



Universitat Autònoma de Barcelona

Facultat de Ciències



Departament de Genètica i de Microbiologia

Caracterización Genómica y Funcional de las Reorganizaciones del Complejo de Genes Hox en Drosophila

Genomic and Functional Characterization
of Hox Gene Complex Rearrangements in Drosophila

Bárbara Negre de Bofarull

Abril de 2005

Memoria presentada por la Licenciada en Biología Bárbara Negre de Bofarull para la obtención del título de Doctora en Ciencias Biológicas.

Bellaterra, veintidós de abril de dos mil cinco.

El Doctor Alfredo Ruiz Panadero, Catedrático del Departament de Genètica i de Microbiologia de la Facultat de Ciències de la Universitat Autònoma de Barcelona,

Certifica que: Bárbara Negre de Bofarull ha llevado a cabo bajo su dirección el trabajo de investigación realizado en el Departament de Genètica i de Microbiologia de la Facultat de Ciències de la Universitat Autònoma de Barcelona que ha llevado a la elaboración de la Tesis Doctoral titulada *Caracterización genómica y funcional de las reorganizaciones del complejo de genes Hox en Drosophila*.

Dr. Alfredo Ruiz Panadero

Al Marc

To Alex

To Nicholas and Ivan

A la Paula

*“The beauty of nature lies in detail;
the message in generality”*

Stephen Jay Gould

Wonderful Life: The Burgess Shale and the Nature of History, 1988.

ÍNDICE

ÍNDICE DE TABLAS	iii
ÍNDICE DE FIGURAS	v
ABREVIATURAS	vii
RESUMEN	ix
ABSTRACT	xi
1. INTRODUCCIÓN	1
1.1. LA UTILIDAD DE LOS ORGANISMOS MODELO	4
1.2. DROSOPHILA COMO MODELO	8
1.3. DROSOPHILA COMO MODELO EN BIOLOGÍA DEL DESARROLLO	12
1.3.1. Ciclo Vital	12
1.3.1.1. Embriogénesis	13
1.3.1.2. Metamorfosis	14
1.3.2. Genética del Desarrollo Embrionario	15
1.4. DROSOPHILA COMO MODELO EN GENÓMICA	17
1.4.1. El Número de Genes	18
1.4.2. Las Secuencias Reguladoras	19
1.4.3. Genómica Comparada	21
1.5. EVOLUCIÓN CROMOSÓMICA	22
1.5.1. Origen de las Reordenaciones Cromosómicas	24
1.5.2. Importancia Evolutiva de los Segmentos Conservados	24
1.6. EL COMPLEJO DE GENES HOX	26
1.6.1. Los Genes Homeóticos	26
1.6.2. Agrupación y Colinealidad	28
1.6.3. Origen del Complejo Hox	29
1.6.4. Los Genes Hox y la Evolución Morfológica	31
1.6.5. Estructura del Complejo en los Metazoos	35
1.7. LOS COMPLEJO DE GENES HOX EN DROSOPHILA	37
1.7.1. Estructura del Complejo	37
1.7.2. Activación y Mantenimiento de los Patrones de Expresión Hox	39
1.7.3. Función de los Genes Hox	40
1.7.4. Cambios Estructurales del HOM-C en el Género Drosophila	41
1.8. OBJETIVOS	44
1.8. OBJECTIVES	45

ii

ÍNDICE DE TABLAS

INTRODUCCIÓN

Tabla 1. Homologías cromosómicas entre algunas especies del género <i>Drosophila</i> .	23
---	----

RESULTADOS

Artículo 1 – Tabla 1. Secuencia de los cebadores utilizados para la amplificación por PCR de diferentes regiones del gen <i>lab</i> en varias especies de <i>Drosophila</i> .	53
Artículo 1 – Tabla 2. Estructura del gen <i>lab</i> en tres especies de <i>Drosophila</i> y <i>Tribolium castaneum</i> .	55
Artículo 1 – Tabla 3. Estimaciones por máxima verosimilitud del número de sustituciones sinónimas (d_S) y no sinónimas (d_N) por posición en regiones codificantes del gen <i>lab</i> en comparaciones a pares entre <i>D. buzzatii</i> , <i>D. virilis</i> y <i>D. melanogaster</i> .	58
Artículo 1 – Tabla 4. Test de la razón de máxima verosimilitud de la variación en la evolución molecular y razón ω del gen <i>lab</i> en el género <i>Drosophila</i> .	59
Artículo 1 – Tabla 5. Logaritmos de los valores de verosimilitud y estimaciones de los parámetros bajo modelos de sitios fijados.	59
Artículo 2 – Tabla 1. Características de las regiones conservadas no codificantes (CNS) detectadas con mVISTA en comparaciones de las regiones de los genes Hox entre <i>D. melanogaster</i> , <i>D. pseudoobscura</i> y <i>D. buzzatii</i> .	70
Artículo 2 – Tabla 2. Características de las regiones conservadas no codificantes (CNS) detectadas con mVISTA en comparaciones de las regiones de los genes no Hox entre <i>D. melanogaster</i> , <i>D. pseudoobscura</i> y <i>D. buzzatii</i> .	71
Artículo 2 – Tabla S1. Análisis estadístico de las características de los CNS entre las tres comparaciones a pares entre especies.	77
Artículo 2 – Tabla S2. Análisis estadístico de las características de los CNS entre parejas de especies para las diferentes regiones.	77

Artículo 2 – Tabla S3. Estadística del proceso de secuenciación de dos 78clones BAC de una genoteca de <i>D. buzzatii</i> .	78
Artículo 2 – Tabla S4. Parejas de cebadores utilizados para la amplificación de cDNA.	78
Artículo 2 – Tabla S5. Listado de CNS encontrados en las comparaciones con <i>D. buzzatii</i> .	79

ÍNDICE DE FIGURAS

INTRODUCCIÓN

Figura 1. Filogenia de los metazoos.	7
Figura 2. Filogenia de los principales grupos de artrópodos e insectos.	7
Figura 3. Árbol filogenético de los principales grupos de <i>Drosophila</i> .	11
Figura 4. Ciclo vital de <i>Drosophila</i> .	12
Figura 5. Fases del desarrollo embrionario de <i>Drosophila</i> .	14
Figura 6. Discos y tejidos imaginales de <i>Drosophila</i> .	15
Figura 7. Cascada de activación de los genes de segmentación.	17
Figura 8. Transformaciones homeóticas en <i>Drosophila</i> .	26
Figura 9. Conservación de la organización genómica y patrones de expresión de los genes Hox.	30
Figura 10. Distribución de los genes Hox en los Bilateria.	32
Figura 11. Jerarquía de cambios evolutivos en los genes Hox según su presunto efecto morfológico.	34
Figura 12. Estructura del HOM-C en <i>D. melanogaster</i> .	38
Figura 13. Genes que intervienen en la regulación del patrón de expresión de los genes Hox.	40
Figura 14. Reconstrucción filogenética de la organización de los genes Hox en el género <i>Drosophila</i> .	42

RESULTADOS

Artículo 1 – Figura 1. Reconstrucción filogenética de la organización de los genes Hox en el género <i>Drosophila</i> .	52
Artículo 1 – Figura 2. Localización cromosómica de los genes <i>lab</i> y <i>abdA</i> en el grupo <i>repleta</i> .	54
Artículo 1 – Figura 3. Mapa de restricción y organización molecular de la región <i>lab</i> en <i>D. buzzatii</i> .	56
Artículo 1 – Figura 4. Alineamiento de la 5' UTR y secuencias 5' de <i>lab</i> en <i>D. buzzatii</i> , <i>D. virilis</i> y <i>D. melanogaster</i> .	57
Artículo 1 – Figura 5. Gráfico VISTA mostrando la similaridad en las secuencias nucleotídicas del gen <i>lab</i> entre <i>D. melanogaster</i> y <i>D. buzzatii</i> .	58

Artículo 1 – Figura 6. Alineamiento múltiple de las proteínas codificadas por los ortólogos de <i>lab</i> de diferentes organismos.	60
Artículo 2 – Figura 1. Posición genómica (A) y filogenética (B) de las tres roturas del HOM-C descritas en el género <i>Drosophila</i> .	68
Artículo 2 – Figura 2. Organización génica de las regiones genómicas <i>lab-abdA</i> y <i>pb</i> en <i>D. buzzatii</i> y comparación con las regiones homólogas de <i>D. pseudoobscura</i> y <i>D. melanogaster</i> .	68
Artículo 2 – Figura 3. Conservación nucleotídica en las regiones <i>lab-abdA</i> y <i>pb</i> entre especies de <i>Drosophila</i> .	69
Artículo 2 – Figura 4. Patrón de expresión de <i>pb</i> en embriones.	71
Artículo 2 – Figura 5. Reconstrucción de la evolución del HOM-C en el subgénero <i>Drosophila</i> .	73
Artículo 2 – Figura S1. Detalle de dos módulos <i>cis</i> -reguladores (CRM).	111
Artículo 2 – Figura S2. Patrón de expresión de <i>pb</i> en discos imaginales.	112
Artículo 2 – Figura S3. Patrón de expresión de <i>lab</i> en embriones.	112
Artículo 2 – Figura S4. Patrón de expresión de <i>lab</i> en discos imaginal de ojo-antena.	113
Artículo 2 – Figura S5. Patrón de expresión de <i>abdA</i> en embriones.	113
Artículo 2 – Figura S6. Patrón de expresión de <i>abdA</i> en discos imaginales.	113
 DISCUSIÓN	
Figura 15. Modelo de evolución del gen <i>Hox3</i> .	122
Figura 16. Expresión del gen <i>pb</i> en embriones de cinco especies de <i>Drosophila</i> .	126
Figura 17. Modelo de evolución del HOM-C en <i>Drosophila</i> .	131
Figura 18. Estructura del HOM-C en los Metazoos.	140

ABREVIATURAS

<i>abdA</i>	gen <i>abdominal-A</i>
<i>AbdB</i>	gen <i>Abdominal-B</i>
<i>ama</i>	gen <i>amalgam</i>
ANT-C	Complejo Antennapedia
<i>Antp</i>	gen <i>Antennapedia</i>
BAC	Cromosoma Artificial de Bacteria (Bacteria Artificial Chromosome)
<i>bcd</i>	gen <i>bicoid</i>
<i>bx</i>	mutación <i>bithorax</i>
<i>bxl</i>	gen no codificante <i>bithoraxoid</i>
BX-C	Complejo Bithorax
<i>Ccp</i>	genes del complejo de proteínas cuticulares (<i>Cuticle Cluster Proteins</i>)
cDNA	DNA complementario al RNA mensajero (complementary DNA)
CNS	Secuencias conservadas no codificantes (Conserved Noncoding Sequences)
CRM	Módulos <i>cis</i> -reguladores (<i>Cis-Regulatory Modules</i>)
<i>Dfd</i>	gen <i>Deformed</i>
DPE	Downstream Promoter Element
<i>ftz</i>	gen <i>fushi tarazu</i>
Hox	gen homeótico
HOM-C	Complejo de genes homeóticos
<i>iab</i>	<i>infraabdominal</i>
kb	kilobases
<i>lab</i>	gen <i>labial</i>
MARs	Regiones de unión a la matriz (Matriz Attachment Regions)
Mb	Megabases
mRNA	RNA mensajero
ORF	marco abierto de lectura (Open Reading Frame)
pb	pares de bases
<i>pb</i>	gen <i>proboscipedia</i>
<i>pbx</i>	<i>postbithorax</i>
PCR	Reacción en Cadena de la Polimerasa (Polimerase Chain Reaction)
PFGE	Electroforesis de campo pulsante (Pulsed Field Gel Electroforesis)

RT-PCR	PCR por transcriptasa inversa (Retro Transcriptase-PCR)
SAGE	Análisis Serial de la Expresión Génica (Serial Analysis of Gene Expression)
SARs	Regiones de unión a la matriz (Scaffolding Attachment Regions)
<i>Scr</i>	gen <i>Sex combs reduced</i>
tRNA	RNA de transferencia
UAS	Secuencia activadora aguas arriba (Upstream Activating Sequence)
<i>Ubx</i>	gen <i>Ultrabithorax</i>
UTR	región transcrita pero no traducida (UnTranslated Region)
<i>zen</i>	gen <i>zerknüllt</i>
<i>zen2</i>	gen <i>zerknüllt2</i>
WGS	Secuenciación genómica aleatoria (Whole Genome Shotgun)

RESUMEN

Los genes homeóticos (Hox) codifican factores de transcripción involucrados en la especificación de la identidad segmental a lo largo del eje anteroposterior en el embrión de los metazoos. Estos genes se presentan habitualmente agrupados y organizados en el mismo orden en el que se expresan a lo largo del eje anteroposterior del cuerpo. La conservación de esta organización genómica a lo largo de la filogenia ha sugerido la existencia de constricciones funcionales actuando sobre ella, pero la desestructuración del complejo en *Caenorhabditis elegans* y las tres roturas descritas en el género *Drosophila* ponen en cuestión que esta organización sea realmente necesaria para su correcto funcionamiento en algunos linajes. En este trabajo se han estudiado las consecuencias genómicas y funcionales de las dos reorganizaciones del complejo de genes Hox presentes en *Drosophila buzzatii*, especie perteneciente al grupo repleta. En la primera parte del trabajo se han clonado y secuenciado las regiones codificantes del gen *labial* (*lab*) en *D. buzzatii* y *D. virilis*. Se han comparado las secuencias de estas dos especies y la de *D. melanogaster* para comprobar si el cambio de posición del gen *lab* ocurrido en el linaje de *D. buzzatii* había producido algún cambio en la estructura del gen o en la evolución de su secuencia. Los resultados muestran una tasa de cambio heterogénea a lo largo del gen pero homogénea a lo largo de la filogenia. La tasa de cambio nucleotídico de *lab* se ha mantenido constante a pesar del cambio de posición. En la segunda parte del trabajo se han secuenciado dos regiones del genoma de *D. buzzatii*, una contiene los genes *lab* y *abdominal A* (*abdA*) y la otra incluye el gen *proboscipedia* (*pb*), y se ha comparado su organización génica con *D. melanogaster* y *D. pseudoobscura* para localizar con precisión los puntos de rotura. También se han comparado las secuencias no codificantes de estas regiones, se ha observado una elevada presencia de bloques conservados en los intrones y regiones adyacentes de los genes Hox. La comparación de los bloques conservados con las regiones reguladoras conocidas de los genes Hox (*lab*, *pb* y *abdA*) en *D. melanogaster* muestra que la posición y orientación de las regiones reguladoras está conservada entre las tres especies, con excepciones menores. Finalmente se analizó el patrón de expresión de los tres genes Hox en embriones y discos imaginales de *D. buzzatii*, *D. repleta*, *D. virilis*, y *D. melanogaster*, cuatro especies con diferentes organizaciones de los genes Hox. Las dos roturas estudiadas se produjeron mediante inversiones paracéntricas, con los puntos de rotura respetando las regiones reguladoras de ambos genes. De manera que las regiones reguladoras y patrones de expresión de los genes Hox adyacentes se han

conservado a pesar de las reorganizaciones. En *Drosophila* el complejo parece tener una estructura formada por módulos (que incluyen los genes y las regiones reguladoras) independientes, cuya agrupación no es necesaria para su correcto funcionamiento. La organización de estos genes es modular y su agrupación parece ser el resultado de la inercia filogenética más que de la necesidad funcional. El descubrimiento de más reorganizaciones en otros linajes y la importancia de la colinealidad temporal en algunos organismos sugieren que la causa funcional de la conservación de la organización genómica podría ser la colinealidad temporal. Las reorganizaciones serían consecuencia de la pérdida de colinealidad temporal en organismos con desarrollo rápido por modificación de la embriogénesis.

ABSTRACT

*Homeotic (Hox) genes code for transcription factors involved in the specification of segmental identity in the anteroposterior axis of the early metazoan embryo. These genes are usually clustered and arranged in the same order as they are expressed along the anteroposterior body axis. The conservation of this Hox gene organization along the phylogeny has suggested the existence of functional constraints. However, the partial disassembly of the *Caenorhabditis elegans* complex and the three splits observed in the *Drosophila* genus question whether this organization is an absolute necessity for proper function in some lineages. In this work, I analysed the genomic and functional consequences of the two splits present in *Drosophila buzzatii*, a member of the *repleta* species group. In the first part, the coding regions of the labial (lab) gene were cloned in *D. buzzatii* and *D. virilis*. The sequences of these two species were compared with that of *D. melanogaster* to test whether the change in position of lab in the *D. buzzatii* lineage produced any change in gene structure or sequence evolution. The results show that the substitution rate is heterogeneous along the gene but homogeneous along the phylogeny. The nucleotide substitution rate of lab has been constant in spite of the positional change. In the second part, two regions of the *D. buzzatii* genome have been sequenced, one including the lab y abdominal A (abdA) genes and the other containing the proboscipedia (pb) gene, and compared with the genic organization of *D. melanogaster* and *D. pseudoobscura* to precisely locate the breakpoints. The noncoding sequences of these regions have also been compared, and a high presence of conserved blocks has been observed in the introns and surrounding regions of Hox genes. The comparison of these conserved blocks with the known regulatory regions of the Hox genes (lab, pb and abdA) in *D. melanogaster* shows that the position and order of the regulatory regions is conserved between the three species, with minor exceptions. Finally the expression pattern of the three Hox genes has been analysed in embryos and imaginal discs of *D. buzzatii*, *D. repleta*, *D. virilis*, and *D. melanogaster*, four species with different Hox gene arrangements. The two splits took place through two paracentric inversions, with their breakpoints between the regulatory regions of adjacent genes. So that the regulatory regions and expression patterns of these Hox genes have been conserved in spite of the reorganizations. In *Drosophila* the Hox gene complex seems to be composed by independent modules (including the gene and its regulatory regions), whose association is not required for proper function. The organization of these genes is modular and their clustering seems the result of*

phylogenetic inertia more than of functional necessity. The discovery of more rearrangements in other lineages and the significance of temporal colinearity in some organisms suggest that the functional cause of the conservation of this genomic organization would be temporal colinearity. Rearrangements would be the consequence of the loss of temporal colinearity in organisms with a very rapid mode of embriogenesis.

1. INTRODUCCIÓN

Existe una diversidad asombrosa de organismos en la tierra, con una variedad morfológica y ecológica increíble. Todos los animales comparten un momento en su ciclo vital en el que no son más que un cigoto unicelular. Resulta sorprendente pensar en la gran variedad de formas con complejos planes corporales adultos que surgen de una única célula por multiplicación y diferenciación, mediante el proceso de desarrollo embrionario. ¿Cómo se forma la arquitectura tridimensional del organismo a partir de la información codificada en una molécula lineal de DNA? La Biología del Desarrollo estudia los mecanismos que median la formación de un organismo, desde que es una célula indiferenciada hasta que alcanza su forma adulta y definitiva. El interés principal de esta disciplina está en identificar y estudiar los genes que ayudan a transformar un cigoto unicelular en un adulto pluricelular. En los últimos años el progreso en la caracterización de estos genes ha sido extraordinario. En algunos casos se ha encontrado que un mismo gen puede iniciar diferentes procesos de desarrollo y, en otros casos, genes o grupos de genes son usados como módulos en el desarrollo de formas biológicas diversas. Para entender cómo se forma un organismo hemos de entender qué genes actúan en cada lugar y cómo se regula su expresión.

La información necesaria para generar el plan corporal de cada ser vivo está cifrada en su genoma, lo que permite que las características básicas de los individuos se transmitan de padres a hijos. Pero esta transmisión no es perfecta. La mutación introduce continuamente nuevas variantes en la población, desde cambios nucleotídicos o puntuales, pasando por inversiones, deleciones y duplicaciones parciales o totales. Estos cambios estables y heredables del material genético son la materia prima de la evolución. La selección natural actúa sobre ella escogiendo las variantes mejor adaptadas al ambiente. La mayor parte de las nuevas variantes introducidas por la mutación son rápidamente eliminadas, pero algunas sobreviven y pueden llegar a incorporarse a todos los individuos de la población. La Biología Evolutiva busca entender los cambios genéticos responsables de las diferencias en el fenotipo, descifrando por ejemplo que cambios han sido fijados por la selección natural y cuáles lo han sido por azar. Uno de sus mayores retos es identificar los cambios moleculares responsables de las diferencias fenotípicas evolutivamente significativas.

En los últimos años se ha afianzado una nueva disciplina que marca la intersección entre la Biología del Desarrollo y la Biología Evolutiva. Esta nueva disciplina, Biología Evolutiva del Desarrollo (Evo-Devo – Evolution and Development), busca entender cómo se originaron los diferentes planes corporales a lo

largo de la evolución (GILBERT 1997). Después de una primera época en la que ha sorprendido la elevada conservación de los genes y redes génicas involucradas en el desarrollo temprano de los metazoos, los estudios se centran ahora en describir cómo estas redes génicas conservadas dan lugar a diferencias morfológicas. ¿Qué genes y redes génicas están implicados? ¿Qué cambios en estos genes son los responsables de los cambios evolutivos? El objetivo de la Evo-Devo es desvelar la secuencia de cambios que llevan a la divergencia morfológica, y, eventualmente, el papel de la selección, la deriva y la contingencia histórica en los cambios macroevolutivos. El problema principal es entender cómo ocurrieron las grandes transiciones entre morfologías diferentes, si fueron grandes mutaciones revolucionarias o una acumulación gradual de mutaciones de pequeños efecto. ¿Qué cambios moleculares son responsables de los cambios fenotípicos? ¿Qué cambios son neutros y cuáles son responsables de la especiación y la diversificación morfológica? ¿Cómo se relacionan los procesos microevolutivos con las diferencias macroevolutivas? ¿Se corresponden los genes identificados por mutaciones que afectan al desarrollo en una especie con los genes responsables de las diferencias entre especies?

Estudios comparativos del desarrollo temprano han mostrado que cambios en el uso spatiotemporal de genes reguladores, así como en la especificidad de proteínas reguladoras, están correlacionados con importantes diferencias morfológicas entre especies filogenéticamente distantes. De todos modos, no se sabe cómo ocurren esos cambios en poblaciones naturales, ni si son consecuencia de un cambio único o de muchos cambios con efecto aditivo. El estudio comparativo del desarrollo en especies filogenéticamente cercanas promete ser decisivo en la resolución de estas cuestiones (GIBSON 1999; SIMPSON 2002).

1.1. LA UTILIDAD DE LOS ORGANISMOS MODELO

A pesar de la gran variedad de formas y complejidad, los animales son mucho más parecidos de lo que parece a simple vista. En 1822, el zoólogo francés Geoffroy St. Hilaire propuso una visión fantástica de los animales pasados y presentes según la cual vertebrados y artrópodos tienen un patrón corporal común pero con el eje dorso-ventral invertido, donde el cordón nervioso se extiende a lo largo de la línea media ventral en insectos y dorsal en vertebrados, mientras el corazón late en la parte dorsal en insectos y

en la región ventral en vertebrados (MÜLLER 1997). Evidencias genéticas y moleculares acumuladas durante los últimos 20 años en el campo de la Biología del Desarrollo han detectado similitudes encubiertas en los diversos patrones de desarrollo. Muchos órganos considerados análogos en diferentes organismos parecen estar controlados por los mismos sistemas de patrón génico, de manera que a nivel del control genético de su formación habría que considerarlos homólogos (HOLLEY *et al.* 1995). Así, el conocimiento de los genes que controlan, por ejemplo, el desarrollo temprano de embriones humanos se podría obtener del estudio detallado de organismos modelo como nemátodos, moscas o ratones (VERAKSA *et al.* 2000).

La mayoría de procesos biológicos son similares en muchos organismos, especialmente los procesos celulares. Por otro lado, cuanta más información se tiene sobre un organismo, más fácil es estudiarlo a otros niveles. La puesta a punto de técnicas conocidas sobre nuevas especies puede ser muy laboriosa, de manera que se tiende siempre a utilizar el organismo dónde más fácil sea abordar el problema. Por estas razones, la investigación se centra en unos pocos organismos que se convierten en “sistemas modelo”. Estas especies son sujetas a un intenso y profundo estudio, con la idea de que lo que es cierto para una especie es cierto para todas, o al menos para muchas otras. Los estudios que se están realizando con diferentes sistemas modelo son de gran importancia, no sólo para entender cómo funcionan estos organismos en particular, sino para extrapolar los resultados obtenidos a un modelo general.

Los principios fundamentales se reconocen sólo cuando los sucesos se estudian en diversos organismos, los cuales operan de forma diferente pero mostrando características comunes. El uso de organismos modelo ha sido muy fructífero, especialmente en cuanto al conocimiento general de los mecanismos genéticos existentes. Por otro lado, la comparación entre estos organismos modelo, alejados filogenéticamente, ha permitido iniciar el estudio de los grandes cambios macroevolutivos. Pero no es posible comprender la evolución de una especie estudiándola de forma aislada o comparándola con especies muy alejadas. Todas las especies forman parte de grupos de especies más o menos cercanas, y cada una solo puede ser entendida en el contexto histórico de grupos unidos en el pasado en un linaje único. Para aprender los principios aplicables a otros organismos es necesario poner en perspectiva evolutiva los conocimientos profundos que tenemos sobre una especie modelo (POWELL 1997). El estudio comparativo entre especies modelo y especies cercanas a estas permiten estudiar mejor que cambios son el motor de la evolución.

La comparación de secuencias de organismos muy alejados filogenéticamente permite identificar grandes homologías conservadas a lo largo de la evolución, pero su distancia dificulta la comparación y la descripción detallada de la función de las diferentes regiones del genoma. Un prerequisite esencial para hacer comparaciones entre grupos es comprender las bases de esa característica dentro de los grupos. La comparación de especies cercanas debería permitir cerrar este vacío, revelando elementos esenciales evolutivamente conservados además de regiones variables que pueden corresponder a elementos no esenciales o a elementos específicos de la especie. Así esta comparación debería ayudarnos a entender mejor la función de cada una de las partes que codifican para la proteína o de las secuencias nucleotídicas adyacentes, hasta ser capaces de predecir si los cambios que observamos en la secuencia se traducen o no en cambios en la estructura del producto génico y si estos cambios pueden o no suponer un cambio en su función o funcionalidad. En definitiva, predecir qué cambios nucleotídicos son responsables de los cambios fenotípicos tanto morfológicos como etológicos (SIMPSON 2002).

Para poder aplicar los conocimientos que se van adquiriendo en una especie modelo a otros organismos o hacer comparaciones entre especies modelo es importante tener una perspectiva evolutiva. Los estudios evolutivos requieren una buena base filogenética para poder interpretar los cambios detectados. La aplicación de métodos moleculares en la elaboración de filogenias en las últimas dos décadas a cambiado notablemente la filogenia de los metazoos. Este hecho ha influido en la interpretación de algunas comparaciones. Especialmente notable es el cambio de posición sufrido por los nemátodos que han pasado de ser considerados un grupo primitivo, rama basal a protóstomos y deuteróstomos, a ser un grupo hermano de los artrópodos (AGUINALDO *et al.* 1997).

Para referencia a lo largo del presente trabajo, se incluyen aquí dos árboles filogenéticos. El primero es una filogenia de los grandes grupos de metazoos (Figura 1). El segundo sitúa al género *Drosophila* dentro de los artrópodos, indicando también la posición de otros géneros mencionados en el presente trabajo (Figura 2).

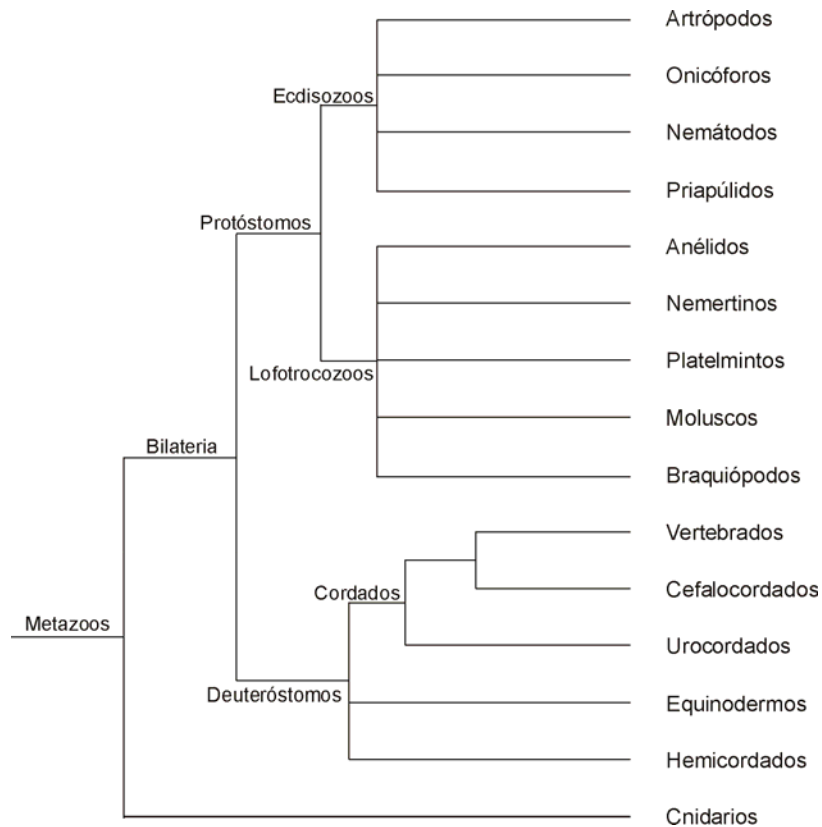


Figura 1. Filogenia de los Metazoos. Modificado de AGUINALDO *et al.* (1997), DE ROSA *et al.* (1999) MARTINEZ *et al.* (1999) y DANCHIN y PONTAROTTI (2004).

Figure 1. Phylogeny of the Metazoan. Modified from AGUINALDO *et al.* (1997), DE ROSA *et al.* (1999) MARTINEZ *et al.* (1999) and DANCHIN and PONTAROTTI (2004).

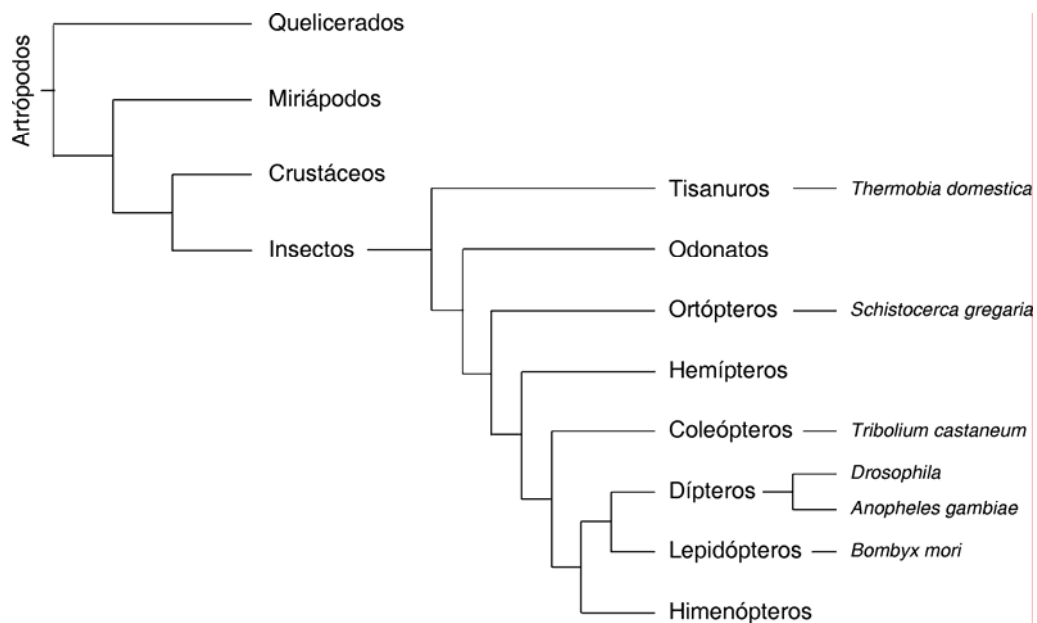


Figura 2. Filogenia de los principales grupos de artrópodos y de insectos, incluyendo la posición de las especies citadas en este trabajo. Modificado de HUGHES y KAUFMAN (2002) y POWERS *et al.* (2000).

Figure 2. Phylogeny of the main groups of arthropods and insects, including the position of species cited along this work. Modified from HUGHES and KAUFMAN (2002) and POWERS *et al.* (2000).

1.2. DROSOPHILA COMO MODELO

Drosophila es uno de los organismos modelo más utilizados, si no el que más. A principios del siglo XX jugó un papel muy importante en los primeros pasos de la genética. A mediados de siglo, fue un importante modelo para estudios empíricos en la síntesis de la teoría genética y evolutiva. En tiempos más recientes, se ha hecho muy popular como sistema modelo en Biología Molecular y del Desarrollo. De *Drosophila* hay información sobre ecología, biología de poblaciones, sistemática, comportamiento, genética, desarrollo y biología molecular. Ningún otro género ha sido tan estudiado a tantos niveles (POWELL 1997).

Hace ya 100 años que T.H. Morgan convirtió *Drosophila* en el modelo líder de la genética clásica. Morgan escogió esta especie por su corto ciclo vital, facilidad de cultivo y elevada fecundidad. En 1910 descubrió el primer mutante, *white eye*. En colaboración con sus estudiantes A.H. Sturtevant, C.B. Bridges y H.J. Muller establecieron la “teoría cromosómica de la herencia”. En 1913, A.H. Sturtevant construyó el primer mapa genético, que incluía seis loci del cromosoma X de *Drosophila melanogaster* y mostraba que los genes están organizados en un orden lineal. En 1914, Bridges, aprovechando la no-disyunción en hembras XXY, demostró que los genes están contenidos en los cromosomas y que no se trata de elementos heredables independientes. En 1918, Muller introdujo el uso de cromosomas balanceadores para el mantenimiento estable de mutaciones letales como heterocigotos. Una muestra y presentación de las fases tempranas de estos trabajos fue publicada por Morgan, Sturtevant, Bridges y Muller en 1915 en el libro *The Mechanism of Mendelian Heredity*. Sus trabajos, realizados con el único método experimental de contar la descendencia de cruces controlados de diferentes mutantes, le valió a Morgan el Premio Nobel de Fisiología o Medicina en 1933. Este premio supuso el primer reconocimiento de que la genética podía beneficiar a la especie humana.

En 1927, Muller mostró que las radiaciones ionizantes causan daño genético, y que por tanto los rayos X podían producir mutaciones, incluidos los reordenamientos cromosómicos, llamando la atención sobre su posible implicación en el riesgo de cáncer. Este descubrimiento le valió a Muller el Premio Nobel de Fisiología o Medicina en 1946, el segundo otorgado a estudios genéticos sobre *Drosophila*.

El descubrimiento de los cromosomas politénicos en las glándulas salivales, en 1933, puso las bases para la cartografía física de los genes. Curiosamente estos

cromosomas fueron descubiertos en la mosca *Bibio hortulanus*, por Heitz y Bauer, aunque rápidamente, en 1934, T.S. Painter publicó los primeros dibujos de cromosomas politénicos de *D. melanogaster*, incluyendo ya la posición de algunos genes. Los dibujos realizados por Bridges en 1935 y 1938 eran tan detallados que siguen utilizándose hoy en día, así como los dibujos de *Drosophila repleta* realizados por WHARTON (1942).

La primera clonación de un fragmento de DNA eucariótico se produjo en *Drosophila* así como su localización en una posición específica del cromosoma. La técnica de la hibridación *in situ* de ácidos nucleicos sobre cromosomas politénicos (PARDUE *et al.* 1970) permite localizar las secuencias homólogas a las contenidas en clones recombinantes, procedentes de la misma o de diferentes especies, proporcionando datos sobre la localización de los genes hasta una resolución de unas 50 kb (HARTL *et al.* 1994). Esta técnica ha permitido elaborar mapas físicos detallados (HARTL *et al.* 1994; HOSKINS *et al.* 2000; GONZÁLEZ *et al.* 2005) así como comparar la organización molecular de los cromosomas de diferentes especies de *Drosophila* (SEGARRA y AGUADE 1992; SEGARRA *et al.* 1995; SEGARRA *et al.* 1996; VIEIRA *et al.* 1997a; VIEIRA *et al.* 1997b; RANZ *et al.* 2001; GONZÁLEZ *et al.* 2002).

Aunque las primeras clonaciones se realizaban al azar, a principios de 1975 ya se habían generado genotecas que representan la totalidad del genoma y se habían desarrollado métodos de hibridación de colonias que permitían rastrear con clones que contienen secuencias específicas (BURTIS *et al.* 2003). El primer “recorrido” cromosómico con fragmentos clonados en fagos lambda cubría más de 200 kb a finales de 1978. Una inversión que ligaba esta región al Complejo Bithorax permitió el primer clonaje posicional de un gen, *Ultrabithorax* (*Ubx*), a principios de 1979 (BENDER *et al.* 1983a; 1983b; RUBIN y LEWIS 2000).

Otro avance importante en la manipulación del genoma se realizó a finales de 1981 con el desarrollo de los métodos para obtener moscas transgénicas. Así se consiguió el primer rescate de un fenotipo mutante en un animal por transferencia génica (RUBIN y SPRADLING 1982). Esta técnica se fue aplicando para diversos usos, como la detección de secuencias *cis*-reguladoras (*enhancers*) (O'KANE y GEHRING 1987) o de genes regulados durante el desarrollo (WILSON *et al.* 1989). Se desarrolló también un sistema versátil de sobreexpresión génica basado en el sistema de levadura Gal4-UAS (Upstream Activating Sequence – UAS) (BRAND *et al.* 1985). También se han desarrollado métodos para llevar a cabo recombinación en lugares específicos,

como el sistema Flp-FRP (XU y RUBIN 1993; GOLIC y GOLIC 1996), que permiten la generación de inversiones o la delección de regiones determinadas. En los últimos años han proliferado los usos de las técnicas de transgénesis gracias a ingenios y variaciones sobre estas técnicas, y se están desarrollando también nuevos vectores para evitar las limitaciones de los sistemas actuales como son el tamaño de las secuencias que se pueden insertar o los efectos de posición (VENKEN y BELLEN 2005). Estas herramientas de manipulación genética se han beneficiado del mantenimiento sencillo y económico de *Drosophila*, que permite mantener de forma rutinaria gran cantidad de cepas “salvajes” o mutantes, tanto espontáneas como inducidas.

Los avances que se han ido produciendo en el conocimiento de la Genética y en las herramientas de manipulación genética van siendo rápidamente aplicados a otros campos como la Biología del Desarrollo. Los trabajos de Lewis (ver apartado 1.6), inicialmente encaminados a contrastar una teoría sobre la duplicación génica, identificaron un grupo de genes responsables de asignar identidad a regiones del tórax y abdomen y de controlar por tanto el desarrollo posterior de órganos especializados en segmentos determinados. En 1980, C. Nusslein-Volhard y E. Wieschaus aplicaron por primera vez a un animal el rastreo mutacional sistemático para identificar genes involucrados en un proceso fundamental, este tipo de estudio sólo se había llevado a cabo con microorganismos. En sus estudios utilizaron la mosca *D. melanogaster* como sistema experimental y fueron capaces de identificar y clasificar un pequeño número de genes clave en la determinación del patrón corporal y la formación de los segmentos. Los principios que descubrieron y describieron para *Drosophila* son aplicables también a organismos “superiores” incluyendo al hombre, y fueron claves para entender muchas malformaciones congénitas. La importancia del estudio de *Drosophila* para la salud humana fue nuevamente reconocida con la concesión del premio Nobel de Medicina y Fisiología a Edward Lewis, Christiane Nusslein-Volhard y Eric Wieschaus en 1995.

La conservación de procesos biológicos de mosca a humanos extiende la influencia de *Drosophila* en la salud humana. Cuando se identifica en *Drosophila* un gen homólogo a un gen de mamífero importante pero poco conocido, el conjunto de técnicas genéticas desarrolladas en *Drosophila* puede ayudar a caracterizarlo. La identificación y caracterización de las redes génicas en *Drosophila* puede elucidar las redes homólogas en mamíferos. Estos métodos se han aplicado para identificar redes génicas

involucradas en cáncer y ciclo celular (KORNBERG y KRASNOW 2000). Actualmente *Drosophila* sigue siendo un modelo clave tanto para la investigación básica como aplicada. Es muy utilizada en Desarrollo, por ejemplo como modelo de la red genética del desarrollo ocular (WAWERSIK y MAAS 2000). Es utilizada también en investigación sobre envejecimiento (ZOU *et al.* 2000; HELFAND y INOUE 2003), en enfermedades humanas como cáncer (POTTER *et al.* 2000) o Parkinson (FEANY y BENDER 2000), en estudios de susceptibilidad y efectos de las drogas adictivas (HEBERLEIN 2000; GUARNIERI y HEBERLEIN 2003; WOLF y HEBERLEIN 2003), y un largo etcétera.

Las características e importancia histórica de *D. melanogaster* la han llevado a ser uno de los organismos mejor caracterizados, hasta convertirse en el tercer eucariota y el segundo metazoo con el genoma secuenciado (ADAMS *et al.* 2000), después de la levadura *Saccharomyces cerevisiae* (GOFFEAU *et al.* 1996) y el nemátodo *Caenorhabditis elegans* (TCES 1998). Actualmente *Drosophila* es también un modelo fundamental en genómica y genómica comparada (ver apartado 1.4).

El género *Drosophila* contiene alrededor de 2000 especies conocidas, aunque seguro hay más especies en regiones aún poco estudiadas como el África Subsahariana, el Sudeste Asiático y las islas del Pacífico Sur (POWELL 1997). Análisis filogenéticos

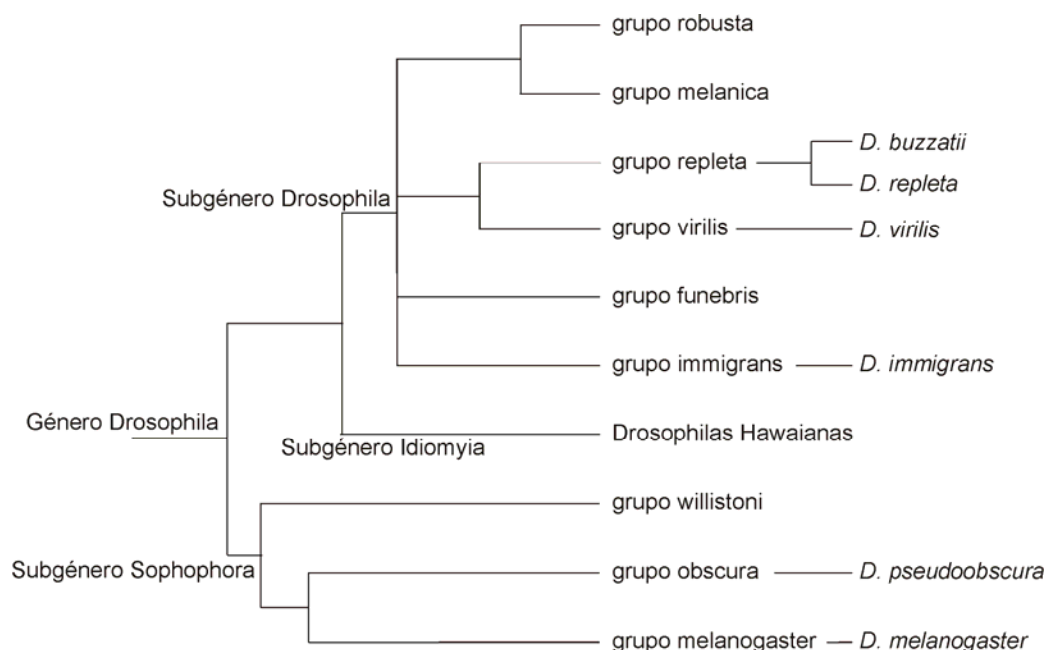


Figura 3. Árbol filogenético de los principales grupos de *Drosophila*. Se muestra la posición de las especies utilizadas en este trabajo. Modificado de POWELL (1997) y Tamura *et al.* (2004).

Figure 3. Phylogenetic tree showing the main groups of *Drosophila*. The position of the species used in this work is shown. Modified from POWELL (1997) and Tamura *et al.* (2004).

indican la existencia de dos linajes principales dentro del género *Drosophila*, los cuales divergieron hace entre 40 y 62 millones de años (Figura 3) (POWELL y DESALLE 1995; RUSSO *et al.* 1995; TAMURA *et al.* 2004). Uno de los linajes dió lugar al subgénero *Sophophora* el cual contiene unas 300 especies, entre ellas se encuentra *D. melanogaster*. El segundo linaje dió lugar a los subgéneros *Drosophila* y *Idiomya* (*Drosophilas* Hawaianas), que incluyen unas 700 y 375 especies descritas respectivamente (POWELL 1997) (<http://taxodros.unizh.ch>).

1.3. DROSOPHILA COMO MODELO EN BIOLOGÍA DEL DESARROLLO

1.3.1. Ciclo Vital

Como todos los insectos holometábolos *Drosophila* ocupa dos hábitats durante su ciclo vital (Figura 4). Las hembras ponen sus huevos en un sustrato blando apto

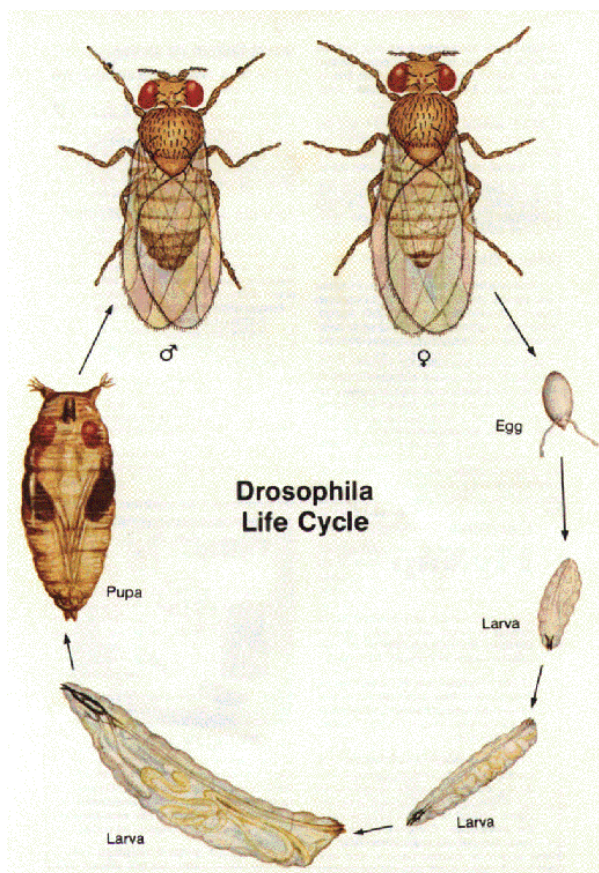


Figura 4. Ciclo vital de *Drosophila*.

Figure 4. Drosophila life cycle.

para el desarrollo de las larvas. Después de la embriogénesis, el huevo eclosiona dando lugar a una larva de primer estadio que pasará su tiempo cavando el sustrato, comiendo casi constantemente. Después de dos mudas, la larva dejará de alimentarse y buscará un lugar seco en el que pupar. La duración del período de embriogénesis y el larvario varía de especie a especie y con factores ambientales tales como la temperatura. En la pupa se produce la histólisis de los órganos y tejidos larvarios y se forman las estructuras adultas a partir de las estructuras y tejidos imaginales de la larva. El proceso de pupación dura al menos cuatro días.

Debido a que un descenso en la temperatura disminuye la velocidad del desarrollo, especies que habitan regiones templadas tienen, generalmente, tiempos de desarrollo más largos que las especies tropicales. Concretamente, *D. melanogaster* requiere un mínimo de 11 días entre huevo y huevo en condiciones ideales de laboratorio, es el tiempo de desarrollo más corto conocido para un Drosophilido (POWELL 1997).

Cuando eclosiona el adulto, éste habita un ambiente muy diferente al de la larva. Son voladores potentes que pasan su tiempo buscando comida y parejas. Generalmente machos y hembras se encuentran en las zonas de alimentación, ya que no se conocen feromonas de largo alcance en *Drosophila*. Además los adultos a menudo se alimentan del mismo sustrato que las larvas, aunque puedan utilizar una variedad más amplia (CARSON 1971). En general los drosophilidos se alimentan de material vegetal y fúngico en descomposición. La nutrición de muchas especies se basa en los hongos y bacterias que descomponen estos materiales (POWELL 1997).

1.3.1.1. Embriogénesis (de cigoto a larva)

Después de la fecundación, las primeras divisiones del cigoto, realizadas de forma sincrónica, darán lugar a un sincitio. A medida que se producen las divisiones, los núcleos migran hacia la superficie donde formarán una monocapa (blastodermo sincital). Después de trece divisiones, extensiones de la membrana se introducen entre los núcleos formando células individuales (blastodermo celular). La región ventral del blastodermo celular forma la llamada banda germinal, que dará lugar a las regiones metaméricas del embrión. A pesar de que toda la superficie del embrión parece homogénea, los futuros metámeros ya están determinados por los patrones de expresión de los genes de formación del patrón (ver apartado 1.3.2).

La gastrulación, o formación de las capas germinales, empieza con la invaginación del surco ventral en la línea media de la banda germinal. La parte central de esta invaginación dará lugar al mesodermo y al cordón nervioso ventral. En el extremo anterior del surco se forma una estria transversal que constituirá el primordio endodérmico anterior. En la parte posterior tiene lugar una compleja remodelación, asociada a la elongación de la banda germinal, que formará el primordio endodérmico posterior (Figura 5).

A diferencia de otros insectos en los que los diferentes segmentos se forman por proliferación a partir de un primordio de la banda germinal (insectos de banda germinal

corta), en *Drosophila* todos los segmentos han quedado ya determinados en el blastodermo celular (banda germinal larga). En este caso, el proceso de elongación de la banda germinal no implica proliferación celular, sino que consiste en una reestructuración de ésta.

La elongación de la banda germinal comporta cambios drásticos en el ectodermo a través de los cuales se hace evidente el patrón corporal. Durante este proceso aparece el surco cefálico que delimitará el procefalon de la banda germinal metamérica. Posteriormente tiene lugar la retracción de la banda germinal, durante la cual se fusionan los primordios endodérmicos anterior y posterior y se forman las gónadas. En este momento empiezan los complejos movimientos morfogénéticos de retracción de la región cefálica (head involution) que darán lugar a la cabeza de la larva.

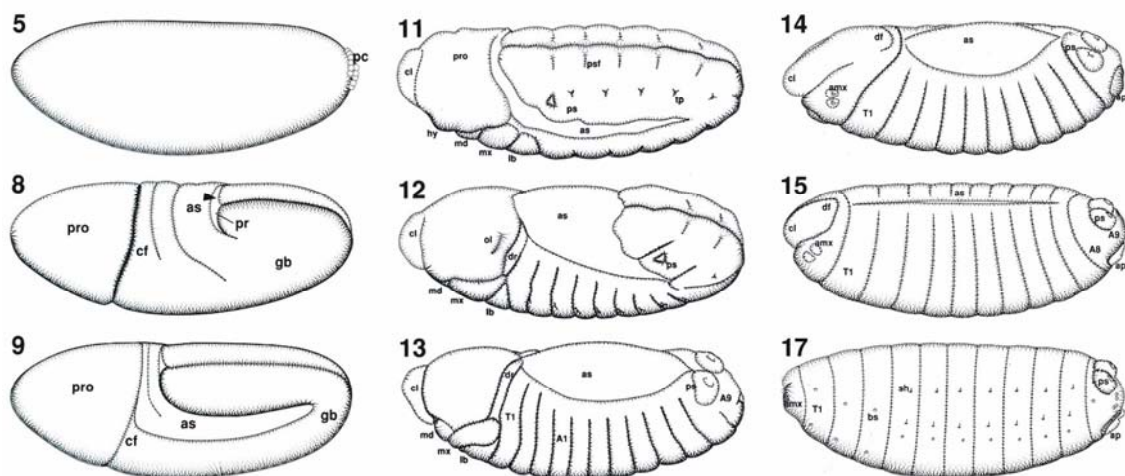


Figura 5. Fases del desarrollo embrionario de *Drosophila*. Los embriones se muestran con la región anterior en la izquierda y la dorsal arriba. El número sobre cada embrión indica el estadio que representa. Tomado de HARTENSTEIN (1993).

Figure 5. Phases of embryonic development in *Drosophila*. The embryos are shown with anterior to the left and posterior up. The number over each embryo corresponds to the stage. Taken from HARTENSTEIN (1993).

1.3.1.2. Metamorfosis (de larva a imago)

Durante el primer estadio larvario, una fracción de células se separa y se mantiene diploide, en vez de poliploide como la mayoría de células somáticas de la larva. La función de estas células será formar los tejidos del imago, insecto adulto. En la cavidad larval, la mayoría de estas células forman pequeñas bolsas planas y redondeadas llamadas discos imaginales, otras células quedan aisladas en el interior de determinados órganos (Figura 6). Durante la metamorfosis, los discos imaginales se evaginarán y expandirán dando lugar a la cutícula del imago (BATE *et al.* 1993).

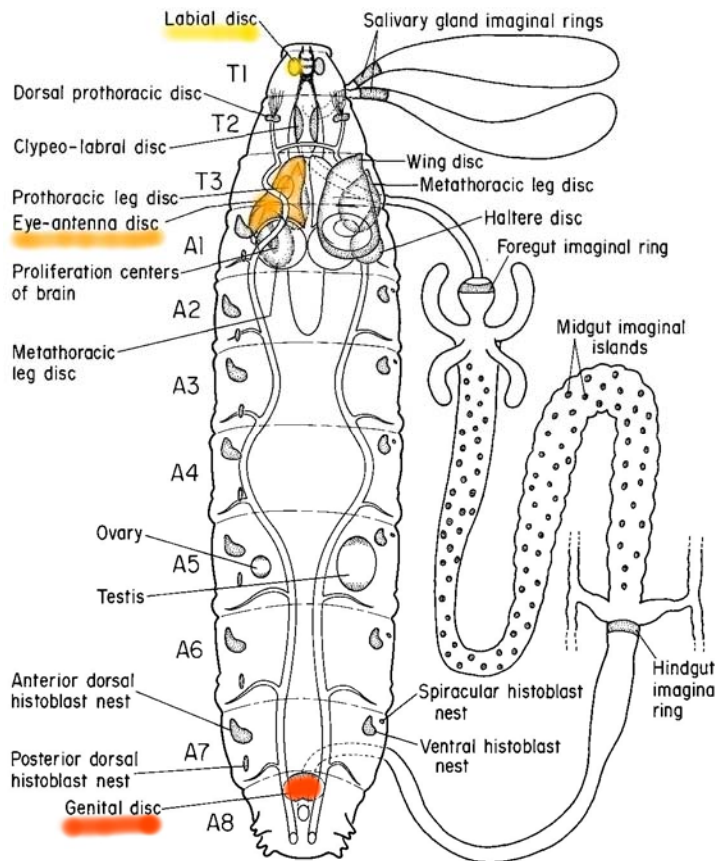


Figura 6. Discos y tejidos imaginales de *Drosophila*. En color se marcan los discos donde se expresan los genes estudiados en este trabajo: disco labial en amarillo, disco de ojo-antena en naranja y disco genital en rojo. Modificado de LAWRENCE (1992).

Figure 6. Imaginal discs and tissues of *Drosophila*. Imaginal discs where the studied genes show expression are in color: labial disc in yellow, eye-antennal disc in orange and genital disc in red. Modified from LAWRENCE (1992).

Muchos de los genes que se expresan durante el desarrollo embrionario muestran un segundo pico de expresión durante el período pupal, mientras aquellos que se expresan en la larva se vuelven a expresar en el adulto (ARBEITMAN *et al.* 2002). A pesar de las diferencias morfológicas entre estadios de desarrollo, diferentes “edades” muestran similitudes moleculares. Los genes Homeóticos (ver apartado 1.6) se encuentran en ese grupo de genes que muestran un pico durante el desarrollo embrionario y un segundo pico de expresión en la pupa, concretamente en los discos imaginales.

1.3.2. Genética del Desarrollo Embrionario

Los ejes corporales quedan preestablecidos ya en el oocito dentro de la cámara ovárica debido al aporte asimétrico de transcritos y proteínas de origen materno. Esta primera asimetría activará diferentes cascadas génicas encargadas de la diferenciación anteroposterior, dorsoventral y terminal. La cascada de interacciones entre factores de transcripción que definen la organización anteroposterior del embrión de *Drosophila* es una de las jerarquías transcripcionales más investigadas (Figura 12). Los trabajos de Nusslein-Volhard y Wieschaus (ver apartado 1.2) fueron capaces de identificar y

clasificar un pequeño número de genes clave en la determinación del patrón corporal y la formación de los segmentos. Su investigación generó el modelo de tres tipos de genes que controlan la subdivisión del embrión en desarrollo: los genes gap, los genes pair-rule, y los genes de polaridad segmental.

En *Drosophila* el núcleo del cigoto es transcripcionalmente inactivo hasta el ciclo 10, cuando las interfases mitóticas empiezan a alargarse se inicia la transcripción génica. En un par de horas, se ha establecido una huella segmental bastante completa a través de los patrones de expresión de los genes de segmentación. Características morfológicas repetidas segmentalmente aparecen en pocas horas, y se completan en menos de un día (NASIADKA *et al.* 2002).

La jerarquía de genes de segmentación empieza con la acción de los genes suministrados de forma materna. Los productos de los genes maternos combinan multitud de mecanismos moleculares para establecer los gradientes de tres factores de transcripción (Bicoid, Hunchback y Caudal) a lo largo del eje anteroposterior del embrión sincital. Durante este mismo período, otros grupos de genes se encargarán de la señalización dorsoventral y terminal del embrión (NASIADKA *et al.* 2002).

Los promotores de los primeros genes expresados por el cigoto, los genes gap, responden a diferentes concentraciones y combinaciones de los tres factores de transcripción maternos. Así cada gen gap se expresa en una banda amplia, que incluye varios de los futuros segmentos, situada en una región diferente del eje del embrión. La mayoría de proteínas codificadas por estos genes son factores de transcripción del tipo dedos de zinc (NASIADKA *et al.* 2002). Al final del blastodermo sincital se inicia la transcripción de los genes pair-rule, los cuales presentan los primeros patrones de expresión con dominios periódicos. De nuevo, los promotores de estos genes responden a las combinaciones y niveles de los factores de transcripción expresados antes, es decir, genes maternos y genes gap. Las siete bandas en segmentos alternos que conforman sus patrones de expresión se refinan gracias a complejas interacciones entre los mismos genes pair-rule. Finalmente, la acción combinatoria de los factores de transcripción pair-rule actúa sobre los promotores de los genes de polaridad segmental, los cuales se expresan en 14 bandas, una por segmento. Estos genes establecen la polaridad del patrón de elementos de cada futuro segmento y actúan para mantener el destino celular a lo largo del resto del desarrollo. Así como la mayor parte de genes gap y pair-rule codifican factores de transcripción, sólo una pequeña parte de los productos de

polaridad segmental son de este tipo. Esta diferencia se debe principalmente a que cuando estos genes se expresan las células ya se han formado, y por tanto es necesario activar, entre otros, los mecanismos de comunicación celular para enviar y recibir señales entre células (NASIADKA *et al.* 2002) lo que permitirá el desarrollo coordinado de las diferentes partes.

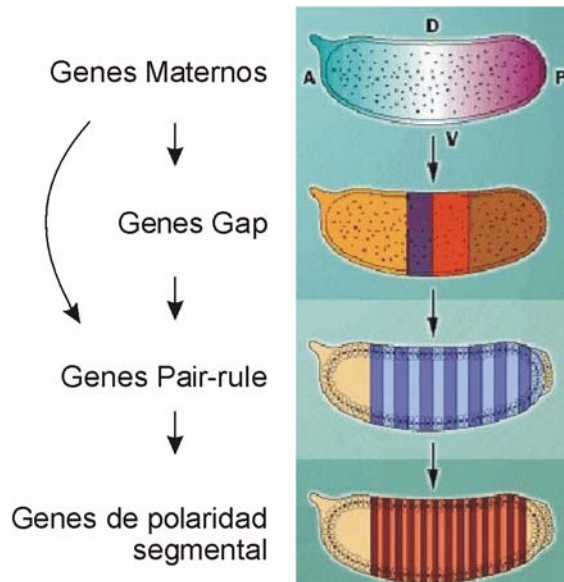


Figura 7. Cascada de activación de genes de segmentación. Los embriones están orientados igual que en la Figura 5. A, anterior; P, posterior; D, dorsal, y V, ventral.

Figure 7. Segmentation genes activation cascade. Embryos are shown as in Figure 5. A, anterior; P, posterior; D, dorsal, y V, ventral.

Así se distinguen cuatro tipos de genes que intervienen en la formación de los segmentos: los genes maternos, los genes gap, los genes pair-rule y los genes de polaridad segmental. La acción de estos cuatro grupos de genes ocurre en una cascada de acontecimientos dependientes encargados de definir la división segmental, es decir, globalmente determinan el número correcto de segmentos en cabeza, tórax y abdomen. Un quinto grupo de genes, los genes Homeóticos (ver apartado 1.6), son necesarios no para formar segmentos sino para asignar una identidad a los segmentos en formación. Estos genes interpretan la información posicional establecida y actúan como interruptores que determinan las identidades de los segmentos.

1.4. DROSOPHILA COMO MODELO EN GENÓMICA

Los datos genómicos sin ordenar son simplemente una retahíla de As, Ts, Gs y Cs. En este estado, los datos de secuencia aportan muy poco conocimiento biológico. Para entender el legado genético de un organismo debemos interpretar su secuencia genómica, traduciendo la información que contiene en forma molecular en anotaciones

leíbles por el hombre (LEWIS *et al.* 2002). El proceso de anotación es la tarea de anclar en la secuencia de ácidos nucleicos notaciones explicativas de sus características o función: transcrito, exón, promotor, elemento transponible, región reguladora, etc. La anotación funcional de las secuencias genómicas de los eucariotas es uno de los principales retos de la biología moderna.

El pequeño tamaño del genoma de *Drosophila* con 176 Megabases (Mb) (HOSKINS *et al.* 2002), la ha convertido en un modelo fundamental en Genómica, disciplina centrada en el estudio de la secuencia, estructura y función del genoma. Aprender cómo identificar, exponer, interrogar e interpretar las características del genoma en un organismo bien caracterizado como *Drosophila*, es crucial para entender los genomas de organismos más complejos, incluyendo humanos. Añadir anotaciones a la secuencia de un genoma de forma rigurosa y consistente es un prerequisite para el uso eficiente de esa secuencia en la investigación biológica (MISRA *et al.* 2002).

La secuencia anotada del genoma de *D. melanogaster* es resultado de la primera aplicación del método “shotgun” para la secuenciación de un genoma animal (ADAMS *et al.* 2000), y supone un modelo para la anotación a gran escala de las secuencias genómicas. En los últimos años se han desarrollado multitud de programas informáticos encaminados a identificar los genes y las regiones *cis*-reguladoras que controlan su expresión, con el objetivo final de automatizar el proceso de anotación de los genomas.

1.4.1. El Número de Genes

Actualmente se define un gen, en términos moleculares, como “un segmento cromosómico completo responsable de producir un producto funcional” (SNYDER y GERSTEIN 2003). Esta definición tiene varios componentes lógicos: la inclusión de regiones codificantes y reguladoras, la expresión del producto génico y la funcionalidad de este. A pesar de que esta definición implica unos criterios concretos para identificar los genes, su aplicación es aún problemática (SNYDER y GERSTEIN 2003) y no disponemos todavía de mecanismos eficientes para predecir de forma directa los genes de la secuencia nucleotídica. Incluso en un organismo tan estudiado como *D. melanogaster* y cuatro años después de la publicación de su genoma (ADAMS *et al.* 2000) aún no hay consenso ni siquiera sobre el número de genes codificantes que contiene (HILD *et al.* 2003; OLIVER y LEBLANC 2003).

La primera anotación del genoma de *D. melanogaster* detectaba 13.601 genes codificantes (ADAMS *et al.* 2000). En la notación del “Release 3” el número disminuyó

ligeramente, 13.379 genes, a pesar de presentar una secuencia más completa e incluir la heterocromatina, ausente en el ensamblaje inicial (MISRA *et al.* 2002). Aunque entre el Release 1 y el Release 3 el número de genes anotados ha variado poco, sí ha variado la estructura de muchos de ellos. Con una predicción menos restrictiva usando Fgenesh, análisis de la expresión por “microarrays” y validación por RT-PCR e hibridación *in situ*, HILD *et al.* (2003) identifican al menos 2,000 nuevos genes en el genoma de *D. melanogaster* no incluidos en la anotación del Release 3. Aunque se ha liberado un Release 4 de la eucromatina, donde se han cerrado 21 gaps de secuencia, la anotación no ha variado respecto al Release 3 (<http://www.fruitfly.org/annot/release4.html>).

1.4.2. Las Secuencias Reguladoras

A pesar de que la mayor parte de la secuencia genómica de los metazoos está formada por DNA no codificante, poco se sabe sobre la función o constricciones evolutivas que actúan sobre estas enigmáticas secuencias. Aunque es probable que algunas porciones de esta fracción del genoma demuestren ausencia de función biológica medible, está ampliamente asumido que gran parte de nuestra complejidad genética es debida a sofisticadas señales reguladoras no codificantes que determinan cuando, dónde y cuánto un gen desarrolla su actividad transcripcional. A pesar de la importancia de las secuencias no codificantes en la regulación génica, nuestra habilidad para identificar y caracterizar computacionalmente estos elementos es muy limitada. La reciente disponibilidad de datos de secuencias de genomas completos ha generado una oportunidad sin precedentes, y un reto, para buscar el significado del DNA no codificante (CLARK 2001).

En los organismos multicelulares, la modulación de la expresión génica se produce a través de la interacción de las proteínas reguladoras (factores de transcripción) y las regiones específicas del DNA (secuencias *cis*-reguladoras), y son los cambios en la expresión génica los responsables de gran parte de la variación fenotípica (MARKSTEIN *et al.* 2002; DE MEAUX *et al.* 2005). Las regiones *cis*-reguladoras están altamente estructuradas y exhiben una elevada organización modular formada por aisladores (insulators o boundary elements), silenciadores (silencers), intensificadores (enhancers), y promotores basales (core promoters) (LEVINE y TJIAN 2003).

Se define un módulo *cis*-regulador (*Cis-Regulatory Module* – CRM) como aquel módulo funcional que produce un aspecto discreto del patrón global de expresión de un

gen cuando es situado al lado de un promotor basal (ARNONE y DAVIDSON 1997; BERMAN *et al.* 2002; WRAY *et al.* 2003). Los CRM son segmentos de DNA (desde unos cientos de pares de bases a una kilobase) que regulan la transcripción génica en los eucariotas superiores. Contienen uno o varios lugares de unión para diferentes factores de transcripción, que actúan de forma cooperativa para activar o reprimir la transcripción del gen diana. Por tanto el término CRM incluye tanto silenciadores (BRAND *et al.* 1985) como intensificadores, ya que su efecto depende del contexto, es decir, su función será una u otra según la región del organismo y el momento en que lo observemos, y de las características del promotor basal. Los CRM pueden estar a distancias diversas del promotor y en cualquier orientación, lo que hace difícil en algunas ocasiones predecir sobre qué promotor y por tanto sobre qué gen actúan. Algunos CRM son específicos de un promotor, es decir, actúan por ejemplo sobre un promotor con caja TATA o con DPE (Downstream Promoter Element), pero otros son promiscuos, actúan sobre cualquier tipo de promotor (BUTLER y KADONAGA 2002). En estos casos la presencia de un aislador evita la acción promiscua del CRM, aislando a uno de los posibles promotores. En otros casos los aisladores se sitúan entre CRM diferentes para garantizar su acción independiente.

Varios trabajos han definido la localización de diferentes regiones reguladoras basándose en delecciones, inserciones o inversiones en mutantes y en el análisis del potencial efecto regulador de fragmentos determinados sobre el promotor lacZ (O'KANE y GEHRING 1987; CHOUINARD y KAUFMAN 1991; GINDHART y KAUFMAN 1995; GINDHART *et al.* 1995; KAPOUN y KAUFMAN 1995; KEYS *et al.* 2005). La mayoría de estos estudios son relativamente groseros en cuanto a la localización y el tamaño de las regiones reguladoras. Los estudios detallados sobre secuencias mínimas necesarias son escasos debido a que la disección funcional de los módulos *cis*-reguladores es muy laboriosa (LUDWIG *et al.* 2000; MICHELSON 2002). Actualmente se está realizando un gran esfuerzo para desarrollar métodos para la identificación y análisis bioinformático de estas regiones a partir de la secuencia.

La agrupación de lugares de unión de uno o varios factores de transcripción puede ser utilizada para la detección de CRM (BERMAN *et al.* 2002; MARKSTEIN *et al.* 2002; LIFANOV *et al.* 2003; BERMAN *et al.* 2004). Pero la ausencia de datos sobre los lugares de unión de la mayoría de factores de transcripción, representa actualmente una

gran limitación para la aplicación generalizada de este método. Como los CRM están sujetos a constricciones funcionales para mantener la expresión del gen diana su secuencia evoluciona más despacio que las secuencias no funcionales. Por lo tanto, la conservación de secuencias no codificantes entre especies filogenéticamente distantes (huella filogenética) puede utilizarse como guía para identificar secuencias reguladoras (TAGLE *et al.* 1988). Varios estudios recientes avalan el uso del análisis comparativo de secuencias y la caracterización de las secuencias no codificantes conservadas (CNS – Conserved Noncoding Sequences) como una aproximación útil para la detección de posibles CRM en *Drosophila* (BERGMAN y KREITMAN 2001; BERGMAN *et al.* 2002), mamíferos (DERMITZAKIS *et al.* 2002; BOFFELLI *et al.* 2003; COOPER y SIDOW 2003; NOBREGA *et al.* 2003; BOFFELLI *et al.* 2004; FRAZER *et al.* 2004b; GAFFNEY y KEIGHTLEY 2004; MARTIN *et al.* 2004), vertebrados (SANTINI *et al.* 2003) y plantas (GUO y MOOSE 2003; BUCHANAN *et al.* 2004). En *Drosophila*, la interpretación de los CNS como probables regiones reguladoras está respaldada por la elevada tasa de sustitución nucleotídica neutra (MORIYAMA y GOJOBORI 1992) y la elevada tasa intrínseca de pérdida de DNA (PETROV *et al.* 1996; SINGH y PETROV 2004). No se espera que las secuencia no codificantes estén conservadas entre especies poco relacionadas a no ser que estén sometidas a constricciones funcionales (O'NEIL y BELOTE 1992).

1.4.3. Genómica Comparada

La comparación de secuencias se está utilizando de forma generalizada para la validación de las predicciones de genes (HUDEK *et al.* 2003). Especialmente la comparación de secuencias de organismos cercanos promete ser una herramienta fundamental para la identificación tanto de genes codificantes como de regiones *cis*-reguladoras. En los últimos años se están desarrollando y mejorando los métodos informáticos para el alineamiento y visualización de secuencias largas como VISTA (MAYOR *et al.* 2000; FRAZER *et al.* 2004a) o PipMaker (SCHWARTZ *et al.* 2000). El género *Drosophila* proporciona un fabuloso sistema modelo para la genómica comparativa y el desarrollo de métodos informáticos (CLARK *et al.* 2003).

En el género *Drosophila* el genoma tiene habitualmente entre unas 100 y 400 Megabases (Mb). Unos tamaños bastante típicos dentro de los Dípteros (por ejemplo el genoma de *Anopheles gambiae* tiene 260 Mb) pero muy pequeño en comparación con

otros eucariotas multicelulares, ya que representan entre el 6% y el 10% del genoma típico de un mamífero (3.000 Mb) (POWELL 1997). Por ejemplo el genoma de *D. melanogaster* contiene 176 Mb (HOSKINS *et al.* 2002), el de *D. simulans* unas 119 Mb (POWELL 1997), el de *D. hydei* y las especies del grupo *repleta* alrededor de 220 Mb (LAIRD 1973; SCHULZE y LEE 1986; RANZ *et al.* 2001), y el de *D. virilis* 313 Mb (HARTL y LOZOVSKAYA 1995). Además el tiempo de divergencia entre diferentes especies o grupos del género *Drosophila* es equiparable a la divergencia entre especies de primates o de mamíferos (NEI *et al.* 2001; GLAZKO y NEI 2003; TAMURA *et al.* 2004).

Las diferencias entre especies se deben tanto al contenido en eucromatina como en heterocromatina: *D. melanogaster* contiene 116.9 Mb (CELNIKER *et al.* 2002) y 58.8 Mb (HOSKINS *et al.* 2002) respectivamente, mientras *D. virilis* contiene ~150 y ~163 Mb (HARTL y LOZOVSKAYA 1995). La variación en el DNA altamente repetitivo explica las diferencias en la heterocromatina, mientras que en la eucromatina se ha propuesto que la diferencia podría ser reflejo de la diferente longitud de los intrones (MORIYAMA *et al.* 1998) y del DNA repetitivo disperso (elementos móviles).

Actualmente se ha secuenciado una segunda especie de *Drosophila*, *D. pseudoobscura* (RICHARDS *et al.* 2005), y se están secuenciando otras 10 (BEGUN y LANGLEY 2003; CLARK *et al.* 2003). Ya están disponibles los ensamblajes provisionales de *D. ananassae*, *D. erecta*, *D. simulans* y *D. yakuba* (subgénero *Sophophora*) y de *D. virilis* y *D. mojavensis* (subgénero *Drosophila*) (<http://rana.lbl.gov/drosophila/multipleflies.html>).

1.5. EVOLUCIÓN CROMOSÓMICA

En las especies del género *Drosophila* el cariotipo es muy variable por lo que refiere al número haploide de cromosomas que está habitualmente entre 3 y 6. Gracias a la información de mapas citológicos y de ligamiento, MULLER (1940) formuló la hipótesis de que el cariotipo ancestral del género estaba formado por seis elementos cromosómicos que se han conservado como cromosomas o brazos cromosómicos en todo el género. Estableció homologías entre las diferentes especies de *Drosophila* elaborando una nomenclatura que permite relacionar los cariotipos. Así los elementos

A, B, C, D y E corresponden a los cinco brazos cromosómicos y el elemento F al *dot* o cromosoma puntiforme (Tabla 1) (MULLER 1940). La conservación del contenido génico de los elementos cromosómicos parece extenderse más allá del género *Drosophila*, habiéndose encontrado en otros dípteros como *Lucilia cuprina* o *Musca domestica* (WELLER y FOSTER 1993). HILLIKER y TRULLIS-COULTER (1987) mostraron que alterar esta agrupación cromosómica no tiene efecto sobre viabilidad o esterilidad, indicando la ausencia de significado funcional de la sinténia a nivel cromosómico. En otros dípteros más alejados como *Anopheles*, cuyo genoma ha sido secuenciado y comparado detalladamente con el de *D. melanogaster*, el contenido génico de los elementos está escasamente conservado (ZDOBNOV *et al.* 2002). Por otro lado, estudios de cartografía física, mediante hibridación *in situ* sobre cromosomas politénicos, han demostrado que a pesar de la gran conservación en el contenido génico de los brazos cromosómicos, el orden de sus genes no está conservado (RANZ *et al.* 1999; GONZÁLEZ *et al.* 2000; RANZ *et al.* 2001; GONZÁLEZ *et al.* 2002; RICHARDS *et al.* 2005). La evolución del cariotipo de *Drosophila* es el resultado de la acción de inversiones paracéntricas (que no incluyen el centrómero) y, en menor medida, de fusiones y fisiones de los centrómeros (CLAYTON y GUEST 1986) generando la variedad de configuraciones del cariotipo actual (POWELL y DESALLE 1995; POWELL 1997). También se han descrito algunos cambios de posición no explicables por inversiones

Tabla 1. Homologías cromosómicas entre algunas especies del género *Drosophila*.

Table 1. Chromosomal homologies between some *Drosophila* species.

		Elemento de Muller					
Especie		A	B	C	D	E	F
<i>D. melanogaster</i>	(S)	X	2L	2R	3L	3R	4
<i>D. pseudoobscura</i>	(S)	XL	IV	III	XR	II	V
<i>D. virilis</i>	(D)	X	4	5	3	2	6
<i>D. repleta</i>	(D)	X	3	5	4	2	6
<i>D. buzzatii</i>	(D)	X	3	5	4	2	6

En los cromosomas metacéntricos L y R indican los brazos cromosómicos izquierdo y derecho respectivamente. (S) Subgénero *Sophophora*, (D) Sugénero *Drosophila*. (SCHAFFER *et al.* 1993; POWELL 1997).

In metacentric chromosomes L and R show the left and right arm respectively. (S) Subgenus Sophophora, (D) Sugenus Drosophila. (SCHAFFER et al. 1993; POWELL 1997).

paracéntricas lo que indica la existencia de transposiciones, aunque a muy baja frecuencia (BETRÁN *et al.* 2002; RANZ *et al.* 2003; GONZÁLEZ *et al.* 2004).

1.5.1. Origen de las Reordenaciones Cromosómicas

La formación de una inversión cromosómica implica la escisión del DNA en dos puntos, el giro del fragmento generado y la unión de los extremos. Entre los posibles mecanismos implicados en este proceso se han descrito las rotura al azar, los puntos frágiles (NOVITSKI 1946; PEVZNER y TESLER 2003) y la recombinación ectópica (LIM y SIMMONS 1994). Aunque la recombinación ectópica se puede producir entre cualquier tipo de secuencia repetida en orientación inversa, hay que destacar la importancia de los elementos móviles. Se han logrado clonar los puntos de rotura de dos inversiones polimórficas naturales demostrando por primera vez la implicación de los elementos transponibles en la generación de inversiones naturales (CACERES *et al.* 1999; CASALS *et al.* 2003). En otros casos no se han encontrado restos de elementos transponibles o secuencias repetidas, ni evidencias de presencia de puntos frágiles (WESLEY y EANES 1994; CIRERA *et al.* 1995).

1.5.2. Importancia Evolutiva de los Segmentos Conservados

A pesar de la gran reorganización intracromosómica que ha habido durante la evolución del género *Drosophila* los estudios de cartografía comparativa permiten identificar segmentos conservados entre las especies que se comparan (VIEIRA *et al.* 1997b; RANZ *et al.* 1999; GONZÁLEZ *et al.* 2000; RANZ *et al.* 2001; GONZÁLEZ *et al.* 2002). Estos fragmentos serán más grandes y frecuentes entre especies cercanas y más pequeños a medida que las especies se alejan filogenéticamente. A menudo se ha interpretado que la conservación de asociaciones génicas entre especies filogenéticamente alejadas podía ser debida a la existencia de restricciones funcionales (MARTINEZ-CRUZADO *et al.* 1988; COLOMBO *et al.* 1992; STATHAKIS *et al.* 1995). De todos modos, cuando se detecta un segmento conservado no se sabe si esta conservación se debe a restricciones funcionales, que impiden la fijación de reordenaciones que rompan el segmento, o si bien es el resultado de la distribución al azar de puntos de rotura de un número finito de reordenaciones cromosómicas producidas desde la divergencia de los linajes comparados (NADEAU y TAYLOR 1984). Complejos génicos cuyos genes presentan elementos *cis*-reguladores compartidos en *D. melanogaster*, como *achaete-scute* (MODELELL y CAMPUZANO 1998), el complejo Iroquois (GOMEZ-

SKARMETA *et al.* 1996) o *knirps* y *knirps related* (LUNDE *et al.* 1998) están conservados en *D. repleta* (GONZÁLEZ *et al.* 2002) lo que sugiere la acción de la selección natural al menos en algunos casos.

Estudios recientes parecen indicar que la organización lineal de los genes en los cromosomas no es aleatoria (KOSAK y GROUDINE 2004). Varios estudios mediante microarrays o “análisis serial de la expresión génica” (SAGE – Serial Analysis of Gene Expression) muestran agrupación genómica de genes coexpresados en *Drosophila* (BOUTANAIEV *et al.* 2002; SPELLMAN y RUBIN 2002), nemátodos, mamíferos (LERCHER *et al.* 2002) o *Arabidopsis* (WILLIAMS y BOWLES 2004). Se ha detectado la agrupación de genes “housekeeping” con nivel de transcripción elevado (LERCHER *et al.* 2002), de genes con patrones de expresión similares (SPELLMAN y RUBIN 2002) o de genes de tejidos específicos (BOUTANAIEV *et al.* 2002). ¿Podría la coexpresión ser un factor en el mantenimiento de la sintenia? Si la expresión coordinada de estos genes es ventajosa, deberíamos encontrar en las regiones coexpresadas más sintenia de la esperada (SPELLMAN y RUBIN 2002).

Las presiones selectivas que promueven la agrupación no están relacionadas simplemente con la tasa de transcripción, sino que favorecen la unión de los genes que son activos en un tejido determinado o en todos los tejidos (LERCHER *et al.* 2002). La correulación transcripcional en eucariotas superiores opera en extensos dominios cromatínicos que comprenden múltiples genes (BOUTANAIEV *et al.* 2002). La estructura de la cromatina, y por tanto, la accesibilidad de diferentes regiones genómicas a la maquinaria transcripcional, puede variar según el tipo celular. En esta situación, podría ser ventajoso agrupar los genes “housekeeping” en una región que mantenga una conformación accesible en cualquier tipo celular o los genes específicos de testículo en una región que se active sólo en esta fase. Aunque todavía falta verificación experimental, la regulación transcripcional coordinada al nivel de la cromatina podría ser un sistema económico de remodelar las estructuras cromosómica (BOUTANAIEV *et al.* 2002; LERCHER *et al.* 2002; UEDA *et al.* 2002). Por otro lado, la agrupación de genes coexpresados podría también ser el resultado de la expresión residual de genes próximos a los verdaderamente activos debido a la descompactación de la cromatina de la región.

Uno de los casos más paradigmáticos de conservación de agrupaciones génicas es el de los genes homeóticos. Estos genes, implicados en la formación del patrón

anteroposterior de los metazoos, se presentan agrupados en el cromosoma y en el mismo orden en el que se expresan a lo largo del eje anteroposterior del embrión. Esta agrupación, o colinealidad, se mantiene en casi todos los grupos estudiados y se ha convertido en un paradigma de la Evo-Devo.

1.6. EL COMPLEJO DE GENES HOX

1.6.1. Los Genes Homeóticos

Ya en el inicio de la genética de *Drosophila*, en el laboratorio de Morgan, se encontraban ocasionalmente ejemplares con malformaciones (Figura 8). En uno de estos mutantes *bithorax* (*bx*), descubierto por Bridges en 1915, los órganos del equilibrio (halterios) aparecían engrosados y con una apariencia parcial de ala. Esta malformación se interpretó como una transformación del tercer segmento torácico (T3) hacia segundo (T2). El término “homeosis”, acuñado por Bateson en 1894 para referirse a la conversión de una estructura en la apariencia de otra homóloga, fue utilizado para describir este tipo de malformaciones en las que células de una región se comportan cómo si estuvieran localizadas en otra región del organismo. Y las mutaciones fueron llamadas homeóticas. Así, *bx* fue el primer mutante homeótico.

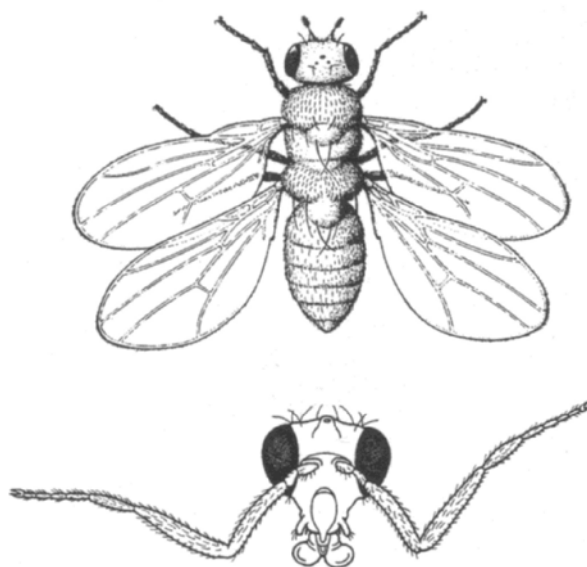


Figura 8. Transformaciones homeóticas en *Drosophila*. Arriba, mutaciones múltiples del BX-C (*abx bx³ pbx*) transforman los halterios en un segundo par de alas (LEWIS 1998). Abajo, una mutación del gen *Antennapedia* en homocigosis transforma las antenas en un par de patas. Dibujos de MÜLLER (1997).

Figure 8. Homeotic transformations in *Drosophila*. Up, multiple mutations of the BX-C (*abx bx³ pbx*) transform halteres in a second pair of wings (LEWIS 1998). Down, a mutation of the *Antennapedia* gene in homozygosis transforms the antenna in a pair of legs. Drawings from MÜLLER (1997).

Desde principios de los 40, Lewis analizó las bases genéticas de las transformaciones homeóticas, en un intento de contrastar la hipótesis de Bridges, basada en evidencias citológicas, de que el genoma contiene duplicaciones génicas naturales, a menudo en tandem. Empezó trabajando con diferentes mutantes ya descritos que causaban transformaciones similares a *bx* y que, en muchos casos, complementaban a esta mutación. Estas mutaciones tenían la particularidad de estar agrupadas en una región del cromosoma, que fue bautizada como Complejo Bithorax (BX-C). Los genes responsables de este fenómeno resultaron ser miembros de una familia génica que controla la identidad de los segmentos a lo largo del eje anteroposterior del cuerpo. Lewis trabajó en estos problemas durante décadas y en 1978 publicó un resumen de sus resultados. En esta revisión formuló teorías sobre cómo interactúan los genes homeóticos, cómo el orden de estos genes se corresponde al orden de los segmentos a lo largo del eje corporal y sobre cómo se expresan los genes individuales.

Paralelamente, Kaufman y colaboradores, estudiando mutantes homeóticos que afectaban a las regiones anteriores del organismo, describieron un segundo complejo, el Complejo Antennapedia (ANT-C) e hicieron predicciones similares a las del BX-C (LEWIS *et al.* 1980b; KAUFMAN *et al.* 1990).

Reglas de la regulación en *cis* de los genes del BX-C (LEWIS 1978; LEWIS 1998):

- 1.- Colinealidad (COL): El orden de los loci en el cromosoma es igual al orden en que esos loci se expresan a lo largo del eje anteroposterior del cuerpo.
- 2.- Inactivación en *cis* (CIN): Las mutaciones de pérdida de función asociados con una región *cis*-reguladora determinada tienden a inactivar la región siguiente del complejo en sentido distal.
- 3.- Sobreexpresión en *cis* (COE): Las mutaciones de pérdida de función asociados con una región *cis*-reguladora determinada tienden a ir acompañados por una sobreexpresión de la función asociada con la región *cis*-reguladora proximal a ella.

Los primeros estudios ya apuntaban a que se trataba de un grupo de genes relacionados y apoyaba la hipótesis de que los nuevos genes se originan por duplicación y divergencia de genes preexistentes. En 1964, y basándose en el modelo bioquímico del operón, Lewis interpretó la función de los genes del BX-C como “represores de ciertos sistemas de diferenciación celular, permitiendo la acción de otros sistemas”. Posteriormente, García-Bellido puntualizó que también podrían funcionar como activadores de estos sistemas (LEWIS 1998).

En 1984, MCGINNIS *et al.* y SCOTT y WEINER descubrieron, de forma independiente, que los genes homeóticos contienen una secuencia de 180 nucleótidos altamente conservada. Esta secuencia, llamada caja homeótica o homeobox, codifica un dominio proteico de 60 aminoácidos, el homeodominio, que forma una estructura angular de tipo hélice-giro-hélice. Este dominio es capaz de reconocer y unirse a motivos nucleotídicos específicos y la proteína que lo contiene puede por tanto actuar como factor de transcripción.

La secuencia del homeobox permitió identificar rápidamente, mediante técnicas de hibridación de ácidos nucleicos o amplificación por PCR con cebadores degenerados, secuencias similares en organismos tan diversos como hidra (SCHUMMER *et al.* 1992; GAUCHAT *et al.* 2000), nemátodos (WANG *et al.* 1993), sanguijuelas (NARDELLI-HAEFLIGER y SHANKLAND 1992), anfioxo (HOLLAND *et al.* 1992), peces (NJOLSTAD y FJOSE 1988; NJOLSTAD *et al.* 1988) o humanos (ACAMPORA *et al.* 1989). Actualmente se conocen gran cantidad de genes con homeobox, que actúan como factores de transcripción habitualmente en procesos de desarrollo. Dentro de este conjunto de genes con homeobox, los genes homeóticos, también llamados genes con homeobox de la clase *Antennapedia* (ZHANG y NEI 1996) o genes Hox, son aquellos que determinan el patrón de desarrollo a lo largo del eje anteroposterior del embrión, especificando la identidad de los segmentos y dando información posicional a lo largo del este eje.

1.6.2. Agrupación y Colinealidad

Cuando se empezaron a caracterizar a nivel molecular los genes de ratón homólogos a los genes homeóticos de *Drosophila*, se encontró que también estaban organizados en complejos génicos, y que sus patrones de expresión en el inicio de la embriogénesis también seguían la regla de colinealidad entre el orden génico y el patrón axial de expresión (Figura 9). Así, no solo los genes estaban conservados sino también su estructura cromosómica en un complejo génico y su patrón de expresión a lo largo del eje corporal principal (AKAM 1989; BOTAS 1993). Una diferencia destacable era que mientras en la mosca hay un único complejo (aunque partido en dos fragmentos) en el ratón se encontraron cuatro copias del complejo situadas en cromosomas diferentes. A pesar de esta diferencia, se hizo evidente que los complejos génicos de ratón que contenían estos genes son verdaderos homólogos a los complejos Hox de *Drosophila* (AKAM 1989).

Los complejos de vertebrados e insectos contienen los descendientes de una familia génica cuyos miembros principales divergieron antes de la separación de estos dos linajes y mantienen en ambos el mismo orden en el cromosoma. Fue inevitable deducir que insectos y vertebrados habían heredado un mecanismo molecular común para indicar las regiones anterior, central y posterior (AKAM 1989). La colinealidad ha llevado a la convicción de que la organización física de los genes en el cromosoma es significativa para su función normal en el desarrollo y se ha convertido en el paradigma de la Evo-Devo (DUBOULE y MORATA 1994).

1.6.3. Origen del Complejo Hox

Los genes Hox forman parte de un sistema de regulación del desarrollo que se ha conservado como una unidad para llevar a cabo una función abstracta necesaria en los animales pluricelulares, asignar valores posicionales en un campo de células (MCGINNIS y KRUMLAUF 1992). No forman estructuras específicas sino que especifican la posición a lo largo del eje anteroposterior del organismo durante el desarrollo embrionario temprano. Esta función, demostrada en artrópodos, cordados y nemátodos, seguramente fue adquirida temprano en la evolución de los metazoos y se ha conservado a pesar de los cambios surgidos en otros mecanismos del desarrollo (MCGINNIS y KRUMLAUF 1992; SLACK *et al.* 1993). Los genes Hox parecen estar estableciendo una representación molecular del plan corporal en estadios tempranos de todos los animales. Lo que ha llevado a considerar a estos genes y sus patrones de desarrollo un carácter definitorio, sinapomorfía, de los metazoos (SLACK *et al.* 1993).

No todos los organismos presentan el mismo número de genes Hox, pero estudios filogenéticos han permitido identificar los genes ortólogos en las diferentes especies estudiadas y establecer la probable dotación génica del ancestro común a los metazoos. Así la mayoría de organismos poseen un único complejo Hox formado por entre 6 y 14 genes, los cuales se pueden agrupar, dentro de cada phylum, en al menos 8 tipos o grupos parálogos. La comparación de los grupos parálogos de los linajes más distantes muestra que esta dotación no se ha adquirido de forma reciente, sino que el último ancestro común de los Bilateria (Urbilateria) poseía un complejo de similar complejidad. Se estima que este complejo poseía ya dos genes anteriores, un gen Hox3, tres genes centrales y un gen posterior (Figuras 9 y 10) (DE ROSA *et al.* 1999; FERRIER y MINGUILLON 2003).

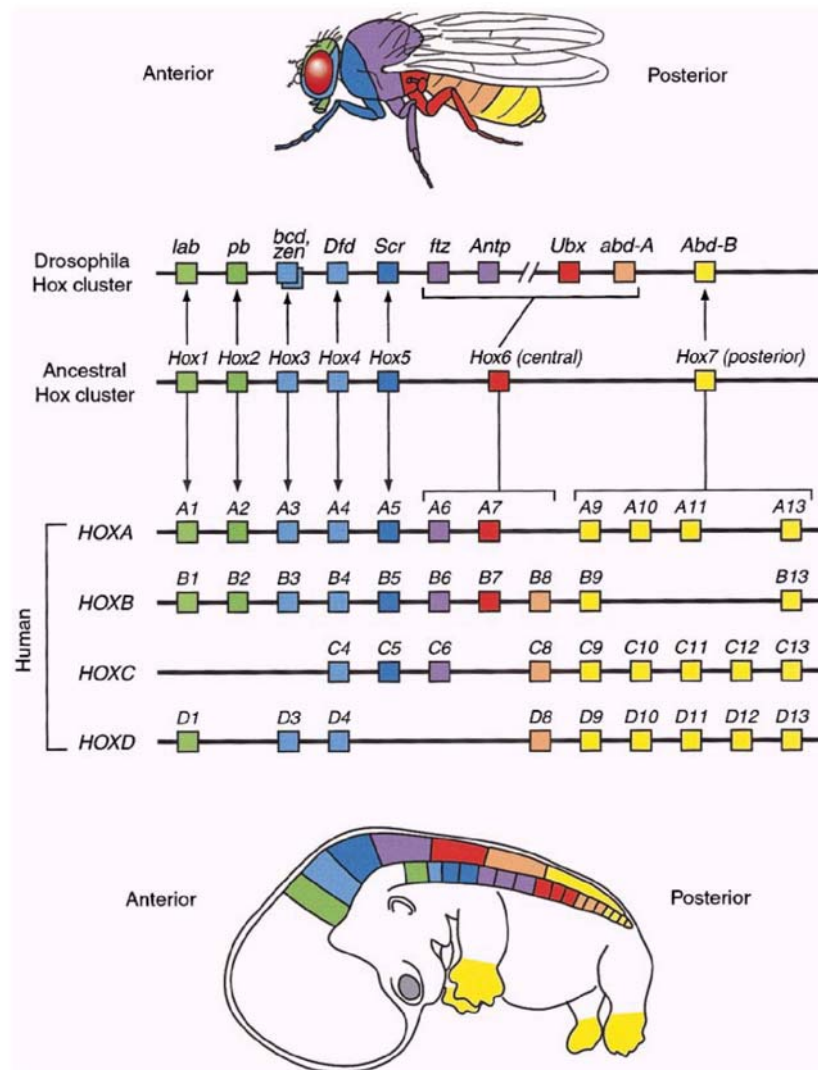


Figura 9. Conservación de la organización genómica y de los patrones de expresión de los genes Hox. Se muestran los complejos de genes Hox de *Drosophila* (arriba) y de mamíferos (abajo), y una representación de sus patrones de expresión. En mamíferos, varios genes posteriores de los complejos *HOXA* y *HOXD* se expresan también en los primordios de las extremidades; indicados en color amarillo. La dotación génica hipotética del complejo en el ancestro se muestra en el centro. Tomado de VERAKSA *et al.* (2000).

Figure 9. Conservation of the genomic structure and expression patterns of Hox genes. The Hox gene complexes of *Drosophila* (up) and mammals (down), and their expression patterns are shown. In mammals, several posterior genes of *HOXA* and *HOXD* complexes are also expressed in limb primordia; shown in yellow. The hypothetical gene complement of the ancestor is shown in the middle. From VERAKSA *et al.* (2000).

El complejo de genes Homeóticos surgió por duplicaciones en tandem y divergencia a partir de un gen Hox ancestral. Pero no solo se duplicaron genes individuales, sino también el complejo entero. Una duplicación del complejo ancestral ProtoHox originó los complejos Hox y ParaHox (BROOKE *et al.* 1998). Esta duplicación ocurrió antes de la divergencia entre los linajes de Cnidarios y animales triploblásticos

(FINNERTY y MARTINDALE 1999). Los genes ParaHox se expresan de forma colinear en el endodermo, al igual que los genes Hox lo hacen en ectodermo y mesodermo.

De las relaciones entre los genes Hox y ParaHox se dedujo la presencia de cuatro genes en el complejo ancestral ProtoHox (BROOKE *et al.* 1998). Pero un reanálisis de los datos existentes muestra otro posible escenario en el que el complejo inicial habría contenido solo dos genes: uno anterior y uno posterior. Después de la duplicación en los complejos Hox y ParaHox se habrían originado los genes Hox3 y centrales en el complejo Hox. En este escenario, el aumento en la diversidad de genes Hox habría coincidido con el origen de los bilateria (GARCIA-FERNÁNDEZ 2004).

Se ha descrito un tercer complejo de genes con homeobox presente en *Drosophila* y vertebrados. Los genes de estos complejos, 93D/E en *Drosophila* o NKL en vertebrados, se expresan en el mesodermo del embrión en desarrollo donde participan en el establecimiento de su patrón y diferenciación. La presencia de genes ortólogos en las dos especies muestra que su agrupación no se debe a una formación reciente por duplicaciones en tándem (HOLLAND 2001; JAGLA *et al.* 2001).

Pero ¿por qué estos genes han permanecido agrupados y en el mismo orden lineal durante más de 500 millones de años de evolución? Se considera que la conservación de estas agrupaciones génicas es reflejo de la necesidad funcional para la correcta activación y función de estos genes. Entre estas necesidades funcionales encontraríamos la existencia de regiones reguladoras compartidas entre genes, los efectos de posición o, más recientemente, la colinealidad temporal.

1.6.4. Los Genes Hox y la Evolución Morfológica

La ausencia de información acerca de genes homeóticos fuera de *Drosophila*, llevó a Lewis 1978, a hipotetizar que los genes del BX-C habían surgido durante la evolución de las especializaciones morfológicas de la mosca. Se creía que la diversidad morfológica a lo largo del organismo era debida a la diversidad de genes Hox, uno para cada tipo de segmento. Pero la presencia en miriápodos, que no presentan tagmosis, de homólogos de todos los genes Hox de *Drosophila* mostró que no fueron necesarios nuevos genes Hox para la diversificación morfológica de los artrópodos (GRENIER *et al.* 1997). La transformación de un par de alas en halterios en el linaje de los dípteros no

fue el resultado de la aparición de un nuevo gen que inhibía la formación del ala, sino la adaptación de un gen existente que a través de pequeños cambios había ido modificando los procesos morfogénicos que llevan a la formación del ala. Poco a poco, el descubrimiento de que estos genes son mucho más antiguos y la presencia de dotaciones complejas en todos los grupos ha ampliado el abanico de hipótesis acerca de su repercusión sobre la evolución morfológica (DENELL *et al.* 1996).

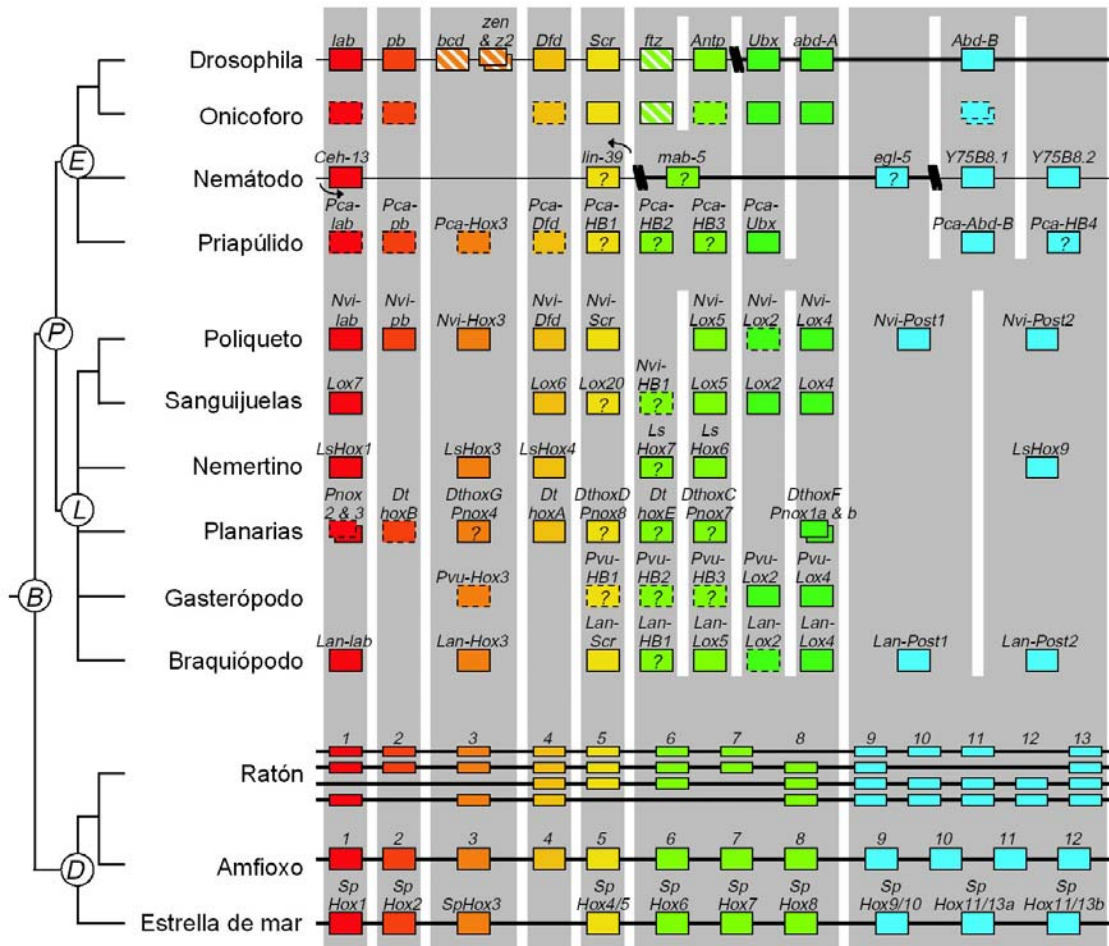


Figura 10. Distribución de los genes Hox en los bilateria. El árbol muestra la filogenia del DNA ribosómico 18S obtenida por AGUINALDO *et al.* (1997). B, ancestro común a los bilateria (Urbilateria); D, deuteróstomos; E, ecdisozoos; L, lofotrocozoos; P, protóstomos. Las líneas horizontales indican datos de cartografía. Las barras blancas verticales delimitan genes o grupos de genes ortólogos. Las relaciones de ortología inciertas están marcadas con un interrogante, los rectángulos punteados indican pequeños fragmentos de PCR, los rectángulos sólidos indican homeobox completos o casi completos, y los rectángulos a rayas indican genes Hox de artrópodos que muestran evolución rápida. Tomado de DE ROSA *et al.* (1999).

Figure 10. Distribution of Hox genes in bilaterians. Phylogenetic tree obtained with ribosomal DNA 18S by AGUINALDO *et al.* (1997). B, common bilaterian ancestor (Urbilateria); D, deuterostomes; E, ecdysozoans; L, lofotrochozoans; P, protostomes. Horizontal black lines indicate mapping data. Vertical white bars delineate orthologous genes or groups of genes. Uncertain orthology relationships are indicated by question marks, dashed boxes indicate short PCR fragment, solid boxes indicate complete or almost complete homeoboxes, and striped colours indicate fast evolving arthropod Hox sequences. From de DE ROSA *et al.* (1999).

Los genes homeóticos desarrollan funciones cruciales durante la embriogénesis y han sido estudiados profundamente desde el punto de vista de la evolución del desarrollo. Parece contradictorio que los genes Hox, que muestran una elevada conservación, puedan ser los responsables de las dramáticas diferencias entre los planes corporales dentro y entre phyla. De todos modos, cambios en número y regulación han sido instrumentos en la evolución y diversificación de los planes corporales a diferentes niveles (Figura 11) (GELLON y MCGINNIS 1998). Se ha propuesto que la divergencia entre clases o phyla ha ido acompañada por cambios moleculares en el complejo de genes Hox o en el uso de sus miembros. Estos cambios incluirían duplicaciones del complejo, pérdida o ganancia de determinados genes, cambios en los niveles de expresión y cambios en las interacciones reguladoras entre los genes Hox y sus genes diana. Análisis genómicos comparativos avalan la hipótesis que las drásticas diferencias en el desarrollo entre diferentes phyla son debidos al uso diferencial de productos génicos conservados estructural y funcionalmente más que de la invención de proteínas *de novo* (CARROLL 1995; GELLON y MCGINNIS 1998).

Aunque la dotación del ancestro de los bilateria era ya compleja (ver apartado 1.6.3) el número y tipo de genes Hox centrales y posteriores difiere entre Ecdisozoos, Lofotrocozoos y Deuteróstomos (DE ROSA *et al.* 1999; CARROLL 2000), lo cual sugiere la relación del número y tipos de genes con las diferencias entre estos grandes grupos. Incluso la aparición de los genes Hox centrales podría haber coincidido con el origen de los animales triploblásticos (GARCIA-FERNÁNDEZ 2004).

Dentro de los Artrópodos, la diferencia entre crustáceos, miriápodos e insectos no es el número o el tipo de genes Hox, sino sus dominios de expresión. El límite anterior de expresión de los genes *Ubx* or *abdominal-A* (*abdA*) está correlacionada con transiciones en la morfología de los apéndices entre estos grupos (GRENIER *et al.* 1997). En insectos, cambios en la morfología del aparato bucal se han relacionado con cambios en el patrón de expresión de *proboscipedia* (*pb*) (ROGERS *et al.* 2002). En cordados, cambios en el área de expresión de genes Hox específicos se han utilizado para explicar la transición de las aletas de los peces a las extremidades de los anfibios (SHUBIN *et al.* 1997), o los cambios en el número relativo de vértebras cervicales, torácicas y lumbares en anfibios, pájaros y mamíferos (BURKE *et al.* 1995).

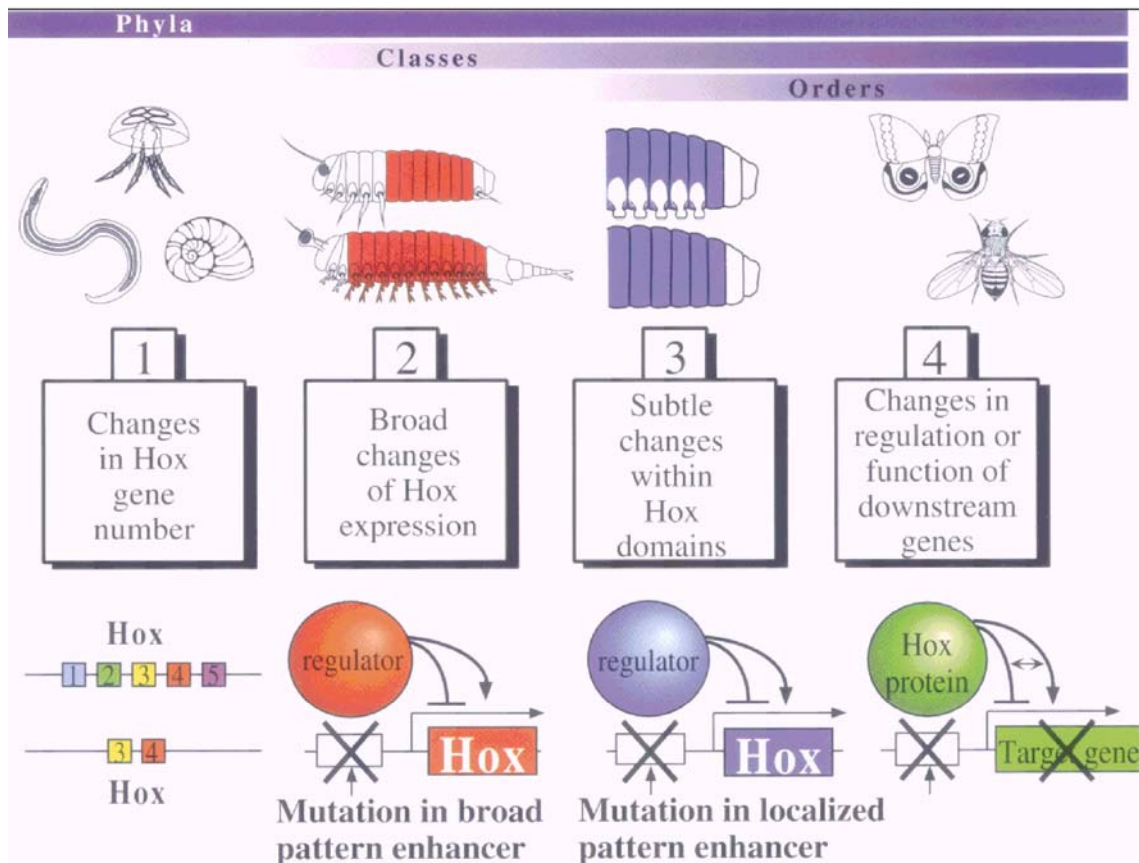


Figura 11. Jerarquía de cambios evolutivos en los genes Hox según su presunto efecto morfológico. Sobre los recuadros, se muestran los cambios en los planes corporales asociados a los cambios en los genes Hox. Bajo los recuadros, probables mecanismos moleculares responsables de estos cambios. Tomado de GELLON y MCGINNIS (1998).

Figure 11. Hierarchy of evolutionary changes in Hox genes arranged by their presumed morphological impact. Above the boxes, modifications of body plans associated with changes in Hox genes. Below the boxes, likely molecular mechanisms responsible of such changes. From GELLON y MCGINNIS (1998).

En otros casos han sido cambios en las proteínas codificadas por los genes Hox lo que ha modificado su acción sobre los genes diana y, en consecuencia su efecto fenotípico. Así, la diversificación morfológica de los segmentos torácicos en insectos se ha producido gracias a la ganancia de nuevos genes diana (WEATHERBEE *et al.* 1999) como resultado de la adquisición de un nuevo dominio en una proteína Hox (GALANT y CARROLL 2002).

Cambios sutiles tienen también efecto fenotípico. El momento y lugar de expresión de un gen Hox dentro de un dominio pueden ser cruciales para su efecto. Por un lado la expresión de un mismo gen en una misma célula tendrá efectos diferentes según el momento en que se produzca, por otro lado pequeños cambios cuantitativos podrán también producir pequeños cambios fenotípicos. Dentro del mismo género,

pequeñas diferencias en el patrón de quetas en el segundo par de patas de *Drosophila* parece ser resultado de diferencias sutiles en la regulación de *Ubx* (STERN 1998). Incluso diferencias morfológicas entre segmentos de un organismo se deben a pequeñas diferencias en el momento en el que un determinado gen se expresa en cada una de las células del segmento (CASTELLI-GAIR y AKAM 1995; CASTELLI-GAIR y LOVEGROVE 2003).

Aunque hace tiempo que sabemos que los genes homeóticos codifican factores de transcripción capaces de cambiar el destino de un segmento entre vías de desarrollo alternativas, todavía sabemos muy poco sobre cómo controlan la morfología detallada de estos segmentos (AKAM 1998). Se creía que los genes Hox determinan los patrones de expresión de los genes realizadores. Ahora se considera que lo que hacen es modificar la acción de las vías de señalización, modificando esos patrones de expresión según la región del organismo. A pesar de la cantidad de trabajos realizados sobre estos genes todavía sabemos muy poco sobre sus genes diana. Los productos de los genes Hox parecen ser generalistas versátiles. Actúan en numerosos tipos celulares y tejidos, donde modulan, unas veces de forma dramática pero habitualmente sutil, una amplia variedad de procesos de desarrollo (AKAM 1998). Pequeños cambios tanto en el homeobox como en la secuencia diana pueden, en teoría, modificar la afinidad y producir pequeños cambios de reconocimiento. Las nuevas herramientas genómicas se están aplicando para la identificación de los genes diana de los genes Hox. Por un lado el uso de microarrays para detectar genes con expresión diferencial en cepas “salvajes” *versus* mutantes (LEEMANS *et al.* 2001). Por otro lado, el estudio del funcionamiento y la evolución de las regiones *cis*-reguladoras donde se unen los genes Hox, para poder predecir los genes diana a partir de las secuencias genómicas (EBNER *et al.* 2005; HERSH y CARROLL 2005).

1.6.5. Estructura del Complejo en los Metazoos

La reconstrucción de la historia evolutiva de la familia de genes Hox es central para entender la evolución de los planes corporales de los bilateria y las relaciones entre la complejidad génica y morfológica. A pesar de que la conservación de la estructura del HOM-C se ha convertido en paradigma de la Evo-Devo, las conclusiones se obtuvieron básicamente comparando *Drosophila* y ratón. Se han clonado fragmentos de genes Hox en multitud de organismos, pero tenemos datos de solo unos pocos grupos en

cuanto a su organización genómica (DE ROSA *et al.* 1999). Aunque en los últimos años han aumentado los datos disponibles a este respecto, me centraré aquí en los datos disponibles cuando se inició este trabajo, en diciembre de 1999 (Figura 10). La organización cromosómica de los genes Hox ha sido caracterizada solamente en insectos, cordados, un equinodermo y un nemátodo. No se ha analizado la organización de estos genes en ningún Lofotrocozoo, a pesar de la gran variedad de phyla incluidos en este grupo.

El complejo del nemátodo *Caenorhabditis elegans* contiene solo seis genes, formando tres pares separados por multitud de genes no relacionados. El orden de uno de los pares está invertido respecto al orden esperado según el patrón de colinealidad. Generalmente se había considerado que los nemátodos habían divergido antes de la separación de protóstomos y deuteróstomos. Como representaban organismos primitivos, esta estructura de los genes Hox se había interpretado como una simplicidad ancestral. Pero la nueva filogenia de los metazoos ha situado a los nemátodos como grupo hermano de los artrópodos lo cual cambia radicalmente su posición filogenética (AGUINALDO *et al.* 1997). Si añadimos los datos sobre la dotación de genes Hox en el ancestro de los bilateria (BALAVOINE y TELFORD 1995; DE ROSA *et al.* 1999), todo indica que esta organización es en realidad una simplificación derivada.

Aparte de *Drosophila*, la organización del complejo ha sido estudiada en algunos insectos no Drosophilidos. El escarabajo *Tribolium castaneum* (BEEMAN 1987), la polilla *Bombyx mori* (UENO *et al.* 1992) o el saltamontes *Schistocerca gregaria* (FERRIER y AKAM 1996) parecen tener un único complejo con homólogos del ANT-C y BX-C de *D. melanogaster*. En *T. castaneum* y *B. mori* los datos provienen del análisis de mutantes homeóticos similares a los descritos en *Drosophila*. Los genes Hox de estos dos insectos se encuentran a 2 o 3 centimorgans, por lo que parece que forman un único complejo indiviso. En el caso de *S. gregaria*, datos obtenidos por hibridación sobre electroforesis de campo pulsante (PFGE – Pulsed Field Gel Electroforesis) sugieren que el complejo Hox no está fragmentado, al menos entre *Sex comb reduced* (*Scr*) y *Abdominal-B* (*AbdB*) (FERRIER y AKAM 1996).

Dentro de los equinodermos se han estudiado los genes Hox en dos erizos de mar. *Heliocidaris erythrogramma* posee al menos ocho genes Hox que parecen estar

agrupados en un fragmento de unas 300 kb (POPODI *et al.* 1996) y *Strongylocentropus purpuratus* presenta 10 genes Hox en un único complejo de algo más de 500 kb (MARTINEZ *et al.* 1999). A pesar de ser organismos con simetría pentaradial poseen un complejo Hox con un contenido génico similar a los complejos de vertebrados.

Los cordados son el grupo más estudiado en cuanto a la organización del complejo Hox, concretamente se ha estudiado el anfioxo y varios vertebrados. El anfioxo posee un único complejo (GARCIA-FERNÁNDEZ y HOLLAND 1994). En vertebrados el complejo homeótico se ha duplicado diversas veces, probablemente con el resto del genoma, de manera que se encuentran cuatro copias del complejo en mamíferos y hasta siete en el pez zebra (AMORES *et al.* 1998). En estos complejos se han definido 13 grupos parálogos, pero no todos los complejos presentan el mismo número de genes, ya que la presencia de copias redundantes ha permitido la pérdida de algunos de ellos.

1.7. EL COMPLEJO DE GENES HOX EN DROSOPHILA

1.7.1. Estructura del Complejo

En *Drosophila* el complejo de genes Hox (HOM-C) está formado por ocho genes Hox, cuatro genes con homeobox y varios genes no relacionados. Los genes Hox fueron descubiertos y descritos inicialmente en *D. melanogaster* (LEWIS 1978). En esta especie el HOM-C se encuentra partido en dos bloques el ANT-C y el BX-C situados a unas 7,5 Mb en el brazo derecho del cromosoma 3 (elemento de Muller E, Tabla 1) (Figura 12). El ANT-C contiene cinco genes Hox: *labial* (*lab*), *pb*, *Deformed* (*Dfd*), *Scr* y *Antennapedia* (*Antp*); cuatro genes con homeobox: *zerknüllt* (*zen*), *zerknüllt related* (*zen2*), *bicoid* (*bcd*) y *fushi tarazu* (*ftz*); y genes sin homeobox como el complejo de genes cuticulares (*Ccp*), *amalgam* (*ama*) y los RNA de transferencia de la lisina (*tRNA-lys*). Los genes *lab*, *pb* y *Dfd* especifican los segmentos cefálicos, mientras que *Scr* y *Antp* contribuyen a dar identidad a segmentos torácicos (KAUFMAN *et al.* 1980; LEWIS *et al.* 1980a; LEWIS *et al.* 1980b; KAUFMAN *et al.* 1990). El BX-C contiene tres genes Hox: *Ubx*, *abdA* y *AbdB*; dos genes no codificantes: *bithoraxoid* (*bxd*) y *infraabdominal 4* (*iab4*), y un gen no relacionado *Glucose transporter type 3* (*Glut*). *Ubx* es requerido para especificar la identidad del tercer segmento torácico y al primer

segmento abdominal, mientras que *abdA* y *AbdB* dan identidad a los restantes segmentos abdominales (LEWIS 1978; SÁNCHEZ-HERRERO *et al.* 1985; DUNCAN 1987; MARTIN *et al.* 1995).

De los genes no homeóticos contenidos en el ANT-C, tres son genes con homeobox que desempeñan funciones importantes en el establecimiento del plan corporal del embrión: el gen materno *bcd* codifica un morfógeno importante para establecer la polaridad anteroposterior (BERLETH *et al.* 1988), *ftz* es un gen pair rule (WEINER *et al.* 1984) y *zen* un gen zigótico involucrado en la diferenciación dorsoventral (RUSHLOW *et al.* 1987). El gen *zen2* contiene un homeobox similar al de *zen*, pero no tiene función conocida (RUSHLOW *et al.* 1987). Genes como *ama* (SEEGER *et al.* 1988) o los *Ccp* (PULTZ *et al.* 1988) no están implicados en el control del desarrollo.

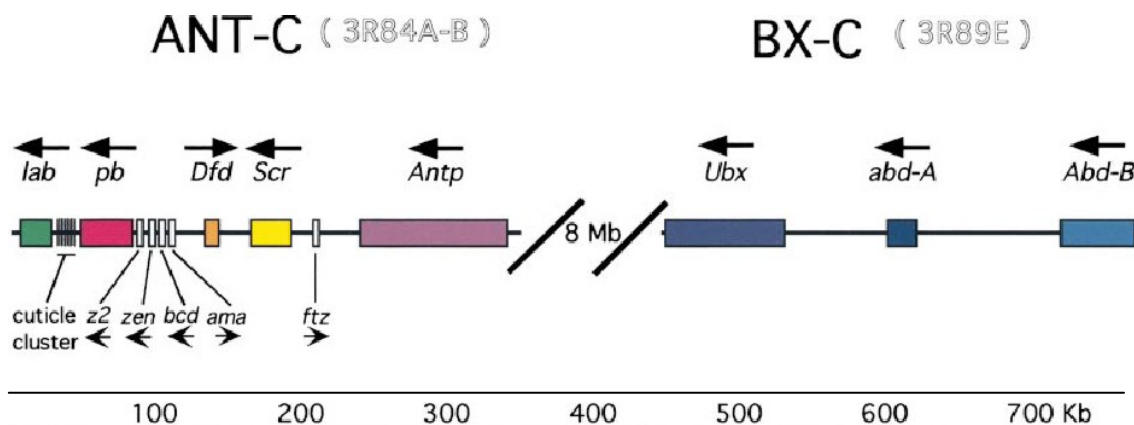


Figura 12. Estructura del HOM-C en *D. melanogaster*.

Figure 12. HOM-C structure in *D. melanogaster*.

El HOM-C de *D. melanogaster* contiene unas 650 kilobases (kb) de secuencia genómica. La mayoría de unidades transcripcionales del HOM-C son largas para lo habitual en *Drosophila*, por ejemplo *Antp* mide unas impresionantes 103 kb. El resto de unidades de transcripción van de 6 a 70 kb (MCGINNIS y KRUMLAUF 1992). La mayoría de los genes homeóticos tienen unidades de transcripción largas con regiones *cis*-reguladoras que se extienden a lo largo de decenas de kb (BENDER *et al.* 1983a; SCOTT *et al.* 1983; KARCH *et al.* 1985; BOULET y SCOTT 1988; CELNIKER *et al.* 1990; SIMON *et al.* 1990; IRVINE *et al.* 1991; HOOPER *et al.* 1992). El elevado tamaño de los genes Hox y del HOM-C se debe al gran tamaño y complejidad de las regiones reguladoras, las cuales incluyen tanto enhancers que promueven la expresión de los transcritos en

dominios específicos, como regiones “barrera” que aíslan las diferentes regiones reguladoras.

1.7.2. Activación y Mantenimiento de los Patrones de Expresión Hox

En *Drosophila* todos los segmentos se forman casi simultáneamente durante el blastodermo sincital gracias a la acción de los genes de la cascada de segmentación (ver apartado 1.3.2). En este momento el embrión posee una distribución asimétrica de numerosos factores de transcripción de origen materno y zigótico, producidos por los genes gap y pair-rule. La serie de gradientes y bandas sobrepuestas que forman estas proteínas exponen cada núcleo en el eje anteroposterior a combinaciones únicas de factores de transcripción. Estos factores se unen a los lugares de unión de las regiones *cis*-reguladoras de los genes Hox y activan o reprimen la transcripción (Figure 13). De manera que cada gen es activado por una combinación única de factores de transcripción generado en una posición particular a lo largo del eje del embrión. Aunque todavía no caracterizados al mismo nivel de detalle, los elementos reguladores que establecen las bandas de expresión de los genes Hox son presumiblemente similares a los enhancers que dirigen las bandas de expresión de los genes pair-rule. Los genes Hox anteriores y posteriores son activados al mismo tiempo, sin dependencia de los genes posteriores en la activación previa de los genes anteriores (AKAM 1987; MCGINNIS y KRUMLAUF 1992; KENNISON 1993; QIAN *et al.* 1993).

Se requiere una combinación específica de proteínas gap y pair-rule para iniciar la transcripción de cada gen Hox en una banda única de células del blastodermo. Una vez establecido el patrón de expresión, este es mantenido por la acción combinada de autoregulación, regulación cruzada entre genes Hox, y acción de los productos de los genes del grupo trithorax y del grupo polycomb (Figura 13). Los genes del grupo trithorax mantienen activa la expresión de los genes Hox. Mientras los genes del grupo polycomb mantienen la represión de los genes Hox, en las regiones donde no ha sido activado con anterioridad. Estos genes parecen actuar mediante cambios en la estructura de la cromatina que la mantendría accesible o inaccesible a la maquinaria transcripcional (KENNISON 1993; NASIADKA *et al.* 2002).

Genes Segmentación

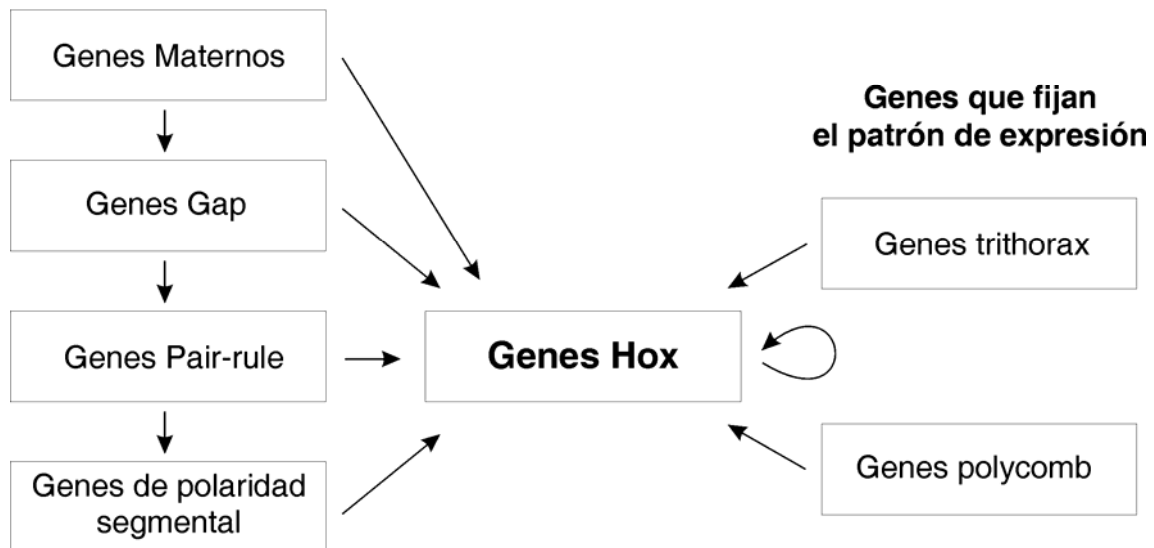


Figura 13. Genes que intervienen en la regulación de los genes Hox: activación por la cascada de segmentación y mantenimiento del patrón por polycomb y trithorax.

Figure 13. Genes involved in Hox gene regulation: activation by the segmentation cascade and maintenance of the expression patterns by polycomb and trithorax.

1.7.3. Función de los Genes Hox

Los genes Hox controlan el desarrollo de estructuras y órganos específicos más que los segmentos en si mismos. Una vez se han establecido los límites segmentales, se encargan de especificar las características estructurales. Los genes del HOM-C especifican la identidad en dominios segmentales o parasegmentales tanto durante la primera fase del desarrollo de *Drosophila* (embriogénesis) como durante la segunda (metamorfosis) (MCGINNIS y KRUMLAUF 1992). Como las estructuras cuticulares de larva y adulto son muy distintas, podemos concluir que la identidad segmental ejercida por los genes del HOM-C tiene caracter abstracto. O bien estos genes actúan sobre diferentes genes diana en estos dos procesos, o bien es el diferente contexto en el que actúan lo que da diferentes funciones a sus genes diana.

La mayoría de estudios de genes Hox se han basado en la observación de alteraciones de la cutícula. Pero en realidad los genes Hox intervienen también en la formación de tejidos internos. Reorganizaciones con puntos de rotura en *iab4*, cuando son viables en heterocigosis aparentan ser “salvajes”, pero internamente carecen de gónadas. Esta fue la primera indicación de que los fenotipos BX-C no se limitan a tejidos ectodérmicos. Siempre es peligroso deducir la función de un gen a partir de mutaciones de pérdida de función, especialmente para genes que afectan a la morfología

(LEWIS 1998). Los fenotipos homeóticos de pérdida de función se producen sólo cuando hay otro gen Hox expresándose en la región, ya sea por que la expresión normal es solapante o porque la ausencia de un gen desreprime la expresión de otro. Cuando esto no ocurre, como en el caso de *lab* y *Dfd*, no se produce una transformación homeótica sino una deficiencia estructural, como es la ausencia de involución cefálica en los mutantes para *lab* (MCGINNIS y KRUMLAUF 1992).

1.7.4. Cambios Estructurales del HOM-C en el Género *Drosophila*

Es remarcable la conservación de la estructura del complejo de genes Hox de *Drosophila* a humanos, y se han realizado muchos intentos para explicarlo. La mayoría de los autores sugiere la existencia de complicadas relaciones reguladoras entre los genes que hacen imposible modificar esta organización durante la evolución. El estudio de la organización de los genes Hox en otras especies del género *Drosophila* debería dar información sobre la conservación de la estructura del complejo y la presencia de genes no homeóticos. La información disponible es muy escasa.

La mayoría de trabajos sobre genes Hox en *Drosophila* se han centrado en estudiar la posición o la secuencia de un gen determinado.

TEROL *et al.* (1991) localizaron por *in situ* la posición de *Antp* en los cromosomas en nueve especies del grupo obscura de *Drosophila*. Posteriormente, localizaron también *lab*, *pb* y *zen*, en una posición citológica adyacente. Intentaron clonar la región *lab-pb* iniciando un recorrido cromosómico desde *pb*, pero sólo lograron llegar a *zen*, en esta especie la distancia entre *zen* y *pb* resulta ser cuatro veces mayor que en *D. melanogaster* (TEROL *et al.* 1995).

MAIER *et al.* (1990) secuenciaron el gen *ftz* y sus regiones reguladoras en *D. hydei*, perteneciente al grupo repleta. En esta especie observaron una orientación del gen inversa respecto a *D. melanogaster*. Posteriormente extendieron el estudio al análisis de las regiones reguladoras de *ftz* en 11 especies del género (MAIER *et al.* 1993). Observaron por ejemplo que, en *D. virilis*, la orientación del gen era igual que en *D. hydei*.

HOOPER *et al.* (1992) clonaron y secuenciaron el gen *Antp* en *D. virilis* y parcialmente en *D. subobscura*. Vieron que la estructura del gen está conservada entre las tres especies. Observaron también la presencia de bloques de secuencia conservados en los intrones largos, y hasta 15 kb en la región 5' y 3 kb en la región 3' del gen.

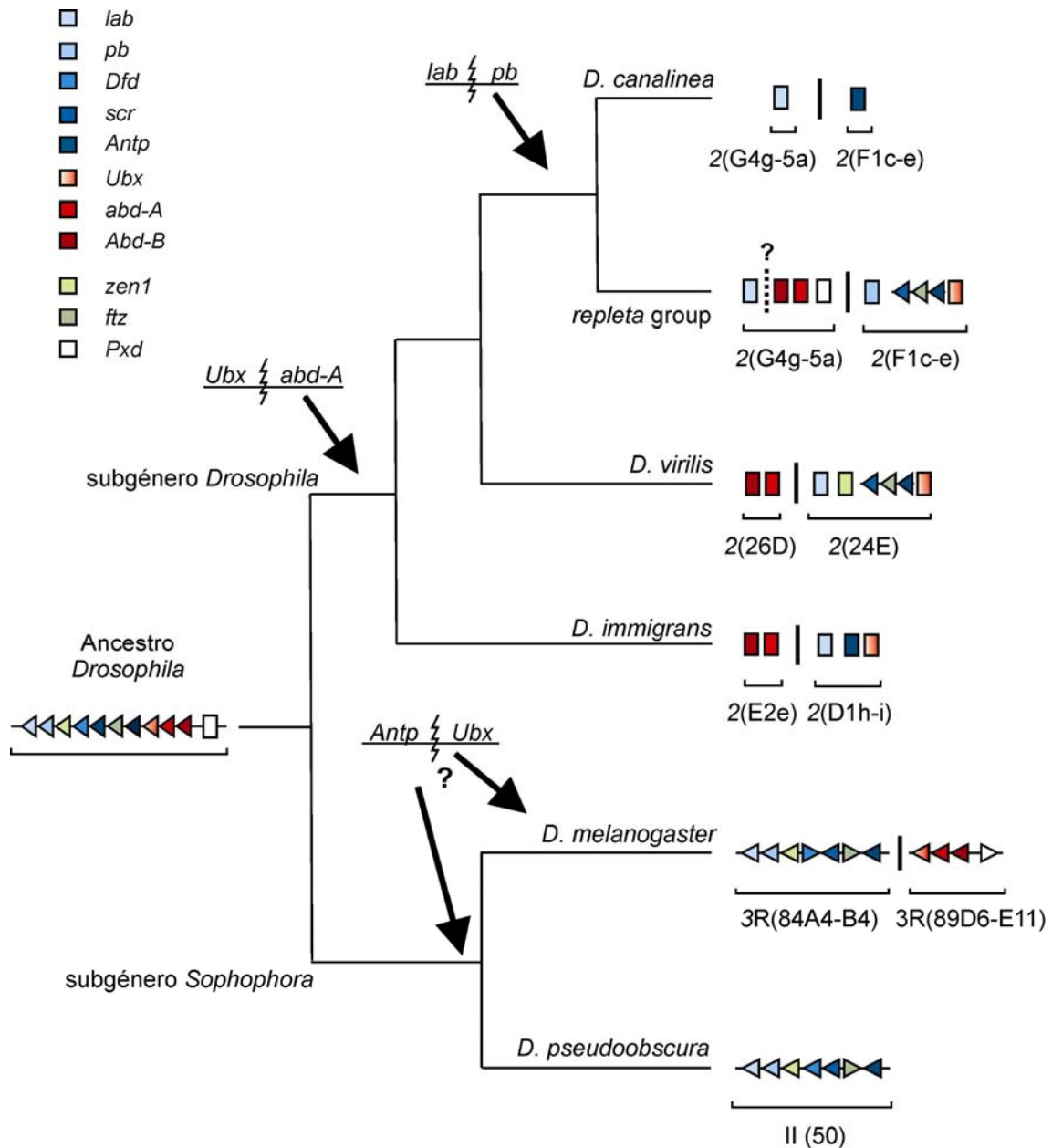


Figura 14. Reconstrucción filogenética de la organización de los genes Hox en el género *Drosophila*. Se ha incluido toda la información disponible (detalles en el texto) sobre la localización de genes Hox. Los genes Hox se muestran como triángulos cuando se conoce su orientación (5'→3'), si no aparecen como rectángulos. Donde no aparecen líneas de conexión, el orden de los genes se ha dibujado como en *D. melanogaster* pero todavía no se ha demostrado. Las flechas indican las roturas conocidas del complejo Hox. Las líneas verticales indican discontinuidades cromosómicas del complejo Hox. De RANZ, CÁCERES y RUIZ (comunicación personal).

Figure 14. Phylogenetic reconstruction of Hox gene organization in the genus *Drosophila*. All known information of Hox gene localization (see text for details) has been plotted. Hox genes are shown as triangles when the orientation (5'→3') is known, they appear as rectangles otherwise. Unless a connecting line appears, the order of the genes has been assumed as in *D. melanogaster* but has not yet been demonstrated. Arrows indicate known splits in the Hox complex. Vertical lines indicate chromosomal discontinuities in the Hox complex. Map position of genes is shown. From RANZ, CÁCERES and RUIZ (personal communication).

BOMZE y LÓPEZ (1994) secuenciaron fragmentos de cDNA de *Ubx* de *D. pseudoobscura*, *D. hydei* y *D. virilis*. Observaron que los diferentes transcritos, fruto del procesamiento alternativo, descritos en *D. melanogaster* estaban también presentes en estas especies. También analizaron el patrón de expresión de *Ubx* en embriones de las cuatro especies, concluyendo que el patrón de expresión de las diferentes isoformas está conservado.

En cuanto a la organización genómica de los genes Hox, RANDAZZO *et al.* (1993) estudiaron la estructura del ANT-C en *D. pseudoobscura*. Concluyeron que su estructura era igual a la descrita para *D. melanogaster* con solo dos diferencias, la orientación del gen *Scr* y la ausencia de *zen2*. Estos autores no estudiaron la región del BX-C. En 1996, VON ALLMEN *et al.* describían una organización diferente en *D. virilis*. Mediante hibridaciones *in situ* sobre cromosomas politénicos observaron que *Ubx* está situado al lado de *Antp* en lugar de adyacente a los genes *abdA* y *AbdB*. Lo que suponía que la organización descrita para *D. melanogaster* no era válida para todo el género como se había supuesto en muchas ocasiones. Esta observación también implica que el ancestro del género *Drosophila* debía tener un complejo único como se observa actualmente en *Tribolium*.

RANZ *et al.* (2001; 2003) mediante hibridación *in situ* sobre cromosomas politénicos localizaron la posición citológica de los genes Hox *pb*, *Antp*, *Ubx*, *abdA* y *AbdB* en *D. repleta* y *D. buzzatii*. En estos trabajos se observó que estas especies presentaban la rotura *Ubx-abdA* descrita en *D. virilis* (VON ALLMEN *et al.* 1996). Posteriormente, la localización del gen *lab* mostró la presencia de una segunda rotura en estas especies. Esta rotura que afecta a la región anterior del complejo, concretamente entre los genes *lab* y *pb*. Esta reorganización situó al gen *lab* en una posición citológica indistinguible de la de los genes *abdA* y *AbdB* (RANZ, CÁCERES y RUIZ comunicación personal). Hibridaciones *in situ* sobre otras especies del subgénero *Drosophila* permitieron delimitar la distribución filogenética de estas roturas (Figura 14).

1.8. Objetivos

Los genes Hox codifican factores de transcripción cuya función es especificar la identidad de los segmentos en el embrión temprano. Estos genes están habitualmente agrupados y organizados en el mismo orden en el que se expresan a lo largo del eje anteroposterior del cuerpo de los metazoos. Aunque la conservación de la organización genómica ha sugerido la existencia de constricciones funcionales actuando sobre ella, la explicación mecanística de la colinealidad ha sido esquiva y bien puede ser que no exista una causa única y universal. La desestructuración del complejo en *C. elegans* y la presencia de tres roturas en dos linajes de *Drosophila* ponen en cuestión que esta organización genómica sea realmente necesaria para su correcto funcionamiento.

Como objetivo de esta tesis nos propusimos investigar las causas y consecuencias de las reorganizaciones naturales del complejo HOM-C de *Drosophila*. Los objetivos concretos son los siguientes:

- Puntos de rotura y mecanismo que las produjo:

1. Determinar en que punto de la región entre *labial* y *proboscipedia* se produjo la rotura presente en las especies del grupo repleta de *Drosophila*.
2. Determinar la distancia entre los genes *lab* y *abdA* que presentan la misma posición citológica en *D. buzzatii*.
3. Identificar el mecanismo responsable de la rotura *lab-pb*. Si se trata de una inversión, determinar el segundo punto de rotura y si este es coincidente con la rotura *Ubx-abdA*. Si se trata de una transposición, delimitar la región transpuesta.

- Efectos funcionales de las roturas del HOM-C sobre los genes Hox adyacentes:

4. Determinar si la separación de *lab* del resto de genes del HOM-C (cambio de posición) ha tenido algún efecto en la estructura del gen o en la evolución de su secuencia.
5. Determinar si la rotura ha separado algún elemento regulador de los genes Hox adyacentes.
6. Determinar si la rotura ha tenido algún efecto en la expresión de los genes Hox adyacentes.

1.8. Objectives

Hox genes code for transcription factors whose function is to specify segment identity in the early embryo. Those genes are usually clustered and arranged in the same order as they are expressed along the antero-posterior body axis of the metazoan. Although the conservation of the genomic organization has suggested the existence of functional constrictions acting over it, the mechanistic explanation of colinearity has been elusive and it can well be that a unique and universal cause does not exist. The disassembly of the complex in C. elegans and the presence of three splits in two lineages of Drosophila call into question if this genomic organization is really necessary for proper functioning.

The objective of this thesis was to investigate the causes and consequences of natural HOM-C rearrangements in Drosophila. The concrete objectives are the following:

- Breakpoints and mechanism responsible:

- 1. Determine the exact position between labial and proboscipedia of the breakpoint present in the species of the repleta group of Drosophila.*
- 2. Determine the distance between lab and abdA, which show the same cytological position in D. buzzatii.*
- 3. Identify the mechanism responsible for the lab-pb split. If it is an inversion, determine the second breakpoint and if it is coincident with the breakpoint of the Ubx-abdA split. If it is a transposition, determine the transposed region.*

- Functional effects of HOM-C splits on the adjacent Hox genes:

- 4. Determine if the separation of the lab gene from the HOM-C (positional change) had any effect in the structure or sequence evolution of this gene.*
- 5. Determine if the split has separated any regulatory element of the adjacent Hox genes.*
- 6. Determine if the split had any effect on the expression of the adjacent Hox genes.*

2. RESULTADOS

2.1. Artículo 1

La primera parte de este trabajo describe el descubrimiento de una nueva rotura del HOM-C en las especies del grupo repleta de *Drosophila*. Este resultado fue obtenido previamente al inicio de esta tesis doctoral. Esta rotura ha recolocado el gen más anterior del complejo, *lab*, en una posición cromosómica distante pero indistinguible de la posición citológica de los genes Hox más posteriores, *abdA* y *AbdB*.

Para investigar las consecuencias evolutivas de las reorganizaciones naturales del HOM-C, se ha clonado y secuenciado el gen *lab* en *D. buzzatii*, un miembro del grupo repleta de *Drosophila*, y parcialmente en *D. virilis*, miembro del grupo virilis, que no presenta la rotura. Los resultados muestran que la estructura de *lab* en *D. buzzatii* está intacta y sitúan la rotura a más de 8 kb del inicio de transcripción. La evolución de la secuencia nucleotídica de *lab* en el género *Drosophila* ha sido investigada mediante métodos de máxima verosimilitud. Aunque la tasa de substitución nucleotídica varía drásticamente a lo largo de la secuencia codificante, no se ha observado variación significativa entre linajes en la tasa de substitución nucleotídica o en la razón entre las substituciones no sinónimas y sinónimas. Por lo tanto, el cambio de posición de *lab* no ha inducido cambios en la tasa de evolución o en el grado de constricciones funcionales.

NEGRE B., J.M. RANZ, F. CASALS, M. CÁCERES y A. RUIZ (2003) A New Split of the Hox Gene Complex in *Drosophila*: Relocation and Evolution of the Gene *labial*. *Molecular Biology and Evolution* **20**(12):2042-2054.

A New Split of the Hox Gene Complex in *Drosophila*: Relocation and Evolution of the Gene *labial*

Bárbara Negre, José María Ranz,¹ Ferran Casals, Mario Cáceres,² and Alfredo Ruiz

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

Hox genes encode transcription factors involved in the specification of segment identity in the early metazoan embryo. These genes are usually clustered and arranged in the same order as they are expressed along the anteroposterior body axis. This conserved genomic organization has suggested the existence of functional constraints acting on the genome organization. Partial disassembly of the Hox gene complex (HOM-C) in *Caenorhabditis elegans* and in two different *Drosophila* lineages, however, calls into question whether this cluster organization is absolutely required for proper function. Here we report a new split of the HOM-C discovered in the species of the *Drosophila repleta* group, which relocated the most anterior gene of the complex, *lab*, to a distant chromosomal site near the two most posterior Hox genes, *abd-A* and *Abd-B*. To investigate the evolutionary consequences of natural rearrangements of the Hox gene complex, the gene *lab* has been cloned and sequenced in *D. buzzatii*, a member of the *D. repleta* group with the split, and in *D. virilis*, a member of a different species group without the split. The results show that the structure of *lab* in *D. buzzatii* is intact and place the breakpoint at least 8 kb from its transcription start site. The nucleotide sequence evolution of *lab* in the genus *Drosophila* has been investigated by means of maximum likelihood methods. No significant variation has been observed among lineages in the rate of nucleotide substitution or in the nonsynonymous/synonymous substitution ratio. Seemingly, the relocation of *lab* has not induced a change in evolution rate or degree of functional constraint. Nevertheless, further work is needed to ascertain whether the *lab-pb* split has had any effects on gene expression.

Introduction

Homeotic (Hox) genes encode transcription factors that specify the future cellular identity along the anteroposterior axis in the early metazoan embryo by activating or repressing their target genes (Gehring and Hiromi 1986). In the genome, Hox genes are found clustered and arranged in the same order as they are expressed along the anteroposterior body axis (Lewis 1978; McGinnis and Krumlauf 1992). This colinearity appeared before the separation between vertebrates and invertebrates, and it is widely conserved across the bilaterian phylogeny, which has suggested the existence of functional constraints acting on their genome organization (Duboule and Morata 1994). Partial disassembly of Hox complexes in *Caenorhabditis elegans* (Ruvkun and Hobert 1998) and *Drosophila* (vonAllmen et al. 1996; Adams et al. 2000; Ranz, Casals and Ruiz 2001), together with reported experimental breaks of the *Drosophila* complexes (Hazelrigg and Kaufman 1982; Struhl 1984; Tiong, Whittle and Gribbin 1987), calls into question whether this organization is absolutely required for proper function.

A single Homeotic gene cluster (HOM-C) is postulated to have existed in the ancestor of the arthropods (Hughes and Kaufman 2002), although most sequences are partial and, usually, there are no data about the chromosomal localization of the genes (Finnerty and Martindale 1998; deRosa et al. 1999; Cook et al. 2001). The 10 Hox genes included in this ancestral complex are (in anteroposterior

order): *labial* (*lab*), *proboscipedia* (*pb*), *Hox3*, *Deformed* (*Dfd*), *Sex comb reduced* (*Scr*), *fushi tarazu* (*ftz*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*). This was probably also the organization in the ancestor of the *Drosophila* genus, although in insects *Hox3* and *ftz* have lost their Homeotic function (Hughes and Kaufman 2002). Two rearrangements of this ancestral Hox complex are known to have occurred during the evolution of the *Drosophila* genus (fig. 1). In the lineage leading to *D. melanogaster* a split between *Antp* and *Ubx* separated the genes in two complexes: the Antennapedia complex (ANT-C) (Kaufman, Seeger, and Olsen 1990) and the Bithorax complex (BX-C) (Lewis 1978; Duncan 1987). These two complexes are located on the right arm of *D. melanogaster* chromosome 3, separated by approximately 9.6 Mb of euchromatic sequences. The ANT-C spans ~400 kb (Adams et al. 2000) and comprises five Hox genes (*lab*, *pb*, *Dfd*, *Scr*, and *Antp*), which are expressed in those segments that will become the head and the anterior thorax. The BX-C spans ~350 kb (Martin et al. 1995; Adams et al. 2000) and contains three Hox genes (*Ubx*, *abd-A*, and *Abd-B*), which are expressed in the posterior thorax and abdomen. A different arrangement of the Hox genes was found in *D. virilis*, a species of the *Drosophila* subgenus (vonAllmen et al. 1996). In this species *Ubx* mapped by in situ hybridization with *Antp* instead of with *abd-A*, indicating that in the lineage leading to this species a disruption of the Hox complex occurred between the genes *Ubx* and *abd-A* (fig. 1).

The diversity in Hox gene arrangement observed among *Drosophila* species along with the reported flexibility of its genome (Stone 1962; Hartl and Lozovskaya 1994; Powell 1997; Ranz, Casals, and Ruiz 2001) prompted us to investigate the organization of Hox genes in *D. buzzatii* and *D. repleta*, two representative species of the *D. repleta* species group of the *Drosophila* subgenus. The in situ hybridization to the salivary gland chromosomes of the *Antp*, *ftz*, *pb*, *Ubx*, *abd-A*, and *Abd-B* genes (Ranz, Segarra,

¹ Present address: Department of Genetics, University of Cambridge, Cambridge, United Kingdom.

² Present address: Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia.

Key words: *labial*, Hox complex, *Drosophila buzzatii*, chromosomal rearrangement, genome evolution.

E-mail: alfredo.ruiz@uab.es.

Mol. Biol. Evol. 20(12):2042–2054. 2003

DOI: 10.1093/molbev/msg238

Molecular Biology and Evolution, Vol. 20, No. 12,

© Society for Molecular Biology and Evolution 2003; all rights reserved.

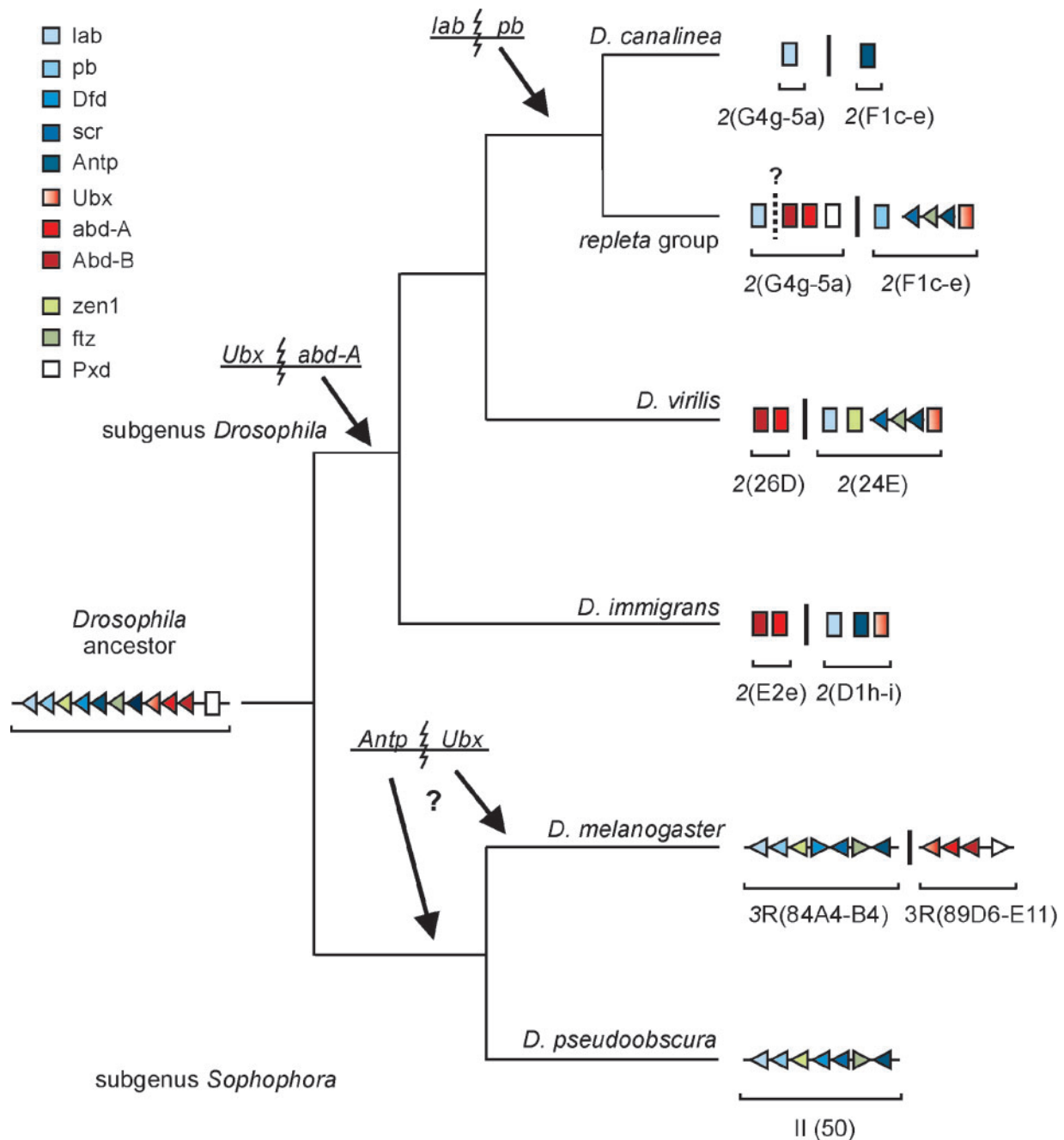


FIG. 1.—Phylogenetic reconstruction of Hox gene organization in the genus *Drosophila*. All known information of Hox gene localization has been plotted in a phylogenetic tree to establish when each reorganization took place. The evolutionary relationships among species and divergence times are from different sources (Powell 1997; Ranz, Segarra, and Ruiz 1997; Kwiatowski and Ayala 1999). Hox genes are shown as triangles when the orientation (5' → 3') is known; they are otherwise shown as rectangles. Hox genes expressed in the anterior and posterior parts of the embryo appear in red and blue gradients, respectively. The genes *ftz*, *zen1*, and *Pxd* are also included (Wharton 1942; leCalvez 1953; vonAllmen et al. 1996; Powell 1997; Adams et al. 2000); the gene *zen2* found in the Hox complex of *D. melanogaster* (Adams et al. 2000) is not represented. Unless a connecting line appears, the order of the genes has been assumed as in *D. melanogaster* but has not yet been demonstrated. Arrows indicate known splits in the Hox complex. Vertical lines indicate chromosomal discontinuities in the Hox complex. When known, map position of genes is shown. Information for *D. canalinea* and *D. immigrans*: this work; *D. melanogaster*: Adams et al. 2000; *D. pseudoobscura*: Terol, Perez-Alonso, and Frutos 1991, Randazzo et al. 1994; and *D. virilis*: Maier, Preiss, and Powell 1990, Maier, Sperlich, and Powell 1993, vonAllmen et al. 1996, and this work. For all the *D. repleta* group species, a representative organization obtained from *D. buzzatii* (Ranz, Segarra and Ruiz 1997; Ranz, Casals, and Ruiz 2001; this work) is shown.

Table 1
Sequence of Oligonucleotide Primers Used for PCR Amplification of Different Regions of the Gene *labial* in Different *Drosophila* Species

Gene Region	Species	Forward (5'–3')	Reverse (5'–3')	Product Size
Exons 2–3 (homeobox)	<i>D. buzzatii</i>	TCCAGCACCGAACTACCTG	TTCACGCGCTTCTTCTGCTTCAT	1,122 bp
	<i>D. hydei</i>			954 bp
	<i>D. mercatorum</i>			725 bp
	<i>D. repleta</i>			727 bp
	<i>D. virilis</i>			1,172 bp
Exon 1	<i>D. buzzatii</i>	ATGATGGACGTAAGCAGCATG	CATTCTCTTGAGTTGCATCCA	1,285 bp
	<i>D. virilis</i>			1,261 bp
cDNA	<i>D. buzzatii</i>	CATCCGTGTCAGCTGCTAAC	TTCACGCGCTTCTTCTGCTTCAT	570 bp
	<i>D. virilis</i>	CCTCGTCATCAACAACATCTG	TTCACGCGCTTCTTCTGCTTCAT	548 bp
Intron 1 fragment	<i>D. buzzatii</i>	CCTGGAGCCATATAATAGGT	GCTCGATTAATTGATGGATG	~3.5 kb
5' UTR	<i>D. virilis</i>	TGCTCGGCCAGCAATCAG	AGGCAGTGGCTGTGGCATGA	818 bp

and Ruiz 1997; Ranz, Casals, and Ruiz 2001; Ranz et al. 2003) indicated that both species lack the *Antp-Ubx* split present in *D. melanogaster* and share the split between *Ubx* and *abd-A* with *D. virilis* (fig. 1). However, the location of the *lab* gene was not yet known in those species. Here, we have mapped the *lab* gene in different *Drosophila* species and show that a new split has separated this gene from the other anterior Hox genes and relocated it near *abd-A* and *Abd-B*, which direct the posterior development of the fly. To determine if the *lab-pb* split has altered in any way the structure or evolution of *lab*, we have cloned and sequenced this gene in *D. buzzatii* (which presents the split) and also partially in *D. virilis* (which lacks the split). The results show that the molecular structure of *lab* is conserved between *D. buzzatii*, *D. melanogaster*, and *D. virilis*, and that the breakpoint must be located at least 8 kb upstream of this gene. In addition, they indicate that *lab* is a relatively fast evolving gene, although there was no evidence for positive or relaxed selection either before or after the split.

Materials and Methods

Drosophila Stocks

The following species and stocks were used (code numbers of the Tucson *Drosophila* Species Stock Center are given in parenthesis): *D. buzzatii* (st-1 and j-19; see Cáceres, Puig, and Ruiz 2001), *D. canalinea* (15050–1201.0), *D. hydei* (Carboneras, Spain), *D. immigrans* (15111–1731.0), *D. mercatorum* (Comarapa, Bolivia), *D. repleta* (15084–1611.2), and *D. virilis* (Tokyo, Japan). All stocks are fixed for the standard chromosomal arrangements except line j-19 of *D. buzzatii*, which is homozygous for inversion 2j.

In Situ Hybridization to Salivary Gland Chromosomes

Clones used as probes to map by in situ hybridization the *Antp*, *Ubx*, *abd-A*, and *Abd-B* genes have been reported elsewhere (Ranz, Segarra, and Ruiz 1997; Ranz, Casals, and Ruiz 2001). To localize the *lab* gene in *D. buzzatii*, *D. hydei*, *D. mercatorum*, *D. repleta*, and *D. virilis*, homologous probes were generated in each species by polymerase chain reaction (PCR) amplification of a 725–1,172-bp segment encompassing exons 2–3 of the gene and including the homeobox (table 1), followed by cloning and sequencing of

the amplified products. The clone of *D. virilis* was used to map *lab* in *D. immigrans* and *D. canalinea*.

Preparations of slides, probe labeling, and hybridization to salivary gland chromosomes were performed as previously described (Ranz, Segarra, and Ruiz 1997). Homologous and heterologous hybridizations were performed at 37°C and 25°C, respectively. Micrographs were taken by phase contrast with a Nikon Optiphot-1 microscope and a Nikon H-III photomicrographic system at 600× magnification using EKTAR-25 Kodak film and a blue filter. The chromosome map of *D. repleta* (Wharton 1942) was used to localize the in situ hybridization signals in all *D. repleta* group species and *D. canalinea*; the cytological maps of *D. immigrans* (leCalvez 1953) or *D. virilis* (Gubenko and Evgen'ev 1984) were used in the remaining cases.

Southern Hybridization and Library Screening

Genomic DNA extraction was performed as described in Piñol et al. (1988). DNA was digested with different restriction enzymes and transferred from the agarose gels to nylon membranes by capillarity transfer under neutral conditions (Sambrook, Fritsch, and Maniatis 1989). Probes were labeled with digoxigenin-11-dUTP by *random primer* with the DIG DNA Labeling and Detection Kit (Roche). Hybridization was performed at 42°C for 16 h using standard hybridization buffer with formamide (50%). Filters were washed twice for 15 min in 2× SSC 0.1% sodium dodecyl sulfate (SDS) at room temperature and twice for 20 min in 0.1× SSC 0.1% SDS at 68°C.

A lambda genomic library derived from *D. buzzatii* line j-19 (Cáceres, Puig, and Ruiz 2001) was amplified following Sambrook, Fritsch, and Maniatis (1989) and used for cloning of *lab*. Screening was done by plaque hybridization under the same conditions as for the Southern hybridization. Two probes were used to clone *lab* in *D. buzzatii*, one corresponding to exons 2–3 and the other to exon 1, separated in *D. melanogaster* by a large intron of nearly 14 kb. The probe containing exons 2 and 3 was the same 1,122 bp fragment used in the in situ hybridization (see above). The exon 1 probe consisted in a 1,285-bp fragment amplified by PCR using specific primers (table 1). Sequencing of the PCR product confirmed its homology with exon 1 of *D. melanogaster lab*.

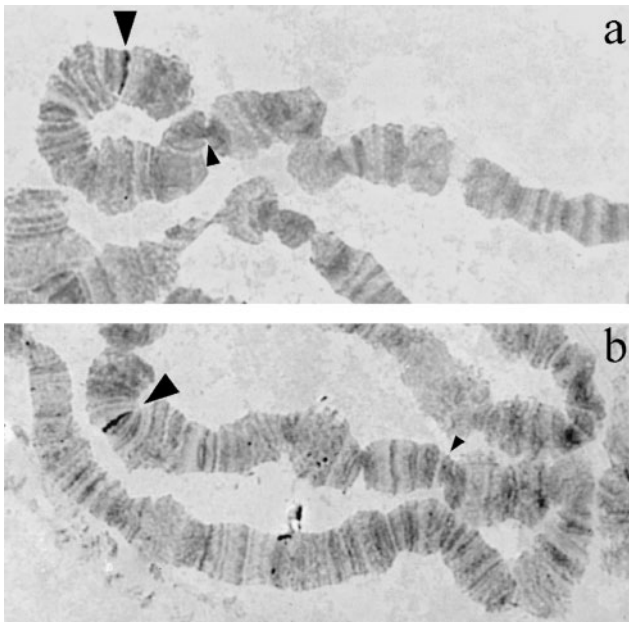


FIG. 2.—Chromosomal localization of the genes *lab* and *abd-A* in the *D. repleta* group. Large arrowheads indicate the hybridization signals of the *abd-A* (a) and *lab* (b) probes at G4g-5a of chromosome 2 of *D. buzzatii*, one representative species of the *D. repleta* group of *Drosophila*. Small arrowheads indicate the chromosomal localization of the Hox genes *Antp*, *pb*, and *Ubx* (Ranz, Segarra, and Ruiz 1997 and Rantze et al. 2003).

Polymerase Chain Reaction Amplification

Primers (table 1) were designed with Primer Designer (1.01—1990 Scientific and Educational Software). Primers used to amplify the *lab* segment comprising exons 2–3 (homeobox) in all species were designed according to the *D. melanogaster lab* sequence. Exon 1 of *D. buzzatii* and *D. virilis* were amplified using the regions conserved between Labial proteins of *D. melanogaster* (AAD19811.1) and *Tribolium castaneum* (AAF64147.1). To amplify the cDNA of *D. buzzatii* and *D. virilis*, forward primers were designed in each species exon 1 sequence. Primers located in the transcription initiator (Inr) sequence of *D. buzzatii* and the beginning of the coding region of *D. virilis* were used to amplify the 5' UTR in the latter species.

Polymerase chain reaction was carried out in a volume of 50 μ l, including 100–200 ng of genomic DNA or cDNA or 1 μ l of phage lysate, 20 pM of each primer, 200 μ M dNTPs, 1.5 mM MgCl₂, and 1 unit Taq DNA polymerase. Temperature cycling conditions were 30 rounds of 30 s at 94°C, 30 s at the annealing temperature, and 60 s at 72°C, with annealing temperature varying between 55° and 61°C depending on the primer pair used.

RNA Isolation and Reverse Transcriptase-Polymerase Chain Reaction

Total RNA from 0–12-h-old embryos from *D. buzzatii* (line j-19) and *D. virilis* was isolated with TRIZOL (GIBCO) following the manufacturer's instructions. Glassware was washed with DEPC-water to eliminate RNases.

cDNA was obtained by *random primer* with the First Strand cDNA Synthesis Kit for reverse transcriptase polymerase chain reaction (RT-PCR; AMV-Roche). Amplification conditions for RT-PCR were as described for PCR.

DNA Sequencing and Sequence Analysis

Polymerase chain reaction products and restriction fragments from lambda phages were cloned into pGEM-T (Promega) and Bluescript II SK (Stratagene) and sequenced with universal primers. In some cases internal specific primers were designed for primer walking. Sequencing was performed in an ALF express DNA automated sequencer (Pharmacia Biotech) and an ABI 373 A (PerkinElmer) automated sequencer. Nucleotide sequences were analyzed with the Wisconsin Package (Genetics Computer Group). Similarity searches were done with algorithms Fasta, BlastN, and BlastX. ClustalW (Thompson, Higgins, and Gibson 1994) was used to perform multiple alignments. AVID was used to align long sequences (>3 kb) and mVISTA to visualize the alignments (Mayor et al. 2000; Dubchak et al. 2000). Untranslated region (UTR) functional elements were searched with UTRscan (Pesole and Liuni 1999).

Maximum likelihood estimates of the number of synonymous and nonsynonymous substitutions per site (d_s and d_N , respectively) were obtained with the program *codeml* of the PAML package (Yang 1997). For hypothesis testing, the base model assumed a rooted tree with constant nucleotide substitution rate (global clock) and a single $\omega = d_N/d_s$ ratio. Codon equilibrium frequencies π_i were estimated from the nucleotide frequencies of the three codon sites (F3X4 option of *codeml*; Yang 1997). The base model included 13 parameters: the transition/transversion rate ratio (κ), the nonsynonymous/synonymous substitution ratio (ω), two relative branch lengths (r_s), and nine nucleotide frequencies.

Results and Discussion

Chromosomal Localization of *labial* in *D. buzzatii* and Other Species in the *Drosophila* Subgenus

To map *labial* in *D. buzzatii*, a fragment encompassing exons 2–3 and including the homeobox was obtained by PCR and used for in situ hybridization to the salivary gland chromosomes. Surprisingly, the *lab* probe hybridized to the dense polytene band G4g-5a of chromosome 2 (fig. 2), separated by about 70 bands from the site (F1c-e) where the other anterior genes of the Hox complex (*pb*, *ftz*, *Antp*) had been previously localized (Ranz, Casals, and Ruiz 2001). Furthermore, *lab* cytological position was almost indistinguishable from that of *abd-A* and *Abd-B* (Ranz, Casals, and Ruiz 2001). The localization of *lab* in *D. buzzatii* chromosomes was later corroborated by hybridization of other probes of the gene (see *Materials and Methods*) and indicates a split of the Hox gene complex between *lab* and *pb*. To date this split and reconstruct the evolutionary history of the Hox gene complex in the *Drosophila* subgenus, the same gene fragment was amplified by PCR and cloned in *D. hydei*, *D. mercatorum*, *D. repleta*, and *D. virilis*. The resulting clones were sequenced and used for

Table 2
Structure of the *labial* Gene in Three *Drosophila* Species and *Tribolium castaneum*

Gene Region	<i>D. buzzatii</i> ^a	<i>D. virilis</i> ^a	<i>D. melanogaster</i> ^b	<i>T. castaneum</i> ^c
Promoter ^d	Inr + DPE	Inr + DPE	Inr + DPE	TATA
5' UTR	757 bp	738 bp	239 bp	181 bp
Exon 1	1,300 bp	1,276 bp	1,216 bp	630 bp
Intron 1	~19 kb	ND	14 kb	11.5 kb
Exon 2	380 bp	380 bp	416 bp	431 bp
Intron 2	687 bp	737 bp	245 bp	—
Exon 3	288 bp	>50 bp	258 bp	—
3' UTR	619–831 bp	ND	647 bp	144 bp

NOTE.—ND: not determined.

^a Data from this work.

^b Accession number NG_000062.

^c Data from Nie et al. (2001).

^d Inr: transcription initiator; DPE: downstream promoter element.

in situ hybridization to the chromosomes of those four species, as well as those of *D. canalinea* and *D. immigrans* (fig. 1). *D. buzzatii*, *D. hydei*, *D. mercatorum*, and *D. repleta* are representative species of four subgroups of the *D. repleta* group (Wasserman 1992), and in all of them *lab* was localized at band G4g-5a of chromosome 2. Likewise, in *D. canalinea*, which belongs to a closely related species group (Wasserman 1992), *Antp* and *lab* mapped to different chromosomal sites. However, in *D. virilis*, a phylogenetically more distant species, *lab* mapped to 24E, the same constricted region of chromosome 2 where the genes *Antp*, *zen*, and *Ubx* have been previously localized (vonAllmen et al. 1996). Finally, five Hox genes (*abd-A*, *Abd-B*, *Antp*, *lab*, and *Ubx*) were mapped by in situ hybridization to the chromosomes of *D. immigrans*, an even more distantly related species. In this species, *lab*, *Antp*, and *Ubx* mapped to the same chromosomal site, whereas *abd-A* and *Abd-B* mapped to another location of the same chromosome (fig. 1). Therefore, we conclude that the *virilis-immigrans* Hox gene organization (including the split between *Ubx* and *abd-A*) appears to be ancestral within the *Drosophila* subgenus. The newly discovered *lab-pb* split occurred before the radiation of the *D. repleta* group, 15–22 MYA, but after the divergence between the *repleta* and *virilis* species groups, 20–30 MYA (Spicer 1988; Russo, Takezaki, and Nei 1995). To our knowledge, the split *lab-pb* is unique among the eukaryotes and provides the opportunity to investigate the evolutionary consequences of natural Hox gene complex rearrangements.

Molecular Structure of *labial* in *D. buzzatii* and *D. virilis*

To determine if the *lab-pb* split has affected in any way the molecular structure of *lab*, this gene was cloned and sequenced in *D. buzzatii* and partially in *D. virilis*. The gene *lab* of *D. melanogaster* (*Dm lab*) has three exons and two introns, intron 1 being rather long (table 2 and fig. 3). Cloning of *lab* in *D. buzzatii* (*Db lab*) was carried out by screening a lambda library of *D. buzzatii* genomic DNA with two probes corresponding to the two distantly located exon 1 and exon 2–3 regions (see *Materials and Methods*). Three positive phages coming from these two regions were subcloned and partially sequenced (fig. 3): λ j-19/21,

including exons 2 and 3, and λ j-19/68 and λ j-19/89, including exon 1. Because those phages did not contain the whole gene, another phage was obtained (λ j-19/77), and the remaining gap (3.5 kb) was closed by PCR amplification (table 1 and fig. 3). Overall, 12,069 bp of *Db lab* were sequenced, corresponding to the entire coding sequence (CDS), intron 2, the 5' UTR, the 3' UTR, and 6,241 bp from the upstream regulatory region (table 2 and fig. 3). Intron 1, with an estimated size of ~19 kb, was only partially sequenced.

In the analysis of the genomic sequence of *Db lab* (12,069 bp) no other conserved coding regions were found apart from those corresponding to the *lab* gene. However, two transposable element insertions were identified: a new copy of the *ISBu2* element (Cáceres, Puig, and Ruiz 2001) of 738 bp and an *ISBu3* element of 993 bp, inserted 5,500 bp and 1,663 bp, respectively, upstream from the *Db lab* transcription start site (fig. 3). Both elements are closely related to the *ISBu1* element of *D. buzzatii* (Cáceres, Puig, and Ruiz 2001), and all three belong to the mini-me class described by Wilder and Hollocher (2001).

In *D. virilis* three genomic fragments of the *lab* gene (*Dv lab*) were amplified by PCR and sequenced (table 2): a 1,172-bp fragment corresponding to exons 2 and 3 and intron 2, a 1,268-bp fragments corresponding to exon 1, and a 1,818-bp fragment containing the 5' UTR. In addition, a 548-bp cDNA fragment comprising the area from the end of exon 1 to the beginning of exon 3 was also cloned and sequenced. Overall, 3,184 bp were sequenced, with only intron 1, part of exon 3, and the 3' UTR remaining.

The molecular structure of *Db lab* and *Dv lab* is similar to that of *Dm lab*, comprising three exons and two introns (table 2 and fig. 3). Therefore, the structure of the *lab* transcription unit has not been altered by the *lab-pb* split in *D. buzzatii* and the exact breakpoint must be located at least 8 kb upstream from the *lab* transcription start point. In contrast, *Db lab* is 32% larger than *Dm lab*, mostly because of changes in noncoding regions. The total length of the CDS is about 4% larger (1,968 bp versus 1,890 bp), whereas intron 1 is ~5 kb longer and intron 2 and the 5' UTR are more than twice as large in *D. buzzatii* (table 2). In general, the *D. virilis* gene is more like that of *D. buzzatii* than that of *D. melanogaster* (table 2), as expected from their phylogenetic relationships (fig. 1). Interestingly, the

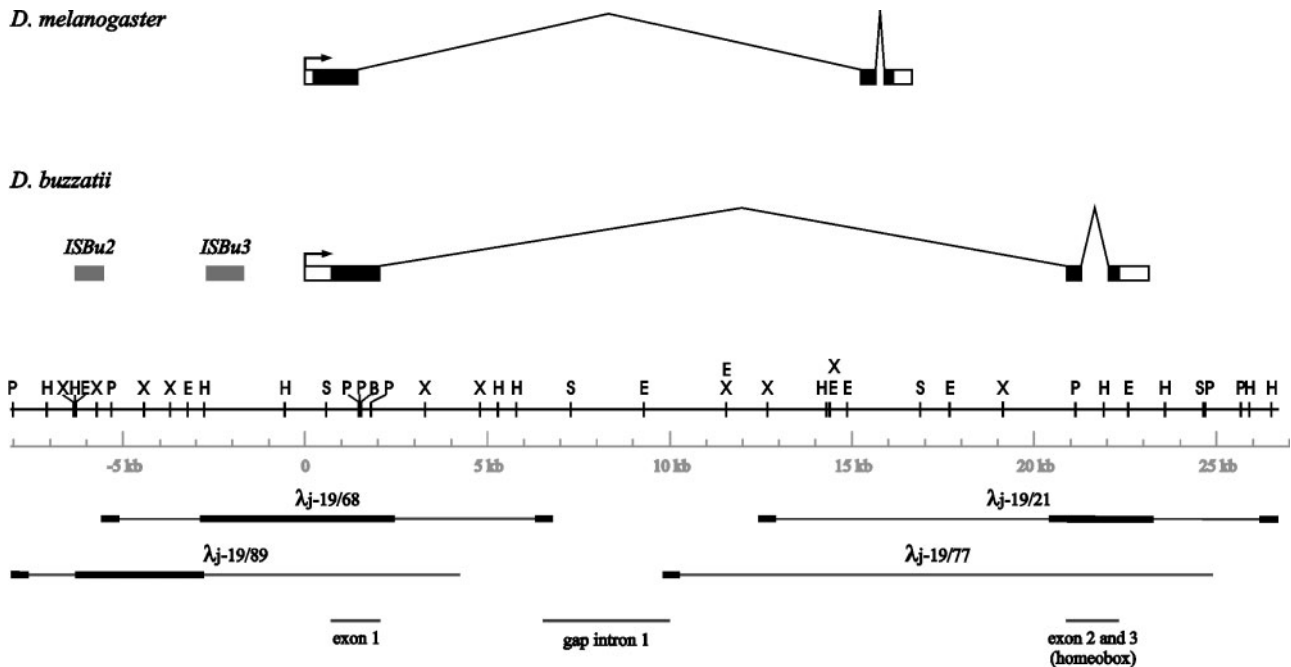


FIG. 3.—Restriction map and molecular organization of the *lab* region in *D. buzzatii*. The lines underneath the map represent lambda clones studied during gene cloning and fragments obtained by PCR amplification. Thick segments of lambda clones represent sequenced regions. The structure of *lab* is shown above the map, for both *D. buzzatii* and *D. melanogaster*. The arrow indicates the start point and direction of transcription, black boxes represent coding regions, and empty boxes show untranslated regions (UTRs). Gray boxes represent transposable elements found in *D. buzzatii*; none has been described in this region in *D. melanogaster*. Restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; P, *Pst*I; S, *Sal*I; X, *Xba*I.

lab structure differs widely between *Drosophila* and *Tribolium castaneum* (table 2). In the latter species, the CDS is 1,061 bp long with two exons only, and intron 1, the 5' UTR, and the 3' UTR are also significantly shorter (Nie et al. 2001).

Expression of *labial* and Identification of Regulatory Sequences

To verify the predicted exon structure of *Db lab*, total RNA from 0–12-h-old embryos was amplified by RT-PCR. Primers were designed on exon 1 and exon 3 (table 1 and fig. 3), allowing us to examine both exon junctions at the same time. The amplified cDNA fragment had the expected size of 570 bp, and its sequence confirmed precisely the predicted exon-intron junctions. In certain *D. melanogaster* lines, an alternative splicing of *lab* intron 1 that produces two mRNAs with 18 nucleotides of difference has been described (Mlodzik, Fjose, and Gehring 1988). This alternative splicing is caused by a small duplication at the end of intron 1 that provides a new functional splice acceptor site and gives rise to the longer form of the mRNA (FlyBase 2002). The analysis of the genomic and cDNA sequences of *Db lab* indicates that in *D. buzzatii* there is no alternative splicing and only the short form of mRNA is present. In *D. virilis*, the sequenced cDNA fragment also corresponds to the short mRNA. To identify the end of the 3' UTR, polyadenylation signals were searched along the sequence 3' from the stop codon. Four possible polyA signals were identified between 623 and 825 bp from the stop codon of *Db lab*. No conserved sequences were found between *Dm lab* and *Db lab* on the 3' UTR.

The open reading frame and translation start site for *Db lab* were predicted by identification of consensus sequences and homology with *Dm lab*. To find the promoter and transcription start sites, 1,308 bp of genomic DNA sequence from the 5' region of *Db lab* were aligned with 630 bp from the 5' region of *Dm lab* (X13104 X12834) including the 5' UTR and 400 additional nucleotides upstream (fig. 4). The most conserved region between the two sequences corresponds to the transcription start site described for *Dm lab* (Mlodzik, Fjose, and Gehring 1988; Diederich et al. 1989). In *Db lab*, we find in this conserved region the sequence TCAGTC that fits well the proposed consensus $TCA_{+1}^G/T^T/C$ of the initiator (Inr) of *Drosophila* genes, where A_{+1} is the transcription start site (Arkhipova 1995). No other sequence similar to the Inr consensus was found in the 5' upstream region of *Db lab*. An additional alignment including the 738-bp upstream sequence of *Dv lab* and using the ClustalW algorithm (Thompson, Higgins, and Gibson 1994) with a gap extension penalty of 0.05 indicated also a high identity around both the transcription (Inr) and the translation (Met) start sites between the three *Drosophila* species (fig. 4). There is a high homology between the Inr and position +32 in the three species (fig. 4), with the sequence CACG located between positions +29 and +32 being fully conserved. This agrees with the characterization of this sequence as the Downstream Promoter Element (DPE) in *Dm lab* (Kutach and Kadonaga 2000). Because no TATA box was found upstream of the transcription start site, we conclude that *Db lab* and *Dv lab* present a TATA-less promoter with Downstream Promoter Element (DPE), as does *Dm lab* and all other homeotic

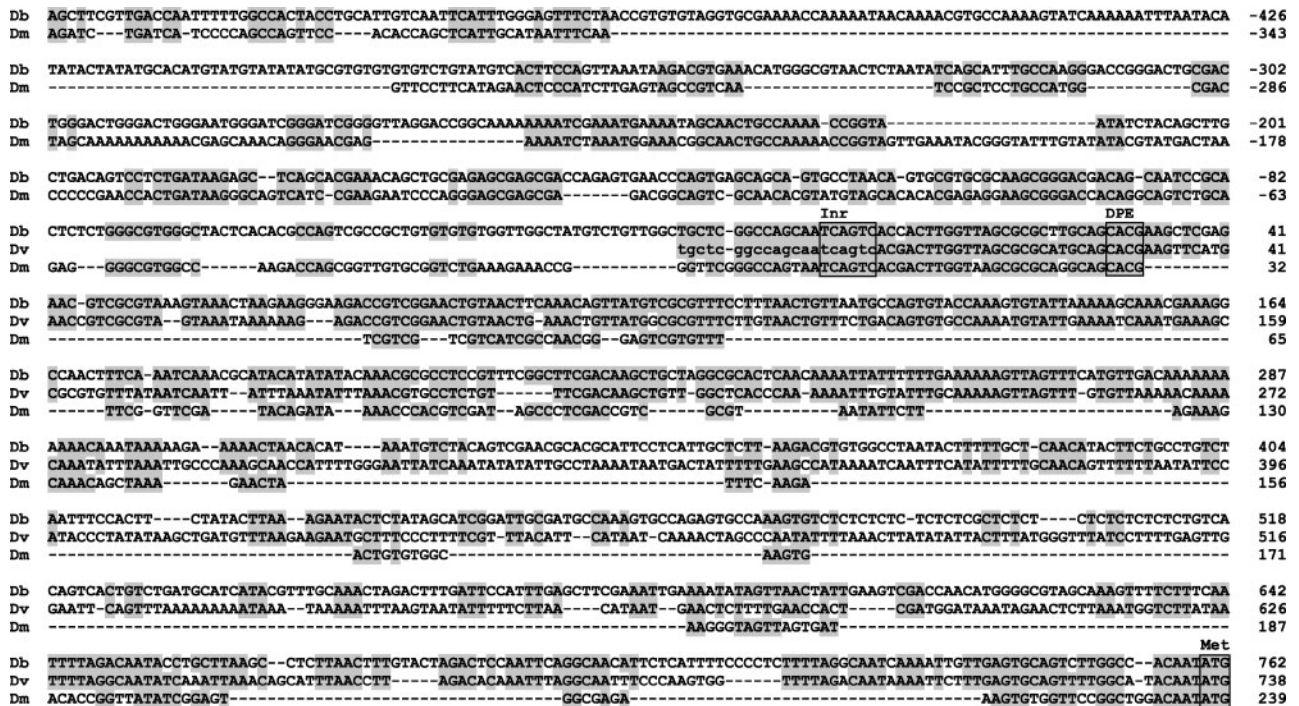


FIG. 4.—Alignment of 5' UTR and upstream sequences of *lab* from *D. buzzatii* (Db), *D. virilis* (Dv) and *D. melanogaster* (Dm). Alignment was performed with ClustalW and revised by hand. The primer used to amplify this region in *D. virilis* is indicated in lowercase. The Initiator (Inr), downstream promoter element (DPE), and first codon of the protein (Met) are shown. Shaded regions show conservation of three nucleotides in a window of four in at least two sequences.

genes described in *Drosophila* (Nie et al. 2001). Analysis with the UTRscan software (Pesole and Liuni 1999) of the 5' UTRs of the three *Drosophila* species identified internal ribosome entry site (IRES) elements in all of them. These IRES elements promote cap-independent translation in different genes (Hellen and Sarnow 2001). In the *Hox* genes *Ubx* and *Antp*, they regulate translation during development (Ye et al. 1997).

Given the time elapsed since the separation of the Sophophora and *Drosophila* subgenera (40–60 Myr), no significant nucleotide identity should be expected in noncoding regions in the absence of selective pressure to conserve the function (O'Neil and Belote 1992). Thus, conserved structures may be interpreted as functionally important for gene expression. The entire sequence for *Db lab* (12,069 bp; fig. 3) was aligned with *Dm lab* (NG_000062) with the AVID algorithm and visualized with the VISTA software (Mayor et al. 2000; Dubchak et al. 2000). The resulting graph (fig. 5) highlights those segments with a nucleotide identity >50% between *Db lab* and *Dm lab*. Conserved regions both in coding and noncoding sequences are shown relative to their position in *Db lab*. There is a high variability in conservation along the coding sequence. Several peaks of conserved sequence appear along exon 1, which could be a consequence of the folding pattern of the protein. Two peaks of conserved sequence at the end of exon 2 and the beginning of exon 3 correspond to the homeobox. In the noncoding regions, two prominent peaks of conserved sequence show up, one segment 180-bp long at the end of intron 1 with an identity

of 82.8%, and a second segment 108 bp long in the 5' region with 84.3% identity. In addition, downstream of this second conserved segment there is a ~750-bp region with 50%–75% identity. A Blast search with those conserved sequences from *D. buzzatii* shows similarity only with the same sequence in *D. melanogaster*.

We have compared the position of those noncoding conserved sequences with the regulatory elements described for *Dm lab* (Chouinard and Kaufman 1991). The conserved regions upstream of *lab* are inside the 2.4-kb fragment that promotes the initiation of procephalon expression. The small fragment with 50%–65% identity located 4,800 bp upstream of *Db lab* corresponds to the 1.2-kb *Dm lab* fragment that drives expression in the anterior midgut via a decapentaplegic (DPP)-regulated enhancer (lab550 – Marty et al. 2001). Finally, the 180-bp conserved segment at the end of intron 1 corresponds to a fragment not tested by Chouinard and Kaufman (1991), which might contain the regulatory element required for embryonic peripheral nervous system (PNS) expression.

Nucleotide Sequence Comparisons

The coding sequences of *Db lab* and *Dv lab* were aligned and compared to that of *Dm lab* (NM_057265). Table 3 shows the number of synonymous and non-synonymous substitutions per site (d_s and d_n) for pairwise comparisons between the three *Drosophila* species. These numbers have been estimated using maximum likelihood methods (Yang and Bielawski 2000; Yang and Swanson

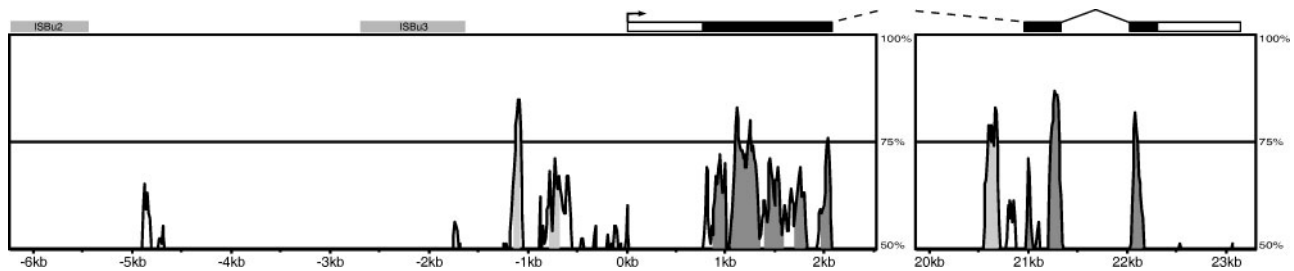


FIG. 5.—VISTA plot showing similarity between the nucleotide sequences of the *lab* gene of *D. melanogaster* and *D. buzzatii*. Conserved regions are shown relative to their position in *Db lab* (horizontal axis), and their percent identities (50%–100%) are indicated on the vertical axes. The location of coding exons (black rectangles) and 5' and 3' UTRs (empty rectangles) are shown above the profile, where the thin arrow indicates the direction of transcription and gray boxes represent transposable elements. Peaks with an identity of at least 70% over 100 bp are shaded.

2002) for the entire coding sequence (CDS) as well as for three separate (non-overlapping) regions: exon 1, exon 2 + 3 (excluding the homeobox) and the homeobox. The ratio between the number of nonsynonymous and synonymous substitutions ($\omega = d_N/d_S$) provides information on the type of selection acting on amino acid sequences. The low average ω values for the entire gene (0.0696–0.0910) suggest a relatively high degree of functional constraint, but there is wide variation among regions (see below).

To investigate if the molecular evolution of *lab* had been affected by the *lab-pb* split in the *D. repleta* species group, three hypotheses related to the molecular evolution of *lab* were tested by comparing the likelihood of different models. First, we tested the constancy of the evolutionary rates by comparing the base model (which assumes a molecular clock and a single ω ratio for all lineages) with a model assuming no clock. The results were nonsignificant for the entire CDS and the three separate coding regions (table 4). Second, we tested for variation in the ω ratio among lineages by comparing the base “one ratio” model with a “free-ratio” model, involving as many ω ratios as branches in the tree (four). The results again were nonsignificant both for the entire CDS and for the three separate coding regions (table 4). Therefore, the evolutionary rate of *lab* has not accelerated or decelerated significantly after its separation from the anterior Hox gene

complex. Likewise, the ω ratio is not different in the phylogenetic branch leading to *D. buzzatii* than in the rest of *Drosophila*, and there is no evidence that positive selection or relaxation of selection has occurred after the relocation of *lab* by the *lab-pb* split.

Finally, we tested for differences among the three separate coding regions, fitting a series of fixed-sites models as described by Yang and Swanson (2002). A detailed list of the models and their results is given in table 5. We were particularly interested in the among-region variation in ω ratio. To test the hypothesis of no variation, model D (which assumes different r_s , κ , and ω) was compared with model B (which assumes different r_s). The log-likelihood difference, $2(l_D - l_B) = 71.65$, is highly significant ($df = 19 - 15 = 4$; $P < 0.001$) indicating differences either in the transition/transversion ratio κ or the ratio ω . A modified D model (designated D') with a fixed κ value was compared to model B, and the resulting log-likelihood difference, $2(l_{D'} - l_B) = 64.36$, showed that the variation in ω ratio was indeed significant ($df = 16 - 15 = 1$; $P < 0.001$). Similar tests can be conducted comparing model E (which assumes different r_s , κ and ω , and π_s) with model C (which assumes different r_s and π_s). The difference in log-likelihood, $2(l_E - l_C) = 85.06$, is highly significant ($df = 37 - 33 = 4$; $P < 0.001$), indicating again significant differences either in the transition/transversion ratio κ or the ratio ω . If a modified E model

Table 3
Maximum Likelihood Estimates of the Number of Synonymous (d_S) and Nonsynonymous (d_N) Substitutions per Site in the Coding Regions of the Gene *labial* for Pairwise Comparisons Between *D. buzzatii*, *D. virilis*, and *D. melanogaster*

Comparison	Region ^a	Codons	t	κ	S	d_S	N	d_N	ω
<i>D. buzzatii</i> - <i>D. virilis</i>	CDS	545	0.7982	1.9050	424.2	0.8140	1,210.8	0.0741	0.0910
	Exon 1	413	0.9009	1.7786	340.3	0.9053	898.7	0.0712	0.0786
	Exon 2 + 3	78	0.7204	2.7602	56.9	0.4878	177.1	0.1605	0.3289
	homeobox	54	0.7578	0.4796	20.7	1.9608	141.3	0.0020	0.0010
<i>D. buzzatii</i> - <i>D. melanogaster</i>	CDS	595	2.4155	1.5918	419.3	2.6722	1,365.7	0.2318	0.0868
	Exon 1	395	3.0631	1.4996	286.7	3.4913	898.3	0.2325	0.0666
	Exon 2 + 3	140	2.0560	1.9818	111.7	1.4648	308.3	0.4028	0.2750
	homeobox	60	0.8506	1.1505	23.3	2.1726	156.7	0.0022	0.0010
<i>D. virilis</i> - <i>D. melanogaster</i>	CDS	525	2.5782	1.4040	345.4	3.1407	1,229.6	0.2186	0.0696
	Exon 1	395	3.6802	1.2985	274.2	4.4807	910.8	0.2470	0.0551
	Exon 2 + 3	76	1.7824	1.8819	49.9	1.5643	178.1	0.3223	0.2061
	homeobox	54	1.7454	4.8507	34.7	2.7077	127.3	0.0027	0.0010

NOTE.—t: branch length; κ : transition/transversion rate ratio; S: number of synonymous sites; N: number of nonsynonymous sites; ω : d_N/d_S ratio.

^a Exon 2 + 3 does not include homeobox.

Table 4
Likelihood Ratio Tests of Variation in the Rate of Molecular Evolution and the ω Ratio of the Gene *labial* in the Genus *Drosophila*

Region ^a	Difference in Log-Likelihood from Base Model ^b			
	Number of Codons	No Clock (df = 1)	Lineage-specific ω ratio (df = 3)	ω Ratio in Base Model
CDS	695	0.98	2.43	0.0971
Exon 1	449	3.66	3.55	0.0766
Exon 2 + 3	186	1.19	3.17	0.3374
Homeobox	60	0.28	0	0.0001

^a Exon 2 + 3 does not include homeobox.

^b Base model assumes a molecular clock and a single ω ratio for all lineages.

(designated E') with a fixed κ value is compared to model C, then the likelihood difference, $2(l_{E'} - l_C) = 81.03$, also provides evidence for significant variation in ω ratio among the three coding regions (df = 34–33 = 1; $P < 0.001$). The homeobox, which is fully conserved between the three *Drosophila* species, shows the lowest ω ratio, which implies a strong constraint. The portion of exons 2–3 after excluding the homeobox shows the highest ω ratio (around 0.3 depending on the model; table 5), which indicates a substantially lower constraint. Finally, the ω ratio is intermediate in exon 1 (around 0.08 depending on the model; table 5).

Amino Acid Sequence Comparisons

Seven different proteins encoded by *lab* orthologous genes were aligned with the ClustalW algorithm (fig. 6). These include the complete Labial sequences from *D. buzzatii* (present work), *D. melanogaster* (Mlodzik, Fjose, and Gehring 1988; Diederich et al. 1989), *T. castaneum* (Nie et al. 2001), and *HoxA1* gene from *Homo sapiens* (Hong et al. 1995) and partial Labial sequences from *D. virilis* (present work), *Anopheles gambiae* (Powers et al. 2000) and *Thermobia domestica* (Peterson et al. 1999). Complete sequences were aligned with default parameters. To avoid artifacts, partial sequences were aligned one by one with *D. buzzatii* increasing gap extension penalty and included in the global alignment afterwards. The most intriguing observation in this comparison is the much larger size of

the *Drosophila* Labial proteins (629–655 aa) compared to those of *T. castaneum* (353 aa) or *H. sapiens* (335 aa). Whereas Labial has almost duplicated its size in *Drosophila* in relation to the other species, other Hox genes do not show this variation. For example, the gene *abd-A* presents a similar size in *T. castaneum* and *D. melanogaster* (Shippy, Brown, and Denell 1998) and it is only slightly smaller in *A. gambiae* (Devenport, Blass, and Eggleston 2000).

The most conserved region between all Labial proteins is the 65-aa homeodomain, located at the C-terminal region, with an identity between 83% and 100%. Outside the homeodomain there are some other conserved regions. The YKWM motif, a variant of the core of the hexapeptide motif, YPWM, involved in the interaction with the cofactor EXD (Gehring, Affolter, and B r glin 1993; Rauskolb and Wieschaus 1994, Johnson, Parker, and Krasnow 1995), presents an identity between 60% and 100% through 15 aa. There is also a short region around the YTNLD motif with an identity between 44% and 100%. Finally the N-terminal region presents between 50% and 100% identity. The conservation of these amino acid sequences across long phylogenetic distances suggests that they suffer functional constraints and are probably involved in highly conserved interactions, such as the YKWM domain (Johnson, Parker, and Krasnow 1995). These results are fully consistent with the previous nucleotide comparison between *Drosophila* species. The homeodomain is the most conserved region of the protein and shows a high level of nucleotide sequence evolution constraint, whereas all other conserved domains are located in exon 1, which is more constrained than exons 2 and 3 (excluding the homeobox).

Outside those conserved domains there is a very low conservation, which is surprising when compared with other proteins. The Labial proteins from *D. buzzatii* and *D. melanogaster* have an overall identity of only 62.8%, much lower than the 95% identity found by Hooper et al. (1992) between the Antp proteins from *D. virilis* and *D. melanogaster* (with the same divergence time). There are very few direct comparisons of homologous genes between *D. melanogaster* and *D. buzzatii*, but *D. melanogaster* and *D. virilis* homologous proteins usually have an identity between 80% and 90%, with values ranging from 36% to 97% (O'Neil and Belote 1992). Interestingly, all regions

Table 5
Log-Likelihood Values and Parameter Estimates Under Fixed-Sites Models

Model	<i>p</i>	<i>l</i>	<i>r</i> ₂	<i>r</i> ₃	κ	ω
A (homogeneous model)	13	−5147.69	1	1	1.6552	0.0971
B (different <i>r</i> s)	15	−5125.17	1.569	0.387	1.6433	0.0862
C (different <i>r</i> s and π s)	33	−5087.45	1.398	0.277	1.6547	0.0813
D (different <i>r</i> s, κ , and ω)	19	−5089.34	0.942	0.983	$\kappa_1 = 1.4036$; $\kappa_2 = 0.0853$; $\kappa_3 = 2.5631$	$\omega_1 = 0.2857$; $\omega_2 = 0.7401$; $\omega_3 = 0.0001$
D' (κ fixed to 1.643)	16	−5092.99	0.888	0.966	$\kappa = 1.643$	$\omega_1 = 0.0883$; $\omega_2 = 0.2718$; $\omega_3 = 0.0001$
E (different <i>r</i> s, κ and (ω), and π s)	37	−5044.93	0.726	0.613	$\kappa_1 = 1.4864$; $\kappa_2 = 0.0809$; $\kappa_3 = 2.3396$	$\omega_1 = 0.3255$; $\omega_2 = 1.1217$; $\omega_3 = 0.0001$
E' (κ fixed to 1.655)	34	−5046.94	0.703	0.670	$\kappa = 1.655$	$\omega_1 = 0.0824$; $\omega_2 = 0.3107$; $\omega_3 = 0.0001$
F (separate analysis)	39	−5042.39			$\kappa_1 = 1.4978$; $\kappa_2 = 2.3044$; $\kappa_3 = 1.0623$	$\omega_1 = 0.0766$; $\omega_2 = 0.3374$; $\omega_3 = 0.0001$

NOTE.—*p*: number of parameters of the model; *l*: log-likelihood value; *r*₂ and *r*₃: relative branch lengths.

	N-terminal		
Db	MMDVSSMYGNPHHHPHATATADGYGSNYFPSP--THHQQQQQQQ---QHQQQQQQQQLQ---QQQLAYNGYDSN	69	
Dv	MMDVSSMYGNPHHHPHATATADGYGSNYFPSP--THHQQQQQQQQQQQQQLHQQQQQLQQVQQHQLAYSGYDSN	76	
Dm	MMDVSSMYGNPHHH-HPHANAIDGYSTTTASAANASSYFAPQHQPHLQLQQQQQHHLQQP---QQHLTYNGYESS	74	
Ag			
Tc	MMMMDVG-MYGHHSQHGTYPSDNYN-----YTSD	29	
Td			
HsHoxA1	MDNARMNSFLEYPIILSSGD-----	19	
Db	SP-NYYQQ---QLTPPPAQHQLOHPLAQHQQQQQQQQQQQQQQLYPHSHLFSPSAAEYGITTTAAVAASNSTPLHPT	145	
Dv	SP-NYYQQQQ-QLTPPPAQHQLOHPLAQ-----QQQQQQQQQLYPHSHLFSPSAAEYGITTTAAVAAPNSTPLHPT	146	
Dm	SPGNYYPQQQAQLTPPTSSSHQVQQHQ-----QQQQAQQQLYPHSHLFSPSAAEYGITTTSTTTGNP-GKPLHPS	144	
Ag			
Tc	IP-N-----AQQMPPHYSNYHYEDP---VGYTGS-----EET	57	
Td			
HsHoxA1	-----SGTCSARAYPSDHRITTFQSCAVSANS CGGDDRFLVGRGVQ	60	
Db	SHSPADSYDSDSVNPFYATAAATVPPANNPVTANATSQQQQQQQHQPTLDAHTPIISSENGMMYTNDCLYS	225	YTNDL
Dv	SHSPADSYYESDSVHSYATAAATVPPANNPVTANATQQQQQQQ---ALDAHAPIISSENGMMYTNDCLYS	222	
Dm	SHSPADSYYESDSVHSYATAAATVAPPSSNPITANASATSNTOQQQQ---QAAIISSENGMMYTNDCLYS	216	
Ag			
Tc	PPSPQEVVYHHQTI-----PQDN-----PIINTETGLSYTNLDYDGS	94	
Td			
HsHoxA1	IGSPHHHHHHHHH-----PQP-----ATYQTSGNLGVSYSHSSCGPS	98	
Db	PNVAQ---TQPHVYAGQMEEKYAAVLHAGYAAAPGMLLEDPSDAMLQLAVGQQQSPQMWHHHQHMGGYALDAGIAMD	302	
Dv	PSVAQQAQTQPHGYGGQMEEKYAAVLHAGYAAPPGMLLEDP-DSMMQLAAGQQQSPQMWHHHQHMSGGYPLDAGIAMD	299	
Dm	TAQAQ---APVHGAYAGIEEKYAAVLHASA---PGMVLED-QDPMMQAT---QS---QMWHHQQHLAGSYALDAMD SLG	284	
Ag			
Tc	-----NFPPINHONY-----N	105	
Td			
HsHoxA1	-----YGSQNFSAF-----Y	108	
Db	YPHAHLHSHSLSHGHISLTSSAQQQQQQQQQQLQQQQQLQQLQAHATGAPAVVVSQHSISPSSNNCLSMANQRGGGV	382	
Dv	YPHAHLHSHSLSHGHISLTSSAQQAQQQQQQQQQQQQQQQ---ANSSQAQVVPQHSISPSSNCTMARNHRSRGV	374	
Dm	-MHAHMHGHLPHGHGLGNLANNPHQQQPVQVQQQQQPHQPHQHPQNSPAAHQHHSVSPNGG---MNRQQRGG-VI	357	
Ag			
Tc	EPIYQRTHEENMLLRHHEEVADGHQHHLHENKYQLEVDSGYHLGASSTPTNSCMEYQH-----LHRYKEEA-LQ	172	
Td			
HsHoxA1	SPYALNQEADVSGGYPQCAPAVYSGNLSSPMVQHSHHHHGGYAG-GAVGSPQYIHHSYG-----QEHQSLALAT	175	
Db	SPGSSSTSSSTASVSAANSAHQANSQSKSPNHATSIPTYKWMQLKRNVPKQPAPKLPAN-MSNMHDYQL-NGLHDMCR---	457	YKWM
Dv	SPGSSSTSSSTASVSAANSTHQASSQSKSPNHAAIPTKYKWMQLKRNVPKQPAPKLPAN-MSNMHDYQL-NGLHDMCRAGN	452	
Dm	SPGSSSTSSST---SASNGAHPASTQSKSPNHSSSIPTYKWMQLKRNVPKQPAPKLPASGIASMDYQM-NGQLDMCRGGG	433	
Ag	GDGRMR-----QHVMHNLSSVPQPQPVLPPTYKWMQVKNVPKPTVEKIPF-----AEFPTT-----	29	
Tc	GDGRMR-----QHVMHNLSSVPQPQPVLPPTYKWMQVKNVPKPTVEKIPF-----AEFPTT-----	223	
Td	GDGRMR-----QHVMHNLSSVPQPQPVLPPTYKWMQVKNVPKPTVEKIPF-----AEFPTT-----	27	
HsHoxA1	YNNLSLPLHA-----SHQEACRSPASETSSPAQTFDWMKVKNPPKT---GKVG-----EYGY-----	225	
Db	-GGSSA-----VLPIQNPLLP-NSATAGVTTNGLAIGSGGGAAAVAGGAGGSPAGVGLSNCSSLSS-TN	517	
Dv	VGGSA-----VLPIQNPLLSNNAAGSASSGLGVG-----LASGNGAAGASAGVGLSGCSSLSS-TN	509	
Dm	GGGSDVGGSPVGVGGNGSPGIGGVLVSQNSLIMANSAAGSAHPNGMG-----VGLSGSGSLSSCSSLSSNTN	501	
Ag	-----PTHVLDPLR--VP-NHTTV-----VAPSAVSPHQSSFMINNN	63	
Tc	-----SSSSALESPNP--ASRSSCLGSN-----VAPSAVSPHQSSFMINNN	254	
Td	-----VNGVG-----AGS-----	35	
HsHoxA1	-----LGQP-----	229	
Db	NSGRNFTNKQLTELEKEFHFNRYLTRARRIEIANTLQNLNETQVKIWFQNRMRKQKKRVKEGLIPADMITQHSAPVLTEK	597	
Dv	NSGRNFTNKQLTELEKEFHFNRYLTRARRIEIANTLQNLNETQVKIWFQNRMRKQKKRVKEGLIPADMITQHSAPVLTEK	568	
Dm	NSGRNFTNKQLTELEKEFHFNRYLTRARRIEIANTLQNLNETQVKIWFQNRMRKQKKRVKEGLIPADILTQHSSTVISEK	581	
Ag	STGRNFTNKQLTELEKEFHFNRYLTRARRIEIANTLQNLNETQVKIWFQNRMRKQKKRVKEGLIPPEIQNQSS--PKHSSL	141	
Tc	NTGRNFTNKQLTELEKEFHFNRYLTRARRIEIASALQNLNETQVKIWFQNRMRKQKKRVKEGLIPPE-----PISASL	327	
Td	NAGRTNFTNKQLTELEKEFHFNRYLTRARRIEIASALQNLNETQVKIWFQNRMRKQKKRVKEGLIPPE-----PISASL	87	
HsHoxA1	NAVRTNFTTKQLTELEKEFHFNRYLTRARRIEIASALQNLNETQVKIWFQNRMRKQKKRVKEGLIPPE-----PISATPPG--NDEK	306	
Db	QPLQHQQQQQQQQQQQQQPQQQQPDLKASEITSNELPATAAAAAASTAAATTTAIAK	655	
Dv			
Dm	PP-----QQQPQPPELQLKSQSGDLGGNEL-ATGAPSTPTTAMTLTAPTSKQS	629	
Ag	IN-----SNDSSASSSTATLTAIPASAGAS-ATNTSTLNE-SNENSRESITM	187	
Tc	ST-----SN-SNSPSS-----TN-LIQNEASNENSRESN	354	
Td			
HsHoxA1	AESSEKSSSSPCVSPSGSSTSDTLTTS	335	

FIG. 6.—Multiple alignment of proteins coded by *lab* orthologs from different organisms. *D. buzzatii* (Db) and *D. virilis* (Dv) from present work. GenBank accession numbers: *D. melanogaster* (Dm, CAB57787), *A. gambiae* (Ag, AAF91398.1 and AAAB01008961), *Tribolium castaneum* (Tc, AAF64147.1 and AAF64148.1), *Thermobia domestica* (Td, AAD50360.1), *Homo sapiens* HoxA1 (HsHoxA1, NP_005513.1). Dv, Ag, and Td are partial sequences. Amino acids in lowercase correspond to degenerate primers. Alignments were performed with ClustalW with default parameters, except for Ag and Td, which have been aligned individually with Db, increasing the gap extension penalty and included in the global alignment

conserved between *Db lab* and *Dm lab* are also more or less conserved in all other sequences studied, including the human gene *HoxA1*. The domain conservation suggests that the protein functionality relies basically on a few domains, leaving the rest of the sequence with more freedom to change. This fact could explain why minigenes carrying the chicken *Hoxb1* gene are able to rescue null mutants for the *lab* gene in *D. melanogaster* (Lutz et al. 1996). Thus, Labial proteins seem to be built on highly conserved blocks joined by rapidly evolving sequences. For example, the spacer or "gene specific region" (Peterson et al. 1999) localized between the YKWM motif and the homeodomain is much more variable, both in sequence and in length, in the *lab* gene than in other Hox genes even between *Drosophila* species.

Concluding Remarks

Although Hox gene colinearity and cluster conservation represents a paradigm in Evo-Devo, in the genus *Drosophila* the Hox cluster has been split at least three times. The *lab-pb* split described here is the most recent example. Even after moving to a new location, *labial* appears to evolve consistently, showing no alteration of its molecular evolutionary patterns. A paracentric inversion as well as a gene transposition event might, in principle, account for the *lab-pb* disruption and the relocation of *lab* near the genes *abd-A* and *Abd-B* observed in the *repleta* group of *Drosophila*. However, the former possibility seems more likely given the extensive reorganization of the *Drosophila* genome produced by chromosomal inversions (Powell 1997; Ranz, Casals, and Ruiz 2001). The molecular characterization of the chromosomal regions 5' of *lab* and 3' of *pb* in *D. buzzatii* should help to determine which was the molecular mechanism that caused such a split and the precise location of the breakpoint between *lab* and *pb*. This study would also elucidate the exact physical relationship, if any, between *lab* and the genes *abd-A* and *Abd-B*. Furthermore, although the rearrangement has not altered the *lab* coding region, and although some of its regulatory sequences seem to be in place, whether the relocation of *lab* had any consequences upon the level or spatiotemporal pattern of expression of this gene remains an open question, one that we are already investigating.

Acknowledgments

We thank S. Celniker, J. E. Hooper, and E. Sánchez-Herrero for probes; the National *Drosophila* Species Resource Center (Bowling Green, Ohio, USA) for fly strains; and J. González and E. Betrán for advice and helpful comments on the manuscript. This work was supported by grant PB98-0900-C02-01 from the Dirección General de Enseñanza Superior e Investigación Científica (MEC, Spain) awarded to A.R. and by a doctoral FI/DGR fellowship from the Generalitat de Catalunya awarded to B.N.

Literature Cited

Adams, M. D., S. E. Celniker, R. A. Holt et al. (192 co-authors). 2000. The genome sequence of *Drosophila melanogaster*. *Science* **287**:2185–2195.

Arkhipova, I. R. 1995. Promoter elements in *Drosophila*

melanogaster revealed by sequence analysis. *Genetics* **139**:1359–1369.

Cáceres, M., M. Puig, and A. Ruiz. 2001. Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. *Genome Res.* **11**:1353–1364.

Chouinard, S., and T. C. Kaufman. 1991. Control of expression of the homeotic *labial* (*lab*) locus of *Drosophila melanogaster*: evidence for positive and negative autogenous regulation. *Development* **113**:1267–1280.

Cook, C. E., M. L. Smith, M. J. Telford, A. Bastianello, and M. Akam. 2001. Hox genes and the phylogeny of arthropods. *Curr. Biol.* **11**:759–763.

DeRosa, R., J. K. Grenier, T. Andreeva, C. E. Cook, A. Adoutte, M. Akam, S. B. Carroll, and G. Balavoine. 1999. Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399**:772–776.

Devenport, M. P., C. Blass, and P. Eggleston. 2000. Characterization of the Hox gene cluster in the malaria vector mosquito, *Anopheles gambiae*. *Evol. Dev.* **2**:326–339.

Diederich, R. J., V. K. L. Merrill, M. A. Pultz, and T. C. Kaufman. 1989. Isolation, structure, and expression of *labial* a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev.* **3**:399–414.

Dubchak, I., M. Brudno, G. G. Loots, L. S. Pachter, C. Mayor, E. M. Rubin, and K. A. Frazer. 2000. Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Res.* **10**:1304–1306.

Duboule, D., and G. Morata. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**:358–364.

Duncan, I. 1987. The *Bithorax* complex. *Annu. Rev. Genet.* **21**:285–319.

Finnerty, J. R., and M. Q. Martindale. 1998. The evolution of the Hox cluster: insights from outgroups. *Curr. Opin. Genet. Dev.* **8**:681–687.

FlyBase 2002. The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res.* **30**:106–108. (<http://flybase.bio.indiana.edu/>)

Gehring, W. J., M. Affolter, and T. Bürglin. 1993. Homeodomain proteins. *Annu. Rev. Biochem.* **63**:487–526.

Gehring, W. J., and Y. Hiromi. 1986. Homeotic genes and the homeobox. *Annu. Rev. Genet.* **20**:147–173.

Gubenko, I. S., and M. B. Evgen'ev. 1984. Cytological and linkage maps of *Drosophila virilis* chromosomes. *Genetics* **65**:127–139.

Hartl, D. L., and E. R. Lozovskaya. 1994. Genome evolution: between the nucleosome and the chromosome. Pp 579–592 in B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds. *Molecular ecology and evolution: approaches and applications*, Birkhauser Verlag, Basel.

Hazlerigg, T., and T. C. Kaufman. 1982. Revertants of dominant mutations associated with the Antennapedia gene complex of *Drosophila melanogaster*: cytology and genetics. *Genetics* **105**:581–600.

Hellen, C. U. T., and P. Samow. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **15**:1593–1612.

Hong, Y. S., S. Y. Kim, A. Bhattacharya, D. R. Pratt, W. K. Hong, and M. A. Tainsky. 1995. Structure and function of the *Hox A1* human homeobox gene cDNA. *Gene* **159**:209–214.

Hooper, J. E., M. Pérez-Alonso, J. R. Bermingham, M. Prout, B. A. Rocklein, M. Wagenbach, J. E. Edstrom, R. de Frutos, and M. P. Scott. 1992. Comparative studies of *Drosophila Antennapedia* genes. *Genetics* **132**:453–569.

- Hughes, C., and T. C. Kaufman. 2002. Hox genes and the evolution of the arthropod body plan. *Evol. Dev.* **4**:459–499.
- Johnson, F. B., E. Parker, and M. A. Krasnow. 1995. Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc. Natl. Acad. Sci. USA* **92**:739–743.
- Kaufman, T. C., M. A. Seeger, and G. Olsen. 1990. Molecular and genetic organization of the *Antennapedia* gene complex of *Drosophila melanogaster*. *Adv. Gen.* **27**:309–362.
- Kutach, A. K., and J. T. Kadonaga. 2000. The downstream promoter element DPE appears to be as widely used as the TATA Box in *Drosophila* core promoters. *Mol. Cell Biol.* **20**:4754–4764.
- Kwiatowski, J., and F. J. Ayala. 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. *Mol. Phylogenet. Evol.* **13**:319–328.
- LeCalvez. 1953. Carte cytologique des chromosomes geants des glandes salivaires de *Drosophila immigrans* Sturt. *Chromosoma* **6**:170–174.
- Lewis, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**:565–570.
- Lutz, B., H. C. Lu, G. Miller, and T. C. Kaufman. 1996. Rescue of the *Drosophila labial* null mutant by chicken ortholog *Hoxb-1* demonstrates that the function of Hox genes is phylogenetically conserved. *Genes Dev.* **10**:176–184.
- Maier, D., A. Preiss, and J. R. Powell. 1990. Regulation of the segmentation gene *fushi tarazu* has been functionally conserved in *Drosophila*. *EMBO J.* **9**:3957–3966.
- Maier, D., D. Sperlich, and J. R. Powell. 1993. Conservation and change of the developmentally crucial *fushi tarazu* gene in *Drosophila*. *J. Mol. Evol.* **36**:315–326.
- Martin, C. H., C. A. Mayeda, C. A. Davis, C. L. Ericsson, J. D. Knafels, D. R. Mathog, S. E. Celniker, E. B. Lewis, and M. J. Palazzolo. 1995. Complete sequence of the bithorax complex of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **92**:8398–8402.
- Marty, T., M. A. Viganò, C. Ribeiro, U. Nussbaumer, N. C. Grieder, and M. Affolter. 2001. A HOX complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer. *Development* **128**:2833–2845.
- Mayor, C., M. Brudno, J. R. Schwartz, A. Poliakov, E. M. Rubin, K. A. Frazer, L. S. Pachter, and I. Dubchak. 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**:1046–1047.
- McGinnis, W., and R. Krumlauf. 1992. Homeobox genes and axial patterning. *Cell* **68**:283–302.
- Mlodzik, M., A. Fjose, and W. J. Gehring. 1988. Molecular structure and spatial expression of a homeobox gene from the *labial* region of the *Antennapedia*-complex. *EMBO J.* **7**:2569–2578.
- Nie, W., B. Stronach, G. Panganiban, T. Shippy, S. Brown, and R. Denell. 2001. Molecular characterization of *Tclabial* and the 3' end of the *Tribolium* homeotic complex. *Dev. Genes Evol.* **211**:244–251.
- O'Neil, M. T., and J. M. Belote. 1992. Interspecific comparison of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics* **131**:113–128.
- Pesole, G., and S. Liuni. 1999. Internet resources for the functional analysis of 5' and 3' untranslated regions of eukaryotic mRNAs. *Trends Genet.* **15**:378.
- Peterson, M. D., B. T. Rogers, A. Popadić, and T. C. Kaufman. 1999. The embryonic expression pattern of *labial*, posterior homeotic complex genes and the *teashirt* homologue in an apterygot insect. *Dev. Genes Evol.* **209**:77–90.
- Piñol, J., A. Francino, A. Fontdevila, and O. Cabré. 1988. Rapid isolation of *Drosophila* high molecular weight DNA to obtain genomic libraries. *Nucleic Acids Res.* **16**:2736–2737.
- Powell, J. R. 1997. Progress and prospects in evolutionary biology: the *Drosophila* model. Oxford University Press, New York.
- Powers, T. P., J. Hogan, Z. Ke, K. Dymbrowski, X. Wang, F. H. Collins, and T. C. Kaufman. 2000. Characterization of the Hox cluster from the mosquito *Anopheles gambiae* (Diptera: Culicidae). *Evol. Dev.* **2**:311–325.
- Randazzo, F. M., M. A. Seeger, C. A. Huss, M. A. Sweeney, J. K. Cecil, and T. C. Kaufman. 1994. Structural changes in the *Antennapedia* complex of *Drosophila pseudoobscura*. *Genetics* **133**:319–330.
- Ranz, J. M., F. Casals, and A. Ruiz. 2001. How maleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* **11**:230–239.
- Ranz, J. M., J. González, F. Casals, and A. Ruiz. 2003. Low occurrence of gene transposition events during the evolution of the genus *Drosophila*. *Evolution* **57**:1325–1335.
- 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.
- Rauskolb, C., and E. Wieschaus. 1994. Coordinate regulation of downstream genes by *extradenticle* and the homeotic selector proteins. *EMBO J.* **13**:3561–3569.
- Russo, C. A. M., N. Takezaki, and M. Nei. 1995. Molecular phylogeny and divergence times of *Drosophilid* species. *Mol. Biol. Evol.* **12**:391–404.
- Ruvkun, G., and O. Hobert. 1998. The taxonomy of developmental control in *Caenorhabditis elegans*. *Science* **282**:2033–2041.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Shippy, T. D., S. J. Brown, and R. E. Denell. 1998. Molecular characterization of the *Tribolium abdominal-A* ortholog and implications for the products of the *Drosophila* gene. *Dev. Genes Evol.* **207**:446–452.
- Spicer, G. S. 1988. Molecular evolution among some *Drosophila* species groups as indicated by two-dimensional electrophoresis. *J. Mol. Evol.* **27**:250–260.
- Stone, W. S. 1962. The dominance of natural selection and the reality of superspecies (species groups) in the evolution of *Drosophila*. *Univ. Texas Publ.* **6205**:507–537.
- Struhl, G. 1984. Splitting the bithorax complex of *Drosophila*. *Nature* **308**:454–457.
- Terol, J., M. Perez-Alonso, and R. Frutos. 1991. Molecular characterization of the *zerknüllt* region of the *Antennapedia* complex of *D. subobscura*. *Hereditas* **114**:131–139.
- Tiong, S. Y., J. R. Whittle, and M. C. Gribbin. 1987. Chromosomal continuity in the abdominal region of the bithorax complex of *Drosophila* is not essential for its contribution to metameric identity. *Development* **101**:135–142.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
- Von Allmen, G., I. Hogga, A. Spierer, F. Karch, W. Bender, H. Gyurkovics, and E. Lewis. 1996. Splits in fruitfly *Hox* gene complexes. *Nature* **380**:116.
- Wasserman, M. 1992. Cytological evolution of the *Drosophila repleta* species group. Pp. 455–552 in C. B. Krimbas and J. R. Powell, eds. *Drosophila* inversion polymorphism. CRC Press, Boca Raton, Fla.
- Wharton, L. T. 1942. Analysis of the *repleta* group of *Drosophila*. *Univ. Texas Publ.* **4228**:23–59.

- Wilder, J., and H. Hollocher. 2001. Mobile elements and the genesis of microsatellites in Dipterans. *Mol. Biol. Evol.* **18**:384–392.
- Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comp. Appl. Biosci.* **13**:555–556.
- Yang, Z., and J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. *Trends Evol.* **15**:496–503.
- Yang, Z., and W. J. Swanson. 2002. Codon substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**:49–57.
- Ye, X., P. Fong, N. Iizuka, D. Choate, and D. R. Cavener. 1997. *Ultrabithorax* and *Antennapedia* 5' untranslated regions promote developmentally regulated internal translation initiation. *Mol. Cell Biol.* **17**:1714–1721.

William R. Jeffery, Associate Editor

Accepted July 14, 2003

2.2. Artículo 2

Este segundo trabajo va dirigido a resolver dos cuestiones. La primera se centra en determinar de forma precisa los puntos en los que se rompió el complejo de genes Hox y el mecanismo que produjo las roturas, y la segunda en determinar si estas roturas han tenido algún efecto funcional sobre los genes Hox adyacentes a ellas.

Primero, hemos secuenciado dos regiones del genoma de *D. buzzatii*. Una contiene los genes *lab* y *abdA* y la otra incluye el gen *pb*, con un total de 257 kb. Se han anotado los genes presentes en estas regiones y se ha comparado su organización con la de *D. melanogaster* y *D. pseudoobscura* para localizar con precisión los puntos de rotura. Se ha observado la presencia de bloques conservados en la secuencia no codificante y se ha comparado con las regiones reguladoras conocidas. Las roturas se han producido respetando los genes completos, es decir, la unidad de transcripción más sus regiones *cis*-reguladoras. La posición de las roturas se ha reducido hasta un margen de unas 3 kb. Todo indica que el mecanismo que produjo las roturas són la inversiones paracéntricas.

La elevada presencia de bloques conservados en la secuencia no codificante y la conservación de las regiones reguladoras apuntan a que los patrones de expresión de estos genes deberían estar conservados. No parece haber habido pérdida de regiones reguladoras con las roturas, pero se podrían haber adquirido nuevas secuencias que modificaran el patrón de expresión. Para contrastar esta posibilidad, se ha analizado la expresión de los tres genes Hox, *lab*, *pb* y *abdA*, en embriones y discos imaginales de cuatro especies de *Drosophila* portadoras de diferentes organizaciones del complejo. Los resultados muestran que las roturas del HOM-C no han alterado los patrones de expresión de los genes Hox adyacentes a ellas. Concluimos que, en *Drosophila*, la agrupación de los genes Hox no es imprescindible para su correcto funcionamiento. Mas bién la organización de los genes Hox parece ser modular y su agrupación genómica el resultado de la inercia evolutiva.

NEGRE B., S. CASILLAS, M. SUZANNE, E. SÁNCHEZ-HERRERO, M. AKAM, A. BARBADILLA, M. NEFEDOV, P. DE JONG y A. RUIZ (2005) Conservation of regulatory sequences and gene expression patterns in the disintegrating *Drosophila* Hox gene complex. *Genome Research* 15: 692-700.

Conservation of regulatory sequences and gene expression patterns in the disintegrating *Drosophila* *Hox* gene complex

Bárbara Negre,¹ Sònia Casillas,¹ Magali Suzanne,² Ernesto Sánchez-Herrero,² Michael Akam,³ Michael Nefedov,⁴ Antonio Barbadilla,¹ Pieter de Jong,⁴ Alfredo Ruiz^{1,5}

¹Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain; ²Centro de Biología Molecular "Severo Ochoa", Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain; ³Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United Kingdom; ⁴Children's Hospital Oakland Research Institute, Oakland 94609, California, USA

Homeotic (*Hox*) genes are usually clustered and arranged in the same order as they are expressed along the anteroposterior body axis of metazoans. The mechanistic explanation for this colinearity has been elusive, and it may well be that a single and universal cause does not exist. The *Hox*-gene complex (HOM-C) has been rearranged differently in several *Drosophila* species, producing a striking diversity of *Hox* gene organizations. We investigated the genomic and functional consequences of the two HOM-C splits present in *Drosophila buzzatii*. Firstly, we sequenced two regions of the *D. buzzatii* genome, one containing the genes *labial* and *abdominal A*, and another one including *proboscipedia*, and compared their organization with that of *D. melanogaster* and *D. pseudoobscura* in order to map precisely the two splits. Then, a plethora of conserved noncoding sequences, which are putative enhancers, were identified around the three *Hox* genes closer to the splits. The position and order of these enhancers are conserved, with minor exceptions, between the three *Drosophila* species. Finally, we analyzed the expression patterns of the same three genes in embryos and imaginal discs of four *Drosophila* species with different *Hox*-gene organizations. The results show that their expression patterns are conserved despite the HOM-C splits. We conclude that, in *Drosophila*, *Hox*-gene clustering is not an absolute requirement for proper function. Rather, the organization of *Hox* genes is modular, and their clustering seems the result of phylogenetic inertia more than functional necessity.

[Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to GenBank under accession nos. AY900631–AY900632 and AY897430–AY897434.]

Homeotic (*Hox*) genes were discovered in *Drosophila melanogaster* as mutations that transform one body part into another. Lewis (1978) and Kaufman et al. (1980) found that these genes are clustered and arranged in the chromosome in the same order as their domains of action in the body of flies. Homologous *Hox* genes were subsequently found in many other animals and their arrangement in complexes (HOM-C) shown to be the general rule (McGinnis and Krumlauf 1992; Ruddle et al. 1994). *Hox* genes encode transcription factors involved in the determination of segment identity along the anteroposterior body axis, and thus, play a fundamental role in animal development. The conserved colinearity between *Hox* gene chromosomal arrangement and expression domain is a basic notion of developmental biology, yet this is an enigmatic phenomenon for which no single satisfactory explanation exists (Kmita and Duboule 2003). Furthermore, HOM-C splits have been observed in *Drosophila* (Von Allmen et al. 1996; Lewis et al. 2003; Negre et al. 2003), *Bombyx* (Yasukochi et al. 2004), nematodes (Aboobaker and Blaxter 2003), and tunicates (Ikuta et al. 2004; Seo et al. 2004).

Ten genes arranged in a single complex comprised the an-

cestral HOM-C of arthropods (Cook et al. 2001; Hughes and Kaufman 2002; Hughes et al. 2004). In winged insects, including *Drosophila*, the genes *Hox3* and *fushi tarazu* (*ftz*) lost their homeotic function, and thus, only eight truly homeotic genes remain. Three different splits of the ancestral HOM-C have been found so far in the *Drosophila* genus (Fig. 1A). In *D. melanogaster*, the complex is split between the genes *Antennapedia* (*Antp*) and *Ultrabithorax* (*Ubx*), leaving two separate gene clusters as follows: the Antennapedia complex, ANT-C (Kaufman et al. 1990) that specifies the identity of the mouth parts and anterior thorax, and the Bithorax complex, BX-C (Duncan 1987; Martin et al. 1995) involved in the development of the posterior thorax and abdomen. In *D. pseudoobscura*, the HOM-C is also similarly divided in the ANT-C and BX-C complexes (Lewis et al. 2003). A different split between *Ubx* and *abdominal A* (*abdA*) occurs in *D. virilis* (Von Allmen et al. 1996), *D. repleta* (Ranz et al. 2001), *D. buzzatii*, and other species of the *Drosophila* subgenus (Negre et al. 2003; Fig. 1B). Finally, an additional split, between *labial* (*lab*) and *proboscipedia* (*pb*), is present in *D. buzzatii* and other species of the *repleta* group (Negre et al. 2003). This third split separated the gene *lab* far from *pb* and the anterior genes of the *Hox* complex and relocated it near the posterior genes *abdA* and *Abdominal B* (*AbdB*) in a flagrant violation of the colinearity rule. The functional consequences of these splits are unknown.

⁵Corresponding author.

E-mail Alfredo.Ruiz@uab.es; fax 0034-93-581-23-87.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.3468605>.

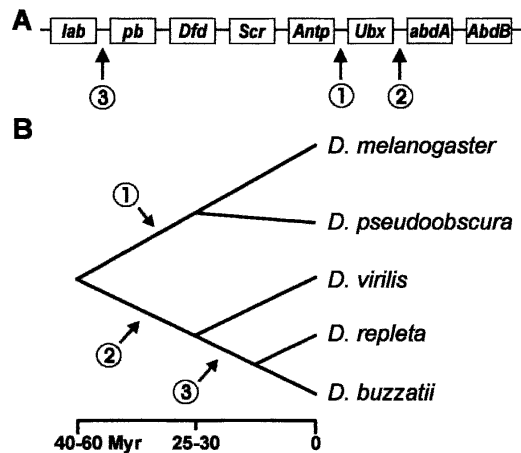


Figure 1. Genomic (A) and phylogenetic (B) localization of the three *Hox* gene complex splits observed in the *Drosophila* genus. (A) Ancestral arrangement of the eight *Hox* genes within the insects is as follows: *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal A* (*abdA*), and *Abdominal B* (*AbdB*). (B) Phylogenetic relationships and divergence times for the five *Drosophila* species included in this study. *D. melanogaster* and *D. pseudoobscura* belong to the *Sophophora* subgenus. *D. repleta* and *D. buzzatii* belong to the *Drosophila* subgenus (see Negre et al. 2003 for details).

In order to ascertain the consequences of *Drosophila* HOM-C splits, we have carried out a genomic and functional characterization of the two splits present in *D. buzzatii*. We isolated and sequenced two BAC clones containing the *lab-abdA* and *pb* chromosomal regions of *D. buzzatii*. The gene organization in these regions is compared with that of the homologous regions in *D. melanogaster* and *D. pseudoobscura* to map the precise site of the two splits. None of the two splits has altered the coding regions of *Hox* genes. We then searched for Conserved Noncoding Sequences (CNS), which are putative regulatory sequences, around the genes *lab*, *pb*, and *abdA*, to find out whether the splits removed or altered any *Hox*-gene enhancer. The position of CNS around *Hox* genes is compared with experimentally identified *Hox*-gene enhancers, and the arrangement of CNS is compared between *Hox* and non-*Hox* genes. Finally, we analyzed the expression patterns of three *Hox* genes, *lab*, *pb*, and *abdA*, in four *Drosophila* species with different *Hox*-gene organizations (with and without the splits) in whole-mount embryos and imaginal discs. The results show that, in *Drosophila* species, *Hox* genes, as well as their regulatory regions and expression patterns, are conserved, despite the *Hox* complex breaks. Thus, the functional significance of the *Hox*-gene clustering in *Drosophila* is questionable.

Results

Molecular characterization of *Hox*-gene complex breakpoints

To characterize the two HOM-C splits present in *D. buzzatii*, we isolated and sequenced two BAC clones, one (5H14, 124,024 bp) containing the *lab-abdA* region, and another (40C11, 132,938 bp) including the *pb* region (see Methods). The organization of the two regions of *D. buzzatii* chromosome 2 is shown in Figure 2 along with the homologous regions of *D. melanogaster* and *D. pseudoobscura* for comparison. *D. melanogaster* and *D. pseudoobscura* are homosequential in the analyzed regions, except where indicated. The sequenced *pb* region (Fig. 2A) contains 16 ORFs including *Dbuz\pb*, *Dbuz\zerknüllt* (*Dbuz\zen*), *Dbuz\zerknüllt-related* (*Dbuz\zen2*), and *Dbuz\bicoid* (*Dbuz\bcd*). These four genes are present in the ANT-C of *D. melanogaster* and also in the homologous region of *D. pseudoobscura* (Fig. 2B,C). The orientation of *Dbuz\zen2* is the same as that of *Dpse\zen2*, but inverted with regard to *Dmel\zen2*. The remaining 12 genes in this region are orthologous to *D. melanogaster* genes from four different regions (84D1–2, 89D2, 84E5, and 91D4–5) of chromosomal arm 3R. One of the genes, *CG14609*, is represented by six copies, in contrast to the single copy present in *D. melanogaster* or *D. pseudoobscura*. A total of four breakpoints are fixed in this region between *D. buzzatii* and *D. melanogaster* beside the *zen2* microrearrangement. That corresponding to the *lab-pb* split is located in the ~3-kb intergenic segment between *Dbuz\pb* and *Dbuz\CG17836* (Fig. 2A).

The sequenced *lab-abdA* region contains 11 ORFs, including *Dbuz\lab*, the cuticular cluster genes (*Dbuz\Ccp*), and *Dbuz\abdA*

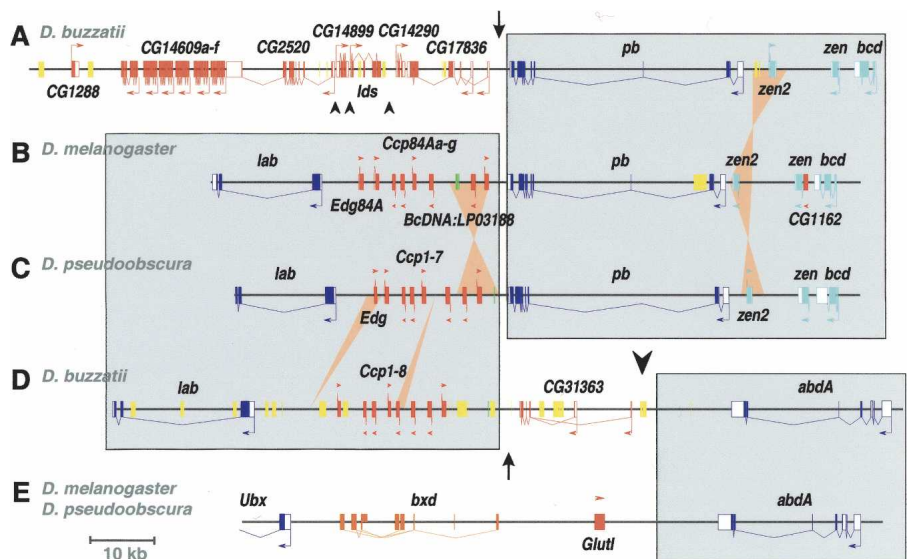


Figure 2. Gene organization of the *lab-abdA* and *pb* genomic regions of *D. buzzatii* compared with the homologous regions of *D. melanogaster* and *D. pseudoobscura*. The localization of the *lab-pb* split (arrow) and the *Ubx-abdA* split (large arrowhead) are indicated. (A) Sequence of *D. buzzatii* BAC 40C11 containing the *pb* region. (B) Organization of the *lab-pb* region in *D. melanogaster*. (C) Idem in *D. pseudoobscura*. (D) Sequence of *D. buzzatii* BAC 5H14 containing the *lab-abdA* region. (E) Organization of the *abdA* region in *D. melanogaster* and *D. pseudoobscura*. Genes are represented as open (UTRs) and filled boxes (coding sequences) with arrows indicating the sense of transcription. *Hox* genes are colored in dark blue, *Hox*-derived genes in light blue, non-*Hox* genes in red, noncoding RNA genes in orange, and the BcDNA:LP03188 and orthologous sequences in green. Transposable element insertions (usually ISBu elements, see Negre et al. 2003) are shown as yellow boxes. Large shaded rectangles include homologous *Hox*-gene regions in different species. Ochre triangles denote small inversions and insertions or deletions. Small arrowheads show breakpoints between *D. buzzatii* and *D. melanogaster* in non-*Hox* regions.

(Fig. 2D). The number of *Ccp* copies (including the gene *Edg*) is eight in the three species, but there is a small inversion encompassing two copies (plus the cDNA *BcDNA:LP03188*) in *D. melanogaster* in comparison to *D. buzzatii* or *D. pseudoobscura*, as well as one gain and one loss (Fig. 2B–D). These 11 genes come from three different regions (84A2–5, 86E11–13, and 89E2) of *D. melanogaster* chromosomal arm 3R, which means two fixed breakpoints between *D. buzzatii* and *D. melanogaster*, beside the small inversion of *Ccp* genes. One breakpoint corresponds to the *lab-pb* split and is found ~40 kb upstream of *Dbuz\lab*, in the 5-kb between the sequence similar to *BcDNA:LP03188* and the gene *Dbuz\CG31363*. The second breakpoint is that of the *Ubx-abdA* split and is located between 11 and 15 kb downstream of *Dbuz\abdA*. The two breakpoints are separated by a DNA segment of only ~22 kb encoding a single gene, *Dbuz\CG31363* (Fig. 2D).

Conserved noncoding sequences in *Hox* gene regions

We analyzed the conservation of noncoding sequences around the three *Hox* genes *lab*, *pb*, and *abdA* by comparing the sequences of the three species *D. buzzatii*, *D. melanogaster*, and *D. pseudoobscura* as done previously by other authors (Bergman and Kreitman 2001; Bergman et al. 2002) (see Methods). Figure 3 shows the VISTA graph, where the conservation between the aligned sequences is plotted (when higher than 50%) and the regions that meet the selected criteria (75% identity in a 25-bp window) are highlighted for both coding and noncoding sequences. A preliminary analysis showed no differences between intergenic and intronic regions, in agreement with previous studies (Bergman and Kreitman 2001). Thus, CNS are defined as intergenic (excluding UTRs) or intronic sequences that meet the above criteria. The characteristics of observed CNS are given in



Figure 3. Nucleotide sequence conservation in the *lab-abdA* and *pb* regions between *Drosophila* species. The three panels in each VISTA plot represent pairwise comparisons between *D. melanogaster* and *D. pseudoobscura* (*mel/pse*), *D. melanogaster* and *D. buzzatii* (*mel/buz*) and *D. pseudoobscura* and *D. buzzatii* (*pse/buz*). The x-axis represents *D. melanogaster* coordinates, and y-axis sequence identity (50%–100%). Gray arrows show the location and orientation of genes. Conservation in exons and UTRs is shown in dark and light blue, respectively. Pink regions represent CNS. Experimentally identified regulatory sequences (solid purple bars) or segments with negative results (empty bars) are indicated on top of each plot. Five microinversions detected in the *lab* or *pb* regions are enclosed in blue frames, and the VISTA graphs generated with the inverted sequences shown to the right of the main plots. VISTA plots for the CG17836–CG14290, CG31363, and CG1288–CG14609–CG2520 regions (adjacent to *Hox* genes) are shown at the bottom of the figure for comparison.

Tables 1 and 2, and the results of statistical analysis are shown in Supplemental Tables S1 and S2.

When *D. buzzatii* is compared with *D. melanogaster* or *D. pseudoobscura*, 395 and 440 CNS are found, respectively, around the three *Hox* genes (Table 1). This gives a density of 4.5 and 5 CNS per kilobase, respectively. These conserved blocks show a mean size of 44 bp with 86.5% nucleotide identity and represent 20%–22% of the analyzed noncoding sequence. When *D. melanogaster* and *D. pseudoobscura* are compared, 563 CNS are detected (6.5/kb) with a mean size of 55 bp and an average identity of 87.4%. In this comparison, the sequence in CNS represents 36% of noncoding sequence. In all three comparisons, the three regions around the *Hox* genes *lab*, *pb*, and *abdA* are homogeneous with little variation either in CNS density, size, or nucleotide identity (Supplemental Table S1). It is worth noting that CNS are coincident in all three comparisons (Fig. 3), which means that all CNS detected when comparing *D. buzzatii* with either *D. melanogaster* or *D. pseudoobscura* are also found in the comparison between the latter two species. Although most CNS keep colinearity (relative position and orientation), we could identify four microinversions, around 1–2 kb in size. One is located within the large intron of *lab* and the other three in introns 2 and 3 of *pb* (Fig. 3).

D. buzzatii is equally distant phylogenetically from either *D. melanogaster* or *D. pseudoobscura* (Fig. 1). The latter two species belong to the same subgenus and are phylogenetically closer. We compared the characteristics of the CNS found in the three pairwise comparisons. As expected, there are no statistical differences between the CNS found when comparing *D. buzzatii* with either *D. melanogaster* or *D. pseudoobscura* (Supplemental Table S2). The CNS density and the proportion of sequence in CNS are significantly higher when comparing the phylogenetically closer species *D. melanogaster* and *D. pseudoobscura*. Increasing divergence time does not seem to affect the nucleotide identity of the CNS, although the size of the CNS detected in the *Hox*-gene regions shows a significant decrease (Supplemental Table S2).

Conserved noncoding sequences in non-*Hox* gene regions

To find out whether the observed pattern of CNS is a particular feature of *Hox* genes, we also analyzed the presence of CNS in regions of the sequenced BACs adjacent, but unrelated, to *Hox* genes. We used the three microsyntenic regions between *D. buzzatii*, *D. melanogaster*, and *D. pseudoobscura* longer than 10 kb, i.e.,

the *CG31363* gene region, between *lab* and *abdA*, and the *CG17836-CG14290* and *CG1288-CG2520* regions, near *pb* (Fig. 2). These regions include one, two, and three genes, respectively. The pattern of CNS detected is shown in Figure 3 and summarized in Table 2. In the comparisons with *D. buzzatii*, we found around 100 CNS (~2/kb), which represents <8% of noncoding sequence. Thus, in these non-*Hox* regions, a much smaller number of CNS is observed and the proportion of sequence in CNS is also significantly lower than in *Hox*-gene regions (Supplemental Table S1). In the *D. melanogaster*–*D. pseudoobscura* comparison, there are 326 CNS (5.7/kb) which represents a 23% of noncoding sequence. Thus, in this case, the density is similar between *Hox* and non-*Hox*-gene regions, but the size of CNS is significantly smaller in the latter regions (Supplemental Table S1). Consequently, the proportion of sequence in CNS is also significantly lower in the non-*Hox*-gene regions. It should be noted that non-*Hox* regions show a significant variation for CNS density and also for the proportion of sequence in CNS that is not observed in *Hox*-gene regions (Supplemental Table S1). The higher variation observed between non-*Hox* regions is probably due to the heterogeneity of the sample from a functional point of view. There is little information available on the function and expression pattern of the six non-*Hox* genes analyzed, which probably represent a mixture of genes with different regulatory needs and number of enhancers.

Conservation of known regulatory sequences

Regulatory sequences of the genes *lab*, *pb*, and *abdA* have been experimentally identified in *D. melanogaster* (Karch et al. 1985; Chouinard and Kaufman 1991; Kapoun and Kaufman 1995; Martin et al. 1995). We compared their position with the pattern of CNS found around *Hox* genes. As shown in Figure 3, the regulatory sequences identified in *D. melanogaster* generally contain or correspond to CNS in *D. buzzatii*. For instance, CNS are found in the sites corresponding to the *iab2 PRE* and *iab2(1.7)* enhancers of *abdA* (Shimell et al. 1994, 2000). Similarly, a prominent conservation peak is observed at the site of the *lab550* enhancer, which directs the expression of *lab* in the embryo midgut (Marty et al. 2001). Also, the inverted segment found in the large intron of the *lab* gene roughly corresponds to the segment responsible for *lab* expression in the posterior midgut. Sequence details of the *lab550* and *iab2(1.7)* enhancer and binding site conservation are shown in Supplemental Figure S1. The Homeotic Response Ele-

Table 1. Characteristics of conserved noncoding sequences (CNS) detected with mVISTA in comparisons of *Hox* gene regions between *D. melanogaster* (mel), *D. pseudoobscura* (pse), and *D. buzzatii* (buz)

Region	Noncoding nucleotides	Species pair	Number of CNS	Density ^a (SD)	Mean size (nt) (SD)	Mean nucleotide identity (%)	Sequence in CNS (%)	
<i>lab</i>	19,227	mel/pse	129	6.71 (0.59)	53.20 (34.03)	87.69	35.69	
		mel/buz	73	3.80 (0.44)	46.70 (24.35)	87.33	17.73	
		pse/buz	84	4.37 (0.48)	44.54 (27.05)	88.08	19.46	
<i>pb</i>	42,056	mel/pse	265	6.30 (0.39)	55.04 (35.79)	87.29	34.68	
		mel/buz	196	4.66 (0.33)	41.88 (22.38)	87.09	19.52	
		pse/buz	215	5.11 (0.35)	42.92 (24.65)	86.38	21.94	
<i>abdA</i>	26,043	mel/pse	169	6.49 (0.50)	59.29 (38.43)	87.44	38.45	
		mel/buz	126	4.84 (0.43)	45.98 (26.15)	86.18	22.25	
		pse/buz	141	5.41 (0.46)	46.11 (24.97)	85.44	24.97	
Total	87,326	mel/pse	563	6.45 (0.27)	55.89 (36.23)	87.42	36.03	
Hox gene regions		mel/buz	395	4.52 (0.23)	44.08 (24.04)	86.83	19.94	
		pse/buz	440	5.04 (0.24)	44.25 (25.53)	86.39	22.30	

^aDensity = number of CNS per kilobase.

Table 2. Characteristics of conserved noncoding sequences (CNS) detected with mVISTA in comparisons of non-*Hox* gene regions between *D. melanogaster* (mel), *D. pseudoobscura* (pse), and *D. buzzatii* (buz)

Region	Noncoding nucleotides	Species pair	Number of CNS	Density ^a (SD)	Mean size (nt) (SD)	Mean nucleotide identity (%)	Sequence in CNS (%)
CG1288-CG2520	18,333	mel/pse	127	6.93 (0.61)	43.48 (28.16)	86.31	30.12
		mel/buz	65	3.55 (0.44)	45.26 (31.51)	86.30	16.05
		pse/buz	67	3.65 (0.45)	42.44 (29.40)	86.81	15.59
CG17836-CG14290	10,921	mel/pse	46	4.21 (0.62)	45.02 (33.27)	82.67	18.96
		mel/buz	18	1.65 (0.39)	39.61 (22.67)	82.88	6.53
		pse/buz	22	2.01 (0.43)	42.09 (24.34)	84.34	8.48
CG31363	27,510	mel/pse	153	5.56 (0.45)	35.51 (14.95)	87.17	19.75
		mel/buz	22	0.80 (0.17)	26.09 (3.94)	82.93	2.09
		pse/buz	23	0.84 (0.17)	28.78 (7.70)	83.23	2.41
Total non- <i>Hox</i> gene regions	56,764	mel/pse	326	5.74 (0.32)	39.96 (24.15)	86.09	22.95
		mel/buz	105	1.84 (0.18)	40.28 (27.50)	85.27	7.45
		pse/buz	112	1.97 (0.19)	39.59 (25.88)	85.76	7.83

^aDensity = number of CNS per kilobase.

ment (HOMRE) of the *lab550* enhancer contains four binding sites; all of them are conserved in the three species. In the *iab2(1.7)* enhancer, there are five Hunchback (HB)-binding sites, three of which are conserved in the three species, whereas the other two vary in position between species. This enhancer also contains a unique Krüppel (KR)-binding site, where point mutations in *D. melanogaster* cause gain-of-expression mutants (Hab1 and Hab2) (Shimell et al. 1994). This binding site is conserved in all three species (Supplemental Fig. S1). The conservation between *D. melanogaster* and *D. buzzatii* around *abdA* ends 9 kb (in *D. melanogaster*) and 11 kb (in *D. buzzatii*) downstream of this gene (Fig. 3). This boundary lies between the *iab2* and *pbx* regulatory sequences, which control the expression of *abdA* and *Ubx*, respectively (Karch et al. 1985). We have shown that the *iab2* region downstream of *abdA* is conserved in *D. buzzatii*. We have not sequenced the *Ubx* region in *D. buzzatii*, but we assume that the *pbx* regulatory sequence will conserve its position upstream of *Ubx*, i.e., there are no rearrangements between *Ubx* and its regulatory sequences (see below).

It is worth noting though, that CNS were also found in fragments not experimentally tested or described as with no effect on expression (Fig. 3). This observation suggests that the regulation of these genes may be even more complex than currently envisaged, and that more regulatory modules may be operative in nature than those experimentally identified in the laboratory.

Hox gene expression patterns

The conservation of regulatory sequences suggests that splits of the HOM-C had no consequences on *Hox*-gene expression. To test this prediction, we compared the expression patterns of the *Hox* genes *lab*, *pb*, and *abdA* between *D. melanogaster*, *D. virilis*, *D. buzzatii*, and *D. repleta*. These four *Drosophila* species represent three different *Hox*-gene organizations (Figs. 1,2). *D. melanogaster* possess the *Antp-Ubx* split only, whereas *D. virilis* has the *Ubx-abdA* split instead. Both *D. buzzatii* and *D. repleta* present the *Ubx-abdA* and *lab-pb* splits. We used in situ hybridization and antibody staining to whole-mount embryos and to imaginal discs from third instar larvae and prepupae (see Methods). Detailed results are given in Figure 4 and Supplemental Figures S2–S6. The expression patterns of the four species closely follow those described for *D. melanogaster* (for review, see Hughes and Kaufman 2002). Interspecific variation was detected only in the *pb* gene, which in *D. virilis* presents an extra domain in the embryo mesoderm (Fig. 4). As this expression domain is not shared by *D. melanogaster*, it is seemingly not related with the *lab-pb* split. Although our analysis is qualitative, and slightly quantitative changes or domain changes of a few cells may remain undetected, it shows that the reorganization of the HOM-C caused no major alterations of the expression patterns of the three *Hox* genes adjacent to the splits, in good agreement with the conservation of regulatory sequences (see above). Likewise, Bomze and

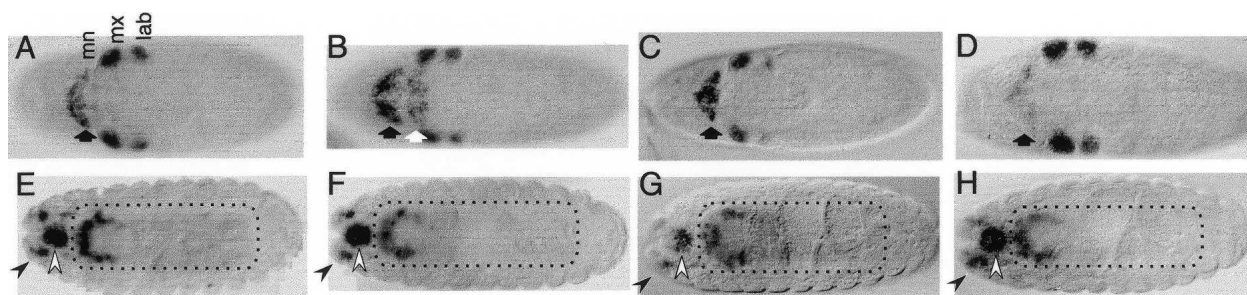


Figure 4. Expression pattern of *pb* in embryos. (A–D) stage 11 embryos, (E–H) stage 17 embryos. (A,E) *D. melanogaster*, (B,F) *D. virilis*, (C,G) *D. buzzatii*, and (D,H) *D. repleta*. (A–D) Expression on the ectoderm of the maxillary and labial lobes. Later in development (E–H) *pb* is detected in the derivatives of the maxillary (white arrowhead) and labial (black arrowhead) lobes, and in the ventral nervous system (boxed area). (A–D) *pb* expression is detected in the mesodermal layer of the mandibular segment (black arrow) in all four species. In *D. virilis* only (B), *pb* is also expressed in the mesodermal layer of the maxillary segment (white arrow). The mandibular (mn), maxillary (mx), and labial (lab) segments are shown in A.

López (1994) found that the expression pattern of *Ubx* in embryos is conserved between *D. melanogaster*, *D. pseudoobscura*, *D. virilis*, and *D. hydei* (a species of the *repleta* group), despite their different *Hox*-gene organization (Figure 1).

Discussion

zen2 predates the *Drosophila* radiation

The *zen* and *bcd* genes come from a duplication of *Hox3* in the ancestor of Cyclorraphan flies (Stauber et al. 2002). A second duplication of *zen* gave birth to *zen2*, which was thought to be a recent event in *D. melanogaster* (Randazzo et al. 1993), where it has no discernible function. However, the existence of *Dpse\zen2* and *Dbuz\zen2* shows that the *zen-zen2* duplication must predate the divergence of the *Sophophora* and *Drosophila* subgenus, and that this gene has been kept during at least 40–60 Myr of evolution. Whether this gene is also present in other flies outside of the *Drosophila* genus is still unknown.

Patterns of conserved noncoding sequence evolution

Cis-Regulatory Modules (CRM) are transcription regulatory DNA segments (from a few hundred base pair to 1 kb in size) that control gene expression in higher eukaryotes (Wray et al. 2003). CRM have a complex structure still not fully understood. They contain one or several binding sites for different transcription factors, which act cooperatively to activate or repress transcription of the target gene. As CRM are functionally constrained to maintain the expression of the target gene, they evolve slower than nonfunctional sequences. Therefore, the conservation of noncoding sequences between phylogenetically distant species may be used as a guide for identification of regulatory sequences. Several recent studies (Bergman and Kreitman 2001; Bergman et al. 2002; Cooper and Sidow 2003; Nobrega et al. 2003; Santini et al. 2003) support the use of comparative sequence analysis and characterization of CNS as a useful approach to detect putative CRM in *Drosophila* and other organisms. The clustering of previously characterized transcription-factor binding sites may be also used for detection of CRM (Berman et al. 2004). However, the absence of high-quality binding data for most *Drosophila* transcription factors represent a great current limitation in the widespread application of this method.

We exhaustively searched for CNS around *lab*, *pb*, and *abdA* and around adjacent non-*Hox* genes by comparing three species pairs. A plethora of highly conserved blocks was found surrounding the three *Hox* genes in the comparison between the phylogenetically distant species *D. buzzatii* and *D. melanogaster* or *D. pseudoobscura* (Fig. 1). The proportion of noncoding sequence included in CNS was 20%–22%. In most cases, these CNS keep their relative position and colinearity, although a few microrearrangements were found. The interpretation of these CNS as regulatory sequences is supported by the high neutral substitution rate (Moriyama and Gojobori 1992) and intrinsic rate of DNA loss (Petrov et al. 1996; Singh and Petrov 2004) in *Drosophila*. Noncoding sequences are not expected to be conserved between such distantly related species unless they are functionally constrained. The coincidence between CNS and known enhancers such as *iab2 PRE* or *lab550* (Supplemental Fig. S1) further supports this interpretation.

A lower CNS density was observed around non-*Hox* genes. This result fits well with previous observations showing that genes with complex developmentally regulated expression show

a higher degree of conservation in noncoding regions than more simple genes with metabolic or housekeeping functions (Bergman and Kreitman 2001; Bergman et al. 2002; Halligan et al. 2004). Moreover, *Hox* genes are associated with larger noncoding regions. *Hox* genes harbor some of the longest introns of any *Drosophila* gene (Moriyama et al. 1998) and mean intron size is significantly greater in the *Hox* than in the non-*Hox* genes analyzed here ($F = 4.69$, $df = 1$, $P < 0.05$). This observation also fits with the notion that the amount of noncoding DNA must be larger in those genes with complex developmental functions in order to harbor the required CRM (Nelson et al. 2004).

HOM-C evolution in *Drosophila*

In *Drosophila*, *Hox* genes are arranged in the same 5'→3' orientation (with only one exception, the *Deformed* gene in *D. melanogaster*). Their regulatory sequences are usually located upstream of each gene and in the introns. If we look at the three HOM-C splits known in *Drosophila*, a common pattern arises. As can be seen in Figure 2, the *lab-pb* split took place close to the 3' end of *pb* and far from the *lab* 5' end. Likewise, the split between the genes *Ubx* and *abdA* took place near the *abdA* 3' end and far from the *Ubx* 5' end, in the short space between their respective regulatory sequences *pbx* and *iab2*. This is approximately the same position where an experimental break that does not affect development has been observed (Struhl 1984), although the deficiencies used in the complementation tests both carry a fraction of the *pbx* and *iab2* regions. Finally, sequence comparison between *D. melanogaster* and *D. virilis* (Lewis et al. 2003) show that both the insertion of the *CG1217* gene and the *Antp-Ubx* split took place close to the *Ubx* 3' end, and far from the *Antp* 5' end (results not shown). Thus, all three splits seem to have occurred far from the 5' end of one gene and much closer to the 3' end of the next one, in such a way as to keep in place the regulatory sequences of both genes. In this way, rearrangements did not alter any of the known regulatory sequences of these *Hox* genes; this would explain the absence of gene expression changes.

In the *repleta* group species, the anterior gene *lab* is located near the posterior genes *abdA* and *AbdB*. The sequence analysis shows that *lab* and *abdA* are only 75 kb apart and show the same orientation. The breakpoint of the *lab-pb* split occurred at ~22 kb from that of the *Ubx-abdA* split. None of those splits seem to have affected the regulatory regions of the *Hox* genes, because the expression patterns of *lab* and *abdA* are unaffected. Although it is intriguing, the proximity between these genes in the *D. buzzatii* genome seems purely accidental and lacking any functional significance.

The most likely mechanisms for the generation of the HOM-C splits are paracentric inversions (Ranz et al. 2001; Gonzalez et al. 2002). A plausible reconstruction of HOM-C evolution in the *Drosophila* subgenus that accounts for the current organization of *Hox* genes in *D. buzzatii* is shown in Figure 5. In lower Dipterans, such as *Anopheles gambiae*, the eight *Hox* genes, plus *Hox3* and *ftz*, are arranged as a single cluster (Powers et al. 2000). Before the radiation of the *Drosophila* genus, two transpositions occurred as follows: the *Ccp* gene cluster between *lab* and *pb*, and the gene *CG1217* between *Antp* and *Ubx* (Lewis et al. 2003). Also, *zen*, *zen2*, and *bcd* evolved from the *Hox3* gene (see above). In the lineage of the *Drosophila* subgenus, an inversion took place with one breakpoint between *Ubx* and *abdA* (split 2 in Fig. 1) and the other one between *CG1363* and an unknown ORF (*X*). This HOM-C structure is now present in species of the *Drosophila* sub-

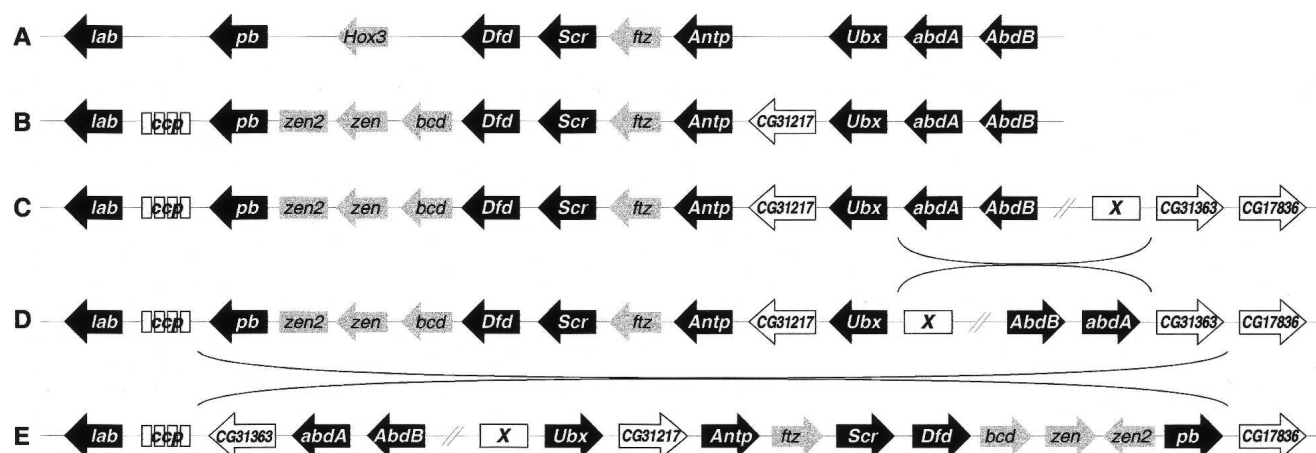


Figure 5. Reconstruction of the *Hox* gene complex evolution in the *Drosophila* subgenus. Genes are shown as arrows when the orientation 5'→3' is known, and as rectangles otherwise. *Hox* genes are in black, *Hox*-related genes in gray, and non-*Hox* genes in white. (A) Lower Dipterans. (B) Before the radiation of the *Drosophila* genus. (C) *Drosophila* subgenus after its separation from that of the *Sophophora* subgenus. (D) Ancestor of the *repleta* group. (E) Present arrangement of *Hox* genes in *Drosophila buzzatii* (cf. Fig. 2A,D).

genus outside the *repleta* group, such as *D. virilis* (see Fig. 1). A second inversion, in the ancestor of the *repleta* group, split the HOM-C between *lab* and *pb* (split 3 in Fig. 1). This inversion, which relocated *lab* close to *abdA*, had one breakpoint between *pb* and the *Ccp* cluster genes and the second breakpoint between *CG31363* and *CG17836*. These two genes are not adjacent in the *D. melanogaster* genome, but we infer that they were so in the ancestor of the *Drosophila* subgenus.

Do flies have a *Hox* gene complex?

Despite the striking conservation of *Hox*-gene clustering in metazoans, if we compare two of the most deeply studied organisms, *Drosophila* and vertebrates, important differences arise (Ferrier and Minguillon 2003; Santini et al. 2003; Wagner et al. 2003). *Drosophila* *Hox*-gene regions (1) are much larger than those of vertebrates, e.g., the human *HoxA* cluster is only 110 kb long, whereas the *D. melanogaster* HOM-C spans 665 kb; (2) contain transposable element insertions, which are remarkably absent in those of vertebrates; (3) contain also non-*Hox* genes that are inserted between the *Hox* genes, and tandem duplications within the complex, such as those of the *zen*-related genes; (4) allow for small inversions of *Hox* genes, such as *Dfd* (Randazzo et al. 1993), and non-*Hox* genes, such as *zen2* (Fig. 2); and (5) are split in three ways in different lineages, apparently without consequences on gene expression. These observations suggest a highly dynamic evolution in *Drosophila* that contrasts with the compact structure seen in vertebrates. Thus, the splits of HOM-C in *Drosophila* indicate a release of functional requirements present in other metazoan.

Moreover, *Drosophila* is not the only organism known to have a split HOM-C. Split *Hox*-gene complexes were also known in nematodes, and recently have been described in *Bombyx* and tunicates. What do those organisms have in common in addition to the split HOM-C? Vertebrate development follows a rostral-to-caudal temporal progression, and the colinearity of *Hox* genes is not only spatial, but also temporal (the *Hox* clock) (Kmita and Duboule 2003). In the tunicate *Oikopleura*, *Hox* gene expression still evokes spatial colinearity but not temporal (Seo et al. 2004), which favors the argument that the constraining force of HOM-C structure conservation is temporal colinearity (Ferrier and Min-

guillon 2003). In nematodes, the pattern of *Hox*-gene evolution seems indicative of the move to a deterministic developmental mode (Aboobaker and Blaxter 2003). *Bombyx* embryogenesis, which is difficult to assign to a short or a long germ insect, is characterized by a quick development (Davis and Patel 2002). *Drosophila* is a long germ insect, where all *Hox* genes are activated almost simultaneously during the cellular blastoderm stage. Thus, none of these organisms seems to show temporal colinearity. A common feature between all organisms shown so far to have a split *Hox* complex seems to be a derived mode of embryogenesis characterized by a fast early development.

The loss of temporal progression in the activation of *Hox* genes in a very rapid mode of embryogenesis could be the ultimate cause for the modular organization of those *Hox* "clusters," where modules can be taken apart without loss of function. Given the high rate of chromosomal rearrangement in the genus *Drosophila* (Ranz et al. 2001; Gonzalez et al. 2002), we anticipate that an even greater variety of *Hox*-gene organizations will be discovered when more species are investigated. It is ironic that *Hox*-gene colinearity was discovered in *Drosophila*, an organism with a partially disassembled complex, which may be the by-product of phylogenetic inertia more than that of functional necessity.

Methods

Flies

D. buzzatii stock st-1 was used for construction of a genomic BAC library (González et al. 2005). The following species and stocks were used for gene expression experiments: *D. buzzatii* (j19), *D. repleta* (1611.2), *D. virilis* (Tokyo-Japan), and *D. melanogaster* (Canton S and Oregon R).

BAC sequencing

The genomic BAC library was screened with probes from the *lab*, *pb*, and *abdA* genes (González et al. 2005). Positive clones were used to build physical maps for the *lab-abdA* and *pb* chromosomal regions, and one BAC clone from each region was chosen for sequencing. Shotgun sublibraries were constructed for each BAC using the vector TOPO, and enough plasmid clones were se-

quenced by both ends to reach an $\sim 6\times$ redundancy. Reads were assembled with the PHRED-PHRAD-CONSED software (Ewing and Green 1998; Ewing et al. 1998; Gordon et al. 1998) and sequences finished with one round of AUTOFINISH (Gordon et al. 2001), followed by PCR to bridge the remaining gaps. A continuous high-quality sequence (PHRED score >40) was obtained for BAC clones 5H14 (124,024 bp), and 40C11 (132,938 bp). Statistical details of the sequencing process are given in Supplemental Table S3.

Sequence annotation

Nucleotide sequences were annotated with the aid of GENESCRIP (Hudek et al. 2003) and ARTEMIS (Berriman and Rutherford 2003). Predicted ORFs were corroborated with GOFigure (Khan et al. 2003) for automatic Gene Ontology (Harris et al. 2004) annotation, and BLAST (McGinnis and Madden 2004) for similarity searches. *D. buzzatii* sequences were compared with those of homologous regions in *D. melanogaster* (Celniker et al. 2002) and *D. pseudoobscura* (Richards et al. 2005) genomes. *D. melanogaster* sequences used were as follows: AE001572 (ANT-C), DMU31961 (BX-C), and AE003692, AE003672, AE003713, AE003676, and AE003724 (other regions). *D. pseudoobscura* contigs AADE01000437 (*lab*), AADE01000149 (*pb*), AADE01000036 (*abdA*), and AADE01000014, AADE000175, AADE01002495, AADE01000322 (non-*Hox* genes) were identified with Genome VISTA (Dubchak et al. 2000) and the regions of interest annotated.

Analysis of regulatory sequences

Pairwise alignments of six homologous genomic regions between *D. buzzatii*, *D. melanogaster*, and *D. pseudoobscura* were performed with the AVID global-alignment tool using default parameters (Bray et al. 2003). CNS were identified in the alignments with mVISTA (Mayor et al. 2000) using a window size of 25 bp and a minimum identity of 75%. Statistical tests were carried out to compare the characteristics of the CNS found in the different regions. Comparisons of CNS size distributions, which depart significantly from normality, were conducted using the G-test (Sokal and Rohlf 1995). The number of CNS and the proportion of sequence within CNS was scored for 1-kb windows along the analyzed regions (masking out exons). The resulting variables (density and percent sequence in CNS) as well as the nucleotide identity (per CNS) were tested using ANOVA (Sokal and Rohlf 1995). A complete list of CNS detected is provided in Supplemental Table S5.

Gene-expression experiments

In situ hybridizations and antibody staining were performed to whole-mount embryos and to imaginal discs from third-instar larvae and prepupae as described (Alonso and Akam 2003; Suzanne et al. 2003). cDNA clones were obtained for *lab* from the four species, *pb* from *D. buzzatii* and *D. melanogaster* and *abdA* from *D. buzzatii*, *D. repleta*, and *D. virilis* as described (Negre et al. 2003) (for primers see Supplemental Table S4). Sense and antisense RNA probes were produced as described (Suzanne et al. 2003). When no species-specific probe was available, at least two different ones were used in independent experiments, and the results were always consistent. Specific antibodies against the protein were used for *abdA* (Macías et al. 1990).

Acknowledgments

We thank Claudio Alonso, Joaquín Ariño, Anna Barceló, James Castelli-Gair, Jose Felix de Celis, Susan Celniker, Ariel Chipman,

Josefa González, Roger Hoskins, Barret Pfeiffer, Theresa Ren, Armand Sánchez, Chung Li Shu, and three anonymous referees for help or comments on the manuscript. This work was supported by grants BMC2002-01708 awarded to A.R. and BMC2002-00300 to E.S.-H. from the Dirección General de Enseñanza Superior e Investigación Científica (MEC, Spain), an Institutional Grant from the Fundación Ramón Areces, a doctoral FI/DGR fellowship from the Generalitat de Catalunya awarded to B.N., and a doctoral FPI fellowship from the Ministerio de Ciencia y Tecnología (BES-2003-0416) awarded to S.C.

References

- Aboobaker, A. and Blaxter, M. 2003. *Hox* gene evolution in nematodes: Novelty conserved. *Curr. Opin. Genet. Dev.* **13**: 593–598.
- Alonso, C.R. and Akam, M. 2003. A *Hox* gene mutation that triggers nonsense-mediated RNA decay and affects alternative splicing during *Drosophila* development. *Nucleic Acids Res.* **31**: 3873–3880.
- Bergman, C.M. and Kreitman, M. 2001. Analysis of conserved noncoding DNA in *Drosophila* reveals similar constraints in intergenic and intronic sequences. *Genome Res.* **11**: 1335–1345.
- Bergman, C., Pfeiffer, B., Rincon-Limas, D., Hoskins, R., Gnirke, A., Mungall, C., Wang, A., Kronmiller, B., Pacleb, J., Park, S., et al. 2002. Assessing the impact of comparative genomic sequence data on the functional annotation of the *Drosophila* genome. *Genome Biol.* **3**: research0086.
- Berman, B.P., Pfeiffer, B.D., Laverty, T.R., Salzberg, S.L., Rubin, G.M., Eisen, M.B., and Celniker, S.E. 2004. Computational identification of developmental enhancers: Conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol.* **5**: R61.
- Berriman, M. and Rutherford, K. 2003. Viewing and annotating sequence data with Artemis. *Brief Bioinform.* **4**: 124–132.
- Bomze, H.M. and López, A.J. 1994. Evolutionary conservation of the structure and expression of alternatively spliced *Ultrabithorax* isoforms from *Drosophila*. *Genetics* **136**: 965–977.
- Bray, N., Dubchak, I., and Pachter, L. 2003. AVID: A global alignment program. *Genome Res.* **13**: 97–102.
- Celniker, S.E., Wheeler, D.A., Kronmiller, B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E., et al. 2002. Finishing a whole-genome shotgun: Release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol.* **3**: research0079.
- Chouinard, S. and Kaufman, T.C. 1991. Control of expression of the homeotic labial (*lab*) locus of *Drosophila melanogaster*: Evidence for both positive and negative autogenous regulation. *Development* **113**: 1267–1280.
- Cook, C.E., Smith, M.L., Telford, M.J., Bastianello, A., and Akam, M. 2001. *Hox* genes and the phylogeny of the arthropods. *Curr. Biol.* **11**: 759–763.
- Cooper, G.M. and Sidow, A. 2003. Genomic regulatory regions: Insights from comparative sequence analysis. *Curr. Opin. Genet. Dev.* **13**: 604–610.
- Davis, G.K. and Patel, N.H. 2002. Short, long, and beyond: Molecular and embryological approaches to insect segmentation. *Annu. Rev. Entomol.* **47**: 669–699.
- Dubchak, I., Brudno, M., Loots, G.G., Pachter, L., Mayor, C., Rubin, E.M., and Frazer, K.A. 2000. Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Res.* **10**: 1304–1306.
- Duncan, I. 1987. The bithorax complex. *Annu. Rev. Genet.* **21**: 285–319.
- Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**: 186–194.
- Ewing, B., Hillier, L., Wendl, M.C., and Green, P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**: 175–185.
- Ferrier, D.E. and Minguillon, C. 2003. Evolution of the *Hox/ParaHox* gene clusters. *Int. J. Dev. Biol.* **47**: 605–611.
- González, J., Ranz, J.M., and Ruiz, A. 2002. Chromosomal elements evolve at different rates in the *Drosophila* genome. *Genetics* **161**: 1137–1154.
- González, J., Nefedov, M., Bosdet, I., Casals, F., Calvete, O., Delprat, A., Shin, H., Chiu, R., Mathewson, C., Wye, N., et al. 2005. A BAC-based physical map of the *Drosophila buzzatii* genome. *Genome Res.* (in press).
- Gordon, D., Abajian, C., and Green, P. 1998. Consed: A graphical tool for sequence finishing. *Genome Res.* **8**: 195–202.

- Gordon, D., Desmarais, C., and Green, P. 2001. Automated finishing with autofinish. *Genome Res.* **11**: 614–625.
- Halligan, D.L., Eyre-Walker, A., Andolfatto, P., and Keightley, P.D. 2004. Patterns of evolutionary constraints in intronic and intergenic DNA of *Drosophila*. *Genome Res.* **14**: 273–279.
- Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., et al. 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**: D258–D261.
- Hudek, A.K., Cheung, J., Boright, A.P., and Scherer, S.W. 2003. Genescript: DNA sequence annotation pipeline. *Bioinformatics* **19**: 1177–1178.
- Hughes, C.L. and Kaufman, T.C. 2002. *Hox* genes and the evolution of the arthropod body plan. *Evol. Dev.* **4**: 459–499.
- Hughes, C.L., Liu, P.Z., and Kaufman, T.C. 2004. Expression patterns of the rogue *Hox* genes *Hox3/zen* and *fushi tarazu* in the apterygote insect *Thermobia domestica*. *Evol. Dev.* **6**: 393–401.
- Ikuta, T., Yoshida, N., Satoh, N., and Saiga, H. 2004. *Ciona intestinalis Hox* gene cluster: Its dispersed structure and residual colinear expression in development. *Proc. Natl. Acad. Sci.* **101**: 15118–15123.
- Kapoun, A.M. and Kaufman, T.C. 1995. A functional analysis of 5', intronic and promoter regions of the homeotic gene *proboscipedia* in *Drosophila melanogaster*. *Development* **121**: 2127–2141.
- Karch, F., Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M., and Lewis, E.B. 1985. The abdominal region of the bithorax complex. *Cell* **43**: 81–96.
- Kaufman, T.C., Lewis, R., and Wakimoto, B. 1980. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: The homeotic gene complex in polytene chromosome interval 84A-B. *Genetics* **94**: 115–133.
- Kaufman, T.C., Seeger, M.A., and Olsen, G. 1990. Molecular and genetic organization of the *antennapedia* gene complex of *Drosophila melanogaster*. *Adv. Genet.* **27**: 309–362.
- Khan, S., Situ, G., Decker, K., and Schmidt, C.J. 2003. GoFigure: Automated Gene Ontology annotation. *Bioinformatics* **19**: 2484–2485.
- Kmita, M. and Duboule, D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**: 331–333.
- Lewis, E.B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**: 565–570.
- Lewis, E.B., Pfeiffer, B.D., Mathog, D.R., and Celniker, S.E. 2003. Evolution of the homeobox complex in the Diptera. *Curr. Biol.* **13**: R587–R588.
- Macias, A., Casanova, J., and Morata, G. 1990. Expression and regulation of the *abd-A* gene of *Drosophila*. *Development* **110**: 1197–1207.
- Martin, C.H., Mayeda, C.A., Davis, C.A., Ericsson, C.L., Knafels, J.D., Mathog, D.R., Celniker, S.E., Lewis, E.B., and Palazzolo, M.J. 1995. Complete sequence of the bithorax complex of *Drosophila*. *Proc. Natl. Acad. Sci.* **92**: 8398–8402.
- Marty, T., Vigano, M.A., Ribeiro, C., Nussbaumer, U., Grieder, N.C., and Affolter, M. 2001. A *Hox* complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer. *Development* **128**: 2833–2845.
- Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S., and Dubchak, I. 2000. VISTA: Visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**: 1046–1047.
- McGinnis, W. and Krumlauf, R. 1992. Homeobox genes and axial patterning. *Cell* **68**: 283–302.
- McGinnis, S. and Madden, T.L. 2004. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **32**: W20–W25.
- Moriyama, E.N. and Gojobori, T. 1992. Rates of synonymous substitution and base composition of nuclear genes in *Drosophila*. *Genetics* **130**: 855–864.
- Moriyama, E.N., Petrov, D.A., and Hartl, D.L. 1998. Genome size and intron size in *Drosophila*. *Mol. Biol. Evol.* **15**: 770–773.
- Negre, B., Ranz, J.M., Casals, F., Cáceres, M., and Ruiz, A. 2003. A new split of the *Hox* gene complex in *Drosophila*: Relocation and evolution of the gene *labial*. *Mol. Biol. Evol.* **20**: 2042–2054.
- Nelson, C.E., Hersh, B.M., and Carroll, S.B. 2004. The regulatory content of intergenic DNA shapes genome architecture. *Genome Biol.* **5**: R25.
- Nobrega, M.A., Ovcharenko, I., Afzal, V., and Rubin, E.M. 2003. Scanning human gene deserts for long-range enhancers. *Science* **302**: 413.
- Petrov, D.A., Lozovskaya, E.R., and Hartl, D.L. 1996. High intrinsic rate of DNA loss in *Drosophila*. *Nature* **384**: 346–349.
- Powers, T.P., Hogan, J., Ke, Z., Dymbrowski, K., Wang, X., Collins, F.H., and Kaufman, T.C. 2000. Characterization of the *Hox* cluster from the mosquito *Anopheles gambiae* (Diptera: Culicidae). *Evol. Dev.* **2**: 311–325.
- Randazzo, F.M., Seeger, M.A., Huss, C.A., Sweeney, M.A., Cecil, J.K., and Kaufman, T.C. 1993. Structural changes in the antennapedia complex of *Drosophila pseudoobscura*. *Genetics* **134**: 319–330.
- Ranz, J.M., Casals, F., and Ruiz, A. 2001. How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* **11**: 230–239.
- Richards, S., Liu, Y., Bettencourt, B.R., Hradscky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P., et al. 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Res.* **15**: 1–18.
- Ruddle, F.H., Bartels, J.L., Bentley, K.L., Kappen, C., Murtha, M.T., and Pendleton, J.W. 1994. Evolution of *Hox* genes. *Annu. Rev. Genet.* **28**: 423–442.
- Santini, S., Boore, J.L., and Meyer, A. 2003. Evolutionary conservation of regulatory elements in vertebrate *Hox* gene clusters. *Genome Res.* **13**: 1111–1122.
- Seo, H.C., Edvardsen, R.B., Maeland, A.D., Bjordal, M., Jensen, M.F., Hansen, A., Flaas, M., Weissenbach, J., Lehrach, H., Wincker, P., et al. 2004. *Hox* cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* **431**: 67–71.
- Shimell, M.J., Simon, J., Bender, W., and O'Connor, M.B. 1994. Enhancer point mutation results in a homeotic transformation in *Drosophila*. *Science* **264**: 968–971.
- Shimell, M.J., Peterson, A.J., Burr, J., Simon, J.A., and O'Connor, M.B. 2000. Functional analysis of repressor binding sites in the *iab-2* regulatory region of the *abdominal-A* homeotic gene. *Dev. Biol.* **218**: 38–52.
- Singh, N.D. and Petrov, D.A. 2004. Rapid sequence turnover at an intergenic locus in *Drosophila*. *Mol. Biol. Evol.* **21**: 670–680.
- Sokal, R.R. and Rohlf, F.J. 1995. *Biometry: The principles and practice of statistics in biological research*. W.H. Freeman and Co., New York.
- Staub, M., Prell, A., and Schmidt-Ott, U. 2002. A single *Hox3* gene with composite *bicoid* and *zerknüllt* expression characteristics in non-Cyclorrhaphan flies. *Proc. Natl. Acad. Sci.* **99**: 274–279.
- Struhl, G. 1984. Splitting the bithorax complex of *Drosophila*. *Nature* **308**: 454–457.
- Suzanne, M., Estella, C., Calleja, M., and Sánchez-Herrero, E. 2003. The *hernandez* and *fernandez* genes of *Drosophila* specify eye and antenna. *Dev. Biol.* **260**: 465–483.
- Von Allmen, G., Hogga, I., Spierer, A., Karch, F., Bender, W., Gyurkovics, H., and Lewis, E. 1996. Splits in fruitfly *Hox* gene complexes. *Nature* **380**: 116.
- Wagner, G.P., Amemiya, C., and Ruddle, F. 2003. *Hox* cluster duplications and the opportunity for evolutionary novelties. *Proc. Natl. Acad. Sci.* **100**: 14603–14606.
- Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., and Romano, L.A. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**: 1377–1419.
- Yasukochi, Y., Ashakumary, L., Wu, C., Yoshida, A., Nohata, J., Mita, K., and Sahara, K. 2004. Organization of the *Hox* gene cluster of the silkworm, *Bombyx mori*: A split of the *Hox* cluster in a non-*Drosophila* insect. *Dev. Genes Evol.* **214**: 606–614.

Received November 15, 2004; accepted in revised form January 26, 2005.

SUPPLEMENTAL MATERIAL

Table S1. Statistical analysis of CNS characteristics between regions for the three pair-wise species comparisons.

Source of variation ⁽¹⁾	Species pair	Density		Size distribution		Nucleotide identity (%)		Sequence in CNS (%)	
		d.f.	F	d.f.	G	d.f.	F	d.f.	F
Between the six regions	mel-pse	5	1.74	60	104.34***	5	4.84***	5	7.17***
	mel-buz	5	15.31***	60	58.19	5	2.13	5	16.16***
	pse-buz	5	18.75***	60	64.82	5	3.30*	5	18.76***
Between Hox regions	mel-pse	2	0.23	24	30.63	2	0.13	2	1.06
	mel-buz	2	1.40	24	26.52	2	0.49	2	1.19
	pse-buz	2	1.18	24	38.04*	2	4.55*	2	1.56
Between non-Hox regions	mel-pse	2	3.34*	24	24.90	2	9.71***	2	3.64*
	mel-buz	2	9.44***	24	16.68	2	2.69	2	11.71***
	pse-buz	2	11.02***	24	9.65	2	0.40	2	10.52***
Hox vs non-Hox regions	mel-pse	1	1.58	12	60.36***	1	4.54*	1	26.48***
	mel-buz	1	54.87***	12	17.41	1	4.29*	1	55.00 ***
	pse-buz	1	69.34***	12	19.59	1	6.63*	1	69.65***

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

⁽¹⁾ Variation between the six regions (df = 5) can be partitioned into three components: variation between the three Hox gene regions (df = 2), variation between the three non-Hox gene regions (df = 2) and variation between Hox and non-Hox gene regions (df = 1).

Table S2. Statistical analysis of CNS characteristics between species pairs for the different regions.

Source of variation ⁽¹⁾	Region	Density		Size distribution		Nucleotide identity (%)		Sequence in CNS (%)	
		d.f.	F	d.f.	G	d.f.	F	d.f.	F
Between the three species pairs	Hox genes	2	15.58***	24	28.23	2	0.70	2	43.97***
	Non-Hox genes	2	49.44***	24	9.59	2	1.89	2	34.10***
Mel/buz versus pse/buz	Hox gene	1	2.09	12	2.96	1	0.08	1	1.69
	Non-Hox gene	1	0.10	12	3.60	1	0.19	1	0.03
Mel/pse versus	Hox gene	1	29.06***	12	27.69**	1	1.32	1	86.24***
Mel/buz or pse/buz	Non-Hox gene	1	98.78***	12	5.28	1	3.59	1	68.17***

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

⁽¹⁾ Variation between the three species pairs, mel/pse, mel/buz and pse/buz (df = 2) can be partitioned into variation between the equally distant species pairs mel/buz and pse/buz (df = 1) and variation between the closely related mel/pse and the other two species pairs (df = 1).

Table S3. Statistics of the sequencing process of two BAC clones from a *Drosophila buzzatii* genomic library.

BAC clone	5H14	40C11
Number of clones in plasmid library (insert size 1.5 kb)	960	1,056
Number of clones sequenced by one end	77	-
Number of clones sequenced by both ends	712	864
Number of reads in shotgun sequencing	1,501	1,728
Coverage	7.7x	6x
Number of contigs (> 2kb) after automatic assembly	7	7
Number of contigs (> 2kb) after manual editing	4	6
Number of contigs (< 2kb) after manual editing	2	0
Number of extra reads from plasmid library (clone ends)	47	43
Number of extra reads from plasmid library (custom primers)	22	31
Number of PCR fragments sequenced (physical gaps)	1	3
Number of reads from PCR fragments	4	11
Number of reads for finishing	73	85
Total number of reads produced	1,574	1,813
Total number of reads in final assembly	1,391	1,396
Average read size (bp)	829	890
Size of BAC clone insert (bp)	124,024	132,938
Error rate (per 10 kb)	< 0.01	0.17

Table S4. Primer pairs used for cDNA amplification.

gene	Species	forward	reverse	Size
<i>lab</i>	<i>D. buzzatii</i>	CATCCGTGTCAGCTGCTAAC		570 bp
	<i>D. repleta</i>		TTCACGCGCTTCTTCTGCTTCAT	400 bp
	<i>D. virilis</i>	TCCAGCACCGAAACTACCTG		400 bp
	<i>D. melanogaster</i>			433 bp
<i>pb</i>	<i>D. buzzatii</i>	TCAATTCGCAGCCGTCGATG	GAGAGCGTTTGCCTCTTGTG	458 bp
	<i>D. melanogaster</i>			475 bp
<i>abdA</i>	<i>D. buzzatii</i>		CTGTGTGCAAGCGTTGC	504 bp
	<i>D. virilis</i>	TCGCAGTACTCGTCTCTG		501 bp
	<i>D. repleta</i>		CTGGTAGAACTGCGCAA	424 bp

Table S5. List of observed CNS.

proboscipedia

<i>D. buzzatii</i>						40C11		<i>D. melanogaster</i>		<i>D. pseudoobscura</i>	
								AE001572		AADE01000149	
inicio	fin		tamaño		%			inicio	fin	inicio	fin
128178	128091	=	90	al	85.6	pse				59361	59272
127928	127889	=	41	al	80.5	mel	302111	302151			
127882	127808	=	76	al	84.2	pse				59157	59083
127881	127808	=	74	al	85.1	mel	302160	302230			
127738	127712	=	27	al	85.2	mel	302271	302295			
127738	127718	=	21	al	81.0	pse				59010	58990
127592	127536	=	58	al	79.3	pse				58880	58827
127569	127536	=	34	al	94.1	mel	302404	302437			
127513	127500	=	14	al	100.0	pse				58779	58766
127365	127350	=	16	al	100.0	pse				58692	58677
125591	125571	=	21	al	81.0	pse				56735	56715
125293	125263	=	31	al	93.5	mel	304260	304290			
125293	125240	=	58	al	84.5	pse				56472	56416
125256	125240	=	17	al	94.1	mel	304300	304316			
125178	125159	=	20	al	95.0	mel	304346	304365			
125175	125159	=	17	al	100.0	pse				56382	56366
125064	125019	=	46	al	93.5	pse				56266	56221
125061	125020	=	43	al	90.7	mel	304497	304539			
124813	124790	=	24	al	95.8	mel	304687	304710			
124813	124790	=	24	al	95.8	pse				56010	55987
124742	124687	=	56	al	82.1	mel	305074	305129			
124735	124671	=	67	al	82.1	pse				55888	55822
124583	124554	=	30	al	80.0	mel	305225	305254			
124580	124563	=	18	al	100.0	pse				55753	55736
124505	124437	=	69	al	82.6	pse				55682	55615
124483	124441	=	43	al	93.0	mel	305351	305393			
124275	124241	=	36	al	91.7	pse				55433	55398
124274	124241	=	34	al	85.3	mel	305559	305592			
124166	124131	=	36	al	97.2	mel	305627	305662			
124166	124126	=	41	al	95.1	pse				55341	55301
124094	124040	=	58	al	82.8	pse				55261	55204
124085	124045	=	41	al	97.6	mel	305692	305732			
123957	123919	=	39	al	74.4	pse				55129	55094
123949	123918	=	32	al	81.3	mel	305847	305878			
123075	122988	=	89	al	93.3	mel	306157	306245			

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
123075	122988	=	91	al	90.1	pse		54802	54713
122714	122670	=	47	al	78.7	mel	306382 306428		
122714	122670	=	47	al	80.9	pse		54541	54495
122549	122504	=	46	al	82.6	pse		54407	54365
122540	122521	=	20	al	90.0	mel	306535 306554		
121839	121815	=	25	al	80.0	mel	306656 306680		
121782	121671	=	114	al	86.0	pse		53961	53848
121775	121671	=	107	al	84.1	mel	307050 307155		
121573	121542	=	32	al	93.8	mel	307227 307258		
121573	121542	=	32	al	90.6	pse		53742	53711
121491	121464	=	28	al	89.3	mel	307302 307327		
121488	121464	=	26	al	92.3	pse		53660	53635
121455	121443	=	13	al	100.0	mel	307339 307351		
121455	121437	=	19	al	100.0	pse		53557	53539
121380	121346	=	35	al	82.9	mel	307396 307428		
121371	121346	=	26	al	100.0	pse		53489	53464
121221	121174	=	53	al	79.2	mel	307468 307520		
121221	121180	=	46	al	78.3	pse		53425	53380
121151	121123	=	30	al	83.3	mel	307576 307605		
121151	121129	=	24	al	83.3	pse		53331	53308
121023	120957	=	70	al	91.4	pse		52887	52818
121021	120957	=	68	al	83.8	mel	308208 308275		
120947	120900	=	48	al	97.9	mel	308299 308346		
120946	120888	=	59	al	86.4	pse		52789	52731
120276	120255	=	24	al	79.2	mel	308771 308794		
120240	120229	=	12	al	100.0	pse		52275	52264
120185	120160	=	26	al	76.9	mel	308863 308888		
120167	120144	=	24	al	79.2	pse		52205	52182
119994	119955	=	40	al	80.0	pse		52017	51978
119932	119826	=	107	al	78.5	pse		51904	51798
119916	119887	=	30	al	73.3	mel	309200 309229		
119883	119856	=	28	al	78.6	mel	309233 309260		
119754	119729	=	26	al	76.9	mel	309347 309372		
119725	119654	=	72	al	77.8	pse		51705	51634
119685	119660	=	26	al	80.8	mel	309416 309441		
119576	119464	=	113	al	80.5	pse		51533	51421
119569	119461	=	109	al	78.0	mel	309543 309651		
119455	119431	=	25	al	76.0	mel	309657 309681		
119455	119403	=	53	al	75.5	pse		51412	51360

inicio	fin	tamaño		%		inicio	fin	inicio	fin
119427	119403	=	25	al	76.0	mel	309685	309709	
119094	119061	=	34	al	88.2	mel	310004	310034	
119078	119043	=	36	al	83.3	pse			51120 51085
118966	118941	=	26	al	80.8	pse			51058 51033
118893	118802	=	92	al	95.7	mel	310160	310251	
118893	118802	=	92	al	95.7	pse			50925 50834
118707	118656	=	59	al	81.4	mel	310655	310713	
118703	118585	=	127	al	79.5	pse			50696 50574
118639	118589	=	51	al	96.1	mel	310725	310775	
117995	117966	=	30	al	80.0	mel	311221	311248	
117982	117966	=	17	al	100.0	pse			50084 50068
117922	117888	=	35	al	80.0	mel	311307	311338	
117922	117888	=	35	al	91.4	pse			50006 49972
117767	117716	=	53	al	90.6	mel	311419	311471	
117767	117714	=	55	al	92.7	pse			49873 49819
117333	117306	=	28	al	78.6	pse			49570 49547
117243	117221	=	23	al	78.3	pse			49500 49478
116764	116725	=	40	al	82.5	pse			49114 49075
116636	116569	=	68	al	77.9	pse			48995 48928
116541	116473	=	69	al	75.4	pse			48900 48832
116260	116231	=	30	al	80.0	pse			48593 48564
116057	116038	=	20	al	100.0	pse			48315 48296
115629	115592	=	38	al	78.9	mel	311999	312036	
115599	115575	=	25	al	80.0	pse			47935 47911
115526	115505	=	23	al	78.3	pse			47852 47830
115420	115400	=	25	al	76.0	pse			47724 47700
113402	113371	=	34	al	82.4	mel	314230	314263	
113380	113326	=	56	al	89.3	pse			46625 46570
113350	113317	=	34	al	85.3	mel	314266	314299	
113308	113280	=	29	al	86.2	mel	314301	314329	
113307	113291	=	17	al	100.0	pse			46558 46542
113206	113137	=	70	al	87.1	pse			46392 46323
113204	113148	=	57	al	89.5	mel	314396	314449	
113060	113031	=	31	al	83.9	pse			46190 46160
113056	113032	=	25	al	76.0	mel	314565	314589	
110974	110950	=	26	al	76.9	mel	315914	315938	
109495	109473	=	23	al	95.7	mel	317519	317541	
109488	109447	=	42	al	92.9	pse			43768 43728
109471	109442	=	30	al	83.3	mel	319277	319305	

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
109292	109269	=	24	al	83.3	pse		43633	43610
109222	109200	=	23	al	100.0	mel	319445	319467	
109217	109200	=	18	al	100.0	pse		43579	43562
109174	109119	=	57	al	91.2	mel	319480	319536	
109174	109130	=	45	al	100.0	pse		43545	43501
109006	108976	=	31	al	83.9	mel	319584	319614	
109006	108972	=	35	al	82.9	pse		43438	43405
108953	108895	=	59	al	91.5	mel	319688	319745	
108953	108894	=	60	al	91.7	pse		43366	43308
108766	108697	=	72	al	88.9	pse		43271	43200
108751	108697	=	56	al	94.6	mel	319777	319832	
108585	108557	=	29	al	82.8	pse		43141	43114
108576	108553	=	26	al	80.8	mel	319882	319907	
107869	107831	=	39	al	84.6	mel	320514	320552	
107862	107825	=	38	al	89.5	pse		42283	42246
107830	107801	=	30	al	80.0	mel	320565	320594	
107781	107759	=	23	al	78.3	mel	320625	320647	
107656	107620	=	37	al	97.3	mel	320742	320778	
107654	107608	=	50	al	90.0	pse		42024	41975
107589	107570	=	20	al	100.0	pse		41966	41947
107587	107555	=	33	al	81.8	mel	320789	320820	
107513	107453	=	61	al	91.8	pse		41895	41835
107505	107451	=	55	al	96.4	mel	320875	320929	
107362	107282	=	82	al	81.7	pse		41802	41722
107348	107278	=	71	al	87.3	mel	320965	321035	
106522	106480	=	43	al	83.7	mel	321340	321382	
106521	106486	=	36	al	83.3	pse		41377	41342
106312	106255	=	60	al	78.3	pse		41247	41195
106276	106255	=	22	al	95.5	mel	321513	321534	
106178	106150	=	29	al	82.8	mel	321585	321613	
106067	106042	=	26	al	80.8	mel	321661	321684	
106066	106042	=	25	al	80.0	pse		41056	41034
105963	105915	=	50	al	90.0	mel	321713	321762	
105937	105910	=	28	al	96.4	pse		40936	40909
105850	105727	=	124	al	89.5	mel	321817	321940	
105845	105709	=	137	al	89.1	pse		40845	40710
105672	105615	=	58	al	98.3	mel	321992	322049	
105671	105583	=	91	al	92.3	pse		40679	40590
105614	105593	=	22	al	100.0	mel	322069	322090	

inicio	fin	tamaño		%		inicio	fin	inicio	fin
105316	105293	=	25	al	76.0	pse		39920	39896
105233	105168	=	66	al	84.8	mel	322469	322533	
105233	105168	=	68	al	80.9	pse		39692	39626
105071	105036	=	36	al	100.0	mel	322614	322649	
105071	105036	=	36	al	100.0	pse		39549	39514
104997	104932	=	70	al	77.1	mel	322669	322738	
104992	104965	=	30	al	90.0	pse		39484	39455
104900	104860	=	41	al	90.2	mel	322780	322820	
104896	104857	=	43	al	90.7	pse		39349	39307
104843	104823	=	21	al	95.2	pse		39288	39269
104788	104755	=	34	al	94.1	mel	322894	322927	
104783	104755	=	29	al	93.1	pse		39219	39191
103960	103908	=	54	al	83.3	pse		38928	38875
103945	103907	=	39	al	97.4	mel	323291	323329	
103848	103803	=	46	al	87.0	pse		38752	38707
103846	103778	=	70	al	82.9	mel	323430	323498	
103800	103772	=	29	al	82.8	pse		38694	38666
103564	103445	=	121	al	89.3	mel	323576	323693	
103563	103447	=	118	al	92.4	pse		38534	38419
103404	103383	=	22	al	86.4	pse		38371	38350
103403	103383	=	21	al	90.5	mel	323728	323748	
103311	103250	=	62	al	98.4	pse		38298	38237
103307	103250	=	58	al	94.8	mel	323787	323843	
103240	103178	=	64	al	96.9	pse		38182	38119
103237	103175	=	63	al	95.2	mel	323879	323941	
103122	103104	=	19	al	94.7	mel	323961	323979	
103113	103101	=	13	al	100.0	pse		38075	38063
102933	102883	=	55	al	76.4	pse		37903	37851
102930	102881	=	53	al	75.5	mel	324103	324154	
102792	102769	=	24	al	83.3	mel	324222	324245	
102788	102769	=	20	al	85.0	pse		37773	37754
102118	102069	=	53	al	92.5	pse		37038	36986
102114	102080	=	35	al	94.3	mel	324724	324758	
101975	101902	=	79	al	81.0	pse		36884	36806
101965	101901	=	66	al	97.0	mel	324834	324899	
101787	101744	=	44	al	88.6	mel	324994	325037	
101781	101744	=	40	al	90.0	pse		36699	36660
101717	101670	=	48	al	83.3	pse		36622	36577
101711	101685	=	27	al	100.0	mel	325085	325111	

RESULTADOS

inicio	fin	tamaño		%			inicio	fin	inicio	fin
101645	101607	=	39	al	89.7	mel	325260	325298		
101645	101592	=	54	al	87.0	pse			36451	36401
101578	101515	=	68	al	77.9	mel	325335	325402		
101578	101535	=	48	al	83.3	pse			36370	36323
101263	101179	=	86	al	84.9	pse			36090	36006
101260	101192	=	69	al	84.1	mel	325606	325674		
101113	101087	=	27	al	81.5	pse			35943	35917
100930	100904	=	27	al	85.2	mel	325815	325841		
100930	100899	=	32	al	81.3	pse			35856	35825
100859	100820	=	40	al	92.5	mel	325904	325942		
100859	100786	=	74	al	81.1	pse			35764	35699
100695	100674	=	23	al	78.3	mel	326033	326055		
100673	100648	=	28	al	78.6	pse			35639	35616
100662	100648	=	15	al	100.0	mel	326098	326112		
100565	100546	=	20	al	85.0	mel	326158	326177		
100539	100512	=	28	al	78.6	pse			35533	35507
100329	100299	=	31	al	71.0	mel	326361	326391		
99852	99824	=	33	al	81.8	pse			35066	35034
99807	99768	=	41	al	82.9	pse			34972	34932
99802	99768	=	35	al	82.9	mel	326603	326637		
99659	99641	=	21	al	81.0	pse			34864	34844
99564	99536	=	29	al	89.7	mel	326806	326834		
99558	99536	=	23	al	87.0	pse			34781	34759
99505	99466	=	40	al	92.5	mel	326905	326944		
99505	99457	=	49	al	87.8	pse			34719	34676
99305	99284	=	22	al	86.4	mel	327049	327070		
99120	99050	=	73	al	95.9	mel	327205	327277		
99120	99051	=	72	al	95.8	pse			34412	34341
99039	99005	=	35	al	85.7	mel	327336	327370		
99039	99005	=	35	al	91.4	pse			34279	34245
98946	98917	=	31	al	74.2	pse			34192	34162
98940	98921	=	21	al	81.0	mel	327423	327443		
98889	98864	=	26	al	88.5	mel	327448	327471		
98889	98866	=	24	al	87.5	pse			34161	34140
98680	98583	=	105	al	89.5	pse			34035	33931
98669	98579	=	93	al	89.2	mel	327895	327987		
98524	98508	=	17	al	94.1	pse			33885	33869
98404	98381	=	24	al	91.7	mel	328168	328191		
98404	98386	=	19	al	89.5	pse			33796	33778

inicio		fin	tamaño			%		inicio	fin	inicio	fin
97783		97757	=	27	al	77.8	mel	328506	328531		
97700		97682	=	19	al	100.0	mel	328586	328604		
97597		97534	=	64	al	76.6	mel	328704	328767		
97255		97223	=	33	al	93.9	mel	328974	329006		
97255	Inv	97229	=	27	al	100.0	pse			33043	33069
97150		97128	=	23	al	95.7	mel	329043	329065		
97145	inver	97121	=	25	al	96.0	pse			33107	33131
96486	sión	96461	=	26	al	76.9	pse			33218	33243
95974		95893	=	83	al	90.4	mel	330006	330088		
95961		95891	=	74	al	83.8	pse			31814	31741
95869		95830	=	40	al	95.0	mel	330181	330220		
95858		95804	=	55	al	81.8	pse			31627	31582
95207		95165	=	44	al	93.2	pse			30649	30606
95204		95162	=	46	al	89.1	mel	330438	330483		
95068		95036	=	33	al	78.8	mel	330605	330637		
95065		95045	=	21	al	90.5	pse			30432	30412
94900		94880	=	25	al	76.0	pse			30299	30275
94850		94830	=	21	al	81.0	pse			30249	30229
94819		94795	=	25	al	76.0	pse			30218	30194
94782		94755	=	28	al	78.6	mel	331268	331294		
94743		94719	=	25	al	96.0	mel	331320	331344		
94741		94716	=	26	al	96.2	pse			30153	30128
94532		94507	=	27	al	81.5	pse			29438	29412
94334		94311	=	24	al	79.2	pse			25970	25949
94185		94161	=	25	al	76.0	pse			27372	27348
94017		93984	=	35	al	80.0	mel	333133	333167		
93974		93957	=	18	al	100.0	mel	333103	333120		
93974		93957	=	18	al	100.0	pse			28012	27995
93952		93906	=	47	al	74.5	mel	333016	333062		
93952		93911	=	42	al	81.0	pse			28090	28049
93825	inver	93802	=	24	al	91.7	mel	332505	332528		
93825	sión	93802	=	24	al	91.7	pse			28391	28368
93776		93677	=	103	al	83.5	pse			28533	28435
93769		93676	=	94	al	86.2	mel	332397	332486		
93629		93601	=	29	al	75.9	pse			28678	28650
93483		93457	=	27	al	81.5	mel	332202	332228		
93473		93463	=	11	al	100.0	pse			28781	28771
93456		93418	=	39	al	82.1	pse			28833	28797
93439		93418	=	22	al	90.9	mel	332174	332195		

RESULTADOS

inicio	fin	tamaño			%		inicio	fin	inicio	fin
93396	93360	=	37	al	83.8	mel	332129	332161		
93386	93360	=	27	al	96.3	pse			28873	28847
92374	92322	=	54	al	90.7	pse			26132	26080
92371	92322	=	50	al	92.0	mel	334667	334715		
92307	92276	=	32	al	96.9	mel	334616	334647		
92307	92270	=	38	al	94.7	pse			26190	26153
92138	92108	=	31	al	93.5	mel	334505	334535		
92136	92108	=	31	al	87.1	pse			26316	26286
92093	92082	=	12	al	100.0	pse			26351	26340
92009	91969	=	41	al	92.7	mel	334387	334427		
92009	91969	=	41	al	97.6	pse			26461	26421
91942	91925	=	18	al	100.0	mel	334341	334358		
91900	91873	=	28	al	78.6	pse			26550	26523
91885	91862	=	24	al	79.2	mel	334293	334316		
91772	91746	=	27	al	88.9	mel	334021	334047		
91772	91749	=	24	al	91.7	pse			26667	26644
91614	91580	=	35	al	97.1	pse			26764	26730
91608	91580	=	29	al	100.0	mel	333921	333949		
91540	91517	=	25	al	80.0	pse			26825	26801
91497	91405	=	93	al	92.5	mel	333751	333842		
91495	91408	=	88	al	95.5	pse			26958	26871
91323	91232	=	94	al	85.1	pse			27212	27120
90743	90693	=	52	al	88.5	mel	335390	335441		
90733	90693	=	41	al	90.2	pse			25469	25429
90687	90665	=	23	al	82.6	mel	335455	335477		
90597	90567	=	34	al	79.4	pse			25349	25316
90563	90502	=	63	al	90.5	mel	335577	335639		
90550	90502	=	50	al	92.0	pse			25283	25234
90469	90404	=	66	al	81.8	pse			25150	25085
90468	90423	=	46	al	82.6	mel	335688	335733		
90345	90323	=	23	al	91.3	pse			25013	24992
90270	90210	=	63	al	79.4	pse			24938	24878
90265	90217	=	52	al	80.8	mel	336091	336140		
90097	90070	=	28	al	82.1	mel	336243	336267		
90032	89984	=	52	al	76.9	pse			24723	24678
90002	89984	=	19	al	100.0	mel	336324	336342		
89979	89967	=	13	al	100.0	pse			24626	24614
88498	88442	=	59	al	81.4	mel	337242	337299		
88498	88437	=	62	al	83.9	pse			23604	23543

inicio	fin	tamaño		%			inicio	fin	inicio	fin
88427	88409	=	19	al	100.0	mel	337334	337352		
88419	88394	=	26	al	92.3	pse			23473	23448
88381	88339	=	44	al	84.1	mel	337385	337428		
88376	88339	=	39	al	79.5	pse			23437	23399
88232	88188	=	45	al	77.8	mel	337486	337528		
88231	88189	=	43	al	79.1	pse			22965	22928
88175	88155	=	21	al	95.2	mel	337571	337591		
88072	87962	=	115	al	88.7	mel	337662	337776		
88072	87955	=	123	al	86.2	pse			22751	22629
87757	87726	=	32	al	84.4	mel	337926	337957		
87757	87727	=	31	al	87.1	pse			22491	22461
87702	87680	=	26	al	76.9	mel	337979	338004		
87592	87565	=	28	al	85.7	pse			22367	22340
87554	87531	=	24	al	87.5	mel	338062	338085		
87552	87520	=	33	al	75.8	pse			22339	22308
86715	86681	=	35	al	91.4	mel	338401	338435		
86715	86681	=	36	al	88.9	pse			21887	21852
86593	86568	=	26	al	84.6	mel	338522	338546		
86587	86569	=	19	al	89.5	pse			21732	21715
86506	86459	=	53	al	81.1	mel	338575	338627		
86489	86453	=	38	al	84.2	pse			21646	21610
86432	86389	=	44	al	86.4	mel	338668	338711		
86432	86400	=	36	al	83.3	pse			21532	21497
86326	86268	=	60	al	78.3	pse			21394	21335
86311	86268	=	44	al	86.4	mel	338848	338891		
86190	86138	=	54	al	87.0	mel	339315	339367		
86169	86112	=	61	al	82.0	pse			20668	20608
85787	85762	=	26	al	80.8	mel	339726	339749		
85780	85759	=	22	al	90.9	pse			20093	20072
85726	85680	=	48	al	83.3	mel	339790	339837		
85726	85680	=	48	al	85.4	pse			20021	19974
85654	85635	=	20	al	85.0	pse			19915	19896
85584	85518	=	72	al	91.7	pse			19877	19806
85583	85523	=	61	al	100.0	mel	339913	339973		
85506	85480	=	27	al	81.5	pse			19722	19696
85379	85357	=	23	al	78.3	pse			19663	19641
85353	85319	=	35	al	80.0	mel	340085	340118		
85353	85336	=	22	al	81.8	pse			19633	19612
85215	85158	=	58	al	84.5	pse			19476	19419

RESULTADOS

inicio	fin	tamaño			%		inicio	fin	inicio	fin
85202	85186	=	17	al	100.0	mel	340241	340257		
84890	84865	=	26	al	76.9	mel	340537	340561		
84578	84545	=	34	al	85.3	mel	340824	340857		
84573	84532	=	42	al	83.3	pse			19039	18998
84438	84356	=	84	al	90.5	mel	340935	341018		
84382	84356	=	27	al	100.0	pse			18938	18912
84221	84179	=	43	al	76.7	mel	341169	341211		
84206	84188	=	19	al	94.7	pse			18778	18760
84154	84096	=	59	al	83.1	mel	341237	341295		
84154	84103	=	53	al	90.6	pse			18719	18667
83851	83791	=	63	al	82.5	mel	341420	341482		
83840	83783	=	60	al	85.0	pse			18496	18437
83738	83689	=	52	al	80.8	mel	341515	341566		
83738	83709	=	30	al	83.3	pse			18406	18377
83699	83671	=	29	al	79.3	pse			18363	18338
83628	83586	=	43	al	79.1	pse			18286	18244
83618	83606	=	13	al	100.0	mel	341629	341641		
83580	83542	=	39	al	74.4	mel	341654	341692		
83567	83539	=	29	al	79.3	pse			18203	18175
83446	83430	=	17	al	100.0	pse			18091	18075
83381	83355	=	27	al	85.2	mel	341804	341827		
83268	83241	=	28	al	100.0	pse			17965	17938
83266	83222	=	47	al	89.4	mel	341919	341965		
83172	83156	=	17	al	100.0	mel	342033	342049		
83170	83156	=	15	al	100.0	pse			17889	17875
83138	83105	=	34	al	94.1	pse			17855	17822
83125	83097	=	32	al	81.3	mel	342064	342095		
82236	82216	=	21	al	81.0	mel	342440	342460		
82172	82126	=	47	al	85.1	pse			17503	17458
82167	82128	=	41	al	95.1	mel	342502	342541		
81969	81903	=	67	al	82.1	pse			17303	17237
81960	81871	=	90	al	81.1	mel	342709	342798		
81895	81871	=	26	al	80.8	pse			17212	17187
81741	81670	=	73	al	90.4	mel	342963	343031		
81740	81670	=	73	al	90.4	pse			17000	16932
81606	81578	=	30	al	83.3	mel	343061	343090		
81606	81588	=	19	al	100.0	pse			16862	16844
81529	81494	=	36	al	94.4	mel	343161	343196		
81529	81489	=	41	al	90.2	pse			16699	16659

inicio	fin	tamaño		%		inicio	fin	inicio	fin
81420	81327	=	95	al	93.7	mel	343278 343372		
81418	81339	=	80	al	97.5	pse		16596	16517
81216	81152	=	67	al	82.1	mel	343392 343456		
81210	81149	=	64	al	82.8	pse		16437	16375
81137	81111	=	29	al	82.8	pse		16367	16339
81047	80936	=	119	al	84.0	mel	343537 343653		
81022	80976	=	47	al	93.6	pse		16218	16173
80974	80939	=	36	al	100.0	pse		16157	16122
80855	80821	=	35	al	82.9	mel	343790 343824		
80855	80832	=	24	al	100.0	pse		15973	15950
80679	80655	=	25	al	76.0	pse		15617	15593
80170	80154	=	17	al	100.0	pse		15035	15019
80121	80065	=	60	al	81.7	pse		14985	14929
80113	80067	=	47	al	93.6	mel	344455 344501		
79896	79862	=	35	al	91.4	mel	344637 344671		
79896	79856	=	41	al	90.2	pse		14794	14754
79620	79581	=	41	al	80.5	pse		14628	14589
79609	79581	=	30	al	83.3	mel	344845 344874		
79249	79227	=	25	al	76.0	pse		14306	14282
79170	79136	=	35	al	80.0	pse		14217	14183
78869	78844	=	27	al	77.8	mel	345407 345433		
78324	78273	=	52	al	86.5	pse		13815	13764
78315	78277	=	42	al	78.6	mel	345638 345679		
76843	76827	=	17	al	94.1	pse		12444	12428
75250	75225	=	26	al	76.9	mel	349109 349134		

RESULTADOS

<i>labial</i>						<i>D.melanogaster</i>		<i>D. pseudoobscura</i>	
<i>D. buzzatii</i>						AE001572		AADE01000437	
inicio	fin	tamaño		%		inicio	fin	inicio	fin
93374	93409	=	36	al	77.8	pse		58888	58853
93856	93884	=	29	al	75.9	mel	374162 374190		
94584	94606	=	23	al	78.3	pse		57871	57849
94627	94656	=	30	al	80.0	mel	374573 374601		
94627	94664	=	38	al	84.2	pse		57813	57776
94675	94692	=	18	al	94.4	pse		57771	57754
94683	94699	=	17	al	94.1	mel	374622 374637		
94749	94764	=	16	al	93.8	pse		57699	57684
94835	94865	=	31	al	87.1	pse		57636	57606
94891	94915	=	25	al	100.0	mel	374807 374831		
94891	94922	=	32	al	96.9	pse		57558	57527
95039	95101	=	66	al	84.8	mel	374880 374945		
95039	95109	=	72	al	84.7	pse		57475	57404
95273	95292	=	20	al	90.0	pse		57242	57224
95274	95296	=	23	al	82.6	mel	375131 375152		
96975	96997	=	23	al	82.6	pse		56750	56729
97580	97604	=	25	al	76.0	mel	375584 375607		
100621	100694	=	75	al	89.3	pse		56160	56086
100635	100728	=	97	al	88.7	mel	376536 376632		
100849	100888	=	40	al	87.5	mel	376678 376717		
100849	100892	=	44	al	88.6	pse		55975	55932
100943	100981	=	39	al	94.9	pse		55885	55847
100947	100981	=	35	al	97.1	mel	376771 376805		
101045	101087	=	43	al	100.0	pse		55755	55713
101050	101087	=	38	al	100.0	mel	376871 376908		
101149	101189	=	43	al	83.7	mel	376973 377014		
101167	101197	=	31	al	83.9	pse		55625	55596
101200	101227	=	29	al	82.8	mel	377019 377047		
101267	101297	=	31	al	87.1	pse		55442	55413
101277	101292	=	16	al	100.0	mel	377163 377178		
101416	101434	=	20	al	90.0	pse		55359	55340
101438	101462	=	26	al	88.5	pse		55325	55300
101545	101581	=	38	al	86.8	mel	377931 377968		
101604	101627	=	26	al	76.9	mel	378013 378037		
101692	101714	=	24	al	83.3	mel	378096 378119		

inicio	fin	tamaño		%		inicio	fin	inicio	fin
101786	101828	=	43	al	88.4	mel	378172	378214	
105496	105535	=	40	al	85.0	mel	380440	380476	
105496	105532	=	37	al	91.9	pse		52438	52402
105556	105621	=	66	al	90.9	mel	380688	380753	
105556	105614	=	59	al	98.3	pse		51931	51873
105676	105811	=	145	al	80.7	mel	380855	380993	
105689	105768	=	81	al	91.4	pse		51756	51677
105774	105803	=	30	al	86.7	pse		51665	51637
106376	106399	=	25	al	80.0	pse		51217	51193
106473	106566	=	94	al	97.9	mel	381608	381701	
106483	106576	=	94	al	100.0	pse		51111	51018
106687	106714	=	28	al	92.9	mel	382145	382172	
106693	106714	=	22	al	100.0	pse		50772	50751
106867	106966	=	100	al	90.0	pse		50638	50542
106877	106921	=	45	al	95.6	mel	382271	382314	
106931	106966	=	38	al	84.2	mel	382322	382357	
107107	107236	=	131	al	90.8	pse		50446	50320
107141	107237	=	98	al	92.9	mel	382478	382575	
107319	107371	=	56	al	87.5	pse		50223	50168
107323	107370	=	50	al	90.0	mel	382607	382656	
108406	108466	=	61	al	85.2	pse		49346	49291
108417	108471	=	55	al	87.3	mel	383359	383413	
108538	108575	=	38	al	86.8	pse		49238	49202
108539	108579	=	41	al	85.4	mel	383464	383503	
108924	108947	=	26	al	76.9	pse		49141	49116
109002	109066	=	65	al	96.9	mel	383649	383712	
109002	109066	=	65	al	96.9	pse		49063	49000
109081	109112	=	32	al	96.9	mel	383750	383781	
109081	109138	=	58	al	84.5	pse		48967	48913
109175	109243	=	69	al	92.8	mel	383858	383926	
109175	109243	=	69	al	95.7	pse		48864	48796
109279	109313	=	35	al	85.7	mel	383991	384022	
109284	109314	=	31	al	83.9	pse		48717	48687
109421	109497	=	78	al	82.1	pse		48279	48202
109424	109497	=	75	al	84.0	mel	384629	384703	
109574	109592	=	19	al	89.5	pse		48079	48061
109600	109624	=	25	al	92.0	mel	384877	384901	
109600	109624	=	25	al	96.0	pse		48051	48027
109801	109840	=	43	al	83.7	pse		47860	47818

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
109806	109846	=	41	al	85.4	mel	384995	385034	
109860	109892	=	35	al	80.0	mel	385064	385098	
109860	109892	=	35	al	85.7	pse		47777	47743
110042	110054	=	13	al	100.0	pse		47555	47543
110178	110223	=	52	al	76.9	pse		47521	47472
110191	110253	=	70	al	80.0	mel	385578	385645	
110226	110262	=	37	al	86.5	pse		47455	47419
110296	110354	=	59	al	88.1	pse		47357	47300
110300	110355	=	56	al	89.3	mel	385732	385787	
110388	110417	=	30	al	80.0	mel	385829	385855	
110395	110417	=	23	al	82.6	pse		47238	47217
110791	110816	=	26	al	76.9	mel	386011	386035	
111279	111307	=	29	al	82.8	pse		46823	46795
111375	111408	=	34	al	85.3	mel	386374	386407	
111375	111411	=	39	al	79.5	pse		46744	46706
111598	111640	=	43	al	86.0	pse		46572	46530
111660	111691	=	32	al	81.3	pse		46510	46479
111830	111899	=	71	al	85.9	pse		46429	46359
111831	111883	=	54	al	81.5	mel	386656	386708	
112384	112410	=	27	al	92.6	mel	386855	386881	
112384	112420	=	39	al	87.2	pse		46215	46177
112434	112459	=	26	al	84.6	mel	386903	386926	
112437	112461	=	25	al	84.0	pse		46162	46139
113018	113047	=	30	al	76.7	pse		44020	43993
113613	113697	=	88	al	88.6	mel	388435	388522	
113615	inver 113639	=	25	al	100.0	pse		44662	44638
113646	sión 113705	=	61	al	93.4	pse		44761	44701
113774	113821	=	48	al	77.1	pse		44818	44774
113788	113814	=	27	al	85.2	mel	388377	388403	
113888	113906	=	19	al	100.0	pse		44906	44888
113889	113919	=	31	al	80.6	mel	388293	388323	
113917	113939	=	23	al	91.3	pse		44958	44936
114013	114042	=	30	al	83.3	pse		45032	45003
114531	inver 114572	=	42	al	78.6	mel	387971	388010	
114537	sión 114563	=	27	al	88.9	pse		45135	45109
114716	114774	=	59	al	86.4	pse		45233	45175
114726	114774	=	49	al	91.8	mel	387887	387935	
114852	114891	=	40	al	100.0	pse		45335	45296
114853	114891	=	39	al	100.0	mel	387785	387823	

inicio	fin	tamaño			%	inicio	fin	inicio	fin
114995	115013	=	21	al	85.7	pse		45404	45384
115106	115132	=	27	al	100.0	mel	387599 387625		
115106	115131	=	26	al	96.2	pse		45528	45503
115232	115302	=	74	al	82.4	mel	387478 387551		
116595	116622	=	28	al	92.9	mel	389264 389291		
116938	116968	=	31	al	80.6	pse		43622	43592
116944	116968	=	25	al	76.0	mel	389589 389613		
116980	117003	=	24	al	95.8	mel	389617 389640		
116980	117003	=	24	al	95.8	pse		43588	43565
117143	117166	=	24	al	83.3	pse		43425	43403
117179	117247	=	72	al	79.2	pse		43390	43319
117210	117247	=	38	al	84.2	mel	389854 389891		
118161	118189	=	31	al	87.1	pse		42992	42962
118239	118334	=	96	al	88.5	mel	390353 390447		
118240	118342	=	103	al	86.4	pse		42820	42719
118376	118397	=	22	al	86.4	mel	390493 390514		
118536	118590	=	55	al	87.3	pse		42551	42499
118537	118590	=	54	al	83.3	mel	390583 390634		
119202	119224	=	23	al	78.3	pse		42064	42046
119262	119326	=	66	al	80.3	mel	391099 391161		
119268	119328	=	62	al	82.3	pse		41998	41939
119456	119516	=	61	al	85.2	mel	391235 391294		
119471	119497	=	27	al	92.6	pse		41874	41848
119669	119756	=	88	al	86.4	pse		41643	41557
119685	119756	=	72	al	88.9	mel	391427 391496		
119777	119827	=	51	al	90.2	mel	391504 391554		
119780	119827	=	48	al	91.7	pse		41546	41499
119981	120046	=	67	al	92.5	pse		41096	41030
119992	120053	=	64	al	79.7	mel	392235 392297		
120072	120091	=	20	al	95.0	pse		41021	41002
120216	120246	=	32	al	84.4	mel	392489 392520		
120216	120235	=	20	al	100.0	pse		40778	40759
121351	121410	=	61	al	95.1	pse		40557	40497
121370	121416	=	48	al	81.3	mel	392759 392806		
121970	122038	=	69	al	94.2	mel	393143 393211		
121971	122132	=	170	al	84.1	pse		40084	39918
122047	122135	=	91	al	91.2	mel	393230 393319		
122230	122267	=	38	al	84.2	mel	393389 393426		
122230	122281	=	53	al	84.9	pse		39874	39822

RESULTADOS

inicio	fin	tamaño			%		inicio	fin	inicio	fin
122393	122413	=	21	al	100.0	mel	393776	393796		
122393	122416	=	24	al	95.8	pse			39757	39734
123457	123522	=	66	al	83.3	pse			39123	39058
123460	123522	=	63	al	85.7	mel	394437	394499		

<i>abd-A</i>						<i>D.melanogaster</i>		<i>D. pseudoobscura</i>	
<i>D. buzzatii</i> 5H14						DMU31961		AADE01000036	
inicio	fin	tamaño		%		inicio	fin	inicio	fin
300	346	=	47	al	80.9	pse		77999	78045
317	343	=	27	al	77.8	mel	152630 152656		
404	464	=	61	al	88.5	pse		78115	78175
407	456	=	50	al	94.0	mel	152739 152788		
514	562	=	49	al	87.8	pse		78297	78344
518	572	=	56	al	76.8	mel	152855 152909		
737	774	=	39	al	79.5	pse		78399	78437
753	780	=	28	al	78.6	mel	152948 152975		
970	1052	=	83	al	84.3	mel	152998 153079		
1211	1235	=	25	al	96.0	pse		79197	79221
1216	1243	=	28	al	85.7	mel	153102 153129		
1363	1389	=	28	al	78.6	pse		79311	79338
1710	1727	=	18	al	100.0	pse		79528	79545
3529	3544	=	17	al	88.2	pse		81189	81205
4112	4168	=	57	al	78.9	mel	155222 155277		
4125	4162	=	38	al	84.2	pse		81627	81664
4321	4344	=	24	al	79.2	pse		81880	81903
5232	5242	=	11	al	100.0	mel	156597 156607		
5964	6007	=	44	al	86.4	mel	156698 156741		
5974	6007	=	34	al	88.2	pse		83110	83143
6068	6129	=	62	al	82.3	pse		83226	83287
6073	6147	=	76	al	84.2	mel	156832 156907		
6136	6152	=	17	al	100.0	pse		83316	83332
7240	7259	=	20	al	100.0	mel	158026 158045		
7241	7259	=	19	al	94.7	pse		83955	83973
7288	7341	=	55	al	80.0	pse		84033	84087
7294	7340	=	48	al	83.3	mel	158116 158163		
7433	7492	=	61	al	83.6	mel	158185 158243		
7443	7480	=	38	al	86.8	pse		84203	84238
7558	7618	=	61	al	88.5	pse		84301	84361
7563	7611	=	49	al	87.8	mel	158270 158318		
7689	7708	=	20	al	95.0	mel	158355 158373		
7727	7763	=	37	al	75.7	mel	158376 158412		
7863	7919	=	57	al	78.9	pse		84467	84523
7870	7919	=	52	al	82.7	mel	158454 158505		

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
7992	8033	=	42	al	76.2	mel	158615	158656	
7992	8022	=	31	al	80.6	pse			84646 84676
8360	8457	=	99	al	84.8	pse			85200 85298
8371	8457	=	88	al	89.8	mel	159191	159278	
8462	8527	=	66	al	90.9	pse			85341 85406
8464	8527	=	64	al	93.8	mel	159292	159355	
8797	8880	=	84	al	97.6	mel	159666	159749	
8797	8880	=	84	al	98.8	pse			85747 85830
8907	8928	=	22	al	90.9	mel	159756	159777	
8917	8930	=	14	al	100.0	pse			85856 85869
8970	9026	=	59	al	89.8	pse			85933 85989
8978	9026	=	49	al	89.8	mel	159853	159898	
9368	9384	=	17	al	100.0	mel	160332	160348	
9393	9496	=	104	al	80.8	mel	160380	160479	
9393	9458	=	66	al	86.4	pse			86395 86460
9477	9506	=	30	al	93.3	pse			86478 86507
9595	9624	=	31	al	80.6	pse			86592 86622
9671	9691	=	21	al	85.7	pse			86675 86695
10072	10088	=	17	al	100.0	pse			86988 87004
10074	10094	=	21	al	85.7	mel	161019	161039	
10134	10153	=	22	al	90.9	mel	161069	161090	
10205	10230	=	26	al	80.8	pse			87107 87132
10224	10249	=	26	al	76.9	mel	161151	161175	
10290	10315	=	26	al	76.9	mel	161239	161260	
10341	10367	=	27	al	77.8	mel	161261	161282	
10400	10432	=	33	al	81.8	pse			87356 87388
10401	10427	=	27	al	85.2	mel	161320	161346	
10636	10664	=	30	al	83.3	mel	161463	161492	
10638	10674	=	38	al	78.9	pse			87538 87575
10712	10723	=	12	al	100.0	pse			87606 87617
10728	10757	=	31	al	80.6	pse			87625 87655
10742	10765	=	25	al	76.0	mel	161529	161553	
10993	11066	=	75	al	85.3	mel	161660	161734	
11004	11070	=	68	al	83.8	pse			87776 87843
11138	11166	=	29	al	79.3	mel	161767	161795	
11151	11174	=	24	al	91.7	pse			87945 87968
11249	11284	=	38	al	78.9	pse			88024 88060
11364	11395	=	35	al	74.3	pse			88123 88157
11647	11665	=	19	al	89.5	pse			88617 88635

inicio	fin	tamaño		%		inicio	fin	inicio	fin
11805	11852	=	48	al	85.4	mel	162573	162620	
11808	11852	=	45	al	84.4	pse		88810	88854
11929	11984	=	56	al	94.6	mel	162713	162768	
11929	11984	=	56	al	92.9	pse		88959	89014
13155	13177	=	23	al	78.3	pse		89371	89393
13379	13398	=	20	al	90.0	mel	163627	163646	
13470	13491	=	22	al	81.8	pse		89549	89570
13503	13523	=	22	al	81.8	pse		89583	89604
13671	13760	=	95	al	90.5	mel	163782	163876	
13676	13749	=	80	al	88.8	pse		89751	89830
13777	13799	=	24	al	87.5	pse		89881	89904
13828	13847	=	20	al	85.0	pse		89925	89944
13901	13946	=	46	al	93.5	pse		89976	90021
13902	13929	=	28	al	100.0	mel	163964	163991	
13964	13985	=	22	al	95.5	mel	164050	164071	
13964	14003	=	42	al	83.3	pse		90035	90075
14103	14180	=	79	al	91.1	mel	164103	164181	
14103	14178	=	76	al	96.1	pse		90088	90163
14189	14211	=	23	al	78.3	pse		90173	90195
14353	14408	=	59	al	86.4	mel	164280	164338	
14356	14408	=	55	al	90.9	pse		90271	90325
14734	14753	=	20	al	90.0	pse		90560	90579
14860	14885	=	26	al	80.8	pse		90692	90717
16251	16271	=	22	al	90.9	mel	165609	165630	
16423	16480	=	58	al	86.2	mel	165701	165757	
16430	16480	=	51	al	82.4	pse		91924	91973
16491	16516	=	26	al	92.3	mel	165761	165786	
16491	16528	=	38	al	78.9	pse		91977	92011
16576	16599	=	24	al	87.5	mel	165854	165877	
16625	16658	=	34	al	91.2	pse		92073	92106
16637	16657	=	21	al	95.2	mel	165899	165919	
16740	16773	=	34	al	97.1	pse		92128	92161
16741	16773	=	33	al	87.9	mel	165951	165983	
16817	16842	=	26	al	76.9	mel	166022	166047	
16817	16845	=	29	al	79.3	pse		92208	92236
16869	16982	=	115	al	87.8	mel	166125	166239	
16877	16954	=	79	al	86.1	pse		92305	92382
16990	17037	=	48	al	91.7	mel	166272	166319	
16990	17081	=	98	al	86.7	pse		92389	92484

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
17052	17080	=	29	al	79.3	mel	166341	166367	
17202	17246	=	45	al	75.6	mel	166709	166752	
17206	17231	=	26	al	92.3	pse		92815	92840
17318	17425	=	115	al	74.8	pse		92925	93039
17322	17431	=	112	al	81.3	mel	166857	166968	
17446	17474	=	32	al	75.0	pse		93130	93161
17460	17485	=	26	al	76.9	mel	167008	167033	
17576	17662	=	87	al	88.5	pse		93248	93333
17586	17666	=	83	al	89.2	mel	167062	167144	
18091	18190	=	100	al	93.0	mel	167344	167443	
18097	18190	=	94	al	91.5	pse		93615	93708
18331	18417	=	87	al	86.2	pse		93750	93834
18346	18417	=	72	al	90.3	mel	167496	167565	
18473	18494	=	22	al	81.8	mel	167591	167612	
18674	18705	=	32	al	71.9	mel	167776	167806	
18728	18750	=	23	al	78.3	mel	167816	167836	
19144	19207	=	66	al	87.9	mel	168099	168164	
19154	19202	=	49	al	87.8	pse		94486	94534
19285	19317	=	33	al	97.0	mel	168252	168284	
19285	19319	=	35	al	94.3	pse		94664	94698
19455	19506	=	55	al	85.5	mel	168332	168386	
19457	19492	=	36	al	86.1	pse		94770	94805
19595	19670	=	78	al	84.6	pse		94939	95016
19598	19655	=	58	al	86.2	mel	168463	168520	
19701	19733	=	33	al	75.8	mel	168566	168598	
19702	19737	=	36	al	80.6	pse		95056	95090
19765	19798	=	35	al	82.9	mel	168656	168687	
19765	19798	=	34	al	91.2	pse		95131	95162
19837	19873	=	37	al	81.1	pse		95182	95218
19849	19873	=	25	al	84.0	mel	168719	168743	
19944	19981	=	40	al	85.0	mel	168853	168892	
19944	19971	=	28	al	96.4	pse		95324	95351
20126	20225	=	100	al	87.0	pse		95488	95587
20142	20225	=	84	al	89.3	mel	169006	169089	
20958	21025	=	68	al	80.9	mel	169736	169803	
20977	21022	=	46	al	80.4	pse		96425	96470
21066	21091	=	26	al	88.5	mel	169892	169917	
21066	21088	=	23	al	91.3	pse		96556	96578
21370	21439	=	70	al	87.1	pse		96748	96817

inicio	fin	tamaño		%		inicio	fin	inicio	fin
21376	21453	=	78	al	87.2	mel	170040	170117	
21477	21515	=	39	al	92.3	mel	170164	170202	
21483	21523	=	41	al	87.8	pse			96881 96921
21531	21560	=	30	al	96.7	mel	170205	170234	
21531	21560	=	30	al	96.7	pse			96922 96951
21838	21876	=	39	al	79.5	mel	170539	170577	
21851	21914	=	65	al	80.0	pse			97328 97390
22541	22576	=	39	al	82.1	pse			97835 97873
22740	22849	=	110	al	84.5	pse			97946 98055
22753	22817	=	67	al	85.1	mel	171457	171522	
22831	22849	=	19	al	94.7	mel	171539	171557	
22866	22970	=	105	al	92.4	pse			98074 98178
22885	22971	=	87	al	96.6	mel	171599	171685	
23001	23032	=	32	al	78.1	pse			98224 98255
23003	23015	=	13	al	100.0	mel	171728	171740	
23284	23308	=	25	al	76.0	pse			98529 98550
23313	23361	=	50	al	80.0	pse			98561 98610
23323	23364	=	42	al	81.0	mel	172198	172239	
23368	23426	=	59	al	76.3	pse			98635 98691
23382	23430	=	49	al	87.8	mel	172277	172323	
23469	23494	=	27	al	85.2	mel	172378	172404	
23506	23550	=	45	al	88.9	pse			98816 98860
23515	23552	=	38	al	84.2	mel	172440	172477	
23608	23661	=	54	al	77.8	mel	172502	172555	
23624	23663	=	40	al	82.5	pse			98881 98920
23802	23813	=	12	al	100.0	mel	172678	172689	
23803	23824	=	22	al	81.8	pse			99042 99062
23875	23899	=	25	al	76.0	mel	172738	172760	
23910	23934	=	25	al	80.0	pse			99140 99163
27534	27558	=	25	al	80.0	pse			102373 102397
27623	27644	=	23	al	87.0	mel	175933	175955	
27625	27642	=	18	al	83.3	pse			102462 102479
27788	27827	=	40	al	82.5	mel	176034	176073	
27788	27824	=	37	al	83.8	pse			102578 102614
28119	28168	=	53	al	77.4	pse			102886 102935
28916	28955	=	40	al	80.0	pse			103592 103630
28940	28962	=	23	al	78.3	mel	176712	176733	
29076	29174	=	99	al	88.9	pse			103735 103833
29081	29169	=	89	al	86.5	mel	176822	176910	

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
29307	29364	=	58	al	87.9	pse		103913	103970
29325	29365	=	41	al	90.2	mel	177045 177085		
29387	29494	=	110	al	84.5	pse		103991	104099
29406	29478	=	73	al	91.8	mel	177111 177183		
29557	29641	=	90	al	85.6	mel	177275 177364		
29557	29596	=	40	al	97.5	pse		104236	104275
29612	29654	=	43	al	81.4	pse		104294	104336
30601	30622	=	22	al	86.4	pse		104874	104895
30625	30659	=	35	al	82.9	mel	177761 177795		
30625	30690	=	66	al	84.8	pse		104920	104985
30672	30690	=	19	al	100.0	mel	177803 177821		
30721	30761	=	41	al	75.6	pse		105034	105074
30723	30761	=	39	al	79.5	mel	177860 177898		
30836	30929	=	94	al	81.9	mel	177951 178044		
30836	30918	=	83	al	89.2	pse		105157	105239
31435	31489	=	55	al	96.4	mel	178192 178246		
31439	31489	=	51	al	98.0	pse		105506	105556
31554	31579	=	26	al	76.9	pse		105615	105640
31601	31635	=	35	al	82.9	mel	178376 178410		
31601	31632	=	32	al	81.3	pse		105674	105705
31645	31681	=	37	al	70.3	pse		105730	105766
31649	31681	=	33	al	75.8	mel	178441 178473		
31966	31995	=	30	al	80.0	mel	178992 179021		
31966	31995	=	30	al	83.3	pse		106145	106174
32156	32276	=	123	al	83.7	pse		106371	106487
32157	32263	=	109	al	88.1	mel	179181 179289		
33687	33705	=	21	al	81.0	pse		106773	106793
33866	33904	=	39	al	74.4	pse		107121	107159
33931	33952	=	22	al	86.4	pse		107231	107252
34014	34052	=	40	al	82.5	pse		107291	107330
34027	34056	=	30	al	86.7	mel	180107 180136		
34088	34110	=	23	al	78.3	pse		107381	107402
34141	34167	=	27	al	77.8	pse		107429	107455
34142	34168	=	27	al	77.8	mel	180196 180222		
34220	34248	=	29	al	79.3	mel	180262 180290		
34232	34248	=	17	al	100.0	pse		107534	107550
34279	34303	=	25	al	88.0	pse		107579	107603
34354	34431	=	78	al	85.9	pse		107702	107776
34373	34433	=	61	al	88.5	mel	180386 180446		

inicio	fin	tamaño		%		inicio	fin	inicio	fin
34474	34515	=	43	al	83.7	pse		107784	107826
34475	34513	=	39	al	84.6	mel	180455 180493		
34657	34688	=	32	al	81.3	mel	180566 180596		
34657	34698	=	42	al	88.1	pse		107843	107884
34782	34829	=	48	al	83.3	pse		107929	107976
34783	34826	=	44	al	81.8	mel	180644 180687		
34867	34892	=	26	al	92.3	mel	180744 180768		
34867	34944	=	82	al	81.7	pse		108046	108122
34897	34919	=	25	al	80.0	mel	180778 180802		
34938	35010	=	73	al	87.7	mel	180847 180919		
34947	35010	=	64	al	87.5	pse		108137	108200
36293	36367	=	75	al	84.0	pse		108918	108988
36299	36367	=	69	al	85.5	mel	181825 181890		
36386	36421	=	36	al	77.8	pse		109009	109044
36401	36419	=	19	al	89.5	mel	181914 181932		
36512	36596	=	87	al	77.0	mel	181967 182053		
36547	36594	=	48	al	85.4	pse		109143	109190
36688	36757	=	70	al	87.1	pse		109355	109420
36690	36740	=	51	al	94.1	mel	182446 182496		
36810	36833	=	24	al	91.7	pse		109443	109466
36816	36826	=	11	al	100.0	mel	182538 182548		
36953	36977	=	25	al	84.0	pse		109542	109565
37041	37078	=	40	al	80.0	pse		109567	109606
37050	37066	=	17	al	94.1	mel	182671 182687		
37204	37219	=	16	al	100.0	mel	182776 182791		
37204	37219	=	16	al	100.0	pse		109709	109724
37267	37343	=	77	al	90.9	mel	182850 182926		
37267	37343	=	77	al	92.2	pse		109768	109844
37422	37518	=	97	al	88.7	pse		109922	110017
37427	37530	=	104	al	82.7	mel	182989 183091		
37543	37556	=	14	al	100.0	mel	183119 183132		
37543	37572	=	30	al	83.3	pse		110048	110077
37594	37627	=	34	al	76.5	pse		110114	110146
37599	37625	=	27	al	81.5	mel	183188 183214		
37678	37803	=	126	al	80.2	pse		110147	110272
37679	37775	=	97	al	82.5	mel	183215 183311		
42033	42056	=	24	al	79.2	pse		112138	112161

RESULTADOS

CG17836-CG14290

<i>D. buzzatii</i> 40C11						<i>D. melanogaster</i> AE003724		<i>D. pseudoobscura</i> AADE001000175	
inicio	fin		tamaño		%	inicio	fin	inicio	fin
69953	69874	=	83	al	81.9	pse		313	395
69794	69762	=	33	al	78.8	pse		475	507
69632	69607	=	26	al	92.3	pse		641	666
69555	69525	=	31	al	80.6	pse		732	762
69512	69448	=	67	al	76.1	pse		769	835
69137	69109	=	29	al	79.3	mel	54961 54989		
68556	68531	=	26	al	76.9	pse		1167	1192
67779	67755	=	25	al	76.0	mel	57456 57479		
67719	67687	=	33	al	90.9	pse		1919	1949
67717	67662	=	56	al	82.1	mel	57532 57587		
67658	67622	=	37	al	78.4	pse		1909	1942
67637	67580	=	62	al	80.6	mel	57597 57658		
67597	67550	=	48	al	83.3	pse		2007	2054
67575	67532	=	48	al	77.1	mel	57663 57710		
67457	67421	=	37	al	75.7	pse		2182	2217
67161	67047	=	117	al	89.7	pse		2502	2616
67148	67047	=	103	al	87.4	mel	58061 58163		
67026	66997	=	30	al	86.7	mel	58187 58216		
67026	66997	=	30	al	93.3	pse		2640	2669
66966	66915	=	56	al	78.6	mel	58240 58295		
66966	66929	=	39	al	89.7	pse		2706	2744
65976	65954	=	24	al	95.8	mel	59263 59286		
65976	65924	=	57	al	82.5	pse		3829	3885
65945	65924	=	22	al	90.9	mel	59300 59321		
65885	65851	=	35	al	97.1	pse		3925	3959
65884	65847	=	38	al	94.7	mel	59370 59407		
65797	65722	=	78	al	84.6	pse		4005	4081
65791	65723	=	71	al	83.1	mel	59444 59514		
64307	64285	=	25	al	76.0	pse		4631	4655
63430	63413	=	18	al	88.9	pse		5181	5198
61461	61450	=	12	al	100.0	mel	62855 62866		
61240	61216	=	26	al	76.9	mel	62991 63016		
61237	61212	=	27	al	88.9	pse		7929	7955
59356	59327	=	30	al	80.0	pse		9792	9821
59264	59247	=	19	al	94.7	pse		9859	9877

inicio	fin	tamaño			%	inicio	fin	inicio	fin
59221	59197	=	25	al	76.0	mel	64919	64943	
58424	58393	=	36	al	75.0	mel	65814	65848	

RESULTADOS

CG31363

<i>D. buzzatii</i>						<i>D. melanogaster</i>		<i>D. pseudoobscura</i>	
5H14						AE003692			
inicio	fin	tamaño		%		inicio	fin	inicio	fin
59041	59017	=	25	al	80.0	mel	70059	70083	<u>AADE01002495</u>
59041	59008	=	38	al	78.9	pse		11971	11934
58917	58891	=	27	al	77.8	mel	70206	70232	
58114	58095	=	21	al	85.7	pse		10251	10231
57090	57064	=	30	al	80.0	mel	75953	75982	
56143	56123	=	21	al	100.0	mel	76122	76142	
56113	56073	=	41	al	90.2	pse		6029	5989
56091	56067	=	25	al	96.0	mel	76215	76239	
56007	55976	=	32	al	81.3	pse		5908	5877
55994	55957	=	38	al	84.2	mel	76302	76339	
55607	55585	=	25	al	76.0	mel	76904	76927	
55332	55299	=	34	al	91.2	pse		4920	4888
55320	55298	=	23	al	100.0	mel	77264	77286	
55297	55275	=	23	al	95.7	pse		4870	4848
55293	55270	=	28	al	78.6	mel	77299	77326	
51694	51682	=	13	al	100.0	pse		2602	2590
51394	51371	=	25	al	76.0	mel	81016	81040	
50844	50812	=	33	al	87.9	mel	81456	81488	
50844	50812	=	33	al	87.9	pse		1595	1563
50778	50759	=	20	al	90.0	mel	81500	81519	
50778	50759	=	20	al	90.0	pse		1507	1488
50561	50534	=	28	al	89.3	mel	81610	81636	
50561	50533	=	31	al	83.9	pse		1293	1263
50510	50484	=	28	al	78.6	mel	81672	81699	
49361	49335	=	27	al	81.5	pse		889	866
35316	35291	=	26	al	80.8	mel	84411	84435	<u>AADE01000322</u>
29277	29253	=	25	al	88.0	pse		884	860
27772	27750	=	23	al	78.3	mel	85276	85298	
26246	26222	=	25	al	76.0	pse		738	762
25320	25295	=	26	al	80.8	mel	90906	90931	
22691	22654	=	38	al	76.3	pse		1026	1063
22568	22545	=	26	al	80.8	pse		1257	1282
22123	22091	=	33	al	78.8	pse		1341	1373
19528	19502	=	28	al	75.0	pse		8256	8283
19390	19368	=	23	al	87.0	pse		8395	8417

inicio	fin	tamaño				%	inicio	fin	inicio	fin
17336	17312	=	26	al	76.9	mel	93578	93602		
17268	17245	=	24	al	87.5	pse			16983	17006
16722	16675	=	48	al	79.2	pse			19213	19260
15114	15091	=	24	al	79.2	mel	94199	94222		
15079	15053	=	27	al	81.5	pse			65030	65056
12663	12639	=	25	al	76.0	mel	94350	94374		
1694	1669	=	26	al	80.8	mel	95860	95885		
1190	1169	=	22	al	81.8	mel	96324	96345		

RESULTADOS

CG1288-CG14609-CG2520

D. buzzatii

40C11

D. melanogaster

AE003672

D. pseudoobscura

AAD000014

inicio	fin	tamaño		%			inicio	fin	inicio	fin
1087	1167	=	82	al	86.6	mel	125136	125056		
1095	1156	=	62	al	93.5	pse			204478	204539
1228	1249	=	22	al	95.5	mel	125042	125021		
1233	1266	=	34	al	82.4	pse			204564	204594
1288	1306	=	19	al	100.0	mel	124977	124959		
1303	1368	=	71	al	81.7	pse			204680	204749
1317	1377	=	67	al	80.6	mel	124936	124870		
2386	2416	=	31	al	93.5	mel	124735	124706		
2673	2702	=	30	al	83.3	pse			205258	205286
2854	2886	=	33	al	84.8	pse			205482	205514
2857	2886	=	30	al	90.0	mel	123983	123954		
2982	3022	=	42	al	85.7	mel	123858	123817		
2983	3022	=	43	al	79.1	pse			205672	205714
3054	3079	=	26	al	96.2	mel	123398	123373		
3054	3081	=	28	al	96.4	pse			205765	205792
3170	3235	=	68	al	91.2	pse			205816	205883
3175	3268	=	103	al	81.6	mel	123342	123241		
3236	3258	=	23	al	91.3	pse			205893	205915
3297	3308	=	12	al	100.0	pse			205965	205976
4378	4398	=	21	al	85.7	pse			206342	206361
4388	4411	=	26	al	84.6	mel	122681	122656		
4488	4529	=	42	al	78.6	pse			206379	206417
4499	4522	=	24	al	91.7	mel	122622	122599		
4720	4779	=	60	al	81.7	mel	122404	122345		
4731	4781	=	51	al	90.2	pse			206679	206729
4792	4818	=	28	al	78.6	pse			206739	206764
8582	8609	=	28	al	78.6	pse			208343	208367
8626	8661	=	36	al	83.3	pse			208396	208431
8628	8672	=	45	al	88.9	mel	120781	120737		
8704	8742	=	39	al	89.7	mel	120713	120675		
8708	8736	=	29	al	96.6	pse			208466	208494
8965	8989	=	25	al	80.0	mel	120566	120542		
9021	9048	=	28	al	82.1	mel	120518	120495		
9024	9048	=	25	al	88.0	pse			208671	208695
10255	10286	=	33	al	90.9	mel	120408	120376		

inicio	fin	tamaño		%		inicio	fin	inicio	fin
10414	10445	=	32	al	84.4	mel	120253	120223	
10485	10528	=	46	al	78.3	mel	120188	120143	
10495	10528	=	34	al	85.3	pse			209019 209052
10628	10648	=	21	al	81.0	pse			209162 209182
10701	10757	=	57	al	86.0	mel	119971	119915	
10701	10751	=	51	al	88.2	pse			209245 209295
10787	10817	=	31	al	87.1	pse			209341 209371
10798	10817	=	20	al	100.0	mel	119853	119834	
11213	11238	=	26	al	100.0	mel	119395	119370	
11249	11299	=	51	al	94.1	pse			209862 209912
11255	11296	=	42	al	97.6	mel	119306	119265	
11478	11548	=	73	al	83.6	pse			210050 210121
11481	11558	=	83	al	83.1	mel	119114	119034	
11854	11884	=	32	al	87.5	pse			210337 210368
11890	11919	=	32	al	78.1	pse			210378 210409
11905	11927	=	24	al	79.2	mel	118736	118713	
11935	11994	=	60	al	90.0	mel	118705	118646	
11935	11994	=	60	al	90.0	pse			210430 210489
12003	12063	=	61	al	88.5	mel	118637	118577	
12009	12063	=	55	al	96.4	pse			210562 210616
12095	12123	=	29	al	89.7	pse			210657 210684
12202	12237	=	36	al	83.3	mel	118476	118441	
12209	12238	=	30	al	86.7	pse			210772 210801
12501	12512	=	12	al	100.0	pse			210992 211003
12953	12977	=	25	al	76.0	mel	118166	118142	
12957	12978	=	22	al	86.4	pse			211203 211224
13051	13084	=	34	al	82.4	mel	118076	118043	
13066	13089	=	25	al	76.0	pse			211309 211333
13314	13353	=	40	al	75.0	mel	117904	117869	
13331	13350	=	22	al	81.8	pse			211476 211497
14000	14018	=	19	al	89.5	mel	117553	117536	
14135	14160	=	27	al	77.8	mel	117419	117393	
14152	14193	=	42	al	76.2	pse			212021 212061
14294	14332	=	39	al	82.1	mel	117249	117211	
14294	14321	=	28	al	92.9	pse			212223 212250
14391	14422	=	34	al	82.4	mel	117177	117146	
31047	31063	=	17	al	100.0	mel	114662	114646	
31047	31063	=	17	al	100.0	pse			214754 214770
33813	33839	=	28	al	85.7	pse			217958 217985

RESULTADOS

inicio	fin	tamaño		%		inicio	fin	inicio	fin
34470	34584	=	115	al	93.9	pse		218563	218675
34477	34617	=	142	al	89.4	mel	111407	111271	
34632	34659	=	29	al	82.8	mel	111259	111231	
34700	34737	=	38	al	78.9	mel	111230	111193	
34759	34785	=	27	al	77.8	mel	111155	111132	
34766	34791	=	26	al	76.9	pse		218763	218788
34824	34916	=	100	al	84.0	mel	111087	110989	
34842	34941	=	108	al	84.3	pse		218800	218904
35079	35103	=	28	al	78.6	pse		219030	219057
35104	35131	=	29	al	75.9	pse		219070	219098
35222	35244	=	23	al	82.6	mel	110757	110736	
35240	35262	=	25	al	80.0	pse		219201	219225
35395	35407	=	13	al	100.0	mel	110534	110522	
35530	35564	=	35	al	77.1	mel	110424	110391	
35712	35749	=	38	al	76.3	pse		219664	219700
35831	35996	=	167	al	92.2	pse		219837	220000
35834	35996	=	163	al	92.6	mel	110153	109993	
36641	36769	=	132	al	85.6	pse		220355	220486
36654	36780	=	127	al	85.8	mel	109640	109514	
36793	36816	=	24	al	79.2	pse		220516	220539
36902	36935	=	34	al	97.1	mel	109431	109398	
36908	36980	=	73	al	83.6	pse		220643	220712
36941	36978	=	38	al	84.2	mel	109386	109349	
37133	37154	=	22	al	86.4	mel	109149	109129	
37215	37249	=	36	al	83.3	mel	109041	109006	
37222	37249	=	30	al	90.0	pse		221103	221132
37292	37360	=	73	al	82.2	pse		221161	221233
37307	37373	=	69	al	85.5	mel	108953	108885	
38385	38399	=	15	al	100.0	pse		221993	222007
38431	38452	=	22	al	100.0	mel	108293	108272	
38508	38538	=	31	al	83.9	pse		222196	222226
38513	38567	=	55	al	81.8	mel	108137	108083	
38549	38579	=	31	al	77.4	pse		222243	222273
38661	38687	=	27	al	88.9	mel	107971	107945	
38755	38781	=	27	al	77.8	mel	107904	107878	
39281	39414	=	134	al	87.3	mel	107430	107297	
39281	39398	=	118	al	88.1	pse		222832	222949
39399	39422	=	24	al	79.2	pse		222961	222984
39516	39573	=	60	al	88.3	mel	107145	107087	

inicio	fin	tamaño			%	inicio	fin	inicio	fin
39517	39572	=	60	al	81.7	pse		223117	223174
39909	39978	=	72	al	90.3	pse		223491	223562
39915	39978	=	64	al	93.8	mel	106770 106707		
40440	40488	=	49	al	77.6	mel	106275 106230		
40440	40488	=	49	al	79.6	pse		224044	224089
40515	40540	=	26	al	80.8	mel	106179 106154		
40669	40694	=	26	al	76.9	pse		224157	224182
41702	41721	=	20	al	90.0	mel	104921 104902		
42098	42134	=	37	al	78.4	pse		225399	225434
42100	42127	=	28	al	78.6	mel	104778 104755		
42189	42223	=	36	al	80.6	mel	104703 104671		
42199	42221	=	23	al	82.6	pse		225502	225523
42853	42875	=	24	al	79.2	pse		226598	226620
45071	45095	=	25	al	76.0	pse		227757	227781
46972	46997	=	26	al	92.3	mel	102031 102006		
46972	46996	=	25	al	92.0	pse		228673	228697
47087	47134	=	48	al	89.6	mel	101953 101906		
47090	47134	=	45	al	95.6	pse		228761	228805
47137	47162	=	26	al	76.9	pse		228812	228833
47431	47458	=	28	al	78.6	pse		229124	229151

Correspondence of regulatory sequences shown in Figure 3 with their domains of expression

labial gene (CHOUINARD y KAUFMAN 1991)

- 1 No Enhancer Activity
- 2 Anterior Midgut
- 3 Maintenance in Procephalon; Eye-Antennal Disc
- 4 Initiation in Procephalon
- 5 Imaginal Disc Repressor
- 6 No Enhancer Activity
- 7 Dorsal Ridge
- 8 Posterior Midgut
- 9 Late Embryonic Head Expression
- 10 No Enhancer Activity

proboscipedia gene (KAPOUN y KAUFMAN 1995)

- 1 Imaginal Antennal and Thoracic Enhancers
- 2 Basal Promoter
- 3 Embryonic Mesodermal, Maxillary, and Labial Lobe Enhancers
- 4 Embryonic Maxillary and Labial Lobe Enhancers and Labial Discs Enhancers
- 5 Late Embryonic Labial Lobe Enhancers (weak) and Labial Discs Enhancers
- 6 Redundant Maxillary and Labial Lobe Enhancers
- 7 “Cryptic” Enhancers

A lab550 HOMRE (Homeotic Response Element)

```
mel  ggagcgggga cagtcgggcg cgggcaatcc atcaatttgg ttgcagtgcc aattacgccca gttact
pse  .....tgc.....a.....c.....
buz  .....ttg.....a.....c.....a.t.....
      HTH  MAD-MED  EA  LAB  BXD  binding sites
```

B iab2(1.7) (fragment) ■ HB ■ KR CNS

```
mel  actgtttcagt ctcataaaaa atgcagcggga aactagaaaa aatcacaaatg attta---c aactgcagcc gagcaacata tacaattctg cacctttca-
pse  ...ccg.a. a.....c.....ta.t ttga----. .cc.gal... ..g....a.t ..a.cgca. cgg.gaatc. ..aga.atgc
buz  ----- .t.tg..t ta.g.tatg. ....a....g .----a..c .g..caaa.....

mel  -----ta atattacgta aaaattgtcc ctttttatt ttgtctgctt gaacagt--- -taccgttgg aaagtcgcaa cttgca-aac cataaaaaat
pse  acctttcac. g.....c.....t.....-g...c ...tgt.gtt g..gt.-t. ....a.-. ....
buz  -----a. ....c.....t.....gca.-g .....g...- -ct.....cc.....

mel  -tgccgttga catgaaaacc ttttgccag- -aaaactcgg cacaaaggca gcagtgcaca aaaga---- -agagacat gaaa----- -gccataaaa
pse  -...t.cc.. .c.....a ...a.ag.a a.....a .....ca. .a.ca.ag.g .g...ggcac ag.....a.....-.....
buz  a.....c... .g.a.....ataa.a- -a.....-c. a.a.. tg..aggtgc t.t.....

mel  cggaagcgg gtat---aa gtga----g aggtgttag gctgggact- ---gggacag caattttag atagcgaatt ttgtatttcc gtaacataaa
pse  ..... tcgaccca.. ----. ....c.g.t c---.gag gat.C.g.....a.....a.....a.....
buz  .a......a----.c.]tgcct. .cca..ca.t tt----- .t.....a. .a.....aa.....

mel  caccgagcct gtgccattcc g----- -agcgcacatc -ccgcacacc gtcgcgcgaaa aaatttcac aggtagcatg
pse  ..... .]----- -agcgcacatc -ccgcacacc gtcgcgcgaaa aaatttcac aggtagcatg
buz  ...a.a.a. aaa.qca.t. ccatttctac tattctgctg caattccctt tgcgtg.... gc..c.d... a.....t.....t.....g..

mel  gaaaattgct aaaaaataaga aacaaaaa-- -actagacaa g---aaagt agagaaaagga atgcggg--- gaaaatcggg gaaaag-cgc
pse  .....c.g. ....a.ga. .t..c.c--- -ag.at.gc ctggc.t.g ..c.t.g.a. ---ac--- ....c..a. tc...a.a.
buz  c.g....taa ....a..a. ]ga.g...ta aaaacagcca g..a.at... c---...a .a.tc...a .a.t.a.aca ....gg.at. ....c-gtt

mel  tgg-----c tgctggccac cagggcgaga cctgccgctg ggcgagtcgg tgcgtgtatt ggta----- -gttccacc Cttt-cacgc actgcgtgcg
pse  a..gccata. ....g... .gc......a. c.g...c... .]----- -t----- -t.....-t.....a
buz  .c----- .-----a .cca...g.. gt..gt...a caaa..ct.c .tct.c.tc. .c.tctgctt ct.....a..t.. cg...agt..
```

Figure S1. Detail of two CRM. Fragments of the AVID alignments (mel/pse and mel/buz) used to generate the VISTA plots of Figure 3. **(A)** lab550-HOMRE (Homeotic Response Element): The four binding sites of this CRM are conserved between all three species in one CNS. **(B)** iab2(1.7) (fragment): All species show five Hunchback (HB) binding sites (three are conserved between all three species; one is shared by *D. melanogaster* and *D. pseudoobscura*; another one is shared by *D. pseudoobscura* and *D. buzzatii*; and there is one unique in *D. melanogaster* and another one in *D. buzzatii*); a unique Krüppel (KR) binding site, where point mutations (red uppercase) in *D. melanogaster* cause gain of expression mutants (Hab1 and Hab2) (Shimell et al. 1994), is conserved in all three species. This region contains 9 CNS in the *D. melanogaster*-*D. pseudoobscura* comparison and 6 CNS between *D. melanogaster*-*D. buzzatii*. PATCH (<http://www.gene-regulation.com/cgi-bin/pub/programs/patch/bin/patch.cgi>) was used for Hunchback binding site prediction in the iab2(1.7) enhancer.

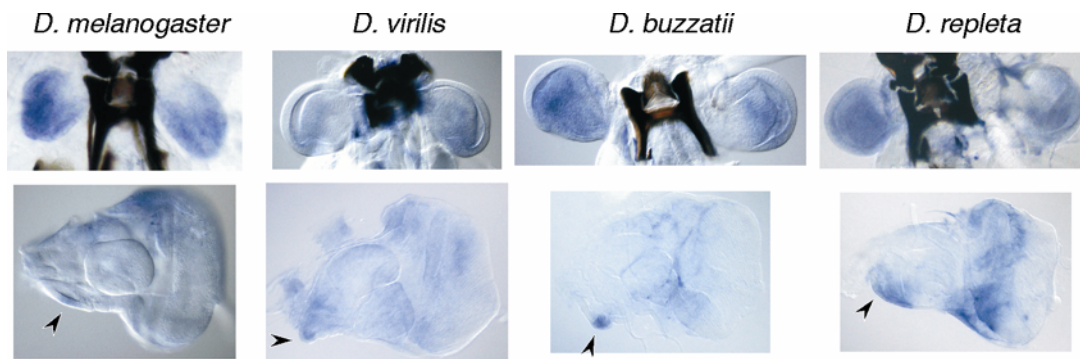


Figure S2. Expression pattern of *pb* in imaginal discs. Labial disc: *pb* is expressed in the labial disc of late third instar larvae and prepupae of the four tested species. Expression is detected over the whole disc, which will give rise to the proboscis. Eye-antennal disc: *pb* is also expressed in the presumptive palpus (arrowhead) of the eye-antennal disc of prepupa and white pupa, which will give rise to the maxillary palps. In overstained discs some *pb* expression can be detected in the presumptive ocellus. Some staining has also been observed in the lateral side of the stalk in white pupae eye-antennal discs of *D. repleta*. However, it has not been determined whether this expression domain is shared by the other species, as this region is easily broken.

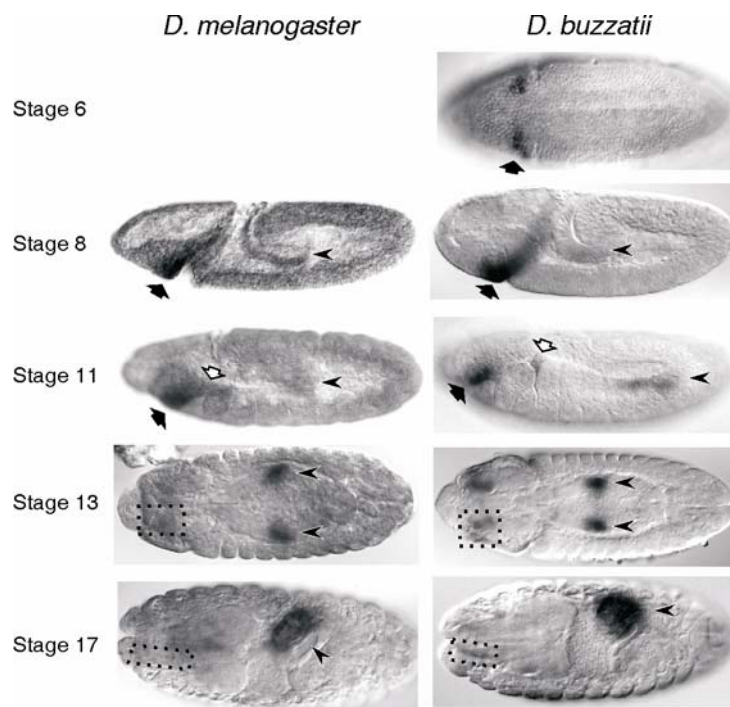


Figure S3. Expression pattern of *lab* in embryos. Cephalic region: *lab* is expressed in a ventral stripe just anterior to the cephalic furrow, which corresponds to the intercalary segment (black arrow). This domain of expression was first detected at stage 6 in *D. buzzatii* and at stage 8 in the other species. Later on *lab* is expressed in the dorsal ridge (white arrow) (stage 11-13) and in sensory structures derived from the intercalary segment (boxed area). Midgut: at stage 8 *lab* expression is detected weakly in the anterior and posterior midgut primordial (arrowheads). When the two primordia fuse (stage 12-13) *lab* expression in this region suddenly increases. Expression is kept at high levels in the second midgut constriction and finally in the second midgut chamber (stage 17).

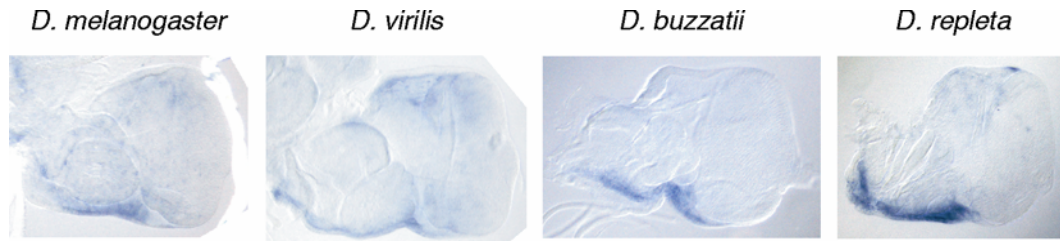


Figure S4. Expression pattern of *lab* in the eye-antennal disc. Expression of *lab* was detected in the peripodial membrane of the eye-antennal disc from late third instar larvae and prepupae. This region will give rise to the lateral and posterior head capsule. No expression has been detected in other imaginal discs.

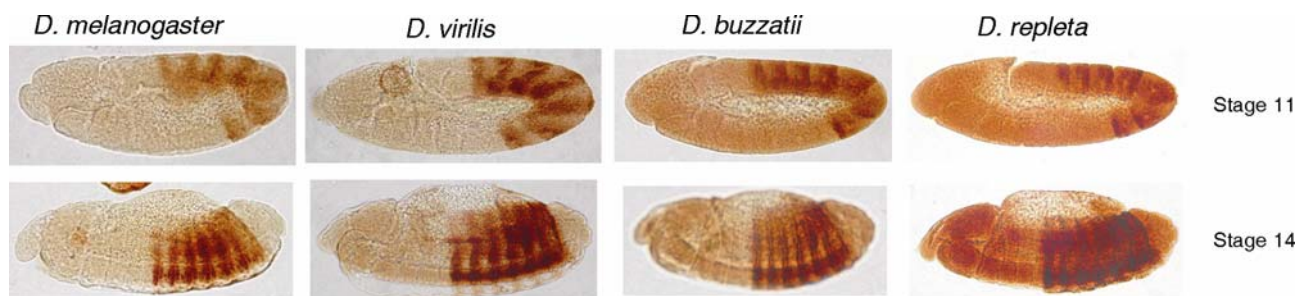


Figure S5. Expression pattern of *abdA* in embryos. ABD-A protein is detected in parasegments 7 through 13 (PS7-13). It is first detected in the ectoderm at stage 10 and later on (stage16-17) is detected on the ventral nervous system.

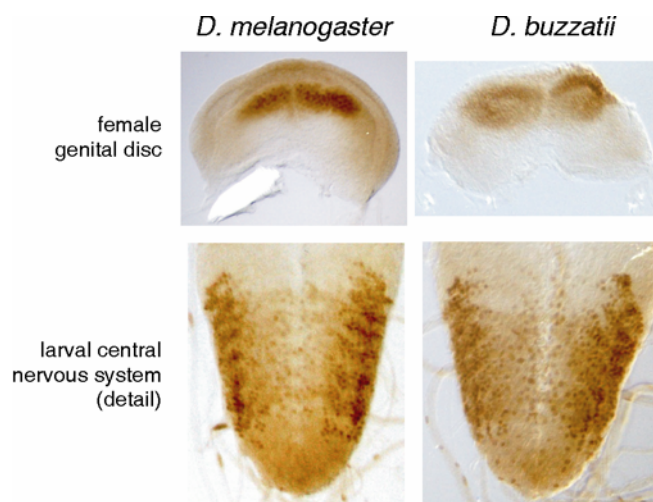


Figure S6. Expression pattern of *abdA* in imaginal discs. Female genital disc: ABD-A protein is observed in the presumptive internal genitalia region (derived from the A8 segment) of the female genital disc. Proper staining with anti-ABD-A was only obtained for *D. melanogaster* and *D. buzzatii*. All imaginal discs were tested for *abdA* expression. Central nervous system of the larva: ABD-A protein is always detected in seven stripes in the larval central nervous system of the four species, although the morphology of the larval central nervous system is slightly different.

3. DISCUSIÓN

3.1 EVOLUCIÓN DE LAS REGIONES CIS-REGULADORAS

Las regiones *cis*-reguladoras controlan el momento y el lugar en el que se expresan los genes, y son los cambios en la regulación génica los responsables de gran parte de la variabilidad fenotípica. Actualmente la comparación de secuencias no codificantes se está convirtiendo en una importante herramienta para la detección de regiones reguladoras, aunque todavía quedan muchos interrogantes sobre como detectar las regiones funcionales (CLARK 2001) o sobre la relación entre la conservación de las secuencias y su función (LUDWIG *et al.* 2000).

Dos estudios con muestras finitas de regiones (BERGMAN y KREITMAN 2001; BERGMAN *et al.* 2002) muestran la utilidad de los estudios comparativos entre *D. pseudoobscura* y *D. melanogaster* para la detección de regiones *cis*-reguladoras. De la comparación del genoma completo, RICHARDS *et al.* (2005) concluyen que estas especies están demasiado alejadas para este tipo de estudio, y sugieren el uso de especies más cercanas a *D. melanogaster*. Teniendo en cuenta estos datos y los resultados aquí obtenidos y descritos (ver Artículo 2), podemos deducir que el tiempo de divergencia “ideal” para estas comparaciones depende del tipo de genes a estudiar. Como hemos mostrado, para genes de desarrollo con patrones complejos de expresión como son los genes Hox, distancias mayores como la que hay entre *D. buzzatii* y *D. melanogaster* son adecuadas. Para obtener una idea global y poder detectar regiones reguladoras de genes bajo diferentes niveles de regulación, sería necesario disponer de secuencias de varias especies situadas a diferente distancia filogenética. La adición de una tercera especie a las comparaciones mejora notablemente la detección de secuencias involucradas en la expresión génica (KOLBE *et al.* 2004). Sería muy informativo disponer de varias especies a esta distancia así como especies cercanas a cada una de ellas. La comparación de varias especies a diferentes tiempos de divergencia es lo que propone precisamente la técnica del sombreado filogenético (phylogenetic shadowing), lo que permite en principio detectar regiones funcionales importantes para diferentes grupos de especies (BOFFELLI *et al.* 2003; CLARK *et al.* 2003). Para regiones del tipo de los genes Hox haría falta añadir especies más alejadas para identificar en más detalle las regiones *cis*-reguladoras. En cambio para genes con menos regulación, las especies a añadir deberían ser más cercanas.

Los genes con expresión compleja durante el desarrollo requieren una región de DNA no codificante mayor para albergar las regiones reguladoras necesarias para

desarrollar el patrón completo (NELSON *et al.* 2004). Esta tendencia se ve claramente en los genes Hox, los cuales poseen algunos de los intrones más largos de los genes de *Drosophila* (MORIYAMA *et al.* 1998). Pero estos genes no solo tienen mayor tamaño de regiones reguladoras sino que además éstas están más conservadas. Los CNS cubren el 20-22% de la secuencia no codificante en estas regiones cuando comparamos especies filogenéticamente distantes como *D. buzzatii* y *D. melanogaster* o *D. pseudoobscura*. Este patrón también es común a los genes con expresión compleja durante el desarrollo, los cuales muestran un nivel de conservación en las regiones no codificantes mucho mayor que los genes con funciones metabólicas o “House-keeping” (BERGMAN y KREITMAN 2001; BERGMAN *et al.* 2002; HALLIGAN *et al.* 2004).

Un factor que dificulta la detección de CRM a partir de la comparación de secuencias es la presencia de reordenaciones. En el caso de los genes *lab* y *pb* hemos detectado pequeñas inversiones de entre 1 y 2 kb que contienen secuencias altamente conservadas pero en orientación inversa en una de las especies comparadas. Como un aspecto de los CRM es que son funcionales en cualquier orientación estas reordenaciones son probablemente neutras cuando los puntos de rotura se sitúan entre CRM. Por otro lado, las regiones invertidas correspondan probablemente a CRM individuales, de manera que la presencia de estas microreordenaciones podría utilizarse como guía para delimitar la posición o los límites entre CRM independientes.

Otra dificultad radica en que la importancia relativa de la arquitectura y composición de los lugares de unión para mantener la funcionalidad de los CRM a lo largo de la evolución es un tema sin resolver. En los casos mirados con detalle, Artículo 2 Figura S1, podemos observar alguna tendencia. Cuando sólo hay un lugar de unión para un determinado factor de transcripción, tanto su posición como su secuencia nucleotídica son mantenidos bajo una fuerte selección purificadora. Cuando hay varios lugares de unión para un mismo factor de transcripción, la presión selectiva que actúa sobre él parece menor y hay por tanto una mayor variabilidad en la secuencia, el número de lugares de unión parece permanecer constante pero no su posición o secuencia exacta. En este segundo caso, parece haber selección estabilizadora para mantener la composición de lugares de unión, pero no siempre el orden u orientación de estos, como se ha mostrado en otros estudios (LUDWIG *et al.* 1998; LUDWIG *et al.* 2000; BERMAN *et al.* 2004). Habitualmente habrá varias secuencias con el potencial de convertirse en lugares de unión activos con un solo cambio nucleotídico (MACARTHUR y BROOKFIELD 2004). CNS y CRM no tienen una relación uno a uno, los CRM son más largos y suelen

contener agrupaciones de CNS, donde cada uno contiene uno o varios lugares de unión. Habría que estudiar más casos para ver cuán general es el patrón observado, ya que la muestra es excesivamente pequeña. En el caso del módulo de la banda 2 de *even-skipped* (*eve* S2E – *even-skipped* Stripe 2 Element) existen múltiples lugares de unión para cada uno de los factores de transcripción involucrados (LUDWIG *et al.* 2000) lo que concuerda con la observación anterior de elevada variación en la secuencia de algunos lugares de unión.

De todos modos, esta dinámica y variabilidad en los lugares de unión no será siempre neutra. Aunque la conservación de las secuencias no codificantes esté demostrando ser un buen método para la detección de los módulos reguladores, no hay que olvidar que la variabilidad fenotípica se debe en muchos casos a la variabilidad de estas regiones (DE MEAUX *et al.* 2005). De manera que el uso de la conservación para la detección de las regiones debería ir seguido de la disección de la variabilidad nucleotídica y su posible relación con la variabilidad en la expresión génica y en último término la variabilidad fenotípica. Debemos centrarnos en los cambios funcionales en vez de en la conservación para entender la dinámica evolutiva a corto plazo de las regiones *cis*-reguladoras.

3.2. EL HOM-C EN DROSOPHILA

3.2.1. Evolución de los genes Hox: Estructura, Regulación y Función

3.2.1.1. *Hox3*: *zen*, *bcd* y *zen2*

Muchos genes del desarrollo están extraordinariamente conservados entre vertebrados e invertebrados, siendo los genes Hox un paradigma de esta conservación. Por otro lado, el desarrollo de nuevas funciones y estructuras no se produce siempre mediante la aparición de nuevos genes, sino que a menudo se basa en la reutilización de genes existentes. Así, a pesar de esta elevada conservación incluso los genes Hox pueden cambiar de forma dramática durante la evolución. Buen ejemplo de ello son los genes *Hox3* y *ftz* que han perdido su función Hox y adquirido nuevas funciones en el desarrollo temprano de *Drosophila* (HUGHES *et al.* 2004).

En *D. melanogaster*, tres genes con homeobox ocupan la posición de *Hox3*, pero ninguno de ellos presenta función o patrón de expresión de tipo homeótico: *bcd* es un morfógeno materno implicado en el establecimiento temprano de la polaridad

anteroposterior (BERLETH *et al.* 1988), *zen* es un gen zigótico que se expresa en el tejido extraembrionario del embrión, la amnioserosa (RUSHLOW *et al.* 1987), y *zen2* es un gen con patrón de expresión similar a *zen* pero sin función conocida ya que la ausencia de este gen no presenta alteraciones en el fenotipo (RUSHLOW *et al.* 1987). Los genes *bcd* y *zen* no son miembros del sistema Hox de determinación del patrón anteroposterior sino que representan genes con homeodominio de evolución rápida (DE ROSA *et al.* 1999).

La clonación de los genes de tipo *Hox3* en *S. gregaria* y *T. castaneum* mostró la homología de este gen con *zen*. Mientras el homeodominio de estos genes era claramente de la familia *Hox3*, tanto motivos de secuencia fuera del homeobox como su patrón de expresión en los tejidos extraembrionarios mostraban que eran homólogos del gen *zen* de *D. melanogaster* (FALCIANI *et al.* 1996). La descripción, en el ácaro *Archegozetes longisetosus*, de un gen homólogo a *zen* pero con patrón de expresión tipo Hox confirmó que en realidad *zen* era el gen *Hox3* modificado (TELFORD y THOMAS 1998). En este organismo *Hox3* presenta un patrón de expresión ampliamente solapado con *pb*, coincidiendo en su límite anterior, lo que llevó a la hipótesis de que deben presentar una cierta redundancia funcional. Esta redundancia entre *pb* y *Hox3* habría sido aprovechada por los insectos para utilizar *Hox3* para una nueva función (FALCIANI *et al.* 1996; TELFORD y THOMAS 1998). Un estudio reciente en el insecto áptero *Thermobia domestica* muestra que *Hox3* todavía retiene el patrón de expresión Hox, pero ha extendido su expresión a los tejidos extraembrionarios (HUGHES *et al.* 2004). En el ancestro de los insectos alados, como *S. gregaria* o *T. castaneum*, este gen perdió su dominio de expresión Hox y reforzó su expresión en los tejidos extraembrionarios (FALCIANI *et al.* 1996). El gen *Hox3* parece haber evolucionado directamente a *zen* durante la diversificación de los insectos (HUGHES *et al.* 2004).

Posteriormente, durante la diversificación de los dípteros, *zen* adquirió un nuevo dominio de expresión que provee un aporte materno en el embrión (STAUBER *et al.* 2002). Una duplicación de este gen en la línea de las moscas ciclorrafas permitió la separación de ambos dominios de expresión, dando lugar a dos genes con funciones diferenciadas: el morfógeno materno *bcd* y el gen zigótico *zen* (STAUBER *et al.* 1999; STAUBER *et al.* 2002).

Una segunda duplicación de *zen* originó el gen *zen2*. Un análisis comparativo de la estructura del ANT-C entre *D. pseudoobscura* y *D. melanogaster* fue incapaz de detectar un homólogo de *zen2* ni en la región estudiada ni en el resto del genoma, de manera que se pensó que *zen2* era una duplicación muy reciente en el linaje de *D.*

melanogaster (RANDAZZO *et al.* 1993). En *D. subobscura* las evidencias también parecían indicar la ausencia de *zen2* (TEROL *et al.* 1995). Pero en el análisis de la secuencia de la región *pb* en *D. buzzatii*, se predijo un gen entre *pb* y *zen* que parecía ser el homólogo de *zen2*. La localización de *zen2* en *D. buzzatii* en la misma posición que en *D. melanogaster*, aunque en orientación invertida, nos llevó a mirar en detalle la secuencia de esta región en *D. pseudoobscura*. En esta especie *zen2* se encuentra también entre *pb* y *zen*, y en la misma orientación que en *D. buzzatii*.

Curiosamente, una búsqueda usando BLASTN con la secuencia del mRNA de *zen2* de *D. melanogaster* (NM_057446) contra las secuencias derivadas de los ensamblajes genómicos (WGS – Whole Genome Shotgun) da un mejor resultado con la secuencia del gen *zen* de *D. pseudoobscura* (49 bp; E=0.002) que con el gen *zen2* de esta misma especie (29 bp; E=6.6). Pero cuando la búsqueda se realiza utilizando las secuencias aminoacídicas (TBLASTN) la proteína *zen2* de *D. melanogaster* identifica a *zen2* de *D. pseudoobscura*, y *zen* de *D. melanogaster* identifica al gen *zen* de *D. pseudoobscura*. Estos resultados, debidos a la rápida divergencia de los genes *zen2*, explicaría los resultados obtenidos por RANDAZZO *et al.* (1993) que hicieron pensar en la ausencia de *zen2*.

La existencia de este gen en *D. pseudoobscura* y *D. buzzatii* indica que es bastante más antiguo de lo supuesto hasta ahora, la duplicación que lo originó fue anterior a la divergencia de los subgéneros *Sophophora* y *Drosophila*. Al pensarse que *zen2* era más reciente, los trabajos de STAUBER *et al.* (1999; 2002) no han intentado identificarlo en otros grupos de dípteros, así que la existencia de este gen en otras especies fuera del género *Drosophila* es todavía una incognita, al igual que la edad de la duplicación.

Así como la primera duplicación de *Hox3-zen* permitió una subdivisión funcional, esta segunda duplicación no parece haber tenido el mismo efecto. De momento *zen2* parece ser un gen superfluo con el mismo patrón de expresión que *zen*, al menos en *D. melanogaster*. De todos modos, el mantenimiento del ORF en las tres especies de *Drosophila* que divergieron hace entre 40 y 62 millones de años (POWELL y DESALLE 1995; RUSSO *et al.* 1995; TAMURA *et al.* 2004) hace pensar que pueda tratarse de un gen funcional, aunque su elevada tasa de substitución podría ser indicativo de un gen en proceso de adquirir una nueva función.

En las tres especies de *Drosophila* analizadas, el gen *zen2* está situado dentro de las regiones reguladoras de *pb*, aunque en orientaciones diferentes. Esta situación se podría interpretar como que la orientación ancestral es la común a *D. pseudoobscura* y *D. buzzatii*, y que el gen se ha invertido en *D. melanogaster*. Pero el hecho que las duplicaciones en tándem suelen dar lugar a dos genes situados en la misma orientación y la posición de *zen2* situado entre el promotor basal de *pb* y sus regiones *cis*-reguladoras, apuntan a una evolución más compleja.

Después de las dos duplicaciones que dieron lugar a *zen*, *bcd* y *zen2* a partir del gen *Hox3* tuvieron lugar dos inversiones (Figura 15). Una primera inversión con un punto de rotura entre el promotor basal de *pb* y sus regiones reguladoras 5' y el otro entre *zen* y *zen2* se habría producido en el ancestro de los subgéneros *Drosophila* y *Sophophora*, dando lugar a la estructura que observamos en *D. pseudoobscura* y *D. buzzatii*. Una segunda inversión, en el linaje de *D. melanogaster*, afectó únicamente al gen *zen2*. Esta segunda inversión habría reutilizado el punto de rotura entre *zen2* y el promotor basal de *pb*, y tendría el segundo punto de rotura entre *zen2* y la región reguladora 5' de *pb*. Este modelo explica, mediante solo dos inversiones, la orientación del gen *zen2* y la posición de las regiones reguladoras de *pb* en las tres especies donde se ha analizado la región.

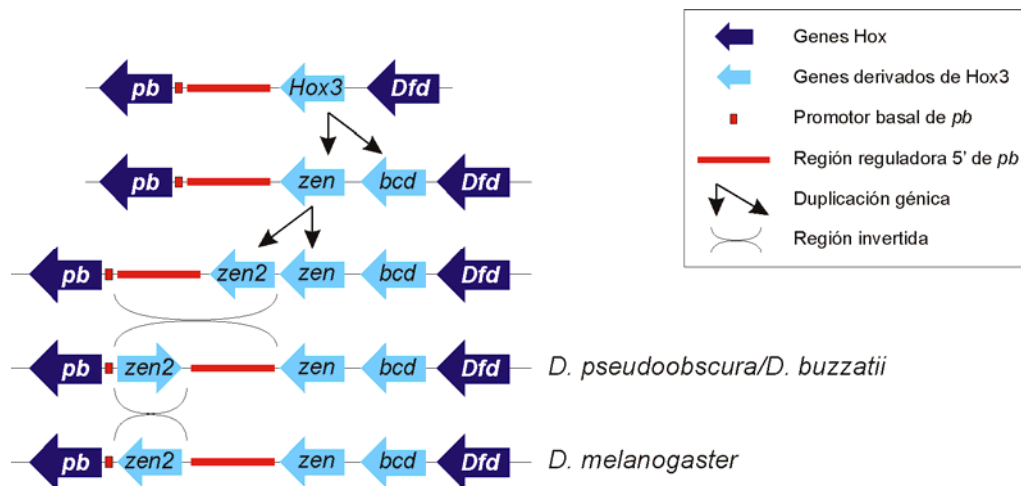


Figura 15. Modelo de evolución del gen *Hox3*.

Figure 15. Model of evolution of the *Hox3* gene.

3.2.1.2. *lab*

labial (*lab*) es el gen más proximal del ANT-C de *D. melanogaster* (WAKIMOTO y KAUFMAN 1981). Corresponde al primer grupo parálogo (PG1). Los mutantes para

este gen presentan letalidad en el embrión, morfológicamente se observa ausencia de estructuras cefálicas pero no transformaciones homeóticas obvias. En el adulto, según observaciones de mutaciones hipomórficas y análisis clonal, se producen ausencia de estructuras ventrales, como los palpos maxilares, y alteraciones en la región dorsal de la cabeza que muestra apariencia de tejido torácico (MERRILL *et al.* 1989). El producto del gen *lab* se acumula en un patrón complejo en tejidos embrionarios e imaginales que podemos agrupar en dos dominios de expresión durante la morfogénesis: el primero en la cabeza, es el gen Hox expresado en una región más anterior, y el segundo en una región del intestino (MLODZIK *et al.* 1988; DIEDERICH *et al.* 1989; DIEDERICH *et al.* 1991). Durante la embriogénesis, *lab* se expresa en células endodérmicas del intestino, en células ectodérmicas del procefalon y el cresta dorsal, y en un pequeño grupo de células progenitoras sensoriales. La expresión imaginal se limita a una región estrecha de la membrana peripodial del disco de ojo-antena (CHOUINARD y KAUFMAN 1991).

Se interpreta que la expresión de *lab* en la cabeza del embrión marca el segmento intercalar, pequeño metámero desprovisto de apéndices y situado entre el lóbulo procefálico y el gnathocephalon, que consiste en los lóbulos mandibulares, maxilares y labiales (DIEDERICH *et al.* 1991; ABZHANOV y KAUFMAN 1999). El gen *lab* es necesario para la correcta involución cefálica. Las células que expresan *lab* en los lóbulos procefálicos parecen estar situadas en puntos clave respecto a los movimientos celulares que conlleva el proceso de involución. Este proceso, que resulta gravemente alterado en los mutante para *lab*, parece ser el responsable de los defectos observados en estos mutantes. La pérdida de las glándulas salivales, la pieza H, y el brazo ventral del aparato cefalofaríngeo parece que no son debidas directamente a la falta de especificación por *lab*, sino una consecuencia secundaria de la disrupción del proceso de involución (MERRILL *et al.* 1989). Dentro del segmento intercalar *lab* también se expresa en el tritocerebro, donde su inactivación produce defectos en la elongación de axones (HIRTH *et al.* 1998). La acción de *lab* es necesaria para la diferenciación neuronal en el cerebro en desarrollo (HIRTH *et al.* 1998).

La expresión de *lab* en el intestino coincide con las células del cobre. *lab* tienen un papel no solo en la determinación y diferenciación de estas células, sino también en el mantenimiento de su estado diferenciado. Estas células parece que absorben iones metálicos del lumen intestinal (HOPPLER y BIENZ 1994) y que están involucradas en la acidificación del intestino. La conservación de la expresión del gen *lab* en el intestino

del embrión en desarrollo en las diferentes especies analizadas muestra que esta función fue adquirida por *lab* por lo menos antes de la diversificación del género *Drosophila*.

La comparación de la secuencia codificante de *lab* realizada entre tres especies de *Drosophila* (ver Artículo 1) muestra una elevada tasa de substitución nucleotídica. Se contrastó la posibilidad de una variación en la tasa de substitución nucleotídica como consecuencia del cambio de posición. Los resultados mostraron una tasa homogénea a lo largo de la divergencia del género, aunque heterogénea a lo largo de la región codificante. La región más conservada corresponde al homeobox, pudiéndose observar la presencia de otras pequeñas regiones que se encuentran también conservadas, a nivel de proteína, en especies alejadas como *T. castaneum* o incluso en Humanos. Esta conservación a nivel de dominio proteicos explica como minigenes de *Hoxb-1* de pollo, expresados bajo el control de regiones reguladoras de *lab* de *D. melanogaster*, eran capaces de rescatar mutantes nulos para *lab* (LUTZ *et al.* 1996). Aunque para algunas funciones del gen los cambios en la proteína pueden ser significativos, para otras funciones es más importante el momento y lugar de expresión, es decir, las regiones *cis*-reguladoras.

Las regiones reguladoras de *lab* fueron estudiadas por CHOUINARD y KAUFMAN (1991). Como hemos visto (Artículo 2) estas regiones están muy conservadas entre especies de *Drosophila*. Las regiones reguladoras todavía no identificadas probablemente están localizadas en las regiones conservadas que no han sido probadas. Dos de las regiones analizadas por CHOUINARD y KAUFMAN (1991) que eran aparentemente no funcionales presentan algunas zonas altamente conservadas, es probable que estas regiones sean funcionales pero que formen parte de un elemento regulador de mayor tamaño, no siendo capaces de inducir expresión por si solas. Por otro lado hay que destacar que, especialmente en la región 5', las regiones reguladoras identificadas son solapantes. Cabe destacar la elevada superposición y la aparente falta de orden de las regiones identificadas respecto a su dominio de expresión en el embrión o discos imaginales.

3.2.1.3. *pb*

proboscipedia (pb) es el segundo gen del HOM-C de *Drosophila*. El primer mutante de *pb* fue descrito en 1933 por Bridges y Dobzhansky. Las mutaciones nulas de *pb* producen transformación homeótica específica del adulto, donde los palpos labiales

se transforman en patas protorácicas. En mutantes de pérdida parcial de función, los palpos labiales resultan transformados en aristas de antena (KAUFMAN 1978). A diferencia de las transformaciones en el adulto, la ausencia de *pb* en el embrión no presentan alteraciones detectables (PULTZ *et al.* 1988). La acción de *pb* es por tanto única entre los genes homeóticos, ya que parece ser innecesario en el embrión.

A pesar de ser necesario solo en el adulto, *pb* se expresa en el embrión además de en los estadios de larva y pupa. En el embrión se expresa en los lóbulos labial y maxilar del gnatocéfalón que dan lugar a estructuras específicas de la cabeza de larva y adulto, en el mesodermo ventral del segmento mandibular, y posteriormente se expresa también en el sistema nervioso central (PULTZ *et al.* 1988). La expresión de *pb* se detecta también en los discos labiales, que dan lugar al labrum del adulto (RANDAZZO *et al.* 1991). Minigenes de *pb* expresados en los discos labiales son capaces de recuperar la transformación homeótica de los mutantes para *pb*, pero no se expresan en el sistema nervioso central, lo que da lugar a moscas con deficiencias a nivel de comportamiento (RANDAZZO *et al.* 1991).

El significado de la expresión embrionaria de *pb* no está claro, ya que la delección del locus tiene consecuencias solo en adultos. KAUFMAN *et al.* (1990) propone varias hipótesis para explicar la aparente ausencia de función de la expresión de *pb* en el embrión. Una posibilidad es que los defectos del embrión sean sutiles y hayan escapado a la detección. Otra posibilidad es que la expresión temprana de *pb* represente un vestigio evolutivo del patrón de expresión en insectos más primitivos donde el desarrollo de la cabeza no está tan modificado como en los dípteros superiores. Alternativamente, esta acumulación temprana podría representar el mecanismo por el cual se asegura la expresión de *pb* posteriormente en los discos imaginales. Los dos segmentos en que se expresa *pb* en el ectodermo son los lugares lógicos de las células progenitoras de las estructuras afectadas del adulto.

Otra peculiaridad del gen *pb* es su patrón de expresión en el sistema nervioso central de la mosca. En esta estructura *pb* se expresa en algunas células más anteriores que *lab*, lo que supone una excepción a la regla de colinearidad. La colinearidad sí se mantiene en el ectodermo, donde la expresión de *pb* es posterior a *lab* (HIRTH *et al.* 1998).

En cuanto al análisis comparativo de expresión, el único cambio que hemos detectado concierne al gen *pb* en *D. virilis*. En esta especie, el gen *pb* presenta un

dominio extra de expresión no detectado en el resto de especies analizadas. Este dominio se sitúa en el mesodermo del embrión en la región ventral del segmento maxilar. El estudio de expresión de *pb* en embriones incluido en el Artículo 2 se ha completado posteriormente con la especie *D. pseudoobscura*. Al analizar la secuencia alrededor de la rotura se observa que en *D. melanogaster* hay una pequeña inversión de dos genes *Ccp* con un punto de rotura muy cerca del extremo 3' de *pb*, en la misma posición que la rotura *lab-pb*. Para poder descartar que el dominio de expresión presente sólo en *D. virilis* no se haya perdido en los otros dos linajes debido a las roturas mencionadas, se estudió la expresión de *pb* en embriones de *D. pseudoobscura*, que parece tener la disposición ancestral en esta región de la secuencia. Los resultados muestran (Figura 16) que *D. pseudoobscura* presenta, en esta fase del desarrollo, el mismo patrón de expresión que *D. melanogaster*. De manera que podemos descartar esta hipótesis, y confirmar los resultados mostrados en el segundo artículo.

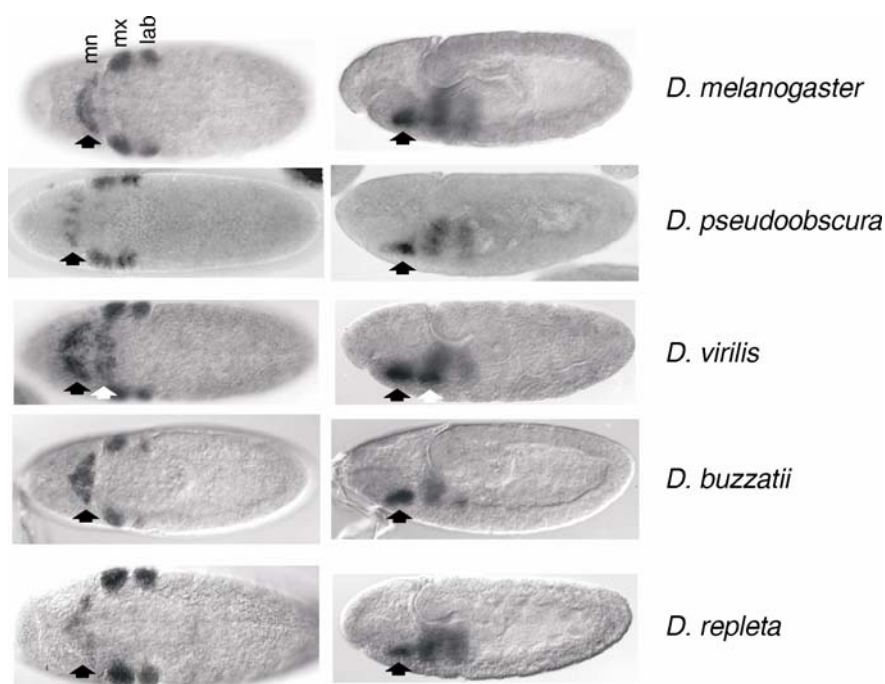


Figura 16. Expresión del gen *pb* en embriones (estadio 11) de cinco especies de *Drosophila*. La flecha negra muestra la expresión en el mesodermo ventral del segmento mandibular, la flecha blanca señala la expresión en el mesodermo ventral del segmento maxilar. El segmento mandibular (mn), maxilar (mx) y labial (lab) se muestran en el primer recuadro.

Figure 16. Expression of *pb* in embryos (stage 11) from five *Drosophila* species. The black arrow shows the expression in the ventral mesoderm of the mandibular segment, the white arrow shows the expression in the ventral mesoderm of the maxillary segment. The mandibular (mn), maxilar (mx) and labial (lab) segments are shown in the first embryo.

Así este cambio no parece relacionado con la rotura *lab-pb*. La expresión de *pb* es prescindible en el embrión, pero la ausencia de función o alteraciones detectables se refiere exclusivamente a la cutícula. No se sabe qué función realiza *pb* en el mesodermo ventral, lo cual dificulta la interpretación de las diferencias observadas.

Aunque es más parsimonioso pensar en una ganancia de un dominio de expresión en el linaje de *D. virilis* que en dos pérdidas en los otros dos linajes, algunas evidencias apoyan la segunda opción. Estudios en otros artrópodos apuntan hacia una reducción en el dominio de expresión de *pb* en la evolución hacia *Drosophila* (HUGHES y KAUFMAN 2002). Por otro lado, se considera que *D. virilis* retiene muchos caracteres ancestrales dentro de los drosophilidos (CLARK *et al.* 2003). Podría por tanto ser que el dominio de expresión de *pb* en el mesodermo del segmento maxilar de *D. virilis* sea en realidad la retención de un carácter ancestral. Representando un vestigio de un dominio de expresión más extenso que se ha perdido de forma independiente en los linajes de *D. buzzatii/D. repleta* y *D. melanogaster/D. pseudoobscura*. El hecho de que en los dos linajes se haya perdido el mismo dominio podría indicar que el dominio retenido es funcional o podría ser una cuestión casual.

Un análisis detallado de la secuencia de *pb* en *D. virilis*, especialmente de sus regiones *cis*-reguladoras, podría quizás ayudar a detectar el elemento *cis*-regulador responsable de este dominio extra. De todos modos, aún es difícil relacionar la secuencia con la función en las regiones *cis*-reguladoras.

3.2.1.4. *abdA*

Cuando Edward Lewis (1978) describió el complejo Bithorax, indicó la presencia de al menos ocho genes, responsable cada uno de dar identidad a un segmento. Posteriormente, el aislamiento y caracterización de mutaciones letales en heterocigosis con deficiencias del BX-C mostraron la existencia de solo tres grupos de complementación (SÁNCHEZ-HERRERO *et al.* 1985; TIONG *et al.* 1985). Estudios moleculares confirmaron la existencia de solo tres genes codificantes: *Ubx*, *abdA* y *AbdB* (CELNIKER *et al.* 1989; KARCH *et al.* 1990). Los “genes” identificados por LEWIS (1978) eran en realidad regiones *cis*-reguladoras de estos tres genes. Cada una de estas regiones *cis*-reguladoras se encarga de activar y mantener la expresión de uno de estos genes en un segmento (o parasegmento) determinado, y están situadas de forma colineal.

El gen *abdominal A* (*abdA*) es el penúltimo gen del HOM-C de *Drosophila*, ortólogo a los genes centrales de otros organismos. Las mutaciones en *abdA* son letales, estos embriones presentan los parasegmentos 7 a 9 transformados en copias del parasegmento 6, y los parasegmentos 10 a 13 también muestran transformación, aunque más leve, a parasegmento 9. La expresión de este gen está regulada por las regiones *cis*-reguladoras *iab*. Mutaciones localizadas en estas regiones *iab* alteran la función del gen únicamente en el segmento en el que guían la expresión de *abdA*. Los números de las regiones *iab* se refieren a los segmentos abdominales donde se observan las alteraciones en los mutantes, aunque en realidad estas regiones dirigen la expresión en parasegmentos. La región *iab2* se refiere al segundo segmento abdominal, que corresponde al parasegmento 7; la región *iab3* al tercer segmento abdominal, parasegmento 8, etc.

De los genes Hox estudiados en este trabajo, *abdA* parece ser el que presenta más constricciones funcionales. La región *abdA* no presenta ninguna microinversión, a diferencia de *lab* y *pb*, ni inserción de elementos móviles, los cuales son muy abundantes en el gen *lab* situado a tan solo 75 kb (Artículo 2 Figura 2). Además los elementos reguladores de *abdA*, a diferencia de *lab* y *pb*, presentan colinealidad.

3.2.2. Evolución de la Estructura del HOM-C en *Drosophila*

3.2.2.1. Estructura del HOM-C

A pesar de que la agrupación y colinealidad de los genes Hox son un paradigma de la Evo-Devo, la poca información disponible sobre la organización de los genes Hox en *Drosophila* (VON ALLMEN *et al.* 1996), ya indicaba que la organización descrita en *D. melanogaster* no era válida para todo el género. Por otro lado la elevada tasa de reordenaciones en este género (GONZÁLEZ *et al.* 2002) hacía probable la existencia de más alteraciones del complejo, hipótesis que se vio confirmada con el análisis de su estructura en especies del grupo *repleta* de *Drosophila* (RANZ, CÁCERES y RUIZ comunicación personal) (Figura 14).

Los proyectos de secuenciación de varias especies del género *Drosophila* van a aportar más información sobre la organización del HOM-C. Las especies más informativas serán probablemente *D. willistoni*, perteneciente al subgénero *Sophophora*, y *D. grimshawi*, única especie del subgénero *Idiomyia* en proceso de secuenciación. Desgraciadamente aún no disponemos de ningún ensamblaje de la secuencia de estas

especies. De momento, además de los datos descritos en la introducción (apartado 1.7.4) disponemos de la secuencia de las regiones *lab-abdA* y *pb* de *D. buzzatii* (Artículo 2), del trabajo de LEWIS *et al.* (2003), donde analizan la rotura *Antp-Ubx*, y de las secuencias genómicas de *D. pseudoobscura* (RICHARDS *et al.* 2005), *D. ananassae*, *D. erecta*, *D. simulans*, *D. yakuba*, *D. virilis* y *D. mojavensis* (<http://rana.lbl.gov/drosophila/multipleflies.html>). He realizado un breve análisis de estas nuevas secuencias genómicas para ver la organización de los genes Hox. Estos datos no indican la existencia de más roturas, pero permiten delimitar cuando se produjeron algunas de las inserciones y reorganizaciones del complejo.

Actualmente se conocen tres organizaciones genómicas diferentes de los genes Hox en el género *Drosophila* (Figura 14). De estas agrupaciones se deduce que el ancestro del género *Drosophila* presentaba un único complejo completo como el que se observa en *Anopheles* (VON ALLMEN *et al.* 1996). Las diferentes roturas del complejo se produjeron después de la separación de los subgéneros *Drosophila* y *Sophophora*.

La organización en dos complejos, el ANT-C y el BX-C, descrita en *D. melanogaster* parece ser común a los diferentes linajes del subgénero *Sophophora* (LEWIS *et al.* 2003). Esta organización es resultado de una inversión que rompió el complejo entre los genes *Antp* y *Ubx*. Habrá que esperar a la secuenciación de *D. willistoni* para discriminar si esta rotura se produjo en el ancestro de todo el subgénero o en el ancestro común a los grupos *obscura* y *melanogaster*.

La segunda organización es la descrita en *D. virilis* y presente también en otras especies como *D. immigrans*. En estas especies, el complejo también se encuentra dividido en dos fragmentos, el primero contiene desde *lab* hasta *Ubx* (es decir ANT-C más *Ubx*) y el segundo los dos genes abdominal (*abdA* y *AbdB*) (VON ALLMEN *et al.* 1996, RANZ *et al.* 2001, RANZ, CÁCERES y RUIZ comunicación personal). La rotura entre *Ubx* y *abdA* está presente en todas las especies estudiadas del subgénero *Drosophila*. Falta saber si se produjo en el ancestro de este subgénero o si está también presente en el subgénero *Idiomyia*.

La tercera organización, descrita en *D. buzzatii* y común a todo el grupo repleta y a *D. canalinea*, presenta dos roturas del complejo. Estas especies presentan la rotura entre *Ubx* y *abdA* común a todo el subgénero *Drosophila* y además una segunda rotura entre *lab* y *pb*. A pesar de esta segunda rotura, los genes siguen agrupados a nivel

citológico en solo dos regiones genómicas, ya que el gen *lab* está a tan solo 75 kb de *abdA*.

3.2.2.2. Reconstrucción de la Evolución de la Estructura del HOM-C

En dípteros basales, como *Anopheles gambiae*, ocho genes Hox más *Hox3* y *ftz* están agrupados en un único complejo (POWERS *et al.* 2000). Antes de la radiación del género *Drosophila*, ocurrieron cuatro transposiciones: el complejo de genes *Ccp* se insertó entre *lab* y *pb* (Artículo 2), el gen *Ama* entre *bcd* y *Dfd*, y el gen *CG31217* entre *Antp* y *Ubx* (LEWIS *et al.* 2003) y *Glut3* entre *Ubx* y *abdA*. Durante este mismo período, *Hox3-zen* dio lugar, a través de dos duplicaciones, a los genes *bcd*, *zen* y *zen2*. Después *zen2* sufrió una inversión, junto con la región reguladora 5' de *pb* (ver apartado 3.2.1.1). Esta organización parece ser la ancestral a todo el género *Drosophila* (Figura 17, 1).

En el linaje que condujo al subgénero *Drosophila*, después de la separación del subgénero *Sophophora*, tuvo lugar una inversión con un punto de rotura entre los genes *abdA* y *Glut3* y el otro entre *CG31363* y un gen *X* (Artículo 2). En la secuencia de *D. virilis* el gen situado en la región 3' de *Ubx* es el *CG6073*, este debería ser el gen *X* siempre que no haya habido más reorganizaciones en esta región en este linaje (Figura 17, 2). Esta estructura del HOM-C es la que encontramos actualmente en especies del subgénero *Drosophila* no pertenecientes al grupo repleta, como por ejemplo *D. virilis*.

Una segunda inversión, en el ancestro del grupo repleta, rompió el HOM-C entre los genes *lab* y *pb*. Esta inversión, que situó a *lab* cerca de *abdA*, tuvo un punto de rotura entre *pb* y el complejo de genes *Ccp* y el segundo entre *CG31363* y *CG17836* (Figura 17, 3). Estos dos genes no son adyacentes en el genoma de *D. melanogaster*, *D. pseudoobscura* ni *D. virilis*, pero inferimos que lo eran en el ancestro del subgénero *Drosophila*.

En el linaje que condujo al subgénero *Sophophora*, el HOM-C se rompió, mediante una inversión, entre los genes *Ubx* y *Antp*, dando lugar a los complejos ANT-C y BX-C. Los puntos de rotura se situaron, concretamente, uno entre *Ubx* y *CG31217* y el otro entre *Sodh* y *CCTg* (Figura 17, 4) (LEWIS *et al.* 2003). Durante este mismo período, una microinversión cambió la orientación del gen *ftz* (Figura 17, 5). Esta estructura parece ser la actual en la especie *D. pseudoobscura*. Posteriormente, en el linaje de *D. melanogaster* ocurrieron otras dos microinversiones. El gen *zen2* volvió a

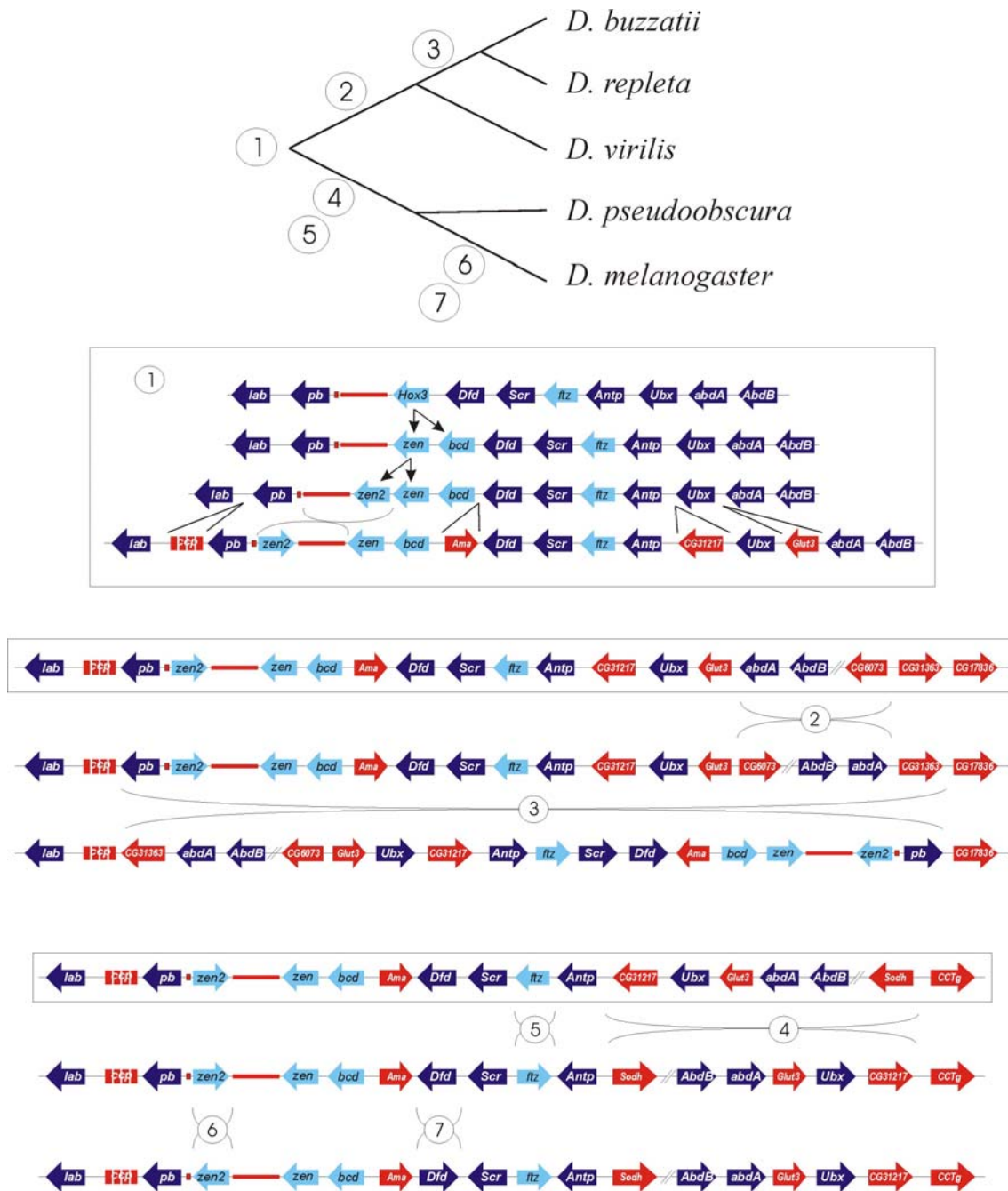


Figura 17. Modelo de evolución del HOM-C en *Drosophila*. Arriba, filogenia de las especies de *Drosophila* estudiadas en este trabajo. Se muestra el lugar de la filogenia en que tuvieron lugar los diferentes eventos. El recuadro con el número (1) muestra los eventos sucedidos antes de la diversificación del género. Los otros dos recuadros muestran el estado del complejo antes de la diversificación, punto de partida donde se produjeron el resto de reorganizaciones. Símbolos y colores igual que en la Figura 15.

Figure 17. Model of the evolution of the HOM-C in *Drosophila*. Up, phylogeny of the *Drosophila* species studied in this thesis. The numbers show where in the phylogeny each event took place. The rectangle with number (1) shows those events that happened before the diversification of the *Drosophila* genus. The other rectangles show the structure of the complex before the diversification, starting point where the rearrangements took place. Symbol and colours as in Figure 15.

invertirse, quedando nuevamente en la misma orientación que la mayoría de genes del complejo (Figura 17, 6). La otra microinversión invirtió la orientación del gen *Dfd* (Figura 17, 7).

Globalmente, en el ancestro del género *Drosophila* se produjeron dos transposiciones (genes *Ccp* y *CG31217*) y una microinversión (*zen2*). Dentro del género *Drosophila* se han producido, en los diferentes linajes, un total de tres inversiones (roturas *Ubx-abdA*, *Antp-Ubx* y *lab-pb*) y otras tres microinversiones (*ftz*, *zen2* y *Dfd*). De todos modos, esta descripción es incompleta, ya que falta determinar cuando se produjeron al menos otra transposición más: la agrupación de tRNA de la lisina (*lys-tRNA*) situados entre *Dfd* y *Scr* en el ANT-C de *D. melanogaster*.

Todos los genes Hox, excepto el gen *Dfd* en *D. melanogaster*, presentan la misma orientación 5' → 3'. Sus regiones reguladoras están habitualmente situadas en la región 5' del gen y en sus intrones. Todas la transposiciones y roturas entre genes Hox se han producido cerca del extremo 3' de un gen y lejos del 5' del otro gen, respetando así las regiones reguladoras de ambos genes. De este modo, las reordenaciones no han alterado ninguna región reguladora conocida de estos genes Hox, lo que explicaría la ausencia de efecto sobre su expresión.

3.2.3. Mecanismo de las Reordenaciones

Una vez determinado que las roturas del complejo de genes Hox han sido debidas a inversiones paracéntricas, se ha analizado en detalle la secuencia de las regiones de los dos puntos de rotura de la inversión *lab-pb* y de uno de la *Ubx-abdA* en *D. buzzatii* (ver Artículo 2) para intentar determinar el mecanismo que las produjo. En dos de estos tres puntos de rotura (*lab-CG31363* y *CG31363-abdA*) existen inserciones de elementos transponibles tipo IsBu. Estos elementos presentan un tamaño pequeño (alrededor de una kilobase), producen duplicaciones de tan solo 2 pb al insertarse y tienden a encontrarse agrupados o incrustados uno dentro de otro. Estas características hacen difícil la identificación de elementos quiméricos que serían el resultado de la recombinación ectópica responsable de una inversión (CACERES *et al.* 1999; CACERES *et al.* 2001; CASALS *et al.* 2003). Aunque no lo podemos descartar completamente, parece poco probable que estos elementos hayan estado relacionados con el origen de estas inversiones.

3.2.4. Causas y Consecuencias de las Roturas del HOM-C

¿Cuál es el significado funcional de la agrupación de los genes Hox? Varias razones se han postulado para explicar la sorprendente conservación de los complejos Hox a lo largo de la filogenia. La similitud de los complejos no se debe a una convergencia evolutiva sino a la conservación de un estado ancestral, pero para que esta estructura se haya mantenido durante más de 600 millones de años de evolución tienen que haber constricciones funcionales actuando sobre el complejo como unidad (MCGINNIS y KRUMLAUF 1992). ¿Cuáles son estas constricciones funcionales? Dos hipótesis podrían explicar la existencia de los complejos Hox: (A) estos genes no pueden funcionar o no funcionan correctamente fuera del dominio del complejo, quizás por requerimientos de la estructura cromatínica o por efectos de posición; (B) la yuxtaposición de los genes miembros y sus complejas regiones *cis*-reguladoras hace la separación por procesos genéticos al azar improbable o incluso imposible (RANDAZZO *et al.* 1993).

La existencia de elementos reguladores compartidos entre varios genes explica en algunos casos la conservación de los complejos génicos. En el caso de los genes Hox, inversiones que rompen el ANT-C de *D. melanogaster* dentro de un gen no afectan a otros genes del complejo, lo que sugiere que estos genes no poseen elementos reguladores compartidos (ABBOTT y KAUFMAN 1986; PULTZ *et al.* 1988; DIEDERICH *et al.* 1989; AVEROF *et al.* 1996). Por otro lado, en la región anterior del HOM-C de *Drosophila*, encontramos dos ejemplos de regiones *cis*-reguladoras intercaladas: *Scr-ftz* y *pb-zen2*. En el primer caso *ftz* se encuentra entre las regiones reguladoras 5' de *Scr* (MAIER *et al.* 1990). De forma similar, *zen2* está situado entre el promotor basal de *pb* y sus regiones reguladoras 5'. En ambos casos el segundo gen, que ha perdido su función Hox, se encuentra entre el primer gen (*Scr* o *pb*) y parte de sus regiones *cis*-reguladoras 5'. Esta organización hace inviable la separación de los dos genes mediante inversión, pero parece indicar que el gen no Hox podría cambiar de posición mediante transposición, aunque la tasa de transposición en el género *Drosophila* es muy baja (RANZ *et al.* 2003).

En cuanto a los genes posteriores, según KARCH *et al.* (1985) el BX-C de *D. melanogaster* está formado por dos unidades separables entre ellas pero no dentro de ellas: la región Ultrabithorax y la región Abdominal. La región Ultrabithorax contiene el gen *Ubx* y su región reguladora que incluye el RNA no codificante *bxd*, el cual presenta

procesamiento alternativo. Esta región se puede separar de la región Abdominal sin efecto fenotípico (STRUHL 1984), y está separada de forma natural en las especies del subgénero *Drosophila* (rotura *Ubx-abdA*) (RANZ *et al.* 2001, VON ALLMEN *et al.* 1996 y Artículo 1). La región Abdominal contiene los genes *abdA* y *AbdB* separados por 100 kb que contienen las regiones *cis*-reguladoras *iab2* a *iab9*. Las regiones *iab*, separadas entre ellas por elementos aisladores, presentan colinealidad entre su orden en el cromosoma y los parasegmentos donde determinan la expresión de *abdA* o *AbdB* (SANCHEZ-HERRERO 1991). Algunas de estas regiones *iab*, concretamente *iab5*, parecen actuar tanto sobre *abdA* como sobre *AbdB*, lo que haría inviable la separación de esta región de alguno de los dos genes. TIONG *et al.* (1987) intentaron demostrar que la unidad de esta región Abdominal no era esencial para su función. Utilizaron una línea que presentaba una translocación de las regiones *Ubx* y *abdA*, una de las roturas se sitúa aproximadamente entre *iab6* e *iab7*; y otra línea que presentaba una delección de las regiones *Ubx* y *abdA*, en este caso una de las roturas se sitúa aproximadamente entre *iab4* e *iab5*. Mostraron que estas dos deficiencias se complementan dando lugar a individuos fenotípicamente salvajes, mostrando que las regiones *abdA* y *AbdB* no necesitan estar en *cis* para su correcto funcionamiento. Pero en realidad no se trata de una rotura “limpia”, sino que hay una región duplicada. Las regiones *iab5* e *iab6* están presentes en *cis* tanto al lado de *abdA* como al lado de *AbdB*. Este trabajo más otros indicios de que *iab5* regula tanto la transcripción de *abdA* como de *AbdB* (DUNCAN 1987; CELNIKER *et al.* 1990) parecen indicar que estas dos regiones podrían separarse físicamente sin efecto funcional siempre y cuando la región *iab5* se duplicara y estuviera presente al lado de ambos genes.

Otra posible explicación para la selección a lo largo de la evolución en contra de la dispersión de los genes homeóticos serían las consecuencias detrimentales de los efectos de posición. Según esta teoría los genes homeóticos podrían estar residiendo en el complejo como prisioneros cuya escapada causaría, en muchos casos, efectos deletéreos para el organismo (RANDAZZO *et al.* 1991). Inversiones dentro del complejo podrían influir en la transcripción de los genes homeóticos acercando secuencias portadoras de nuevos elementos reguladores o modificando la estructura global de la cromatina. Mientras algunos genes han mostrado ausencia de efecto cuando se expresan de forma ectópica en *Drosophila* (BONNER *et al.* 1984; BASLER y HAFEN 1989), la expresión ectópica de los genes homeóticos *Antp*, *Ubx*, *abdA* o *pb* tienen efectos

deletéreos graves sobre el organismo (GONZALEZ-REYES *et al.* 1990; CRIBBS *et al.* 1992; SÁNCHEZ-HERRERO *et al.* 1994).

Los genes de los grupos trithorax y polycomb mantienen la activación y represión, respectivamente, de los genes Hox. Estos genes parecen realizar su función actuando sobre la estructura de la cromatina manteniendo la región accesible o inaccesible a la maquinaria transcripcional. Este efecto sobre la cromatina podría ser el responsable del mantenimiento de la agrupación de los genes Hox en *Drosophila*. CHIANG *et al.* (1995) mostraron que los elementos que mantienen la represión de los genes Hox por acción de polycomb actúan de forma cooperativa. Un reordenamiento de estas regiones podría prevenir esta acción cooperativa, disminuyendo la acción represiva de estos elementos o haciendo al individuo más sensible a alteraciones en el sistema polycomb. El mantenimiento de la acción cooperativa de estos elementos represores podría ser una de los factores que mantienen la integridad de los complejos. El complejo de genes Iroquois está también regulado por los genes polycomb y trithorax a través de la estructura cromatínica, de forma similar a los genes del HOM-C (NETTER *et al.* 1998).

Por otro lado, existen evidencias de la necesidad de apareamiento de los cromosomas homólogos para la correcta regulación de algunos genes, así como la existencia de regiones de unión a la matriz (SARs – Scaffolding Attachment Regions, o MARs – Matrix Attachment Regions) en diferentes posiciones del complejo, lo que sugiere la importancia de la arquitectura cromosómica en el correcto funcionamiento de estas regiones (RANDAZZO *et al.* 1993).

Hay que destacar que a pesar de la importancia del complejo de genes Hox, en *Drosophila* es posible romper este complejo y tener una mosca viable. Trabajos de roturas inducidas o rescates por minigenes parecían indicar la posibilidad de estas roturas en condiciones de laboratorio, pero cuestionaban si serían capaces de sobrevivir a las presiones selectivas de la naturaleza (RANDAZZO *et al.* 1991). Hay que recalcar que las roturas del HOM-C descritas en el apartado 3.2.2.1 son naturales. No se trata de roturas experimentales cuyos portadores sobreviven en el laboratorio sin aparentes malformaciones, sino que se trata de inversiones fijadas en grupos de especies con elevado éxito evolutivo. Incluso la rotura más reciente, entre *lab* y *pb*, está fijada en todo el grupo repleta, uno de los grupos más exitosos con más de 100 especies.

3.2.5. ANT-C *versus* BX-C

Tradicionalmente los estudios de genes Hox en *Drosophila* se han centrado o en un gen en concreto o en uno de los complejos, lo que hace que los estudios existentes sobre los genes del ANT-C y del BX-C sean bastante diferentes y que no existan trabajos globales. Quizás por este motivo, algunas diferencias entre los genes anteriores y posteriores del complejo pueden haber pasado desapercibidas.

En el análisis de la conservación de las regiones no codificantes (ver Artículo 2), se han detectado cuatro microinversiones en las regiones reguladoras de *lab* y *pb*, en cambio no se ha encontrado ninguna inversión semejante en *abdA*. Es en la región de los genes abdominal donde las regiones reguladoras de cada parasegmento mantienen la colinealidad. Es posible que en estas regiones la colinealidad siga siendo un requisito funcional. El efecto de una mutación en la que *pbx* se sitúa aguas abajo en vez de aguas arriba de *Ubx* sugiere que el orden de las regiones reguladoras del complejo tiene significado funcional (DUNCAN 1987). Además las regiones *iab* se transcriben, también de forma ordenada, para activar la transcripción de los genes abdominal en cada uno de los parasegmentos. La transcripción de estas regiones reguladoras activa los genes codificantes a través de la apertura de la cromatina que efectúa la maquinaria transcripcional (SÁNCHEZ-HERRERO y AKAM 1989; SANCHEZ-HERRERO 1991; BAE *et al.* 2002; DREWELL *et al.* 2002).

También hemos observado diferencias en la presencia de inserciones de TE entre *lab* y *abdA*, a pesar de que ambos genes están muy próximos en el genoma de *D. buzzatii*, a tan solo 75 kb. El gen *lab* presenta un elevado número de inserciones de elementos tipo ISBu, tanto en la región 5' como en los intrones. En cambio los otros dos genes Hox no presentan casi inserciones. Teniendo en cuenta la mayor abundancia de TE en regiones pericentroméricas (CASALS *et al.* 2005) destaca la ausencia de TE en la región *abdA*, a pesar de estar muy próximo al centrómero y adyacente a *lab*.

En *Drosophila* los genes Hox posteriores mantienen la colinealidad entre genes y entre regiones reguladoras de estos genes (especialmente los genes abdominal), pero en la región anterior del complejo la situación es diferente. La cabeza de *Drosophila* es una estructura compleja con un proceso de desarrollo altamente derivado. Esta puede ser la causa de que algunos de los genes Hox que se expresan en esta región anterior hayan

perdido la colinealidad en los dominios de expresión. Por ejemplo *pb* tiene una expresión más anterior que *lab* en el sistema nervioso central (HIRTH *et al.* 1998).

Pero no solo están perdiendo la colinealidad, sino que están perdiendo su función Hox. Aparte de *zen* y *ftz*, que han perdido completamente esta función, los genes *lab*, *pb* y *Dfd* solo presentan fenotipo Hox en el adulto, pero no en el embrión. Los complejos patrones de expresión de los genes Hox en estadios tardíos del desarrollo sugieren que estos genes son utilizados en otros procesos de desarrollo no relacionados con la diferenciación axial (AKAM 1989). Efectivamente, estos genes están adquiriendo nuevas funciones. El gen *lab* realiza tres funciones diferentes: (1) función Hox, especifica la identidad de la región posterior de la cabeza del imago; (2) involución cefálica, la ausencia de expresión de *lab* en el embrión produce alteraciones en el proceso de involución cefálica, y (3) diferenciación de un tipo celular, la expresión de *lab* en el intestino interviene en la diferenciación de las células del cobre y en el mantenimiento de su estado diferenciado. El gen *pb* realiza también dos funciones diferentes: (1) función Hox, especifica la identidad de los palpos maxilares y labiales, y (2) sistema nervioso central de la larva, al final del desarrollo embrionario se expresa en grupos de células a lo largo de todo el sistema nervioso central de la larva, su ausencia produce alteraciones de comportamiento.

Estos dos genes podrían representar una situación preliminar que podría llevar a un caso similar a *Hox3*, donde se ha ido perdiendo la función Hox y adquiriendo nuevas funciones que, como en el caso de *zen-bcd*, pueden llegar a subdividirse mediante duplicación génica. AVEROF *et al.* (1996) sugieren que la diversificación funcional de los genes Hox empieza con la adquisición de múltiples y distintos elementos *cis*-reguladores responsables de diferentes aspectos de la expresión génica.

Globalmente queda patente el diferente comportamiento de los genes Hox anteriores y posteriores de *Drosophila*. Mientras los genes anteriores del complejo parecen estar en proceso de perder su función Hox y adquirir nuevas funciones, los genes posteriores parecen mantener la colinealidad a nivel de genes y de regiones reguladoras.

3.3. CONSERVACIÓN Y DESINTEGRACIÓN DEL HOM-C EN LOS METAZOOS

3.3.1. *Drosophila* versus Vertebrados

La presencia de genes Hox agrupados en un complejo se ha observado en el genoma de muchos metazoos, y cuando se ha probado, se ha mostrado que estos genes especifican la diferenciación en diferentes posiciones a lo largo del eje anteroposterior del organismo. De todos modos, a pesar de esta conservación, la estructura del complejo de genes Hox difiere considerablemente en cada uno de los grupos en los que se ha examinado. Varias características diferencian el HOM-C de *Drosophila* del de vertebrados (FERRIER y MINGUILLON 2003; SANTINI *et al.* 2003; WAGNER *et al.* 2003).

Los complejos de vertebrados son compactos (100-200 kb) y las unidades de transcripción cortas, aunque múltiples complejos están presentes en el genoma (MCGINNIS y KRUMLAUF 1992). El complejo de *Drosophila* es extenso (unas 700 kb), con intrones y regiones reguladoras intergénicas excepcionalmente largas comparado con lo compacto que es el resto del genoma. Las regiones reguladoras parecen estar organizadas en dominios discretos, específicos de un segmento, y ampliamente autónomos (AVEROF *et al.* 1996).

En *Drosophila* hay multitud de inserciones de elementos transponibles en estas regiones, mientras que estos son prácticamente inexistentes en los complejos de vertebrados (LANDER *et al.* 2001). También hay inserciones de genes no Hox entre los genes Hox (además de los que han perdido su función Hox) y duplicaciones en tandem dentro del complejo. *Drosophila* permite la inversión de genes Hox, como *Dfd*, y de genes no Hox, como *zen2*. Además el HOM-C de *Drosophila* se ha roto en diferentes puntos sin aparente efecto en su función. Estas observaciones sugieren una evolución dinámica en *Drosophila* que contrasta con la económica estructura y la conservación observada en vertebrados.

Estas diferencias sugieren que los complejos de vertebrados y *Drosophila* están actualmente sometidos a constricciones estructurales y reguladoras muy diferentes. La rotura del complejo de genes Hox en *Drosophila* parece el proceso final de desintegración, hecho posible sólo después del resto de cambios observados, los cuales han ido esponjando el complejo.

De todos modos, muchas diferencias entre especies deben reflejar accidentes históricos congelados que no tienen ningún significado funcional (p.e. algunas reorganizaciones cromosómicas) o no tienen una relevancia directa en la función de los genes Hox (p.e. un incremento en el tamaño génico por acumulación y deterioro de elementos transponibles). No hay modo seguro de diferenciar cambios significativos de cambios triviales, pero una manera de comprobar relaciones propuestas entre la estructura y la función de los complejos Hox es determinar si la distribución de características particulares es consistente con la función propuesta (FERRIER y AKAM 1996). ¿Existen reorganizaciones del complejo Hox en muchos linajes de artrópodos, pero rara vez o nunca en cordados? ¿Son los complejos grandes una característica única de insectos con desarrollo rápido o se encuentran también en insectos que invierten días en vez de horas en establecer el patrón de regulación de los genes Hox?

3.3.2. Estructura del HOM-C en los Metazoos.

Los datos sobre la organización genómica de los genes Hox han aumentado desde el inicio de este trabajo (ver apartado 1.6.4). De todos modos siguen siendo pocos y muy irregulares a lo largo de la filogenia. En la Figura 18 se muestra la estructura de los complejos en los organismos donde se ha estudiado. A medida que se estudia la organización genómica de los genes Hox en más organismos, están apareciendo más casos de roturas del complejo. En el último año se han descrito roturas del HOM-C en tres organismos.

Aunque se han clonado muchos fragmentos de genes Hox en los Lofotrocozoos, sigue sin haber datos sobre su organización genómica. Únicamente en el nemertino *Lineus sanguineus* hay alguna evidencia de agrupación de dos genes Hox, obtenida por hibridación sobre electroforesis de campo pulsante (KMITA-CUNISSE *et al.* 1998).

En cuanto a los Ecdisozoos, los datos siguen limitados al nemátodo *C. elegans* y a insectos. Como ya se ha comentado, *C. elegans* posee solo seis genes Hox agrupados en parejas. Esta dotación génica se interpreta como una reducción derivada debido a su desarrollo determinista (WANG *et al.* 1993; BALAVOINE y TELFORD 1995).

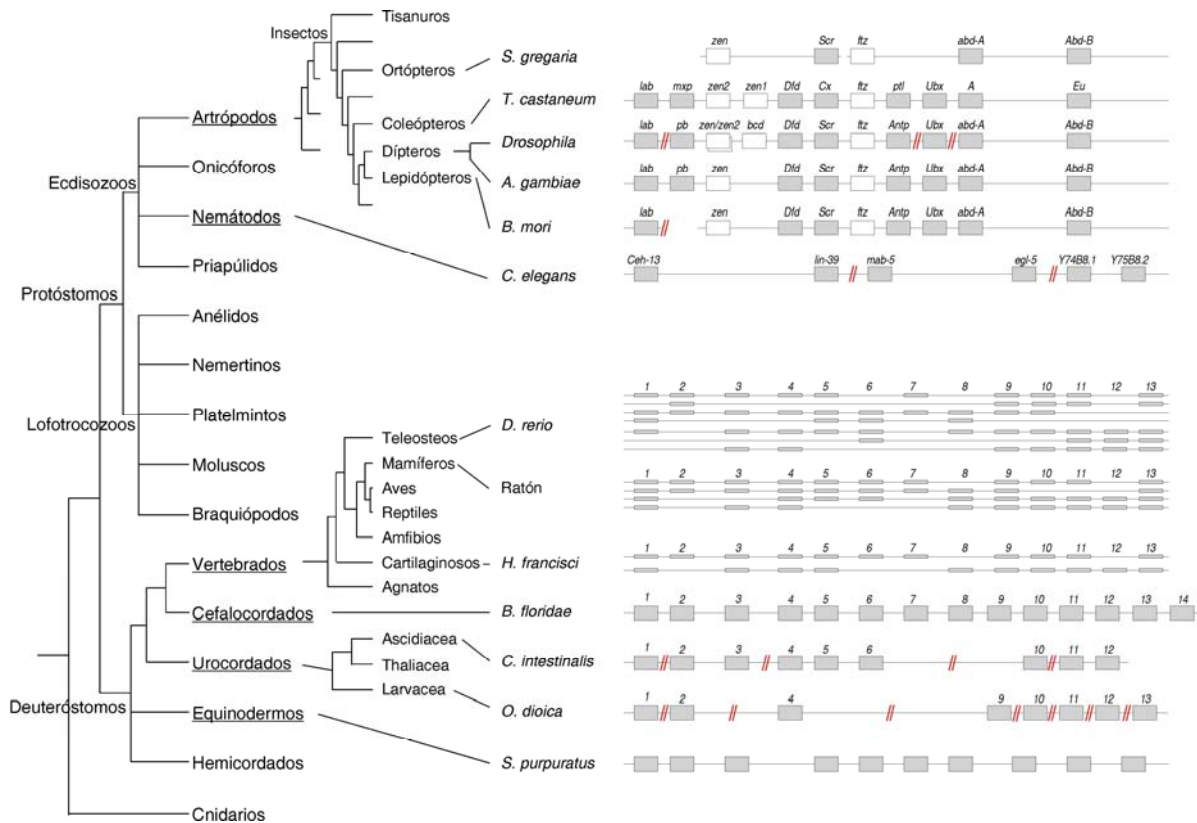


Figura 18. Estructura del HOM-C en los Metazoos. Árbol filogenético con la información disponible sobre la organización genómica de los genes Hox. Información obtenida de: *S. gregaria* (FERRIER y AKAM 1996); *T. castaneum* (DENELL et al. 1996; BROWN et al. 2002); *A. gambiae* (POWERS et al. 2000; LEWIS et al. 2003); *B. mori* (YASUKOCHI et al. 2004); *C. elegans* (DE ROSA et al. 1999); *D. rerio* (AMORES et al. 2004); *H. francisci* de (KIM et al. 2000; CHIU et al. 2002); *B. floridae* (MINGUILLÓN et al. 2004); *C. intestinalis* (SPAGNUOLO et al. 2003; IKUTA et al. 2004); *O. Dioica* (SEO et al. 2004), y *S. purpuratus* (MARTINEZ et al. 1999).

Figure 18. HOM-C structure in the Metazoan. Phylogenetic tree with the available information on the genomic organization of Hox genes. Information from: *S. gregaria* (FERRIER y AKAM 1996); *T. castaneum* (DENELL et al. 1996; BROWN et al. 2002); *A. gambiae* (POWERS et al. 2000; LEWIS et al. 2003); *B. mori* (YASUKOCHI et al. 2004); *C. elegans* (DE ROSA et al. 1999); *D. rerio* (AMORES et al. 2004); *H. francisci* (KIM et al. 2000; CHIU et al. 2002); *B. floridae* (MINGUILLÓN et al. 2004); *C. intestinalis* (SPAGNUOLO et al. 2003; IKUTA et al. 2004); *O. Dioica* (SEO et al. 2004), y *S. purpuratus* (MARTINEZ et al. 1999).

De nuevo, aunque hay muchos datos de presencia de genes Hox en diversos Artrópodos y, en este caso, también patrones de expresión, solamente en insectos se ha determinado la organización genómica de estos genes.

El complejo del saltamontes es al menos tan largo como el ANT-C y el BX-C de *Drosophila* combinados (alrededor de 700 kb), y posiblemente comprende hasta 2 Mb. No se conoce la longitud de las unidades de transcripción, pero estas no pueden ser tan cortas como las de vertebrados. La gran longitud de los genes, y del complejo, sugiere

que *S. gregaria* contiene extensos elementos reguladores específicos de cada gen, más que elementos compactos y sobrepuestos como en vertebrados. (FERRIER y AKAM 1996)

El escarabajo *T. castaneum* también posee un HOM-C (BEEMAN 1987). En algunos aspectos, este complejo es más parecido al HOM-C de vertebrados que al de *Drosophila*, por ejemplo en que incluye unidades de transcripción grandes y distancias intergénicas largas importantes para la elaborada regulación de los genes (DENELL *et al.* 1996). El HOM-C de *Tribolium* forma un bloque contiguo y no presenta genes accesorios (MCGINNIS y KRUMLAUF 1992). Se ha clonado y secuenciado la región equivalente al ANT-C (BROWN *et al.* 2002).

En el 2000 dos trabajos estudiaron la organización genómica de los genes Hox en *A. gambiae* (DEVENPORT *et al.* 2000; POWERS *et al.* 2000). Posteriormente, la secuencia genómica ha completado los datos restantes confirmando la presencia de un complejo único y entero. Las transposiciones en el ancestro de *Drosophila*, así como las duplicaciones génicas (Hox3), se produjeron después de la separación de los linajes de *Drosophila* y *Anopheles* (LEWIS *et al.* 2003).

En cambio, en *Bombyx mori* se había identificado, mediante análisis de mutantes con fenotipo homeótico, un sólo complejo Hox equivalente a los complejos ANT-C y BX-C de *D. melanogaster* (UENO *et al.* 1992). Pero un estudio reciente, en el que se ha clonado en clones BAC todo el complejo, se ha visto que en realidad el gen *labial* se encuentra fuera del complejo y por tanto el complejo está en esta especie partido en dos fragmentos (YASUKOCHI *et al.* 2004).

En los deuteróstomos, además de los datos existentes sobre equinodermos, cefalocordados y vertebrados, recientemente se ha descrito la organización de los genes Hox en dos urocordados. La ascidea *Ciona intestinalis* presenta nueve genes Hox, pero no están agrupados en un complejo sino en cinco fragmentos con entre uno y tres genes (SPAGNUOLO *et al.* 2003; IKUTA *et al.* 2004). El larváceo *Oikopleura dioica* presenta nueve genes Hox que no muestran ningún tipo de agrupación (SEO *et al.* 2004). A diferencia del resto de deuteróstomos estudiados, estos dos organismos muestran pérdida de genes Hox y desestructuración del complejo.

3.3.3. ¿Que Tienen en Común los Organismos con Complejos Fragmentados?

Como hemos visto, *Drosophila* no es el único organismo el que se han descrito roturas del HOM-C. La presencia de un complejo de genes Hox fragmentado se conocía ya en el nemátodo *C. elegans*, y recientemente se ha sido descrito en el lepidóptero *B. mori* y en dos tunicados, *C. intestinalis* y *O. dioica*. ¿Que tienen en común los organismos con el HOM-C fragmentado?

El desarrollo de los vertebrados sigue una progresión antero-posterior, y se ha visto que la colinealidad de los genes Hox no es solo espacial sino también temporal (el reloj Hox) (KMITA y DUBOULE 2003). En el tunicado *Oikopleura* la expresión de los genes Hox todavía mantiene la colinealidad espacial, pero no la temporal (SEO *et al.* 2004), lo que favorece la idea de que la fuerza que mantiene conservada la estructura del HOM-C es la colinealidad temporal (FERRIER y MINGUILLON 2003). En los nemátodos, el patrón evolutivo de los genes Hox se ha relacionado con el cambio hacia un tipo de desarrollo determinístico (ABOUBAKER y BLAXTER 2003). En el caso de *Bombyx*, su desarrollo embrionario es bastante particular. Se han hecho diversos intentos de caracterizarlo, y según que características se han usado se ha clasificado a veces como un insecto de banda germinal corta (formación de los segmentos por proliferación) otra veces como un insecto de banda germinal larga (todos los segmentos se forman de manera simultánea). De todos modos, en lo que coinciden todos los estudios es en que presenta un desarrollo complejo y rápido (DAVIS y PATEL 2002). Finalmente, *Drosophila* es un insecto de banda germinal larga, con un desarrollo rápido donde todos los genes Hox se activan casi simultáneamente durante el estadio de blastodermo celular.

Una característica común a todos los organismos donde se ha descrito un complejo Hox fragmentado parece ser la presencia de un tipo de embriogénesis derivada caracterizada por un desarrollo temprano muy rápido. Ninguno de estos organismos parece mostrar colinealidad temporal. La pérdida de la progresión temporal en la activación de los genes Hox en una embriogénesis rápida podría ser la causa primera de la organización modular de los “complejos” Hox, donde los módulos pueden separarse sin que se produzca una pérdida de función.

4. CONCLUSIONES

1. Las secuencias nucleotídicas obtenidas de las regiones *lab-abdA* y *pb* de *D. buzzatii* y su comparación con las secuencias de *D. melanogaster* y *D. pseudoobscura* muestran que la rotura del HOM-C entre los genes *lab* y *pb* se produjo a menos de 3 kb del extremo 3' de *pb*, entre este gen y la agrupación de genes cuticulares (*Ccp*).

2. En *D. buzzatii* los genes *lab* y *abdA*, situados en la misma posición citológica, se encuentran a tan solo 75 kb. En esta región encontramos un solo gen, *CG31363*, que no está relacionado con los genes Hox. A pesar de su proximidad, la posición de estos dos genes Hox parece ser casual y sin significado funcional.

3. La secuencia obtenida de las regiones *lab* y *pb* en *D. buzzatii* y su comparación con *D. melanogaster* y *D. pseudoobscura* muestran que la rotura *lab-pb* se produjo mediante una inversión paracéntrica y no por transposición. El segundo punto de rotura se produjo entre los genes *CG31363* y *CG17836*. De todos modos, no parecen quedar evidencias del mecanismo de la inversión.

4. El análisis comparativo de la secuencia de la región codificante del gen *lab* entre *D. buzzatii*, *D. virilis* y *D. melanogaster* muestra una evolución heterogénea a lo largo del gen, pero homogénea a lo largo de la filogenia. El cambio de posición no parece haber afectado a su tasa de substitución nucleotídica.

5. La comparación de las secuencias no codificantes alrededor de los tres genes Hox (*lab*, *pb* y *abdA*) entre *D. buzzatii*, *D. melanogaster* y *D. pseudoobscura* muestra una elevada presencia de bloques conservados. Estos bloques conservados, que llegan a ocupar un 22% de la secuencia no codificante en la comparación entre las especies más distantes, representan probablemente regiones reguladoras de la expresión génica

6. La comparación de los bloques conservados con las regiones reguladoras conocidas de estos genes Hox en *D. melanogaster* muestra que la posición y orientación de las regiones reguladoras está conservada entre las tres especies, excepto pequeñas regiones invertidas que probablemente corresponden con CRM individuales.

7. Las tres roturas del HOM-C conocidas en el género *Drosophila* se han producido respetando las regiones reguladoras de los genes adyacentes. En los tres casos la rotura se ha producido cerca del extremo 3' de un gen y lejos del extremo 5' del otro.

8. Los patrones de expresión de los genes Hox *lab* y *abdA* están conservados entre *D. melanogaster*, *D. virilis*, *D. repleta* y *D. buzzatii*. El gen *pb* presenta un dominio extra de expresión en los embriones de *D. virilis*, pero esta diferencia no parece estar relacionada con la rotura *lab-pb*. Por lo tanto las roturas del HOM-C no han tenido efecto sobre el patrón de expresión de los genes Hox adyacentes (*lab*, *pb* o *abdA*).

9. En *Drosophila* los genes Hox presentan una estructura modular donde su agrupación es prescindible para su correcto funcionamiento. En estas especies el HOM-C se ha ido desestructurado a través de la inserción de genes no relacionados mediante transposición y la rotura del complejo mediante inversiones paracéntricas. Además los genes Hox de la región anterior (*lab* y *pb*) están en proceso de modificar su función o adquirir nuevas, como ya ha sucedido con *Hox3* y *ftz*.

10. Además de *Drosophila* actualmente se conocen cuatro especies que presentan roturas en el HOM-C: el nemátodo *C. elegans*, el lepidóptero *B. mori* y los tunicados *C. intestinalis* y *O. dioica*. Estos organismos parecen mantener la colinealidad espacial pero no la temporal. La presencia de un desarrollo embrionario rápido y derivado podría ser la causa de la pérdida de colinealidad temporal y la desestructuración del HOM-C.

1. *The nucleotide sequences from the regions lab-abdA and pb of D. buzzatii and their comparison with the sequences of D. melanogaster and D. pseudoobscura shows that the split of the HOM-C between the genes lab and pb was produced at less than 3 kb from the 3' end of pb, between this gene and the cluster of cuticular genes (Ccp).*

2. *In D. buzzatii the genes lab and abdA, located at the same cytological position, are at 75 kb only. In the intervening region there is only one gene, CG31363, which is not related with the Hox genes. Despite their proximity, the position of these two genes seems fortuitous and without functional meaning.*

3. *The sequences from the lab and pb regions of D. buzzatii and their comparison with those of D. melanogaster and D. pseudoobscura show that the lab-pb split was produced through a paracentric inversion and not by transposition. The second breakpoint was located between the CG31363 and CG17836 genes. However, there is no firm evidence about the mechanism that produced this inversion.*

4. *The comparative analysis of the coding regions of the lab gene between D. buzzatii, D. virilis and D. melanogaster shows that the evolution rate is heterogeneous along the gene, but homogeneous along the phylogeny. Thus, the change in position had no effect on the nucleotide substitution rate of lab.*

5. *The comparison between D. buzzatii, D. melanogaster and D. pseudoobscura of the noncoding sequences around the three Hox genes (lab, pb and abdA) shows a high number of conserved blocks. These blocks represent up to a 22% of noncoding sequences in the comparisons between the more distant species and are putative enhancers.*

6. *The comparison between the conserved blocks and the known regulatory regions in D. melanogaster shows that the position and orientation of the regulatory regions is conserved between the three species, except small inverted regions that probably correspond to individual CRM.*

7. *The three HOM-C splits known in the Drosophila genus have been produced between the regulatory regions of the adjacent genes. In all three cases the splits took place near the 3' end of one gene and far from the 5' end of the other gene.*

8. *The expression patterns of the Hox genes lab and abdA are conserved between D. melanogaster, D. virilis, D. repleta and D. buzzatii. The gene pb presents an extra domain of expression in the embryos of D. virilis, but this difference seems unrelated to the lab-pb split. Thus, the HOM-C splits had no effect on the expression pattern of the adjacents Hox genes (lab, pb or abdA).*

9. *In Drosophila the Hox genes present a modular structure where their clustering is unnecessary for their correct function. In these species the HOM-C is being disassembled by the insertion of nonrelated genes through transposition and split through paracentric inversions. Moreover, the anterior Hox genes (lab and pb) are in process of modifying their function or acquiring new ones, as it has happened with Hox3 and ftz.*

10. *Besides Drosophila, four species are known to have a split HOM-C: the nematode C. elegans, the lepidopteran B. mori and the tunicates C. intestinalis and O. dioica. These organisms seem to keep spatial but not temporal colinearity. The presence of a derived and quick mode of embryogenesis could be the ultimate cause of the loss of temporal colinearity and the HOM-C disassembly.*

5. REFERENCIAS

- ABBOTT, M. K. y T. C. KAUFMAN. 1986. The relationship between the functional complexity and the molecular organization of the Antennapedia locus of *Drosophila melanogaster*. *Genetics* **114**: 919-942.
- ABOUBAKER, A. y M. BLAXTER. 2003. Hox gene evolution in nematodes: novelty conserved. *Curr Opin Genet Dev* **13**: 593-598.
- ABZHANOV, A. y T. C. KAUFMAN. 1999. Homeotic genes and the arthropod head: expression patterns of the *labial*, *proboscipedia*, and *Deformed* genes in crustaceans and insects. *Proc Natl Acad Sci U S A* **96**: 10224-10229.
- ACAMPORA, D., M. D'ESPOSITO, A. FAIELLA, M. PANNESE, E. MIGLIACCIO, F. MORELLI, A. STORNAIUOLO, V. NIGRO, A. SIMEONE y E. BONCINELLI. 1989. The human HOX gene family. *Nucleic Acids Res* **17**: 10385-10402.
- 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.
- AGUINALDO, A. M., J. M. TURBEVILLE, L. S. LINFORD, M. C. RIVERA, J. R. GAREY, R. A. RAFF y J. A. LAKE. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**: 489-493.
- AKAM, M. 1987. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**: 1-22.
- AKAM, M. 1989. Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell* **57**: 347-349.
- AKAM, M. 1998. Hox genes: from master genes to micromanagers. *Curr Biol* **8**: R676-678.
- AMORES, A., A. FORCE, Y. L. YAN, L. JOLY, C. AMEMIYA, A. FRITZ, R. K. HO, J. LANGELAND, V. PRINCE, Y. L. WANG, M. WESTERFIELD, M. EKKER y J. H. POSTLETHWAIT. 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**: 1711-1714.
- AMORES, A., T. SUZUKI, Y.-L. YAN, J. POMEROY, A. SINGER, C. AMEMIYA y J. H. POSTLETHWAIT. 2004. Developmental Roles of Pufferfish Hox Clusters and Genome Evolution in Ray-Fin Fish. *Genome Res.* **14**: 1-10.
- ARBEITMAN, M. N., E. E. M. FURLONG, F. IMAM, E. JOHNSON, B. H. NULL, B. S. BAKER, M. A. KRASNOW, M. P. SCOTT, R. W. DAVIS y K. P. WHITE. 2002. Gene

- Expression During the Life Cycle of *Drosophila melanogaster*. *Science* **297**: 2270-2275.
- ARNONE, M. I. y E. H. DAVIDSON. 1997. The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**: 1851-1864.
- AVEROF, M., R. DAWES y D. FERRIER. 1996. Diversification of arthropod Hox genes as a paradigm for the evolution of gene functions. *Seminars in Cell & Developmental Biology* **7**: 539-551.
- BAE, E., V. C. CALHOUN, M. LEVINE, E. B. LEWIS y R. A. DREWELL. 2002. Characterization of the intergenic RNA profile at *abdominal-A* and *Abdominal-B* in the *Drosophila bithorax* complex. *Proc Natl Acad Sci U S A* **99**: 16847-16852.
- BALAVOINE, G. y M. J. TELFORD. 1995. Identification of planarian homeobox sequences indicates the antiquity of most Hox/homeotic gene subclasses. *Proc Natl Acad Sci U S A* **92**: 7227-7231.
- BASLER, K. y E. HAFEN. 1989. Ubiquitous expression of sevenless: position-dependent specification of cell fate. *Science* **243**: 931-934.
- BATE, M., V. HARTENSTEIN y A. MARTINEZ ARIAS. 1993. *The Development of Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
- BEEMAN, R. W. 1987. A homoeotic gene cluster in the red flour beetle. *Nature* **327**: 247-249.
- BEGUN, D. J. y C. H. LANGLEY. 2003. Proposal for the Sequencing of *Drosophila yakuba* and *D. simulans*. *White Paper*.
- BENDER, W., M. AKAM, F. KARCH, P. A. BEACHY, M. PEIFER, P. SPIERER, E. B. LEWIS y D. S. HOGNESS. 1983a. Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science* **221**: 23-29.
- BENDER, W., P. SPIERER y D. S. HOGNESS. 1983b. Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the bithorax complex in *Drosophila melanogaster*. *J Mol Biol* **168**: 17-33.
- BERGMAN, C., B. PFEIFFER, D. RINCON-LIMAS, R. HOSKINS, A. GNIRKE, C. MUNGALL, A. WANG, B. KRONMILLER, J. PACLEB, S. PARK, M. STAPLETON, K. WAN, R. GEORGE, P. DE JONG, J. BOTAS, G. RUBIN y S. CELNIKER. 2002. Assessing the impact of comparative genomic sequence data on the functional annotation of the *Drosophila* genome. *Genome Biol* **3**: research0086.

- BERGMAN, C. M. y M. KREITMAN. 2001. Analysis of conserved noncoding DNA in *Drosophila* reveals similar constraints in intergenic and intronic sequences. *Genome Res* **11**: 1335-1345.
- BERLETH, T., M. BURRI, G. THOMA, D. BOPP, S. RICHSTEIN, G. FRIGERIO, M. NOLL y C. NUSSLEIN-VOLHARD. 1988. The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo. *Embo J* **7**: 1749-1756.
- BERMAN, B. P., Y. NIBU, B. D. PFEIFFER, P. TOMANCAK, S. E. CELNIKER, M. LEVINE, G. M. RUBIN y M. B. EISEN. 2002. Exploiting transcription factor binding site clustering to identify *cis*-regulatory modules involved in pattern formation in the *Drosophila* genome. *Proc Natl Acad Sci U S A* **99**: 757-762.
- BERMAN, B. P., B. D. PFEIFFER, T. R. LAVERTY, S. L. SALZBERG, G. M. RUBIN, M. B. EISEN y S. E. CELNIKER. 2004. Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Genome Biol* **5**: R61.
- BETRÁN, E., K. THORNTON y M. LONG. 2002. Retroposed new genes out of the X in *Drosophila*. *Genome Res* **12**: 1854-1859.
- BOFFELLI, D., J. MCAULIFFE, D. OVCHARENKO, K. D. LEWIS, I. OVCHARENKO, L. PACHTER y E. M. RUBIN. 2003. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* **299**: 1391-1394.
- BOFFELLI, D., M. A. NOBREGA y E. M. RUBIN. 2004. Comparative genomics at the vertebrate extremes. *Nat Rev Genet* **5**: 456-465.
- BOMZE, H. M. y A. J. LOPEZ. 1994. Evolutionary conservation of the structure and expression of alternatively spliced *Ultrabithorax* isoforms from *Drosophila*. *Genetics* **136**: 965-977.
- BONNER, J. J., C. PARKS, J. PARKER-THORNBURG, M. A. MORTIN y H. R. PELHAM. 1984. The use of promoter fusions in *Drosophila* genetics: isolation of mutations affecting the heat shock response. *Cell* **37**: 979-991.
- BOTAS, J. 1993. Control of morphogenesis and differentiation by HOM/Hox genes. *Curr Opin Cell Biol* **5**: 1015-1022.
- BOULET, A. M. y M. P. SCOTT. 1988. Control elements of the P2 promoter of the *Antennapedia* gene. *Genes Dev* **2**: 1600-1614.

- BOUTANAEV, A. M., A. I. KALMYKOVA, Y. Y. SHEVELYOV y D. I. NURMINSKY. 2002. Large clusters of co-expressed genes in the *Drosophila* genome. *Nature* **420**: 666-669.
- BRAND, A. H., L. BREEDEN, J. ABRAHAM, R. STERNGLANZ y K. NASMYTH. 1985. Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. *Cell* **41**: 41-48.
- BROOKE, N. M., J. GARCIA-FERNANDEZ y P. W. HOLLAND. 1998. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* **392**: 920-922.
- BROWN, S. J., J. P. FELLERS, T. D. SHIPPY, E. A. RICHARDSON, M. MAXWELL, J. J. STUART y R. E. DENELL. 2002. Sequence of the *Tribolium castaneum* homeotic complex: the region corresponding to the *Drosophila melanogaster* antennapedia complex. *Genetics* **160**: 1067-1074.
- BUCHANAN, C. D., P. E. KLEIN y J. E. MULLET. 2004. Phylogenetic Analysis of 5'-Noncoding Regions From the ABA-Responsive rab16/17 Gene Family of Sorghum, Maize and Rice Provides Insight Into the Composition, Organization and Function of *cis*-Regulatory Modules. *Genetics* **168**: 1639-1654.
- BURKE, A. C., C. E. NELSON, B. A. MORGAN y C. TABIN. 1995. Hox genes and the evolution of vertebrate axial morphology. *Development* **121**: 333-346.
- BURTIS, K. C., R. S. HAWLEY y H. D. LIPSHITZ. 2003. The 2003 Thomas Hunt Morgan Medal. *Genetics* **164**: 1243-1245.
- BUTLER, J. E. F. y J. T. KADONAGA. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* **16**: 2583-2592.
- CÁCERES, M., M. PUIG y A. RUIZ. 2001. Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. *Genome Res* **11**: 1353-1364.
- CÁCERES, M., J. M. RANZ, A. BARBADILLA, M. LONG y A. RUIZ. 1999. Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**: 415-418.
- CARROLL, R. L. 2000. Towards a new evolutionary synthesis. *Trends in Ecology & Evolution* **15**: 27-32.
- CARROLL, S. B. 1995. Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**: 479-485.
- CARSON, H. L. 1971. *The ecology of Drosophila breeding sites*. University of Hawaii. Harold L. Lyon Arboretum lecture Number Two, Honolulu.

- CASALS, F., M. CÁCERES, M. H. MANFRIN, J. GONZÁLEZ y A. RUIZ. 2005. Molecular characterization and chromosomal distribution of Galileo, Kepler and Newton, three foldback transposable elements of the *Drosophila buzzatii* species complex. *Genetics* **169**:2047-2059.
- CASALS, F., M. CÁCERES y A. RUIZ. 2003. The foldback-like transposon Galileo is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*. *Mol Biol Evol* **20**: 674-685.
- CASTELLI-GAIR, J. y M. AKAM. 1995. How the Hox gene *Ultrabithorax* specifies two different segments: the significance of spatial and temporal regulation within metameres. *Development* **121**: 2973-2982.
- CASTELLI-GAIR, J. y B. LOVEGROVE. 2003. Beyond homeosis---HOX function in morphogenesis and organogenesis. *Differentiation* **71**: 461-476.
- CELNIKER, S. E., D. J. KEELAN y E. B. LEWIS. 1989. The molecular genetics of the bithorax complex of *Drosophila*: characterization of the products of the *Abdominal-B* domain. *Genes Dev* **3**: 1424-1436.
- CELNIKER, S. E., S. SHARMA, D. J. KEELAN y E. B. LEWIS. 1990. The molecular genetics of the bithorax complex of *Drosophila*: *cis*-regulation in the *AbdominalB* domain. *Embo J* **9**: 4277-4286.
- CELNIKER, S. E., D. A. WHEELER, B. KRONMILLER, J. W. CARLSON, A. HALPERN, S. PATEL, M. ADAMS, M. CHAMPE, S. P. DUGAN, E. FRISE, A. HODGSON, R. A. GEORGE, R. A. HOSKINS, T. LAVERTY, D. M. MUZNY, C. R. NELSON, J. M. PACLEB, S. PARK, B. D. PFEIFFER, S. RICHARDS *et al.* 2002. Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol* **3**: research0079.
- CHIANG, A., M. B. O'CONNOR, R. PARO, J. SIMON y W. BENDER. 1995. Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. *Development* **121**: 1681-1689.
- CHIU, C. H., C. AMEMIYA, K. DEWAR, C. B. KIM, F. H. RUDDLE y G. P. WAGNER. 2002. Molecular evolution of the HoxA cluster in the three major gnathostome lineages. *Proc Natl Acad Sci U S A* **99**: 5492-5497.
- CHOUINARD, S. y T. C. KAUFMAN. 1991. Control of expression of the homeotic *labial* (*lab*) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. *Development* **113**: 1267-1280.

- CIRERA, S., J. M. MARTIN-CAMPOS, C. SEGARRA y M. AGUADE. 1995. Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. *Genetics* **139**: 321-326.
- CLARK, A., G. GIBSON, T. KAUFMAN, B. MCALLISTER, E. MYERS y P. O'GRADY. 2003. Proposal for *Drosophila* as a Model System for Comparative Genomics. *White Paper*.
- CLARK, A. G. 2001. The Search for Meaning in Noncoding DNA. *Genome Res.* **11**: 1319-1320.
- 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*, editado por M. ASHBURNER y E. NOVITSKI. Academic Press, London ; New York.
- COLOMBO, P., J. YON, K. GARSON y M. FRIED. 1992. Conservation of the organization of five tightly clustered genes over 600 million years of divergent evolution. *Proc Natl Acad Sci U S A* **89**: 6358-6362.
- COOPER, G. M. y A. SIDOW. 2003. Genomic regulatory regions: insights from comparative sequence analysis. *Curr Opin Genet Dev* **13**: 604-610.
- CRIBBS, D. L., A. M. PATTATUCCI, M. A. PULTZ y T. C. KAUFMAN. 1992. Ectopic expression of the *Drosophila* homeotic gene *proboscipedia* under Antennapedia P1 control causes dominant thoracic defects. *Genetics* **132**: 699-711.
- DANCHIN, E. G. y P. PONTAROTTI. 2004. Towards the reconstruction of the bilaterian ancestral pre-MHC region. *Trends Genet* **20**: 587-591.
- DAVIS, G. K. y N. H. PATEL. 2002. SHORT, LONG, AND BEYOND: Molecular and Embryological Approaches to Insect Segmentation. *Annual Review of Entomology* **47**: 669-699.
- DE MEAUX, J., U. GOEBEL, A. POP y T. MITCHELL-OLDS. 2005. Allele-Specific Assay Reveals Functional Variation in the Chalcone Synthase Promoter of *Arabidopsis thaliana* That Is Compatible with Neutral Evolution. *Plant Cell* **17**: 676-690.
- DE ROSA, R., J. K. GRENIER, T. ANDREEVA, C. E. COOK, A. ADOUTTE, M. AKAM, S. B. CARROLL y G. BALAVOINE. 1999. Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399**: 772-776.
- DENELL, R. E., S. J. BROWN y R. W. BEEMAN. 1996. Evolution of the organization and function of insect homeotic complexes. *Seminars in Cell & Developmental Biology* **7**: 527-538.

- DERMITZAKIS, E. T., A. REYMOND, R. LYLE, N. SCAMUFFA, C. UCLA, S. DEUTSCH, B. J. STEVENSON, V. FLEGEL, P. BUCHER, C. V. JONGENEEL y S. E. ANTONARAKIS. 2002. Numerous potentially functional but non-genic conserved sequences on human chromosome 21. *Nature* **420**: 578-582.
- DEVENPORT, M. P., C. BLASS y P. EGGLESTON. 2000. Characterization of the Hox gene cluster in the malaria vector mosquito, *Anopheles gambiae*. *Evol Dev* **2**: 326-339.
- DIEDERICH, R. J., V. K. MERRILL, M. A. PULTZ y T. C. KAUFMAN. 1989. Isolation, structure, and expression of *labial*, a homeotic gene of the Antennapedia Complex involved in Drosophila head development. *Genes Dev* **3**: 399-414.
- DIEDERICH, R. J., A. M. PATTATUCCI y T. C. KAUFMAN. 1991. Developmental and evolutionary implications of *labial*, *Deformed* and *engrailed* expression in the Drosophila head. *Development* **113**: 273-281.
- DREWELL, R. A., E. BAE, J. BURR y E. B. LEWIS. 2002. Transcription defines the embryonic domains of cis-regulatory activity at the Drosophila bithorax complex. *Proc Natl Acad Sci U S A* **99**: 16853-16858.
- DUBOULE, D. y G. MORATA. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* **10**: 358-364.
- DUNCAN, I. 1987. The bithorax complex. *Annu Rev Genet* **21**: 285-319.
- EBNER, A., C. CABERNARD, M. AFFOLTER y S. MERABET. 2005. Recognition of distinct target sites by a unique Labial/Extradenticle/Homothorax complex. *Development* **132**: 1591-1600.
- FALCIANI, F., B. HAUSDORF, R. SCHRODER, M. AKAM, D. TAUTZ, R. DENELL y S. BROWN. 1996. Class 3 Hox genes in insects and the origin of *zen*. *Proc Natl Acad Sci U S A* **93**: 8479-8484.
- FEANY, M. B. y W. W. BENDER. 2000. A Drosophila model of Parkinson's disease. *Nature* **404**: 394-398.
- FERRIER, D. E. y M. AKAM. 1996. Organization of the Hox gene cluster in the grasshopper, *Schistocerca gregaria*. *Proc Natl Acad Sci U S A* **93**: 13024-13029.
- FERRIER, D. E. y C. MINGUILLON. 2003. Evolution of the Hox/ParaHox gene clusters. *Int J Dev Biol* **47**: 605-611.
- FINNERTY, J. R. y M. Q. MARTINDALE. 1999. Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. *Evol Dev* **1**: 16-23.

- FRAZER, K. A., L. PACTER, A. POLIAKOV, E. M. RUBIN y I. DUBCHAK. 2004a. VISTA: computational tools for comparative genomics. *Nucl. Acids Res.* **32**: W273-279.
- FRAZER, K. A., H. TAO, K. OSOEGAWA, P. J. DE JONG, X. CHEN, M. F. DOHERTY y D. R. COX. 2004b. Noncoding Sequences Conserved in a Limited Number of Mammals in the SIM2 Interval are Frequently Functional. *Genome Res.* **14**: 367-372.
- GAFFNEY, D. J. y P. D. KEIGHTLEY. 2004. Unexpected conserved non-coding DNA blocks in mammals. *Trends Genet* **20**: 332-337.
- GALANT, R. y S. B. CARROLL. 2002. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**: 910-913.
- GARCIA-FERNANDEZ, J. 2004. Hox, ParaHox, ProtoHox: facts and guesses. *Heredity* **94**: 145-152.
- GARCIA-FERNÁNDEZ, J. y P. W. HOLLAND. 1994. Archetypal organization of the amphioxus Hox gene cluster. *Nature* **370**: 563-566.
- GAUCHAT, D., F. MAZET, C. BERNEY, M. SCHUMMER, S. KREGER, J. PAWLOWSKI y B. GALLIOT. 2000. Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc Natl Acad Sci U S A* **97**: 4493-4498.
- GELLON, G. y W. MCGINNIS. 1998. Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *Bioessays* **20**: 116-125.
- GIBSON, G. 1999. Insect evolution: Redesigning the fruitfly. *Curr Biol* **9**: R86-89.
- GILBERT, S. F. 1997. *Developmental biology*. Sinauer Associates, Sunderland, Mass.
- GINDHART, J. G., JR. y T. C. KAUFMAN. 1995. Identification of Polycomb and trithorax group responsive elements in the regulatory region of the Drosophila homeotic gene *Sex combs reduced*. *Genetics* **139**: 797-814.
- GINDHART, J. G., JR., A. N. KING y T. C. KAUFMAN. 1995. Characterization of the cis-regulatory region of the Drosophila homeotic gene *Sex combs reduced*. *Genetics* **139**: 781-795.
- GLAZKO, G. V. y M. NEI. 2003. Estimation of divergence times for major lineages of primate species. *Mol Biol Evol* **20**: 424-434.
- GOFFEAU, A., B. G. BARRELL, H. BUSSEY, R. W. DAVIS, B. DUJON, H. FELDMANN, F. GALIBERT, J. D. HOHEISEL, C. JACQ, M. JOHNSTON, E. J. LOUIS, H. W. MEWES, Y. MURAKAMI, P. PHILIPPSEN, H. TETTELIN y S. G. OLIVER. 1996. Life with 6000 genes. *Science* **274**: 546, 563-547.

- GOLIC, K. G. y M. M. GOLIC. 1996. Engineering the *Drosophila* Genome: Chromosome Rearrangements by Design. *Genetics* **144**: 1693-1711.
- GOMEZ-SKARMETA, J. L., R. DIEZ DEL CORRAL, E. DE LA CALLE-MUSTIENES, D. FERRE-MARCO y J. MODOLELL. 1996. *Araucan* and *caupolican*, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell* **85**: 95-105.
- GONZÁLEZ, J., E. BETRAN, M. ASHBURNER y A. RUIZ. 2000. Molecular organization of the *Drosophila melanogaster* *Adh* chromosomal region in *D. repleta* and *D. buzzatii*, two distantly related species of the *Drosophila* subgenus. *Chromosome Res* **8**: 375-385.
- GONZÁLEZ, J., F. CASALS y A. RUIZ. 2004. Duplicative and conservative transpositions of *larval serum protein 1* genes in the genus *Drosophila*. *Genetics* **168**: 253-264.
- GONZÁLEZ, J., M. NEFEDOV, I. BOSDET, F. CASALS, O. CALVETE, A. DELPRAT, H. SHIN, C. CHIU, C. MATHEWSON, N. WYE, R. A. HOSKINS, J. E. SCHEIN, P. DE JONG y A. RUIZ. 2005. A BAC-based physical map of the *Drosophila buzzatii* genome. *Genome Res.*: en prensa.
- GONZÁLEZ, J., J. M. RANZ y A. RUIZ. 2002. Chromosomal elements evolve at different rates in the *Drosophila* genome. *Genetics* **161**: 1137-1154.
- GONZALEZ-REYES, A., N. URQUIA, W. J. GEHRING, G. STRUHL y G. MORATA. 1990. Are cross-regulatory interactions between homoeotic genes functionally significant? *Nature* **344**: 78-80.
- GRENIER, J. K., T. L. GARBER, R. WARREN, P. M. WHITINGTON y S. CARROLL. 1997. Evolution of the entire arthropod Hox gene set predated the origin and radiation of the onychophoran/arthropod clade. *Curr Biol* **7**: 547-553.
- GUARNIERI, D. J. y U. HEBERLEIN. 2003. *Drosophila melanogaster*, a genetic model system for alcohol research. *Int Rev Neurobiol* **54**: 199-228.
- GUO, H. y S. P. MOOSE. 2003. Conserved Noncoding Sequences among Cultivated Cereal Genomes Identify Candidate Regulatory Sequence Elements and Patterns of Promoter Evolution. *Plant Cell* **15**: 1143-1158.
- HALLIGAN, D. L., A. EYRE-WALKER, P. ANDOLFATTO y P. D. KEIGHTLEY. 2004. Patterns of evolutionary constraints in intronic and intergenic DNA of *Drosophila*. *Genome Res* **14**: 273-279.
- HARTENSTEIN, V., 1993 *Atlas of Drosophila development*. Cold Spring Harbor Laboratory Press, Plainview, N.Y.

- HARTL, D. L. y E. R. LOZOVSKAYA. 1995. *The Drosophila genome map: a practical guide*. Springer-Verlag, New York.
- HARTL, D. L., D. I. NURMINSKY, R. W. JONES y E. R. LOZOVSKAYA. 1994. Genome structure and evolution in Drosophila: applications of the framework P1 map. *Proc Natl Acad Sci U S A* **91**: 6824-6829.
- HEBERLEIN, U. 2000. Genetics of alcohol-induced behaviors in Drosophila. *Alcohol Res Health* **24**: 185-188.
- HELFAND, S. L. y S. K. INOUE. 2003. Aging, life span, genetics and the fruit fly. *Clinical Neuroscience Research* **2**: 270-278.
- HERSH, B. M. y S. B. CARROLL. 2005. Direct regulation of *knot* gene expression by *Ultrabithorax* and the evolution of *cis*-regulatory elements in Drosophila. *Development* **132**: 1567-1577.
- HILD, M., B. BECKMANN, S. A. HAAS, B. KOCH, V. SOLOVYEV, C. BUSOLD, K. FELLEBERG, M. BOUTROS, M. VINGRON, F. SAUER, J. D. HOHEISEL y R. PARO. 2003. An integrated gene annotation and transcriptional profiling approach towards the full gene content of the Drosophila genome. *Genome Biol* **5**: R3.
- HILLIKER, A. J. y S. N. TRUSIS-COULTER. 1987. Analysis of the Functional Significance of Linkage Group Conservation in Drosophila. *Genetics* **117**: 233-244.
- HIRTH, F., B. HARTMANN y H. REICHERT. 1998. Homeotic gene action in embryonic brain development of Drosophila. *Development* **125**: 1579-1589.
- HOLLAND, P. W. 2001. Beyond the Hox: how widespread is homeobox gene clustering? *J Anat* **199**: 13-23.
- HOLLAND, P. W., L. Z. HOLLAND, N. A. WILLIAMS y N. D. HOLLAND. 1992. An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development* **116**: 653-661.
- HOLLEY, S. A., P. D. JACKSON, Y. SASAI, B. LU, E. M. DE ROBERTIS, F. M. HOFFMANN y E. L. FERGUSON. 1995. A conserved system for dorsal-ventral patterning in insects and vertebrates involving *sog* and *chordin*. *Nature* **376**: 249-253.
- HOOPER, J. E., M. PEREZ-ALONSO, J. R. BERMINGHAM, M. PROUT, B. A. ROCKLEIN, M. WAGENBACH, J. E. EDSTROM, R. DE FRUTOS y M. P. SCOTT. 1992. Comparative studies of Drosophila *Antennapedia* genes. *Genetics* **132**: 453-469.
- HOPPLER, S. y M. BIENZ. 1994. Specification of a single cell type by a Drosophila homeotic gene. *Cell* **76**: 689-702.

- HOSKINS, R. A., C. R. NELSON, B. P. BERMAN, T. R. LAVERTY, R. A. GEORGE, L. CIESIOLKA, M. NAEEMUDDIN, A. D. ARENSON, J. DURBIN, R. G. DAVID, P. E. TABOR, M. R. BAILEY, D. R. DESHAZO, J. CATANESE, A. MAMMOSER, K. OSOEGAWA, P. J. DE JONG, S. E. CELNIKER, R. A. GIBBS, G. M. RUBIN *et al.* 2000. A BAC-based physical map of the major autosomes of *Drosophila melanogaster*. *Science* **287**: 2271-2274.
- HOSKINS, R. A., C. D. SMITH, J. W. CARLSON, A. B. CARVALHO, A. HALPERN, J. S. KAMINKER, C. KENNEDY, C. J. MUNGALL, B. A. SULLIVAN, G. G. SUTTON, J. C. YASUHARA, B. T. WAKIMOTO, E. W. MYERS, S. E. CELNIKER, G. M. RUBIN y G. H. KARPEN. 2002. Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol* **3**: RESEARCH0085.
- HUDEK, A. K., J. CHEUNG, A. P. BORIGHT y S. W. SCHERER. 2003. Genescript: DNA sequence annotation pipeline. *Bioinformatics* **19**: 1177-1178.
- HUGHES, C. L. y T. C. KAUFMAN. 2002. Hox genes and the evolution of the arthropod body plan. *Evol Dev* **4**: 459-499.
- HUGHES, C. L., P. Z. LIU y T. C. KAUFMAN. 2004. Expression patterns of the rogue Hox genes *Hox3/zen* and *fushi tarazu* in the apterygote insect *Thermobia domestica*. *Evol Dev* **6**: 393-401.
- IKUTA, T., N. YOSHIDA, N. SATOH y H. SAIGA. 2004. *Ciona intestinalis* Hox gene cluster: Its dispersed structure and residual colinear expression in development. *Proc Natl Acad Sci U S A* **101**: 15118-15123.
- IRVINE, K. D., S. L. HELFAND y D. S. HOGNESS. 1991. The large upstream control region of the *Drosophila* homeotic gene *Ultrabithorax*. *Development* **111**: 407-424.
- JAGLA, K., M. BELLARD y M. FRASCH. 2001. A cluster of *Drosophila* homeobox genes involved in mesoderm differentiation programs. *Bioessays* **23**: 125-133.
- KAPOUN, A. M. y T. C. KAUFMAN. 1995. A functional analysis of 5', intronic and promoter regions of the homeotic gene *proboscipedia* in *Drosophila melanogaster*. *Development* **121**: 2127-2141.
- KARCH, F., W. BENDER y B. WEIFFENBACH. 1990. *abdA* expression in *Drosophila* embryos. *Genes Dev* **4**: 1573-1587.
- KARCH, F., B. WEIFFENBACH, M. PEIFER, W. BENDER, I. DUNCAN, S. CELNIKER, M. CROSBY y E. B. LEWIS. 1985. The abdominal region of the bithorax complex. *Cell* **43**: 81-96.

- KAUFMAN, T. C. 1978. Cytogenetic Analysis of Chromosome 3 in *Drosophila melanogaster*: Isolation and Characterization of Four New Alleles of the *proboscipedia (pb)* Locus. *Genetics* **90**: 579-596.
- KAUFMAN, T. C., R. LEWIS y B. WAKIMOTO. 1980. Cytogenetic Analysis of Chromosome 3 in *Drosophila melanogaster*: the Homoeotic Gene Complex in Polytene Chromosome Interval 84A-B. *Genetics* **94**: 115-133.
- KAUFMAN, T. C., M. A. SEEGER y G. OLSEN. 1990. Molecular and genetic organization of the *Antennapedia* gene complex of *Drosophila melanogaster*. *Adv Genet* **27**: 309-362.
- KENNISON, J. A. 1993. Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. *Trends Genet* **9**: 75-79.
- KEYS, D. N., B.-I. LEE, A. DI GREGORIO, N. HARAFUJI, J. C. DETTER, M. WANG, O. KAHSAI, S. AHN, C. ZHANG, S. A. DOYLE, N. SATOH, Y. SATOU, H. SAIGA, A. T. CHRISTIAN, D. S. ROKHSAR, T. L. HAWKINS, M. LEVINE y P. M. RICHARDSON. 2005. A saturation screen for *cis*-acting regulatory DNA in the Hox genes of *Ciona intestinalis*. *Proc Natl Acad Sci U S A* **102**: 679-683.
- KIM, C. B., C. AMEMIYA, W. BAILEY, K. KAWASAKI, J. MEZEY, W. MILLER, S. MINOSHIMA, N. SHIMIZU, G. WAGNER y F. RUDDLE. 2000. Hox cluster genomics in the horn shark, *Heterodontus francisci*. *Proc Natl Acad Sci U S A* **97**: 1655-1660.
- KMITA, M. y D. DUBOULE. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**: 331-333.
- KMITA-CUNISSE, M., F. LOOSLI, J. BIERNE y W. J. GEHRING. 1998. Homeobox genes in the ribbonworm *Lineus sanguineus*: evolutionary implications. *Proc Natl Acad Sci U S A* **95**: 3030-3035.
- KOLBE, D., J. TAYLOR, L. ELNITSKI, P. ESWARA, J. LI, W. MILLER, R. HARDISON y F. CHIAROMONTE. 2004. Regulatory Potential Scores From Genome-Wide Three-Way Alignments of Human, Mouse, and Rat. *Genome Res.* **14**: 700-707.
- KORNBERG, T. B. y A. M. A. KRASNOW. 2000. The *Drosophila* Genome Sequence: Implications for Biology and Medicine. *Science* **287**: 2218-2220.
- KOSAK, S. T. y M. GROUDINE. 2004. Form follows function: the genomic organization of cellular differentiation. *Genes Dev.* **18**: 1371-1384.
- LAIRD, C. D. 1973. DNA of *Drosophila* chromosomes. *Annu Rev Genet* **7**: 177-204.

- LANDER, E. S., L. M. LINTON, B. BIRREN, C. NUSBAUM, M. C. ZODY, J. BALDWIN, K. DEVON, K. DEWAR, M. DOYLE, W. FITZHUGH, R. FUNKE, D. GAGE, K. HARRIS, A. HEAFORD, J. HOWLAND, L. KANN, J. LEHOCZKY, R. LEVINE, P. MCEWAN, K. MCKERNAN *et al.* 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- LAWRENCE, P. A. 1992. *The making of a fly : the genetics of animal design*. Blackwell Science, Oxford [England] ; Cambridge, Mass., USA.
- LEEMANS, R., T. LOOP, B. EGGER, H. HE, L. KAMMERMEIER, B. HARTMANN, U. CERTA, H. REICHERT y F. HIRTH. 2001. Identification of candidate downstream genes for the homeodomain transcription factor Labial in *Drosophila* through oligonucleotide-array transcript imaging. *Genome Biol* **2**.
- LERCHER, M. J., A. O. URRUTIA y L. D. HURST. 2002. Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat Genet* **31**: 180-183.
- LEVINE, M. y R. TJIAN. 2003. Transcription regulation and animal diversity. *Nature* **424**: 147-151.
- LEWIS, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**: 565-570.
- LEWIS, E. B. 1998. The bithorax complex: the first fifty years. *Int J Dev Biol* **42**: 403-415.
- LEWIS, E. B., B. D. PFEIFFER, D. R. MATHOG y S. E. CELNIKER. 2003. Evolution of the homeobox complex in the Diptera. *Curr Biol* **13**: R587-588.
- LEWIS, R. A., T. C. KAUFMAN, R. E. DENELL y P. TALLERICO. 1980a. Genetic Analysis of the Antennapedia Gene Complex (ANT-C) and Adjacent Chromosomal Regions of *Drosophila melanogaster*. I. Polytene Chromosome Segments 84B-D. *Genetics* **95**: 367-381.
- LEWIS, R. A., B. T. WAKIMOTO, R. E. DENELL y T. C. KAUFMAN. 1980b. Genetic Analysis of the Antennapedia Gene Complex (ANT-C) and Adjacent Chromosomal Regions of *Drosophila melanogaster*. II. Polytene Chromosome Segments 84A-84B1,2. *Genetics* **95**: 383-397.
- LEWIS, S. E., S. M. SEARLE, N. HARRIS, M. GIBSON, V. LYER, J. RICHTER, C. WIEL, L. BAYRAKTAROGLIR, E. BIRNEY, M. A. CROSBY, J. S. KAMINKER, B. B. MATTHEWS, S. E. PROCHNIK, C. D. SMITHY, J. L. TUPY, G. M. RUBIN, S. MISRA,

- C. J. MUNGALL y M. E. CLAMP. 2002. Apollo: a sequence annotation editor. *Genome Biol* **3**: RESEARCH0082.
- LIFANOV, A. P., V. J. MAKEEV, A. G. NAZINA y D. A. PAPATSENKO. 2003. Homotypic Regulatory Clusters in Drosophila. *Genome Res.* **13**: 579-588.
- LIM, J. K. y M. J. SIMMONS. 1994. Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* **16**: 269-275.
- LUDWIG, M. Z., C. BERGMAN, N. H. PATEL y M. KREITMAN. 2000. Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**: 564-567.
- LUDWIG, M. Z., N. H. PATEL y M. KREITMAN. 1998. Functional analysis of *eve stripe 2* enhancer evolution in Drosophila: rules governing conservation and change. *Development* **125**: 949-958.
- LUNDE, K., B. BIEHS, U. NAUBER y E. BIER. 1998. The *knirps* and *knirps-related* genes organize development of the second wing vein in Drosophila. *Development* **125**: 4145-4154.
- LUTZ, B., H. C. LU, G. EICHELE, D. MILLER y T. C. KAUFMAN. 1996. Rescue of Drosophila *labial* null mutant by the chicken ortholog *Hoxb-1* demonstrates that the function of Hox genes is phylogenetically conserved. *Genes Dev* **10**: 176-184.
- MACARTHUR, S. y J. F. BROOKFIELD. 2004. Expected Rates and Modes of Evolution of Enhancer Sequences. *Mol Biol Evol* **21**: 1064-1073.
- MAIER, D., A. PREISS y J. R. POWELL. 1990. Regulation of the segmentation gene *fushi tarazu* has been functionally conserved in Drosophila. *Embo J* **9**: 3957-3966.
- MAIER, D., D. SPERLICH y J. R. POWELL. 1993. Conservation and change of the developmentally crucial *fushi tarazu* gene in Drosophila. *J Mol Evol* **36**: 315-326.
- MARKSTEIN, M., P. MARKSTEIN, V. MARKSTEIN y M. S. LEVINE. 2002. Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo. *Proc Natl Acad Sci U S A* **99**: 763-768.
- MARTIN, C. H., C. A. MAYEDA, C. A. DAVIS, C. L. ERICSSON, J. D. KNAFELS, D. R. MATHOG, S. E. CELNIKER, E. B. LEWIS y M. J. PALAZZOLO. 1995. Complete sequence of the bithorax complex of Drosophila. *Proc Natl Acad Sci U S A* **92**: 8398-8402.

- MARTIN, N., S. PATEL y J. A. SEGRE. 2004. Long-range comparison of human and mouse Sprr loci to identify conserved noncoding sequences involved in coordinate regulation. *Genome Res.* **14**: 2430-2438.
- MARTINEZ, P., J. P. RAST, C. ARENAS-MENA y E. H. DAVIDSON. 1999. Organization of an echinoderm Hox gene cluster. *Proc Natl Acad Sci U S A* **96**: 1469-1474.
- MARTINEZ-CRUZADO, J. C., C. SWIMMER, M. G. FENERJIAN y F. C. KAFATOS. 1988. Evolution of the autosomal chorion locus in *Drosophila*. I. General organization of the locus and sequence comparisons of genes *s15* and *s19* in evolutionary distant species. *Genetics* **119**: 663-677.
- MAYOR, C., M. BRUDNO, J. R. SCHWARTZ, A. POLIAKOV, E. M. RUBIN, K. A. FRAZER, L. S. PACHTER y I. DUBCHAK. 2000. VISTA : visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**: 1046-1047.
- MCGINNIS, W. y R. KRUMLAUF. 1992. Homeobox genes and axial patterning. *Cell* **68**: 283-302.
- MCGINNIS, W., M. S. LEVINE, E. HAFEN, A. KUROIWA y W. J. GEHRING. 1984. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* **308**: 428-433.
- MERRILL, V. K., R. J. DIEDERICH, F. R. TURNER y T. C. KAUFMAN. 1989. A genetic and developmental analysis of mutations in *labial*, a gene necessary for proper head formation in *Drosophila melanogaster*. *Dev Biol* **135**: 376-391.
- MICHELSON, A. M. 2002. Deciphering genetic regulatory codes: a challenge for functional genomics. *Proc Natl Acad Sci U S A* **99**: 546-548.
- MINGUILLÓN, C., GARDENYES, J. E. SERRA, F. C. CASTRO, A. HILL-FORCE, P. W. H. HOLLAND, C. T. AMEMIYA y J. GARCIA-FERNÁNDEZ. 2004. No more than 14: the end of the amphioxus Hox cluster. *Int. J. Biol. Sci* **1**: 19-23.
- MISRA, S., M. CROSBY, C. MUNGALL, B. MATTHEWS, K. CAMPBELL, P. HRADECKY, Y. HUANG, J. KAMINKER, G. MILLBURN, S. PROCHNIK, C. SMITH, J. TUPY, E. WHITFIELD, L. BAYRAKTAROGU, B. BERMAN, B. BETTENCOURT, S. CELNIKER, A. DE GREY, R. DRYSDALE, N. HARRIS *et al.* 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biology* **3**: research0083.0081 - 0083.0022.
- MLODZIK, M., A. FJOSE y W. J. GEHRING. 1988. Molecular structure and spatial expression of a homeobox gene from the *labial* region of the Antennapedia-complex. *Embo J* **7**: 2569-2578.

- MODELELL, J. y S. CAMPUZANO. 1998. The achaete-scute complex as an integrating device. *Int J Dev Biol* **42**: 275-282.
- MORIYAMA, E. N. y T. GOJOBORI. 1992. Rates of synonymous substitution and base composition of nuclear genes in *Drosophila*. *Genetics* **130**: 855-864.
- MORIYAMA, E. N., D. A. PETROV y D. L. HARTL. 1998. Genome size and intron size in *Drosophila*. *Mol Biol Evol* **15**: 770-773.
- MULLER, J. H. 1940. Bearings of the *Drosophila* work on systematics, pp. 185-268 en *New systematics*, editado por J. HUXLEY. Clarendon Press, Oxford.
- MÜLLER, W. A. 1997. *Developmental Biology*. Springer-Verlag, New York.
- NADEAU, J. H. y B. A. TAYLOR. 1984. Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc Natl Acad Sci U S A* **81**: 814-818.
- NARDELLI-HAEFLIGER, D. y M. SHANKLAND. 1992. *Lox2*, a putative leech segment identity gene, is expressed in the same segmental domain in different stem cell lineages. *Development* **116**: 697-710.
- NASIADKA, A., B. H. DIETRICH y H. M. KRAUSE. 2002. Anterior-posterior patterning in the *Drosophila* embryo. *Advances in Developmental Biology and Biochemistry* **12**: 155-204.
- NEI, M., P. XU y G. GLAZKO. 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proc Natl Acad Sci U S A* **98**: 2497-2502.
- NELSON, C. E., B. M. HERSH y S. B. CARROLL. 2004. The regulatory content of intergenic DNA shapes genome architecture. *Genome Biol* **5**: R25.
- NETTER, S., M.-O. FAUVARQUE, R. DIEZ DEL CORRAL, J.-M. DURA y D. COEN. 1998. white⁺ Transgene Insertions Presenting a Dorsal/Ventral Pattern Define a Single Cluster of Homeobox Genes That Is Silenced by the Polycomb-group Proteins in *Drosophila melanogaster*. *Genetics* **149**: 257-275.
- NJOLSTAD, P. R. y A. FJOSE. 1988. *In situ* hybridization patterns of zebrafish homeobox genes homologous to *Hox-2.1* and *En-2* of mouse. *Biochem Biophys Res Commun* **157**: 426-432.
- NJOLSTAD, P. R., A. MOLVEN, H. G. EIKEN y A. FJOSE. 1988. Structure and neural expression of a zebrafish homeobox sequence. *Gene* **73**: 33-46.
- NOBREGA, M. A., I. OVCHARENKO, V. AFZAL y E. M. RUBIN. 2003. Scanning human gene deserts for long-range enhancers. *Science* **302**: 413.

- NOVITSKI, E. 1946. Chromosome Variation in *Drosophila athabasca*. *Genetics* **31**: 508-524.
- O'KANE, C. J. y W. J. GEHRING. 1987. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc Natl Acad Sci U S A* **84**: 9123-9127.
- OLIVER, B. y B. LEBLANC. 2003. How many genes in a genome? *Genome Biology* **5**: 204.
- O'NEIL, M. T. y J. M. BELOTE. 1992. Interspecific comparison of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics* **131**: 113-128.
- PARDUE, M. L., S. A. GERBI, R. A. ECKHARDT y J. G. GALL. 1970. Cytological localization of DNA complementary to ribosomal RNA in polytene chromosomes of Diptera. *Chromosoma* **29**: 268-290.
- PETROV, D. A., E. R. LOZOVSKAYA y D. L. HARTL. 1996. High intrinsic rate of DNA loss in *Drosophila*. *Nature* **384**: 346-349.
- PEVZNER, P. y G. TESLER. 2003. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *PNAS* **100**: 7672-7677.
- POPODI, E., J. C. KISSINGER, M. E. ANDREWS y R. A. RAFF. 1996. Sea urchin Hox genes: insights into the ancestral Hox cluster. *Mol Biol Evol* **13**: 1078-1086.
- POTTER, C. J., G. S. TURENCHALK y T. XU. 2000. *Drosophila* in cancer research. An expanding role. *Trends Genet* **16**: 33-39.
- POWELL, J. R. 1997. *Progress and Prospects in Evolutionary Biology: the Drosophila Model*. Oxford University Press, New York.
- POWELL, J. R. y R. DESALLE. 1995. *Drosophila* molecular phylogenies and their uses. *Evol Biol* **28**: 87-138.
- POWERS, T. P., J. HOGAN, Z. KE, K. DYMBROWSKI, X. WANG, F. H. COLLINS y T. C. KAUFMAN. 2000. Characterization of the Hox cluster from the mosquito *Anopheles gambiae* (Diptera: Culicidae). *Evol Dev* **2**: 311-325.
- PULTZ, M. A., R. J. DIEDERICH, D. L. CRIBBS y T. C. KAUFMAN. 1988. The proboscipedia locus of the Antennapedia complex: a molecular and genetic analysis. *Genes Dev* **2**: 901-920.
- QIAN, S., M. CAPOVILLA y V. PIRROTTA. 1993. Molecular mechanisms of pattern formation by the BRE enhancer of the Ubx gene. *Embo J* **12**: 3865-3877.

- RANDAZZO, F. M., D. L. CRIBBS y T. C. KAUFMAN. 1991. Rescue and regulation of *proboscipedia*: a homeotic gene of the Antennapedia Complex. *Development* **113**: 257-271.
- RANDAZZO, F. M., M. A. SEEGER, C. A. HUSS, M. A. SWEENEY, J. K. CECIL y T. C. KAUFMAN. 1993. Structural changes in the Antennapedia complex of *Drosophila pseudoobscura*. *Genetics* **134**: 319-330.
- 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**: 32-43.
- RANZ, J. M., F. CASALS y A. RUIZ. 2001. How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res* **11**: 230-239.
- RANZ, J. M., J. GONZÁLEZ, F. CASALS y A. RUIZ. 2003. Low occurrence of gene transposition events during the evolution of the genus *Drosophila*. *Evolution Int J Org Evolution* **57**: 1325-1335.
- RICHARDS, S., Y. LIU, B. R. BETTENCOURT, P. HRADECKY, S. LETOVSKY, R. NIELSEN, K. THORNTON, M. J. HUBISZ, R. CHEN, R. P. MEISEL, O. COURONNE, S. HUA, M. A. SMITH, P. ZHANG, J. LIU, H. J. BUSSEMAKER, M. F. VAN BATENBURG, S. L. HOWELLS, S. E. SCHERER, E. SODERGREN *et al.* 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and *cis*-element evolution. *Genome Res.* **15**: 1-18.
- ROGERS, B. T., M. D. PETERSON y T. C. KAUFMAN. 2002. The development and evolution of insect mouthparts as revealed by the expression patterns of gnathocephalic genes. *Evol Dev* **4**: 96-110.
- RUBIN, G. M. y E. B. LEWIS. 2000. A Brief History of *Drosophila*'s Contributions to Genome Research. *Science* **287**: 2216-2218.
- RUBIN, G. M. y A. C. SPRADLING. 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348-353.
- RUSHLOW, C., H. DOYLE, T. HOEY y M. LEVINE. 1987. Molecular characterization of the *zerknüllt* region of the Antennapedia gene complex in *Drosophila*. *Genes Dev* **1**: 1268-1279.
- RUSSO, C. A., N. TAKEZAKI y M. NEI. 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol* **12**: 391-404.

- SÁNCHEZ-HERRERO, E. 1991. Control of the expression of the bithorax complex genes abdominal-A and abdominal-B by *cis*-regulatory regions in *Drosophila* embryos. *Development* **111**: 437-449.
- SÁNCHEZ-HERRERO, E. y M. AKAM. 1989. Spatially ordered transcription of regulatory DNA in the bithorax complex of *Drosophila*. *Development* **107**: 321-329.
- SÁNCHEZ-HERRERO, E., I. GUERRERO, J. SAMPEDRO y A. GONZALEZ-REYES. 1994. Developmental consequences of unrestricted expression of the *abd-A* gene of *Drosophila*. *Mech Dev* **46**: 153-167.
- SÁNCHEZ-HERRERO, E., I. VERNOS, R. MARCO y G. MORATA. 1985. Genetic organization of *Drosophila* bithorax complex. *Nature* **313**: 108-113.
- SANTINI, S., J. L. BOORE y A. MEYER. 2003. Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. *Genome Res* **13**: 1111-1122.
- SCHAFER, D. J., D. K. FREDLINE, W. R. KNIBB, M. M. GREEN y J. S. BARKER. 1993. Genetics and linkage mapping of *Drosophila buzzatii*. *J Hered* **84**: 188-194.
- SCHULZE, D. H. y C. S. LEE. 1986. DNA sequence comparison among closely related *Drosophila* species in the mulleri complex. *Genetics* **113**: 287-303.
- SCHUMMER, M., I. SCHEURLEN, C. SCHALLER y B. GALLIOT. 1992. HOM/HOX homeobox genes are present in hydra (*Chlorohydra viridissima*) and are differentially expressed during regeneration. *Embo J* **11**: 1815-1823.
- SCHWARTZ, S., Z. ZHANG, K. A. FRAZER, A. SMIT, C. RIEMER, J. BOUCK, R. GIBBS, R. HARDISON y W. MILLER. 2000. PipMaker---A Web Server for Aligning Two Genomic DNA Sequences. *Genome Res.* **10**: 577-586.
- SCOTT, M. P. y A. J. WEINER. 1984. Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of *Drosophila*. *Proc Natl Acad Sci U S A* **81**: 4115-4119.
- SCOTT, M. P., A. J. WEINER, T. I. HAZELRIGG, B. A. POLISKY, V. PIRROTTA, F. SCALENGHE y T. C. KAUFMAN. 1983. The molecular organization of the Antennapedia locus of *Drosophila*. *Cell* **35**: 763-776.
- SEEGER, M. A., L. HAFFLEY y T. C. KAUFMAN. 1988. Characterization of *amalgam*: a member of the immunoglobulin superfamily from *Drosophila*. *Cell* **55**: 589-600.
- SEGARRA, C. y M. AGUADE. 1992. Molecular organization of the X chromosome in different species of the obscura group of *Drosophila*. *Genetics* **130**: 513-521.
- SEGARRA, C., E. R. LOZOVSKAYA, G. RIBO, M. AGUADE y D. L. HARTL. 1995. P1 clones from *Drosophila melanogaster* as markers to study the chromosomal evolution

- of Muller's A element in two species of the obscura group of *Drosophila*. *Chromosoma* **104**: 129-136.
- SEGARRA, C., G. RIBO y M. AGUADE. 1996. Differentiation of Muller's chromosomal elements D and E in the obscura group of *Drosophila*. *Genetics* **144**: 139-146.
- SEO, H. C., R. B. EDVARSEN, A. D. MAELAND, M. BJORDAL, M. F. JENSEN, A. HANSEN, M. FLAAT, J. WEISSENBAACH, H. LEHRACH, P. WINCKER, R. REINHARDT y D. CHOURROUT. 2004. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* **431**: 67-71.
- SHIMELL, M. J., J. SIMON, W. BENDER y M. B. O'CONNOR. 1994. Enhancer point mutation results in a homeotic transformation in *Drosophila*. *Science* **264**: 968-971.
- SHUBIN, N., C. TABIN y S. CARROLL. 1997. Fossils, genes and the evolution of animal limbs. *Nature* **388**: 639-648.
- SIMON, J., M. PEIFER, W. BENDER y M. O'CONNOR. 1990. Regulatory elements of the bithorax complex that control expression along the anterior-posterior axis. *Embo J* **9**: 3945-3956.
- SIMPSON, P. 2002. Evolution of development in closely related species of flies and worms. *Nat Rev Genet* **3**: 907-917.
- SINGH, N. D. y D. A. PETROV. 2004. Rapid sequence turnover at an intergenic locus in *Drosophila*. *Mol Biol Evol* **21**: 670-680.
- SLACK, J. M., P. W. HOLLAND y C. F. GRAHAM. 1993. The zootype and the phylotypic stage. *Nature* **361**: 490-492.
- SNYDER, M. y M. GERSTEIN. 2003. Genomics. Defining genes in the genomics era. *Science* **300**: 258-260.
- SPAGNUOLO, A., F. RISTORATORE, A. DI GREGORIO, F. ANIELLO, M. BRANNO y R. DI LAURO. 2003. Unusual number and genomic organization of Hox genes in the tunicate *Ciona intestinalis*. *Gene* **309**: 71-79.
- SPELLMAN, P. T. y G. M. RUBIN. 2002. Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J Biol* **1**: 5.
- STATHAKIS, D. G., E. S. PENTZ, M. E. FREEMAN, J. KULLMAN, G. R. HANKINS, N. J. PEARLSON y T. R. WRIGHT. 1995. The genetic and molecular organization of the *Dopa decarboxylase* gene cluster of *Drosophila melanogaster*. *Genetics* **141**: 629-655.

- STAUBER, M., H. JACKLE y U. SCHMIDT-OTT. 1999. The anterior determinant *bicoid* of *Drosophila* is a derived Hox class 3 gene. *Proc Natl Acad Sci U S A* **96**: 3786-3789.
- STAUBER, M., A. PRELL y U. SCHMIDT-OTT. 2002. A single *Hox3* gene with composite *bicoid* and *zerknüllt* expression characteristics in non-Cyclorrhaphan flies. *Proc Natl Acad Sci U S A* **99**: 274-279.
- STERN, D. L. 1998. A role of Ultrabithorax in morphological differences between *Drosophila* species. *Nature* **396**: 463-466.
- STRUHL, G. 1984. Splitting the bithorax complex of *Drosophila*. *Nature* **308**: 454-457.
- TAGLE, D. A., B. F. KOOP, M. GOODMAN, J. L. SLIGHTOM, D. L. HESS y R. T. JONES. 1988. Embryonic *epsilon* and *gamma* globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *J Mol Biol* **203**: 439-455.
- TAMURA, K., S. SUBRAMANIAN y S. KUMAR. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol* **21**: 36-44.
- TCES, C. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium. *Science* **282**: 2012-2018.
- TELFORD, M. J. y R. H. THOMAS. 1998. Of mites and *zen*: expression studies in a chelicerate arthropod confirm *zen* is a divergent Hox gene. *Dev Genes Evol* **208**: 591-594.
- TEROL, J., M. PEREZ-ALONSO y R. DE FRUTOS. 1991. *In situ* localization of the *Antennapedia* gene on the chromosomes of nine *Drosophila* species of the obscura group. *Hereditas* **114**: 131-139.
- TEROL, J., M. PEREZ-ALONSO y R. DE FRUTOS. 1995. Molecular characterization of the *zerknüllt* region of the Antennapedia complex of *D. subobscura*. *Chromosoma* **103**: 613-624.
- TIONG, S., L. M. BONE y J. R. WHITTLE. 1985. Recessive lethal mutations within the bithorax-complex in *Drosophila*. *Mol Gen Genet* **200**: 335-342.
- TIONG, S. Y., J. R. WHITTLE y M. C. GRIBBIN. 1987. Chromosomal continuity in the abdominal region of the bithorax complex of *Drosophila* is not essential for its contribution to metameric identity. *Development* **101**: 135-142.

- UEDA, H. R., A. MATSUMOTO, M. KAWAMURA, M. IINO, T. TANIMURA y S. HASHIMOTO. 2002. Genome-wide Transcriptional Orchestration of Circadian Rhythms in *Drosophila*. *J. Biol. Chem.* **277**: 14048-14052.
- UENO, K., C. C. HUI, M. FUKUTA y Y. SUZUKI. 1992. Molecular analysis of the deletion mutants in the E homeotic complex of the silkworm *Bombyx mori*. *Development* **114**: 555-563.
- VENKEN, K. J. T. y H. J. BELLEN. 2005. Emerging Technologies for Gene Manipulation in *Drosophila melanogaster*. *Nat Rev Genet* **6**: 167-178.
- VERAKSA, A., M. DEL CAMPO y W. MCGINNIS. 2000. Developmental patterning genes and their conserved functions: from model organisms to humans. *Mol Genet Metab* **69**: 85-100.
- VIEIRA, J., C. P. VIEIRA, D. L. HARTL y E. R. LOZOVSKAYA. 1997a. Discordant rates of chromosome evolution in the *Drosophila virilis* species group. *Genetics* **147**: 223-230.
- VIEIRA, J., C. P. VIEIRA, D. L. HARTL y E. R. LOZOVSKAYA. 1997b. A framework physical map of *Drosophila virilis* based on P1 clones: applications in genome evolution. *Chromosoma* **106**: 99-107.
- VON ALLMEN, G., I. HOGGA, A. SPIERER, F. KARCH, W. BENDER, H. GYURKOVICS y E. LEWIS. 1996. Splits in fruitfly Hox gene complexes. *Nature* **380**: 116.
- WAGNER, G. P., C. AMEMIYA y F. RUDDLE. 2003. Hox cluster duplications and the opportunity for evolutionary novelties. *Proc Natl Acad Sci U S A* **100**: 14603-14606.
- WAKIMOTO, B. T. y T. C. KAUFMAN. 1981. Analysis of larval segmentation in lethal genotypes associated with the antennapedia gene complex in *Drosophila melanogaster*. *Dev Biol* **81**: 51-64.
- WANG, B. B., M. M. MULLER-IMMERGLUCK, J. AUSTIN, N. T. ROBINSON, A. CHISHOLM y C. KENYON. 1993. A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. *Cell* **74**: 29-42.
- WAWERSIK, S. y R. L. MAAS. 2000. Vertebrate eye development as modeled in *Drosophila*. *Hum Mol Genet* **9**: 917-925.
- WEATHERBEE, S. D., H. F. NIJHOUT, L. W. GRUNERT, G. HALDER, R. GALANT, J. SELEGUE y S. CARROLL. 1999. Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Curr Biol* **9**: 109-115.

- WEINER, A. J., M. P. SCOTT y T. C. KAUFMAN. 1984. A molecular analysis of *fushi tarazu*, a gene in *Drosophila melanogaster* that encodes a product affecting embryonic segment number and cell fate. *Cell* **37**: 843-851.
- WELLER, G. L. y G. G. FOSTER. 1993. Genetic maps of the sheep blowfly *Lucilia cuprina*: linkage-group correlations with other dipteran genera. *Genome* **36**: 495-506.
- 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 U S A* **91**: 3132-3136.
- WHARTON, L. T. 1942. Analysis of the repleta group of *Drosophila*. *Univ. Texas Publ.* **4228**: 23-52.
- WILLIAMS, E. J. B. y D. J. BOWLES. 2004. Coexpression of Neighboring Genes in the Genome of *Arabidopsis thaliana*. *Genome Res.* **14**: 1060-1067.
- WILSON, C., R. K. PEARSON, H. J. BELLEN, C. J. O'KANE, U. GROSSNIKLAUS y W. J. GEHRING. 1989. P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes Dev* **3**: 1301-1313.
- WOLF, F. W. y U. HEBERLEIN. 2003. Invertebrate models of drug abuse. *J Neurobiol* **54**: 161-178.
- WRAY, G. A., M. W. HAHN, E. ABOUHEIF, J. P. BALHOFF, M. PIZER, M. V. ROCKMAN y L. A. ROMANO. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* **20**: 1377-1419.
- XU, T. y G. RUBIN. 1993. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117**: 1223-1237.
- YASUKOCHI, Y., L. A. ASHAKUMARY, C. WU, A. YOSHIDO, J. NOHATA, K. MITA y K. SAHARA. 2004. Organization of the Hox gene cluster of the silkworm, *Bombyx mori*: a split of the Hox cluster in a non-*Drosophila* insect. *Dev Genes Evol* **214**: 606-614.
- ZDOBNOV, E. M., C. VON MERING, I. LETUNIC, D. TORRENTS, M. SUYAMA, R. R. COPLEY, G. K. CHRISTOPHIDES, D. THOMASOVA, R. A. HOLT, G. M. SUBRAMANIAN, H. M. MUELLER, G. DIMOPOULOS, J. H. LAW, M. A. WELLS, E. BIRNEY, R. CHARLAB, A. L. HALPERN, E. KOKOZA, C. L. KRAFT, Z. LAI *et al.* 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* **298**: 149-159.

ZHANG, J. y M. NEI. 1996. Evolution of Antennapedia-class homeobox genes. *Genetics* **142**: 295-303.

ZOU, S., S. MEADOWS, L. SHARP, L. JAN y Y. JAN. 2000. Genome-wide study of aging and oxidative stress response in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **97**: 13726 - 13731.

ANEXO 1. PROTOCOLOS

PROTOCOLOS

1. PROCEDIMIENTOS BÁSICOS	181
1.1. Medios de Cultivo	181
1.2. Soluciones Estoc	182
1.3. Electroforesis en Gel de Agarosa	183
1.4. Digestión con Enzimas de Restricción	184
2. OBTENCIÓN DE EMBRIONES DE DROSOPHILA	185
2.1. Puesta	185
2.2. Recolección de Embriones	186
2.3. Decorionización	186
2.4. Almacenaje	187
3. EXTRACCIÓN Y PURIFICACIÓN DE ÁCIDOS NUCLEICOS	188
3.1. Extracción de DNA Plasmídico por Lisis Alcalina	188
3.2. Extracción de DNA Genómico de Adultos de Drosophila	190
3.3. Extracción de DNA de Fagos	192
3.4. Extracción de DNA a Partir de Geles de agarosa	194
3.5. Extracción de RNA	196
4. AMPLIFICACIÓN POR REACCIÓN EN CADENA DE LA POLIMERASA – PCR	197
4.1. Obtención de cDNA (RT-PCR)	197
4.1.1. Tratamiento con DNasa	197
4.1.2. Síntesis de cDNA	197
4.2. Amplificación de DNA (PCR)	199
5. CLONACIÓN DE FRAGMENTOS DE DNA	200
5.1. Obtención de DNA del Inserto	200
5.2. Ligación	200
5.3. Obtención de Células Competentes	201
5.4. Transformación	202
5.5. Selección de Colonias Transformantes	202

6. HIBRIDACIÓN DE ÁCIDOS NUCLEICOS	203
7. HIBRIDACIÓN POR TRANSFERENCIA A FILTRO – <i>SOUTHERN BLOT</i>	204
7.1. Digestión del DNA y Electroforesis de Transferencia	204
7.2. Transferencia	204
7.3. Marcage y Comprovación de las Sondas	206
7.4. Hibridación	207
7.4.1. Prehibridación e Hibridación	207
7.4.2. Lavados Posthibridación	208
7.5. Revelado e Interpretación de los Resultados	208
8. HIBRIDACIÓN <i>IN SITU</i> SOBRE CROMOSOMAS POLITÉNICOS	210
8.1. Preparaciones Cromosómicas	210
8.1.1. Siliconizado de los Cubreobjetos	210
8.1.2. Tratamiento de los Portaobjetos	210
8.1.3. Disección de las Glándulas Salivares	211
8.2. Marcage y Tratamiento de las Sondas	212
8.3. Hibridación	215
8.3.1. Prehibridación	215
8.3.2. Hibridación	215
8.3.3. Posthibridación	216
8.3.4. Revelado	216
8.4. Localización e Interpretación de las Señales de Hibridación	217
9. HIBRIDACIÓN <i>IN SITU</i> SOBRE EMBRIONES DE <i>DROSOPHILA</i>	218
9.1. Marcage de Sondas RNA	218
9.2. Fijación de Embriones	219
9.3. Hibridación	220
9.4. Lavados	221
9.5. Revelado	221
9.6. Montaje	222

10. HIBRIDACIÓN <i>IN SITU</i> SOBRE DISCOS IMAGINALES DE <i>DROSOPHILA</i>	223
10.1. Marcaje de Sondas RNA	223
10.2. Disección de Discos Imaginales	223
10.3. Fijación de Discos Imaginales	224
10.4. Hibridación	224
10.5. Revelado	224
10.6. Montaje de Discos en Glicerol	226
11. TINCIÓN CON ANTICUERPOS	227
11.1. Tinción de Embriones	227
11.1.1. Fijación de Embriones	227
11.1.2. Anticuerpo, Revelado y Montaje	227
11.2. Tinción de Discos Imaginales	228
11.2.1. Disección y Fijación	228
11.2.2. Anticuerpo	228
11.2.3. Revelado y Montaje	228
12. GENOTECA DE FAGOS LAMBDA	230
12.1. Infección de Fagos	230
12.2. Obtención de Lisados	231
12.3. Amplificación de Genotecas	232
12.4. Rastreo	232
12.4.1. Infección y Transferencia a Filtro	232
12.4.2. Hibridación y Revelado	234
12.4.3. Selección de Fagos Positivos	234
13. SECUENCIACIÓN	235
13.1. Secuenciación de Productos de PCR	235
13.2. Secuenciación Mediante Subclonación	235
13.3. Secuenciación Aleatoria	235
14. PROGRAMAS PARA EL ANÁLISIS DE SECUENCIAS	236

1. PROCEDIMIENTOS BÁSICOS

1.1. Medios de Cultivo

LB (Luria-Bertani)

Preparación	1 g triptona
	0,5 g extracto de levadura
	0,5 g NaCl
	Enrasar a 100 mL con agua destilada
	Autoclavar

TB (Terrific Broth)

Preparación	1,2 g triptona	0,231g KH_2PO_4
	2,4 g extracto de levadura	1,254 g K_2HPO_4
	0,4 mL glicerol	10 mL agua
	90 mL agua destilada	Autoclavar
	Autoclavar	
	Dejar enfriar y mezclar las dos soluciones	

X-gal 20 mg/mL

Preparación	200 mg 5-bromo-4-cloro-3-indolil- β -galactopiranosido
	10 mL dimetilformamida (en tubo de polipropileno)
	Hacer alícuotas de 1 mL
	Guardar a -20°C .
	Usar 2 μL por mL de cultivo

Ampicilina 25 mg/mL

Preparación	250 mg Ampicilina
	10 mL agua destilada
	Esterilizar por filtración
	Guardar a 4°C .
	Usar 2 μL por mL de cultivo

1.2. Soluciones Estoc

20xSSC	3 M NaCl; 0,3 M Na ₃ C ₆ H ₅ O ₇ x 2H ₂ O
Preparación	175,32 g NaCl 88,23 g Na ₃ C ₆ H ₅ O ₇ x 2H ₂ O Ajustar a pH 7 con HCl Enrasar a 1 L, Autoclavar
10xPBS	1,3 M NaCl, 0,07 M Na ₂ HPO ₄ x12H ₂ O, 0,03 M NaH ₂ PO ₄ H ₂ O
Preparación	75,972 g NaCl 25,07 g Na ₂ HPO ₄ x12H ₂ O 4,14 g NaH ₂ PO ₄ H ₂ O Enrasar a 1 L, Autoclavar
EDTA 0,5 M	
Preparación	186,1 g etilendiamintetraacetatox2H ₂ O 800 mL agua destilada Añadir unos 20 g de NaOH Ajustar a pH 8 Enrasar a 1 L, Autoclavar
TrisHCl 1M	
Preparación	121,1 g Tris 800 mL agua destilada Añadir unos 65 mL de HCl 35% Ajustar al pH deseado (7, 7,5 ó 8) Enrasar a 1 L, Autoclavar
SDS 20%	
Preparación	20 g SDS 100 mL agua destilada Calentar a 68°C

1.3. Electroforesis en Gel de Agarosa

La comprobación del estado del DNA, la cuantificación y la separación de fragmentos se realizan mediante electroforesis horizontal en minigeles de agarosa. Los geles se hacen a una concentración de agarosa del 0,7 % en 1xTAE, excepto cuando se quieren separar bandas menores a 2 kb que se utiliza 1 ó 1'5%. Las electroforesis se realizan a 70-80V, controlando la migración de las bandas mediante los colorantes azul de bromofenol y cianol xileno contenidos en el tampón de carga. Los geles se tiñen una vez finalizada la electroforesis para no interferir con la migración del DNA. La tinción se realiza sumergiendo los geles durante 20 minutos en una solución de bromuro de etidio en H₂O destilada. El DNA se visualiza en un transiluminador de 312 nm y se fotografía mediante Polaroid (película Polaroid 667, 3.000 ASA) o cámara digital.

Tampón de electroforesis	40mM Tris, 1mM EDTA pH 8, 1M àcid acètic
1xTAE(Tris-acetate-EDTA)	

Preparación 50xTAE	242 g Tris base
	100 mL EDTA 0,5 M
	50 mL àcid acètic
	750 mL agua destilada
	Ajustar a pH 7,5
	Enrasar a 1 L

Tampón de carga	30% glicerol, azul de bromofenol, cianol xileno
------------------------	---

Preparación	1,5 mL glicerol
	3,5 mL agua destilada
	Añadir una punta de espátula de cada colorante
	Guardar a 4°C.

El azul de bromofenol migra de forma similar al DNA lineal de 300 pb, y el cianol xileno como el DNA lineal de unas 4 kb.

Patrones de peso molecular: DNA del fago λ digerido con HindIII
 100 bp DNA Ladder (GIBCOBRL®)
 pBSK circular
 pBSK lineal (2.961 pb)

1.4. Digestión con Enzimas de Restricción

Poner en un eppendorf:	0'5 μ L de enzima (10 unidades/ μ L)
	1 μ L de tampón de enzima
	X μ L de DNA
	Y μ L de agua destilada
	10 μ L totales

Mezclar, hacer un pulso e incubar a la temperatura indicada para cada enzima (Habitualmente 37°C).

Tiempo de incubación:

DNA genómico	Toda la noche
DNA plasmídicos	De 2 horas a toda la noche
DNA fagos lambda	1-2 horas
DNA BACs	3 horas

Con DNA plasmídico o de fagos se pueden realizar digestiones de comprobación rápidas incubando 1 minuto en el microondas a máxima potencia, en estas condiciones algunos enzimas producen fácilmente digestiones parciales.

2. OBTENCIÓN DE EMBRIONES DE DROSOPHILA

2.1. Puesta

1. Preparar los medios de puesta.

(Las placas con el medio se pueden guardar a 4°C varios días).

2. Precalentar las placas y poner la levadura justo antes de usar.

3. Poner la placa de medio con levadura en la Caja o Falcon de puesta (Figura A1).

Dejar a la temperatura óptima de la cepa.

4. Cambiar los medios de puesta después del tiempo deseado.

5. Los embriones se pueden recolectar inmediatamente o dejar las placas hasta un par de días en la nevera (adecuado cuando las puestas contienen muy pocos embriones).

Es conveniente empezar siempre con dos puestas de 1 hora, descartar estas puestas.

Medios de puesta:

<i>D. melanogaster</i>	Agar al 2% con zumo de manzana, y una pizca de levadura.
<i>D. pseudoobscura</i>	
<i>D. virilis</i>	Agar con zumo de manzana y una capa fina de levadura (para evitar que los huevos queden demasiado undidos en el medio).
<i>D. buzzatii</i>	Agar al 2% con ácido acético y una capa fina de levadura
<i>D. repleta</i>	(sólo ponen embriones sobre la levadura).

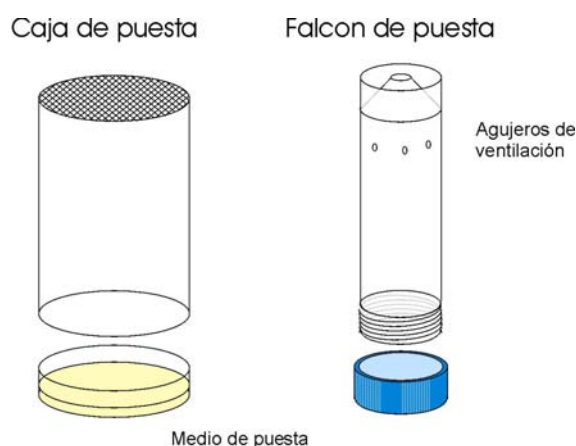


Figura A1. Cajas de puesta. Cuando no se dispone de cajas de puesta, se utilizan tubos Falcon de 50 mL con agujeros de ventilación. El medio de puesta se pone en una placa o tapón.

2.2. Recolección de Embriones

Para medios de puesta con una pizca de levadura

1. Poner el filtro Falcon (Figura A2) sobre un embudo y este sobre un erlenmeyer.
2. Coger los embriones con un pincel bajo la lupa y pasarlos al filtro.
3. Lavar con solución de lavado o con agua del grifo (No usar agua destilada).

Para medios de puesta con una capa de levadura

2. Con solución de lavado a cierta presión (utilizar una botella como la de limpiar el pHmetro) soltar la levadura del medio sobre el filtro Falcon. Lavar la levadura con los embriones sobre el filtro, la levadura se irá disolviendo y los embriones quedarán retenidos en el filtro.

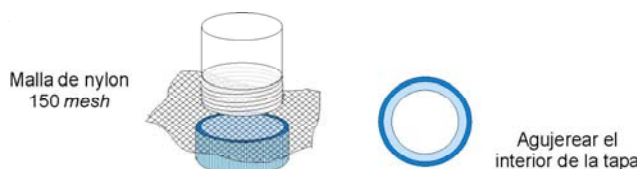


Figura A2. Filtro Falcon. El filtro se realiza con un trozo de tubo Falcon de 50 mL y un trozo de malla de nylon. Desmontar y lavar bien después de cada uso.

Solución de lavado	0,03% Triton-X100, 0,7% NaCl
Preparación 10x	300 μ L Triton-X100
	7 g NaCl
	100 mL agua destilada

2.3. Decorionización

Decorionizar los embriones mediante inmersión en lejía.

1. Introducir el filtro con los embriones en un vaso de precipitados pequeño con lejía 50%.
2. Mantener entre 2 y 5 minutos.
3. Controlar el proceso bajo la lupa.
4. Lavar con agua del grifo o con solución de lavado.
5. Secar ligeramente.

2.4. Almacenaje

Para extracción de RNA:

1. Colocar el filtro boca abajo sobre una placa de Petri.
2. Hacer caer los embriones en la placa de petri con solución de lavado.
3. Rotar lentamente la placa de forma que los embriones se acumulen en el centro.
4. Recoger los embriones con una pipeta Pasteur y pasarlos a un eppendorf
(poner la pipeta vertical para que los embriones no se enganchen a la pared).
5. Centrifugar brevemente y eliminar el líquido.
6. Guardar los embriones a -20°C .

Para hibridación in situ o tinción con anticuerpo fijar según protocolo (ver Apartados 9.2 y 11.1.1).

3. EXTRACCIÓN Y PURIFICACIÓN DE ÁCIDOS NUCLEICOS

3.1. Extracción de DNA Plasmídico por Lisis Alcalina

(Partiendo de glicerinado)

1. Sembrar en placa para obtener colonias aisladas.
Incubar toda la noche a 37°C. Guardar a 4°C.
2. Inocular un tubo con 3 mL de LB ó TB + antibiótico con una colonia aislada de la bacteria. Incubar toda la noche a 37°C en agitación.
3. Pasar 1,5 mL del cultivo a un eppendorf. (En caso de haber poco pellet se pueden
4. Centrifugar 30 segundos. añadir otros 1,5 mL de cultivo)
5. Eliminar el sobrenadante por aspiración.
6. Centrifugar 10 segundos.
7. Eliminar los restos de medio con punta amarilla.
8. Resuspender en 100 µL de **solución I** fría (4°C).
9. Vortear hasta homogeneizar.
10. Poner en hielo.
11. Añadir 200 µL de **solución II** (reciente).
12. Agitar por inversión hasta que esté transparente (lisis celular).
13. 5 minutos en hielo (no pasarse).
14. Añadir 150 µL de **solución III** fría (4°C).
15. Agitar por inversión hasta que se forme precipitado.
16. 5-10 minutos en hielo.
17. Centrifugar 5 minutos.
18. Pasar el sobrenadante a otro eppendorf, NO arrastrar sedimento.
19. Desproteinizar con fenol:cloroformo (1:1)
 - a. Añadir 1 volumen (450 µL) de fenol:cloroformo.
 - b. Agitar hasta obtener emulsión homogénea.
 - c. Centrifugar 5 minutos.
 - d. Transferir la fase acuosa (superior) a otro eppendorf.
20. Desproteinizar con cloroformo isoamílico (24:1, v/v)
 - a. Añadir 1 volumen de cloroformo isoamílico.
 - b. Mezclar por inversión.
 - c. Centrifugar 5 minutos.
 - d. Transferir la fase acuosa (superior) a otro eppendorf.

Opcional: segundo lavado con cloroformo isoamílico.

21. Precipitar con etanol absoluto

- a. Añadir 2 volúmenes (900 μ L) de etanol absoluto frío (-20°C).
- b. Mezclar por inversión.
- c. Dejar a -20°C durante 20 minutos (hasta un fin de semana).
- d. Centrifugar 10 minutos.
- e. Eliminar el sobrenadante por aspiración.

22. Lavar el sedimento con 1 volumen de etanol 70%.

23. Secar: con espaditas de papel Whatman + vacío.

24. Resuspender en 15-20 μ L de H₂O mQ.

25. Añadir 0,5 μ L de RNasa. Dar un pulso.

26. Incubar 30 minutos a 37°C.

27. Guardar a 4°C. # *Utilizar 1 μ L para el gel de comprobación.*

Solución I	50 mM glucosa, 10 mM EDTA pH 8, 25 mM Tris HCl pH 8
Preparación	0,9 g de glucosa 2 mL de EDTA 0,5 M 2,5 mL de Tris 1M Agua hasta 90 mL Ajustar pH a 8,0 Enrasar a 100 mL con agua destilada
Solución II	NaOH 0,2 M, SDS 1%
Preparación	500 μ L SDS 20% 200 μ L NaOH 10N Enrasar a 10 ml con agua destilada
Solución III	
Preparación	11,5 mL ácido acético glacial 60 mL acetato potásico 28,5 mL agua destilada

3.2. Extracción de DNA Genómico de Adultos de *Drosophila*

La extracción de DNA genómico de *Drosophila* se basa en el protocolo descrito por Piñol *et al.* (1988).

- Trabajar a 4°C.
1. Cargar el homogeneizador con 0,2 g de moscas anestesiadas en hielo.
 2. Añadir 5 mL de tampón A (a 4°C).
 3. Homogeneizar la mezcla con 10 emboladas del pistón del taladro a 2.000 rpm.
 4. Filtrar a través de una malla de nylon de 150 *mesh* sobre un tubo de 10 mL.
 5. Añadir 5 mL de tampón A.
 6. Centrifugar 10 minutos a 4.000 rpm.
 7. Decantar el sobrenadante.
 8. Lavar 3 veces con 10 mL de tampón B.
 - a. añadir 10 mL tampón B.
 - b. resuspender.
 - c. centrifugar 10 minutos a 4.000 rpm.
 - d. decantar sobrenadante.
 9. Resuspender los núcleos en 1,5 mL de tampón B.
 10. Añadir 1 mL de perclorato sódico 4 M. Mezclar bien.
 11. Añadir 1 mL de SDS 4%. Mezclar bien.
 12. Desproteinizar con 1 volumen de fenol-cloroformo (1:1).

(centrifugar cada vez 5 minutos a 4.000 rpm)
 13. Desproteinizar dos veces con 1 volumen de cloroformo-isoamilalcohol (24:1, v/v).
 14. Añadir 2,5 volúmenes de etanol absoluto frío.

Mezclar por inversión con cuidado. (Deberían verse los “hilos” de DNA.)
 15. Dejar a -20°C durante un mínimo de 30 minutos.
 16. Pasar los “hilos” de DNA con cuidado a un eppendorf de 1,5 mL.
 17. Centrifugar 10 minutos a 12.000 rpm.
 18. Lavar con etanol 70%.
 19. Secar al vacío.
 20. Añadir 50 µL de agua destilada miliQ, resuspender unas horas a 37 ó 42°C.
 21. Guardar a 4°C.

Tampón A	100 mM NaCl, 10 mM EDTA, 0'5% Triton X-100, 20 mM Tris-HCl pH 7'5
Preparación	100 mL Tampón B 500 µL Triton X100 (sacar de la nevera para descongelar)

Tampón B	100 mM NaCl, 10 mM EDTA, 20 mM Tris-HCl pH 7'5
Preparación	20 mL de NaCl 5 M 20 mL de EDTA 0'5 M 20 mL de Tris 1 M 900 mL de agua destilada Ajustar pH a 7,5 Enrasar a 1 L

3.3. Extracción de DNA de Fagos

El método utilizado para la extracción de DNA de fagos se basa en el *Plate Lysate Method* descrito por Sambrook *et al.* (1989).

- (Partir de un lisado previamente titulado)
1. Infectar una placa de LBosa a una concentración suficiente para producir una lisis total.
 2. Incubar toda la noche a 37°C.
 3. Añadir 3 mL de λ diluent.
 4. Dejar a 4°C toda la noche (se puede prescindir de este paso, (o un fin de semana). pero se obtiene un rendimiento menor)
 5. 1 ó 2 horas en agitación a temperatura ambiente.
 6. Recuperar el líquido en un tubo de 10 mL.
 7. Añadir 0,5 mL de λ diluent en la placa.
 8. Mantener las placas inclinadas 15 minutos a temperatura ambiente.
 9. Recuperar el resto del líquido.
 10. Centrifugar 10 minutos a 4.000 rpm a 4°C.
 11. Transferir el sobrenadante, por decantación, a un tubo nuevo.
 12. Añadir 0,5 μ L de RNasa (0,5 μ g/ μ L) y 0,5 μ L de DNasa (1 μ g/ μ L).
 13. Incubar 30 minutos a 37°C.
 14. Añadir un volumen (3 mL) de solución de precipitación de fagos.
 15. Mezclar bien e incubar 1 hora a 0°C. (Poner los tubos en un vaso de precipitados con agua y hielo, dentro de un recipiente con hielo)
 16. Repartir el contenido de cada tubo de 10 mL en 3 eppendorf de 1,5 mL.
 17. Centrifugar 10 minutos a 12.000 rpm.
 18. Eliminar completamente el sobrenadante.
 19. Resuspender el sedimento en 170 μ L de agua destilada miliQ, con ayuda del vortex.
 20. Juntar de nuevo el contenido de los 3 eppendorf correspondientes a una misma extracción.
 21. Añadir 5 μ L de SDS 10%.
 22. Incubar 5 minutos a 65°C.

23. Añadir 10 μL de NaCl 5M.
24. Desproteínizar con 1 volumen de (mismo procedimiento que en la fenol-cloroformo (1:1). extracción de DNA plasmídico.)
25. Desproteínizar 2 veces con 1 volumen de cloroformo-isoamilalcohol (24:1, v/v).
26. Precipitar con 500 μL de isopropanol frío (-20°C).
27. Dejar 1 hora a -20°C .
28. Centrifugar 15 minutos a 12.000 rpm.
29. Lavar el sedimento con etanol 70%.
30. Resuspender en 50 μL de agua destilada miliQ.

En las primeras extracciones se realizaba un tratamiento con RNasa similar al de la extracción de DNA plasmídico, este tratamiento se ha eliminado del protocolo para reducir la degradación de los fagos en las digestiones posteriores con enzimas de restricción.

λ diluent	10 mM TrisHCl pH 7,5, 10 mM MgSO_4
-------------------------------------	---

Preparación	1 mL de Tris HCl 1 M pH 7,5 1 mL de MgSO_4 1 M 98 mL de agua destilada
-------------	---

Solución de precipitación de fagos	20% w/v polietilenglicol, 2M NaCl en λ diluent
---	--

Preparación	10 g PEG8000 (20%) Disolver en λ diluent (hace muchas burbujas) Dejar reposar 5,844 g NaCl Enrasar a 50 mL en λ diluent
-------------	---

3.4. Extracción de DNA a Partir de Geles de Agarosa

El DNA de bandas de restricción o de productos de PCR se ha purificado mediante extracción a partir de gel de agarosa. Para fragmentos de > 500 pb se ha utilizado el Kit GeneClean[®] (BIO101 Inc.) basado en la utilización del GLASSMILK[®], o la variante Kit GeneClean SPIN[®] si se quería secuenciar el producto. Para fragmentos < 500 pb se ha utilizado el kit Wizard[®] PCR Preps DNA Purification System (Promega) basado en la utilización de resinas.

1. Separar la muestra en un gel de agarosa (preferiblemente al 0,7 %).

Teñir el gen con bromuro de etidio y fotografiar (Cuidado con el transiluminador)

2. Cortar la banda. Poner en un eppendorf (GeneClean) o tubo de 5 mL (Wizard).
3. Añadir 3 μ L de NaI 6 M por mg de agarosa.
4. Dejar 5 minutos en un baño a 55°C, o hasta que se deshaga la agarosa.

Kit GeneClean[®]

5. Resuspender bien el GLASSMILK[®] con ayuda del vortex, si está sólido o muy denso añadir agua destilada.
6. Poner 5 μ L de GLASSMILK[®]. Resuspender bien.
7. Dejar 20 minutos en hielo. Mezclar cada 5 minutos.
8. Centrifugar 1 minuto.
9. Lavar el sedimento 3 veces con 500 μ L de Solución NEW.
10. Secar el sedimento con “espaditas”, y 2 minutos al vacío.
11. Añadir 7-15 μ L de agua miliQ y resuspender bien.
12. Incubar 5 minutos a 55°C.
13. Centrifugar 1 minuto.
14. Pasar el sobrenadante a otro eppendorf. (Controlar que no quede GLASSMILK[®])

Opcional: Para aumentar el rendimiento, repetir pasos 11 a 14 con un volumen menor.

Kit GeneClean SPIN[®]

8. Transferir la suspensión a una columna. Centrifugar 1 minuto.
 10. Centrifugar para eliminar el resto del líquido. Dejar el tubo abierto 2 minutos.
 11. Añadir 20-30 μ L de agua miliQ precalentada a 55°C.
 12. Centrifugar sobre un eppendorf nuevo.
- # *Opcional*: volver a pasar el agua por la columna para aumentar el rendimiento.

Wizard[®] PCR Preps DNA Purification System

5. Resuspender bien la resina en el vórtex
 6. Añadir 1 mL de resina. Mezclar bien.
 7. Masar la mezcla por una Wizard Minicolumn con ayuda de una jeringa.
 8. Hacer dos lavados de la columna con 3 mL de isopropanol (con la jeringa).
 9. Pasar la minicolumna a un eppendorf de 1,5 mL.
 10. Centrifugar 2 minutos para eliminar el líquido de la columna.
 11. Dejar secar a temperatura ambiente entre 5 y 15 minutos.
 12. Transferir la minicolumna a un eppendorf nuevo
 13. Añadir 25-50 µL de agua miliQ precalentada a 50-55°C sobre la resina
y esperar 2 minutos.
 14. Centrifugar 1 minuto.
- # Opcional:* volver a pasar el agua por la columna para aumentar el rendimiento.
15. Descartar la columna.
 16. Guardar a 4°C.

3.5. Extracción de RNA

Todo el proceso de extracción debe hacerse con material libre de RNasa y con guantes. Cambiar los guantes regularmente o cuando toquemos alguna fuente de contaminación con RNasa. Usar el reactivo TRIZOL y el cloroformo en la campana de gases. Todos los volúmenes de reactivos se dan para muestras de 50-100 mg de tejido.

Homogeneización: 1. Añadir a la muestra 1 mL de TRIZOL. Homogeneizar con la pipeta de 1 mL o con una varilla autoclavada.

(Especialmente los embriones son difíciles de homogeneizar, pero aunque parezcan bastante enteros siempre se obtiene RNA).

Separación de fases: 2. Incubar las muestras 5 minutos a temperatura ambiente (entre 15 y 30°C).

3. Añadir 200 µL de Cloroformo (SIGMA C2432).

4. Agitar vigorosamente durante 15 segundos.

5. Incubar a temperatura ambiente 2 ó 3 minutos (o más).

6. Centrifugar a $\leq 12.000x$ g a 4°C durante 15 minutos.

Precipitación del RNA: 7. Transferir la fase acuosa (superior) a otro eppendorf.

8. Añadir 500 µL de isopropyl alcohol.

9. Mezclar por inversión.

10. Incubar 15 minutos a -20°C.

11. Centrifugar a $\leq 12.000x$ g a 4°C durante 15 minutos.

Lavado del RNA: 12. Eliminar el sobrenadante.

13. Añadir 1 mL de etanol 75%.

14. Mezclar con el Vórtex (para soltar el pellet).

15. Centrifugar a 7.500x g a 4°C durante 5 minutos.

Redisolución del RNA: 16. Eliminar el sobrenadante y secar el pellet al aire o en la cámara de vacío 1 ó 2 minutos.

17. Disolver en 50 µL de DEPC-H₂O. Incubar a 55°C durante 10 minutos.

18. 5 minutos en hielo.

19. Guardar a -20°C.

Protocolo para la “Transcriptor Reverse Transcriptasa” (Roche).

1. Poner en un eppendorf: 10 μ L de RNA tratado con DNasa.
4 μ L tampón de enzima 5x
2 μ L dNTP Mix
2 μ L Random primers
0,5 μ L RNase inhibitor
0,5 μ L transcriptasa (20 U/ μ L)
1 μ L H₂O miliQ estéril

2. Hacer un pulso.

3. Incubar 10 minutos a 25°C.

4. Incubar 30 minutos a 55°C.

5. Incubar 5 minutos a 85°C.

6. Poner en hielo.

#Si no se usa inmediatamente, guardar a -4°C unas horas o a -20°C.

4.2. Amplificación de DNA (PCR)

Las reacciones de PCR se realizan en un volumen de 25 μL si se trata de una comprobación o de 50 μL cuando se quiere obtener cantidad para usos posteriores (en este caso hay que duplicar los volúmenes indicados). Habitualmente se amplifica DNA o cDNA, pero también se puede amplificar directamente fagos en SM sin purificación previa del DNA.

Poner en un eppendorf: 0,5 ó 1 μL de muestra
 (de 0,5 ó 0,2 μL) 2,5 μL de tampón de enzima
 0,75 μL de MgCl_2 50mM*
 2,5 μL de mezcla de nucleótidos
 0'1 μL de cada oligonucleótido (100 pmol/ μL)
 ó 2,5 μL (4 pmol/ μL)
 0,1 μL Taq polimerasa
 Hasta 25 μL de agua miliQ estéril

*Si no está incluido en el tampón de enzima.

Incorporar siempre una reacción de resultado conocido como control positivo y una reacción sin DNA como control negativo.

Las condiciones de amplificación se especifican en la Tabla siguiente.

Tabla A1. Condiciones de la amplificación por PCR

	Nº de ciclos		Temperatura	Tiempo
Desnaturalitzación	1		94°	5' genómico 1' cDNA
Amplificación	30	Desnaturalitzación	94°	30''
		Apareamiento	Según reacción	30''
		Polimeritzación	72°	1' 30''
Finalización de reacciones	1	Polimeritzación	72°	7'

5. CLONACIÓN DE FRAGMENTOS DE DNA

Los fragmentos de DNA provenientes de digestión con enzimas de restricción se han clonado en pBSK y los productos de PCR en pGEM-T o pGEM-Teasy.

5.1. Obtención de DNA del Inserto

Para clonar el DNA procedente de digestión o de PCR se purifica según los protocolos del Apartado 3.4. El DNA obtenido se resuspende en el volumen mínimo que permite el método de purificación utilizado.

5.2. Ligación

- En pBSK:
1. Poner en un eppendorf: 7 μ L DNA inserto
1 μ L DNA vector linealizado
 2. Incubar 5 minutos a 65°C.
 3. 10 minutos en hielo
 4. Añadir: 1 μ L de tampón de Ligasa 10x
1 μ L de T4 DNA ligasa (1 unidad/ μ L)
 5. Incubar toda la noche a 14-16°C,
ó 3-4 horas a temperatura ambiente.

En pGEM-T o pGEM-Teasy:

- Tampón 10x
1. Poner en un eppendorf:
7 μ L DNA inserto
1 μ L vector
1 μ L tampón ligasa 10x
1 μ L ligasa
 2. Incubar a 4°C durante toda la noche.
- Tampón 2x
1. Poner en un eppendorf:
3 μ L DNA inserto
1 μ L vector
5 μ L tampón ligasa 2x
1 μ L ligasa
 2. Incubar 2 horas a temperatura ambiente

5.3. Obtención de Células Competentes

El método de obtención de células competentes se basa en que el tratamiento con cationes bivalentes y a baja temperatura desestructura la membrana celular favoreciendo la incorporación de DNA plasmídico.

1. Inocular 3 mL de medio LB con una colonia de DH5 α f⁻.
2. Crecer a 37°C en agitación durante toda la noche.
3. Inocular 20 mL de medio LB con 200 μ L del cultivo anterior.
4. Incubar a 37°C en agitación hasta que la OD₅₅₀ esté a 0,5 (unas 3 horas).
5. Mantener el cultivo en hielo durante 15-20 minutos.
6. Repartir el cultivo en 2 tubos de 10 mL.
7. Centrifugar 10 minutos a 4.000 rpm a 4°C.
8. Descartar el sobrenadante.
9. Resuspender las células en $\frac{1}{2}$ del volumen inicial de CaCl₂ 50 mM frío y estéril (5 mL por tubo).
10. Mantener 1 hora en hielo.
11. Centrifugar 10 minutos a 4.000 rpm a 4°C.
12. Descartar el sobrenadante.
13. Resuspender, muy suavemente, las células en $\frac{1}{10}$ del volumen inicial de CaCl₂ 50 mM (1 mL por tubo).
14. Mantener 1 hora en hielo.
15. Añadir un 15% de glicerol (176 μ L por tubo) .
16. Hacer alícuotas de 50 ó 100 μ L.
17. Guardar a -70°C hasta su utilización.

5.4. Transformación

La transformación se ha realizado por el método de choque térmico. Rutinariamente se han utilizado 5 µL de ligación para transformar una alícuota de células competentes, pero en casos en que es parte de poco DNA o de mala calidad se utilizan los 10 µL.

1. Descongelar una alícuota de células competentes en hielo durante 10 – 15 minutos.
2. Añadir 5-10 µL de ligación, mezclar con suavidad.
3. Dejar 20 minutos en hielo.
4. Someter a un choque térmico a 42°C durante 2 minutos.
5. Añadir 500 µL de medio LB precalentado a 37°C, mezclar suavemente.
6. Incubar 1 hora a 37°C.
7. Sembrar, con el asa de Digralsky, 200 µL en una placa de LB Amp-XGal.
8. Incubar a 37°C con agitación durante toda la noche.

5.5. Selección de Colonias Transformantes

La selección de colonias transformantes se realiza por α -complementación mediante la observación de la actividad β -galactosidasa en placas con Ampicilina y X-Galactosa. Las colonias transformantes aparecen de color blanco, pero en el caso de insertos pequeños (menores de 2 kb) el plásmido puede producir proteína parcialmente funcional dando lugar a colonias azuladas que pueden resultar difíciles de distinguir de las colonias con plásmidos no recombinantes.

Una vez identificados los posibles positivos se resiembran en placa y se realiza un cultivo de noche para extraer DNA plasmídico y analizar los clones con enzimas de restricción para verificar que contienen el inserto deseado. La comprobación de los clones se realiza por digestión con los mismos enzimas de la clonación, en el caso de pBSK, con EcoRI para pGEM-Teasy, o con otros enzimas según la información previa disponible. De los clones seleccionados se guardan dos alícuotas de cultivo de 750 µL en glicerol al 30% a -20°C.

6. HIBRIDACIÓN DE ÁCIDOS NUCLEICOS

Las técnicas de Hibridación de ácidos nucleicos se basan en la hibridación específica de un DNA o RNA marcado con la secuencia o secuencias complementarias existentes en una determinada muestra. El esquema de la Figura A3 muestra los pasos básicos de cualquier proceso de hibridación. Se denomina hibridación *in situ* cuando la muestra consiste en una sección de tejido, células o cromosomas preservados morfológicamente (Wilkinson 1992).

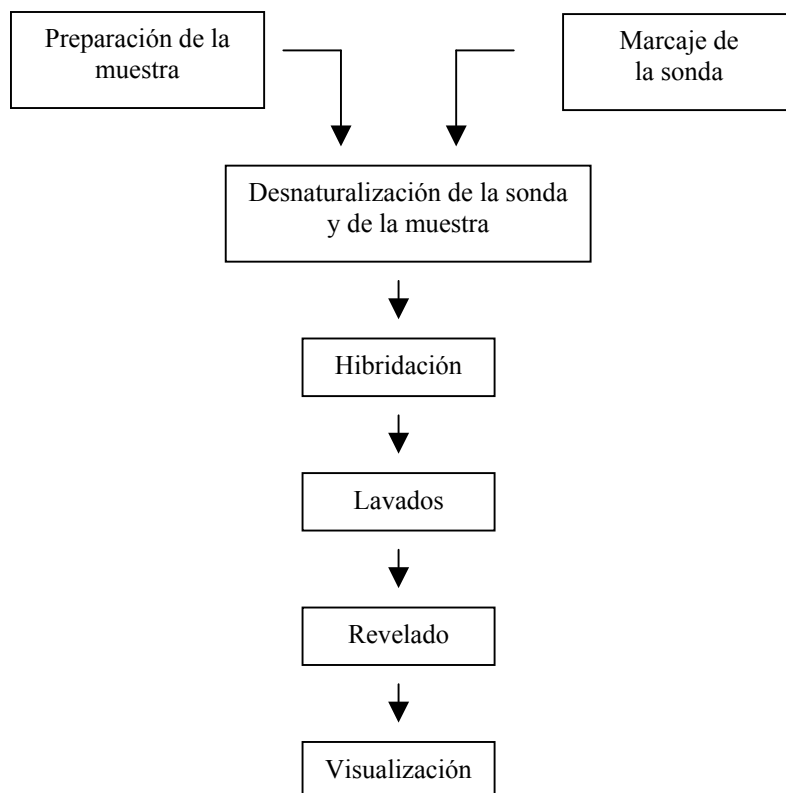


Figura A3. Esquema de los principales pasos de las técnicas de hibridación.

7. HIBRIDACIÓN POR TRANSFERENCIA A FILTRO – *SOUTHERN BLOT*

La técnica de transferencia descrita por Southern (1975) permite la localización de una determinada secuencia en DNA genómico, clonado o producto de PCR. El DNA se corta con enzimas de restricción y los fragmentos obtenidos se separan mediante un gel de agarosa. El DNA se desnaturaliza *in situ* y se transfiere a un filtro donde se hibrida con una sonda marcada y se revela de manera que podemos localizar las posiciones de las bandas complementarias a la sonda.

7.1. Digestión del DNA y Electroforesis de Transferencia

Realizar las digestiones en un volumen total de 25 μ L (ver Apartado 1.3. para proporciones y tiempos de incubación). Comprobar 5 μ L de digestión en un minigel. Realizar la electroforesis de transferencia en una cubeta larga, correr el gel despacio y fotografiarlo con una regla fluorescente al lado del marcador de peso molecular, situando el cero a la altura de los pocillos. Irradiar la mitad superior del gel durante un minuto para fragmentar el DNA de gran tamaño y facilitar su transferencia.

7.2. Transferencia

Se utiliza el protocolo de transferencia por capilaridad a membrana de nylon en condiciones neutras descrito en Sambrook *et al.* (1989).

1. Desnaturalizar el gel 1 hora en solución de desnaturalización fresca, a temperatura ambiente y agitación suave.
2. Neutralizar el gel en solución de neutralización durante 1 hora y en agitación suave.
3. Montar la transferencia:
 - a. Lavar bien con agua destilada todo el material incluidos los gastos.
 - b. Cortar un trozo de filtro de nylon y tres de papel Whatman del tamaño del gel y un trozo de papel Whatman largo.
 - c. Poner 200 mL de 10xSSC nuevo en la fiambarrera de transferencia.
 - d. Colocar la placa de la transferencia perpendicular sobre la fiambarrera.
 - e. Humedecer el papel Whatman largo en el 10xSSC y colocar perpendicular a la placa con los extremos sumergidos en el 10xSSC.
 - f. Eliminar las burbujas del papel Whatman con una pipeta de vidrio.

- g.** Poner en la fiambarrera hasta 1 L de 10xSSC reutilizado.
 - h.** Colocar el gel boca abajo sobre el papel Whatman (el DNA queda hacia arriba).
 - i.** Eliminar el aire con la pipeta de vidrio.
 - j.** Preparar 100 mL de 2xSSC, poner en una fiambarrera pequeña.
 - k.** Humedecer el filtro de nylon en el 2xSSC, y colocarlo sobre el gel.
 - l.** Eliminar el aire con la pipeta de vidrio.
 - m.** Humedecer los tres papeles Whatman y colocar uno a uno sobre el filtro.
 - n.** Eliminar el aire con la pipeta de vidrio.
 - o.** Poner parafilm alrededor del gel.
 - p.** Poner 8-10 cm de papel absorbente (servilletas) sobre los papeles Whatman.
 - q.** Colocar un peso sobre el papel absorbente.
 - r.** Dejar transferir por capilaridad durante unas 16 horas.
- 4.** Retirar el peso y los papeles absorbentes.
 - 5.** Invertir los papeles Whatman con el gel y el filtro en su interior.
(Queda el gel encima y el filtro debajo)
 - 6.** Marcar los pocillos en el filtro y
marcar una esquina para reconocer la posición del DNA.
 - 7.** Lavar el filtro en 2xSSC (guardado del día anterior) y dejar secar.
 - 8.** Irradiar el filtro 3 minutos por cara para fijar el DNA.
 - 9.** Teñir de nuevo el gel para comprobar como ha ido la transferencia.

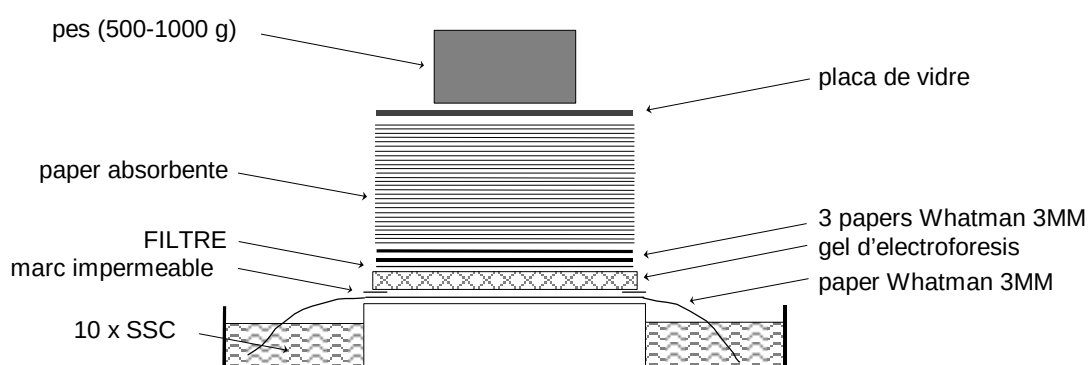


Figura A4. Esquema del dispositivo de transferencia.

Solución de desnaturalización	1,5M NaCl, 0,5M NaOH
Preparación	5 g NaOH 21,9 g NaCl 250 mL agua destilada
Solución de neutralización	1,5M NaCl, 0,5M TrisHCl pH 7,5
Preparación	500 mL TrisHCl 1 M pH 7,5 87,7 g NaCl Enrasar a 1 L con agua destilada Autoclavar

7.3. Marcaje y Comprobación de la Sonda

La sonda se marcan de forma no radioactiva con digoxigenina (digoxigenin-11-dUTP) utilizando el kit “DIG DNA Labeling and Detection Kit” (ROCHE), siguiendo el protocolo de “random primer”. Para una mayor eficiencia en el marcaje y para evitar posibles inespecificidades, se realiza una extracción del DNA a partir de gel de agarosa (ver Apartado 3.4.), eliminando así los restos de PCR o el DNA del vector.

1. Desnaturalizar 15 μ L de DNA hirviendo 10 minutos.
2. Poner en hielo.
3. Añadir en hielo 2 μ L de hexanucleótidos
 2 μ L de dNTP de marcaje
 1 μ L de Klenow
4. Mezclar bien y hacer un pulso.
5. Incubar a 37°C de 1 a 24 horas.
6. Añadir 2 μ L de EDTA 0,2 M pH 8 para parar la reacción.
7. Añadir 2,5 μ L de CILi 4 M y 75 μ L de etanol absoluto frío.
8. Precipitar 2 horas a -20°C.
9. Centrifugar 15 minutos.
10. Lavar el pellet con 50 μ L de etanol 70% frío.
11. Secar el pellet.
12. Resuspender en 50 μ L de agua miliQ.
13. Guardar a -20°C.

Comprobar el macaje mediante “dot blot”.

1. Hacer 4 diluciones de la sonda (dil -1: 1 μ L de sonda + 9 μ L de agua;
dil -2: 1 μ L dil -1 + 9 μ L de agua...)
2. En un fragmento de filtro de nylon poner 1 μ L de cada dilución de sonda y de DNA control.
3. Dejar secar
4. Fijar el DNA al filtro.
5. Revelar el filtro. (ver Apartado 6.5.) (5 mL solución de anticuerpo)
(2 mL solución de coloración)

7.4. Hibridación

7.4.1. Prehibridación e Hibridación

1. Incubar el filtro 1-2 horas en 10 mL de solución de hibridación,
a la temperatura de hibridación (42°C para homólogas y 37°C para heterólogas)
2. Mientras desnaturalizar la sonda: Hervir 5 minutos
Poner en hielo
3. Poner la sonda desnaturalizada en 10 mL de solución de hibridación
precalentada a temperatura de hibridación.
Las sondas reutilizadas se desnaturalizan 10 minutos a 65°C.
4. Descartar la solución de prehibridación y
añadir la solución de hibridación con sonda.
5. Incubar a la temperatura de hibridación durante mínimo 16 horas para DNA
genómico ó 12 horas para DNA de clones.

	50% formamida, 0'02% SDS,
Solución de hibridación	0'1% N-laurilsarcosina, 2'5% reactivo de bloqueo, 25% 20xSSC
Preparación	50 mL formamida 0,1 mL SDS 20% 0,1 g N-laurilsarcosina 2,5 g reactivo de bloqueo (ó 25 mL al 10%) Enrasar a 10 mL con agua destilada Guardar a -20°C

7.4.2. Lavados Posthibridación

1. Dos lavados de 15 minutos en 2xSSC 0'1% SDS a temperatura ambiente.
2. Dos lavados de 20 minutos en 0'1xSSC 0'1% SDS a 68°C (homólogo) ó 50°C (heterólogo).

2xSSC 0'1% SDS

Preparación	100 mL 20xSSC
	5 mL SDS 20%
	Enrasar a 1 L
	Autoclavar

0,1xSSC 0'1% SDS

Preparación	5 mL 20xSSC
	5 mL SDS 20%
	Enrasar a 1 L
	Autoclavar

7.5. Revelado e Interpretación de los Resultados

1. Lavar 5 minutos en Tampón 1, en agitación.
2. Incubar 30 minutos en 100 mL de solución de bloqueo a temperatura ambiente, con agitación suave.
3. Incubar 30 minutos en 20 mL de solución de anticuerpo a temperatura ambiente.
4. Lavar 2 veces 15 minutos en Tampón 1, en agitación.
5. Equilibrar 5 minutos en Tampón 3.
6. Revelar entre 2 y 20 horas, a oscuras, en 10 mL solución de coloración recién preparada.
7. Lavar con abundante agua destilada.
8. Dejar secar.

Para estimar el tamaño de las bandas reveladas se utiliza como referencia la fotografía del gel con la regla milimetrada al lado del marcador de peso molecular.

Tampón 1	0,1 m ácido maleico, 0,15 M NaCl, pH 7,5
Preparación	11,607 g ácido maleico 8,766 g NaCl 7,5 g NaOH Ajustar a pH 7,5 Enrasar a 1 L con agua destilada

Tampón 3	
Preparación	100 mL TrisHCl 1 M 5,844 g NaCl 10,17 g MgCl ₂ Ajustar a pH 9,5 con NaOH (26-30 lentejas) Enrasar a 1 L con agua destilada

Solución de bloqueo	
Preparación	100 mL Tampón 1 1 g leche en polvo
ó	90 mL Tampón 1 10 mL Reactivo de bloqueo 10%

Solución de anticuerpo	
Preparación	20 mL Solución de bloqueo 4 µL anticuerpo (Antidigoxigenin-AP)

Solución de coloración	
Preparación	10 mL Tampón 3 200 µL NBT/BCIP Stock Solution

8. HIBRIDACIÓN *IN SITU* SOBRE CROMOSOMAS POLITÉNICOS

El protocolo de hibridación *in situ* utilizado ha sido el descrito por Montgomery *et al.* (1987) con pequeñas modificaciones.

8.1. Preparaciones Cromosómicas

Para la obtención de preparaciones cromosómicas es necesario en primer lugar siliconizar los cubreobjetos y tratar los portaobjetos para aumentar la adherencia del material biológico y evitar así la pérdida de material durante el proceso de hibridación *in situ*.

8.1.1. Siliconizado de los Cubreobjetos

1. Hervir ácido clorhídrico 1 N en un vaso de precipitados.
2. Colocar 50-100 cubreobjetos y moverlos con una varilla mientras hierven durante 5 minutos.
3. Aclarar los cubreobjetos con H₂O destilada.
4. Dejar secar 20 minutos.
5. Sumergir los cubreobjetos uno a uno en etanol.
6. Dejar secar.
7. En una campana de gases pasar los cubreobjetos uno a uno por una solución Repel-Silane (20 g/l dimetilclorosilano en 1,1,1-triclorometano).
8. Dejar secar dentro de la campana.
9. Lavar los cubreobjetos uno a uno en etanol.
10. Dejar secar.
11. Guardar a temperatura ambiente en una caja protegidos del polvo.

8.1.2. Tratamiento de los Portaobjetos

1. Hervir ácido clorhídrico 1 N en un vaso de precipitados.
2. Colocar los portaobjetos de 20 en 20 y dejar hervir 5 minutos.
3. Poner los portaobjetos en una cubeta con H₂O destilada.
4. Sacarlos y pasarlos uno a uno bajo un chorro de H₂O destilada.
Opcionalmente se pueden dejar toda la noche en una cubeta con H₂O destilada.
5. Secar los portaobjetos durante una hora.

6. Poner los portaobjetos en una cubeta con solución SSC Denhart en una baño a 65°C.
Dejarlos en esta solución entre 2 y 3 horas.
7. Poner los portaobjetos en una cubeta con etanol:ácido acético (3:1)
durante 20 minutos.
8. Sumergir los portaobjetos durante 2 segundos unas 10 veces en una cubeta
con etanol al 95%.
9. Dejar secar y guardarlos a 4°C.

SSC Denhart

Preparación	0,012 g polivinilpirrolidona
	0,012 g ficoll
	0,012 g albúmina de suero bovino
	90 ml 20xSSC
	600 ml H ₂ O destilada

8.1.3. Disección de las Glándulas Salivares

Pera obtener un buen material que maximice la probabilidad de éxito en la hibridación se utilizan larvas de tercer estadio crecidas a baja densidad.

1. Lavar la zona de trabajo con alcohol.
2. Colocar una gota de ácido acético al 45% (recién preparado) sobre un portaobjetos limpio.
3. Poner sobre el acético una larva de tercer estadio y diseccionarla bajo la lupa con dos pinzas: unas sujetando la larva por la parte distal y las otras por la parte proximal.
Estirar en línea recta de modo que queden a la vista las glándulas salivares.
4. Con dos alfileres finos eliminar la grasa que queda adherida a las glándulas.
5. Poner sobre un cubreobjetos siliconizado 8 µl de solución
ácido láctico:H₂O destilada:ácido acético (1:2:3).
6. Transferir las glándulas a esta gota donde se dejan 8-10 minutos.
7. Dejar caer con cuidado un portaobjetos tratado sobre el cubreobjetos.
8. Girar el portaobjetos y mirarlo al microscopio. Se han de ver los núcleos y dentro de ellos los cromosomas separados por la acción del ácido láctico.

9. Picar sobre el cubreobjetos con una punta de pipeta para romper los núcleos y separar bien los cromosomas.
10. Poner un pliegue de papel absorbente sobre el cubreobjetos y presionar con el pulgar. Hay que evitar el desplazamiento del cubreobjetos para que los cromosomas no se rompan.
11. Observar las preparaciones al microscopio, seleccionar las mejores y colocarlas en una carpeta plana que se dejará en posición horizontal toda una noche.
12. Coger un portaobjetos y sumergirlo en nitrógeno líquido hasta que este deje de burbujear.
13. Colocar el filo de una cuchilla entre una esquina del cubreobjetos y el portaobjetos y tirar hacia arriba hasta que salte el cubreobjetos.
14. Colocar el portaobjetos en una cubeta con etanol al 95% un mínimo de 15 minutos.
15. Dejar secar los portaobjetos y mirar al microscopio para seleccionar las mejores preparaciones que se guardarán a 4°C hasta su hibridación.

8.2. Marcaje y Tratamiento de las Sondas

Las sondas se marcan con dUTP-biotina mediante el método de desplazamiento de mella o “nick translation” descrito por Montgomery *et al.* (1987) cuando se trata de fagos o clones o por “random primer” en el caso de productos de PCR sin clonar.

La reacción de desplazamiento de mella se basa en la utilización de dos enzimas: DNasa I y DNA polimerasa. La DNasa I introduce cortes de cadena sencilla en la molécula de DNA. La actividad exonucleasa 5'→3' de la DNA polimerasa elimina mononucleótidos en el extremo 5' del nick. Esta enzima cataliza también la incorporación de nucleótidos marcados y no marcados a las dos cadenas del DNA. El parámetro más crítico en la reacción es la actividad DNasa I. Si la actividad es demasiado baja la incorporación de nucleótidos marcados es muy baja y las sondas serán demasiado largas. Si es demasiado alta las sondas serán demasiado pequeñas (Leitch *et al.* 1994).

1. Poner en un eppendorf: 500 ng de sonda
hasta 16 μ L de H₂O estéril
2. Añadir 2,5 μ L de tampón de la reacción de marcaje.
3. Añadir 3,75 μ L de nucleótidos no marcados.
4. Añadir 1,25 μ L de 16-biotina-UTP.
5. Añadir 1 μ L de DNasa que previamente hemos disuelto en SM gel.
6. Añadir 0,5 μ L de DNA polimerasa I.
7. Hacer un pulso.
8. Colocar en un baño que esté entre 12°C –14°C durante 90 minutos.
9. Hacer un pulso.
10. Añadir 1,5 μ L de EDTA 0,5 M.
11. Incubar en un baño de 65°C durante 5 minutos.
12. Poner en hielo y añadir 2,9 μ L de NaOAc 3 M y 58 μ L de etanol absoluto frío.
13. Agitar y guardar a -20° C.

tampón de la reacción	100 mM DTT; 100 mM Cl ₂ Mg;
de marcaje	500 mM Tris HCl pH 8; 1 mg/mL de gelatina
Preparación	100 μ L DTT 1 M
	100 μ L Cl ₂ Mg 1 M
	500 μ L Tris HCl pH 8
	100 μ L gelatina 10 mg/mL
	200 μ L agua miliQ

SM gel	6 mM NaSO ₄ x 7H ₂ O; 0,099 M NaCl;
	0,05 M Tris HCl 1 M pH 7,5; 0,1 μ g/ml gelatina
Preparación	0,58 g NaCl
	0,2 g NaSO ₄ x 7H ₂ O
	5 mL de TrisHCl 1 M pH 7,5
	1 mL de gelatina 10 mg/mL
	Enrasar a 100 mL con agua destilada

El marcaje por “random primer” de basa en la incorporación de nucleótidos marcados durante la amplificación del DNA molde por la enzima Klenow, que actúa a partir de cebadores hexanucleótidos al azar.

1. Desnaturalizar el DNA (13 μ L) hirviendo 10 minutos
2. Poner en hielo.
3. Añadir a cada reacción 2 μ L de hexanucleótidos
 1 μ L de dCTP
 1 μ L de dATP
 1 μ L de dGTP
 1 μ L de biotina dUTP
 1 μ L de Klenow
4. Incubar a 37°C durante 15-20 horas.
5. Añadir 2 μ L de EDTA pH 8 para parar la reacción.
6. Añadir 2,5 μ L de LiCl 4 M y 75 μ L de Etanol absoluto frío, para precipitar el DNA.
7. Guardar a -20°C hasta su utilización (mínimo 2 horas).

Antes de hibridar se ha de precipitar la sonda:

1. Centrifugar la sonda marcada a 4°C durante 15 minutos.
2. Eliminar el sobrenadante con micropasteurs.
3. Secar con espaditas y 2 minutos en la bomba de vacío.
4. Resuspender el DNA en 30 ó 60 μ L de solución de hibridación.
5. Vortear 1 minuto.
6. Mantener en hielo.

Solución de hibridación	2xSSC; 10% dextrán sulfato; 50% formamida; DNA de esperma de salmón
Preparación	100 μ L de 20xSSC 200 μ L dextrán sulfato al 50% 500 μ L formamida 80 μ L DNA de esperma de salmón (5mg/ml) 120 μ L agua miliQ Vortear 1 minuto y guardar a -20°C.

8.3. Hibridación

8.3.1. Prehibridación

1. Incubar las preparaciones cromosómicas 30 minutos en 2xSSC en un baño a 65°C.
2. Poner las preparaciones en una cubeta con etanol al 70% durante 5 minutos.
3. Pasar las preparaciones a otra cubeta de etanol 70% donde se dejarán otros 5 minutos.
4. Pasar las preparaciones a otra cubeta con etanol 95% durante 5 minutos.
5. Dejar secar las preparaciones cromosómicas (mínimo 15 minutos).

8.3.2. Hibridación

Se han realizado dos tipos de hibridaciones: homólogas, cuando el DNA de la sonda proviene de la misma especie que el material donde se hibrida, y heterólogas, cuando el DNA de la sonda proviene de una especie diferente.

1. Poner las preparaciones en una cubeta con solución de desnaturalización (NaOH 0,07 N) durante 2 minutos.
2. Pasar las preparaciones a una cubeta con 2xSSC durante 5 minutos.
Repetir 2 veces.
3. Poner las preparaciones en una cubeta con etanol 70% durante 5 minutos.
Repetir la operación con otra cubeta de etanol 70% y una tercera con etanol al 95%.
4. Dejar secar las preparaciones al menos 15 minutos.
5. Desnaturalizar la sonda incubándola a 65°C durante 15 minutos.
6. Hacer un pulso para que baje el líquido evaporado y mantener la sonda en hielo.
7. Añadir 20 µL de sonda a la preparación cromosómica.
8. Poner un cubreobjetos sobre la zona donde se ha añadido la sonda.
Si quedan burbujas eliminarlas con la punta de la pipeta.
9. Incubar la preparación cromosómica en una cámara húmeda un mínimo de 14 horas a 25°C (hibidaciones heterólogas) o a 37°C (hibridaciones homólogas).

solución I	10 mL Tris 50 mM pH 7,6; 4% de BSA
Preparación	Utilizar material de vidrio autoclavado 0,4 g de BSA (albúmina de suero bovino) 10 mL Tris 50 mM pH 7,6 autoclavado Hacer alícuotas de 210 µL Guardar a -20°C
solución III	
Preparación	5 µg de diaminobencidina (DAB) 1 mL de Tris HCl 50 mM pH 7,6 1 µL de H ₂ O ₂
tampón P	
Preparación	40% NaH ₂ PO ₄ H ₂ O 50 mM 60% Na ₂ HPO ₄ 12 H ₂ O 50 mM Ajustar a pH 7 con las soluciones anteriores

8.4. Localización e Interpretación de las Señales de Hibridación

Las preparaciones se han observado con un microscopio Nikon Optiphot-2. Se tomaron fotografías de las preparaciones cromosómicas utilizando las películas EKTAR-25 y Royal Gold 25 ASA de Kodak. Las señales de hibridación muestran un color rosado o azul intenso que contrasta con el color azul-liláceo de los cromosomas. El resultado de una hibridación se considera positivo cuando la señal o señales se observan en la misma posición de un determinado cromosoma en varios núcleos de la preparación. En casos de señales múltiples a veces se distingue una señal más intensa que corresponde a la señal primaria. Para poder determinar con precisión la banda o bandas donde se observan las señales de hibridación se han utilizado los siguientes mapas: la reconstrucción del mapa de *D. buzzatii* de Ruiz y Wasserman (1993) basada en el mapa de *D. repleta* (Warthon 1942) y para *D. virilis* los mapas de Gubenko y Evgen'ev (1984) y Kress (1993).

9. HIBRIDACIÓN *IN SITU* SOBRE EMBRIONES DE *DROSOPHILA*

9.1. Marcaje Sondas RNA

- 1. Extracción DNA plasmídico**
- 2. Linealización del DNA (utilizar una enzima que corte el plásmido en el extremo del inserto opuesto al promotor que queramos utilizar):**
 - a. Poner en un eppendorf:** 20 μ L DNA
5 μ L tampón de la enzima
5 μ L enzima de restricción
20 μ L H₂O miliQ
 - b. Incubar 2-3 horas a 37°C (o temperatura optima de la enzima).**
 - c. Hacer la electroforesis, cortar la banda del gel y purificar el DNA.**
 - d. Resuspender en 10 μ L de H₂O miliQ estéril.**
- 3. Reacción de transcripción**
 - a. Poner en un eppendorf:** 10 μ L DNA linealizado
2 μ L 10x transcription buffer
2 μ L DIG RNA labelling mix (10x rNTP)
1 μ L inhibidor RNasa (40 U/ μ L)
2 μ L RNA polimerasa (20 U/ μ L) T7 o SP6
3 μ L H₂O miliQ
 - b. Incubar 2 horas a 37° (o toda la noche)**
- 4. Hidrólisis**
 - a. Añadir a cada marcaje:** 5 μ L H₂O miliQ
25 μ L 2x carbonate buffer.
 - b. Incubar 40 minutos a 65°C.**
- 5. Parar la reacción y precipitar el RNA**
 - a. Añadir a cada tubo:** 50 μ L stop solution
10 μ L LiCl 4M
5 μ L tRNA (20 mg/ μ L)
300 μ L EtOH 100%
 - b. Precipitar el RNA 15 minutos a -20°C.**
 - c. Centrifugar 15 minutos a 4°C.**

d. Lavar con Etanol 70%.

6. Resuspender RNA marcado en 150 μ L de solución de hibridación.

7. Guardar a -20°C .

2x carbonate buffer	120 mM Na_2CO_3 , 80 mM NaHCO_3
Stop solution	0,2 M NaAc pH 6
Preparación	0,82 g NaAc 50 mL agua destilada Ajustar a pH 6 con ácido acético

9.2. Fijación de Embriones

1. Recolectar y decorionizar los embriones (ver Apartado 2).
2. Poner 1 mL de heptano en un eppendorf de 2 mL.
3. Recoger los embriones del filtro Falcon con el pincel bajo la lupa y pasarlos al epp.
4. Añadir 1 mL de PAF 4% (Los embriones deberían quedar en la interfase).
5. Dejar en agitación durante 20 minutos
6. Retirar el fijador (PAF- fase inferior)
7. Añadir 1 mL de metanol .
8. Agitar vigorosamente durante 1 minuto. (Se forma una emulsión que se separa lentamente mientras los embriones desvitelinizados caen al fondo del tubo.)
9. Retirar todo el líquido y substituirlo por metanol.
10. Repetir 3 veces.
11. Pasar los embriones a un eppendorf nuevo.
12. Cambiar el metanol 3 veces.
13. Guardar a -20°C de forma indefinida.

PAF 4%

Preparación	10 mL 10xPBS 70 mL H ₂ O destilada Calentar a 55°C. Añadir 4 g de paraformaldehído. (Al disolverse queda transparente) Poner 1 gota de NaOH 1M. Ajustar el pH a 7,3 a Temperatura ambiente. Enrasar a 100 mL con agua destilada. Alicuotar y guardar a -20°C.
paraformaldehído	Fluka-Chemika 76240 (CH ₂ O) _n M _r 30.03 _n

9.3. Hibridación

1. Rehidratar los embriones con baños de 5 minutos en 1 mL de metanol/PBT:
70:50, 50:50, 25:75, 10:90.
2. Lavar 3 veces con 1 mL de PBT 5 minutos.
3. Incubar 30 minutos en 1 mL de PBT.
4. Reemplazar 0,5 mL de PBT con 0,5 mL de solución de hibridación.
5. Incubar 10 minutos.
6. Prehibridar en 1 mL de solución de hibridación 1 hora a 60°C.
7. Mientras desnaturalizar la sonda: Hervir 10 minutos
Poner en hielo
8. Poner la sonda desnaturalizada en 300 µL de solución de hibridación precalentada a temperatura de hibridación.
Las sondas reutilizadas se desnaturalizan 10 minutos a 80°C.
9. Descartar la solución de prehibridación y añadir la solución de hibridación con sonda.
10. Incubar toda la noche a 60°C.

PBT PBS, 0,1% TritonX-100

Preparación	10 mL 10xPBS 90 mL H ₂ O destilada 100 µL TritonX-100
-------------	--

Solución de Hibridación

Preparación	50 mL Formamida
	25 mL 20x SSC
	200 µL tRNA (50 mg/ml)
	1 mL Heparina (50 mg/ml)
	1 mL Tween 10%
	23 mL H ₂ O destilada

9.4. Lavados

1. Retirar la solución de hibridación con sonda. (Guardar la sonda a -20°C).
2. Añadir 1,5 mL de solución de hibridación.
3. Incubar 1 hora a 60°C .
4. Retirar 1 mL de solución y 750 µL de PBT prescalentado a 60°C .
5. Incubar 5 minutos a 60°C .
6. Retirar el líquido. Añadir 1 mL de solución de hibridación/PBT 1/1 a 60°C .
7. 15 minutos en agitación a temperatura ambiente.
8. Lavar 5 veces con PBT durante 5 minutos.

9.5. Revelado

1. Poner 100 µL de anti-DIG-AP 1/2000 en PBT.
2. Incubar 2 horas a temperatura ambiente o toda la noche a 4°C .
3. Lavar 5 veces con PBT durante 5 minutos.
4. Lavar 3 veces con APB durante 5 minutos.
5. Colorear en un pocillo con 1 mL de APB + 10 µL NBT + 5 µL BCIP.
6. Parar la coloración con 3 lavados de 15 minutos con PBT.

APB 100 mM NaCl, 50 mM MgCl₂, 0,01% Tween-20,
100 mM TrisHCl pH 9,5

Preparación	100 mL Tampón 3	2 mL NaCl
	(Revelado Southern)	5 mL MgCl ₂
	10 µL Tween-20	10 mL TrisHCl pH 9,5
		10 mL Tween-20
		Hasta 100 mL agua dest.

9.6. Montaje

1. Lavar 2 veces con PBS.
2. Deshidratar los embriones con lavados sucesivos en etanol/PBS:
30:70, 50:50, 70:30 y etanol.
3. Lavar 3 veces con etanol 100%.
4. Montar con DURCUPAN
 - a. Poner una gota de medio DURCUPAN en el centro de un portaobjetos.
 - b. Colocar a cada lado medio cubreobjetos (del n° cero).
 - c. coger los embriones (con un poquito de etanol)
con una pipeta de 200 µL con la punta cortada.
 - d. poner los embriones sobre la gota de medio.
 - e. colocar un cubreobjetos de 24x, con cuidado de que no quede aire.
 - f. moviendo con cuidado el cubreobjetos podemos girar los embriones
para verlos en diferente orientación.

Guardar a -20°C mientras queramos poder girar los embriones. Siempre en horizontal.

Para endurecer el medio dejar un par de días a temperatura ambiente.

DURCUPAN ACM FLUKA

Preparación	5,4 g Resin A
	4,45 g Hardener B
	0,25 g Accelerator C
	1,109 g Plasticiser D
	Mezclar vigorosamente, alicuotar y guardar a -20°C

Para montar en GLICEROL

4. Rehidratar con lavados sucesivos en etanol/PBS:
70:30, 50:50, 30:70 y PBS.
5. Substituir el PBS por glicerol en lavados sucesivos en agitación suave.
6. Poner los embriones en glicerol 100% en un portaobjetos, colocar un cubreobjetos
(con cuidado de que no queden burbujas) y sellar con pintañas (mejor transparente).
8. Guardar a 4°C .

10. HIBRIDACIÓN *IN SITU* SOBRE DISCOS IMAGINALES DE *DROSOPHILA*

10.1. Marcaje Sondas RNA Ver apartado 9.1.

10.2. Disección de Discos Imaginales

Para obtener buen material, crecer las moscas a baja densidad.

Utilizar dos placa de tres pocillos.

1. Poner una placa sobre hielo para mantener el material frío durante la disección.
2. Poner 1xPBS en tres pocillos de cada placa.
3. Colocar las larvas del estadio deseado o las pupas a diseccionar en el primer pocillo de la placa en hielo.
4. Coger una larva o pupa y situarla en el primer pocillo de la segunda placa o en un porta para realizar la disección.

Bajo la lupa:

5. Eliminar los restos de medio adheridos a la larva o pupa y pasarla a otro pocillo.
6. Partir la larva o pupa por la mitad con dos pinzas finas (larvas)
o con tijeras de disección (pupas).
7. Coger cada mitad y dar la vuelta a la cutícula como si fuera un guante, colocando una pinza en el extremo y moviendo la parte abierta de la cutícula con la otra pinza. (si solo se quiere un tipo determinado de disco, la otra mitad se desecha directamente)
8. Eliminar con cuidado la grasa y las estructuras internas de la larva.
9. Los discos quedan unidos a la cutícula por las tráqueas, separarlos un poco para que queden accesibles pero sin soltarlos de la cutícula.
10. Pasar las mitades invertidas y limpias al tercer pocillo.
11. Eliminar los restos adheridos y pasarlas a un pocillo de la placa en hielo.

La disección debe durar un máximo de 20 minutos, para evitar la degradación del RNA.

12. Fijar según el protocolo a usar (hibridación *in situ* o tinción con anticuerpo).
(Si se ha obtenido poco material, después de fijadas se pueden juntar dos o más muestras)

10.3. Fijación de Discos Imaginales

1. Disección en PBS. Guardar discos en hielo (4°C).
2. Fijar 20 minutos con PAF 4% (en balancín).
3. Lavar 3 veces 5 minutos con PBS.
4. Fijar 20 minutos con PAF 4% 0,1% Tween.
5. Lavar 3 veces 5 minutos con PBT. Pasar a eppendorf.
6. Lavar en 70:30, 50:50, 30:70 PBT/solución de hibridación.
7. Guardar en solución de hibridación a -20°C (mínimo 3 días).

10.4. Hibridación

1. Prehibridar 1 hora en solución de hibridación a 55° (en bloque).
2. Desnaturalizar la sonda 10 minutos a 80°C.
(5 µL de sonda en 50 µL de solución de hibridación)
(Sondas reutilizadas suelen dar mejor resultado, se reduce el fondo.)
3. Mantener la sonda en hielo.
4. Descartar la solución de prehibridación.
5. Poner la solución de hibridación con sonda desnaturalizada.
Incubar toda la noche a 55°C (en bloque).

10.5. Revelado

1. Precalentar (5 minutos) 3 mL de solución de hibridación por cada muestra.
2. Quitar la sonda de cada muestra, sin que quede completamente seco
(guardar la sonda a -20°C para reutilizar).
3. Lavar 2 minutos a 55°C (en el bloque) con 1 mL de solución de hibridación precalentada.
4. Dos lavados de 15 minutos a 55°C con 1 mL de solución de hibridación precalentada.
5. Lavar 15 minutos con 50% solución de hibridación:50% PBT en balancín.
(Quitar 500 µL de sol hibridación y añadir 500 µL PBT).
6. Cinco lavados de 10 minutos en PBT (o más lavados más cortos p.e. 7 x 5 minutos).
7. Incubar con anti-DIG 1/2000 1,5 horas a Temperatura ambiente.
Usar 5 µL de anti-DIG (1/50) preabsorbido en 200 µL de PBT.
8. Cuatro lavados de 20 minutos en PBT.
9. Preparar el tampón de coloración.

10. Dos lavados de 5 minutos con tampón de coloración recién preparado.
11. Preparar la solución de coloración.
12. Calentar la solución de coloración 5 minutos a 37°C.
13. Pasar los discos del eppendorf a un pocillo.
14. Poner 500 µL de solución de coloración caliente.
15. Controlar la tinción a la lupa.
16. Parar el revelado con 3 lavados de 10 minutos con PBT.
17. Deshidratar los discos con lavados sucesivos con
30, 50, 70, 90 y 100% etanol en PBS.
18. Guardar las muestras en 100% etanol absoluto a 4°C durante toda la noche.
19. Rehidratar los discos con lavados sucesivos con
100, 90, 70, 50 y 30% etanol en PBS.
20. Lavar con PBS

Solución de Hibridación

Preparación	50 mL Formamida
	25 mL 20x SSC
	200 µL tRNA (50 mg/ml)
	1 mL Heparina (50 mg/ml)
	1 mL Tween 10%
	23 mL H ₂ O destilada
tRNA	109525 ROCHE

Tampón de coloración

Preparación	500 µL NaCl 1M (o 125 µL 4M)
	250 µL MgCl ₂ 1M
	200 µL tRNA (50 mg/ml)
	500 µL Tris 1M pH 9,5)
	50 µL Tween 10%
	3,75 mL H ₂ O destilada

Solución de coloración

Preparación	1 mL de tampón coloración
	9 μ L de NBT
	7 μ L de BCIP (llamado también XPhos)

10.6. Montaje de Discos en Glicerol

1. Después de la hibridación (o la tinción) los discos vuelven a estar en 1xPBS.
2. Sustituir la mitad del PBS por glicerol 50% (en PBS).
Mezclar bien mediante agitación suave.
3. Sustituir tres cuartos del volumen por glicerol 50% (en PBS).
Mezclar bien mediante agitación suave.
4. Sustituir tres cuartos del volumen por glicerol.
Mezclar bien mediante agitación suave.
5. Colocar una de las mitades sobre una gota de glicerol sobre un portaobjetos.

Bajo la lupa:

6. Diseccionar los discos uno a uno y colocarlos sobre una gota de glicerol en un portaobjetos nuevo.
(Se pueden levantar con las pinzas o agujas de disección o con una pipeta de 20 μ L con la punta cortada, en este segundo caso tener cuidado de lo coger otros restos)
7. Una vez diseccionada toda la muestra, colocar un cubreobjetos (con cuidado de que no queden burbujas) y sellar con pintura (mejor si es transparente).
8. Guardar a 4°C.

11. TINCIÓN CON ANTICUERPOS

11.1. Tinción de Embriones

11.1.1. Fijación de Embriones

1. Recolectar y decorionizar los embriones (ver Apartado 2).
2. Fijar los embriones según Apartado 9.2, excepto:

Fijador (preparar en el momento)	3,2 mL agua destilada
	500 μ L PBS 10X
	1,3 mL formaldehído 37%

Fijar durante 25 minutos.

11.1.2. Anticuerpo, Revelado y Montaje

1. Lavar 3 veces con PBT
2. Poner el anticuerpo (dilución según anticuerpo).
3. Incubar toda la noche a 4°C.
4. Quitar el anticuerpo (guardar para reutilizar)
5. Lavar 3 veces con PBT.
6. Poner el anticuerpo secundario biotinilado.
7. Incubar 1,5 horas en rotación.
8. Lavar 3 veces con PBT.
9. Preparar el AB complex

500 μ L PBT
5 μ L A
5 μ L B
10. Añadir el AB complex e incubar durante 45 minutos.
11. Lavar 3 veces con PBT.
12. Revelar con DAB (ver Apartado 11.2.3).
13. Lavar 3 veces con PBT.
14. Montar en glicerol.

11.2. Tinción de Discos Imaginales

11.2.1. Disección y fijación

1. Disección en PBS (ver Apartado 10.2.). Guardar los discos en hielo.
2. Fijar 20 minutos con PAF 4%.
3. Fijar 20 minutos con 500 μ L PAF 4% + 5 μ L DOC + 5 μ L Triton 10%.
4. Lavar 3 veces 20 minutos con PBT/BSA.

11.2.2. Anticuerpos

1. Incubar con **anticuerpo primario** a 4°C toda la noche (o 4h a temperatura ambiente).
Siempre en balancín.
2. Lavar 4 veces 20 minutos con PBT/BSA.
3. Incubar con **anticuerpo secundario** 2 h a Temperatura ambiente.
4. Lavar 3 minutos 20 minutos con PBT/BSA.

Al empezar el segundo lavado preincubar el AB complex

300 μ L PBT + 3 μ L A + 3 μ L B

30 minutos a temperatura ambiente

5. Incubar los discos con el AB complex 45 minutos.

11.2.3. Revelado y Montaje

1. Lavar 3 veces 10 minutos en PBT.
2. Poner 2 μ L H₂O₂ al 30% en 1 mL de DAB (0,5 mg/ml en PBS)
3. En una placa de 4 x 6 pocillos poner un “colador” en el primer pocillo,
0,5 ml de DAB con H₂O₂ en el segundo y PBT en los 4 siguientes.
4. Con una pipeta de 1 mL con la punta cortada pasar las larvas del eppendorf
al colador del primer pocillo.
5. Controlando bajo la lupa pasar el colador al DAB. Máximo 3 minutos.
Si se oscurecen muy rápidamente retirar antes.
6. Hacer un lavado rápido en el primer PBT, mover un poco en el segundo
y dejar un rato en el tercero. Pasar al cuarto PBT y pasar a eppendorf.
7. Lavar 3 veces 5 minutos en PBT.
8. Diseccionar discos y montar en glicerol.

PBT/BSA	1xPBT, 0,3% triton, 1% BSA
Preparación	200 mL PBS
	0,6 mL triton 100%
	2 g BSA

Anticuerpos secundarios

Anti-rabbit Ig, biotinylated species-specific whole antibody (from donkey) 2 ml. RPN 1004. Amersham Pharmacia Biotech.

Anti-mouse Ig, biotinylated species-specific whole antibody (from sheep) 2 ml. RPN 1001. Amersham.

12. GENOTECA DE FAGOS LAMBDA

12.1. Infección de Fagos

1. Inocular 3 mL de LB enriquecido con 30 μ L $MgSO_4$ 1 M y 30 μ L Maltosa 20% con una colonia aislada de LE392. El cultivo es adecuado durante 24 h.
2. Preparar las diluciones pertinentes de fagos en SM buffer.
3. En tubos de 5 mL mezclar 100 μ L de células (LE392) con 100 μ L de fago.
(Poner primero las células para evitar contaminaciones)
4. Incubar la mezcla 30 minutos a 37°C.
5. Calentar tubos de 5 mL a 46°C. Tantos como infecciones queramos hacer.
6. Poner 3 mL de TopAgar en cada tubito a 46°.
(Poner el TopAgar con pipeta de vidrio, hay que limpiarla al momento para evitar que el agar solidifique en su interior).
7. Precalentar a 37°C placas LB ligeramente secas.
8. Volcar los 3 mL de TopAgar en el tubo de células con fagos mezclar mediante 2 ó 3 inversiones suaves.
9. Volcar la mezcla sobre una placa de LB, repartir bien y dejar solidificar 15 minutos.
10. Incubar a 37°C durante 8-10 horas o durante la noche.

LB TopAgar

Preparación	1 g Tryptona
	0,5 g Yeast extract
	0,5 g NaCl
	0,6 g Agarosa
	Enrasar a 100 mL con agua destilada
	Autoclavar
	Guardar a 45°
	Antes de usar añadir 1 mL de $MgSO_4$ 1 M

SM buffer

Preparación	1 mL gelatina 10 mg/ml
	5 mL Tris-HCl 1 M
	0,58 g NaCl
	0,8 mL MgSO ₄ 1 M
	Enrasar a 100 mL con agua destilada
	Autoclavar

12.2. Obtención de Lisados

1. Infectar una placa de LB con el fago a concentración suficiente para obtener una lisis total (no pasarse con la concentración del fago, ya que demasiada cantidad produce una lisis demasiado rápida y poco crecimiento del fago).
2. Añadir 2,5 mL de SM buffer por placa y dejar a 4°C.
3. Poner en agitación 1 ¹/₂-2 horas.
4. Recuperar el líquido con la pipeta de 1 mL y ponerlo en un tubo de 10 mL.
5. Añadir 0,5 mL de SM buffer a cada placa.
6. Dejar las placas inclinadas 15 minutos (para que escurra el líquido).
7. Recuperar el líquido.
8. Añadir 50 µL de cloroformo (Triclorometano) en cada tubo.
9. Mezclar por inversión.
10. Centrifugar 10 minutos a 4000 rpm.
11. Recuperar el sobrenadante con pipeta pasteur.
12. Añadir 50 µL de cloroformo por tubo.
13. Mezclar por inversión.
14. Guardar a 4°.

12.3. Amplificación de Genotecas

Protocolo de Sambrook *et al.* (1989) con pequeñas modificaciones.

1. Añadir a un vial de genoteca tampón SM hasta un volumen final de 1 mL.
2. Infectar 1 mL de cultivo de células LE392 reciente y repartir en 10 placas.
(Ver infección de fagos)
3. Repartir en 10 placas (200 µL por placa).
4. Incubar a 37°C durante solo 7 horas.
5. Añadir 2,5 mL de SM buffer por placa. Dejar a 4°C toda la noche.
6. Recuperar el lisado. Juntar las placas.
7. Guardar a 4°C.

12.4. Rastreo

12.4.1. Infección y Transferencia a Filtro

Para tener una probabilidad de 0,99 de obtener un fago positivo, se rastrean unos 75.000 fagos distribuidos en 10 placas (~7.500 calvas/placa). Los cálculos se muestran en la Figura A5.

$$N = \frac{\ln(1-P)}{\ln(1-f)}$$

N - nombre de clons que cal hibridar
P - probabilitat d'obtenir un clon positiu (0'99)

$$N = \frac{\ln(1 - 0'99)}{\ln(1 - 13'5 / 2'2 \times 10^5)} = 75.045 \text{ clons}$$

$$f = \frac{\text{longitud mitjana de l'insert } (\sim 13'5 \text{ kb})}{\text{longitud del genoma } (\sim 220 \text{ Mb} = 2'2 \times 10^5 \text{ kb})}$$

Figura A5. Cálculo del numero de fagos que hay que rastrear para obtener como mínimo un positivo (Sambrook *et al.* 1989, Klug y Cummings 1999).

1. Sembrar los fagos a hibridar en placas LB, a baja densidad para poder identificar luego los fagos positivos. Crecer a 37°C.
2. Guardar las placas 1 hora a 4°C (para que el agar se endurezca).
3. Cortar filtros circulares del tamaño de las placas, uno por placa.
4. Numerar las placas y los filtros (estos últimos a lápiz), y hacer tres marcas asimétricas en la placa que servirán de referencia para posicionar los positivos.
5. Abrir la placa y colocar el filtro correspondiente con el número hacia arriba y haciéndolo coincidir con el de la placa. Dejar caer el filtro poco a poco para que no queden burbujas, haciendo que toque primero el centro.
6. Dejar transferir 3 minutos.
7. Pinchar el filtro con una aguja al lado de cada una de las tres marcas hechas en la placa, aprovechando el último pinchazo para levantar el filtro.
8. Coger el filtro con las pinzas y colocarlo boca arriba sobre papel Whatmann.
9. Dejar secar.
10. En dos cápsulas de petri de vidrio de mayor tamaño que los filtros poner dos papeles Whatmann impregnados uno con solución de Desnaturalización (recién preparada) y el otro con solución de Neutralización.
11. En una fiambarrera grande de vidrio poner otro papel Whatmann con 2xSSC.
12. Poner los filtros, siempre boca arriba y controlando que la solución no los cubra ni que se seque el papel:

5 minutos en sol. de desnaturalización
5 minutos en sol. de neutralización
15 minutos en 2xSSC
13. Dejar secar.
14. Fijar el ADN irradiando 3 minutos por cara con UV.

12.4.2. Hibridación y Revelado

Las sondas se han de probar siempre en Southern sobre DNA genómico antes de usarlas para rastrear la genoteca.

El proceso de prehibridación, hibridación y revelado de los filtros es el mismo que en el Southern, únicamente cambian los volúmenes de las soluciones.

Soluciones de lavado posthibridación	1 L de cada
Solución de bloqueo	100 mL
Solución de anticuerpo	10 mL/filtro
Solución de coloración	5 mL/filtro

12.4.3. Selección de Fagos Positivos

Una vez revelados los filtros, los fagos positivos se ven en el filtro como una manchita circular con una pequeña estela. Hacer una fotocopia de los filtros en transparencia. Se coloca la transparencia boca abajo bajo la placa y se orienta con ayuda de las marcas asimétricas hechas durante la transferencia. Así se identifican las calvas correspondientes a los positivos, en cas de duda se recuperan varias calvas. Las calvas identificadas se recuperan con una pipeta Pasteur esterilizada y se deposita el agar recuperado en un eppendorf con 500 μ L de tampón SM y 20 μ L de triclorometano. Se deja una hora a temperatura ambiente y se guarda a 4°C.

Para descartar falsos positivos y eliminar la posible contaminación con otros fagos, se vuelven a hibridar los fagos recuperados. Se infecta una placa de LB agar amb cada fago recuperado, esta vez a baja densidad, y se repite el proceso de transferencia e hibridación. Esta vez los fagos se recupera en 300 μ L de SM y 15 μ L de triclorometano.

13. SECUENCIACIÓN

Las reacciones de secuenciación se han realizado en un secuenciador automático ABI 373 A (Upgrade Stretch Kit) en el Servei de Seqüenciació de DNA de la UAB.

13.1. Secuenciación de Productos de PCR

Para secuenciar productos de PCR sin clonar, se extraer el DNA a partir de gel de agarosa (ver Apartado 3.4) y se secuencia con los mismos primers utilizados para la PCR. Cuando los productos de PCR se quieren utilizar posteriormente como sondas, se clonan en pGEM-T y se secuencian con primers universales M13.

13.2. Secuenciación mediante Subclonación

Para secuenciar los fagos lambda se realiza un mapa de restricción mediante restricciones con varios enzimas (especialmente EcoRI, BamHI y HindIII) y con ayuda de Southern blots. Los fragmentos de interés se subclonan en pBSK y se secuencian con primers universales M13. Se intenta obtener clones de un tamaño aproximado de 0,9 kb ligeramente solapantes para facilitar el ensamblaje. En caso necesario se diseñan primers específicos para completar la secuencia o verificar el ensamblaje. Las secuencias se corrigen y ensamblan manualmente con la ayuda del programa Genetool.

13.3. Secuenciación Aleatoria

Para secuenciar completamente clones BAC se realiza una subclonación aleatoria del clon (inserto + vector) con un tamaño promedio de los subclones de 1,5 kb. Se secuencian por los dos extremos (con primers universales) el número de clones necesarios para obtener una secuencia 6x. Las secuencias se ensamblan con el paquete Fred-Frap-Consed. Se realiza un edición manual del ensamblaje y se realiza una ronda de autofinish. Se completan las secuencias propuestas por este programa (con primers universales o específicos) y se unen al ensamblaje. En los dos BAC secuenciados la secuencia se ha finalizado cerrando los gaps restantes mediante PCR.

14. PROGRAMAS PARA EL ANÁLISIS DE SECUENCIAS

GENETOOL: Editor de secuencias.

FRED-FRAP-CONSED y AUTOFINISH: Ensamblaje, edición y finalización de secuencias “shotgun”.

ARTEMIS: Para la anotación de secuencias.

FASTA, BLASTN y BLASTX: Búsquedas por similitud en las bases de datos.

ClustalW: Alineamientos múltiples de secuencias nucleotídicas y aminoácidas.

mVISTA: Alineamiento de secuencias largas (>3 kb).

DnaSP y PAML: Análisis de divergencia y evolución nucleotídica.

ANEXO 2. MAPAS FÍSICOS

1. MAPAS DE LAS REGIONES *labial-abdominalA* Y *proboscipedia* EN *D. buzzatii*.

Estos mapas se han realizado a partir de los clones BAC obtenidos de cada región al rastrear la genoteca CHORI-225 de *D. buzzatii*. Los mapas se han utilizado para escoger un BAC de cada región para secuenciar.

Actualmente se dispone de un mapa físico detallado de *D. buzzatii* realizado a partir de esta misma genoteca. Los clones de la región *labial-abdominalA* se encuentran en el contig 976 de dicho mapa y la región *proboscipedia* en el 968.

1.1. Región *labial-abdominalA*

El mapa de esta región (Figura A5) se ha realizado utilizando la información obtenida de las tres sondas con las que se rastreó la genoteca (307, 310.5 y 3'abdA), dos Southern blots realizados con sondas adicionales (302.1.1 y abdA2) y las estimas del tamaño de los clones realizadas mediante electroforesis de campo pulsante de digestiones con la enzima NotI.

1.2. Región *proboscipedia*

Para realizar el mapa físico de esta región (Figura A6) se han secuenciado los extremos de los clones BAC obtenidos. De los 11 clones BAC se han obtenido 20 secuencias (de dos clones solo se ha obtenido un extremo). Con estas secuencias se realizaron búsquedas en las bases de datos (con blastn) para obtener una primera idea de la región que contenían y para poder descartar aquellas secuencias que contuvieran elementos repetitivos. Cinco secuencias contienen elementos transponibles tipo IsBu, otras cinco contienen regiones de homología con el complejo ANT-C de *D. melanogaster*, concretamente con las regiones *zen-zen2* (2), *ama-Dfd* (1) y *Scr-ftz* (2), una secuencia corresponde al gen *CG14606* (localizado en *D. melanogaster* en 84D2) y el resto de secuencias no muestran homología con regiones conocidas. Se han escogido diez secuencias, aparentemente alejadas, sobre las que se han diseñado primers para amplificar sobre el resto de clones. Han funcionado seis de las diez STS. La información de las secuencias, la presencia de estas STS y de digestiones con las enzimas NotI y AvrII han permitido realizar un mapa bastante preciso de la región. Uno de los 11 BAC iniciales no presenta ninguno de los STS ni un mapa de restricción similar al resto de clones, por lo que se trata de un falso positivo.

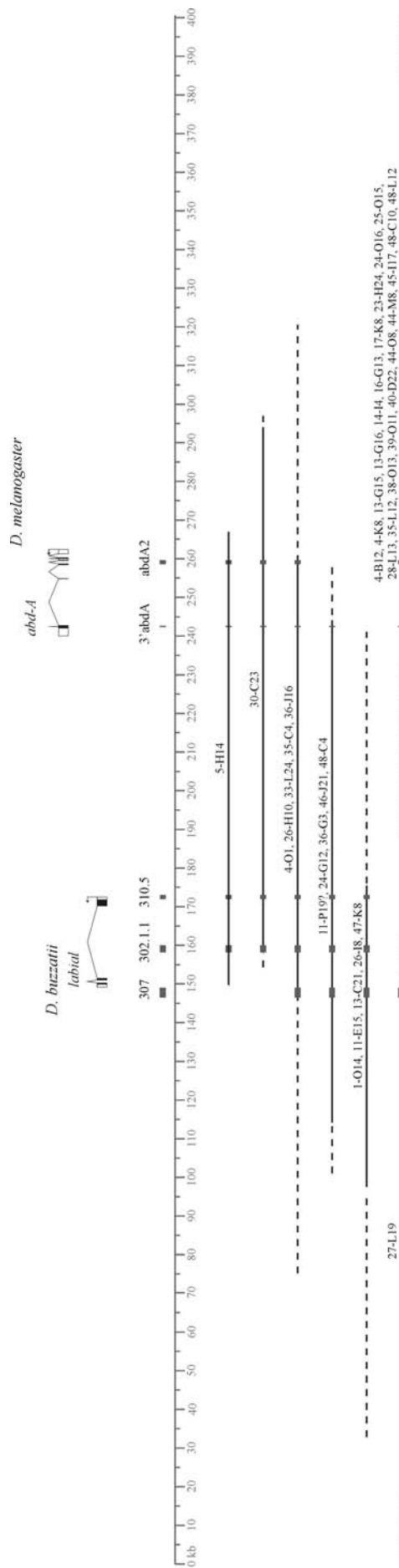


Figura A5. Mapa de la región *lab-abdA*. Arriba se muestra la posición de las sondas utilizadas y el gen al que corresponden. Las sondas de *abdA* se muestran en relación a la estructura y tamaño del gen de *D. melanogaster*, única información disponible al realizar el mapa. Abajo se muestran los clones BAC, se agrupan aquellos que contienen los mismos marcadores. Para la región del gen *lab* se han utilizado como sondas 3 clones en pBSK: pGPE307 (región 3'), pGEP 302.1.1 (fragmento del intrón 1) y pGPE320.5 (extremo 5'). Para el gen *abdA* se ha utilizado una *oligoprove* diseñada sobre secuencia de *D. virilis* (3'abdA) y un fragmento del gen (exones 2-4) amplificado a partir de DNA genómico de *D. melanogaster* (abdA2).

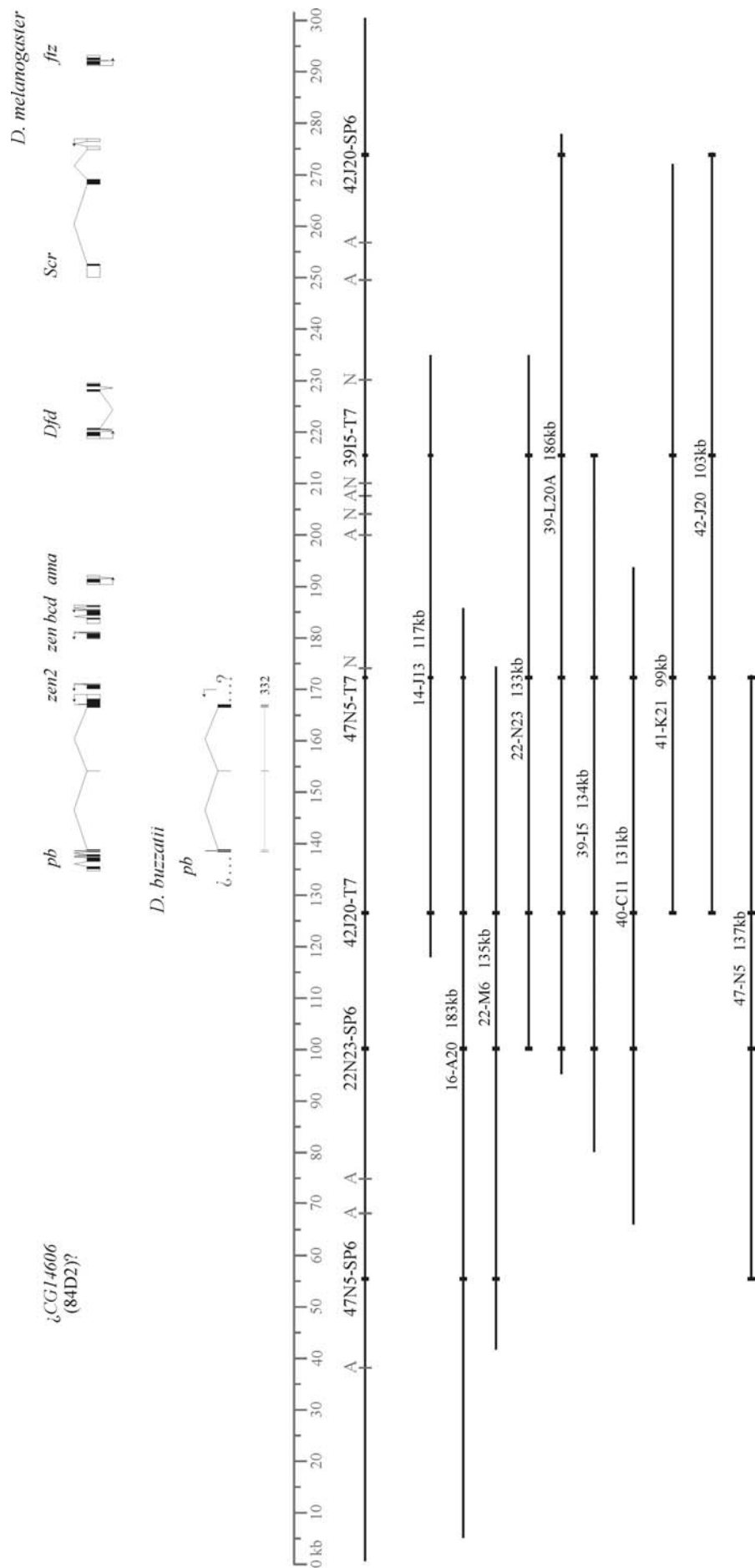


Figura A6. Mapa de la región *pb*. Arriba se muestra, como referencia, la estructura de parte del complejo ANT-C de *D. melanogaster*. Sobre los clones BAC se muestra el mapa de restricción de NotI y AvrII de la región. Los puntos negros muestran la posición de los STS. La sonda utilizada para el reastreo de la genoteca es un clon de cDNA de *D. buzzatii* (pGPE 332).

ANEXO 3. SECUENCIAS

1. NÚMEROS DE ACCESO DE LAS SECUENCIAS DEPOSITADAS EN GENBANK

Número de acceso	Especie	Gen o región
AY035293	<i>D. repleta</i>	labial gene, exons 2 and 3 and partial cds
AY035294	<i>D. mercatorum</i>	labial gene, exons 2 and 3 and partial cds
AY035295	<i>D. hydei</i>	labial gene, exons 2 and 3 and partial cds
AY035296	<i>D. buzzatii</i> (st1)	labial gene, exons 2 and 3 and partial cds
AY035297	<i>D. virilis</i>	labial gene, exons 2 and 3 and partial cds
AY313767	<i>D. virilis</i>	labial (lab) gene, cDNA fragment
AY313768	<i>D. virilis</i>	labial (lab) gene, exon 1
AY313769	<i>D. virilis</i>	labial (lab) gene, exons 2 and 3 and partial cds
AY313770	<i>D. buzzatii</i> (j19)	labial (lab) gene, cDNA fragment
AY313771	<i>D. buzzatii</i> (j19)	labial gene, exon 1
AY313772	<i>D. buzzatii</i> (j19)	labial gene, exons 2 and 3 and CDS
AY897430	<i>D. buzzatii</i> (j19)	proboscipedia (pb) mRNA, partial cds
AY897431	<i>D. buzzatii</i> (j19)	abdominal A (abdA) mRNA, partial cds
AY897432	<i>D. virilis</i>	abdominal A (abdA) mRNA, partial cds
AY897433	<i>D. repleta</i>	abdominal A (abdA) mRNA, partial cds
AY897434	<i>D. repleta</i>	labial (lab) mRNA, partial cds
AY900631	<i>D. buzzatii</i> (st1)	Drosophila buzzatii clone BAC 5H14, complete sequence; labial - abdominalA gene region
AY900632	<i>D. buzzatii</i> (st1)	Drosophila buzzatii clone BAC 40C11, complete sequence; proboscipedia gene region

