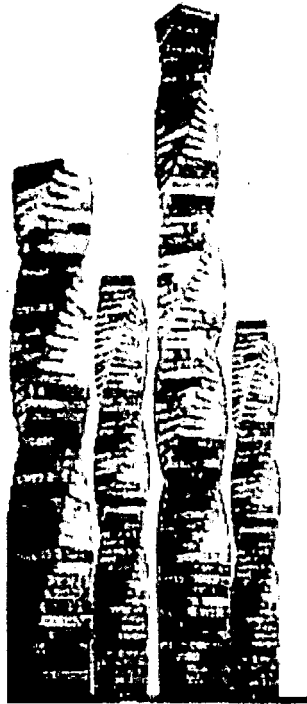




Universitat Autònoma de Barcelona
Facultat de Ciències
Departament de Genètica i de Microbiologia

Inversiones Cromosómicas en *Drosophila*:
Origen Molecular y Significado Evolutivo de su Tamaño



T UAB

5411

Universitat Autònoma de Barcelona
Servei de Biblioteques



1500758568

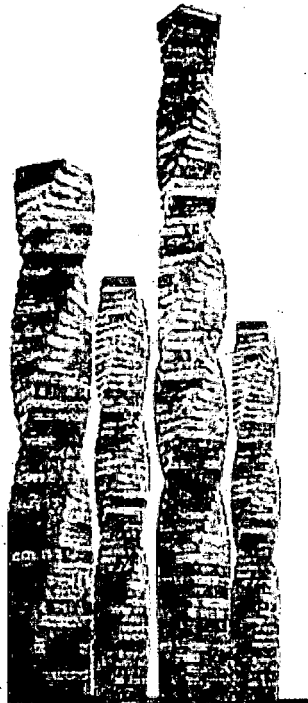
Mario Cáceres Aguilar

Octubre 2000



**Universitat Autònoma de Barcelona
Facultat de Ciències
Departament de Genètica i de Microbiologia**

**Inversiones Cromosómicas en *Drosophila*:
Origen Molecular y Significado Evolutivo de su Tamaño**



Mario Cáceres Aguilar

Octubre 2000



Inversiones Cromosómicas en *Drosophila*:
Origen Molecular y Significado Evolutivo de su Tamaño

Memoria presentada por el licenciado en Biología

MARIO CÁCERES AGUILAR

para optar al grado de

Doctor en Ciencias Biológicas

Tesis Doctoral realizada bajo la dirección de los doctores

Dr. Alfredo Ruiz Panadero

Dr. Antonio Barbadilla Prados

en el Departament de Genètica i de Microbiologia

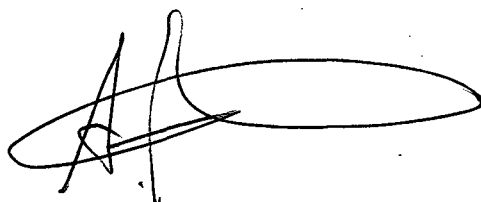
Facultat de Ciències

Universitat Autònoma de Barcelona

Programa de Doctorado en Genética 1995-1997



Dr. Alfredo Ruiz



Dr. Antonio Barbadilla



Mario Cáceres

Bellaterra, 16 de octubre del 2000

*"Caminante, no hay camino,
se hace camino al andar.
Al andar se hace camino,
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar."*

Antonio Machado

ÍNDICE

1. INTRODUCCIÓN	1
1.1 Cambios cromosómicos	4
1.1.1 El polimorfismo de inversiones en el género <i>Drosophila</i>	6
1.1.2 Distribución de las inversiones en el género <i>Drosophila</i>	8
1.2 Origen de las inversiones cromosómicas	14
1.2.1 Generación de las inversiones por recombinación ectópica	15
1.2.2 Elementos genéticos transponibles	18
1.2.2.1 Elementos transponibles de Clase I	19
1.2.2.2 Elementos transponibles de Clase II	20
1.2.2.3 Elementos transponibles de tipo <i>Foldback</i>	21
1.2.3 Elementos transponibles y alteraciones cromosómicas	23
1.2.4 Origen de las inversiones en las poblaciones naturales	26
1.3 Mantenimiento de las inversiones en la naturaleza	29
1.3.1 Carácter adaptativo del polimorfismo de inversiones	29
1.3.2 Efecto de las inversiones sobre la eficacia biológica	32
1.3.3 Estudios sobre la longitud de las inversiones	34
1.4 El organismo de estudio: la especie <i>D. buzzatii</i>	37
1.4.1 El polimorfismo de inversiones en <i>D. buzzatii</i>	38
1.4.2 Elementos transponibles en <i>D. buzzatii</i>	41
1.5 Objetivos	43

2. ARTÍCULOS	47
2.1 Artículo I. <i>Generation of a widespread Drosophila inversion by a transposable element.</i>	49
2.2 Artículo II. <i>Natural hotspots induced by transposable element insertion at the breakpoints of a Drosophila inversion.</i>	55
2.3 Artículo III. <i>Inversion length and breakpoint distribution in the Drosophila buzzatii species complex: Is inversion length a selected trait?</i>	85
2.4 Artículo IV. <i>Recombination rate predicts inversión size in diptera.</i>	95
3. RESULTADOS Y DISCUSIÓN	107
3.1 Origen molecular de las inversiones naturales	109
3.1.1 Clonación de los puntos de rotura de la inversión <i>2j</i>	109
3.1.2 Secuencia de los puntos de rotura de la inversión <i>2j</i>	113
3.1.3 Estructura del elemento transponible <i>Galileo</i>	116
3.1.4 Mecanismo de transposición del elemento <i>Galileo</i>	119
3.1.5 Mecanismo de generación de la inversión <i>2j</i>	124
3.1.6 ¿Son calientes los puntos de rotura de la inversión <i>2j</i> ?	131
3.2 Significado adaptativo de las inversiones	138
3.2.1 Posible efecto de posición de la inversión <i>2j</i>	138
3.2.2 Relación entre la longitud y el éxito de las inversiones	140
3.2.3 ¿Existe una longitud genética óptima para las inversiones?	144
4. CONCLUSIONES	149
5. REFERENCIAS	153
6. AGRADECIMIENTOS	179

1. INTRODUCCIÓN

A lo largo de la evolución se han generado una enorme variedad de seres vivos, desde los más simples, las primeras moléculas orgánicas capaces de autoreplicarse, a los millones de especies que habitan el planeta en la actualidad. Una de las claves de este fenómeno es la existencia de variación genética, la variación en caracteres heredables, sobre la que actúa la selección natural y que constituye la materia prima imprescindible a partir de la cual el proceso evolutivo ha dado lugar a la diversidad biológica existente. Tradicionalmente, se ha distinguido entre dos tipos de variación genética:

- a) Cambios puntuales que afectan a la secuencia nucleotídica de un único gen.
- b) Cambios a nivel cromosómico, que afectan al número o a la estructura de los cromosomas.

El estudio de la evolución a nivel molecular se ha centrado sobre todo en los genes, que son los principales determinantes de las características hereditarias de los organismos (Kimura 1983; Li 1997). No obstante, los genes no se hallan de forma independiente por el genoma, si no que están dispuestos linealmente sobre la molécula de DNA constituyendo los cromosomas. Los cambios cromosómicos alteran la distribución de los genes en el genoma y modifican el contexto en el que se encuentran. Por tanto, podrían tener consecuencias importantes sobre la función de los genes. Se ha postulado así que este tipo de cambios podrían jugar un papel crucial en la adaptación de los individuos al ambiente y en la aparición de nuevas especies (White 1978; King 1993).

Un aspecto fundamental para entender la evolución biológica es explicar cómo se producen los cambios en los cromosomas y qué repercusiones tienen. En este trabajo se ha intentado aportar información y aumentar nuestro conocimiento sobre ambas cuestiones.

1.1 Cambios cromosómicos

El genoma es dinámico y ha experimentado cambios considerables a lo largo de la evolución. Históricamente, los cambios a nivel cromosómico fueron uno de los primeros tipos de variación genética en ser estudiado. Tienen la ventaja de poder detectarse citológicamente mediante el uso de técnicas sencillas, como la microscopía óptica. Dentro de los cambios que pueden afectar al conjunto de cromosomas de una especie, se distinguen las alteraciones en el número de cromosomas, o cambios cromosómicos numéricos, y las alteraciones en la estructura de los cromosomas, o cambios cromosómicos estructurales (Figura 1.1).

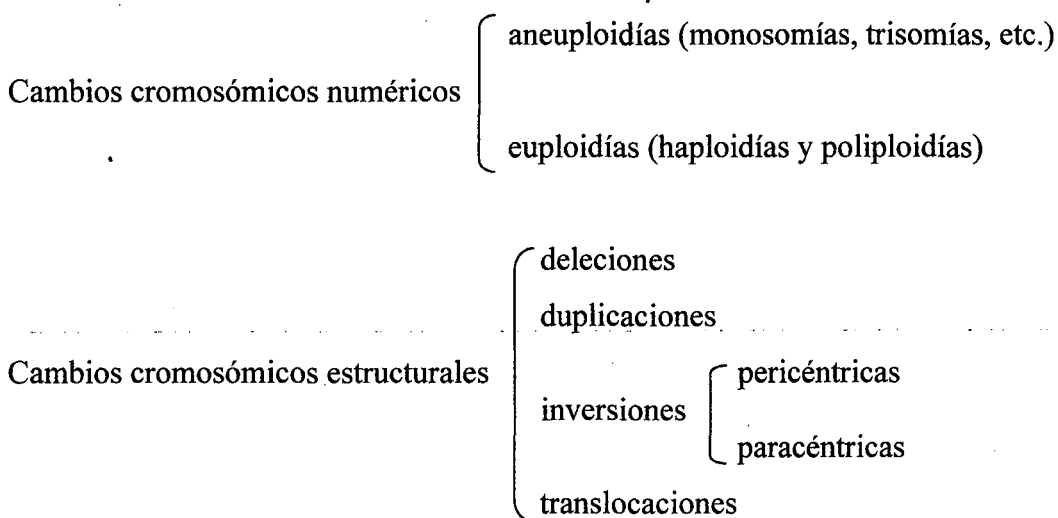


Figura 1.1. Clases de cambios cromosómicos.

Los cambios en la dotación cromosómica o cariotipo son una constante en el proceso evolutivo. Cuando se compara el genoma de especies distintas se observa una gran variación. Ciertamente, no todas las especies difieren en su cariotipo, pero la mayoría sí que lo hacen, y se estima que un 90-99% de especies muestran algún tipo de diferencias en sus cromosomas (White 1978). Por tanto, una gran parte de

los eventos de especiación parecen haber ido acompañados de diferenciación a nivel cromosómico, aunque eso no implica necesariamente que estos cambios hayan tenido un papel esencial en el proceso.

Uno de los cambios cromosómicos numéricos más comunes en la naturaleza es la poliploidía, que podría haber contribuido al incremento del tamaño del genoma y al desarrollo de nuevas funciones génicas (Ohno 1970). En concreto, la poliploidía está ampliamente distribuida en un gran número de grupos vegetales y se cree que aproximadamente la mitad de las angiospermas son poliploides o derivan de ancestros poliploides (Briggs y Walters 1997). Además, también es muy frecuente la presencia de cambios en la organización interna de los cromosomas entre especies cercanas. Los cambios cromosómicos estructurales consisten principalmente en inversiones, pericéntricas y paracéntricas, y translocaciones, incluyendo tanto translocaciones recíprocas como fusiones y fisiones céntricas. Por ejemplo, se estima que dentro de los primates han ocurrido más de 200 inversiones y translocaciones, que han conducido desde un teórico cariotipo ancestral hasta la gran cantidad de cariotipos diferentes que existen hoy en día (Dutrillaux *et al.* 1981).

El número y la morfología de los cromosomas suelen ser constantes dentro de una misma especie. Sin embargo, en algunos casos, existen de manera natural diferencias cromosómicas entre los individuos pertenecientes a una especie. Esta variación es lo que se conoce como polimorfismo cromosómico, en el que uno o varios cromosomas están presentes en dos o más formas alternativas. Un ejemplo clásico de polimorfismo cromosómico es el de las especies de plantas del género *Oenothera*, en las que se encuentran múltiples translocaciones recíprocas en heterocigosis (Cleland 1972). Asimismo, ciertas especies de saltamontes (especialmente de los géneros *Circotettix* y *Trimerotropis*) presentan polimorfismo en sus cromosomas para inversiones pericéntricas (White 1973). De todos modos, el ejemplo más extraordinario y mejor estudiado de variación cromosómica en la naturaleza es el polimorfismo de inversiones existente en el género *Drosophila* y

algunos otros géneros de dípteros como *Anopheles*, *Chironomus* y *Sciara*. Buena parte de las especies de estos géneros son polimórficas para inversiones paracéntricas y sus cromosomas están presentes en varias ordenaciones diferentes (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997).

1.1.1 El polimorfismo de inversiones en el género *Drosophila*

Las inversiones son un tipo de alteración estructural por la cual un segmento cromosómico cambia de orientación dentro del propio cromosoma. Se denominan pericéntricas o paracéntricas, respectivamente, según si el centrómero del cromosoma está incluido o no dentro del segmento invertido. La existencia de inversiones cromosómicas en *Drosophila* fue propuesta por primera vez por Sturtevant (1917), como posible explicación a la supresión de la recombinación que se producía en determinadas regiones de los cromosomas en el cruce de ciertas cepas de *D. melanogaster*, conocidas como cepas supresoras del entrecruzamiento. Más tarde, la elaboración de mapas genéticos permitió demostrar la existencia de una inversión en el orden de los marcadores en los cromosomas de estas cepas (Sturtevant 1926). Además, puso de manifiesto que el cromosoma 3 de *D. melanogaster* y *D. simulans* difería por una inversión fijada entre ambas especies (Sturtevant y Plunkett 1926). Finalmente, el descubrimiento de los cromosomas politénicos en las glándulas salivares de *Drosophila* (Painter 1933) hizo posible observar citológicamente la presencia de inversiones naturales en *D. pseudoobscura* y *D. persimilis* (Tan 1935; Koller 1936).

Los cromosomas politénicos son típicos de las células de tejidos con una alta actividad metabólica, como las glándulas salivares de muchas especies de dípteros (Lacadena 1996). Se trata de un tipo particular de cromosomas interfásicos, formados por un apareamiento de las dos moléculas de DNA homólogas seguido de repetidos ciclos de replicación del DNA no acompañados de división celular. En las

glándulas salivares de *Drosophila*, estos cromosomas se forman por 10 rondas de replicación sucesivas y están constituidos por la yuxtaposición de unas 1.024 moléculas de DNA (Sorsa 1988). Su tamaño es así mucho mayor que el de los cromosomas metafásicos mitóticos y no en vano reciben el nombre de cromosomas gigantes. Además, se caracterizan por presentar un patrón de bandas constante, con bandas oscuras e interbandas claras, fácilmente distinguible al microscopio óptico. Las especiales características de los cromosomas politénicos los han hecho un material muy apropiado para el estudio de las inversiones, pudiendo observarse éstas de forma relativamente sencilla mediante cambios en el patrón de bandas o la aparición de bucles en los heterocigotos.

El desarrollo de las técnicas de estudio de los cromosomas politénicos facilitó enormemente la detección de los cambios cromosómicos y puso en marcha un análisis sin precedentes de la variación cromosómica existente en las poblaciones naturales a nivel de todo el género *Drosophila* (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997). En contraposición con la escasa incidencia de otros tipos de cambios estructurales, este análisis reveló un sorprendente grado de polimorfismo cromosómico para inversiones paracéntricas, que serían con mucho el tipo de cambio cromosómico más frecuente en las especies del género *Drosophila*. Se ha visto que aproximadamente dos tercios de las especies de *Drosophila* estudiadas con detalle presentan inversiones en sus cromosomas (Sperlich y Pfriem 1986; Powell 1997) y se ha propuesto que habría entre 18.000 y 28.000 inversiones polimórficas en total (Stone *et al.* 1960). Además, las inversiones paracéntricas constituirían también el cambio cromosómico fijado más frecuente entre las especies de *Drosophila*. Se estima que globalmente se habrían fijado entre 22.000 y 56.000 inversiones durante la evolución del género (Stone *et al.* 1960). De este modo, cuando se compara la tasa de evolución cromosómica entre diferentes especies de *Drosophila* con la de otros grupos de organismos, se observa que la tasa de fijación de cambios cromosómicos en *Drosophila* es como mínimo un orden de magnitud mayor, aunque en la mayoría de casos es varios cientos de veces más grande (Tabla 1.1).

Tabla 1.1. Tasas de evolución cromosómica en diferentes taxones.

Comparación	Disrupciones cromosómicas (por Mb y Ma)
<i>D. melanogaster</i> vs. <i>D. repleta</i>	0.05837
<i>D. melanogaster</i> vs. <i>D. pseudoobscura</i>	0.08485
<i>Homo sapiens</i> vs. <i>Mus musculus</i>	0.00028
<i>Homo sapiens</i> vs. <i>Sus scrofa</i>	0.00007
<i>Mus musculus</i> vs. <i>Sus scrofa</i>	0.00013
<i>Homo sapiens</i> vs. <i>Capra hircus</i>	0.00018
<i>Mus musculus</i> vs. <i>Capra hircus</i>	0.00029
<i>Arabidopsis thaliana</i> vs. <i>Brassica nigra</i>	0.00536
<i>Lycopersicon</i> vs. <i>Capsicum</i> ssp.	0.00014
<i>Zea mays</i> vs. <i>Sorghum bicolor</i>	0.00019

Debido a la variación en los tipos de cambios cromosómicos fijados en los diferentes organismos, la tasa de evolución se ha expresado como el número de disrupciones cromosómicas por megabase de DNA (Mb) y por millón de años (Ma). La cantidad de DNA se ha calculado como el promedio del tamaño de la región estudiada en las dos especies comparadas. Tomado de Ranz *et al.* (2000).

1.1.2 Distribución de las inversiones en el género *Drosophila*

El polimorfismo de inversiones de las especies del género *Drosophila* se ha venido estudiando intensamente a lo largo de los últimos 60 años. Una de las principales características de este polimorfismo es la distribución no al azar de las inversiones entre especies y grupos de especies, entre cromosomas, y dentro de los cromosomas.

En primer lugar, las inversiones no se distribuyen uniformemente entre las especies. Al considerar las diferentes especies de *Drosophila* estudiadas, se observa que existe una gran variación en el número de inversiones en cada especie. En

comparación con lo esperado de acuerdo con una distribución al azar de las inversiones, parece haber demasiadas especies con muchas inversiones y pocas especies que tengan un número pequeño de inversiones (Sperlich y Pfriem 1986; Powell 1997). Parte de estas diferencias podrían ser debidas al distinto grado con el que se ha estudiado el polimorfismo cromosómico en cada especie. Por ejemplo, en *D. melanogaster*, la especie emblemática del género, se han descrito un total de 339 inversiones paracéntricas (Lemeunier y Aulard 1992), pero sólo 21 de éstas podrían considerarse realmente polimórficas. El resto serían extremadamente raras y se han encontrado una única vez debido al estudio exhaustivo que se ha llevado a cabo. Sin embargo, otras especies como *D. simulans* y *D. virilis*, a pesar de que han sido también intensamente estudiadas, han resultado ser monomórficas para sus cromosomas (Sperlich y Pfriem 1986).

Dentro de las especies que acumulan un mayor número de inversiones destacan las del grupo *willistoni*, en particular *D. willistoni* y *D. paulistorum* con 50 y 63 inversiones, respectivamente (Dobzhansky y Powell 1975), y las del grupo *obscura*, concretamente *D. subobscura* con 66 inversiones (Krimbas 1992). Ambos grupos pertenecen al subgénero *Sophophora* del género *Drosophila* y este hecho podría no ser casual. En comparación con los otros dos subgéneros, *Drosophila* y *Idiomya*, el subgénero *Sophophora* parece contener una proporción muy superior de especies polimórficas para inversiones (Tabla 1.2) y un mayor número promedio de inversiones por especie (Sperlich y Pfriem 1986; Carson 1992).

Por otra parte, dentro de las especies que presentan polimorfismo de inversiones, existen grandes diferencias en cuanto a la distribución de las inversiones en los cromosomas. En algunas especies como *D. subobscura* o *D. willistoni*, las inversiones parecen distribuirse de forma aleatoria y se observa un elevado grado de polimorfismo en todos los cromosomas, aún cuando el número de inversiones concreto en cada cromosoma pueda variar (Sperlich y Pfriem 1986). Sin embargo, en otras especies, no todos los cromosomas parecen poseer la misma probabilidad de tener inversiones y éstas se concentran predominantemente en un

único cromosoma. Un caso extremo es *D. pseudoobscura*, que presenta 34 inversiones polimórficas en el cromosoma 3, tres en el cromosoma X (correspondientes a la ordenación *sex ratio*) y ninguna en el resto de cromosomas (Powell 1992).

Tabla 1.2. Número de especies polimórficas para inversiones.

Subgénero	Especies	
	polimórficas	monomórficas
<i>Drosophila</i>	41 (48%)	44 (52%)
<i>Sophophora</i>	38 (95%)	2 (5%)
<i>Idiomya</i>	27 (47%)	31 (53%)
Total	106	77

Sólo se han incluido especies en las que se han estudiado al menos 10 cepas independientes. Tomado de Powell (1997)

Asimismo, la acumulación de inversiones en un determinado cromosoma se puede producir incluso en especies pertenecientes a diferentes grupos. Los elementos cromosómicos básicos, presentes en todas las especies de *Drosophila*, reciben el nombre de *A*, *B*, *C*, *D*, *E* y *F* (Muller 1940). Estos elementos se habrían conservado a lo largo de la evolución del género debido a que la mayoría de los cambios cromosómicos han consistido simplemente en inversiones paracéntricas y fusiones céntricas (Clayton y Guest 1986). En el grupo *repleta* de *Drosophila*, un 70% de las inversiones, tanto polimórficas como fijadas entre especies, se concentran en el cromosoma 2 (Wasserman 1992). Este cromosoma corresponde al elemento *E* de Muller (Schafer *et al.* 1993; Ranz *et al.* 1997) y representa únicamente un 23% de la longitud total de la fracción eucromática del genoma (Wharton 1942). En el subgénero *Idiomya* y en el grupo *melanica*, se observa también una acumulación similar de inversiones polimórficas en el elemento

cromosómico *E* (Tabla 1.3). De esta forma, en el conjunto de los tres grupos, un 65% del total de 462 inversiones descritas se localizan en este elemento.

Tabla 1.3. Distribución de las inversiones entre diferentes elementos cromosómicos en varios grupos de especies de *Drosophila*.

	Elementos cromosómicos					Inversiones totales
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	
<i>grupo repleta</i>						
inversiones polimórficas	9	21	8	7	132	177
inversiones fijadas	12	18	11	2	76	119
<i>grupo melanica</i>	12	4	1	7	56	80
subgénero <i>Idiomya</i>	13	18	9	11	35	86
Total	46	61	29	27	299	462

Datos de Tonzetich *et al.* (1988) y Wasserman (1992)

Por último, la distribución de los puntos de rotura de las inversiones dentro de los cromosomas tampoco suele ser al azar. En algunos casos, los puntos de rotura de las inversiones tienden a acumularse en las regiones cercanas al centrómero (Tonzetich *et al.* 1988; Krimbas y Powell 1992). Además, en los cromosomas de la mayoría de especies estudiadas parece haber posiciones en las que se acumulan los puntos de rotura de múltiples inversiones y que han recibido el nombre de puntos calientes para roturas cromosómicas (Krimbas y Powell 1992). Por ejemplo, en *D. subobscura* hay una banda cromosómica que se rompe hasta seis veces y tres posiciones que participan en cuatro inversiones cada una (Krimbas 1992). Como siempre, el caso mejor estudiado es el de *D. melanogaster*. En esta especie se ha contabilizado el número de puntos de rotura de inversiones en cada subdivisión cromosómica y la distribución obtenida se ha comparado con la distribución de Poisson esperada si los puntos de rotura se distribuyesen al azar. Se ha podido demostrar así que, exceptuando el cromosoma *X* que presenta muy pocas

inversiones, en el resto de cromosomas existen diferencias significativas entre las dos distribuciones (Lemeunier y Aulard 1992). Se observa un exceso de subdivisiones en las que no hay ningún punto de rotura y de subdivisiones que acumulan más de tres puntos de rotura. Un resultado similar se ha obtenido para la distribución de los puntos de rotura de las inversiones fijadas y polimórficas en 105 especies de *Drosophila* endémicas de Hawai. En este caso, a excepción del cromosoma 5, en todos los otros cromosomas la distribución de puntos de rotura observada difiere significativamente de la esperada por azar (Tonzetich *et al.* 1988). Por su parte, en *D. pseudoobscura* la distribución de los puntos de rotura de las 34 inversiones descritas en el cromosoma 3 no parece diferir significativamente de una distribución uniforme (Olvera *et al.* 1979).

La distribución no aleatoria de las inversiones entre especies, entre cromosomas y dentro de los cromosomas, podría obedecer a dos tipos de procesos diferentes: el mecanismo de generación de las inversiones y la supervivencia diferencial de las inversiones que se generan. Además, en el caso de la distribución de inversiones entre especies, se ha recurrido también a razones históricas (como la edad de la especie), demográficas (como el tamaño de población) o ecológicas (como la diversidad de hábitats) para explicar las diferencias observadas. Para discernir entre las posibles explicaciones alternativas, se ha tratado de simular cómo aparecen las inversiones mediante el uso de radiaciones ionizantes. A diferencia de lo que ocurre con las inversiones naturales, tanto los puntos de rotura de las inversiones inducidas por rayos X (Federer *et al.* 1967) como las roturas producidas mediante rayos γ (Tonzetich *et al.* 1988), tienden a distribuirse al azar por los cromosomas. Estos resultados parecerían indicar que la no aleatoriedad es debida al proceso de mantenimiento de las inversiones en las poblaciones naturales. No obstante, no está del todo claro hasta que punto los experimentos de inducción de roturas cromosómicas artificialmente son representativos de los mecanismos por los que surgen las inversiones en la naturaleza.

Por tanto, a la vista de la riqueza, la diversidad y la distribución de las inversiones presentes en el género *Drosophila*, se plantean de inmediato dos cuestiones fundamentales:

- a) ¿Cómo se originan este tipo de cambios cromosómicos en la naturaleza?
- b) ¿Cuáles son los factores implicados en el mantenimiento de las inversiones en las poblaciones naturales?

1.2 Origen de las inversiones cromosómicas

Los mecanismos implicados en la reestructuración de los genomas tienen un papel muy importante en el cambio evolutivo a largo plazo y, en su mayor parte, son aún desconocidos. Por tanto, desde un principio, el origen de la gran cantidad de inversiones polimórficas y fijadas presentes en el género *Drosophila* ha sido una cuestión muy debatida.

Desde el punto de vista clásico, las inversiones surgen debido a la aparición de dos roturas aleatorias y simultáneas en un mismo cromosoma, seguidas de un proceso de reparación incorrecta de las roturas, en el que el fragmento cromosómico generado se vuelve a unir en orientación invertida (Krimbas y Powell 1992). En este proceso, el repliegamiento que sufre la molécula de DNA sobre si misma dentro del núcleo celular podría facilitar que puedan aparecer dos roturas a la vez y que se reparen incorrectamente (Figura 1.2). El cromosoma se podría romper simplemente de forma espontánea, por ejemplo a causa de tensiones físicas, o por la acción de diferentes agentes mutagénicos que inducen lesiones a nivel del DNA, como ciertas sustancias químicas o las radiaciones ionizantes. Concretamente, la capacidad de las radiaciones ionizantes de producir roturas y reordenaciones cromosómicas se había puesto de manifiesto en múltiples experimentos en el laboratorio (Federer *et al.* 1967; Tonzetich *et al.* 1988).

Una de las principales consecuencias del modelo clásico es que las inversiones deberían tener un origen único y ser monofiléticas (Krimbas y Powell 1992; Powell 1997). Esto es debido a que si las roturas en los cromosomas aparecen al azar, es muy poco probable que surjan dos inversiones con puntos de rotura exactamente en las mismas posiciones. Además, en el caso de que surgiesen, la probabilidad de que ambas inversiones sobrevivan en las poblaciones sería baja. De esta forma, tradicionalmente, la posesión de una misma inversión en dos especies se ha considerado evidencia de un origen común y las inversiones cromosómicas se

han utilizado con gran éxito como carácter filogenético en diversos grupos (Carson 1992; Wasserman 1992).

1.2.1 Generación de las inversiones por recombinación ectópica

No obstante, cada vez más, el anterior modelo está siendo reemplazado por un modelo de generación de las inversiones y otras reorganizaciones cromosómicas mediante recombinación ectópica. Por recombinación ectópica se entiende la recombinación entre segmentos de DNA con homología de secuencia que se encuentran en distintas posiciones de los cromosomas. Aparentemente, tiene lugar por un mecanismo semejante al de la recombinación alélica entre los cromosomas homólogos durante la meiosis (Virgin y Bailey 1998). La principal diferencia es que, en este caso, la recombinación se produce entre secuencias similares situadas en regiones no homólogas de los cromosomas. El primer paso sería el apareamiento ectópico entre las dos secuencias homólogas. El siguiente paso sería la rotura de las cadenas de DNA apareadas y su unión de forma intercambiada, produciéndose así el entrecruzamiento entre ellas (Figura 1.3). Como en el modelo clásico, tiene lugar también un proceso de rotura y unión de los cromosomas, pero la propia maquinaria enzimática celular sería la encargada de llevarlo a cabo de forma muy precisa a nivel molecular.

En principio, la recombinación ectópica podría dar lugar a todo tipo de alteraciones en la estructura de los cromosomas. El resultado del proceso depende de la localización y la orientación de las secuencias homólogas entre las que se produce la recombinación. Si originalmente ambas secuencias se encuentran dispuestas en orientación contraria dentro de un mismo cromosoma, el apareamiento hace que se forme un bucle y el entrecruzamiento genera una inversión (Figura 1.3). Por el contrario, si las dos secuencias homólogas están en la misma orientación, el entrecruzamiento genera una deleción. Finalmente, si los

segmentos que recombinan pertenecen a cromosomas diferentes, se produce una translocación.

Por otra parte, la recombinación ectópica podría tener lugar entre secuencias de naturaleza diversa repetidas por el genoma. Yoon y Richardson (1978) fueron unos de los primeros en proponer la importancia del apareamiento ectópico en la generación de reorganizaciones cromosómicas en *Drosophila*, al observar asociaciones ectópicas entre regiones de heterocromatina intercalar de los cromosomas. No obstante, actualmente se cree que los elementos transponibles desempeñan un papel preponderante en este proceso. Se ha demostrado que muchas alteraciones cromosómicas estructurales que aparecen en *Drosophila* en el laboratorio son causadas por los elementos transponibles (Lim y Simmons 1994). Además, en la mayoría de casos, el mecanismo por el que se generan estos cambios parece consistir en la recombinación ectópica entre copias de los elementos insertadas en diferentes sitios del genoma (Schneuwly *et al.* 1987; Eggleston *et al.* 1996). Dos características intrínsecas de los elementos transponibles promoverían su implicación en todo tipo de procesos de recombinación ectópica. En primer lugar, suelen estar presentes en múltiples copias y proporcionan pequeños segmentos de homología móviles dispersos por el genoma. En segundo lugar, la actividad de los elementos transponibles conlleva cortes de cadena sencilla y de doble cadena en la molécula de DNA (Craig 1997), que podrían inducir la recombinación a través de los mecanismos celulares de reparación de las roturas. En consecuencia, los elementos transponibles se han constituido actualmente como los principales sospechosos de la generación de la gran cantidad de inversiones cromosómicas existentes en el género *Drosophila* y de llevar a cabo la reestructuración del genoma.

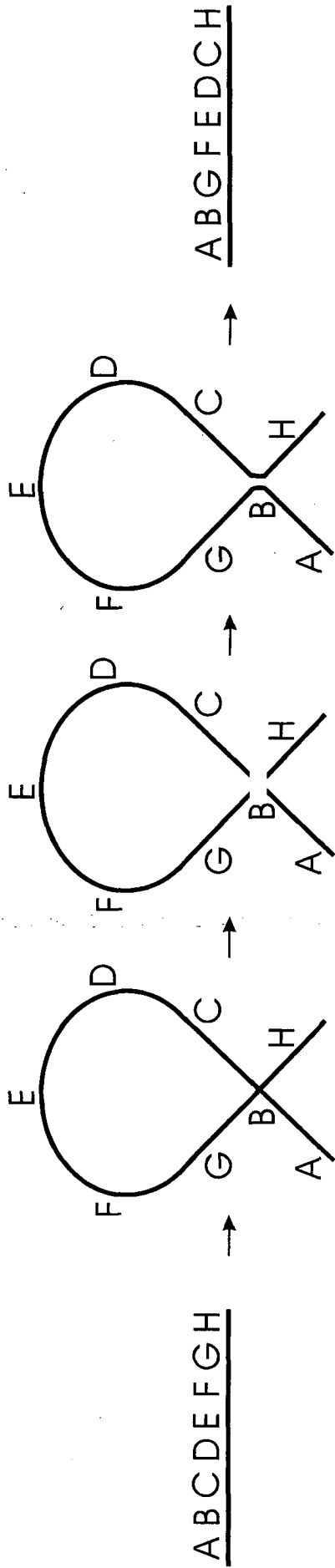


Figura 1.2. Modelo de generación de las inversiones cromosómicas por la aparición simultánea de dos roturas en un mismo cromosoma y la reparación incorrecta de éstas.

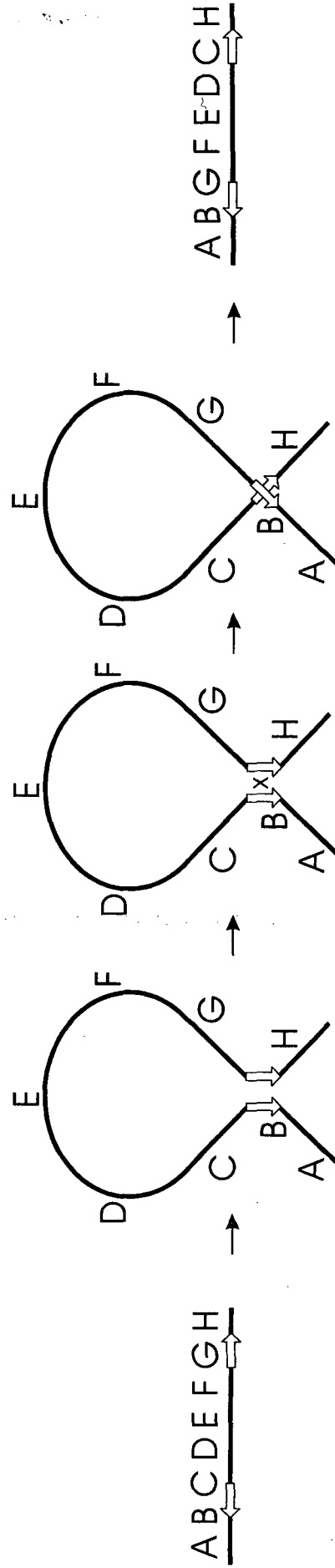


Figura 1.3. Modelo de generación de las inversiones cromosómicas por recombinación ectópica entre segmentos de DNA homólogos dispuestos en orientación contraria en posiciones diferentes de un mismo cromosoma.

1.2.2 Elementos genéticos transponibles

Bárbara McClintock fue capaz de deducir la existencia de los elementos transponibles en el maíz combinando una adecuada serie de experimentos con una prodigiosa intuición genética (McClintock 1984). Desde entonces, y en contra del escepticismo inicial, el desarrollo de las técnicas moleculares ha permitido demostrar la presencia de elementos transponibles en todos los organismos que se han estudiado con detalle, desde procariotas a eucariotas superiores (Berg y Howe 1989; Capy *et al.* 1998), y no hay razón alguna para no pensar que son ubicuos en los seres vivos.

Los elementos transponibles se definen por su capacidad de cambiar de posición dentro del genoma en el que residen y se conocen también como elementos móviles. Se trata de segmentos discretos de DNA que pueden replicarse y propagarse por el genoma, insertándose en nuevos lugares de la secuencia nucleotídica. En el proceso de inserción generan unas pequeñas duplicaciones características de la secuencia del lugar en el que se insertan. Estas duplicaciones reciben el nombre de duplicaciones del sitio diana y aparecen siempre como repeticiones directas de unas pocas bases a ambos lados del elemento. La movilidad de los elementos transponibles hace que suelen mostrar variabilidad en su localización entre los individuos de una misma especie. Además, su capacidad de replicarse hace que normalmente estén presentes en un elevado número de copias en el genoma, formando una parte importante del DNA moderadamente repetitivo.

A pesar de que los elementos transponibles parecen ser constituyentes intrínsecos del genoma de la mayoría de organismos, las razones de su presencia no acaban de estar claras. En general, se han considerado meros parásitos intracelulares egoístas, preocupados tan sólo de su propia supervivencia (Doolittle y Sapienza 1980; Orgel y Crick 1980). Como fruto de su actividad parasitaria, tendrían principalmente un efecto negativo sobre los organismos en los que residen y su presencia en el genoma sería debida únicamente a su capacidad de replicarse, que

impide que sean eliminados (Charlesworth *et al.* 1994). Sin embargo, como muchos otros parásitos, podrían ocupar cualquier lugar dentro de la frontera difusa que separa el parasitismo de la simbiosis. Así, se ha postulado también que podrían resultar beneficiosos para el organismo hospedador y que su actividad podría haber tenido importantes consecuencias en la evolución de las especies (McDonald 1993). Los elementos transponibles son una fuente de mutaciones espontáneas y, bajo ciertas condiciones, podrían contribuir a incrementar la variabilidad genética y el potencial adaptativo de la especie. En este sentido, en la región reguladora de algunos genes se han encontrado inserciones de elementos transponibles fijadas, que parecen haber promovido cambios en la expresión génica (McDonald 1995; Britten 1996; Kidwell y Lisch 1997; Labrador y Corces 1997).

La diversidad de elementos transponibles descritos y la variedad de efectos que pueden ejercer sobre el genoma son muy grandes. No obstante, las diferencias en la estructura de los elementos transponibles y en su mecanismo de transposición han permitido clasificarlos en dos grandes grupos: los elementos de Clase I o retrotransposones y los elementos de Clase II, denominados también simplemente transposones (Finnegan 1989).

1.2.2.1 Elementos transponibles de Clase I

Los elementos de Clase I parecen ser exclusivos de eucariotas y constituyen el tipo mayoritario de elementos transponibles en *Drosophila* (Arkhipova *et al.* 1995). Estos elementos están estrechamente relacionados con los retrovirus y reciben por eso el nombre de retrotransposones. Se caracterizan principalmente por que su mecanismo de transposición alterna fases de DNA y de RNA (Craig 1997). Primero, la forma de DNA integrada en el genoma es transcrita a un intermediario de RNA por los enzimas celulares. El RNA se utiliza luego como molde en un proceso de transcripción inversa para sintetizar un DNA de doble cadena. Finalmente, el DNA generado será el que se integrará en un nuevo lugar del

genoma. La secuencia de los retrotransposones suele contener marcos de lectura abiertos (ORFs) con homología a los genes *gag* y *pol* de retrovirus (Finnegan 1989; Capy *et al.* 1998). Estos ORFs serían los responsables de codificar las proteínas encargadas de llevar a cabo el proceso anteriormente descrito, que son básicamente la transcriptasa inversa, la integrasa y diversas proteínas estructurales.

Dentro de los retrotransposones se distinguen dos tipos, según presenten o no repeticiones terminales largas en sus extremos (Finnegan 1989). Los elementos del primer tipo son los más similares a los retrovirus. Suelen tener un tamaño relativamente grande, entre 5 y 9 kb, y terminan en repeticiones directas largas de entre 250 y 600 bp (Arkhipova *et al.* 1995; Capy *et al.* 1998). Los ejemplos más conocidos son los elementos *Ty1* y *Ty3* de *Saccharomyces cerevisiae* y los elementos *copia* y *gypsy* de *D. melanogaster*. El segundo tipo de elementos de Clase I tiene normalmente un tamaño de entre 4 y 8 kb, aunque con frecuencia se encuentran copias incompletas truncadas en su extremo 5' (Arkhipova *et al.* 1995; Capy *et al.* 1998). Se caracterizan por no poseer repeticiones terminales. En su lugar, tienen una secuencia rica en adenosinas en el extremo 3' de la cadena de DNA codificante, que recuerda a la cola poli(A) de los RNA mensajeros eucarióticos. A este tipo de elementos pertenecen los elementos *LINE* de mamíferos y los elementos *I*, *F* y *jockey* de *Drosophila*.

1.2.2.2 Elementos transponibles de Clase II

Los elementos de Clase II se han encontrado tanto en procariotas como en eucariotas y reciben también el nombre de transposones. Su característica principal es que todos ellos presentan repeticiones terminales invertidas, que delimitan los extremos del elemento respecto al DNA adyacente (Finnegan 1989; Capy *et al.* 1998). Estos elementos parecen transponerse directamente en forma de DNA y se cree que en su mayoría utilizan un mecanismo de transposición no replicativo de corte y empalme. En este proceso, el elemento simplemente se excindiría de su

posición original y se volvería a insertar en algún otro lugar del genoma (Plasterk 1995). La reacción de transposición es catalizada por una enzima denominada transposasa, con actividad de endonucleasa e integrasa (Plasterk 1995; Polard y Chandler 1995). La transposasa reconocería ciertas secuencias situadas en las repeticiones terminales invertidas y cortaría la molécula de DNA a ambos lados del transposón. Además, se encargaría también de realizar un corte escalonado en el sitio diana, en el que posteriormente tendría lugar la integración. Esta enzima está codificada normalmente en la secuencia del propio elemento. No obstante, entre los elementos de Clase II es muy común la existencia de copias defectivas, generadas mediante deleciones internas, que conservan los extremos pero han perdido la capacidad de codificar transposasa (O'Hare y Rubin 1983; Streck *et al.* 1986). Son los denominados elementos no autónomos, que sin embargo pueden transponerse utilizando la transposasa producida por los elementos completos.

Los transposones presentan en su mayoría una organización más variable y son más heterogéneos que los retrotransposones. En función de similitudes estructurales, homologías en la secuencia de la transposasa y tamaño de las duplicaciones del sitio diana generadas durante la inserción, se han distinguido varias superfamilias de elementos de Clase II. Dentro de estas, destacan la superfamilia *hAT*, que incluye los elementos *hobo* de *D. melanogaster*, *Ac* de *Zea mays* y *Tam3* de *Antirrhinum majus* (Calvi *et al.* 1991), la superfamilia *mariner-Tc1*, que incluye elementos descritos en un amplio espectro de organismos diferentes (Doak *et al.* 1994), y la superfamilia *P*, cuyos únicos representantes son los elementos *P* de *Drosophila* (Engels 1996).

1.2.2.3 Elementos transponibles de tipo *Foldback*

Un tipo especial de elementos de Clase II son los llamados elementos *Foldback* o *FB* (Finnegan 1989; Capy *et al.* 1998). Fueron descubiertos por primera vez en *D. melanogaster* (Potter *et al.* 1980; Truett *et al.* 1981), pero actualmente se

han encontrado elementos similares en muchas otras especies, tanto de animales como de plantas. Así, entre los elementos *Foldback* también se incluyen el elemento *BS* de *D. melanogaster* (Campuzano *et al.* 1986), *TFB1* del díptero *Chironomus thummi* (Hankeln y Schmidt 1990), *Tc4* del nematodo *Caenorhabditis elegans* (Yuan *et al.* 1991), los elementos *TU* del erizo de mar *Strongylocentrotus purpuratus* (Hoffman-Liebermann *et al.* 1989) y los elementos *SoFT* de las solanáceas (Rebatchouk y Narita 1997).

Los elementos *Foldback* tienen una estructura peculiar y se caracterizan por poseer repeticiones terminales invertidas muy largas, que en muchos casos constituyen la práctica totalidad del elemento. Estas repeticiones les confieren la propiedad que les da su nombre. Cuando la doble cadena de DNA se desnaturaliza, los extremos invertidos del elemento se aparean entre si y se pliegan rápidamente, generando estructuras secundarias estables en forma de tallo y lazo que son visibles por microscopía electrónica (Truett *et al.* 1981). Además, estos elementos presentan una gran heterogeneidad estructural. Tanto la longitud de las repeticiones invertidas como el tamaño del segmento central situado entre ellas suelen ser muy variables entre las copias de un mismo elemento. Asimismo, la longitud de las dos repeticiones terminales invertidas de cada elemento puede variar.

En general, los elementos *Foldback* no parecen poseer capacidad codificadora y no está todavía totalmente claro cómo se transponen (Capy *et al.* 1998). En *D. melanogaster*, se ha visto que aproximadamente un 10% de los elementos *FB* contienen una secuencia conservada de 4 kb en su región central, que ha recibido el nombre de secuencia *NOF* (Harden y Ashburner 1990; Smith y Corces 1991). Esta secuencia posee marcos de lectura abiertos de gran tamaño que podrían codificar las proteínas necesarias para la transposición del elemento (Templeton y Potter 1989; Harden y Ashburner 1990). De hecho, a partir de extractos de embriones, se ha podido aislar una proteína de 71 kDa perteneciente a uno de los marcos de lectura abiertos (Templeton y Potter 1989). Aunque esta proteína no presenta homología con las transposasas de los otros elementos de Clase

II, está formada por aminoácidos básicos y parece capaz de unirse a DNA. Estos resultados sugieren que, de forma análoga a lo que ocurre en otros elementos de Clase II (O'Hare y Rubin 1983; Streck *et al.* 1986), los elementos *FB* con secuencia *NOF* serían los elementos completos, mientras que el conjunto de copias heterogéneas sin secuencia *NOF* corresponderían a elementos defectivos no autónomos.

1.2.3 Elementos transponibles y alteraciones cromosómicas

El inesperado descubrimiento de los elementos transponibles en el maíz en los años 50 (McClintock 1984), proporcionó nuevas perspectivas para el problema de la generación de las reorganizaciones cromosómicas. A pesar de la aparente falta de función de los elementos transponibles, su presencia en el genoma podría tener importantes consecuencias. Así, desde un primer momento, los elementos transponibles mostraron tener una gran capacidad de producir roturas cromosómicas y alteraciones en la estructura de los cromosomas. Primero en el maíz (Fedoroff 1983) y más tarde en bacterias (Iida *et al.* 1983; Heffron 1983; Kleckner 1983), levaduras (Roeder y Fink 1983) y *Drosophila* (Green 1980; Rubin 1983), se encontraron múltiples ejemplos de mutaciones causadas por elementos transponibles consistentes en deleciones, inversiones y, con menor frecuencia, duplicaciones y translocaciones.

El genoma de *Drosophila* contiene un elevado número de familias diferentes de elementos transponibles (Powell 1997). Dentro de éstas, sólo una pequeña parte han mostrado ser capaces de mediar en la producción de cambios en la estructura de los cromosomas en el laboratorio (Lim y Simmons 1994). Sin embargo, incluyen los principales tipos de elementos transponibles descritos. Entre los elementos que se han implicado en la aparición de reorganizaciones cromosómicas se encuentran elementos de Clase I, como *I* (Busseau *et al.* 1989), *BEL* (Goldberg *et al.* 1983),

Doc (Schneuwly *et al.* 1987) y *roo* (Montgomery *et al.* 1991), y elementos de Clase II, como *P* (Engels y Preston 1984), *hobo* (Lim 1988) y *FB* (Collins y Rubin 1984). Las reorganizaciones causadas por estos elementos suelen ser compatibles con un mecanismo de generación mediante recombinación ectópica, pero hay excepciones.

Uno de los primeros en los que se demostró este tipo de actividad fueron los elementos *P* de *D. melanogaster*. Los elementos *P* se descubrieron como consecuencia de las inusuales características fenotípicas que mostraban los individuos resultantes de ciertos cruces entre machos de cepas naturales y hembras de cepas de laboratorio. Este fenómeno recibió el nombre de disgénesis de los híbridos y consistía en un síndrome complejo, que incluía esterilidad, presencia de recombinación en machos y una elevada tasa de mutación (Bregliano y Kidwell 1983). Además, iba acompañado por la aparición de numerosas reorganizaciones cromosómicas y todos estos efectos parecían ser debidos a un incremento de la actividad de transposición de los elementos *P*. El estudio de una gran cantidad de reorganizaciones inducidas durante la disgénesis híbrida, permitió determinar que al menos uno de los puntos de rotura tiene lugar normalmente en posiciones en las que previamente se habían localizado elementos *P*, dentro o muy cerca de éstos (Engels y Preston 1981, 1984). No obstante, en este caso, los cambios en la estructura de los cromosomas parecían surgir por rotura y reunión al azar de los fragmentos generado, y no por recombinación ectópica entre los elementos.

Un caso similar es el de los elementos *I*, que como los *P* inducen un fenómeno de disgénesis híbrida al cruzar ciertas cepas de *D. melanogaster* (Bregliano y Kidwell 1983). En este proceso, además de mutaciones causadas por la inserción de los elementos *I*, se estimula también la producción de reorganizaciones cromosómicas. La secuenciación de los puntos de rotura de algunas de estas reorganizaciones ha mostrado que poseen elementos *I* con una disposición compatible con un origen debido a recombinación ectópica entre ellos (Busseau *et al.* 1989). No obstante, los elementos *I* no estaban presentes en los cromosomas

originales en los puntos entre los que se produjeron las reorganizaciones, por lo que éstas hubieron de tener lugar durante o poco después de la transposición.

Otro ejemplo bien estudiado es el del elemento *hobo*, que es un eficiente productor de deleciones e inversiones en cepas portadoras de cromosomas *X* inestables (Lim 1988). En este caso, se ha llevado a cabo un análisis molecular detallado de los puntos de rotura de varias de las reorganizaciones causadas por el elemento *hobo* y se ha podido demostrar claramente el mecanismo por el que surgen estos cambios. De acuerdo con el modelo propuesto, las deleciones tienen lugar por recombinación ectópica entre dos elementos *hobo* preexistentes situados en la misma orientación en diferentes posiciones del cromosoma, mientras que las inversiones se producen por recombinación ectópica entre dos elementos *hobo* preexistentes situados en orientación contraria en diferentes posiciones del cromosoma (Lim 1988; Eggleston *et al.* 1996).

Por último, entre los elementos capaces de inducir la producción de reorganizaciones cromosómicas destacan los elementos *Foldback*. A partir del estudio de las mutaciones w^c y w^{DZL} , causadas por la inserción de elementos *FB* en el locus *white* de *D. melanogaster*, se ha visto que estos elementos generan inestabilidad genética y promueven todo tipo de procesos de recombinación (Bingham y Zachar 1989; Smith y Corces 1991). Así, los elementos *FB* producen espontáneamente deleciones, inversiones y translocaciones recíprocas con una elevada frecuencia, aparentemente por recombinación ectópica entre dos de ellos (Collins y Rubin 1984). Además, aparte de promover reorganizaciones en el DNA del hospedador, la recombinación entre las repeticiones invertidas largas de los elementos podría ser uno de los principales mecanismos responsables de la generación de la heterogeneidad estructural característica de los elementos de tipo *Foldback*.

1.2.4 Origen de las inversiones en las poblaciones naturales

A pesar de las numerosas evidencias que implicaban a los elementos transponibles en la aparición de cambios cromosómicos en el laboratorio, su papel como causantes de las inversiones presentes en la naturaleza no estaba del todo claro. La investigación sobre cómo se originan las inversiones naturales y la posible participación de los elementos transponibles en el proceso se ha centrado principalmente en dos aspectos: (1) intentar relacionar la distribución de los puntos de rotura de las inversiones con la distribución de diferentes familias de elementos transponibles; y (2) tratar de clonar los puntos de rotura de inversiones para buscar huellas del mecanismo molecular que las ha originado.

En el primer tipo de estudios, mediante hibridación *in situ* sobre los cromosomas politénicos de *D. melanogaster* se ha podido localizar el elemento transponible *hobo* cerca de los puntos de rotura de cuatro inversiones endémicas de origen relativamente reciente (Lyttle y Haymer 1992). Esto sugiere que *hobo* no tan sólo produciría inversiones en el laboratorio, sino también en la naturaleza. La ausencia del elemento de los puntos de rotura de cuatro inversiones cosmopolitas de la misma especie podría ser debida simplemente a que estas inversiones son más antiguas y el mayor tiempo transcurrido desde su generación ha permitido que los elementos se hayan extinguido. De forma similar, en *D. willistoni* se ha observado la existencia de cierta relación entre los lugares de inserción en los cromosomas de los elementos *P* y la localización de los puntos de rotura de varias de las inversiones naturales descritas (Regner *et al.* 1996). No obstante, ni en *D. subobscura* ni en *D. pseudoobscura* se ha podido encontrar una asociación análoga entre la distribución sobre los cromosomas de diferentes secuencias moderadamente repetitivas, que incluyen los elementos transponibles, y los puntos de rotura de las inversiones (Felger y Sperlich 1989; Powell 1997).

Por otra parte, hasta el momento, el análisis directo de los puntos de rotura de las inversiones naturales ha proporcionado resultados escasos y más bien

contradictorios. Esto ha sido debido en gran parte a las dificultades técnicas que supone la clonación de los puntos de rotura de las inversiones, sobre todo en la gran mayoría de especies de *Drosophila* en las que la información citogenética y cartográfica disponible es escasa. Los primeros en clonar los puntos de rotura de una inversión polimórfica fueron Wesley y Eanes (1994). Utilizando un método muy ingenioso y bastante rápido, basado en la microdissección de regiones específicas de los cromosomas (Wesley *et al.* 1990), fueron capaces de aislar los dos puntos de rotura de la inversión *In(3L)Payne* de *D. melanogaster*. En la secuencia de los puntos de rotura no se halló evidencia alguna de la presencia ni de la acción de elementos transponibles. Además, no se observaron inserciones ni deleciones superiores a 4 bp entre los cromosomas con y sin la inversión. Un resultado similar se obtuvo en la caracterización de los puntos de rotura de una inversión fijada que diferencia el cromosoma *X* de *D. melanogaster* y *D. subobscura* (Cirera *et al.* 1995). En este caso, el tiempo transcurrido desde la generación de la inversión y la divergencia entre las secuencias de las dos especies no permitió localizar con total precisión los puntos de rotura de la inversión. No obstante, tampoco se pudo apreciar rastro alguno de elementos transponibles completos ni defectivos en las regiones que los contenían. Ambos estudios se tomaron como evidencias en contra de la implicación de los elementos transponibles en la generación de las inversiones (Powell 1997).

Recientemente se han clonado y secuenciado los puntos de rotura de otra inversión polimórfica de *D. melanogaster*, la inversión *In(2L)t* (Andolfatto *et al.* 1999). Además, dentro del género *Anopheles*, donde también existe un considerable polimorfismo de inversiones, se han clonado los puntos de rotura de la inversión *2Rd'* de *A. arabiensis* (Mathiopoulos *et al.* 1998). En ambas inversiones se encontraron secuencias repetitivas en los puntos de rotura, que parecían presentar similitudes con elementos transponibles. No obstante, en ninguno de los dos casos se ha podido determinar el mecanismo de generación de la inversión y el papel real que podrían haber desempeñado las secuencias repetitivas presentes en los puntos de rotura. Por tanto, la cuestión del origen de las inversiones naturales y de la

posible participación de los elementos transponibles en el proceso seguía sin estar resuelta.

1.3 Mantenimiento de las inversiones en la naturaleza

Aparte del origen de las inversiones, otra cuestión fundamental es cómo se mantienen las inversiones en las poblaciones naturales y qué papel desempeña el polimorfismo de inversiones, característico de tantas especies de dípteros, en la adaptación de los organismos al ambiente. En un principio, la ausencia de diferencias visibles en la apariencia externa de las moscas con y sin la inversión hizo creer que las inversiones eran variantes genéticas neutras, sin efectos positivos ni negativos para los individuos portadores (Dobzhansky 1970). Sin embargo, un análisis más detallado y el descubrimiento de la gran cantidad de inversiones presentes en las poblaciones naturales de numerosas especies de *Drosophila*, hicieron que se abandonase esta visión simplista. Desde entonces, se ha considerado que el polimorfismo de inversiones es mantenido activamente por selección natural y se han dedicado una gran cantidad de esfuerzos a desentrañar los procesos que gobiernan su dinámica en las poblaciones. Las inversiones cromosómicas del género *Drosophila* se habrían constituido como uno de los ejemplos paradigmáticos de polimorfismos equilibrados y se han utilizado extensamente como modelo para el estudio de los mecanismos selectivos implicados en el mantenimiento de la variación genética.

1.3.1 Carácter adaptativo del polimorfismo de inversiones

Tras muchos años de investigación, son pocos los estudiosos del polimorfismo de inversiones que dudan aún de su carácter adaptativo y de que la selección natural es la fuerza principal que determina la frecuencia de las inversiones en la naturaleza. Dos tipos distintos de evidencias han permitido llegar a esta conclusión: (1) las variaciones temporales y espaciales regulares en la composición del polimorfismo de inversiones en poblaciones naturales; y (2) los

experimentos de laboratorio con cajas de poblaciones (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997).

Las primeras pruebas de la acción de la selección sobre las inversiones provienen del descubrimiento de fluctuaciones estacionales recurrentes en las frecuencias relativas de varias ordenaciones cromosómicas de *D. pseudoobscura* en poblaciones de California (Dobzhansky 1943, 1956, 1971). Difícilmente se podían explicar estos cambios regulares si no era mediante la acción de la selección. Más adelante, variaciones estacionales similares en la frecuencia de las inversiones se observaron en especies como *D. persimilis*, *D. robusta*, *D. melanica*, *D. willistoni* y *D. funebris* (Sperlich y Pfriem 1986), e incluso en otras especies, como por ejemplo *D. subobscura* (Fontdevila *et al.* 1983; Rodríguez-Trelles *et al.* 1996), en las que se suponía que no existían tales oscilaciones.

La acción de la selección se infirió también de la existencia de patrones de distribución espaciales de las inversiones correlacionados con la variación de algún factor ambiental, conocidos también como clinas. En este sentido, uno de los mejores ejemplos del carácter adaptativo del polimorfismo de inversiones se encuentra en *D. subobscura*. Esta especie es originaria del continente euroasiático y presenta una considerable variación geográfica en la frecuencia de sus ordenaciones cromosómicas. No obstante, recientemente ha colonizado Norteamérica y Sudamérica. En ambos casos se han establecido de forma independiente gradientes latitudinales norte-sur en la frecuencia de las diferentes ordenaciones, que son equivalentes a los encontrados previamente en Europa y el norte de África (Prevosti *et al.* 1988; Ayala *et al.* 1989). Asimismo, en *D. melanogaster* existe un patrón de clinas latitudinales para la frecuencia de las cuatro inversiones cosmopolitas principales similar en ambos hemisferios (Lemeunier y Aulard 1992). Por último, aparte de clinas latitudinales, se han encontrado también cambios graduales en la frecuencia de las inversiones en función de la altitud de la población estudiada, como en el caso de *D. pseudoobscura* (Dobzhansky 1948, 1971). Además, *D. robusta* muestra clinas altitudinales semejantes en distintas montañas con una

amplia distribución geográfica (Levitan 1992). La existencia de estos patrones estacionales, latitudinales y altitudinales en la distribución de las inversiones se consideran evidencias inequívocas de su carácter adaptativo y se pueden interpretar como que las ordenaciones cromosómicas estarían mejor adaptadas a diferentes condiciones ambientales (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997).

Por otra parte, los experimentos de laboratorio con cajas de poblaciones, en las que se seguía la dinámica de diversas ordenaciones cromosómicas a lo largo de varias generaciones, contribuyeron a reforzar aún más la idea del carácter adaptativo del polimorfismo de inversiones. Estos experimentos se han llevado a cabo con numerosas especies de *Drosophila*. No obstante, la especie mejor estudiada con diferencia es *D. pseudoobscura*. Los primeros experimentos con poblaciones artificiales de esta especie datan de 1946 y en ellos se observó que, tras varios meses de muestreo, se obtenía un equilibrio más o menos estable para las dos ordenaciones consideradas, independientemente de cual fuese su frecuencia inicial (Wright y Dobzhansky 1946). Posteriormente, en poblaciones experimentales de otras especies se observó que se establecían polimorfismos de inversiones estables análogos (Sperlich y Pfriem 1986). A partir de los resultados anteriores se aceptó de forma generalizada que el polimorfismo de inversiones estaba mantenido en las poblaciones naturales por selección equilibradora debida a superioridad de los heterocariotipos. Además, los importantes cambios observados en la frecuencia de las inversiones en tan sólo unas pocas generaciones sugerían que, al menos en algunos casos, la presión selectiva sobre las inversiones debía ser bastante intensa (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997).

1.3.2 Efecto de las inversiones sobre la eficacia biológica

Como hemos visto, las inversiones parecían conferir, bajo ciertas condiciones y en ciertas combinaciones, una ventaja selectiva a los individuos portadores. Quedaba por dilucidar cómo se producía este efecto. En principio, las inversiones no suponen pérdida ni ganancia de material genético, sino simplemente una reorganización de éste. Así, únicamente cuando uno de los puntos de rotura afecta directamente a un gen, se espera que la inversión tenga consecuencias fenotípicas importantes. No obstante, igual que otros tipos de alteraciones cromosómicas estructurales, las inversiones, tanto paracéntricas como pericéntricas, suelen producir una disminución de la fertilidad debido a la formación de gametos no funcionales durante la meiosis de los heterocariotipos. En el caso concreto de las inversiones paracéntricas, un único entrecruzamiento sencillo dentro de la zona invertida produce cromosomas dicéntricos y acéntricos que pueden dar lugar a gametos desequilibrados no viables.

Ciertas particularidades de la meiosis en *Drosophila* y otros dípteros impiden que esto ocurra y hacen que no se produzca selección en contra de las inversiones paracéntricas en estos organismos (Sturtevant y Beadle 1936; Navarro *et al.* 1997). En primer lugar, en la mayoría de especies de *Drosophila* no hay recombinación durante la espermatogénesis en los machos (White 1973). En segundo lugar, en las hembras la meiosis tiene lugar de forma ordenada y, probablemente debido al retraso sufrido en su separación, las cromátidas aberrantes producidas por recombinación siempre quedan relegadas a los corpúsculos polares y no llegan al óvulo (Sturtevant y Beadle 1936; Carson 1946). De este modo, sólo se producen gametos desequilibrados y cigotos inviables cuando ocurren entrecruzamientos dobles en los que intervienen las cuatro cromátidas, que dan lugar únicamente a cromosomas dicéntricos y acéntricos, (Sturtevant y Beadle 1936; Roberts 1976). Por tanto, la letalidad asociada a recombinación en inversiones paracéntricas en

Drosophila es mínima y se puede decir que sólo tienen efectos negativos apreciables sobre la fertilidad cuando son grandes (Navarro *et al.* 1997).

No obstante, el hecho de que las inversiones paracéntricas no sean desfavorables para los individuos portadores no basta para explicar la abundancia de inversiones en las poblaciones y la existencia de un polimorfismo equilibrado. Es necesario que la selección natural intervenga activamente. Se han propuesto varias teorías, basadas en los efectos genéticos de las inversiones, para tratar de determinar cómo opera la selección natural sobre éstas y cuáles podrían ser las causas de su carácter adaptativo.

Por una parte, las inversiones conllevan roturas en la molécula de DNA y modifican la disposición de los genes en los cromosomas. Se ha postulado así que los efectos de posición resultantes del cambio en la localización física de los genes podrían ser beneficiosos y resultar en una ventaja de los individuos portadores de la inversión (Sperlich 1963, 1966). El cambio de posición de un gen en relación con las secuencias adyacentes y ciertas zonas de los cromosomas podría alterar su expresión. Sobre todo, sería importante el efecto mutacional producido en el momento de aparición de la inversión en la zona de los puntos de rotura, cuando éstos se sitúan próximos a la secuencia de un gen.

Por otra parte, la otra consecuencia principal de la presencia de las inversiones es la reducción de la recombinación dentro de la zona invertida en los heterocariotipos, que se produce probablemente tanto por la no incorporación a los gametos de los productos resultantes del entrecruzamiento (Sturtevant y Beadle 1936; Carson 1946), como por problemas mecánicos en el apareamiento meiótico entre las regiones homólogas (Coyne *et al.* 1993). Dentro del segmento invertido, sólo se podría producir intercambio de material genético entre los cromosomas homólogos con y sin la inversión por entrecruzamientos dobles entre las mismas dos cromátidas no hermanas y por conversión génica (Navarro *et al.* 1997). De esta forma, la cuestión del mantenimiento de las inversiones se incluye dentro de la

problemática de la evolución de los modificadores de la recombinación y la ventaja selectiva de éstos. La recombinación es un mecanismo imprescindible de intercambio genético y flexibilidad evolutiva, pero bajo unas condiciones estables estaría favorecida la existencia de sistemas que la limiten (Mather 1943; Charlesworth 1993).

En concreto, las inversiones tienden a mantener unido el conjunto de genes incluido originalmente dentro del segmento invertido y su efecto podría provenir de la protección respecto a la recombinación de combinaciones favorables de alelos. Se han propuesto básicamente dos modelos según si se consideran interacciones epistáticas entre los loci o no. El modelo de selección génica supone que los genes contenidos dentro de la inversión actúan independientemente y tienen efectos aditivos (Nei *et al.* 1967). Una inversión se verá favorecida si atrapa un menor número de alelos deletéreos o un mayor número de alelos favorables que el promedio de la población en dicho segmento. En este caso, la ventaja de los heterocigotos para la inversión sería debida a la simple dominancia de los alelos favorables sobre los desfavorables. En la teoría de la coadaptación, cada ordenación contiene y conserva conjuntos de alelos complementarios, que actúan eficazmente de forma combinada y dan lugar a una superioridad de los heterocariotipos (Dobzhansky 1970). Estas combinaciones favorables de alelos reciben el nombre de complejos génicos coadaptados y suponen interacciones epistáticas entre los loci. Estas interacciones epistáticas podrían existir antes de que se genere la inversión (Charlesworth y Charlesworth 1973) o formarse una vez se ha producido ésta (Kojima y Schaffer 1964; Wasserman 1968).

1.3.3 Estudios sobre la longitud de las inversiones

En la observación de los cromosomas politénicos al microscopio, las inversiones se identifican normalmente por la presencia de bucles o cambios en el

patrón de bandas. Nada se puede saber a simple vista de su contenido. No obstante, a partir de la descripción de las inversiones sobre los cromosomas politénicos, se puede determinar de manera sencilla una característica importante: su longitud. Por lo general, se considera que la gran mayoría de inversiones que surgen por mutación se pierden, y sólo una pequeña parte sobreviven y llegan a establecerse en las poblaciones (Krimbas y Powell 1992; Powell 1997). Las inversiones presentes actualmente en las poblaciones naturales podrían ser, en consecuencia, una muestra no al azar de las que han ocurrido y su longitud podría estar sesgada por el proceso selectivo. El análisis de la longitud de las inversiones proporcionaría así información sobre los diversos mecanismos que podrían intervenir en su mantenimiento.

A lo largo de muchos años, diferentes estudios han tratado la cuestión de la longitud de las inversiones naturales. Wallace (1954) y Crumpacker y Kastritsis (1967) fueron los primeros en proponer que, de acuerdo con la teoría de la coadaptación, debía existir una relación entre la longitud y el éxito de las inversiones. Las inversiones endémicas serían mayoritariamente cortas, mientras que las más frecuentes tendrían un tamaño intermedio. Posteriormente, Olvera *et al.* (1979) encontraron evidencias de esta relación analizando la longitud de las 34 inversiones descritas en el cromosoma 3 de *D. pseudoobscura*. Clasificaron las inversiones en dos grupos en función de su éxito evolutivo, utilizando como criterio su distribución geográfica y su frecuencia en las poblaciones: cosmopolitas y frecuentes por un lado y endémicas y raras por el otro. No detectaron diferencias significativas entre la distribución de las longitudes de los dos tipos de inversiones, pero observaron que las inversiones pequeñas (de menos del 20% del cromosoma) eran predominantemente raras y endémicas. Un resultado similar se obtuvo para 15 inversiones de *D. buzzatii* (Ruiz *et al.* 1984). En este caso sí que se encontraron diferencias significativas entre las longitudes de los dos grupos de inversiones. Todas las inversiones cosmopolitas y frecuentes tenían un tamaño intermedio (entre un 20% y un 40% del cromosoma), mientras que las inversiones pequeñas eran endémicas y raras.

Por su parte, Van Valen y Levins (1968) estudiaron la longitud de 101 inversiones pertenecientes a un gran número de especies de *Drosophila* y otras familias de insectos (Agromyzidae, Cecidomyidae, Chironomidae, Sciaridae y Simuliidae). A partir de aquí, probaron el ajuste de la distribución de la longitud de las inversiones en los diversos grupos a un modelo de generación de las inversiones por rotura al azar de los cromosomas. Todos los grupos mostraron distribuciones de la longitud de las inversiones claramente diferentes y, mientras en Agromyzidae y Chironomidae había un exceso de inversiones largas, las inversiones de la familia Simuliidae eran mayoritariamente de tamaño pequeño. En el caso de *Drosophila* se observó únicamente una deficiencia de inversiones pequeñas, de menos del 10% del cromosoma, respecto a la distribución esperada. Un análisis más exhaustivo de las inversiones presentes en *D. subobscura* (Krimbas y Loukas 1980) y *D. obscura* Fallen (Brehm y Krimbas 1991) mostró una deficiencia de inversiones pequeñas y grandes de acuerdo con el modelo de rotura al azar.

De estos estudios sobre la longitud de las inversiones se derivaron dos conclusiones principales. En primer lugar, la distribución de longitudes de las inversiones difería considerablemente entre diversos grupos de especies (Van Valen y Levins 1968). En segundo lugar, los resultados obtenidos parecían apuntar a que las inversiones de tamaño intermedio estaban favorecidas por la selección natural (Krimbas y Powell 1992).

1.4 El organismo de estudio: la especie *D. buzzatii*

La especie *D. buzzatii* fue descrita por Patterson y Wheeler (1942) y recibió su nombre del Dr. A. Buzzati-Traverso, que fue el primero en descubrirla. Es una especie que se viene estudiando intensamente desde hace varias décadas y que, poco a poco, ha ido ganando una mayor atención dentro de la comunidad de *Drosophila*. Una de sus principales ventajas para los estudios evolutivos es que su nicho ecológico es bien conocido. Se trata de una especie cactófila y se suele encontrar asociada sobre todo con cactus del género *Opuntia*, pero también del género *Trichocereus* (Hasson *et al.* 1992a), en los cladodios y frutos en descomposición de los cuales se alimenta y pone sus huevos. Ocupa hábitats semiáridos y, aunque es originaria del continente americano, actualmente tiene una distribución subcosmopolita y se ha extendido por Sudamérica, el sur de Europa, el norte de África y Australia (Fontdevila *et al.* 1981, 1982; Barker 1982). La colonización de la región mediterránea se produjo probablemente en el siglo XVIII, con la apertura de los puertos argentinos y el comercio con Europa (Fontdevila *et al.* 1981), mientras que su aparición en Australia habría ido ligada a la introducción de la polilla *Cactoblastis cactorum* en los años 30 con el objetivo de controlar la expansión de *Opuntia* en ese continente (Barker *et al.* 1985).

D. buzzatii pertenece al complejo *buzzatii* del subgrupo *mulleri*, que a su vez se incluye dentro del grupo *repleta* del subgénero *Drosophila* de *Drosophila*. El grupo *repleta* es uno de los más numerosos de todo el género (Wasserman 1982, 1992). Comprende un total de más de 90 especies, distribuidas en cinco subgrupos, entre los que destaca el subgrupo *mulleri*, con 46 especies. El complejo *buzzatii* está formado por un conjunto monofilético de 12 especies cactófilas, estrechamente relacionadas, originarias de Sudamérica (Ruiz y Wasserman 1993; Ruiz *et al.* 1997). Estas especies han sido muy bien estudiadas a nivel citogenético. Mediante la comparación del patrón de bandas de los cromosomas politénicos de las diferentes especies y de diversos híbridos interespecíficos se ha podido elaborar una

filogenia cromosómica completa de todas ellas (Figura 1.4). El complejo consta de tres clusters, *stalkeri*, *buzzatii* y *martensis*, con dos, seis y cuatro especies, respectivamente, que se diferencian por una serie de inversiones fijadas. Posteriormente, se han llevado a cabo filogenias moleculares a partir de la secuenciación de los genes mitocondriales de las subunidades I, II y III de la citocromo oxidasa (Spicer 1995) y del gen de la xantina deshidrogenasa (Rodríguez-Trelles *et al.* 2000), que han permitido corroborar y resolver mejor algunas de las relaciones filogenéticas propuestas.

1.4.1 El polimorfismo de inversiones en *D. buzzatii*

D. buzzatii ha sido sometida a un extenso análisis citogenético en numerosas localidades, tanto del Viejo como del Nuevo Mundo (Wasserman 1954; Fontdevila *et al.* 1981, 1982; Ruiz *et al.* 1984; Barker *et al.* 1985; Hasson *et al.* 1995). El cariotipo de esta especie consiste en cinco pares de cromosomas telocéntricos y un par de cromosomas puntiformes. En total, se han descrito 16 inversiones paracéntricas en los cromosomas de *D. buzzatii*, de las cuales nueve se pueden considerar polimórficas y siete se han encontrado sólo en un único individuo (Tabla 1.4). No obstante, las inversiones no se distribuyen al azar por los cromosomas. La mayoría se concentran en el cromosoma 2, que igual que en el resto del grupo *repleta* (Wasserman 1992), sería con mucho el más dinámico.

Entre las inversiones polimórficas presentes en *D. buzzatii* destaca la inversión $2j$ (Wasserman 1954). Esta inversión se encuentra prácticamente en toda el área de distribución de la especie, en una proporción en muchos casos superior a la de la ordenación 2 *standard* ($2st$). Su frecuencia varía ampliamente en las poblaciones de esta especie, de 0 en diversas localidades del noreste de Brasil, en las que está ausente (Figueiredo y Sene 1992), a 1 en algunas localidades del norte de Argentina, en las que está fijada (Hasson *et al.* 1995), con un valor promedio de

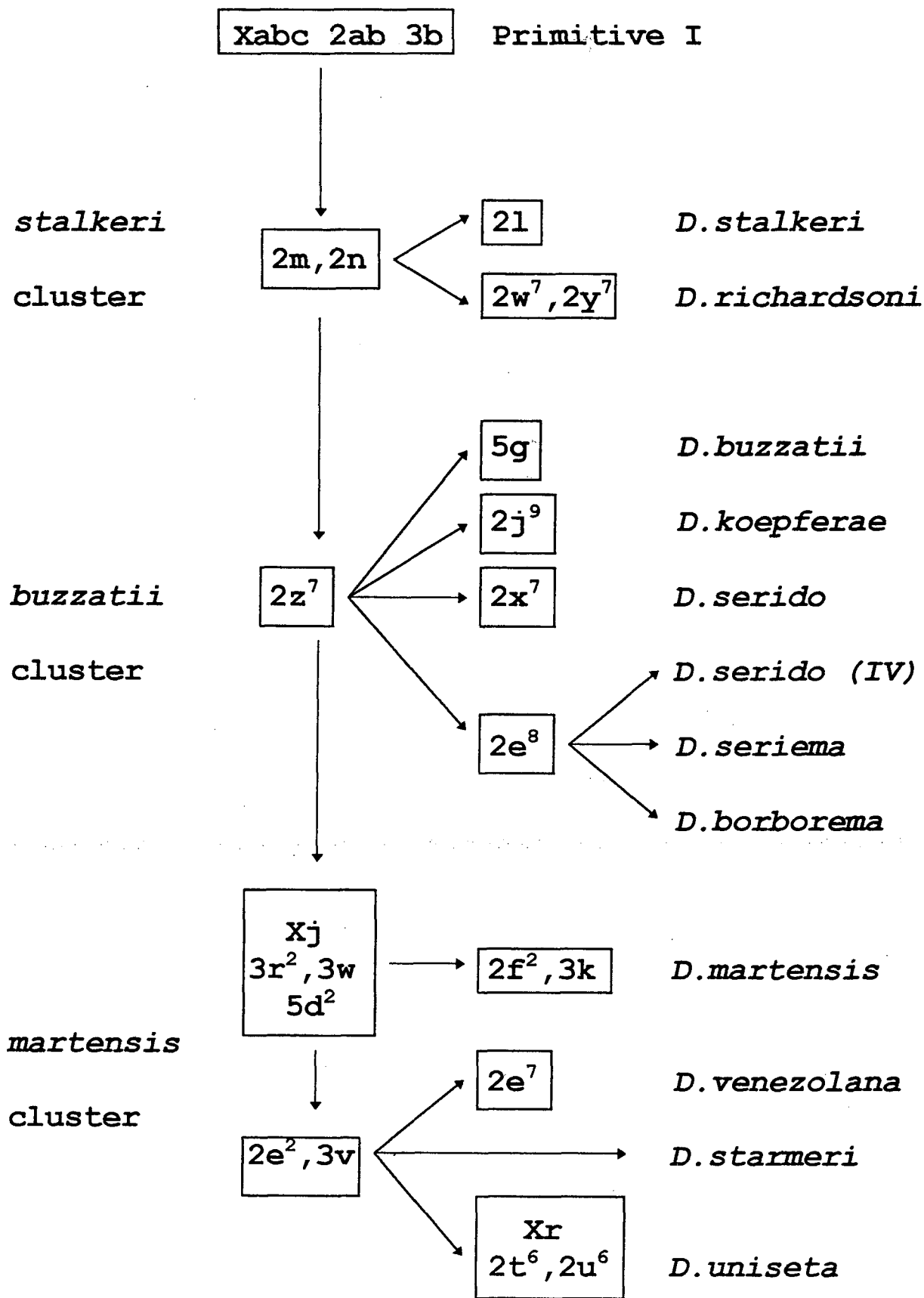


Figura 1.4. Filogenia cromosómica de las 12 especies descritas en el complejo *buzzatii* de *Drosophila*. Tomado de Ruiz *et al.* (1997).

aproximadamente 0,6. En Sudamérica, esta variación sería debida en parte a la presencia de una clina en la frecuencia de la inversión $2j$ relacionada con la altitud (Hasson *et al.* 1995).

Tabla 1.4. Inversiones cromosómicas en *Drosophila buzzatii*.

Inversión	Ordenación original	Distribución geográfica
$2j$	$2st$	cosmopolita
$2z^3$	$2j$	cosmopolita
$2q^7$	$2j$	moderadamente frecuente
$4s$	$4st$	moderadamente frecuente
$2y^3$	$2st$	endémica
$2c^9$	$2j$	endémica
$2r^9$	$2st$	endémica
$2s^9$	$2j$	endémica
$2i^9$	$2j$	única
$2d^9$	$2jc^9$	única
$2e^9$	$2st$	única
$2f^9$	$2st$	única
$2g^9$	$2st$	única
$2h^9$	$2j$	única
$3j^2$	$3st$	única
$5c^2$	$5st$	endémica

Datos de Ruiz *et al.* (1984) y Barker *et al.* (1985)

Diversos estudios han mostrado la existencia de una relación entre el polimorfismo de inversiones del cromosoma 2 de *D. buzzatii* y varios componentes de la eficacia biológica (Ruiz *et al.* 1986; Hasson *et al.* 1991). Además, el efecto de las diferentes ordenaciones del cromosoma 2 sobre la eficacia biológica ha sido

investigado intensamente (Ruiz y Santos 1989; Ruiz *et al.* 1991; Hasson *et al.* 1992b; Barbadilla *et al.* 1994; Norry *et al.* 1995; Betrán *et al.* 1998; Fernández Iriarte y Hasson 2000). Los cromosomas con la inversión $2j$ parecen formar un polimorfismo equilibrado con los cromosomas $2st$, en el que cada ordenación tendría ventaja en una fase diferente del ciclo vital. Los cromosomas $2j$ están asociados con un mayor tamaño corporal de los adultos y conferirían ventaja en los componentes tardíos de la eficacia biológica. Por su parte, los cromosomas $2st$ están relacionados con un menor tiempo de desarrollo de las larvas y conferirían ventaja en los componentes tempranos de la eficacia biológica. Por tanto, el polimorfismo de inversiones del cromosoma 2 de *D. buzzatii* estaría mantenido activamente en las poblaciones por selección natural y la inversión $2j$ tendría efectos adaptativos importantes sobre los individuos portadores.

1.4.2 Elementos transponibles en *D. buzzatii*

Al comienzo de este trabajo, la información a nivel molecular disponible en *D. buzzatii* y el resto de especies del complejo era bastante escasa. Pocos eran también, por tanto, los elementos transponibles que se habían descrito hasta ese momento en estas especies. El principal estudio encaminado a tratar de identificar elementos transponibles en el complejo *buzzatii* es el de Marín *et al.* (1992), estimulado por el descubrimiento de un fenómeno de inestabilidad cromosómica en los híbridos entre *D. buzzatii* y *D. koepferae* similar al síndrome de disgénesis híbrida producido de forma intraespecífica en *D. melanogaster* (Naveira y Fontdevila 1985). En este estudio, se analizó el patrón de hibridación de más de 200 clones de DNA genómico provenientes de *D. buzzatii* y *D. koepferae*. En ambas especies se observó un alto contenido de DNA repetitivo (19-26% en *D. buzzatii* y 27-32% en *D. koepferae*) y se seleccionaron para su caracterización posterior aquellos clones con un patrón de hibridación consistente en bandas euromáticas

dispersas en la hibridación *in situ* y bandas discretas en la hibridación de Southern, que parecían incluir DNA moderadamente repetitivo.

De esta forma se pudieron caracterizar dos nuevos elementos transponibles no descritos previamente: el retrotransposón *Oswaldo* en *D. buzzatii* y el transposón *Gandalf* en *D. koepferae*. *Oswaldo* es un retrotransposón con repeticiones terminales largas que está relacionado con el grupo de retrotransposones *gypsy/Ty3* (Labrador y Fontdevila 1994). El elemento completo tiene un tamaño de 9 kb y contiene tres marcos de lectura grandes que podrían codificar todos los elementos necesarios para que se comporte como un retrovirus funcional (Pantazidis *et al.* 1999). *Gandalf* es un elemento transponible de Clase II de pequeño tamaño, alrededor de 1 kb, con repeticiones terminales invertidas de 12 bp (Marín y Fontdevila 1995). Se han clonado y secuenciado tres copias de este elemento en *D. koepferae* y, a pesar de que muestran pequeñas diferencias en su secuencia, en ningún caso se han encontrado marcos de lectura abiertos de más de 103 aminoácidos. Además, no se han encontrado tampoco regiones con homología a las transposasas de los transposones conocidos. Por tanto, se desconoce si este elemento tiene capacidad codificadora y en algunos casos se ha puesto en duda su naturaleza transponible. No obstante, la hibridación de Southern reveló que *Gandalf* está presente en muchas especies del complejo *buzzatii*, entre ellas *D. buzzatii*, aunque en este caso en un número de copias muy bajo.

Asimismo, el estudio anterior permitió detectar fragmentos de DNA genómico de *D. koepferae* con homología a elementos transponibles previamente descritos en otras especies, como *gypsy* de *D. melanogaster* y *T1Ag* y *Q* de *A. gambiae* (Marín y Fontdevila 1996). Además, en *D. buzzatii* se han clonado y secuenciado dos retrotransposones que están presentes en un elevado número de especies de *Drosophila*: *copia* (Francino *et al.* 1993; O. Francino y O. Cabré, resultados no publicados) y *R1* (Lathe III *et al.* 1995).

1.5 Objetivos

El presente trabajo se enmarca dentro del estudio de uno de los principales tipos de variación genética, la variación a nivel cromosómico. Específicamente, se ha centrado en dos cuestiones fundamentales sobre el polimorfismo de inversiones en el género *Drosophila*: (1) el origen molecular de las inversiones; y (2) los factores implicados en su mantenimiento en las poblaciones naturales.

Uno de los objetivos básicos de esta tesis era tratar de determinar los mecanismos moleculares mediante los que se generan los cambios cromosómicos estructurales presentes en la naturaleza. Se ha pretendido investigar cómo surgen los cambios cromosómicos evolutivamente exitosos y, en concreto, cómo se han originado las inversiones presentes en el género *Drosophila*. Los elementos transponibles habían mostrado ampliamente que eran capaces de producir inversiones en *Drosophila* en el laboratorio, en la mayoría de casos por recombinación ectópica entre copias situadas en orientación opuesta en distintas posiciones de un mismo cromosoma (Lim y Simmons 1994). Sin embargo, las radiaciones ionizantes también habían mostrado ser capaces de inducir la producción de inversiones (Federer *et al.* 1967; Tonzetich *et al.* 1988). Poca era la información proveniente de inversiones naturales y no estaba claro hasta que punto estos diferentes mecanismos eran relevantes en la naturaleza. Los únicos dos ejemplos existentes en los que se habían clonado y secuenciado los puntos de rotura de inversiones naturales no habían mostrado evidencias de la acción de elementos transponibles y no habían proporcionado pista alguna sobre su mecanismo de generación (Wesley y Eanes 1994; Cirera *et al.* 1995). Por tanto, se desconocía aún que causa las inversiones en las poblaciones naturales de *Drosophila* y su origen seguía siendo un misterio.

De esta forma, como modelo, nos propusimos caracterizar molecularmente los puntos de rotura de la inversión *2j* de *D. buzzatii*, que es la inversión polimórfica más

ampliamente distribuida y más frecuente de esta especie. En relación con el origen de las inversiones, los objetivos concretos eran:

- a) Determinar la secuencia nucleotídica de los puntos de rotura de la inversión 2j de *D. buzzatii* en la ordenación invertida y no invertida.
- b) Tratar de averiguar cuál fue el mecanismo molecular por el que se originó esta inversión.
- c) Analizar la variabilidad molecular en la región de los puntos de rotura para determinar si la inversión tiene un origen filogenético único o múltiple.
- d) Identificar posiciones cromosómicas que acumulan un gran número de puntos de rotura de inversiones.

La mayor parte de estos objetivos se llevaron a cabo en dos artículos estrechamente relacionados. En el primero de los artículos presentados, "*Generation of a widespread Drosophila inversion by a transposable element*" (M. Cáceres, J. M. Ranz, A. Barbadilla, M. Long y A. Ruiz; *Science* **285**: 415-418, 1999), se describe el proceso que condujo a la clonación y secuenciación de los puntos de rotura de la inversión 2j de *D. buzzatii*. Además, se propone el posible mecanismo por el que surgió la inversión, mediante recombinación ectópica entre dos copias en orientación invertida de lo que parece ser un elemento transponible nuevo.

En el segundo de los artículos presentados, "*Natural hotspots induced by transposable element insertion at the breakpoints of a Drosophila inversion*" (M. Cáceres, M. Puig y A. Ruiz; enviado), se profundiza en el estudio de las inserciones encontradas en los puntos de rotura de la inversión en diferentes líneas 2j. En este trabajo se describe la existencia de una gran variación estructural en los puntos de rotura que sugiere que podría tratarse de puntos genéticamente inestables. Además, el análisis de la variabilidad nucleotídica en los puntos de rotura confirma el carácter monofilético de la inversión y proporciona una primera estima de su edad.

Por otra parte, en esta tesis se ha intentado también aportar información sobre los procesos implicados en el mantenimiento de las inversiones en las poblaciones naturales de *Drosophila*. A pesar de los muchos años de estudio y de la cantidad de estudios realizados sobre el polimorfismo de inversiones en *Drosophila*, quedaban importantes cuestiones por resolver. Varias evidencias indicaban que algunas inversiones parecían tener un carácter adaptativo, pero no estaba totalmente claro cuál era la base de su efecto sobre la eficacia biológica de los individuos portadores (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997). Se consideraba que el efecto principal de las inversiones es la reducción de la recombinación en el segmento invertido y la conservación del conjunto de alelos atrapado en su interior (Dobzhansky 1970). En este sentido, ciertos estudios sobre la longitud de las inversiones en unas pocas especies se habían tomado como evidencia a favor de esta idea (Krimbas y Powell 1992). Sin embargo, faltaba una confirmación de tipo más general. Además, en ninguno de los estudios previos se había tenido en cuenta el efecto de la recombinación sobre la longitud de las inversiones. Por último, la posibilidad de que las mutaciones causadas en los puntos de rotura en el momento de la generación de la inversión pudiesen desempeñar un papel importante en determinados casos o en las fases iniciales de la vida de la inversión no estaba completamente descartada (Wesley y Eanes 1994).

Así, en primer lugar, a partir de la secuenciación de los puntos de rotura de la inversión *2j*, nos propusimos investigar si la aparición de esta inversión había afectado de forma directa a algún gen. Paralelamente, nos propusimos realizar un estudio completo sobre la longitud de las inversiones naturales y su relación con el éxito evolutivo de las mismas. Este estudio se llevó a cabo primero a nivel del complejo *buzzatii* y posteriormente se extendió a todo el género *Drosophila*. En concreto, los objetivos eran:

- a) Investigar si la inversión *2j* de *D. buzzatii* ha afectado a algún gen cercano a sus puntos de rotura.

- b) Analizar la distribución de la longitud de las inversiones descritas en el complejo *buzzatii* y comprobar si existen diferencias en la longitud entre inversiones con distinto grado de éxito evolutivo.
- c) Determinar si existen diferencias en la distribución de la longitud de las inversiones exitosas entre especies y si estas diferencias pueden ser explicadas por la variación en la tasa de recombinación entre ellas.

La consecución de los anteriores objetivos se describe principalmente en dos artículos más. El estudio detallado de la longitud de las inversiones de las especies del complejo *buzzatii* y de la distribución de sus puntos de rotura se describe en el tercero de los artículos presentados, "*Inversion length and breakpoint distribution in the Drosophila buzzatii species complex: Is inversion length a selected trait?*" (M. Cáceres, A. Barbadilla y A. Ruiz; *Evolution* **51**: 1149-1155, 1997). En este trabajo se muestra que existen diferencias claras en la distribución de la longitud de varios grupos de inversiones con distinto éxito evolutivo y que las inversiones evolutivamente exitosas tienden a tener un tamaño intermedio.

Por último, en el cuarto de los artículos presentados, "*Recombination rate predicts inversion size in Diptera*" (M. Cáceres, A. Barbadilla y A. Ruiz; *Genetics* **153**: 251-259, 1999), se lleva a cabo un análisis conjunto del tamaño de las inversiones y de la longitud genética del genoma en diferentes especies de *Drosophila* y *Anopheles*. Este análisis permite postular la existencia de una longitud genética óptima para las inversiones evolutivamente exitosas y apunta a que de forma generalizada la ventaja selectiva de las inversiones radica en su efecto sobre la recombinación.

2. ARTÍCULOS

ARTÍCULO I

Generation of a Widespread Drosophila Inversion by a Transposable Element

Mario Cáceres, José María Ranz, Antonio Barbadilla, Manyuan Long
y Alfredo Ruiz

1999

Science **285**: 415-418

would have to have developed only in the $\beta_2M^{-/-}$ →B6 and C3H.SW→B6 chimeras, and not in the B6→B6 controls. In any event, if suppressor cells were involved, host APCs would be required for their development.

These results contrast with those of Sprent and colleagues (4). They found that in a CD8⁺ T cell-dependent MHC-compatible but miHA-incompatible GVHD model donor→host bone marrow chimeras that were reirradiated and injected with donor bone marrow and lymph node cells developed GVHD. They suggested that either the chimeras were not devoid of host APCs because of insufficient radiation, or that host antigens were processed by donor marrow cells. Our data support the former explanation. Sprent and colleagues also found that heavily irradiated allogeneic parent→F₁ bone marrow chimeras developed GVHD in response to very high doses of parental T cells (up to 8×10^7 T cells per recipient) and concluded that nonhematopoietic cells functioned as APCs. Given the increased precursor frequency of T cells recognizing allogeneic MHC molecules in comparison to miHA on self MHC and the large dose of T cells used, they may have unmasked the presence of small numbers of residual host APCs. Alternatively, nonhematopoietic cells when challenged with very large numbers of allogeneic T cells may cause T cell activation sufficient to induce GVHD.

Our results suggest that depleting host APCs before the conditioning regimen should abrogate GVHD without the need for prolonged T cell-targeted immunosuppression. Such an approach, perhaps using toxin-conjugated or radiolabeled antibodies, could expand the range of diseases treated with alloBMT. To test the feasibility of in vivo antibody-mediated depletion of host dendritic cells, we injected mice with N418, a hamster monoclonal antibody to the β integrin CD11c expressed on murine dendritic cells (17). The injected anti-CD11c bound to all CD11c-expressing dendritic cells in both lymph node and spleen (15), supporting the feasibility of antibody-mediated APC depletion.

A subset of alloBMT recipients have self-limited GVHD, which was presumed to reflect acquired T cell tolerance. Our data suggest another explanation: replacement of host with donor APCs abrogates T cell activation. Infusions of T cells from original bone marrow donors given to relapsed leukemia patients months to years after the initial alloBMT (18) cause less GVHD than has been observed when T cells are given at the time of transplantation (19). Although there may be other explanations (20), we suggest that the replacement of host with donor APCs reduces the chance of a donor CD8⁺ T cell interacting with a GVHD-inducing host APC. In addition to suggesting explanations for these clinical observations, our data provide the

foundation for a different strategy for reducing GVHD-host APC depletion. This approach may avoid the problems associated with T cell depletion of marrow allografts: failure of engraftment, poor immune reconstitution, and lack of immunoreactivity against the tumor.

References and Notes

1. M. M. Horowitz et al., *Blood* **75**, 555 (1990); D. Gaziev et al., *Transplantation* **63**, 854 (1997).
2. J. M. M. den Haan et al., *Science* **268**, 1476 (1995); W. Wang et al., *ibid.* **269**, 1588 (1995).
3. N. J. Chao and H. J. Deeg, in *Graft-Vs.-Host Disease*, J. L. Ferrara, H. J. Deeg, S. J. Burakoff, Eds. (Dekker, New York, 1996), pp. 639–666.
4. R. Korngold and J. Sprent, *J. Exp. Med.* **155**, 872 (1982); H. Kosaka, C. D. Surh, J. Sprent, *ibid.* **176**, 1291 (1992).
5. J. W. Streilein and R. E. Billingham, *ibid.* **132**, 163 (1977); *ibid.*, p. 181.
6. P. L. Weiden et al., *J. Immunol.* **116**, 1212 (1976).
7. R. L. Truitt and A. A. Atasoylu, *Blood* **77**, 2515 (1991).
8. R. N. Germain, *Cell* **76**, 287 (1994).
9. P. Matzinger and M. J. Bevan, *Cell. Immunol.* **33**, 92 (1977); M. J. Bevan, *J. Exp. Med.* **182**, 639 (1995); P. K. Srivastava, H. Udono, N. E. Blachere, Z. Li, *Immunogenetics* **39**, 93 (1994); C. Kurts, J. F. Miller, R. M. Subramaniam, F. R. Carbone, W. R. Heath, *J. Exp. Med.* **188**, 409 (1998).

10. M. L. Albert, B. Sauter, N. Bhardwaj, *Nature* **392**, 86 (1998).
11. B6 mice received 200 μ g of antibody to NK1.1 (anti-NK1.1) on consecutive days (21) followed by two 500-cGy fractions and injection with 10^7 T^{CR α} BM cells from B6 $\beta_2M^{-/-}$ or B6 mice. Four months later, chimeras were reirradiated with two 375-cGy fractions, then injected with 7×10^6 C3H.SW T^{CR α} BM cells with or without 10^6 or 2×10^6 C3H.SW CD8⁺ T cells.
12. B. H. Koller, P. Marrack, J. W. Kappler, O. Smithies, *Science* **248**, 1227 (1990).
13. T cells were depleted with anti-Thy1.2-labeled magnetic microbeads (Miltenyi Biotec) according to the manufacturer's protocol.
14. CD8⁺ T cells were isolated by staining lymph node cells with biotin-labeled antibodies to CD4, B220, and CD11b, followed by depletion with streptavidin-conjugated magnetic beads (Miltenyi Biotec). Ninety-five percent of cells were CD8⁺; less than 0.25% of cells were CD4⁺.
15. W. Shlomchik, data not shown.
16. T. Maier and H. N. Claman, in (3), pp. 75–94; H. J. Deeg and M. Cottler-Fox, in (3), pp. 311–335.
17. J. P. Metlay et al., *J. Exp. Med.* **171**, 1753 (1990).
18. H. J. Kolb et al., *Blood* **86**, 2041 (1995).
19. K. M. Sullivan et al., *N. Engl. J. Med.* **320**, 828 (1989).
20. J. L. Ferrara, *Bone Marrow Transplant.* **14**, 183 (1994).
21. M. Bix et al., *Nature* **349**, 329 (1991).
22. Supported by P50HL-54516 from the National Heart, Lung and Blood Institute and the Leukemia Society of America.

3 March 1998; accepted 16 June 1999

Generation of a Widespread *Drosophila* Inversion by a Transposable Element

Mario Cáceres,^{1,2*} José María Ranz,¹ Antonio Barbadilla,¹ Manyuan Long,² Alfredo Ruiz¹

Although polymorphic inversions in *Drosophila* are very common, the origin of these chromosomal rearrangements is unclear. The breakpoints of the cosmopolitan inversion 2j of *D. buzzatii* were cloned and sequenced. Both breakpoints contain large insertions corresponding to a transposable element. It appears that the two pairs of target site duplications generated upon insertion were exchanged during the inversion event, and that the inversion arose by ectopic recombination between two copies of the transposon that were in opposite orientations. This is apparently the mechanism by which transposable elements generate natural inversions in *Drosophila*.

Most *Drosophila* species are naturally polymorphic for inversions in one or more chromosomes (1, 2), but little is known about the molecular mechanisms underlying the generation of these inversions. Indirect evidence suggests that transposable elements (TEs) cause *Drosophila* inversions: TEs mediated chromosomal rearrangements in laboratory populations (3), and in situ hybridization studies detected the transposon *hobo* around the breakpoints of four endemic inversions of *D. melanogaster*

(4). However, in *D. subobscura* and *D. pseudoobscura* there was no cytological association between middle repetitive sequences and inversion breakpoints (1, 2), and the two direct studies that sequenced the breakpoints of naturally occurring inversions did not detect any TE (5).

The species *D. buzzatii* belongs to the *D. repleta* group of the *Drosophila* subgenus (6). Two chromosomal arrangements are commonly observed in chromosome 2 of *D. buzzatii*: the ancestral one or 2 *standard* (2st), and the 2j, which derived from the 2st by inversion 2j and is distributed throughout the species range at high frequencies (7). Here, we cloned and sequenced the breakpoints of inversion 2j of *D. buzzatii*. First, the region of the proximal breakpoint in the 2st chromosome (designated as

¹Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. ²Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA.

*To whom correspondence should be addressed (in Spain). E-mail: mariocs@cc.uab.es

REPORTS

CD) was cloned from a line homozygous for the 2st arrangement (st-1). Then the proximal and distal breakpoints in the 2j chromosome (designated as BD and AC, respectively) were isolated from a line homozygous for the 2j arrangement (j-1). Finally, the region of the distal breakpoint in the 2st chromosome (named AB) was obtained from the st-1 line.

The proximal breakpoint of inversion 2j in the 2st arrangement lies between the *nAcR β -96A* and *Pp1 α -96A* genes (Fig. 1A), which are 4 kb apart in *D. melanogaster* (8). A cDNA clone of the *D. melanogaster nAcR β -96A* gene (9) was used to screen a st-1 lambda genomic library (10). Three positive clones were identified, and one of these, λ st9, contained the *Pp1 α -96A* gene (Fig. 1A) and spanned the breakpoint (Fig. 1B). By sequential subcloning of λ st9 and in situ hybridization of the different subclones to 2st and 2j chromosomes, the CD breakpoint was further located in a 2-kb Hind III fragment, pGPE102.2, and a 0.9-kb Dra I fragment, pGPE102.2.2 (Fig. 1A). The AC and BD breakpoints in the 2j arrangement were recovered by screening a j-1 lambda genomic library with pGPE102.2 used as a probe. Five positive phages were identified, and the two with the most divergent restriction patterns, λ j3.1 and λ j10, were selected. In situ hybridization of both phages to st-1 and j-1 chromosomes yielded multiple signals (Fig. 1B), showing that they include repetitive sequences. However, the stronger signals on 2j chromosomes at the proximal breakpoint with λ j3.1 and at the distal breakpoint with λ j10 indicated that the inserts of these phages came from the BD and AC regions, respectively. Fragments of λ j3.1 containing D sequences (pGPE105.4.1) and of λ j10 containing C sequences (pGPE107.2.1.1) were identified by hybridization of digested DNA from each phage with a probe of the CD breakpoint (Fig. 1A). Next, a series of contiguous fragments of λ j3.1 adjacent to pGPE105.4.1 and of λ j10 adjacent to pGPE107.2.1.1 were subcloned. Most showed a repetitive in situ hybridization pattern, similar to that of the entire phages. A 1.2-kb Pst I–Bam HI fragment of λ j3.1 (pGPE109.2), which contained B sequences and gave a strong signal at the AB region of st-1 chromosomes, was found 3.3 kb away from pGPE105.4.1 (Fig. 1A). Likewise, a 2.4-kb Pst I–Xba I fragment of λ j10 (pGPE107.3), containing A sequences and giving a unique signal in the AB region of st-1 chromosomes, was found 1.9 kb away from pGPE107.2.1.1 (Fig. 1A). All these fragments were sequenced. A 95% identity at the amino acid level was found between the proximal end of pGPE107.3 and the 44 COOH-terminal residues of the *D. melanogaster* Rox8 protein (11), thus revealing the location of the *D. buzzatii* *rox8* gene (Fig. 1A). The AB region in the 2st arrangement

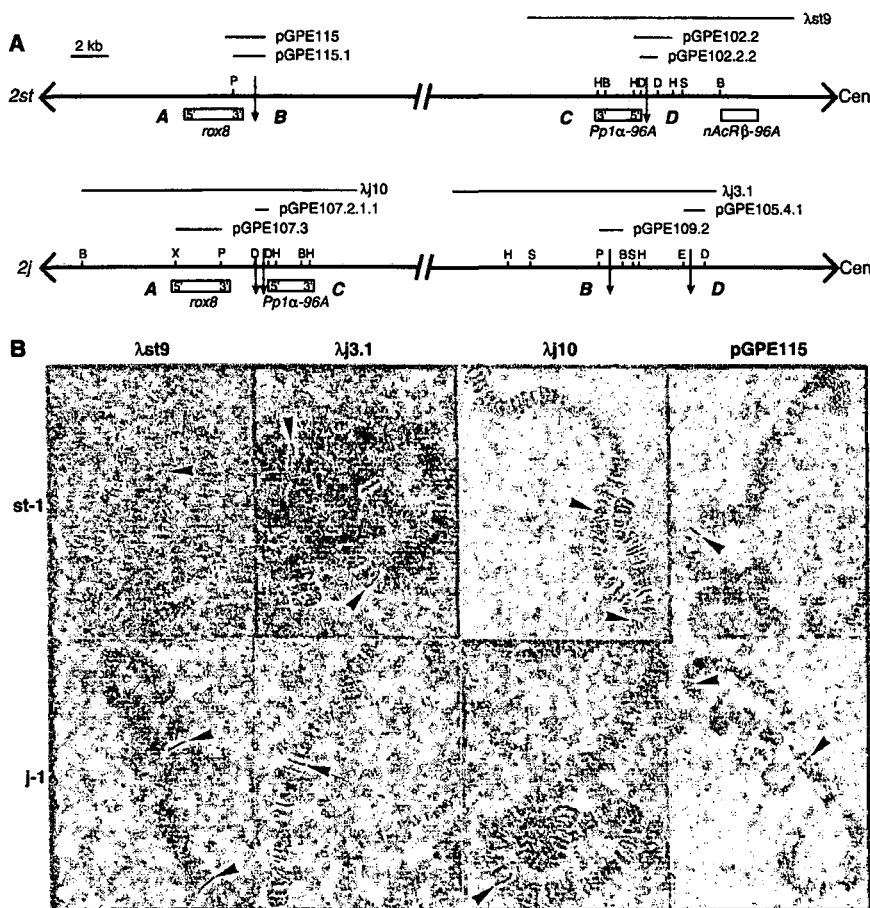


Fig. 1. Cloning strategy and in situ hybridization of the regions spanning the breakpoints of inversion 2j. (A) Physical map of 2st and 2j chromosomes showing the breakpoint regions (AB and CD in 2st and AC and BD in 2j) and different clones obtained (thin lines shown above map). Vertical arrows mark the limits of the A, B, C, and D regions (in 2j chromosomes there are insertions between A and C and between B and D). Empty boxes below map represent the genes flanking the breakpoints according to the sizes of the respective cDNAs in *D. melanogaster*. The direction of transcription, when known, is indicated as 5' and 3'. Cen, centromere; B, Bam HI; D, Dra I; E, Eco RI; H, Hind III; P, Pst I; S, Sal I; X, Xba I. (B) In situ hybridization to the salivary gland chromosomes of the st-1 and j-1 *D. buzzatii* lines of the four clones containing the breakpoint regions of inversion 2j (arrowheads). Clones spanning the breakpoints in the 2st arrangement (λ st9 and pGPE115) give one signal in st-1 and hybridize to both breakpoints in j-1. Clones spanning the breakpoints in the 2j arrangement (λ j3.1 and λ j10) give two signals in st-1 and one signal in j-1.

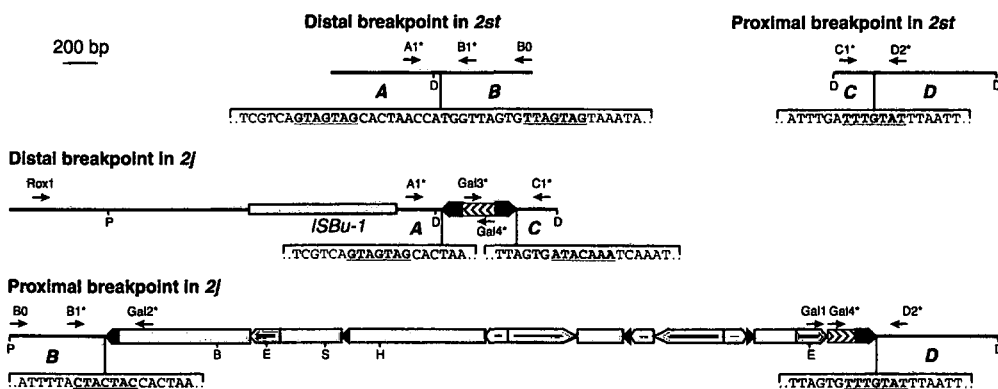
was finally obtained by polymerase chain reaction (PCR) amplification from st-1 genomic DNA using primers Rox1 and B0, located at the 3' end of the *rox8* coding region and the distal end of pGPE109.2, respectively. In situ hybridization to st-1 and j-1 chromosomes confirmed that the resulting 2.1-kb PCR product (pGPE115) spanned the 2j distal breakpoint (Fig. 1B).

Alignment of AB and CD sequences in st-1 (2086 nucleotides) and AC and BD sequences in j-1 (8668 nucleotides) showed that there are large insertions at both 2j inversion breakpoints that are not present in the 2st arrangement (Fig. 2). The insertion between A and C is 392 base pairs (bp) long and has inverted terminal repeats (ITRs) of 106 bp. The insertion between B and D is 4319 bp long, with a complex internal structure. It has

ITRs like those of the AC insertion, of 106 and 47 bp, the latter corresponding to the beginning of the 106-bp repeat. The central 180 bp of the AC insertion and the BD sequence that follows the proximal ITR are homologous (95% identity) but are oppositely oriented (Fig. 2). Flanking each insertion are 7-bp sequences that resemble target site duplications that were produced during the insertion event. However, the duplicated sequences are separated and inverted (Fig. 2), which strongly suggests that inversion 2j appeared by intrachromosomal pairing and recombination between the two homologous sequences inserted in distant sites in opposite orientations (Fig. 3). Although the precise structure of the original insertions is not known, they were likely homologous at least over a span of 274 bp, and this relatively

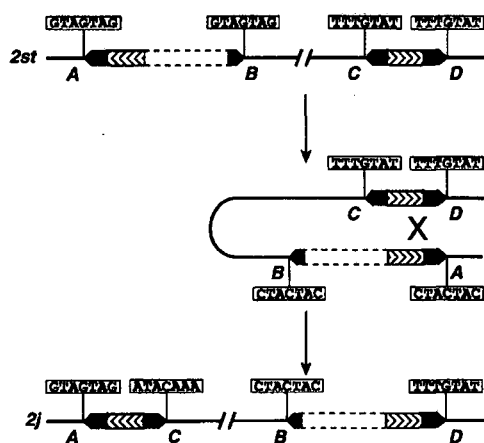
REPORTS

Fig. 2. Features of the sequenced regions of the distal and proximal breakpoints of inversion *2j* in the *2st* and *2j* arrangements. Horizontal lines correspond to the A, B, C, and D sequences, and insertions within them are represented as colored boxes. The insertions at the breakpoints are delimited by black blocks corresponding to the ITRs. Orientation of the homologous sequences of the two insertions is indicated by angle bracket patterns (<<<< >>>>). Internal repeats within the BD insertion are represented as sharp-pointed blocks with different colors. *ISBu-1* is a nondescribed, highly dispersed repeated sequence inserted in the *2j* arrangement in the A region, 256 bp away from the insertion at the breakpoint. Details of the breakpoint sequences in st-1 and j-1 are shown below each diagram (A and B sequences in blue, C and D sequences in red, and insertions in black). The target site duplications generated during insertion are underlined and in boldface. The distal breakpoint in *2st* shows a structure with 7-bp imperfect (one nucleotide difference) direct repeats flanking two adjacent 9-bp inverted repeats, which resembles



the footprints left by class II TEs after excision and incomplete repair of the double strand break (30). The A sequence in the *2j* arrangement is that of the j-14 line, because this sequence is absent in the j-1 line as the result of a 17-bp deletion. Horizontal arrows represent different PCR primers; those used for sequencing are marked with an asterisk. (Sequences of all primers can be found at Science Online, www.sciencemag.org/feature/data/1039043.shl.) Codes for restriction sites are as in Fig. 2.

Fig. 3. Model for the generation of inversion *2j*. First, the two homologous sequences became inserted between the A and B sequences and between the C and D sequences of a *2st* chromosome, generating direct duplications of seven nucleotides (boxed sequences). Then, ectopic recombination between the homologous sequences in opposite orientation generated the inversion. As a result of this process, target site duplications were exchanged between the insertions and acquired their present orientation in *2j* chromosomes, with each insertion flanked by its original target site and an inverted target site of the other element. Dashed lines refer to sequences not drawn to scale.



short sequence appears to be long enough to sustain ectopic recombination in *Drosophila* (12).

The duplication of a 7-bp sequence upon insertion, the presence of ITRs, and the moderately repetitive in situ hybridization pattern identify the breakpoint insertions as copies of a class II TE (13) that we designated as *Galileo* (14). Furthermore, we have observed variation among *D. buzzatii* lines both in the number of in situ hybridization signals and in their chromosomal location, which suggests recent transpositional activity of this TE (15). The AC insertion is probably a defective copy that derived from a complete element by an internal deletion but retained the terminal sequences needed for transposition (16). The BD copy may have degenerated partially, but it contains open reading frames (ORFs) that could code for a protein involved in transposition. The largest ORF has a predicted sequence of 172 amino acids, displaying 32% identity with eight gaps (53% similarity) to the putative transposase of the transposon *Gandalf* of *D. koepferae*, a

sibling species of *D. buzzatii* (17). Ectopic recombination between repeated DNA sequences has been implicated in inversion generation in diverse organisms, such as yeast (18) and humans (19), and TE activity could stimulate this recombination (20). However, precedents of TE involvement in the generation of naturally occurring inversions were restricted to bacteria (21). A TE-like sequence was found at one breakpoint of a polymorphic inversion of *Anopheles arabiensis*, but its implication in the origin of the inversion was not clear (22). Evidence from inversion *2j* confirms the role of TEs in *Drosophila* chromosomal evolution.

The previous results were corroborated by PCR amplification of the breakpoint regions and sequencing of the PCR products in *2st* lines and *2j* lines of different origin (23). Primer pairs A1-B1 and C1-D2 (Fig. 2) were used with genomic DNA of four *2st* lines, whereas primer pairs A1-C1, B1-Gal2, and Gal1-D2 (Fig. 2) were used with DNA from 21 *2j* lines. In each case PCR products of the appropriate sizes

were obtained, indicating that the breakpoint insertions are present in all *2j* lines and absent in all *2st* lines (in the amplification of the AC segment only, j-9 and jq⁷-2 showed larger PCR products than expected, resulting apparently from additional insertions, and j-10 yielded no PCR product). We sequenced 521 bp of the A, B, C, and D regions in three *2st* lines and five *2j* lines and 800 bp of the *Galileo* insertions in the *2j* lines (24). All the *2j* chromosomes in this geographically diverse set had the same insertions and identical breakpoints with regard to *2st* chromosomes. Nucleotide diversity (25) over the 489 aligned sites in the eight sequences (excluding gaps) was 0 and 0.0279 for the *2j* and *2st* arrangements, respectively, and *2j* chromosomes shared 22 fixed differences when compared to *2st* chromosomes (including six small insertions of 1 to 4 bp). *Galileo* sequences showed more variation among the *2j* lines, with a nucleotide diversity over the 794 aligned sites of 0.0053. Thus, it appears that despite the transposon-mediated origin, all *2j* arrangements descend from the product of a single ectopic recombination event: They are monophyletic.

Finally, we tested the effect of the inversion on the nearby genes *rox8* and *Pp1 α -96A* (Fig. 1A). The *Pp1 α -96A* transcriptional startpoint (26) is just 156 bp away from the proximal breakpoint in the *2st* arrangement, and a semi-quantitative RT-PCR was carried out to compare *Pp1 α -96A* expression between the st-1 and j-1 lines at four developmental stages (27). No apparent differences in expression between st-1 and j-1 were seen. As in *D. melanogaster* (26), the *Pp1 α -96A* transcript was present in embryos, was detected with a lower expression in adults, and was barely visible in larvae and pupae. We also performed Northern (RNA) analyses (28) using probes of the AB (pGPE115.1) and CD regions (pGPE102.2).

REPORTS

Both probes hybridized to the transcripts of *Pp1 α -96A* [2.3 kb (26)] and *rox8* [3 to 3.3 kb (11)] only, and no detectable differences in the expression of these genes between st-1 and j-1 were observed. It seems that inversion 2j neither disrupted any transcriptional unit nor affected the expression of the closest genes, ruling out a relation between the mutational effect of the inversion and its adaptive value (29).

References and Notes

- C. B. Krimbas and J. R. Powell, Eds., *Drosophila Inversion-Polymorphism* (CRC Press, Boca Raton, FL, 1992).
- J. R. Powell, *Progress and Prospects in Evolutionary Biology: The Drosophila Model* (Oxford Univ. Press, New York, 1997).
- W. R. Engels and C. R. Preston, *Genetics* **107**, 657 (1984); S. Schneuwly, A. Kuroiwa, W. J. Gehring, *EMBO J.* **6**, 201 (1987); E. A. Montgomery, S.-M. Huang, C. H. Langley, B. H. Judd, *Genetics* **129**, 1085 (1991); J. K. Lim and M. J. Simmons, *Bioessays* **16**, 269 (1994); V. Ladeveze, S. Aulard, N. Chaminade, G. Periquet, F. Lemeunier, *Proc. R. Soc. London Ser. B* **265**, 1157 (1998).
- T. W. Lyttle and D. S. Haymer, *Genetica* **86**, 113 (1992).
- C. S. Wesley and W. F. Eanes, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3132 (1994); S. Cirera, J. M. Martín-Campos, C. Segarra, M. Aguadé, *Genetics* **139**, 321 (1995).
- M. Wasserman, in (7), pp. 455-541; A. Ruiz and M. Wasserman, *Heredity* **70**, 582 (1993).
- A. Fontdevila, A. Ruiz, G. Alonso, J. Ocaña, *Evolution* **35**, 148 (1981); A. Ruiz, H. Naveira, A. Fontdevila, *Genet. Iber.* **36**, 13 (1984); E. Hasson et al., *J. Evol. Biol.* **8**, 369 (1995).
- J. M. Ranz, C. Segarra, A. Ruiz, *Genetics* **145**, 281 (1997); J. M. Ranz, M. Cáceres, A. Ruiz, *Chromosoma* **108**, 32 (1999).
- E. Sawruk, C. Udri, H. Betz, B. Schmitt, *FEBS Lett.* **273**, 177 (1990).
- For the construction of genomic libraries, ~30 μ g of high molecular weight genomic DNA were partially digested with *Sau* 3A1 and electrophoresed on a 0.4% agarose gel. Fragments in the size range 15 to 23 kb were eluted from the gel, ligated with LambdaGEM-11 Bam HI arms, and packaged in vitro using Packaging extracts (Promega).
- S. Brand and H.-M. Bourbon, *Nucleic Acids Res.* **21**, 3699 (1993).
- I. Busseau, A. Pellisson, A. Bucheton, *Mol. Gen. Evol.* **218**, 222 (1989); N. Nassif and W. R. Engels, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1262 (1993).
- D. J. Finnegan, *Trends Genet.* **5**, 103 (1989); P. Cappy, C. Bazin, D. Higuier, T. Langin, *Dynamics and Evolution of Transposable Elements* (Springer-Verlag, Heidelberg, 1998).
- The insertions at the 2j inversion breakpoints do not show homology at the nucleotide level to any known sequence in the databases and thus appear to correspond to a previously undescribed TE. We have named this TE after the Italian astronomer, mathematician, and physicist Galileo Galilei (1564-1642).
- A 5-kb probe of the λ j3.1 proximal end that contains most of the 8D insertion was in situ hybridized to the polytene chromosomes of seven *D. buzzatii* lines. The total number of euchromatic signals found per genome was 55 in st-1, 44 in st-3, 47 in st-4, 34 in j-1, 34 in j-8, 37 in j-11, and 28 in j-13 (23). On average, 11.8% of each line's signals were exclusive.
- K. O'Hare and G. M. Rubin, *Cell* **34**, 25 (1983); R. D. Streck, J. E. MacGaffey, S. K. Beckendorf, *EMBO J.* **5**, 3615 (1986).
- I. Marín and A. Fontdevila, *Mol. Gen. Evol.* **248**, 423 (1995).
- G. S. Roeder and C. R. Fink, in *Mobile Genetic Elements*, J. Shapiro, Ed. (Academic Press, Orlando, FL, 1983), pp. 299-328; T. D. Petes and C. W. Hill, *Annu. Rev. Genet.* **22**, 147 (1988).
- D. Lakich, H. H. Kazanian, S. E. Antonarakis, J. Gitschier, *Nature Genet.* **5**, 236 (1993); J. A. Naylor et al., *Hum. Mol. Genet.* **4**, 1217 (1995); K. Small, J. Iber, S. T. Warren, *Nature Genet.* **16**, 96 (1997).
- W. B. Eggleston, N. R. Rim, J. K. Lim, *Genetics* **144**, 647 (1996).
- Y. Komoda, M. Enomoto, A. Tominaga, *ibid.* **129**, 639 (1991); M.-L. Daveran-Mingot, N. Campo, P. Ritzenhaller, P. le Bourgeois, *J. Bacteriol.* **180**, 4834 (1998).
- K. D. Mathiopoulos, A. della Torre, V. Predazzi, V. Petrarca, M. Coluzzi, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12444 (1998).
- The *D. buzzatii* lines used are homokaryotypic for four different chromosome arrangements: 2st, 2j, 2jz³, and 2jq⁷ (2jz³ and 2jq⁷ derive from the 2j arrangement and bear inversions 2z³ and 2q⁷, respectively). Their geographic origins cover a major part of the species distribution; lines that were sequenced are indicated in boldface. Spain: st-1, st-2, j-1, j-2, j-3, j-4, j-5, j-6, jz³-1, jz³-2, and jq⁷-1 (Carboneras), j-7 and jq⁷-3 (Caldetas), and jq⁷-2 (Mogan, Canary Islands). Tunisia: jz³-3 (Kariouan). Argentina: st-3 (Vipos), j-8 (San Luis), j-9 (Quilmes), j-10 (Palo Labrado), and jq⁷-4 (Otamendi). Bolivia: j-11 (Los Negros). Brazil: st-4, j-12, and j-13 (Guaritas, RS). Australia: j-14. All lines (except jq⁷-3 and jq⁷-4) were made isogenic by repeated sib-mating or by using the second chromosome balancer stock *Antp/Δ5* [J. S. F. Barker, *Genetica* **92**, 165 (1994)].
- The PCR products were purified from 1% agarose gels, reamplified with one of the sequencing primers shown in Fig. 2, and sequenced in an ALFexpress DNA automated sequencer (Pharmacia Biotech). Both strands of each fragment were sequenced completely using primers separated by 323 to 466 nucleotides.
- M. Nei, *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York, 1987).
- V. Dombrádi et al., *Eur. J. Biochem.* **194**, 739 (1990).
- Two sets of primers were used to amplify simultaneously a 271-bp 5' fragment of the *Pp1 α -96A* cDNA and 438 bp of the *Gapdh* cDNA (internal control)

from the total cDNA. In both cases the primer sites were selected to span introns (of 61 bp and 69 bp, respectively). *Gapdh* primers were designed on the basis of the sequence of *D. hydei* [K. M. Wojtas, L. von Kalm, J. R. Weaver, D. T. Sullivan, *Genetics* **132**, 789 (1992)]. PCRs were performed with varying amounts of the reverse transcription reaction product as template and different concentrations of the two sets of primers.

- The Northern blot hybridization was done with ~30 μ g of total RNA from embryos, larvae, pupae, and adults of the st-1 and j-1 lines as described [A.-K. Rost, in *Quantitation of mRNA by Polymerase Chain Reaction: Nonradioactive PCR Methods*, Th. Köhler et al., Eds. (Springer-Verlag, Berlin, 1995), pp. 93-114].
- A. Ruiz et al., *Genetics* **128**, 739 (1991); E. Betrán, M. Santos, A. Ruiz, *Evolution* **52**, 144 (1998).
- W. R. Engels, D. M. Johnson-Schlitz, W. B. Eggleston, J. Sved, *Cell* **62**, 515 (1990); R. H. A. Plasterk, *EMBO J.* **10**, 1919 (1991); E. Takasu-Ishikawa, M. Yoshihara, Y. Hotta, *Mol. Gen. Evol.* **232**, 17 (1992).
- Supported by a doctoral fellowship and a travel grant from the Comissionat per a Universitats i Recerca (Generalitat de Catalunya, Spain) (M.C.) and grant PB95-0607 from the Dirección General de Investigación Científica y Técnica (Ministerio de Educación y Cultura, Spain) (A.R.). We thank J. Zhang for his help in the RNA work; J. Barbé, O. Cabré, and A. Tapias for material and technical help; B. Schmitt for the *nACR-96A* clone; J. S. F. Barker, E. Betrán, A. Fontdevila, and F. M. Sene for *D. buzzatii* stocks; and M. Aguadé, M. Ashburner, A. Berry, F. Brunet, B. Charlesworth, and R. de Frutos for valuable comments and helpful discussions.

4 February 1999; accepted 9 June 1999

Identification of a Vertebrate Sister-Chromatid Separation Inhibitor Involved in Transformation and Tumorigenesis

Hui Zou, Thomas J. McGarry, Teresita Bernal, Marc W. Kirschner*

A vertebrate securin (vSecurin) was identified on the basis of its biochemical analogy to the Pds1p protein of budding yeast and the Cut2p protein of fission yeast. The vSecurin protein bound to a vertebrate homolog of yeast separins Esp1p and Cut1p and was degraded by proteolysis mediated by an anaphase-promoting complex in a manner dependent on a destruction motif. Furthermore, expression of a stable *Xenopus* securin mutant protein blocked sister-chromatid separation but did not block the embryonic cell cycle. The vSecurin proteins share extensive sequence similarity with each other but show no sequence similarity to either of their yeast counterparts. Human securin is identical to the product of the gene called pituitary tumor-transforming gene (*PTTG*), which is overexpressed in some tumors and exhibits transforming activity in NIH 3T3 cells. The oncogenic nature of increased expression of vSecurin may result from chromosome gain or loss, produced by errors in chromatid separation.

The metaphase to anaphase transition is the final discrete event in duplication and separation of the genetic material of a cell. Its

timing is regulated by the activation of the anaphase-promoting complex (APC), which mediates selective proteolysis of various mitotic regulators (1-3). Experiments with *Xenopus* egg extracts indicated that a putative protein factor might exist whose degradation was required for the onset of sister-chromatid separation (4). Proteins with such an activity

Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA.

*To whom correspondence should be addressed. E-mail: marc@hms.harvard.edu

ARTÍCULO II

Natural Hotspots Induced by Transposable Element Insertion at the Breakpoints of a Drosophila Inversion

Mario Cáceres, Marta Puig y Alfredo Ruiz

2000

(Enviado)

Biological Sciences: Evolution

**Natural hotspots induced by transposable element insertion
at the breakpoints of a *Drosophila* inversion**

Mario Cáceres, Marta Puig, and Alfredo Ruiz

Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona,
08193 Bellaterra (Barcelona), Spain.

Corresponding author: Mario Cáceres
Departament de Genètica i de Microbiologia
Facultat de Ciències
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona), SPAIN
Phone: 34-93-581-2729
Fax: 34-93-581-2387
E-mail: marioes@cc.uab.es

Manuscript information: 21 text pages, 4 figures, 3 tables.

Abstract words: 202 words

Character count: 55,038 characters

Abbreviations footnote: TE, transposable element; ITR, inverted terminal repeat;
FB, *Foldback* element; Myr, million years.

ABSTRACT

Transposable elements (TEs) have been implicated in the origin of chromosomal changes, but their causal role in genome restructuring could be greater than previously thought. Here, we describe the structural and nucleotide variation at the breakpoints of a natural inversion generated by ectopic recombination between two oppositely oriented copies of a TE named *Galileo*. A detailed molecular analysis by Southern hybridization, PCR amplification, and DNA sequencing of 7.1 kb surrounding the inversion breakpoints in 39 *Drosophila buzzatii* lines revealed an unprecedented degree of structural variation, consisting of 22 insertions of 10 previously undescribed TEs, 12 deletions, and one inversion. All these changes are exclusively found in inverted chromosomes and seem to have occurred after the insertion of the two *Galileo* elements, within or close to them. Nucleotide variation at the breakpoint regions is six times lower in inverted chromosomes than in *standard* chromosomes, and suggests that most of the observed changes originated in less than 0.35 Myr. Thus, *Galileo* insertions appear to induce the generation of genetically unstable hotspots. The particular features of two new *Galileo* copies indicate that this TE belongs to the *Foldback* family. Together, these observations strengthen the importance of TEs, especially *Foldback* elements, as motors of evolutionary change.

INTRODUCTION

The generation of genetic variation is a prerequisite for evolution. Several types of DNA alterations are possible, from single nucleotide substitutions to chromosomal rearrangements. Transposable elements (TEs) are found in the genomes of all living organisms (1, 2) and are likely to play an evolutive role as a major source of new genetic variation (3-5). However, this view has been mainly based on laboratory observations, and the evolutionary significance of TE activity in natural populations awaits to be fully elucidated (6). TEs can exert a great variety of effects over host genomes. The most common is the production of deleterious mutations by their insertion within or near a gene's coding sequences, as happens in many *Drosophila* phenotypic mutants (7) and in human genetic diseases (8, 9). Nevertheless, some TE insertions induce favorable gene expression changes, which could have important evolutionary consequences (10). TEs have been also implicated in the shaping of genomes. Gross chromosomal rearrangements, such as deletions or inversions, appear by recombination between scattered TE copies (11, 12). Likewise, TEs comprise a major fraction of the genome in some plants and might account for most of genome size variation (13). Finally, TE insertions can have other more subtle effects. For example, TEs have been shown to increase the rate of recombination in the region where they insert (14-16).

One of the most outstanding examples of natural variation in chromosome structure is the extraordinarily rich inversion polymorphism of *Drosophila* species (17). A non-random distribution of inversion breakpoints, with some sites harboring many repeated breaks, has been found in most species (17, 18), suggesting the existence of mutational hotspots. Since TEs were implicated in the generation of *Drosophila* inversions in the laboratory (12), they could be responsible for the hotspots observed in nature. Recently, we sequenced the breakpoints of a highly successful *Drosophila* inversion, *D. buzzatii* 2j inversion, and demonstrated that it was caused by ectopic recombination between the homologous sequences of two oppositely oriented TEs (19). At both inversion breakpoints we found large

insertions which fulfilled all characteristic features of TEs and were considered copies of a new TE named *Galileo*. A preliminary study indicated that the insertions were absent from all 2 *standard* (*2st*) chromosomes and present in all chromosomes with the *2j* inversion, but showed that some structural variability among *2j* chromosomes exists (19). Besides, the insertion at the proximal breakpoint exhibited a very complex structure, with several internal repeats resembling additional TE insertions. To better characterize these TEs and to determine the level of structural and nucleotide variation at the breakpoints, we carried out a detailed molecular analysis of the *2j* breakpoint regions in nine lines with *2st* chromosomes and 30 lines with the *2j* inversion. A low level of nucleotide variation and a contrasting high degree of naturally occurring structural variation, caused mostly by insertions and deletions, have been found among *2j* chromosomes.

MATERIALS AND METHODS

***Drosophila* Stocks.** Thirty-nine lines of *D. buzzatii* and one of *D. martensis* were used. *D. buzzatii* lines (except *jq*⁷-3 and *jq*⁷-4) are isogenic for chromosome 2 and bear one of four different 2 chromosome arrangements: *2st*, *2j*, *2jz*³, and *2jq*⁷ (*2jz*³ and *2jq*⁷ derive from the *2j* arrangement and carry inversions *2z*³ and *2q*⁷, respectively). These lines were isolated from different natural populations covering the whole range of the species distribution. Geographic origins of *2st* lines are: *st*-1 and *st*-2, Carboneras (Spain); *st*-3, Vipos (Argentina); *st*-4, Guaritas (Brazil); *st*-5, Catamarca (Argentina); *st*-6, Salta (Argentina); *st*-7, Termas de Rio Hondo (Argentina); *st*-8, Ticucho (Argentina); and *st*-9, Trinkey (Australia). Geographic origin of *2j* lines is given in Table 1. The *D. martensis* line (Ma-4) is from Guaca (Venezuela).

Southern Hybridization, PCR Amplification, and Construction of Genomic Libraries. Southern hybridization was carried out by standard methods as previously described (20). Two probes were used for the analysis of the *2j*

breakpoint regions (Figure 1). The *AB* probe consists of a 1.7-kb Pst I fragment containing 1178 bp of *A* and 510 bp of *B* sequences. The *CD* probe consists of a 0.9-kb Dra I fragment containing 242 bp of *C* and 715 bp of *D* sequences. For the PCR amplification, different pairs of oligonucleotide primers covering the entire regions of study were designed (Figure 2). In order to specifically amplify the breakpoint insertions, primers that anneal to inserted repetitive sequences were always used in combination with primers located on the flanking non-repetitive DNA. Difficult templates, that were not amplified with the normal PCR conditions, were assayed with the GC-Rich PCR System (Roche), using 0.5-2 M GC-Rich resolution solution and an elongation temperature of 68°. Construction and screening of genomic libraries were as in Cáceres et al. (19).

DNA Sequencing and Sequence Analysis. DNA fragments coming from restriction enzyme digestion or PCR amplification were cloned into Bluescript II SK (Stratagene) or pGEM-T (Promega) vectors, respectively, and sequenced on an ALFexpress automated DNA sequencer (Amersham Pharmacia Biotech), using M13 universal and reverse primers. Nucleotide sequences were analyzed with the Wisconsin Package (Genetics Computer Group). Bestfit was used to align pairs of homologous sequences in different lines to detect inserted or deleted segments. Similarity searches through the GenBank/EMBL databases using Fasta, Blastx and Tblastx were carried out with the inserted sequences. For population genetic studies, we sequenced the same breakpoint regions as in Cáceres et al. (19) in six additional *2st* lines and seven additional *2j* lines. Both strands of PCR-generated templates were sequenced completely with the appropriate primers. Sequences were multiply aligned with Clustal W (21). Polymorphism and phylogenetic analysis were performed using the DnaSP program (22) and the PHYLIP software package (J. Felsenstein).

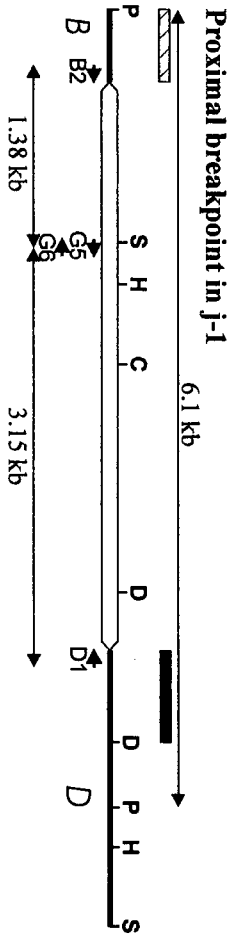
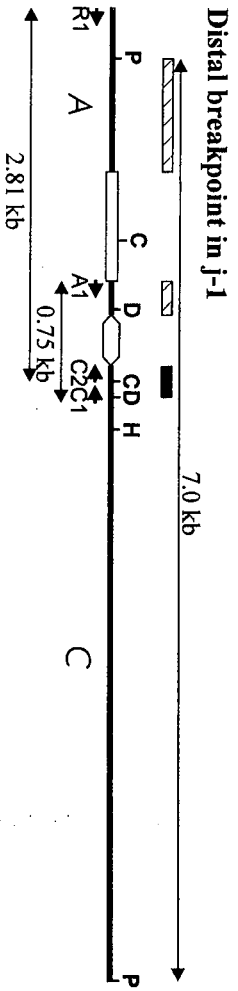
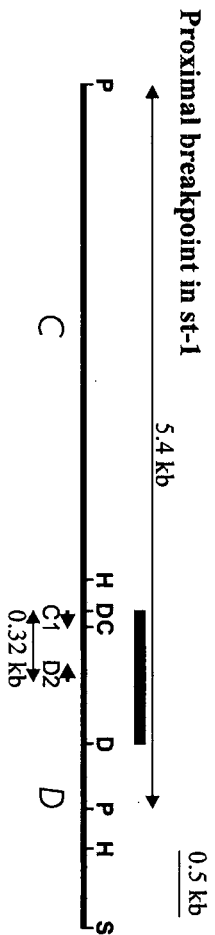
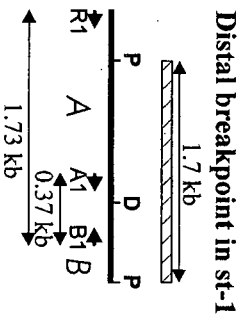


Figure 1 Physical map of the distal and proximal $2j$ breakpoint regions in the st-1 and j-1 lines. Thick lines represent the single copy *A*, *B*, *C*, and *D* sequences. TE insertions are represented as empty boxes. Hatched and black rectangles correspond, respectively, to the *AB* and *CD* probes used for the Southern hybridization analysis. Small arrows are primers used in the PCR amplification. Some of the restriction sites found in these regions are shown: C, Cla I; D, Dra I; H, Hind III; P, Pst I; S, Sal I.

RESULTS

Structural Variation at Inversion *2j* Breakpoints. Figure 1 shows the breakpoint regions in *st-1* and *j-1*, respectively, the *2st* and *2j* lines originally characterized (19). The breakage and inversion event occurred between *A* and *B* sequences and between *C* and *D* sequences. In inverted chromosomes large insertions are found between *A* and *C* (distal breakpoint) and *B* and *D* (proximal breakpoint). To examine the structural variation at those regions, different molecular techniques with increasing resolution power and accuracy were sequentially used: Southern blot hybridization, PCR amplification, restriction mapping of the PCR products, and DNA sequencing. No differences were found at the *AB* or *CD* regions among nine geographically diverse *2st* lines. In all cases Southern hybridization of Pst I digested DNA with *AB* and *CD* probes revealed the same bands of 1.7 kb and 5.4 kb, corresponding to the distal and proximal breakpoints regions, respectively (Figure 1). Similarly, PCR amplification of the 1.73-kb R1-B1 segment, 0.37-kb A1-B1 segment or 0.32-kb C1-D2 segment did not show any size variation among the *2st* lines either.

Clearly contrasting results were found in *2j* chromosomes. First, genomic DNA of 30 *2j* lines was digested with Pst I and hybridized with a *CD* probe. Two hybridization bands were observed in all *2j* lines and remarkable variation was detected among them: there were 11 bands of different sizes corresponding to the proximal breakpoint and five different bands for the distal breakpoint (Table 1). More detailed restriction maps of the breakpoint regions in those lines showing a Pst I hybridization pattern different from that of *j-1* (Figure 1) were obtained using additional restriction enzymes (Dra I, Hind III, Sal I, and Sal I-Cla I) and *AB* and *CD* probes. In the PCR analysis, a smaller region, containing the breakpoint insertions and the adjacent single-copy DNA, was studied. Primer pairs B2-G6 and G5-D1 (proximal breakpoint) and R1-C2 and A1-C1 (distal breakpoint) were used with genomic DNA of all *2j* lines (Figure 1). PCR products were compared by gel-electrophoresis and were digested with restriction enzymes to detect and map any

Table 1. Molecular analysis by Southern hybridization and PCR amplification of the 2*j* breakpoint regions of the 30 2*j* lines used in the study.

Name	Geographic origin	Hybridization bands (kb)		PCR products (kb)			
		Proximal	Distal	B2-G6	G5-D1	R1-C2	A1-C1
j-1	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-2	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-3	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-4	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-5	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-6	Carboneras (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-7	Caldetas (Spain)	6.1	7.0	1.38	1.92	2.81	0.75
j-8	San Luis (Argentina)	8.5	7.0	4.15	2.13	2.83	0.77
j-9	Quilmes (Argentina)	5.0	8.5	1.32	2.07	4.34	2.28
j-10	Palo Labrador (Argentina)	5.1	8.9	1.38	2.13	-	-
j-11	Los Negros (Bolivia)	8.8	7.0	1.32	2.07	2.83	0.77
j-12	Guaritas (Brazil)	8.8	7.0	1.32	2.07	2.83	0.77
j-13	Guaritas (Brazil)	8.8	7.0	1.32	2.07	2.81	0.75
j-14	Laboratory (Australia)	6.1	7.0	1.38	1.92	2.83	0.77
j-15	Catamarca (Argentina)	6.1	7.0	1.38	1.92	2.81	0.75
j-16	Salta (Argentina)	12.1	7.0	1.38	-	2.83	0.77
j-17	Tilcara (Argentina)	6.0	7.0	1.38	-	2.83	0.77
j-18	Termas Rio Hondo (Arg.)	5.0	7.0	1.32	2.07	2.83	0.77
j-19	Ticucho (Argentina)	10.3	8.9	1.32	2.25	-	-
j-20	Hemmant (Australia)	6.1	7.0	1.38	1.92	2.81	0.75

j-21	Hemmant (Australia)	6.1	7.0	1.38	1.92	2.81	0.75
j-22	Trinkey (Australia)	8.8	7.0	1.32	2.07	2.83	0.77
jz ³ -1	Carboneras (Spain)	9.9	7.0	1.32	3.11	2.83	0.77
jz ³ -2	Carboneras (Spain)	9.9	7.0	1.32	3.11	2.81	0.75
jz ³ -3	Kariouan (Tunisia)	9.9	7.0	1.32	3.11	2.83	0.77
jz ³ -4	Tilcara (Argentina)	8.3	9.2	1.34	-	-	-
jq ⁷ -1	Carboneras (Spain)	7.5	7.0	1.36	-	2.81	0.75
jq ⁷ -2	Mogan, Canary Is. (Spain)	7.5	11.0	1.36	-	3.62	1.56
jq ⁷ -3	Caldetas (Spain)	7.5	7.0	1.36	-	2.81	0.75
jq ⁷ -4	Otamendi (Argentina)	6.1	7.0	1.38	1.92	2.83	0.77

Hybridization bands are those obtained by Southern hybridization of Pst I digested genomic DNA of each line with the *CD* probe. Proximal and distal refer to the proximal and distal breakpoint, respectively. Proximal breakpoint bands indicated in boldface include a 3.8 kb extra segment due to a polymorphism in a Pst I site. Products of each PCR were digested with different restriction enzymes: BamH I-EcoR I (B2-G6), Dra I (G5-D1, R1-C2, and A1-C1).

variation existing among $2j$ lines (Table 1). The PCR results mostly agreed with the restriction maps from the Southern hybridizations. However, two problems arose in the PCR amplifications. First, Taq DNA polymerase sometimes jumped between distant parts of certain templates, excising the intervening segment. Sequencing of G5-D1 PCR products of $j-1$ and $j-19$ lines showed that a deletion occurred between short homologous sequences repeated in direct orientation, contained within long inverted repeats. This process resembles the mechanism of spontaneous deletion by slippage during DNA replication (23), that seems to be stimulated by the formation of stem-loop secondary structures (24). Second, no amplification occurred in some of the $2j$ lines, neither using specific combinations of primers nor with PCR conditions specially designed for difficult templates. For these segments (or those that underwent loss of fragments during PCR) we turned to traditional cloning. Thus, two lambda genomic libraries of the $j-19$ and jz^3-4 lines were constructed and both breakpoints in each line were isolated. As a final step, DNA sequencing was employed to obtain the nucleotide sequence of the parts where differences between $2j$ lines were found and to fully identify the changes that have occurred. Varying restriction fragments were cloned and completely sequenced (Figure 2). When two or more $2j$ lines did not show any variation in the restriction map of a particular region, only the DNA of one of them was sequenced as representative. All segments found to differ among $2j$ lines were characterized.

The Southern hybridization, PCR and DNA sequence data allowed to infer the structures present at the breakpoints of the 30 $2j$ lines studied. Ten structural variants were found in the proximal breakpoint and six in the distal breakpoint (Figure 2). Most of these different structures are related to each other by insertions or deletions of DNA segments. Five hypothetical variants (Hyp) have been postulated to illustrate the sequence of changes. We have considered that all insertions occurred independently, although a few insertions could have originated in a single event. In the proximal breakpoint, the simplest structure is that of Hyp-P1. It contains a *Galileo* insertion between *B* and *D* sequences with three other TEs inserted inside (Figure 2A). The TEs inside *Galileo* are flanked by direct repeats,

presumably generated by the duplication of the target site during the insertion event. The only exception is that of *BuT1*. In this case, the absence of the outermost nucleotide of the right inverted terminal repeat (ITR) suggests that a deletion after *BuT1* insertion removed the last base pair of the element, the right target site duplication, and part of the left *Galileo* ITR. From Hyp-P1, eight large insertions, eight deletions, and the inversion of an internal segment could generate the actual structural diversity (Figure 2A). In the distal breakpoint, the simplest structure is that of j-12, formed by a small *Galileo* insertion between *A* and *C* sequences and an *ISBu1* insertion in *A* (Figure 2B). From here, eight insertions and four deletions should have occurred to explain the other five structural variants observed (Figure 2B).

Table 2 summarizes the main features of the 22 insertions found at inversion *2j* breakpoints. The target site duplications flanking most insertions, the presence of multiple copies, and the variation found among lines, identify the inserted DNA sequences as TEs (2). In addition, sequence similarities suggest that they belong to ten different previously undescribed TEs. Besides the original *Galileo-1* and *Galileo-2* elements, that were implicated in the generation of inversion *2j* (19), two other *Galileo* elements have been found. The new *Galileo* copies are basically composed of very long ITRs and do not seem to code any protein involved in their transposition. There are also five insertions corresponding to two new TEs named *Kepler* and *Newton*, that are closely related to each other (average sequence identity 84%) and to *Galileo* (average sequence identity 73%). Neither of the three elements shows homology to any known sequence in the available databases, but they display significant structural similarities to *Foldback* (FB) elements (25-29). They share the terminal 40 bp of their ITRs (one single nucleotide difference) which are usually very long, duplicate 7 bp of the target site upon insertion, and have been included in the same family. Insertions i10 to i17 belong to four different TEs that can be ascribed to Class II (2, 3). We have designated them as *D. buzzatii* transposons or BuTs. Based on sequence similarities they have been included in the *hAT* superfamily (2): *BuT1* and *BuT2* show homology to the element *Gandalf* of *D.*

Figure 2 Structures found at the proximal (A) and distal (B) breakpoints of inversion $2j$ in the 30 $2j$ lines studied. All different structures are shown, except for that of j-16, that differs from jz^3-4 by the absence of d6 deletion. Thick lines represent the single-copy *A*, *B*, *C*, and *D* sequences. TEs are represented as colored boxes and sharp ends correspond to the ITRs. Insertions and deletions are delimited, respectively, by green and red lines and named with an *i* and a *d* followed by a number. Target site duplications flanking insertions are shown above them. Blue lines indicate the inversion of an internal segment. Arrows below the diagrams inform on the orientation of some homologous segments. Segments sequenced are enclosed within clear rectangles. Only the *D. buzzatii* lines representative of each structural variant are shown. Lines sharing the same structure in the proximal breakpoint are: jq^7-1 , jq^7-2 , and jq^7-3 ; j-1, j-2, j-3, j-4, j-5, j-6, j-7, j-14, j-15, j-20, j-21, and jq^7-4 ; j-9, j-11, j-12, j-13, j-18, and j-22 (deletion d2 was detected by j-12 sequencing and it is not known whether is present in other lines or not); and jz^3-1 , jz^3-2 , and jz^3-3 . Lines sharing the same structure in the distal breakpoint are: j-1, j-2, j-3, j-4, j-5, j-6, j-7, j-13, j-15, j-20, j-21, jz^3-2 , jq^7-1 , and jq^7-3 ; j-8, j-11, j-12, j-14, j-16, j-17, j-18, j-22, jz^3-1 , jz^3-3 , and jq^7-4 ; and j-10 and j-19. Hyp are hypothetical structures not found in our sample of $2j$ lines. Small black arrows are PCR primers used in the study.

koepferae (30), whereas *BuT3* and *BuT4* are related to the element *Hopper* of *Bactrocera dorsalis* (31). Finally, five insertions could not be classified into any of the previously known types of TEs. *BuT5* ends in 3-bp ITRs, generates 9-bp duplications during insertion, shows a moderately repetitive pattern by in situ hybridization (J. M. Ranz, personal communication), and has been tentatively considered a Class II TE. The other four insertions belong to a new class of highly repetitive mobile elements, that do not possess ITRs and duplicate 2 bp upon insertion. They have been called *ISBu* elements because of their structural and sequence similarity to the IS elements of the *D. obscura* group (32).

We have also detected 12 deletions bigger than 10 bp at the *2j* breakpoints (Figure 2). The length of the observed deletions is: d1, 93 bp; d2, 24 bp; d3, 238 bp; d4, 32 bp; d5, 179 bp; d6, 41 bp; d7, 536 bp; d8, 20 bp; d9, 17 bp; d10, 248 bp; d11, 649 bp; and d12, 1023 bp (the lengths of d7 and d11 are minimum estimates, since the real size of the deleted fragments is not known). Four of these deletions could originate by the well established mechanism of slipped-strand mispairing (23): d2, d3, and d6 took place between 3-4 bp repeated sequences, and eliminated one of them and the intervening DNA; d8 removed one copy of a 20 bp sequence duplicated in tandem. Finally, we have found a change of orientation of a 55 bp *Galileo-1* internal fragment in some of the lines (Figure 2A). This suggests that an inversion has occurred inside the proximal breakpoint insertion, probably by recombination between the oppositely oriented ITRs of *Kepler-1* and *Kepler-2* in Hyp-P2.

Nucleotide Variation at Inversion *2j* Breakpoints. We have sequenced 596 bp corresponding to the *A*, *B*, *C*, and *D* single-copy DNA in the nine *2st* lines, 12 of the *2j* lines comprising the diversity of structural types found, and *D. martensis*; and 839 bp of the distal breakpoint insertion and the ends of the proximal breakpoint insertion in the *2j* lines only. Figure 3 shows the 89 polymorphic sites found. Nucleotide diversity, π (33), was calculated ignoring sites with alignment gaps or missing data only in pairwise comparisons (Table 3).

Table 2. TE insertions at the breakpoint regions of inversion 2j of *D. buzzatii*.

Insertion	TE	Size (bp)	ITRs (bp)	Target site (bp)	BP
<i>Foldback-like elements</i>					
i1	<i>Galileo-1</i>	1589	228/443	7	P
i2	<i>Galileo-2</i>	392	106	7	D
i3	<i>Galileo-3</i>	2204	683/684	7	P
i4	<i>Galileo-4</i>	1948	782/917	ND	D
i5	<i>Kepler-1</i>	722	150	5	P
i6	<i>Kepler-2</i>	735	ND	7	P
i7	<i>Kepler-3</i>	692	20	ND	D
i8	<i>Newton-1</i>	1510	572/575	7	P
i9	<i>Newton-2</i>	1512	575/574	7	D
<i>hobo, Activator, Tam3 (hAT) elements</i>					
i10	<i>BuT1</i>	801	15/14	ND	P
i11	<i>BuT2</i>	2775	12	8	P
i12	<i>BuT3-1</i>	413	23	8	P
i13	<i>BuT3-2</i>	844	23	8	P
i14	<i>BuT3-3</i>	803	23	8	P
i15	<i>BuT3-4</i>	795	23	8	D
i16	<i>BuT3-5</i>	147	ND	ND	D
i17	<i>BuT4</i>	719	24/23	8	D
Unclassified elements					
i18	<i>BuT5</i>	1039	3	9	P
i19	<i>ISBu1-1</i>	841	-	2	D
i20	<i>ISBu1-2</i>	1467	-	2	P
i21	<i>ISBu1-3</i>	852	-	2	D
i22	<i>ISBu2</i>	725	-	2	D

Elements have been classified by structural and sequence similarities with described TEs according to Copy et al. (2). When different, the size of the left and right inverted terminal repeats (ITRs) are indicated. BP refers to the location of the element in the proximal (P) or distal (D) breakpoint. ND, data that could not be determined due to deletions.

In the single-copy DNA, the Kreitman and Hudson's homogeneity test (34) did not detect differences in polymorphism level across *A*, *B*, *C* and *D* ($X^2_L = 2.86$, $df = 3$, $P = 0.41$) and the data of the four regions have been pooled. Overall, nucleotide variation is six times higher in *2st* chromosomes than in *2j* chromosomes (Table 3), and *2st* and *2j* chromosomes exhibit a great number of fixed differences, including 17 nucleotide substitutions and six indels of 1 to 4 bp (TE insertions and target site duplications excluded). A neighbor-joining tree (35) of the single-copy sequences is shown in Figure 4. All *2j* sequences form a monophyletic cluster of high bootstrap value, clearly separated from that of *2st* sequences, confirming the proposed unique origin of the inversion (19). No significant departures from the neutral model were found with the Tajima and Fu and Li tests (36, 37), and nucleotide variation was used to date the origin of the inversion and of the sampled *2st* and *2j* alleles. The age of the inversion was estimated from the fixed differences between *2st* and *2j* chromosomes. The average number of nucleotide differences, d_{xy} (33), between *2st* and *2j* chromosomes is 0.0353, and between *D. buzzatii* and *D. martensis* is 0.1094. Subtracting from both figures the intraspecific polymorphism (0.0197), the net average number of nucleotide substitutions is obtained (33). Combining available information (38, 39), we have estimated the divergence time between *D. buzzatii* and *D. martensis* as 5.8 million years (Myr) and a rate of 7.7×10^{-9} nucleotide substitutions/site/year for this region. The *2j* inversion must be then ~1 Myr old, which is consistent with its widespread distribution through all *D. buzzatii* populations (19). The coalescence time of *2st* and *2j* alleles was estimated from the number of nucleotide substitutions per site between the two most divergent alleles within each chromosomal arrangement (40). Therefore, the sampled *2st* alleles should be 0.92 Myr old and the sampled *2j* alleles 0.35 Myr old.

The TE sequences inserted at the proximal breakpoint in *2j* chromosomes accumulate strikingly higher nucleotide variation than the single-copy regions and the distal breakpoint insertion ($X^2_L = 8.61$, $df = 2$, $P = 0.01$). The difference between the two insertions ($X^2_L = 4.00$, $df = 1$, $P = 0.04$), which are expected to be equally

Table 3. Nucleotide variation in the breakpoint regions of inversion $2j$ of *D. buzzatii*.

Region	Total (N = 21)			$2st$ (N = 9)		$2j$ (N = 12)	
	m	S	π	S	π	S	π
<i>ABCD</i>	596	35	0.0197	15	0.0075	3	0.0013
<i>A</i>	179	13	0.0251	5	0.0063	2	0.0189
<i>B</i>	143	9	0.0320	2	0.0076	0	0
<i>C</i>	155	9	0.0167	6	0.0104	0	0
<i>D</i>	119	4	0.0045	2	0.0055	1	0.0015
Insertions	839	-	-	-	-	13	0.0066
proximal	447	-	-	-	-	11	0.0096
distal	392	-	-	-	-	2	0.0007

Positions A65 to A101 of st-3 and st-8 lines, probably originated by some sort of genetic exchange, have been excluded from the estimation of the nucleotide diversity. N, number of sequences considered; m, maximum number of nucleotides sequenced in each region; S, number of segregating sites; π , average number of pairwise differences between sequences per nucleotide.

selectively constrained, suggests an intrinsic increased rate of nucleotide change at the proximal breakpoint insertion.

DISCUSSION

Our detailed study of the inversion *2j* breakpoints has revealed a great amount of naturally occurring structural variation, consisting of 22 TE insertions, 12 deletions, and an internal inversion. Together with the inversion *2j* itself, it represents an unprecedented level of chromosomal change and defines these regions as genetically unstable hotspots. The observed frequency of insertions in *Drosophila* varies from 0 to 0.016 insertions/kb/chromosome (41). Considering the single-copy region analyzed by Southern hybridization of Pst I digested DNA (7.1 kb), a frequency of 0 and 0.601 insertions/kb/chromosome is obtained for *2st* and *2j* chromosomes, respectively (only those insertions known to have occurred independently were computed). That of *2j* chromosomes is 40 times bigger than the highest frequency of insertions ever found in *Drosophila* and 100 times greater than the 0.006 TEs/kb density estimated from the 2.9-Mb sequence of the *D. melanogaster Adh* region (42). Also, insertions and deletions are not distributed uniformly within *2j* chromosomes (Figure 2). Most TEs are nested inside preexisting elements and seem to accumulate after the initial *Galileo-1* and *Galileo-2* insertions, within or very close to them; and the deletions occurred inside or at the ends of TEs. Furthermore, the degree of structural variation in *2j* chromosomes contrasts with the relatively low level of nucleotide variation. The coalescence time of the sampled *2j* alleles sets an upper limit for the origin of the changes that differentiate them. Thus, 16 TE insertions, the 12 deletions, and the internal inversion should have occurred in less than 0.35 Myr. In the case of TEs, this results in an insertion rate even faster than the explosive accumulation of 23 retrotransposons in the originally 80-kb long *adh-1* region of maize over the last 6 Myr (43).

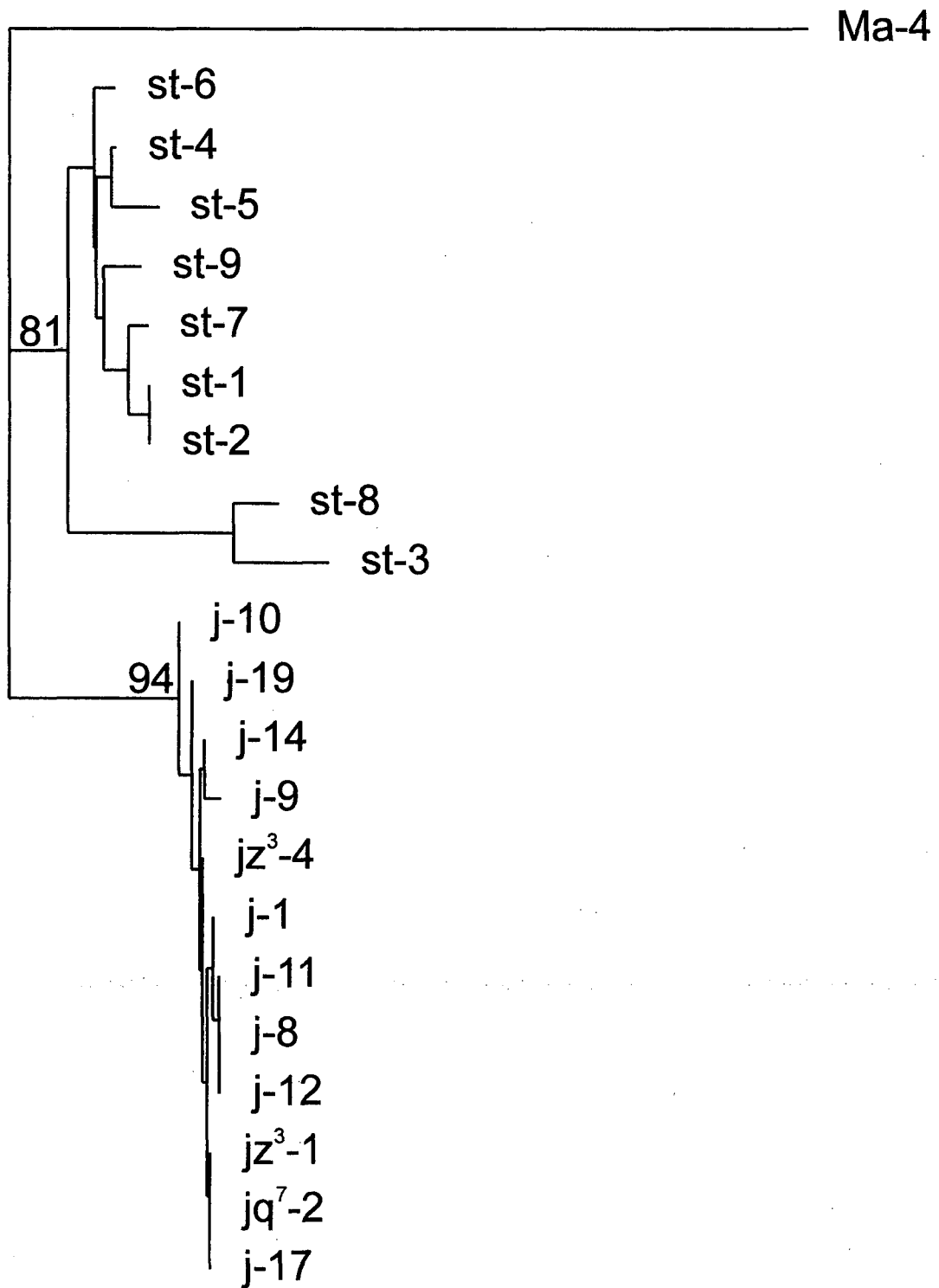


Figure 4 Neighbor-joining phylogenetic tree of the breakpoint sequences of inversion $2j$, based on the *A*, *B*, *C*, and *D* sequence data for the nine $2st$ and 12 $2j$ *D. buzzatii* lines. The Ma-4 *D. martensis* line was used as outgroup. Bootstrap values in percentage out of 500 replicates are indicated for the main nodes.

What is the cause of the hotspot? The clear differences between the homologous regions in *2st* and *2j* chromosomes point to an effect of the inversion or of the initial *Galileo* insertions as most likely explanations. It has been argued that TEs could accumulate around inversion breakpoints because the reduction of recombination protects TEs from being eliminated by deleterious ectopic exchanges (6). This, however, does not fit with our observations. First, given the intermediate frequency of inversion *2j*, the reduction in recombination would affect in a similar way *2st* and *2j* chromosomes. Second, it is noteworthy the complete absence of retrotransposons, which constitute the majority of TEs in *Drosophila* (7). And third, our insertions lie in very small regions around the inversion breakpoints. In previous studies, no TE insertions were found at a 2.5-kb region flanking *In(3L)P* breakpoints (44) and only two were found in a 5-kb region surrounding the *In(2L)t* proximal breakpoint (45).

All *2j* chromosomes share three TE insertions and one hypothetical deletion inside the *Galileo-1* element, and *ISBu1-1* insertion at the distal breakpoint. It could seem then that the hotspot predates the *2j* inversion and stems from the *Galileo-1* and *Galileo-2* insertions. There are several examples of preferential insertion of retrotransposons inside preexisting elements. This is thought to be a strategy to reduce the damage caused to the host by TE mobilization (5). Ty1, Ty2, Ty3 and Ty4 elements of *Saccharomyces cerevisiae* show a strong preference for regions upstream of genes transcribed by RNA polymerase III, and clusters of elements inserted within elements are formed in these regions (46, 47). Blocks of nested retrotransposons are also found in intergenic regions of the maize genome: 14 of the 23 retrotransposons at *adh1-F* region were inserted within other retrotransposons (43). Retrotransposons aside, just *P* elements show some tendency to insert into each other (48), and only *Drosophila* heterochromatin might harbor similar reiterative insertions of transposons as in the *2j* breakpoints (49, 50). On the other hand, transposons have been found to stimulate the production of deletions in laboratory studies. *P* elements tend to produce internal deletions (51), whereas deletions recovered from *mariner* elements usually affect the ITR of the element

and the flanking DNA (16). In both cases, sometimes extra DNA appears between the deletion endpoints, as happens in our d4 and d5 deletions which were accompanied by the introduction of a new nucleotide. Finally, *Galileo* insertions mediated the generation of inversion $2j$ (19), and TEs are probably implicated in the origin of other naturally occurring inversions in Diptera (45, 52).

Genetic instability might result from the mechanism of transposition of the TEs inserted at the $2j$ breakpoints. Excluding *ISBu1* and *ISBu2*, all the other elements are thought to transpose by a cut-and-paste mechanism (2, 3). Transposase-induced DNA breaks at the transposon ends serve as substrate for host repair functions, and many different types of alterations can be produced by aberrant repair of those lesions (16). Breaks stimulate branch-migration and recombination processes (14-16), that can lead to the generation of inversions. Either an increased mutation rate due to repeated repair events or an increased frequency of genetic exchange could account for the high nucleotide variation at the proximal breakpoint insertion. It is particularly remarkable the presence of FB elements as most-likely inducers of the hotspots. FB elements are characterized by the production of extremely unstable mutations and chromosomal rearrangements at unusually high frequencies in *D. melanogaster* laboratory populations (25). Thus, despite being one of the least noted groups of TEs, FB elements could have played a major role in chromosomal evolution. The possibility that *Galileo* elements could be involved in similar hotspots at other locations of the *D. buzzatii* genome is a very interesting question for further investigation.

ACKNOWLEDGEMENTS

We are deeply indebted to J. M. Ranz for the data on the repetitive nature of *BuT5* and general advice at all stages of this work. J. S. F Barker kindly provided us with 15 of the *D. buzzatii* stocks used. We also thank A. Barbadilla for helpful discussion of results, and M. Ashburner, A. Berry, P. Capy, F. Casares, and D.

Petrov for valuable comments and suggestions. Work was supported by grant PB98-0900-C02-01 from the Dirección General de Investigación Científica y Técnica (Ministerio de Educación y Cultura, Spain) awarded to A. R. and a doctoral FI fellowship from the Comissionat per a Universitats i Recerca (Generalitat de Catalunya, Spain) awarded to M. C.

REFERENCES

1. Berg, D. E. & Howe, M. M. (1989) *Mobile DNA* (American Society for Microbiology, Washington, D.C.).
2. Capy, P., Bazin, C., Higuete, D. & Langin, T. (1998) *Dynamics and Evolution of Transposable Elements* (Springer-Verlag, Heidelberg).
3. Finnegan, D. J. (1989) *Trends Genet.* **5**, 103-107.
4. McDonald, J. F. (1995) *Trends Ecol. Evol.* **10**, 123-126.
5. Kidwell, M. G. & Lisch, D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7704-7711.
6. Charlesworth, B., Sniegowski, P. & Stephan, W. (1994) *Nature* **371**, 215-220.
7. Lindsley, D. L. & Zimm, G. G. (1992) *The Genome of Drosophila melanogaster* (Academic Press, San Diego).
8. Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W. & Collins, F. S. (1991) *Nature* **353**, 864-866.
9. Holmes, S. E., Dombrowski, B. A., Krebs, C. M., Boehm, C. D. & Kazazian, H. G. (1994) *Nature Genet.* **7**, 143-148.
10. Britten, R. J. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9374-9377.
11. Roeder, G. S. & Fink, G. R. (1983) in *Mobile Genetic Elements*, ed. Shapiro, J. (Academic Press, Orlando, Fl.), pp. 299-328.
12. Lim, J. K. & Simmons, M. J. (1994) *BioEssays* **16**, 269-275.
13. SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z. & Bennetzen, J. L. (1996) *Science* **274**, 765-768.
14. Sved, J. A., Eggleston, W. B. & Engels, W. R. (1990) *Genetics* **124**, 331-337.
15. McCarron, M., Duttaroy, A., Doughty, G. & Chovnick, A. (1994) *Genetics* **136**, 1013-1023.
16. Lohe, A. R., Timmons, C., Beerman, I., Lozovskaya, E. R. & Hartl, D. L. (2000) *Genetics* **154**, 647-656.
17. Krimbas, C. B. & Powell, J. R. (1992) in *Drosophila Inversion Polymorphism*, eds. Krimbas, C. B. & Powell, J. R. (CRC Press, Boca Raton, FL), pp. 1-52.

18. Cáceres, M., Barbadilla, A. & Ruiz, A. (1997) *Evolution* **51**, 1149-1155.
19. Cáceres, M., Ranz, J. M., Barbadilla, A., Long, M. & Ruiz, A. (1999) *Science* **285**, 415-418.
20. Ranz, J. M., Cáceres, M. & Ruiz, A. (1999) *Chromosoma* **108**, 32-43.
21. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673-4680.
22. Rozas, J. & Rozas, R. (1999) *Bioinformatics* **15**, 174-175.
23. Farabaugh, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. (1978) *J. Mol. Biol.* **126**, 847-863.
24. Egner, C. & Berg, D. E. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 459-463.
25. Bingham, P. M. & Zachar, Z. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (American Society for Microbiology, Washington, D.C.), pp. 485-502.
26. Hoffman-Liebermann, B., Liebermann, D. & Cohen, S. N. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (American Society for Microbiology, Washington, D.C.), pp. 575-592.
27. Hankeln, Th. & Schmidt, E. R. (1990) *J. Mol. Biol.* **215**, 477-482.
28. Yuan, J., Finney, M., Tsung, N. & Horvitz, H. R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3334-3338.
29. Rebatchouk, D. & Narita, J. O. (1997) *Plant. Mol. Biol.* **34**, 831-835.
30. Marín, I. & Fontdevila, A. (1995) *Mol. Gen. Genet.* **248**, 423-433.
31. Handler, A. M. & Gomez, S. P. (1997) *Gene* **185**, 133-135.
32. Hagemann, S., Miller, W. J., Haring, E. & Pinsker, W. (1998) *Chromosoma* **107**, 6-16
33. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia University Press, New York).
34. Kreitman, M. & Hudson, R. R. (1991) *Genetics* **127**, 565-582.
35. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406-425.
36. Tajima, F. (1989) *Genetics* **123**, 585-595.
37. Fu, Y.-X. & Li, W.-H. (1993) *Genetics* **133**, 693-709.

38. Russo, C. A. M., Takezaki, N. & Nei, M. (1995) *Mol. Biol. Evol.* **12**, 391-404.
39. Rodríguez-Trelles, F., Alarcón, L. & Fontdevila, A. (2000) *Mol. Biol. Evol.* **17**, 1112-1122.
40. Rozas, J. & Aguadé, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11517-11521.
41. Aquadro, C. F. (1993) in *Molecular Approaches to Fundamental and Applied Entomology*, eds. Oakeshott, J. & Whitten, M. J. (Springer-Verlag, New York), pp. 222-266.
42. Ashburner, M., Misra, S., Roote, J., Lewis, S., Blazej, R., Davis, T., Doyle, C., Galle, R., George, R., Harris, N., et al. (1999) *Genetics* **153**, 179-219.
43. SanMiguel, P., Gaut, B. S., Tikhonov, A., Nakajima, Y. & Bennetzen, J. L. (1998) *Nature Genet.* **20**, 43-45.
44. Hasson, E. & Eanes, W. F. (1996) *Genetics* **144**, 1565-1575.
45. Andolfatto, P., Wall, J. D. & Kreitman, M. (1999) *Genetics* **153**, 1297-1311.
46. Voytas, D. F. (1996) *Science* **274**, 737-738.
47. Boeke, J. D. & Devine, S. E. (1998) *Cell* **93**, 1087-1089.
48. Engels, W. R. (1996) in *Transposable Elements*, eds. Saedler, H. & Gierl, A. (Springer-Verlag, Berlin), pp. 103-123.
49. Pimpinelli, S., Berloco, M., Fanti, L., Dimitri, P., Bonaccorsi, S., Marchetti, E., Caizzi, R., Caggese, C. & Gatti, M. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3804-3808.
50. Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., et al. (2000) *Science* **287**, 2185-2195.
51. Staveley, B. E., Heslip, T. R., Hodgetts, R. B. & Bell, J. B. (1995) *Genetics* **139**, 1321-1329.
52. Mathiopoulos, K. D., della Torre, A., Predazzi, V., Petrarca, V. & Coluzzi, M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 12444-12449.

ARTÍCULO III

*Inversion Length and Breakpoint Distribution in the *Drosophila buzzatii* Species
Complex: Is Inversion Length a Selected trait?*

Mario Cáceres, Antonio Barbadilla y Alfredo Ruiz

1997

Evolution **51**: 1149-1155

INVERSION LENGTH AND BREAKPOINT DISTRIBUTION IN THE *DROSOPHILA BUZZATII* SPECIES COMPLEX: IS INVERSION LENGTH A SELECTED TRAIT?

MARIO CÁCERES,^{1,2} ANTONIO BARBADILLA, AND ALFREDO RUIZ

Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

¹E-mail: IBGB4@CC.UAB.ES

Abstract.—Length and position of breakpoints are characteristics of inversions that can be precisely determined on the polytene chromosomes of *Drosophila* species, and they provide crucial information about the processes that govern the origin and evolution of inversions. Eighty-six paracentric inversions described in the *Drosophila buzzatii* species complex and 18 inversions induced by introgressive hybridization in *D. buzzatii* were analyzed. In contrast to previous studies, inversion length and breakpoint distribution have been considered simultaneously. We conclude that: (1) inversion length is a selected trait; rare inversions are predominantly small while evolutionarily successful inversions, polymorphic and fixed, are predominantly intermediate in length; a nearly continuous variation in length, from small to medium sized, is found between less and more successful inversions; (2) there exists a significant negative correlation between length and number of polymorphic inversions per species which explains 39% of the inversion length variance; (3) natural selection on inversion length seems the main factor determining the relative position of breakpoints along the chromosomes; (4) the distribution of breakpoints according to their band location is non-random, with chromosomal segments that accumulate up to eight breakpoints.

Key words.—*Drosophila buzzatii* complex, inversion breakpoints, inversion length, inversion polymorphism, polymorphic inversion richness.

Received September 23, 1996. Accepted April 14, 1997.

There is a long history of theoretical models and empirical tests dealing with the relationship between inversion length and the processes generating and maintaining paracentric inversions in the *Drosophila* genus. Wallace (1954) and Crumacker and Kastiris (1967) first proposed a relationship between length and success of inversions and postulated that rare inversions were generally short while well-established inversions were at least moderately long. Olvera et al. (1979) tried to confirm this relationship in *D. pseudoobscura*, and showed that rare inversions were more often short than polymorphic and widespread ones (see also Ruiz et al. 1984 for similar results in *D. buzzatii*). Moreover, Van Valen and Levins (1968) fit length data from a selected subset of inversions of the *Drosophilidae* family to different theoretical models. They found a deficiency of small inversions with respect to a uniform breakage probability model. Krimbas and Loukas (1980) in *D. subobscura* and Brehm and Krimbas (1991) in *D. obscura* found a deficiency of small and large inversions when compared with the expectations of the same model. All these results seem to indicate that moderately sized inversions are favored by natural selection.

Distribution of inversion breakpoints has also been subjected to extensive study. Position of breakpoints can be considered in two different ways: (1) as the chromosomal band in which the break occurs (band location); and (2) as the relative position along the chromosome. According to band location, non-random distribution of breakpoints and hot spots with repeated breaks have been observed in several *Drosophila* species (Krimbas and Loukas 1980; Tonzetich et al. 1988; Lemeunier and Aulard 1992), but not in *D. pseudoobscura* (Olvera et al. 1979). The distribution of breakpoints with regard to relative position seems also to be non-random in *Drosophila*, with regions near the centromere

showing more breaks (Tonzetich et al. 1988; Krimbas and Powell 1992).

Obviously, length and position of breakpoints are related and must be studied simultaneously. Information on both is essential to understand the processes underlying the origin and success of chromosomal inversions. However, none of the previous studies has considered these two variables at the same time. In addition, all studies have been carried out either using a single species (*D. pseudoobscura*, Olvera et al. 1979; *D. subobscura*, Krimbas and Loukas 1980; *D. obscura*, Brehm and Krimbas 1991; *D. melanogaster*, Lemeunier and Aulard 1992), or lumping together heterogeneous data from many species (Van Valen and Levins 1968; Tonzetich et al. 1988). In the former case, fixed inversions cannot be analyzed and compared. In the latter, the observed patterns may be difficult to interpret. Furthermore, none of the previous studies has tested for the effect of the species (or its phylogenetic affinities) on the length of inversions.

The *Drosophila buzzatii* species complex is a closely related set of 12 species belonging to the *mulleri* subgroup of the *Drosophila repleta* species group. Eleven formally described species have been ascribed to the *buzzatii* complex (Wasserman 1992; Ruiz and Wasserman 1993; Tidon-Sklorz and Sene 1995). In addition, the populations of *D. serido* of central-western Brazil (*D. serido* IV) very likely represent a separate, and as yet undescribed, species (Tosi and Sene 1989; Silva and Sene 1991). All *buzzatii* complex species have similar basic karyotypes consisting of four pairs of almost equal-length acrocentric autosomes, one pair of dot autosomes, a long acrocentric X, and a Y chromosome. Its chromosomal phylogeny has been established by comparing the banding patterns of the various species and also by the observation of the polytene chromosomes in many interspecific hybrids (Ruiz and Wasserman 1993). Results fully congruent with the proposed chromosomal phylogeny have been obtained by sequencing the mitochondrial cytochrome oxidase genes of

² Corresponding author.

eight of the 12 species (Spicer 1995) and by comparative gene mapping between *D. buzzatii* (one of the species in the complex) and *D. repleta* (the reference species for all cytological studies in the group) (Ranz et al. 1997). The complete and reliable record of its extant inversion polymorphism and chromosomal evolution, and the availability of a sample of inversions induced in *D. buzzatii* by introgressive hybridization (Naveira and Fontdevila 1985), makes this complex a suitable species set for the study of the evolutionary sequence from the origin to the fixation of inversions.

MATERIALS AND METHODS

This study is based on the 86 inversions described so far in the *D. buzzatii* species complex (Ruiz and Wasserman 1993; Kuhn et al. 1996) and 18 inversions induced in a laboratory strain of *D. buzzatii* by introgressive hybridization with *D. koepferae* (Naveira and Fontdevila 1985). According to their evolutionary success, naturally occurring inversions were grouped as fixed, polymorphic, or rare. Only those inversions found in all individuals of one or more species were considered fixed. Polymorphic inversions are those segregating within a species in at least two different localities. Inversions fixed in one species but segregating in a different one ($2f^2$, $2e^7$, and $2r^6$) were also considered polymorphic. Finally, inversions found in only one locality, and usually in very low frequency, were considered rare (Ruiz and Wasserman 1993). Physical length and position of breakpoints were determined using maps of the polytene chromosomes of the *buzzatii* complex species (Ruiz and Wasserman 1993) and the original descriptions of the inversions. These maps are a cut-and-paste reconstruction of the *D. repleta* polytene chromosomes (Wharton 1942) from which these species differ by a small number of inversions (Ruiz and Wasserman 1993). The full data set is available from the authors upon request.

Relative length of each inversion (l) was measured as the distance between the two breakpoints divided by the total length of the chromosome. Two different kinds of analysis were carried out. First, different classes of inversions were compared using analysis of variance and non-parametric tests. Second, observed length distributions were compared to those predicted by four theoretical models that assume different breakpoint distributions. The four models are: (1) Uniform Breakage Probability Model: assumes that distribution of breakpoints of inversions is uniform along the chromosome (Federer et al. 1967; Van Valen and Levins 1968); (2) Van Valen and Levins' Model: a truncated version of the previous model, with very short inversions ($l < 0.1$) being excluded (Federer et al. 1967; Van Valen and Levins 1968); (3) Non-uniform Breakage Probability Model: derived from the first model by relaxing the assumption that the probability of a break is uniform along the chromosome; it allows for unequal break probabilities in different regions of the chromosome, although it assumes that breaks distribute uniformly inside each region (see Appendix for its derivation); (4) Truncated Non-uniform Breakage Probability Model: derived from the third model, it excludes very short inversions ($l < 0.1$).

The positions of inversion breakpoints were analyzed as

the relative distance from each break to the centromere (relative position) and as the chromosomal band where the break is located (band location). For the study of the relative position of breakpoints, chromosomes were divided into three regions (proximal, central, and distal), each one-third the total length; and the observed number of breakpoints in each region was compared to a uniform distribution. The study of the band location of breakpoints was limited to chromosome 2, which harbors 61 out of the 86 inversions described in the *buzzatii* complex. Using the usual procedure (Tonzetich et al. 1988; Lemeunier and Aulard 1992), chromosome 2ab, the putative ancestral sequence of the *buzzatii* complex (Ruiz and Wasserman 1993), was divided into 74 equal-length segments, and the number of breakpoints in each segment was recorded. The number of segments is arbitrary, but large enough for the expected number of breakpoints per segment to approximate a Poisson distribution.

RESULTS

Length of Inversions

To evaluate the factors associated with inversion length we carried out a log linear analysis (Sokal and Rohlf 1995) with length, chromosome, and evolutionary success as variables. Only the interaction between length and success was significant ($G = 29.10$; $df = 5$; $P < 0.001$). The chromosome does not have an effect on inversion length, i.e. inversions on different chromosomes have similar lengths. Interaction between chromosome and success was not significant either: fixed, polymorphic, and rare inversions have a similar distribution in the different chromosomes, with a marked accumulation of inversions in chromosome 2 in all three cases.

Figure 1 shows the length distribution of fixed, polymorphic, and rare inversions. The differences between them are obvious. Analysis of variance showed significant differences in length between the three types of inversions ($F = 5.74$; $df = 2, 83$; $P = 0.005$). Post hoc comparisons using Newman-Keuls test showed that mean length of rare inversions is significantly lower than that of polymorphic ($P = 0.004$) and fixed inversions ($P = 0.025$). Polymorphic and fixed inversions do not have significantly different means ($P = 0.59$), but very different variances (Levene's test: $F = 9.12$; $df = 1, 60$; $P = 0.004$).

Polymorphic and rare inversions can be assigned to each species and their length can be compared among them. Analysis of variance showed significant differences in length between species for polymorphic inversions ($F = 5.29$; $df = 9, 33$; $P < 0.001$) but not for rare inversions ($F = 0.66$; $df = 7, 16$; $P = 0.70$). To ascertain how the total variation among species is distributed among and within the three species clusters (*stalkeri*, *martensis*, and *buzzatii*) of the *buzzatii* complex (Ruiz and Wasserman 1993), we carried out a nested analysis of variance (Harvey and Pagel 1991). Results were unequivocal: the between-clusters level accounts for no variance, and all variation in inversion length is explained by the between-species (within clusters) level (Table 1). When the number of polymorphic inversions per species, rather than species, was used as the independent variable, a significant negative correlation between length and number of polymorphic inversions was found (Fig. 2). This correlation ex-

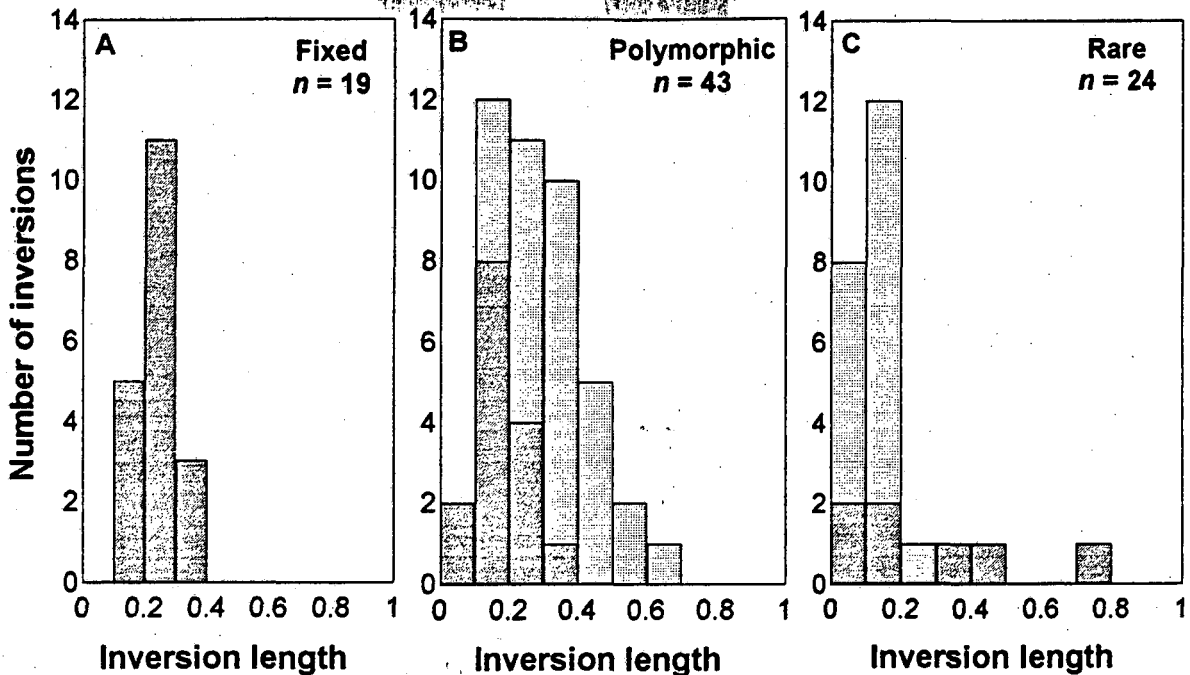


FIG. 1. Length distribution of the three evolutionary success classes of inversions. Mean length (\pm SD) for fixed (A), polymorphic (B), and rare inversions (C) is 0.259 (\pm 0.067), 0.279 (\pm 0.135), and 0.169 (\pm 0.151). (B) Polymorphic inversions are subdivided in polymorphic endemic (hatched bars) and polymorphic widespread (stippled bars). Mean length (\pm SD) for polymorphic endemic ($n = 15$) and widespread inversions ($n = 28$) is 0.177 (\pm 0.074) and 0.333 (\pm 0.129). (C) Rare inversions are subdivided in unique inversions (hatched bars) and nonunique rare inversions (stippled bars). Mean length (\pm SD) for unique ($n = 7$) and nonunique inversions ($n = 17$) is 0.280 (\pm 0.245) and 0.124 (\pm 0.054).

plains 39% of the variance in inversion length and persists when considering only inversions of chromosome 2 ($r = -0.59$; $P < 0.001$). Species with many polymorphic inversions tend to have smaller inversions than those with fewer inversions.

Length distributions of the three types of inversions were compared to expected length distributions predicted by the four theoretical models (Table 2). Naturally occurring inversions (fixed, polymorphic, and rare) do not fit the uniform breakage probability model, not even after truncation of very short inversions (Van Valen and Levins' model). Except for rare inversions, the elimination of inversions of length less than 0.1 significantly improved fit of observed distributions to expected ones (Table 2). Fixed and polymorphic inversions, besides the deficiency of very short inversions, show a paucity of long and an excess of medium-sized inversions. Rare inversions present an excess of short inversions. When the non-uniform breakage probability model is considered, length distribution of polymorphic inversions shows an ex-

cellent fit to the expected truncated distribution (Table 2). Nevertheless, although different probabilities of breaks are considered, fixed inversions do not fit the non-uniform breakage probability model.

Distribution of Inversion Breakpoints

The distribution of breakpoints according to relative position is given in Table 3. While rare inversions fit a uniform

TABLE 1. Nested analysis of variance to test for differences between clusters and between species (within clusters) in length of polymorphic inversions.

Source of variation	SS	df	MS	F	Variance component
Between clusters	0.0775	2	0.0387	0.73	0%
Between species (within clusters)	0.3715	7	0.0531	5.63***	58.9%
Within species	0.3114	33	0.0094		41.1%

*** $P < 0.001$.

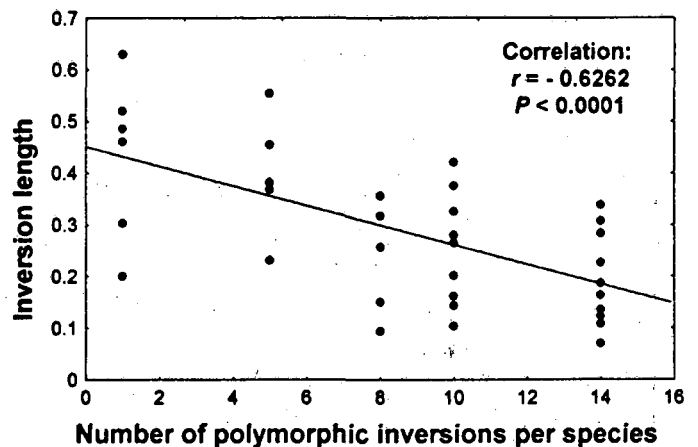


FIG. 2. Linear regression of polymorphic inversion length on number of polymorphic inversions per species. Species with one polymorphic inversion are *Drosophila richardsoni*, *D. borborema*, *D. serido* IV, *D. martensis*, *D. venezolana*, and *D. uniseti*. The other species are *D. serido* (5 inversions), *D. buzzatii* (8 inversions), *D. koepferae* (10 inversions) and *D. starneri* (14 inversions).

TABLE 2. Goodness of fit test of observed length distributions to those predicted by four different theoretical models.

Models tested	Fixed inversions		Polymorphic inversions		Rare inversions		Induced inversions	
	G	df	G	df	G	df	G	df
Uniform probability	35.2***	5	22.5***	5	24.5***	5	6.9	5
Truncated uniform	27.2***	4	14.8**	4	21.7***	4	6.8	4
Non-uniform probability ^a	28.2***	3	16.7***	3	—	—	—	—
Truncated non-uniform ^a	17.1***	2	3.0	2	—	—	—	—

** $P < 0.01$; *** $P < 0.001$.

^aIn non-uniform breakage probability models, probabilities of a break to occur in each chromosomal region (p_1 , p_2 , and p_3) were estimated from the observed frequency of breaks of each type of inversion (Table 3). Breaks of rare and induced inversions did not depart from the uniform distribution (Table 3). Thus, their length distributions were not fitted to the non-uniform breakage probability models.

distribution, significant differences between observed and expected distributions of breaks are found for fixed and polymorphic inversions. Breaks tend to accumulate in central regions of chromosomes. Further, for fixed inversions, breaks occur significantly more often in proximal regions than in distal regions (Table 3).

The distribution of the 122 inversion breakpoints according to their band location among the 74 segments established in chromosome 2 is shown in Table 4. When compared to the Poisson distribution a significant departure was found. Breakpoints are clustered, with segments that accumulate up to eight breaks and many segments showing no breaks at all. In several cases breaks fall apparently in the same chromosomal band. There are four bands with four (C6a, D1g, D5a, and E2e), one band with five (F6a), and one band with eight breakpoints (F2a). Considering that there are approximately 280 bands in chromosome 2, given a random distribution of breaks, the probability of a band having four, five, or eight breaks is 0.0014, 0.0001, or $4 \cdot 10^{-8}$, respectively. When inversions were grouped according to evolutionary success, the breakpoint distribution of fixed inversions deviated significantly from the Poisson distribution, but those of polymorphic and rare inversions did not. However, when polymorphic and rare inversions were grouped, their distribution of breakpoints also departed significantly from the Poisson distribution (Table 4).

TABLE 3. Distribution of inversion breakpoints among the three chromosomal regions according to their relative position and goodness of fit tests for comparison with the expected uniform distribution: G_T = all three regions; $G_{C \text{ vs } D+P}$ = central versus distal + proximal regions; $G_{D \text{ vs } P}$ = distal versus proximal region.

Chromosomal region	Number of breakpoints per region				
	All inversions	Fixed inversions	Polymorphic inversions	Rare inversions	Induced inversions
Proximal	58	17	24	17	10
Central	80	19	43	18	15
Distal	34	2	19	13	11
Total breaks	172	38	86	48	36
G_T (df = 2)	19.11***	18.03***	10.71**	0.90	1.13
$G_{C \text{ vs } D+P}$ (df = 1)	12.78***	4.47*	10.12**	0.37	1.09
$G_{D \text{ vs } P}$ (df = 1)	6.33*	13.55***	0.58	0.53	0.05

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Inversions Induced by Introgressive Hybridization

The distribution of the 18 induced inversions among the five chromosomes is homogeneous ($G = 7.29$; $df = 4$; $P = 0.12$), and contrasts with that of naturally occurring inversions (Ruiz and Wasserman 1993). The mean length (\pm SD) of induced inversions, 0.327 (\pm 0.183), differs significantly from that of rare inversions ($F = 9.29$; $df = 1, 40$; $P = 0.004$) but not from that of polymorphic ($F = 1.31$; $df = 1, 59$; $P = 0.26$) or fixed inversions ($F = 2.28$; $df = 1, 35$; $P = 0.14$). Induced inversions, unlike natural inversions, fit the uniform probability model, with or without truncation of small inversions (Table 2). In addition, the relative position of their breakpoints did not show differences among the distal, central, and proximal region of chromosomes (Table 3).

DISCUSSION

Our statistical analysis of the paracentric inversions of the *D. buzzatii* complex corroborates that inversion length is not a neutral trait, and provides strong support for the hypothesis that medium-sized inversions are favored by natural selection (Krimbas and Powell 1992). Successful inversions (fixed and polymorphic) show an accumulation of medium-sized inversions ($0.1 \leq l \leq 0.4$), while rare inversions are predominantly short. Furthermore, when polymorphic inversions are subdivided into two groups, polymorphic widespread and polymorphic endemic, according to whether they are present in more than or less than 25% of sampled localities of each species, significant differences in length are found ($F = 18.53$; $df = 1, 41$; $P < 0.001$). Widespread are closer to fixed inversions, whereas endemic are closer to rare inversions (Fig. 1B). Selection seems to generate a nearly gradual change in length, from small to medium sized, between less and more successful inversions.

Neither fixed nor polymorphic inversions fit the uniform breakage probability model (Table 2). Polymorphic inversions fit the truncated non-uniform breakage probability model. Fixed inversions do not, although fit improves significantly when the distribution of breakpoints is considered. Therefore, there exists a strong relationship between the non-uniform distribution of breakpoints and the observed length distribution. Figure 3 shows the expected proportion of breakpoints in the central region of chromosomes for inversions of different lengths. Medium-length inversions will show more breaks in the central region of chromosomes than ex-

TABLE 4. Observed distribution of inversion breakpoints according to their band location in the 74 segments defined in chromosome 2 and goodness of fit test for comparison with the expected number of segments according to the Poisson distribution.

Breaks/segment	Number of segments									
	All inversions		Fixed inversions		Polymorphic inversions		Rare inversions		Polymorphic and rare	
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp
0	24	14.2	61	55.0	34	31.2	46	45.5	26	19.2
1	20	23.5	8	16.3	23	26.9	21	22.1	21	25.9
2	12	19.3	3	2.4	12	11.7	6	5.4	11	17.5
3	7	10.6	1	0.24	3	3.4	1	0.87	9	7.9
4	3	4.4	0	0.02	2	0.73	0	0.11	5	2.7
5	5	1.4	1	0.00	0	0.13	0	0.02	2	0.72
6	2	0.40			0	0.02			0	0.16
7	0	0.09							0	0.04
8	1	0.03								
Total breaks	122		22		64		36		100	
G	21.64***		7.47**		1.00		0.12		8.66*	
df ^a	4		1		2		1		3	

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

^a Classes with expected values lower than 1 were pooled with adjacent classes in order to obtain a class with an expected value higher than 1.

pected at random, as observed in our fixed and polymorphic inversions. On the other hand, for small inversions only a slight excess of breaks in the central region of chromosomes is expected, as observed in our rare inversions. So, two sources of evidence—that of length differences and that of different proportions of breakpoints in the central region of chromosomes among inversion classes with different evolutionary success—reinforce the hypothesis that selection is actively determining the length distribution of inversions.

Our sample of inversions induced by introgressive hybridization fit the uniform breakage probability model and their breakpoints did not show a regional clustering along the chromosomes. These results agree with that of Federer et al. (1967) with inversions induced by X-ray treatment and those of Tonzetich et al. (1988) with breaks induced by γ -irradiation. One could reasonably conceive that rare inversions

would represent a sample of newly arisen inversions. However, they differ significantly in length from our sample of induced inversions and do not fit the uniform breakage probability model. Most rare inversions are probably not recently arisen, but rather have been floating in nature for a long time (Olvera et al. 1979). Seven of our rare inversions have been detected only once in the intensively studied species *D. buzzatii* and can be described as unique (Ruiz et al. 1984). Their mean length approximates to that of induced inversions and their length distribution does not differ from that of the uniform breakage probability model ($G = 1.46$; $df = 4$; $P = 0.23$) (Fig. 1C). Thus, these could be representative of new inversions. The remaining 17 inversions are endemic to a single or a few localities, but probably have persisted there for many generations.

How could the relationship between length of inversions and their success be explained? First, the longer an inversion, the higher the expected frequency of double crossing over within the inverted segment. Thus, selection is expected to operate against long inversions due to the semisterility of heterokaryotypes (Sturtevant and Beadle 1936; Navarro et al. 1997) and also due to the disruption of coadapted gene complexes. Second, the initial selective advantage of a newly arisen inversion is proportional to the epistasis between the trapped genes and their recombination rate (Charlesworth and Charlesworth 1973). To be highly successful, an inversion must usually have a minimum size, both to capture at least two epistatic genes and to capture genes separated by enough recombination rate. These two counteracting factors could explain the observed selection favoring medium-length inversions. Accordingly, small inversions would have only a relatively slight selective advantage that allows them to persist for a long time in the populations but not to increase in frequency or spread to a larger portion of the species range.

In our study a significant negative correlation between inversion length and number of polymorphic inversions per species was found. This pattern, not reported previously in the literature, seems to hold true over a wide taxonomic range of *Drosophila* species (M. Cáceres, A. Barbadilla, and A.

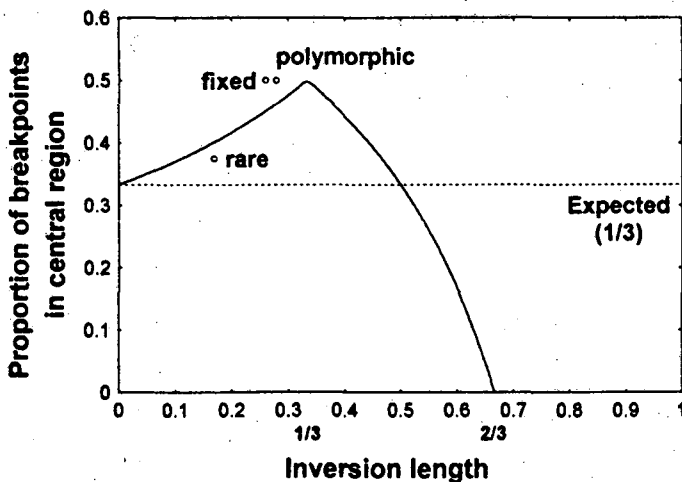


FIG. 3. Expected proportion of breakpoints in the central region of chromosomes (P) for inversions of different lengths. P -values were obtained following the reasoning of Appendix: for $0 \leq l \leq 1/3$, $P = (1/3)/1 - l$; for $1/3 \leq l \leq 2/3$, $P = [(2/3) - l]/1 - l$; for $2/3 \leq l \leq 1$, $P = 0$. Dashed line is the expected P -value assuming uniform distribution of breaks. Observed values for fixed, polymorphic, and rare inversions are shown.

Ruiz, unpubl. data). How can we account for it? In some cases, in *D. subobscura* and other species, two nonoverlapping inversions in the same chromosome almost completely inhibit recombination in the region between them (Krimbas and Loukas 1980; Krimbas 1992). Thus, the recombination-reducing effect of an inversion occurring in a chromosome already segregating for another inversion may be much greater than expected by its length and span over the entire segment encompassing both inversions. In such cases, small inversions could be favored by selection and the average length of polymorphic inversions would decrease.

The distribution of inversion breakpoints was studied as band location and as relative position. This allowed us to differentiate between content properties (those related with size, organization, or particular DNA sequence of each band) and positional properties (those related with the overall organization of chromosomes and selection for specific inversion length) as factors affecting the distribution of breaks (Tonzetich et al. 1988). Selection of small- and medium-length inversions could explain the accumulation of breakpoints in the central region of the chromosomes (Fig. 3). In addition, natural inversions breakpoints tend to concentrate in the proximal more than in the distal third of the chromosome. Such a pattern, found also in other species (Tonzetich et al. 1988; Krimbas and Powell 1992), could be explained by the deleterious effect of simultaneous crossovers in the inverted and proximal segments of an heterokaryotype, which favors inversions located near the centromere (Navarro et al. 1997). On the other hand, the non-random distribution of breakpoints according to band location is not surprising. Wasserman (1992) already noted that within the *repleta* group, although 208 inversions have been described, only 323 different breakpoints have been recorded. Moreover, in other species of *Drosophila* similar results have been found (Tonzetich et al. 1988; Krimbas and Loukas 1980; Lemeunier and Aulard 1992). Our analysis suggests that content differences between chromosomal segments cause the non-random distribution of breakpoints. Either the particular organization of DNA or target sites of break-producing agents could be involved. Transposable elements have been shown to induce chromosomal rearrangements in laboratory experiments with *Drosophila* (Lim and Simmons 1994). However, their implication in the generation of natural inversions remains unclear. To date, the molecular characterization of natural inversion breakpoints has provided no evidence of the action of transposable elements (Wesley and Eanes 1994; Cirera et al. 1995). In the *buzzatii* complex, bands C6a, D1g, D5a, E2e, F6a, and F2a are clear candidates for breakage hot spots. Further studies at the molecular level will allow their characterization and the distinction between the processes proposed for the generation of inversions.

ACKNOWLEDGMENTS

We wish to thank sincerely A. Navarro for his computer assistance and H. Naveira for his introgressive hybridization data. We also thank A. Berry and A. Caballero for discussion of results and two anonymous referees for their valuable comments. Work was supported by a FI fellowship from the Commissionat per a Universitats i Recerca (Generalitat de Catal-

unya, Spain) to MC and grant PB95-0607 from the Dirección General de Investigación Científica y Técnica (Ministerio de Educación y Ciencia, Spain) awarded to AR.

LITERATURE CITED

- BREHM, A., AND C. B. KRIMBAS. 1991. Inversion polymorphism in *Drosophila obscura*. *J. Hered.* 82:110-117.
- CHARLESWORTH, B., AND D. CHARLESWORTH. 1973. Selection of new inversions in multilocus genetic systems. *Genet. Res.* 21: 167-183.
- CIRERA, S., J. M. MARTÍN-CAMPOS, C. SEGARRA, AND M. AGUADÉ. 1995. Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. *Genetics* 139:321-326.
- CRUMPACKER, D. W., AND C. D. KASTRITSIS. 1967. A new gene arrangement in the third chromosome of *Drosophila pseudoobscura*. *J. Hered.* 58:2-6.
- FEDERER, W. T., R. G. D. STEEL, AND B. WALLACE. 1967. Mathematical models for chromosomal inversions. *Genetics* 55:783-795.
- HARVEY, P. H., AND M. D. PAGEL. 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford.
- KRIMBAS, C. B. 1992. The inversion polymorphism of *Drosophila subobscura*. Pp. 127-220 in C. B. Krimbas and J. R. Powell, eds. *Drosophila inversion polymorphism*. CRC Press, Boca Raton, FL.
- KRIMBAS, C. B., AND M. LOUKAS. 1980. The inversion polymorphism of *Drosophila subobscura*. Pp. 163-234 in M. K. Hecht, W. C. Steere, and B. Wallace, eds. *Evolutionary biology*. Vol. 12. Plenum Press, New York.
- KRIMBAS, C. B., AND J. R. POWELL. 1992. Introduction. Pp. 1-52 in C. B. Krimbas and J. R. Powell, eds. *Drosophila inversion polymorphism*. CRC Press, Boca Raton, FL.
- KUHN, G. C. S., A. RUIZ, M. A. R. ALVES, AND F. M. SENE. 1996. The metaphase and polytene chromosomes of *Drosophila sericea* (*repleta* group; *mulleri* subgroup). *Rev. Brasil. Genét.* 19: 209-216.
- LEMEUNIER, F., AND S. AULARD. 1992. Inversion polymorphism in *Drosophila melanogaster*. Pp. 339-405 in C. B. Krimbas and J. R. Powell, eds. *Drosophila inversion polymorphism*. CRC Press, Boca Raton, FL.
- LIM, J. K., AND M. J. SIMMONS. 1994. Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *BioEssays* 16:269-275.
- NAVARRO, A., E. BETRÁN, A. BARBADILLA, AND A. RUIZ. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* 146:695-709.
- NAVEIRA, H., AND A. FONTDEVILA. 1985. The evolutionary history of *Drosophila buzzatii*. IX: High frequencies of new chromosome rearrangements induced by introgressive hybridization. *Chromosoma* 91:87-94.
- OLVERA, O., J. R. POWELL, M. E. DE LA ROSA, V. M. SALCEDA, M. I. GASO, J. GUZMAN, W. W. ANDERSON, AND L. LEVINE. 1979. Population genetics of Mexican *Drosophila*. VI. Cytogenetic aspects of the inversion polymorphism in *Drosophila pseudoobscura*. *Evolution* 33:381-395.
- RANZ, J. M., C. SEGARRA, AND A. RUIZ. 1997. Chromosomal homology and molecular organization of Muller's elements D and E in the *Drosophila repleta* species group. *Genetics* 145:281-295.
- RUIZ, A., AND M. WASSERMAN. 1993. Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity* 70:582-596.
- RUIZ, A., H. NAVEIRA, AND A. FONTDEVILA. 1984. La historia evolutiva de *Drosophila buzzatii*. IV. Aspectos citogenéticos de su polimorfismo cromosómico. *Genética Ibérica* 36:13-35.
- SILVA, A. F. G., AND F. M. SENE. 1991. Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). *Rev. Brasil. Entomol.* 35:455-468.

SOKAL, R. R., AND F. J. ROHLF. 1995. *Biometry*. Freeman, New York.

SPICER, G. S. 1995. Phylogenetic utility of the mitochondrial cytochrome oxidase gene: molecular evolution of the *Drosophila buzzatii* species complex. *J. Mol. Evol.* 41:749-759.

STURTEVANT, A. H., AND G. W. BEADLE. 1936. The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics* 21:544-604.

TIDON-SKLORZ, R., AND F. M. SENE. 1995. *Drosophila seriema*: A new member of the *Drosophila serido* (Diptera, Drosophilidae) superspecies taxon. *Ann. Entomol. Soc. Am.* 88:139-142.

TONZETICH, J., T. W. LYTTLE, AND H. L. CARSON. 1988. Induced and natural break sites in the chromosomes of Hawaiian *Drosophila*. *Proc. Nat. Acad. Sci. USA* 85:1717-1721.

TOSI, T., AND F. M. SENE. 1989. Further studies on chromosomal variability in the complex taxon *Drosophila serido* (Diptera, Drosophilidae). *Rev. Brasil. Genét.* 12:729-745.

VAN VALEN, L., AND R. LEVINS. 1968. The origins of inversion polymorphism. *Am. Nat.* 102:5-23.

WALLACE, B. 1954. Coadaptation and the gene arrangements of *Drosophila pseudoobscura*. *Int. Union Biol. Sci. Symp. Genet. Popul. Struct. Ser. B* 15:67-94.

WASSERMAN, M. 1992. Cytological evolution of the *Drosophila repleta* species group. Pp. 455-541 in C. B. Krimbas and J. R. Powell, eds. *Drosophila inversion polymorphism*. CRC Press, Boca Raton, FL.

WESLEY, C. S., AND W. F. EANES. 1994. Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L)Payne* in *Drosophila melanogaster*. *Proc. Nat. Acad. Sci. USA* 91:3132-3136.

WHARTON, L. P. 1942. Analysis of the *repleta* group of *Drosophila*. *Univ. Texas Publ.* 4228:23-52.

Corresponding Editor: W. Rice

APPENDIX

Here we derive the formulae of the non-uniform breakage probability model. Consider a chromosome of length *N* sites

(e.g., base pairs). Let *p*₁, *p*₂, and *p*₃ be the probabilities of a break to occur in the distal, central, and proximal regions, respectively, each of one third of the chromosome ($\sum p_i = 1$). The probability of a break to occur at a given site within each of these regions is *p*₁/(*N*/3), *p*₂/(*N*/3), and *p*₃/(*N*/3). The density probability function of an inversion of length *L* depends on the length of the inversion as follows. The total number of inversions of length *L* that can occur in the chromosome is (*N* - *L* + 1). All inversions of length between 2*N*/3 and *N* will have their breakpoints in the two extreme regions. [(2*N*/3) - *L* + 1] inversions of length between *N*/3 and 2*N*/3 will have their breakpoints in adjacent regions (distal and central or central and proximal), and the rest in the two extreme regions. Finally, [(*N*/3) - *L* + 1] inversions of length between 0 and *N*/3 will have their breakpoints within each region, and the rest in adjacent regions. When length of inversions is expressed relative to the total length of the chromosome (*l*) and *N* is assumed to be very large, the following continuous density functions are obtained:

For $0 \leq l \leq \frac{1}{3}$

$$f(l) = 18 \left[\left(\frac{1}{3} - l \right) p_1^2 + \left(\frac{1}{3} - l \right) p_2^2 + \left(\frac{1}{3} - l \right) p_3^2 + l p_1 p_2 + l p_2 p_3 \right]$$

For $\frac{1}{3} \leq l \leq \frac{2}{3}$

$$f(l) = 18 \left[\left(l - \frac{1}{3} \right) p_1 p_3 + \left(\frac{2}{3} - l \right) p_1 p_2 + \left(\frac{2}{3} - l \right) p_2 p_3 \right]$$

For $\frac{2}{3} \leq l \leq 1$

$$f(l) = 18[(1 - l)p_1 p_3]$$

ARTÍCULO IV

Recombination Rate Predicts Inversion Size in Diptera

Mario Cáceres, Antonio Barbadilla y Alfredo Ruiz

1999

Genetics **153**: 251-259

Recombination Rate Predicts Inversion Size in Diptera

Mario Cáceres, Antonio Barbadilla and Alfredo Ruiz

Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Manuscript received November 15, 1998

Accepted for publication May 3, 1999

ABSTRACT

Most species of the *Drosophila* genus and other Diptera are polymorphic for paracentric inversions. A common observation is that successful inversions are of intermediate size. We test here the hypothesis that the selected property is the recombination length of inversions, not their physical length. If so, physical length of successful inversions should be negatively correlated with recombination rate across species. This prediction was tested by a comprehensive statistical analysis of inversion size and recombination map length in 12 Diptera species for which appropriate data are available. We found that (1) there is a wide variation in recombination map length among species; (2) physical length of successful inversions varies greatly among species and is inversely correlated with the species recombination map length; and (3) neither the among-species variation in inversion length nor the correlation are observed in unsuccessful inversions. The clear differences between successful and unsuccessful inversions point to natural selection as the most likely explanation for our results. Presumably the selective advantage of an inversion increases with its length, but so does its detrimental effect on fertility due to double crossovers. Our analysis provides the strongest and most extensive evidence in favor of the notion that the adaptive value of inversions stems from their effect on recombination.

OVER 60 years of research on *Drosophila* inversion polymorphism have yielded a vast amount of empirical information. However, besides the notion that some kind of balancing selection is responsible for its maintenance in natural populations (SPERLICH and PFRIEM 1986; KRIMBAS and POWELL 1992; POWELL 1997), very few genus-wide generalizations have been made to settle which property, if any, is selected for in evolutionarily successful inversions. A consistent relationship has been found between size and frequency of inversions in different species: rare or endemic (unsuccessful) inversions are usually small, while polymorphic or fixed (successful) inversions are predominantly medium-sized (OLVERA *et al.* 1979; BREHM and KRIMBAS 1991; KRIMBAS 1992; CÁCERES *et al.* 1997). This suggests that natural selection discriminates among inversions of different sizes, likely favoring those of intermediate physical length (KRIMBAS and POWELL 1992; POWELL 1997).

Because the most conspicuous effect of paracentric inversions is the substantial reduction of recombination within the inverted chromosomal segment in heterozygous individuals (STURTEVANT and BEADLE 1936; NAVARRO *et al.* 1997), it appears that natural selection operates upon polymorphic inversions through their effect on recombination (KOJIMA and SCHAFFER 1964; DOBZHANSKY 1970; CHARLESWORTH and CHARLESWORTH 1973; CHARLESWORTH 1974; ÁLVAREZ and

ZAPATA 1997). Accordingly, selection would act on recombination length of inversions, not physical length, and the observed relationship between evolutionary success and physical length of inversions would merely be a reflection of the actual correlation between physical and recombination length. Interspecific comparisons are crucial to test this hypothesis. *Drosophila* species seem to differ both in physical length of inversions (CÁCERES *et al.* 1997) and in their recombination map length (TRUE *et al.* 1996). If recombination length of inversions, rather than physical length, were the property selected for, differences in physical length of inversions among species would be accounted for by differences in recombination map length among species.

On the basis of the observed intermediate size of successful inversions, a simple selective model can be considered, by which natural selection favors a constant optimal recombination length (γ) of inversions across species (a less constrained and more realistic model would assume an optimal interval, but the conclusions would not differ qualitatively from those of this simpler model). This would be the net result of the selective advantage gained by the reduction of recombination and the detrimental effect on fertility of the inversions (see DISCUSSION). If species vary in their genome recombination map length (G), selection for optimal recombination length of inversions will result in variable optimal physical lengths (L) among species. Let L_i and G_i be, respectively, the optimal physical length of inversions and the recombination map length of species i , then

$$L_i = \gamma \frac{1}{G_i}$$

Corresponding author: Mario Cáceres, Departament de Genètica i de Microbiologia, Facultat de Ciències, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain.
E-mail: marioacs@cc.uab.es

Taking physical length of inversions in relative units, as a proportion of the total physical length of the genome, makes our model independent of the genome length of each species, which could vary among them. Thus, this model predicts a positive linear correlation across species between the relative physical length of selectively favored or successful inversions (that should be distributed around the species optimal physical length) and $1/G_i$. Conversely, the physical length of unsuccessful inversions, *i.e.*, those with zero or low selective advantage, is expected to be unrelated to the species recombination map length. We can test this model via its predictions, with unsuccessful inversions being used as the control data set.

In this article we carry out a comprehensive statistical analysis of inversion size and recombination map length in Diptera species for which appropriate data are available. The results corroborate the above predictions. We have found (1) that recombination map length varies greatly across species; (2) significant differences in length among species in successful inversions, but not in unsuccessful inversions; and (3) contrasting results, depending on the class of inversion considered, for the relationship between the two variables. In successful inversions, physical length of inversions and species recombination map length are strongly correlated, but this is not so in unsuccessful inversions.

MATERIALS AND METHODS

An exhaustive search of relevant data was made in the extant literature. We studied Diptera species with (1) detailed cytological maps of the salivary gland chromosomes and unambiguous descriptions of chromosomal inversions and (2) recombination (linkage) maps of at least one chromosome. Ten *Drosophila* species (*Drosophila ananassae*, *D. buzzatii*, *D. funebris*, *D. hydei*, *D. mediopunctata*, *D. melanogaster*, *D. persimilis*, *D. prosaltans*, *D. pseudoobscura*, and *D. subobscura*) and two *Anopheles* species (*A. gambiae* and *A. stephensi*) met our requirements.

Physical length of inversions was measured as the distance between the two breakpoints relative to the total physical length of the genome (the euchromatic portion represented in the polytene chromosome maps) in percentage. For each inversion, the length was obtained by comparison with the arrangement from which it originated (OLVERA *et al.* 1979), assuming that the major determinant of the fate of an inversion is its interaction with the parental chromosome (because this will be the one with which any new chromosomal arrangement will be combined as heterozygote most frequently). In *D. buzzatii*, *D. pseudoobscura*, and *D. subobscura* the data were directly taken from previous studies (OLVERA *et al.* 1979; KRIMBAS 1992; CÁCERES *et al.* 1997). Because they are scarce and have distinct genetic dynamics, neither pericentric nor sex-ratio inversions were included in the data set. According to their evolutionary success, paracentric inversions were classified as "successful" and "unsuccessful," and both groups were analyzed separately. An *a priori* criterion, based on the geographical distribution and frequency of the different chromosomal arrangements, was adopted to maximize the proportion of selectively favored inversions in the successful class. Only those inversions described previously as common and widespread (OLVERA *et al.* 1979; MOORE and TAYLOR 1986; KRIM-

BAS 1992; LEMEUNIER and AULARD 1992) or that have been reported as present in a significant fraction of the species distribution, namely at least 25% of the sampled localities (CÁCERES *et al.* 1997), were considered successful. All other inversions, which are rare or restricted to a few localities, were considered unsuccessful. The 207 inversions used in the study are listed in Table 1.

Recombination map data available for the 12 species are given in Table 2. When necessary, the published linkage map of each chromosome was corrected with the widely used Kosambi mapping function (CROW 1990). Then, its length was multiplied by $(n + 1)/(n - 1)$ to account for the different number of markers (n) per chromosome (CHAKRAVARTI *et al.* 1991). To estimate the total recombination map length of a given species, G_i , the lengths of all chromosomes in the haploid set were summed. When no data were available for a given chromosome, its recombination length was inferred from the recombination lengths of other chromosome(s) for which data were available by assuming proportionality with its physical length (in the polytene chromosome maps). To test this assumption of proportionality, we computed the correlation between the relative physical length and the relative recombination length of chromosomes. We used the recombination data of species with at least two mapped chromosomes (8 out of 12 species; Table 2). Because the variables take relative values, only $a - 1$ of the a pairs of values of a species are independent, and we omitted one data pair (chosen at random) of each of the 8 species. The Pearson r value was 0.76 (d.f. 13; $P = 0.0010$), which supports our estimation procedure. In *Drosophila*, recombination is limited to females, while in *Anopheles gambiae* and *A. stephensi*, males and females have similar recombination rates (PARVEZ *et al.* 1985; ZHENG *et al.* 1996). Accordingly, to make recombination values directly comparable between different chromosomes and different genera, the estimated G_i were multiplied by $1/2$ for *Drosophila* autosomal inversions and by $2/3$ for X-linked inversions (BEGUN and AQUADRO 1992). However, the different recombination value for the X chromosome with regard to the autosomes, as well as the absence of X-chromosome inversions in 9 of the 12 species, renders the statistical analysis of inversion length and recombination unbalanced and makes it intractable. Therefore, X-linked inversions (4 successful and 14 unsuccessful) were omitted from the correlation analysis, although it should be noted that their behavior is very similar to that of autosomal inversions (see Figure 1).

RESULTS

While the two *Anopheles* species have a relatively small recombination map length, there is remarkable variation in recombination map length among *Drosophila* species, from 285.4 cM in *D. prosaltans* to 1007.6 cM in *D. subobscura* (Table 3). Mean length of successful inversions also varies greatly among species, from 3.64% in *D. mediopunctata* to 11.45% in *D. prosaltans*, and the differences are statistically significant as shown by the ANOVA ($F = 5.28$; d.f. 11, 70; $P < 0.0001$) or the non-parametric Kruskal-Wallis test ($H = 27.92$; d.f. 11; $P = 0.0033$). To test the effect of species recombination map length on physical length of successful inversions, we performed an analysis of variance and regression, where the variation among species in physical length of inversions is partitioned into linear and nonlinear components (Table 4). The F -test showed a very significant

TABLE 1
Inversions used in the study classified according to their evolutionary success

Species	Successful inversions	Unsuccessful inversions	Reference ^a
<i>D. ananassae</i>	2LA, 3LA, 3RA	2LB, 2LC, 2RA, 3RD	(1, 2)
<i>D. buzzatii</i>	2j, 2z, 2q, 4s	2y, 2c, 2d, 2e, 2g, 2h, 2f, 2j, 2r, 2s, 3f, 5c	(3, 4)
<i>D. funebris</i>	II-1, II-2, II-3, III-1, IV-1		(5)
<i>D. hydei</i>	2a ²		(6)
<i>D. mediopunctata</i>	2 AB, 2 AC		(7)
<i>D. melanogaster</i>	(2L)NS, (2L)t, (2R)NS, (3L)M, (3L)P, (3R)C, (3R)K, (3R)M, (3R)Mo, (3R)P	(1)12A;18D, (1)16D;18D, (2L)A, (2L)W, (2L)22A;26B, (2R)NC, (2R)O, (2R)49B;56A, (3L)L, (3L)Y, (3L)62D;68A	(8)
<i>D. persimilis</i>	CO, KL, MD, RD, SE, WT	HU, MA, MR, NA, TP, TU, VI, WA, WE	(9-13)
<i>D. prosaltans</i>	PXL _a , PXL _d , PIII _a , PIII _a	PXL _b , PXL _c , PXL _e , PXL _f , PXL _g , PXR _a , PIII _b , PIII _c , PIII _a , PIII _b	(14, 15)
<i>D. pseudoobscura</i>	AR, CH, CU, EP, HI, OA, OL, PP, SC, TA, TL	AF, AM, BE, CC, EB, FC, HY, IZ, MA, MI, OZ, PA, PI, PO, SA, SB, SJ, SO, TE, TH, UR, VA, ZI	(16)
<i>D. subobscura</i>	A ₁ , A ₂ , J ₁ , J ₃ , J ₄ , U ₁ , U ₂ , U ₃ , U ₄ , U ₅ , U ₇ , U ₈ , E ₁ , E ₂ , E ₃ , E ₄ , E ₈ , E ₉ , E ₁₂ , O ₁ , O ₂ , O ₃ , O ₄ , O ₅ , O ₆ , O ₇ , O ₈ , O ₂₂	A ₃ , A ₄ , A ₅ , A ₆ , A ₇ , A ₈ , J ₅ , U ₅ , U ₉ , U ₁₀ , U ₁₁ , U ₁₂ , E ₅ , E ₁₃ , E ₁₄ , E ₁₅ , E ₁₆ , E ₁₇ , E ₁₈ , E ₁₉ , E ₂₀ , O ₉ , O ₁₀ , O ₁₁ , O ₁₂ , O ₁₃ , O ₁₄ , O ₁₅ , O ₁₆ , O ₁₇ , O ₁₈ , O ₁₉ , O ₂₀ , O ₂₁ , O ₂₃ , O ₂₄ , O ₈	(17)
<i>A. gambiae</i>	2La, 2Rb, 2Rc, 2Rd, 2Rj, 2Rk, 2Ru	2Rt	(18, 19)
<i>A. stephensi</i>	2Rb	2Rc, 2Rd, 2Re, 2Rf, 2Lc, 2Ld, 3Ra, 3Rb, 3Rc, 3Lb, 3Lc, 3Ld, 3Le, 3Lf, 3Lg, 3Lh, 3Li, 3Lj	(20, 21)

^aReferences: (1) FUTCH (1996); (2) HINTON and DOWNS (1975); (3) RUIZ *et al.* (1984); (4) BARKER *et al.* (1985); (5) DUBININ and TINIAKOV (1946); (6) WASSERMAN (1962); (7) KASTRITSIS (1966); (8) LEMEUNIER and AULARD (1992); (9) DOBZHANSKI and EPLING (1944); (10) DOBZHANSKI (1948); (11) SPIESS (1950); (12) SPIESS (1965); (13) BECKENBACH (1986); (14) BICUDO (1973); (15) BICUDO *et al.* (1978); (16) OLVERA *et al.* (1979); (17) KRIMBAS (1992); (18) COLUZZI and SABATINI (1967); (19) COLUZZI *et al.* (1979); (20) COLUZZI *et al.* (1973); (21) MAHMOOD and SAKAI (1984).

correlation between physical length of inversions and $1/G_i$. However, the variances within species were not homogeneous, even after log transformation. Thus, we also performed a resampling test as follows. A random sample with replacement is drawn of the inversion length data (78 inversions in our case). Let N_i be the number of inversions analyzed in species i . N_1, N_2, \dots, N_{12} sampled values are assigned randomly to species 1, 2, \dots , 12. To the whole random sample, the same analysis of variance and regression shown in Table 4 is applied, and the F value computed. The F distribution of 10,000 replicates obtained with this procedure is compared with the F empirical value, and its significance is estimated. This analysis corroborated the parametric significant probabilities (Table 4). As shown in Figure 1A, the average physical length of successful inversions decreases as species recombination map length increases.

Clearly contrasting results were obtained for unsuccessful inversions. First, they are smaller than successful inversions (mean length 4.72% vs. 6.40%) as shown by a t -test ($t = 4.10$; d.f. 205; $P < 0.0001$) or a sign test ($P = 0.0391$). Second, the weighted mean variance in

inversion length within species (computed dividing each inversion length value by the species average of its class to take out the length differences) is larger for the unsuccessful inversions than for the successful ones ($F = 2.33$; d.f. 70, 116; $P = 0.0001$). Third, no significant differences in mean length of unsuccessful inversions among species are found (ANOVA: $F = 1.31$; d.f. 8, 116; $P = 0.25$; Kruskal-Wallis test: $H = 13.16$; d.f. 8; $P = 0.11$). Finally, the correlation between physical length of inversions and species recombination map length is not significant (Table 4). Therefore, there is no relationship between physical length of unsuccessful inversions and recombination map length (Figure 1B).

So far, we have considered the mean values of species as independent data points. This implicitly assumes that polymorphic inversions are not inherited across species but arise *de novo* in each species (see DISCUSSION). It could be argued, nevertheless, that closely related species share their trait values because of common ancestry (FELSENSTEIN 1985; HARVEY and PAGEL 1991). Figure 2 shows the phylogenetic relationships of the 12 species. To test for a phylogenetic clustering of recombination values, we performed a nested ANOVA at four taxo-

TABLE 2
Recombination data for the 12 species of Diptera

Species	Chromosome	Number of markers (<i>n</i>)	Linkage map (cM)	Kosambi's correction	(<i>n</i> + 1)/(<i>n</i> - 1) correction	Reference ^a
<i>D. ananassae</i>	X	31	106.9	107.6	114.8	(1)
	2	27	113.1	114.1	122.9	(1)
	3	26	103.5	104.1	112.4	(1)
<i>D. buzzatii</i>	X	14	108.9	108.9	125.7	(2)
	2	11	138.5	138.5	166.2	(2)
<i>D. funebris</i>	X	13	158.0	172.2	200.9	(3)
<i>D. hydei</i>	X	25	115.7	117.6	127.4	(4)
<i>D. mediopunctata</i>	X	6	89.0	94.4	132.2	(5)
<i>D. melanogaster</i>	X	868	73.1	73.1	73.3	(6)
	2	624	110.0	110.0	110.4	(6)
	3	573	110.9	110.9	111.3	(6)
<i>D. persimilis</i>	X	11	164.5	191.6	229.9	(7)
<i>D. prosaltans</i>	X	14	100.0	105.5	121.7	(8)
	2	7	59.0	61.3	81.7	(8)
	3	5	52.0	54.7	82.0	(8)
<i>D. pseudoobscura</i>	X	34	228.2	228.2	242.0	(9)
	2	11	101.3	113.8	136.6	(10)
	3	12	68.0	69.4	82.0	(10)
	4	7	69.2	72.4	96.5	(10)
<i>D. subobscura</i>	A	14	150.0	167.6	193.4	(11)
	O	17	228.3	226.1	254.4	(12)
<i>A. gambiae</i>	E	6	107.4	108.4	151.8	(12)
	X	46	48.9	48.9	51.1	(13)
	2	57	72.4	72.4	75.0	(13)
<i>A. stephensi</i>	3	28	93.7	93.7	100.6	(13)
	2	7	98.4	128.4	171.2	(14)
	3	6	77.2	89.1	124.7	(14)

^a References: (1) TOBARI (1993); (2) SCHAFER *et al.* (1993); (3) PERJE (1955); (4) HESS (1976); (5) SAMPAIO *et al.* (1996); (6) LINDSLEY and ZIMM (1992); (7) BEERS (1937); (8) SPASSKY *et al.* (1950); (9) ORR (1995); (10) ANDERSON (1990); (11) SPURWAY (1945); (12) LOUKAS *et al.* (1979); (13) ZHENG *et al.* (1996); (14) PARVEZ *et al.* (1985).

onomic levels: genus, subgenus, group, and species (BELL 1989; HARVEY and PAGEL 1991). Most of the variation was due to the "between groups within subgenera" level, which explained 68% of the total variation (Table 5). A nested ANOVA was also applied to the physical length of inversions. For successful inversions, we found again that the group level explained most of the variation in inversion length (50%). On the other hand, for unsuccessful inversions, nearly all the length variation (~96%) was found within species and none in the other taxonomic levels (Table 5). To account for the possible phylogenetic effect on the observed correlation between physical length of successful inversions and recombination map length, we repeated the same analysis of variance and regression of Table 4 on the eight species groups. Because little additional variance remains at higher levels for both variables, the groups can be considered as statistically independent. As in Table 4, the *F* value for the regression with $1/G_i$ (19.92) was highly significant either by the parametric (d.f. 1, 6; $P = 0.0043$) or the resampling test ($P = 0.0060$). We also applied Felsenstein's (1985) independent contrasts

method to our data (using the CAIC program of PURVIS and RAMBAUT 1995). The method requires the prior knowledge of the phylogeny and the branching times of species. We used the time estimates available in the literature either for our species or for closely related species (Figure 2; RUSSO *et al.* 1995; POWELL 1997). The correlation between physical length of successful inversions and $1/G_i$ for the independent contrasts was 0.47, which is marginally significant (d.f. 9; $P = 0.07$).

DISCUSSION

Our analysis shows that the physical length of evolutionarily successful inversions differs among species and that there is a negative correlation between inversion length and species recombination map length (Figure 1A) that explains a sizable part of inversion length variance among species (61%). For both variables, species recombination map length and physical length of successful inversions, the group level accounts for a significant proportion of the variance (Table 5). Hence, two alternative explanations are possible. First, the correla-

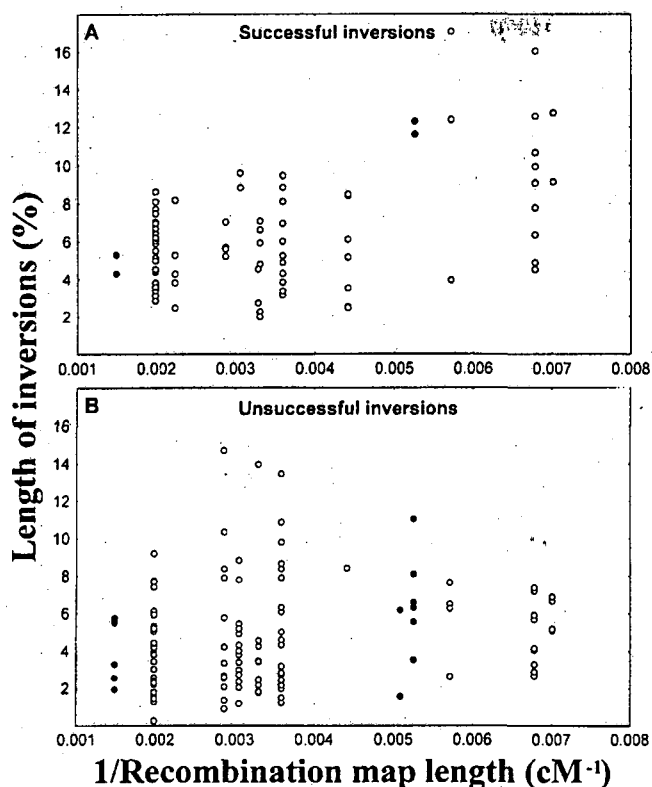


FIGURE 1.—Scatterplot of physical length of inversions (in percentage of the total physical length of the genome) vs. the inverse of the species recombination map length. (A) Successful inversions ($N = 82$) and (B) unsuccessful inversions ($N = 125$). X-linked inversions (4 successful and 14 unsuccessful) are shown as solid circles and display similar behavior as autosomal inversions (open circles).

tion could have a purely phylogenetic origin, *i.e.*, species could share both characters due to common ancestry. In this case, it should disappear when appropriate comparative methods that take into account the phylogeny are used. Second, there could be a real causal relationship between recombination map length and inversion length, which would result in a pseudophylogenetic correlation as a consequence of the phylogenetic clustering of recombination values. In this second case, the correlation could diminish but should not disappear entirely when comparative methods are used.

Two different arguments allow us to exclude a phylogenetic explanation for the observed correlation. First, when we obviate the phylogenetic effect, considering just the eight independent species groups (BELL 1989; HARVEY and PAGEL 1991), the same highly significant correlation is still found. Likewise, despite the reduction of statistical power, the independent contrasts method (FELSENSTEIN 1985) also yields a marginally significant correlation. Second, the phylogenetic explanation assumes that the species trait values are inherited from common ancestors. Polymorphic inversions are nevertheless very recent in evolutionary terms, and none of our inversions are shared between closely related species. One can still imagine that closely related species could share a mutational propensity to produce inversions within a given length range. However, this possibility can be safely excluded because then we would expect to observe the same effect on both successful and unsuccessful inversions, and this was certainly not the case.

The comparison between successful and unsuccessful inversions is critical for the interpretation of our data.

TABLE 3

Total recombination map length and mean length of inversions for the 12 species of Diptera

Species	Recombination map (cM)	Successful inversions			Unsuccessful inversions		
		\bar{L} (%)	SD	N	\bar{L} (%)	SD	N
<i>D. ananassae</i>	350.1	11.14	6.64	3	5.76	2.17	4
<i>D. buzzatii</i>	696.5	5.90	0.79	4	5.35	4.21	12
<i>D. funebris</i>	895.6	4.82	2.14	5	—	—	—
<i>D. hydei</i>	655.2	9.61	—	1	—	—	—
<i>D. mediopunctata</i>	609.6	3.64	1.28	2	—	—	—
<i>D. melanogaster</i>	294.9	9.09	3.52	10	4.64	1.95	11
<i>D. persimilis</i>	605.1	4.79	2.20	6	4.21	3.80	9
<i>D. prosaltans</i>	285.4	11.45	1.60	4	6.49	2.02	10
<i>D. pseudoobscura</i>	557.1	5.84	2.22	11	4.96	3.41	23
<i>D. subobscura</i>	1007.6	5.40	1.62	28	4.07	2.10	37
<i>A. gambiae</i>	226.7	5.24	2.55	7	8.40	—	1
<i>A. stephensi</i>	327.1	8.84	—	1	4.18	1.87	18
Total		6.40	3.06	82	4.72	2.77	125

\bar{L} (%), mean physical length of inversions expressed as the percentage of the total physical length of the genome; SD, standard deviation; N , number of inversions for each species.



TABLE 4
Analysis of variance and regression to test the effect of species recombination map length (G_i) on physical length of inversions

Source of variation	Successful inversions				Unsuccessful inversions			
	d.f.	MS	F	P^a	d.f.	MS	F	P^a
Among species	11	25.19	4.05	0.0002 (0.0011)	8	6.32	0.79	0.61 (0.53)
Regression $1/G_i$	1	167.84	15.36	0.0029 (0.0054)	1	16.38	3.36	0.11 (0.13)
Deviations	10	10.92	1.76	0.08 (0.08)	7	4.88	0.61	0.75 (0.67)
Within species	66	6.22			102	7.95		

X-linked inversions (4 successful and 14 unsuccessful) were excluded from this analysis. MS, mean square.

^a The probability was assessed by the parametric F -test and by a resampling test with 10,000 replicas (shown in parentheses).

Inversions that have survived and flourished, becoming common and widespread in natural populations (successful inversions), are a highly selected subset of all occurring inversions: those with the highest selective advantage. On the other hand, unsuccessful inversions are a mixture comprising unique (recently arisen) inversions and also some inversions currently endemic or restricted to a small portion of the species distribution. They are scattered over the entire range of recombination values and taxons, and, although some of them might have been partially sieved by natural selection, it is clear that their selective advantage cannot be large. As previously noted (OLVERA *et al.* 1979; CÁCERES *et al.* 1997), unsuccessful inversions are smaller than successful inversions and have a larger length variance within

species. Furthermore, their length does not differ across species and does not show either phylogenetic clustering around the group level or correlation with recombination map length. These clear contrasting results rule out any mutational cause for the length distribution of successful inversions among species and point to natural selection as the only possible explanation.

To explain the correlation between physical length of successful inversions and species recombination map length, we must assume then that natural selection is acting either (i) directly through recombination length of inversions, or (ii) indirectly through a trait correlated with recombination. Perhaps species with similar recombination values within the same taxon also share the same selective pressure on inversion length because of an unknown cause independent of recombination. This explanation does not seem quite parsimonious, and, given the results of the comparative analyses and the expected connection of inversion length with recombination, we believe that recombination-mediated selection is indeed responsible.

Two opposite selective forces seem to be acting on inversion length. On one side, three alternative but nonmutually exclusive theories predict a positive relationship between the selective advantage of an inversion and its length, *i.e.*, that long inversions should be favored. Under the genic selection model (NEI *et al.* 1967), the longer the inversion, the greater its selective advantage, provided that it contains few or no deleterious alleles (SANTOS 1986). The coadaptation theory proposes that inversions reduce recombination in the heterokaryotypes, allowing the capture of favorable allelic combinations (CHARLESWORTH and CHARLESWORTH 1973; CHARLESWORTH 1974) and the building up of coadapted gene complexes (KOJIMA and SCHAFER 1964; DOBZHANSKY 1970; ÁLVAREZ and ZAPATA 1997). Accordingly, the probability of "catching" two or more genes with epistatic effects on fitness increases with the size of the inversion, and the selective advantage gained

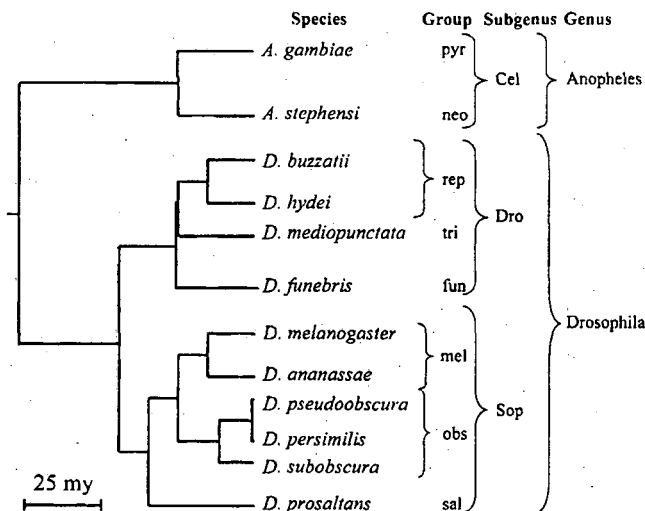


FIGURE 2.—Phylogenetic tree and branching times of the 12 species of Diptera used in the study. To test for a phylogenetic clustering of the variables, four taxonomic levels were considered (genus, subgenus, group, and species). Abbreviations: pyr, pyretophorus; neo, neocellia; rep, repleta; tri, tripunctata; fun, funebris; mel, melanogaster; obs, obscura; sal, saltans; Cel, Cellia; Dro, Drosophila; Sop, Sophophora.

TABLE 5
Nested analysis of variance to test for a phylogenetic clustering of species recombination map length and physical length of successful and unsuccessful inversions

Source of variation	Physical length of inversions														
	Recombination map length ($1/G_i$)					Successful inversions					Unsuccessful inversions				
	d.f.	MS $\times 10^6$	F_5	Variance component (%)	d.f.	MS	F_5	Variance component (%)	d.f.	MS	F_5	Variance component (%)			
Between genera	1	0.10	0.01	0.00	1	3.22	0.23	0.00	1	1.47	0.24	0.00			
Between subgenera	1	8.40	2.03	18.14	1	14.20	0.30	0.00	1	5.98	0.60	0.00			
Between groups	5	4.14	8.04*	68.02	5	46.98	7.39*	50.21	3	10.00	2.24	4.15			
Between species	4	0.51	—	13.84	4	6.36	1.02	0.16	3	4.47	0.56	0.00			
Within species	—	—	—	—	66	6.22	—	49.63	102	7.95	—	95.85			

* $P < 0.05$.

by the inversion increases with recombination distance between them (CHARLESWORTH and CHARLESWORTH 1973). Finally, stabilizing selection on a quantitative trait should favor those modifiers that reduce recombination, e.g., inversions, because they reduce the genetic variance of the trait and increase the mean fitness of the population (MATHER 1943; CHARLESWORTH 1993). If the modifier reduces the map length of a chromosome, its selective advantage would be proportional to the reduction of recombination (CHARLESWORTH 1993). In other words, if the modifier is an inversion, its selective advantage would be proportional to its recombination length.

On the other side, inversion length also has negative consequences on fertility. Because of the ordered oogenesis of females in *Drosophila* and other Diptera, the unbalanced chromosomes resulting from single crossovers within the inverted region of heterokaryotypes are always set into the polar bodies and no inviable zygotes are formed (STURTEVANT and BEADLE 1936; CARSON 1946). In *Drosophila*, there is no chiasma formation in males, and, in *Anopheles* spermatogenesis, the bridge between the two nuclei at anaphase I prevents their separation and their development into sperm (WHITE 1973). The real problem arises with double crossovers, which produce one-fourth of unbalanced gametes because four-strand double crossovers yield only unbalanced gametes (STURTEVANT and BEADLE 1936; ROBERTS 1976). This fertility effect selects against long inversions because of their increased probability of double crossovers (NAVARRO *et al.* 1997). The operation of the two opposite factors is evident from Figure 1A. Large inversions are totally absent in species with long recombination maps, while there is a relative paucity of small inversions in species with short recombination maps.

According to the previous discussion, selection favors larger inversions in species with a low recombination rate than in species with a high recombination rate. One might predict, consequently, that species with larger recombination maps should accumulate more inversions than species with shorter recombination maps, because the smaller inversions block only a relatively small part of the chromosome and leave space for new inversions to settle. Our results agree with this prediction. It seems that there is a positive correlation between the number of successful inversions per species and the recombination map length ($r = 0.47$), although, because we have only 12 species, the correlation was not statistically significant ($P = 0.12$). There are nevertheless differences among species in the presence or absence of inversions that do not fit into this explanation. For instance, *D. simulans* and *D. mauritiana*, with a recombination map similar to that of their close relative *D. melanogaster*, exhibit no inversion polymorphism. Likewise, *D. virilis*, with no known polymorphic inversions, has one of the largest recombination maps of the *Drosophila* genus. Thus other factors, such as the age of

the species or differences in the molecular mechanisms that generate inversions, e.g., the smaller level of middle repetitive DNA in *D. simulans* relative to *D. melanogaster* (DOWSETT and YOUNG 1982), should also be important determinants of the number of polymorphic inversions per species. In the *D. buzzatii* species complex, a significant negative correlation between length of polymorphic inversions and number of polymorphic inversions per species was observed (CÁCERES *et al.* 1997). However, in our genus-wide data set such correlation was not significant ($F = 1.14$; d.f. 1, 10; $P = 0.31$), which suggests that recombination, rather than number of inversions, is the main determinant of inversion size in different species.

The significant correlation observed between physical length of successful inversions and recombination map length is striking if one considers that the several sources of error underlying the diverse data used would tend to hamper our ability to detect a trend. Although we have estimated each species' recombination map length as accurately as possible, our estimates are inevitably approximate in some cases. Moreover, our model assumes that the distribution of crossovers along the chromosome is uniform, which is not always the case (TRUE *et al.* 1996). Relative inversion length was also measured approximately, because cytological distances are only rough estimates of real physical distances and to determine the exact location of the breakpoints is usually difficult. Finally, in some cases, inversions are found associated on the same chromosome (overlapping or in strong linkage disequilibrium), with more complicated recombination-reducing effects, depending on the chromosomal arrangements combined in each population, than just the recombination length of the inversion (KRIMBAS 1992). Thus, that the trend's signal could emerge so strongly is remarkable and suggests a consistent and general phenomenon. The population genetics of polymorphic inversions has long been one of the most powerful and mysterious examples of natural selection in action. Our statistical analysis of the available empirical evidence corroborates the notion that inversions are selected by their effects on recombination and adds further evidence to the importance of recombination differences among species in their evolutionary dynamics.

We appreciate very much the comments on the manuscript of A. Berry, B. Charlesworth, J. Coyne, R. C. Lewontin, A. Navarro, J. M. Ranz, and two anonymous referees. A. Berry considerably improved the article readability. We thank E. Betrán, M. Coluzzi, A. Fontdevila, A. M. de Ron, J. Rozas, C. Segarra, V. P. Sharma, S. K. Subbarao, and C. Zapata for providing us articles with linkage maps and/or descriptions of chromosomal inversions. Work was supported by a FI fellowship from the Comissionat per a Universitats i Recerca (Generalitat de Catalunya, Spain) awarded to M.C. and grant PB95-0607 from the Dirección General de Investigación Científica y Técnica (Ministerio de Educación y Cultura, Spain) awarded to A.R.

LITERATURE CITED

ÁLVAREZ, G., and C. ZAPATA, 1997 Conditions for protected inversion polymorphism under supergene selection. *Genetics* **146**: 717-722.

- ANDERSON, W. W., 1990 Linkage map of *Drosophila pseudoobscura*, pp. 3188-3189 in *Genetic Maps: Locus Maps of Complex Genomes, Vol. 3, Lower Eukaryotes*, edited by S. J. O'BRIEN. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- BARKER, J. S. F., F. M. SENE, P. D. EAST and M. A. Q. R. PEREIRA, 1985 Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. *Genetica* **67**: 161-170.
- BECKENBACH, A. T., 1986 The third chromosome inversion polymorphisms in northern populations of *Drosophila persimilis*. *Can. J. Genet. Cytol.* **28**: 401-408.
- BEERS, C. V., 1937 Linkage groups in *Drosophila pseudoobscura*, race B. *Genetics* **22**: 577-586.
- BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**: 519-520.
- BELL, G., 1989 A comparative method. *Am. Nat.* **133**: 553-571.
- BICUDO, H. E. M. C., 1973 Chromosomal polymorphism in the saltans group of *Drosophila*. I. The saltans subgroup. *Genetica* **44**: 520-552.
- BICUDO, H. E. M. C., M. K. HOSAKI, J. MACHADO and M. C. N. MARQUES, 1978 Chromosomal polymorphism in the saltans group of *Drosophila*. II. Further study on *D. prosaltans*. *Genetica* **48**: 5-15.
- BREHM, A., and C. B. KRIMBAS, 1991 Inversion polymorphism in *Drosophila obscura*. *J. Hered.* **82**: 110-117.
- CÁCERES, M., A. BARBADILLA and A. RUIZ, 1997 Inversion length and breakpoint distribution in the *Drosophila buzzatii* species complex: is inversion length a selected trait? *Evolution* **51**: 1149-1155.
- CARSON, H. L., 1946 The selective elimination of inversion dicentric chromatids during meiosis in the eggs of *Sciara impatiens*. *Genetics* **31**: 95-113.
- CHAKRAVARTI, A., L. K. LASHER and J. E. REEFER, 1991 A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* **128**: 175-182.
- CHARLESWORTH, B., 1974 Inversion polymorphism in a two-locus genetic system. *Genet. Res.* **23**: 259-280.
- CHARLESWORTH, B., 1993 Directional selection and the evolution of sex and recombination. *Genet. Res.* **61**: 205-224.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1973 Selection of new inversions in multilocus genetic systems. *Genet. Res.* **21**: 167-183.
- COLUZZI, M., and A. SABATINI, 1967 Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. *Parassitologia* **9**: 73-88.
- COLUZZI, M., M. DI DECO and G. CANCRINI, 1973 Chromosomal inversions in *Anopheles stephensi*. *Parassitologia* **15**: 129-136.
- COLUZZI, M., A. SABATINI, V. PETRARCA and M. DI DECO, 1979 Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* **73**: 483-497.
- CROW, J. F., 1990 Mapping functions. *Genetics* **125**: 669-671.
- DOBZHANSKY, TH., 1948 Genetics of natural populations. XVI. Altitudinal and seasonal changes produced by natural selection in certain populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. *Genetics* **33**: 158-176.
- DOBZHANSKY, TH., 1970 *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- DOBZHANSKY, TH., and C. EPLING, 1944 Contributions to the genetics, taxonomy, and ecology of *Drosophila pseudoobscura* and its relatives. *Carnegie Inst. Washington Publ.* **554**: 1-183.
- DOWSETT, A. P., and M. W. YOUNG, 1982 Differing levels of dispersed repetitive DNA among closely related species of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **79**: 4570-4574.
- DUBININ, N. P., and G. G. TINIAKOV, 1946 Structural chromosome variability in urban and rural populations of *Drosophila funebris*. *Am. Nat.* **80**: 393-396.
- FELSENSTEIN, J., 1985 Phylogenies and the comparative method. *Am. Nat.* **125**: 1-15.
- FUTCH, D. G., 1966 A study of speciation in South Pacific populations of *Drosophila ananassae*. *Univ. Texas Publ.* **6615**: 79-120.
- HARVEY, P. H., and M. D. PAGEL, 1991 *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- HESS, O., 1976 Genetics of *Drosophila hydei* Sturtevant, pp. 1343-1363 in *The Genetics and Biology of Drosophila*, Vol. 1c, edited by M. ASHBURNER and E. NOVITSKI. Academic Press, New York.
- HINTON, C. W., and J. E. DOWNS, 1975 The mitotic, polytene, and meiotic chromosomes of *Drosophila ananassae*. *J. Hered.* **66**: 353-361.

- KASTRITSIS, C. D., 1966 Cytological studies on some species of the tripunctata group of *Drosophila*. Univ. Texas Publ. 6615: 413-473.
- KOJIMA, K., and H. E. SCHAFER, 1964 Accumulation of epistatic gene complexes. *Evolution* 18: 127-131.
- KRIMBAS, C. B., 1992 The inversion polymorphism of *Drosophila subobscura*, pp. 127-220 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Boca Raton, FL.
- KRIMBAS, C. B., and J. R. POWELL, 1992 Introduction, pp. 1-52 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Boca Raton, FL.
- LEMEUNIER, F., and S. AULARD, 1992 Inversion polymorphism in *Drosophila melanogaster*, pp. 339-405 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Boca Raton, FL.
- LINDSLEY, D. L., and G. G. ZIMM, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.
- LOUKAS, M., C. B. KRIMBAS, P. MAVRAGANI-TSIPIDOU and C. D. KASTRITSIS, 1979 Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. *J. Hered.* 70: 17-26.
- MAHMOOD, F., and R. K. SAKAI, 1984 Inversion polymorphisms in natural populations of *Anopheles stephensi*. *Can. J. Genet. Cytol.* 26: 538-546.
- MATHER, K., 1943 Polygenic inheritance and natural selection. *Biol. Rev.* 18: 32-64.
- MOORE, B. C., and C. E. TAYLOR, 1986 *Drosophila* of southern California. III. Gene arrangements of *Drosophila persimilis*. *J. Hered.* 77: 313-323.
- NAVARRO, A., E. BETRÁN, A. BARBADILLA and A. RUIZ, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* 146: 695-709.
- NEI, M., K. KOJIMA and H. E. SCHAFER, 1967 Frequency changes of new inversions in populations under mutation-selection equilibria. *Genetics* 57: 741-750.
- OLVERA, O., J. R. POWELL, M. E. DE LA ROSA, V. M. SALCEDA, M. I. CASO *et al.*, 1979 Population genetics of Mexican *Drosophila*. VI. Cytogenetic aspects of the inversion polymorphism in *Drosophila pseudoobscura*. *Evolution* 33: 381-395.
- ORR, H. A., 1995 A new linkage map of the *Drosophila pseudoobscura* X chromosome. *Dros. Inf. Serv.* 76: 127-128.
- PARVEZ, S. D., K. AKHTAR and R. K. SAKAI, 1985 Two new mutations and a linkage map of *Anopheles stephensi*. *J. Hered.* 76: 205-207.
- PERJE, A. M., 1955 Genetic and cytological studies of *Drosophila funebris*. Some sex-linked mutations and their standard order. *Acta Zool.* 36: 51-66.
- POWELL, J. R., 1997 *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University Press, New York.
- PURVIS, A., and A. RAMBAUT, 1995 Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analyzing comparative data. *Comput. Appl. Biosci.* 11: 247-251.
- ROBERTS, P. A., 1976 The genetics of chromosomal aberration, pp. 67-184 in *The Genetics and Biology of Drosophila*, Vol. 1a, edited by M. ASHBURNER and E. NOVITSKI. Academic Press, London.
- RUSO, C. A. M., N. TAKEZAKI and M. NEI, 1995 Molecular phylogeny and divergence times of *Drosophila* species. *Mol. Biol. Evol.* 12: 391-404.
- RUIZ, A., H. NAVEIRA and A. FONTDEVILA, 1984 La historia evolutiva de *Drosophila buzzatii*. IV. Aspectos citogenéticos de su polimorfismo cromosómico. *Genét. Ibér.* 36: 13-35.
- SAMPAIO, M. C., F. R. VARANDAS, S. C. VAZ and A. B. CARVALHO, 1996 Mapeamento do cromossomo X de *Drosophila mediopunctata*. *Rev. Brasil. Genét.* 19 (Suppl.): 268.
- SANTOS, M., 1986 The role of genetic selection in the establishment of inversion polymorphism in *Drosophila subobscura*. *Genetica* 69: 35-45.
- SCHAFER, D. J., D. K. FREDLINE, W. R. KNIBB, M. M. GREEN and J. S. F. BARKER, 1993 Genetics and linkage mapping of *Drosophila buzzatii*. *J. Hered.* 84: 188-194.
- SPASSKY, B., S. ZIMMERING and TH. DOBZHANSKY, 1950 Comparative genetics of *Drosophila prosaltans*. *Heredity* 4: 189-200.
- SPELICH, D., and P. PFRIEM, 1986 Chromosomal polymorphism in natural experimental populations, pp. 257-309 in *The Genetics and Biology of Drosophila*, Vol. 3e, edited by M. ASHBURNER, H. L. CARSON and J. N. THOMPSON, JR. Academic Press, London.
- SPIESS, E. B., 1950 Experimental populations of *Drosophila persimilis* from an altitudinal transect of the Sierra Nevada. *Evolution* 4: 14-33.
- SPIESS, E. B., 1965 A discovery and rediscovery of third chromosome arrangements in *Drosophila persimilis*. *Am. Nat.* 99: 423-425.
- SPURWAY, H., 1945 The genetics and cytology of *Drosophila subobscura*. I. Element A. Sex-linked mutants and their standard order. *J. Genet.* 46: 268-286.
- STURTEVANT, A. H., and G. W. BEADLE, 1936 The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics* 21: 544-604.
- TOBARI, Y. N., 1993 Linkage maps, pp. 49-51 in *Drosophila ananassae—Genetical and Biological Aspects*, edited by Y. N. TOBARI. Japan Scientific Societies Press, Tokyo.
- TRUE, J. R., J. M. MERCER and C. C. LAURIE, 1996 Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* 142: 507-523.
- WASSERMAN, M., 1962 Cytological studies of the repleta group of the genus *Drosophila*. IV. The hydei subgroup. Univ. Texas Publ. 6205: 73-84.
- WHITE, M. J. D., 1973 *Animal Cytology and Evolution*. Cambridge University Press, London.
- ZHENG, L., M. Q. BENEDICT, A. J. CORNELL, F. H. COLLINS and F. C. KAFATOS, 1996 An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* 143: 941-952.

Communicating editor: R. R. HUDSON

3. RESULTADOS Y DISCUSIÓN

3.1 Origen molecular de las inversiones naturales

El primero de los objetivos de esta tesis era caracterizar molecularmente los puntos de rotura de la inversión *2j* de *D. buzzatii*, para tratar de determinar cómo se había originado. En el momento de iniciar este trabajo la información disponible sobre los mecanismos moleculares implicados en la aparición de la gran cantidad de inversiones presentes en el género *Drosophila* era más bien escasa. El estudio de los puntos de rotura de la inversión *2j* ha proporcionado el primer caso en el que se ha podido desentrañar el mecanismo de generación de una inversión natural altamente exitosa. Además, ha permitido corroborar el papel activo que se suponía que desempeñaban los elementos transponibles en el origen de las inversiones en la naturaleza. Finalmente, ha puesto de manifiesto que los puntos de rotura de esta inversión son regiones de inestabilidad genética, presumiblemente causada por la inserción de un nuevo elemento transponible de tipo *Foldback*. La relevancia de estos resultados no se limita al género *Drosophila*, sino que contribuyen a incrementar el conocimiento sobre cómo surgen las reorganizaciones cromosómicas en los eucariotas. Por otra parte, constituyen un buen ejemplo de la importancia que podría tener la actividad de los elementos transponibles en la evolución en general, y en la reestructuración del genoma en particular.

3.1.1 Clonación de los puntos de rotura de la inversión *2j*

El estudio de los puntos de rotura de las inversiones a nivel molecular se ha visto normalmente dificultado por la ausencia de marcadores cercanos que permitan acceder a ellos. En 1994, Wesley y Eanes describieron un método novedoso para clonar puntos de rotura de inversiones, que hizo posible aislar el primero de los puntos de rotura de la inversión *In(3L)Payne* de *D. melanogaster* en tan sólo cuatro semanas. El método se basaba en la obtención de sondas de las regiones de los puntos de rotura

de los cromosomas con la inversión mediante microdissección. Posteriormente, utilizando las sondas obtenidas, al rastrear una genoteca de una línea con la ordenación *standard* se podrían identificar los clones que contienen los puntos de rotura como aquellos que hibridan con ambas sondas.

No obstante, la estrategia seguida para la clonación de los puntos de rotura de la inversión *2j* de *D. buzzatii* difiere considerablemente de la metodología puesta a punto por Wesley y Eanes (1994). En este caso, se utilizaron las técnicas del aterrizaje cromosómico y la clonación posicional, partiendo de la información cartográfica previa generada en nuestro grupo. El aterrizaje cromosómico consiste en hacer uso de los marcadores moleculares disponibles en una especie bien estudiada para poder tener acceso a una determinada región homóloga de una especie de interés (Tanksley *et al.* 1995). Esta técnica se basa en la existencia de regiones cromosómicas conservadas entre especies próximas, que hacen posible trasladar la información cartográfica de una a otra. A lo largo de varios años, se han venido realizando amplios estudios de cartografía comparada mediante hibridación *in situ* de sondas provenientes principalmente del brazo cromosómico *3R* de *D. melanogaster* sobre los cromosomas politénicos de *D. buzzatii*. Estos estudios han permitido construir un mapa bastante detallado del cromosoma 2 de *D. buzzatii* y definir una serie de regiones conservadas entre ambas especies (Ranz *et al.* 1997, 1999, 2000). La información obtenida ha puesto así al alcance la caracterización molecular de casi cualquier región que se desee del cromosoma 2 de *D. buzzatii*, donde se localizan la mayoría de inversiones de esta especie, partiendo de los marcadores acumulados en *D. melanogaster*.

El punto de partida para la clonación de los puntos de rotura de la inversión *2j*, fue el descubrimiento de una región cromosómica presumiblemente conservada entre *D. buzzatii* y *D. melanogaster*, que incluía el punto de rotura proximal de la inversión (Ranz *et al.* 1997). Esta región estaba delimitada por los genes *nAcR α -96Aa* (*ALS*) y *Pp1 α -96A*, situados fuera y dentro del segmento invertido, respectivamente. Ambos genes hibridaban en la misma banda del cromosoma 2 de *D. buzzatii* y parecían estar muy cercanos, pero la distancia real entre ellos era desconocida. Por tanto, el primer

paso fue tratar de estimar la distancia entre los dos genes e intentar encontrar otros marcadores situados entre ellos que pudiesen estar aún más próximos al punto de rotura. Con este fin, se elaboró un mapa físico preliminar de la región cromosómica 95A-96A de *D. melanogaster*, donde se encontraba el segmento *nAcR α -96Aa* y *Pp1 α -96A*, utilizando 13 genes y 28 cósmidos pertenecientes a esta región (Ranz *et al.* 1999). Gracias a este análisis, se pudo encontrar un nuevo marcador más cercano al punto de rotura proximal de la inversión 2j, *nAcR β -96A* (SBD) y se pudo localizar el punto de rotura con mayor precisión entre los genes *nAcR β -96A* y *Pp1 α -96A*. Además, la elaboración de un mapa de restricción de los cósmidos 148H6 y 172B8, que incluían ambos genes, permitió estimar la distancia entre ellos en tan sólo 4 kb (Figura 3.1). A partir de los datos de la secuenciación del genoma de *D. melanogaster* (Adams *et al.* 2000), se ha podido determinar *a posteriori* que la distancia exacta entre las regiones codificadoras de los genes *nAcR β -96A* y *Pp1 α -96A* es de 3.689 bp, muy inferior a las 96,5 kb que en realidad separan a los dos marcadores iniciales, *nAcR α -96Aa* y *Pp1 α -96A*. De esta forma, el descubrimiento de dos marcadores situados muy cerca del punto de rotura proximal de la inversión 2j hizo posible su clonación mediante métodos convencionales (Cáceres *et al.* 1999a). La distancia entre los genes *nAcR β -96A* y *Pp1 α -96A* en la ordenación 2st de *D. buzzatii* resultó ser muy similar a la encontrada en *D. melanogaster* y, paradójicamente, uno de ellos, *Pp1 α -96A*, resultó estar situado a sólo unos cientos de pares de bases del punto de rotura.

La estrategia seguida para aislar los puntos de rotura de la inversión 2j era en principio más lenta y costosa que la utilizada en el caso de la inversión *In(3L)Payne*. No obstante, presentaba diversas ventajas y a la postre ha demostrado ser una de las claves del éxito del proceso. En primer lugar, la información cartográfica acumulada en el cromosoma 2 ha proporcionado importantes estimas de la tasa de evolución cromosómica en el género *Drosophila* (Ranz *et al.* 1997, 1999, 2000) y está sirviendo para clonar los puntos de rotura de otras inversiones. Además, aparte de la dificultad técnica de la microdissección, el método empleado por Wesley y Eanes (1994) tiene graves inconvenientes cuando en ambos puntos de rotura se encuentran copias de un

1 Kb

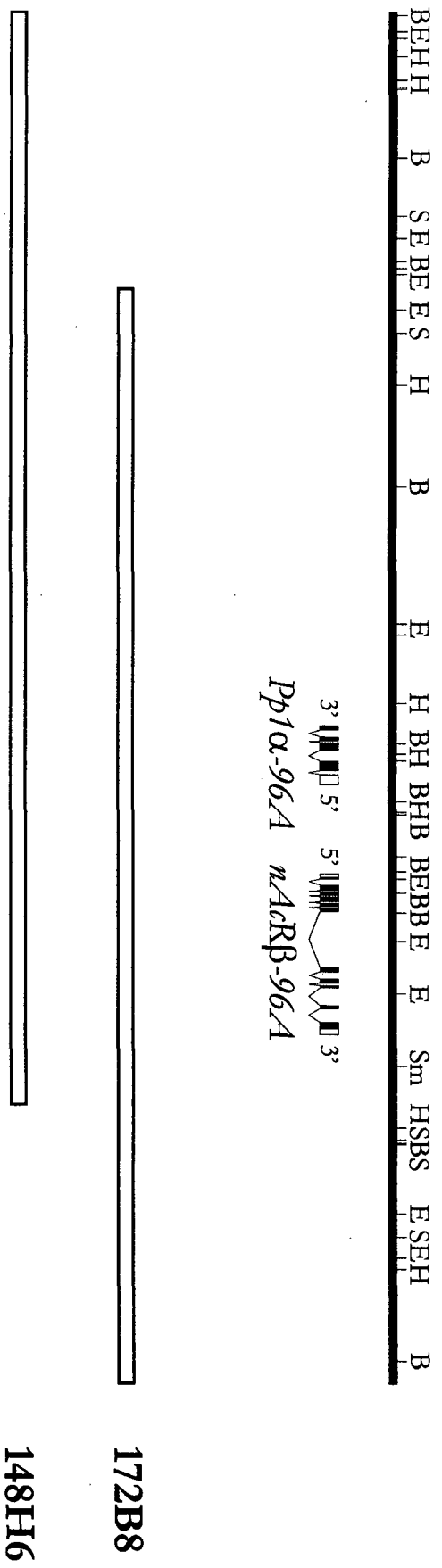


Figura 3.1. Mapa de restricción de la región comprendida por los cósmidos 148H6 y 172B8 de *D. melanogaster*, mostrando la localización de los genes *Pp1α-964* y *nAcRB-964* de acuerdo con los datos de la secuencia del genoma (Adams *et al.* 2000). B, *Bam* HI; E, *Eco* RI; H, *Hind* III; S, *Sal* I; Sm, *Sma* I.

mismo elemento transponible, algo muy probable si éste está implicado en el origen de la inversión. La presencia de un elemento transponible haría que al rastrear por duplicado la genoteca de la ordenación *standard* con las sondas de uno y otro punto de rotura se obtuviesen muchos clones positivos y la mayoría de ellos fuesen comunes. Por último, la construcción y el rastreo de genotecas ha resultado ser una técnica más segura que la PCR inversa, que se había utilizado en otros estudios para el aislamiento de los puntos de rotura de la inversión en la ordenación invertida (Wesley y Eanes 1994; Andolfatto *et al.* 1999). La PCR inversa tiene limitaciones en cuanto al tamaño del fragmento que se puede amplificar (Ochman *et al.* 1988). Esta limitación, junto con la presencia de una secuencia repetitiva, habría impedido completar la clonación de los puntos de rotura de la inversión *In(2L)t* de *D. melanogaster* (P. Andolfatto, comunicación personal).

3.1.2 Secuencia de los puntos de rotura de la inversión *2j*

Durante el proceso de clonación de los puntos de rotura de la inversión *2j* de *D. buzzatii* se detectó la existencia de grandes bloques de DNA repetitivo en los fragmentos clonados provenientes de los cromosomas con la inversión, que sin embargo no estaban presentes en los cromosomas con la ordenación no invertida (Cáceres *et al.* 1999a). El alineamiento de la secuencia nucleotídica de estos fragmentos en una línea *2st* y una línea *2j* reveló que el DNA repetitivo estaba situado exactamente en los puntos de rotura de la inversión. Además, el análisis de las secuencias extras dio a entender que eran el resultado de la inserción, tanto en el punto de rotura proximal como en el distal, de dos copias de lo que parecía ser un nuevo elemento transponible, no descrito previamente, y que fue denominado *Galileo*. Este elemento transponible localizado en los puntos de rotura era el presunto responsable de la generación de la inversión.

En un principio, la secuencia completa de las inserciones presentes en los puntos de rotura se obtuvo en una única línea con la inversión $2j$. Estas inserciones, en especial la del punto de rotura proximal, resultaron tener una estructura extremadamente compleja y difícil de interpretar (Cáceres *et al.* 1999a). Un estudio completo posterior de los puntos de rotura de la inversión en 30 líneas $2j$ de orígenes geográficos diversos, demostró que las inserciones estaban presentes en todas ellas. No obstante, puso de manifiesto la existencia de una considerable variación tanto en la longitud como en la estructura de las inserciones de las diferentes líneas. En total, se encontraron seis variantes estructurales en el punto de rotura distal y diez variantes estructurales en el punto de rotura proximal en las 30 líneas caracterizadas (Cáceres *et al.* 2000). Además, la información obtenida permitió entender mejor el origen de la inserción del punto de rotura proximal de la línea secuenciada inicialmente.

Las secuencias que se insertaron en cada punto de rotura originalmente son desconocidas. Sin embargo, la comparación de las diferentes variantes encontradas en los puntos de rotura ha permitido inferir la estructura de las teóricas inserciones ancestrales. De acuerdo con el principio de parsimonia, que supone el mínimo número posible de cambios en una secuencia evolutiva, se puede postular que las variantes estructurales más sencillas serían más ancestrales y estarían más cercanas a las inserciones originales. Así, en el punto de rotura distal, la inserción ancestral estaría formada simplemente por una copia del elemento *Galileo* de 0,4 kb, con repeticiones terminales invertidas de 106 bp. A 256 bp de esta inserción, se encontraría otra de 841 bp perteneciente a un elemento denominado *ISBul*. Por su parte, en el punto de rotura proximal, la inserción ancestral tendría un tamaño aproximado de 3,5 kb y estaría formada por hasta cuatro elementos transponibles diferentes: un elemento *Galileo* de 1.589 bp y otros tres elementos insertados en distintas posiciones dentro de él, designados como *Kepler*, *BuT1* y *BuT3*.

A partir de las inserciones ancestrales se habría generado toda la diversidad observada mediante la acumulación gradual de una serie de cambios. Las diferentes variantes estructurales poseen las mismas inserciones de *Galileo* en los puntos de

rotura y comparten la mayor parte de las secuencias insertadas, pero divergen entre sí principalmente por la inserción y deleción de fragmentos de DNA. De esta forma, en el punto de rotura distal se habrían producido otras ocho inserciones de siete elementos transponibles diferentes, incluyendo un nuevo elemento *Galileo*, y cuatro deleciones. En el punto de rotura proximal habrían ocurrido también ocho inserciones de diferentes elementos transponibles, incluyendo otra copia de *Galileo*, y ocho deleciones. Además, en algunas de las líneas habría tenido lugar la inversión de un pequeño segmento interno de la secuencia insertada en el punto de rotura. De esta forma, se habrían generado estructuras tan complejas como la de la línea jq⁷-2 en el punto de rotura proximal, con un tamaño de 5,8 kb y seis elementos transponibles insertados sobre el elemento *Galileo* original (Cáceres *et al.* 2000).

En total, la secuenciación de los puntos de rotura de la inversión *2j* ha permitido caracterizar diez elementos transponibles no descritos previamente y ha incrementado considerablemente nuestro conocimiento sobre los elementos transponibles de *D. buzzatii* (Cáceres *et al.* 2000). Muchos de ellos presentan repeticiones terminales invertidas y parecen corresponder a elementos transponibles de Clase II. Dentro de esta clase se incluyen los elementos *Galileo*, *Kepler* y *Newton*, que están relacionados entre sí y poseen en su mayoría repeticiones terminales invertidas largas, los elementos *BuT1* y *BuT2*, que tienen repeticiones terminales invertidas similares de 12-15 bp y muestran homología con el elemento *Gandalf* de *D. koepferae* (Marín y Fontdevila 1995), y los elementos *BuT3* y *BuT4*, que tienen repeticiones terminales invertidas de 23-24 bp y se relacionan con el elemento *Hopper* de *Bactrocera dorsalis* (Handler y Gomez 1997). Además, se han encontrado también elementos que no parecen pertenecer a ninguno de los tipos descritos previamente. Éstos serían *BuT5*, que posee repeticiones terminales invertidas de sólo 3 bp seguidas de repeticiones subterminales imperfectas de 17 bp, y los elementos *ISBu1* e *ISBu2*, que no presentan repeticiones en sus extremos y podrían corresponder a una nueva clase de elementos móviles (Hagemann *et al.* 1998; Amador y Juan 1999; Vivas *et al.* 1999).

3.1.3 Estructura del elemento transponible *Galileo*

Inicialmente, la estructura del elemento *Galileo* se intentó determinar a partir de las copias insertadas en los puntos de rotura de la inversión $2j$ en todas las líneas, *Galileo-1* y *Galileo-2*. No obstante, el análisis de la secuencia de estas dos primeras copias proporcionó poca información acerca de la organización real del elemento (Cáceres *et al.* 1999a). Afortunadamente, la caracterización de los puntos de rotura de la inversión en un mayor número de líneas $2j$ hizo posible descubrir otras dos copias de *Galileo*, *Galileo-3* y *Galileo-4*, insertadas de forma más reciente en estas regiones (Cáceres *et al.* 2000). Además, paralelamente, se amplió el estudio a más elementos *Galileo* originarios de otros lugares del genoma y se secuenciaron por completo varios de ellos. *Galileo-5* y *Galileo-6* fueron obtenidos rastreando una genoteca de *D. buzzatii* con una sonda de *Galileo-2*, mientras que *Galileo-7* y *Galileo-8* se amplificaron por PCR utilizando oligonucleótidos complementarios a los extremos del elemento. A partir de la información de las ocho copias de *Galileo* disponibles, se ha podido definir de forma más o menos precisa la estructura de este elemento. Asimismo, nos ha permitido clasificarlo dentro de uno de los tipos principales de elementos transponibles de Clase II, los elementos *Foldback*, que se caracterizan por la presencia de repeticiones terminales invertidas de gran tamaño (Finnegan 1989; Capy *et al.* 1998).

Galileo-3 y *Galileo-4* son las copias de *Galileo* más grandes y corresponden a inserciones eucromáticas más recientes que las insertadas originalmente en los puntos de rotura. Por tanto, son las que se han utilizado para la descripción de este transposón. La principal característica de ambas copias es la presencia de repeticiones invertidas largas en sus extremos, de entre 683 y 917 bp (Figura 3.2). Igual que en muchos de los elementos de tipo *Foldback* descritos (Bingham y Zachar 1989; Hoffman-Liebermann *et al.* 1989; Hankeln y Schmidt 1990; Rebatchouk y Narita 1997), la longitud de estas repeticiones terminales invertidas varía entre las diferentes copias de *Galileo*. Además, la longitud de las dos repeticiones terminales invertidas varía incluso dentro de un

mismo elemento, como en el caso de *Galileo-4*. Estas repeticiones tienen una organización compleja. Aparentemente, presentan una región terminal de 479 bp conservada, seguida de una región interna más variable, formada por una serie de repeticiones imperfectas de una secuencia de 136 bp dispuestas en tándem (Figura 3.3). Esta organización de los extremos es típica de los elementos *Foldback* (Rebatchouk y Narita 1997). Por ejemplo, los elementos *FB* de *D. melanogaster* también contienen una región terminal más o menos constante y múltiples repeticiones en tándem de una secuencia de 155 bp, que a su vez está formada por cinco unidades diferentes de 31 bp (Potter 1982).

Las repeticiones terminales invertidas parecen formar la mayor parte del elemento *Galileo*. El segmento comprendido entre ellas suele ser pequeño y es heterogéneo entre diferentes copias de *Galileo*. Además, puede ser debido a que la repetición de un extremo es más larga que la otra. En el caso de *Galileo-3*, la mayoría de la región no repetida entre los dos extremos corresponde a secuencias que están formando parte de las repeticiones terminales invertidas de *Galileo-4* y únicamente 483 bp corresponden a secuencias que no se han encontrado en las repeticiones (Figure 3.2). En *Galileo-4*, la secuencia comprendida entre las dos repeticiones terminales invertidas es de sólo 249 bp. No obstante, esta estructura es similar a la de muchos de los elementos de tipo *Foldback* descritos, que consisten en dos copias de las repeticiones terminales invertidas sin prácticamente nada entre medio (Truett *et al.* 1981; Yuan *et al.* 1991; Rebatchouk y Narita 1997).

El resto de copias de *Galileo* analizadas comparten la mayor parte de su secuencia y su estructura con *Galileo-3* y *Galileo-4*, aunque ninguna de ellas es tan completa como las anteriores. *Galileo-5*, *Galileo-6*, *Galileo-7* y *Galileo-8* no poseen repeticiones terminales invertidas y contienen en su mayoría sólo uno de los extremos del elemento. Además, algunas de estas copias presentan deleciones internas respecto a *Galileo-3* y *Galileo-4* (Figura 3.2). Por tanto, serían probablemente copias defectivas de *Galileo*, provenientes quizás de la heterocromatina centromérica, como parece ser el caso de *Galileo-5* (J. M. Ranz, comunicación personal).

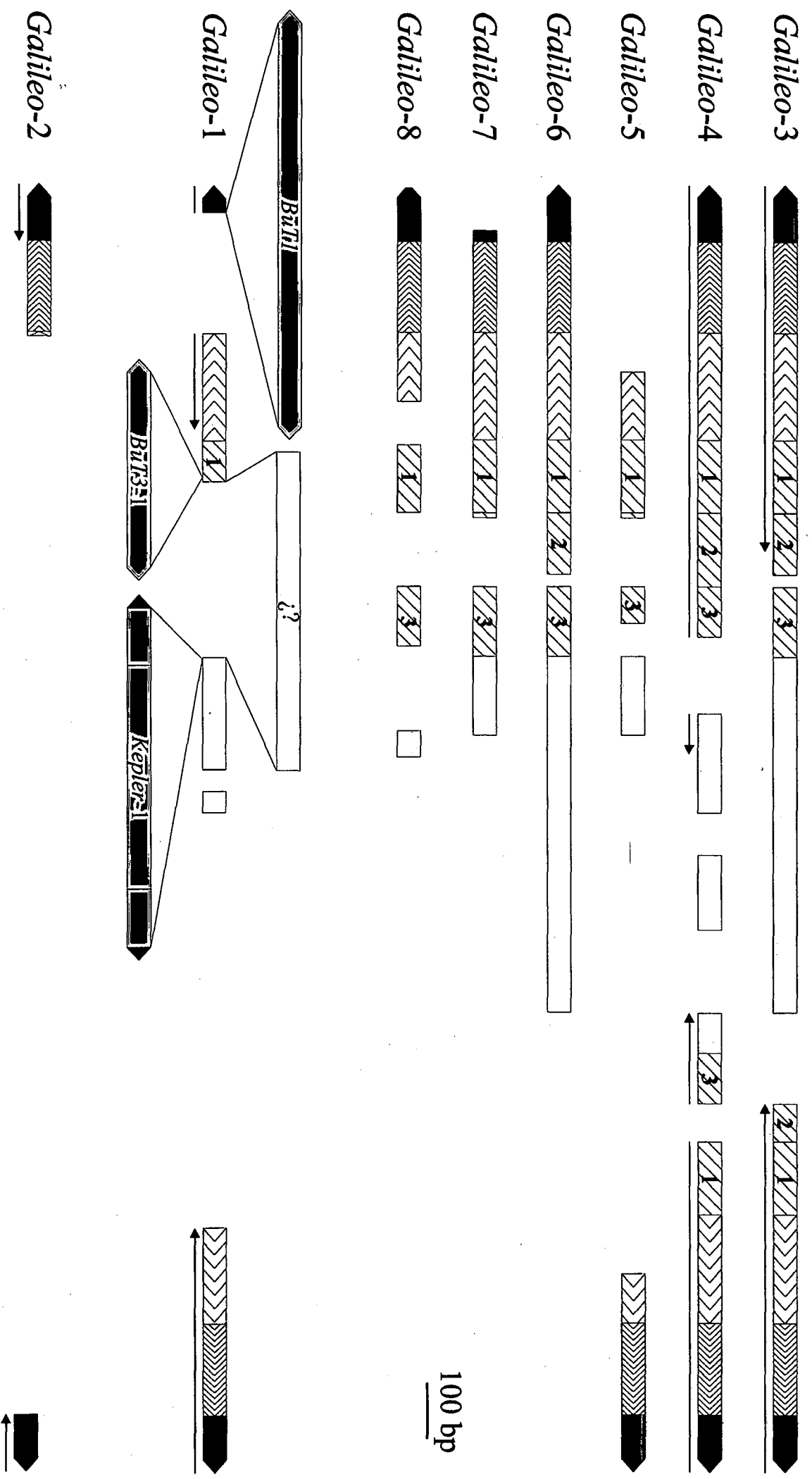


Figura 3.2. Diagrama con la estructura de los diversos elementos *Galileo* secuenciados, en el que las secuencias homólogas se representan alineadas. Las flechas indican las repeticiones terminales invertidas de cada elemento. Los segmentos numerados corresponden a las diferentes copias de la secuencia repetida de 136 bp (Figura 3.3). Las secuencias presentes en *Galileo-1*, pero no en ninguna de las otras copias de *Galileo*, se representan encima y debajo de la estructura de éste.

Dentro de las copias de *Galileo* insertadas en los puntos de rotura de la inversión 2j, *Galileo-1* y *Galileo-2*, la del punto de rotura proximal es la que más se diferencia respecto a la estructura teórica del elemento. No obstante, la presencia de diversos fragmentos con homología a los otros elementos *Galileo*, permiten postular cómo se ha originado su compleja estructura (Figura 3.2). En primer lugar, contiene al menos otros tres elementos transponibles diferentes en su interior: *Kepler-1*, *BuT1* y *BuT3-1*. Además, parece haber sufrido una deleción de 237 bp en uno de sus extremos, que habría eliminado una parte de la repetición terminal izquierda, la supuesta duplicación de 8 bp generada en la inserción de *BuT1* y el último par de bases de este elemento. Finalmente, no posee ninguna copia completa de la secuencia repetida de 136 bp y, en su lugar, tendría una secuencia de 589 bp de origen desconocido, que no se encuentra en ninguno de los otros elementos *Galileo*. Por su parte, la copia del punto de rotura distal tiene una estructura mucho más sencilla. Está formada tan sólo por 286 bp y 106 bp de los extremos de las dos repeticiones terminales invertidas y podría ser simplemente el producto de una deleción interna de todo el resto del elemento (Figura 3.2).

3.1.4 Mecanismo de transposición del elemento *Galileo*

La naturaleza transponible del elemento *Galileo* y su capacidad de movilizarse e insertarse en nuevos lugares del genoma está fuera de duda. Cuatro de las copias de *Galileo* caracterizadas son claramente inserciones *de novo*, presentes sólo en algunas líneas de *D. buzzatii* y flanqueadas por duplicaciones directas del sitio de inserción (Cáceres *et al.* 1999a, 2000). Además, la hibridación *in situ* de una sonda de *Galileo* sobre los cromosomas politénicos de diferentes líneas de *D. buzzatii* ha mostrado la existencia de una gran variación, tanto en el número de copias del elemento como en la posición de éstas (J. M. Ranz, comunicación personal). Sin embargo, todavía no está del todo claro cómo se transponen exactamente los elementos *Galileo*.

Gal-3 1 CATACACACACACATTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CAC-ATTAGAACCCCGACGACGTTGANTGTTCCGTTGGTGGGTAAC GATGTGACCTTG
2 CATACACACACACACATTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCCGCCCTGTAGACACACATTTAGAACCCCGACGACTGA----TCGTTG-----TG
3 TACACATACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCA TGTAG CAATTCATAATTCGGAGCGAATTCGTTGTT-----GGCTAGCCAAATGTACAGCATG

Gal-4 1 CATACACACACACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CACTATTAAGAAATCCGAGCAGTTGANTGTTCCGTTGGTGGGTAAC GATGTGACCTTG
2 CATACACACACACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CACTATTAAGAAATCCGAGCAGTTGANTGTTCCGTTGGTGGGTAAC GATGTGACCTTG
3 CATACACACAAACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCA TGTAG CAATTCATAATTT

Gal-5 1 CATACACACACACACATTTACACTGCCCTACATAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CACTATTA-ACATAAGCAGTTGANTGTTG-----GGTTGGGTAACATAATGATGAGTTTG
2 AATAGCGACACA
3 CCTACACATA-ACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCG-----AATACATAATG TGCA TA

Gal-6 1 CATACACACACACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CAC-ATTAGAACCCCGACGACGTTGANTGTTCCGTTGGTGGGTAAC GATGTGACCTTG
2 CATACACACACACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCCGCCCTGTAGACACACATTTAGAACCCCGACGACTGA---TCGTTG-----TG
3 TACACATACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCA TGTAG CAATTCATAATTCGGAGCCAAATTCGTTGTT-----GGCTAGCCAAATGTACAGCATG

Gal-7 1 CATACACACACACACATTTGCACTGCCCTTACATAA GHTAGAAATTTAAGTGCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CACTAATA-ACATAAGCAGTTTAAATGTTG-----GGTTGGGTAACCAATGTGAGTTTG
2 CATAGCGACACA
3 CATACAAATACACATTTACACTTCCCTGCAACATAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCA TATAG CAATTTAATTTCCGACGCAATTCGTTGTT-----TGTATGCTAATGTGACAGCATG

Gal-8 1 AAAAAATTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG CACTATTAAGAACCCCGACGACTGANTGTTCCGTTGGGTAACGCTGCTGAGCCTTG
3 CATACACATACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCCA TGTAG CAATTCATAATTTCCGACGCAATTCGTTGTT-----GGG

Gal-1 1 CATACACACACACACATTTACACTTCCCTGGCCTAATATGATGATGAGGCAATTAATGTCCTGTGTGCAAAATATGCTTGTCTGCC TGTAG

Figura 3.3. Secuencia nucleotídica de las diferentes copias de la repetición de 136 bp presentes en los ocho elementos *Galileo* caracterizados. Los nucleótidos que varían entre las diversas copias y los diversos elementos se representan en diferentes colores. En azul se indican los cambios únicos. En rojo se indican cambios que parecen caracterizar a las copias de tipo 3. Los guiones representan deleciones internas.

La falta de información sobre el mecanismo de transposición y el desconocimiento de las proteínas implicadas en el proceso son habituales en los elementos de tipo *Foldback* descritos (Capy *et al.* 1998). La práctica totalidad de estos elementos no parecen poseer capacidad codificadora (Hoffman-Liebermann *et al.* 1989; Hankeln y Schmidt 1990; Yuan *et al.* 1991; Rebatchouk y Narita 1997). Únicamente en *D. melanogaster*, la especie más y mejor estudiada, se han encontrado copias de elementos *FB* que contienen en su región central marcos de lectura abiertos que podrían codificar las proteínas encargadas de llevar a cabo las funciones necesarias para la transposición (Templeton y Potter 1989; Harden y Ashburner 1990). No obstante, estas proteínas no se parecen a las otras transposasas conocidas. Además, los elementos con capacidad codificadora son escasos y constituyen tan sólo un 10% del total de elementos *FB*. Por tanto, puede ser que, en las diferentes familias de elementos *Foldback*, las copias funcionales capaces de expresar transposasa estén en una proporción baja respecto a una gran mayoría de copias defectivas y sean así difíciles de detectar.

El caso de *Galileo* no sería en consecuencia muy distinto al del resto de elementos *Foldback*. La posesión de repeticiones terminales invertidas sugiere que la transposición tiene lugar por un mecanismo de corte y empalme similar al de los otros elementos de Clase II (Plasterk 1995). Además, la presencia de una secuencia formada por dos repeticiones invertidas de 9 bp correspondientes a los extremos del elemento, en la región del punto de rotura distal de varios de los cromosomas con la ordenación *2st* (Figura 3.4), recuerda a la huella que dejan algunos elementos de Clase II tras su excisión (Engels *et al.* 1990; Plasterk 1991). Esta secuencia se podría haber generado durante el proceso de reparación de una rotura de doble cadena en la molécula de DNA provocada por la movilización del elemento e indicaría que la transposición transcurre por un mecanismo no replicativo (Plasterk 1995).

Por otra parte, las secuencias nucleotídicas de las diversas copias de *Galileo* encontradas hasta el momento no muestran homología con las transposasas de los otros elementos transponibles de Clase II. Además, el segmento central de la mayoría

de las copias de *Galileo* secuenciadas es pequeño, tiene un elevado contenido en AT (71%) y no parece codificar proteína alguna. No obstante, una posibilidad es que *Galileo* sea similar a algunos transposones bacterianos, como Tn5 o Tn10. Estos elementos presentan en los extremos dos secuencias de inserción de 1,5 kb y 1,3 kb, respectivamente, dispuestas en orientación invertida a ambos lados de una región central que contiene un gen de resistencia a antibióticos, y las secuencias de inserción se encargan de codificar las proteínas necesarias para la transposición (Berg 1989; Kleckner 1989). Son varias las evidencias que sugieren que la transposasa de *Galileo* podría estar codificada dentro de las repeticiones terminales invertidas. En primer lugar, los 600 nucleótidos terminales de las repeticiones invertidas de *Galileo* están muy conservados entre las diferentes copias. En segundo lugar, esta secuencia contiene un marco de lectura abierto con capacidad para codificar hasta 168 aminoácidos. Aunque las transposasas eucarióticas más pequeñas son considerablemente mayores, 345 aminoácidos en el caso de la transposasa del elemento *mariner* de *D. mauritiana* (Hartl 1989), las proteínas codificadas por algunas secuencias de inserción bacterianas pueden tener tamaños equivalentes (Galas y Chandler 1989). Por último, al comparar los elementos relacionados *Galileo*, *Kepler* y *Newton*, esta zona sería la única que presenta homología de secuencia entre los tres elementos y su conservación sugiere que podría tener una importancia funcional.

El hecho de que los 40 bp terminales de *Galileo*, *Kepler* y *Newton* sean prácticamente idénticos, plantea la posibilidad de que todos ellos puedan ser reconocidos y movilizados por una misma transposasa común. Algo similar ocurre en el elemento *Tc7* de *Caenorhabditis elegans*, que tiene repeticiones invertidas de 345 bp separadas por una secuencia de sólo 230 bp, y se transpone mediante la transposasa codificada por el elemento *Tc1*, con el que comparte 36 de los últimos 38 nucleótidos de cada extremo (Rezsohazy *et al.* 1997). Una de las principales características del mecanismo de acción de la transposasa es el tamaño de las duplicaciones del sitio diana generadas durante la inserción del elemento. El tamaño de estas duplicaciones viene determinado por la separación entre los cortes escalonados que realiza la transposasa en cada cadena del DNA y es una propiedad intrínseca de ésta (Plasterk

1995). En nuestro caso, se ha podido observar que los elementos *Galileo*, *Kepler* y *Newton* generan todos ellos duplicaciones de 7 bp en la inserción (Cáceres *et al.* 2000).

				←	→		
<i>D. buzzatii</i>	st-1	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGGTTAGTG	<u>TTAGTAG</u>	TAAA
	st-2	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGGTTAGTG	<u>TTAGTAG</u>	TAAA
	st-3	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
	st-4	GTCA	<u>GTAGTAGTACT</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
	st-5	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
	st-6	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
	st-7	GTCA	<u>GTAGTAG</u>	CACTAACCA	TGGTTAGTG	<u>TTAGTAG</u>	TAAA
	st-8	GTCA	<u>GTA-TAG</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
	st-9	GTCA	<u>GTAGTAGTACT</u>	CACTAACCA	TGG-----	<u>TTAGTAG</u>	TAAA
<i>D. martensis</i>	Ma-4	GTCA	-TAGCAA	-----	-----	-----	TAAA

Figura 3.4. Secuencia de la región del punto de rotura distal (*AB*) en diferentes líneas de *D. buzzatii* con la ordenación *2st*. En negrita se indican las secuencias con homología a los extremos de *Galileo*. Las secuencias supuestamente duplicadas durante la inserción se representan subrayadas. Como comparación se muestra también la secuencia de esta región en una línea de *D. martensis*.

Además, la transposasa podría tener también un papel fundamental en la elección del sitio de inserción del elemento (Craig 1997). Se ha visto que algunos elementos transponibles muestran una marcada preferencia por insertarse en sitios diana que comparten una secuencia consenso común. Un ejemplo bien estudiado es el del elemento *P* (O'Hare y Rubin 1983). El alineamiento de los sitios de inserción de las diferentes copias de *Galileo*, *Kepler* y *Newton* sugiere que los tres elementos muestran preferencia por una misma secuencia consenso de 7 bp (Figura 3.5). Esta secuencia parece ser palindrómica y recuerda a los extremos del elemento (CAC....GTG; Figura 3.4), lo que permite hipotetizar tal vez algún tipo de función de las últimas bases en el reconocimiento y el corte por parte de la transposasa. No obstante, serían necesarios más estudios para poder determinar realmente cómo se transponen estos elementos.

<i>Galileo-1</i>	GTAGTAG
<i>Galileo-2</i>	TTTGTAT
<i>Galileo-3</i>	GAAGAAC
<i>Galileo-4</i>	GTGATAC
<i>Galileo-5</i>	GTAATAT
<i>Galileo-6</i>	GTAGTAC
<i>Galileo-9</i>	GTGATAC
<i>Kepler-1</i>	CATATAT
<i>Kepler-2</i>	GTAGTAT
<i>Kepler-3</i>	GTAGTAG
<i>Kepler-4</i>	CTAAAAC
<i>Newton-1</i>	GTAGTAT
<i>Newton-2</i>	GTGATAC
Consenso	G₁₀T₁₁a₈g₇T₁₁A₁₃C₆

Figura 3.5. Secuencias adyacentes a las diferentes copias de *Galileo*, *Kepler* y *Newton* de las que se han caracterizado los extremos. Se han considerado las dos orientaciones de cada secuencia y se ha representado la que se ajusta mejor al consenso (O'Hare y Rubin 1983).

3.1.5 Mecanismo de generación de la inversión $2j$

La información obtenida a partir de la secuenciación de los puntos de rotura de la inversión $2j$ ha permitido proponer un modelo plausible para la generación de esta inversión (Cáceres *et al.* 1999a). La inversión habría surgido por recombinación ectópica entre las secuencias homólogas de las dos copias de *Galileo* dispuestas en orientación contraria, mediante un mecanismo similar al descrito en la Figura 1.3. Este modelo había sido propuesto previamente para explicar la aparición de inversiones causadas por elementos transponibles en poblaciones de laboratorio de *Drosophila* (Lim y Simmons 1994) y su validez se había demostrado en algunos casos (Schneuwly *et al.* 1987; Eggleston *et al.* 1996). El primer paso habría sido la inserción de las dos copias de *Galileo* en las regiones de los puntos de rotura de un cromosoma

2st. Estas inserciones de *Galileo* no están presentes en ninguno de los cromosomas 2st estudiados, pero en la región del punto de rotura distal de algunos de ellos se ha encontrado una secuencia formada por dos repeticiones invertidas de 9 bp correspondientes a los extremos del elemento (Figura 3.4), que podría ser la huella dejada tras su excisión (Engels *et al.* 1990; Plasterk 1991). Posteriormente, una vez insertados los elementos, el siguiente paso habría sido el apareamiento intracromosómico entre las secuencias homólogas y el evento de recombinación que habría invertido todo el segmento cromosómico comprendido entre ellas.

Son varias las evidencias que apuntan a que realmente se produjo una recombinación entre las dos copias de *Galileo* insertadas en los puntos de rotura de la inversión. La más importante es la de la disposición de las duplicaciones del sitio diana generadas en el momento de la inserción. En los cromosomas con la inversión, las duplicaciones se encuentran intercambiadas entre ambos elementos y cada uno está flanqueado por dos secuencias de 7 bp diferentes. Además, una de las secuencias de cada par se halla en orientación invertida (Cáceres *et al.* 1999a). Esta misma disposición de las duplicaciones se ha observado en el caso de la inversión generada por el elemento *Doc* responsable de la mutación *Antp*^{73b} (Schneuwly *et al.* 1987). Por otra parte, otro indicio de la recombinación proviene del análisis de la secuencia de las repeticiones terminales invertidas de los elementos *Galileo*. La comparación de estas secuencias ha permitido inferir la existencia de dos grupos de elementos *Galileo*, claramente diferenciados por toda una serie de cambios nucleotídicos (Figura 3.6). *Galileo-1*, *Galileo-5* y *Galileo-7* pertenecerían a un grupo, mientras que *Galileo-3*, *Galileo-4*, *Galileo-6* y *Galileo-8* pertenecerían a otro. *Galileo-2*, la copia insertada en el punto de rotura distal, sería la única que presenta aparentemente una secuencia quimérica, con una parte idéntica a cada tipo. Esta inusual estructura de *Galileo-2* podría ser el resultado de un proceso no recíproco de conversión génica, que habría acompañado al entrecruzamiento entre *Galileo-1* y *Galileo-2* durante la recombinación ectópica, en el que una parte de este último habría pasado de un tipo a otro. Según el modelo de Holliday, el desplazamiento del punto de intersección entre las cadenas apareadas durante la recombinación genera una región de heterodúplex,

que puede resolverse en forma de conversión génica (Lewin 1997). De acuerdo con esta hipótesis, el inicio del proceso de intercambio genético en los puntos de rotura se puede localizar dentro de la secuencia de 8 bp situada entre las posiciones 101 y 110 de *Galileo* (Figura 3.6).

La recombinación entre secuencias situadas en la misma posición de los cromosomas (recombinación alélica) y en diferentes posiciones de los cromosomas (recombinación ectópica) poseen muchas características en común y se ha postulado que tendrían lugar por un mecanismo similar (Virgin y Bailey 1998). En ambos casos se ha de producir un apareamiento entre las cadenas de las moléculas de DNA que recombinan y es necesario que exista cierto nivel de homología de secuencia para que el proceso tenga lugar de forma eficiente. Por tanto, una posible objeción al modelo anteriormente propuesto para la generación de la inversión *2j*, es que la secuencia homóloga que comparten las dos copias de *Galileo* de los puntos de rotura y que tuvo que intervenir en el proceso de recombinación ectópica está restringida a sólo 274 bp de los extremos (Cáceres *et al.* 1999a). Como se desconoce la estructura exacta de los elementos implicados en la generación de la inversión, se podría postular que la región de homología era originalmente mayor y se perdió parte después, quizás en la delección que dio lugar a *Galileo-2*. De todas formas, el pequeño segmento homólogo encontrado parece ser más que suficiente para que se produzca la recombinación ectópica. Dentro de los eucariotas, la longitud de homología mínima requerida para la recombinación se ha estimado en 63-89 bp en levaduras (Sugawara y Haber 1992) y en 200 bp en células de mamíferos (Liskay *et al.* 1987). En *Drosophila*, Nassif y Engels (1993) han observado que sólo hacen falta regiones de homología muy pequeñas, inferiores incluso a 115 bp, para que pueda darse la recombinación ectópica. Así, en *D. melanogaster*, se ha encontrado una inversión cromosómica producida durante la disgénesis híbrida I-R que se habría originado aparentemente por recombinación ectópica entre un elemento *I* completo y un fragmento de únicamente 344 bp (Busseau *et al.* 1989).

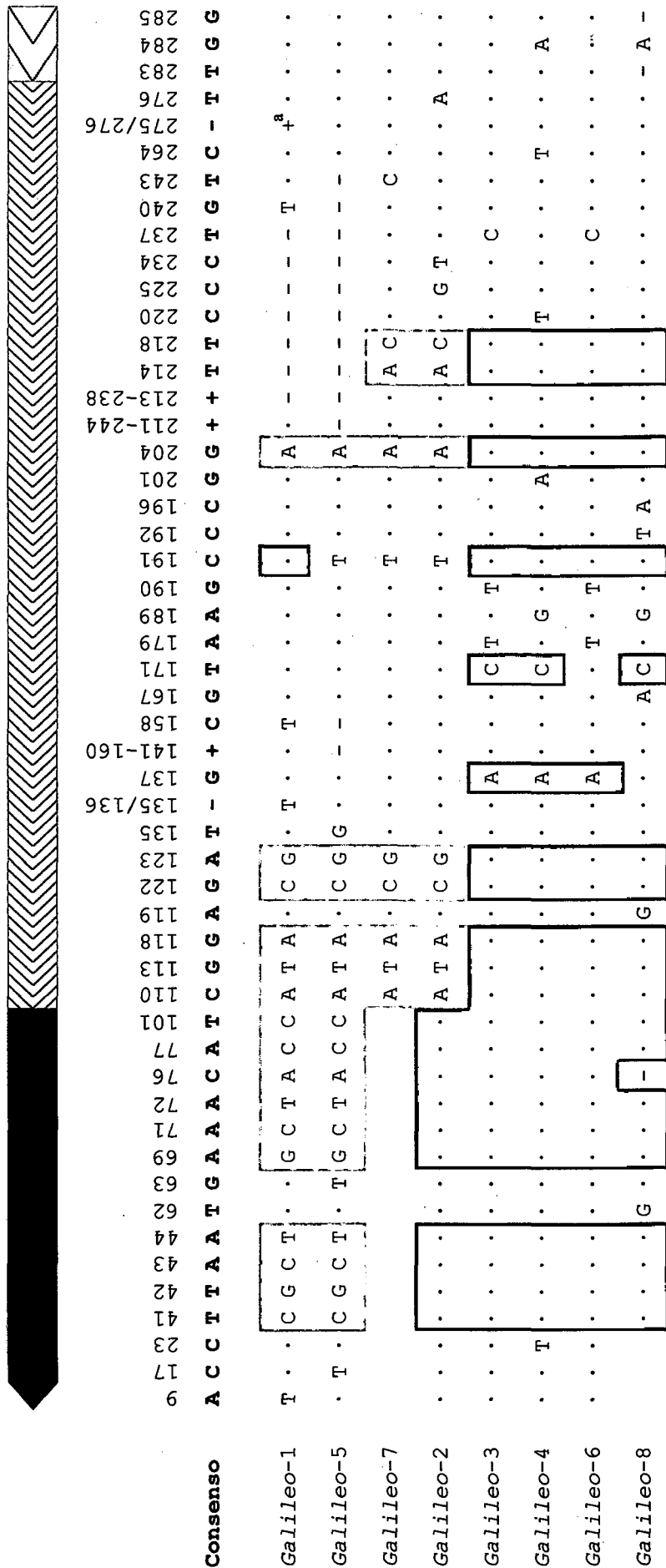


Figura 3.6. Cambios nucleotídicos presentes en los extremos de las ocho copias de *Galileo* secuenciadas. En color se muestran las posiciones en las que difieren los dos grupos de elementos *Galileo* (en color más claro los cambios que están presentes en un 75% de los elementos de un tipo). Los puntos indican identidad con la secuencia consenso y los guiones deletiones respecto a ésta. ^aInserción de 12 bp en *Galileo-1*.

La recombinación ectópica entre copias de un mismo elemento transponible dispersas por el genoma podría ser un mecanismo universal de aparición de inversiones y otras alteraciones cromosómicas en la naturaleza. Con anterioridad, se había demostrado su papel en la generación de inversiones cromosómicas naturales en bacterias (Daveran-Míngot *et al.* 1998). En este trabajo, se ha demostrado por primera vez su implicación en una inversión natural del género *Drosophila*, que se puede considerar que ha tenido un gran éxito evolutivo. De hecho, se ha recurrido a las diferencias en el contenido de elementos transponibles en el genoma, para explicar algunas de las diferencias que hemos visto en la Introducción en la distribución de las inversiones entre las especies. Así, por ejemplo, se ha visto que *D. simulans*, que es monomórfica, tendría un menor contenido de DNA repetitivo que *D. melanogaster*, que es polimórfica para varias inversiones (Dowsett y Young 1982).

La reacción de transposición de los distintos tipos de elementos transponibles conlleva roturas en la molécula de DNA (Craig 1997) y esta característica, junto con la presencia de múltiples copias, los hace posiblemente substratos preferentes de la recombinación ectópica. En los elementos transponibles de Clase I son necesarios cortes en el DNA para que el cDNA proveniente de la transcripción inversa del RNA del elemento se pueda integrar en nuevas posiciones del genoma. Por su parte, además, para que se produzca la transposición de los elementos transponibles de Clase II, el DNA del elemento se ha de separar del DNA adyacente mediante cortes de doble cadena a ambos lados. En cualquier caso, los extremos de DNA generados por ambas vías podrían promover la búsqueda de homología y el intercambio de las cadenas, que constituyen las primeras fases de la recombinación (Eggleston *et al.* 1996). El modelo actualmente más aceptado de recombinación en los eucariotas supone que el proceso se inicia por la reparación de una rotura de DNA de doble cadena creada por una topoisomerasa (Lewin 1997). En *Drosophila*, se ha visto así que, de una manera dependiente de la presencia de transposasa, la inserción de elementos *P* estimula localmente la recombinación en machos (Sved *et al.* 1990; McCarron *et al.* 1994) y el elemento *mariner* aumenta la frecuencia de recombinación meiótica en la región en la que se inserta (Lohe *et al.* 2000). Un fenómeno similar podría tener lugar también en

el caso de la recombinación ectópica. Por tanto, los elementos transponibles podrían no ser simplemente sujetos pasivos sobre los que actúan los sistemas de recombinación celulares, sino que podrían promover de forma activa la producción de reorganizaciones como consecuencia más o menos directa de su mecanismo de acción.

De momento, la *2j* es la única inversión natural de *Drosophila* que avala la validez del modelo de recombinación ectópica entre elementos transponibles. Aparte de ésta, hay otras dos inversiones polimórficas en las que se han encontrado secuencias repetitivas relacionadas con elementos transponibles en los puntos de rotura: la inversión *In(2L)t* de *D. melanogaster* (Andolfatto *et al.* 1999) y la inversión *2Rd'* de *A. arabiensis* (Mathiopoulos *et al.* 1998). Esto elevaría a un 60% el porcentaje de inversiones caracterizadas a nivel molecular que podrían tener elementos transponibles implicados en su origen. Lamentablemente, en ninguno de los dos casos anteriores se ha podido determinar el mecanismo concreto por el que se generó la inversión ni el papel real que habrían desempeñado las secuencias repetitivas. En la inversión *In(2L)t*, se ha encontrado un ORF de 279 bp adyacente a uno de los puntos de rotura, que podría pertenecer a una transcriptasa inversa de un elemento no descrito de tipo *LINE* (Andolfatto *et al.* 1999). En el otro punto de rotura en principio no parece haber ninguna secuencia homóloga, aunque el pequeño trozo secuenciado no permite descartar que se trate de parte del mismo elemento dispuesto en orientación contraria. En el cromosoma *Y* humano se ha descubierto una inversión fijada que se habría producido por recombinación ilegítima entre elementos *LINE-1* (Schwartz *et al.* 1998), y éste podría ser quizás un caso similar. En la inversión *2Rd'*, en uno de los puntos de rotura se ha encontrado una secuencia repetitiva que podría corresponder a un nuevo elemento transponible denominado *Odysseus* (Mathiopoulos *et al.* 1998). No obstante, no se ha clonado el otro punto de rotura de la inversión, por lo que no se ha podido comprobar si el elemento está presente también en éste o no.

En la inversión *In(3L)Payne* (Wesley y Eanes 1994) y la inversión fijada entre *D. melanogaster* y *D. subobscura* (Cirera *et al.* 1995) no se han hallado indicios de la presencia de elementos transponibles en los puntos de rotura. Una posibilidad es que hubiese elementos transponibles implicados en el origen de ambas inversiones, pero, al ser móviles, en un momento posterior los elementos se excindiesen y desapareciesen. Esto podría ser cierto para la inversión fijada, en la que el tiempo transcurrido desde su generación ha impedido determinar exactamente donde se produjeron las roturas y bien podría haber borrado la huellas de la acción de los elementos transponibles. No obstante, la edad de la inversión *In(3L)Payne* se ha estimado en sólo 360.000 años (Hasson y Eanes 1996) y no se han detectado prácticamente diferencias en la secuencia de los puntos de rotura entre los cromosomas con y sin la inversión (Wesley y Eanes 1994). Por tanto, este hecho nos lleva a sugerir que la inversión no habría sido causada por elementos transponibles y que quizás se produjo simplemente por la reparación incorrecta de dos roturas espontáneas del cromosoma, sin causa conocida.

La desaparición de los elementos transponibles de los puntos de rotura de una inversión tras su generación sin dejar rastro podría no ser tan fácil. Los elementos de Clase I se transponen por un mecanismo replicativo y en principio la copia original se conserva sin cambios (Craig 1997). Los elementos de Clase II sí que suelen excindirse durante la transposición, pero dejan tras de sí las duplicaciones del sitio diana producidas en el momento de la inserción y, en muchos casos, pequeños restos del elemento (Engels *et al.* 1990; Plasterk 1991). La excisión precisa podría darse en ambos tipos de elementos por recombinación o por deslizamiento de la cadena durante la síntesis de DNA entre las duplicaciones directas del sitio diana (Egner y Berg 1981; Collins y Rubin 1983). No obstante, como hemos visto, en el caso de los puntos de rotura de una inversión, las secuencias a cada lado de los elementos serían diferentes (Schneuwly *et al.* 1987; Cáceres *et al.* 1999a). Además, los elementos de Clase II pueden desaparecer de forma precisa cuando se utiliza la secuencia de un cromosoma homólogo sin la inserción como molde para la reparación de la rotura de doble cadena producida en

la excisión (Engels *et al.* 1990). Sin embargo, una vez el elemento ha causado la inversión, no existe ninguna región idéntica a los puntos de rotura que no presente el elemento.

De esta forma, los mecanismos mediante los que han surgido las inversiones descritas en *Drosophila* podrían ser variados. Aparte de la rotura espontánea del cromosoma que podría haber causado la *In(3L)Payne* y de la recombinación ectópica entre elementos transponibles que ha dado lugar a la inversión *2j*, las inversiones también podrían generarse de otras formas. Por ejemplo, la recombinación ectópica se puede producir entre cualquier tipo de secuencias repetidas por el genoma, como genes pertenecientes a familias multigénicas, genes de RNA ribosómicos y genes de RNA de transferencia. Así, se han encontrado inversiones causadas por recombinación entre genes de RNA ribosómicos en bacterias (Hill y Harnish 1981) y una inversión polimórfica descubierta en humanos se habría producido por recombinación entre dos repeticiones invertidas de 11.3 kb de origen desconocido (Small *et al.* 1997). Harían falta más estudios para poder generalizar sobre la importancia que tienen los diferentes mecanismos en el origen de las inversiones naturales y el papel real de los elementos transponibles en la reestructuración del genoma.

3.1.6 ¿Son calientes los puntos de rotura de la inversión *2j*?

El análisis estadístico de la localización de los puntos de rotura de las 61 inversiones descritas en el cromosoma 2 en las 12 especies del complejo *buzzatii* ha mostrado que éstos no se distribuyen aleatoriamente (Cáceres *et al.* 1997). De forma similar a lo que sucede en la mayoría de las otras especies de *Drosophila* estudiadas (Tonzetich *et al.* 1988; Krimbas 1992; Lemeunier y Aulard 1992), se observa un mayor número de segmentos cromosómicos que no presentan ninguna rotura y de segmentos en los que se localizan más de cinco roturas de lo esperado por azar. Esto sería debido a que a nivel citológico existen, aparentemente, puntos calientes que

participan en múltiples inversiones y podrían tener una mayor predisposición a sufrir roturas. En concreto, dentro del complejo *buzzatii*, se han encontrado seis posiciones del cromosoma 2 que podrían ser interesantes para un posterior estudio: las bandas C6a, D1g, D5a y E2e, que acumulan cuatro roturas cada una, la banda F6a, que acumula cinco roturas, y la banda F2a, en la que se han localizado los puntos de rotura de hasta ocho inversiones diferentes (Cáceres *et al.* 1997).

Según la visión tradicional sobre el origen de las inversiones, estos presuntos puntos calientes se han explicado por la existencia de regiones en el DNA con mayor tendencia a romperse, bien por la propia naturaleza de la secuencia nucleotídica o bien por que la organización de la molécula de DNA dentro del núcleo celular hace que estén más expuestas a la acción de agentes clastógenos, como las radiaciones ionizantes (Krimbas y Powell 1992). Sin embargo, la demostración de que los elementos transponibles desempeñaban un papel fundamental en la generación de las inversiones en el laboratorio (Lim y Simmons 1994), hizo sospechar que los puntos calientes podían ser debidos a la acción de éstos. De acuerdo con el modelo de generación de las inversiones por recombinación ectópica entre elementos transponibles, tanto la preferencia de los elementos por insertarse en una misma región como la presencia de un elemento insertado de forma permanente en una posición, podrían provocar que un determinado lugar del cromosoma intervenga en varias inversiones y se constituya en un punto caliente. Asimismo, la retención de los elementos responsables de la aparición de una inversión en los puntos de rotura, haría que éstos puedan volver a recombinar ectópicamente y producir otras inversiones.

La inversión $2j$ es un ejemplo de inversión originada por recombinación ectópica entre copias de un elemento transponible y puede servir para estudiar esta cuestión. Los puntos de rotura proximal y distal se localizan, respectivamente, en las bandas E5a y C6c del cromosoma 2 de *D. buzzatii* (Ruiz *et al.* 1984). En el complejo *buzzatii* se habían descrito cuatro inversiones más que parecían tener al menos uno de sus puntos de rotura en una de estas posiciones. En las inversiones $2w^9$ y $2x^9$ de *D. koepferae*, uno de los puntos de rotura se situaba aparentemente en la banda E5a (Ruiz y

Wasserman 1993). Las inversiones $2z^3$ de *D. buzzatii* (Ruiz *et al.* 1984) y $2g^8$ de *D. borborema* (Wasserman y Richardson 1987) compartían supuestamente el punto de rotura situado en la banda C6c con la inversión $2j$. En el caso de la inversión $2z^3$, el hecho que surgiese sobre un cromosoma con la inversión $2j$, hacía posible que el elemento *Galileo-2* situado en esa posición pudiese estar también implicado en su origen. No obstante, un estudio más preciso mediante hibridación *in situ* del gen *Pp1 α -96A* sobre cromosomas $2jz^3$, mostró que el punto de rotura de la inversión $2z^3$ no coincidía en realidad con el de la $2j$, sino que estaba situado más internamente (Ranz *et al.* 1997).

La existencia de puntos calientes para la generación de inversiones y la implicación de los elementos transponibles en su aparición podría tener consecuencias importantes a nivel evolutivo. En primer lugar, abre la posibilidad de que las inversiones puedan ocurrir de forma recurrente. La probabilidad de que surjan inversiones diferentes con idénticos puntos de rotura sería mayor de lo que se espera por simple rotura al azar de los cromosomas. Por tanto, una determinada inversión se podría producir varias veces de forma independiente, por ejemplo por recombinación ectópica entre elementos insertados en las mismas posiciones, y algunas inversiones podrían tener un origen polifilético. Un estudio detallado de la variabilidad nucleotídica y estructural en la región de los puntos de rotura de la inversión $2j$ ha permitido descartar esta posibilidad (Cáceres *et al.* 2000). En principio, todos los cromosomas $2j$ analizados poseen puntos de rotura idénticos respecto a los cromosomas $2st$, pero esto no es garantía suficiente de un origen único de la inversión. El análisis de las inserciones presentes en los puntos de rotura de la inversión en diversos cromosomas $2j$ reveló la existencia de una sorprendente variación. No obstante, aparentemente las diferentes variantes estructurales encontradas en cada punto de rotura derivan de una única inserción ancestral común. Además, el análisis de la variabilidad nucleotídica en la región de los puntos de rotura de 9 cromosomas $2st$ y 12 cromosomas $2j$, de procedencias diversas y pertenecientes a los diversos tipos estructurales caracterizados, ha mostrado que todos los cromosomas $2j$ forman un grupo claramente monofilético. Los cromosomas $2j$ comparten una gran cantidad de

diferencias fijadas respecto a los cromosomas *2st*, incluyendo tres inserciones y tres deleciones pequeñas. Al mismo tiempo, los cromosomas *2j* muestran muy poca variabilidad nucleotídica entre sí. Por tanto, varias evidencias apuntan a que la inversión *2j* tendría un origen único y que los cromosomas *2j* actuales serían el resultado del mismo evento de recombinación ectópica entre *Galileo-1* y *Galileo-2*.

La participación de los elementos móviles en la generación de las inversiones por recombinación ectópica no parece estar así necesariamente refrendada con el carácter monofilético que se les había supuesto (Krimbas y Powell 1992; Powell 1997). Aunque una misma inversión pueda producirse varias veces, la probabilidad de que las diferentes réplicas sobrevivan y lleguen a establecerse en las poblaciones sigue siendo baja. De hecho, en general los datos moleculares han confirmado el origen único de las inversiones estudiadas. Estos estudios se han basado en el análisis de la variabilidad nucleotídica y la realización de árboles filogenéticos de cromosomas con y sin la inversión, ya sea utilizando genes incluidos dentro del segmento invertido (Aquadro *et al.* 1991; Popadic y Anderson 1994; Rozas y Aguadé 1994; Navarro-Sabaté *et al.* 1999) o bien las propias secuencias de los puntos de rotura de la inversión (Hasson y Eanes 1996; Andolfatto *et al.* 1999). Hasta el momento, la única excepción es el caso de la inversión *2L^a*, que a nivel citológico parece haber surgido y haberse retenido de forma independiente en dos especies del complejo *gambiae* de *Anopheles*, *A. merus* y *A. arabiensis*, y apoya el posible origen polifilético de algunas inversiones cromosómicas (Caccone *et al.* 1998).

Por otra parte, el mantenimiento de las secuencias causantes de la inversión en sus puntos de rotura, haría que, en teoría, se pueda producir la reinversión y que, de igual forma que se generó la inversión, los cromosomas invertidos puedan revertir con cierta frecuencia a la ordenación no invertida. Este proceso podría tener importancia al incrementar el flujo genético entre las ordenaciones invertida y no invertida. No obstante, a pesar de que teóricamente podría producirse por una simple recombinación ectópica entre los mismos elementos que generaron la inversión, en el caso de la inversión *2j* de momento un estudio preliminar no ha detectado evidencias de

reinversión. Aunque no se ha llevado a cabo un estudio específico encaminado a investigar si es posible que en las poblaciones naturales la inversión revierta con cierta frecuencia a la ordenación *2st*, ninguno de los nueve cromosomas *2st* analizados parece ser el resultado de un proceso de este tipo (Cáceres *et al.* 2000). En primer lugar, ninguno de ellos posee elementos *Galileo* insertados en los puntos de rotura de la inversión. En segundo lugar, el análisis de la variabilidad nucleotídica en los puntos de rotura muestra una clara separación entre los cromosomas *2st* y *2j*.

Sin embargo, en los puntos de rotura de la inversión *2j* se ha detectado un grado de variación estructural natural sin precedentes en *Drosophila* (Aquadro 1993), que permite calificarlos al menos como puntos genéticamente inestables (Cáceres *et al.* 2000). En una región de tan sólo 7,1 kb en los cromosomas *2st*, se han producido un total de 22 inserciones de elementos transponibles, 12 deleciones y dos inversiones, la inversión *2j* y la inversión de un segmento interno de la inserción presente en el punto de rotura proximal. Además, el elevado número de cambios estructurales acumulados en esta región, contrasta considerablemente con los 16 cambios nucleotídicos encontrados entre los 12 cromosomas *2j* en los que se han secuenciado las regiones de los puntos de rotura. La edad de la inversión *2j* se ha calculado en aproximadamente un millón de años. De acuerdo con la estima del tiempo de coalescencia para las secuencias del DNA de copia única de los puntos de rotura de los 12 cromosomas *2j* correspondientes a las diferentes variantes estructurales, las 16 inserciones de elementos transponibles, las 12 deleciones y la inversión interna, habrían tenido lugar en menos de 350.000 años (Cáceres *et al.* 2000).

En principio, en los cromosomas con la ordenación *2st* no tiene lugar ninguno de los cambios anteriores. El claro contraste de la variación estructural observada en los cromosomas *2j* con la total ausencia de variación en los cromosomas *2st*, es clave para poder determinar cuales podrían ser las causas de este fenómeno. A primera vista, los cromosomas *2st* y *2j* muestran únicamente dos claras diferencias comunes a las regiones de los dos puntos de rotura, a las que se pueda atribuir tan desigual comportamiento: la propia inversión del segmento comprendido entre los puntos de

rotura y las inserciones de *Galileo* en ambos puntos de rotura. Es difícil imaginar un mecanismo por el cual la presencia de la inversión pueda estimular la acumulación del tipo de cambios observado y, a la vez, hacerlo de forma preferente en los cromosomas con la inversión. Además, esta variación no se ha observado en los puntos de rotura de las otras inversiones naturales caracterizados (Wesley y Eanes 1994; Hasson y Eanes 1996; Andolfatto *et al.* 1999). Se ha propuesto que los elementos transponibles podrían acumularse en regiones de baja recombinación, como pueden ser los puntos de rotura de las inversiones (Montgomery *et al.* 1987; Eanes *et al.* 1992; Sniegowski y Charlesworth 1994). No obstante, esto no parece ser suficiente para explicar nuestros resultados (Cáceres *et al.* 2000).

La mayor parte de los cambios encontrados son posteriores a la inserción de los elementos *Galileo-1* y *Galileo-2* y ocurren en el interior o cerca de los extremos de éstos. En consecuencia, todo apuntaría a que la causa más probable de la inestabilidad en los puntos de rotura de la inversión *2j* es la inserción de los elementos *Galileo*. En cierto modo, no es sorprendente que estos elementos puedan estar implicados en la generación de la inversión y los otros cambios estructurales observados. Los elementos *Galileo* parecen ser elementos de tipo *Foldback* y los elementos *FB* de *D. melanogaster* han dado muestras de estar asociados con múltiples fenómenos de inestabilidad genética y de ser muy recombinogénicos en el laboratorio (Bingham y Zachar 1989; Smith y Corces 1991). En primer lugar, las inserciones de los elementos *FB* dan lugar a mutaciones altamente inestables, que revierten con una elevada frecuencia al fenotipo salvaje. La reversión sería debida a una excisión precisa tanto del elemento como de la secuencia duplicada en la inserción, aparentemente por un mecanismo de deslizamiento de la cadena durante la síntesis de DNA (Collins y Rubin 1983). En segundo lugar, estos elementos parecen estimular la producción de deleciones, inversiones y translocaciones recíprocas por recombinación ectópica entre dos de ellos (Collins y Rubin 1984). Por último, los elementos *FB* producen también translocaciones insercionales, en las que dos o más elementos movilizan conjuntamente grandes bloques de DNA situados entre ellos (Goldberg *et al.* 1982). Los transposones gigantes generados se ha visto que sufren

frecuentemente reorganizaciones internas, por recombinación entre los elementos *FB* que contienen (Lovering *et al.* 1991).

Los puntos de rotura de la inversión $2j$ son el primer caso en que se estudia la variabilidad nucleotídica y estructural en las poblaciones naturales asociada a una inserción de un elemento transponible. Por tanto, sería interesante conocer si otros elementos transponibles producen fenómenos de inestabilidad similares. Además, en los cromosomas de *D. buzzatii* se han localizado otras inserciones eucromáticas de *Galileo* (J. M. Ranz, comunicación personal) y sería interesante ver si en estos casos también hay un incremento de la variación estructural en la región. La inversión $2j$ podría ser en última instancia el resultado de un fenómeno de inestabilidad genética provocado por la actividad del elemento *Galileo*. En principio, sólo se ha podido caracterizar otra inversión en la que están implicados estos elementos, la causada por una recombinación entre dos copias del elemento *Kepler* situadas dentro de *Galileo-1*. No obstante, a la vista de la inestabilidad observada, los mismos elementos podrían haber participado en muchas otras reorganizaciones cromosómicas que no han tenido éxito y no han sobrevivido hasta nuestros días. Se puede sospechar que, muy probablemente, los elementos de tipo *Foldback* de *D. buzzatii* están implicados en la aparición de alguna otra de las inversiones descritas en esta especie.

3.2 Significado adaptativo de las inversiones

En relación con el polimorfismo de inversiones, existe la creencia generalizada de que tiene un carácter adaptativo y está sometido a fuertes presiones selectivas (Sperlich y Pfriem 1986; Krimbas y Powell 1992; Powell 1997). No obstante, hay discrepancias en cuanto a cuáles son las características de las inversiones sobre las que actúa la selección. En esta tesis hemos tratado de discriminar entre los dos principales efectos de las inversiones: las mutaciones causadas en los puntos de rotura en el momento de su aparición y la reducción de la recombinación dentro del segmento invertido en los heterocariotipos. Por una parte, la clonación de los puntos de rotura de la inversión *2j* ha permitido determinar que ésta no parece haber afectado a ninguno de los genes cercanos a sus puntos de rotura. Por otra parte, el análisis de la distribución de la longitud de las inversiones descritas en las especies del complejo *buzzatii* ha mostrado que en la naturaleza existe una selección a favor de las inversiones de tamaño intermedio. Además, se ha comprobado que el tamaño promedio de las inversiones evolutivamente exitosas está relacionado con la frecuencia de recombinación en diversas especies de dípteros y que parece existir una longitud genética óptima para las inversiones. Todos estos resultados constituyen poderosas evidencias a favor de la hipótesis de que la ventaja selectiva de las inversiones deriva principalmente de su efecto sobre la recombinación.

3.2.1 Posible efecto de posición de la inversión *2j*

La posibilidad de que las inversiones afecten a los genes próximos a sus puntos de rotura y las importantes consecuencias fenotípicas que esto puede tener queda de manifiesto por la existencia de mutaciones morfológicas como la *Antp*^{73b}, en la que una inversión provoca el intercambio del primer exón y las regiones reguladoras entre el gen *Antennapedia* y otro gen de función desconocida (Schneuwly *et al.* 1987).

Además, en bacterias, se han descrito múltiples sistemas encargados de controlar la transición entre fenotipos alternativos, que se basan en la inversión específica de ciertos segmentos de DNA, provocando la activación o la inhibición de la expresión de determinados genes (Glasgow *et al.* 1989). Ejemplos bien conocidos de estos sistemas son el que controla la expresión de los pili en *Escherichia coli* y el que controla el cambio de fase de los antígenos del flagelo de *Salmonella typhimurium*. Finalmente, en el caso de las inversiones naturales de *Drosophila*, se ha visto que uno de los puntos de rotura de la inversión *In(3L)Payne* interrumpe tres transcritos que se expresan normalmente en los individuos con la ordenación *standard*, aunque la función de estos transcritos y las consecuencias de su disrupción se desconocen (Wesley y Eanes 1994).

En el caso de la inversión *2j* parecen existir diferencias entre los individuos portadores de los cromosomas *2st* y *2j* en el tiempo de desarrollo de las larvas y el tamaño corporal de los adultos (Ruiz y Santos 1989; Ruiz *et al.* 1991; Hasson *et al.* 1992b; Barbadilla *et al.* 1994; Norry *et al.* 1995; Betrán *et al.* 1998; Fernández Iriarte y Hasson 2000). No obstante, las causas genéticas de estas diferencias no se conocen. Una posibilidad es que sean debidas al efecto de posición producido en los cromosomas *2j* con la aparición de la inversión, que podría haber afectado la expresión de alguno de los genes situados cerca de sus puntos de rotura.

Principalmente, son dos los genes que se han encontrado próximos a los puntos de rotura de la inversión *2j*: *rox8* y el conocido *Pp1 α -96A* (Cáceres *et al.* 1999a). El gen *rox8* está situado fuera de la inversión y, en función de la homología con el gen de *D. melanogaster*, su extremo 3' parece localizarse a 649 bp del punto de rotura distal. Este gen codifica aparentemente una nucleolisina con capacidad de unirse al RNA, que podría participar en la degradación del DNA celular (Brand y Bourbon 1993). Por su parte, el gen *Pp1 α -96A* cambia de posición con la inversión y en los cromosomas *2st* su extremo 5' está muy cerca del punto de rotura proximal. Concretamente, el inicio de la transcripción parece estar a tan sólo 156 bp de éste. Este gen codifica una proteína fosfatasa de tipo I (Dombrádi *et al.* 1990) y cabe

destacar que estas enzimas están implicadas en la regulación de muchas funciones celulares, incluyendo la división y el ciclo celular (Orr-Weaver 1994; Cohen 1997).

A pesar de la proximidad a los puntos de rotura, en ninguno de los dos casos se han podido detectar diferencias en la expresión de los genes entre una línea *2st* y una línea *2j* con la metodología empleada (Cáceres *et al.* 1999a). Por tanto, de acuerdo con estos resultados, el efecto de la inversión *2j* sobre los individuos portadores no parece ser debido a las mutaciones producidas en el momento de su aparición. Sin embargo, muy recientemente, la secuenciación del genoma de *D. melanogaster* (Adams *et al.* 2000), ha permitido descubrir la existencia de un gen de función desconocida en la región de los puntos de rotura de la inversión *2j*. El codón de finalización de la traducción de este gen está únicamente a 12 bp del punto de rotura proximal. Los estudios que se están llevando a cabo en nuestro grupo permitirán determinar si la expresión del nuevo gen se ha visto afectada por la inversión.

3.2.2 Relación entre la longitud y el éxito de las inversiones

La alternativa al efecto mutacional producido en el momento de su aparición como explicación del carácter adaptativo de las inversiones, es el efecto reductor de la recombinación en los heterocariotipos. La reducción de la recombinación está íntimamente relacionada con el tamaño de las inversiones (Navarro *et al.* 1997). En consecuencia, el análisis de la longitud de las inversiones presentes en las poblaciones naturales proporciona una manera sencilla de investigar la importancia del efecto sobre la recombinación en su supervivencia. Este tipo de estudio permite aprovechar el enorme trabajo citogenético de búsqueda de inversiones y descripción de las mismas que se viene realizando en distintas especies de *Drosophila* durante mucho tiempo (Sperlich y Pfriem 1986). Además, tiene la ventaja de que, al considerarse conjuntamente muchas inversiones de diferentes especies, permite extraer conclusiones de tipo general. Por último, la longitud de las inversiones se puede

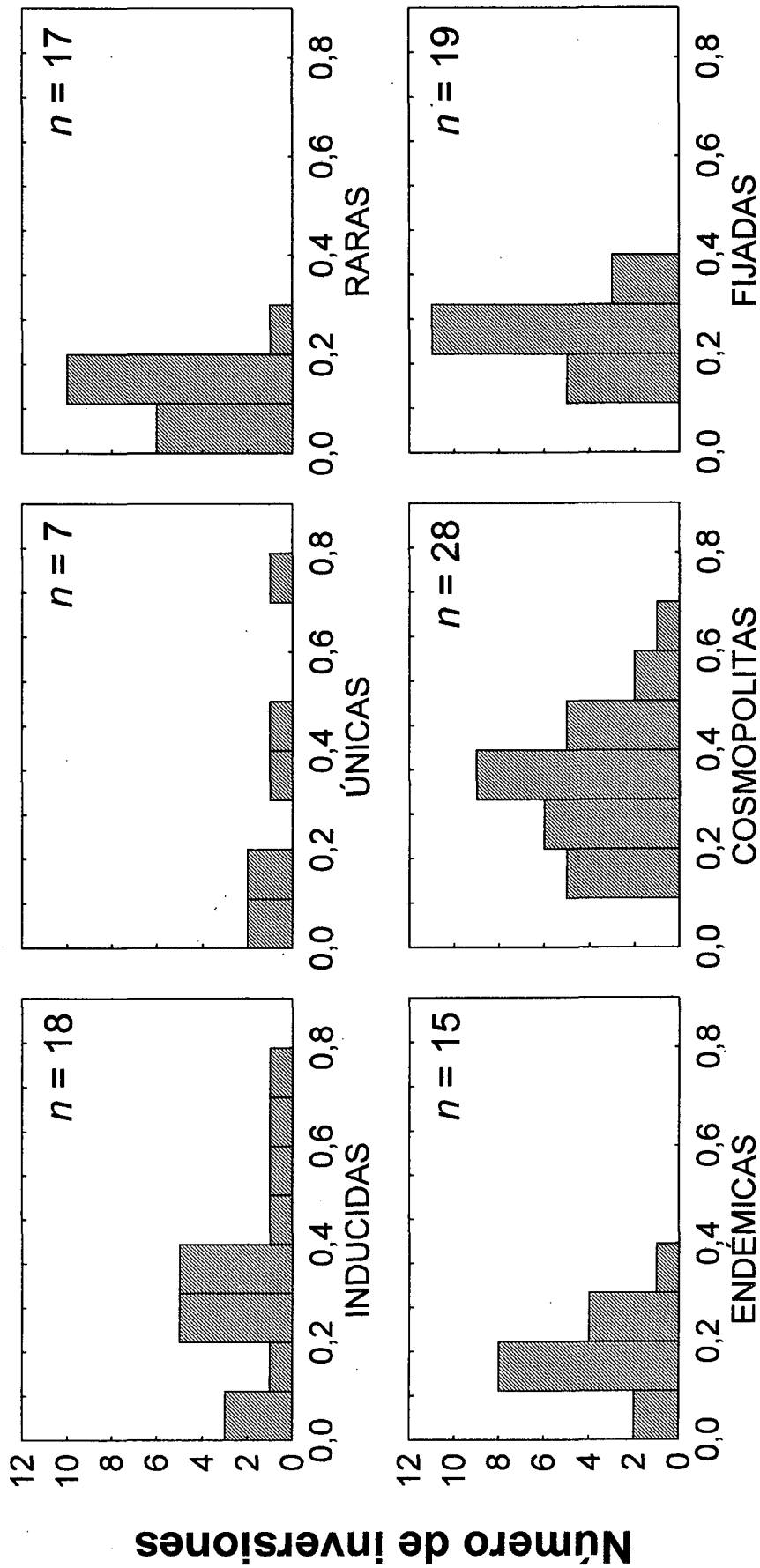
estimar fácilmente a partir de los mapas de los cromosomas politénicos y, a pesar de que los límites de las inversiones están sujetos a cierto grado de incertidumbre por la dificultad en la localización exacta de los puntos de rotura, el error cometido en la determinación de la longitud suele ser proporcionalmente pequeño.

Considerando que la ventaja selectiva de las inversiones proviene de la supresión de la recombinación en el segmento invertido (Kojima y Schaffer 1964; Wasserman 1968; Nei *et al.* 1967; Dobzhansky 1970; Charlesworth y Charlesworth 1973), se pueden hacer dos predicciones en cuanto a su longitud. En primer lugar, las inversiones cortas reducirán la recombinación tan sólo en una pequeña proporción y se espera que su ventaja selectiva sea igualmente pequeña. En segundo lugar, las inversiones largas tienen una mayor probabilidad de sufrir dobles entrecruzamientos y se espera que estén desfavorecidas, tanto por el incremento del flujo genético entre las ordenaciones como por el efecto detrimental que tienen una parte de los dobles entrecruzamientos (Navarro *et al.* 1997). Por tanto, si el anterior razonamiento es cierto, la conclusión es que las inversiones que han pasado el filtro selectivo y han llegado a establecerse en las poblaciones, ya sea en forma de polimorfismos o como diferencias fijadas entre especies, tenderían a tener un tamaño intermedio.

El complejo *buzzatii* constituye un material idóneo para el análisis de la longitud de las inversiones y las variables relacionadas. La información acumulada sobre la evolución cromosómica y el polimorfismo de inversiones de las especies que forman este complejo a lo largo de varios años de intenso estudio es muy completa (Ruiz y Wasserman 1993). Así, ha permitido disponer de un total de 104 inversiones, pertenecientes a varios grupos con diferente grado de éxito evolutivo, que se pueden comparar para probar el efecto de la selección sobre la longitud. Los tipos de inversiones considerados van concretamente desde las recién aparecidas en un individuo hasta las que se encuentran en toda una especie: (1) inversiones inducidas en el laboratorio en una cepa de *D. buzzatii* mediante hibridación introgresiva de un segmento cromosómico de *D. koepferae* (Naveira y Fontdevila 1985); (2) inversiones únicas, que se han encontrado en un único individuo de la especie *D. buzzatii* (Ruiz *et*

al. 1984); (3) inversiones raras, que se han encontrado en una única localidad; (4) inversiones endémicas, que están presentes en menos de un 25% de las localidades muestreadas de la especie; (5) inversiones cosmopolitas, que están presentes en más de un 25% de las localidades muestreadas de la especie; y (6) inversiones fijadas, que se encuentran en todos los individuos de una o más especies.

De esta forma, se ha podido apreciar un cambio casi gradual en la longitud de las inversiones, expresada como fracción del cromosoma, en función del éxito evolutivo de las mismas (Figura 3.7). Las inversiones inducidas por hibridación introgresiva y las inversiones únicas podrían representar bastante bien el proceso de surgimiento de las inversiones en la naturaleza. Ambos grupos incluyen toda la gama posible de longitudes de las inversiones, desde las pequeñas a las más grandes de todo el conjunto de datos, y su distribución no diferiría de lo esperado según un modelo de rotura al azar de los cromosomas (Federer *et al.* 1967; Van Valen y Levins 1968). Por tanto, podrían ser un buen modelo de la distribución de longitudes inicial de las inversiones, antes de que se produzca ningún tipo de selección. A partir de aquí, las inversiones raras, con una longitud promedio de 0,124, y las inversiones endémicas, con una longitud promedio de 0,177, comprenden sólo inversiones inferiores al 30% y al 40% del cromosoma, respectivamente. Estas inversiones podrían tener un carácter casi neutro. Su pequeño tamaño haría que no tengan efectos detrimentales sobre los individuos portadores, pero, al mismo tiempo, no les conferiría una ventaja selectiva suficiente como para incrementar de frecuencia y extenderse por las poblaciones. Por su parte, las inversiones cosmopolitas y las inversiones fijadas tienen una longitud promedio de 0,333 y 0,259, respectivamente, y corresponden únicamente a inversiones intermedias. Las inversiones fijadas tienen un tamaño muy restringido y parecen consistir en un subconjunto dentro de las inversiones cosmopolitas. Esto puede interpretarse como que sólo las inversiones polimórficas con una mayor capacidad de extenderse y mantenerse en las poblaciones de la especie, quizás aquellas con un tamaño más adecuado, llegan a fijarse. De acuerdo con esto, el proceso que lleva a la fijación de las inversiones no sería muy diferente del implicado en el establecimiento del polimorfismo.



Longitud de las inversiones

Figura 3.7. Distribución de la longitud de las inversiones descritas en el complejo *buzzatii* de *Drosophila*, clasificadas en seis grupos en función de su éxito evolutivo.

La comparación de la distribución de las longitudes de las inversiones de estos diferentes grupos ha proporcionado una prueba clara de la acción de la selección natural sobre las inversiones y de que una de las características claves en la supervivencia de las inversiones es el tamaño (Cáceres *et al.* 1997). Previamente ya se había observado la existencia de una selección a favor de las inversiones de tamaño intermedio (Krimbas y Powell 1992). No obstante, en ningún otro caso se habían obtenido evidencias tan completas y significativas.

3.2.3 ¿Existe una longitud genética óptima para las inversiones?

Hemos visto como las inversiones de tamaño intermedio están favorecidas en el complejo *buzzatii* (Cáceres *et al.* 1997). Además, esta observación parece ser extensiva a otras especies de *Drosophila*, como *D. pseudoobscura* (Olvera *et al.* 1979) y *D. subobscura* (Krimbas y Loukas 1980). No obstante, la hipótesis de la importancia de la reducción de la recombinación en el mantenimiento de las inversiones sugiere que la acción de la selección se ejerce a nivel de la longitud genética de las inversiones. Por tanto, siguiendo el mismo razonamiento que antes, debería existir una especie de longitud genética óptima de las inversiones y el tamaño intermedio observado para las inversiones evolutivamente exitosas sería causado por la selección de inversiones de una determinada longitud genética. La longitud genética óptima de las inversiones resultaría, por un lado, del incremento de la ventaja selectiva de la inversión al proteger de la recombinación una mayor parte del cromosoma (Kojima y Schaffer 1964; Wasserman 1968; Nei *et al.* 1967; Dobzhansky 1970; Charlesworth y Charlesworth 1973; Charlesworth 1993) y, por otro lado, del incremento de la probabilidad de sufrir dobles entrecruzamientos y el efecto negativo que tienen éstos (Sturtevant y Beadle 1936; Navarro *et al.* 1997).

Para investigar la existencia de un óptimo en la longitud genética de las inversiones, se ha llevado a cabo un análisis conjunto de la longitud física de las

inversiones y la tasa de recombinación en diversas especies de dípteros (Cáceres *et al.* 1999b). La intención era tratar de extender el análisis al mayor número posible de especies y géneros diferentes, pero al final sólo se pudieron incluir diez especies de *Drosophila* y dos de *Anopheles*. En este caso, a diferencia de los estudios previos (Van Valen y Levins 1968; Olvera *et al.* 1979; Krimbas y Loukas 1980; Brehm y Krimbas 1991; Cáceres *et al.* 1997), la longitud física de las inversiones se ha expresado con respecto al tamaño total de la fracción eucromática presente en los cromosomas politénicos. Además, como estima de la tasa de recombinación en cada especie se ha utilizado la longitud genética total del genoma. De esta forma, al expresar las dos variables en relación al tamaño del genoma, se evita el problema de la existencia de variación en el tamaño del genoma entre especies, que por ejemplo es de 180 Mb en *D. melanogaster* (Adams *et al.* 2000), mientras que en *A. gambiae* alcanza las 260 Mb (Besansky y Powell 1992).

Dos son los principales resultados de este análisis. Primero, se ha observado que existen diferencias muy significativas entre la longitud física promedio de las inversiones evolutivamente exitosas en las diversas especies. Segundo, se ha visto que las diferencias en la longitud física de las inversiones se pueden explicar en gran parte por las diferencias en la longitud genética total del genoma de cada especie. Las especies con una longitud genética del genoma mayor tenderían a tener inversiones más pequeñas, mientras que las especies con una longitud genética del genoma menor tenderían a tener inversiones más grandes. Además, se han utilizado inversiones de las mismas especies con un menor grado de éxito evolutivo como control. Al contrario de lo que ocurre con las exitosas, estas inversiones no muestran diferencias de tamaño entre especies ni relación entre su tamaño y la longitud genética del genoma de la especie.

El contraste entre las inversiones exitosas y no exitosas pone de manifiesto que estos resultados no son debidos a cómo surgen las inversiones, sino a que en cada especie la selección natural está seleccionando inversiones de diferente tamaño en función de la longitud genética total del genoma. Asimismo, apunta a que,

posiblemente a nivel de todo el género *Drosophila* y quizás en los dípteros en general, las inversiones exitosas tienden a tener una longitud genética intermedia. Esto no quiere decir, sin embargo, que las inversiones se seleccionen exclusivamente en función de su longitud. Lo principal sería el conjunto concreto de genes contenido en el segmento invertido, pero la longitud constituiría un factor muy importante para que una inversión pueda tener éxito. Tanto los resultados de las inversiones del complejo *buzzatii* como los obtenidos con las 12 especies diferentes de dípteros, serían hasta el momento las pruebas más claras de que a nivel global la ventaja selectiva de las inversiones nace de su efecto sobre la recombinación.

La recombinación es un factor crucial en la adaptación de los organismos a su entorno (Prevosti 1978). Permite el intercambio entre las dotaciones genéticas de los individuos de una misma especie y, en consecuencia, la aparición de nuevas combinaciones génicas más favorables. No obstante, supondría un lastre una vez se ha conseguido la adaptación a unas condiciones estables. El nivel de recombinación en cada especie sería así el resultado del compromiso entre flexibilidad y estabilidad, y las diferencias entre especies podrían tener importantes consecuencias en su dinámica evolutiva. En nuestro caso, se ha encontrado una gran variación en la longitud genética total del genoma de 12 especies de dípteros (Cáceres *et al.* 1999b). Esta variación tiene un componente filogenético y parte podría ser atribuible a diferencias en el tamaño del genoma de cada especie, aunque este factor no parece ser suficiente para explicar las discrepancias observadas (Schafer *et al.* 1993). A lo largo de la evolución se han desarrollado diversos mecanismos para regular el nivel de recombinación en el genoma, como la existencia de reproducción sexual o la variación en el número de cromosomas. Además, dentro del género *Drosophila*, se han observado diferencias en la frecuencia de recombinación entre *D. melanogaster*, *D. simulans* y *D. mauritiana* (True *et al.* 1996). Las inversiones cromosómicas podrían ser un tipo de solución de compromiso, que permite modificar el nivel de recombinación en una especie de forma reversible.

Como ya hemos visto, la recombinación interacciona con los elementos transponibles para generar las inversiones presentes en la naturaleza. Además, de acuerdo con los resultados anteriores, desempeña también un papel fundamental en su mantenimiento en las poblaciones.

4. CONCLUSIONES

Del presente trabajo se pueden extraer principalmente las siguientes conclusiones.

En relación con el origen de las inversiones cromosómicas en el género *Drosophila*:

1. El mecanismo molecular por el que se ha generado la inversión cosmopolita *2j* de *D. buzzatii* es la recombinación ectópica entre dos secuencias homólogas de DNA moderadamente repetitivo insertadas en orientación opuesta en la región de los puntos de rotura.
2. Las secuencias repetitivas responsables de la generación de la inversión *2j* son copias defectivas de un nuevo elemento transponible de Clase II denominado *Galileo*. Este elemento transponible se caracteriza por la presencia de repeticiones terminales invertidas largas y por la duplicación de 7 bp del sitio en el que se inserta, y parece corresponder a un elemento de tipo *Foldback*.
3. El análisis de la variabilidad nucleotídica presente en las regiones de los puntos de rotura en los cromosomas con la ordenación *2 standard* y en los cromosomas con la inversión, muestra que todos los cromosomas *2j* estudiados provienen de un único evento de recombinación ectópica entre las copias de *Galileo* y que la inversión *2j* tiene un origen monofilético.
4. Los puntos de rotura de la inversión *2j* son regiones de inestabilidad genética en las que se han producido un elevado número de inserciones de elementos transponibles, deleciones de segmentos de DNA y otros tipos de cambios estructurales en un corto período de tiempo. Esta inestabilidad es debida supuestamente a la inserción de los elementos *Galileo* en ambos puntos de rotura.
5. Los puntos de rotura de las inversiones descritas en el complejo *buzzatii* de *Drosophila* no se distribuyen uniformemente sobre el cromosoma 2 y hay seis posiciones cromosómicas que acumulan un número de roturas muy alto.

En relación con el mantenimiento de las inversiones en las poblaciones naturales de *Drosophila*:

6. La aparición de la inversión *2j* no parece haber afectado a la expresión del gen *rox8* ni del gen *Ppl α -96A*, que se sitúan muy próximos a sus puntos de rotura. El análisis de la variabilidad nucleotídica en las regiones de los puntos de rotura de esta inversión no ha permitido detectar indicios de la acción de la selección.
7. En el complejo *buzzatii* existen diferencias claras en la distribución de la longitud de las inversiones con distinto grado de éxito evolutivo. Las inversiones evolutivamente más exitosas, las fijadas y las polimórficas, tienden a tener un tamaño intermedio, mientras que las inversiones raras y endémicas tienden a tener un tamaño pequeño.
8. La selección natural favorece inversiones de diferente longitud física en distintas especies de *Drosophila* y *Anopheles* en función de la longitud genética total del genoma de cada especie. Parece existir una longitud genética óptima de las inversiones como consecuencia de su efecto sobre la recombinación.
9. El fenómeno de la recombinación desempeña un papel clave tanto en el origen de las inversiones cromosómicas, conjuntamente con los elementos transponibles, como en su mantenimiento en las poblaciones naturales.

5. REFERENCIAS

- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185-2195.
- Amador, A. y E. Juan (1999) Nonfixed duplication containing the *Adh* gene and a truncated form of the *Adhr* gene in the *Drosophila funebris* species group: Different modes of evolution of *Adh* relative to *Adhr* in *Drosophila*. *Mol. Biol. Evol.* **16**: 1439-1456.
- Andolfatto, P., J. D. Wall y M. Kreitman (1999) Unusual haplotype structure at the proximal breakpoint of *In(2L)t* in a natural population of *Drosophila melanogaster*. *Genetics* **153**: 1297-1311.
- Aquadro, C. F. (1993) Molecular population genetics of *Drosophila*, pp. 222-266 en *Molecular Approaches to Fundamental and Applied Entomology*. Eds. J. Oakeshott y M. J. Whitten. Springer-Verlag, New York.
- Aquadro, C. F., A. L. Weaver, S. W. Schaeffer y W. W. Anderson (1991) Molecular evolution of inversions in *Drosophila pseudoobscura*: The amylase gene region. *Proc. Natl. Acad. Sci. USA* **88**: 305-309.
- Arkhipova, I. R., N. V. Lyubomirskaya y Y. V. Ilyin. (1995) *Drosophila* retrotransposons. Springer-Verlag, Heidelberg.
- Ayala, F. J., L. Serra y A. Prevosti (1989) A grand experiment in evolution: The *Drosophila subobscura* colonization of the Americas. *Genome* **31**: 246-255.
- Barbadilla, A., A. Ruiz, M. Santos y A. Fontdevila (1994) Mating pattern and fitness-component analysis associated with inversion polymorphism in a natural population of *Drosophila buzzatii*. *Evolution* **48**: 767-780.

- Barker, J. S. F. (1982) Population genetics of *Opuntia* breeding *Drosophila* in Australia, pp. 209-224 en *Ecological Genetics and Evolution. The Cactus-Yeast-Drosophila Model System*. Eds. J. S. F. Barker y W. T. Starmer. Academic Press, Sydney.
- Barker, J. S. F., F. M. Sene, P. D. East y M. A. Q. R. Pereira (1985) Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. *Genetica* **67**: 161-170.
- Berg, D. E. (1989) Transposon Tn5, pp. 185-210 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Berg, D. E. y M. M. Howe (1989) *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Besansky, N. J. y J. R. Powell (1992) Reassociation kinetics of *Anopheles gambiae* (Diptera: Culicidae) DNA. *J. Med. Entomol.* **29**: 125-128.
- Betrán, E., M. Santos y A. Ruiz (1998) Antagonistic pleiotropic effect of second-chromosome inversions on body size and early life-history traits in *Drosophila buzzatii*. *Evolution* **52**: 144-154.
- Bingham, P. M. y Z. Zachar (1989) Retrotransposons and the *FB* transposon from *Drosophila melanogaster*, pp. 485-502 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Brand, S. y H.-M. Bourbon (1993) The developmentally-regulated *Drosophila* gene *rox8* encodes an RRM-type RNA binding protein structurally related to human TIA-1-type nucleolysins. *Nucleic Acids Res.* **21**: 3699-3704.
- Bregliano, J.-C. y M. G. Kidwell (1983) Hybrid dysgenesis determinants, pp. 363-410 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.

- Brehm, A. y C. B. Krimbas (1991) Inversion polymorphism in *Drosophila obscura*. *J. Hered.* **82**: 110-117.
- Briggs, D. y S. M. Walters (1997) *Plant Variation and Evolution*. Cambridge University Press, Cambridge, Gran Bretaña.
- Britten, R. J. (1996) DNA sequence insertion and evolutionary variation in gene regulation. *Proc. Natl. Acad. Sci. USA* **93**: 9374-9377.
- Busseau, I., A. Pelisson y A. Bucheton (1989) *I* elements of *Drosophila melanogaster* generate specific chromosomal rearrangements during transposition. *Mol. Gen. Genet.* **218**: 222-228.
- Caccone, A., G.-S. Min y J. R. Powell (1998) Multiple origins of cytologically identical chromosome inversions in the *Anopheles gambiae* complex. *Genetics* **150**: 807-814.
- Cáceres, M., A. Barbadilla y A. Ruiz (1997) Inversion length and breakpoint distribution in the *Drosophila buzzatii* species complex: Is inversion length a selected trait? *Evolution* **51**: 1149-1155.
- Cáceres, M., J. M. Ranz, A. Barbadilla, M. Long y A. Ruiz (1999a) Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**: 415-418.
- Cáceres, M., A. Barbadilla y A. Ruiz (1999b) Recombination rate predicts inversion size in Diptera. *Genetics* **153**: 251-259.
- Cáceres, M., M. Puig y A. Ruiz (2000) Natural hotspots induced by transposable element insertion at the breakpoints of a *Drosophila* inversion. (enviado)

- Calvi, B. R., T. J. Hong, S. D. Findley y W. M. Gelbart (1991) Evidence for a common evolutionary origin of inverted repeat transposons in *Drosophila* and plants: *hobo*, *Activator*, and *Tam3*. *Cell* **66**: 465-471.
- Campuzano, S., L. Balcells, R. Villares, L. Carramolino, L. García-Alonso y J. Modolell (1986) Excess function *hairy-wing* mutations caused by *gypsy* and *copia* insertions within structural genes of the *achaete-scute* locus of *Drosophila*. *Cell* **44**: 303-312.
- Capy, P., C. Bazin, D. Higuete y T. Langin (1998) *Dynamics and Evolution of Transposable Elements*. Springer-Verlag, Heidelberg.
- Carson, H. L. (1946) The selective elimination of inversion dicentric chromatids during meiosis in the eggs of *Sciara impatiens*. *Genetics* **31**: 95-113.
- Carson, H. L. (1992) Inversions in Hawaiian *Drosophila*, pp. 407-439 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Cirera, S., J. M. Martín-Campos, C. Segarra y M. Aguadé (1995) Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. *Genetics* **139**: 321-326.
- Clayton, F. E. y W. C. Guest (1986) Overview of chromosomal evolution in the family Drosophilidae, pp. 1-38 en *The Genetics and Biology of Drosophila*, Vol. 3e. Eds. M. Ashburner, H. L. Carson y J. N. Thompson, Jr. Academic Press, London.
- Cleland, R. E. (1972) *Oenothera. Cytogenetics and Evolution*. Academic Press, New York.
- Cohen, P. T. W. (1997) Novel protein serine/threonine phosphatases: variety is the spice of life. *Trends Biochem. Sci.* **22**: 245-250.

- Collins, M. y G. M. Rubin (1983) High-frequency precise excision of the *Drosophila* foldback transposable element. *Nature* **303**: 259-260.
- Collins, M. y G. M. Rubin (1984) Structure of chromosomal rearrangements induced by the *FB* transposable element in *Drosophila*. *Nature* **308**: 323-327.
- Coyne, J. A., W. Meyers, A. P. Crittenden y P. Sniegowski (1993) The fertility effects of pericentric inversions in *Drosophila melanogaster*. *Genetics* **134**: 487-496.
- Craig, N. L. (1997) Target site selection in transposition. *Annu. Rev. Biochem.* **66**: 437-474.
- Crumpacker, D. W. y C. D. Kastriasis (1967) A new gene arrangement in the third chromosome of *Drosophila pseudoobscura*. *J. Hered.* **58**: 2-6.
- Charlesworth, B. (1993) Directional selection and the evolution of sex and recombination. *Genet. Res.* **61**: 205-224.
- Charlesworth, B. y D. Charlesworth (1973) Selection of new inversions in multilocus genetic systems. *Genet. Res.* **21**: 167-183.
- Charlesworth, B., P. Sniegowski y W. Stephan (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215-220.
- Daveran-Mingot, M.-L., N. Campo, P. Ritzenthaler y P. le Bourgeois (1998) A natural large chromosomal inversion in *Lactococcus lactis* is mediated by homologous recombination between two insertion sequences. *J. Bacteriol.* **180**: 4834-4842.
- Doak, T. G., F. P. Doerder, C. L. Jahn y G. Herrick (1994) A proposed superfamily of transposase genes: Transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc. Natl. Acad. Sci. USA* **91**: 942-946.

- Dobzhansky, Th. (1943) Genetics of natural populations. IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura*. *Genetics* **28**: 162-186.
- Dobzhansky, Th. (1948) Genetics of natural populations. XVI. Altitudinal and seasonal changes produced by natural selection in certain populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. *Genetics* **33**: 158-176.
- Dobzhansky, Th. (1956) Genetics of natural populations. XXV. Genetic changes in populations of *Drosophila pseudoobscura* and *Drosophila persimilis* in some localities in California. *Evolution* **10**: 82-92.
- Dobzhansky, Th. (1970) *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- Dobzhansky, Th. (1971) Evolutionary oscillations in *Drosophila pseudoobscura*, pp. 109-133 en *Ecological Genetics and Evolution*. Ed. R. Creed. Blackwell, Oxford.
- Dobzhansky, Th. y J. R. Powell (1975) The *willistoni* group of sibling species of *Drosophila*, pp. 589-622 en *Handbook of Genetics*, Vol. 3. Ed. R. C. King. Plenum Press, New York.
- Dombrádi, V., J. M. Axton, N. D. Brewis, E. F. Da Cruz e Silva, L. Alphey y P. T. W. Cohen (1990) *Drosophila* contains three genes that encode distinct isoforms of protein phosphatase 1. *Eur. J. Biochem.* **194**: 739-745.
- Doolittle, W. F. y C. Sapienza (1980) Selfish genes, the phenotype paradigm, and genome evolution. *Nature* **284**: 601-603.

- Dowsett, A. P. y M. W. Young (1982) Differing levels of dispersed repetitive DNA among closely related species of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **79**: 4570-4574.
- Dutrillaux, B., J. Couturier y E. Viegas-Pequignot (1981) Chromosomal evolution in primates, pp. 176-191 en *Chromosomes Today*, Vol. 7. Eds. M. D. Bennett, M. Bobrow y G. Hewitt. George Allen & Unwin, London.
- Eanes, W. F., C. Wesley y B. Charlesworth (1992) Accumulation of *P* elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genet. Res.* **59**: 1-9.
- Eggleston, W. B., N. R. Rim y J. K. Lim (1996) Molecular characterization of *hobo*-mediated inversions in *Drosophila melanogaster*. *Genetics* **144**: 647-656.
- Egner, C. y D. E. Berg (1981) Excision of transposon Tn5 is dependent on the inverted repeats but not on the transposase function of Tn5. *Proc. Natl. Acad. Sci. USA* **78**: 459-463.
- Engels, W. R. (1996) *P* elements in *Drosophila*, pp.103-123 en *Transposable Elements*. Eds. H. Saedler y A. Gierl. Springer-Verlag, Berlin.
- Engels, W. R. y C. R. Preston (1981) Identifying *P* factors in *Drosophila* by means of chromosome breakage hotspots. *Cell* **26**: 421-428.
- Engels, W. R. y C. R. Preston (1984) Formation of chromosome rearrangements by *P* factors in *Drosophila*. *Genetics* **107**: 657-678.
- Engels, W. R., D. M. Johnson-Schlitz, W. B. Eggleston y J. Sved (1990) High-frequency *P* element loss in *Drosophila* is homolog dependent. *Cell* **62**: 515-525.

- Federer, W. T., R. G. D. Steel y B. Wallace (1967) Mathematical models for chromosomal inversions. *Genetics* **55**: 783-795.
- Fedoroff, N. V. (1983) Controlling elements in maize, pp. 1-63 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.
- Felger, I. y D. Sperlich (1989) Cytological localization and organization of dispersed middle repetitive DNA sequences of *Drosophila subobscura*. *Chromosoma* **98**: 342-350.
- Fernández Iriarte, P. y E. Hasson (2000) The role of the use of different host plants in the maintenance of the inversion polymorphism in the cactophilic *Drosophila buzzatii*. *Evolution* **54**: 1295-1302.
- Figueiredo, V. L. C. y F. M. Sene (1992) Chromosome variability in Brazilian populations of *Drosophila buzzatii* (Diptera, Drosophilidae). *Rev. Brasil. Biol.* **52**: 555-561.
- Finnegan, D. J. (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet.* **5**: 103-107.
- Fontdevila, A., A. Ruiz, G. Alonso y J. Ocaña (1981) The evolutionary history of *Drosophila buzzatii*. I. Natural chromosomal polymorphism in colonized populations of the Old World. *Evolution* **35**: 148-157.
- Fontdevila, A., A. Ruiz, J. Ocaña y G. Alonso (1982) The evolutionary history of *Drosophila buzzatii*. II. How much has chromosomal polymorphism changed in colonization? *Evolution* **36**: 843-851.
- Fontdevila, A., C. Zapata, G. Álvarez, L. Sánchez, J. Méndez e I. Enríquez (1983) Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. *Genetics* **105**: 935-955.

- Francino, O., O. Cabré y A. Fontdevila (1993) Distribution of the *copia* transposable element in the *repleta* group of *Drosophila*. *Genet. Sel. Evol.* **25**: 501-516.
- Galas, D. J. y M. Chandler (1989) Bacterial insertion sequences, pp. 109-162 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Glasgow, A. C., K. T. Hughes y M. I. Simon (1989) Bacterial DNA inversion systems, pp. 637-659 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Goldberg, M. L., R. Paro y W. J. Gehring (1982) Molecular cloning of the *white* locus region of *Drosophila melanogaster* using a large transposable element. *EMBO J.* **1**: 93-98.
- Goldberg, M. L., J. Y. Sheen, W. J. Gehring y M. M. Green (1983) Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the *Drosophila melanogaster* genome. *Proc. Natl. Acad. Sci. USA* **80**: 5017-5021.
- Green, M. M. (1980) Transposable elements in *Drosophila* and other Diptera. *Annu. Rev. Genet.* **14**: 109-120.
- Hagemann, S., W. J. Miller, E. Haring y W. Pinsker (1998) Nested insertions of short mobile sequences in *Drosophila P* elements. *Chromosoma* **107**: 6-16.
- Handler, A. M. y S. P. Gomez (1997) A new *hobo*, *Ac*, *Tam3* transposable element, *hopper*, from *Bactrocera dorsalis* is distantly related to *hobo* and *Ac*. *Gene* **185**: 133-135.

- Hankeln, Th. y E. R. Schmidt (1990) New *Foldback* transposable element *TFB1* found in histone genes of the midge *Chironomus thummi*. *J. Mol. Biol.* **215**: 477-482.
- Harden, N. y M. Ashburner (1990) Characterization of the *FB-NOF* transposable element of *Drosophila melanogaster*. *Genetics* **126**: 387-400.
- Hartl, D. L. (1989) Transposable element *mariner* in *Drosophila* species, pp. 531-536 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Hasson, E. y W. F. Eanes (1996) Contrasting histories of three gene regions associated with *In(3L)Payne* of *Drosophila melanogaster*. *Genetics* **144**: 1565-1575.
- Hasson, E., J. C. Vilardi, H. Naveira, J. J. Fanara, C. Rodríguez, O. A. Reig y A. Fontdevila (1991) The evolutionary history of *Drosophila buzzatii*. XVI. Fitness component analysis in an original natural population from Argentina. *J. Evol. Biol.* **4**: 209-225.
- Hasson, E., H. Naveira y A. Fontdevila (1992a) The breeding sites of Argentinian cactophilic species of the *Drosophila mulleri* complex (subgenus *Drosophila-repleta* group). *Rev. Chilena de Historia Natural* **65**: 319-326.
- Hasson, E., J. J. Fanara, C. Rodríguez, J. C. Vilardi, O. A. Reig y A. Fontdevila (1992b) The evolutionary history of *Drosophila buzzatii*. XXIV. Second chromosome inversions have different average effects on thorax length. *Heredity* **68**: 557-563.
- Hasson, E., C. Rodríguez, J. J. Fanara, H. Naveira, O. A. Reig y A. Fontdevila (1995) The evolutionary history of *Drosophila buzzatii*. XXVI. Macrogeographic patterns of inversion polymorphism in New World populations. *J. Evol. Biol.* **8**: 369-384.

- Heffron, F. (1983) Tn3 and its relatives, pp. 223-260 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.
- Hill, C. W. y B. W. Harnish (1981) Inversions between ribosomal RNA genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78**: 7069-7072.
- Hoffman-Liebermann, B., D. Liebermann y S. N. Cohen (1989) *TU* elements and *Puppy* sequences, pp. 575-592 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Iida, S., J. Meyer y W. Arber (1983) Prokaryotic IS elements, pp. 159-221 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.
- Kidwell, M. G. y D. Lisch (1997) Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA* **94**: 7704-7711.
- Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, Gran Bretaña.
- King, M. (1993) *Species Evolution: The Role of Chromosome Change*. Cambridge University Press, Cambridge, Gran Bretaña.
- Kleckner, N. (1983) Transposon Tn10, pp. 261-298 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.
- Kleckner, N. (1989) Transposon Tn10, pp. 227-268 en *Mobile DNA*. Eds. D. E. Berg y M. M. Howe. American Society for Microbiology, Washington, D.C.
- Kojima, K. y H. E. Schaffer (1964) Accumulation of epistatic gene complexes. *Evolution* **18**: 127-131.
- Koller, P. C. (1936) Structural hybridity in *Drosophila pseudoobscura*. *J. Genet.* **32**: 79-102.

- Krimbas, C. B. (1992) The inversion polymorphism of *Drosophila subobscura*, pp. 127-220 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Krimbas, C. B. y M. Loukas (1980) The inversion polymorphism of *Drosophila subobscura*, pp. 163-234 en *Evolutionary biology*, Vol. 12. Eds. M. K. Hecht, W. C. Steere y B. Wallace. Plenum Press, New York.
- Krimbas, C. B. y J. R. Powell (1992) Introduction, pp. 1-52 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Labrador, M. y A. Fontdevila (1994) High transposition rates of *Osvaldo*, a new *Drosophila buzzatii* retrotransposon. *Mol. Gen. Genet.* **245**: 661-674.
- Labrador, M. y V. G. Corces (1997) Transposable element-host interactions: Regulation of insertion and excision. *Annu. Rev. Genet.* **31**: 381-404.
- Lacadena, J.-R. (1996) *Citogenética*. Editorial Complutense, Madrid.
- Lathe III, W. C., W. D. Burke, D. G. Eickbush y T. H. Eickbush (1995) Evolutionary stability of the *R1* retrotransposable element in the genus *Drosophila*. *Mol. Biol. Evol.* **12**: 1094-1105.
- Lemeunier, F. y S. Aulard (1992) Inversion polymorphism in *Drosophila melanogaster*, pp. 339-405 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Levitan, M. (1992) Chromosomal variation in *Drosophila robusta* Sturtevant, pp. 221-338 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Lewin, B. (1997) *Genes VI*. Oxford University Press, New York.

- Li, W.-H. (1997) *Molecular Evolution*. Sinauer Associates, Sunderland, MA.
- Lim, J. K. (1988) Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **85**: 9153-9157.
- Lim, J. K. y M. J. Simmons (1994) Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *BioEssays* **16**: 269-275.
- Liskay, R. M., A. Letsou y J. L. Stachelek (1987) Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics* **115**:161-167.
- Lohe, A. R., C. Timmons, I. Beerman, E. R. Lozovskaya y D. L. Hartl (2000) Self-inflicted wounds, template-directed gap repair and a recombination hotspot: Effects of the *mariner* transposase. *Genetics* **154**: 647-656.
- Lovering, R., N. Harden y M. Ashburner (1991) The molecular structure of *TE146* and its derivatives in *Drosophila melanogaster*. *Genetics* **128**: 357-372.
- Lyttle, T. W. y D. S. Haymer (1992) The role of the transposable element *hobo* in the origin of endemic inversions in wild populations of *Drosophila melanogaster*. *Genetica* **86**: 113-126.
- Marín, I. y A. Fontdevila (1995) Characterization of *Gandalf*, a new inverted-repeat transposable element of *Drosophila koepferae*. *Mol. Gen. Genet.* **248**: 423-433.
- Marín, I. y A. Fontdevila (1996) Evolutionary conservation and molecular characteristics of repetitive sequences of *Drosophila koepferae*. *Heredity* **76**: 355-366.

- Marín, I., M. Labrador y A. Fontdevila (1992) The evolutionary history of *Drosophila buzzatii*. XXIII. High content of nonsatellite repetitive DNA in *D. buzzatii* and in its sibling *D. koepferae*. *Genome* **35**: 967-974.
- Mather, K. (1943) Polygenic inheritance and natural selection. *Biol. Rev.* **18**: 32-64.
- Mathiopoulos, K. D., A. della Torre, V. Predazzi, V. Petrarca y M. Coluzzi (1998) Cloning of inversion breakpoints in the *Anopheles gambiae* complex traces a transposable element at the inversion junction. *Proc. Natl. Acad. Sci. USA* **95**: 12444-12449.
- McCarron, M., A. Duttaroy, G. Doughty y A. Chovnick (1994) *Drosophila P* element transposase induces male recombination additively and without a requirement for *P* element excision or insertion. *Genetics* **136**: 1013-1023.
- McClintock, B. (1984) The significance of responses of the genome to challenge. *Science* **226**: 792-801.
- McDonald, J. F. (ed.) (1993) *Transposable Elements and Evolution*. Kluwer Academic Publishers, Dordrecht.
- McDonald, J. F. (1995) Transposable elements: Possible catalysts of organismic evolution. *Trends Ecol. Evol.* **10**: 123-126.
- Montgomery, E. A., S.-M. Huang, C. H. Langley y B. H. Judd (1991) Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: Genome structure and evolution. *Genetics* **129**: 1085-1098.
- Montgomery, E. A., B. Charlesworth y C. H. Langley (1987) A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet. Res.* **49**: 31-41.

- Muller, J. H. (1940) Bearings of the "*Drosophila*" work on systematics, pp. 185-268 en *The New Systematics*. Ed. J. Huxley. Clarendon Press, Oxford.
- Nassif, N. y W. Engels (1993) DNA homology requirements for mitotic gap repair in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **90**: 1262-1266.
- Navarro, A., E. Betrán, A. Barbadilla y A. Ruiz (1997) Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695-709.
- Navarro-Sabaté, A., M. Aguadé y C. Segarra (1999) The relationship between allozyme and chromosomal polymorphism inferred from nucleotide variation at the *AcpH-1* gene region of *Drosophila subobscura*. *Genetics* **153**: 871-889.
- Naveira, H. y A. Fontdevila (1985) The evolutionary history of *Drosophila buzzatii*. IX. High frequencies of new chromosome rearrangements induced by introgressive hybridization. *Chromosoma* **91**: 87-94.
- Nei, M., K. Kojima y H. E. Schaffer (1967) Frequency changes of new inversions in populations under mutation-selection equilibria. *Genetics* **57**: 741-750.
- Norry, F. M., J. C. Vilarde, J. J. Fanara, E. Hasson y C. Rodríguez (1995) An adaptive chromosomal polymorphism affecting size-related traits, and longevity selection in a natural population of *Drosophila buzzatii*. *Genetica* **96**: 285-291.
- Ochman, H., A. S. Gerber y D. L. Hartl (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621-623.
- O'Hare, K. y G. M. Rubin (1983) Structures of *P* transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* **34**: 25-35.
- Ohno, S. (1970) *Evolution by Gene Duplication*. Springer-Verlag, Berlin.

- Olvera, O., J. R. Powell, M. E. de la Rosa, V. M. Salceda, M. I. Gaso, J. Guzman, W. W. Anderson y L. Levine (1979) Population genetics of Mexican *Drosophila*. VI. Cytogenetic aspects of the inversion polymorphism in *Drosophila pseudoobscura*. *Evolution* **33**: 381-395.
- Orgel, L. E. y F. H. C. Crick (1980) The ultimate parasite. *Nature* **284**: 604-607.
- Orr-Weaver, T. L. (1994) Developmental modification of the *Drosophila* cell cycle. *Trends Genet.* **10**: 321-327.
- Painter, T. S. (1933) A new method for the study of chromosome rearrangements and the plotting of chromosome maps. *Science* **78**: 585-586.
- Pantazidis, A., M. Labrador y A. Fontdevila (1999) The retrotransposon *Osvaldo* from *Drosophila buzzatii* displays all structural features of a functional retrovirus. *Mol. Biol. Evol.* **16**: 909-921.
- Patterson, J. T. y M. R. Wheeler (1942) Description of new species of the subgenera *Hirtodrosophila* and *Drosophila*. *Univ. Texas Publ.* **4213**: 67-109.
- Plasterk, R. H. A. (1991) The origin of footprints of the *Tc1* transposon of *Caenorhabditis elegans*. *EMBO J.* **10**: 1919-1925.
- Plasterk, R. H. A. (1995) Mechanisms of DNA transposition, pp. 18-37 en *Mobile Genetic Elements*. Ed. D. J. Sherratt. Oxford University Press, Oxford.
- Polard, P. y M. Chandler (1995) Bacterial transposases and retroviral integrases. *Mol. Microbiol.* **15**: 13-23.
- Popadic, A. y W. W. Anderson (1994) The history of a genetic system. *Proc. Natl. Acad. Sci. USA* **91**: 6819-6823.

- Potter, S. S. (1982) DNA sequence of a *foldback* transposable element in *Drosophila*. *Nature* **297**: 201-204.
- Potter, S. S., M. Truett, M. Phillips y A. Maher (1980) Eukaryotic transposable genetic elements with inverted terminal repeats. *Cell* **20**: 639-647.
- Powell, J. R. (1992) Inversion polymorphisms in *Drosophila pseudoobscura* and *Drosophila persimilis*, pp. 73-126 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Powell, J. R. (1997) *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University Press, New York.
- Prevosti, A. (1978) Polimorfismo cromosómico y evolución. *Investigación y Ciencia* **26**: 90-103.
- Prevosti, A., G. Ribó, L. Serra, M. Aguadé, J. Balaña, M. Monclús y F. Mestres (1988) Colonization of America by *Drosophila subobscura*: Experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proc. Natl. Acad. Sci. USA* **85**: 5597-5600.
- Ranz, J. M., C. Segarra y A. Ruiz (1997) Chromosomal homology and molecular organization of Muller's elements *D* and *E* in the *Drosophila repleta* species group. *Genetics* **145**: 281-295.
- Ranz, J. M., M. Cáceres y A. Ruiz (1999) Comparative mapping of cosmids and gene clones from a 1.6 Mb chromosomal region of *Drosophila melanogaster* in three species of the distantly related subgenus *Drosophila*. *Chromosoma* **108**: 32-45.
- Ranz, J. M., F. Casals y A. Ruiz (2000) How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* (en prensa)

- Rebatchouk, D. y J. O. Narita (1997) *Foldback* transposable elements in plants. *Plant. Mol. Biol.* **34**: 831-835.
- Regner, L. P., M. S. O. Pereira, C. E. V. Alonso, E. Abdelhay y V. L. S. Valente (1996) Genomic distribution of *P* elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. *J. Hered.* **87**: 191-198.
- Rezsohazy, R., H. G. A. M. Van Luenen, R. M. Durbin y R. H. A. Plasterk (1997) *Tc7*, a *Tc1*-hitch hiking transposon in *Caenorhabditis elegans*. *Nucleic Acids Res.* **25**: 4048-4054.
- Roberts, P. A. (1976) The genetics of chromosomal aberration, pp. 67-184 en *The Genetics and Biology of Drosophila*, Vol. 1a. Eds. M. Ashburner y E. Novitski. Academic Press, London.
- Rodríguez-Trelles, F., G. Álvarez y C. Zapata (1996) Time-series analysis of seasonal changes of the *O* inversion polymorphism of *Drosophila subobscura*. *Genetics* **142**: 179-187.
- Rodríguez-Trelles, F., L. Alarcón y A. Fontdevila (2000) Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Mol. Biol. Evol.* **17**: 1112-1122.
- Roeder, G. S. y G. R. Fink (1983) Transposable elements in yeast, pp. 299-328 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.
- Rozas, J. y M. Aguadé (1994) Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **91**: 11517-11521.
- Rubin, G. M. (1983) Dispersed repetitive DNAs in *Drosophila*, pp. 329-361 en *Mobile Genetic Elements*. Ed. J. A. Shapiro. Academic Press, Orlando.

- Ruiz, A. y M. Santos (1989) Mating probability, body size, and inversion polymorphism in a colonizing population of *Drosophila buzzatii*, pp. 96-113 en *Evolutionary Biology of Transient Unstable Populations*. Ed. A. Fontdevila. Springer-Verlag, Berlin.
- Ruiz, A. y M. Wasserman (1993) Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity* **70**: 582-596.
- Ruiz, A., H. Naveira y A. Fontdevila (1984) La historia evolutiva de *Drosophila buzzatii*. IV. Aspectos citogenéticos de su polimorfismo cromosómico. *Genét. Ibér.* **36**: 13-35.
- Ruiz, A., A. Fontdevila, M. Santos, M. Seoane y E. Torroja (1986) The evolutionary history of *Drosophila buzzatii*. VIII. Evidence for endocyclic selection acting on the inversion polymorphism in a natural population. *Evolution* **40**: 740-755.
- Ruiz, A., M. Santos, A. Barbadilla, J. E. Quezada-Díaz, E. Hasson y A. Fontdevila (1991) Genetic variance for body size in a natural population of *Drosophila buzzatii*. *Genetics* **128**: 739-750.
- Ruiz, A., J. M. Ranz, M. Cáceres, C. Segarra, A. Navarro y A. Barbadilla (1997) Chromosomal evolution and comparative gene mapping in the *Drosophila repleta* species group. *Braz. J. Genet.* **20**: 553-565.
- Schafer, D. J., D. K. Fredline, W. R. Knibb, M. M. Green y J. S. F. Barker (1993) Genetics and linkage mapping of *Drosophila buzzatii*. *J. Hered.* **84**: 188-194.
- Schneuwly, S., A. Kuroiwa y W. J. Gehring (1987) Molecular analysis of the dominant homeotic *Antennapedia* phenotype. *EMBO J.* **6**: 201-206.

- Schwartz, A., D. C. Chan, L. G. Brown, R. Alagappan, D. Pettay, C. Disteché, B. McGillivray, A. de la Chapelle y D. C. Page (1998) Reconstructing hominid *Y* evolution: *X*-homologous block, created by *X-Y* transposition, was disrupted by *Yp* inversion through *LINE-LINE* recombination. *Hum. Mol. Genet.* **7**: 1-11.
- Small, K., J. Iber y S. T. Warren (1997) Emerin deletion reveals a common *X*-chromosome inversion mediated by inverted repeats. *Nat. Genet.* **16**: 96-99.
- Smith, P. A. y V. G. Corces (1991) *Drosophila* transposable elements: Mechanisms of mutagenesis and interactions with the host genome. *Advances in Genetics* **29**: 229-300.
- Sniegowski, P. y B. Charlesworth (1994) Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. *Genetics* **137**: 815-827.
- Sorsa, V. (1988) *Polytene Chromosomes in Genetic Research*. John Wiley & Sons, New York.
- Sperlich, D. (1963) Chromosomal polymorphism and position effect in *Drosophila subobscura*. *Proc. 11th Int. Congr. Genet.* **1**: 166.
- Sperlich, D. (1966) Equilibria for inversions induced by X-rays in isogenic strains of *Drosophila pseudoobscura*. *Genetics* **53**: 835-842.
- Sperlich, D. y P. Pfriem (1986) Chromosomal polymorphism in natural and experimental populations, pp. 257-309 en *The Genetics and Biology of Drosophila*, Vol. 3e. Eds. M. Ashburner, H. L. Carson y J. N. Thompson, Jr. Academic Press, London.
- Spicer, G. S. (1995) Phylogenetic utility of the mitochondrial *cytochrome oxidase* gene: Molecular evolution of the *Drosophila buzzatii* species complex. *J. Mol. Evol.* **41**: 749-759.

- Stone, W. S., W. C. Guest y F. D. Wilson (1960) The evolutionary implications of the cytological polymorphism and phylogeny of the *virilis* group of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **46**: 350-361.
- Streck, R. D., J. E. MacGaffey y S. K. Beckendorf (1986) The structure of *hobo* transposable elements and their insertion sites. *EMBO J.* **5**: 3615-3623.
- Sturtevant, A. H. (1917) Genetic factors affecting the strength of linkage in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **3**: 555-558.
- Sturtevant, A. H. (1926) A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biol. Zentralbl.* **46**: 697-702.
- Sturtevant, A. H. y C. R. Plunkett (1926) Sequence of corresponding third chromosome genes in *Drosophila melanogaster* and *Drosophila simulans*. *Biol. Bull.* **50**: 56-60.
- Sturtevant, A. H. y G. W. Beadle (1936) The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics* **21**: 544-604.
- Sugawara, N. y J. E. Haber (1992) Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol. Cell. Biol.* **12**: 563-575.
- Sved, J. A., W. B. Eggleston y W. R. Engels (1990) Germ-line and somatic recombination induced by *in vitro* modified P elements in *Drosophila melanogaster*. *Genetics* **124**: 331-337.
- Tan, C. C. (1935) Salivary gland chromosomes in the two races of *Drosophila pseudoobscura*. *Genetics* **20**: 392-402.

- Tanksley, S. D., M. W. Ganai y G. B. Martin (1995) Chromosome landing: A paradigm for map-based gene cloning in plants with large genomes. *Trends Genet.* **11**: 63-68.
- Templeton, N. S. y S. S. Potter (1989) Complete *foldback* transposable elements encode a novel protein found in *Drosophila melanogaster*. *EMBO J.* **8**: 1887-1894.
- Tonzetich, J., T. W. Lyttle y H. L. Carson (1988) Induced and natural break sites in the chromosomes of Hawaiian *Drosophila*. *Proc. Natl. Acad. Sci. USA* **85**: 1717-1721.
- True, J. R., J. M. Mercer y C. C. Laurie (1996) Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* **142**: 507-523.
- Truett, M. A., R. S. Jones y S. S. Potter (1981) Unusual structure of the *FB* family of transposable elements in *Drosophila*. *Cell* **24**: 753-763.
- Van Valen, L. y R. Levins (1968) The origins of inversion polymorphism. *Am. Nat.* **102**: 5-23.
- Virgin, J. B. y J. P. Bailey (1998) The *M26* hotspot of *Schizosaccharomyces pombe* stimulates meiotic ectopic recombination and chromosomal rearrangements. *Genetics* **149**: 1191-1204.
- Vivas, M. V., J. García-Planells, C. Ruiz, G. Marfany, N. Paricio, R. González-Duarte y R. de Frutos (1999) *GEM*, a cluster of repetitive sequences in the *Drosophila subobscura* genome. *Gene* **229**: 47-57.
- Wallace, B. (1954) Coadaptation and the gene arrangements of *Drosophila pseudoobscura*. *Inter. Union Biol. Sci. Symp. Genet. Popul. Struct. Ser. B* **15**: 67-94.

- Wasserman, M. (1954) Cytological studies of the *repleta* group. *Univ. Texas Publ.* **5422**: 130-152.
- Wasserman, M. (1968) Recombination-induced chromosomal heterosis. *Genetics* **58**: 125-139.
- Wasserman, M. (1982) Evolution of the *repleta* group, pp. 61-139 en *The Genetics and Biology of Drosophila*, Vol. 3b. Eds. M. Ashburner, H. L. Carson y J. N. Thompson, Jr. Academic Press, London.
- Wasserman, M. (1992) Cytological evolution of the *Drosophila repleta* species group, pp. 455-552 en *Drosophila Inversion Polymorphism*. Eds. C. B. Krimbas y J. R. Powell. CRC Press, Boca Raton, FL.
- Wasserman, M. y R. H. Richardson (1987) Evolution of Brazilian *Drosophila mulleri* complex species. *J. Hered.* **78**: 282-286.
- Wesley, C. S. y W. F. Eanes (1994) Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L)Payne* in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **91**: 3132-3136.
- Wesley, C. S., M. Ben, M. Kreitman, N. Hagag y W. F. Eanes (1990) Cloning regions of the *Drosophila* genome by microdissection of polytene chromosome DNA and PCR with nonspecific primer. *Nucleic Acids Res.* **18**: 599-603.
- Wharton, L. P. (1942) Analysis of the *repleta* group of *Drosophila*. *Univ. Texas Publ.* **4228**: 23-52.
- White, M. J. D. (1973) *Animal Cytology and Evolution*. Cambridge University Press, Cambridge, Gran Bretaña.
- White, M. J. D. (1978) *Modes of Speciation*. W. H. Freeman and Company, San Francisco.

- Wright, S. y Th. Dobzhansky (1946) Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*. *Genetics* **31**: 125-156.
- Yoon, J. S. y R. H. Richardson (1978) A mechanism of chromosomal rearrangements: The role of heterochromatin and ectopic joining. *Genetics* **88**: 305-316.
- Yuan, J., M. Finney, N. Tsung y R. Horvitz (1991) *Tc4* a *Caenorhabditis elegans* transposable element with an unusual fold-back structure. *Proc. Natl. Acad. Sci. USA* **88**: 3334-3338.

6. AGRADECIMIENTOS

Siempre he creído que los agradecimientos eran una de las partes más difíciles de escribir en una tesis. Es momento de hacer repaso a varios años de trabajo y son muchas las personas que, en mayor o menor medida, me han ayudado y han contribuido a que pudiese llevarlo a buen término. Es difícil resumir en unas pocas líneas mi inmensa gratitud hacia ellas. No obstante, me gustaría expresar mi más sincero agradecimiento a los que han participado más directamente en los cuatro artículos que forman parte de la tesis, así como reiterar mi agradecimiento a las personas mencionadas en ellos. Además, me gustaría agradecer:

A mis directores de tesis, Alfredo Ruiz y Antonio Barbadilla, por su ayuda en la realización de esta tesis y la oportunidad de llevarla a cabo.

A Sebastián Calero, por enseñarme prácticamente todo lo que sé sobre técnicas moleculares y sentar las bases para que pudiese aprender lo que no tuvo tiempo de enseñarme. Sin eso no creo que hubiese podido realizar este trabajo.

A mi amigo José María Ranz, con el que empecé esta aventura que no sabíamos como podía acabar y con el que he compartido tantos buenos y malos momentos en el laboratorio. Contigo los buenos momentos fueron mejores y los malos, menos malos.

Al resto de mis compañeros de laboratorio, que, de una forma u otra, me han estimulado en mi trabajo. A Arcadi Navarro, que con su ayuda llenó mis lagunas informáticas en múltiples ocasiones. A Marta Puig, que ha contribuido a mejorar este escrito y ha hecho más agradable mi estancia en el laboratorio este último año. A Cristina Santa, por la magnífica portada que ilustra la tesis.

A mis compañeros de la Universidad de Chicago, Fred Brunet, Mark Jensen, Soojin Yi, Wen Wang y Jianming Zhang, y principalmente a Manyuan Long, que me acogieron como uno de los suyos.

A Oriol Cabré, por su generosidad dejándonos utilizar su laboratorio y su consejo técnico en las primeras fases de la clonación de los puntos de rotura. A Jordi Barbé, por proporcionarnos reactivos y material, de los que, desafortunadamente, no disponíamos.

Al Personal d'Administració i Serveis del Departament de Genètica i de Microbiologia, que han facilitado mi trabajo, y sobre todo a Julia Provecho y Montse Sales, siempre dispuestas a solucionar cualquier problema.

A mis otros amigos, que han sido una válvula de escape a la siempre absorbente investigación: Jorge, Marta Sol., Jordi y los compañeros del equipo de fútbol, Félix, Agus, Jose, Salva, Ignaci, Óscar, . . . A mis compañeros de la central, Sebas y Álex, por esos seminarios tan informales y espontáneos.

A Angels, por acompañarme en todo el camino y estar siempre a mi lado, personal y científicamente. Esta tesis es casi tan tuya como mía.

A mi familia, y en especial a mi madre, que es la que más ha tenido que soportarme durante estos años.

Por último, al Comissionat per a Universitats i Recerca de la Generalitat de Catalunya y a la Dirección General de Investigación Científica y Técnica del antiguo Ministerio de Educación y Ciencia, que con su financiación, aunque muchas veces considerada insuficiente, han hecho posible este trabajo.

A todos ellos, gracias. Espero que, igual que yo, piensen que ha valido la pena.



o UAB
Universitat Autònoma de Barcelona
Servei de Biblioteques
Reg. 170058568
Sig. UVAB / 5411
Ref. 12500