes de copia única o en bajo número de copias. Además, debido a la dispensabilidad, que hace que la presión selectiva para la conservación

de estas secuencias sea casi nula, y a la ausencia de recombinación, el modelo de evolución molecular de los cromosomas B predice que las secuencias génicas localizadas en éstos, que serán fundamentalmente residuos del cromosoma A ancestral del que se originó el cromosoma B, mostrarán un gran número de mutaciones respecto a sus homólogas en los cromosomas A y estarán muy fragmentadas por inserciones de transposones (Green, 1990).

Las primeras evidencias de que un cromosoma B contenía genes, que además eran activos, vinieron de una asociación entre la presencia de cromosomas B y un cambio fenotípico observable, como en el caso de los genes para la resistencia a la roya en la planta *Avena sativa* (Dherawattana y Sadanaga, 1973) y a la pisatina en el hongo *Nectria haematococca* (Miao et al., 1991).

Pero la primera demostración molecular de la presencia de genes para proteínas en los cromosomas B tuvo que esperar hasta que Graphodatsky et al. (2005) observaron, mediante FISH, la existencia de copias del oncogen C-KIT en los cromosomas B de varias especies de cánidos. Posteriormente se observó, también mediante FISH, la presencia de genes para histonas en el cromosoma B de *Locusta migratoria* (Teruel et al., 2010). En el cromosoma B del maíz se ha confirmado recientemente, también mediante FISH, la existencia de al menos 4 genes (Huang et al., 2016).

Durante los últimos años, el desarrollo de las técnicas de secuenciación de nueva generación (Next Generation Sequencing, NGS) ha permitido una caracterización mucho más profunda del contenido de los cromosomas B. Martis et al. (2012) describieron el contenido del cromosoma B del

centeno como un mosaico de secuencias derivadas de genes procedentes del genoma hospedador (identifican casi 5.000), fragmentos de ADN mitocondrial y cloroplastidial y una gran cantidad de ADN repetido, en su mayoría invadido por elementos transponibles Ty1/copia y un satélite específico del B de monómeros extraordinariamente largos (entre 900 y 4000 pares de bases), probablemente originado por quimerismo entre otras secuencias repetidas.

Por otra parte, Valente et al. (2014) analizaron el contenido de los cromosomas B del pez *Astatotilapia latifasciata* comparando secuencias Illumina de genomas 0B y 2B, complementado por una secuenciación 454 de un cromosoma B microdiseccionado. Mapearon estas secuencias contra scaffolds del genoma de *Metriaclima zebra*, y detectaron bloques donde la relación entre la cobertura B+/B- era mayor que el promedio que se observaba en los scaffolds que no estaban localizados en el cromosoma B. Estos bloques fueron anotados con el genoma de *M. zebra*, y se encontró entre ellos más de 5.000 secuencias que podrían ser genes, además de muchos elementos transponibles y ADN repetitivo. Algunos de los genes localizados en el cromosoma B estaban asociados con funciones como la organización de los microtúbulos y la estructura del cinetocoro, recombinación y progresión del ciclo celular, apuntando a que parte del contenido del B podría estar involucrado en la transmisión y mantenimiento del cromosoma supernumerario.

Actividad transcripcional

La heterocromatinización se puede considerar como una manifestación citológica del silenciamiento génico (Zhimulev y Beliaeva, 2003). Esto, junto con el alto contenido en ADN repetido y la dispensabilidad de los cromosomas B, ha llevado durante mucho tiempo a su consideración como elementos genéticos transcripcionalmente inertes. Los experimentos de incorporación de uridina tritiada en espermatocitos del ratón *Apodemus peninsulae* (Ishak et al., 1991) y en los saltamontes *Myrmeleotettix maculatus* y *Chorthippus parallelus* (Fox et al., 1974) apoyaban esta idea, indicando que el nivel de transcripción en los cromosomas B es muy bajo o inexistente.

Son varios los casos donde se ha reseñado que la presencia de cromosomas B puede ejercer una cierta influencia sobre la expresión de los A, aunque ésta no se derive de la actividad génica del propio cromosoma B. Por ejemplo, en *Scilla autumnalis*, las plantas con cromosoma B expresan un gen que codifica para una variante de esterasas que no está activa en las plantas sin B (Ruiz-Rejón et al., 1980; Oliver et al., 1982). Algo similar ocurre en *Allium schoenoprasum*, donde la presencia de cromosomas B influye sobre la expresión de un gen que codifica para una proteína del endospermo (Plowman y Bougourd, 1994). En *Apodemus flavicolis* se ha observado la expresión diferencial de algunos genes en relación con la presencia de cromosomas B, entre ellos una chaperona molecular (Tanić et al., 2005). En el caso de que el cromosoma B estuviera totalmente silenciado, estos efectos podrían explicarse por alteraciones en la organización del genoma hospedador y los cromosomas B en los núcleos interfásicos, ya que la expresión génica parece estar influenciada

por la posición de los genes y los cromosomas en el núcleo (Misteli, 2007) y por la estructura de la cromatina (Wallrath y Elgin, 1995; Zhimulev y Beliaeva, 2003). De hecho, en el centeno, se ha visto que la presencia de cromosomas B induce cambios de compactación en ciertas regiones de los cromosomas A (Delgado et al., 2004).

En estudios recientes, comparando transcriptomas de individuos con y sin cromosoma B, se ha visto, al menos en dos casos, que la presencia de cromosomas B está asociada con cambios de expresión génica. Sin embargo, los resultados son dificiles de relacionar con los conocimientos que se tienen sobre los cromosomas B. Por ejemplo, en *Nasonia vitripennis*, un análisis comparativo de transcriptomas de testículos de machos WT (wild type) y PSR+ (portador de cromosoma PSR), asoció la presencia del cromosoma supernumerario con algunos cambios de expresión génica, pero ninguno de ellos estaba relacionado con los efectos descritos previamente para este cromosoma, tales como la condensación cromosómica, la estructura de la cromatina y los elementos transponibles (Akbari et al., 2013). En el maíz, un estudio similar ha concluido que la presencia del cromosoma B provoca una serie de cambios en la expresión génica del hospedador, con efectos dependientes del número de cromosomas B. La mayoría de estos cambios estaban relacionados con el metabolismo celular y proteínas de unión a nucleótidos (Huang et al., 2016).

Sin embargo, se han descrito algunos casos donde sí se observa actividad en estos cromosomas. La primera evidencia indirecta de expresión de genes localizados en un cromosoma B viene de los casos en los que el cromosoma supernumerario está asociado con un cambio fenotípico observable, como los ya comentados genes de resistencia a la roya en

Avena sativa (Dherawattana y Sadanaga, 1973) y los de resistencia a la pisatina del hongo *Nectria haematococca* (Miao et al., 1991).

Se ha demostrado actividad génica en varios casos como, por ejemplo, el cromosoma B en estado politénico del mosquito *Simulium juxtacrenobium* (Brockhouse et al., 1989), y un neo-B en la avispa *Nasonia vitripennis* (Perfectti y Werren, 2001), y se ha observado también la transcripción de secuencias específicas de los cromosomas B de centeno y maíz, que podrían estar implicada en el mecanismo de acumulación de estos cromosomas B (Lamb et al., 2007; Carchilan et al., 2007, 2009), y de secuencias similares a elementos móviles en el cromosoma B del centeno (Carchilan et al., 2007). Más recientemente, la secuenciación del genoma de *Drosophila albomicans* ha permitido identificar un scaffold procedente del cromosoma B, que contiene secuencias que se expresan activamente (Zhou et al., 2012), y el análisis de transcriptomas en *Nasonia vitripennis* ha permitido identificar transcritos procedentes del cromosoma PSR (Akbari et al., 2013). También se ha detectado transcripción del ADN ribosómico de los cromosomas B de la planta *Crepis capillaris* (Leach et al., 2005), la avispa *Trichogramma kaykai* (Van Vugt et al., 2003) y el saltamontes *Eyprepocnemis plorans* (Cabrero et al., 1987; Teruel et al., 2009b; Ruiz-Estévez et al., 2012).

Hasta ahora, la única evidencia de funcionalidad inequívoca la aportan los genes para ARN ribosómico localizados en el cromosoma B del saltamontes *Eyprepocnemis plorans*, ya que su actividad está asociada con la presencia de un nucleolo asociado al cromosoma B, que constituye el fenotipo de estos genes. Aun así, en la mayoría de las poblaciones analizadas, el ADNr del B se encuentra activo en un porcentaje pequeño de machos (Ruiz-Estévez et al., 2013), y la contribución del B, sobre el

total de rRNA que se produce en la célula, es insignificante comparado con la de los cromosomas A (Ruiz-Estévez et al., 2014), lo que sugiere un alto grado de represión del cromosoma B.

Aunque son crecientes las evidencias de transcripción de secuencias localizadas en el cromosoma B, la presencia y actividad de genes funcionales siguen muy cuestionadas. El hecho de que los genes localizados en los cromosomas B estén fragmentados no implica que su expresión no tenga ningún impacto. Se ha observado, en numerosas ocasiones, que los pseudogenes que se transcriben pueden alterar la expresión génica de sus genes parentales actuando a muchos niveles, tales como modificaciones epigenéticas, generación de ARN de interferencia (siRNA, small interfering RNA) (Pink et al., 2011), inhibición competitiva en la unión de micro ARNs (Poliseno et al., 2010), del complejo de traducción o de otras proteínas de unión a ARN (Poliseno et al., 2015). Si llegan a traducirse, las proteínas pseudogénicas pueden afectar al funcionamiento de las proteínas parentales, incluso si no son completamente funcionales (para revisión, ver Poliseno et al., 2015). Por ello, en trabajos recientes se ha sugerido que la expresión del contenido del cromosoma B puede modular la expresión de genes localizados en el genoma estándar, sobre todo como reguladores en trans de la actividad génica, tal como se ha visto en centeno (Carchilan et al., 2009; Banaei-Moghaddam et al., 2013). Banaei-Moghaddam et al. (2013) compararon los fragmentos génicos que encontraron en el cromosoma B de centeno con sus correspondientes genes ancestrales en los A, confirmando su procedencia y la pseudogenización de los fragmentos génicos del B, y estudiaron los patrones de expresión génica, que eran, en algunos casos, dependientes de tejido y genotipo, concluyendo que es

posible que la expresión de genes localizados en el B esté afectando a la expresión regular de las células de centeno.

Efectos de los cromosomas B

En raras ocasiones la presencia de los cromosomas supernumerarios lleva asociado un cambio fenotípico observable. El efecto más drástico y llamativo es, sin duda, el que produce la presencia del cromosoma PSR de las avispas *Nasonia vitripennis* (Werren, 1991) y *Trichogramma kaykai* (Van Vugt et al., 2003), que debe su nombre al hecho de que su presencia en los cigotos diploides de esta avispa causa la condensación y disgregación del juego cromosómico paterno completo, convirtiendo lo que serían hembras diploides en machos haploides portadores de PSR. Otros casos, bastante claros, de efectos fenotípicos asociados a cromosomas B los encontramos, por ejemplo, en la planta *Happlopappus gracilis*, donde la presencia de un cromosoma B está asociada a un cambio en el color del fruto (Jackson y Newmark, 1960), en el maíz, donde se ha asociado la presencia de cromosomas B a la aparición de rayas blancas en las hojas (Staub, 1987) y en los anteriormente comentados genes de resistencia a la roya en *A. sativa* (Dherawattana y Sadanaga, 1973) y a la pisatina en *N. haematococca* (Miao et al., 1991). Más recientemente, se ha visto que la presencia de cromosomas B influye sobre la determinación del sexo en peces cíclidos (Yoshida et al., 2011).

Aunque existen varios ejemplos de cromosomas B cuya presencia en números bajos, aporta algún beneficio para el hospedador (Dherawattana y Sadanaga, 1973; Holmes y Bougourd, 1989; Miao et al., 1991), la

mayoría de los cromosomas B tienden a ser neutrales cuando están en números bajos y perjudiciales en números altos (Jones y Rees, 1982; Jones y Houben, 2003). Algunas veces presentan la peculiaridad de causar efectos en relación con su presencia en número par o impar (Jones y Rees, 1967; Camacho et al., 2004), siendo, en general, más perjudiciales los números impares. En ambos casos, los efectos suelen aparecer en caracteres asociados con el vigor, fertilidad y fecundidad, siendo este efecto negativo sobre la eficacia biológica de los individuos portadores. Sin embargo, la transmisión vertical obligada del cromosoma B como parásito, implica que su fitness está íntimamente ligada con la del hospedador, por lo que se espera que exista atenuación temprana de la virulencia (Camacho, 2005).

Muchos de los efectos de los cromosomas B se han relacionado con características como el vigor, la fertilidad, la fecundidad o la tasa de crecimiento. Así, en centeno, las plantas con más cromosomas B muestran germinación tardía, disminución en el crecimiento y vigor, en número de flores y en la fertilidad (Müntzing, 1943; Moss et al., 1966; Cruz-Pardilla et al., 1989; Romera et al., 1989; Jones, 1991b; Jiménez et al., 1994; Puertas et al., 1985). En el saltamontes *Myrmeleotettix maculatus* se ha observado una reducción en la tasa de crecimiento (Hewitt y East, 1978) y en *Eyprepocnemis plorans* una reducción en la tasa de fertilidad y fecundidad (Muñoz et al., 1998; Bakkali et al., 2010). También en saltamontes se ha descrito la formación de espermátidas anómalas en relación con la presencia y número de cromosomas B, lo que probablemente representará una reducción en fertilidad de los individuos portadores (Bidau, 1987; Hewitt et al., 1987; Teruel et al., 2009a).

Los efectos de los cromosomas B derivan, en ocasiones, de su

mera presencia en el genoma, lo que se conoce como efectos nucleotípicos (Bennett, 1972). Los cromosomas B producen un incremento intraespecífico en la cantidad de DNA nuclear y producen efectos que pueden manifestarse al nivel celular. En este tipo de efectos se enmarcan, por ejemplo, aumentos en el tamaño de la célula como consecuencia del aumento de la cantidad de ADN (Cavalier-Smith, 1985; Jones y Rees, 1982) y división celular más lenta, como consecuencia del aumento de tamaño del núcleo y de la cantidad de ADN (Bennett, 1971; Evans et al., 1972).

El efecto de los cromosomas B sobre la frecuencia de quiasmas y, por tanto, en la tasa de recombinación potencial ha sido estudiado en multitud de plantas y animales (véase Jones y Rees, 1982; Bell y Burt, 1990; Camacho, 2005). En muchas especies, la presencia de cromosomas B se asocia con un incremento de la frecuencia de quiasmas entre los cromosomas A durante la meiosis que, en la mayoría de las ocasiones, se ha visto que es dependiente del número de cromosomas B y, a veces, del grado de parasitismo de éstos (Camacho et al., 2002). En otros casos, el incremento en la frecuencia de quiasmas muestra efecto par-impar en relación al número de Bs (Ghaffari y Bidmeshkipoor, 2002). Se han visto también alteraciones de la distribución de los quiasmas (Jones y Rees, 1967; Fletcher y Hewitt, 1980). En la mayoría de los casos este efecto se explica bajo la teoría de la recombinación inducible, que postula que un incremento en la frecuencia de quiasmas se favorece en presencia de los cromosomas B, ya que generará un mayor número de descendientes recombinantes y con ello aumenta la probabilidad de que aparezcan individuos resistentes a la acumulación de cromosomas B en la línea germinal (Bell y Burt, 1990).

Los cambios en la frecuencia y distribución de los quiasmas se han visto implicados en la respuesta a diversas condiciones estresantes (Korol, 1999). Un cromosoma B parasítico podría actuar como factor intrínseco de estrés en las poblaciones invadidas, ya que puede alterar el estado y los procesos normales de la célula. En relación a este argumento, además del incremento en la frecuencia de quiasmas (Camacho et al., 2002) en *Eyprepocnemis plorans* se han detectado alteraciones, asociadas con la presencia de cromosomas B, en otros marcadores de estrés, tales como el tamaño nucleolar (Teruel et al., 2007) y los niveles de expresión de la proteína Hsp70 (Teruel et al., 2011).

Comportamiento mitótico y meiótico: mecanismos de acumulación

La característica evolutivamente más interesante de los cromosomas B es su capacidad de transmitirse a tasas superiores a las establecidas por las leyes de Mendel (el denominado impulso meiótico) incrementando así su frecuencia en las poblaciones naturales. Los mecanismos de acumulación en virtud de los cuales el cromosoma B presenta ese comportamiento pueden darse antes, durante o después de la meiosis (Jones, 1991a).

La mayoría de los cromosomas B son mitóticamente estables, es decir, aparecen en número constante en todos los tejidos somáticos que se conocen. Cuando hay inestabilidad mitótica, ésta puede llevar a la ausencia de cromosomas B en ciertos tejidos, como se ha descrito en algunas especies de plantas que carecen de cromosomas B en las raíces pero los tienen en las partes aéreas (Jones, 1975), o simplemente a

la variación entre células para el número de cromosomas B, como se ha descrito en los saltamontes *Calliptamus palaestinensis* (Nur, 1963), *Camnula pellucida* (Nur, 1969) y *Locusta migratoria* (Cabrero et al., 1984), en el pez *Prochilodus lineatus* (Cavallaro et al., 2000) y en la planta *Crepis capillaris* (Rutishauser y Rothlisberger, 1966). Este mosaicismo para la presencia de cromosomas B en las células del hospedador suele resultar de la existencia de mecanismos de acumulación pre-meiótica. Este tipo de acumulación se produce mediante no disyunción mitótica en las células embrionarias de la línea germinal, de modo que, antes de la gametogénesis, las células que darán lugar a los gametos ya portan un número medio de cromosomas B más alto que el resto de células del hospedador, por lo que sus gametos llevarán un incremento en cromosomas B. Se ha sugerido que la estabilización mitótica de este tipo de cromosomas B podría estar relacionada con la neutralización de su mecanismo de acumulación por no disyunción mitótica (Viseras et al., 1990; Cavallaro et al., 2000).

La acumulación meiótica ocurre normalmente durante la meiosis femenina. El mecanismo consiste en la segregación preferente del cromosoma B hacia la célula meiótica que formará el oocito, en vez de al corpúsculo polar, que degenera. Este tipo de acumulación lo presentan, por ejemplo, los cromosomas B de *Lillium callosum* (Kimura y Kayano, 1961) y los saltamontes *Myrmeleotettix maculatus* (Hewitt, 1976) y *Eyprepocnemis plorans* (Zurita et al., 1998). Existe otra variante de acumulación meiótica, durante la espermatogénesis en *Pseudococcus affinis* (Nur, 1962). En este sistema, las hembras se desarrollan a partir de embriones en los que los dos juegos cromosómicos son eucromáticos, mientras que, en embriones macho, el juego cromosómico de origen

paterno se heterocromatiniza. Durante la espermatogénesis, ambos juegos segregan a polos opuestos y sólo los productos meióticos que contienen el juego eucromático serán funcionales. Durante la profase I de la espermatogénesis, el cromosoma B, que hasta este momento era heterocromático, pasa a estar incluso menos condensado que el juego eucromático, y este cambio de condensación hace posible que segregue preferencialmente con el juego funcional, consiguiendo así el cromosoma B una tasa de transmisión cercana al 90 % (Nur, 1962).

La acumulación post-meiótica se da, sobre todo, en plantas, y se produce mediante no-disyunción mitótica de los cromosomas B, que determina que las dos cromátidas del B migren juntas al núcleo generativo durante la formación del gametofito masculino. La no-disyunción puede ocurrir en la primera mitosis de formación del grano de polen, o en la segunda. Los cromosomas B del centeno (Jones y Puertas, 1993) y del maíz (Roman, 1947) se comportan así.

Por último, en sistemas de determinación sexual haplodiploide puede darse un tipo de acumulación ameiótica. Este sistema consistiría en la eliminación de un juego cromosómico completo, lo que hace que los cigotos fecundados, destinados a ser hembras, se conviertan en machos haploides portadores del cromosoma B, lo que supondría para éste una tasa de transmisión cercana al 100 %. A estos cromosomas B se les denomina PSR, por “paternal sex ratio”, ya que alteran drásticamente la proporción de hembras y machos en las poblaciones. Este mecanismo se ha observado en las avispas *Nasonia vitripennis* (Werren, 1991) y *Trichogramma kaykai* (Van Vugt et al., 2003).

Control de la transmisión

Una de las más enigmáticas cuestiones que se plantean en el estudio de los cromosomas supernumerarios es cómo éstos son capaces de llevar a cabo los mecanismos de acumulación que les permiten invadir y establecerse en las poblaciones. Parece probable que, en cierta medida, esto dependa del contenido genético de dichos cromosomas. Al menos en dos casos se ha descrito que la transmisión de los cromosomas B puede verse favorecida por la existencia de genes o secuencias localizadas en estos cromosomas, que favorecen su propia transmisión. Este el caso de los cromosomas B del centeno y del maíz.

En el centeno (*Secale cereale*), se ha demostrado que los factores que controlan la tasa de transmisión de los cromosomas B mediante la no-disyunción de los cromosomas B durante la formación de los granos de polen está controlada por algún tipo de factor localizado en la región distal del brazo largo de éste (Puertas et al., 1998; Puertas, 2002). Cuando todos los Bs de la célula en división carecen de esta región, la no-disyunción no se produce. Sin embargo, la presencia de un B con la región distal intacta hace que la no-disyunción ocurra incluso en los Bs que carecen de esta región (Lima-de Faria, 1962), sugiriendo algún tipo de interacción en trans entre el cromosoma B capaz de transmitirse y el defectivo. Se han caracterizado dos secuencias específicas del cromosoma B, denominadas E3900 y D1100 (Langdon et al., 2000), que están repetidas en la región distal del B del centeno. Difieren en el grado de condensación, son activas transcripcionalmente, y producen una colección heterogénea de ARNs no codificantes (Carchilan et al., 2007). E3900 parece derivar de un retrotransposon Gypsy, mientras que D1100 es de

origen indeterminado (Puertas, 2002). Estas secuencias parecen estar implicadas en el proceso de transmisión del B del centeno (para revisión, ver Jones et al., 2008).

En el cromosoma B del maíz (*Zea mays*), existen varios mecanismos de acumulación que, mediante estudios de translocaciones entre cromosomas A y B, se han relacionado con cuatro regiones diferentes del B. Así, el mecanismo de no disyunción en la segunda división meiótica del polen requiere un factor localizado en la región eucromática distal del brazo largo, que además actúa en trans (Lamb et al., 2006), induciendo una fuerte unión de la región centromérica de los cromosomas B, de forma que las cromátidas no se separan (Han et al., 2007). El tercer bloque heterocromático del brazo largo (H3) permite a los univalentes B migrar a uno de los polos durante la meiosis, a pesar de la ausencia de un homólogo (Carlson y Roseman, 1992). Se han caracterizado varios elementos específicos del maíz, uno de ellos se ha denominado ZmBs y se encuentra organizado en tandem con repeticiones muy largas (~1400 pb) en la zona del centrómero, y parece estar implicado en la función centromérica (Jin et al., 2005), en la no disyunción (Han et al., 2007) y en la correcta segregación de los cromosomas B (González-Sánchez et al., 2008). Se ha descrito otra secuencia, que se identificó inicialmente como marcador RAPD entre líneas isogénicas con y sin cromosoma B (Stark et al., 1996) y contiene una región codificante homóloga al elemento LTR Prem2 del maíz. También se ha encontrado una secuencia llamada StarkB, de unos 22kb, que se distribuye entre el tercer y cuarto bloque de heterocromatina, no está organizada en arrays, y se compone de secuencias repetitivas procedentes del genoma A y secuencias específicas del B, frecuentemente interrumpida

por retrotransposones. Esta secuencia se transcribe, y podría estar involucrada en los mecanismos de acumulación y mantenimiento del B (Lamb et al., 2007).

En el transcurso de la coevolución de los cromosomas B con su genoma hospedador, la selección favorece que este último desarrolle mecanismos de resistencia frente a la invasión de los cromosomas parásitos. La posibilidad de la existencia de genes capaces de reducir la tasa de transmisión de los cromosomas B fue propuesta incialmente por Shaw (1984). Al año siguiente se publicaron tres trabajos concernientes a esta hipótesis, dos de ellos que demostraban experimentalmente la existencia de estos genes y de la coevolución entre el genoma A y el cromosoma B en *Myrmeleotettix maculatus* (Shaw y Hewitt, 1985) y *Pseudococcus affinis* (= *obscurus*) (Nur y Brett, 1985), y el otro demostraba su ausencia en *Secale cereale*, argumentando que el mantenimiento de los B en esta especie se debe principalmente a sus mecanismos de acumulación (Puertas et al., 1985), aunque más tarde se observarían mecanismos de supresión de la transmisión de los cromosomas B en esta especie (Romera et al., 1991; Jimenez et al., 1995). En maíz, se han identificado algunos loci que se asocian a una alta o baja probabilidad de transmisión de los cromosomas B (González-Sánchez, 2003).

Una de las principales evidencias de la existencia de genes supresores de la acumulación de los cromosomas B se describe en dos trabajos de Nur y Brett (1985, 1987), que relacionan la tasa de transmisión de los cromosomas B con el genotipo de origen materno de los machos de la cochinilla *P. affinis* (= *obscurus*), demostrando que en algunas poblaciones se podía reducir la tasa de transmisión del B desde 0.9 a menos de 0.1 en poco tiempo. Poco después descubrieron que este efecto

se debía a que algunos genotipos alteraban el grado de condensación del cromosoma B y su posición en la placa metafásica, que son esenciales para su acumulación, de forma que su segregación no se producía correctamente (Nur y Brett, 1988).

Modelos de mantenimiento de los cromosomas B en las poblaciones naturales

Existen varios modelos para explicar el mantenimiento de los cromosomas B en las poblaciones naturales.

El modelo parasítico sostiene que los cromosomas B se mantienen por sus mecanismos de acumulación, a pesar de los efectos deletéreos que causen a los organismos hospedadores. Como se ha comentando anteriormente, la mayoría de los sistemas de cromosomas B estudiados hasta ahora se ajustan mejor a este modelo (Camacho, 2005).

El modelo heterótico propone que los cromosomas B no tienen mecanismos de acumulación y se mantienen debido a los efectos beneficiosos que confieren a los individuos portadores cuando se encuentran en bajo número, equilibrados por los perjuicios que causan en alto número (Darlington, 1958; White, 1973).

Ambos modelos conllevan la existencia de un equilibrio entre dos fuerzas contrapuestas, que conduciría a una frecuencia de cromosomas B más o menos estable en las poblaciones actuales. En el primero, estas fuerzas son la acumulación del B, que incrementa su frecuencia, y la selección fenotípica contra los portadores, que la disminuye. En el segundo, el equilibrio resulta de la selección fenotípica a favor de los individuos con

pocos B y la que se produce contra los individuos con muchos B.

En cualquiera de los casos, el incremento en frecuencia del cromosoma B llega a ser una carga para el hospedador. De esta manera, si consideramos la existencia de respuestas defensivas por parte del genoma hospedador, entramos en la típica dinámica coevolutiva hospedador-parásito: una "carrera de armamentos" intragenómica.

Sobre esta base, Camacho et al. (1997) propusieron el modelo casi-neutro, que complementa al modelo parasítico al incluir la respuesta del genoma hospedador, y que no asume que el sistema tenga que alcanzar un equilibrio, sino que pasa por una sucesión en el tiempo y en el espacio de diferentes relaciones coevolutivas.

Las primera etapa de esta sucesión sería la de invasión, en la que un cromosoma B parasítico muestra mecanismos de acumulación y consigue aumentar rápidamente su frecuencia en la población. Cuando el incremento en frecuencia del cromosoma B llega a ser una carga para el hospedador, se produce un conflicto intragenómico entre el genoma hospedador y los cromosomas B. El genoma hospedador desarrollaría mecanismos de resistencia a los cromosomas B como, por ejemplo, un aumento en la frecuencia de alelos en los cromosomas A que supriman la acumulación de los cromosomas B. El sistema estaría entrando entonces en una etapa de neutralización: si los genes de resistencia que suprimen la acumulación del cromosoma B tienen un coste insignificante o menor a la virulencia de los B, el cromosoma B puede llegar a neutralizarse, dando lugar a la supresión de la acumulación.

El B neutralizado puede seguir tres caminos: desaparecer de la población por acción principalmente de la deriva genética y de la selección natural

contra individuos con alto número de Bs, establecerse en la población, convertidos en elementos neutros segregando a la misma tasa que los cromosomas A, o bien regenerar su capacidad de acumulación. Dada la alta tasa de mutabilidad de estos cromosomas, es probable que se produzca una mutación y se origine una nueva variante de cromosoma B que presente mecanismos de acumulación, de manera que pueda invadir de nuevo la población y repetir el ciclo una y otra vez.

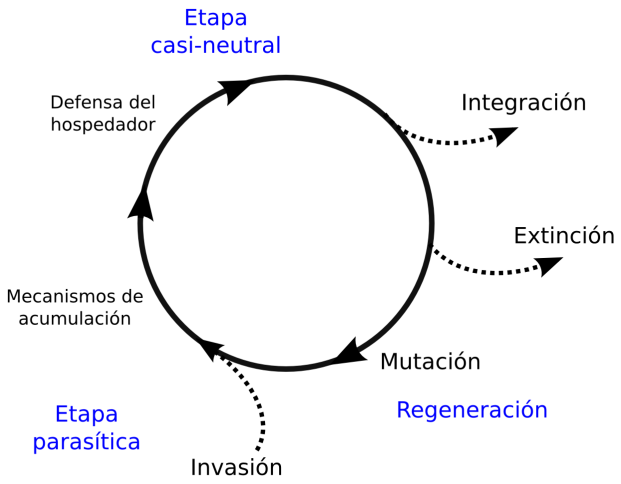

Figura 1: Dinámica evolutiva de un cromosoma B parasítico según el modelo de Camacho et al. (1997).

El sistema de cromosomas B de *Eyprepocnemis plorans plorans*

La especie de saltamontes *Eyprepocnemis plorans* (Orthoptera, Acrididae) fue descrita por Charpentier en 1825. La distribución geográfica de la subespecie *E. plorans plorans* comprende todo el litoral Mediterráneo, el Cáucaso, Turquía, Turkmenistán, Irán y el sudoeste de la Península Arábiga (Dirsh, 1958).

La dotación cromosómica de *E. plorans* es la típica de la familia Acrididae: 22 autosomas y un cromosoma X telocéntrico, que se clasifican en tres grupos según su tamaño: cromosomas largos (L1-L2), medianos (M3-M8) y pequeños (S9-S11), siendo el X de tamaño intermedio entre los cromosomas L2 y M3. El sistema cromosómico de determinación del

sexo en esta especie es X0/XX, siendo los machos X0 y las hembras XX. En la mayoría de las poblaciones estudiadas se ha descrito la presencia adicional de cromosomas B. El sistema de cromosomas B es mitóticamente estable, por lo que el mismo número de cromosomas B está presente en todas las células del individuo.

El sistema de cromosomas B de *E. plorans plorans* ha sido ampliamente estudiado desde su descripción (para revisión, ver Camacho et al., 2003). Los cromosomas B de este sistema son altamente polimórficos, con más de 50 variantes citogenéticas descritas (López-León et al., 1993). La frecuencia de cromosomas B varía ampliamente, tanto entre poblaciones como entre individuos, pudiendo encontrar desde individuos no portadores de cromosoma B hasta individuos con seis cromosomas supernumerarios (Camacho et al., 2003).

Los cromosomas B de *E. plorans* son heterocromáticos y están enriquecidos en dos clases de secuencias de ADN repetido, que además son compartidas con los cromosomas A: un ADN satélite de 180 bp y ADN ribosómico (López-León et al., 1994; Cabrero et al., 1999). La presencia en el cromosoma X de estas dos secuencias en una ordenación, con respecto al centrómero, similar a la observada en el cromosoma B2 permitió hipotetizar a López-León et al. (1994) que el cromosoma B2 había derivado del cromosoma X. La ordenación de estas secuencias repetidas está conservada en la mayoría de los cromosomas B de la región Mediterránea occidental (Bakkali et al., 1999; Cabrero et al., 1999, 2014), lo que sugiere un origen común para todas las variantes (Cabrero et al., 2014).

La proporción relativa de ADNr y ADN satélite que llevan los

cromosomas B es diferente en las poblaciones del Mediterráneo oriental (Abdelaziz et al., 2007; López-León et al., 2008), que además contienen ADN ribosómico 5S, que no es detectado en las poblaciones occidentales (Cabrero et al., 2003), lo que hizo considerar un origen multiregional de los cromosomas B. Esta posibilidad fue, sin embargo, descartada posteriormente al observar la alta similitud entre poblaciones muy distantes para una secuencia específica del cromosoma B (Muñoz-Pajares et al., 2011).

En cuanto a la transcripción de las secuencias de ADN contenidas en los cromosomas B y sus efectos fenotípicos, tal como hemos comentado anteriormente, algunas variantes de B observadas en *E. plorans* se han relacionado con la eficacia biológica de las hembras, la frecuencia de quiasmas, el área nucleolar y la expresión de la proteína de estrés Hsp70. Además, se ha observado la funcionalidad en los genes para ARNr localizados en los cromosomas B de *E. plorans* (Ruiz-Estévez et al., 2012), aunque esta expresión se produce en un porcentaje muy bajo de individuos (Ruiz-Estévez et al., 2013) y su contribución al rRNA total es mínima (Ruiz-Estévez et al., 2014), lo que apunta a un alto grado de silenciamiento, tal y como sugiere el estado hipoacetilado de los cromosomas B durante la meiosis, en las células en interfase de testículos y embriones (Cabrero et al., 2007).

En el saltamontes *E. plorans* se han documentado los tres estadios de evolución contemplados en el modelo de evolución a largo plazo de los cromosomas B parasíticos, propuesto por Camacho et al. (1997). La acumulación de los cromosomas B de esta especie tiene lugar durante la meiosis femenina, y está basada en su segregación preferencial hacia el polo anafásico que formará el oocito secundario (Zurita et al., 1998). Las

principales variantes de cromosomas B observadas en *E. plorans* (B1, B2 y B5) se encuentran en un estadio de evolución neutro o cercano a la neutralidad, ya que muestran tasas de transmisión mendelianas y no producen efectos deletéreos sobre el hospedador (López-León et al., 1992). Esta neutralización es presumiblemente el resultado de la coevolución de los cromosomas parásitos con el genoma hospedador, como se comprobó más tarde mediante experimentos que demostraron que una misma hembra con 1B podía transmitir su cromosoma B a tasa mendeliana si copulaba con un macho 0B de su misma población, pero a tasa mayor (no-mendeliana) si copulaba con un macho de una población sin cromosomas B (Herrera et al., 1996). Posteriormente, se observó la acumulación de la variante B24 en la población malagueña de Torrox, con una tasa media de transmisión de 0,7, y un significativo descenso en la fertilidad de los huevos (Zurita et al., 1998), y también la acumulación de algunas variantes de cromosomas B de poblaciones marroquíes (Bakkali et al., 2002), que demostraron que los cromosomas B de *E. plorans* tienden a mostrar tasas de transmisión superiores a la tasa mendeliana, a pesar de que, en la mayoría de los casos, estos cromosomas no muestran acumulación (López-León et al., 1992, 1993). La doble cara de los cromosomas B de esta especie, mostrando unas veces acumulación y otras no, quedó en evidencia al observar que la tasa de transmisión de B24 en Torrox se redujo a cerca de 0,5 en sólo seis años (Perfectti et al., 2004).

Bibliografía

Abdelaziz, M., Teruel, M., Chobanov, D., Camacho, J. P. M., y Cabrero, J. (2007). Physical mapping of rDNA and satDNA in A and B chromosomes of the grasshopper *Eyprepocnemis plorans* from a Greek population. *Cytogenetic and Genome Research*, 119(1-2):143–146.

Akbari, O. S., Antoshechkin, I., Hay, B. A., y Ferree, P. M. (2013). Transcriptome profiling of *Nasonia vitripennis* testis reveals novel transcripts expressed from the selfish B chromosome, paternal sex ratio. *G3: Genes| Genomes| Genetics*, 3(9):1597–1605.

Bakkali, M., Cabrero, J., López-León, M. D., Perfectti, F., y Camacho, J. P. M. (1999). The B chromosome polymorphism of the grasshopper *Eypepocnemis plorans* in North Africa. I. B variants and frequency. *Heredity*, 83(4):428–434.

Bakkali, M., Manrique-Poyato, M. I., López-León, M. D., Perfectti, F., Cabrero, J., y Camacho, J. P. M. (2010). Effects of B chromosomes on egg fertility and clutch size in the grasshopper *Eyprepocnemis plorans*. *Journal of Orthoptera Research*, 19(2):197–203.

Bakkali, M., Perfectti, F., y Camacho, J. P. M. (2002). The B-chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa: II. Parasitic and neutralized B1 chromosomes. *Heredity*, 88(1):14–18.

Banaei-Moghaddam, A. M., Meier, K., Karimi-Ashtiyani, R., y Houben, A. (2013). Formation and expression of pseudogenes on the B chromosome of rye. *The Plant Cell*, 25(7):2536–2544.

Battaglia, E. (1964). *Cytogenetics of B-chromosomes. Caryologia, 17: 245–299*. Taylor & Francis Online.

Baverstock, P. R., Gelder, M., y Jahnke, A. (1982). Cytogenetic studies of the Australian rodent, *Uromys caudimaculatus*, a species showing extensive heterochromatin variation. *Chromosoma*, 84(4):517–533.

Bell, G. y Burt, A. (1990). B-chromosomes: germ-line parasites which induce changes in host recombination. *Parasitology*, 100(S1):S19–S26.

Bennett, M. D. (1971). The Duration of Meiosis. *Proceedings of the Royal Society of London B: Biological Sciences*, 178(1052):277–299.

Bennett, M. D. (1972). Nuclear DNA Content and Minimum Generation Time in Herbaceous Plants. *Proceedings of the Royal Society of London B: Biological Sciences*, 181(1063):109–135.

Bidau, C. J. (1987). Influence of a rare unstable B-chromosome on chiasma frequency and nonhaploid sperm production in *Dichroplus pratensis* (Melanoplinae, Acrididae). *Genetica*, 73(3):201–210.

Brockhouse, C., Bass, J., Feraday, R., y Straus, N. (1989). Supernumerary chromosome evolution in the *Simulium vernum* group (Diptera: Simuliidae). *Genome*, 32(4):516–521.

Burt, A. y Trivers, R. (2006). B chromosomes. *Genes in conflict: The biology of selfish genetic elements*, pages 325–380.

Cabrero, J., Alché, J. D., y Camacho, J. P. M. (1987). Effects of B chromosomes on the activity of nucleolar organizer regions in the grasshopper *Eyprepocnemis plorans*: activation of a latent nucleolar organizer region on a B chromosome fused to an autosome. *Genome*, 29(1):116–121.

Cabrero, J., Bakkali, M., Bugrov, A., Warchalowska-Sliwa, E., López-León, M. D., Perfectti, F., y Camacho, J. P. M. (2003). Multiregional origin of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Chromosoma*, 112(4):207–211.

Cabrero, J., López-León, M., Bakkali, M., y Camacho, J. (1999). Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. *Heredity*, 83(4):435–439.

Cabrero, J., López-León, M. D., Gómez, R., Castro, A. J., Martín-Alganza, A., y Camacho, J. P. M. (1997). Geographical distribution of B chromosomes in the grasshopper *Eyprepocnemis plorans*, along a river basin, is mainly shaped by non-selective historical events. *Chromosome Research*, 5(3):194.

Cabrero, J., López-León, M. D., Ruiz-Estévez, M., Gómez, R., Petitpierre, E., Rufas, J. S., Massa, B., Kamel Ben Halima, M., y Camacho, J. P. M. (2014). B1 was the ancestor B chromosome variant in the western Mediterranean area

in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 142(1):54–58.

Cabrero, J., Teruel, M., Carmona, F. D., Jiménez, R., y Camacho, J. P. M. (2007). Histone H3 lysine 9 acetylation pattern suggests that X and B chromosomes are silenced during entire male meiosis in a grasshopper. *Cytogenetic and Genome Research*, 119(1-2):135–142.

Cabrero, J., Viseras, E., y Camacho, J. P. M. (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(3):155–164.

Camacho, J. P. M. (2005). B chromosomes. In Gregory, T. R., editor, *The Evolution of the Genome*, pages 223–286. Academic Press.

Camacho, J. P. M., Bakkali, M., Corral, J. M., Cabrero, J., López-León, M. D., Aranda, I., Martín-Alganza, A., y Perfectti, F. (2002). Host recombination is dependent on the degree of parasitism. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1505):2173–2177.

Camacho, J. P. M., Cabrero, J., López-León, M. D., Bakkali, M., y Perfectti, F. (2003). The B chromosomes of the grasshopper *Eyprepocnemis plorans* and the intragenomic conflict. *Genetica*, 117(1):77–84.

Camacho, J. P. M., López-León, M. D., Pardo, M. C., Cabrero, J., y Shaw, M. W. (1997). Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper *Eyprepocnemis plorans*. *The American Naturalist*, 149(6):1030–1050.

Camacho, J. P. M., Perfectti, F., Teruel, M., López-León, M. D., y Cabrero, J. (2004). The odd-even effect in mitotically unstable B chromosomes in grasshoppers. *Cytogenetic and Genome Research*, 106(2-4):325–331.

Camacho, J. P. M., Sharbel, T. F., y Beukeboom, L. W. (2000). B-chromosome evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355(1394):163–178.

Carchilan, M., Delgado, M., Ribeiro, T., Costa-Nunes, P., Caperta, A., Morais-Cecílio,

L., Jones, R., Viegas, W., y Houben, A. (2007). Transcriptionally active heterochromatin in rye B chromosomes. *The Plant Cell Online*, 19(6):1738.

Carchilan, M., Kumke, K., Mikolajewski, S., y Houben, A. (2009). Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma*, 118(5):607–616.

Carlson, W. R. y Roseman, R. R. (1992). A new property of the maize B chromosome. *Genetics*, 131(1):211–223.

Cavalier-Smith, T. (1985). Cell volume and the evolution of eukaryotic genome size. *The evolution of genome size*, pages 105–184.

Cavallaro, Z. I., Bertollo, L. A. C., Perfectti, F., y Camacho, J. P. M. (2000). Frequency increase and mitotic stabilization of a B chromosome in the fish *Prochilodus lineatus*. *Chromosome Research*, 8(7):627–634.

Cruz-Pardilla, M., Vences, F. J., Garcia, P., y Perez De La Vega, M. (1989). The effect of B chromosomes on outcrossing rate in a population of rye *Secale cereale*. *L. Heredity*, 62:319–326.

Darlington, C. D. (1958). *Evolution of genetic systems*. Oliver and Boyd, Edinburgh and London.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Delgado, M., Caperta, A., Ribeiro, T., Viegas, W., Jones, R. N., y Morais-Cecilio, L. (2004). Different numbers of rye B chromosomes induce identical compaction changes in distinct A chromosome domains. *Cytogenetic and Genome Research*, 106(2-4):320–324.

Dhar, M. K., Friebe, B., Koul, A. K., y Gill, B. S. (2002). Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma*, 111(5):332–340.

Dherawattana, A. y Sadanaga, K. (1973). Cytogenetics of a crown rust-resistant hexaploid oat with 42+ 2 fragment chromosomes. *Crop Science*, 13(6):591–594.

Dimitri, P., Junakovic, N., y Arcà, B. (2003). Colonization of heterochromatic genes by transposable elements in Drosophila. *Molecular Biology and Evolution*, 20(4):503–512.

Dirsh, V. M. (1958). Revision of the genus Eyprepocnemis Fieber, 1853 (orthoptera: Acridoidea). *Proceedings of the Royal Entomological Society of London. Series B, Taxonomy*, 27(3-4):33–45.

Doolittle, W. F. y Sapienza, C. (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature*, 284(5757):601–3.

Evans, G. M., Rees, H., Snell, C. L., y Sun, S. (1972). The relationship between nuclear DNA amount and the duration of the mitotic cycle. *Chromosomes today*, 3(1):24–31.

Fletcher, H. L. y Hewitt, G. (1980). Effect of a "B" chromosome on chiasma localisation and frequency in male *Euthystira brachyptera*. *Heredity*, 44(3):341–347.

Fox, D. P., Hewitt, G. M., y Hall, D. J. (1974). DNA replication and RNA transcription of euchromatic and heterochromatic chromosome regions during grasshopper meiosis. *Chromosoma*, 45(1):43–62.

Ghaffari, S. M. y Bidmeshkipoor, A. (2002). Presence and behaviour of B-chromosomes in *Acanthophyllum laxiusculum*. *Genetica*, 115(3):319–323.

González-Sánchez, M., González-García, M., Vega, J. M., Rosato, M., Cuacos, M., y Puertas, M. J. (2008). Meiotic loss of the B chromosomes of maize is influenced by the B univalent co-orientation and the TR-1 knob constitution of the A chromosomes. *Cytogenetic and Genome Research*, 119(3-4):282–290.

Graphodatsky, A. S., Kukekova, A. V., Yudkin, D. V., Trifonov, V. A., Vorobieva, N. V., Beklemisheva, V. R., Perelman, P. L., Graphodatskaya, D. A., Trut, L. N., Yang, F., Ferguson-Smith, M. A., Acland, G. M., y Aguirre, G. D. (2005). The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Research*, 13(2):113–122.

Green, D. M. (1990). Muller's Ratchet and the evolution of supernumerary chromosomes. *Genome*, 33(6):818–824.

Gutknecht, J., Sperlich, D., y Bachmann, L. (1995). A species specific satellite DNA family of *Drosophila subsilvestris* appearing predominantly in B chromosomes. *Chromosoma*, 103(8):539–544.

Han, F., Lamb, J. C., Yu, W., Gao, Z., y Birchler, J. A. (2007). Centromere function

and nondisjunction are independent components of the maize B chromosome accumulation mechanism. *The Plant Cell*, 19(2):524–533.

Herrera, J., López-León, M. D., Cabrero, J., Shaw, M. W., y Camacho, J. P. M. (1996). Evidence for B chromosome drive suppression in the grasshopper *Eyprepocnemis plorans*. *Heredity*, 76(6):633–639.

Hewitt, G. (1976). Meiotic drive for B-chromosomes in the primary oocytes of *Myrmeleotettix maculatus* (Orthopera: Acrididae). *Chromosoma*, 56(4):381.

Hewitt, G. M. (1979). Animal Cytogenetics. Volume 3. Insecta 1: Orthoptera, Grasshoppers and crickets.

Hewitt, G. M. y East, T. M. (1978). Effects of B chromosomes on development in grasshopper embryos. *Heredity*, 41:347–356.

Hewitt, G. M., East, T. M., y Shaw, M. W. (1987). Sperm dysfunction produced by B-chromosomes in the grasshopper *Myrmeleotettix maculatus*. *Heredity*, 58(5):68.

Holmes, D. y Bougourd, S. (1989). Heredity - Abstract of article: B-chromosome selection in *Allium schoenoprasum*. I. Natural populations. *Heredity*, 63(1):83–87.

Houben, A., Thompson, N., Ahne, R., Leach, C. R., Verlin, D., y Timmis, J. N. (1999). A monophyletic origin of the B chromosomes of *Brachycome dichromosomatica* (Asteraceae). *Plant Systematics and Evolution*, 219(1-2):127–135.

Huang, W., Du, Y., Zhao, X., y Jin, W. (2016). B chromosome contains active genes and impacts the transcription of A chromosomes in maize (*Zea mays* L.). *BMC Plant Biology*, 16:88.

Hurst, L. D. (1992). Intragenomic conflict as an evolutionary force. *Proceedings of the Royal Society of London B: Biological Sciences*, 248(1322):135–140.

Ishak, B., Jaafar, H., Maetz, J., y Rumpler, Y. (1991). Absence of transcriptional activity of the B-chromosomes of *Apodemus peninsulae* during pachytene. *Chromosoma*, 100(4):278–281.

Jackson, R. C. y Newmark, P. (1960). Effects of supernumerary chromosomes on production of pigment in *Haplopappus gracilis*. *Science*, 132(3436):1316–1317.

Jamilena, M., Garrido-Ramos, M., Ruiz Rejón, C., y Ruiz Rejón, M. (1994). Molecular

relationship between the A and B chromosomes of *Crepis capillaris*. *Heredity*, 73(5):527–531.

Jiménez, M. M., Romera, E., Puertas, M. J., y Jones, R. N. (1994). B-chromosomes in inbred lines of rye (*Secale cereale* L.). *Genetica*, 92(3):149–154.

Jimenez, M. M., Romera, F., Gallego, A., y Puertas, M. J. (1995). Genetic control of the rate of transmission of rye B chromosomes. II. 0b times 2b crosses. *Heredity*, 74:518–523.

Jin, W., Lamb, J. C., Vega, J. M., Dawe, R. K., Birchler, J. A., y Jiang, J. (2005). Molecular and functional dissection of the maize B chromosome centromere. *The Plant Cell*, 17(5):1412–1423.

John, B. y Hewitt, G. M. (1965). The B-chromosome system of *Myrmeleotettix maculatus* (Thunb.). I. The mechanics. *Chromosoma*, 16(5):548–578.

Jones, N. y Houben, A. (2003). B chromosomes in plants: escapees from the A chromosome genome? *Trends in Plant Science*, 8(9):417–423.

Jones, R. N. (1975). B-chromosome systems in flowering plants and animal species. volume 40, pages 1–100.

Jones, R. N. (1985). Are B chromosomes selfish? In Cavalier-Smith, T., editor, *The evolution of genome size*, volume 30, pages 397–425. Wiley, london edition.

Jones, R. N. (1991a). B-chromosome drive. *American Naturalist*, pages 430–442.

Jones, R. N. (1991b). Cytogenetics of B chromosomes in crops. *Chromosome Engineering in Plants: Genetics, Breeding, Evolution*, 2:141.

Jones, R. N. (1995). Tansley review no. 85. B chromosomes in plants. *New Phytologist*, pages 411–434.

Jones, R. N. y Puertas, M. J. (1993). The B-chromosomes of rye (*Secale cereale* L.). *Frontiers in plant science research*, pages 81–112.

Jones, R. N. y Rees, H. (1967). Genotypic control of chromosome behaviour in rye. XI. The influence of B chromosomes on meiosis. *Heredity*, 22:333–347.

Jones, R. N. y Rees, H. (1982). *B chromosomes.* Academic Press.

Jones, R. N., Viegas, W., y Houben, A. (2008). A century of B chromosomes in plants: so what? *Annals of botany*, 101(6):767–775.

Kimura, M. y Kayano, H. (1961). The maintenance of supernumerary chromosomes in wild populations of *Lilium callosum* by preferential segregation. *Genetics*, 46(12):1699.

Korol, A. B. (1999). Selection for adaptive traits as a factor of recombination evolution: evidence from natural and experimental populations (a review). In *Evolutionary theory and processes: modern perspectives*, pages 31–53. Springer.

Lamb, J., Riddle, N., Cheng, Y., Theuri, J., y Birchler, J. (2007). Localization and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Research*, 15(3):383–398.

Lamb, J. C., Han, F., Auger, D. L., y Birchler, J. A. (2006). A trans-acting factor required for non-disjunction of the B chromosome is located distal to the TB-4lb breakpoint on the B chromosome. *Maize Genetics Cooperation Newsletter*, 80:4.

Langdon, T., Seago, C., Jones, R. N., Ougham, H., Thomas, H., Forster, J. W., y Jenkins, G. (2000). De novo evolution of satellite DNA on the rye B chromosome. *Genetics*, 154(2):869–884.

Leach, C. R., Houben, A., Field, B., Pistrick, K., Demidov, D., y Timmis, J. N. (2005). Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. *Genetics*, 171(1):269–278.

Lima-de Faria, A. (1962). Genetic interaction in rye expressed at the chromosome phenotype. *Genetics*, 47(10):1455.

López-León, M. D., Cabrero, J., Camacho, J. P. M., Cano, M. I., y Santos, J. L. (1992). A widespread B chromosome polymorphism maintained without apparent drive. *Evolution*, pages 529–539.

López-León, M. D., Cabrero, J., Dzyubenko, V. V., Bugrov, A. G., Karamysheva, T. V., Rubtsov, N. B., y Camacho, J. P. M. (2008). Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of the grasshopper *Eyprepocnemis plorans plorans*. *Cytogenetic and Genome Research*, 121(3-4):260–265.

López-León, M. D., Cabrero, J., Pardo, M. C., Viseras, E., Camacho, J., y Santos, J. (1993). Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). *Heredity*, 71:352–352.

López-León, M. D., Neves, N., Schwarzacher, T., Heslop-Harrison, J. S. P., Hewitt, G. M., y Camacho, J. P. M. (1994). Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Research*, 2(2):87–92.

Martis, M. M., Klemme, S., Banaei-Moghaddam, A. M., Blattner, F. R., Macas, J., Schmutzer, T., Scholz, U., Gundlach, H., Wicker, T., Šimková, H., Novák, P., Neumann, P., Kubaláková, M., Bauer, E., Haseneyer, G., Fuchs, J., Doležel, J., Stein, N., Mayer, K. F. X., y Houben, A. (2012). Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences*, 109(33):13343–13346.

McAllister, B. y Werren, J. (1997). Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma*, 106(4):243–253.

Mestriner, C. A., Galetti, P. M., Valentini, S. R., Ruiz, I. R., Abel, L. D., Moreira-Filho, O., y Camacho, J. P. M. (2000). Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity*, 85 (Pt 1):1–9.

Miao, V. P., Covert, S. F., y VanEtten, H. D. (1991). A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. *Science*, 254(5039):1773–1776.

Misteli, T. (2007). Beyond the sequence: Cellular organization of genome function. *Cell*, 128(4):787–800.

Moss, J. P., DARLINGTON, C., y LEWIS, K. (1966). The adaptive significance of B-chromosomes in rye. *Chromosomes today*, pages 15–23.

Muñoz, E., Perfectti, F., Martín-Alganza, Á., y Camacho, J. P. M. (1998). Parallel effects of a B chromosome and a mite that decrease female fitness in the grasshopper *Eyprepocnemis plorans*. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1408):1903.

Muñoz-Pajares, A. J., Martínez-Rodríguez, L., Teruel, M., Cabrero, J., Camacho, J.

P. M., y Perfectti, F. (2011). A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics*, 187(3):853–863.

Müntzing, A. (1943). Genetical effects of duplicated fragment chromosomes in rye. *Hereditas*, 29(1-2):91–112.

Nur, U. (1962). A supernumerary chromosome with an accumulation mechanism in the lecanoid genetic system. *Chromosoma*, 13(3):249–271.

Nur, U. (1963). A mitotically unstable supernumerary chromosome with an accumulation mechanism in a grasshopper. *Chromosoma*, 14(4):407–422.

Nur, U. (1969). Mitotic instability leading to an accumulation of B-chromosomes in grasshoppers. *Chromosoma*, 27(1):1–19.

Nur, U. y Brett, B. L. (1985). Genotypes suppressing meiotic drive of a B chromosome in the mealybug, *Pseudococcus obscurus*. *Genetics*, 110(1):73–92.

Nur, U. y Brett, B. L. H. (1987). Control of meiotic drive of B chromosomes in the mealybug, *Pseudococcus affinis (obscurus)*. *Genetics*, 115(3):499–510.

Nur, U. y Brett, B. L. H. (1988). Genotypes affecting the condensation and transmission of heterochromtic B chromosomes in the mealybug *Pseudococcus affinis*. *Chromosoma*, 96(3):205–212.

Nur, U., Werren, J. H., Eickbush, D. G., Burke, W. D., y Eickbush, T. H. (1988). A "selfish"B chromosome that enhances its transmission by eliminating the paternal genome. *Science (New York, N.Y.)*, 240(4851):512–514.

Oliver, J. L., Posse, F., Martinez-Zapater, J. M., Enriquez, A. M., y Ruiz-Rejón, M. (1982). B-Chromosomes and E-1 isozyme activity in mosaic bulbs of *Scilla autumnalis* (Liliaceae). *Chromosoma*, 85(3):399–403.

Östergren, G. (1945). Parasitic nature of extra fragment chromosomes. *Botaniska Notiser*, 2:157–163.

Palestis, B. G., Burt, A., Jones, R. N., y Trivers, R. (2004). B chromosomes are more frequent in mammals with acrocentric karyotypes: support for the theory of centromeric drive. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(Suppl 3):S22–S24.

Peppers, J. A., Wiggins, L. E., y Baker, R. J. (1997). Nature of B chromosomes in the harvest mouse *Reithrodontomys megalotis* by fluorescence in situ hybridization (FISH). *Chromosome Research*, 5(7):475–479.

Perfectti, F., Corral, J. M., Mesa, J. A., Cabrero, J., Bakkali, M., López-León, M. D., y Camacho, J. P. M. (2004). Rapid suppression of drive for a parasitic B chromosome. *Cytogenetic and Genome Research*, 106(2-4):338–343.

Perfectti, F. y Werren, J. (2001). The interspecific origin of B chromosomes: experimental evidence. *Evolution*, 55(5):1069–1073.

Pink, R. C., Wicks, K., Caley, D. P., Punch, E. K., Jacobs, L., y Carter, D. R. F. (2011). Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA*, 17(5):792–798.

Plowman, A. B. y Bougourd, S. M. (1994). Selectively advantageous effects of B chromosomes on germination behaviour in. *Heredity*, 72:587–593.

Poliseno, L., Marranci, A., y Pandolfi, P. P. (2015). Pseudogenes in human cancer. *Frontiers In Medicine*, 2.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., y Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, 465(7301):1033–1038.

Puertas, M. J. (2002). Nature and evolution of B chromosomes in plants: a non-coding but information-rich part of plant genomes. *Cytogenetic and Genome Research*, 96(1-4):198–205.

Puertas, M. J., González-Sánchez, M., Manzanero, S., Romera, F., y Jiménez, M. M. (1998). Genetic control of the rate of transmission of rye B chromosomes. IV. Localization of the genes controlling B transmission rate. *Heredity*, 80(2):209–213.

Puertas, M. J., Romera, F., y de la Peña, A. (1985). Comparison of B chromosome effects on *Secale cereale* and *Secale vavilovii*. *Heredity*, 55(2):229–234.

Randolph, L. F. (1928). Types of supernumerary chromosomes in maize. *The Anatomical Record*, 41:102.

Roman, H. (1947). Mitotic nondisjunction in the case of interchanges involving the B-type chromosome in maize. *Genetics*, 32(4):391.

Romera, F., Jimenez, M. M., y Puertas, M. J. (1991). Genetic control of the rate of transmission of rye B chromosomes. I. Effects in 2b times 0b crosses. *Heredity*, 66:61–65.

Romera, F., Vega, J. M., Díez, M., y Puertas, M. J. (1989). B chromosome polymorphism in Korean rye populations. *Heredity*, 62(1):117–121.

Ruiz-Estévez, M., Badisco, L., Broeck, J. V., Perfectti, F., López-León, M. D., Cabrero, J., y Camacho, J. P. M. (2014). B chromosomes showing active ribosomal RNA genes contribute insignificant amounts of rRNA in the grasshopper *Eyprepocnemis plorans*. *Molecular Genetics and Genomics*, 289(6):1209–1216.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., y Camacho, J. P. M. (2012). B-chromosome ribosomal DNA is functional in the grasshopper *Eyprepocnemis plorans*. *PLoS One*, 7(5):e36600.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., y Camacho, J. P. M. (2013). Ribosomal DNA is active in different B chromosome variants of the grasshopper *Eyprepocnemis plorans*. *Genetica*, 141(7-9):337–345.

Ruiz-Rejón, M., Posse, F., y Oliver, J. L. (1980). The B chromosome system of *Scilla autumnalis* (Liliaceae): effects at the isozyme level. *Chromosoma*, 79(3):341–348.

Rutishauser, A. y Rothlisberger, E. (1966). Boosting mechanism of B-chromosomes in *Crepis capillaris*. In *Chromosomes today*, volume 1, pages 28–30.

Sandery, M. J., Forster, J. W., Blunden, R., y Jones, R. N. (1990). Identification of a family of repeated sequences on the rye B chromosome. *Genome*, 33(6):908–913.

Shaw, M. y Hewitt, G. (1985). The genetic control of meiotic drive acting on the B chromosome of *Myrmeleotettix maculatus* (Orthoptera: Acrididae). *Heredity*, 54(2):187–194.

Shaw, M. W. (1984). The population genetics of the B-chromosome polymorphism of *Myrmeleotettix maculatus* (Orthoptera: Acrididae). *Biological Journal of the Linnean Society*, 23(1):77–100.

Silva, M. J. J. y Yonenaga-Yassuda, Y. (1998). Heterogeneity and meiotic behaviour of B and sex chromosomes, banding patterns and localization of (TTAGGG)˜n sequences by fluorescence in situ hybridization in the neotropical water rat Nectomys (Rodentia, Cricetidae). *Chromosome Research*, 6:455–462.

Stark, E. A., Connerton, I., Bennett, S. T., Barnes, S. R., Parker, J. S., y Forster, J. W. (1996). Molecular analysis of the structure of the maize B-chromosome. *Chromosome Research*, 4(1):15–23.

Staub, R. W. (1987). Leaf striping correlated with the presence of B chromosomes in maize. *Journal of Heredity*, 78(2):71–74.

Tanić, N., Vujošević, M., Dedović-Tanić, N., y Dimitrijević, B. (2005). Differential gene expression in yellow-necked mice *Apodemus flavicollis* (Rodentia, Mammalia) with and without B chromosomes. *Chromosoma*, 113(8):418–427.

Teruel, M., Cabrero, J., Perfectti, F., Alché, J. D., y Camacho, J. P. M. (2009a). Abnormal spermatid formation in the presence of the parasitic B24 chromosome in the grasshopper *Eyprepocnemis plorans*. *Sexual Development*, 3(5):284–289.

Teruel, M., Cabrero, J., Perfectti, F., y Camacho, J. P. M. (2007). Nucleolus size variation during meiosis and NOR activity of a B chromosome in the grasshopper *Eyprepocnemis plorans*. *Chromosome Research*, 15(6):755–765.

Teruel, M., Cabrero, J., Perfectti, F., y Camacho, J. P. M. (2009b). Quantitative analysis of NOR expression in a B chromosome of the grasshopper *Eyprepocnemis plorans*. *Chromosoma*, 118(3):291–301.

Teruel, M., Cabrero, J., Perfectti, F., y Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, 119(2):217–225.

Teruel, M., Sørensen, J. G., Loeschcke, V., Cabrero, J., Perfectti, F., y Camacho, J. P. M. (2011). Level of heat shock proteins decreases in individuals carrying B-chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 132(1-2):94–99.

Valente, G. T., Conte, M. A., Fantinatti, B. E. A., Cabral-de Mello, D. C., Carvalho, R. F., Vicari, M. R., Kocher, T. D., y Martins, C. (2014). Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution*, 31(8):2061–2072.

Van Vugt, J. F., Salverda, M., de Jong, J. H., y Stouthamer, R. (2003). The paternal sex ratio chromosome in the parasitic wasp *Trichogramma kaykai* condenses the paternal chromosomes into a dense chromatin mass. *Genome*, 46(4):580–587.

Viseras, E., Camacho, J. P. M., Cano, M. I., y Santos, J. L. (1990). Relationship between mitotic instability and accumulation of B chromosomes in males and females of *Locusta migratoria*. *Genome*, 33(1):23–29.

Wallrath, L. L. y Elgin, S. C. (1995). Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes and Development*, 9(10):1263–1277.

Werren, J. H. (1991). The paternal-sex-ratio chromosome of Nasonia. *American Naturalist*, 137(3):392–402.

Werren, J. H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10863–10870.

Werren, J. H., Nur, U., y Wu, C.-I. (1988). Selfish genetic elements. *Trends in Ecology and Evolution*, 3(11):297–302.

White, M. J. D. (1973). *Animal Cytology and Evolution*. Cambridge University Press, London, 3rd ed. edition.

Wilson, E. B. (1907). The supernumerary chromosomes of Hemiptera. *Science*, 26:870–871.

Wolf, K. W., Mertl, H. G., y Traut, W. (1991). Structure, mitotic and meiotic behaviour, and stability of centromere-like elements devoid of chromosome arms in the fly Megaselia scalaris (Phoridae). *Chromosoma*, 101(2):99–108.

Wurster-Hill, D. H., Ward, O. G., Davis, B. H., Park, J. P., Moyzis, R. K., y Meyne, J. (1988). Fragile sites, telomeric DNA sequences, B chromosomes, and DNA content in raccoon dogs, *Nyctereutes procyonoides*, with comparative notes on foxes, coyote, wolf, and raccoon. *Cytogenetic and Genome Research*, 49(4):278–281.

Yoshida, K., Terai, Y., Mizoiri, S., Aibara, M., Nishihara, H., Watanabe, M., Kuroiwa, A., Hirai, H., Hirai, Y., Matsuda, Y., y Okada, N. (2011). B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS Genetics*, 7(8):e1002203.

Zhimulev, I. F. y Beliaeva, E. S. (2003). Heterochromatin, gene position effect and gene silencing. *Genetika*, 39(2):187–201.

Zhou, Q., Zhu, H.-m., Huang, Q.-f., Zhao, L., Zhang, G.-j., Roy, S. W., Vicoso, B., Xuan, Z.-l., Ruan, J., Zhang, Y., Zhao, R.-p., Ye, C., Zhang, X.-q., Wang, J., Wang, W., y Bachtrog, D. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genomics*, 13:109.

Zurita, S., Cabrero, J., López-León, M. D., y Camacho, J. P. M. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, pages 274–277.

Objetivos

El objetivo general de la presente Tesis Doctoral es desvelar los mecanismos moleculares de la interacción entre los cromosomas B y el genoma hospedador en el saltamontes *Eyprepocnemis plorans*. Para ello, pondremos a punto diferentes técnicas de PCR cuantitativa, que serán esenciales para abordar los principales objetivos de esta tesis, que se resumen en averiguar si los cromosomas B de esta especie contienen genes para proteínas, si estos genes está transcripcionalmente activos y si esto afecta a los niveles de transcripción de los genes de los cromosomas A.

Para ello, nos planteamos los siguientes objetivos específicos:

1. Averiguar si la presencia de cromosomas B está asociada con mayores niveles de dobles roturas y reparación en el ADN, mediante inmunofluorescencia con anticuerpos para varias proteínas relacionadas con estas funciones.

2. Comprobar si la disminución de los niveles de proteína Hsp70 observados previamente en esta especie, en presencia de cromosomas B, tiene lugar al nivel transcripcional.

3. Poner a prueba la hipótesis que sostiene que los cromosomas B son genéticamente inertes, mediante la búsqueda de genes codificadores de proteínas localizados en los cromosomas B, y el análisis de su posible actividad transcripcional.

4. Averiguar si la presencia de uno de los genes detectados en el cromosoma B24, concretamente el gen que codifica para la subunidad CAP-G de la condensina I, es una peculiaridad de esta variante o bien es un hecho compartido con otras variantes de cromosomas B en *E. plorans*.

5. Estudiar en mayor profundidad el gen CAP-G, analizando el impacto que puede tener la expresión de las copias del cromosoma B sobre la expresión de otros genes relacionados funcionalmente con éste.

6. Trazar un perfil molecular de la respuesta del genoma hospedador a la presencia de los cromosomas B, mediante el estudio comparativo de los cambios de expresión génica observados entre individuos portadores y no portadores de cromosomas B.

Metodología general

Material biológico

Los especímenes utilizados en esta Tesis Doctoral pertenecen mayoritariamente a la especie de ortóptero *Eyprepocnemis plorans*. Para los diferentes experimentos realizados con esta especie hemos utilizado machos y hembras adultos capturados en las poblaciones de Torrox (Málaga) y Salobreña (Granada), durante los años 2011, 2012 y 2013, así

Tabla 1: Resumen del material biológico de utilizado en esta Tesis Doctoral

Eyprepocnemis plorans			
Año	Población	Sex	Técnica
2011	Torrox	Machos	Inmunofluorescencia/ RNAi
	Socovos	Machos	Inmunofluorescencia
	El Gallego	Machos	Inmunofluorescencia
	Laboratorio	Embriones	Inmunofluorescencia
2012	Torrox	Machos	Illumina WGS
		Hembras	Illumina RNA-seq
	Salobreña	Hembras	Microarrays
2013	Torrox	Hembras	qPCR
		Machos	qPCR
	Salobreña	Hembras	qPCR
		Machos	qPCR
2014	Laboratorio	Embriones	FISH-TSA
Otras especies			Inmunofluorescencia
Heteracris adspersa		*Aiolopus strepens*	*Chorthippus vagans*
H. litoralis		*Acrotylus insubricus*	*Ch. jacobsi*
Sphingonotus azurescens		*Calliptamus barbarus*	*Omocestus panteli*
Locusta migratoria		*Dociostaurus jagoi*	*Schistocerca gregaria*
Oedipoda coerulescens		*D. maroccanus*	*Mus musculus*

como embriones en desarrollo obtenidos en nuestro laboratorio a partir de cruzamientos controlados. La Tabla 1 muestra un resumen del material biológico utilizado. Adicionalmente, hemos estudiado machos adultos de 14 especies de saltamontes, dos de ellas estrechamente relacionadas con *E. plorans*, y de ratón (Tabla 1).

Obtención y fijación del material de estudio

Los individuos adultos capturados en las poblaciones naturales fueron preparados para las técnicas citogenéticas y moleculares al llegar al laboratorio, o bien se mantuvieron en cultivo, en cajas de madera, a una temperatura de 28^{o}C para la obtención de embriones. A los machos, se les extrajeron los testículos, uno de los cuales se fijó en etanol:ácido acético (3:1) y se almacenó a 4^{o}C hasta su utilización en los estudios citogenéticos. El cuerpo, conteniendo el otro testículo, se sumergió en nitrógeno líquido para su congelación y se mantuvo a -80^{o}C para la posterior extracción del ADN genómico y el ARN total.

Las hembras se diseccionaron y se les extrajo una parte de las ovariolas que fue sumergida en colchicina al 2 % (en solución salina) durante 2 horas, con objeto de favorecer la visualización de cromosomas en metafase mitótica. Posteriormente, fueron fijadas en etanol:ácido acético (3:1) y se conservaron a 4^{o}C hasta su uso para los estudios citogenéticos. El resto de las ovariolas y el cuerpo se congelaron en nitrógeno líquido, y se mantuvieron a -80^{o}C hasta el momento de las extracciones de ácidos

nucleicos.

Los embriones en desarrollo se obtuvieron incubando puestas de huevos de hembras grávidas a 28^{o}C en una cámara húmeda durante diez días. Transcurrido ese tiempo, diseccionamos los huevos para extraer los embriones que fueron luego sumergidos en una solución de colchicina al 0.05 % en solución salina, durante 2 horas. Posteriormente, fueron sometidos a un choque osmótico con agua destilada y fijados de forma similar a las gónadas de adultos, para su posterior estudio citogenético.

Técnicas citogenéticas

Obtención de preparaciones cromosómicas

Las preparaciones cromosómicas de gónadas de machos y hembras adultas fueron realizadas disgregando un par de folículos testiculares u ovariolas sobre un portaobjetos en una gota de ácido acético al 50 %, con un macerador de extremo plano. Luego se coloca un cubreobjetos, y tras eliminar las burbujas de aire, presionando sobre el cubreobjetos con la punta de una aguja enmangada, se aplasta el material presionando con el dedo pulgar para disponer las células en un mismo plano. Posteriormente, las preparaciones se introducen en nitrógeno líquido durante unos segundos y se elimina el cubreobjetos con ayuda de una cuchilla de afeitar.

Para inmunofluorescencia, las preparaciones se realizaron con material

fresco que se fijó en paraformaldehido al 2 % en PBS con 0,05 % de Tween 20, disgregando dos folículos testiculares sobre un porta en una gota del mismo fijador.

Las preparaciones de cromosomas mitóticos de embriones se realizaron siguiendo el protocolo descrito por Meredith (1969). En un tubo eppendorff de 1.5 ml, se homogeniza un embrión en 20µl de ácido acético al 70 % y se depositan varias gotas del homogeneizado en un portaobjetos colocado sobre un termoblock a 60^{o}C, para conseguir la fijación del material al portaobjetos. Estas preparaciones se utilizaron en fresco o fueron congeladas para su uso posterior.

Determinación del número de cromosomas B

El número de cromosomas B que portaban los machos se determinó estudiando preparaciones de folículos testiculares teñidas con orceína lactopropiónica al 2 % (**?**), aunque en algunos casos fue necesario realizar bandeo C (Fig. 2a) (Camacho et al., 1991) para distinguir el tipo de B. Esta última técnica permite poner de manifiesto la presencia de la heterocromatina constitutiva en los cromosomas, pudiendo así caracterizar el número y tipo (variante) de cromosoma B que porta cada individuo, ya que muestran una mayor intensidad de tinción debido a la naturaleza heterocromática de los mismos.

La presencia y número de cromosomas B en las hembras de 2012 fue determinada mediante bandeo C en preparaciones de hemolinfa de especímenes vivos, siguiendo la metodología propuesta por Cabrero et al. (2006). Este protocolo permite determinar fácilmente la presencia

de cromosomas B en esta especie ya que aparecen como un bloque intensamente teñido en los hemocitos interfásicos, como puede apreciarse en las Figuras 2b y 2c, que corresponden, respectivamente, a una hembra no portadora de cromosomas B y a una hembra con un cromosoma B. Durante 2013, el número de cromosomas B en las hembras se determinó mediante bandeo C en las ovariolas tratadas con colchicina, observándose el cromosoma B como un cromosoma supernumerario muy heterocromático (Figura 2d).

Los embriones que portaban cromosomas B fueron identificados tiñendo las preparaciones cromosómicas con 2 µg/ml 4',6 diamidino-2-phenylindole (DAPI) lo que permite revelar la presencia del cromosoma B como el que muestra la mayor superficie $DAPI^{+}$.

Hibridación *in situ* fluorescente con amplificación de la señal por tiramidas

La localización cromosómica del gen de la subunidad *CAP-G* del complejo de condensina I (capítulo 4) se ha llevado a cabo mediante hibridación in situ fluorescente con amplificación de la señal por tiramidas (FISH-TSA) sobre preparaciones cromosómicas de embriones, siguiendo los protocolos descritos previamente por Krylov et al. (2008, 2007) y Funkhouser-Jones et al. (2015), con ligeras modificaciones que incluyen un tratamiento previo de las preparaciones con RNasa (100 µg/ml in 2XSSC) durante 2 horas a 37^{o}C y la utilización de una mezcla de dNTPs, que contenía 1mM dATP, 1mMdCTP, 1mMdGTP, 0,65mM dTTP and 0.01mM digoxigenin -11-dUTP, en el marcado de la sonda

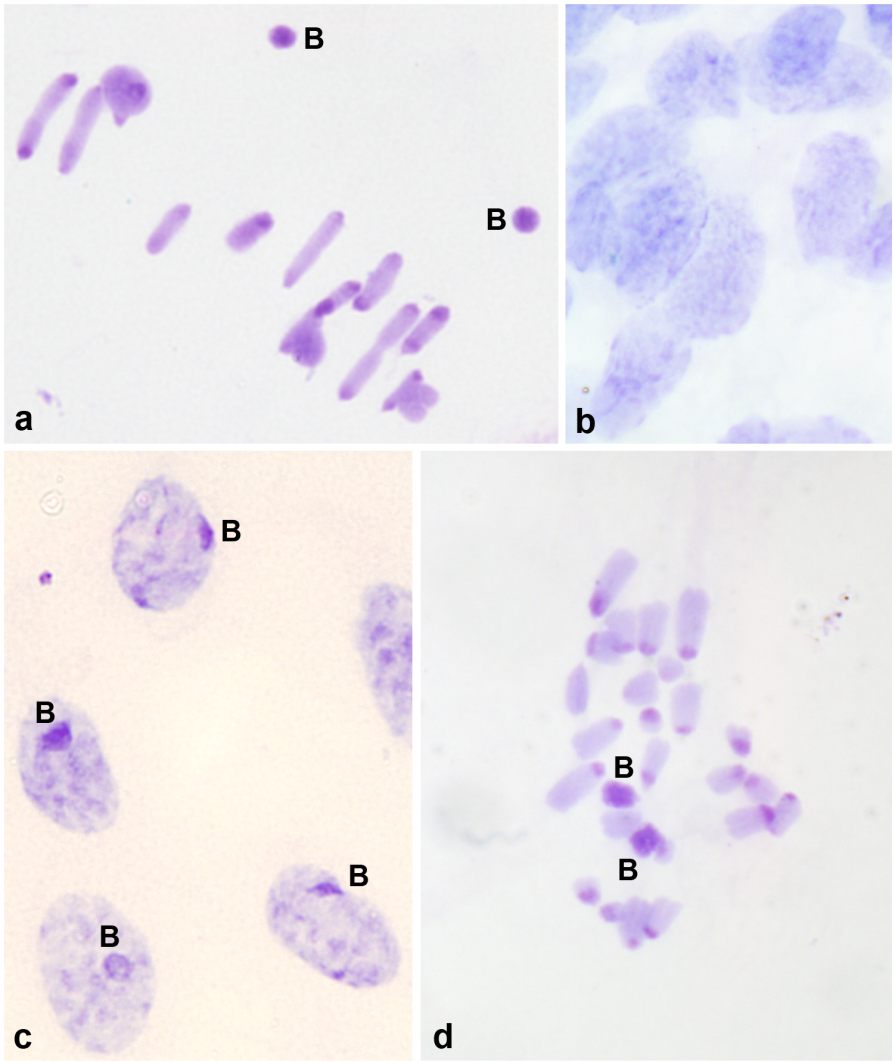

Figura 2: Determinación del número de cromosomas B en especímenes de *E. plorans* mediante bandeo C de preparaciones de folículos testiculares (a), frotis de hemolinfa (b, c) y ovariolas tratadas con colchicina (d).

que fue realizado a 37^{o}C durante 20h con el kit Decalabel DNA labeling kit (Thermo Scientific). (Ver el capítulo 4 para más detalles).

Inmunofluorescencia

Para la inmunofluorescencia, utilizamos el protocolo descrito en (Cabrero et al., 2007b,a). Brevemente, éste consiste en sumergir las preparaciones en PBT y realizar varios lavados, tras lo que se realiza un bloqueo con BSA al 10 % en PBT, incubando durante 1 hora en cámara húmeda, tras lo que se añade el anticuerpo primario. Tras incubar durante toda la noche, se realizan varios lavados con PBT y se añade el anticuerpo secundario y, después de 1 hora de incubación en la oscuridad, se vuelve a lavar en PBT y se tiñe con DAPI para identificar los cromosomas. En el capítulo 1 se utilizan anticuerpos para las proteínas γH2AX, Ku70, Ku80 y RAD51, implicadas en la reparación de roturas del ADN de doble cadena y para α-tubulina, que permite marcar el citoesqueleto. Los anticuerpos utilizados y la zona de la proteína con la que reaccionan pueden consultarse en el capítulo 1.

Tinción de Feulgen

La tinción de Feulgen se utilizó en el capítulo 1 para estimar la cantidad de ADN en las espermátidas normales y en las macroespermátidas mediante la técnica de análisis densitométrico de imagen de Feulgen (FIAD), con el protocolo descrito en Ruiz-Ruano et al. (2011), ya que tiñe el material estequiométricamente en relación a la cantidad de ADN.

Brevemente, las preparaciones se tratan con HCl 5N durante 20 minutos y con reactivo de Schiff durante 2 horas, después de lo cual se realizan lavados con agitación en agua sulfurosa.

Técnicas moleculares

Extracción de ácidos nucleicos

Para la secuenciación Illumina realizada (capítulo 3) el ADN que se utilizó fue extraído de una pata saltadora, para evitar contaminaciones procedentes del aparato digestivo. El material congelado se pulverizó en nitrógeno líquido, con ayuda de un mortero de porcelana. En los demás casos, la extracción de ácidos nucleicos de cuerpos de saltamontes se realizó pulverizando un cuerpo o hemicuerpo en nitrógeno líquido, con ayuda de un mortero de porcelana. Aproximadamente 1/4 del material pulverizado obtenido se utilizó para extraer ADN, usando el kit GenElute(R) Mammalian Genomic DNA minipreps (Sigma) y los restantes 3/4 para extraer ARN con el kit Real Total RNA Spin Plus kit (Durviz). Las gónadas se disgregaron con un mini homogeneizador en tubos eppendorf de 1,5 mL, sumergidos en Quiazol, para su extracción con el kit RNeasy Lipid Tissue Mini Kit (Qiagen).

En ambos protocolos de extracción de ARN se realizó un tratamiento con DNasa (DNase Amplification Grade I, Sigma) en la membrana de la columna de extracción. Posteriormente, las muestras de ARN de

cuerpos fueron sometidas a un segundo tratamiento con DNasa (kit REALSTAR, Durviz), que no fue necesario en el caso de las muestras de gónada. La integridad del ARN y la ausencia de contaminación con ADN genómico se comprobó en un gel desnaturalizante MOPS-agarosa (1 %) con formaldehido 2M. La ausencia de ADN contaminante se comprobó también mediante un control por PCR, basado en la ausencia de amplificación de genes de ARN ribosómico o histonas en las muestras del ARN aislado. En el caso del ARN con el que se realizó la hibridación de microarrays (capítulo 5), la calidad y ausencia de contaminación se chequeó con un bioanalizador Agilent 2100.

La cuantificación de las muestras de ácidos nucleicos, así como la obtención de medidas de calidad de las mismas mediante la medición de los ratios de absorbancia a 260 y 280 nm, se realizó en un nanoespectrofotómetro Tecan's Infinite 200 NanoQuant.

Retrotranscripción

Obtención del ADN complementario (ADNc) utilizado en las técnicas de PCR y qPCR

En el capítulo 1, la obtención de ADNc de cadena simple a partir de ARN (retrotranscripción) se realizó a partir de 100 ng de ARN total con el kit SuperScript III First-Strand Synthesis SuperMix kit (Invitrogen), usando cebadores oligo dT y siguiendo las instrucciones proporcionadas por el fabricante en el manual de usuario. En los capítulos 2, 3 y 4, la retrotranscripción se realizó a partir de 100 ng de ARN total por muestra, con el kit PrimeScript™ RT reagent Kit, Perfect Real Time

(Takara), usando una combinación de hexámeros random y oligo dT. En cada reacción de retrotranscripción se incluía un control negativo (sin ARN) para asegurar la ausencia de contaminación en los reactivos.

Reamplificación y obtención del ADNc de cadena doble utilizado para la hibridación sobre microarrays

Para obtener el ADNc de doble cadena necesario para la hibridación sobre microarrays (capítulo 5) se utilizó el kit TransPlex Whole Transcriptome Amplification Kit (Sigma), siguiendo las instrucciones del fabricante y las recomendaciones para la hibridación en NimbleGen Arrays User's Guide (Gene Expression Arrays v6.0). La cantidad de ARN que se utilizó por muestra fue de 62.5 ng, que era la cantidad máxima disponible para la muestra menos concentrada.

Reacción en cadena de la polimerasa (PCR)

La PCR se ha utilizado en esta Tesis Doctoral para amplificar fragmentos de ADN y ADNc para la obtención de la sonda empleada en la FISH-TSA y para la secuenciación Sanger. Estas amplificaciones se han realizado con el kit Horse-PowerTaq DNA polymerase (Canvax). Las secuencias de los cebadores utilizados y el método con el que se han diseñado pueden consultarse en el apartado de Material y Métodos de cada capítulo.

En general, los programas consisten en una fase inicial de desnaturalización (94^{o}C durante 5 minutos), entre 30 y 40 ciclos de desnaturalización (94^{o}C, 30 segundos), hibridación (45-60^{o}C,

dependiendo de la temperatura específica de cada pareja de cebadores, durante 30 segundos) y extensión (72^{o}C, con tiempo variable en función de la longitud del fragmento a amplificar) y una fase final de extensión a 72^{o}C durante 5 minutos. La visualización del producto de PCR se ha realizado en geles de agarosa/TBE al 1,5 % con SYBR® Safe (Invitrogen). Cuando ha sido necesario, las bandas obtenidas se han recortado y limpiado con el kit GenElute® Gel Extraction (Sigma) y/o el producto de PCR se ha limpiado con el kit GenElute® PCR Clean-Up (Sigma).

Clonación

En el capítulo 2, se detalla cómo los fragmentos obtenidos por PCR con cebadores degenerados se clonaron antes de la secuenciación, usando TOPO TA cloning kit (Invitrogen).

PCR cuantitativa en tiempo real (qPCR)

En las reacciones de qPCR realizadas para estimar la abundancia de un determinado gen en el genoma de *E. plorans* (capítulos 3 y 4), el ADN genómico se utilizó diluido a una concentración de 5ng/µl. Para las reacciones de qPCR llevadas a cabo para cuantificar la expresión relativa de un determinado gen en relación al número de cromosomas B (capítulos 2, 3 y 4), el ADNc obtenido a partir de 100 ng de ARN era diluido en proporción 1:10 en agua libre de nucleasas. En el capítulo 1, el ADNc se utilizó diluido en una proporción 1:5. Las reacciones contenían 5 µl de ADNc, 5 µl de SensiMix™ SYBR Kit (Bioline) y 2.5 µl de cada

primer, a una concentración de 2.5 µM en un volumen total de 15 µl. Para minimizar la variación técnica en las reacciones se usaron pipetas electrónicas (Eppendorf Research®Pro) y además se hicieron dos réplicas técnicas de cada reacción. En cada experimento se incluyó un control negativo para cada pareja de cebadores y una muestra de referencia, llamada calibrador que consiste en una mezcla de ADN/ADNc que es siempre el mismo y que se usa para normalizar la variación entre placas.

El programa de qPCR consta de una fase de desnaturalización inicial (95^{o}C, 10 minutos), seguida de 40 ciclos de 94^{o}C/15'', 58-60^{o}C/15'' y 72^{o}C/15'', con una lectura de placa en cada ciclo. Al final del 40^o ciclo, se incluyó una curva de disociación desde 72 hasta 95^{o}C para asegurar la especificidad de la reacción. Los datos de fluorescencia se midieron y procesaron con Opticon Monitor 3.1 (Bio-Rad Laboratories, Inc).

Estima de la eficiencia de la reacción

Para cada pareja de cebadores, se realizó una curva estándar a partir de 5 alícuotas de diluciones seriadas 1:10 de ADNc. El factor de amplificación en cada ciclo (i.e., el número de moléculas que se generan a partir de una en cada ciclo de PCR, normalmente cercano a 2) y el porcentaje de eficiencia de la reacción se calculó a partir de la pendiente de la curva estándar.

Genes de referencia

Para la selección de los genes de referencia más adecuados en cada caso, amplificamos un total de 7 genes de expresión constitutiva en *E. plorans*, utilizando los cebadores descritos en Van Hiel et al. (2009) y Chapuis et al. (2011). La selección de los más adecuados para cada tipo de muestra se hizo con el programa GeNorm (Vandesompele et al., 2002).

La selección de genes de referencia para la comparación entre las muestras de 2013 mediante qPCR aparece detallada en el capítulo 2, tabla 4. Para los capítulos 3 y 4 se utilizaron los mismos genes de referencia.

Cuantificación relativa de los resultados de la amplificación de las muestras

En el capítulo 1, el cálculo de la cuantificación relativa se hizo en relación a las muestras control, y normalizada por la media geométrica de dos genes de referencia. En lo demás casos (capítulos 2, 3 y 4), los cálculos de cuantificación relativa tanto de expresión como de abundancia en el genoma se han hecho referidos a la muestra calibradora, con un script de R propio, basados en el método de Pfaffl (2001)y, en el caso de los valores de expresión, normalizados respecto a la media geométrica de los valores de los genes seleccionados por GeNorm (Vandesompele et al., 2002) en cada tipo de muestra (ver capítulo 2 tabla 4).

Secuenciación Sanger

Las secuenciaciones se han realizado en el Servicio de Análisis Genético del Departamento de Genética de la Universidad de Granada y en Macrogen (Macrogen Europe, Amsterdam, Holanda).

Secuenciación Illumina

Las secuenciaciones Illumina de ARN (una hembra 0B y una hembra 1B) y ADN (un macho 0B y un macho 4B) de los individuos de *E. plorans* de la población de Torrox que se han utilizado en los capítulos 3, 4 y 5, se han secuenciado en Macrogen (Macrogen Inc., Seoul, Corea del Sur) con tecnología Illumina HiSeq2000. Estas secuencias están disponibles en la base de datos NCBI-SRA con los siguientes números de acceso: SRR2970625 (gDNA_0B), SRR2970627 (gDNA_4B), SRR2969416 (RNA_0B) and SRR2969417 (RNA_1B).

Hibridación de microarrays

El diseño de los arrays custom (3x1.4M Gene Expression arrays) utilizados en el capítulo 4 fue realizado por Nimblegen. El marcado, la hibridación, el lavado y el escaneado de los arrays se realizó siguiendo el protocolo proporcionado por la empresa (NimbleGen Arrays user's guide for gene expression arrays) en un escáner de microarray MS 200.

Knockdown mediante ARN interferente (iARN)

El ARN de doble cadena para el knockdown del gen de la proteína Ku70 (capítulo 1) se sintetizó a partir de un producto de PCR siguiendo el protocolo del kit Ambion T7 RNA polymerase transcription. Las inyecciones con el ARNi se realizaron intra-abdominalmente, a nivel de los testículos.

Bioinformática

Diseño de cebadores

Las herramientas para el diseño de los cebadores que se han utilizado en esta Tesis Doctoral han sido Primer3 (Koressaar y Remm, 2007; Untergasser et al., 2012), FastPCR (Kalendar et al., 2009) y NetPrimer (Premier Biosoft).

Alineamiento y edición de secuencias

Se ha utilizado BLAST (Altschul et al., 1990) para alineamientos pairwise, MAFFT (Katoh et al., 2002) para alineamiento múltiple y Geneious 4.8 Drummond et al. (2009) para la edición de las secuencias y la elaboración de figuras.

Herramientas de análisis de datos obtenidos por secuenciación de alto rendimiento

El control de calidad y la estadísticas de las lecturas Illumina se realizó con fastqc (Andrews, 2010), el *trimming* se realizó con Trimmomatic (Bolger et al., 2014) y la identificación y eliminación de posibles contaminaciones se realizó con deconseq (Schmieder y Edwards, 2011).

Las lecturas Illumina RNAseq de hembras de *E. plorans* se ensamblaron con Trinity (versión 20131111, Grabherr et al., 2011) siguiendo las directrices propuestas por Haas et al. (2013) y obteniendo un transcriptoma *de novo*.

La predicción computacional de secuencias codificantes en el transcriptoma *de novo* se realizó con Transdecoder (Haas et al., 2013). La anotación funcional de estas secuencias se realizó siguiendo el flujo de trabajo de Trinotate (Grabherr et al., 2011), haciendo una búsqueda por homología con BLASTX y BLASTP (Altschul et al., 1990, 1997) en las bases de datos UniProtKB/Swiss-Prot y UniProt Reference Clusters (UniRef) con los parámetros por defecto. También se buscaron dominios proteicos con HMMER (Finn et al., 2011) y PFAM (Finn et al., 2013). La clasificación funcional de estas secuencias se realizó mediante dos métodos estandarizados: Gene Ontology (GO; Ashburner et al., 2000) y Eukaryotic Orthologous Groups (KOG; Tatusov et al., 2003). Se realizó una búsqueda de elementos transponibles en el ensamblaje con RepeatMasker (Smit et al., 1996) en una base de datos obtenida de Repbase (Jurka et al., 2005) con elementos transponibles descritos en *Locusta migratoria*. Para el estudio de expresión diferencial (capítulo 5), se mapearon sobre este transcriptoma ambas librerías de ARN (0B y 1B

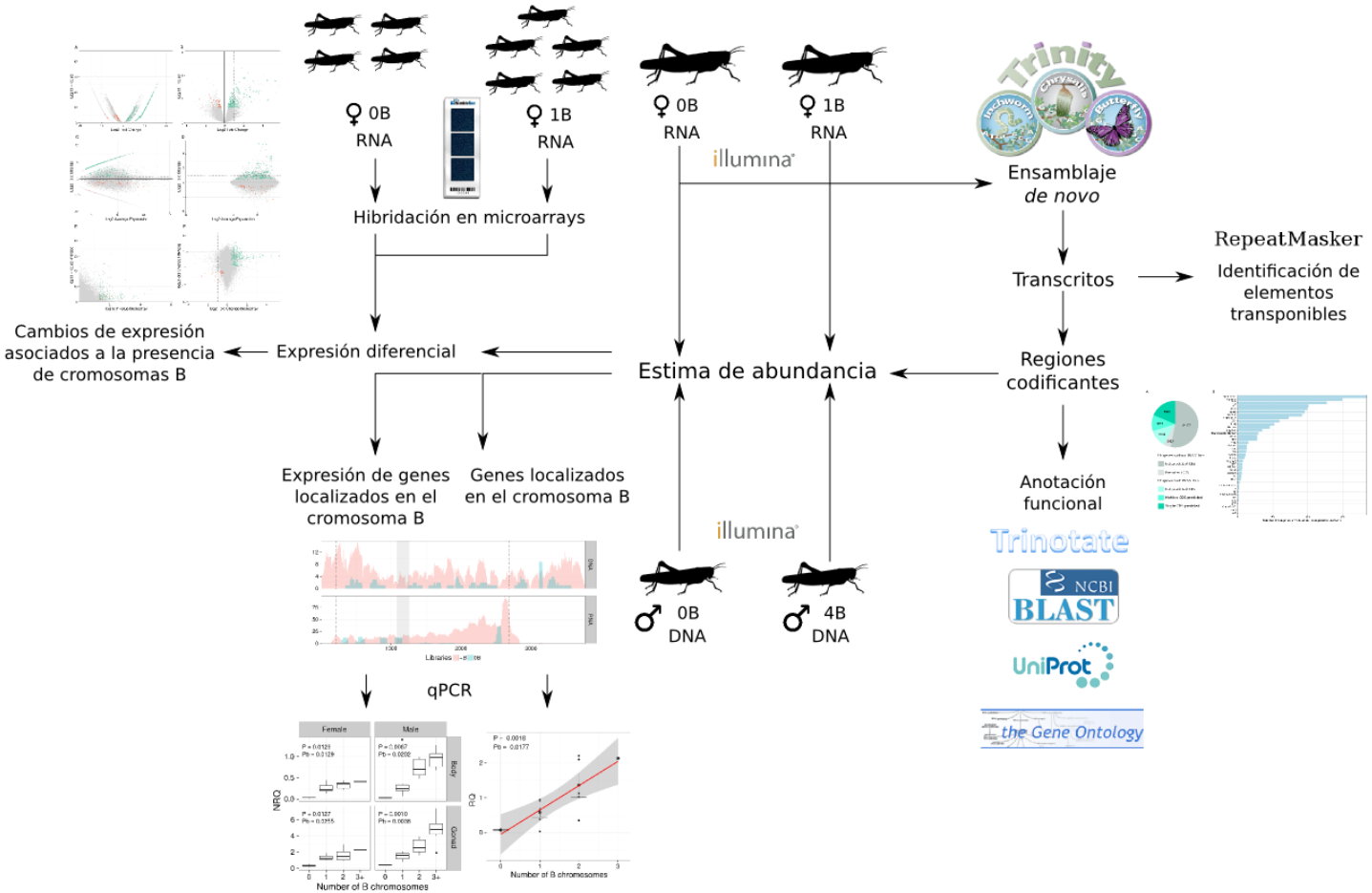

Figura 3: Resumen gráfico de los métodos informáticos utilizados para el análisis de datos de secuenciación de alto rendimiento (capítulos 3 y 5) e hibridación de microarrays, utilizados para la localización de genes codificadores de proteínas en los cromosomas B y para detectar cambios de expresión génica asociada a la presencia de cromosomas B.

de hembras, tabla 1) con Bowtie2 (Langmead et al., 2009). La estima de abundancia por transcrito se hizo con RSEM (Li y Durbin, 2009) y la estadística para la expresión diferencial se hizo a partir de los valores de FPKM con edgeR (Robinson et al., 2010).

Para la identificación de genes en los cromosomas B (capítulo 3), se mapearon sobre el transcriptoma *de novo* las librerías de DNA de machos de Torrox (0B y 4B) con SSAHA2 (Ning et al., 2001), y la cobertura por posición nucleotídica se calculó con un script custom y se normalizó como se describe en el capítulo 3.

La predicción de los sitios de unión exón-exón en los transcritos de los genes potencialmente localizados en el cromosoma B (capítulo 3) se realizó con el programa exonerate (Slater y Birney, 2005).

En la figura 3 se puede encontrar un resumen gráfico de los métodos informáticos utilizados para el análisis de datos de secuenciación de alto rendimiento en los capítulos 3 y 5 para localizar genes codificadores de proteínas en los cromosomas B y para detectar cambios de expresión génica.

Normalización y expresión diferencial con microarrays

Para la normalización RMA (robust multi-array average) de los datos de microarray se utilizó el paquete oligo (Carvalho y Irizarry, 2010). Las diferencias de expresión se estudiaron con el paquete limma (Smyth, 2005).

Bibliografía

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., y Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., y Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., y Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25(1):25–29.

Bolger, A. M., Lohse, M., y Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, page btu170.

Cabrero, J., Manrique-Poyato, M. I., y Camacho, J. P. M. (2006). Detection of B chromosomes in interphase hemolymph nuclei from living specimens of the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 114(1):66–69.

Cabrero, J., Palomino-Morales, R. J., y Camacho, J. P. M. (2007a). The DNA-repair Ku70 protein is located in the nucleus and tail of elongating spermatids in grasshoppers. *Chromosome Research*, 15(8):1093–1100.

Cabrero, J., Teruel, M., Carmona, F. D., y Camacho, J. P. M. (2007b). Histone H2AX phosphorylation is associated with most meiotic events in grasshopper. *Cytogenetic and Genome Research*, 116(4):311–315.

Camacho, J. P. M., Cabrero, J., Viseras, E., López-León, M. D., Navas-Castillo, J., y Alche, J. D. (1991). G banding in two species of grasshopper and its relationship to C, N, and fluorescence banding techniques. *Genome*, 34(4):638–643.

Carvalho, B. S. y Irizarry, R. A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics (Oxford, England)*, 26(19):2363–2367.

Chapuis, M.-P., Tohidi-Esfahani, D., Dodgson, T., Blondin, L., Ponton, F., Cullen, D., Simpson, S. J., y Sword, G. A. (2011). Assessment and validation of a suite of reverse transcription-quantitative PCR reference genes for analyses of density-dependent behavioural plasticity in the Australian plague locust. *BMC molecular biology*, 12:7.

Drummond, A. J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., y Wilson, A. (2009). Geneious 4.8. *Biomatters, Auckland, New Zealand*.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., y Mistry, J. (2013). Pfam: the protein families database. *Nucleic Acids Research*, page gkt1223.

Finn, R. D., Clements, J., y Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*, 39(Web Server issue):W29–37.

Funkhouser-Jones, L. J., Sehnert, S. R., Martínez-Rodríguez, P., Toribio-Fernández, R., Pita, M., Bella, J. L., y Bordenstein, S. R. (2015). Wolbachia co-infection in a hybrid zone: discovery of horizontal gene transfers from two Wolbachia supergroups into an animal genome. *PeerJ*, 3:e1479.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., y Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7):644–652.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., Leduc, R. D., Friedman, N., y Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8):1494–1512.

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., y Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, 110(1-4):462–467.

Kalendar, R., Lee, D., y Schulman, A. H. (2009). FastPCR software for PCR primer and probe design and repeat search. *Genes, Genomes and Genomics*, 3(1):1–14.

Katoh, K., Misawa, K., Kuma, K.-i., y Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14):3059–3066.

Koressaar, T. y Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics (Oxford, England)*, 23(10):1289–1291.

Krylov, V., Tlapakova, T., y Macha, J. (2007). Localization of the single copy gene Mdh2 on *Xenopus tropicalis* chromosomes by FISH-TSA. *Cytogenetic and Genome Research*, 116(1-2):110–112.

Krylov, V., Tlapakova, T., Macha, J., Curlej, J., Ryban, L., y Chrenek, P. (2008). Localization of human coagulation factor VIII (hFVIII) in transgenic rabbit by FISH-TSA: identification of transgene copy number and transmission to the next generation. *Folia Biologica*, 54(4):121.

Langmead, B., Trapnell, C., Pop, M., y Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3):R25.

Li, H. y Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754–1760.

Meredith, R. (1969). A simple method for preparing meiotic chromosomes from mammalian testis. *Chromosoma*, 26(3):254–258.

Ning, Z., Cox, A. J., y Mullikin, J. C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Research*, 11(10):1725–1729.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29(9):e45.

Robinson, M. D., McCarthy, D. J., y Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, 26(1):139–140.

Ruiz-Ruano, F. J., Ruiz-Estévez, M., Rodríguez-Pérez, J., López-Pino, J. L., Cabrero, J., y Camacho, J. P. M. (2011). DNA amount of X and B chromosomes in the grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenetic and Genome Research*, 134(2):120–126.

Schmieder, R. y Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PloS One*, 6(3):e17288.

Slater, G. S. y Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6(1):31.

Smit, A. F., Hubley, R., y Green, P. (1996). *RepeatMasker Open-3.0.* Available at http://www.repeatmasker.org/.

Smyth, G. K. (2005). Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor*, pages 397–420. Springer.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., y Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, 4:41.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., y Rozen, S. G. (2012). Primer3–new capabilities and interfaces. *Nucleic Acids Research*, 40(15):e115.

Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. V., y Simonet, G. (2009). Identification and validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. *BMC Molecular Biology*, 10:56.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., y Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7):RESEARCH0034.

1

The Ku70 DNA-repair protein is involved in centromere function in a grasshopper species[I]

Josefa Cabrero, Mohammed Bakkali, Beatriz Navarro-Domínguez, Francisco J. Ruíz-Ruano, Rubén Martín-Blázquez, María Dolores López-León, Juan Pedro M. Camacho

Departamento de Genética, Universidad de Granada

Abstract The Ku70 protein is involved in numerous cell functions, the non-homologous end joining (NHEJ) DNA repair pathway being the best known. Here we report a novel function for this protein in the grasshopper *Eyprepocnemis plorans*. We observed the presence of large Ku70 *foci* on the centromeres of meiotic and mitotic chromosomes during the cell cycle stages showing the highest centromeric activity (i.e. metaphase and anaphase). The fact that colchicine treatment prevented centromeric location of Ku70, suggests a microtubule-dependent

[I]Published: Chromosome Research, 21(4), 393-406 (2013).

centromeric function for Ku70. Likewise, the absence of Ku70 at metaphase-anaphase centromeres from three males whose *Ku70* gene had been knocked down using RNAi, and the dramatic increase in the frequency of polyploid spermatids observed in these males, suggest that the centromeric presence of Ku70 is required for normal cytokinesis in this species. The centromeric function of Ku70 was not observed in 14 other grasshopper and locust species, nor in the mouse, thus suggesting that it is an autapomorphy in *E. plorans*.

Introduction

Eukaryotic cells repair double stranded DNA breaks (DSB) via two pathways: homologous recombination (HR) and non-homologous end-joining (NHEJ), both playing a role in maintaining chromosomal DNA throughout the cell cycle (Takata et al., 1998). Among the several proteins involved in these pathways, Ku is a heterodimer of two polypeptides of about 69 and 83 kD, called Ku70 and Ku80 (or Ku86) respectively. It is a multifunctional protein involved directly or indirectly in many important cellular metabolic processes, such as DNA DSB repair through the NHEJ pathway, V(D)J recombination of immunoglobulins and T-cell receptor genes, immunoglobulin isotype switching, DNA replication, transcription regulation, heat-shock response, telomere maintenance, and cell-cycle regulation (Tuteja and Tuteja, 2000). Such a variety of functions is derived from the ability of Ku to bind to DNA,

RNA, and proteins. The DNA repair function of this protein is performed in conjunction with the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), with which it forms the DNA-dependent protein kinase (Gell and Jackson, 1999). It has been suggested that the monomer Ku70 (on which this work focuses) acts as a switch between the HR and NHEJ DSB repair pathways, so that downregulation of Ku70 in early meiosis (pre-leptotene-zygotene) may ensure that meiotic DSBs are processed by homologous recombination (Goedecke et al., 1999). Phylogenetic analysis of Ku70 orthologs in fungi, plants and animals has revealed that the Ku70 protein is highly conserved among eukaryotes (Pöggeler and Kück, 2006).

Being components of the nuclear matrix, Ku proteins are predominantly nuclear, where they appear diffusely distributed throughout the nucleoplasm (Koike et al., 1999). However, in addition to its DNA-repair function in the nucleus, Ku70 has been located in the cytoplasm, where it can inhibit Bax-mediated apoptosis (Sawada et al., 2003). It has also been found in the plasma membrane, where it can participate in heterologous and homologous cell adhesion (Koike, 2002) as well as fibronectin binding (Monferran et al., 2004a), and can interact with metalloprotease 9 (MMP-9) (Monferran et al., 2004b). Furthermore, a binding site for Ku70/80 has been found in LINE-1 retrotransposons, suggesting that these proteins may also participate in transposition (Katz et al., 2005).

Ku moves rapidly throughout the nucleus, showing a transient high-flux association with nuclear substrates, including both DNA and the nuclear matrix —the latter associating only with the Ku70 monomer (Rodgers et al., 2002). As if the Ku functional picture wasn't complex enough,

several reports suggest that this protein may have additional functions that are independent of the DNA-dependent protein kinase holoenzyme (Gao et al., 1998; Gu et al., 2000). Evenmore, the Ku70 monomer was suggested to possibly have functions that are independent of Ku80 (Koike et al., 2001). Consistently, it has been reported that Ku70 and Ku80 monomers can show independent location in the nucleus (Bertinato et al., 2001). It has also been shown that Ku70, but not Ku80, interacts with heterochromatin protein 1α (HP1α), a protein known to be associated with telomeres and heterochromatin (Song et al., 2001). Likewise, Ku70 seems to be involved in the internalization of the obligate intracellular bacterium *Rickettsia conorii* (Martinez et al., 2005). It has also been suggested that Ku80 has a Ku70-independent DNA DSB repair function, in addition to the one dependent on Ku70 (Koike and Koike, 2005b).

Throughout our several studies on supernumerary (B) chromosomes (a kind of parasitic chromosomes), we consistently came across several effects of these chromosomes in the grasshopper *Eyprepocnemis plorans*, e.g. decreasing egg fertility (Zurita et al., 1998; Muñoz et al., 1998) or increasing the frequency of chiasmata (the cytological manifestation of meiotic recombination) in the host genome (Camacho et al., 1980, 2002). To determine the molecular causes of chiasma frequency changes, we analyzed whether the presence of B chromosomes is paralleled by an increase in the frequency of double-strand breaks (DSBs) during meiotic prophase I, since these are required for HR. We thus investigated the presence, in B-carrying and B-lacking individuals, of the Ser-139 phosphorylated form of the H2AX histone (γH2AX), a modification that occurs in the vicinity of DSBs and thus serves as signal for the DNA-repair events (Zhou and Elledge, 2000; Cabrero et al., 2007b).

Although we did not observe qualitative differences for γH2AX between +B and –B individuals at prophase I cells, we noticed the presence of conspicuous *foci* of this protein at centromeric regions of all chromosomes during metaphase I. This hinted at the interesting possibility that centromeres might be experiencing frequent DSBs during chromosome segregation. To answer this question, we engaged in the current work in which we analyze the location of several proteins involved in the two main pathways of DNA DSB repair. i.e. Rad51 (involved in the HR pathway) and Ku70 and Ku80 (involved in the NHEJ pathway). We especially focused on any possible location of these proteins at the centromeric region and whether, in addition to their DNA DSB repair role, such centromeric location means an involvement of these proteins in the centromeric function. We initially based our work on immunofluorescence microscopy and, in order to verify the immunofluorescence results, we carried out RNAi-mediated knockdown of the *Ku70* gene in *E. plorans*.

In this paper we show that, in the grasshopper *E. plorans*, Rad51 and Ku80 are not found in centromeric regions, but Ku70 is conspicuously present at centromeres during diakinesis, metaphase I, anaphase I, metaphase II and anaphase II, as well as at spermatogonial and embryonic mitotic metaphases. However, these large centromeric *foci* of Ku70 failed to appear after RNAi-mediated knockdown of the *Ku70* gene, and also after colchicine treatment, suggesting a microtubule-dependent centromere-related function of Ku70. Unexpectedly, the centromeric location of Ku70 seems to be exclusive to *E. plorans*, since 14 other grasshopper and locust species, as well as the domestic mouse, failed to show it.

Materials and methods

Materials

Adult males of the grasshopper *Eyprepocnemis plorans* were collected at a B-carrying population (Torrox, Málaga, Spain), and two B-lacking populations (Socovos and El Gallego, Albacete, Spain). Developing embryos of *E. plorans* were obtained by incubating eggs laid by gravid females bred in the laboratory. We also analyzed the grasshopper species *Heteracris adspersa* (Agramon, Albacete, Spain) and *H. litoralis* (Torrox, Málaga, Spain), which are close relatives of *E. plorans* also belonging to the subfamily Eyprepocneminae. In addition, we analyzed other distantly related grasshopper and locust species collected in the Granada province (Spain), including *Aiolopus strepens*, *Locusta migratoria*, *Oedipoda coerulescens*, *Sphingonotus azurescens* and *Acrotylus insubricus* (subfamily Oedipodinae), *Calliptamus barbarus* (subfamily Calliptaminae), *Dociostaurus jagoi*, *D. maroccanus*, *Chorthippus vagans*, *Ch. jacobsi*, *Omocestus panteli* (subfamily Gomphocerinae) and *Schistocerca gregaria* (subfamily Cyrtacanthacridinae) (this latter from our laboratory colony). Furthermore, we analyzed the domestic mouse. No specific permits were required for the field studies. The locations sampled were not privately owned or protected in any way, and this study did not involve endangered or protected species.

Slide preparations and inmunofluorescence

Fixation and squash preparations of grasshopper testis follicles and embryos as well as the mouse seminiferous tubules, were performed as described in (Cabrero et al., 2007b). In brief, all materials were fixed in freshly prepared 2% paraformaldehyde in PBS (phosphate-buffered saline) containing 0.05% Tween 20 for 15 min. Testis preparations were made by squashing two testis tubules in a small drop of the paraformaldehyde fixative. After immersion into liquid nitrogen and coverslip removal, the slides were immediately transferred into PBS. Cytological preparations of mouse spermatocytes were made following the protocol for grasshopper embryos described in (Camacho et al., 1991). Immunofluorescence analysis of γH2AX, Rad51, Ku70, Ku80 and α-tubulin was performed following the methods described in (Cabrero et al., 2007a,b). For γH2AX immunofluorescence, we used a monoclonal mouse antibody (Upstate) raised against amino acids 134-142 of the human histone H2AX (Cabrero et al., 2007b; Paull et al., 2000). For Ku70 we used H-308 (Santa Cruz Biotechnology Inc., CA, USA), a rabbit antibody raised against aminoacids 302-609 of the human Ku70. For Ku80, we used sc-1485 (Santa Cruz Biotechnology Inc., CA, USA) a goat antibody raised against a peptide mapping at the C-terminus of the Ku80 of mouse origin. For Rad51 we employed R-1528 (Sigma), a rabbit antibody raised against aminoacids 24-41 of the human Rad51. For α-tubulin, we used (B-5-1-2):sc-23948 (Santa Cruz Biotechnology Inc., CA, USA), a mouse monoclonal antibody raised against sperm axonemes of the sea urchin *Strongylocentrotus purpuratus*. Preparations were washed in PBT, mounted in DAPI-Vectashield mounting medium,

and observed in an epifluorescence microscope. The images were recorded with a cooled CCD camera and optimized for contrast and brightness with The Gimp freeware. In all cases, a negative control was performed by adding PBT instead of the primary antibody. Images were acquired in an Olympus BX41 fluorescence microscope with a DP70 camera, and were optimized for levels, brightness and contrast with The Gimp freeware.

RNAi knockdown

To verify the Ku70 immunostaining results and discard false positives, three males were injected with $5\mu g$ of a double stranded RNA corresponding to the *E. plorans Ku70* gene region coding for the same amino acids (302-609) used for generating the H-308 antibody. Using our local blast library of all the arthropodan sequences found in the public databases we identified the *Ku70* regions that are conserved among arthropoda. Two pairs of primers were then used for two nested-design PCR reactions, being GGAGTCATACAATGCGGGCCTA and CTCAGCCATAGCATCCACGTC the forward and reverse primers for the outer PCR and CAAAGGCATATGATGGATTTCA and GAAATTTAGTGCCATAGCTTCCA the forward and reverse primers for the inner PCR. As template, the outer PCR used *E. plorans* cDNA, whilst the inner PCR used the amplicons yielded by the outer primers. Both reactions were performed for 35 cycles at 55^oC annealing temperature and 30 s extension time. DNA from the inner PCR reaction was cleaned, using the Qiagen PCR Clean Kit, sequenced for confirmation (GenBank Accession Number: JX979131), and transcribed from both

ends using the standard Ambion T7 RNA polymerase transcription protocol. To ensure knockdown and significant decrease in protein levels, animals were subjected to two injections with the double stranded RNA. The second injection was carried out 8 days after the first one and, to ensure efficient delivery of the double stranded RNA, both were made intra-abdominally at testis level. One male was sacrificed and immunostaining was carried out on its testis follicles 18 days from the first injection (dffi). Twelve days later (i.e., 30 dffi), two other males were vivisected to extract a few testis follicles for immunostaining without sacrificing the animals. These were not sacrificed until the 53rd dffi. This way we obtained two temporally consecutive samples for immunostaining from the same RNAi individuals.

Quantitative PCR

To test the efficiency of the RNAi knockdown, we analyzed the expression level of the *Ku70* gene by means of quantitative PCR (qPCR) in three males which total RNA was extracted, using the REAL Total RNA spin plus kit (Durviz S.L.U.), at 8, 18 and 53 dffi. Genomic DNA was removed by DNAse I treatment (REAL Star kit, Durviz S.L.U.), and the quantity and quality of the total RNA was determined using Tecan's Infinite 200 NanoQuant as well as by gel electrophoresis. 100 *ng* of total RNA were reverse transcribed using the oligo(dT) primer and SuperScript III First-Strand Synthesis SuperMix kit (Invitrogen). For each analysis, a negative control devoid of reverse transcriptase was included to test for contaminating DNA. PCR amplification of a glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) sequence was carried out on each cDNA and

control as described in Van Hiel et al. (2009). This way we ensured that the cDNA was properly synthesized and free of contaminating DNA. Prior to using it for the qPCR, each cDNA was diluted with sterile, nuclease free, double distilled water to a 1:5 ratio.

qPCR testing was carried out on different samples and for each of the seven candidate genes suggested for qPCR normalization by Van Hiel et al. (2009). Of these, two genes, namely glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and ribosomal protein 49 (*RP49*), were highlighted by the geNorm algorithm (Vandesompele et al., 2002) as the most stably expressed. Consequently, we used them as reference genes for our qPCRs. Primers for the qPCR on the target gene (Ku70-F: CACGTCTGGCATTGTTGA, Ku70-R: TCCAGAAGATGCGACAGC) was designed by using the Primer3 web utility (Rozen and Skaletsky, 2000), and for the reference genes, we used the primers GADPH-F: GTCTGATGACAACAGTGCAT, GADPH-R: GTCCATCACGCCACAACTTTC, RP49-F: CGCTACAAGAAGCTTAAGAGGTCAT, and RP49-R: CCTACGGCGCACTCTGTTG, as in Van Hiel et al. (2009). QPCR was conducted using the Chromo 4 System CFB-3240 (BIO-RAD) and Sensimix SYBR kit (BIOLINE). Each reaction was run in duplicate and contained $5\mu l$ of Sensimix, $5\mu l$ of cDNA template (equivalent to 5ng of total RNA) along with $5\mu l$ (10pmol) of each primer and $35\mu l$ of dd$H_2$0. Cycling parameters were 95^oC for 10 min, then 40 cycles of 94^oC for 15s, 56^oC for 15s and 72^oC for 15s, with a plate-read after each extension (i.e., 72^oC) step. To verify that only a single product was amplified in every reaction, melting curves were performed between each 72^oC and 95^oC step, with a plate-read every 1^oC. A standard

curve was used to estimate the efficiency of each primer pair as 2.00 for *Ku70*, 1.97 for *GADPH* and 1.99 for *RP49*. *Ku70* expression levels were calculated according to the Pfaffl's method (Pfaffl, 2001) and normalized as described in Vandesompele et al. (2002).

Colchicine treatment

To analyze a possible relationship of *Ku70* with centromeric function, two additional *E. plorans* males were used for testing the effect of cochicine treatment on the immunostaining results. One was injected with 0.3 ml of 0.1% colchicine in insect saline solution 6 h prior to immunofluorescence analysis while the other was vivisected to extract few testis follicles prior to similar colchicine injection. This enabled an analysis of meiotic cells from the same male both with (after) and without (before) the colchicine treatment.

DNA content in macrospermatids

The amount of DNA in normal and macrospermatids was measured using the Feulgen Image Analysis (FIA) technique described in Ruiz-Ruano et al. (2011).

Results

Ku70 is located at the centromeres of *E. plorans* chromosomes

Immunofluorescence analysis showed centromeric colocalization of Ku70 and γH2AX at meiotic metaphase I (Figure 1). During most first meiotic prophase, however, Ku70 was scarce (Fig. 2, a-h), in accordance with previous results in yeast (Åström et al., 1999; Lee et al., 1999) and mouse (Goedecke et al., 1999). The centromeric location of Ku70 was first seen at diakinesis (Fig. 2, i-j), and remained apparent through metaphase and anaphase stages during both meiotic divisions (Fig. 2, k-p). The centromeric location of these *foci* is inferred from the fact that all chromosomes are telocentric (rod-shaped) and the presence, in anaphase II cells, of Ku70 signal only in the pulling ends of the segregating chromosomes (see Fig. 2, p). The transient presence of Ku70 at centromeric regions, at stages where the nuclear envelope has already been broken down, suggests that this protein might be located on the kinetochore. Immunofluorescence analysis also showed the absence of centromeric *foci* for Ku80 at metaphase I cells (Fig. 2, q-r). Although we observed abundance of Rad51 during early first prophase stages, in agreement with the involvement of this protein in HR (Viera et al., 2004; Calvente et al., 2005), no trace of this protein was observed at metaphase I (Fig. 2, s-t).

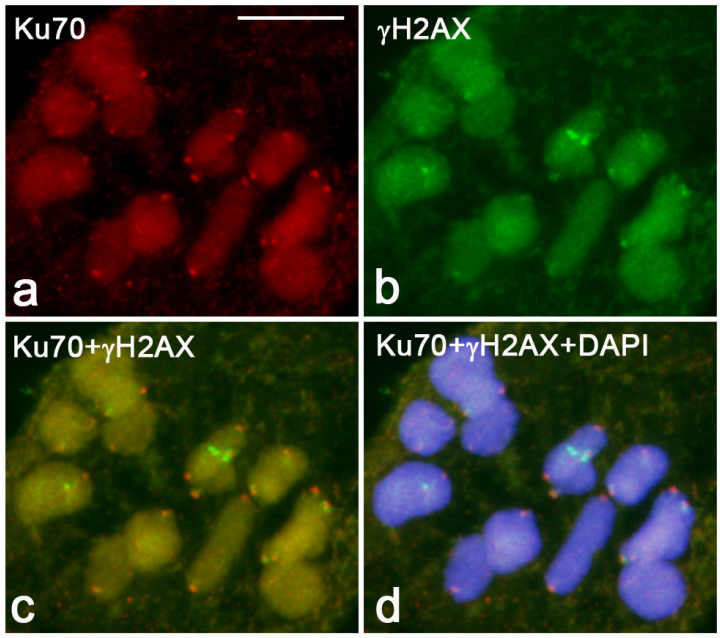

Figure 1: Colocalization of Ku70 with γH2AX in *E. plorans*. Metaphase I cell showing centromeric immunofluorescence patterns for Ku70 (a), γH2AX (b), a merge of the patterns for both antibodies without (c) and with DAPI staining (d). Bar= 10μm.

The centromeric location of Ku70 in *E. plorans* is associated with centromere function

The presence of one centromeric Ku70-*focus* per "segregating unit" is remarkable. In fact, we observe one centromeric Ku70-*focus* per homologous chromosome at metaphase I and one per sister chromatid at metaphase II. This coincides with the fact that centromeres separate two homologous chromosomes at metaphase I and two sister chromatids at metaphase II, thus suggesting the functional presence of Ku70 at the centromere. Colchicine effect on Ku70 immunofluorescence pattern was analyzed in two *E. plorans* males, one of them tested both before

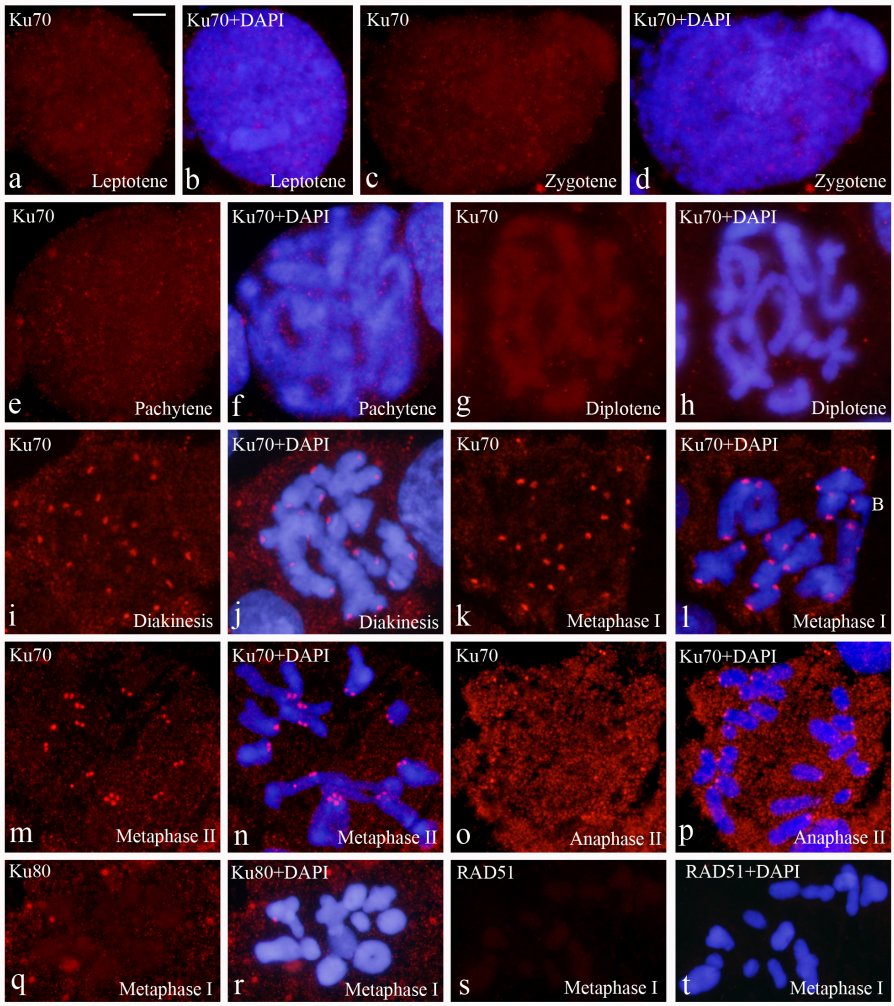

Figure 2: Centromeric location of Ku70 in *E. plorans* spermatocytes. Cells at several meiotic stages are depicted twice, one (on the left) after immunofluorescence with a Ku70 antibody (red) and other (on the right) also including DAPI staining to show chromatin. Note the presence in (i-p) of one large centromeric Ku70 focus per segregating unit, i.e. one per homologous chromosome in diakinesis (i,j) and metaphase I (k,l) cells, and one per sister chromatid in metaphase II cells (m,n), with the *foci* of sister chromatids oriented towards opposite poles in anaphase II cells (o,p). The centromeric *foci* of Ku70 (shown in i-p) were observed in 100% of the cells analyzed from 30 untreated males from three different populations. Also note the absence of centromeric *foci* for both Ku80 (q,r) and RAD51 (s,t) in metaphase I cells. Bar= 10µm. B in (l) points the B chromosome.

and after colchicine injection. With no colchicine, the results were similar to those in non-injected males, with large Ku70 *foci* located on all centromeres (Fig. 3, a-b). In both males, the centromeric *foci* of Ku70 disappeared after colchicine injection (Fig.3, c-d). Since colchicine impedes microtubule formation (for review, see Rieder and Palazzo, 1992), this result suggests that the presence of Ku70 in the centromere depends on the presence of microtubules. An association between the presence of Ku70 on the centromeres and the latter´s microtubule-dependent function can therefore be established.

RNAi knockdown of the *Ku70* gene abolishes the centromeric *foci* of Ku70

Since the antibody employed (H-308, from Santa Cruz Biotechnology) is a rabbit polyclonal antibody raised against human Ku70, and to rule out the possibility that our immunofluorescence results might be false positives caused by potential H-308 cross-reactivity with an epitope from a centromeric protein instead of Ku70, we designed two primer pairs for a nested PCR amplification of the region of the *E. plorans Ku70* gene homologous to the 302-609 amino acids region of the human Ku70, and we synthesized double stranded RNA for RNA interference and *Ku70* gene knockdown in *E. plorans.* Quantitative PCR analysis showed that the expression of the *Ku70* gene was very low in RNAi males analyzed at 8 or 18 days from first injection (dffi) and it was considerably higher in the male analyzed at 53 dffi (Fig. 3, e), suggesting that males whose *Ku70* gene was knocked-down were not only viable, likewise Ku-deficient mice males (Nussenzweig et al., 1996; Gu et al., 1997), but also that

they were able to recover this gene expression —as expected given the transient nature of the RNAi-mediated gene knockdown technique.

In contrast to control males (Fig.3, f-g), none of the examined cells from the testis of the three RNAi males showed immunostaining signal for Ku70 on the centromeres (Fig. 3, h-i), indicating that the large *foci* observed in normal cells actually correspond to genuine detection of the Ku70 protein on the centromeres. Still, the RNAi males showed some immunostaining signals, indicating the presence of Ku70 in the cytosol and also in the spermatids, as previously shown for normal spermatids in this species (Cabrero et al., 2007a). The detectable signals in the cytosol of the RNAi cells are due to the knockdown, not silencing, effect of the RNAi technique, whereas the spermatid signals highlight the refractory nature of the mature germ-line cells to RNAi (Singson, 2001). Remarkably, we did not find anaphase cells in any of the three RNAi males shown in Table 1, even after analyzing their complete set of testis tubules.

Ku70 knockdown males show polyploid spermatids

To look for possible adverse effects resulting from potential impairment of the correct functioning of the centromere after inactivation of the *Ku70* gene, we analyzed 811-1715 spermatids from each of the three RNAi males and the control males employed for our experiments (Table 1). All RNAi males showed the presence of macrospermatids of several sizes (Fig. 4). Estimations of the DNA amount, by means of the Feulgen Image Analysis (FIA) technique, indicated the presence of 2C, 4C, 8C

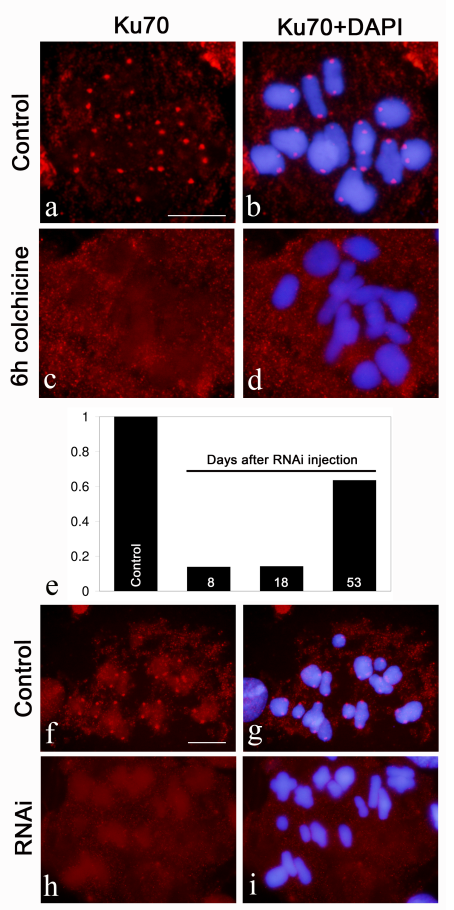

Figure 3: Disappearance of the centromeric Ku70 *foci* by both colchicine injection and *Ku70* gene knockdown in *E. plorans* males. Metaphase I cells from control and experimental males from the colchicine (a-d) and RNAi (f-i) experiments. Note the presence of centromeric Ku70-*foci* before colchicine injection (a,b) and their absence after it (c,d). This pattern was observed in 100% of the cells analyzed from the two males analyzed. (e) Quantitative PCR analysis of the expression level of the *Ku70* gene in RNAi males after 8, 18 and 53 days from first injection (dffi). Observe the sharp decrease in expression level (RQ) at 8 and 18 dffi, and the subsequent partial recovery at 53 dffi. (f,g) Metaphase I cell from a control male (with no RNAi injection), showing presence of large Ku70 *foci* on all centromeres. (h,i) Metaphase I cell after RNAi injection showing absence of centromeric Ku70 *foci*. The patterns shown in (f,g) and (h,i) did not show intraindividual variation either in the two control or the three experimental males analyzed. RQ= Relative quantification, Bar= 10μm.

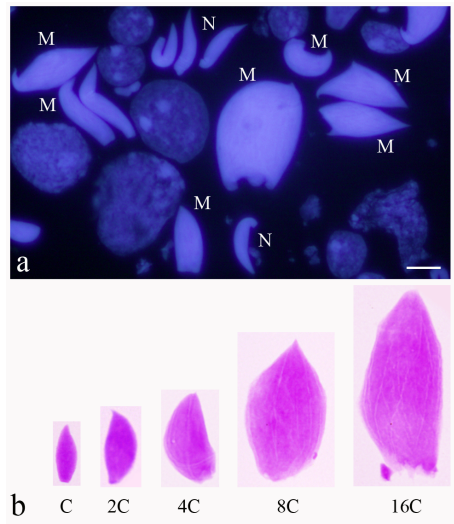

Figure 4: Formation of polyploid spermatids in RNAi males. (a) A microscopic field showing DAPI staining of some normal spermatids (N) along with many macrospermatids (M). (b) DNA content of spermatids, measured by the Feulgen Image Analysis (FIA) technique, and showing to be 1C for normal spermatids, and 2C, 4C, 8C or 16C for macrospermatids (for the number of spermatids analyzed, see Table 1). Bar= $10\mu m$.

and 16C spermatids. In addition, microspermatids were also observed, although at much lower frequency than macrospermatids (Table 1).

The quantitative analysis of the frequency of macrospermatids, shown in Table 1, indicated the existence of a basal proportion of macrospermatids in control males (0.21-1.16%). In RNAi males, however, we found a significant increase in this proportion which varied with time since RNAi injection: it increased 2.27-fold in the male analyzed at 18 dffi, 75.87-fold and 87-fold in the two males analyzed at 30 dffi, and 11.86-fold and 3.07-fold in these same males analyzed at 53 dffi. This suggests that the effects of *Ku70* knockdown on macrospermatid formation are not observed until one meiotic cycle has been completed. The basal frequency of microspermatids in the control males was very low (0.12-0.31%), and its frequency was about similar in the 18 dffi male (0.15%), 3.7% and 1.3% in the two 30 dffi males, and 0.18% and 2.8% in these same males analyzed at 53 dffi. Therefore, the most conspicuously observable change

associated with RNAi treatment was an increase in the frequency of polyploid spermatids.

Centromeric Ku70 *foci* are also present during mitosis

To analyze whether the centromeric location of the Ku70 protein is exclusive to meiosis or it also occurs during mitosis, we analyzed mitotic metaphases in spermatogonia from adult males as well as in two types of embryonic cells. In addition to the small *foci* scattered throughout the cytosol, large *foci* of Ku70 were present on the centromeric regions of the chromosomes of all three types of cells analyzed (Fig. 5, a-d; Fig. S1). We therefore conclude that Ku70 is associated with centromeric function in both meiotic and mitotic cells.

The centromeric *foci* of Ku70 are dependent on tubulin presence

To ascertain whether Ku70 presence at centromere regions during metaphase and anaphase is dependent on the presence of microtubules, we performed double immunofluorescence analysis with antibodies against Ku70 and α-tubulin on colchicine and control males. As Fig. 5 shows, in the absence of colchicine, mitotic anaphase cells show conspicuous *foci* of Ku70 on the chromosome ends that locate close to the cell poles, in close association with α-tubulin (Fig. 5, a-d). Likewise, meiotic metaphase I cells showed abundant α-tubulin across the cytosol, rounding all bivalents, and Ku70 centromeric *foci* were

Table 1: Frequency of macro- and microspermatids found in the RNAi and control males. Exp= experiment, Bs= number of B chromosomes, Dffi= days from first injection, N= normal spermatids, M= macrospermatids, m= microspermatids, RM= ratio of macrospermatids between RNAi and control males, Rm= ratio of microspermatids between RNAi and control males. Contingency chi-squared tests comparing the frequency of macrospermatids between control and experimental males indicated significantly higher proportions in RNAi males: ** indicate P<0.01, *** indicate P<0.001

				Type of spermatids							
Exp	**Type**	**Bs**	**Dffi**	**N**	**M**	**m**	**Total**	**% M**	$\mathbf{R}_M$	**% m**	$\mathbf{R}_m$
1	Control	1	18	1613	19	5	1637	1.16		0.31	
	RNAi	1	18	1329	36	2	1367	2.63	2.27 **	0.15	0.48
2	Control	2	30	1415	3	2	1420	0.21		0.14	
	RNAi_1	3	30	651	130	30	811	16.03	75.87 ***	3.70	26.26 ***
	RNAi_2	1	30	1498	323	24	1845	17.51	82.87 ***	1.30	9.24 ***
	Control	2	53	1700	13	2	1715	0.76		0.12	
	RNAi_1	3	53	1495	148	3	1646	8.99	11.86 ***	0.18	1.56
	RNAi_2	1	53	1590	39	47	1676	2.33	3.07 ***	2.80	24.05 ***

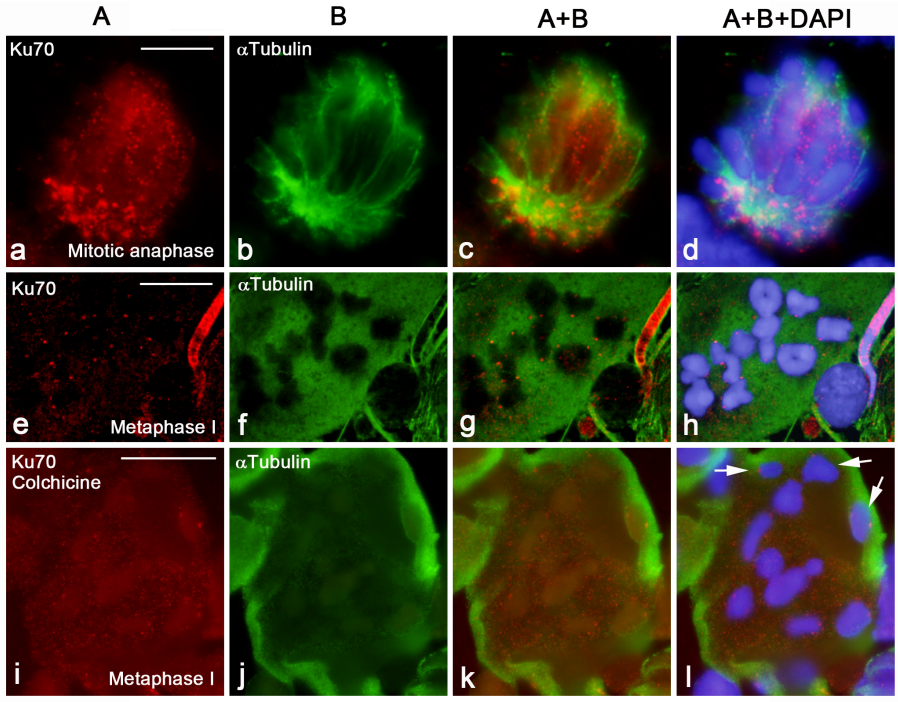

Figure 5: Association of the centromeric *foci* of Ku70 (red) with α-tubulin (green) in a mitotic anaphase (a-d) and meiotic metaphase I (e-h) cells, without colchicine treatment, and in a metaphase I cell (i-l) from a male injected with colchicine. Note the lower amount of α-tubulin in the latter and the presence of several centromeric *foci* of Ku70 in the periphery of this cell in coincidence with higher amounts of α-tubulin (arrows in l). Also note the abundance of Ku70 in spermatid nuclei (arrowhead in h), as previously reported by Cabrero et al. (2007a). Bar= 10µm.

adjacent to α-tubulin (Fig 5, e-h). However, in males injected with colchicine, α-tubulin was very scarce or absent in the most part of the cytosol, and Ku70 centromeric *foci* were absent, with the remarkable few exceptions in those bivalents that were placed at the periphery of the cell where the amount of α-tubulin was higher (Fig. 5, i-l).

The novel centromeric function of Ku70 appears to be an autapomorphy in *E. plorans*

To test whether the newly discovered centromeric function of Ku70 is a general characteristic that also occurs in other species, we analyzed the Ku70 immunocytochemical pattern in 14 additional grasshopper and locust species (including *Heteracris adspersa* and *H. litoralis*, as close relatives of *E. plorans*, and the swarming locusts *Locusta migratoria* and *Schistocerca gregaria*), as well as the mouse as a vertebrate representative. In all cases, Ku70 was present at metaphase I, forming numerous small *foci* in the cytosol (Fig. S2) but, opposed to what we report for *E. plorans*, no concentration of Ku70 at centromeric regions was observed in any of the other species. This indicates that the centromere-related function of Ku70, which we report here, might be exclusive to *E. plorans*. However, in the 14 other grasshopper and locust species, as in *E. plorans*, Ku70 was abundant in the nucleus of the elongating spermatids as well as in the sperm tail (Cabrero et al., 2007a), suggesting that, aside from the centromeric function, Ku70 seems to fulfill similar functions in *E. plorans* and other grasshopper species.

The centromeric function of Ku70 is not an adaptation to supernumerary chromosomes

Since *E. plorans* shows a very widespread polymorphism for supernumerary (B) chromosomes, the new function we report here for Ku70 might conceivably be a byproduct of the presence of these supernumerary chromosomes and/or of the arms race between them and the chromosomes of the standard genome. To assess this possibility, we analyzed the immunofluorescence pattern of Ku70 in individuals with and without B chromosomes from a B-carrying population as well as individuals from two non-B-carrying populations. In all cases, we observed large *foci* of Ku70 on the centromeres of all chromosomes, including the B chromosomes, suggesting that the centromeric location and function of Ku70 is a general characteristic for this species, at least in the populations analyzed in this work. Hence the centromeric function of Ku70 does not seem to be an adaptation to the presence of supernumerary chromosomes in *E. plorans*.

Discussion

RNAi efficiently knocks down the *Ku70* gene

The first report proving the possibility of RNAi mediated gene knockdown and the presence of the *Sid1* gene in an orthopteran

species was in *Schistocerca americana* (Dong and Friedrich, 2005). Our experiment with another orthopteran species further proves that RNAi is a valid technique for inactivating genes in orthoptera. Indeed, we observed a close correspondence between double stranded RNA administration and a decrease of the *Ku70* gene expression followed by the disappearance of this gene product from centromeres and a subsequent high increase in the frequency of polyploid spermatids. This gene knockdown effect must be due to the presence in orthopteran species of the double stranded RNA uptake membrane protein SID1 as well as the DICER and all the protein components of the RNA-Induced Silencing Complex. Interestingly, RNAi did not inhibit the presence of Ku70 in mature and developing spermatozoa, with spermatids and spermatozoa of injected males showing the same normal (Ku70+) pattern previously shown in Cabrero et al. (2007a). At first sight, Ku70 presence in sperm from RNAi males might appear to contradict the validity of the gene silencing technique. However, this result is logical if one bears in mind that spermatozoa seem to be refractory to RNAi even in *Caenorrhabditis elegans*, the model species for RNAi experimentation (see Singson, 2001).

The Ku70 protein gathers at the centromeric region of the chromosomes in *E. plorans*

We have observed that the Ku70 monomer of the Ku heterodimer is conspicuously present on the centromeres during the mitotic and meiotic stages of active chromosome segregation whereas the Ku80 monomer fails to show this pattern, suggesting that Ku70 might play a Ku80-independent centromeric function in the grasshopper *E. plorans*.

The disappearance of the centromeric *foci* of Ku70 after colchicine injection, in parallel with a decrease in α-tubulin amount, suggests that the putative centromeric function of Ku70 might have something to do with microtubule association to kinetochore regions, since colchicine is a microtubule-targeted agent that binds to α and β subunits of tubulin heterodimers to inhibit its polymerization (for review, see Bhattacharyya et al., 2008). In fact, the close association between Ku70 and α-tubulin presence in metaphase I cells from males injected with colchicine indicates that the centromeric function of Ku70 is tubulin dependent (see Fig. 5, i-l). Bearing this in mind, the disappearance of Ku70 from the centromeric region in all chromosomes of RNAi males, along with the absence of anaphase cells, suggest a general failure in chromosome segregation leading to cytokinesis failure and the appearance of polyploid spermatids. In fact, completion of cytokinesis requires accurate chromosome segregation in human cells since even non-disjunction for a single chromosome may result in tetraploid instead of aneuploid cells (Shi and King, 2005). Remarkably, also in human cells, it has been shown that Ku70 interacts with ARF6, a protein localized in the cleavage furrow and midbody of mitotic cells, and whose activity is regulated during cytokinesis (Schweitzer and D'Souza-Schorey, 2005).

The exclusive presence of even ploidy levels in spermatids, and the absence of anaphase cells, indicate the occurrence of cytokinesis failures and restitution processes, an expected result of an impairment of kinetochore-microtubule association leading to cytokinesis failure (Normand and King, 2010). Bearing in mind that a primary spermatocyte is 4C, the presence of polyploid spermatids with ploidy levels higher than 4C suggests that these cytokinesis failures can also

occur during preceding spermatogonial mitoses (see Fig. S3, a). For instance, 2C spermatids would result from failure in the second meiotic division, whereas 4C ones would result from failures in both meiotic divisions, 8C from failures in the last spermatogonial mitosis plus both meiotic divisions, and 16C in the two last spermatogonial mitoses plus both meiotic divisions. The differences in ploidy level of the resulting spermatids thus rest on the timing of double stranded RNA uptake and processing by the different dividing cell types (spermatogonium, primary spermatocyte or secondary spermatocyte) (see some examples in Fig. S3).

As to the dynamics of the above mentioned spermatid malformations, our results show only a slight increase in the frequency of macrospermatids during the first 18 days post-RNAi treatment, then a much steeper rise of the prevalence of these mis-formed spermatids during the following 12 days (*i.e.*, at 30 dffi; see Table 1). This is consistent with the fact that the complete meiotic cycle in grasshoppers lasts about 20 days (Fox et al., 1974), so that at 18 dffi only a few RNAi spermatocytes had completed meiosis and thus were affected by the *Ku70* gene knockdown, whereas, 12 days later, many more cells had had time to finish meiosis and manifest cell division problems.

Ku70 is a transient centromeric passenger presumably involved in microtubule anchoring

The absence of colocalization of Ku80 in the centromeric *foci* of Ku70 indicates that the latter is not there for DNA repair purposes.

Likewise, its centromeric presence does not appear to be a consequence of interactions with HP1α, since RNAi knockdown of HP1α fails to eliminate the centromeric *foci* of Ku70 (Ruiz-Estévez et al., in preparation). In order to look for other conceivable functions of Ku70 at centromeric level, diverse biological functions can be inferred from its ability to bind other proteins. One of these, clusterin (Yang et al., 1999), is a chaperone-like molecule that interacts with the microtubule-destabilizing protein SCLIP of the stathmin family (Kang et al., 2005). The fact that colchicine prevents the centromeric location of Ku70 (see Fig. 3) suggests that Ku70 may have a thus far unreported microtubule-dependent function at *E. plorans* centromeres. Additional support to this conclusion comes from our observation that Ku70 seems to also have a microtubule-dependent function in the formation of the sperm tail both in *E. plorans* and all the 14 other orthopteran species studied in this work (Cabrero et al., 2007a), so that the Ku70 function at both centromeres and sperm tail could be rather similar.

The presence of polyploid spermatids in *Ku70* RNAi-knocked-down males indicates that many spermatocytes finished meiosis despite the absence of Ku70 at centromeres and the presumed problems in microtubule anchoring, thus surpassing the spindle assembly checkpoint (SAC) control. Among the dynamic protein components of the kinetochore, the SAC components reach their maximum concentration in the absence of microtubules (Howell et al., 2004), just the opposite pattern of Ku70 in *E. plorans*, which is not present 6h after colchicine injection. This indicates that the Ku70 centromeric function reported here is different from the canonical SAC function.

A more suggestive possibility is an interaction with some kinetochore

proteins, e.g. the APC (adenomatous polyposis coli) and/or EB1 (end binding 1) proteins, which associate to kinetochores only when microtubules are anchored (Kaplan et al., 2001; Tirnauer et al., 2002) thus resembling the transient presence of Ku70 on the *E. plorans* centromeres. Remarkably, siRNA-mediated inhibition of APC, EB1, or APC and EB1 together gives rise to spindle defects without mitotic arrest (Green et al., 2005), and APC inhibition produces polyploidy (Dikovskaya et al., 2007), i.e. exactly the same consequences as *Ku70* knockdown in *E. plorans*. A possible microtubule anchoring function of Ku70 at the kinetochores of *E. plorans* is thus conceivable through interaction with APC and EB1. This microtubule-dependent function of Ku70 could operate at kinetochores (exclusively in *E. plorans*) but also at the centriolar adjuncts of spermatids in all grasshopper species (Cabrero et al., 2007a).

APC and EB1 belong to a group of proteins showing affinity for the plus-ends of growing microtubules (+TIPs). It is thus conceivable that Ku70 binds the +TIPS and it concentrates in the kinetochore when bundles of microtubules reach it. Indeed, the fact that the spindle microtubules are sensitive to colchicine disruption (Bhattacharyya et al., 2008), supports our observation that colchicine treatment in *E. plorans* prevents the presence of Ku70 at centromeres and is consistent with Ku70 involvement in a microtubule-dependent centromeric function. Therefore, the Ku70 function that better explains its transient presence at centromeric regions is probably related to the kinetochore-microtubule anchoring that allows accurate chromosome segregation.

The centromeric function of Ku70 is an autapomorphy

In this work we analyzed a high number of grasshopper and locust species (14 including *E. plorans*, two of its closely related species and the well known pest locusts *Locusta migratoria* and *Schistocerca gregaria*), as well as the mouse. Among these, only *E. plorans* showed centromeric presence of Ku70. Still, we cannot categorically rule out that it might be present in other organisms not yet examined. In fact, in human cells, it has recently been shown that DNA-PKcs, with which Ku70 and Ku80 form the DNA-dependent protein kinase (Gell and Jackson, 1999), is activated and phosphorylated in close association with the spindle apparatus at centrosomes and kinetochores, thus suggesting that, in addition to its role in DSB repair, DNA-PKcs (and perhaps Ku too) is a critical regulator of mitosis and could modulate microtubule dynamics in chromosome segregation (Lee et al., 2011).

Whatever the case, the absence of Ku70 at the chromosome centromeres of 14 orthopteran and 1 vertebrate species strongly suggests that Ku70 binding to the centromere is most likely an autapomorphy in *E. plorans* and that the ancestral Ku70 state did not include a centromeric function. Since most of the functions of Ku are not directly related to each other, establishing the evolutionary history of this protein might be challenging (Tuteja and Tuteja, 2000), and the centromeric location of Ku70 during mitotic and meiotic metaphase and anaphase in *E. plorans* is a clear evidence of the complex evolutionary pathways that this protein may have followed in parallel to the acquisition of its multiple functions. Interestingly, Scherthan and Trelles-Sticken (2008) showed that the absence of Ku interferes with the ordered occurrence of chromosome

dynamics during first meiotic prophase, and Shuaib et al. (2010) found the presence of both Ku70 and Ku80 among the proteins associated with e-CENP-A in prenucleosomal complexes purified by immunoprecipitation from HeLa cell extracts, suggesting that Ku70 might perform thus far undiscovered and previously unrecognized centromeric functions not only in the grasshopper *E. plorans* but in other organisms too.

Acknowledgements

We thank R. Jiménez for providing us with mouse spermatocytes, M. Ruiz-Estévez for giving us an *E. plorans* cDNA aliquote, and T. López for technical assistance. This study was supported by grants from the Spanish Ministerio de Ciencia e Innovación (CGL2009-11917 and BFU2010-16438) and Plan Andaluz de Investigación (CVI-6649), and was partially performed by FEDER funds. M. Bakkali was supported by a Ramón y Cajal fellowship from the Spanish Ministerio de Ciencia e Innovación.

References

Åström, S. U., Okamura, S. M., and Rine, J. (1999). Yeast cell-type regulation of DNA repair. *Nature*, 397(6717):310–310.

Bertinato, J., Schild-Poulter, C., and Hache, R. J. (2001). Nuclear localization of Ku antigen is promoted independently by basic motifs in the Ku70 and Ku80 subunits. *Journal of Cell Science*, 114(1):89–99.

Bhattacharyya, B., Panda, D., Gupta, S., and Banerjee, M. (2008). Anti-mitotic activity of colchicine and the structural basis for its interaction with tubulin. *Medicinal Research Reviews*, 28(1):155–183.

Cabrero, J., Palomino-Morales, R. J., and Camacho, J. P. M. (2007a). The DNA-repair Ku70 protein is located in the nucleus and tail of elongating spermatids in grasshoppers. *Chromosome Research*, 15(8):1093–1100.

Cabrero, J., Teruel, M., Carmona, F. D., and Camacho, J. P. M. (2007b). Histone H2ax phosphorylation is associated with most meiotic events in grasshopper. *Cytogenetic and Genome Research*, 116(4):311–315.

Calvente, A., Viera, A., Page, J., Parra, M. T., Gómez, R., Suja, J. A., Rufas, J. S., and Santos, J. L. (2005). DNA double-strand breaks and homology search: inferences from a species with incomplete pairing and synapsis. *Journal of Cell Science*, 118(13):2957–2963.

Camacho, J. P. M., Bakkali, M., Corral, J. M., Cabrero, J., López-León, M. D., Aranda, I., Martín-Alganza, A., and Perfectti, F. (2002). Host recombination is dependent on the degree of parasitism. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1505):2173–2177.

Camacho, J. P. M., Cabrero, J., Viseras, E., López-León, M. D., Navas-Castillo, J., and Alche, J. D. (1991). G banding in two species of grasshopper and its relationship to C, N, and fluorescence banding techniques. *Genome*, 34(4):638–643.

Camacho, J. P. M., Carballo, A. R., and Cabrero, J. (1980). The B-chromosome system of the grasshopper *Eyprepocnemis plorans* subsp. *plorans* (Charpentier). *Chromosoma*, 80(2):163–176.

Dikovskaya, D., Schiffmann, D., Newton, I. P., Oakley, A., Kroboth, K., Sansom, O., Jamieson, T. J., Meniel, V., Clarke, A., and Näthke, I. S. (2007). Loss of APC induces polyploidy as a result of a combination of defects in mitosis and apoptosis. *The Journal of Cell Biology*, 176(2):183–195.

Dong, Y. and Friedrich, M. (2005). Nymphal RNAi: systemic RNAi mediated gene knockdown in juvenile grasshopper. *BMC Biotechnology*, 5:25.

Fox, D. P., Hewitt, G. M., and Hall, D. J. (1974). DNA replication and RNA transcription of euchromatic and heterochromatic chromosome regions during grasshopper meiosis. *Chromosoma*, 45(1):43–62.

Gao, Y., Chaudhuri, J., Zhu, C., Davidson, L., Weaver, D. T., and Alt, F. W. (1998). A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. *Immunity*, 9(3):367–376.

Gell, D. and Jackson, S. P. (1999). Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic Acids Research*, 27(17):3494–3502.

Goedecke, W., Eijpe, M., Offenberg, H. H., Aalderen, M. v., and Heyting, C. (1999). Mre11 and Ku70 interact in somatic cells, but are differentially expressed in early meiosis. *Nature Genetics*, 23(2):194–198.

Green, R. A., Wollman, R., and Kaplan, K. B. (2005). APC and EB1 function together in mitosis to regulate spindle dynamics and chromosome alignment. *Molecular Biology of the Cell*, 16(10):4609–4622.

Gu, Y., Seidl, K. J., Rathbun, G. A., Zhu, C., Manis, J. P., van der Stoep, N., Davidson, L., Cheng, H.-L., Sekiguchi, J. M., Frank, K., Stanhope-Baker, P., Schlissel, M. S., Roth, D. B., and Alt, F. W. (1997). Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity*, 7(5):653–665.

Gu, Y., Sekiguchi, J., Gao, Y., Dikkes, P., Frank, K., Ferguson, D., Hasty, P., Chun, J., and Alt, F. W. (2000). Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America*, 97(6):2668–2673.

Howell, B. J., Moree, B., Farrar, E. M., Stewart, S., Fang, G., and Salmon, E. D. (2004). Spindle checkpoint protein dynamics at kinetochores in living cells. *Current Biology*, 14(11):953–964.

Kang, S.-W., Shin, Y.-J., Shim, Y.-J., Jeong, S.-Y., Park, I.-S., and Min, B.-H. (2005). Clusterin interacts with SCLIP (SCG10-like protein) and promotes neurite outgrowth of PC12 cells. *Experimental Cell Research*, 309(2):305–315.

Kaplan, K. B., Burds, A. A., Swedlow, J. R., Bekir, S. S., Sorger, P. K., and Näthke, I. S. (2001). A role for the Adenomatous Polyposis Coli protein in chromosome segregation. *Nature Cell Biology*, 3(4):429–432.

Katz, D. J., Beer, M. A., Levorse, J. M., and Tilghman, S. M. (2005). Functional characterization of a novel Ku70/80 pause site at the H19/Igf2 imprinting control region. *Molecular and Cellular Biology*, 25(10):3855–3863.

Koike, M. (2002). Dimerization, translocation and localization of Ku70 and Ku80 proteins. *Journal of Radiation Research*, 43(3):223–236.

Koike, M., Awaji, T., Kataoka, M., Tsujimoto, G., Kartasova, T., Koike, A., and Shiomi, T. (1999). Differential subcellular localization of DNA-dependent protein kinase components Ku and DNA-PKcs during mitosis. *Journal of Cell Science*, 112 (Pt 22):4031–4039.

Koike, M. and Koike, A. (2005b). The Ku70-binding site of Ku80 is required for the stabilization of Ku70 in the cytoplasm, for the nuclear translocation of Ku80, and for Ku80-dependent DNA repair. *Experimental Cell Research*, 305(2):266–276.

Koike, M., Shiomi, T., and Koike, A. (2001). Dimerization and nuclear localization of Ku proteins. *Journal of Biological Chemistry*, 276(14):11167–11173.

Lee, K.-J., Lin, Y.-F., Chou, H.-Y., Yajima, H., Fattah, K. R., Lee, S.-C., and Chen, B. P. C. (2011). Involvement of DNA-dependent protein kinase in normal cell cycle progression through mitosis. *Journal of Biological Chemistry*, 286(14):12796–12802.

Lee, S. E., Pâques, F., Sylvan, J., and Haber, J. E. (1999). Role of yeast SIR genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths. *Current Biology*, 9(14):767–770.

Martinez, J. J., Seveau, S., Veiga, E., Matsuyama, S., and Cossart, P. (2005). Ku70, a component of DNA-dependent protein kinase, is a mammalian receptor for *Rickettsia conorii*. *Cell*, 123(6):1013–1023.

Monferran, S., Muller, C., Mourey, L., Frit, P., and Salles, B. (2004a). The Membrane-associated Form of the DNA Repair Protein Ku is Involved in Cell Adhesion to Fibronectin. *Journal of Molecular Biology*, 337(3):503–511.

Monferran, S., Paupert, J., Dauvillier, S., Salles, B., and Muller, C. (2004b). The

membrane form of the DNA repair protein Ku interacts at the cell surface with metalloproteinase 9. *The EMBO Journal*, 23(19):3758–3768.

Muñoz, E., Perfectti, F., Martín-Alganza, Á., and Camacho, J. P. M. (1998). Parallel effects of a B chromosome and a mite that decrease female fitness in the grasshopper *Eyprepocnemis plorans*. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1408):1903–1909.

Normand, G. and King, R. W. (2010). Understanding Cytokinesis Failure. In MA, R. Y. C. P., editor, *Polyploidization and Cancer*, pages 27–55. Springer New York.

Nussenzweig, A., Chen, C., da Costa Soares, V., Sanchez, M., Sokol, K., Nussenzweig, M. C., and Li, G. C. (1996). Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature*, 382(6591):551–555.

Paull, T. T., Rogakou, E. P., Yamazaki, V., Kirchgessner, C. U., Gellert, M., and Bonner, W. M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Current Biology*, 10(15):886–895.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, 29(9):e45–e45.

Pöggeler, S. and Kück, U. (2006). Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. *Gene*, 378:1–10.

Rieder, C. L. and Palazzo, R. E. (1992). Colcemid and the mitotic cycle. *Journal of Cell Science*, 102 (Pt 3):387–392.

Rodgers, W., Jordan, S. J., and Capra, J. D. (2002). Transient association of Ku with nuclear substrates characterized using fluorescence photobleaching. *Journal of Immunology (Baltimore, Md.: 1950)*, 168(5):2348–2355.

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology (Clifton, N.J.)*, 132:365–386.

Ruiz-Ruano, F. J., Ruiz-Estévez, M., Rodríguez-Pérez, J., López-Pino, J. L., Cabrero, J., and Camacho, J. P. M. (2011). DNA amount of X and B chromosomes in the grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenetic and Genome Research*, 134(2):120–126.

Sawada, M., Sun, W., Hayes, P., Leskov, K., Boothman, D. A., and Matsuyama, S. (2003). Ku70 suppresses the apoptotic translocation of Bax to mitochondria. *Nature Cell Biology*, 5(4):320–329.

Scherthan, H. and Trelles-Sticken, E. (2008). Absence of yKu/Hdf1 but not myosin-like proteins alters chromosome dynamics during prophase I in yeast. *Differentiation; Research in Biological Diversity*, 76(1):91–98.

Schweitzer, J. K. and D'Souza-Schorey, C. (2005). A requirement for ARF6 during the completion of cytokinesis. *Experimental Cell Research*, 311(1):74–83.

Shi, Q. and King, R. W. (2005). Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. *Nature*, 437(7061):1038–1042.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proceedings of the National Academy of Sciences of the United States of America*, 107(4):1349–1354.

Singson, A. (2001). Every sperm is sacred: fertilization in *Caenorhabditis elegans*. *Developmental Biology*, 230(2):101–109.

Song, K., Jung, Y., Jung, D., and Lee, I. (2001). Human Ku70 interacts with heterochromatin protein 1alpha. *The Journal of Biological Chemistry*, 276(11):8321–8327.

Takata, M., Sasaki, M. S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A., and Takeda, S. (1998). Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *The EMBO Journal*, 17(18):5497–5508.

Tirnauer, J. S., Canman, J. C., Salmon, E. D., and Mitchison, T. J. (2002). EB1 targets to kinetochores with attached, polymerizing microtubules. *Molecular Biology of the Cell*, 13(12):4308–4316.

Tuteja, R. and Tuteja, N. (2000). Ku autoantigen: a multifunctional DNA-binding protein. *Critical Reviews in Biochemistry and Molecular Biology*, 35(1):1–33.

Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. V., and Simonet, G. (2009). Identification and

validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. *BMC Molecular Biology*, 10:56.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7).

Viera, A., Santos, J. L., Page, J., Parra, M. T., Calvente, A., Cifuentes, M., Gómez, R., Lira, R., Suja, J. A., and Rufas, J. S. (2004). DNA double-strand breaks, recombination and synapsis: the timing of meiosis differs in grasshoppers and flies. *EMBO Reports*, 5(4):385–391.

Yang, C. R., Yeh, S., Leskov, K., Odegaard, E., Hsu, H. L., Chang, C., Kinsella, T. J., Chen, D. J., and Boothman, D. A. (1999). Isolation of Ku70-binding proteins (KUBs). *Nucleic Acids Research*, 27(10):2165–2174.

Zhou, B. B. and Elledge, S. J. (2000). The DNA damage response: putting checkpoints in perspective. *Nature*, 408(6811):433–439.

Zurita, S., Cabrero, J., López-León, M. D., and Camacho, J. P. M. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, 52(1):274–277.

Supplementary material

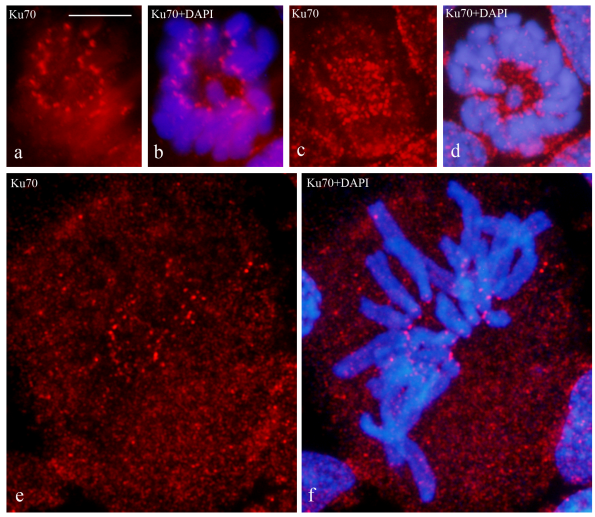

Figura S1: The centromeric location of Ku70 also occurs during mitosis in *E. plorans*. (a,b) Spermatogonial mitotic metaphase. (c,d) Embryo mitotic metaphase cell. (e,f) Embryo mitotic metaphase from a neuroblast cell (of giant size). Note the presence of centromeric *foci* of Ku70 in the three cells. This was observed in 100 % of the analyzed cells. Bar= $10\mu m$.

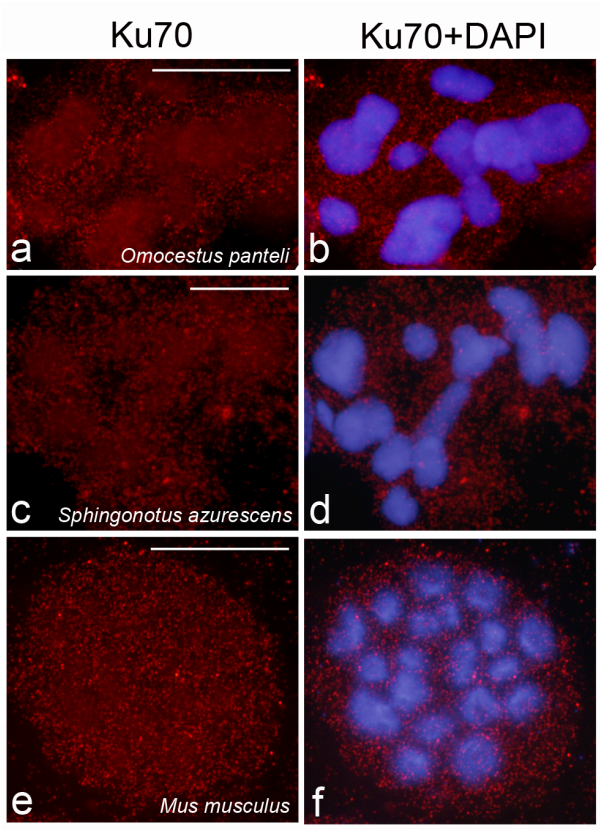

Figura S2: Absence of centromeric immunofluorescence signals for Ku70 at meiotic metaphase I in 14 grasshopper species and the mouse. Examples are shown for Ku70 (a,c,e) and DAPI+Ku70 (b,d,f) in the grasshopper species *Omocestus panteli* (a, b) and *Sphingonotus azurescens* (c, d), and the mouse (e,f). Bar= $10\mu m$.

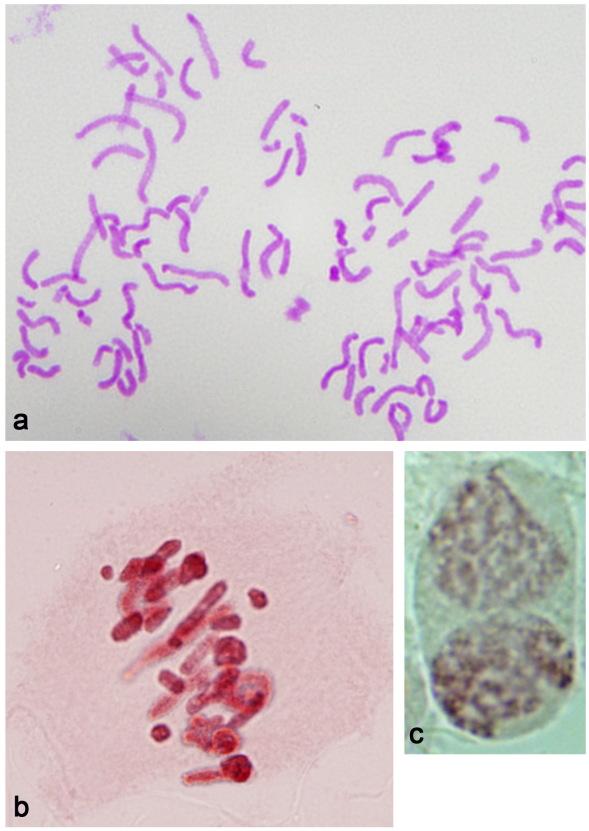

Figura S3: Presence of polyploid and binucleate cells in RNAi males. a) Two tetraploid spermatogonia at mitotic prometaphase. b) Tetraploid metaphase I cell. c) Binucleate cell.

2

B chromosome effects on *Hsp70* gene expression does not occur at transcriptional level in the grasshopper *Eyprepocnemis plorans*[I]

Beatriz Navarro-Domínguez, Josefa Cabrero, Juan Pedro M. Camacho, María Dolores López-León

Departamento de Genética, Universidad de Granada

Abstract As intragenomic parasites, B chromosomes can elicit stress in the host genome, thus inducing a response for host adaptation to this kind of continuous parasitism. In the grasshopper *Eyprepocnemis plorans*, B chromosome presence has been previously associated with a decrease in the amount of the heat shock protein 70 (HSP70). To investigate whether this effect is already apparent at transcriptional level, we analyze the expression levels of the *Hsp70* gene in gonads and somatic tissues of males and females with and without B chromosomes from

[I]In press. Molecular Genetics and Genomics, June 2016

two populations, where the predominant B chromosome variants (B2 and B24) exhibit different levels of parasitism, by means of quantitative real-time PCR (qPCR) on complementary DNA (cDNA). The results revealed the absence of significant differences for *Hsp70* transcripts associated with B chromosome presence in virtually all samples, except for testis in the Torrox population where B-carrying individuals showed higher levels than B-lacking ones. This indicates that the decrease in HSP70 protein levels, formerly reported in this species, may be not a consequence of transcriptional down-regulation of *Hsp70* genes but the result of post-transcriptional regulation. These results will help to design future studies oriented to identifying factors modulating Hsp70 expression, and will also contribute to uncover the biological role of B chromosomes in eukaryotic genomes.

Introduction

Heat shock proteins (HSPs) are a family of highly conserved proteins which act as molecular chaperones in stabilization and refolding of denatured proteins (Clark and Worland, 2008). These proteins bind to other proteins exhibiting wrong conformations, due either to protein-denaturing stress or because peptides have not been properly synthesized, folded, assembled, or localized to its correct place (Feder and Hofmann, 1999). They are known to be induced by heat and other stresses, but they are also essential for normal cell functioning, being

involved in many cellular processes. The stress inducible HSPs action has been shown in a wide variety of organisms as adaptive response to a number of different environmental stresses, including not only heat or cold stress, but also radiation, cellular energy depletion, osmotic and oxidative stress, toxic substances, crowding, hypoxia, parasitism, bacterial and viral infections, senescence, inbreeding, and deleterious mutations (for review, see Sørensen et al., 2003). They also influence processes, such as transcription regulation, signal transduction, apoptotic pathways, and even show immunogenic properties (Gambill et al., 1993; Maresca and Kobayashi, 1994; Schmitt et al., 2007). Families of HSPs have been described and named according to their molecular mass, the largest and most conserved being the 70 kDa heat shock proteins (HSP70s) (Roberts et al., 2010). Members of this protein family comprise those which are constitutively expressed during non-stressful conditions, with little or no response to cellular stress (HSC70: heat shock cognate 70) and those which are either not expressed, or else expressed at low rate, during normal conditions, but are rapidly induced by stress (HSP70: heat shock protein 70) (Denlinger et al., 2001).

The mechanisms of *Hsp70* gene activation have been a classical and fascinating model for understanding how the cell recognizes and responds to acute and chronic exposures to environmental and physiological stress (Morimoto et al., 1992). Therefore, studies on regulation of *Hsp70* gene have been conducted at pre-transcriptional, post-transcriptional, pre-translational, and post-translational levels in a wide variety of animal and plant species. Pre-transcriptional activation of the *Hsp70* gene, via binding of heat shock factors (HSF) to heat shock gene promoters, has been profusely analyzed and seems to exert the greatest control over

stress response. Even though information about downstream *Hsp70* gene regulation is less abundant, multiple mechanisms of regulation have been described in addition to transcription activation (see Silver and Noble, 2012 for review), in some cases with sharp discrepancies between *Hsp70* mRNA level and protein abundance (Metzger et al., 2012).

In addition to these general functions of Hsp genes, up-regulation of *Hsp70* has been observed in the presence of viral infections (Wang et al., 2006; Santoro et al., 2010), genetics stress factors, such as aging (Wheeler et al., 1999), mutations (Zhao et al., 2002), and inbreeding (Kristensen et al., 2002), even at not stressful temperatures. Other possible source of genetic stress could be the presence of additional genomic elements, such as supernumerary (B) chromosomes (Camacho et al., 2002; Houben et al., 2013). B-chromosomes are additional dispensable chromosomes which are frequently viewed as intragenomic parasites decreasing carrier fitness and showing drive mechanisms (Camacho, 2005). The presence of B-chromosomes can cause genetic stress on the individuals carrying them, or may affect carriers' ability to deal with stress, especially with high B numbers (Teoh and Jones, 1978; Teruel et al., 2011). Signs of stress associated with B-chromosome presence have been reported in the grasshopper *Eyprepocnemis plorans*, manifested through a decrease in eggs fertility and clutch size (Zurita et al., 1998; Muñoz et al., 1998; Bakkali et al., 2010) and nucleolus size (Teruel et al., 2007, 2009), increase of chiasma frequency (Camacho et al., 2002), and a general enhancement in transposable element transcription (Navarro-Domínguez et al. in preparation; see chapter 5) which could be associated with higher rates of transposition, a typical genomic reaction to stress (Arnault and Dufournel, 1994). Likewise, transcriptome and microarray analyses

have revealed that B-chromosome presence is associated with gene expression changes related with protein modification and response to stress (Navarro-Domínguez et al. in preparation; see chapter 5).

In the grasshopper *E. plorans*, Teruel et al. (2011) found that B-carrying males and females from two Spanish populations (Salobreña and Torrox) showed lower levels of HSP70 protein than B-lacking ones, during two consecutive years. These populations harbor the B2 and B24 variants, respectively, the latter being more parasitic and derived from the former (Zurita et al., 1998). These authors found a decrease in the level of HSP70 protein in male and female gonads of B-carrying individuals from both populations, but HSP70 levels were lower in Torrox than Salobreña. The decrease showed a dose-dependent pattern in Torrox, i.e. the population harboring the most parasitic B variant, and this was interpreted as a decrease in the ability of the host to deal with stress Teruel et al. (2011), as it could keep some relationship with the reduction in nucleolus area previously reported by (Teruel et al., 2007), given that HSP70 acts as nucleolus maintainer under stressful conditions.

Here, we investigate whether the formerly observed decrease in HSP70 protein level in B-carrying individuals of *E. plorans* is a consequence of previous regulation at the transcriptional level, by analyzing *Hsp70* gene expression by means of qPCR on individuals from the two same populations.

Materials and methods

Experimental materials

A total of 80 adult males and females of the grasshopper *Eyprepocnemis plorans* were collected in two natural populations (Salobreña and Torrox, Spain) in 2013. Animals were anesthetized in the laboratory and dissected under a stereomicroscope. Half of the testis tubules and ovarioles were removed from males and females, respectively, fixed in freshly prepared 3:1 ethanol-acetic acid and stored at 4^oC for cytogenetic analysis. Prior to fixation, ovarioles were immersed in 2% colchicine in isotonic insect saline solution for 2h. The remaining half gonads of males and females were frozen in liquid nitrogen and stored at -80^oC, likewise every individual somatic body (whole body minus the gonads), for molecular studies. B chromosome number was determined in males by squashing two testis tubules in 2% lactopropionic orcein. In females, colchicine-treated ovarioles were submitted to C-banding for B chromosome identification in mitotic metaphases, where they can be visualized as highly heterochromatic supernumerary chromosomes.

RNA extractions and cDNA synthesis

Total RNA extractions from frozen somatic bodies were performed using the Real Total RNA Spin Plus kit (Durviz), whereas RNA isolation from frozen hemi-gonads was performed using the RNeasy Lipid Tissue Mini Kit (Qiagen), following manufacturer's recommendations plus an

additional DNase treatment in the column membrane consisting in 20 units DNase Amplification Grade DNase I (Sigma) for 30 minutes. The absence of genomic DNA (gDNA) contamination was checked in agarose gels and corroborated by the lack of PCR amplification for ribosomal DNA (rDNA) or histone genes in the extracted RNA. When needed, a second DNase treatment was performed after RNA extractions, using the REALSTAR kit (Durviz).

Quality and quantity of RNA were assessed with a Tecan's Infinite 200 NanoQuant spectrophotometer and in a denaturing MOPS-agarose gel with formaldehyde (0.05%) to ensure the absence of RNA degradation. cDNA synthesis was performed by retro-transcription of 100 ng per sample of total RNA using a combination of random and oligo dT hexamers (PrimeScript™ RT reagent Kit, Perfect Real Time, Takara). The cDNA obtained was diluted in RNase-DNase free water for a 1:10 working solution.

Amplification and sequencing of a fragment of Hsp70 coding sequence

Degenerate primers for the amplification of part of the *Hsp70* gene coding region in *E. plorans* were designed based on its sequence in *Locusta migratoria* described in (Qin et al., 2003) (GenBank accession number AY178988). Primers were designed according to CODEHOP strategy (Rose 2003) using the program GENtle (Manske, 2006). Degenerated primer sequences and amplicon length are detailed in Table 1. The sequence obtained can be found in GenBank under accession number

Table 1: Primer pairs used for degenerated amplification of *Hsp70* (deg.Hsp70), quantification of *Hsp70* expression by means of qPCR (q.Hsp70) and housekeeping genes used for qPCR normalization. Housekeeping genes primer sequences were taken from Van Hiel et al. (2009) and Chapuis et al. (2011). AL = Amplicon Length; AT = Anneling temperature.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	AL	AT
deg.Hsp70	AGGCCAGCATTGARATHGAY	CAACCACCAGCAGTYTCDATDC	391	45
q.Hsp70	GTGTTGGTGGGTGGGTCAACTCG	ACGTCAAGCAGCAGCAGGTCC	185	58
ACT	AATTACCATTGGTAACGAGCGATT	TGCTTCCATACCCAGGAATGA	73	60
Arm	ACTTCTTATGAGAGCATTCCAGGAT	GCTCCTTCTTTGTCTGCTGCT	114	60
EF1a	GATGCTCCAGGCCACAGAGA	TGCACAGTCGGCCTGTGAT	66	60
GAPDH	GTCTGATGACAACAGTGCAT	GTCCATCACGCCACAACTTTC	81	60
RP49	CGCTACAAGAAGCTTAAGAGGTCAT	CCTACGGCGCACTCTGTTG	66	60
TubA1	TGACAATGAGGCCATCTATG	CGCAAAGATGCTGTGATTGA	117	60
Ubi	GACTTTGAGGTGTGGCGTAG	GGATCACAAACACAGAACGA	78	60

KX034180.

We amplified the *Hsp70* gene from *E. plorans* cDNA by means of the PCR technique. Low astringency conditions were used to facilitate degenerate primers hybridization. PCR experiments were performed in a 25 µl reaction containing 2 µl of cDNA, 0.4 µM of each primer, 0.1mM dNTPs, 3mM $MgCl_2$, 1x Taq polymerase buffer, and 2.5 units of Horse-Power Taq polymerase (Canvax). PCR experiments were carried out in a Eppendorf Mastercycler ep Gradient with the following conditions: initial denaturation for15 s at 94^{o}C followed by 40 cycles of 1 min at 94^{o}C, 1 min at 44.5^{o}C, 1 min at 72^{o}C and a final extension for 10 min at 72^{o}C. PCR on cDNA resulted in a single fragment of the expected size. PCR product was cleaned using GenElute™ PCR Clean-up kit (Sigma) and then cloned using TOPO TA cloning kit (Invitrogen) and Sanger sequenced by Macrogen Inc. We ran BLASTX (Altschul et al., 1990) on NCBI's Non-redundant protein sequences (nr) database to

confirm the specificity of the amplified sequence.

Quantitative Real Time PCR (qPCR)

Primers design and specificity

On the basis of the partial DNA sequence obtained for the *Hsp70* coding region, we designed a primer pair (named q.Hsp70; Table 1) suitable for qPCR which amplified a fragment of 185bp. Primer design was carried out using the FastPCR (Kalendar et al., 2009) software, and the absence of secondary structures was examined with NetPrimer (Premier Biosoft International).

To select appropriate house-keeping genes (HKGs) for reference in the gene expression experiments, we initially amplified seven genes using the primers provided by Van Hiel et al. (2009) and Chapuis et al. (2011) for *Actin 5C* (*Act*), *Armadillo* (*Arm*), *Elongation factor 1α* (*EF1a*), *Glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*), *Ribosomal protein 49* (*RP49*), *α-tubulin 1A* (*Tub*) and *Ubiquitin conjugating enzyme 10* (*Ubi*) (primer sequences and amplicon length are detailed in Table 1).

qPCR amplification for both *Hsp70* and the seven reference genes was performed on a Chromo4 Real Time PCR thermocycler (BioRad). Reactions contained $5\mu l$ of cDNA, $5\mu l$ of SensiMix™ SYBR Kit (Bioline) and $2.5\mu l$ of each $2.5\mu M$ primer in a total $15\mu l$ volume. Electronic pipettes (Eppendorf Research®Pro) were used to minimize pipetting errors, and each reaction was carried out in duplicate. The coefficient of variation between technical replicates was in all cases lower than 5%.

Quantitative PCR protocol consisted in an initial denaturation step for 10 minutes at 95^{o}C, followed by 40 cycles of 15 s at 94^{o}C, 15 s at 60^{o}C and 15 s at 72^{o}C, with a plate read at the end of every cycle. Specificity of reaction was assessed by means of a dissociation curve (from 72^{o}C to 95^{o}C with plate reading every 1^{o}C) after the 40th cycle. The fluorescence data were measured and processed using Opticon Monitor 3.1 (Bio-Rad Laboratories, Inc). Negative controls for each primer pair were included in all reactions.

qPCR products were also visualized by electrophoresis in a 1.5% agarose gel. To further confirm specificity, they were also cleaned with the GenElute™ PCR Clean-up kit and sequenced by Macrogen Inc. Sequences for all housekeeping genes amplified in *E. plorans* can be found in GenBank under accession numbers KX034173 to KX034179.

Determination of amplification efficiency for *Hsp70* and reference genes

To determine reaction efficiency for *Hsp70* and reference genes amplification, we performed a standard curve with 5 aliquots of 1:10 serial dilutions of *E. plorans* cDNA for each primer pair used in this study.

Reference gene selection

We determined the most stable reference genes, and their optimal number, using GeNorm (Vandesompele et al., 2002). Stability analysis was performed using 10 samples chosen randomly from each sex, body part

and population, and measuring expression levels of the five potential reference genes. Stability analyses were performed according to GeNorm manual.

Relative quantification of Hsp70 mRNA expression and normalization

We quantified the expression of the *Hsp70* gene in all 80 *E. plorans* individuals collected. Relative quantities (RQs) were calculated in reference to a calibrator sample, consisting in a mixture of cDNAs which was amplified in every run to make them comparable, and were normalized by the geometrical average of the most stable reference genes for each kind of sample. Calculations were done with a custom R package (Navarro-Domínguez et al., in preparation), based on (Pfaffl, 2001), but considering the geometrical average of multiple reference genes for normalization (Vandesompele et al., 2002).

Statistical analysis

Statistical analysis was performed in R. We first performed an ANCOVA analysis with four factors (population, sex, body part and B-chromosome presence) and log10 of RQ values as dependent variable, to accomplish the assumptions of the model. We also analyzed the effect of B chromosome presence and population separately in male bodies, female bodies, testis and ovaries of *E. plorans*, by means of two-way ANOVA.

Table 2: Individuals of *Eyprepocnemis plorans* analyzed per population, sex and number of B chromosomes. Bs = number of B chromosomes. N = Number of individuals.

Population	Sex	Bs	N
Salobreña	Female	0	2
		1	5
		2	2
		+2	5
		Total	*14*
	Male	0	4
		1	10
		2	4
		+2	3
		Total	*21*
Torrox	Female	0	4
		1	9
		2	8
		+2	1
		Total	*22*
	Male	0	2
		1	7
		2	8
		+2	6
		Total	23
Total Samples			80

Results

The number of B chromosomes carried by males and females from both populations is shown in Table 2. The B chromosomes observed in Salobreña and Torrox corresponded to the B2 and B24 variants, respectively, previously reported in these same populations (Cabrero et al., 2014).

PCR amplification of the *Hsp70* gene on *E. plorans* cDNA using degenerate primers yielded a single product which was visualized as a 391 bp fragment in electrophoresis gels. Cloning and Sanger sequencing indicated that the product showed homology with several *Hsp70* sequences of other insects found in the GenBank database, including the *Locusta migratoria* sequence used for primer design (ac.no. AY178988). *E. plorans Hsp70* partial sequence is available in GenBank under accession number KX034180. We then designed primers for qPCR to amplify a 185 bp fragment, from the nucleotide 169 to 353 of the *E. plorans* Hsp70 partial sequence, Table 3 shows the amplification factor per cycle of the qPCR (E) and the qPCR reaction efficiency (in percentage) for all primer pairs used in this study, calculated from the standard curve slope of a ten-fold serial dilution. All E values ranged from 1.81 to 2.05 (81% to 105%), with linear correlation coefficients (R^2) being higher than 0.9 (from 0.9760 to 0.9990). Therefore, all primer pairs showed acceptable efficiency, clearly higher than the minimum recommend efficiency (80%) (Wang and Seed, 2006). Some efficiency values were higher than 100% (Table 3), which are not theoretically possible, probably indicating that either this efficiency calculation method is not optimal and slightly overestimates the "real efficiency" (Pfaffl, Michael W., 2004), or else the presence of inhibitory factors in the PCR reaction.

After preliminary screening of seven house-keeping genes for reference, two of them (*EF1α* and *Ubi*) were excluded due to non-specific amplification, and the five remainder were analyzed with GeNorm (Vandesompele et al., 2002). We selected the most stable reference genes for each population, sex and body part according to GeNorm M

Table 3: Efficiency of qPCR amplification for each primer pair. Slope = Slope of the standard curve of serial dilutions; E = PCR reaction efficiency (fold-increase per cycle); Eff. = Percentage of efficiency

Primer pair	Slope	R^2	E	Eff.
q.Hsp70	-3.234	0.994	2.04	104
ACT	-3.650	0.984	1.88	88
Arm	-3.868	0.976	1.81	81
GAPDH	-3.216	0.997	2.05	105
RP49	-3.323	0.941	2.00	100
TubA1	-3.413	0.999	1.96	96

values, and 2 or 3 genes depending on the V value (Vandesompele et al., 2002). Selected genes were: *Act-RP49* for female bodies and ovaries from Salobreña, *GAPDH-Tub* for bodies and *Act-Tub-Arm* for testes from Salobreña, *Act-Arm-RP49* for female bodies from Torrox, *Act-Arm* for ovaries from Torrox, *Act-Tub* for male bodies and *Act-RP49* for testes from Torrox (Table 4).

We performed qPCR experiments on the gonad and body samples from all 80 males and females from the two populations. ANCOVA analysis, with population, sex, body part (gonad or somatic body) and B chromosome presence as factors, and log10RQ values (transformed) as response variable, revealed significant differences in Hsp70 expression between populations (Torrox>Salobreña in most samples except testis), sexes (females>males) and body parts (somatic body>gonad), but not in respect to the presence of B chromosomes (Table 5). Given the huge differences in Hsp70 expression observed between sexes and between body and gonad, we also analyzed the effect of B chromosome presence (B-carrying vs. B-lacking) and population separately in male bodies,

Table 4: Selection of reference genes based on GeNorm algorithm

Population	Sex	Body Part	Reference genes Rank (M value)				V value
Salobreña	Female	Body	*TUB* (0.75)	*GADPH* (0.62)	*ARM* (0.42)	*ACT-RP49* (0.28)	V2/3 (0.15)
		Ovary	*GADPH* (0.76)	*TUB* (0.68)	*ARM* (0.57)	*ACT-RP49* (0.31)	V2/3 (0.18)
	Male	Body	*ARM* (3.03)	*ACT* (0.54)	*RP49* (0.42)	*GADPH-TUB* (0.30)	V2/3 (0.15)
		Testicle	*GADPH* (0.80)	*RP49* (0.51)	*ACT* (0.42)	*TUB-ARM* (0.21)	V3/4 (0.13)
Torrox	Female	Body	*GADPH* (0.81)	*TUB* (0.58)	*RP49* (0.45)	*ACT-ARM* (0.16)	V3/4 (0.16)
		Ovary	*GADPH* (1.23)	*RP49* (0.47)	*TUB* (0.46)	*ACT-ARM* (0.42)	V2/3 (0.14)
	Male	Body	*GADPH* (1.03)	*RP49* (0.47)	*ARM* (0.33)	*ACT-TUB* (0.24)	V2/3 (0.01)
		Testis	*TUB* (1.45)	*GADPH* (1.13)	*ARM* (0.49)	*ACT-RP49* (0.31)	V2/3 (0.18)

Table 5: ANCOVA testing for the influence of population, sex body part and B chromosome presence on *Hsp70* transcription. SS = Sum of squares; RSS = Residual sum of squares; AIC = Akaike information criterion.

	df	SS	RSS	AIC	F	P
Population	1	1.1236	5.4078	-468.89	36.7178	**1,20e-05**
Sex	1	5.9945	10.2786	-375.77	195.8895	**<2.2e-16**
Body part	1	3.1005	7.3847	-423.71	101.3207	**<2.2e-16**
B presence	1	0.0005	4.2846	-502.65	0.0147	0.9036

female bodies, testis and ovaries of *E. plorans*, by means of two-way ANOVA. This revealed significant differential expression for the *Hsp70* gene associated with the presence of B chromosomes in testes from the Torrox population, but not in all remaining samples (Fig. 1, Table 6), with B-carrying individuals showing higher expression level than B-lacking ones, with no dosage effect (Fig. 2).

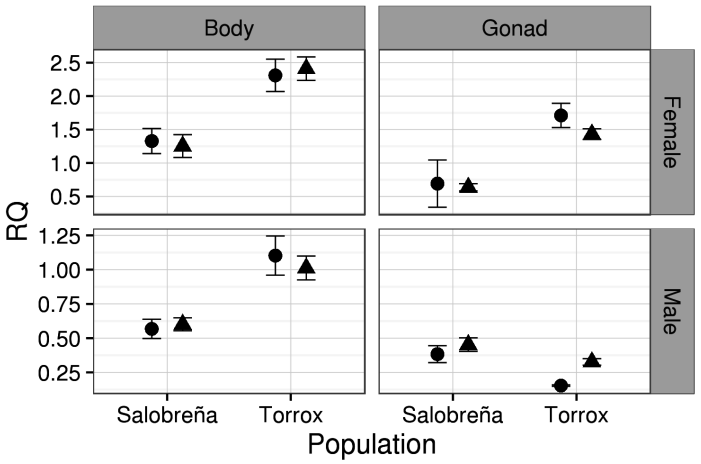

Figure 1: *Hsp70* expression (mean ± SE) per sex, body part (somatic body and gonad) and population (Salobreña and Torrox) in B chromosome lacking (circles) and B chromosome carrying (triangles) individuals. RQ= Relative quantity (expression level)

Table 6: Two-way ANOVA testing population and B-chromosome effects on *Hsp70* gene expression, per sex and body part. SS = Sum of squares; RSS = Residual sum of squares; AIC = Akaike information criterion; Item: Pop = Population, Bpres = Presence of the B chromosome.

Sex	Body Part	Item	df	SS	RSS	AICc	F	P
Male	Body	Pop.	1	0.299556	0.98695	-132.73	14.8166	**0.0005**
		B.pres	1	0.002142	0.68954	-146.35	0.1060	0.7468
		Pop+B.pres	1	0.006687	0.69408	-146.10	0.3307	0.5690
	Testis	Pop.	1	0.29459	1.02973	-131.12	13.6246	**0.0008**
		B.pres	1	0.15323	0.88837	-136.73	7.0870	**0.0118**
		Pop+B.pres	1	0.07059	0.80573	-140.44	3.2647	**0.0796**
Female	Body	Pop.	1	0.32950	0.95351	-120.10	16.3690	**0.0003**
		B.pres	1	0.00276	0.62677	-134.79	0.1371	0.7137
		Pop+B.pres	1	0.00485	0.62887	-134.67	0.2411	0.6269
	Ovary	Pop.	1	0.73439	1.18987	-107.99	48.3702	**0.0001**
		B.pres	1	0.00535	0.46083	-140.24	0.3523	0.5573
		Pop+B.pres	1	0.00962	0.46510	-139.92	0.6337	0.4323

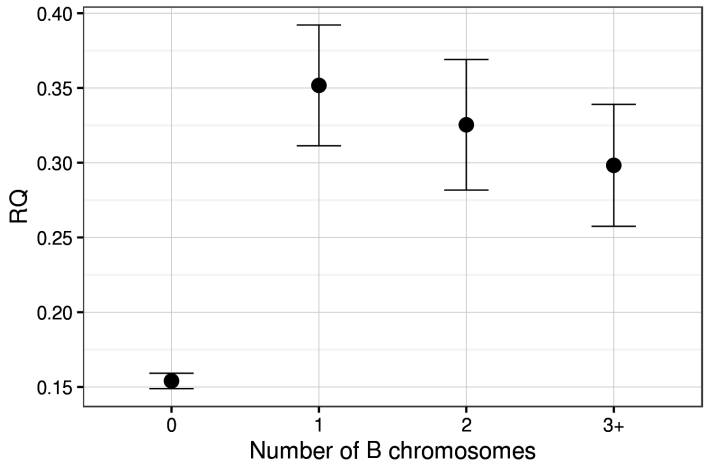

Figure 2: *Hsp70* expression (mean ± SE) in testis of males from Torrox with different number of B-chromosomes. RQ= relative quantity (expression level)

Discussion

Our results have shown that the decrease in HSP70 protein levels observed by Teruel et al. (2011) in B-carrying male and female gonads of *E. plorans* from the Torrox and Salobreña populations is not a consequence of transcriptional down-regulation of the *Hsp70* gene. In fact, only testes of B-carrying males from Torrox showed significantly different amounts of Hsp70 transcripts in respect to B-lacking ones, but it was an increase, thus in the opposite direction than previous observations at protein level. The remaining samples showed no significant difference associated with the presence of B chromosomes, suggesting that the most important regulation events of *Hsp70*, at this respect, could take place at post-transcriptional level. This may include any of the battery of mRNA modification processes (e.g., 5' cap addition, 3' polyadenylation, splicing, editing, etc.) which protect mature mRNA from degradation and make it suitable to be transported to the cytosol and translated into functional proteins (Tranter et al., 2011). Such a diversity of processes could make this regulation level prone to be affected by the presence of B chromosomes. Opposite mRNA/protein *Hsp70* levels have also been observed in the bivalve mollusk *Crassotrea gigas* exposed to copper (Metzger et al., 2012). Likewise, Foth et al. (2008) found, in *Plasmodium falciparum*, increased Hsp70-1 isoform transcript levels which were not correlated with protein abundance. However, examples of concordance between *Hsp70* transcript and protein levels have also been reported (Zhao et al., 2005).

A combination of factors, operating by delaying protein synthesis and/or

influencing mRNA lifespan, could explain the lack of correspondence between transcript and protein levels (Renaut et al., 2006; Gedeon and Bokes, 2012). It has been shown that when the stress source is continuous or prolonged, *Hsp70* mRNAs remain in the cell, and HSP70 protein supply takes place without the need for activation of transcription promoters (Kaarniranta et al., 1998). This could be the case for a genome burdened by intragenomic parasites as B chromosomes.

In Drosophila, it has been reported that cells with continuous expression of *Hsp70* showed an initially reduced growth rate, which lately recovers due to a cellular response consisting in sequestering and inactivating HSP70 proteins (Feder et al., 1992). It is thus conceivable that B chromosome presence might induce some changes in how the host genome manages protein maintenance. We have performed the transcriptome analysis in whole bodies of B-carrying and B-lacking females of *E. plorans*, as well as microarray analysis in ovaries, which also failed to show differential expression of the *Hsp70* gene. Remarkably, expression changes were observed for several genes belonging to the ubiquitin-proteasoma pathway involved in degradation of misfolded proteins (Navarro-Domínguez et al., in preparation, see chapter 5). This pathway is functionally related with heat shock response to stress (Hofmann and Somero, 1995; Wang et al., 2010) and, therefore, it could have something to do with the low levels of HSP70 proteins observed by (Teruel et al., 2011). We have no comparable analyses in males, but it is reasonable to argue that the decrease in HSP70 protein levels observed by the former authors in testes might also be a consequence of changes in other genes expression, especially bearing in mind the prominent role of HSP70 and other chaperones, cochaperones and the

ubiquitination-proteasome system during spermatogenesis (Meccariello et al., 2014). In addition, some members of the HSP70 family have been shown to play an important role in the completion of meiosis during spermatocyte differentiation (Govin et al., 2006). In any case, possible implication of alternative stress regulatory genes pathway to *Hsp70*, such as ubiquitin-proteasoma system, which could be affected by B chromosomes in the cellular response to stress, demands additional research.

We cannot rule out other possible scenarios explaining our results, since the decrease of HSP70 protein levels reported by Teruel et al. (2011) was based on samples collected during 2006 and 2007, whereas the present study has been carried out on samples collected in 2013. Therefore, the lack of B chromosome effects on Hsp70 gene expression found here might be a consequence of B24 neutralization (i.e., a reduction in the level of parasitism) resulting from the coevolution between A and B chromosomes during last years. Indeed, the transmission rate of the B24 chromosome has decreased significantly over time (Perfectti et al., 2004).

Alternatively, the presence of genetic content in the B chromosome that might be modulating gene expression in B-carrying genomes cannot be ruled out. For instance, Banaei-Moghaddam et al. (2013) have shown that B chromosomes in rye carry pseudogenes with transcriptional activity which may influence the expression of their A genome counterparts. Likewise, the presence and activity of the *Hsp70* pseudogenes in the B chromosome of *E. plorans* might compensate a decrease in Hsp70 mRNA induced by B-presence, thus yielding similar levels in B-carrying and B-lacking individuals. Although this possibility

is conceivable, we have no evidence for the presence of *Hsp70* genes or pseudogenes in the B chromosome, even though we have found nine protein-coding genes in the B24 chromosome, some of which were actively transcribed (Navarro-Domínguez et al., submitted; see chapter 3).

A final comment merits the population differences observed for *Hsp70* gene expression in all samples analyzed here, with higher values in Torrox than Salobreña for three samples and the reverse pattern in the remainder (testes). These results are intriguing, and it is difficult to know whether they have something to do with the presence of a more parasitic B variant in Torrox (B24) than in Salobreña (B2) (Zurita et al., 1998)or else these are population differences associated with environmental differences, since *Hsp70* expression levels frequently respond to environmental conditions (Sørensen et al., 2003). This last possibility seems more likely as the differences between populations are also appreciated in 0B individuals. Alternatively, it would be conceivable some kind of genomic imprinting conditioning *Hsp70* expression level in B-lacking individuals to the B chromosome type carried by their parents, in resemblance to the maternal imprinting effect previously shown for rye B chromosomes Puertas et al. (1990). The analysis of *Hsp70* expression level in experimental crosses, under the common uniform environmental conditions for individuals from both populations, would help to solve this puzzle.

Acknowledgements

We thank two anonymous reviewers for useful comments.

This study was funded by the Spanish Ministerio de Economía y Competitividad (CGL2015-70750-P) and Plan Andaluz de Investigación (CVI-6649).

Navarro-Domínguez was supported by a FPI fellowship from the Spanish Ministerio de Ciencia e Innovación (BES-2010-030016).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Arnault, C. and Dufournel, I. (1994). Genome and stresses: Reactions against aggressions, behavior of transposable elements. *Genetica*, 93(1-3):149–160.

Bakkali, M., Manrique-Poyato, M., López-León, M., Perfectti, F., Cabrero, J., and Camacho, J. (2010). Effects of B chromosomes on egg fertility and clutch size in the grasshopper *Eyprepocnemis plorans*. *Journal of Orthoptera Research*, 19(2):197–203.

Banaei-Moghaddam, A. M., Meier, K., Karimi-Ashtiyani, R., and Houben, A. (2013). Formation and expression of pseudogenes on the B chromosome of rye. *Plant Cell*, 25(7):2536–2544.

Cabrero, J., López-León, M., Ruíz-Estévez, M., Gómez, R., Petitpierre, E., Rufas, J., Massa, B., Kamel Ben Halima, M., and Camacho, J. (2014). B1 was the ancestor B chromosome variant in the western Mediterranean area in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 142(1):54–58.

Camacho, J. P. M. (2005). B chromosomes. In *The Evolution of the Genome*, pages 223–286. Academic Press.

Camacho, J. P. M., Bakkali, M., Corral, J. M., Cabrero, J., López-León, M. D., Aranda, I., Martín-Alganza, A., and Perfectti, F. (2002). Host recombination is dependent on the degree of parasitism. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1505):2173–2177.

Chapuis, M.-P., Tohidi-Esfahani, D., Dodgson, T., Blondin, L., Ponton, F., Cullen, D., Simpson, S. J., and Sword, G. A. (2011). Assessment and validation of a suite of reverse transcription-quantitative PCR reference genes for analyses of density-dependent behavioural plasticity in the Australian plague locust. *BMC Molecular Biology*, 12:7.

Clark, M. S. and Worland, M. R. (2008). How insects survive the cold: molecular mechanisms—a review. *Journal of Comparative Physiology B*, 178(8):917–933.

Denlinger, D. L., Rinehart, J. P., and Yocum, G. D. (2001). Stress proteins: a role in insect diapause. In *Insect timing: Circadian rhythmicity to seasonality*, pages 155–171. Gulf Professional Publishing.

Feder, J. H., Rossi, J. M., Solomon, J., Solomon, N., and Lindquist, S. (1992). The consequences of expressing hsp70 in Drosophila cells at normal temperatures. *Genes Dev.*, 6(8):1402–1413.

Feder, M. E. and Hofmann, G. E. (1999). Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annual Review of Physiology*, 61(1):243–282.

Foth, B. J., Zhang, N., Mok, S., Preiser, P. R., and Bozdech, Z. (2008). Quantitative protein expression profiling reveals extensive post-transcriptional regulation and post-translational modifications in schizont-stage malaria parasites. *Genome Biology*, 9(12):1–18.

Gambill, B. D., Voos, W., Kang, P. J., Miao, B., Langer, T., Craig, E. A., and Pfanner, N. (1993). A dual role for mitochondrial heat shock protein 70 in membrane translocation of preproteins. *The Journal of Cell Biology*, 123(1):109–117.

Gedeon, T. and Bokes, P. (2012). Delayed protein synthesis reduces the correlation between mRNA and protein fluctuations. *Biophysical Journal*, 103(3):377–385.

Govin, J., Caron, C., Escoffier, E., Ferro, M., Kuhn, L., Rousseaux, S., Eddy, E. M., Garin, J., and Khochbin, S. (2006). Post-meiotic shifts in HSPA2/HSP70.2 chaperone activity during mouse spermatogenesis. *The Journal of Biological Chemistry*, 281(49):37888–37892.

Hofmann, G. and Somero, G. (1995). Evidence for protein damage at environmental temperatures: seasonal changes in levels of ubiquitin conjugates and hsp70 in the intertidal mussel *Mytilus trossulus*. *Journal of Experimental Biology*, 198(7):1509–1518.

Houben, A., Banaei-Moghaddam, A. M., Klemme, S., and Timmis, J. N. (2013). Evolution and biology of supernumerary B chromosomes. *Cellular and Molecular Life Sciences*, 71(3):467–478.

Kaarniranta, K., Elo, M., Sironen, R., Lammi, M. J., Goldring, M. B., Eriksson, J. E., Sistonen, L., and Helminen, H. J. (1998). Hsp70 accumulation in chondrocytic cells

exposed to high continuous hydrostatic pressure coincides with mRNA stabilization rather than transcriptional activation. *Proceedings of the National Academy of Sciences*, 95(5):2319–2324.

Kalendar, R., Lee, D., and Schulman, A. H. (2009). FastPCR software for PCR primer and probe design and repeat search. In *Focus on Bioinformatics*. Mansour, a. edition.

Kristensen, T. N., Dahlgaard, J., and Loeschcke, V. (2002). Inbreeding affects Hsp70 expression in two species of Drosophila even at benign temperatures. *Evolutionary Ecology Research*, 4(8):1209–1216.

Manske, M. (2006). *GENtle, a free multi-purpose molecular biology tool*. Phd. thesis, Universität zu Köln.

Maresca, B. and Kobayashi, G. S. (1994). Hsp70 in parasites: as an inducible protective protein and as an antigen. *Experientia*, 50(11-12):1067–1074.

Meccariello, R., Chianese, R., Ciaramella, V., Fasano, S., and Pierantoni, R. (2014). Molecular chaperones, cochaperones, and ubiquitination/deubiquitination system: Involvement in the production of high quality spermatozoa. *BioMed Research International*, 2014:e561426.

Metzger, D. C., Pratt, P., and Roberts, S. B. (2012). Characterizing the effects of heavy metal and vibrio exposure on Hsp70 expression in *Crassostrea gigas* gill tissue. *Journal of Shellfish Research*, 31(3):627–630.

Morimoto, R. I., Sarge, K. D., and Abravaya, K. (1992). Transcriptional regulation of heat shock genes. A paradigm for inducible genomic responses. *The Journal of Biological Chemistry*, 267(31):21987–21990.

Muñoz, E., Perfectti, F., Martín-Alganza, Á., and Camacho, J. P. M. (1998). Parallel effects of a B chromosome and a mite that decrease female fitness in the grasshopper *Eyprepocnemis* plorans. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1408):1903–1909.

Perfectti, F., Corral, J., Mesa, J., Cabrero, J., Bakkali, M., López-León, M., and Camacho, J. (2004). Rapid suppression of drive for a parasitic B chromosome. *Cytogenetic and Genome Research*, 106(2-4):338–343.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, 29(9):e45–e45.

Pfaffl, Michael W. (2004). Quantification strategies in real-time PCR. In *AZ of quantitative PCR*, pages 89–113. International University Line La Jolla, CA.

Puertas, M., Jimenez, M., Romera, F., Vega, J., and Diez, M. (1990). Maternal imprinting effect on B chromosome transmission in rye. *Heredity*, 64(2):197–204.

Qin, W., Tyshenko, M. G., Wu, B. S., Walker, V. K., and Robertson, R. M. (2003). Cloning and characterization of a member of the hsp70 gene family from Locusta migratoria, a highly thermotolerant insect. *Cell Stress and Chaperones*, 8(2):144–152.

Renaut, J., Hausman, J.-F., and Wisniewski, M. E. (2006). Proteomics and low-temperature studies: bridging the gap between gene expression and metabolism. *Physiologia Plantarum*, 126(1):97–109.

Roberts, R. J., Agius, C., Saliba, C., Bossier, P., and Sung, Y. Y. (2010). Heat shock proteins (chaperones) in fish and shellfish and their potential role in relation to fish health: a review. *Journal of Fish Diseases*, 33(10):789–801.

Santoro, M. G., Amici, C., and Rossi, A. (2010). Role of Heat Shock Proteins in Viral Infection. In Pockley, A. G., Calderwood, S. K., and Santoro, M. G., editors, *Prokaryotic and Eukaryotic Heat Shock Proteins in Infectious Disease*, pages 51–84. Springer Netherlands.

Schmitt, E., Gehrmann, M., Brunet, M., Multhoff, G., and Garrido, C. (2007). Intracellular and extracellular functions of heat shock proteins: repercussions in cancer therapy. *Journal of Leukocyte Biology*, 81(1):15–27.

Silver, J. T. and Noble, E. G. (2012). Regulation of survival gene hsp70. *Cell Stress and Chaperones*, 17(1):1–9.

Sørensen, J. G., Kristensen, T. N., and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecology Letters*, 6(11):1025–1037.

Teoh, S. and Jones, R. (1978). B chromosome selection and fitness in rye. *Heredity*, 41(1):35–48.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2007). Nucleolus size variation during meiosis and NOR activity of a B chromosome in the grasshopper *Eyprepocnemis* plorans. *Chromosome Research*, 15(6):755–765.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2009). Quantitative analysis of NOR expression in a B chromosome of the grasshopper *Eyprepocnemis* plorans. *Chromosoma*, 118(3):291–301.

Teruel, M., Sørensen, J. G., Loeschcke, V., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2011). Level of heat shock proteins decreases in individuals carrying B-chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 132(1-2):94–99.

Tranter, M., Helsley, R. N., Paulding, W. R., McGuinness, M., Brokamp, C., Haar, L., Liu, Y., Ren, X., and Jones, W. K. (2011). Coordinated post-transcriptional regulation of Hsp70.3 gene expression by microRNA and alternative polyadenylation. *The Journal of Biological Chemistry*, 286(34):29828–29837.

Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. V., and Simonet, G. (2009). Identification and validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. *BMC Molecular Biology*, 10:56.

Vandesompele, J., Preter, K. D., Pattyn, F., Poppe, B., Roy, N. V., Paepe, A. D., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7):1–12.

Wang, A. M., Morishima, Y., Clapp, K. M., Peng, H.-M., Pratt, W. B., Gestwicki, J. E., Osawa, Y., and Lieberman, A. P. (2010). Inhibition of Hsp70 by methylene blue affects signaling protein function and ubiquitination and modulates polyglutamine protein Degradation. *The Journal of Biological Chemistry*, 285(21):15714–15723.

Wang, B., Li, F., Dong, B., Zhang, X., Zhang, C., and Xiang, J. (2006). Discovery of the genes in response to white spot syndrome virus (WSSV) infection in *Fenneropenaeus chinensis* through cDNA microarray. *Marine Biotechnology*, 8(5):491–500.

Wang, X. and Seed, B. (2006). High-throughput primer and probe design. In Tevfik-Dorak, M., editor, *Real-time PCR*, volume 1, pages 93–106. Taylor and Francis Group.

Wheeler, J. C., King, V., and Tower, J. (1999). Sequence requirements for upregulated expression of Drosophila hsp70 transgenes during aging. *Neurobiology of Aging*, 20(5):545–553.

Zhao, J., Xie, J., Xu, Y., Zhang, Z., and Zhang, N. (2005). Expression of heat stress protein 70 mRNA in patients with chronic obstructive pulmonary disease and its significance. *Journal of Huazhong University of Science and Technology [Medical Sciences]*, 25(1):20–23.

Zhao, M., Tang, D., Lechpammer, S., Hoffman, A., Asea, A., Stevenson, M. A., and Calderwood, S. K. (2002). Double-stranded RNA-dependent protein kinase (pkr) is essential for thermotolerance, accumulation of HSP70, and stabilization of ARE-containing HSP70 mRNA during stress. *The Journal of Biological Chemistry*, 277(46):44539–44547.

Zurita, S., Cabrero, J., López-León, M. D., and Camacho, J. P. M. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, 52(1):274–277.

3

Not so silenced: B chromosomes contain active protein-coding genes

Beatriz Navarro-Domínguez[1], Francisco J. Ruíz-Ruano[1], Josefa Cabrero[1], Jose María Corral[2,3], María Dolores López-León[1], Timothy F. Sharbel[2,4], Juan Pedro M. Camacho[1]

[1]*Departamento de Genética, Universidad de Granada*

[2]*Leibniz Institute for Plant Genetics and Crop Plant Research (IPK), Gatersleben (Germany)*

[3]*Department of Bioanalytics, Coburg University of Applied Sciences and Arts*

[4]*Global Institute for Food Security (GIFS), University of Saskatchewanx*

Abstract. For many years, parasitic B chromosomes have been considered genetically inert elements. Here we show the presence of nine protein-coding genes in the B chromosome of the grasshopper *Eyprepocnemis plorans*. Four of these genes (*CIP2A*, *GTPB6*, *KIF20A*, and *MTG1*) were complete in the B chromosome whereas the five remaining (*CKAP2*, *CND3*, *HYI*, *MYCB2* and *SLIT*) were truncated.

Five of these genes (*CIP2A*, *CKAP2*, *CND3*, *KIF20A*, and *MYCB2*) were significantly up-regulated in B-carrying individuals, as expected if they were actively transcribed from the B chromosome. This conclusion is supported by three truncated genes (*CKAP2*, *CND3* and *MYCB2*) which showed up-regulation only in the regions being present in the B chromosome. Our results indicate that B chromosomes are not so silenced as was hitherto believed. Interestingly, the five active genes in the B chromosome code for functions related with cell division, which is the main arena where B chromosome destiny is played. This suggests that B chromosome evolutionary success can lie on its gene content.

Introduction

Supernumerary (B) chromosomes are mainly composed of repetitive satellite and ribosomal DNA sequences, and their unusual meiotic dynamics and dispensable nature makes them an easy target for the integration and expansion of transposable elements (TEs) (Puertas et al., 2000; Camacho, 2005). Many other kinds of DNA sequences could be interspersed among these repeats, mostly coming from the standard (A) genome, making B chromosomes a complex mosaic of DNA.

However, the high enrichment of repetitive DNA in B chromosomes has hindered the detection of protein-coding genes in them, with only few exceptions. For instance, B chromosomes in the red fox (*Vulpes vulpes*) and the racoon dog (*Nyctereutes procyonoides*) carry the proto-oncogene

c-kit (Graphodatsky et al., 2005); those in the migratory locust (*Locusta migratoria*) carry H3 and H4 histone genes (Teruel et al., 2010); and those in the cyclid fish *Lithochromis rubripinnis* carry five different protein-coding genes (Yoshida et al., 2011). Likewise, it has been found that B chromosomes of rye (*Secale cereale*) bear gene-derived sequences, as well as specific repeats and insertions of organellar DNA (Martis et al., 2012). These authors compared B chromosome sequences with the ancestral DNA sequences in two standard (A) chromosomes, and identified fragmentation and pseudogenization of the gene-like sequences located in the Bs. Recently, Valente et al. (2014) analyzed the gene content of B chromosomes by comparing Illumina sequences from 0B and 2B genomes in the cichlid fish *Astatotilapia latifasciata*, with subsequent confirmation by quantitative PCR and FISH mapping, complemented by Roche 454 sequencing of a microdissected B chromosome. They mapped those sequences to the scaffolds of the genome of the cichlid *Metriaclima zebra*, and detected blocks where the B+/B- coverage was higher than the average calculated for scaffolds not located in the B. Those blocks with high coverage in the +B library where further annotated with the *M. zebra* genome. Among them, they found more than 5,000 sequences putatively identified as genes, besides repeats and transposable elements. Most of the genes found in the B chromosome were fragmented, but a few of them were complete and associated with functions such as microtubule organization, kinetochore structure, recombination and progression through the cell cycle, which may be involved in the transmission and maintenance of the parasitic chromosome.

The general belief that B chromosomes are genetically inactive was enforced by experiments using tritiated uridine in the grasshoppers

Myrmeleotettix maculatus and *Chorthippus paralelus* (Fox et al., 1974), and in the rodent *Apodemus peninsulae* (Ishak et al., 1991). Nevertheless, gene expression in B chromosomes is probably behind the remarkable effects reported in some cases, such as the fungus *Nectria haematococca*, whose B chromosome carries a gene which confers resistance to pisatin, an antibiotic produced by the pea host plant (Miao et al., 1991), or the paternal sex ratio (PSR) chromosome of the parasitic wasp *Nasonia vitripennis*, which causes the conversion of diploid zygotes (destined to be females) to haploid males (Werren, 1991). More recently, it has been shown that B chromosomes influence sex determination in cichlid fishes (Yoshida et al., 2011). Consistently, gene expression has recently been found in B chromosomes of several species. For instance, the presence of rRNA transcripts specifically coming from B chromosomes has been shown in the plant *Crepis capillaris* (Leach et al., 2005), the parasitic wasp *Trichogramma kaykai* (Van Vugt et al., 2005) and the grasshopper *Eyprepocnemis plorans* (Ruiz-Estévez et al., 2012). Also, Carchilan et al. (2009) showed, in rye, the presence of B-specific transcribed DNA sequences belonging to high-copy number families with similarity to mobile elements, while Zhou et al. (2012) have characterized a B chromosome linked scaffold that contains an actively transcribed unit. Recently, it has been suggested that B chromosome content might influence the expression of genes located in the standard A genome (Banaei-Moghaddam et al., 2013; Huang et al., 2016).

The B chromosome system of *E. plorans* is very widespread and highly polymorphic, with Bs being present in almost all populations from the circum- Mediterranean region (López-León et al., 2008). The presence of a same type of B in populations from Spain, Morocco, Tunisia and

Sicilia suggests a recent invasion into these areas (Cabrero et al., 2014). The high success of these B chromosomes, which are present in almost all populations hitherto analyzed, except those in the headwaters area of the Segura River basin in Spain (Cabrero et al., 1997), has arisen from the transmission advantage they show during female meiosis (Zurita et al., 1998). This drive is suppressed by the A chromosomes (Herrera et al., 1996; Perfectti et al., 2004) as a consequence of the arms race between A and B chromosomes predicted by the near-neutral model of B chromosome evolution (Camacho et al., 1997).

B chromosomes of *E. plorans* are able to transcribe their rDNA and organize a nucleolus (Ruiz-Estévez et al., 2012), although this occurs only in a minority of males from most populations (Ruiz-Estévez et al., 2013), and the rRNA contribution by the B to total rRNA is insignificant compared to that from A chromosomes (Ruiz-Estévez et al., 2014), suggesting that B chromosomes in this species are highly repressed. Here we analyze the possible presence of protein-coding genes in the B chromosome of this species by means of NGS analysis of 0B and 4B male genomes and their mapping against the coding sequence (CDS) regions of a *de novo* transcriptome built with all the reads obtained from 0B and 1B female RNAs. This approach has revealed the presence of nine protein-coding genes, five of which are actively transcribed in males and females. While previous studies have provided evidence for gene fragments in B chromosomes, this is the first demonstration of full-length functional protein coding sequences in them.

Materials and methods

Materials

E. plorans individuals were collected in Torrox (Málaga) (Table 1), a population where the prevalent B chromosome variant is B24 (Cabrero et al., 2014). Males were anaesthetised before dissecting out testes, one of which was fixed in 3:1 ethanol-acetic acid for cytological analysis, while the other testis and body remains were frozen in liquid nitrogen for nucleic acid extraction. The number of B chromosomes in males was determined by C-banding of testicular follicles. In the case of the two females used for transcriptome analysis, the number of B chromosomes was determined in interphase hemolymph nuclei (Cabrero et al., 2006). The full bodies of the two females selected for transcriptome analysis (one 0B and one with 1B) were frozen in liquid nitrogen. The remaining females were anaesthetized before dissecting out the ovaries. A few ovarioles were incubated in 2% colchicine in isotonic insect saline solution for 2h, and then fixed in 3:1 ethanol-acetic acid for cytological analysis. The remaining ovarioles were frozen in liquid nitrogen for gene expression analysis. Body remains were frozen in liquid nitrogen for later nucleic acid extraction. In these females, the number of B chromosomes was determined by C-banding on the colchicine-treated ovarioles. Additionally, B chromosome presence/absence was corroborated in both sexes by PCR amplification of the B-specific SCAR marker, described by Muñoz-Pajares et al. (2011), on genomic DNA.

Table 1: Biological samples used for each experiment in the current work. Bs= Number of B chromosomes, N= Number of individuals.

Experiment	Technique	Sex	Body part	Bs	N
B chromosome gene content	Illumina WGS	Male	Hind leg	0	1
				4	1
	qPCR (DNA)	Male	Body	0	2
				1	6
				2	5
				3	1
Validation of B chromosome expression	Illumina RNA-seq	Female	Full body	0	1
				1	1
	qPCR (cDNA)	Male	Body/Testes	0	2
				1	7
				2	8
				3	6
		Female	Body/Ovary	0	4
				1	8
				2	8
				3	1

Nucleic acid isolation

Genomic DNA (gDNA) was extracted from hind legs using the GenElute Mammalian Genomic DNA Miniprep kit (Sigma). Quality was checked by TBE-agarose gel electrophoresis and also by measuring 260/280 and 260/230 ratios with an Infinite M200 Pro NanoQuant (Tecan).

Total RNA extractions from frozen bodies were performed using a Real Total RNA Spin Plus kit (Durviz), whereas RNA extractions from gonads were done using the RNeasy Lipid Tissue Mini Kit (Qiagen), following manufacturer's recommendations. In both protocols, we carried out a DNAse treatment on the column membrane (20 units during 30 minutes incubation, DNAse Amplification Grade I, Sigma), to avoid gDNA

contamination, which was validated by PCR amplification of rDNA or histone genes in the extracted RNA and subsequent visualization on an agarose gel. In body samples, a second DNAse treatment was performed after the extractions, using the REALSTAR kit (Durviz). Quality and quantity of RNA was assessed with a Tecan's Infinite 200 NanoQuant spectrophotometer and in a denaturing MOPS-agarose gel to ensure the absence of RNA degradation.

Illumina sequencing

Total RNA was extracted from each of two females, one lacking B chromosomes (0B) and the other carrying 1B, whereas gDNA was collected from two males (0B and 4B; Table 1). Each of the four libraries (gDNA_0B, gDNA_4B, RNA_0B, RNA_1B) was sequenced on an Illumina Hiseq2000 platform, each yielding about 5Gb of paired-end reads (2x101 nucleotides). Illumina sequences are available in NCBI SRA database under accession numbers SRR2970625 (gDNA_0B), SRR2970627 (gDNA_4B), SRR2969416 (RNA_0B) and SRR2969417 (RNA_1B).

Analysis of abundance and integrity of protein-coding genes located putatively in the B chromosome

We used an *E. plorans de novo* transcriptome assembly (available at Figshare, Navarro-Domínguez, 2016) as reference for read mapping. We reduced redundancies with CD-HIT-EST (Fu et al., 2012) with local

alignment and greedy algorithm, and grouped those sequences showing 80% or higher similarity in at least 80% of length (options -M 0 -aS 0.8 -c 0.8 -G 0 -g 1). Potential CDSs were predicted with Transdecoder, considering open reading frames (ORFs) longer than 300 nucleotides (Haas et al., 2013).

To search for protein-coding genes residing in the B chromosome, we performed the following analysis. Against the *de novo* transcriptome, we mapped: 1) the reads obtained from the 0B and 4B genomes (gDNA_0B and gDNA_4B libraries), and 2) those obtained from RNA sequencing in 0B and 1B females (RNA_0B and RNA_1B libraries). We mapped the reads using SSAHA2 (Ning et al., 2001) with a minimum alignment score of 40 and 80% minimum identity. We used a custom script[I] to count the number of mapped reads as a measure of abundance. Coverage for each nucleotide position was normalized per library size, haploid genome size and mean coverage in the 0B library (to compensate for the coverage underestimation derived from mapping genomic reads onto a transcriptomic reference), by means of the following equation:

$$NCP = \frac{L_S \times CP_L}{G_S \times CS_{0B}}$$

where NCP= Normalized Coverage per Position for a transcript in gDNA; L_S= Library Size; CP_L= Coverage per Position in the Library; G_S= Genome Size; CS_{0B}= mean Coverage of the Sequence in the 0B library. We calculated a haploid genome size (GS) of 10.525 Gb for the 0B individual and 11.885Gb for the 4B individual, bearing in mind the C-value and DNA content of X and B chromosomes reported in *E. plorans* by Ruiz-Ruano et al. (2011).

[I]https://github.com/fjruizruano/ngs-protocols/blob/master/count_reads_bam.py

We selected coding sequences (CDS) putatively being located in the B chromosome on the basis of the following criteria: 1) log_2 of the quotient between the number of mapped reads in the 4B and 0B gDNA libraries (4B/0B) was equal to or higher than 1.58. This figure was inferred by assuming that a single-copy gene would have two copies in a diploid 0B genome, whereas, if there is one copy of that gene in each B chromosome, the 4B genome would carry six copies, i.e. three times more copies than the 0B one, so that $log_2(3) = 1.58$. 2) Coverage (i.e. the sum of normalized mapped reads on 0B and 4B gDNA libraries) should be higher than 40.

We analyzed the integrity and structure of the B-located genes using a transcriptome assembly from 12 RNA-Seq libraries of *Locusta migratoria* (Ruiz-Ruano et al., in preparation) and the *L. migratoria* genome assembled by Wang et al. (2014) (accession number AVCP000000000). We searched for homologous sequences in the *L. migratoria* transcriptome with BLASTN (Altschul et al., 1990). Using the sequence with the lowest E-value as reference for a second BLASTN (Altschul et al., 1990), we searched for hologous sequences in the *L. migratoria* genome. We aligned our transcripts from *E. plorans* to the selected genomic sequences with the exonerate software (Slater and Birney, 2005) to search for exon junction sites. If an *E. plorans* transcriptome sequence was incomplete with respect to the *L. migratoria* one, we used the *L. migratoria* transcriptome as reference for mapping the *E. plorans* genomic reads on the above mentioned sequences with SSAHA2 software (Ning et al., 2001).

Gene function classifications were performed according to two standarized methods: Gene Ontology (GO) (Ashburner et al., 2000)

and Eukaryotic Orthologous Groups (KOG) (Tatusov et al., 2003). GO assignments to predicted proteins were performed with Blast2GO (Conesa et al., 2005), and KOG classification was performed with the WebMGA software (Wu et al., 2011), searching the KOG database of NCBI.

Relative quantification of genomic abundance and transcription analysis of the nine protein-coding genes residing in the B chromosome

Quantification of relative copy number for the B chromosome genes was performed in 16 males from Torrox carrying 0-3 B chromosomes (Table 1). Transcription of B-located genes was analyzed in 23 males and 21 females from Torrox (Table 1), separately in the gonads and the rest of the body. cDNA synthesis was performed by retro-transcription of 100 ng per sample of total RNA using a combination of random and oligo dT hexamers (PrimeScript™ RT reagent Kit, Perfect Real Time, Takara). The cDNA obtained was diluted in RNase-DNase free water for a 1:10 working solution.

Quantitative PCR was performed on a Chromo 4 Real Time PCR thermocycler (Biorad). Primers were designed with Primer3 (Untergasser et al., 2012) (see sequences in Table 2). Each reaction mixture contained 5 µl of gDNA at 5 ng/µl (25 ng gDNA per reaction) or 5 µl of cDNA working solution obtained as described above, 5µl of SensiMix™ SYBR Kit (Bioline) and 2.5 µl of each 2.5 µM primer. Reactions were carried out in duplicate and the coefficient of variation was lower than 8% in all

Table 2: Sequences of all the primers used for qPCR experiments in this work.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
CIP2A	TGGCGCTGGTACTGAGTATG	GATCCACCTGAAGAGCTTGG
CKAP2_1	CAAAATGGCGTGCTGAAAG	CGTCTTTTGATTTAATAGTGGAATTTG
CKAP2_2	TCTTCGATGTTTTGGCCTTC	TGGTCATCATTTGCCAGAGA
CND3_1	GAGGTATGGAACACGCACAA	AGTGGCACGTTTCGTCTTCT
CND3_2	CAACAGCGCCTGTCACTAAA	GCTGAGGTGTCTGCTCACAA
GTPB6	CACTGGAGGACGCGATGT	CCGGAGTGAGATCAAAAGACC
HYI_1	TGTCCGGACGAGTTGACA	CGGAACAGAATGTGGATTGA
HYI_2	CCACATCCAGATTGCACAAG	ACTCCAAACCAATCCAACCA
KIF20A	CAGGGCACAAATGAAAATCC	TTGCTGCTTCTCTTCATCCA
MTG1	AGCTCCAGTAGGTGCAAAGG	GGCCTGCTTCAACATCTCT
MYCB2_1	ACCCGTCACATACACAACGA	CCATCACCATTGCTTGTACG
MYCB2_2	GCAAGGAAGAAGAGGAAGCA	CCAGTGCCATAACCCAGAAC
SLIT1	AACATGCTGCATTGCGACT	CACTTGAAGTCGTGGTCGTG
SLIT2	AACCTGTCGGAAAGAGCAAA	TTCGCACAATGTCAACATCC

cases. We estimated the amplification efficiency (E) of each primer pair in gDNA or cDNA experiments by means of a standard curve performed on a 10-fold dilution series of *E. plorans* gDNA or cDNA mixture from several individuals with different numbers of B chromosomes, which was also used as an external calibrator. Then the relative abundance of each gene in 0B, 1B, 2B and 3B genomes was calculated according to $RQ = E^{Ct_C - Ct_S}$, where RQ= Relative quantity, E= Amplification efficiency (fold increase per cycle), Ct_C= Ct value of the calibrator sample and Ct_S= Ct value of each sample.

RQs of the expression analysis were calculated according to the same formula, but RQ values were normalized by the geometrical average of several housekeeping genes, i.e. Ribosomal Protein 49 (*RP49*), Actin 5C (*ACT*), Armadillo (*ARM*), and α-tubulin 1A (*TUB*), selected using GeNorm (Vandesompele et al., 2002). We used *RP49+ACT+ARM* for female bodies, *ACT+ARM* for ovaries, *ACT+TUB* for male bodies and *ACT+RP49* for testis. Amplification, sequencing, efficiency and stability analysis of these housekeeping genes in *E. plorans* was previously performed in Navarro-Domínguez et al. (in press; see Chapter 2).

Statistical analysis

For qPCR validation of gene presence in the B chromosome, we should expect that those genes actually residing in the B chromosome would show RQ values on gDNA linearly increasing with the number of B chromosomes. This relationship was tested by the Spearman's rank correlation test. Differential gene expression was tested by means of

Kruskal-Wallis tests comparing individuals with different number of B chromosomes. In all cases, the sequential Bonferroni test was applied to minimize type I errors.

Results

At least nine protein-coding genes reside in the parasitic chromosome

Clustering and identification of potential CDSs in the *de novo* transcriptome yielded 13,190 sequences that were used as reference for mapping and comparative coverage analysis in 0B and 4B gDNA libraries.

The comparative analysis of read mapping for 0B and 4B genomes on the CDS regions of the *de novo* transcriptome detected nine protein-coding genes putatively residing in the B chromosome, as they were highly over-represented in the 4B genome (Fig. 1). Orthologues of the nine *E. plorans* contigs in the *L. migratoria* genome (Wang et al., 2014) included *CIP2A* (CIP2A protein), *CKAP2* (Cytoskeleton-associated protein 2), *CND3* (Condensin complex subunit 3), *GTPB6* (GTP-binding protein 6), *HYI* (Hydroxypiruvate isomerase), *KIF20A* (Kinesin-like protein KIF20A), *MTG1* (Mitochondrial GTPase 1), *MYCB2* (E3 ubiquitin-protein ligase MYCBP2) and *SLIT* (protein SLIT). The sequences of these nine transcripts can be found in GenBank under

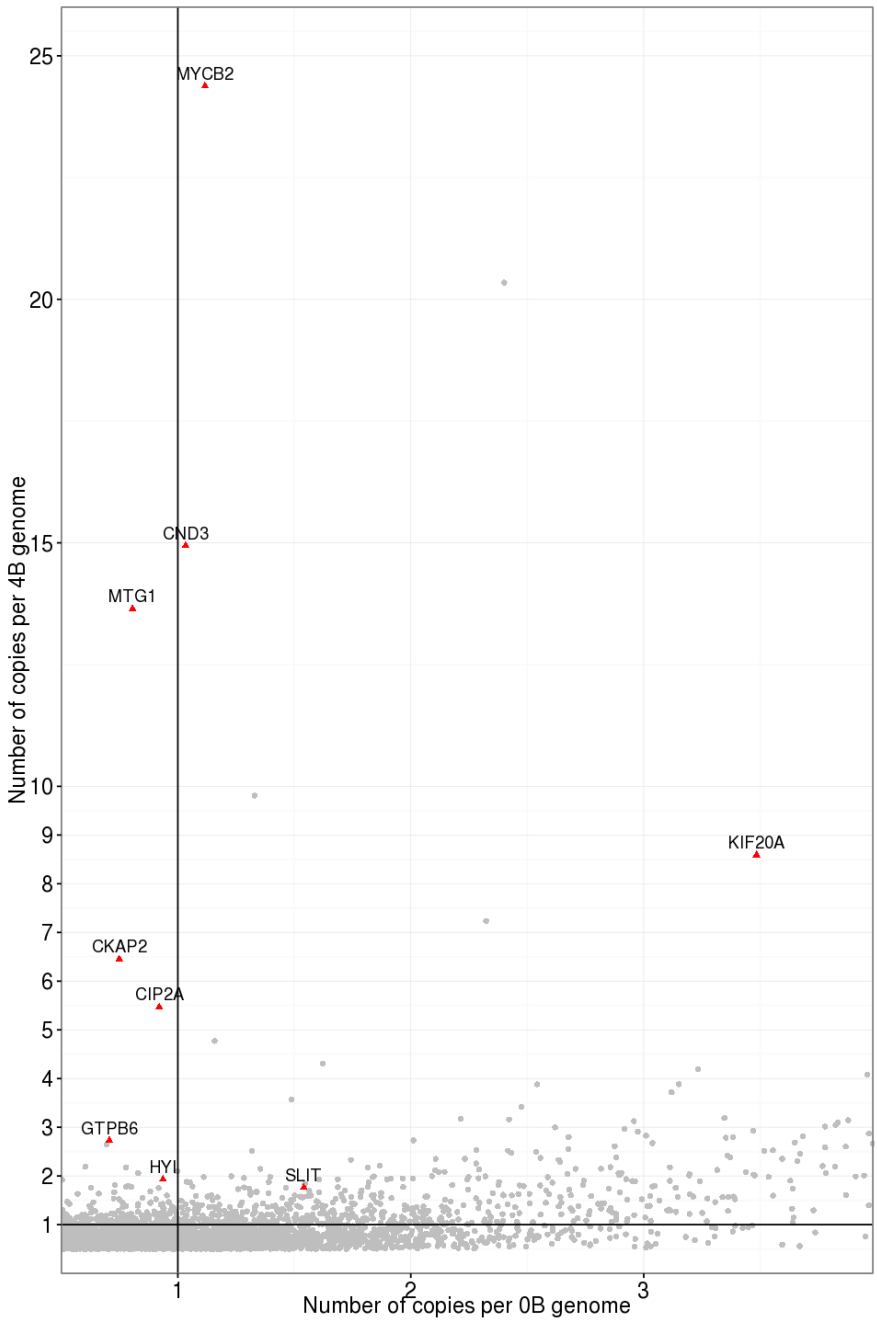

Figure 1: Identification of protein-coding genes located on B chromosomes of the grasshopper *E. plorans*, using copy number estimation in the 0B (X axis) and 4B (Y axis) genomes for DNA sequences in the *de novo* assembled transcriptome. Labeled red triangles indicate the 9 genes found in the B chromosome. Note that all of them are single-copy in the 0B genome, except *KIF20A* which shows three copies.

accession numbers KX034164 to KX034172 (Table 3). Based on the relative coverage in 0B and 4B libraries, we estimate that all of them have one copy per haploid standard genome, except *KIF20A*, which has 3 in the standard genome, and diverse number of copies on the B chromosome, ranging from one to ten (Fig. 1, Table 3). Six of these genes (*HYI*, *CIP2A*, *CKAP2*, *CND3*, *MTG1* and *MYCB2*) showed high 1B/0B ratios when we compared read counts between the two transcriptome libraries, suggesting that they could be actively transcribed on the B chromosome (Fig. 2).

According to the coverage pattern observed in 0B and 4B genomes, full length coding sequences (CDS) were found in the 4B genome for four genes: *CIP2A*, *GTPB6*, *KIF20A* and *MTG1*, suggesting that these genes might be complete in the B chromosome. As shown in Fig. 3A and Suppl. Figs. S1A, S2A and S3A, coverage for these genes in the gDNA_4B library was uniformly high along all CDS length. The five remaining genes (*MYCB2*, *CND3*, *CKAP2*, *HYI*, and *SLIT*), however, appeared to be incomplete in the B chromosome. *MYCB2* was clearly truncated (Fig. 4A), showing only the last 5,764 nucleotides of the 3' side, out of of the 14,434 nucleotides reported for the homologous gene in *L. migratoria*. In *CND3*, the last five exons (exons 20-24) of the CDS reported for the *L. migratoria* genome show low coverage in the *E. plorans* 4B genome (Fig. S4A). In *CKAP2*, the 5' UTR and 396 nucleotides of the 5' end of the CDS seem to be missing (Fig. S5A). For *HYI*, we only observed 85 nucleotides in the 3' end of the CDS plus the 3' UTR, but this gene actually showed very low coverage in both gDNA libraries, and its integrity was difficult to assess (Fig. S6A). In the case of the *SLIT* gene, only the last two exons of the 5' end were in the B

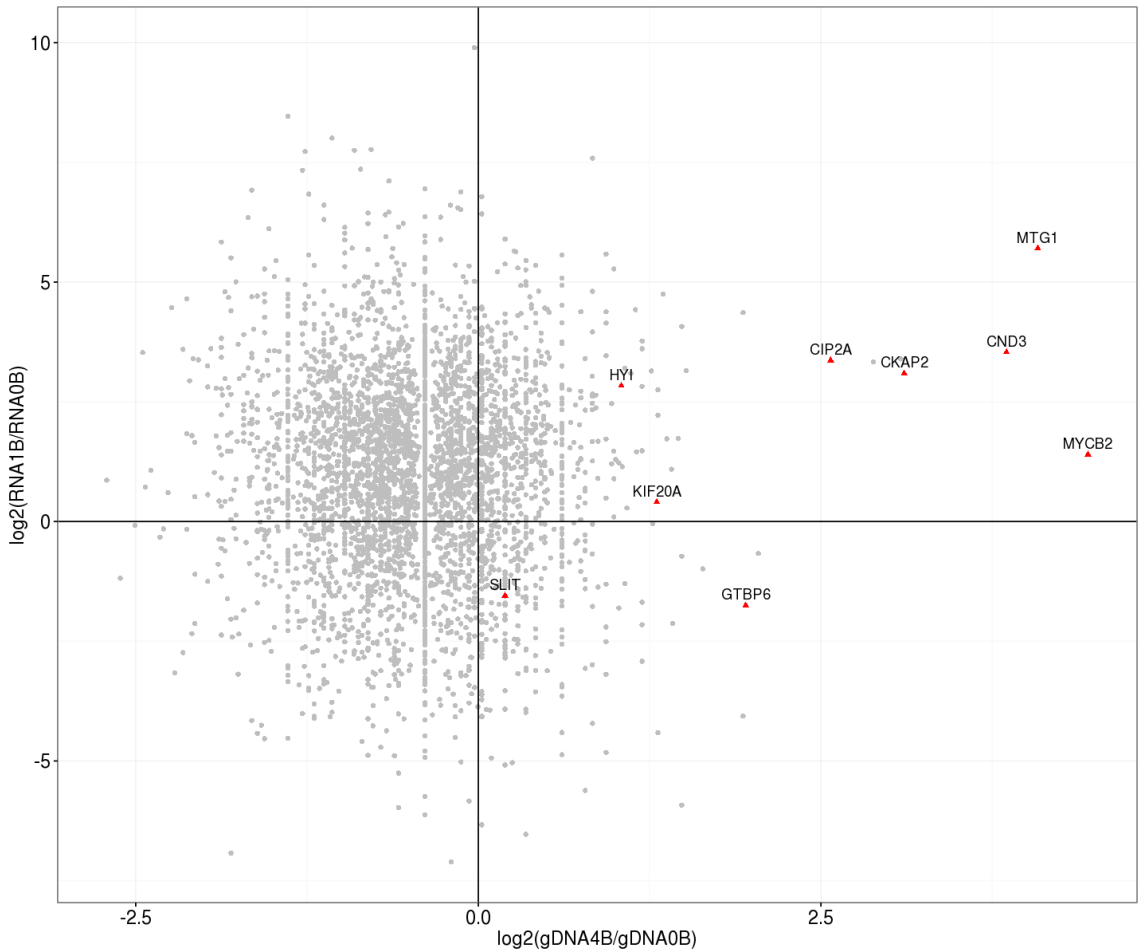

Figure 2: Graphical representation of a comparison between the genomic 4B/0B (X axis) and transcriptomic 1B/0B (Y axis) ratios. Each dot corresponds with a DNA sequence in the *de novo* transcriptome. Labeled red triangles indicate the 9 genes located in the B chromosome of *E. plorans*.

Table 3: Short description of the genes located in the B chromosome. Acc. No.= GenBank accession number for *E. plorans* transcript. B-copy may be complete (C) or truncated (T). B-activity refers to the qPCR experiments on cDNA (see Table S2). B chr. = B chromosome.

Gene	Protein	Acc.No.	B-copy	B-activity	Copy number per	
					Haploid A set	B chr.
CIP2A	Protein CIP2A	KX034164	C	Yes	1	3
CKAP2	Cytoskeleton-associated protein 2	KX034165	T	Yes	1	3
CND3	Condensin complex subunit 3	KX034166	T	Yes	1	6
GTPB6	GTP-binding protein 6	KX034167	C	No	1	1
HYI	Hydroxypiruvate isomerase	KX034168	T	No	1	4
KIF20A	Kinesin-like protein KIF20A	KX034169	C	Yes	3	4
MTG1	Mitochondrial GTPase 1	KX034170	C	No	1	8
MYCB2	E3 ubiquitin-protein ligase MYCBP2	KX034171	T	Yes	1	10
SLIT	Protein slit	KX034172	T	No	1	1

chromosome, out of the 30 exons reported in *L. migratoria* (Fig. S7A).

Consistently, qPCR experiments on gDNA from males carrying 0-3 B chromosomes showed that genomic abundance for these nine genes increased linearly with the number of B chromosomes (Figs. 3B and 4B, Table S1 and Figs. S1B-S7B). Remarkably, in the truncated genes this linear relationship was observed in the regions with high coverage in the gDNA_4B library, indicating their location in the B chromosome, and no relationship with B number was observed for the low coverage region of all five incomplete genes (*MYCB2*, *CND3*, *CKAP2*, *HYI* and *SLIT*), confirming the absence of this region in the B chromosome and the incompleteness of the genes (see Fig. 4B and Figs. S4B-S7B).

The parasitic chromosome is transcriptionally active

The analysis of differential gene expression between B-carrying and B-lacking individuals, by means of qPCR, revealed that five out of the nine genes located on the B chromosome (*CIP2A*, *CKAP2*, *CND3*, *KIF20A*, and *MYCB2*) showed significant up-regulation in B-carrying males and females from Torrox (Table S2), suggesting that some of the B-located gene copies are transcribed. Remarkably, three of these genes (*MYCB2*, *CND3* and *CKAP2*) are truncated in the B chromosome and showed differential expression only for the gene regions contained in the B chromosome but not for those regions being missing (see Fig. 4C, Figs. S4C and S5C, and Table S2). This strongly supports that the up-regulation of these genes is due to the activity of the B chromosome copies and not simply to up-regulation of the A chromosome gene copies.

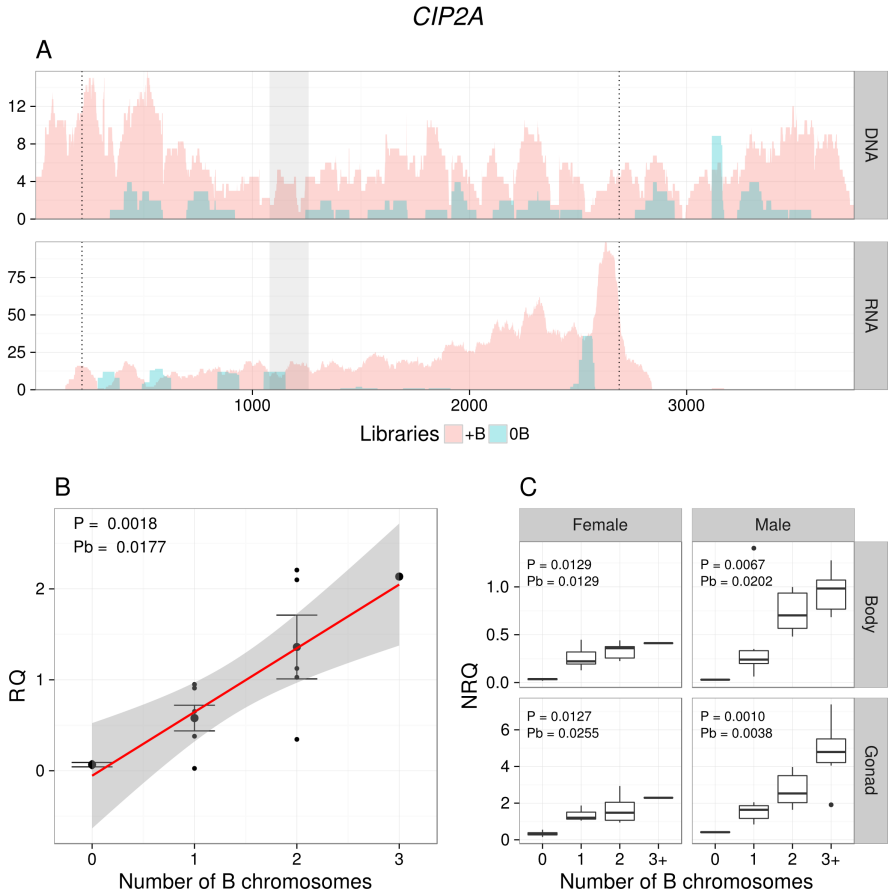

Figure 3: Coverage for the *CIP2A* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Note that coverage was higher in the 4B library across the complete sequence length, including the full CDS (delimited by the dotted vertical lines), the 5' UTR (from the 5' end to the first dotted line) and 3' UTR (from the second dotted line to the 3' end). Likewise, note the higher coverage for this transcript in the B-carrying RNA library. The shaded zone in A marks the region amplified by qPCR. qPCR on gDNA (B) revealed that genomic copy number for the *CIP2A* gene increases with B chromosome number, following a dose-dependent pattern, thus supporting its presence in the B chromosome. qPCR on cDNA (C) revealed that *CIP2A* is expressed in all tissues and sexes analyzed, also following a dose-depending pattern and suggesting the active transcription of B chromosome gene copies. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

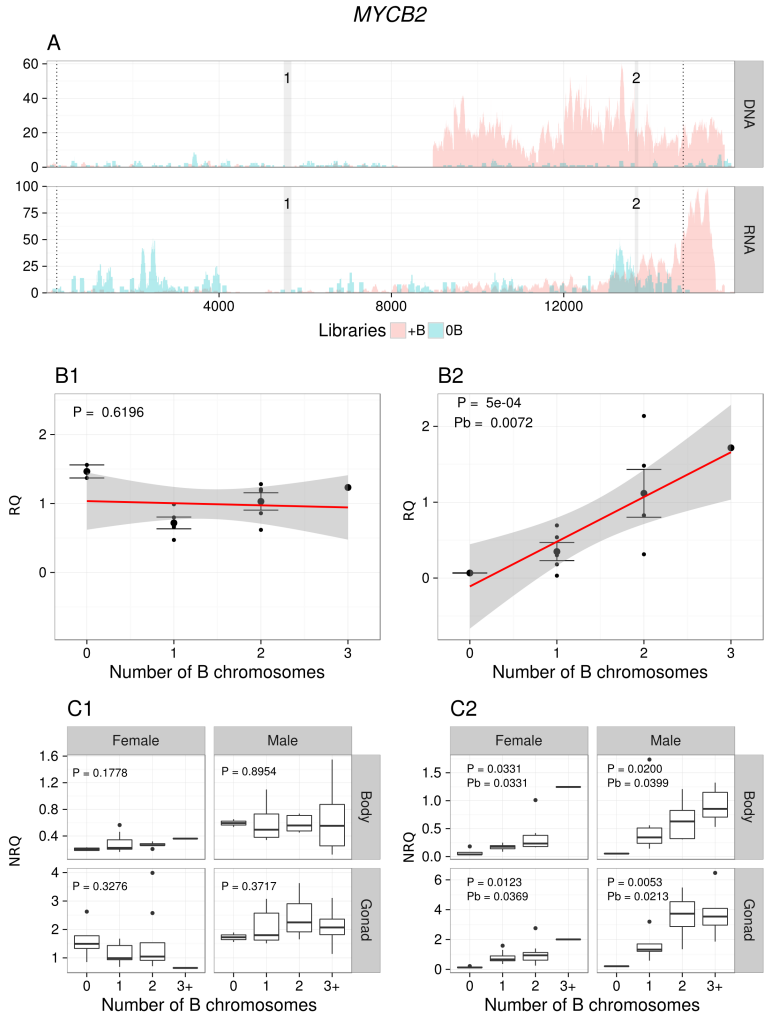

Figure 4: Coverage for the *MYCB2* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Note that less than half of *MYCB2* CDS (between the dotted vertical lines) showed high coverage in the 4B library, specifically from position 8961 to the 3' end (A), suggesting that this B chromosome gene is truncated. Two regions were selected for qPCR amplification of this gene, one within the region being apparently absent in the B chromosome (shaded zone 1) and the other within the region being present in it (shaded zone 2). qPCR on gDNA with zone 1 primers showed that copy number for this gene region was independent on the number of B chromosomes (B1). However, qPCR on gDNA with zone 2 primers showed that abundance of this *MYCB2* gene region increased with B chromosome number in a dose-dependent pattern (B2). Likewise, qPCR on cDNA showed that *MYCB2* expression was independent of B chromosome number when probed with zone 1 primers (C1) but it increased in a dosage-dependent pattern with zone 2 primers (C2), suggesting the active transcription of B chromosome truncated gene copies. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

For *CIP2A* (Fig. 3C) and *KIF20A* (only in males, Fig. S2C) genes, expression level increased with B chromosome number, and a similar dosage effect was observed for *MYCB2* (Fig. 4C), *CND3* (Fig. S4C) and *CKAP2* (Fig. S5C) in the case of gene regions being present in the B chromosome. No significant differential expression was found for the remaining genes and neither for the B chromosome missing regions of *MYCB2* (Fig. 4C), *CND3* (Fig. S4C) and *CKAP2* (Fig. S5C). Taken together, these results reinforce the conclusion that most B chromosome genes are actively transcribed.

Gene Ontology (GO) analysis for these nine genes revealed their implication in some biological processes likely profitable for a parasitic B chromosome, such as the regulation of mitotic cell cycle (*KIF20A*, *CND3*, and *CKAP2*), DNA replication and regulation of transcription (*CKAP2* and *MYCB2*), apoptotic processes and regulation of cell death (*CKAP2*), chromosome condensation and organization (*CND3*), cell-cell signaling and cellular response to stimulus (*SLIT*), and reproductive structure development (*SLIT*) (see Table S3). The EuKaryotic Orthologous Groups (KOG) classification of these nine genes also gave interesting indications of their potential functions (Table 4), some of which being highly valuable for the advantageous transmission of this parasitic chromosome. For instance, *CIP2A* and *KIF20A* have functions related with cytoskeleton and thus microtubule dynamics, and *CND3* is related with chromosome condensation, chromatin structure and dynamics and thus cell cycle control. The two former genes appear to be complete and active in the B chromosome, but the latter is truncated thus probably rendering non-translated or inactive transcripts.

Table 4: KOG classification of genes located on the B chromosome

Gen	Hit	E-Value	Description	Class	Class Description
CND3	KOG2025	7,00E-031	Chromosome condensation complex Condensin, subunit G	B/D	Chromatin structure and dynamics Cell cycle control and mitosis
CIP2A	KOG0161	4,00E-010	Myosin class II heavy chain	Z	Cytoskeleton
GTPB6	KOG0410	2,00E-085	Predicted GTP binding protein	R	General function prediction only
HYI	KOG4518	2,00E-075	Hydroxypyruvate isomerase	G	Carbohydrate transport and metabolism
KIF20A	KOG0247	3,00E-084	Kinesin-like protein	Z	Cytoskeleton
MTG1	KOG2485	4,00E-086	Conserved ATP/GTP binding protein	R	General function prediction only
MYCB2	KOG1428	7,00E-084	Neuronal presynaptic protein Highwire/PAM/RPM-1	T	Signal transduction mechanisms
SLIT	KOG4237	0.0	Extracellular matrix protein slit	W/T	Extracellular structures Signal transduction mechanism

Discussion

As intranuclear parasites, B chromosomes mimic A chromosomes in many respects, such as the structure and organization of the DNA sequences contained in them, although they are usually heterochromatic and, as such, assumed to be genetically inert elements (for review, see Camacho, 2005). Recently, transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (*Capreolus pygargus*) has been shown (Trifonov et al., 2013), while Banaei-Moghaddam et al. (2013) used NGS to demonstrate that about 15% of the pseudogene-like fragments on B chromosomes are transcribed following a pattern related to genotype and tissue type, with some of them apparently playing a role in trans-regulation of genes located in the A chromosomes. Likewise, Valente et al. (2014) analyzed transcriptome sequences from the cichlid fish, *Pundamilia nyererei*, and some of them showed high sequence similarity with the B-encoded variants for the Separin, *TUBB1* and *KIF11* genes found in *Astatotilapia latifasciata*, thus suggesting that *P. nyererei* might have B chromosomes expressing these genes. These findings clearly contradict the "B genetic inertness" hypothesis, although whether B chromosomes express truly functional genes (i.e. generate proteins) or regulatory factors (i.e. small RNAs) remains to be tested. In fact, the only B chromosome genes which have been shown to yield functional transcripts are those for rRNA in the grasshopper *E. plorans*, as they give rise to the expected phenotype, i.e. a nucleolus (Ruiz-Estévez et al., 2012).

Here we show that B chromosomes contain at least 9 protein-coding

genes which are actively transcribed in *E. plorans*, five of which were active in the population analyzed. Many of the B-derived transcripts might be functionless because the gene copies on the B chromosome are incomplete, so that their translation would yield anomalous polypeptides posing metabolic stress on cells. Alternatively, these gene fragments could interfere A chromosome gene expression by competitively binding transcription factors (Seo et al., 2012; Poliseno et al., 2015; Hirotsune et al., 2003). However, we cannot rule out that the transcripts from some B-located genes, being apparently complete, can be functional, as previously observed for 45S rRNA transcripts (Ruiz-Estévez et al., 2012). Of course, the possibility that a B chromosome can contribute gene products which are useful for its own survival (e.g. through cell division) is a new and interesting prospect in B chromosome research.

Three genes which are truncated in the B chromosome (*CKAP2*, *CND3* and *MYCB2*) show up-regulation only for the region being present in the B chromosome, indicating that gene copies located on the B chromosome are transcribed. This might rule out the alternative possibility that their up-regulation is due to over-expression of the gene copies located in the A chromosomes instead of transcription of the B-copies, except in the case that the B copy is somehow post transcriptionally affecting gene regulation in the As. *CKAP2* codes for a cytoskeleton-associated protein which localizes to spindle poles and microtubules from prophase to anaphase (Seki and Fang, 2007), and seems to play an important role in chromosome segregation and stability (Hong et al., 2009; Case et al., 2013). *CND3* codes for a subunit of the condensin chromosome condensation complex, subunit G (Ono et al., 2003), whereas *MYCB2* encodes a component protein of the anaphase promoting complex (APC)

governing the exit from mitosis (Sivakumar and Gorbsky, 2015). If these transcripts were translated, they would presumably yield non functional truncated polypeptides, especially *MYCB2*, where more than half of the CDS is missing in the B chromosome.

It is also conceivable that, as suggested by Banaei-Moghaddam et al. (2013), these gene fragments present in the B could act as trans modulators, affecting the activity of its counterparts located in the A genome. A possible example of this could be the down-regulation shown by the *GTPB6* gene in the 1B RNA library (see Fig. S1A), meaning that it would be repressed in presence of the B chromosome, and a similar trend was observed in qPCR experiments on ovary (Fig. S1C), although it was not significant (Table S2).

Remarkably, the two remaining up-regulated protein-coding genes (*CIP2A* and *KIF20A*), which were complete in the B chromosome, coded for potentially interesting functions for a parasitic chromosome. *CIP2A* codes for an oncoprotein that inhibits protein phosphatase 2A (*PP2A*), promoting anchorage-independent cell growth and tumor formation, and its overexpression causes premature chromosome segregation and aneuploidy (Pallai et al., 2015). Bearing in mind that B chromosomes most likely originate as a kind of aneuploidy (Hewitt, 1979; Camacho, 2005), it is conceivable that the excess of *CIP2A* gene products, provided by transcription from the B chromosome copies, might be advantageous for B chromosome maintenance. Likewise, *KIF20A* codes for a mitotic kinesin required for chromosomal passenger complex (CPC) transport during cytokinesis (Nguyen et al., 2014) and for appropriate assembly of microtubules at anaphase and metaphase-anaphase transition (Cesario et al., 2006). In Xenopus, Takemoto et al. (2009) showed that *PP2A*

plays a role in the recruitment and targeting of Condensin II and kinesin protein *KIF4a* to chromosomes during mitosis, and *PP2A* inhibition causes Condensin II and *KIF4a* dissociation from assembled chromosomes. In *E. plorans*, B chromosomes carrying active *CIP2A* and *KIF20A* genes could potentially influence the course of cell division for their own benefit, thus revealing their true parasitic nature. In rye, the presence of repetitive DNA sequences in the short arm of the B chromosome promotes mitotic non-disjunction which is the basis for its drive mechanism (Banaei-Moghaddam et al., 2012). In *E. plorans*, B chromosome drive takes place during female meiosis (Herrera et al., 1996; Zurita et al., 1998), and the possibility of manipulating it through gene expression might be the basis for the high success of B chromosomes in this species, as they are present in almost all natural populations hitherto analyzed (López-León et al., 2008).

Taken together, our results show, for the first time, that the secret for B chromosome success may lie on its gene content, as suggested by the active transcription of the complete CDS of *CIP2A* and *KIF20A* and the fragments of *CKAP2*, *CND3* and *MYCB2*, all being genes with functions related with cell division. This hypothesis will be tested, in the near future, as more B chromosomes are analyzed by the present approach.

Acknowledgements

This study was funded by the Spanish Ministerio de Economía y Competitividad (CGL2015-70750-P) and Plan Andaluz de Investigación (CVI-6649). Navarro-Domínguez was supported by a FPI fellowship from the Spanish Ministerio de Ciencia e Innovación (BES-2010-030016).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25(1):25–29.

Banaei-Moghaddam, A. M., Meier, K., Karimi-Ashtiyani, R., and Houben, A. (2013). Formation and expression of pseudogenes on the B chromosome of rye. *The Plant Cell*, 25(7):2536–2544.

Banaei-Moghaddam, A. M., Schubert, V., Kumke, K., Weiβ, O., Klemme, S., Nagaki, K., Macas, J., González-Sánchez, M., Heredia, V., Gómez-Revilla, D., González-García, M., Vega, J. M., Puertas, M. J., and Houben, A. (2012). Nondisjunction in favor of a chromosome: the mechanism of rye B chromosome drive during pollen mitosis. *The Plant Cell*, 24(10):4124–4134.

Cabrero, J., López-León, M. D., Gómez, R., Castro, A. J., Martín-Alganza, A., and Camacho, J. P. (1997). Geographical distribution of B chromosomes in the

grasshopper *Eyprepocnemis plorans*, along a river basin, is mainly shaped by non-selective historical events. *Chromosome Research*, 5(3):194–198.

Cabrero, J., López-León, M. D., Ruíz-Estévez, M., Gómez, R., Petitpierre, E., Rufas, J. S., Massa, B., Kamel Ben Halima, M., and Camacho, J. P. M. (2014). B1 was the ancestor B chromosome variant in the western Mediterranean area in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 142(1):54–58.

Cabrero, J., Manrique-Poyato, M. I., and Camacho, J. P. M. (2006). Detection of B chromosomes in interphase hemolymph nuclei from living specimens of the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 114(1):66–69.

Camacho, J. P., Shaw, M. W., López-León, M. D., Pardo, M. C., and Cabrero, J. (1997). Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper *Eyprepocnemis plorans*. *The American Naturalist*, 149(6):1030–1050.

Camacho, J. P. M. (2005). B chromosomes. In Gregory, T. R., editor, *The Evolution of the Genome*, pages 223–286. Academic Press.

Carchilan, M., Kumke, K., Mikolajewski, S., and Houben, A. (2009). Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma*, 118(5):607–616.

Case, C. M., Sackett, D. L., Wangsa, D., Karpova, T., McNally, J. G., Ried, T., and Camps, J. (2013). CKAP2 ensures chromosomal stability by maintaining the integrity of microtubule nucleation sites. *PloS One*, 8(5):e64575.

Cesario, J. M., Jang, J. K., Redding, B., Shah, N., Rahman, T., and McKim, K. S. (2006). Kinesin 6 family member Subito participates in mitotic spindle assembly and interacts with mitotic regulators. *Journal of Cell Science*, 119(Pt 22):4770–4780.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2go: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18):3674–3676.

Fox, D. P., Hewitt, G. M., and Hall, D. J. (1974). DNA replication and RNA

transcription of euchromatic and heterochromatic chromosome regions during grasshopper meiosis. *Chromosoma*, 45(1):43–62.

Graphodatsky, A. S., Kukekova, A. V., Yudkin, D. V., Trifonov, V. A., Vorobieva, N. V., Beklemisheva, V. R., Perelman, P. L., Graphodatskaya, D. A., Trut, L. N., Yang, F., Ferguson-Smith, M. A., Acland, G. M., and Aguirre, G. D. (2005). The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Research*, 13(2):113–122.

Herrera, J., Lopez-Leon, M., Cabrero, J., Shaw, M., and Camacho, J. (1996). Evidence for B chromosome drive suppression in the grasshopper *Eyprepocnemis plorans*. *Heredity*, 76(6):633–639.

Hewitt, G. M. (1979). Animal Cytogenetics. Volume 3. Insecta 1: Orthoptera, Grasshoppers and crickets. *Animal Cytogenetics. Volume 3. Insecta 1: Orthoptera, Grasshoppers and crickets.*

Hirotsune, S., Yoshida, N., Chen, A., Garrett, L., Sugiyama, F., Takahashi, S., Yagami, K.-i., Wynshaw-Boris, A., and Yoshiki, A. (2003). An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature*, 423(6935):91–96.

Hong, K. U., Kim, E., Bae, C.-D., and Park, J. (2009). TMAP/CKAP2 is essential for proper chromosome segregation. *Cell Cycle (Georgetown, Tex.)*, 8(2):314–324.

Huang, W., Du, Y., Zhao, X., and Jin, W. (2016). B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). *BMC Plant Biology*, 16:88.

Ishak, B., Jaafar, H., Maetz, J. L., and Rumpler, Y. (1991). Absence of transcriptional activity of the B-chromosomes of *Apodemus peninsulae* during pachytene. *Chromosoma*, 100(4):278–281.

Leach, C. R., Houben, A., Field, B., Pistrick, K., Demidov, D., and Timmis, J. N. (2005). Molecular evidence for transcription of genes on a B chromosome in Crepis capillaris. *Genetics*, 171(1):269–278.

López-León, M. D., Cabrero, J., Dzyubenko, V. V., Bugrov, A. G., Karamysheva, T. V., Rubtsov, N. B., and Camacho, J. P. M. (2008). Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of

the grasshopper *Eyprepocnemis plorans plorans*. *Cytogenetic and Genome Research*, 121(3-4):260–265.

Martis, M. M., Klemme, S., Banaei-Moghaddam, A. M., Blattner, F. R., Macas, J., Schmutzer, T., Scholz, U., Gundlach, H., Wicker, T., Šimková, H., Novák, P., Neumann, P., Kubaláková, M., Bauer, E., Haseneyer, G., Fuchs, J., Doležel, J., Stein, N., Mayer, K. F. X., and Houben, A. (2012). Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences*, 109(33):13343–13346.

Miao, V. P., Covert, S. F., and VanEtten, H. D. (1991). A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. *Science*, 254(5039):1773–1776.

Muñoz-Pajares, A. J., Martínez-Rodríguez, L., Teruel, M., Cabrero, J., Camacho, J. P. M., and Perfectti, F. (2011). A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics*, 187(3):853–863.

Nguyen, P. A., Groen, A. C., Loose, M., Ishihara, K., Wühr, M., Field, C. M., and Mitchison, T. J. (2014). Spatial organization of cytokinesis signaling reconstituted in a cell-free system. *Science (New York, N.Y.)*, 346(6206):244–247.

Ning, Z., Cox, A. J., and Mullikin, J. C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Research*, 11(10):1725–1729.

Ono, T., Losada, A., Hirano, M., Myers, M. P., Neuwald, A. F., and Hirano, T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell*, 115(1):109–121.

Pallai, R., Bhaskar, A., Barnett-Bernodat, N., Gallo-Ebert, C., Nickels, J. T., and Rice, L. M. (2015). Cancerous inhibitor of protein phosphatase 2a promotes premature chromosome segregation and aneuploidy in prostate cancer cells through association with shugoshin. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, 36(8):6067–6074.

Perfectti, F., Corral, J. M., Mesa, J. A., Cabrero, J., Bakkali, M., López-León, M. D., and Camacho, J. P. M. (2004). Rapid suppression of drive for a parasitic B chromosome. *Cytogenetic and Genome Research*, 106(2-4):338–343.

Poliseno, L., Marranci, A., and Pandolfi, P. P. (2015). Pseudogenes in human cancer. *Pathology*, page 68.

Puertas, M. J., Jiménez, G., Manzanero, S., Chiavarino, A. M., Rosato, M., Naranjo, C. A., and Poggio, L. (2000). Genetic control of B chromosome transmission in maize and rye. In Olmo, P. D. E. and Redi, P. C. A., editors, *Chromosomes Today*, pages 79–92. Birkhäuser Basel.

Ruiz-Estévez, M., Badisco, L., Broeck, J. V., Perfectti, F., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2014). B chromosomes showing active ribosomal RNA genes contribute insignificant amounts of rRNA in the grasshopper *Eyprepocnemis plorans*. *Molecular Genetics and Genomics*, 289(6):1209–1216.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2012). B-chromosome ribosomal DNA is functional in the grasshopper *Eyprepocnemis plorans*. *PloS One*, 7(5):e36600.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2013). Ribosomal DNA is active in different B chromosome variants of the grasshopper *Eyprepocnemis plorans*. *Genetica*, 141(7-9):337–345.

Ruiz-Ruano, F. J., Ruiz-Estévez, M., Rodríguez-Pérez, J., López-Pino, J. L., Cabrero, J., and Camacho, J. P. M. (2011). DNA amount of X and B chromosomes in the grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenetic and Genome Research*, 134(2):120–126.

Seki, A. and Fang, G. (2007). CKAP2 is a spindle-associated protein degraded by APC/C-Cdh1 during mitotic exit. *The Journal of Biological Chemistry*, 282(20):15103–15113.

Seo, P. J., Hong, S.-Y., Ryu, J. Y., Jeong, E.-Y., Kim, S.-G., Baldwin, I. T., and Park, C.-M. (2012). Targeted inactivation of transcription factors by overexpression of their truncated forms in plants. *The Plant Journal*, 72(1):162–172.

Sivakumar, S. and Gorbsky, G. J. (2015). Spatiotemporal regulation of the anaphase-promoting complex in mitosis. *Nature Reviews. Molecular Cell Biology*, 16(2):82–94.

Slater, G. S. C. and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6:31.

Takemoto, A., Maeshima, K., Ikehara, T., Yamaguchi, K., Murayama, A., Imamura, S., Imamoto, N., Yokoyama, S., Hirano, T., Watanabe, Y., Hanaoka, F., Yanagisawa,

J., and Kimura, K. (2009). The chromosomal association of condensin II is regulated by a noncatalytic function of PP2A. *Nature Structural and Molecular Biology*, 16(12):1302–1308.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., and Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, 4:41.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, 119(2):217–225.

Trifonov, V. A., Dementyeva, P. V., Larkin, D. M., O'Brien, P. C. M., Perelman, P. L., Yang, F., Ferguson-Smith, M. A., and Graphodatsky, A. S. (2013). Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (*Capreolus pygargus*). *BMC Biology*, 11:90.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40(15):e115.

Valente, G. T., Conte, M. A., Fantinatti, B. E. A., Cabral-de Mello, D. C., Carvalho, R. F., Vicari, M. R., Kocher, T. D., and Martins, C. (2014). Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution*, 31(8):2061–2072.

Van Vugt, J. J. F. A., de Nooijer, S., Stouthamer, R., and de Jong, H. (2005). NOR activity and repeat sequences of the paternal sex ratio chromosome of the parasitoid wasp *Trichogramma kaykai*. *Chromosoma*, 114(6):410–419.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C., Wang, Y., He, J., Luo, Y., Wang, Z., Guo, X., Guo, W., Wang, X., Zhang, Y., Yang, M., Hao, S., Chen, B., Ma, Z., Yu, D., Xiong, Z., Zhu, Y., Fan, D., Han, L., Wang, B., Chen, Y., Wang, J., Yang, L., Zhao, W., Feng, Y., Chen, G., Lian, J., Li, Q., Huang, Z., Yao, X., Lv, N., Zhang, G., Li, Y., Wang, J., Wang, J., Zhu, B., and Kang, L. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, 5:2957.

Werren, J. H. (1991). The Paternal-Sex-Ratio chromosome of Nasonia. *The American Naturalist*, 137(3):392–402.

Wu, S., Zhu, Z., Fu, L., Niu, B., and Li, W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, 12:444.

Yoshida, K., Terai, Y., Mizoiri, S., Aibara, M., Nishihara, H., Watanabe, M., Kuroiwa, A., Hirai, H., Hirai, Y., Matsuda, Y., and Okada, N. (2011). B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS genetics*, 7(8):e1002203.

Zhou, Q., Zhu, H.-m., Huang, Q.-f., Zhao, L., Zhang, G.-j., Roy, S. W., Vicoso, B., Xuan, Z.-l., Ruan, J., Zhang, Y., Zhao, R.-p., Ye, C., Zhang, X.-q., Wang, J., Wang, W., and Bachtrog, D. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genomics*, 13:109.

Zurita, S., Cabrero, J., López-León, M. D., and Camacho, J. P. M. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, 52(1):274–277.

Supplementary material

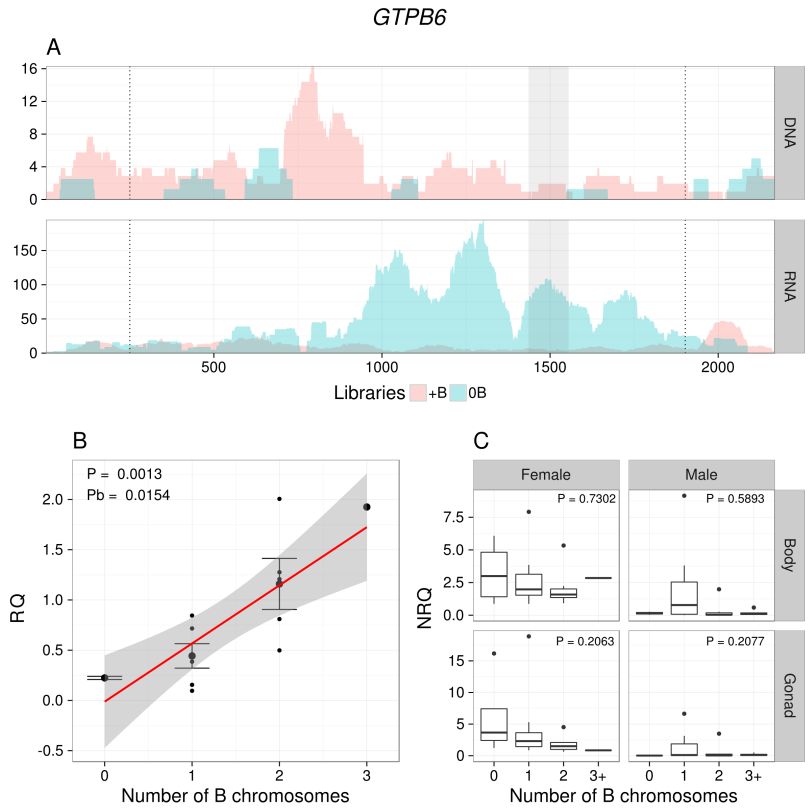

Figura S1: Coverage for the *GTPB6* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Note that coverage was higher in the 4B gDNA library along all sequence length, but it was very low in the B-carrying RNA library (A). The shaded zone in A marks the region amplified by qPCR. qPCR on gDNA (B) revealed that genomic copy number for the *GTPB6* gene increases with B chromosome number following a dose-dependent pattern, thus confirming its presence in the B chromosome. qPCR on cDNA (C) revealed that *GTPB6* expression was independent of B chromosome number in all tissues and sexes analyzed, suggesting that B chromosome gene copies are silenced. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

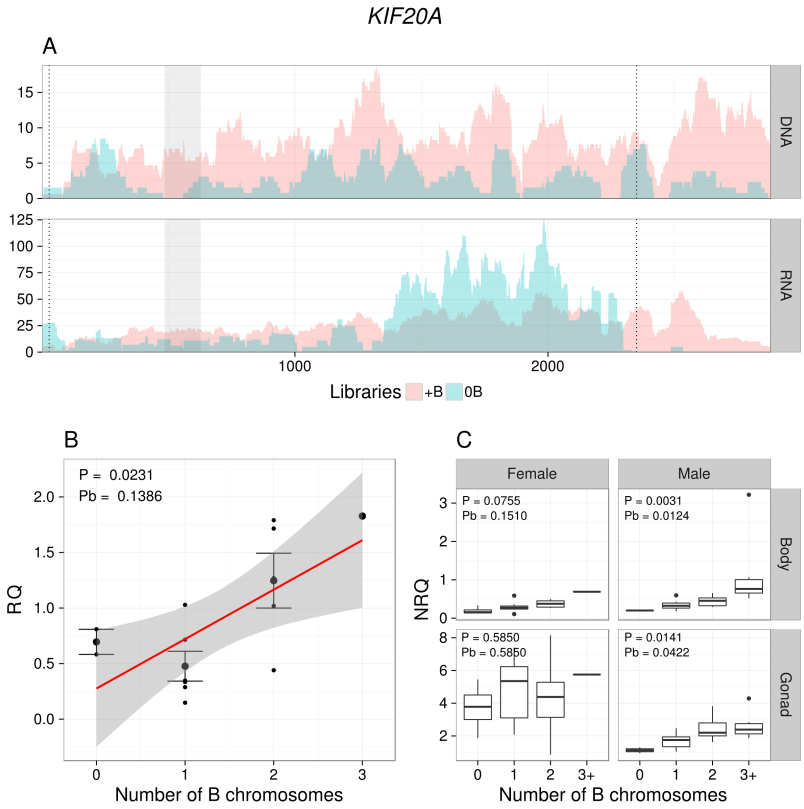

Figura S2: Coverage for the *KIF20A* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Note that coverage was higher in the 4B gDNA library along all sequence length, but not in the B-carrying RNA library (A). The shaded zone in A marks the region amplified by qPCR. qPCR on gDNA (B) revealed that genomic copy number for the *KIF20A* gene increases with B chromosome number following a dose-dependent pattern, thus confirming its presence in the B chromosome, although this association was only marginally significant. qPCR on cDNA (C) showed that *KIF20A* expression tended to increase with B chromosome number only in males. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

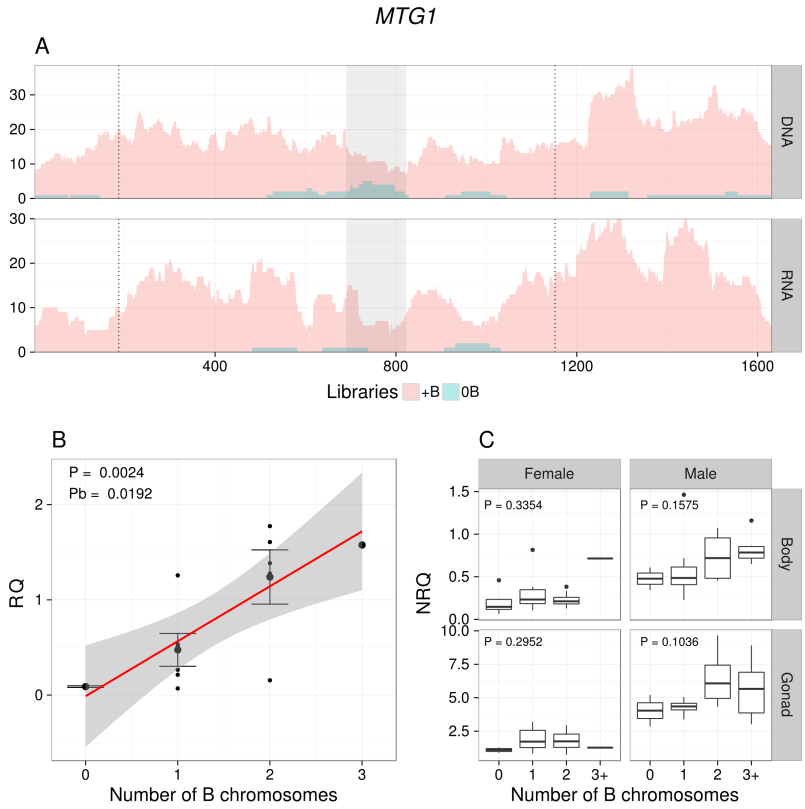

Figura S3: Coverage for the *MTG1* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Note that coverage was higher in the 4B gDNA library along all sequence length, and also in the B-carrying RNA library (A). The shaded zone in A marks the region amplified by qPCR. qPCR on gDNA (B) revealed that genomic copy number for the *MTG1* gene increases with B chromosome number following a dose-dependent pattern, thus confirming its presence in the B chromosome. qPCR on cDNA (C) showed that *MTG1* expression was independent on the number of B chromosomes in all samples, suggesting that this B chromosome gene is silenced. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

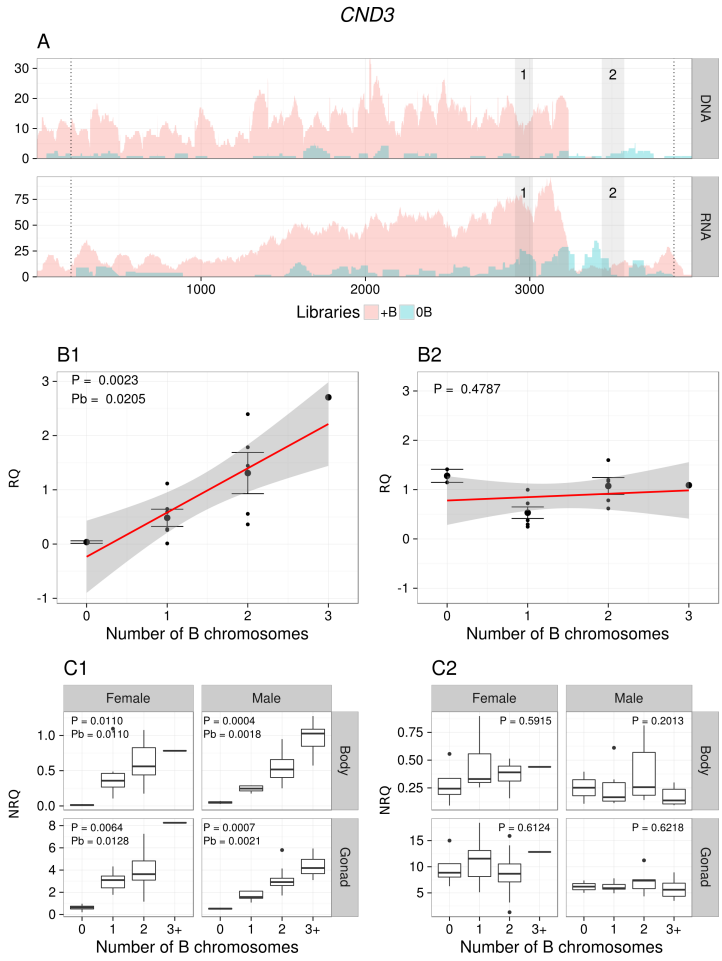

Figura S4: Coverage for the *CND3* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Low coverage in the 4B genome for the last five exons (20th-24th) suggests that the CDS of this B chromosome gene (delimited by dotted vertical lines) is truncated. Two regions were selected for qPCR amplification of this gene, one within the region being present in the B chromosome (shaded zone 1) and the other within the region being absent in it (shaded zone 2). qPCR on gDNA with zone 1 primers showed that copy number for this gene region increased with the number of B chromosomes (B1), whereas qPCR on gDNA with zone 2 primers was independent on B number (B2). Likewise, qPCR on cDNA showed that *CND3* expression increased in a dosage-dependent pattern for zone 1 primers (C1) but it was independent of B chromosome number for zone 2 primers (C2), suggesting the active transcription of B chromosome truncated gene copies. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

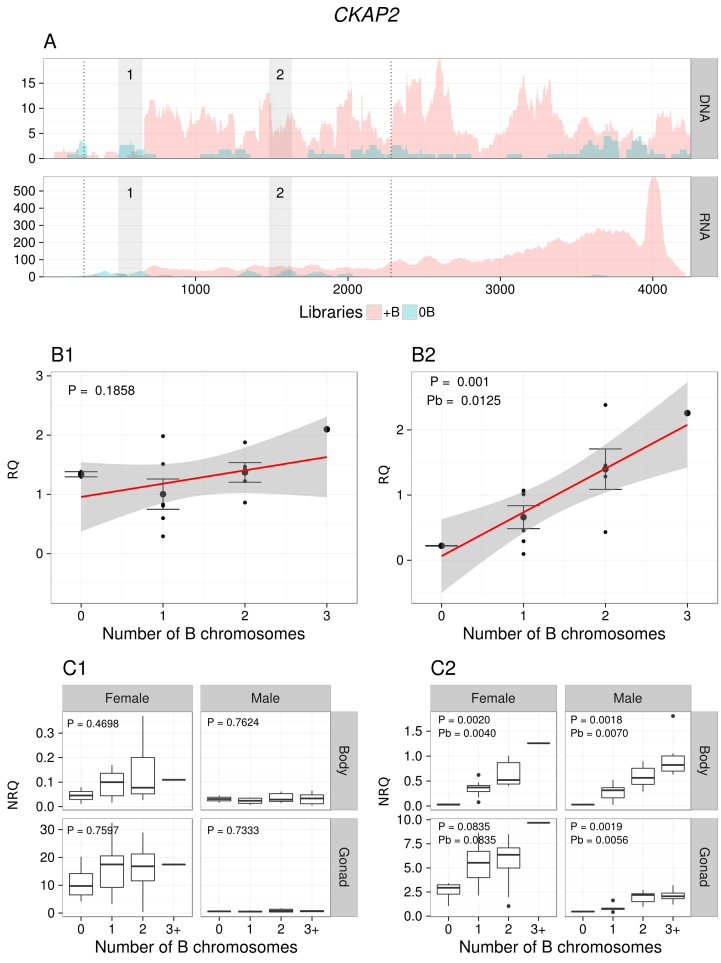

Figura S5: Coverage for the *CKAP2* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Low coverage in the 4B genome for the 5' region of the CDS (till nucleotide 663), suggests that this gene is truncated in the B chromosome. Two regions were selected for qPCR amplification of this gene, one within the region being missing in the B chromosome (shaded zone 1) and the other within the region being present in it (shaded zone 2). qPCR on gDNA with zone 1 primers showed that copy number for this gene was independent on the number of B chromosomes (B1), whereas qPCR on gDNA with zone 2 primers showed that it increased with the number of B chromosomes (B2), showing that this B chromosome gene is truncated. qPCR on cDNA showed that *CKAP2* expression was independent of B chromosome number for region 1 (C1) but increased in a dosage-dependent pattern for region 2 (C2), suggesting the active transcription of B chromosome truncated gene copies. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

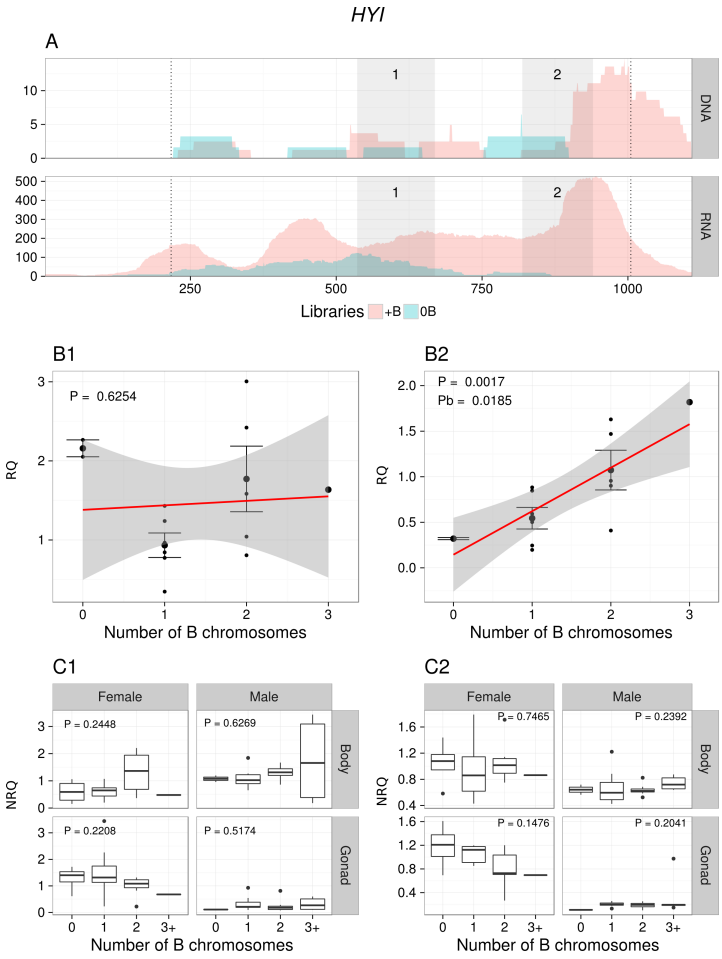

Figura S6: Coverage for the *HYI* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Low coverage in the 4B genome for the 5' region of the CDS (till nucleotide 901) suggests that this gene is truncated in the B chromosome. Two regions were selected for qPCR amplification of this gene, one within the region being missing in the B chromosome (shaded zone 1) and the other within the region being present in it (shaded zone 2). qPCR on gDNA with zone 1 primers showed that copy number for this region was independent on the number of B chromosomes (B1), whereas qPCR on gDNA with zone 2 primers showed that it increased with the number of B chromosomes (B2), thus showing that this B chromosome gene is truncated. qPCR on cDNA showed that *HYI* expression was independent of B chromosome number for both regions (C1 and C1), suggesting B chromosome truncated copies are inactive. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

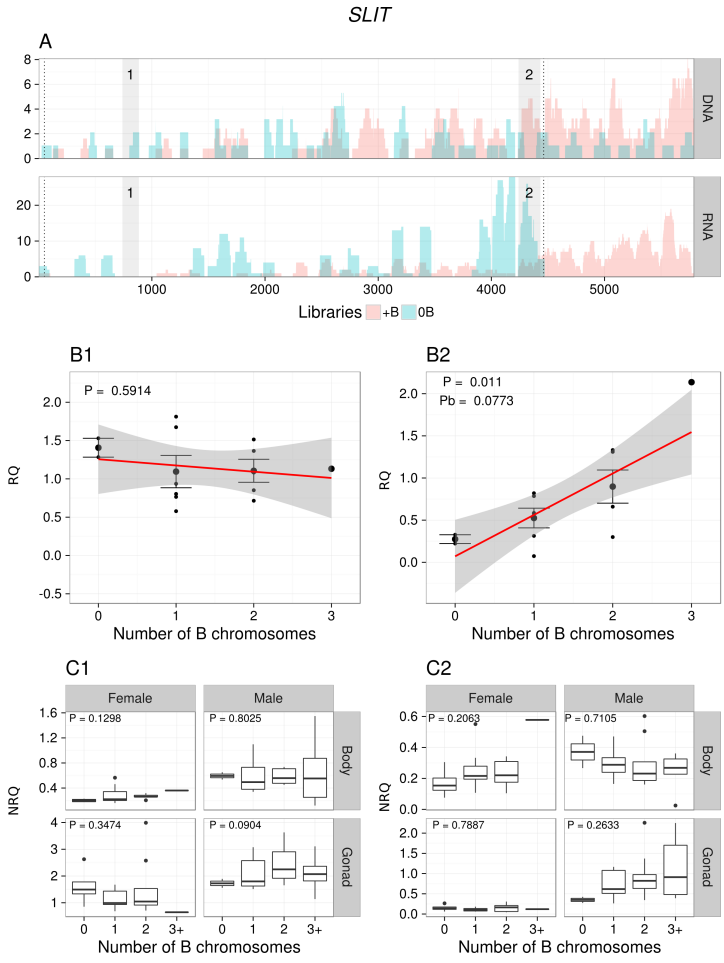

Figura S7: Coverage for the *SLIT* transcript in the gDNA (0B and 4B) and RNA (0B and 1B) Illumina reads (A), and qPCR on gDNA (B) and cDNA (C). Coverage in the 4B genome was high for only a small part of the 5' region of the CDS, suggesting that this gene is truncated in the B chromosome (A). Two regions were selected for qPCR amplification of this gene, one within the region being absent in the B chromosome (shaded zone 1) and the other within the region being present in it (shaded zone 2). qPCR on gDNA with zone 1 primers showed that copy number for this gene was independent on the number of B chromosomes (B1), whereas qPCR on gDNA with zone 2 primers showed that it increased with the number of B chromosomes (B2), thus demonstrating that B chromosome copies are truncated. qPCR on cDNA showed that *SLIT* expression was independent of B chromosome number for both regions (C1 and C1), suggesting the B chromosome truncated copies are inactive. RQ = Relative quantity. NRQ = Normalized relative quantity. P= P-value and Pb= Sequential Bonferroni P-value for Spearman rank correlation (B) and Kruskal-Wallis (C) analyses.

Cuadro S1: Spearman rank correlation (rS) between genomic abundance of nine genes and the number of B chromosomes, found in qPCR experiments. Pb = Sequential Bonferroni corrected P. N= Number of individuals.

Gene	N	r_S	P	Pb
CIP2A	14	0.7557	0.0018	**0.0177**
CKAP2_1	14	0.3755	0.1858	
CKAP2_2	14	0.7815	0.001	**0.0125**
CND3_1	14	0.744	0.0023	**0.0205**
CND3_2	14	0.2065	0.4787	
GTPB6	14	0.7698	0.0013	**0.0154**
HYI_1	14	0.1432	0.6254	
HYI_2	14	0.7581	0.0017	**0.0185**
KIF20A	14	0.6008	0.0231	0.1386
MTG1	14	0.7416	0.0024	**0.0192**
MYCB2_1	14	-0.1455	0.6196	
MYCB2_2	13	0.8253	0.0005	**0.0072**
SLIT_1	14	-0.1572	0.5914	
SLIT_2	14	0.6548	0.011	0.0773

Cuadro S2: Kruskal-Wallis analysis of qPCR gene expression experiments in males and females with 0-3 B chromosomes. Sex: M = Male, F = Female; Item: B = Body, G = Gonad; N= Number of individuals, df= degrees of freedom, Pb = Sequential Bonferroni P-value.

Gene	Sex	Item	N	H	df	P	Pb
CIP2A	F	B	18	10.80	3	0.0129	**0.0129**
		G	21	10.82	3	0.0127	**0.0255**
	M	B	23	12.20	3	0.0067	**0.0202**
		G	23	16.36	3	0.0010	**0.0038**
CKAP2_1	F	B	20	2.53	3	0.4698	
		G	21	1.17	3	0.7597	
	M	B	22	1.16	3	0.7624	
		G	22	1.28	3	0.7333	
CKAP2 _2	F	B	21	14.78	3	0.0020	**0.0040**
		G	21	6.66	3	0.0835	**0.0835**
	M	B	23	15.07	3	0.0018	**0.0070**
		G	23	14.94	3	0.0019	**0.0056**
CND3_1	F	B	21	11.15	3	0.0110	**0.0110**
		G	21	12.31	3	0.0064	**0.0128**
	M	B	23	17.98	3	0.0004	**0.0018**
		G	23	17.00	3	0.0007	**0.0021**
CND3_2	F	B	21	1.91	3	0.5915	
		G	21	1.81	3	0.6124	
	M	B	23	4.63	3	0.2013	
		G	23	1.77	3	0.6218	
GTPB6	F	B	21	1.30	3	0.73021	
		G	21	4.57	3	0.20629	
	M	B	22	1.92	3	0.58933	
		G	23	4.55	3	0.20766	
HYI_1	F	B	21	4.16	3	0.24479	
		G	21	4.41	3	0.22084	
	M	B	21	1.75	3	0.62688	
		G	23	2.27	3	0.51743	
HYI _2	F	B	21	1.23	3	0.74647	
		G	21	5.35	3	0.14764	
	M	B	23	4.21	3	0.23922	
		G	23	4.59	3	0.20412	

Continue in next page

Continued from previous page

Gene	Sex	Item	N	H	df	P	Pb
KIF20A	F	B	19	6.89	3	0.07551	0.15101
		G	20	1.94	3	0.585	0.585
	M	B	23	13.85	3	0.00311	**0.01244**
		G	23	10.61	3	0.01406	**0.04218**
MTG1	F	B	21	3.39	3	0.33536	
		G	21	3.70	3	0.29519	
	M	B	23	5.20	3	0.15749	
		G	23	6.17	3	0.10364	
MYCB2_1	F	B	21	4.92	3	0.17784	
		G	20	3.45	3	0.32764	
	M	B	23	0.60	3	0.89539	
		G	23	3.13	3	0.3717	
MYCB2_2	F	B	19	8.73	3	0.03306	**0.03306**
		G	21	10.90	3	0.01228	**0.03685**
	M	B	23	9.84	3	0.01995	**0.03989**
		G	23	12.71	3	0.00531	**0.02125**
SLIT_1	F	B	21	5.65	3	0.12979	
		G	21	3.30	3	0.34737	
	M	B	23	0.99	3	0.80247	
		G	23	6.48	3	0.09035	
SLIT_2	F	B	21	4.57	3	0.20629	
		G	21	1.05	3	0.78868	
	M	B	23	1.38	3	0.71046	
		G	23	3.98	3	0.26326	

4

A condensin subunit *CAP-G* pseudogene located in a B chromosome is actively transcribed

Beatriz Navarro-Domínguez, Francisco J. Ruíz-Ruano, Juan Pedro M. Camacho, Josefa Cabrero, María Dolores López-León.

Departamento de Genética, Universidad de Granada

Abstract Parasitic B chromosomes invade and persist in natural populations through several mechanisms for transmission advantage (drive). Contrary to the general belief that B chromosomes are genetically inert, recent work has provided evidence for the existence of gene-derived sequences in B chromosomes which, in some cases, are actively transcribed. A further interesting question is, however, whether B-derived transcripts actually become into functional products. In the grasshopper *Eyprepocnemis plorans*, one of the B-derived gene sequences shows homology with the gene coding for the *CAP-G* subunit of condensin I. We show here, by means of fluoresence *in situ* hybridization

coupled with tyramide signal amplification (FISH-TSA), that this gene is located in the distal region of the B24 chromosome variant. The DNA sequence located in the B chromosome lacks five exons of the 3' end and shows a non-synonymous substitution in a highly conserved region and a premature stop codon, indicating that it is a pseudogenic version of the *CAP-G* gene (*B-CAP-G*). In two Spanish populations, we found active transcription of *B-CAP-G*, but it did not influence the expression of *CAP-D2* and *CAP-D3* genes coding for the corresponding subunits of condensin I and II, respectively. Our results indicate that some of the B chromosome known effects may be related with its gene content and transcriptional activity, thus changing the current paradigm of B chromosome inertness and opening new exciting avenues for research.

Introduction

Chromatin condensation and chromosome segregation during cellular division are two essential events for maintenance and transmission of genetic information from cell to cell and through generations. Therefore, it is not surprising that these crucial processes involve many proteins providing structural support or regulation during mitosis and meiosis. Among these, condensin complexes play an important role for the timely compaction and segregation of chromosomes during cell divisions (Wood et al., 2010), which explains the growing interest in their study.

Condensins are highly conserved heteropentameric complexes,

constituted by a V-shaped dimer of SMC (structural maintenance of chromosomes) ATPase core subunits and an additional set of non-SMC regulatory CAP (chromosomal associated protein) subunits. Most eukaryotes contain two different types of condensin complexes, known as condensins I and II, which accomplish different functions and are subjected to different regulation during mitosis and meiosis. Condensins I and II are pentameric protein complexes sharing a same heterodimeric pair of SMC subunits, i.e. SMC2 and SMC4, and composed of a kleisin subunit (CAP-H in condensin I and CAP-H2 in condensin II) and a pair of HEAT-repeat subunits (CAP-D2 and CAP-G in Condensin I and CAP-D3 and CAP-G2 in Condensin II) (Hirano, 2005). HEAT repeats are bihelical structures organized in tandem which can enfold target substrates and act as a scaffold for the assembly of other molecular components (Neuwald and Hirano, 2000). They have been found, among others, in the four proteins included in the HEAT acronymous: huntingtin, elongation factor 3, the A subunit of the protein phosphatase 2A and TOR1, a target of rapamycin that seems to be essential for progression of the G1 phase of cell cycle (Andrade and Bork, 1995).

In fungi, condensin II has been lost during evolution, and one of the components of condensin II (CAP-G2) has not been found in the *Drosophila melanogaster* genome (Hirano, 2012). Therefore, in this latter species, it is not clear the division of labor between condesin I and II, and the existence of the condensin II complex has been even questioned (Herzog et al., 2013). However, in *Caenorhabditis elegans*, condensin I has a minor contribution compared with condensin II, and there is a third condensin I-like complex (condensin I^{DC}), with a modificated SMC4 subunit, which is involved in dosage compensation (Csankovszki

et al., 2009; Kruesi et al., 2013). Many bacterial and archaeal species have primitive types of condensin (Hirano, 2016).

The most conspicuous phenotype of condensin mutants is the massive formation of bridges between chromosomes during mitotic and meiotic anaphases, as a consequence of defective resolution of concatenations between sister chromatids and incorrect compaction of chromosomes prior to mitotic anaphase separation (Wood et al., 2010). This last condition leads to anaphase bridging and lagging chromatids. Anaphase bridges are associated with DNA double strand breaks (McClintock, 1941), which often result in chromosome rearrangements such as translocations or deletions (Acilan et al., 2007), and also gene amplification through the breakage-fusion-bridge (BFB) cycle model (McClintock, 1941; Shimizu et al., 2005). The absence of condensin also entails fuzzy chromosome appearance (see Hirano, 2012 for review). Furthermore, condensins have a role in multiple other cellular processes, as suggested by the fact that, in *C. elegans*, mutants for any condensin I subunit show changes in DNA double-strand break distribution and frequency, causing a boost in the frequency of crossover events (Mets and Meyer, 2009). In *D. melanogaster*, condensin II limits transvection and is required for the dissasembly of polytene chromosomes, which occurs at interphase during ovarian development (Hartl et al., 2008), and mutations on several subunits of condensin alter position-effect variegation in X and Y chromosomes in regions next to ribosomal DNA (rDNA) (Cobbe et al., 2006). There are many evidences for a functional relationship beetween condensins and rDNA, since the condensin complex associates with rDNA loci (Freeman et al., 2000; Wang et al., 2004; Kim et al., 2013), and even rDNA transcription regulates condensin function

in *S. cerevisiae* (Wang et al., 2006; Wang and Strunnikov, 2008). Also, several components of condensin have been found to locate in interphase nucleoli of Xenopus (Cabello et al., 2001; Uzbekov et al., 2003) and also regulating rDNA silencing in *S. cerevisiae*, thus maintaining silent chromatin domains (Machín et al., 2004).

Effects of mutations on the *CAP-G* subunit gene in Drosophila have been profusely described in the literature, thus providing many insights on its role in the overall function of the condensin complex. Some of these effects are embryonic lethality, female infertility, delays in chromosome condensation during prophase, failures in sister chromatid separation and resolution, lagging chromosomes, bridging, and ultimately aneuploidy. *CAP-G* also has a role in gene expression regulation during interphase, possibly associated with the suppression of position-effect variegation (Dej et al., 2004). Furthermore, the CAP-G subunit has been shown to interact not only genetically but physically with *CID*, the Drosophila homolog for *CENP-A*, evidencing a link between condensin and kinetochore structure (Jäger et al., 2005).

Many eukaryote genomes harbor special, additional and dispensable chromosomes called supernumerary (B) chromosomes (for review, see (Jones and Rees, 1982; Camacho et al., 2000; Camacho, 2005). They are considered intragenomic parasites taking advantage of cell replication machinery for its accumulation and persistence in natural populations, as they do not obey Mendelian segregation law by managing to incorporate into gametes at rates higher than 0.5. B chromosomes can originate from the same genome harboring them, or else they can derive from a different species through hybridization. They do not recombine with standard (A) chromosomes, for which reason it is said that they follow their own

evolutionary pathway. Most B chromosomes are heterochromatic and mainly composed of non-coding repetitive DNA. These features, along with their dispensable nature, have led to the long-lasting view that B chromosomes are gene lacking and transcriptionally inert chromosomes. Nevertheless, in the last years, protein-coding genes have been found in B chromosomes from several animal and plant systems (Graphodatsky et al., 2005; Teruel et al., 2010; Fantinatti et al., 2011; Martis et al., 2012; Valente et al., 2014; Huang et al., 2016; Navarro-Domínguez et al., submitted; see chapter 3; Ruiz-Ruano et al. in preparation). In the grasshopper *Eyprepocnemis plorans*, the B chromosome variant named B24 harbors ribosomal RNA genes expresed at low rate (Cabrero et al., 1987; Ruiz-Estévez et al., 2012, 2014), and at least four protein-coding genes showing their complete coding sequence (CDS), and five other genes with fragmented CDS (Navarro-Domínguez et al., submitted; see chapter 3). One of these fragmented genes shows homology with the *CAP-G* subunit of condensin I (also known as condensin complex subunit 3 or CND3). In contrast to the *CAP-G* sequence described in the *Locusta migratoria* draft genome (Wang et al., 2014), the *CAP-G* sequence found in the B24 chromosome (from here onwards, *B-CAP-G*) of *E. plorans* lacks the last five exons (20-24) of the 3' end. Remarkably, the *B-CAP-G* gene is actively transcribed in B24-carrying adults males and females from the Torrox population (see Navarro-Domínguez et al., submitted; see chapter 3).

As intragenome parasites, B chromosomes could take advantage of alterations in the correct functioning of cell division allowing them bypassing the mitotic and/or meiotic checkpoints thus facilitating B chromosome drive. Several works have suggested the possible implication

of condensins in cellular responses to chromosome damage (Aono et al., 2002) and in the spindle assembly checkpoint (SAC) (Yong-Gonzalez et al., 2007; Murillo-Pineda et al., 2014; Xu et al., 2015). In fact, contrary to previous beliefs claiming that condensin I access to nucleus occurs after nuclear envelop breakdown, it is now known that a small fraction of condensin I remains in cell nucleus during interphase acting on gene regulation, and their mutants fail in activating cellular control pathways (Aono et al., 2002; Uhlmann, 2002; Zhang et al., 2016).

Due to the high relevance of the condensin complex, and its possible impact on B chromosome maintenance in natural populations, we analyze here the changes in DNA sequence shown by the *B-CAP-G* pseudogene found in B chromosomes from two populations of the grasshopper *E. plorans*, its transcriptional activity and its possible influence on the activity of other condensin complex protein subunit genes (*CAP-D2* and *CAP-D3*), as an indirect indication for its possible functional role.

Materials and methods

Experimental materials and karyotypic characterization

This study was carried out in 80 adult individuals of the grasshopper *Eyprepocnemis plorans*, collected in Salobreña (Granada, Spain) and Torrox (Málaga, Spain) in October 2013, and showing different numbers of B chromosomes (Table 1). Embryos were obtained from egg pods

Table 1: Individuals of *Eyprepocnemis plorans* analyzed per population, sex and number of B chromosomes. Bs = number of B chromosomes. N = Number of individuals.

Population	Sex	Bs	N
Salobreña	Female	0	2
		1	5
		2	2
		+2	5
		Total	*14*
	Male	0	4
		1	10
		2	4
		+2	3
		Total	*21*
Torrox	Female	0	4
		1	9
		2	8
		+2	1
		Total	*22*
	Male	0	2
		1	7
		2	8
		+2	6
		Total	*23*
Total Samples			80

dissected in insect saline solution after ten days of incubation at 28^{o}C, which were used for *CAP-G* physical mapping.

Testes and ovaria were dissected out from anesthetized animals. One gonad (testis or ovary) and the somatic bodies were immediately frozen in liquid nitrogen and stored at 80^{o}C until DNA and RNA extraction. The other testis was fixed in 3:1 ethanol-acetic acid and stored at 4^{o}C for cytological analysis. The remaining ovary was immersed in

2% colchicine in isotonic insect saline solution for 2 hours, fixed in 3:1 ethanol-acetic acid, and stored at 4^{o}C for cytological analysis. In males, the number of B chromosomes was determined by visualizing them in primary spermatocytes at diplotene or metaphase I obtained by squashing two testis tubules in a drop of 2% lactopropionic orcein (Camacho et al., 2015). In females, the number of B-chromosomes was analysed in squash preparations of two ovarioles submitted to C-banding, a technique which shows B chromosomes much darker than the A chromosomes (Camacho et al., 2015). For physical mapping of the *CAP-G* gene, we used 0.05% colchine-treated embryos prepared as described in Camacho et al. (2015). In embryos, B chromosome presence was determined by staining chromosome slides with 2µg/µl 4',6 diamidino-2-phenylindole (DAPI) which reveals the presence of several large DAPI^{+} bands on the B chromosome.

Chromosomal location of *B-CAP-G* using Tyramide-coupled FISH

Physical mapping of the *B-CAP-G* pseudogene was performed on embryo chromosomes by means of fluorescent in situ hybridization coupled with tyramide signal amplification (FISH-TSA), following the protocols described previously by Krylov et al. (2007, 2008) and Funkhouser-Jones et al. (2015), with minor changes.

To prepare the probe for FISH, a 110 bp fragment was amplified on *E. plorans* gDNA with primers anchoring on exon 18 of the *CAP-G* sequence (Navarro-Domínguez et al., submitted; see chapter 3). PCR

reaction was performed with the Horse-PowerTaq DNA polymerase (Canvax) kit, and contained 1X PCR buffer, 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 0.4 µM of each primer, 2.5 units of DNA polymerase and about 2.5 µl/µl of DNA per reaction. The thermocycler program was as follows: an initial denaturation step at 94^{o}C for 5 min, 40 cycles at 94^{o}C (15 s), 60^{o}C (15 s) and 72^{o}C (15 s), and a final elongation step at 72^{o}C for 5 min.

200 ng probe was labeled with digoxigenin-11-dUTP (Roche) by random primers using Decalabel DNA labeling kit (Thermo Scientific) according to manufacturer's recommendation but using 9 µl of a labeled dNTP mix containing 1mM dATP, 1mM dCTP, 1mM dGTP, 0,65mM dTTP and 0.01mM digoxigenin -11-dUTP. Probe was labelled at 37^{o}C for 20 hours and then purified and resuspended in 10 µl ultrapure water.

Chromosome preparations were dehydrated in 70%, 90% and 100% ethanol series for 3, 3 and 5 min, respectively. Afterwards, they were air dried and stored at 60^{o}C overnight. After incubation with 200 µl RNase (100 µg/ml in 2XSSC, 300mM NaCl, 30 mM sodium citrate, pH 7.0) for 2 hours at 37^{o}C, they were washed three times in 2XSSC for 5 min and once in PBS for 5 min. Metaphase spreads were then treated with 2% paraformaldehyde at room temperature followed by three washes in PBS (phosphate buffer saline, pH=7.3) for 5 min each. Endogenous peroxidases were quenched by treating slides for 30 min with 1% H_2O_2, followed by three washes in PBS for 5 min and one in distilled water. Chromosome preparations were then dehydrated in a series of 70%, 90% and 100% ethanol for 3, 3 and 5 min, respectively. Chromosomal DNA was denatured and hybridized with 50 µl of hybridization mixture, composed of 50%

deionized formamide, 2XSSC and 40 ng of labeled probe, under a plastic coverslip in a hot plate at 80^{o}C for 6 min. Slides were then incubated in a humid chamber overnight at 37^{o}C. Posthybridization washing and signal detection of FISH-TSA were performed as reported by Krylov et al. (2007, 2008). For detection of digoxigenin-labeled *CAP-G* gene probe, we used an anti-digoxigenin antibody conjugated with horse-radish peroxidase (Sigma) at a concentration of 1:500 in TNB (Tris-NaCl-blocking buffer), by incubating slides for 1 hour in a humid chamber at room temperature. Tyramide signal amplification reaction was performed using TSA Plus Fluorescence kit (Perkin Elmer), applying 100 µl of 1:50 diluted tyramide solution per slide with dark-incubation for 10 min at room temperature. Slides were counterstained with 2 µg/ml DAPI for 18 min, washed for 5 min in PBS, 3 min in distilled water and mounted in antifading solution (Vectashield H-100). Hybridized metaphase cells were analyzed under a BX41 Olympus epifluorescence microscope and photographs were taken with a DP70 cooled camera. Images were merged and optimized for brightness and contrast with the Gimp software.

Prediction of *B-CAP-G* sequence functionality

To analyze sequence variation in *B-CAP-G*, pseudogene in respect to the *CAP-G* gene located in the A genome, we used SSAHA2 (Ning et al., 2001) to map Illumina reads obtained from 0B and 1B *E. plorans* female RNA, and those obtained from male genomic DNA with 0B and 4B from the Torrox population (NCBI-SRA accession numbers SRR2969416, SRR2969417, SRR2970625 and SRR2970627; Navarro-Domínguez et al.,

in preparation; see chapters 3 and 5). To search for B-specific sequence variants, we quantified the number of reads carrying a given nucleotide at each position using pysamstats[1], and selected the positions where the variants were shared by the two +B libraries but were absent in the two 0B libraries. For further confirmation, we also mapped reads obtained from three different 0B females, three 0B males and three 4B males, also from the Torrox population (Martín-Peciña et al., personal communication).

Although *B-CAP-G* lacks five exons on the 3' end (Navarro-Domínguez et al., submitted see chapter 3), it has been reported that condensin can be functional even with a C-terminal deletion (Herzog et al., 2013), for which reason we compared the predicted aminoacid sequence for the *B-CAP-G* transcript with with homologous proteins available in the databases. We got 18 sequences of CAP-G condensin I subunit in the NCBI-GenBank database from many different organisms (seven mammals, a bird, a frog, a fish and eight insects; see Table 2 for accession numbers). To compare with other grasshopper species, we used the *CAP-G* transcript from a *de novo* assembled transcriptome of *Locusta migratoria* (Ruiz-Ruano et al. in preparation) and from a Trinity-based de novo assembly performed by us with reads of *Chortippus mollis* found in the databases (NCBI-SRA accession number SRR2051368; Berdan et al., 2015), and predicted the protein coded by those transcripts. Alignments were performed using Geneious 4.8 (Drummond et al., 2009). The possible impact of aminoacidic changes and C-terminus deletion was predicted with PROVEAN (Choi and Chan, 2015).

[1]https://github.com/alimanfoo/pysamstats

Table 2: NCBI-GenBank accession numbers of CAP-G protein sequences used for prediction of the B-CAP-G sequence functionality

Organism	Acc. No
Homo sapiens	NP_071741
Pan troglodytes	XP_526535
Macaca mulatta	XP_001102882
Bos taurus	NP_001095846
Canis lupus familiaris	XP_536278
Rattus norvergicus	XP_223468
Mus musculus	NP_062311
Gallus gallus	XP_420769
Xenopus tropicalis	NP_989190
Danio rerio	NP_001243134
Drosophila melanogaster	NP_001163135
Anopheles gambiae	XP_564680
Acromyrmex echinatior	XP_011050712
Solenopsis invicta	XP_011166263
Megachile rotundata	XP_003705431
Bombus impatients	XP_003488800
Apis florea	XP_003696218

DNA extractions

Genomic DNA (gDNA) from Salobreña males was extracted using GenElute Mammalian Genomic DNA Miniprep kit (Sigma). Absence of degradation was checked in a 1% TBE-agarose gel, and quantification and assessment of 260/280 and 260/230 ratios was performed with an Infinite M200 Pro NanoQuant (Tecan). For qPCR working solution, gDNA samples were diluted to 5ng/µl.

RNA extractions and complementary DNA (cDNA) synthesis

Total RNA extractions were performed using Real Total RNA Spin Plus kit (Durviz) for somatic bodies and RNeasy Lipid Tissue Mini Kit (Qiagen) for gonads, complementing both protocols with a DNase treatment on the column membrane (20 units of Sigma DNAse Amplification Grade DNase I for 30 minutes). gDNA contamination on gonad extracted RNA was negligible or nonexistent, but the RNA extracted from bodies needed an additional DNase treatment with the REALSTAR kit (Durviz). Quality check and quantification of total RNA was performed with a Tecan's Infinite 200 NanoQuant and in a denaturing MOPS-agarose gel to assure the absence of degradation and DNA contamination, which was further corroborated by lack of amplification of ribosomal DNA (rDNA) and histone genes in the extractions. Retrotranscription was performed on 100 *ng* total RNA combined with random and oligo-dT hexamers from the PrimeScriptTM RT reagent -Perfect Real Time- Kit (Takara), and 1:10 diluted to get the working solution.

Quantitative PCR (qPCR)

In order to test whether *B-CAP-G* is located and fragmented in the B2 chromosome from the Salobreña population, we performed relative quantification of the abundance of the *CAP-G* gene by means of qPCR analysis, using two primer pairs, one anchored on exon 18, which is expected to amplify on gDNA from both A and B chromosomes, and

the other anchored on exon 22, which is expected to amplify only on A chromosomes. This is based on previous observations in the Torrox population indicating that *B-CAP-G* lacks the five last 3' exons (20 to 24) (Navarro-Dominguez et al., see chapter 3). We also PCR amplified these regions on cDNA from gonads and somatic bodies of males and females with different numbers of B2 chromosomes from the Salobreña population, in order to ascertain whether *B-CAP-G* is transcribed in this population. Primers for exon 18 and exon 22 of *E. plorans CAP-G* gene were the same described in Navarro-Domínguez et al., submitted; see chapter 3).

In addition, we analyzed in both populations (Torrox and Salobreña) whether the expression of the *B-CAP-G* pseudogene in B-carrying individuals influences the transcription of genes for other subunits of the condensin complex, specifically *CAP-D2* (Condensin I) and *CAP-D3* (Condensin II). To search for these gene sequences in the *E. plorans* assembled transcriptome (available in Figshare, Navarro-Domínguez, B. (2016)), we retrieved CAP-D2 and CAP-D3 protein sequences described for *Zootermopsis nevadensis* from NCBI-GenBank (accession numbers KDR15738.1 and KDR16504.1) and then performed local TBLASTN (Altschul et al., 1990) on the *de novo* assembled transcriptome of *E. plorans* (Navarro-Domínguez et al., in preparation; see chapter 5). *E. plorans* transcriptome sequences were also blasted with the BLASTX algorithm over NCBI-NR database, to further confirm homology with CAP-D2 and CAP-D3 proteins. Sequences of *E. plorans* CAP-D2 and CAP-D3 transcripts can be found in NCBI-GenBank under accesion numbers KX376471 and KX376472, respectively.

Primer sequences for *E. plorans CAP-D2* and *CAP-D3* were designed

with Primer3 software (Koressaar and Remm, 2007; Untergasser et al., 2012). They were F: 5'-GGTGACATTGCTTTCCGATT-3' and R: 5'-CTCCGGATGTGCATCTGTTA-3' for *CAP-D2*, and F: 5'-CCCAGAAAGAAGCTGAGGTG-3' and R: 5'-TCAAAACATGCCTACCAGCA-3' for *CAP-D3*.

Selection of reference genes

Primer sequences for the amplification of reference genes were those described in Van Hiel et al. (2009) and Chapuis et al. (2011). Amplification, validation and stability of *Actin 5C* (*Act*), *Armadillo* (*Arm*), *Glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*), *Ribosomal protein 49* (*RP49*) and $\alpha-$*tubulin 1A* (*Tub*) in somatic body and gonad of males and females of *E. plorans* from Torrox and Salobreña populations, were formerly performed by Navarro-Domínguez et al. (in press; see chapter 2). Two or three reference genes were employed here, specifically *Act-RP49* for bodies and ovaries from Salobreña females, *GAPDH-Tub* for bodies and *Act-Tub-Arm* for testes from Salobreña males, *RP49-Act-Arm* for bodies and *Act-Arm* for ovaries from Torrox females, and *Act-Tub* for bodies and *Act-RP49* for testes from Torrox males.

qPCR reaction

qPCR amplification was performed in a Chromo4 Real Time PCR thermocycler (BioRad). Reaction mixture contained 5µl gDNA or cDNA,

5µl SensiMix™SYBR Kit and 2.5µl each 2.5µM primer in a total 15µl volume. Electronic pipettes (Eppendorf Research®Pro) were used in order to minimize pipetting errors, and each reaction was carried out in duplicate. Samples were discarded when the coefficient of variation between technical replicates was higher than 5%. Quantitative PCR protocol consisted in an initial denaturation step for 10 min at 95^{o}C, followed by 40 cycles of 15 s at 94^{o}C, 15 s at annealing temperature (60^{o}C for *CAP-G* and 58^{o}C for *CAP-D2* and *CAP-D3*) and 15 s at 72^{o}C, with plate reading at the end of every cycle. Specificity of reaction was assessed for each primer pair by means of a dissociation curve (from 72^{o}C to 95^{o}C with plate reading every 1^{o}C) after the 40th cycle. Fluorescence was measured and processed using Opticon Monitor 3.1 (Bio-Rad Laboratories, Inc). Negative controls for each primer pair were included in all reactions.

Data analysis

Efficiency for all primer pairs was calculated by a standard curve performed with 1:10 serial dilutions. Relative Quantities (RQs) were calculated referred to a calibrator sample, included as an inter-plate variation control, consisting in a mixture of DNAs or cDNAs from several individuals. RQs were calculated following the Pfaffl (2001) method. In cDNA reactions, RQ values were normalized by the geometrical average of the most stable reference genes for each sample type (Vandesompele et al., 2002).

In order to meet the normality requirements of parametric analyses,

we transformed the qPCR data to natural logarithms. We tested the linear relationship between B chromosome number and abundance of the *CAP-G* gene (with primer pairs anchored on exons 18 and 22), in *E. plorans* males from Salobreña, using Pearson's correlation analysis. *CAP-G* expression level in bodies and gonads of males and females with different numbers of B chromosomes, was tested by one-way ANOVA and post-hoc sequential Bonferroni correction. To test whether *CAP-G* transcription is associated with *CAP-D2* and *CAP-D3* transcription, we performed an analysis of covariance (ANCOVA) including population, sex, body part and B chromosome presence as discrete independent variables, *CAP-G18* (measured in exon 18) and *CAP-G22* (measured in exon 22) as continuous independent variables, and *CAP-D2* or *CAP-D3* as dependent variables. We finally calculated partial correlations of *CAP-G18* and *CAP-G22* with *CAP-D2* or *CAP-D3* by means of multiple regression analysis.

Results

Tyramide-coupled FISH reveals *B-CAP-G* pseudogene localization in the distal region of the B24 chromosome

The FISH-TSA experiments showed that the B24 chromosome in the Torrox population carries *B-CAP-G* pseudogenes in a subdistal location, specifically in the inner part of the rDNA distal region (Fig. 1). No consistent signal was found on A chromosomes. This is an

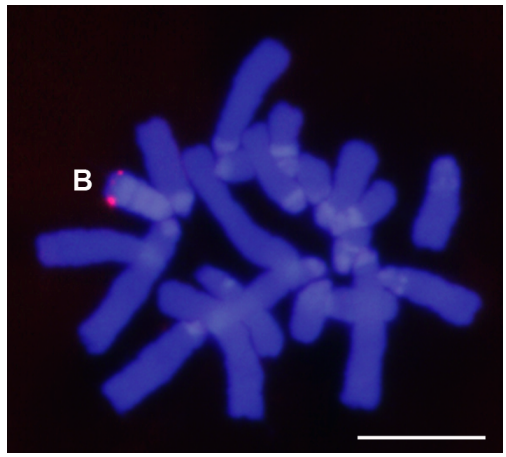

Figure 1: FISH-TSA showing the physical location of the B-CAP-G pseudogene in the distal region of the B24 chromosome. The absence of signal on A chromosomes suggests a tandem organization of several copies of this pseudogene in the B chromosome. Bar=10µM.

indirect evidence that the B chromosome contains several copies of the pseudogene, for which reason it was visualized on the B but not on the A chromosomes.

The *B-CAP-G* pseudogene shows specific sequence changes

The CDS of the complete *CAP-G* transcript (discarding the UTR regions) found in B-lacking *E. plorans* individuals was 3,669 bp long, thus showing the same length as that found in *L. migratoria* (3,669 bp), and it was 15 nucleotides shorter than that found in *Ch. mollis* (3,684 bp). *E. plorans* standard *CAP-G* gene is presumably structured in 24 exons, likewise the gene found in the *L. migratoria* genome (Wang et al., 2014). The *B-CAP-G* pseudogene, however, lacks the five last 3' exons (20-24), as shown by coverage pattern in the genomic reads from a 4B male

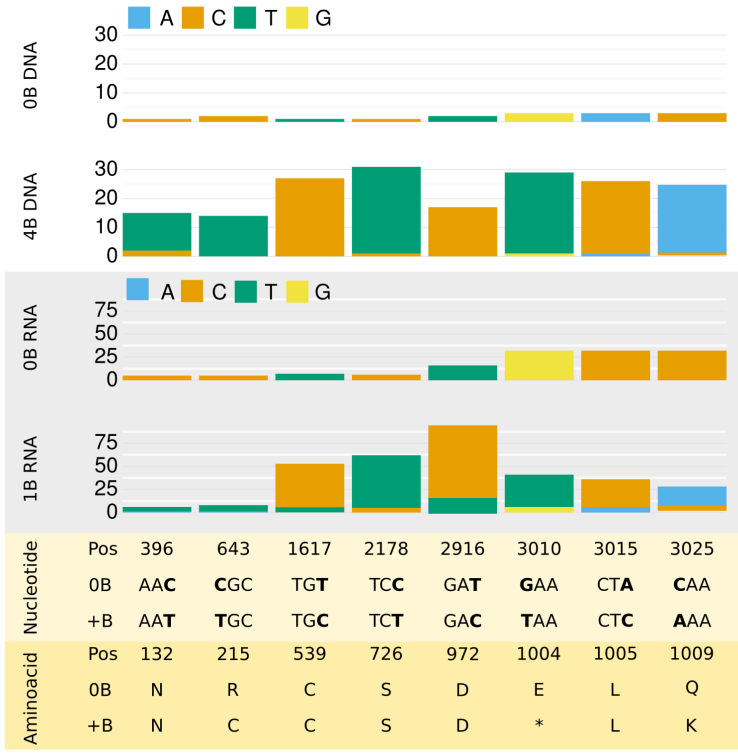

Figure 2: Nucleotidic variation in the *CAP-G* gene associated with B chromosome presence in *E. plorans* from Torrox population. The upper panel (white background) shows gDNA counts for 8 nucleotidic positions in 0B and 4B males from Torrox. Note that the 0B reference is usually at lower frequency because of the presence of several gene copies in the B chromosome. The middle panel (grey background) shows RNA counts for these same 8 nucleotidic positions in 0B and 1B females from Torrox. Note that the 0B female has the same nucleotides as the 0B gDNA from Torrox, and that the 1B female has esentially the same nucleotide composition as the B-carrying gDNA libraries, indicating the expression of B chromosome gene copies. The lower panel shows codon (light yellow background) and aminoacidic (yellow background) changes provoked by the 8 substitutions. Note that only nucleotidic changes in positions 643 and 3010 provoke alterations on the predicted CDS of the B chromosome gene copies (R for C in the 215 residue and E for stop codon in the 1004 residue, respectively), and that the substitution of nucleotide 3025 is beyond the stop codon.

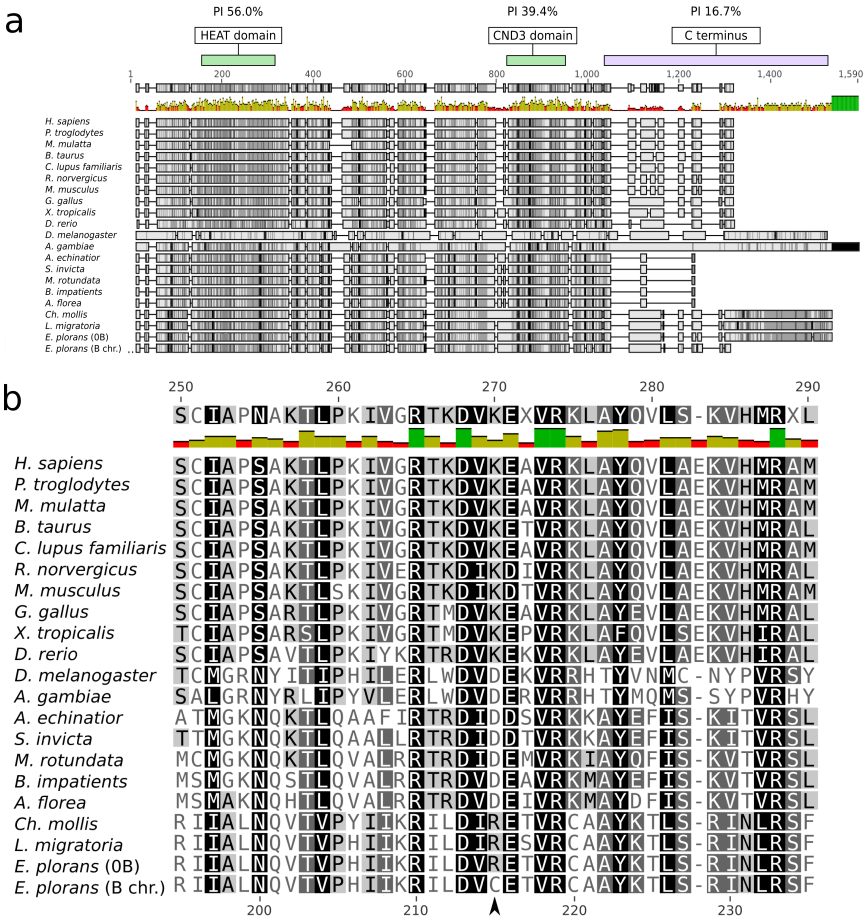

Figure 3: Protein alignment of the complete sequence (a) and the conserved HEAT domain (b) of CAP-G from several organisms. a) Note that the predicted protein generated if the B chromosome transcript were translated, marked at the bottom of the alignment, would include the two conserved domains (HEAT and CND3) but lack the less conserved C-terminus region, the latter being absent in some species. b) Alignment of the HEAT domain region. The arrowhead points aminoacid 215 which is R in *E. plorans* 0B, likewise in the grasshoppers *L. migratoria* and *Ch. mollis*, but C in the *E. plorans* B chromosome sequence.

(Navarro-Domínguez et al., submitted; see chapter 3). Therefore, the *B-CAP-G* pseudogene is about 18% shorter than the *CAP-G* gene in A chromosomes. In addition, B-carrying individuals from this same population showed that *B-CAP-G* transcript amount increased with B chromosome number when probed for exon 18, but not for exon 22, thus indicating active transcription of the *B-CAP-G* pseudogene (Navarro-Domínguez et al., submitted; see chapter 3). Sequence analysis showed the presence of eight nucleotide changes being exclusive of B-carrying libraries from Torrox (Fig. 2). Five of them were synonymous and thus lacked impact on the predicted protein. However, one of the three remaining substitutions was a nonsynonymous transition of C to T in position 643 of the CDS (exon 5), changing a CGC codon (Arg) for TGC (Cys). Another substitution was a transversion from G to T at position 3010 (exon 19 i.e. the last codon before gene truncation on the B chromosome), changing GAA (Glu codon) to TAA (stop codon), thus shortening the predicted protein in 199 aminoacids from the C-terminus. Finally, we found another nonsynonymous substitution in position 3025 (CAA for AAA, changing Gln for Lys), but it is located beyond the premature stop codon thus probably lacking impact on the predicted protein. Therefore, the predicted protein for the B24 transcript is 1,004 aminoacids long, instead of the 1,223 of the A chromosome transcript.

It was remarkable that, in spite of the premature stop codon leading to transcript truncation, the predicted *B-CAP-G* protein for the B24 gene copies would include all conserved regions described for this protein, i.e. the HEAT domain, from exon 3 to exon 5, and the CND3 domain, from exon 11 to exon 17. An alignment of CAP-G sequences from several organisms showed that the C-terminus is the least conserved

region and is even absent in most species, with the exception of the three grasshoppers (*Ch. mollis*, *L. migratoria* and *E. plorans*) and the two diptera (*D. melanogaster* and *A. gambiae*) species (Fig. 3a). In all cases, the HEAT domain was highly conserved (Fig. 3b). Evaluation of the possible effect of the two putatively impairing substitutions being specific to B-carrying individuals, i.e. the Arg-Cys substitution and the stop codon (see above), with the PROVEAN software scored -3.1 (i.e. under the -2.5 threshold) for the substitution and -10.1 for the stop codon. Therefore, both changes are probably deleterious for the canonical *CAP-G* function, reinforcing the pseudogenical character of the *B-CAP-G* sequence.

The *B-CAP-G* pseudogene is fragmented and transcribed also in the B2 chromosome

qPCR analysis of *CAP-G* abundance on gDNA from Salobreña males showed that it increased linearly with B chromosome number when we used primers anchored on exon 18 (Fig. 4a) but not when primers were anchored on exon 22 (Fig. 4b). This suggests the presence of truncated versions of the *CAP-G* gene (*B-CAP-G* pseudogene) in the B2 chromosome.

Similar experiments on cDNA on males and females from the same population showed, in both sexes and body parts analyzed, significant positive association between *CAP-G* transcript abundance and B chromosome number for exon 18 primers (Fig. 5a) but not for exon 22 ones (Fig. 5b). This pattern suggests that the excess of

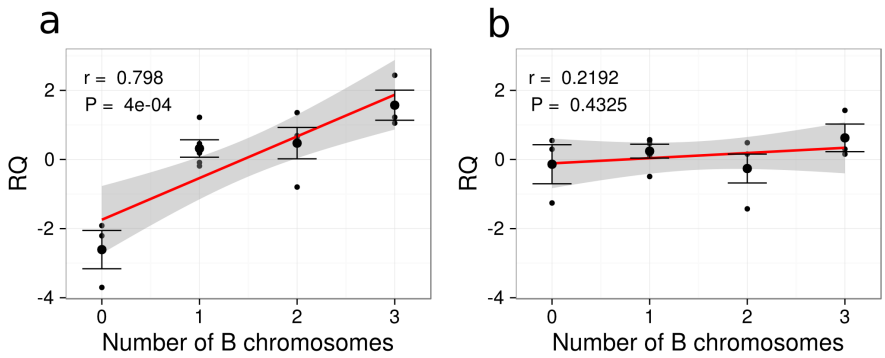

Figure 4: Relative quantification of CAP-G gene abundance by means of qPCR analysis on gDNA from males with 0-3 B2 chromosomes collected at the Salobreña population, using primers anchored on exon 18 (a) and exon 22 (b). Note that CAP-G abundance increases linearly with the number of B chromosomes when assayed for exon 18 (a), indicating that this exon is present in the B chromosome gene copies, but not for exon 22 (b), suggesting the absence of this exon in the B chromosome copies. Taken together, both experiments reveal the presence of truncated *CAP-G* gene copies in the B2 chromosome. RQ = relative quantity, r = Pearson's linear correlation coefficient, P = P-value for Pearson's correlation.

CAP-G transcripts observed in B-carrying individuals is due to active transcription of the *B-CAP-G* pseudogene.

Transcription of *CAP-D2* and *CAP-D3* is not altered by B chromosome presence or *B-CAP-G* transcription

ANCOVA revealed that *CAP-D2* and *CAP-D3* transcription levels show significant differences between populations and sexes, but they are not influenced by the presence of B chromosomes (Tables 3 and 4). In addition, it showed that transcription levels of both *CAP-D2* and *CAP-D3* are significantly associated with CAP-G expression when measured at exon 22 (which is present only in A chromosomes) but

Table 3: ANCOVA for *CAP-D2* transcription level (dependent variable), *CAP-G* expression at exons 18 and 22 (continuous independent variables), and population (pop), sex, body part and B chromosome presence (Bpre) (discrete independent variables). SS= sum of squares, df = degrees of freedom, MS = mean sum of squares, p = p-value. Significant effects are noted in bold-type letter.

Item	SS	df	MS	F	p
Intercept	36.69109	1	36.69109	198.5950	**0.000000**
CAP-G18	0.52497	1	0.52497	2.8415	0.094152
CAP-G22	15.64933	1	15.64933	84.7039	**5.55E-16**
pop	3.63684	1	3.63684	19.6848	**1.87E-05**
sex	15.10622	1	15.10622	81.7642	**1.33E-15**
bodypart	0.01699	1	0.01699	0.0920	0.762150
Bpre	0.34359	1	0.34359	1.8597	0.174911
pop*sex	1.48797	1	1.48797	8.0538	**0.005237**
pop*bodypart	0.64508	1	0.64508	3.4916	0.063833
sex*bodypart	0.18564	1	0.18564	1.0048	0.317938
pop*Bpre	0.01003	1	0.01003	0.0543	0.816112
sex*Bpre	0.00329	1	0.00329	0.0178	0.894086
bodypart*Bpre	0.10585	1	0.10585	0.5729	0.450418
pop*sex*bodypart	0.21997	1	0.21997	1.1906	0.277136
pop*sex*Bpre	0.19324	1	0.19324	1.0459	0.308261
pop*bodypart*Bpre	0.16818	1	0.16818	0.9103	0.341725
sex*bodypart*Bpre	0.15000	1	0.15000	0.8119	0.369149
pop*sex*bodypart*Bpre	0.03587	1	0.03587	0.1942	0.660164
Error	25.12645	136	0.18475		

Table 4: ANCOVA for *CAP-D3* transcription level (dependent variable), *CAP-G* expression at exons 18 and 22 (continuous independent variables), and population (pop), sex, body part and B chromosome presence (Bpre) (discrete independent variables). SS= sum of squares, df = degrees of freedom, MS = mean sum of squares, p = p-value. Significant effects are noted in bold-type letter.

Item	SS	df	MS	F	p
Intercept	0,00007	1	0,000072	0,00015	0,990379
CAP-G18	0,75388	1	0,753884	1,53676	0,217204
CAP-G22	8,12576	1	8,125758	16,56405	**7,88E-05**
pop	0,22069	1	0,220692	0,44987	0,503517
sex	5,08332	1	5,083322	10,36216	**0,001604**
bodypart	0,02955	1	0,029545	0,06023	0,806503
Bpre	0,19611	1	0,196109	0,39976	0,528259
pop*sex	0,98455	1	0,984547	2,00696	0,158832
pop*bodypart	1,08372	1	1,083718	2,20912	0,139478
sex*bodypart	0,01597	1	0,015974	0,03256	0,857063
pop*Bpre	0,45883	1	0,458833	0,93531	0,335178
sex*Bpre	0,00739	1	0,007387	0,01506	0,902513
bodypart*Bpre	0,18477	1	0,184772	0,37665	0,540412
pop*sex*bodypart	0,08094	1	0,080939	0,16499	0,685231
pop*sex*Bpre	0,19129	1	0,191285	0,38993	0,533368
pop*bodypart*Bpre	0,00325	1	0,003251	0,00663	0,935232
sex*bodypart*Bpre	0,00000	1	0,000000	0,00000	1,000000
pop*sex*bodypart*Bpre	0,20048	1	0,200479	0,40867	0,523706
Error	67,69811	138	0,490566		

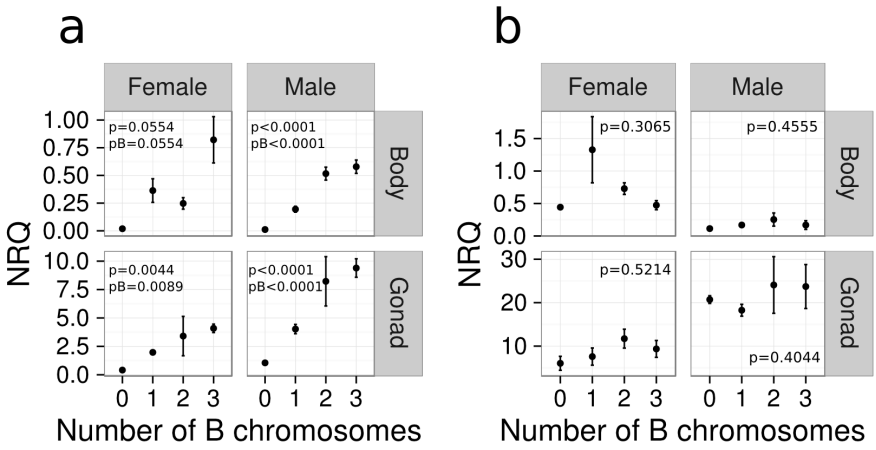

Figure 5: Relative quantification of *CAP-G* transcript abundance by means of qPCR analysis on cDNA from males and females carrying 0-3 B2 chromosomes collected at the Salobreña population, using primers anchored on exon 18 (a) and exon 22 (b). Note that *CAP-G* expression increased with B chromosome number when assayed for exon 18 (a) but not for exon 22 (b), suggesting the active transcription of B chromosome truncated gene copies. NRQ = normalized relative quantities; P = P-value for one-way ANOVA analysis; pB = Sequential Bonferroni P-value.

not when measured at exon 18 (which is present in both A and B chromosomes). Likewise, multiple linear regression analysis showed that the relationship of *CAP-G* transcription with *CAP-D2* (condensin I) and *CAP-D3* (condensin II) transcription was only significant for *CAP-G22*, i.e. for A chromosome *CAP-G* transcripts (Fig. 6; Tables S3 and S4). Taken together, these results indicate that the excess of *CAP-G*-like transcripts derived from the transcription of *B-CAP-G* pseudogenes does not alter the transcription rate of other condensin subunit genes, althought it does not discard possible influences of these transcripts at post-transcriptional level.

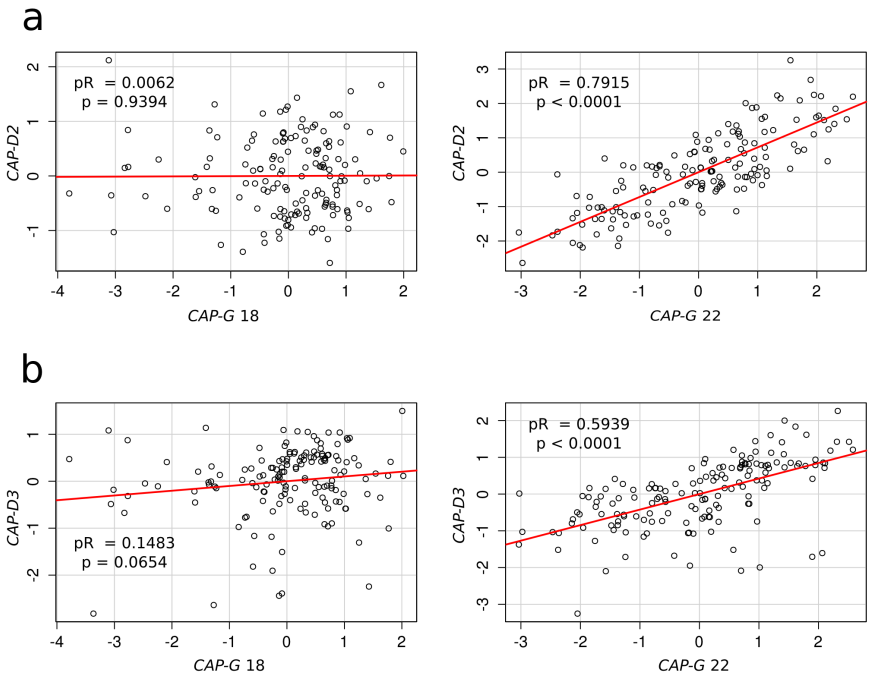

Figure 6: Partial regression plot showing the relationship of *CAP-G* transcription with the expression of *CAP-D2* (a) and *CAP-D3* (b) condensin I and II, respectively, subunit genes. Note that *CAP-G* transcription measured at exon 22 was positively correlated with *CAP-D2* and *CAP-D3* transcription levels, whereas that measured at exon 18 (which includes the *B-CAP-G* transcripts) was not. pR = partial correlation coefficient, p= p-value.

Discussion

Our present results, along with previous findings by Navarro-Domínguez et al. (submitted; see chapter 3), have shown that B chromosomes from two different populations (Torrox and Salobreña) carry several copies of the *B-CAP-G* pseudogene which are actively transcribed. Given the importance of the *CAP-G* gene for cell division, i.e. the main arena for B chromosome destiny, it is conceivable that the transcription of B chromosome *B-CAP-G* pseudogenes might change some gene regulation equilibria altering some cell division functions which might facilitate B chromosome transmission advantage and maintenance in natural populations. Some effects hitherto described for B chromosome presence in *E. plorans* are reminiscent of the effects described for condensin mutants and could thus be derived from the expression of the *B-CAP-G* pseudogene. For instance, B chromosome presence decreases egg fertility (i.e. the proportion of fertilized eggs) (Muñoz et al., 1998; Zurita et al., 1998; Bakkali et al., 2010) whereas, in Drosophila, some *CAP-G* mutants result in female sterile phenotypes (Resnick et al., 2009), and others provoke early mortality in embryos (Dej et al., 2004). Moreover, it has been observed that B chromosome presence leads to an increase in chiasma frequency, and thus recombination, on A chromosomes (Camacho et al., 2002), and this effect could actually be a byproduct of an anomalous activity of the condensin complex in B carrying individuals, likewise condensin mutants show higher rates of crossover in *C. elegans* (Mets and Meyer, 2009).

In addition, transcriptome and microarray analyses comparing gene

expression between B-carrying and B-lacking *E. plorans* females, have shown upregulation of topoisomerase IIα (TOP2A) in presence of the B chromosome (Navarro-Domínguez et al., in preparation; see chapter 5). This gene shows a strong functional relationship with condensins (Baxter and Aragón, 2012) and is involved in the resolution of the ultrafine anaphase bridges (Broderick et al., 2015). The up-regulation of TOP2A in B-carrying individuals could thus constitute a response to poor chromatid resolution caused by non proper function of the CAP-G protein. The high abundance of anomalous *CAP-G*-like transcripts, which may or may not be translated, could thus lead to a failure in the normal regulation of condensin function, causing mutant-like effects in a lesser extent than with total loss of function.

The possibility that the *CAP-G* transcript excess derived from *B-CAP-G* pseudogene expression causes an enhancement in condensin function is actually remote because 1) the *B-CAP-G* transcripts carry a severe non-synonymous substitution and a premature stop codon, and 2) transcript levels of other condensin subunit genes (*CAP-D2* and *CAP-D3*) are independent on *B-CAP-G* expression level, according to our ANCOVA and multiple regression analysis. More indirectly, previous work has reported that an excess of CAP-H2 leads to altered chromosome structure, dispersal of centromeres, chromosome unpairing and separation of salivary gland polytene chromosomal components (Hartl et al., 2008; Nguyen et al., 2014), but no such symptoms have been observed in B carrying individuals of *E. plorans*. A general hypercondensation of mitotic chromosomes and prophase shortening, as described for a gain of function mutation of *CAP-D3* (Bakhrebah et al., 2015), has neither been observed in *E. plorans*.

The aminoacid change in position 215 (within the HEAT domain) of the predicted protein for the *B-CAP-G* pseudogene could have a high impact on the functionality of the predicted protein product, since most of the loss-of-function mutations hitherto described for this protein took place within the HEAT domain, which is also the most conserved region (Cobbe et al., 2006). It is thus highly likely that the predicted protein for the *B-CAP-G* transcript is not fully functional. Even in this case, it is known that pseudogenic proteins could affect the activity of the parental proteins (for review, see Poliseno et al., 2015b). Moreover, the lack of several exons does not necessarily mean that the predicted B-CAP-G protein cannot perform the CAP-G function since, in Drosophila, the N-terminal two-thirds of CAP-G are sufficient for assemblying with the condensin I complex and efficient chromatin localization during mitosis, whereas the C-terminus is dispensable for condensin I function during cell cycle and development, although it is required for nuclear location and heterochromatinization during interphase (Herzog et al., 2013).

Even if the *B-CAP-G* transcripts were not translated, and considering that transcription rates of *CAP-D2* and *CAP-D3* are not altered by *B-CAP-G* transcription, the presence of anomalous *CAP-G* transcripts could influence post-transcriptionally the expression of the canonical *CAP-G* gene and, consequently, its normal functioning. There is growing evidence for a role of pseudogenes in the regulation of parental gene expression by means of several post-transcriptional levels of regulation, e.g. via epigenetic modification (Poliseno et al., 2015a), through the generation of endogenous siRNA (Tam et al., 2008) or else acting as competitive inhibitors for binding to microRNAs (Pink et al., 2011), the translational complex or other RNA-binding proteins (Poliseno et al.,

2015a).

We cannot exclude the possibility that *B-CAP-G* expression is simply incidental and lack further consequences other than the waste of energy employed to yield useless transcripts or polypeptides, i.e. it is simply transcriptional noise being part of the burden imposed by B chromosomes. Interestingly, the B24 variant in *E. plorans* harbors genes involved in cell division control and checkpoints, some of them showing a full-length CDS and being actively transcribed, thus suggesting a possible implication of B chromosome gene content in its own evolutionary success (Navarro-Domínguez et al., submitted; see chapter 3). Among the possible implications of the presence of *B-CAP-G* transcripts mentioned above, the partial inhibition of normal *CAP-G* function through some kind of A chromosomes mRNA neutralization, is highly consistent with the parasitic model of B chromosome evolution, as a decrease in CAP-G function might avoid the complete silencing of the B chromosome, thus allowing the expression of those B chromosome genes being important for its own survival. Of course, the mechanisms by which this A and B chromosome crosstalk takes place remain to be uncovered, but our present research opens new avenues for future research on such an interesting prospect.

Acknowledgements

We thank José L. Bella and Miguel Pita for sharing with us their protocol of the Tyramide-coupled FISH tecnique.

This study was funded by the Spanish Ministerio de Economía y Competitividad (CGL2015-70750-P) and Plan Andaluz de Investigación (CVI-6649). Navarro-Domínguez was supported by a FPI fellowship from the Spanish Ministerio de Ciencia e Innovación (BES-2010-030016).

References

Acilan, C., Potter, D. M., and Saunders, W. S. (2007). DNA repair pathways involved in anaphase bridge formation. *Genes, Chromosomes and Cancer*, 46(6):522–531.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Andrade, M. A. and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. *Nature genetics*, 11(2):115–116.

Aono, N., Sutani, T., Tomonaga, T., Mochida, S., and Yanagida, M. (2002). Cnd2 has dual roles in mitotic condensation and interphase. *Nature*, 417(6885):197–202.

Bakhrebah, M., Zhang, T., Mann, J. R., Kalitsis, P., and Hudson, D. F. (2015). Disruption of a conserved CAP-D3 threonine alters condensin loading on mitotic chromosomes leading to chromosome hypercondensation. *The Journal of Biological Chemistry*, 290(10):6156–6167.

Bakkali, M., Manrique-Poyato, M. I., López-León, M. D., Perfectti, F., Cabrero, J., and Camacho, J. P. M. (2010). Effects of B chromosomes on egg fertility and clutch

size in the grasshopper *Eyprepocnemis plorans*. *Journal of Orthoptera Research*, 19(2):197–203.

Baxter, J. and Aragón, L. (2012). A model for chromosome condensation based on the interplay between condensin and topoisomerase II. *Trends in genetics: TIG*, 28(3):110–117.

Berdan, E. L., Mazzoni, C. J., Waurick, I., Roehr, J. T., and Mayer, F. (2015). A population genomic scan in Chorthippus grasshoppers unveils previously unknown phenotypic divergence. *Molecular Ecology*, 24(15):3918–3930.

Broderick, R., Nieminuszczy, J., Blackford, A. N., Winczura, A., and Niedzwiedz, W. (2015). TOPBP1 recruits TOP2a to ultra-fine anaphase bridges to aid in their resolution. *Nature Communications*, 6:6572.

Cabello, O. A., Eliseeva, E., He, W. G., Youssoufian, H., Plon, S. E., Brinkley, B. R., and Belmont, J. W. (2001). Cell cycle-dependent expression and nucleolar localization of hCAP-H. *Molecular Biology of the Cell*, 12(11):3527–3537.

Cabrero, J., Alché, J. D., and Camacho, J. P. M. (1987). Effects of B chromosomes on the activity of nucleolar organizer regions in the grasshopper *Eyprepocnemis plorans*: activation of a latent nucleolar organizer region on a B chromosome fused to an autosome. *Genome*, 29(1):116–121.

Camacho, J. P., Cabrero, J., López-León, M. D., Cabral-de Mello, D. C., and Ruiz-Ruano, F. J. (2015). Grasshoppers (Orthoptera). In Sharakhov, I. V., editor, *Protocols for Cytogenetic Mapping of Arthropod Genomes*. Crc press edition.

Camacho, J. P. M. (2005). B chromosomes. In Gregory, T. R., editor, *The Evolution of the Genome*, pages 223–286. Academic Press.

Camacho, J. P. M., Bakkali, M., Corral, J. M., Cabrero, J., Lopez-Leon, M. D., Aranda, I., Martin-Alganza, A., and Perfectti, F. (2002). Host recombination is dependent on the degree of parasitism. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1505):2173–2177.

Camacho, J. P. M., Sharbel, T. F., and Beukeboom, L. W. (2000). B-chromosome evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355(1394):163–178.

Chapuis, M.-P., Tohidi-Esfahani, D., Dodgson, T., Blondin, L., Ponton, F., Cullen, D., Simpson, S. J., and Sword, G. A. (2011). Assessment and validation of a suite of reverse transcription-quantitative PCR reference genes for analyses of density-dependent behavioural plasticity in the Australian plague locust. *BMC Molecular Biology*, 12:7.

Choi, Y. and Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics (Oxford, England)*, 31(16):2745–2747.

Cobbe, N., Savvidou, E., and Heck, M. M. S. (2006). Diverse mitotic and interphase functions of condensins in Drosophila. *Genetics*, 172(2):991–1008.

Csankovszki, G., Collette, K., Spahl, K., Carey, J., Snyder, M., Petty, E., Patel, U., Tabuchi, T., Liu, H., McLeod, I., Thompson, J., Sarkeshik, A., Sarkesik, A., Yates, J., Meyer, B. J., and Hagstrom, K. (2009). Three distinct condensin complexes control *C. elegans* chromosome dynamics. *Current biology: CB*, 19(1):9–19.

Dej, K. J., Ahn, C., and Orr-Weaver, T. L. (2004). Mutations in the Drosophila condensin subunit dCAP-G: defining the role of condensin for chromosome condensation in mitosis and gene expression in interphase. *Genetics*, 168(2):895–906.

Drummond, A. J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., and Wilson, A. (2009). Geneious 4.8. *Biomatters, Auckland, New Zealand*.

Fantinatti, B. E. A., Mazzuchelli, J., Valente, G. T., Cabral-de Mello, D. C., and Martins, C. (2011). Genomic content and new insights on the origin of the B chromosome of the cichlid fish *Astatotilapia latifasciata*. *Genetica*, 139(10):1273–1282.

Freeman, L., Aragon-Alcaide, L., and Strunnikov, A. (2000). The condensin complex governs chromosome condensation and mitotic transmission of rDNA. *The Journal of Cell Biology*, 149(4):811–824.

Funkhouser-Jones, L. J., Sehnert, S. R., Martínez-Rodríguez, P., Toribio-Fernández, R., Pita, M., Bella, J. L., and Bordenstein, S. R. (2015). Wolbachia co-infection in a

hybrid zone: discovery of horizontal gene transfers from two Wolbachia supergroups into an animal genome. *PeerJ*, 3:e1479.

Graphodatsky, A. S., Kukekova, A. V., Yudkin, D. V., Trifonov, V. A., Vorobieva, N. V., Beklemisheva, V. R., Perelman, P. L., Graphodatskaya, D. A., Trut, L. N., Yang, F., Ferguson-Smith, M. A., Acland, G. M., and Aguirre, G. D. (2005). The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 13(2):113–122.

Hartl, T. A., Smith, H. F., and Bosco, G. (2008). Chromosome alignment and transvection are antagonized by condensin II. *Science (New York, N.Y.)*, 322(5906):1384–1387.

Herzog, S., Nagarkar Jaiswal, S., Urban, E., Riemer, A., Fischer, S., and Heidmann, S. K. (2013). Functional dissection of the *Drosophila melanogaster* condensin subunit Cap-G reveals its exclusive association with condensin I. *PLoS genetics*, 9(4):e1003463.

Hirano, T. (2005). Condensins: organizing and segregating the genome. *Current biology: CB*, 15(7):R265–275.

Hirano, T. (2012). Condensins: universal organizers of chromosomes with diverse functions. *Genes and Development*, 26(15):1659–1678.

Hirano, T. (2016). Condensin-based chromosome organization from bacteria to vertebrates. *Cell*, 164(5):847–857.

Huang, W., Du, Y., Zhao, X., and Jin, W. (2016). B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). *BMC Plant Biology*, 16:88.

Jäger, H., Rauch, M., and Heidmann, S. (2005). The *Drosophila melanogaster* condensin subunit Cap-G interacts with the centromere-specific histone H3 variant CID. *Chromosoma*, 113(7):350–361.

Jones, R. N. and Rees, H. (1982). *B chromosomes.* Academic Press.

Kim, J. H., Zhang, T., Wong, N. C., Davidson, N., Maksimovic, J., Oshlack, A., Earnshaw, W. C., Kalitsis, P., and Hudson, D. F. (2013). Condensin I associates

with structural and gene regulatory regions in vertebrate chromosomes. *Nature Communications*, 4:2537.

Koressaar, T. and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics (Oxford, England)*, 23(10):1289–1291.

Kruesi, W. S., Core, L. J., Waters, C. T., Lis, J. T., and Meyer, B. J. (2013). Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. *eLife*, 2:e00808.

Krylov, V., Tlapakova, T., and Macha, J. (2007). Localization of the single copy gene Mdh2 on *Xenopus tropicalis* chromosomes by FISH-TSA. *Cytogenetic and genome research*, 116(1-2):110–112.

Krylov, V., Tlapakova, T., Macha, J., Curlej, J., Ryban, L., and Chrenek, P. (2008). Localization of human coagulation factor VIII (hFVIII) in transgenic rabbit by FISH-TSA: identification of transgene copy number and transmission to the next generation. *Folia Biologica*, 54(4):121.

Machín, F., Paschos, K., Jarmuz, A., Torres-Rosell, J., Pade, C., and Aragón, L. (2004). Condensin regulates rDNA silencing by modulating nucleolar Sir2p. *Current biology: CB*, 14(2):125–130.

Martis, M. M., Klemme, S., Banaei-Moghaddam, A. M., Blattner, F. R., Macas, J., Schmutzer, T., Scholz, U., Gundlach, H., Wicker, T., Šimková, H., Novák, P., Neumann, P., Kubaláková, M., Bauer, E., Haseneyer, G., Fuchs, J., Doležel, J., Stein, N., Mayer, K. F. X., and Houben, A. (2012). Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences*, 109(33):13343–13346.

McClintock, B. (1941). Spontaneous alterations in chromosome size and form in Zea mays. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 9, pages 72–81. Cold Spring Harbor Laboratory Press.

Mets, D. G. and Meyer, B. J. (2009). Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure. *Cell*, 139(1):73–86.

Muñoz, E., Perfectti, F., Martín-Alganza, Á., and Camacho, J. (1998). Parallel effects of a B chromosome and a mite that decrease female fitness in the grasshopper

Eyprepocnemis plorans. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1408):1903.

Murillo-Pineda, M., Cabello-Lobato, M. J., Clemente-Ruiz, M., Monje-Casas, F., and Prado, F. (2014). Defective histone supply causes condensin-dependent chromatin alterations, SAC activation and chromosome decatenation impairment. *Nucleic Acids Research*, 42(20):12469–12482.

Navarro-Domínguez, B. (2016). *E. plorans* transcriptome files. *Figshare.* DOI: http://dx.doi.org/10.6084/m9.figshare.3408580.v3.

Neuwald, A. F. and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. *Genome Research*, 10(10):1445–1452.

Nguyen, P. A., Groen, A. C., Loose, M., Ishihara, K., Wühr, M., Field, C. M., and Mitchison, T. J. (2014). Spatial organization of cytokinesis signaling reconstituted in a cell-free system. *Science (New York, N.Y.)*, 346(6206):244–247.

Ning, Z., Cox, A. J., and Mullikin, J. C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Research*, 11(10):1725–1729.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucl. Acids Res.*, 29(9):e45–e45.

Pink, R. C., Wicks, K., Caley, D. P., Punch, E. K., Jacobs, L., and Carter, D. R. F. (2011). Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA*, 17(5):792–798.

Poliseno, L., Marranci, A., and Pandolfi, P. P. (2015a). Pseudogenes in human cancer. *Frontiers in medicine*, 2.

Poliseno, L., Marranci, A., and Pandolfi, P. P. (2015b). Pseudogenes in human cancer. *Pathology*, page 68.

Resnick, T. D., Dej, K. J., Xiang, Y., Hawley, R. S., Ahn, C., and Orr-Weaver, T. L. (2009). Mutations in the chromosomal passenger complex and the condensin complex differentially affect synaptonemal complex disassembly and metaphase I configuration in Drosophila female meiosis. *Genetics*, 181(3):875–887.

Ruiz-Estévez, M., Badisco, L., Broeck, J. V., Perfectti, F., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2014). B chromosomes showing active ribosomal RNA genes contribute insignificant amounts of rRNA in the grasshopper *Eyprepocnemis plorans*. *Molecular genetics and genomics*, 289(6):1209–1216.

Ruiz-Estévez, M., Lopez-Leon, M. D., Cabrero, J., and Camacho, J. P. M. (2012). B-chromosome ribosomal DNA is functional in the grasshopper *Eyprepocnemis plorans*. *PLoS One*, 7(5):e36600.

Shimizu, N., Shingaki, K., Kaneko-Sasaguri, Y., Hashizume, T., and Kanda, T. (2005). When, where and how the bridge breaks: anaphase bridge breakage plays a crucial role in gene amplification and HSR generation. *Experimental Cell Research*, 302(2):233–243.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., and Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453(7194):534–538.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, 119(2):217–225.

Uhlmann, F. (2002). Cell biology: keeping the genome in shape. *Nature*, 417(6885):135–136.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40(15):e115.

Uzbekov, R., Timirbulatova, E., Watrin, E., Cubizolles, F., Ogereau, D., Gulak, P., Legagneux, V., Polyakov, V. J., Le Guellec, K., and Kireev, I. (2003). Nucleolar association of pEg7 and XCAP-E, two members of *Xenopus laevis* condensin complex in interphase cells. *Journal of Cell Science*, 116(Pt 9):1667–1678.

Valente, G. T., Conte, M. A., Fantinatti, B. E. A., Cabral-de Mello, D. C., Carvalho, R. F., Vicari, M. R., Kocher, T. D., and Martins, C. (2014). Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution*, 31(8):2061–2072.

Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. V., and Simonet, G. (2009). Identification and validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. *BMC Molecular Biology*, 10:56.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7):RESEARCH0034.

Wang, B.-D., Butylin, P., and Strunnikov, A. (2006). Condensin function in mitotic nucleolar segregation is regulated by rDNA transcription. *Cell Cycle (Georgetown, Tex.)*, 5(19):2260–2267.

Wang, B.-D. and Strunnikov, A. (2008). Transcriptional homogenization of rDNA repeats in the episome-based nucleolus induces genome-wide changes in the chromosomal distribution of condensin. *Plasmid*, 59(1):45–53.

Wang, B.-D., Yong-Gonzalez, V., and Strunnikov, A. V. (2004). Cdc14p/FEAR pathway controls segregation of nucleolus in S. cerevisiae by facilitating condensin targeting to rDNA chromatin in anaphase. *Cell Cycle (Georgetown, Tex.)*, 3(7):960–967.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C., and others (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature communications*, 5.

Wood, A. J., Severson, A. F., and Meyer, B. J. (2010). Condensin and cohesin complexity: the expanding repertoire of functions. *Nature Reviews. Genetics*, 11(6):391–404.

Xu, X., Nakazawa, N., and Yanagida, M. (2015). Condensin HEAT subunits required for DNA repair, kinetochore/centromere function and ploidy maintenance in fission yeast. *PloS One*, 10(3):e0119347.

Yong-Gonzalez, V., Wang, B.-D., Butylin, P., Ouspenski, I., and Strunnikov, A. (2007). Condensin function at centromere chromatin facilitates proper kinetochore tension and ensures correct mitotic segregation of sister chromatids. *Genes to Cells: Devoted to Molecular and Cellular Mechanisms*, 12(9):1075–1090.

Zhang, T., Paulson, J. R., Bakhrebah, M., Kim, J. H., Nowell, C., Kalitsis, P., and Hudson, D. F. (2016). Condensin I and II behaviour in interphase nuclei and cells undergoing premature chromosome condensation. *Chromosome Research*, 24(2):243–269.

Zurita, S., Cabrero, J., López-León, M., and Camacho, J. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, pages 274–277.

Supplementary material

Table S1: Pearson's linear correlation analysis between the number of B chromosomes and *CAP-G* abundance analyzed by qPCR in gDNA of *E. plorans* males (N=15) from Salobreña. Note that CAP-G abundance was measured separately for exon 18 (which is present in both A and B chromosomes) and exon 22 (only present in A chromosomes). N= number of individuals, df= degrees of freedom, r= linear correlation coefficient, p= p-value. Significant effects are noted in bold-type letter.

Item	r	t	df	p
CAP-G18	0.7980	47.741	13	**0.0004**
CAP-G22	0.2192	0.81011	13	0.4325

Table S2: One-way ANOVA analyses comparing *CAP-G* transcription level between *E. plorans* individuals with different number of B chromosomes. Note that transcription level was separately measured at exon 18 (being present in both A and B chromosomes) and exon 22 (only present in A chromosomes). df= degrees of freedom, SS= sum of squares, MS= mean sum of squares, p= p-value, pB= sequential Bonferroni p-value. Significant effects are noted in bold-type letter.

Exon	Body part	df	SS	MS	F	p	pB
CAP-G18	Female Body	3	1,1788	0,39294	3,5546	5,54E-02	5,54E-02
	Female Gonad	3	23,7264	7,9088	8,3638	4,44E-03	**8,88E-03**
	Male Body	3	0,83729	0,27909	29,813	4,00E-07	**1,60E-06**
	Male Gonad	3	159,939	53,313	18,015	1,00E-05	**3,00E-05**
CAP-G22	Female Body	3	2,1851	0,72837	1,374	0,3065	
	Female Gonad	3	40,459	13,486	0,8005	0,5214	
	Male Body	3	0,04574	0,01525	0,9105	0,4555	
	Male Gonad	3	114,72	38,239	1,0375	0,4044	

Table S3: Partial and semipartial correlation between CAP-G transcription levels, measured at exon 18 and 22, on the expression of the CAP-D2 subunit gene of condensin I. Beta= regression coefficient, t= Student t test (degrees of freedom in brackets), p = p-value. Significant effects are noted in bold-type letter.

Exon	Beta	Partial	R-square	t(151)	p
CAP-G18	0,004251	0,006195	0,565875	0,07612	0,939423
CAP-G22	0,888739	0,791505	0,565875	15,91422	**4,18E-34**

Table S4: Partial and semipartial correlation between CAP-G transcription levels, measured at exon 18 and 22, on the expression of the CAP-D3 subunit gene of condensin II. Beta= regression coefficient, t= Student t test (degrees of freedom in brackets), p = p-value. Significant effects are noted in bold-type letter.

Exon	Beta	Partial	R-square	t(153)	p-level
CAP-G18	0,138042	0,148355	0,550874	1,855581	0,065437
CAP-G22	0,679302	0,593917	0,550874	9,131272	**3,78E-16**

5

The presence of a parasitic chromosome elicits gene expression changes in the host genome

Beatriz Navarro-Domínguez[1], Francisco J. Ruíz-Ruano[1], Josefa Cabrero[1], Jose María Corral[2,3], María Dolores López-León[1], Timothy F. Sharbel[2,4], Juan Pedro M. Camacho[1]

[1]*Departamento de Genética, Universidad de Granada*

[2]*Leibniz Institute for Plant Genetics and Crop Plant Research (IPK), Gatersleben (Germany)*

[3]*Department of Bioanalytics, Coburg University of Applied Sciences and Arts*

[4]*Global Institute for Food Security (GIFS), University of Saskatchewan*

Abstract. Parasitism evokes adaptive physiological changes in the host, many of which take place through gene expression changes. This response can be more or less local, depending on the organ or tissue affected by the parasite, or else systemic when the parasite is present in all host cells.

The most extreme of the latter cases is intragenomic parasitism, where the parasite is present in all host nuclei as any other genomic element. Here we show the molecular crosstalk between a parasitic chromosome (also named B chromosome) and the host genome, manifested through gene expression changes. The transcriptome analysis of 0B and 1B females of the grasshopper *Eyprepocnemis plorans*, validated by a microarray experiment performed on four B-lacking and five B-carrying females, revealed changes in gene expression for 188 unigenes being consistent in both experiments. Once discarded B-derived transcripts, there were 46 differentially expressed genes (30 up- and 16 downregulated) related with the adaptation of the host genome to the presence of the parasitic chromosome. Interestingly, the functions of these genes could explain some of the most important effects of B chromosomes, such as nucleotypic effects derived from the additional DNA they represent, chemical defense and detoxification, protein modification and response to stress, ovary function and regulation of gene expression. Collectively, these changes uncover an intimate host-parasite crosstalk during crucial steps of gene expression and protein function.

Introduction

Parasitism is very frequent in nature, and can be disguised in multiple forms and at several levels, from genes to species. As part of the molecular crosstalk between the different partners, parasites evoke gene

expression changes in the host (Soumana et al., 2014). The most conspicuous parasites usually focus on a discrete organ or stage of host life cycle, and as such host response can be limited to the affected parts. However, intragenomic parasites are present in all host cell nuclei, hence host response needs to be systemic. Supernumerary (B) chromosomes are a frequent genomic component of a wide variety of eukaryotes, behaving as parasitic elements whose spread in natural populations is based on transmissional advantage (drive), following the typical arms race dynamics of host-parasite symbioses (Camacho, 2005). As intragenomic parasites, B chromosomes can trigger gene expression changes at any stage of host ontogeny. On this basis, B chromosomes constitute an excellent model to shed light upon intragenomic adaptive changes in gene expression during parasitism.

Up to 11 years ago, the only DNA sequences known on B chromosomes were repetitive DNA sequences such as satellite DNA, ribosomal DNA and transposable elements (TEs) (Camacho, 2005). However, from 2005 onwards, protein-coding genes or pseudogenes have been found in the B chromosomes of several species (Graphodatsky et al., 2005; Teruel et al., 2010; Yoshida et al., 2011; Martis et al., 2012; Trifonov et al., 2013; Banaei-Moghaddam et al., 2013; Valente et al., 2014; Huang et al., 2016, Navarro-Dominguez et al., submitted; see chapter 3). The general belief of genic inactivity of B chromosomes (Camacho et al., 2000) has also changed during last years by the findings of transcription of B chromosome sequences (Leach et al., 2005; Van Vugt et al., 2003; Ruiz-Estévez et al., 2012; Carchilan et al., 2009; Zhou et al., 2012; Trifonov et al., 2013; Banaei-Moghaddam et al., 2013; Valente et al., 2014; Huang et al., 2016; Navarro-Dominguez et al., submitted; see

chapter 3)

Gene activity of B chromosomes can potentially elicit a gene regulation response by the host genome (i.e. A chromosomes), but comparative transcriptome analysis of B-carrying and B-lacking individuals has hitherto been performed only in the parasitic wasp *Nasonia vitripennis* (Akbari et al., 2013) and maize (Huang et al., 2016). In *N. vitripennis*, transcriptome analysis in testes led to the identification of nine transcripts lacking homology to any known DNA sequence, which were expressed only in the B-carrying transcriptome and, by means of fluorescence in situ hybridization, they showed that at least three of them were physically located in the B chromosome (i.e. the Paternal Sex Ratio chromosome, PSR). Contrary to their expectations, they did not find any traces of gene expression changes associated to PSR chromosome presence in relation with the known effects of this B chromosome, such as chromatin structure or condensation, transposable elements or small RNA regulation pathways (Akbari et al., 2013). In maize, B chromosome presence influences A-genome transcription, with 130 differentially expressed genes mainly involved in cell metabolism and nucleotide binding (Huang et al., 2016).

Almost all natural populations of the grasshopper *Eyprepocnemis plorans* in the circum-Mediterranean region carry B chromosomes (López-León et al., 2008). The presence of a same B variant (B1) in populations from Spain, Morocco, Tunisia and Sicilia suggests a recent spread of B chromosomes into these areas (Cabrero et al., 2014). The high success of B chromosomes in this species results from their transmission advantage during female meiosis (Zurita et al., 1998) and the scarcity of phenotypic effects, as B-carrying individuals do not show differences in body size in

comparison to B-lacking ones (Camacho et al., 1980). However, some endophenotypic effects of B presence have been reported with respect to NOR activity (Cabrero et al., 1987; López-León et al., 1995; Teruel et al., 2007), chiasma frequency (Camacho et al., 2002), egg fertility and clutch size (Zurita et al., 1998; Bakkali et al., 2010), spermatid formation (Teruel et al., 2009a) and heat shock protein 70 (Hsp70) levels (Teruel et al., 2011).

B chromosomes in *E. plorans* are able to transcribe their ribosomal DNA and organize a nucleolus (Ruiz-Estévez et al., 2012), but this occurs only in a minority of males in most populations (Ruiz-Estévez et al., 2013) and the relative rRNA contribution of the B chromosome is insignificant compared to that of A chromosomes (Ruiz-Estévez et al., 2014), suggesting that B chromosomes in this species are highly repressed. However, our recent finding of nine protein-coding genes residing in the B chromosome of this species, six of which were actively expressed (Navarro-Dominguez et al.,submitted; see chapter 3), indicates that B chromosomes are not so silenced as previously thought, and suggests the possibility that a transcriptomic crosstalk may be operating between A and B chromosomes in B-carrying individuals.

To investigate this interesting possibility, we examine here changes in genome-wide gene expression patterns associated with the presence and absence of B-chromosomes. For this purpose, we used two different techniques (RNAseq and microarrays) on two types of samples (whole body and ovaries) from two populations harbouring two different B chromosome variants, namely B2 and B24, the latter being derived from the former (Henriques-Gil and Arana, 1990).

Materials and methods

Materials

E. plorans individuals were collected in the Torrox (Málaga) and Salobreña (Granada) populations harbouring the B24 and B2 variants, respectively. The number of B chromosomes was determined by C-banding of interphase hemolymph nuclei in females (Cabrero et al., 2006), and by C-banding of spermatocytes in males. B chromosome presence/absence was also confirmed by PCR amplification of the B-specific SCAR marker described in Muñoz-Pajares et al. (2011) on genomic DNA. Complete bodies of two females from Torrox (0B and 1B) were used for the RNA Illumina sequencing experiment, and 9 females from Salobreña (4 with 0B and 5 with 1B chromosome) were dissected in order to obtain their ovaries, whose RNA was used for the microarray experiment.

Illumina sequencing

Total RNA was extracted from each of two females, one lacking B chromosomes (0B) and the other carrying 1B. Both libraries were sequenced on an Illumina Hiseq2000 platform, each yielding about 5Gb of paired-end reads (2x101 nt). Illumina sequences are available in NCBI SRA database under accession numbers SRR2969416 (RNA_0B) and SRR2969417 (RNA_1B).

Transcriptome assembly, annotation and differential expression

De novo transcriptome assembly was carried out with Trinity software release 20131111 (Grabherr et al., 2011). Read preparation prior to assembly and downstream analysis of the *de novo* transcriptome were performed following the guidelines provided by Haas et al. (2013). We used Trimmomatic (Bolger et al., 2014) to remove adapters, low quality or N bases with quality lower than Q3 at the beginning and the end of the reads, nucleotides with an average quality lower than Q15 in a sliding window of four bases, and those reads which remained smaller than 36 bases long.

Prior to assembly, we normalized the libraries by kmer coverage in order to reduce redundant information from the deep sequencing reads, with the aim of making assembly easier, shortening computational time, increasing the chance for detection of rare transcripts, and avoiding bias derived from differences in expression between genes (Haas et al., 2013). For this purpose, we used the normalize_by_kmer_coverage.pl script provided by Trinity with the –JM 50G, –max_cov 30, –pairs_together and –PARALLEL_STATS parameters. This approach extracts K-mers with the Jellyfish algorithm (Marçais and Kingsford, 2011) so that each RNAseq read is probabilistically selected based on its median k-mer coverage value and the targeted maximum coverage value (Haas et al., 2013). Reads from the two libraries (0B and 1B) were assembled as a pooled data set, in order to assemble a reference *de novo* transcriptome. Assembled sequences being shorter than 200 base pairs were discarded.

Contamination in the *de novo* assembled transcriptome was examined

using the standalone version of DeconSeq (Schmieder and Edwards, 2011). This program aligns the sequences of our transcriptome to a database containing sequences from possible sources of contamination, using BWA (Li and Durbin, 2009) and discarding matched sequences. We ran this software using the viruses, bacteria and human databases provided by the developers, plus two custom databases for fungi and plants, built from the genomes of *Saccharomyces cerevisiae* and *Arabidopsis thaliana*, which were downloaded from the NCBI FTP server[I]. Subsequent analyses were carried out with the decontaminated assembly.

Functional annotation was done following the Trinotate pipeline (release 20140708; Grabherr et al., 2011). Protein coding sequences (CDS) were predicted using TransDecoder(Haas et al., 2013), detecting open reading frames (ORFs) with 300 bp minimun length. Sequence homology search was performed with BLASTX of the transcripts and BLASTP (Altschul et al., 1990, 1997) of the predicted proteins against UniProtKB/Swiss-Prot and UniProt Reference Clusters (UniRef) databases, using default settings. In addition, protein domains were analyzed with HMMER (Finn et al., 2011) and PFAM (Finn et al., 2013).

Gene function classifications were performed according to two standarized methods: Gene Ontology (GO) (Ashburner et al., 2000) and Eukaryotic Orthologous Groups (KOG) (Tatusov et al., 2003). GO assignments to predicted proteins were performed with Trinotate, and KOG classification was performed with the WebMGA software (Wu et al., 2011), searching the KOG database of NCBI.

In addition, we screened the *de novo* assembly transcriptome for known

[1]ftp://ftp.ncbi.nlm.nih.gov/genomes/

transposable elements (TEs), by means of RepeatMasker (Smit et al., 1996) in a database including TEs described in *Locusta migratoria* (data obtained from Repbase, Jurka et al., 2005).

Trinity output consists of a set of sequences (called "isoforms" in Trinity release 20131111, but we will call them simply "sequences" to avoid confusion) grouped into clusters ("c"). In the case of protein-coding sequences, all sequences included in the same cluster and coding for homologous proteins will be considered as sequence variants of the same gene, i.e. a unigene. The sequences within unigenes being most similar to the homologous sequences in the anotation database were used as representative when needed. In case of sequences showing homology with transposable elements, we grouped those showing 80% or higher similarity using CD-HIT-EST (Fu et al., 2012), with default options except -c 0.8, in order to remove redundancy and assembly artifacts.

Differential expression analysis based on Illumina sequencing

For differential expression analysis, Illumina reads were mapped against the reference transcriptome in order to get an estimate of read abundance for each gene or isoform. Prior to mapping, libraries were trimmed with Trimmomatic (Bolger et al., 2014), in order to remove adapters, bases with quality lower than Q3 in the beginning and the end of the read, and bases with an average quality lower than Q15 in a sliding window of four bases. After trimming, reads that became smaller than 36 bases long were discarded. Mapping was performed using the Bowtie algorithm (Langmead et al., 2009) and read abundance per gene and sequence

was estimated by RSEM (Li and Dewey, 2011) and expressed in FPKM (Fragments Per Kilobase Per Million Fragments Mapped). Differential expression was analyzed using edgeR (Robinson et al., 2010). Bowtie2, RSEM and edgeR were used as implemented in the Trinity pipeline (Haas et al., 2013).

Microarray differential expression analysis

For microarray analysis, we used nine females (four 0B and five 1B) from Salobreña. We extracted total RNA from the ovaries using the RNeasy Lipid Tissue Mini Kit (Qiagen). Samples were treated with DNAse I in a column as described above. Quality and absence of DNA contamination was verified in an Agilent 2100 Bioanalyzer using RNA Nanochips (Agilent Technologies). RNA quantity was determined in a Nanodrop ND-1000 spectrophotometer. 62.5 ng of RNA, which was the maximum quantity available from the least concentrated sample, was reamplified and retrotranscribed with a TransPlex Whole Transcriptome Amplification Kit (Sigma), following manufacturer's instructions and the recommendations provided in the NimbleGen Arrays User's Guide (Gene Expression Arrays v6.0).

NimbleGen Custom 3x1.4M Gene Expression arrays were designed using 434,838 *E. plorans* sequences, obtained by the assembly of the Illumina reads with TransABySS software (Robertson et al., 2010). Probe length was 60bp. Layout was outsourced to NimbleGen. Labelling, hybridation, washing and scanning of the arrays was performed following the instructions in the NimbleGen Arrays user's guide for gene expression

arrays on an MS 200 microarray scanner.

To allocate expression values from the microarrays to the Trinity-assembled transcriptome, we searched for homology between the 434,838 TransABySS- assembled sequences used for building the microarrays and the 73,889 sequences assembled by Trinity, by means of BLASTN (Altschul et al., 1990). For each Trinity sequence, the microarray expression values from up to 20 sequences showing high similarity (E-value<1e-100) were averaged.

RMA normalization of the microarray data was performed with the oligo R package (Carvalho and Irizarry, 2010) and the differences in expression between the two groups were assessed by a t-test performed with the limma R package (Smyth, 2005). Data analysis was performed following the steps described for Nimblegen arrays in the related vignette of the oligo package. Both packages and the vignettes are available in Bioconductor (Gentleman et al., 2004).

Results

Illumina sequencing and *de novo* transcriptome assembly

Illumina sequencing generated 35,345,561 reads for the RNA_0B library and 27,247,068 for the RNA_1B library.

After removing potential contaminants, the *de novo* assembled transcriptome included 73,889 sequences grouped into 45,555 unigenes.

Table 1: Summary statistics of Trinity assembly before and after removal of potential contaminations with DeConSeq

Item		Before deconseq	After deconseq
Sequences	Transcripts	74378	73889
	Genes (Isoform clusters)	45633	45555
N50	Transcripts	2330	2302
	Longest isoform per gene	1703	1702
Max. length	Transcripts	21010	19141
	Longest isoform per gene	21010	19141
Min. length	Transcripts	201	201
	Longest isoform per gene	201	201
Average length	Transcripts	1325.27	1314.80
	Longest isoform per gene	914.75	914.65
Median length	Transcripts	793	788
	Longest isoform per gene	485	485
Total assembled bases	Transcripts	98571226	97149070
	Longest isoform per gene	41742580	41666905
Percent GC		40.37	40.24

Sequence length ranged from 200 to 19,141 bp, with 788 bp median sequence length and 2,302 N50. For the unigenes, median length was 485 bp and N50 was 1,702. A comparison of the assembly before and after decontamination demonstrated a low contamination level (~0.65%). A summary of Illumina sequencing results and assembly and decontamination statistics are described in Table 1.

Functional annotation of *E. plorans* transcriptome

About 42% of the 45,555 unigenes showed significant similarity (Evalue < 10e-5) with proteins in Uniprot, Uniref90 or both. In most cases (~45%) a single potential coding sequence (CDS) was found, and a high percentage (~81%) of the sequences with a single CDS could be

assigned to one or more Gene Ontology terms based on BLAST matches to sequences with known function. In other cases, more than one CDS (~27%) or no potential CDS (~28%) were found. This could be due to the expression of pseudogenes with fragmented CDS or to sequencing, assembling or ORF prediction artifacts, and this was taken into account for further analysis.

The remaining 58% of the unigenes did not show significant homology with known coding proteins, being 91% of them apparently non protein-coding transcripts (i.e. lacking a CDS).

Summarizing, we found 18,999 unigenes with BLAST hits, 13,570 of which carried one or more CDSs, and 2,429 unigenes carrying a putative CDS but failing to show BLAST hits, whereas a high number of transcripts (24,127) did not bear any predicted CDS and did not show BLAST hits to known proteins (Fig. 1A).

In addition, we screened the *de novo* assembled transcriptome for TEs using RepeatMasker. The number of unigenes showing homology with the TEs described in *Locusta migratoria* (data obtained from Repbase, Jurka et al., 2005) was 5,555 (i.e. 12% of total unigenes), the most frequent being Mariner/Tc1 (923), Penelope (749) and RTE (638)(Fig. 1B).

B chromosome presence triggers differential expression for 188 unigenes

According to edgeR results, we found that 24,462 sequences (grouped in 16,013 unigenes) showed significant differential expression between

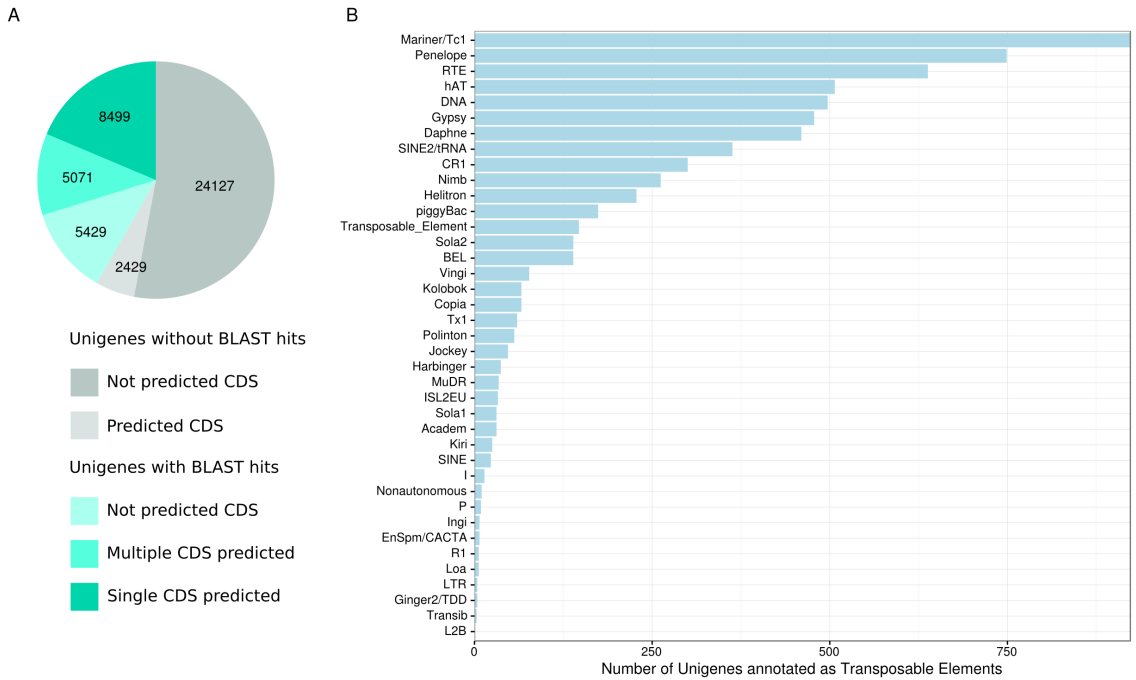

Figure 1: Overview of *E. plorans* transcriptome annotation. (A) Results of BLAST searches in UniProtKB/Swiss-Prot and UniProt Reference Clusters (UniRef) databases. (B) Number of Trinity unigenes annotated by RepeatMasker as different families of Transposable Elements in the *E. plorans* transcriptome

0B and 1B females ($P < 0.05$). We validated these transcriptome results with a microarray experiment performed on females from a different population (Salobreña) and using RNA from the ovary instead of the whole body. The t-test performed on microarray data resulted in 1,614 differentially expressed sequences (grouped in 1,202 unigenes) ($P < 0.05$). Statistical corrections for multiple tests (e.g. Bonferroni) would erase significance for all but four non-annotated sequences. Alternatively, we intersected the significant results obtained in the transcriptome and microarray experiments and considered as significant gene expression changes only those being present in both cases and in the same direction. This means that we actually tested differential expression in ovary. Since the B-carrying females used for the transcriptome and microarray experiments carried different B chromosome variants (B24 and B2, respectively), we can infer that the changes observed following our approach are common to both kinds of B chromosomes.

A comparative analysis of the identity of these sequences between the transcriptome and microarray samples revealed that 535 sequences showed significant differential expression, 258 of which showed opposite expression patterns, i.e. they were upregulated according to one of the analyses and downregulated according to the other. The remaining 277 sequences (grouped in 188 unigenes) showed matching expression patterns in both analyses, thus representing gene expression changes associated with the presence of B chromosomes taking place in both analyzed populations (Fig. 2). Remarkably, 246 of these differentially expressed sequences (161 unigenes, i.e. 89%) were upregulated in the presence of the B chromosome, whereas only 31 sequences (27 unigenes, i.e. 11%) were downregulated (Table 2).

Table 2: Summary of the anotation of the differentially expressed unigenes

	Annotated CDS	Non-annotated	Transposable elements	Total
Up-regulated	35	86	40	161
Down-regulated	16	9	2	27
Total	51	95	42	188

Among the 188 unigenes corresponding to the differentially expressed sequences, 95 unigenes (122 sequences) failed to show homology with any protein described in Uniprot or Uniref90 for any organism or repetitive element described for *L. migratoria* in RepBase. However, 42 out of the 93 remaining unigenes (66 sequences) showed homology with transposable elements, and the remaining 51 unigenes (89 sequences) showed homology with protein-coding genes described in Uniprot, Uniref90 or both (Table 2). This low proportion of annotations was undoubtedly due to the absence of a fully annotated genome in *E. plorans* or other grasshopper species, since the recently published genome of *L. migratoria* is still in the draft stage (Wang et al., 2014). Our subsequent interpretation of these results is thus conditioned by this partial annotation. Anyway, we will concentrate efforts on possible functional meaning.

Host genome adaptation for B chromosome presence stands on differential expression for 46 unigenes

Assuming that the five differentially expressed unigenes located in the B chromosome (i.e. CIP2A, CKAP2, CND3, KIF20A and MYCB2) showed up-regulation due to the transcription of B-located copies

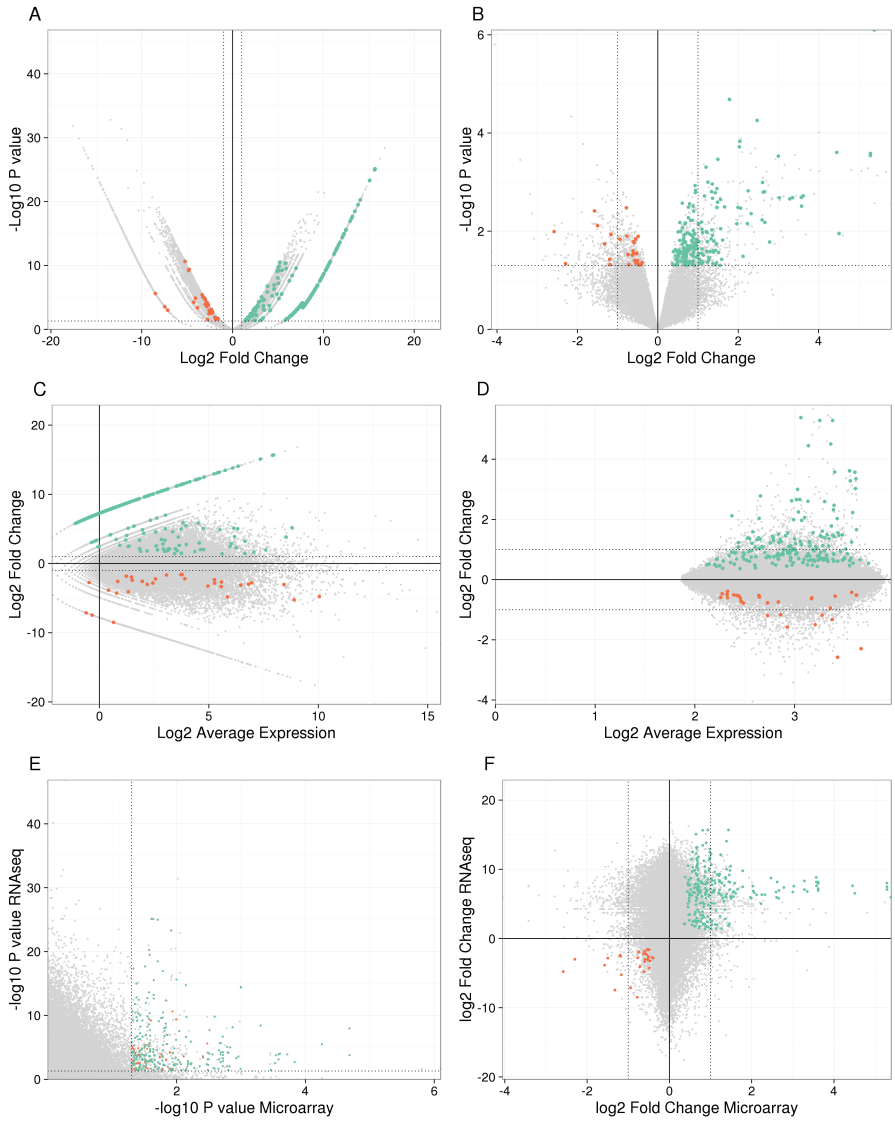

Figure 2: Differential expression analysis of B-carrying and B-lacking samples. Volcano plots (Fold Change vs. Statistical Significance) and MA plots (Fold Change vs. Mean Expression) for RNAseq (A and C) and microarray (B and D) data in *E. plorans*. Coincident differentially expressed sequences (277) between RNAseq and microarray analyses, in terms of statistical significance (E) and fold change (F). Green and orange dots represent up- and downregulations, respectively. Note the higher number of green dots, indicating that B chromosome presence causes a general increase in gene expression.

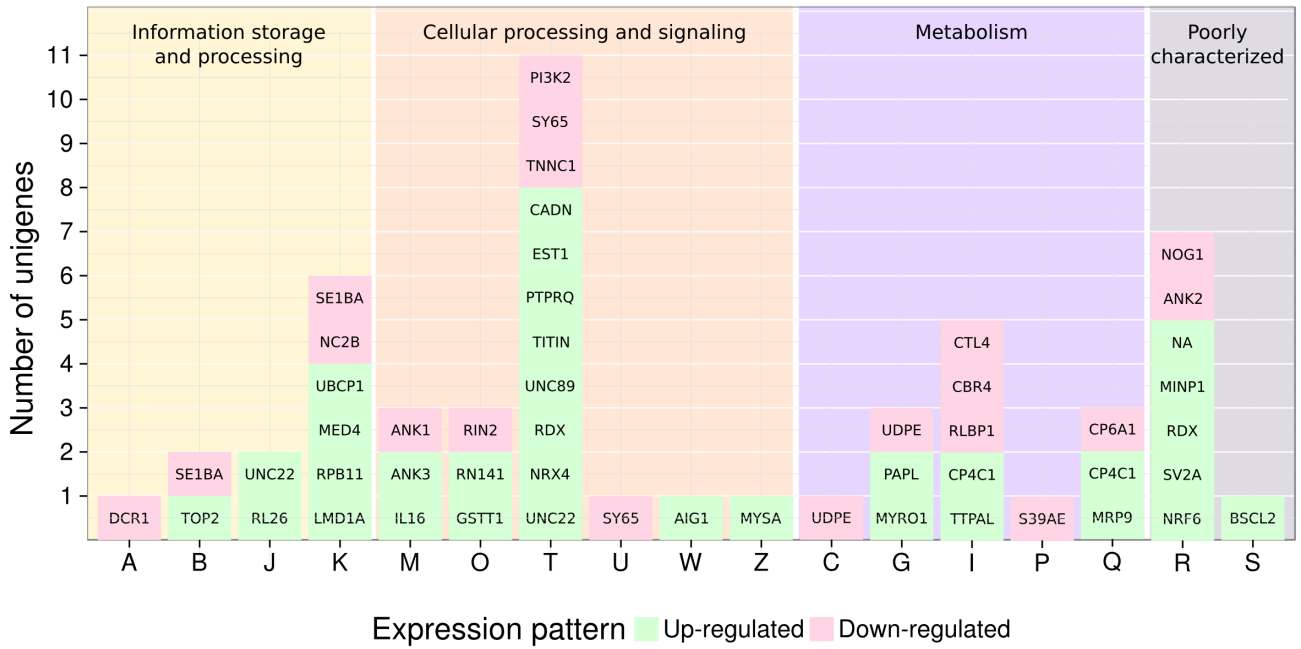

Figure 3: Number of differentially expressed unigenes, in the host genome, within each KOG class, defined as: A: RNA processing and modification, B: Chromatin Structure and dynamics, J: Translation, ribosomal structure and biogenesis, K: Transcription, M: Cell wall/membrane/envelope biogenesis, O: Post-translational modification, protein turnover, chaperone functions, T: Signal Transduction Mechanisms, U : Intracellular trafficing and secretion, W: Extracellular structures, Z: Cytoskeleton, C: Energy production and conversion, G: Carbohydrate metabolism and transport, I: Lipid transport and metabolism, P: Inorganic ion transport and metabolism, Q: Secondary Structure, R: General Functional Prediction only, S: Function Unknown.

(Navarro-Dominguez et al., submitted; see chapter 3), we can delimit gene expression changes in the host genome, associated with the presence of the parasitic chromosome, to 46 unigenes (30 up- and 16 down-regulated), after discarding these five unigenes. We then analyzed the GO terms associated with these 46 unigenes and interpreted them as gene expression changes that the parasitic chromosome triggers on the host genome. They were related with post-transcriptional gene expression regulation, endoplasmic reticulum unfolded protein response and endoplasmic reticulum stress, post-transcriptional gene silencing via small RNA, histone-methyltransferase activity, DNA conformation change, protein kinase activity and regulation of cell death (Table 3, Supplementary table S1 available in Figshare; Navarro-Domínguez, B., 2016). Remarkably, some of these terms define expected functional consequences of the crossfire between the genome and the parasitic chromosome (see below). Most these functions were also apparent in terms of the KOG classification (Fig. 3). Surprisingly, no gene expression changes in the host genome were related with cell cycle and cell division (KOG class D), in contrast with some of the B-located genes (Navarro-Domínguez et al., submitted; chapter 3).

Drammatic upregulation of TEs in the B-carrying transcriptome

About 22% of the 188 unigenes showing differential expression in B-carrying ovaries were annotated as TEs (Table 2). Among them we found Mariner/Tc1 (12), RTE (7), SINE (4), Unknown (3), Daphne (3), Penelope (3), DNA (2), Gypsy (2), CR1 (1), hAT (1), L2 (1), R1

Table 3: Gene Ontology terms for some of the differentially expressed unigenes

Expression	DEG	GO.Id	GO. Term
Down regulated	RIN2	GO:0035966	Response to topologically incorrect protein
		GO:0035967	Cellular response to topologically incorrect protein
		GO:0030968	Endoplasmic reticulum unfolded protein response
		GO:0034976	Response to endoplasmic reticulum stress
	DCR1	GO:0016441	Posttranscriptional gene silencing
		GO:0031047	Gene silencing by RNA
		GO:0035194	Posttranscriptional gene silencing by RNA
		GO:0035196	Production of mirnas involved in gene silencing by miRNA
		GO:0070918	Production of small RNA involved in gene silencing by RNA
Up regulated	SE1BA	GO:0018024	Histone-Lysine N-methyltransferase activity
	TITIN	GO:0071103	DNA conformation change
	TOP2		
	ANKK1	GO:0004672	Protein kinase activity
	UNC89		
	UNC22		
	RDX	GO:0010942	Positive regulation of cell death
	LMD1A	GO:0060548	Negative regulation of cell death

(1), Sola (1) and Vingi (1). Remarkably, 40 out of these 42 unigenes showed upregulation, and only two (one Daphne and one Mariner/Tc1 sequences) showed downregulation.

Discussion

B chromosomes are intragenomic parasites being in intimate contact with host genes within the same nuclei, for which reason they have the chance to elicit a true transcriptomic arms race with A chromosomes. Our present results show that this is the case. Recent transcriptomic analyses in interspecific parasitisms have unveiled gene expression changes mostly focused on parasite attack and host resistance (for instance, see Daurelio et al., 2011; Foth et al., 2014; Prasopdee et al., 2014; Robledo et al., 2014; Nishimura et al., 2015). The transcriptomic signature of an intragenomic parasitism, shown here for B-carrying ovaries, shows that the presence of a parasitic chromosome evokes a response in the host genome implying gene expression changes associated with the presence of additional DNA (nucleotypic effects), sensing cellular stress, detoxifying and immune defense, ovary function and regulation of gene expression. These changes can be viewed as a manifestation of a molecular arms race between the parasitic and host counterparts of a same genome, which parallels population dynamics pathways (Camacho et al., 1997).

The presence of additional DNA in the form of B chromosomes can passively derive in a series of sequence-independent physiological changes

known as nucleotypic effects (Bennett, 1971). In several cases, B chromosomes have been shown to increase cell volume and slow cell division (Randolph, 1941; Müntzing and Akdik, 1948; Commoner, 1964; John and Jones, 1970; Evans and Rees, 1971; Cavalier-Smith, 1980; Nurse, 1985). As remarked by Gregory (2000), "the addition of only one or two B chromosomes appears not to have a noticeable effect on cell size (John and Jones, 1970), perhaps indicating that cells are able to compensate for the effects of these selfish elements". In *E. plorans*, cell or nucleus size has not been measured, but body size, which is frequently positively correlated with cell size (Gregory, 2002), has been measured twice (Camacho et al., 1980; Martín-Alganza et al., 1997) and no association with the number of B chromosomes was found. It is thus conceivable that some of the observed gene expression changes are actually revealing how B-carrying cells cope with the presence of this extra DNA in terms of signal and molecule transport, intercellular communication and trafficking, membranes, ion transport, signal transduction mechanisms, intracellular trafficking, secretion, and vesicular transport.

For instance, the observed downregulation of Phosphatidylinositol 3-kinase 2 (*PI3K2*) in B-carrying *E. plorans* females could be a response to maintain cell size in spite of the extra DNA added by the B chromosome, since this gene is involved in the regulation of cell morphogenesis by promoting cell growth (Leevers et al., 1996; Weinkove et al., 1999), and also in the regulation of exaggerated trait growth in insects (Lavine et al., 2015). Likewise, Neural-cadherin (CADN), Neurexin-4 (*NRX4*), Synaptic vesicle glycoprotein 2A (*SV2A*) and Androgen-induced gene 1 protein (*AIG1*) show functions related with cell

growth, cell adhesion and secretion, transport and cell communication (Harden, 2002; Krizaj et al., 2002; Halbleib and Nelson, 2006; Humbert et al., 2003; Sudhof, 2004; Yi et al., 2008; Kurusu et al., 2008), and their up-regulation could represent a response of the host genome to cope with the presence of the 600 Mb extra DNA per one B chromosome (Ruiz-Ruano et al., 2011). It is conceivable that the presence of the extra B-DNA elicits some changes in how cells communicate themselves, and this is reinforced by the observed expression changes in genes with functions related with membranes and ion transport, such as the Zinc transporter *ZIP14* (*S39AE*) and Pro-interleukin-16 (*IL16*), which play a role in parasitic infections (Pappu et al., 2011), in strong parallelism with other types of parasitism (for instance, see Lemaire et al., 2013; Tako et al., 2013; Aufauvre et al., 2014; Foth et al., 2014; Robledo et al., 2014; Telleria et al., 2014; Yang et al., 2014; Nishimura et al., 2015).

The presence of extra DNA to replicate can explain the up-regulation of Topoisomerase II (*TOP2*) in the ovaries of B-carrying adult females, since this protein is involved in DNA conformational changes during replication (Saavedra, 1990). It has been frequently argued that B chromosome presence represents a burden to the host cell in terms of DNA replication (see Evans et al., 1972), but we show here the first evidence at this respect. The fact that *TOP2* is also involved in chromosome condensation (Coelho et al., 2003; Baxter and Aragón, 2012) opens an interesting prospect for future research, especially bearing in mind the presence of active genes for the condensin complex subunit CAP-G in the B chromosome (Navarro-Domínguez et al., in preparation; see chapter 4).

It was highly remarkable the abundance of gene expression changes

dealing with chemical defense and detoxification observed in presence of the B chromosome in *E. plorans*. Metazoan genomes contain many genes involved in responses to environmental stressors. Chemical defense genes include cytochromes P450 and other oxidases, various conjugating enzymes, ATP-dependent efflux transporters, oxidative detoxification proteins, and transcription factors that regulate these genes which, as a whole, account for more than 400 genes in the sea urchin genome (Goldstone et al., 2006). Cytochrome P450 4C1 (*CP4C1*) is a chemical defense gene involved in the metabolism of insect hormones and in the breakdown of synthetic insecticides, thus being a defensive factor against stress (Goldstone et al., 2006). P450 up-regulation in B-carrying *E. plorans* females could thus be a response of the host genome against parasitic chromosomes. However, a related gene, Cytochrome P450 6A1 (*CP6A1*), involved in the same functions was downregulated in the same females, so it is impossible to know the net effect of both changes.

Other gene expression changes dealing with chemical defense were those for Myrosinase 1 (*MYRO1*), which is involved in chemical defense by crucifer plants but is also present in some herbivores like aphids (Kazana et al., 2007) or beetles (Beran et al., 2014), Liver carboxylesterase 1 (*EST1*), which is involved in the detoxification of xenobiotics, alcohol metabolic process and regulation of cell morphogenesis (Matsumoto et al., 2014), Multiple inositol polyphosphate phosphatase 1 (*MINP1*), which is located in the endoplasmic reticulum and is involved in a wide array of biological and pathophysiological responses, including tumorigenesis, invasion and metastasis, as well as defense signaling in plants (Chi et al., 2000; Windhorst et al., 2013; Kilaparty et al., 2014), or Glutathione S-transferase 1-1 (*GSTT1*), which has DDT dehydrochlorinase activity

(Lumjuan et al., 2005). These four gene expression changes, all being up-regulated in B carrying females, could constitute part of the resistance response to the parasitic chromosome. However, the down-regulation of UDP-glucuronosyl and UDP-glucosyl transferase (*UDPE*), which plays a central role in the detoxication and elimination of a wide range of endogenous and exogenous compounds (Burchell and Coughtrie, 1989), is more difficult to interpret.

It has been suggested that gene expression in host–parasite interactions is likely to evolve toward greater immunological surveillance and reduced parasite conspicuousness (Nuismer and Otto, 2005), in a Red Queen dynamics (Barribeau et al., 2014). For instance, endoparasites such as Trypanosoma, Nosema and Trichuris elicit host responses increasing the expression of genes related with immunity, stress tolerance and protein translation (Telleria et al., 2014; Foth et al., 2014; Aufauvre et al., 2014). Remarkably, some of the gene expression changes found in B-carrying ovaries can be viewed in a similar way. This is the case for four up-regulated gene products: the Roadkill protein (*RDX* gene), which promotes protein ubiquitination and degradation by the proteasome (Liu et al., 2009), N-alpha-acetyltransferase 38-A, NatC auxiliary subunit (*LMD1A*), which is involved in protein degradation (Kim et al., 2014), Iron/zinc purple acid phosphatase-like protein (*PAPL*), which has acid phosphatase activity, metal ion binding and dephosphorylation, and RING finger protein 141 (*RN141*) which may simultaneously bind ubiquitination enzymes and their substrates (Lee et al., 2014). However, the down-regulation of the *RIN2* gene, which codes for an E3 ubiquitin protein ligase involved in protein polyubiquitination, response to endoplasmic reticulum (ER) stress, ER-nucleus signaling pathway

and response to unfolded proteins (Chow et al., 2015; Chaubey et al., 2014), in B-carrying ovaries is difficult to explain. Although we cannot know the precise mechanism of action, it is tempting to speculate that these expression changes regarding ubiquitinization and endoplasmic reticulum stress might constitute a host genome reponse to the presence of topologically anomalous polypeptides derived from B chromosome gene or pseudogene transcripts (see Navarro-Dominguez et al., submitted; see chapter 3). The observed changes for ankyrins 1-3 (*ANK1*, *ANK2* and *ANK3*), which are involved in protein localization to cell surface, and intracellular signal transduction (Goellner and Aberle, 2012), could have also something to do with control of biological quality against B-derived defective polypeptides. However, B chromosome effects are also contradictory since *ANK1* and *ANK2* are down-regulated whereas *ANK3* is up-regulated (see Fig. 3).

Several DEGs were involved in lipid transport and metabolism, with conceivable consequences on egg fertility. Six of them were up-regulated in B-carrying females: Alpha-tocopherol transfer protein-like (*TTPAL*), which is presumed to function in the intracellular transport of alpha-tocopherol (vitamin E), thus influencing the amount of eggs laid (Zwołińska-Śniatałowa, 1976), Multidrug resistance-associated protein 9 (*MRP9*), which plays a biologically important role as membrane transporter and ion channel modulator (Higgins, 1992), Seipin (*BSCL2*), a regulator of lipid catabolism being essential for adipocyte differentiation and necessary for correct lipid storage, as well as three other genes involved in lipid transport and regulation of fat cell differentiation, namely Nose resistant to fluoxetine protein 6 (*NRF6*), border follicle cell migration (*MYSA*) and phosphatidylinositol phosphatase (*PTPRQ*).

However, another five DEGs involved in lipid metabolism and transport, regulation of oocyte development and contraction of ovarian muscle, were down-regulated perhaps explaining the lower egg fertility shown by B-carrying females (Zurita et al., 1998; Muñoz et al., 1998). This is the case of Retinaldehyde-binding protein 1 (*RLBP1*), Choline transporter-like protein 4 (*CTL4*), Carbonyl reductase family member 4 (*CBR4*), Synaptotagmin 1 (*SY65*) and Troponin C, isoform 3 (*TNNC3*). Taken together, these gene expression changes might reflect the harmful effects of the parasitic chromosome on yolk production in the eggs and molecule transport between cells.

Four other DGEs related with muscle function were up-regulated in B-carrying females: Muscle M-line assembly protein unc-89 (*UNC89*), (Gautel, 2011; Katzemich et al., 2012), Twitchin (*UNC22*) (Riddell et al., 2014), Titin (*TITIN*) (Fabian et al., 2007) and Myosin heavy-chain (*MYSA*) (Polyák et al., 2003). Interestingly, the two latter genes also play a role during mitosis (Machado and Andrew, 2000; Fabian and Forer, 2007; Fabian et al., 2007). Collectively, these four gene expression changes could be a consequence of higher effort by B-carrying females in muscle function, probably related with ovary function, and/or higher costs of building mitotic spindles in ovaries due to the higher number of chromosomes to move apart during cell division.

Among the most interesting gene expression changes observed in ovaries carrying the parasitic chromosome were those associated with the regulation of gene expression. Two DEGs were involved in the biogenesis of the 60S ribosomal subunit: the L26 protein (*RL26*), which is a structural constituent of this subunit, and the Nucleolar GTP-binding protein 1 (*NOG1*) which has a regulatory role in its biogenesis. The

former is up-regulated and the latter down-regulated in B-carrying females.

In addition, Histone-lysine N-methyltransferase SETD1B-A (*SE1BA*) is a histone methyltransferase that specifically methylates Lys-4 of histone H3 providing a specific tag for epigenetic transcriptional activation (Bernstein et al., 2002; Santos-Rosa et al., 2002), which is down-regulated B-carrying ovarioles. It is thus tempting to suggest that down-regulation of *H3K4* methyltransferase complex in B-carrying females of *E. plorans* might have something to do with the observed transcription of B chromosome genes (see above).

Three up-regulated DEGs might reflect the extra transcription effort due to the higher gene expression observed in B-carrying females, with double number of up- than down-regulations for protein-coding genes and an extraordinary up-regulation for many TEs. These were the DNA-directed RNA polymerase II subunit RPB11 (*RPB11*), which is a component of RNA polymerase II itself, Mediator of RNA polymerase II transcription subunit 4 (*MED4*), which plays a role in the regulation of transcription of nearly all RNA polymerase II-dependent genes, and Ubiquitin-like domain-containing CTD phosphatase 1 (*UBCP1*), which plays a role in the regulation of phosphorilation state of RNA polimerase II C-terminal domain (Zheng et al., 2005). In consistency with these three up-regulations, we observed down-regulation for Protein Dr1 homolog (*NC2B*), which is a negative regulator of transcription by controlling the rate of RNA polymerase II transcription. Taken together, the changes observed for these four genes would reflect an increase in transcription effort probably due to the fact that the B chromosome is not completely silenced (Navarro-Domínguez et al., in preparation; see chapters 3 and

4).

Finally, B-carrying females showed down-regulation of Dicer 1 (*DCR1*), an RNA polymerase III being essential for RNA interference (RNAi) and microRNA (miRNA) gene silencing. Hence, down-regulation of Dicer 1 implies higher susceptibility to pathogens (Fablet, 2014). Remarkably, the down-regulation of Dicer 1 in B-carrying females of *E. plorans* is associated with an explosive up-regulation for many transposable elements, in high consistency with the role of RNA interference as an important defense against viruses and transposable elements (Obbard et al., 2009).

Little is known on how Dicer 1 is regulated (Levy et al., 2010). Mouse primary oocytes induced to be deficient in Dicer 1, are able to reach full size and be ovulated, but they have meiotic defects, such as first polar body extrusion failing (Murchison et al., 2007; Tang et al., 2007), due to the presence of DICER1-dependent factors in the ooplasm of primary oocytes which are crucial for chromosome segregation and meiotic maturation (Mattiske et al., 2009). In B-carrying females of *E. plorans*, however, no meiotic defects have been reported in oocytes (Henriques-Gil et al., 1986; Cano et al., 1987; Henriques-Gil et al., 1989), suggesting that Dicer 1 down-regulation does not negatively affect grasshopper oogenesis. Dicer 1 down-regulation in B-carrying females could keep some relationship with the molecular stress generated by B chromosome transcripts (Navarro-Dominguez et al., submitted; see chapter 3), as Dicer 1 down-regulation has been reported during hypoxia (Ho et al., 2012) and other types of stress (Wiesen and Tomasi, 2009), and promotes cellular senescence (Zhao et al., 2015) and cancer (Kitagawa et al., 2013). At first sight, Dicer 1 down-regulation in

B-carrying females might appear puzzling since RNAi is necessary for heterochromatinization (Hall et al., 2002), and B chromosomes in *E. plorans* are mostly heterochromatic. However, it is conceivable that Dicer 1 down-regulation is achieved by the B chromosome itself to resist silencing by the host genome, since about half of the protein-coding genes located in the B chromosome were actively transcribed.

Taken together, the gene expression changes observed here in B-carrying ovaries of *E. plorans* appear to constitute a logical response of the host genome to counteract gene expression of B chromosome genes and pseudogenes reported by Navarro-Domínguez et al., in preparation (see chapter 3). It is remarkable that all the active genes and pseudogenes in the B chromosome, but almost none of the observed gene expression changes in the host genome, code for cell division regulation. However, most changes in host genome gene expression, had to do with chemical defense and detoxification, protein modification and response to stress, ovary function and regulation of gene expression. suggesting an intimate host-parasite crosstalk during crucial steps of gene expression and protein function. As a whole, our present results illuminate a broad spectrum of future molecular research on this evolutionarily interesting intragenomic parasitism, and the common molecular arena for the two counterparts makes this a good model to analyze gene regulation pathways.

Acknowledgements

We thank Rubén Martín-Blázquez and Mohammed Bakkali for their collaboration on the Illumina transcriptome reads sequencing and first assembly. We also thank Marco Pellino for his help on the microarray hybridization.

This study was funded by the Spanish Ministerio de Economía y Competitividad (CGL2015-70750-P) and Plan Andaluz de Investigación (CVI-6649). Navarro-Domínguez was supported by a FPI fellowship from the Spanish Ministerio de Ciencia e Innovación (BES-2010-030016).

References

Akbari, O. S., Antoshechkin, I., Hay, B. A., and Ferree, P. M. (2013). Transcriptome profiling of *Nasonia vitripennis* testis reveals novel transcripts expressed from the selfish B chromosome, paternal sex ratio. *G3: Genes| Genomes| Genetics*, 3(9):1597–1605.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald,

M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25(1):25–29.

Aufauvre, J., Misme-Aucouturier, B., Viguès, B., Texier, C., Delbac, F., and Blot, N. (2014). Transcriptome analyses of the honeybee response to *Nosema ceranae* and insecticides. *PloS One*, 9(3):e91686.

Bakkali, M., Manrique-Poyato, M. I., López-León, M. D., Perfectti, F., Cabrero, J., and Camacho, J. P. M. (2010). Effects of B chromosomes on egg fertility and clutch size in the grasshopper *Eyprepocnemis plorans*. *Journal of Orthoptera Research*, 19(2):197–203.

Banaei-Moghaddam, A. M., Meier, K., Karimi-Ashtiyani, R., and Houben, A. (2013). Formation and expression of pseudogenes on the B chromosome of rye. *The Plant Cell*, 25(7):2536–2544.

Barribeau, S. M., Sadd, B. M., du Plessis, L., and Schmid-Hempel, P. (2014). Gene expression differences underlying genotype-by-genotype specificity in a host-parasite system. *Proceedings of the National Academy of Sciences of the United States of America*, 111(9):3496–3501.

Baxter, J. and Aragón, L. (2012). A model for chromosome condensation based on the interplay between condensin and topoisomerase II. *Trends in genetics: TIG*, 28(3):110–117.

Bennett, M. D. (1971). The Duration of Meiosis. *Proceedings of the Royal Society of London B: Biological Sciences*, 178(1052):277–299.

Beran, F., Pauchet, Y., Kunert, G., Reichelt, M., Wielsch, N., Vogel, H., Reinecke, A., Svatoš, A., Mewis, I., Schmid, D., Ramasamy, S., Ulrichs, C., Hansson, B. S., Gershenzon, J., and Heckel, D. G. (2014). Phyllotreta striolata flea beetles use host plant defense compounds to create their own glucosinolate-myrosinase system. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20):7349–7354.

Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T., and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(13):8695–8700.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, 30(15):2114–2120.

Burchell, B. and Coughtrie, M. W. (1989). UDP-glucuronosyltransferases. *Pharmacology and Therapeutics*, 43(2):261–289.

Cabrero, J., Alché, J. D., and Camacho, J. P. M. (1987). Effects of B chromosomes on the activity of nucleolar organizer regions in the grasshopper *Eyprepocnemis plorans*: activation of a latent nucleolar organizer region on a B chromosome fused to an autosome. *Genome*, 29(1):116–121.

Cabrero, J., López-León, M. D., Ruiz-Estévez, M., Gómez, R., Petitpierre, E., Rufas, J. S., Massa, B., Kamel Ben Halima, M., and Camacho, J. P. M. (2014). B1 was the ancestor B chromosome variant in the western Mediterranean area in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 142(1):54–58.

Cabrero, J., Manrique-Poyato, M. I., and Camacho, J. P. M. (2006). Detection of B chromosomes in interphase hemolymph nuclei from living specimens of the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 114(1):66–69.

Camacho, J. P. M. (2005). B chromosomes. In Gregory, T. R., editor, *The Evolution of the Genome*, pages 223–286. Academic Press.

Camacho, J. P. M., Bakkali, M., Corral, J. M., Cabrero, J., López-León, M. D., Aranda, I., Martín-Alganza, A., and Perfectti, F. (2002). Host recombination is dependent on the degree of parasitism. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1505):2173–2177.

Camacho, J. P. M., Carballo, A. R., and Cabrero, J. (1980). The B-chromosome system of the grasshopper *Eyprepocnemis plorans* subsp. plorans (Charpentier). *Chromosoma*, 80(2):163–176.

Camacho, J. P. M., López-León, M. D., Pardo, M. C., Cabrero, J., and Shaw, M. W. (1997). Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper *Eyprepocnemis plorans*. *The American Naturalist*, 149(6):1030–1050.

Camacho, J. P. M., Sharbel, T. F., and Beukeboom, L. W. (2000). B-chromosome evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 355(1394):163–178.

Cano, M. I., Jones, G. H., and Santos, J. L. (1987). Sex differences in chiasma frequency and distribution in natural populations of *Eyprepocnemis plorans* containing B-chromosomes. *Heredity*, 59(2):237–243.

Carchilan, M., Kumke, K., Mikolajewski, S., and Houben, A. (2009). Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma*, 118(5):607–616.

Carvalho, B. S. and Irizarry, R. A. (2010). A framework for oligonucleotide microarray preprocessing. *Bioinformatics (Oxford, England)*, 26(19):2363–2367.

Cavalier-Smith, T. (1980). r- and K-tactics in the evolution of protist developmental systems: cell and genome size, phenotype diversifying selection, and cell cycle patterns. *Biosystems*, 12(1-2):43–59.

Chaubey, S., Grover, M., and Tatu, U. (2014). Endoplasmic reticulum stress triggers gametocytogenesis in the malaria parasite. *Journal of Biological Chemistry*, 289(24):16662–16674.

Chi, H., Yang, X., Kingsley, P. D., O'Keefe, R. J., Puzas, J. E., Rosier, R. N., Shears, S. B., and Reynolds, P. R. (2000). Targeted deletion of Minpp1 provides new insight into the activity of multiple inositol polyphosphate phosphatase in vivo. *Molecular and Cellular Biology*, 20(17):6496–6507.

Chow, C. Y., Wang, X., Riccardi, D., Wolfner, M. F., and Clark, A. G. (2015). The genetic architecture of the genome-wide transcriptional response to ER stress in the mouse. *PLoS Genetics*, 11(2):e1004924.

Coelho, P. A., Queiroz-Machado, J., and Sunkel, C. E. (2003). Condensin-dependent localisation of topoisomerase II to an axial chromosomal structure is required for sister chromatid resolution during mitosis. *Journal of Cell Science*, 116(Pt 23):4763–4776.

Commoner, B. (1964). Roles Of Deoxyribonucleic Acid in Inheritance. *Nature*, 202(4936):960–968.

Daurelio, L. D., Petrocelli, S., Blanco, F., Holuigue, L., Ottado, J., and Orellano, E. G. (2011). Transcriptome analysis reveals novel genes involved in nonhost response to bacterial infection in tobacco. *Journal of Plant Physiology*, 168(4):382–391.

Evans, G. M. and Rees, H. (1971). Mitotic Cycles in Dicotyledons and Monocotyledons. *Nature*, 233(5318):350–351.

Evans, G. M., Rees, H., Snell, C. L., and Sun, S. (1972). The relationship between nuclear DNA amount and the duration of the mitotic cycle. *Chromosomes today*, 3(1):24–31.

Fabian, L. and Forer, A. (2007). Possible roles of actin and myosin during anaphase chromosome movements in locust spermatocytes. *Protoplasma*, 231(3-4):201–213.

Fabian, L., Xia, X., Venkitaramani, D. V., Johansen, K. M., Johansen, J., Andrew, D. J., and Forer, A. (2007). Titin in insect spermatocyte spindle fibers associates with microtubules, actin, myosin and the matrix proteins skeletor, megator and chromator. *Journal of Cell Science*, 120(13):2190–2204.

Fablet, M. (2014). Host control of insect endogenous retroviruses: small RNA silencing and immune response. *Viruses*, 6(11):4447–4464.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., and Mistry, J. (2013). Pfam: the protein families database. *Nucleic Acids Research*, page gkt1223.

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*, 39(Web Server issue):W29–37.

Foth, B. J., Tsai, I. J., Reid, A. J., Bancroft, A. J., Nichol, S., Tracey, A., Holroyd, N., Cotton, J. A., Stanley, E. J., and Zarowiecki, M. (2014). Whipworm genome and dual-species transcriptome analyses provide molecular insights into an intimate host-parasite interaction. *Nature Genetics*, 46(7):693–700.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23):3150–3152.

Gautel, M. (2011). The sarcomeric cytoskeleton: who picks up the strain? *Current Opinion in Cell Biology*, 23(1):39–46.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y. H., and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80.

Goellner, B. and Aberle, H. (2012). The synaptic cytoskeleton in development and disease. *Developmental Neurobiology*, 72(1):111–125.

Goldstone, J. V., Hamdoun, A., Cole, B. J., Howard-Ashby, M., Nebert, D. W., Scally, M., Dean, M., Epel, D., Hahn, M. E., and Stegeman, J. J. (2006). The chemical defensome: environmental sensing and response genes in the Strongylocentrotus purpuratus genome. *Developmental biology*, 300(1):366–384.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7):644–652.

Graphodatsky, A. S., Kukekova, A. V., Yudkin, D. V., Trifonov, V. A., Vorobieva, N. V., Beklemisheva, V. R., Perelman, P. L., Graphodatskaya, D. A., Trut, L. N., Yang, F., Ferguson-Smith, M. A., Acland, G. M., and Aguirre, G. D. (2005). The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosome Research*, 13(2):113–122.

Gregory, T. R. (2000). Nucleotypic effects without nuclei: genome size and erythrocyte size in mammals. *Genome / National Research Council Canada = Génome / Conseil National De Recherches Canada*, 43(5):895–901.

Gregory, T. R. (2002). A bird's-eye view of the C-value enigma: genome size, cell size, and metabolic rate in the class aves. *Evolution; International Journal of Organic Evolution*, 56(1):121–130.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N.,

Henschel, R., Leduc, R. D., Friedman, N., and Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8):1494–1512.

Halbleib, J. M. and Nelson, W. J. (2006). Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes and Development*, 20(23):3199–3214.

Hall, I. M., Shankaranarayana, G. D., Noma, K.-i., Ayoub, N., Cohen, A., and Grewal, S. I. (2002). Establishment and maintenance of a heterochromatin domain. *Science*, 297(5590):2232–2237.

Harden, N. (2002). Signaling pathways directing the movement and fusion of epithelial sheets: lessons from dorsal closure in Drosophila. *Differentiation*, 70(4-5):181–203.

Henriques-Gil, N. and Arana, P. (1990). Origin and substitution of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Evolution*, 44(3):747–753.

Henriques-Gil, N., Cano, M. I., and Santos, J. L. (1989). Sex-dependent meiotic behaviour of B chromosomes in the grasshopper *Eyprepocnemis plorans*. *Heredity*, 62:217–221.

Henriques-Gil, N., Jones, G. H., Cano, M. I., Arana, P., and Santos, J. L. (1986). Female meiosis during oocyte maturation in *Eyprepocnemis plorans* (Orthoptera: Acrididae). *Canadian Journal of Genetics and Cytology*, 28(1):84–87.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annual review of cell biology*, 8(1):67–113.

Ho, J. J. D., Metcalf, J. L., Yan, M. S., Turgeon, P. J., Wang, J. J., Chalsev, M., Petruzziello-Pellegrini, T. N., Tsui, A. K. Y., He, J. Z., Dhamko, H., Man, H. S. J., Robb, G. B., Teh, B. T., Ohh, M., and Marsden, P. A. (2012). Functional importance of Dicer protein in the adaptive cellular response to hypoxia. *The Journal of Biological Chemistry*, 287(34):29003–29020.

Huang, W., Du, Y., Zhao, X., and Jin, W. (2016). B chromosome contains active genes and impacts the transcription of A chromosomes in maize (*Zea mays* L.). *BMC Plant Biology*, 16:88.

Humbert, P., Russell, S., and Richardson, H. (2003). Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 25(6):542–553.

John, P. C. L. and Jones, R. N. (1970). Molecular heterogeneity of soluble proteins and histones in relationship to the presence of B-chromosomes in rye. *Experimental Cell Research*, 63(2-3):271–276.

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, 110(1-4):462–467.

Katzemich, A., Kreisköther, N., Alexandrovich, A., Elliott, C., Schöck, F., Leonard, K., Sparrow, J., and Bullard, B. (2012). The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of Drosophila. *Journal of Cell Science*, 125(14):3367–3379.

Kazana, E., Pope, T. W., Tibbles, L., Bridges, M., Pickett, J. A., Bones, A. M., Powell, G., and Rossiter, J. T. (2007). The cabbage aphid: a walking mustard oil bomb. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1623):2271–2277.

Kilaparty, S. P., Singh, A., Baltosser, W. H., and Ali, N. (2014). Computational analysis reveals a successive adaptation of multiple inositol polyphosphate phosphatase 1 in higher organisms through evolution. *Evolutionary Bioinformatics Online*, 10:239–250.

Kim, H.-K., Kim, R.-R., Oh, J.-H., Cho, H., Varshavsky, A., and Hwang, C.-S. (2014). The N-terminal methionine of cellular proteins as a degradation signal. *Cell*, 156(1-2):158–169.

Kitagawa, N., Ojima, H., Shirakihara, T., Shimizu, H., Kokubu, A., Urushidate, T., Totoki, Y., Kosuge, T., Miyagawa, S., and Shibata, T. (2013). Downregulation of the microRNA biogenesis components and its association with poor prognosis in hepatocellular carcinoma. *Cancer Science*, 104(5):543–551.

Krizaj, D., Demarco, S. J., Johnson, J., Strehler, E. E., and Copenhagen, D. R. (2002). Cell-specific expression of plasma membrane calcium ATPase isoforms in retinal neurons. *Journal of Comparative Neurology*, 451(1):1–21.

Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., and Zinn, K. (2008). A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection. *Neuron*, 59(6):972–985.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3):R25.

Lavine, L., Gotoh, H., Brent, C. S., Dworkin, I., and Emlen, D. J. (2015). Exaggerated trait growth in insects. *Annual Review of Entomology*, 60:453–472.

Leach, C. R., Houben, A., Field, B., Pistrick, K., Demidov, D., and Timmis, J. N. (2005). Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. *Genetics*, 171(1):269–278.

Lee, Y.-I., Giovinazzo, D., Kang, H. C., Lee, Y., Jeong, J. S., Doulias, P.-T., Xie, Z., Hu, J., Ghasemi, M., Ischiropoulos, H., Qian, J., Zhu, H., Blackshaw, S., Dawson, V. L., and Dawson, T. M. (2014). Protein microarray characterization of the S-nitrosoproteome. *Molecular and Cellular Proteomics: MCP*, 13(1):63–72.

Leevers, S. J., Weinkove, D., MacDougall, L. K., Hafen, E., and Waterfield, M. D. (1996). The Drosophila phosphoinositide 3-kinase Dp110 promotes cell growth. *The EMBO journal*, 15(23):6584.

Lemaire, J., Mkannez, G., Guerfali, F. Z., Gustin, C., Attia, H., Sghaier, R. M., Dellagi, K., Laouini, D., Renard, P., and Sysco-Consortium (2013). MicroRNA expression profile in human macrophages in response to Leishmania major infection. *PLoS Neglected Tropical Diseases*, 7(10):e2478.

Levy, C., Khaled, M., Robinson, K. C., Veguilla, R. A., Chen, P.-H., Yokoyama, S., Makino, E., Lu, J., Larue, L., and Beermann, F. (2010). Lineage-specific transcriptional regulation of DICER by MITF in melanocytes. *Cell*, 141(6):994–1005.

Li, B. and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12:323.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754–1760.

Liu, J., Ghanim, M., Xue, L., Brown, C. D., Iossifov, I., Angeletti, C., Hua, S., Nègre, N., Ludwig, M., Stricker, T., Al-Ahmadie, H. A., Tretiakova, M., Camp, R. L., Perera-Alberto, M., Rimm, D. L., Xu, T., Rzhetsky, A., and White, K. P. (2009).

Analysis of Drosophila segmentation network identifies a JNK pathway factor overexpressed in kidney cancer. *Science (New York, N.Y.)*, 323(5918):1218–1222.

López-León, M. D., Cabrero, J., and Camacho, J. P. M. (1995). Changes in DNA methylation during development in the B chromosome NOR of the grasshopper *Eyprepocnemis plorans*. *Heredity*, 74(3):296–302.

López-León, M. D., Cabrero, J., Dzyubenko, V. V., Bugrov, A. G., Karamysheva, T. V., Rubtsov, N. B., and Camacho, J. P. M. (2008). Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of the grasshopper *Eyprepocnemis plorans plorans*. *Cytogenetic and Genome Research*, 121(3-4):260–265.

Lumjuan, N., McCarroll, L., Prapanthadara, L.-a., Hemingway, J., and Ranson, H. (2005). Elevated activity of an Epsilon class glutathione transferase confers DDT resistance in the dengue vector, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology*, 35(8):861–871.

Machado, C. and Andrew, D. J. (2000). D-Titin: a giant protein with dual roles in chromosomes and muscles. *The Journal of Cell Biology*, 151(3):639–652.

Marçais, G. and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics (Oxford, England)*, 27(6):764–770.

Martín-Alganza, A., Cabrero, J., López-León, M. D., Perfectti, F., and Camacho, J. P. M. (1997). Supernumerary heterochromatin does not affect several morphological and physiological traits in the grasshopper *Eyprepocnemis plorans*. *Hereditas*, 126(2):187–189.

Martis, M. M., Klemme, S., Banaei-Moghaddam, A. M., Blattner, F. R., Macas, J., Schmutzer, T., Scholz, U., Gundlach, H., Wicker, T., Šimková, H., Novák, P., Neumann, P., Kubaláková, M., Bauer, E., Haseneyer, G., Fuchs, J., Doležel, J., Stein, N., Mayer, K. F. X., and Houben, A. (2012). Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences*, 109(33):13343–13346.

Matsumoto, Y., Suetsugu, Y., Nakamura, M., and Hattori, M. (2014). Transcriptome

analysis of the salivary glands of *Nephotettix cincticeps* (Uhler). *Journal of Insect Physiology*, 71:170–176.

Mattiske, D. M., Han, L., and Mann, J. R. (2009). Meiotic maturation failure induced by DICER1 deficiency is derived from primary oocyte ooplasm. *Reproduction*, 137(4):625–632.

Muñoz, E., Perfectti, F., Martín-Alganza, Á., and Camacho, J. P. M. (1998). Parallel effects of a B chromosome and a mite that decrease female fitness in the grasshopper *Eyprepocnemis plorans*. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1408):1903.

Muñoz-Pajares, A. J., Martínez-Rodríguez, L., Teruel, M., Cabrero, J., Camacho, J. P. M., and Perfectti, F. (2011). A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics*, 187(3):853–863.

Müntzing, A. and Akdik, S. (1948). Cytological disturbances in the first inbred generations of rye. *Hereditas*, 34(4):485–509.

Murchison, E. P., Stein, P., Xuan, Z., Pan, H., Zhang, M. Q., Schultz, R. M., and Hannon, G. J. (2007). Critical roles for Dicer in the female germline. *Genes & Development*, 21(6):682–693.

Navarro-Domínguez, B. (2016). *E. plorans* transcriptome files. *Figshare.* DOI: http://dx.doi.org/10.6084/m9.figshare.3408580.v3.

Nishimura, M., Tanaka, S., Ihara, F., Muroi, Y., Yamagishi, J., Furuoka, H., Suzuki, Y., and Nishikawa, Y. (2015). Transcriptome and histopathological changes in mouse brain infected with *Neospora caninum*. *Scientific Reports*, 5:7936.

Nuismer, S. L. and Otto, S. P. (2005). Host-parasite interactions and the evolution of gene expression. *PLoS Biology*, 3(7):e203.

Nurse, P. (1985). The genetic control of cell volume. In Cavalier-Smith, T., editor, *The evolution of genome size*, pages 185–196.

Obbard, D. J., Gordon, K. H., Buck, A. H., and Jiggins, F. M. (2009). The evolution of RNAi as a defence against viruses and transposable elements. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1513):99–115.

Pappu, R., Ramirez-Carrozzi, V., and Sambandam, A. (2011). The interleukin-17 cytokine family: critical players in host defence and inflammatory diseases. *Immunology*, 134(1):8–16.

Polyák, E., Standiford, D. M., Yakopson, V., Emerson, C. P., and Franzini-Armstrong, C. (2003). Contribution of myosin rod protein to the structural organization of adult and embryonic muscles in Drosophila. *Journal of Molecular Biology*, 331(5):1077–1091.

Prasopdee, S., Sotillo, J., Tesana, S., Laha, T., Kulsantiwong, J., Nolan, M. J., Loukas, A., and Cantacessi, C. (2014). RNA-Seq reveals infection-induced gene expression changes in the snail intermediate host of the carcinogenic liver fluke, *Opisthorchis viverrini*. *PLoS Neglected Tropical Diseases*, 8(3):e2765.

Randolph, L. F. (1941). Genetic characteristics of the B chromosomes in maize. *Genetics*, 26(6):608–631.

Riddell, C. E., Garces, J. D. L., Adams, S., Barribeau, S. M., Twell, D., and Mallon, E. B. (2014). Differential gene expression and alternative splicing in insect immune specificity. *BMC Genomics*, 15(1):1.

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., Mungall, K., Lee, S., Okada, H. M., and Qian, J. Q. (2010). De novo assembly and analysis of RNA-seq data. *Nature Methods*, 7(11):909–912.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, 26(1):139–140.

Robledo, D., Ronza, P., Harrison, P. W., Losada, A. P., Bermúdez, R., Pardo, B. G., Redondo, M. J., Sitjà-Bobadilla, A., Quiroga, M. I., and Martínez, P. (2014). RNA-seq analysis reveals significant transcriptome changes in turbot (*Scophthalmus maximus*) suffering severe enteromyxosis. *BMC Genomics*, 15:1149.

Ruiz-Estévez, M., Badisco, L., Broeck, J. V., Perfectti, F., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2014). B chromosomes showing active ribosomal RNA genes contribute insignificant amounts of rRNA in the grasshopper *Eyprepocnemis plorans*. *Molecular Genetics and Genomics*, 289(6):1209–1216.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2012). B-chromosome ribosomal DNA is functional in the grasshopper *Eyprepocnemis plorans*. *PLoS One*, 7(5):e36600.

Ruiz-Estévez, M., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2013). Ribosomal DNA is active in different B chromosome variants of the grasshopper *Eyprepocnemis plorans*. *Genetica*, 141(7-9):337–345.

Ruiz-Ruano, F. J., Ruiz-Estévez, M., Rodríguez-Pérez, J., López-Pino, J. L., Cabrero, J., and Camacho, J. P. M. (2011). DNA amount of X and B chromosomes in the grasshoppers *Eyprepocnemis plorans* and *Locusta migratoria*. *Cytogenetic and Genome Research*, 134(2):120–126.

Saavedra, R. A. (1990). Environmental stimuli and transcriptional activity generate transient changes in DNA torsional tension. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, 12(3):125–128.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*, 419(6905):407–411.

Schmieder, R. and Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PloS One*, 6(3):e17288.

Smit, A. F., Hubley, R., and Green, P. (1996). *RepeatMasker Open-3.0*. Available at http://www.repeatmasker.org/.

Smyth, G. K. (2005). Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor*, pages 397–420. Springer.

Soumana, I. H., Tchicaya, B., Simo, G., and Geiger, A. (2014). Comparative gene expression of Wigglesworthia inhabiting non-infected and Trypanosoma brucei gambiense-infected *Glossina palpalis gambiensis* flies. *Frontiers in microbiology*, 5.

Sudhof, T. C. (2004). The synaptic vesicle cycle. *Annual Review of Neuroscience*, 27:509.

Tako, E. A., Hassimi, M. F., Li, E., and Singer, S. M. (2013). Transcriptomic analysis of the host response to *Giardia duodenalis* infection reveals redundant mechanisms for parasite control. *mBio*, 4(6):e00660–00613.

Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S. C., Sun, Y. A., Lee, C., Tarakhovsky, A., Lao, K., and Surani, M. A. (2007). Maternal microRNAs are essential for mouse zygotic development. *Genes and Development*, 21(6):644–648.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., and Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, 4:41.

Telleria, E. L., Benoit, J. B., Zhao, X., Savage, A. F., Regmi, S., e Silva, T. L. A., O'Neill, M., and Aksoy, S. (2014). Insights into the trypanosome-host interactions revealed through transcriptomic analysis of parasitized tsetse fly salivary glands. *PLoS Neglected Tropical Diseases*, 8(4):e2649.

Teruel, M., Cabrero, J., Perfectti, F., Alché, J. D., and Camacho, J. P. M. (2009a). Abnormal spermatid formation in the presence of the parasitic B24 chromosome in the grasshopper *Eyprepocnemis plorans*. *Sexual Development*, 3(5):284–289.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2007). Nucleolus size variation during meiosis and NOR activity of a B chromosome in the grasshopper *Eyprepocnemis plorans*. *Chromosome Research*, 15(6):755–765.

Teruel, M., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, 119(2):217–225.

Teruel, M., Sørensen, J. G., Loeschcke, V., Cabrero, J., Perfectti, F., and Camacho, J. P. M. (2011). Level of heat shock proteins decreases in individuals carrying B-chromosomes in the grasshopper *Eyprepocnemis plorans*. *Cytogenetic and Genome Research*, 132(1-2):94–99.

Trifonov, V. A., Dementyeva, P. V., Larkin, D. M., O'Brien, P. C. M., Perelman, P. L., Yang, F., Ferguson-Smith, M. A., and Graphodatsky, A. S. (2013). Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (*Capreolus pygargus)*. *BMC Biology*, 11:90.

Valente, G. T., Conte, M. A., Fantinatti, B. E. A., Cabral-de Mello, D. C., Carvalho, R. F., Vicari, M. R., Kocher, T. D., and Martins, C. (2014). Origin and evolution

of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular Biology and Evolution*, 31(8):2061–2072.

Van Vugt, J. F., Salverda, M., de Jong, J. H., and Stouthamer, R. (2003). The paternal sex ratio chromosome in the parasitic wasp *Trichogramma kaykai* condenses the paternal chromosomes into a dense chromatin mass. *Genome*, 46(4):580–587.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., and Ma, C. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, 5.

Weinkove, D., Neufeld, T. P., Twardzik, T., Waterfield, M. D., and Leevers, S. J. (1999). Regulation of imaginal disc cell size, cell number and organ size by Drosophila class I(A) phosphoinositide 3-kinase and its adaptor. *Current biology: CB*, 9(18):1019–1029.

Wiesen, J. L. and Tomasi, T. B. (2009). Dicer is regulated by cellular stresses and interferons. *Molecular Immunology*, 46(6):1222–1228.

Windhorst, S., Lin, H., Blechner, C., Fanick, W., Brandt, L., Brehm, M. A., and Mayr, G. W. (2013). Tumour cells can employ extracellular Ins(1,2,3,4,5,6)P(6) and multiple inositol-polyphosphate phosphatase 1 (MINPP1) dephosphorylation to improve their proliferation. *The Biochemical Journal*, 450(1):115–125.

Wu, S., Zhu, Z., Fu, L., Niu, B., and Li, W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, 12:444.

Yang, Z., Wafula, E. K., Honaas, L. A., Zhang, H., Das, M., Fernandez-Aparicio, M., Huang, K., Bandaranayake, P. C., Wu, B., and Der, J. P. (2014). Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. *Molecular Biology and Evolution*, page msu343.

Yi, P., Johnson, A. N., Han, Z., Wu, J., and Olson, E. N. (2008). Heterotrimeric G proteins regulate a noncanonical function of septate junction proteins to maintain cardiac integrity in Drosophila. *Developmental cell*, 15(5):704–713.

Yoshida, K., Terai, Y., Mizoiri, S., Aibara, M., Nishihara, H., Watanabe, M., Kuroiwa, A., Hirai, H., Hirai, Y., Matsuda, Y., and Okada, N. (2011). B chromosomes have a

functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS Genetics*, 7(8):e1002203.

Zhao, Y., Wu, D., Fei, C., Guo, J., Gu, S., Zhu, Y., Xu, F., Zhang, Z., Wu, L., Li, X., and Chang, C. (2015). Down-regulation of Dicer1 promotes cellular senescence and decreases the differentiation and stem cell-supporting capacities of mesenchymal stromal cells in patients with myelodysplastic syndrome. *Haematologica*, 100(2):194–204.

Zheng, H., Ji, C., Gu, S., Shi, B., Wang, J., Xie, Y., and Mao, Y. (2005). Cloning and characterization of a novel RNA polymerase II C-terminal domain phosphatase. *Biochemical and Biophysical Research Communications*, 331(4):1401–1407.

Zhou, Q., Zhu, H.-m., Huang, Q.-f., Zhao, L., Zhang, G.-j., Roy, S. W., Vicoso, B., Xuan, Z.-l., Ruan, J., Zhang, Y., Zhao, R.-p., Ye, C., Zhang, X.-q., Wang, J., Wang, W., and Bachtrog, D. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genomics*, 13:109.

Zurita, S., Cabrero, J., López-León, M. D., and Camacho, J. P. M. (1998). Polymorphism regeneration for a neutralized selfish B chromosome. *Evolution*, pages 274–277.

Zwołińska-Śniatałowa, Z. (1976). Research on the effect of alpha-tocopherol (vit. E) on the growth and reproduction of the Colorado potato beetle (*Leptinotarsa decemlineata* Say). In Jermy, T., editor, *The Host-Plant in Relation to Insect Behaviour and Reproduction*, pages 303–306. Springer US.

Conclusiones

1. La presencia de cromosomas B no está asociada con cambios aparentes en frecuencia de dobles roturas y reparación del ADN, y la localización centromérica de la proteína Ku70 en *E. plorans* está relacionada con la función centromérica pero no con la presencia de cromosomas B.

2. Los niveles de transcritos del gen *Hsp70* son independientes de la presencia de cromosomas B, por lo que la disminución de la cantidad de proteína Hsp70, observada previamente en esta especie, es probablemente el resultado de un mecanismo de regulación post-transcripcional.

3. Es posible descubrir la presencia de genes para proteínas presentes en los cromosomas B, mediante el análisis comparativo del patrón de cobertura en librerías Illumina de ADN genómico de individuos sin cromosomas B e individuos con muchos cromosomas B. Este método puede ser aplicable para la búsqueda de genes en situaciones en las que haya un cromosoma en número variable, como es el caso de los cromosomas sexuales heteromórficos y las polisomías.

4. El cromosoma B de *E. plorans* contiene, al menos, nueve genes que codifican para proteínas. En cuatro de ellos, la secuencia codificadora del gen se encuentra completa en el cromosoma B, mientras que, en cinco, ésta se encuentra truncada.

5. Dos de los genes completos (*CIP2A* y *KIF20A*) y tres de los truncados (*CKAP2*, *CND3* and *MYCB2*) se expresan activamente,

demostrando que los cromosomas B no son genéticamente inertes.

6. Estos cinco genes que están activos en el cromosoma B tienen funciones relacionadas con la división celular. Esto sugiere que el secreto del éxito evolutivo de los cromosomas B de esta especie puede residir en la actividad de su contenido génico.

7. Tal como se esperaba, dado su origen común, tanto el cromosoma B2 como el B24 contienen varias copias del gen *CAP-G*. En ambos casos, las copias del B están activas transcripcionalmente, pero no afectan a los niveles de expresión de los genes para otras subunidades de condensina, por lo que no queda claro si los productos del gen *CAP-G* del B afectan de alguna otra manera.

8. Nuestros análisis transcriptómicos y de microarrays han revelado la existencia de 46 genes diferencialmente expresados en los ovarios de hembras con B y sin B, una vez descontados los que se encuentran en los cromosomas B. Las funciones de estos genes guardaron mucha relación con los efectos descritos previamente para los cromosomas B, tales como los efectos nucleotípicos, la respuesta al estrés y la detoxificación, la modificación de proteínas, la función ovárica y la regulación de la expresión génica. Concluimos, por tanto, que la presencia de cromosomas parásitos evoca una respuesta en el genoma hospedador mediada por cambios en expresión génica.

9. Uno de los genes reprimidos en presencia de los cromosomas B es *Dicer1*, un enzima clave para regular la actividad de los elementos transponibles mediante el ARNi. La represión de este gen explica la sobrexpresión de muchos transposones y retrotransposones,

de diferentes familias, observada en las hembras portadoras de cromosomas B.

Conclusions

1. B chromosome presence is not associated with obvious changes in double strand breaks and DNA repair, but the Ku70 protein is probably involved in centromeric function in *Eyprepocnemis plorans*.

2. *Hsp70* transcription levels do not depend on B chromosome presence. Therefore, the decrease in Hsp70 protein levels, previously observed in this species, was probably the result of postranscriptional regulation.

3. The presence of protein-coding genes in B chromosomes can be evidenced through the comparative analysis of coverage patterns for reads obtained from B-carrying and B-lacking individuals on the CDSs of a *de novo* transcriptome of the same species. This method can also be applied to other situations where the number of chromosomes is variable, such as heteromorphic sexual chromosomes and polysomy.

4. The B24 chromosome of *E. plorans* contains at least nine protein-coding genes, four of them showing full length coding sequence and the remaining five being truncated.

5. Two complete (*CIP2A* and *KIF20A*) and three truncated (*CKAP2*, *CND3* and *MYCB2*) genes are actively expressed, thus demonstrating that B chromosomes are not genetically inert.

6. All five genes being actively transcribed in the B chromosome carry out functions that are related with cellular division, suggesting

that the secret of the evolutionary success of B chromosomes in this species might lie on its gene content and transcriptional activity.

7. As expected from their common origin, both B2 and B24 chromosomes contain several copies of the *CAP-G* gene. In both cases, B chromosome copies are transcriptionally active, but do not influence transcription levels of other condensin subunit genes. From this pattern, we can rule out a general enhancenment of condensin activity, but not that the excess of anomalous *CAP-G* transcripts is perturbing the system in other ways.

8. Our transcriptome and microarray analysis revealed the existence of 46 differentially expressed genes in ovaries with and without B chromosomes, once discarded the expression of genes located in the B chromosome. Functions of these genes were related with previously known effects of B chromosomes, such as nucleotypic effects, stress response and detoxification, protein modification and turnover, ovarian function and gene expression regulation. Therefore, we conclude that parasitic B chromosomes evoke gene-expression responses in the host genome.

9. One of the down-regulated genes associated with B chromosome presence is *Dicer1*, which have a key role in controlling transposable elements activity through the RNAi pathway. Down-regulation of this gene explains the up-regulation of many kinds of transposons and retrotransposons from different families, observed in B-carrying females.

Perspectivas

Durante el desarrollo de esta Tesis Doctoral, centrada en estudiar el impacto que puede tener a nivel molecular la presencia de un cromosoma supernumerario (B) en *Eyprepocnemis plorans*, han salido a la luz varios aspectos de los cromosomas B que transforman el concepto general que se tenía de ellos como elementos genéticos transcripcionalmente inertes, carentes de genes y repletos de lo que se ha venido denominando "basura genómica". Este cambio de paradigma abre nuevos interrogantes, que discutiremos en este capítulo.

En el capítulo 1 descubrimos, que la localización centromérica de la proteína de reparación de dobles roturas de ADN, Ku70, no está relacionada con la presencia de cromosomas B, y que el estrés que pueden generar los cromosomas B no está asociado a un incremento en la frecuencia ni en la localización de las roturas de doble cadena en el ADN. Sin embargo, uno de los resultados más intrigantes de esta Tesis Doctoral es el papel que desempeña la proteína de reparación Ku70 en el centrómero de *E. plorans*. La respuesta a cómo y por qué se encuentra esta autapomorfía en *E. plorans* podría obtenerse mediante experimentos de inmunoprecipitación de la cromatina (ChIP), que permitieran identificar las proteínas y el ADN satélite con las que Ku70 interacciona en el centrómero. Además, es posible que comparando secuencias centroméricas de *E. plorans* con las de otras especies relacionadas en las que Ku70 no tiene localización centromérica, que podemos obtener precipitando con CENPA o alguna otra proteína centromérica, pudieramos averiguar si es debido a una peculiaridad de

los centrómeros en esta especie.

La mayoría de los efectos de la presencia de cromosomas B en *E. plorans* descritos hasta ahora, como por ejemplo el aumento en la frecuencia de quiasmas, la disminución del área nucleolar y la disminución de los niveles de la proteína de estrés Hsp70, podían interpretarse en relación a cómo la célula responde al estrés causado por el parasitismo intragenómico. En el capítulo 2, descubrimos que la expresión del gen *Hsp70* no presenta cambios, al nivel transcripcional, asociados a la presencia de cromosomas B, por lo que esta disminución debe ser fruto de regulación post-transcripcional. Sin embargo, al ser una proteína cuya expresión es inducible por cambios ambientales, sería interesante cuantificar los niveles tanto de mRNA como de proteínas en individuos criados en condiciones homogéneas, de forma que lo único que los diferenciara fuera el número de cromosomas B. Otra posibilidad a explorar sería la existencia de algún tipo de impronta en individuos 0B hijos de padres con cromosoma B, por lo que sería interesante hacer cruzamientos controlados y comparar patrones entre 0B hijos de 0B y 0B hijos de +B, y también comparar los niveles de expresión de *Hsp70* al nivel de transcritos y de proteínas entre individuos 0B de poblaciones con alta frecuencia de B y en individuos de poblaciones donde no haya cromosomas B.

En el capítulo 3, demostramos la presencia y expresión de genes y fragmentos génicos codificadores de proteínas en los cromosomas B de *E. plorans*, y lo más interesante es que de los nueve que hemos encontrado, los cinco que están transcripcionalmente activos tienen funciones relacionadas con la división celular. Esto indica que estos parásitos genómicos son capaces de utilizar la maquinaria celular para

transcribir sus genes, lo que, en el mejor de los casos, supone para el hospedador un gasto energético inútil y, en el peor, puede alterar la regulación de algún proceso de la división celular que facilite la transmisión del cromosoma B. La pregunta más evidente que surge de los resultados que se presentan en este capítulo versa sobre la posible funcionalidad de estos genes, algo que adquiere especial relevancia en el caso de los dos genes completos que presentan expresión activa. Sin embargo, en el caso de los genes completos, no tenemos la total seguridad de que la sobre-expresión que detectamos en individuos portadores de cromosomas B provenga en efecto de la expresión de éstos, o bien se trata simplemente de la sobre-expresión de los genes homólogos localizados en los cromosomas A. Por tanto, un primer acercamiento a la solución de este problema debe ser buscar polimorfismos (SNPs) que aparezcan específicamente en las secuencias del cromosoma B y cuantificar la frecuencia de los diferentes alelos al nivel de transcritos. Después de comprobar si estos SNPs pueden causar sustituciones no sinónimas en la secuencia de la proteína, sería interesante saber si estos transcritos se traducen, localizando estas proteínas anómalas mediante secuenciación.

En el caso de los genes truncados, la evidencia de que las copias que se están expresando son las que se localizan en los cromosomas B, viene de la detección de diferentes patrones cuando comparamos, en individuos con diferente número de Bs, los niveles de expresión con cebadores situados en la zona que está presente en el cromosoma B y con cebadores que no lo están, observando que el nivel de expresión sólo aumenta con el número de Bs cuando utilizamos cebadores anclados en la zona que está presente en el cromosoma B. Un experimento interesante para comprobar cómo afectan estos genes a la expresión de otros genes, o a

alguna característica de los cromosomas A o de la célula, podría consistir en realizar su "knockdown" mediante ARN de interferencia. De esta forma podríamos comparar qué consecuencias tiene la inhibición de la expresión de estos transcritos usando dianas presentes y ausentes en el gen del cromosoma B. Comparando los efectos de la interferencia sobre individuos sin cromosomas B (donde es de esperar efectos similares con la dos dianas) con los observados en individuos con B (donde las dos dianas pudieran tener efectos diferentes) podríamos obtener respuestas sobre la posible funcionalidad de los transcritos de los genes truncados del cromosoma B. Otro posible experimento sería detectar, mediante inmunofluorescencia, posibles cambios en la abundancia y localización celular de las proteínas que codifican estos genes en individuos con cromosomas B.

Por otra parte, los nueve genes que detectamos en el cromosoma B podrían usarse también como marcadores para determinar el origen del cromosoma B de *E. plorans*. Si, como sospechamos, todas las variantes de cromosomas B de esta especie comparten un origen común, sería una evidencia indiscutible a favor de esta teoría que todos estos genes se encontraran en todas las variantes de cromosomas B que conocemos en diferentes poblaciones, más aún si tuvieran la misma estructura truncada o si compartieran polimorfismos. Otro punto muy interesante a comprobar sería si estos nueve genes se localizan en el mismo cromosoma A, ya que esto nos permitiría averiguar de qué cromosoma derivó el B en sus orígenes. Si encontráramos variantes alélicas de estos genes se podría estudiar su herencia en cruzamientos de individuos 0B, para comprobar si se heredan en bloque, es decir, en desequilibrio de ligamiento. Esto no resuelve la pregunta de qué cromosoma es el que dio origen al B, pero

sí indicaría, en el caso de que hubiera sintenia, que el B derivó de un único autosoma.

Entre los genes localizados en el cromosoma B, elegimos el gen truncado de la subunidad *CAP-G* de la condensina I para estudiar más a fondo su estructura y las consecuencias de su sobrexpresión. La existencia de un codón de stop prematuro y de un cambio no conservado en el centro activo de la proteína, que se predicen a partir del transcrito, nos lleva a concluir que el B expresa copias pseudogénicas. Esto no excluye la posibilidad de que la expresión de este pseudogén afecte, de alguna manera, a la función normal de la condensina. Sin embargo, para averiguar a qué nivel lo hace, sería necesario analizar mediante *Western-blot*, primero, la existencia de polipéptidos de menor tamaño a los normales (correspondientes al transcrito truncado del B) en los individuos portadores de cromosomas B, lo que indicaría que las secuencias truncadas que se transcriben en el B también se traducen. En segundo lugar, habría que comprobar si la expresión del pseudogen *B-CAP-G* conduce a una disminución de la cantidad de proteína *CAP-G* estándar, ya que los transcritos pseudogénicos pueden actuar como inhibidores competitivos de la traducción o pueden ser fuente de microARNs que neutralizan los transcritos del gen normal. En este sentido, podría ser interesante realizar una secuenciación Illumina de miRNAs en individuos con y sin cromosoma B, y comprobar si efectivamente se están generando miRNAs a partir de los transcritos pseudogénicos del B. Esto también nos daría información similar para los demás genes que se expresan en el cromosoma B.

En cuanto al impacto que la presencia y expresión de los transcritos del B pueden tener sobre la expresión de otros genes del genoma hospedador,

quedan muchas cosas por hacer. Una de las posiblidades es estudiar transcriptomas con individuos en diferentes estadios de desarrollo, desde embriones a ninfas, y con adultos de ambos sexos. Hemos encontrado efectos del cromosoma B sobre genes implicados en procesos relacionados con la función ovárica, las rutas de modificación y degradación de proteínas y la regulación de la expresión génica por la vía de RNA de interferencia. El estudio más profundo de estas rutas de regulación puede ser importante para entender cómo funcionan los mecanismos de defensa del hospedador frente a los cromosomas B, especialmente en el caso de la inhibición de Dicer1, enzima clave en la generación de miRNAs y siRNAs. Este caso merece un estudio más detallado, ya que esta vía de regulación es la mejor baza celular para el control de los elementos transponibles y donde se desarrolla la acción en las interacciones virus-célula, por lo que los mecanismos de virulencia y defensa del cromosoma B como parásito y del genoma A como hospedador pueden también interaccionar a este nivel.