

T.D. 326

B

UNIVERSIDAD DE VALENCIA
FACULTAD DE CIENCIAS BIOLÓGICAS

**AISLAMIENTO Y CARACTERIZACION
MOLECULAR DEL GEN Antennapedia
DE *Drosophila subobscura***

Trabajo presentado por

MANUEL PEREZ ALONSO

para optar al grado de Doctor

BIBLIOTECA
FACULTAT DE C. BIOLÒGUES
VALÈNCIA

UMI Number: U607651

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607651

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Doña ROSA DE FRUTOS ILLAN, Doctora en Ciencias Biológicas y Profesora Titular del Departamento de Genética de la Universidad de Valencia,

CERTIFICA: Que D. Manuel Pérez Alonso ha realizado bajo su dirección en este Departamento el trabajo de investigación con el título "Aislamiento y caracterización molecular del gen Antennapedia de Drosophila subobscura", para optar al grado de Doctor.

Y para que así conste, en cumplimiento de la legislación vigente, expide el presente Certificado en Valencia, a catorce de noviembre de mil novecientos ochenta y nueve.

Rosa de Frutos

Fdo.: Rosa de Frutos Illán

INDICE

INTRODUCCION	1
1. Embriología de Drosophila	
1.1 Descripción del desarrollo	3
1.2 Destino de las células del blastodermo	8
2. Genética del Desarrollo de Drosophila	
2.1 El genoma de Drosophila	11
2.2 Análisis genético del desarrollo en Drosophila	12
3. Genes implicados en el desarrollo de Drosophila	
3.1 Genes de efecto materno	14
3.2 Genes de segmentación	18
3.3 Genes homeóticos	20
3.3.1 El Complejo Bithorax	21
3.3.2 El Complejo Antennapedia	23
3.4 Genes de mantenimiento	26
4. Dominios de expresión de los genes del desarrollo	27
5. Patrones estructurales en los genes del desarrollo	30
6. Unión al DNA de las proteínas con dominio homeótico	32
7. Cómo funciona el programa de desarrollo	34
8. Estructura del gen Antennapedia	39
9. Aspectos esenciales y aspectos irrelevantes de la estructura: una aproximación evolutiva	42
FIGURAS Y TABLAS	45
OBJETIVOS	54

RESULTADOS Y DISCUSION

1. Detección de fragmentos homólogos en el DNA genómico de <i>Drosophila subobscura</i>	58
2. Aislamiento y caracterización de clones genómicos de la librería de <i>Drosophila subobscura</i>	62
2.1 Identidad de los clones: localización cromosómica y secuenciación parcial	63
2.2 Localización de los segmentos que contienen los exones	67
2.3 Mapas de restricción de los clones aislados	73
3. Extensión de la región clonada: chromosomal walk	75
4. Mapa de restricción de la región clonada	78
5. Secuenciación de los exones de la región	80
5.1 Estrategia de secuenciación	80
5.2 Asignación de intrones y exones: descripción de la ORF	85
5.3 Secuencia del producto protéico: dominios constantes y dominios variables	91
5.4 Segmentos conservados en el líder y en el trailer del transcrito	98
5.5 ¿Conservación en los intrones?	101
6. Aspectos conservados y características diferenciales en la estructura del gen	103
7. Otras consideraciones	111
FIGURAS Y TABLAS	116

CONCLUSIONES	135
APENDICE: MATERIALES Y METODOS	138
REFERENCIAS	169

INTRODUCCION

INTRODUCCION

Una de las cuestiones fundamentales de la Biología es conocer cómo una sola célula (zigoto) interpreta la información que posee (su DNA junto con ciertos productos maternos) y se desarrolla en un organismo adulto. El objeto de la Genética del Desarrollo es buscar qué genes están implicados en este proceso de desarrollo y determinar qué tipo de relación hay entre la actividad de dichos genes y el fenotipo observable (el cuerpo de un animal adulto, así como los estadios intermedios). ¿Cuántos genes participan en la elaboración de cada uno de los rasgos de un organismo? ¿Cuál es la función de estos genes?

El organismo en el que se ha alcanzado el más alto grado de conocimiento sobre los mecanismos genéticos del desarrollo es *Drosophila melanogaster*. Los estudios genéticos en este organismo han empezado a revelar la lógica de algunas de las partes del programa de desarrollo. Puede decirse que los genes del desarrollo controlan la forma final del organismo.

Durante las primeras fases de este programa, el embrión de *Drosophila* adquiere la polaridad axial como resultado de la actividad de los productos de varios genes maternos (que se expresan durante la ovogénesis); el número y la polaridad de los segmentos están determinados por los productos de unos genes de expresión zigótica: los genes de segmentación. Posteriormente, estos segmentos adquieren su individualidad

por medio de la actividad de otra serie de genes de expresión zigótica: los genes homeóticos.

La importancia de los genes homeóticos radica en que sus productos conducen a las células por determinadas rutas de diferenciación. Los genes homeóticos son considerados genes selectores: sus mutaciones producen, a menudo, la transformación de un segmento en otro. Actualmente se está llegando a descifrar cómo se regulan los genes homeóticos, pero todavía no se han descrito los genes que están por debajo de la cascada de interacciones (genes efectores).

La contribución de los estudios genéticos al estado actual de conocimiento ha sido decisiva. El estudio de mutantes letales ha revelado la existencia de numerosos genes implicados en el desarrollo. Cuando su actividad se elimina por mutación, se producen defectos específicos en el desarrollo, cuyo momento y lugar de aparición se pueden determinar con precisión.

Sin embargo, es posible estudiar detalladamente el tipo de defectos que produce un determinado mutante y el momento exacto en que estos efectos comienzan a producirse (hasta llegar a la muerte del organismo); pero con esto no se explica cuál es la actividad del gen en el desarrollo normal del individuo. En cualquier caso, los análisis genéticos proporcionan una información muy valiosa, que permite abordar posteriormente los estudios a nivel molecular.

A pesar de sus limitaciones, la Genética clásica ha revelado muchos aspectos de gran interés. Uno de ellos es la

evidencia de que no hay un único gen para cada carácter, sino que no existe tal exclusividad. En la mayoría de los casos, los caracteres vienen determinados por varios genes y a este efecto se le denomina pleiotropía. Los caracteres fenotípicos que no están sujetos a pleiotropía son la excepción y no la regla.

Con los análisis a nivel genético se han detectado numerosas interacciones entre muchos genes del desarrollo. Por su parte, los estudios a nivel molecular están permitiendo analizar los mecanismos de dichas interacciones. Con el análisis de la estructura molecular y de los patrones de expresión de estos genes se ha puesto de manifiesto que la estructura y la regulación de los genes del desarrollo, especialmente la de los genes homeóticos, son muy complejas.

1. EMBRIOLOGIA DE *Drosophila*

1.1. Descripción del desarrollo

El huevo de *Drosophila* responde al tipo centrolecito (común a todos los insectos). Este huevo está rodeado por una membrana coriónica muy resistente, que es segregada por las células foliculares durante la oogénesis. El huevo se desarrolla en una larva vermiforme que, posteriormente experimenta una metamorfosis que transforma totalmente su

aspecto (*Drosophila* es un insecto holometábolo). A continuación describiremos brevemente el desarrollo de este insecto. Los tiempos de desarrollo en las distintas especies de este género son diferentes. Aquí nos referiremos a *Drosophila melanogaster*, cuando se cultiva a 25°C.

Hasta el momento de la fertilización, el oocito permanece en la metafase I de la meiosis. La fertilización se produce conforme salen los huevos maduros por el oviducto: la entrada del espermatozoide libera el proceso de la meiosis del oocito. Los pronúcleos masculino y femenino no se fusionan inmediatamente, sino que se dividen una vez. Tras ésto, se produce la fusión de ambos núcleos, dando lugar a dos núcleos diploides (PARKS, citado en WILKINS, 1986).

El huevo de *Drosophila* comienza su desarrollo con una serie de divisiones nucleares sincrónicas y muy rápidas (cada 8 min aproximadamente) que no vienen seguidas de división celular. Las nueve primeras divisiones tienen lugar en el interior del huevo, dando lugar a una nube de núcleos que migran progresivamente hacia la periferia. Las últimas cuatro divisiones se producen en el citoplasma periférico, originando un sincitio embrionario, con unos 6.000 núcleos que forman una monocapa (este número varía notablemente, entre 4.600 y 6.300). A esta fase se le denomina blastodermo sincitial. En la décima división, unos pocos núcleos se celularizan en el extremo posterior del embrión, dando lugar a las células polares, que son el origen de los futuros gametos (línea germinal). En la figura 1A presentamos un esquema de las

primeras divisiones nucleares del huevo.

Finalmente, a las 3 h de desarrollo, la mayoría de estos núcleos migran a la periferia produciéndose células por invaginación de la membrana del embrión. Así, se forma una lámina uniforme de células, ninguna de las cuales se ha formado realmente por división celular: se trata del blastodermo celular (3.5 h a partir de la fecundación). Esta fase de blastodermo celular marca la transición entre la fase controlada por productos maternos y la fase controlada por productos zigóticos. Los ciclos tan rápidos de replicación del DNA dificultan, de alguna forma, la transcripción del genoma del cigoto. Hasta el blastodermo celular, el desarrollo depende, en gran parte, de mRNAs y proteínas de origen materno, que se han acumulado en el huevo durante la oogénesis. Tras la celularización del blastodermo, las divisiones celulares se producen con menor frecuencia, de tal forma que la tasa de transcripción aumenta enormemente.

En el desarrollo, la única contribución del gameto masculino es su núcleo; el gameto femenino, además de aportar su núcleo, establece el esquema de las primeras etapas del desarrollo. El oocito no sólo tiene funciones de almacenamiento sino que, lo que es más importante, lleva una serie de instrucciones morfogénicas.

El oocito de *Drosophila* contiene mayoritariamente la yema (lípidos y proteínas) y una reserva de enzimas y nucleótidos para la replicación del DNA, así como tubulina para la formación de los husos mitóticos. Además contiene una enorme

cantidad de ribosomas y de otros RNAs. Sorprendentemente, este RNA posee una gran complejidad informacional de secuencia (alrededor del 10% del total del DNA de copia única), si bien esto parece ser simplemente una reserva de nucleótidos: sólo una pequeña parte de este RNA aporta información para el desarrollo.

Como hemos señalado, al principio los núcleos del embrión están bajo la influencia de productos maternos y es en la fase del blastodermo celular cuando se produce la activación del genoma zigótico. Cuando el genoma zigótico y el materno son normales (no mutantes), esta transición entre la fase controlada por productos maternos y la fase controlada por el genoma zigótico pasa inadvertida a nivel morfológico. Sin embargo, cuando hay alteraciones en el genoma materno, éstas se manifiestan en las primeras etapas del desarrollo. Cuando hay alteraciones en el genoma del cigoto, éstas se manifiestan posteriormente a dicha transición.

Inmediatamente después de la fase de blastodermo, se produce una serie de invaginaciones y movimientos celulares: la gastrulación. Los primeros signos morfológicos de segmentación aparecen a las 5 h, en la fase de extensión de la banda germinativa. Una descripción detallada del desarrollo posterior puede encontrarse en WILKINS (1986), así como en WIESCHAUS y NÜSSLEIN-VOLHARD (1986).

La larva del primer estadio de *Drosophila* tiene, tras la eclosión, unas 10.000 células. Durante los tres estadios larvarios se produce un considerable aumento de tamaño. La

mayoría de las células larvarias, normalmente politenizadas en mayor o en menor grado, serán destruidas durante la histólisis pupal. Sin embargo, la larva contiene unas células diploides que posteriormente darán lugar a la mosca adulta (imago): se trata de los discos imaginales y de las células histoblásticas. Ambos tipos de células se diferencian de las células larvarias circundantes durante la primera mitad de la embriogénesis, pero sólo se agruparán al final de la pupación para formar el adulto.

Las células de los discos imaginales se dividen durante el período larvario, mientras que las células histoblásticas sólo empiezan a dividirse al principio de la metamorfosis, cuando las células de los discos ya han terminado de dividirse.

La mosca adulta (más de un millón de células) surge, principalmente, de estos dos grupos de células. Las células imaginales para las estructuras externas de la cabeza y del tórax, toda la genitalia y la mayor parte de la musculatura están organizadas en los discos imaginales. La superficie externa del abdomen del imago se origina a partir de las células histoblásticas. En la figura 2B mostramos la correspondencia entre los discos imaginales y las estructuras del adulto.

Una hormona, la ecdisona, promueve el proceso de diferenciación de los discos, así como la proliferación de las células histoblásticas. Sin embargo, la ecdisona es simplemente un "revelador" de los cambios que ya se han producido en las células imaginales.

Uno de los fenómenos que han recibido mayor atención es el de la segmentación del embrión, que da lugar a la aparición de una serie de estructuras repetidas (metámeros). Más adelante nos referiremos a los genes implicados en este proceso (genes de segmentación), así como a los genes que, posteriormente, especifican la identidad de cada segmento (genes homeóticos).

1.2. Destino de las células del blastodermo

Es posible seguir el destino de muchas de las células del blastodermo a lo largo del desarrollo. Es decir, puede conocerse de dónde deriva cada una de las estructuras de la larva y de la mosca adulta. Para ello, se dispone de una serie de técnicas que permiten "marcar" una célula (o varias) y ver qué originan.

La forma más directa de analizar el destino celular es la observación directa hasta la eclosión del huevo. Otro método consiste en inducir defectos en el embrión y observar sus consecuencias. Por último, el medio más resolutivo es el mapeo por métodos genéticos.

Para estudiar el origen (en el embrión) de las estructuras de la larva y del imago el método genético apropiado es el de análisis de mosaicos ginandromorfos. Para estudiar el origen de las estructuras de la mosca adulta también ha sido de gran utilidad la técnica de recombinación mitótica. Con esta última, es posible inducir clones marcados en cualquier

momento del desarrollo y analizar su destino. A la representación del origen de cada estructura sobre la superficie del embrión (blastodermo) se le denomina mapa de destino (fate map). En la figura 1B mostramos un fate map de *Drosophila*. Ya que el blastodermo consiste en una monocapa de células, su superficie puede representarse proyectada en un plano. Si se representa el destino de cada célula sobre este plano, puede verse que el destino de cada segmento es una función de la posición en los ejes D-V y A-P.

Las regiones que originan los segmentos torácicos y abdominales de la larva se sitúan entre el 15 y el 60 % de la longitud del huevo (EL: 100 % en el extremo anterior y 0 % en el extremo posterior). Los segmentos de la cabeza se originan entre el 60 y el 70 % EL. Esto supone 1/3 de las células del blastodermo, unas 2200 células. El resto de las células blastodérmicas forman tejidos internos. Siendo que la circunferencia del embrión mide unas 72 células, se calcula que cada segmento tiene, en el blastodermo celular, un "grosor" de 3 ó 4 células (en la dimensión A-P).

Estos experimentos de fate mapping han demostrado que los mapas de destino del imago y de la larva están estrechamente relacionados. Esta concordancia implica que las células larvarias y las del imago tienen mucho en común durante el desarrollo temprano.

Ningún disco imaginal parece tener un origen estrictamente clonal, a juzgar por el hecho de que todos los discos imaginales han mostrado mosaicismo en alguna ocasión. Un

número medio de 10 células del blastodermo son el origen de cada disco imaginal. Esto se ha calculado asumiendo que la mancha masculina más pequeña que se encuentre en un ginandromorfo (para una estructura imaginal dada) corresponde a un clon originado a partir de una única célula precursora. De esta forma, puede estimarse el número de células del primordio de esa estructura en el blastodermo, en función del área relativa que ocupa en el adulto.

El análisis clonal también ha evidenciado que las células del primordio para cada estructura son prácticamente equivalentes, pudiendo todas ellas generar cualquier parte de la estructura del adulto. Las sucesivas restricciones del potencial de desarrollo (dentro del contexto fisiológico) no llegan a determinar los tipos concretos de tejido hasta el final del desarrollo embrionario.

Un clon determinado nunca ocupa, a la vez, la parte anterior y posterior de una estructura adulta. Al proceso de asignar regiones concretas a grupos de células se le llama compartimentalización (GARCIA-BELLIDO y otros, 1979). Cada compartimento es un policlón, que consiste en varios grupos completos de clones celulares.

Se considera que los compartimentos son las unidades fundamentales de expresión de los genes que especifican la identidad de los segmentos (los genes homeóticos). Cada segmento del cuerpo de *Drosophila* (bien sea el embrión, la larva o el imago) está dividido en un compartimento anterior y uno posterior. Así, por ejemplo, el segundo segmento torácico

(T2) se divide en un compartimento anterior (T2a) y uno posterior (T2p) (ver fig. 2A).

2. GENETICA DEL DESARROLLO DE Drosophila

2.1. El genoma de Drosophila

El genoma de *Drosophila melanogaster* consta de 165.000 kb, aproximadamente 1/20 del genoma humano y 50 veces el genoma de *E.coli* (RUBIN, 1988). El 75 % es DNA de copia única; el 15 % es DNA altamente repetitivo, situado mayoritariamente en los centrómeros; y el 10 % restante es DNA moderadamente repetitivo, que corresponde a elementos transponibles y secuencias dispersas. Es probable que muchas de estas secuencias dispersas (no transponibles) correspondan a patrones estructurales de productos génicos de función similar: repeticiones M, secuencias homeo box, secuencias "paired box", etc.

Cabe plantearse cual es el número de genes contenidos en el DNA de *Drosophila*. Las primeras estimaciones (5.000 genes) parecen haberse quedado cortas, pues ya se han descrito más de 3.000 genes distintos (RUBIN, 1988). Una estima razonable parece la cifra de 15.000-20.000 genes (WILKINS, 1986), lo que da un tamaño medio de 6.000-8.000 pb por gen. La cuestión inmediata es cómo se organizan los genes en este genoma. La

respuesta más cauta sea tal vez decir que no hay datos suficientes, salvo en varios complejos génicos, que han sido analizados con detalle. En los próximos apartados nos referiremos a los complejos génicos Antennapedia y Bithorax (ANT-C y BX-C).

2.2. Análisis genético del desarrollo en Drosophila

El análisis de mutantes que producen defectos en el desarrollo constituye una poderosa herramienta para el estudio de los mecanismos que dirigen el desarrollo embrionario. En el estudio del desarrollo se asume que los distintos tipos celulares son distintos porque tienen una composición (protéica) diferente. En el fondo, esto es el resultado de la expresión de genes de forma diferenciada. Buscar los mecanismos de diferenciación equivale a buscar las diferencias entre grupos de células. Sin embargo, el estudio del perfil temporal de síntesis protéica a lo largo del desarrollo (pulsos de aminoácidos marcados y electroforesis bidimensional) no ha revelado diferencias significativas. Esto se debe, probablemente, a que las proteínas que se sintetizan con una tasa suficiente para que sean detectables con estas técnicas son, en su mayoría, proteínas con funciones de mantenimiento (housekeeping).

Los métodos genéticos son mucho más sensibles a la hora de detectar la existencia de productos celulares. La existencia

de un mutante revela la producción de un producto alterado, lo que permite conocer la existencia de este producto en los individuos normales.

A la hora de interpretar los efectos de un mutante dado es imprescindible determinar con qué tipo de mutación se está tratando. No es suficiente con determinar la dominancia o recesividad del mutante. Pueden considerarse dos tipos de mutantes recesivos: hipomorfos y amorfos. Los mutantes recesivos amorfos son aquellos en los que se suprime la actividad génica normal. En los genes esenciales, estas mutaciones suelen ser letales. Normalmente se deben a una delección. A estos mutantes se les denomina mutantes por pérdida de función (loss-of-function).

Los mutantes recesivos hipomorfos son aquellos en los que queda una actividad parcial, pudiendo, en muchos casos, permitir la supervivencia del mutante.

Igualmente, se pueden considerar tres tipos de mutantes dominantes: neomorfos, antimorfos, o por haplo-insuficiencia. Un dominante neomorfo es aquél que produce una actividad génica nueva. También se les denomina mutantes por ganancia de función (gain-of-function). Por su parte, un mutante antimorfo es el que produce una actividad que compite frente a la del gen normal. La haplo-insuficiencia tiene lugar cuando el gen mutante produce una actividad génica disminuida y, a la vez, un único alelo normal no es suficiente para la producción del fenotipo normal.

Hay dos tipos de screening de mutantes implicados en la

determinación del embrión: búsqueda de mutantes zigóticos y búsqueda de mutantes de efecto materno. Los mutantes de efecto materno corresponden a genes que se transcriben durante la oogénesis; los mutantes zigóticos son los que se expresan a partir de la fase de blastodermo (previamente no hay expresión zigótica en el embrión). Para un mutante dado, la fase en la que aparecen los defectos en el desarrollo, suele indicar de qué tipo de mutante se trata (aunque no siempre).

Los genes que afectan al desarrollo se clasifican en: genes de efecto materno; genes de segmentación; y genes homeóticos. Los genes de efecto materno establecen gradientes morfogenéticos en el huevo (información posicional). Los genes de segmentación interpretan la información suministrada por los gradientes y la traducen en una serie de divisiones del embrión que son las unidades básicas de los insectos: los segmentos. Por último, los genes selectores homeóticos mantienen las diferencias entre segmentos y especifican la identidad de cada uno de ellos (SCOTT y CARROLL, 1987; INGHAM, 1988).

3. GENES IMPLICADOS EN EL DESARROLLO DE *Drosophila*

3.1. Genes de efecto materno

Son aquellos en los que el genotipo de la madre determina

el fenotipo del embrión. Entre las mutaciones que tienen efecto materno, las producidas en los genes maternos estrictos son las que aportan realmente información sobre la contribución de los genes de la madre al desarrollo del huevo. Un gen materno estricto es aquél que se expresa exclusivamente durante la ovogénesis. Sus mutaciones producen, en todos los casos, embriones que no son viables, por lo que provocan esterilidad en las hembras (ya que la letalidad de estas mutaciones se produce en el embrión). Por ello, los mutantes maternos estrictos están incluidos en un amplio grupo de genes denominados female sterility (fs).

Cabe distinguir tres tipos de genes maternos que afectan al eje anteroposterior (LEHMANN, 1988):

1. Genes anteriores. Sus mutaciones provocan que la parte anterior no aparezca y, en su lugar, se formen estructuras abdominales. Un ejemplo claro es bicoid.

2. Genes posteriores. La mutación de estos genes da lugar a que la parte posterior no aparezca y, en su lugar, se formen estructuras torácicas. Se conocen, al menos, siete genes diferentes. Un representante típico es oskar.

3. Genes terminales que, en su estado mutado, provocan una delección de los dos extremos del embrión. Se conocen, al menos, cinco de estos genes.

En la tabla 1, mostramos una lista de genes de tipo materno que afectan a la polaridad dorsoventral. Hay que señalar que, además de estos genes maternos, también hay otros que afectan a la polaridad dorsoventral del embrión, tales

como Toll, dorsal y easter (ANDERSON y NÜSSLEIN-VOLHARD, 1984).

La determinación de las posiciones en ambos ejes son procesos independientes: los mutantes que alteran la polaridad dorsoventral dejan intacto el número de segmentos.

En los mutantes maternos, los cambios que se producen afectan a todas las células del blastodermo: suelen ser cambios generalizados. Una hipótesis que explicaría estos efectos sería la producción de un gradiente a lo largo de cada uno de los dos ejes.

La existencia de un gradiente anteroposterior ha sido totalmente demostrada al estudiar la distribución de la proteína bicoid en embriones (DRIEVER y NÜSSLEIN-VOLHARD, 1988a). El gen de efecto materno bicoid es el mejor caracterizado a nivel molecular. Su aislamiento y su secuencia han sido descritos recientemente (BERLETH y otros, 1988). Se trata de un gen perteneciente al complejo Antennapedia (ANT-C) que incluye una secuencia homeobox (dominio de unión al DNA). El mRNA del gen bicoid se sintetiza sólo durante la oogénesis (gen materno estricto) en las células nodriza. Este RNA pasa al oocito a través de los canales intercelulares que unen al oocito con las células nodriza. Al entrar en el oocito, el mRNA de bicoid queda atrapado en la parte anterior, anclado en algunos componentes del citoesqueleto (codificados, al parecer por los genes exuperantia y swallow). Este RNA permanece inactivo en el huevo hasta el momento de la puesta.

Cuando la hembra pone el huevo, se inicia la traducción

del transcrito, formándose un gradiente exponencial de concentración de la proteína bicoid, con una concentración máxima en el extremo anterior y mínima en el posterior (la proteína difunde libremente en el embrión, pues todavía no hay membranas celulares). Además, se ha comprobado que todos los mutantes maternos del grupo anterior (Tabla 1) afectan al gradiente protéico de bicoid. Esto es una prueba muy significativa de que bicoid es el determinante anterior que especifica la posición de cada punto del embrión (en el eje A-P) con las distintas concentraciones protéicas producidas (DRIEVER y NÜSSLEIN-VOLHARD, 1988b; LAWRENCE, 1988). Dicho de otra forma bicoid determina las coordenadas del eje A-P en cada punto del huevo (ver fig. 1B).

También se ha demostrado la existencia de un gradiente protéico similar, aunque en el eje dorsoventral, para el producto protéico del gen dorsal (STEWART y otros, 1988). Por analogía con bicoid se supone que este gradiente especifica las coordenadas del eje D-V, aunque ésto todavía no ha sido demostrado a nivel molecular.

Surge la cuestión de cómo estos gradientes protéicos especifican las coordenadas de los ejes. Esto equivale a plantearse cuál es la actividad de estas proteínas. En el caso de la proteína bicoid, su actividad se conoce con bastante detalle. El hecho de que incluye un dominio homeótico (ver más adelante) sugiere que es un gen que puede regular a otros genes del desarrollo. En el apartado 7 resumiremos las evidencias que existen de que, en efecto, éste es el caso: el

zigoto lee e interpreta la información facilitada por los genes maternos, y la traduce en forma de expresión diferenciada de otros genes en distintas regiones del embrión.

3.2. Genes de segmentación

NÜSSLEIN-VOLHARD y WIESCHAUS (1980) llevaron a cabo un estudio muy detallado de las funciones genéticas que determinan la segmentación y la polaridad de los segmentos. Estos autores aislaron una gran cantidad de mutaciones recesivas de estos genes. Todas ellas resultan ser letales, por lo que sus efectos se estudiaron en embriones y larvas muertas. En la Tabla 2 mostramos una lista no exhaustiva de genes de segmentación. Su clasificación responde al tipo de efectos fenotípicos que producen.

En estas mutaciones, el fenotipo del embrión depende (al menos parcialmente) del genotipo del cigoto. Esto difiere de los genes a los que nos hemos referido en el apartado anterior, en los que sólo el genotipo de la madre determina el efecto fenotípico en el embrión. Por eso, los genes de segmentación son genes de expresión cigótica.

Se ha descrito un primer grupo de genes cuya mutación produce la delección de grupos de segmentos. Debido a que producen huecos se les ha llamado "gap genes". Un segundo grupo lo constituyen las mutaciones que producen la desaparición de segmentos completos alternos. Por esta

característica se les llama genes de la regla par, "pair rule genes". Por último, un tercer grupo lo constituyen las mutaciones que provocan una desaparición repetitiva de una parte de cada segmento, junto con una duplicación simétrica de lo que queda. Son los genes de polaridad de los segmentos, "segment polarity genes".

Esta regionalidad en los lugares de aparición de los defectos sugiere que la expresión de estos genes tiene lugar en zonas específicas del embrión. En efecto, utilizando sondas de estos genes para la detección de sus transcritos en los embriones (con genotipo normal), se ha comprobado que cada uno de estos genes tiene un dominio de expresión muy preciso que, en general, coincide con los lugares donde se detectan los defectos fenotípicos. Se interpreta que los genes de segmentación subdividen el embrión en dominios cada vez más reducidos.

Los genes del tipo gap se expresan en dominios que abarcan varios segmentos; los genes del tipo pair rule se expresan en bandas alternas, cuyo "grosor" es de cuatro células (el grosor de un segmento); los genes de la categoría segment polarity se expresan una parte de cada uno de los segmentos.

Por ejemplo, el gen Krüppel (tipo gap) se transcribe en una región que comprende los segmentos T1 a A1 (aunque parece ser que la proteína difunde hasta el segmento A5) (JÄCKLE y otros, 1985; GAUL y otros, 1987). El gen ftz (pair rule) se transcribe en la mitad posterior de un segmento y en la mitad anterior del siguiente segmento, produciendo 7 bandas de

expresión. Por último, un ejemplo de expresión de genes del tipo segment polarity es engrailed, que se expresa en 14 bandas que corresponden a una célula (de las cuatro que forman un segmento) de la parte posterior de cada segmento.

Los genes de segmentación son los primeros genes zigóticos que se transcriben. Los genes del tipo gap (como Krüppel) son los que primero se expresan, justo después de la división nuclear 11, cuando el embrión es todavía un sincitio; inmediatamente después se expresan los genes del tipo pair-rule (como fushi tarazu y hairy); por último, durante la división nuclear 13, se expresan los genes del tipo segment polarity, tales como engrailed y wingless (SCOTT y CARROL, 1987; INGHAM, 1988). La expresión de los genes de tipo gap y del tipo pair-rule es sólo durante estos periodos mencionados; mientras que la expresión de los genes del tipo segment polarity se mantiene a lo largo de toda la vida del animal.

3.3. Genes homeóticos

Los genes homeóticos difieren de los descritos anteriormente en que sus dominios de expresión no incluyen regiones tan amplias del embrión, sino que sus dominios son más reducidos. Las mutaciones homeóticas se caracterizan porque cambian el destino de un segmento y lo transforman en otro. El fenotipo más llamativo de las mutaciones homeóticas es el de los alelos dominantes, como Antennapedia^{73b}, que

produce la transformación de las antenas en patas mesotorácicas. Se considera que los genes homeóticos son genes selectores (GARCIA-BELLIDO y otros, 1979), implicados en la especificación de la identidad de determinados segmentos.

La expresión de los genes homeóticos empieza a detectarse a partir de la formación del blastodermo celular, después de la expresión de los genes de los tipos gap y pair-rule. Al analizar la expresión de estos genes, se observa que el embrión se subdivide en dominios de expresión para cada uno de los genes homeóticos conocidos.

Los genes homeóticos se encuentran agrupados en el genoma de *Drosophila* en dos complejos génicos: el Complejo Antennapedia (ANT-C) y el Complejo Bithorax (BX-C). Las unidades de transcripción incluidas en cada uno de estos dos complejos están enumeradas en la Tabla 3. El ANT-C incluye cinco genes homeóticos, mientras que el BX-C incluye tres. Los genes del ANT-C participan en la diferenciación de la cabeza, el primer segmento torácico (T1) y de la parte anterior del segundo segmento torácico (T2a). Por su parte, los genes del BX-C participan en la diferenciación del resto del tórax (T2p y T3) y de los 8 segmentos abdominales (A1-A8).

3.3.1 El Complejo Bithorax

El desarrollo del tórax requiere el gen Ultrabithorax. Cada uno de los segmentos del abdomen necesita la actividad de

un gen del BX-C (un gen infra-abdominal, desde iab-2 hasta iab-8). Sin embargo, el análisis de la estructura molecular de este complejo ha revelado éste sólo contiene tres regiones que codifican proteínas: Ultrabithorax, abdominal-A y Abdominal-B.

El gen Ultrabithorax (un gen del BX-C) especifica principalmente el segmento T3. El análisis detallado de las mutaciones Ubx indicó que lo que en éstas se produce no es un cambio de un segmento en otro, sino que la transformación tiene lugar en la mitad posterior de un segmento y la mitad anterior del siguiente segmento, que se transforman en estructuras anteriores. Así, formalmente, la transformación Ubx consiste en que T2p/T3a dan lugar a T1p/T2a, mientras que T3p/A1a dan lugar también a T1p/T2a. Esta observación, junto con otras evidencias, ha conducido al concepto del parasegmento como unidad primaria del desarrollo (MARINEZ-ARIAS y LAWRENCE, 1985). Un parasegmento consta de la parte posterior de un segmento junto con la parte anterior del siguiente segmento. Así, es posible definir la transformación Ubx como el cambio de los parasegmentos PS 5 y PS 6 en el parasegmento PS 4. En la figura 2B mostramos la correspondencia entre los segmentos y los parasegmentos en el embrión, en la larva y en la mosca adulta.

Todos los datos sobre la localización in situ de los transcritos de los genes de segmentación y de los genes homeóticos confirman que, en efecto, los parasegmentos son las unidades de expresión génica durante el desarrollo.

Al principio de este punto nos hemos referido al hecho de

que el BX-C sólo codifica para tres productos génicos, siendo que los estudios genéticos habían descrito un número mayor de regiones reguladoras (revisado en PEIFER y otros, 1987). Con los estudios a nivel molecular, se ha comprobado que cada región reguladora es un elemento que funciona en forma de enhancer de la transcripción, específico para cada parasegmento. Así, el gen Ubx está regulado por un enhancer (región abx/bx) que le confiere su expresión en el parasegmento 5 y por otro enhancer (región bxd/pbx) que le confiere la expresión en el parasegmento 6. Asimismo, el gen abd-A está regulado por las regiones (enhancers) iab-2 a iab-7, que confieren la expresión desde el PS 7 al PS 12 (KARCH, comun. pers.). El gen Abd-B está regulado por la región iab-8.

3.3.2. El Complejo Antennapedia

Si el BX-C especifica desde el compartimento T2p hasta el segmento A8, el ANT-C especifica el resto de los compartimentos: desde el extremo anterior hasta el compartimento T2a. La actividad normal del gen Antp es necesaria para el desarrollo del tórax, pero no para el de la cabeza. Las mutaciones por pérdida de función indican que la actividad del gen Antp se requiere en el tórax para promover el desarrollo de las patas y de las alas, pero no tiene relación con el desarrollo normal de la cabeza.

El fenotipo de la mutación dominante $Antp^{73b}$, a la que nos hemos referido al principio del apartado 3.3, corresponde a un mutante neomorfo por ganancia de función; los mutantes por pérdida de función se producen por la delección del locus $Antp$, y conducen a un fenotipo recesivo letal, que conlleva la transformación de estructuras torácicas en estructuras más anteriores (STRUHL, 1981). Se trata de una transformación en dirección anterior, mientras que el fenotipo dominante conlleva una transformación posterior.

La clonación de este complejo génico (SCOTT y otros, 1983; GARBER y otros, 1983) y su análisis molecular demuestran el agrupamiento de ocho genes en su interior. Cinco de éstos son genes homeóticos: labial (DIEDERICH y otros, 1989; MLODZIK y otros, 1988), proboscipedia (PULTZ y otros, 1988), Deformed (REGULSKI y otros, 1987), Sex combs reduced (KUROIWA y otros, 1985; RILEY y otros, 1987) y Antennapedia (SCHNEUWLY y otros, 1986; LAUGHON y otros, 1986). Los otros tres genes del ANT-C son: el gen materno bicoid (BERLETH, 1988); el gen de segmentación fushi tarazu (del tipo pair rule) (LAUGHON y SCOTT, 1984; KUROIWA y otros, 1984); y un gen implicado en la diferenciación dorsoventral, zerknüllt (RUSHLOW y otros, 1987).

La localización cromosómica de este complejo génico, así como las posiciones relativas de estos genes, vienen señaladas en la figura 3. En la figura 4 hemos dibujado un esquema de la organización de todas estas unidades de transcripción en el complejo. En este esquema puede verse la gran diferencia de

tamaño existente entre unos genes del complejo y otros. El gen *Antp* se representa sobre la línea gruesa (el genoma de *D.melanogaster*), mientras que los genes *bcd*, *Dfd*, *ftz*, *zen* y *lab* están "ampliados" en este esquema. La estructura completa de los genes *Scr* y *pb* no ha sido descrita todavía, por lo que la región que codifica para los productos protéicos (marcada en negro en los demás genes) no viene indicada.

Los experimentos de hibridación *in situ* para la detección de transcritos han permitido demostrar que la expresión génica de *Antp*, durante el desarrollo embrionario temprano, coincide con el lugar donde se observan los defectos en los mutantes recesivos de *Antp* (parasegmentos 4 y 5, con mayor intensidad en el PS 4) (MARTINEZ-ARIAS, 1986). Utilizando anticuerpos específicos, se ha comprobado que la proteína *Antp*, al igual que otros productos protéicos derivados de genes con homeobox, se localiza mayoritariamente en los núcleos de las células donde este gen se expresa (CARROL y otros, 1986).

Una característica muy llamativa de los genes homeóticos es que existe una correlación entre el orden anteroposterior de los dominios de expresión en el embrión y la disposición (organización molecular) de estos genes sobre el cromosoma. Así, en *Drosophila* el orden cromosómico de genes *lab-pb-Dfd-Scr-Antp-//-Ubx-abd.A-Abd.B* se corresponde con el orden de los dominios de expresión dentro del embrión. Se desconoce totalmente el fundamento de dicha correlación, si bien se especula que debe tener alguna función fundamental, ya

que los homólogos de estos genes en el ratón (también organizados en complejos denominados Hox) mantienen esta correlación (GRAHAM y otros, 1989; revisado en AKAM, 1989).

3.4. Genes de mantenimiento

Se ha descrito un gran número de genes (fuera de los dos complejos que hemos mencionado) que, en su estado mutado, transforman la identidad de los segmentos. No se trata, sin embargo, de genes que participen en el establecimiento de las identidades de los segmentos sino que son genes con funciones de mantenimiento, tales como el gen Polycomb.

Un aspecto que no hemos tratado todavía son las funciones de mantenimiento de los patrones de expresión de los genes de tipo segment-polarity y de los genes homeóticos. A diferencia de los otros genes del desarrollo, dichos genes continúan expresándose durante toda la vida del animal. En la regulación de esta función de mantenimiento, los genes del grupo Polycomb (unos 40) parecen desempeñar un papel clave. A juzgar por la evidencia genética, todos ellos parecen tener una función represora de la expresión de los genes homeóticos (allí donde no deben expresarse). En los embriones Pc^- la expresión de los genes homeóticos se dispara en todos los segmentos.

La función represora de Pc se ha puesto de manifiesto al detectar la presencia de la proteína Pc en unas 60 posiciones de los cromosomas politénicos, incluyendo el ANT-C y el BX-C

(ZINK y PARO, 1989). A pesar de eso, no hay evidencia de que esta proteína se una directamente al DNA.

Algunos de los genes mejor caracterizados de este grupo son: Polycomb (Pc), extra sex combs (esc), Additional sex combs (Asc), extra denticles (exd), Polycomb-like (Pcl), polyhomeotic (ph), Posterior sex combs (Psc), Sex combs on midleg (Scm) y super sex combs (sxc). Todos estos genes manifiestan un fenotipo que parece un mutante débil de Pc (transformación parcial de la mayoría de los segmentos en segmentos abdominales posteriores). PARO y colaboradores han demostrado que la proteína Pc se une in vivo al promotor P1 del gen Antp (com.pers). Estos autores han sugerido que la función del producto protéico de Pc podría ser la de definir los límites anteriores de la expresión de los genes homeóticos.

4. DOMINIOS DE EXPRESION DE LOS GENES DEL DESARROLLO

En el apartado 3.3 nos hemos referido a la estructura de los complejos ANT-C Y BX-C, así como a los genes maternos y de segmentación. La clonación de estos genes ha permitido obtener sondas para la detección in situ de los transcritos en embriones y determinar, así, los dominios de expresión de dichos genes (AKAM, 1983; HAFEN et al, 1983).

La detección de transcritos no sólo se ha llevado a cabo

en individuos normales. Además, se ha hecho una gran cantidad de trabajo encaminado a determinar la distribución de transcritos de un gen a en mutantes de un gen b. Este tipo de estudios ha permitido evidenciar múltiples interacciones entre los distintos tipos de genes que hemos enumerado.

El estudio de la distribución de transcritos ha evidenciado varias características importantes, a las que ya nos hemos referido en parte: a) los genes del desarrollo se expresan en dominios espaciales muy precisos; y b) existe un orden temporal en la expresión y en la función.

En general, se observa que los genes homeóticos y de segmentación se expresan en niveles elevados en aquellas regiones del embrión donde se observan los efectos fenotípicos de las mutaciones por pérdida de función.

Se demuestra que los genes de segmentación se expresan muy temprano durante la embriogénesis. Los primeros genes que se transcriben en el cigoto son los gap genes. Así, el RNA de Krüppel se detecta inmediatamente después de la división nuclear 11 (sincitio). Su dominio de expresión es una banda gruesa, en el centro del embrión, que corresponde al primordio de la región torácica.

Los transcritos de los pair-rule genes, tales como hairy, runt, fushi tarazu y even-skipped se detectan poco después de la expresión de Kr. Así el gen ftz se expresa únicamente desde el blastodermo temprano hasta el estadio de gástrula. Su dominio espacial de expresión son siete bandas con un grosor igual al primordio de un segmento (4 células), aunque estas

bandas están desplazadas con respecto a los primordios de los segmentos. En términos de parasegmentos, el gen *ftz* se expresa en todos los parasegmentos pares; el dominio de expresión de *eve* es exactamente el complementario: los parasegmentos impares.

El mRNA de los segment polarity genes, como *wingless*, *engrailed* y *gooseberry* aparece después, durante la división nuclear 13. El gen *en* se expresa en la parte anterior de cada uno de los 14 parasegmentos; el gen *wg* se expresa en la parte posterior de los parasegmentos.

Puede verse que los genes de segmentación tienen dominios de expresión cada vez más restringidos, lo cual subdivide al embrión en regiones que poseen diferentes combinaciones de productos génicos.

Por último, la primera expresión de los genes homeóticos se detecta en la división nuclear 14 (blastodermo celular, a las 3 h de desarrollo). Cada uno de los cinco genes homeóticos del ANT-C (apartado 3.3.2) tiene un dominio de expresión (en parasegmentos) de tal forma que estos genes "se reparten" el control de la identidad de la cabeza y de la parte anterior del tórax (parasegmentos 1 a 4).

Los estudios de hibridación *in situ* están poniendo de manifiesto también la existencia de unas redes de interacción complejas. En general, se concluye que los genes de una cierta clase vienen influenciados por la acción de los genes que actúan previamente (los cuales controlan zonas más extensas) y por algunos genes de su misma clase. En el apartado 7 nos

referiremos a estos aspectos.

5. PATRONES ESTRUCTURALES EN LOS GENES DEL DESARROLLO

Después de la clonación de los ANT-C y BX-C, pronto se descubrió que varias regiones de ambos complejos génicos tenían una secuencia de 180 pb en común. A juzgar por las intensidades de hibridación dichas secuencias eran muy similares entre ellas, aunque no idénticas (McGINNIS y otros, 1984; SCOTT y WEINER, 1984). Pronto se comprobó que estas secuencias, que se denominaron homeo box (HB), se encuentran en la práctica totalidad del reino animal (RUDDLE y otros, 1985; McGINNIS, 1985) y que codifican para un dominio protéico de 60 aminoácidos.

El hecho de encontrar una secuencia en común entre genes que tienen en común la función de controlar el desarrollo llevó a la proposición de que el dominio protéico codificado tendría una actividad similar en todos estos genes. Las primeras especulaciones sobre su posible actividad de unión al DNA se basaban en una cierta similitud con algunos represores procarióticos de la clase helix-turn-helix (SCOTT y WEINER, 1984). Por otra parte, la localización subcelular de las proteínas que contienen HB es, mayoritariamente, nuclear.

En efecto se ha comprobado que la proteína de cualquier gen que contenga HB (expresada en vectores bacterianos) se une

de forma específica al DNA (DESPLAN y otros, 1985, 1988). Hay toda una serie de trabajos recientes que demuestran que los genes con dominio homeótico se unen a secuencias específicas y regulan la transcripción de los genes donde se unen.

Se estima que en *Drosophila* hay entre 30 y 40 genes que contienen HB: sólo el Complejo Antennapedia contiene 9 de estas secuencias. Muchos de los genes que contienen homeobox han sido clonados utilizando hibridaciones de estrictancia reducida durante la búsqueda en librerías genómicas. Ha sido posible, incluso, la clonación de genes de los que no había mutaciones disponibles, tales como *caudal* (MLODZIK, y otros, 1985; MLODZIK y GEHRING, 1987). Hasta la fecha, todos los genes con HB que se han clonado en este organismo son genes de expresión temprana en el desarrollo: genes homeóticos (ej. *Antp*), genes de segmentación (ej. *ftz*) y, al menos, un gen de efecto materno (*bcd*). La única excepción es el gen *rough*, que está implicado en la diferenciación terminal del ojo (TOMLINSON y otros, 1988). En el siguiente apartado nos referiremos a las propiedades de unión al DNA del dominio homeótico.

Otro patrón estructural común a casi todos los genes de segmentación del tipo *gap* son los "zink fingers". De esta forma, los genes *Krüppel* y *hunchback* no sólo pertenecen a una misma familia de genes en cuanto a su función, sino que también están relacionados a través de su estructura. Las proteínas con la estructura de zink finger son reguladores de la transcripción, que reconocen el surco mayor del DNA. El

dominio de unión al DNA del otro gen del tipo gap, knirps, no contiene ningún zink finger, sino que contiene una región que parece pertenecer a la familia de los receptores de esteroides y glucocorticoides.

A diferencia de los dominios que hemos citado, hay otra característica común a muchos genes homeóticos y de segmentación cuya función se desconoce todavía. Se trata de la abundancia extrema de codones para un cierto aminoácido. Normalmente son repeticiones del codón CAG, que codifica el aminoácido glutamina. A estas repeticiones se les denomina: repeticiones opa, repeticiones M o, simplemente, repeticiones CAG.

Este tipo de repeticiones se encuentra en unas 200 copias en el genoma de Drosophila y su presencia se ha descrito en los genes Antp, Notch, hairy, hunchback, engrailed, bicoid y otros. En otros capítulos nos referiremos extensamente a estas secuencias.

6. UNION AL DNA DE LAS PROTEINAS CON DOMINIO HOMEOTICO

En el apartado anterior nos hemos referido al dominio homeótico como una región de unión al DNA. Hasta hace muy poco no se conocía con exactitud cuál era la actividad de las proteínas que contienen este dominio. Por la función genética de los genes homeóticos y de segmentación, se suponía que sus

productos serían reguladores de la transcripción. Esto ya ha sido demostrado, tal como veremos en el apartado 7. Una confirmación adicional de la implicación de las proteínas HB en la regulación transcripcional de genes específicos radica en el reciente hallazgo de que un factor de la transcripción en el hombre, el OTF-2, es una proteína que contiene HB (SCHEIDEREIT y otros, 1988; LEVINE y HOEY, 1988).

La estructura tridimensional del dominio homeótico (HD) del gen *Antp* en solución acuosa ha sido determinada mediante resonancia magnética nuclear (OTTING y otros, 1988). De esta forma se ha demostrado que, tal como se había especulado, el HD tiene una estructura que se encuadra dentro del patrón helix-turn-helix. Los 60 aminoácidos del HD forman tres α -hélices, siendo las hélices 2 y 3 las implicadas en la interacción con el DNA. A la hélice 3 se le atribuye la función de reconocimiento de secuencia.

Dado que las hélices 3 de los HD de los genes *Antp* y *ftz* son idénticas, se pensó que ambos productos protéicos debían reconocer las mismas secuencias de DNA. Esta posibilidad se ha confirmado al comprobar las propiedades de unión al DNA del HD de *Antp* expresado en *E. coli* y de la proteína del gen *ftz* purificada. Ambos productos se unen con idéntica afinidad a secuencias clonadas del gen *engrailed*. Todos los sitios de unión tienen en común la secuencia ANNNNCATTA. Esta secuencia consenso también se encuentra en la región promotora del gen *ftz*; el dominio homeótico de *Antp* y la proteína *ftz* también se unen en esta posición (MÜLLER y otros, 1988b).

7. COMO FUNCIONA EL PROGRAMA DE DESARROLLO

Muchos de los productos de los genes a los que nos hemos referido hasta ahora tienen en común patrones protéicos, lo que indica posibles similitudes funcionales. Esto es cierto, al menos, en lo que se refiere al dominio homeótico y a los zink-fingers de los genes del tipo gap.

Es un hecho demostrado que las proteínas que contienen HB son factores de la transcripción con especificidad de secuencia. Hay evidencias firmes de que el producto protéico de bicoid activa la transcripción del gen hunchback (DRIEVER y NÜSSLEIN-VOLHARD, 1989), lo que demuestra, a nivel molecular, las evidencias genéticas de que bicoid regula positivamente a hunchback. Esto es un claro ejemplo de interacción entre los productos del genoma materno y los genes del cigoto. También se ha demostrado que los productos protéicos de los genes fushi tarazu y Antennapdia son activadores de la transcripción de cualquier gen que contenga un sitio de unión para su dominio homeótico (JAYNES y O'FARRELL, 1988; WINSLOW y otros, 1989).

No todas las proteínas con dominio homeótico son activadores de la transcripción: se ha comprobado que los productos de los genes engrailed y even-skipped no tienen actividad activadora. Al contrario, estas proteínas tienen la

capacidad de inhibir la activación que puedan producir otros activadores de la transcripción (JAYNES y O'FARRELL 1988; BIGGIN y TJIAN, 1989). A nivel genético, ya se conocía que el gen *even-skipped* es un represor de los genes *fushi tarazu*, *wingless* y *Ubx*. Ahora se conoce el mecanismo de represión: el producto del gen *eve* inhibe la activación de los promotores de estos tres genes. Esto constituye un ejemplo de tres tipos de interacciones génicas: (1) entre genes del grupo *pair rule* (*eve* con *ftz*); (2) entre genes del grupo *pair rule* y del grupo *segment polarity* (*eve* con *wg*); y (3) entre genes del grupo *pair rule* y los genes homeóticos (*eve* con *Ubx*).

También se han determinado, a nivel molecular, algunas de las interacciones entre los genes homeóticos. Así, se ha demostrado que el producto protéico de *Ubx* se une al promotor distal (P1) del gen *Antp* y reprime su transcripción (BEACHY y otros, 1988; KRASNOW y otros, 1989). Estos datos proporcionan una prueba de las evidencias genéticas que indicaban la existencia de interacciones reguladoras entre ambos genes. Además de unirse al promotor P1 de *Antp*, la proteína de *Ultrabithorax* también se une al promotor de su propio gen, si bien en este caso se produce una activación de la transcripción.

El que un producto de un gen active su propio promotor parece ser un hecho general en todos los genes homeóticos, lo que podría ser uno de los mecanismos que dan lugar a la expresión continuada de estos genes. Así, también se ha descrito la autoactivación del gen *Deformed* (KUZIORA y

McGINNIS, 1988).

Ya nos hemos referido, a grandes rasgos, a los momentos y lugares donde se expresan los genes del desarrollo. En los tres últimos párrafos hemos presentado algunos ejemplos de interacciones entre genes cuyos mecanismos se han determinado a nivel molecular. Estos trabajos han confirmado las hipótesis originales de que existe un circuito de órdenes jerarquizadas y de interacciones.

Los genes maternos proporcionan una información (en forma de gradientes) que es interpretada por los genes de segmentación del tipo gap. Los genes gap se expresan de acuerdo con esta información, y subdividen el embrión en varias regiones. Estas regiones delimitadas por los genes gap marcan los límites de expresión de los genes homeóticos. Además, la expresión de los genes gap regula la de los genes que se expresan a continuación: los de la clase pair rule.

Los genes hairy y runt parecen ser los dos genes de la regla par cuya actividad depende directamente de la expresión de los genes gap (INGHAM, 1988); el resto de los genes de esta clase están regulados por la actividad de hairy y runt. Así, se sabe que hairy y runt reprimen, respectivamente, a los genes de su grupo ftz y eve, lo que da lugar a la formación de siete bandas complementarias de expresión. Cada una de estas bandas de expresión coincide exactamente con lo que van a ser los parasegmentos (LAWRENCE y otros, 1987).

En el momento de la formación del blastodermo, todas estas interacciones han dado lugar a que las distintas células del

embrión expresen distintas combinaciones de genes de la clase pair rule. Esta información será interpretada por los genes de la clase segment polarity, con lo que se marcarán los límites de los parasegmentos. Los genes engrailed y wingless son los que responden directamente a los genes de la regla par (INGHAM y otros, 1988).

Cada una de las 14 bandas de expresión de en marca el límite anterior de cada parasegmento (INGHAM y otros, 1985); mientras que las bandas de expresión de wg delimitan los límites posteriores. El incio de estas bandas de expresión de los genes de polaridad de los segmentos requiere distintas combinaciones de las producidas por los genes que se han expresado anteriormente (pair rule) (INGHAM, 1988). La expresión de los genes de polaridad de los segmentos se mantiene permanentemente y, junto con la actividad de los genes homeóticos, contribuye a mantener la identidad de los parasegmentos.

Por último, los promotores de los genes homeóticos se expresan de una forma que parece ser dependiente de determinadas combinaciones de expresión de genes de la clase gap y de la clase pair rule (IRISH y otros, 1989; INGHAM y MARTINEZ-ARIAS, 1986). Así, hay evidencias de que el promotor P1 del gen Antennapedia se activa por el producto del gen Krüppel (clase gap); el promotor P2 de dicho gen depende, al menos, de los genes hunchback y ftz.

Nos hemos referido principalmente a las interacciones entre genes de una clase con los genes de otra clase. Junto

con esto, hay otra serie de interacciones entre genes de una misma clase, lo que da lugar a un "refinamiento" de los dominios de expresión. Por ejemplo, se sabe que el producto protéico del gen homeótico Ubx reprime la expresión del gen Antennapedia, lo que limita la expresión de este último gen al parasegmento 4 (y, parcialmente, el 5).

La forma de "leer" la información proporcionada por el gradiente materno parece ser por la afinidad que los sitios de unión tengan por la proteína materna del gradiente. Un gen que, en su promotor, contenga un sitio de baja afinidad de unión a la proteína bcd será activado sólo en la parte donde la concentración de esta proteína sea mayor (la parte anterior del embrión); un promotor que contenga un sitio de unión de alta afinidad por bcd se expresará en un dominio anterior mucho más amplio (DRIEVER y otros, 1989).

Por analogía, puede pensarse que la información proporcionada por los genes del grupo posterior (como oskar) será interpretada de una forma similar al gradiente producido por bicoid.

Nuestra exposición sobre algunos datos del programa de desarrollo se ha referido a la diferenciación a lo largo de un sólo eje: el anteroposterior. Como dijimos en el apartado 2.2, hay otro programa de desarrollo independiente, que especifica las diferencias a lo largo del eje dorsoventral. Establecer la polaridad D-V equivale a hablar de determinación de las hojas embrionarias, es decir, de la organización de los tipos tisulares en el embrión. Este proceso se conoce con menor

detalle, de modo que no nos vamos a referir al mismo en esta introducción.

8. ESTRUCTURA DEL GEN Antennapedia

En este capítulo vamos a describir algunos aspectos de la estructura de este gen, con el fin de facilitar la comparación de nuestros resultados con lo que ya se conoce en *Drosophila melanogaster*. La organización transcripcional del gen Antennapedia es una de las más complejas que se han descrito en *Drosophila*. Ya en los primeros trabajos de clonación, se observó que el transcrito primario de Antp podría ser extremadamente largo, unas 100 kb. Se llegó a sugerir, incluso, que el transcrito final se produjera tras una reordenación cromosómica (GARBER y otros, 1983).

En la figura 5 mostramos un esquema de la estructura molecular de este gen. El análisis de esta unidad de transcripción ha revelado que el gen consta de ocho exones, separados por intrones extremadamente largos. El mayor de los exones tiene un tamaño de unas 60 kb e incluye al exon 3. El gen contiene dos promotores, separados por 60 kb kb, y dos regiones de poliadenilación (SCHNEUWLY y otros, 1986; LAUGHON y otros, 1986). El promotor distal (P1) tiene una TATA box situada a una distancia apropiada; el promotor interno (P2) no contiene ninguna. Ambos promotores son funcionales en el

embrión (JORGENSEN y GARBER, 1987).

Esta unidad de transcripción produce, así, cuatro transcritos poliadenilados, todos ellos con secuencias líder y trailer anormalmente largas. Estas secuencias son posibles candidatos que porten señales que regulen la estabilidad del RNA. Sin embargo, ninguna de estas diferencias en los transcritos afecta a la ORF.

Todos los límites intrón/exón están de acuerdo con la secuencia consenso, excepto el segundo sitio donador (D2) del intrón 7 donde, al principio del intrón, la secuencia es GA, en vez de GT. El primer sitio donador, en cambio, sí se corresponde con la secuencia consenso.

Recientemente se ha descrito la existencia de un patrón complejo de splicing, que dá lugar a la producción de cuatro productos protéicos distintos (BERMINGHAM y SCOTT, 1988). Estos cuatro productos difieren, por un lado, en la posesión o no de los 13 aminoácidos correspondientes al exón 6 y, por otro lado, en la posesión o no de los cuatro últimos aminoácidos del exón 7. Estas variaciones en el procesado no tienen relación alguna con la existencia de cuatro distintos tipos de transcrito (atendiendo a las secuencias líder y trailer). Realmente puede hablarse de la existencia de 16 transcritos diferentes de Antp.

Se desconoce totalmente la función (si la hay) de la existencia de estas variantes protéicas, si bien hay motivos para suponer que dichas variaciones no son triviales. En un análisis exhaustivo de las preferencias temporales de cada

variante, se observa que hay distintos dominios temporales (BERMINGHAM y SCOTT, 1988), lo cual no sería de esperar si las variaciones carecieran de significado biológico. Además, también se ha descrito un patrón de procesado complejo en el gen Ubx, que igualmente produce varias (cinco, al menos) alternativas protéicas a lo largo del desarrollo (O'CONNOR et al, 1988).

En ambos casos, la heterogeneidad en los productos protéicos dá lugar a variaciones de la secuencia en las cercanías del dominio homeótico, por lo que uno de los efectos podría ser la alteración de la afinidad y/o especificidad de unión al DNA (O'CONNOR et al, 1988).

A pesar de las preferencias temporales en la formación de los distintos patrones de splicing, varias de las especies coexisten en muchas ocasiones. Queda abierta la posibilidad de que cada especie tenga un subdominio de expresión dentro del dominio Antp. Se ha sugerido que las diferencias en la utilización de las distintas alternativas podrían ser mucho más sutiles de lo que es posible detectar con extractos de embriones enteros.

El análisis de la secuencia de este gen ha revelado la existencia de numerosos codones CAG en tándem, lo que dá lugar a una proteína con un contenido muy elevado en Glutamina. Esta característica también se da en otros genes de Drosophila (ver apartado 5 de esta introducción).

La secuencia homeo-box está situada en el exón 8, de modo que el dominio homeótico se sitúa prácticamente en el extremo

carboxiterminal de la proteína. Otras características de la proteína se discutirán en el contexto de su comparación con la secuencia de *Drosophila subobscura*.

La mayoría de los mutantes dominantes que producen la transformación de la antena en pata tienen inversiones grandes, visibles a nivel citológico. La localización de los puntos de dichas inversiones dentro de la unidad de transcripción demostró que ninguno de éstos afecta a la ORF, evidenciando así que, en los mutantes dominantes por ganancia de función, la región del gen que codifica el producto protéico no está alterada, pero sí separada de sus regiones propias de control (JORGENSEN y GARBER, 1987).

9. ASPECTOS ESENCIALES Y ASPECTOS IRRELEVANTES DE LA ESTRUCTURA: UNA APROXIMACION EVOLUTIVA

La simple descripción de la estructura molecular de los genes implicados en el desarrollo ha revelado, en muchos casos, ciertas particularidades: promotores múltiples, patrones de splicing alternativo, intrones gigantescos, etc. Esto ha despertado la siguiente cuestión: ¿es necesaria esa estructura para que el gen lleve a cabo sus funciones?

Una aproximación a la cuestión consiste en la comparación de la estructura de un mismo gen en varios organismos en los que se asume una identidad de función para dicho gen. Aquellos

aspectos no conservados son, en principio, características no esenciales; análogamente, puede concluirse que las características conservadas corresponden a lo que es esencial para la función. Estas afirmaciones son totalmente cuestionables. Se puede argumentar que no hay motivo para suponer que todos los genes del desarrollo desempeñen la misma función en cualquier organismo, ni siquiera dentro del género *Drosophila*.

Si, en efecto, uno de los mecanismos básicos de la evolución es el cambio en la regulación de los patrones de expresión génica durante el desarrollo (BONNER, 1983), las diferencias observadas entre dos especies podrían corresponder a un cambio en la regulación del gen o en la función de su producto, y no a variaciones en regiones no esenciales. En principio, este tipo de diferencias debería quedar minimizado estudiando especies no muy distantes en la escala evolutiva (el autor se declara totalmente incapaz de medir la distancia óptima).

Por otra parte, si las especies son demasiado próximas (aceptemos una medida relativa de la proximidad óptima para el estudio), es probable que encontremos muchas similitudes en la estructura. Pero este parecido se deberá, simplemente, al corto período de tiempo que ha transcurrido entre la divergencia de las especies en estudio, y no a un carácter esencial de ese aspecto.

Este tipo de discusión no se ha llevado a cabo en profundidad: en casi todas las comparaciones interespecíficas

de genes del desarrollo publicadas en la actualidad, los autores han escogido una especie de *Drosophila* "distante" (*D. virilis*). Desde luego una aproximación ideal sería la comparación en varias especies, incluyendo otros género, como hicieron WILDE y AKAM (1987) al comparar parte del gen *Ubx* en *Drosophila pseudoobscura*, *D. funebris* y *Musca domestica*.

A pesar de que algunos puntos sean discutibles, la comparación entre especies es tal vez la única estrategia disponible para determinar las estructuras esenciales en un gen del tamaño de *Antennapedia*.

La especie a comparar debe ser lo suficientemente distante de *D. melanogaster* para que las secuencias de DNA que no esten sujetas a restricciones evolutivas hayan divergido lo suficiente como para permitir que los elementos funcionales puedan identificarse en forma de secuencias conservadas.

De la comparación, cabe esperar varios tipos de resultados. Por un lado, la identificación de aquellos dominios del producto protéico de *Antp* que están sujetos a restricciones (además de la homeo box) y de aquellos elementos necesarios para la expresión correcta de *Antp* (en el entorno de los promotores). Por otro lado, es de esperar la identificación de secuencias del mRNA no traducidas (líder y trailer) que tengan alguna funcionalidad. También se espera determinar qué aspectos de la organización general del gen se han conservado.

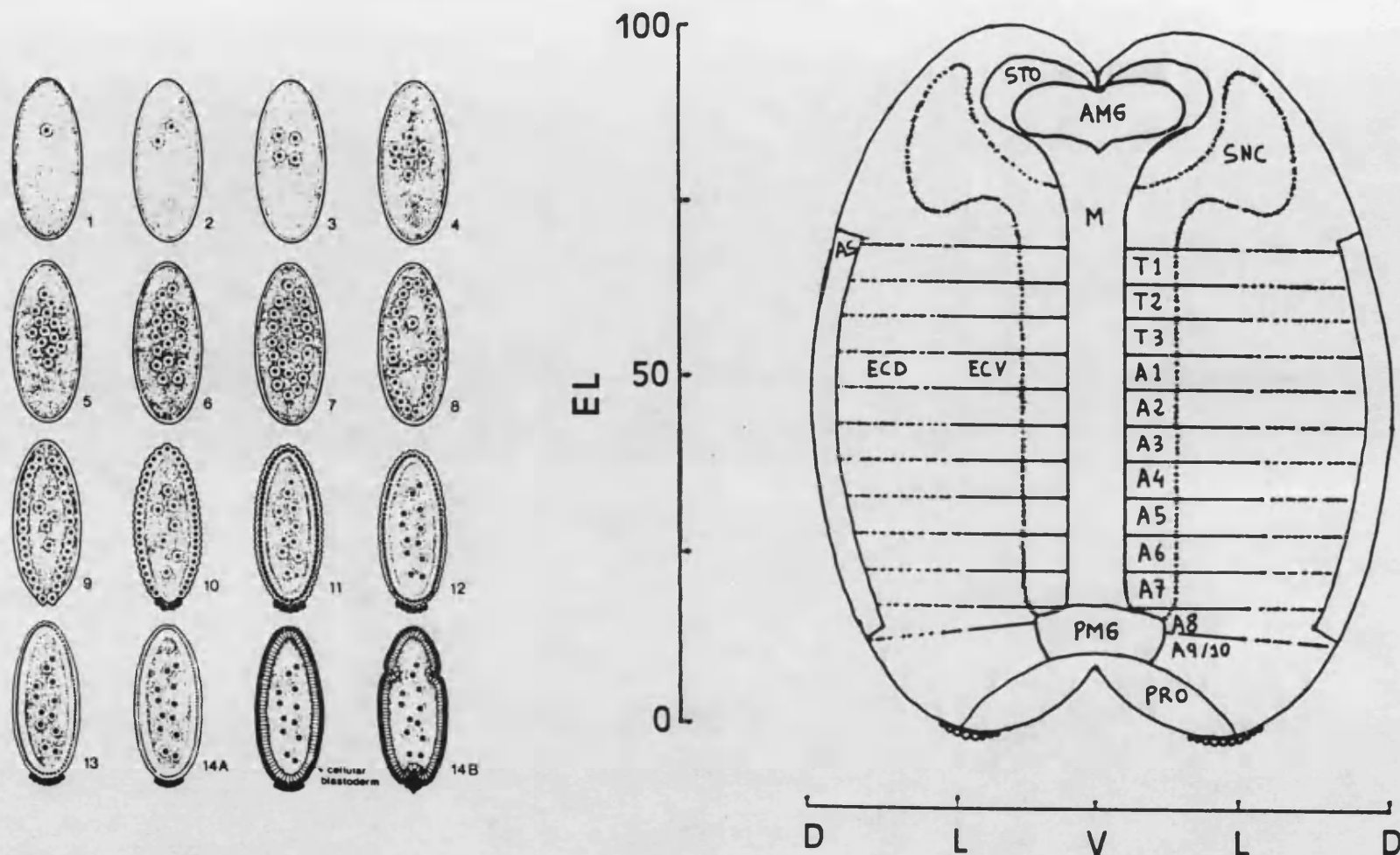
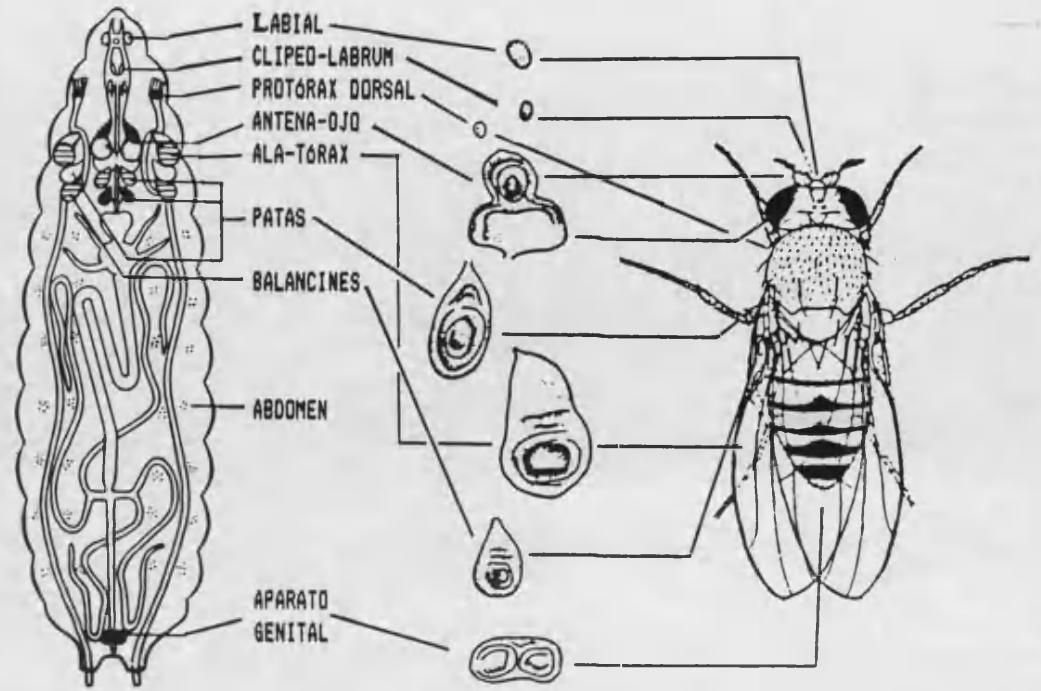
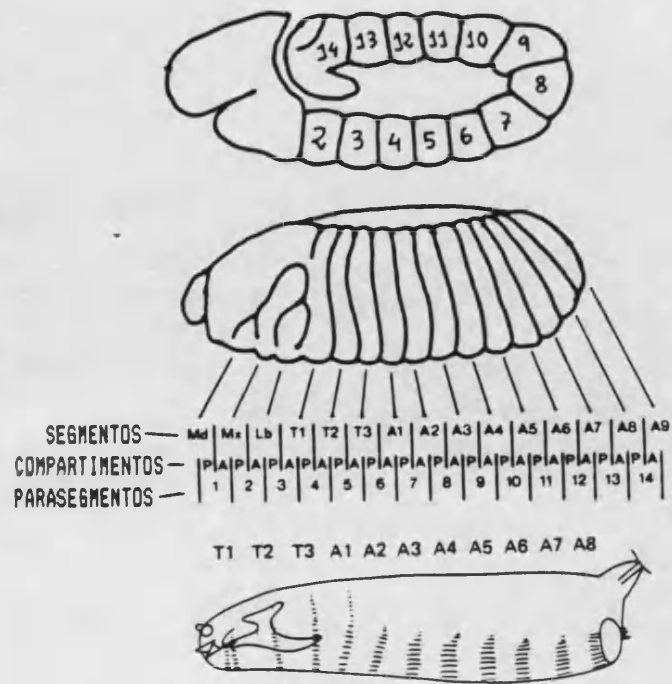


Figura 1. (A) Estadios de divisiones nucleares durante el desarrollo temprano de *D. melanogaster*. La parte anterior del embrión está orientada hacia arriba. Los puntos negros representan los núcleos, que sólo se separan por membranas celulares en la última fase: el blastodermo celular (según FOE y ALBERTS, citados en WILKINS, 1986). (B) Mapa de destino del blastodermo para las estructuras larvarias de *Drosophila*. El eje de ordenadas representa la distancia relativa desde el extremo posterior (egg length, EL=0), hasta el extremo anterior (EL=100). El eje de abscisas representa la posición a lo largo del eje dorsoventral. T, segmentos torácicos; A segmentos abdominales; M, mesodermo; SNC, sistema nervioso central; STO, estomodeo; PRO, proctodeo; AMG, midgut anterior; PM6, midgut posterior; ECD, ectodermo dorsal; ECV, ectodermo ventral; AS, membrana extraembriónica (amnioserosa); las células del extremo posterior son las células polares de la línea germinal. (según NÜSSLEIN VOLHARD, citado en WILKINS, 1986, con modificaciones de FRENCH, 1988).



A B

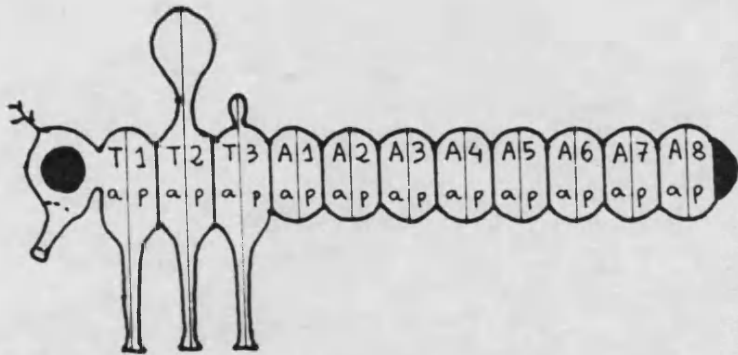


Figura 2. (A) Correspondencia entre segmentos, compartimentos y parasegmentos en el embrión, la larva, y el imago. El embrión superior corresponde a la fase de extensión de la banda germinativa (germ band extension unas 5 horas después de la fecundación); los números indicados denotan los parasegmentos en que se divide. El embrión inferior corresponde a las 10 horas de desarrollo. (B) Discos imaginales y células histoblásticas y su localización en la larva de tercer estadio. Su correspondencia con las estructuras del adulto se muestra en el esquema.

3L

3R

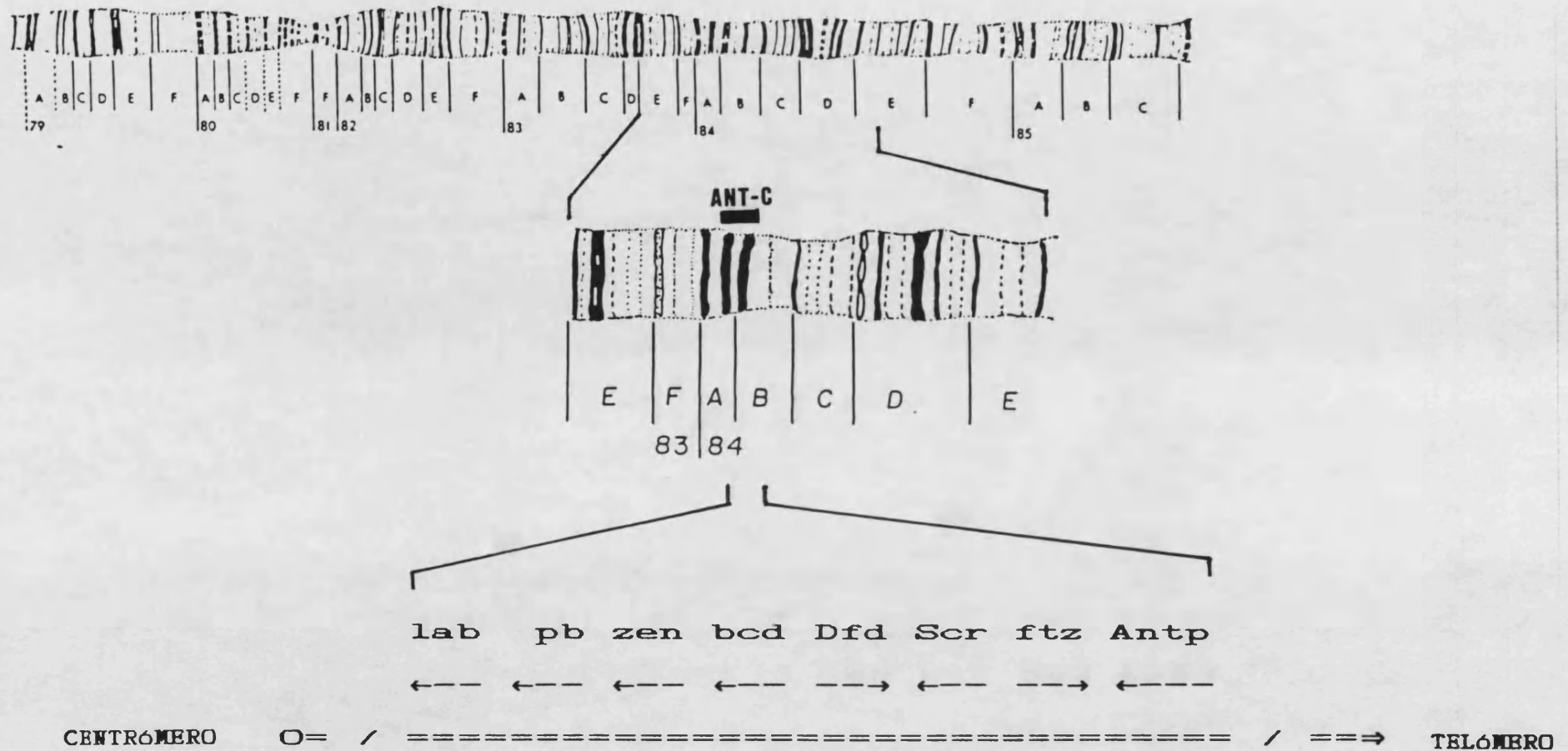


Figura 3. Localización del Complejo Antennapedia (ANT-C) en el cromosoma 3R de *Drosophila melanogaster*. La orientación relativa de los genes respecto del centrómero y del telómero se indica en los diagramas inferiores. Las flechas señalan la dirección de la transcripción de los genes incluidos en el complejo. *lab*, labial; *pb*, proboscipedia; *zen*, zerknullt; *bcd*, bicoid; *Dfd*, Deformed; *Scr*, Sex combs reduced; *ftz*, fushi-*tarazu*; *Antp*, Antennapedia. En la figura 4 se muestra un mapa de la estructura del complejo a nivel molecular.

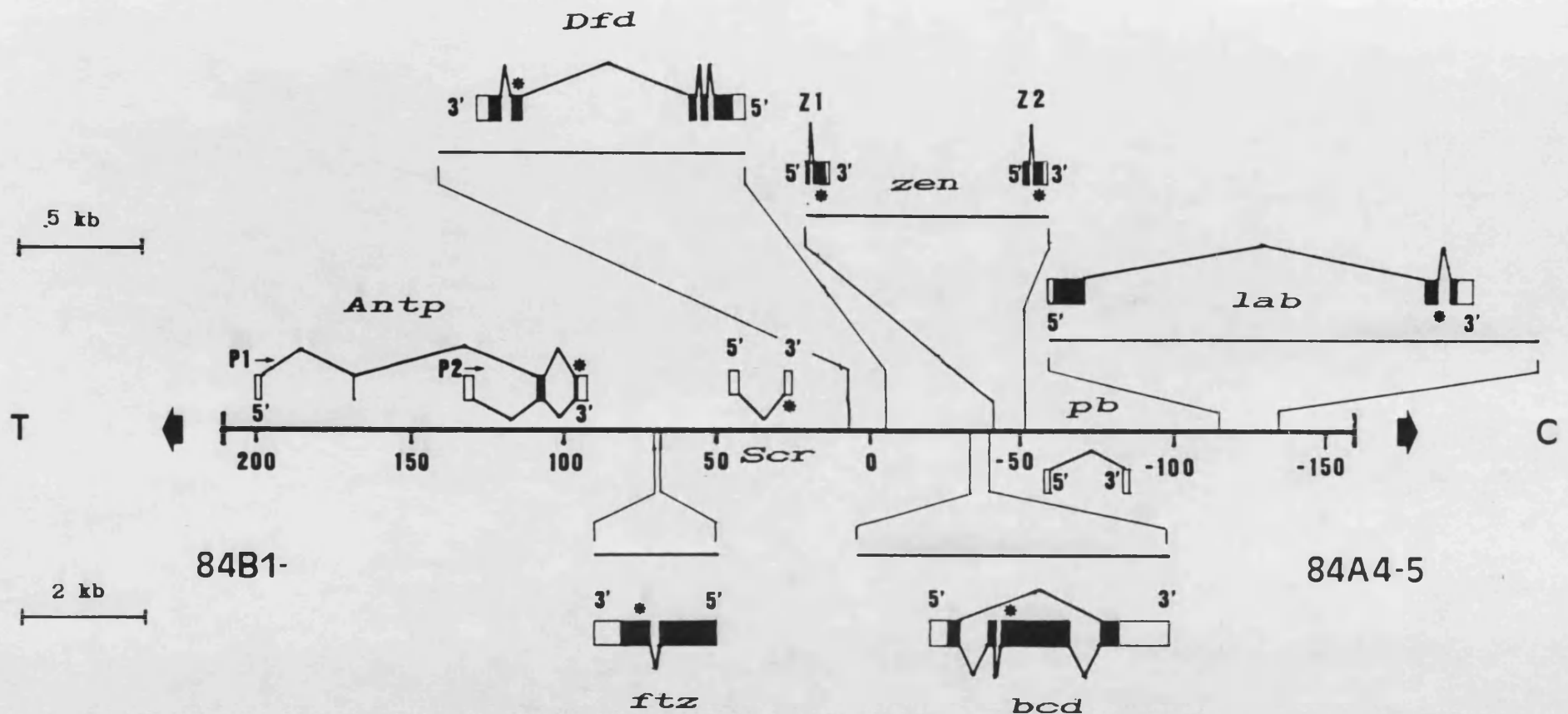


Figura 4. Estructura del Complejo Antennapadia a nivel molecular. Las coordenadas de la región (desde +200 kb hasta -150 kb) corresponden a las descritas por SCOTT y otros (1983). La orientación del complejo es la opuesta a la indicada en la figura 3. T marca la dirección del telómero y C la del centrómero. Los símbolos 84A4-5 y 84B1-2 señalan la correspondencia aproximada con el patrón de bandas de los cromosomas politénicos (fig. 3). La orientación de cada unidad de transcripción se ha indicado con 5' y 3'. Los recuadros señalan los exones (en blanco, regiones no traducidas; en negro, regiones traducidas). Los intrones los hemos representado con líneas que forman ángulos. Obsérvese que hemos representado las distintas unidades de transcripción con diferentes escalas; los genes *Antp*, *Scr* y *pb* están dibujados con la propia escala del genoma. Los genes *Dfd*, *zen*, y *lab* están ampliados con la escala indicada en la parte superior; el resto de los genes están todavía más ampliados, con la otra escala indicada. Las regiones traducidas y no traducidas de los genes *Scr* y *pb* no están indicadas debido a que la estructura de sus transcritos no ha sido descrita totalmente. Las posiciones de las secuencias homeobox están señaladas con asteriscos. La estructura del gen de mayor tamaño (*Antp*) se muestra con mayor detalle en la figura 5.

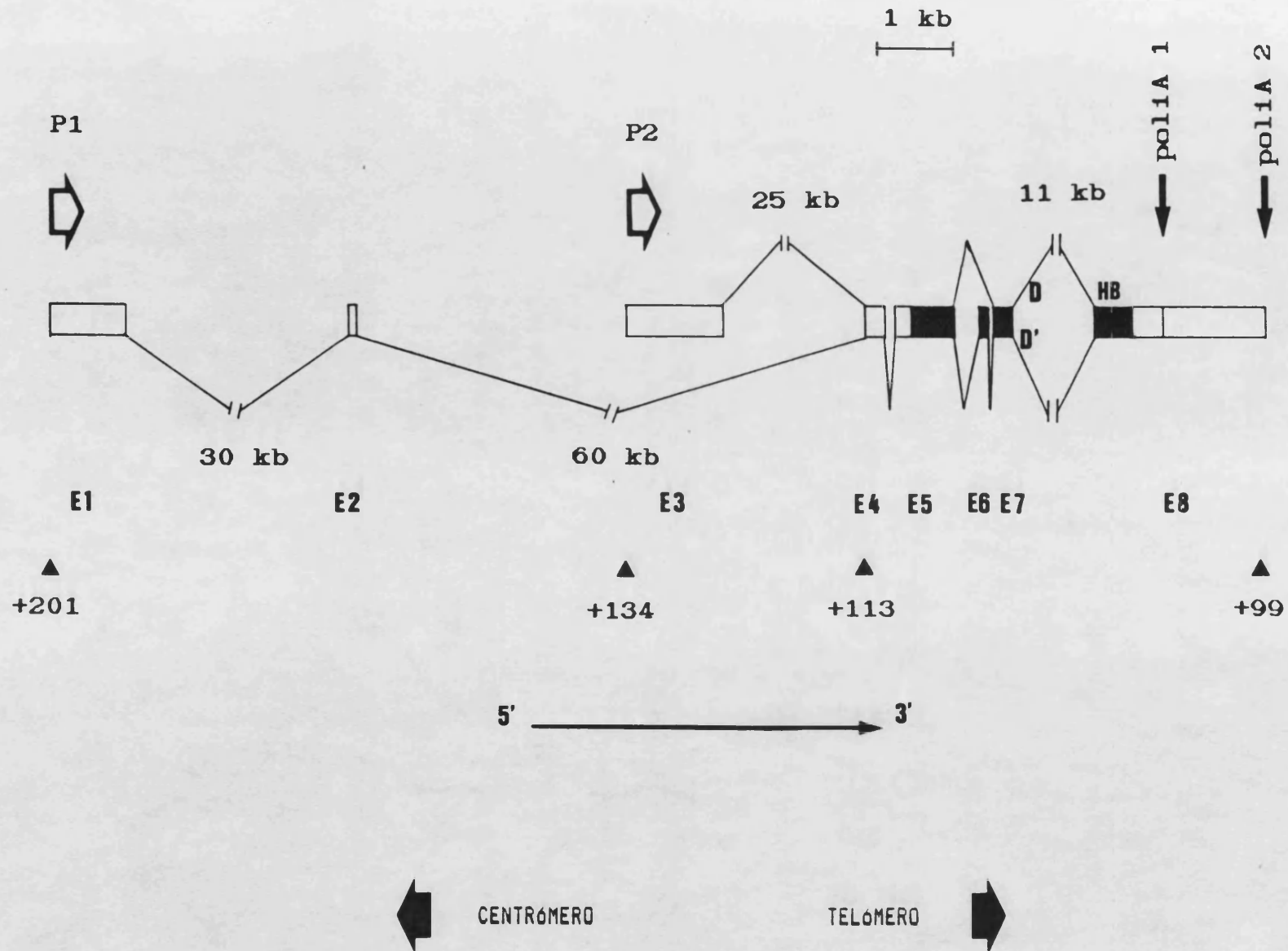


Figura 5. Estructura del gen *Antennapedia*. Se ha seguido la misma simbología que en la fig.4. Los exones están numerados con E1, E2,...; los intrones están interrumpidos y su tamaño está indicado junto a las interrupciones. P1 y P2 son los dos promotores del gen. D y D' son los dos sitios dadores alternativos. Los triángulos señalan las coordenadas genómicas, como en la fig.4 (según BERMINGHAM y SCOTT, 1988; y SCHNEUWLY y otros, 1986).

Genes de expresión materna implicados en la polaridad anteroposterior

GRUPO

LOCUS

Anterior

<i>bicoid</i>	<i>bcd</i>
<i>exuperantia</i>	<i>exu</i>
<i>swallow</i>	<i>swa</i>

Posterior

<i>oskar</i>	<i>osk</i>
<i>vasa</i>	<i>vas</i>
<i>valois</i>	<i>val</i>
<i>nanos</i>	<i>nos</i>
<i>tudor</i>	<i>tud</i>
<i>staufen</i>	<i>stau</i>
<i>pumilio</i>	<i>pum</i>

Terminal

<i>torso</i>	<i>tor</i>
<i>trunk</i>	<i>trk</i>
<i>torso-like</i>	<i>tsl</i>
<i>fs(1) polehole</i>	<i>fs(1)ph</i>
<i>fs(1) Nasrat</i>	<i>fs(1)N</i>
<i>l(1) polehole</i>	<i>l(1)ph</i>

Tabla 1. Genes de efecto materno que afectan al patrón del eje anteroposterior (según LEHMANN, 1988). Descripción en el texto.

Genes de expresión zigótica implicados en la segmentación

CATEGORÍA	LOCUS	Dominio	DNA-binding
-----------	-------	---------	-------------

Gap

	<i>Krüppel</i>	<i>Kr</i>	Zn-finger
	<i>knirps</i>	<i>kni</i>	Receptor esteroides
	<i>hunchback</i>	<i>hb</i>	Zn-finger

Pair-rule

	<i>hairy</i>	<i>h</i>	homología con <i>N-myc</i>
	<i>runt</i>	<i>run</i>	
	<i>fushi tarazu</i>	<i>ftz</i>	Homeobox
	<i>paired</i>	<i>prd</i>	Homeobox
	<i>odd-paired</i>	<i>opa</i>	
	<i>sloppy-paired</i>	<i>slp</i>	
	<i>even-skipped</i>	<i>eve</i>	Homeobox
	<i>odd-skipped</i>	<i>odd</i>	
	<i>barrel</i>	<i>brr</i>	

Segment Polarity

	<i>engrailed</i>	<i>en</i>	Homeobox
	<i>wingless</i>	<i>wg</i>	
	<i>patched</i>	<i>ptc</i>	
	<i>patch</i>	<i>pat</i>	
	<i>hedgehog</i>	<i>hh</i>	
	<i>gooseberry</i>	<i>gsb</i>	Homeobox
	<i>fused</i>	<i>fu</i>	
	<i>cubitus interruptus</i>	<i>ci</i>	
	<i>naked</i>	<i>nkd</i>	
	<i>armadillo</i>		

Tabla 2. Genes de segmentación (según NÜSSLEIN-VOLHARD y WIESCHAUS, 1980; SCOTT y CARROLL, 1987; INGHAM, 1988). Descripción en el texto.

Genes homeóticos

COMPLEJO GÉNICO

LOCUS

ANT-C

<i>Antennapedia</i>	<i>Antp</i>
<i>Deformed</i>	<i>Dfd</i>
<i>Sex combs reduced</i>	<i>Scr</i>
<i>proboscipedia</i>	<i>pb</i>
<i>labial</i>	<i>lab</i>

BX-C

<i>Ultrabithorax</i>	<i>Ubx</i>
<i>abdominal A</i>	<i>abd-A</i>
<i>Abdominal B</i>	<i>Abd-B</i>

Tabla 3. Genes homeóticos de los Complejos Antennapedia y Bithorax. El ANT-C contiene, además de estos cinco genes homeóticos, otros tres genes implicados en el desarrollo (ver fig. 4).

OBJETIVOS

OBJETIVOS

El trabajo que aquí presentamos forma parte de un proyecto inciado con un objetivo final: conocer qué aspectos de la estructura del gen Antennapedia (y su producto) son esenciales para que este gen lleve a cabo su función.

En el punto 8 de la Introducción hemos descrito algunas de las características estructurales más llamativas del gen Antp de *D.melanogaster*: el tamaño de la unidad de transcripción; el complejo patrón de procesamiento de los transcritos, que da lugar a variaciones en el producto protéico; la gran abundancia de repeticiones de un mismo aminoácido; etc. Una vez conocidos todos estos aspectos, se plantea la cuestión: ¿es todo éso necesario para el funcionamiento del gen Antp?

No es fácil diseñar experimentos que permitan destacar los aspectos de la estructura de un gen que sí son necesarios para su función, y menos cuando se trata de una unidad de transcripción de 100 kb. Una aproximación, a la que nos hemos referido en el punto 9 de la Introducción es comparar la estructura del gen en cuestión entre dos o más especies. Esta aproximación se ha utilizado para el estudio de tres genes del desarrollo, que previamente se habían descrito en *D.melanogaster*: engrailed, Ultrabithorax y hunchback (KASSIS y otros, 1986; WILDE y AKAM, 1987; TREIER y otros, 1989). Todos estos análisis han proporcionado información importante sobre la conservación (posible carácter esencial) de determinados

aspectos en otras especies.

Los genes *engrailed* y *hunchback* han sido clonados y secuenciados en su totalidad en *Drosophila virilis* (posteriormente a su caracterización en *D.melanogaster*). Esto ha sido posible gracias a su tamaño, relativamente pequeño. Sin embargo, para la comparación del gen *Ubx* en otras especies sólo se ha clonado y secuenciado una pequeña parte (el exón que codifica la porción aminoterminal de la proteína y la región 5' adyacente). El gran tamaño del gen *Ubx* hace difícil su estudio incluso en *D.melanogaster*, donde no se ha determinado la secuencia de toda la unidad de transcripción (sólo se han secuenciado los exones).

En este contexto, y dado que no se ha descrito la estructura de *Antp* en ninguna otra especie (además de *D.melanogaster*), decidimos iniciar la clonación de dicho gen en *D.subobscura*. El objetivo a largo plazo es determinar la estructura completa de este gen en dicha especie, y comparar el resultado con la estructura que se ha descrito en *D.melanogaster* (SCHNEUWLY y otros, 1986; LAUGHON y otros, 1986).

En el plazo de esta Tesis Doctoral, nos hemos planteado los siguientes puntos:

1. Aislar clones genómicos de la región del gen *Antp* de *Drosophila subobscura* que codifica el producto protéico.
2. Analizar la organización de los intrones y de los exones en la región clonada.
3. Determinar qué regiones del producto protéico de este

gen (en las dos especies) están sujetas a restricciones evolutivas y cuales son susceptibles de variación.

4. Buscar secuencias de DNA no traducidas que se hayan conservado en la evolución.

5. Completar, en la medida de lo posible, la clonación del resto de la unidad de transcripción (exones no traducidos e intrones gigantes).

RESULTADOS Y DISCUSION

RESULTADOS Y DISCUSION

1. DETECCION DE FRAGMENTOS HOMOLOGOS EN EL DNA GENOMICO DE *D.subobscura*

Como un primer paso antes de iniciar la clonación del gen *Antp* de *Drosophila subobscura*, hemos considerado necesario detectar, en esta especie, las secuencias homólogas al gen *Antp* de *D.melanogaster*. El objetivo no es confirmar la presencia de un gen homólogo en nuestra especie, cuya existencia nos parece obvia, sino determinar unas condiciones óptimas de hibridación para la posterior búsqueda de clones homólogos (screening) en la genoteca.

Con esta finalidad, nos hemos valido de una sonda de cDNA de *D.melanogaster*, *pcDm903*, que contiene los exones 1-2-4-5-6-7-8 (SCHNEUWLY y otros, 1986). Los exones 1, 2 y 4 forman parte del líder no traducido del transcrito; los exones 5, 6, 7 y 8 contienen la ORF de los productos protéicos. Un segmento del exón 5 forma parte del líder del transcrito. Este exón contiene, además, la señal de inicio de la traducción y la región 5' de la ORF, que codifica la mitad aminoterminal de la proteína. Por su parte, el exón 8 contiene el trailer del transcrito y, además, una pequeña parte de la ORF: la homeobox (SCHNEUWLY y otros, 1986; LAUGHON y otros, 1986).

El DNA de *D.subobscura*, digerido con *EcoRI* y transferido a

un filtro de nylon siguiendo el procedimiento de SOUTHERN, tiene una serie de fragmentos que hibridan, en condiciones de estringencia moderada, con el inserto del plásmido pcDm903, marcado con ^{32}P . La hibridación se ha llevado a cabo en una mezcla con un 46 % de formamida y una concentración salina de 5 x SSC, a una temperatura de 42°C. Todos los lavados se han realizado en 2 x SSC, 0.1% SDS (los dos primeros a temperatura ambiente; los dos siguientes a 50°C).

La autorradiografía que presentamos en la figura 6A demuestra la presencia, en el genoma de *Drosophila subobscura*, de un fragmento de aproximadamente 1.7 kb, que es el que proporciona la intensidad de hibridación más fuerte (banda marcada con una flecha en la figura). Además, se observa la presencia de, al menos, 10 fragmentos de menor intensidad de hibridación, todos ellos de mayor tamaño que el anterior. Interpretamos que este elevado número de bandas no se ha producido por una falta de especificidad de la sonda, ni tampoco a que las condiciones de hibridación sean demasiado permisivas. Dado que la sonda contiene dos tipos de secuencias repetitivas, que se encuentran en numerosos genes, puede comprenderse el motivo de esta hibridación en muchas posiciones del genoma. Estas secuencias repetitivas son la homeobox, de la que hay alrededor de 40-50 copias en el genoma de *Drosophila melanogaster* y las repeticiones del codón CAG, presentes en varios centenares de veces (McGINNIS, y otros, 1984). Por tanto, concluimos que, en estas condiciones de hibridación, la sonda utilizada no sólo hibrida con los

fragmentos correspondientes al gen Antp de nuestra especie sino, también, con fragmentos de otros genes que contengan homeobox y/o repeticiones del codón CAG.

El control con DNA de *D.melanogaster* (fig. 6A, muestra M) revela, en las mismas condiciones de hibridación, una mancha que corresponde a la hibridación de la sonda en numerosos lugares del genoma. El mayor número de fragmentos (que dan lugar a una mancha), así como la mayor intensidad de hibridación, los atribuimos a la mayor homología general de la sonda con el DNA del genoma de donde procede.

Cabe plantearse si la información que hemos obtenido es suficiente para proceder al screening de la genoteca. A continuación presentamos las consideraciones que nos han llevado a pensar que sí es suficiente.

Antes de abordar el screening, sería deseable determinar un nivel de estringencia de hibridación que únicamente permitise detectar las secuencias del gen Antennapedia y eliminase estas hibridaciones con otras regiones del genoma. Sin embargo, no hemos intentado conseguir tal "limpieza" de señal porque pensamos que sería imposible encontrar un punto en que la sonda únicamente hibride con los fragmentos que contienen los exones de Antp en nuestra especie. Nuestro fundamento es que, en *D.melanogaster*, algunas secuencias homeobox son muy similares a la homeobox del gen Antennapedia, como es el caso de la homeobox del gen fushi tarazu. No hay ningún motivo para pensar que la situación sea diferente en *D.subobscura*, de tal forma que debe ser

prácticamente imposible eliminar la hibridación de la sonda con el fragmento genómico que contenga la homeobox del homólogo de ftz en *D.subobscura* sin eliminar, a la vez, la hibridación con el fragmento que incluya la homeobox de *Antp*. A esta consideración se une un razonamiento similar sobre la hibridación de la sonda con las regiones que incluyen repeticiones del codón CAG (repeticiones M).

En vez de buscar ese punto, tal vez inexistente, nos hemos decidido a valorar las diferentes intensidades de hibridación de las señales obtenidas en el Southern genómico. En la figura 6A se aprecia que el fragmento de 1.7 kb tiene una intensidad claramente superior al resto de los fragmentos. Hemos considerado que, precisamente, es este fragmento *EcoRI* el que probablemente contenga la mayor parte de los exones que codifican el producto protéico del gen *Antp* de *D.subobscura*. Nuestro esfuerzo se ha centrado, principalmente, en la clonación de este fragmento.

Decidimos, pues, seleccionar de la genoteca aquellos clones que produjeran las señales de hibridación más intensas, en comparación con otros que, en el mismo screening, muestren señales más débiles. Los clones positivos de intensidad moderada corresponderían, mayoritariamente, a otros genes con homeobox, así como a genes que incluyan repeticiones del codon CAG en su ORF.

Cabe plantearse si es posible detectar alguna señal de hibridación que corresponda a los exones no traducidos 1, 2 y 4. Nosotros nos inclinamos por pensar que no es posible, dado

el bajo nivel de conservación esperable en las regiones no traducidas, en comparación con las secuencias que codifican el producto protéico. En cualquier caso, la débil señal de hibridación que pudiera existir no sería diferenciable de las señales debidas a la hibridación de la sonda con otros genes (a través de la HB y de las repeticiones M).

2. AISLAMIENTO Y CARACTERIZACION DE CLONES GENOMICOS DE LA LIBRERIA DE D.subobscura

Teniendo en cuenta el resultado y las consideraciones del apartado anterior, hemos llevado a cabo la búsqueda de clones en una librería genómica de D.subobscura, preparada en el vector lambda EMBL 4 (tamaño medio del inserto: 16 kb) por G. Marfany y R. González-Duarte (Universidad de Barcelona). (Se trata de una librería amplificada a partir de una librería primaria que contenía 200.000 clones recombinantes). Con este fin, hemos infectado células de la cepa NM538 de E.coli con esta genoteca. A partir de cada placa de cultivo con las placas de lisis, preparamos dos replicas en filtros de nylon. La hibridación de estos filtros se llevó a cabo con la misma sonda que se empleó para la detección de los fragmentos homólogos en el DNA genómico.

Las condiciones de hibridación durante el screening han sido idénticas a las descritas en el apartado anterior y el

número total de clones analizados en el screening es de 100.000. Asumiendo que el tamaño del genoma de *D.subobscura* es el mismo que el de *D.melanogaster* (165.000 kb), este número de clones corresponde a unos 10 equivalentes genómicos. El cultivo de los clones para el primer screening se llevó a cabo con una densidad de 15.000 clones por placa (de 90 mm de diámetro). Únicamente las señales que aparecieron en los dos filtros de una placa dada fueron aisladas para una segunda ronda de hibridación, donde la densidad del cultivo fue de unos 200 clones por placa.

El resultado de esta segunda hibridación ha sido la obtención de un total de 12 clones con señal positiva reproducible, de los cuales 3 destacan por tener una intensidad de hibridación superior a los demás. Los tres clones positivos de señal más fuerte (λ DsA1, λ DsB1 y λ DsC3) han sido el principal objeto de nuestro interés.

2.1. Identidad de los clones: localización cromosómica y secuenciación parcial

Con el fin de determinar cuantos tipos de clones han sido aislados, hemos llevado a cabo un análisis preliminar de los clones obtenidos. Efectuamos el aislamiento de DNA de los mismos, partiendo de minipreparaciones de lisados líquidos, y realizamos un primer análisis de restricción. Aunque más

adelante nos referiremos in extenso al resultado de este análisis, resaltaremos aquí que pudimos detectar, en el clon lambda DsC3, la presencia de dos fragmentos EcoRI de tamaños muy similares (1.8 y 1.7 kb) al que habíamos detectado en el Southern con DNA genómico de *Drosophila* (cuyo tamaño lo habíamos estimado en 1.7 kb). Consideramos, pues, que uno de ellos podría ser el mismo fragmento, por lo que decidimos centrar nuestro primer análisis en este clon.

Dado que la sonda utilizada en el screening es un clon de cDNA del gen *Antp* de *D.melanogaster*, sería esperable que estos clones positivos, especialmente el lambda DsC3, contengan exones del gen *Antp* de *D.subobscura*. Sin embargo, caben otras posibilidades, que no pueden descartarse a priori. Por ejemplo, un fragmento de *D.subobscura* que sea muy rico en repeticiones CAG podría tal vez dar una señal muy fuerte, tanto en la hibridación con DNA genómico como en la hibridación de las réplicas en filtro de la genoteca. Por este motivo, antes de iniciar un análisis exhaustivo de los clones obtenidos, hemos considerado conveniente efectuar una comprobación que permita, en poco tiempo, confirmar que los clones de mayor intensidad corresponden, en efecto, al gen *Antennapdia*.

Decidimos efectuar dos tipos de experimentos para confirmar la identidad de los clones, al menos la del clon lambda DsC3: (1) localización cromosómica y (2) secuenciación parcial.

En primer lugar, se llevó a cabo una hibridación in situ

en los cromosomas politénicos de *D.subobscura*, utilizando como sonda el clon de *D.melanogaster* pDm903. Con este experimento pudimos determinar que la región homóloga de *D.subobscura* se encuentra en la posición 80D del cromosoma 0 (el homólogo al cromosoma 3R de *D.melanogaster*; ver fig. 3). Un experimento similar utilizando como sonda nuestro clon aislado lambda DsC3 demostró que corresponde exactamente a la misma región cromosómica, lo que sugiere que, en efecto, se trata del gen Antennapedia. Sin embargo, el complejo Antennapedia (ANT-C) incluye 9 genes que contienen secuencias homeobox, así como repeticiones M. Por ello, no se puede descartar que la sonda de *D.melanogaster* hibridase en la posición cromosómica del ANT-C de *D.subobscura* y que nuestro clon DsC3 fuera cualquier gen del complejo.

El análisis más concluyente es el de secuenciación, por lo que decidimos seguir una estrategia para determinar parte de la secuencia del clon lambda DsC3, por ser éste el que contiene fragmentos EcoRI que pudieran corresponder al fragmento que habíamos detectado en el genoma. Para ello, utilizamos el DNA purificado a partir de minipreparaciones de lisados líquidos de este clon. Efectuamos una digestión Sau3A, seguida de transferencia Southern e hibridación con la sonda original, con el fin de determinar la presencia de fragmentos de hibridación positiva que tuviesen un tamaño apropiado para la subclonación en el vector M13.

Una vez determinada la presencia de varios fragmentos positivos, preparamos una minilibrería de fragmentos Sau3A

insertados en el sitio BamHI del vector M13mp10 (cultivada sobre JM105). Esta minilibrería fue preparada mediante la técnica de la perdigonada. Con el fin de seleccionar los fragmentos responsables de la hibridación con la sonda, aislamos ssDNA de 18 clones, escogidos al azar, y efectuamos una hibridación en dot blot de dichos DNAs con la sonda de *D.melanogaster*. El resultado fue la identificación de dos clones positivos: tDsC.001 y tDsC.013.

Estos subclones fueron secuenciados mediante el método de SANGER, siguiendo el protocolo habitual de secuenciación de moldes de cadena simple. La comparación de las secuencias obtenidas con la secuencia publicada del gen *Antp* de *D.melanogaster* (SCHNEUWLY y otros, 1986; LAUGHON y otros, 1986) demostró una similitud evidente con la secuencia de dicho gen, descrita por estos autores. Con la ayuda del ordenador, la secuencia del clon tDsC.001 puede alinearse frente al exón 7 del gen *Antp* de *D.melanogaster*. Igualmente, la secuencia del clon tDsC.013 se puede alinear frente al extremo 3' del exón 5 (ver fig. 8; flechas sin ninguna marca en su origen).

Así pues, con estos resultados, consideramos definitivamente comprobado que al menos el clon lambda DsC3 contiene parte del gen de *D.subobscura* homólogo al *Antp* de *D.melanogaster*. Tras esta comprobación, decidimos iniciar la caracterización de este clon y de todos los que presentasen regiones solapadas con éste.

2.2. Localización de los segmentos que contienen los exones

Una vez conocida la identidad de, al menos, el clon lambda DsC3, hemos procedido a purificar a gran escala el DNA de este clon, así como el de los clones DsA1 y DsB1, con el fin de proceder a un análisis de restricción más detallado y a la identificación de los segmentos codificantes.

El análisis con EcoRI (fig. 6B, izquierda) revela que los clones lambdaA1 y lambdaB1 tienen tres fragmentos EcoRI comunes, cuyos tamaños son 10.5, 2.2 y 1.0 kb. Cada uno de estos dos clones contiene, además, un fragmento adicional de 0.5 (clon A1) y 0.8 kb (clon B1). La digestión del clon lambdaC3 con EcoRI produce un patrón que no tiene ninguna banda en común con A1 ni con B1: 6.0, 3.6, 1.8, 1.7, y 0.4 kb. En esta figura mostramos una separación electroforética de las digestiones EcoRI de estos tres clones. Además de los fragmentos que hemos enumerado, aparecen también dos fragmentos comunes a todos los clones, de 20.3 y 9.2 kb, que corresponden a los brazos del vector EMBL 4.

El DNA de estos tres clones, digerido con EcoRI, fue transferido a un filtro de nylon por el método de Southern e hibridado con la misma sonda utilizada en los experimentos antes descritos. El resultado de esta hibridación se presenta en la figura 6B (derecha). Este análisis demuestra que los fragmentos de 2.2 y de 1.0 kb (marcados con puntas de flecha),

comunes a los clones lambdaA1 y lambdaB1, dan una fuerte intensidad de hibridación con esta sonda de cDNA de *Drosophila melanogaster*. Además, el fragmento de 10.5 kb (marcado en la figura), también común a los dos clones produce una señal muy débil, sólo visible en exposiciones autorradiográficas prolongadas. En cuanto al DNA del clon lambdaC3, el fragmento de 1.8 kb produce una señal de gran intensidad, pudiendo también observarse una hibridación muy débil en el fragmento de 6.0 kb (ambos fragmentos están marcados en la figura).

Interpretar el resultado de esta hibridación y determinar el contenido de cada uno de estos fragmentos de hibridación positiva ha sido un objetivo fundamental para conocer cómo se organiza la unidad de transcripción *Antennapedia* en *D. subobscura*. Las páginas que vienen a continuación describen los experimentos y razonamientos que nos han permitido asignar exones específicos a cada uno de los fragmentos. Para llegar a algunas de las conclusiones que presentamos a continuación es inevitable recurrir a algunos datos de experimentos que describiremos en otras secciones de este capítulo. Por este motivo, a veces no se encontrará en la lectura un orden perfecto de experimentos que conduzcan directamente a las conclusiones dadas.

El clon DsC3 contiene los exones 5, 7 y, probablemente, el 6

La notable coincidencia entre, por un lado, el tamaño del

fragmento EcoRI del clon lambdaC3 que hibrida fuertemente con la sonda (1.8 kb) y, por otro lado, el tamaño del fragmento detectado en el DNA genómico (que habíamos estimado en 1.7 kb, ver apartado 1 de RESULTADOS) hace pensar que se trata realmente del mismo fragmento. Por otra parte, los datos parciales sobre la secuencia del clon lambda DsC3 nos llevan a interpretar que es este fragmento clonado de 1.8 kb el que contiene, al menos, los exones 5 y 7 del gen (y, presumiblemente, el 6).

Además de los exones 5, 6 y 7, la sonda de cDNA contiene otro exón que codifica para el producto protéico. Se trata del exón 8 que, en *D.melanogaster* se encuentra separado del exón 7 por un intrón de 10-11 kb. A continuación presentamos evidencias de que esta organización también se presenta en *D.subobscura*.

Los clones DsA1 y DsB1 contienen el exón 8, que incluye la homeobox

Si el clon lambda DsC3 contiene los exones 5-6-7, que codifican para la mayor parte del producto protéico, y siendo que no parece que este clon tenga ninguna región en común con los clones DsA1 y DsB1, cabe plantearse cual es el contenido de los dos fragmentos positivos de estos dos clones. Dado que, aparentemente, estos clones no comparten ningún fragmento de restricción con el clon lambda DsC3, sólo caben dos

alternativas: o bien (1) contienen otros exones del gen Antennapedia; o bien (2) contienen secuencias que, aún hibridando con la sonda, no corresponden a este gen, sino que la hibridación se produciría con uno de los motivos repetitivos antes mencionados (repeticiones CAG y secuencias homeobox).

En caso de tratarse de la alternativa 1, consideramos que debería tratarse de un exón de los que codifican el producto protéico. Pensamos que los exones no traducidos, aunque tengan algún segmento en común entre las dos especies, esto no es suficiente para producir una señal con la intensidad que observamos. Dado que ya hemos asignado la hibridación de los exones 5 y 7 (y, presumiblemente, también el pequeño exón 6) de la sonda al clon lambda DsC3, el único exón que codifica producto protéico que queda en la sonda es el exón 8. Ya que en el gen Antp de *D.melanogaster* el exón 8, que contiene la secuencia homeobox, está separado de los exones 4-5-6-7 por 10-11 kb de DNA intrónico, hemos supuesto que uno (o ambos) de los fragmentos positivos de los clones lambda DsA1 y DsB1, se deba a la hibridación con este exón de la sonda. Esto explicaría también la aparente distancia física entre estos dos clones y el lambda DsC3 (a juzgar por las diferencias en sus patrones de restricción con EcoRI).

Con el fin de verificar este extremo, hemos llevado a cabo una hibridación, en un dot blot, de muestras de DNA de estos tres clones de lambda utilizando como sonda específica un plásmido que contiene la secuencia homeobox del gen Antp de

D.melanogaster (sonda pDmA2015). El resultado de este experimento fue la detección de la presencia de una secuencia homeobox en los clones lambda DsA1 y lambda DsB1, si bien todavía no se puede excluir que esto corresponda a la homeobox de un gen distinto de Antp.

Los clones DsA1 y DsB1 contienen, además un fragmento común con el clon DsC3)

Los experimentos llevados a cabo posteriormente (mapas de restricción y secuenciación) demostraron que uno de los fragmentos positivos de los clones DsA1 y DsB1 (fragmento de 2.2 kb) contiene el exón 8 del gen; mientras que el otro (fragmento de 1.0 kb) contiene los exones 6, 7, y parte del exón 5. Así, aunque inesperado, los clones lambda DsA1 y DsB1 sí poseen regiones solapadas con lambda DsC3. Los apartados que vienen a continuación muestran los experimentos que hemos llevado a cabo para llegar a estas conclusiones, así como otras relativas a la organización de Antp en la región de *D.subobscura* que hemos clonado.

Nos planteamos si existe correspondencia entre los segmentos de DNA que hemos clonado y el genoma de *Drosophila subobscura*. Ya hemos mencionado que la banda de 1.8 kb del clon lambda DsC3 se corresponde, según nuestra interpretación, con aquella banda genómica cuyo tamaño habíamos estimado en 1.7 kb. Nos preguntamos si, de la misma forma, hay un

paralelismo entre todos los demás fragmentos positivos clonados y los fragmentos positivos que observamos al analizar el DNA genómico de *D.subobscura* (figs. 6A y 6B). En efecto, hay una correspondencia casi perfecta, salvo una excepción. La banda positiva del clon DsC3 (fragmento EcoRI de 6 kb), que proporciona una señal de hibridación muy débil, corresponde probablemente a una de las bandas detectadas en la región de alto peso molecular del DNA genómico.

En lo que respecta a los fragmentos EcoRI de los clones de lambda DsA1 y DsB1, la situación no es la misma. El fragmento de 2.2 kb, común a estos dos clones, sí puede corresponder a un fragmento visible en el Southern con DNA genómico de *D.subobscura*. Sin embargo, el fragmento de 1.0 kb de estos clones (que produce una señal de hibridación muy fuerte) no aparece, de modo alguno, en el DNA genómico. La explicación a esto no se ha podido encontrar hasta que hemos elaborado los mapas de restricción de los tres clones. Con los mapas de restricción elaborados, se encuentra que el fragmento EcoRI de 1.0 kb corresponde al extremo izquierdo de los clones lambda DsA1 y DsB1. Por este motivo, este fragmento sólo contiene un sitio EcoRI genómico, siendo el segundo un sitio EcoRI que pertenece al vector. El análisis de restricción aclarará definitivamente la situación: el fragmento de 1.0 kb de los clones DsA1 y DsB1 no es más que una parte del "auténtico" fragmento EcoRI genómico de 1.8 kb (incluido en el clon lambda DsC3).

2.3. Mapas de restricción de los clones aislados

Con el fin de disponer de suficiente DNA de cada uno de los fragmentos EcoRI de los tres clones de lambda, hemos subclonado todos estos fragmentos en el sitio EcoRI de los plásmidos pUC18 o pUC19. Con este fin, una mezcla de la digestión EcoRI de cada clon de lambda fue ligada con el vector defosforilado en un experimento de "perdigonada" (aunque la mezcla de digestión contiene los brazos del vector EMBL 4, estos fragmentos no son clonables). El resultado de esta ligación se utilizó para la transformación de células competentes de E.coli. Preparando de 12 a 18 clones recombinantes (tomados al azar) hemos conseguido, salvo en casos excepcionales, todos los subclones necesarios. Este proceso lo hemos aplicado, sistemáticamente, a cada uno de los clones del fago lambda que hemos aislado y cuya pertenencia al gen Antp hemos podido demostrar.

El mapa de restricción de los tres clones de lambda del screening original lo hemos obtenido combinando digestiones simples y dobles con los enzimas EcoRI, BamHI y Sall. Para asignar los sitios de restricción a los fragmentos específicos se han realizado digestiones con los enzimas BamHI y Sall de aquellos fragmentos subclonados en que ha sido necesario. Cuando ha sido preciso para eliminar incertidumbres, hemos purificado fragmentos BamHI o Sall y los hemos digerido con otro enzima de restricción.

Estos análisis de restricción han detectado la presencia de un fragmento BamHI en el clon lambda DsC3 cuyo tamaño (2.1 kb) es idéntico a otro fragmento BamHI del clon lambda DsA1. La digestión doble con los enzimas BamHI y SalI demuestra la presencia de otro fragmento común a estos dos clones (fragmento BamHI/SalI de 0.6 kb). Esto sugiere fuertemente que los clones lambda DsC3 y lambda DsA1 comparten una región común, aunque esto no sea aparente con las digestiones EcoRI.

Dicha posibilidad ha sido verificada utilizando fragmentos clonados del clon DsC3 como sonda e hibridándolos frente a digestiones de el DNA del clon DsA1 (en condiciones de la máxima estringencia). Con estas experiencias, hemos demostrado la identidad de secuencia entre esos fragmentos de tamaños idénticos. Con ello, demostramos también el solapamiento entre los clones lambda DsC3 y DsA1. Esto demuestra, así, que la homeobox cuya presencia habíamos detectado en los clones DsA1 y DsB1 (en el experimento de dot blot, con la sonda pDmA2015) es la del gen Antp de *D. subobscura*.

En la figura 7 presentamos un esquema que muestra la extensión de las regiones comunes a los clones lambda DsC3, DsA1 y DsB1. La línea central representa el DNA genómico de *D. subobscura* y sobre ésta hemos señalado los sitios de restricción genómicos, deducidos a partir de los mapas de restricción de cada uno de los clones. Las regiones que muestran hibridación frente a la sonda de cDNA de *D. melanogaster* vienen señaladas con un trazo grueso debajo de la línea de el DNA genómico. La línea inferior tiene marcados

unos rectángulos que representan las regiones donde hemos asignado exones determinados. Los fragmentos incluidos en estos rectángulos han sido el objeto de nuestro trabajo de secuenciación, que se describe más adelante.

En esta figura, además de los tres clones que ya hemos descrito, se describe el clon lambda DsU2, obtenido mediante chromosomal walk, tal como se describe en la siguiente sección.

3. EXTENSION DE LA REGION CLONADA: CHROMOSOMAL WALK

Aunque la extensión de la región clonada la hemos iniciado después de determinar la secuencia de nucleótidos de la mayoría de los exones del gen (apartado 5.1), nos referiremos antes a ella con el fin de aportar agrupados todos los datos sobre el análisis de restricción de la región.

Una vez determinado el grado de solapamiento entre varios clones aislados del screening original, y una vez conocida la organización de los exones en el segmento clonado (apartado 5.2), parece necesario extender nuestra clonación en dirección 5', con el fin de alcanzar la región que debe incluir las secuencias promotoras del gen.

En principio, dado que la sonda utilizada en el screening original contiene los exones 1 y 2, se podría esperar que algunos de los clones positivos de señal moderada que hemos

purificado pudieran contener los homólogos de estos exones en *D. subobscura*. Esto nos ha parecido improbable, dado que ninguno de estos dos exones está incluido en la ORF (el codón de iniciación se encuentra en el exón 5). Los exones que no codifican para el producto protéico no están sujetos a una presión de selección tan fuerte como los que sí codifican la proteína. Es de esperar, por tanto, que estos exones sean relativamente distintos entre las dos especies (conservando sólo los elementos esenciales). Esta ausencia de conservación (salvo algunos elementos conservados) en las regiones no traducidas de un gen es el caso habitual en cualquier comparación interespecífica.

En cualquier caso, decidimos comprobar si alguno de los clones purificados (distintos de los ya mapeados por análisis de restricción, cuyo contenido ya conocemos) contiene el exón 1. Para ello, utilizamos DNAs de estos clones, digeridos con *EcoRI* y transferidos a un filtro por la técnica de Southern, y los hibridamos con una sonda de DNA genómico de *D. melanogaster* que contiene este exón 1, derivado del promotor P1 (esta sonda es un subclón plasmídico de un fragmento *EcoRI* que corresponde a las coordenadas 199 a 203 del ANT-C; ver figs. 4 y 5). El resultado fué negativo. Pensamos que los clones de intensidad moderada pueden ser otros genes con homeobox (y/o con repeticiones M).

Es probable que si en la genoteca existen clones positivos que contengan los exones 1 y 2, la señal que estos clones proporcionen sea menor que la producida por centenares de

otros clones que tengan mayor similitud con la sonda de Antp (clones que contengan secuencias homeobox o secuencias repetidas CAG).

Así pues, dado que no disponemos de clones que correspondan a la región 5', decidimos buscar otro sistema para purificarlos. En principio, aunque la homología esperable a nivel de los exones no traducidos no sea suficiente para que los clones que los contengan resalten del ruido de fondo, una sonda específica con el exón 1 marcado (promotor P1) o bien con el exón 3 (promotor P2) debería permitir aislar los homólogos de *D. subobscura*. La situación, así, sería distinta del screening original.

Con esta idea, decidimos efectuar una búsqueda en la genoteca utilizando sondas de los exones no traducidos 1 y 3. El resultado, sin embargo, no ha sido satisfactorio: los clones aislados no han proporcionado señales reproducibles.

Con estos resultados, la única alternativa posible para el aislamiento de la región 5' del gen, y extender así la región clonada inicialmente, es la técnica del chromosomal walk.

Hemos conseguido iniciar esta extensión efectuando un nuevo rastreo de la genoteca, utilizando ahora como sonda un fragmento subclonado del clon lambda DsC3. Se trata del fragmento EcoRI de 1.7 kb que, en el mapa de restricción de DsC3 queda más a la izquierda (ver fig. 7). Este fragmento fué marcado con digoxigenina y se utilizó para la búsqueda de clones positivos en la genoteca, siguiendo la misma metodología del screening original, salvo que las condiciones

de hibridación fueron de máxima estringencia.

Esta búsqueda proporcionó tres clones positivos que, al analizar su DNA con varios enzimas de restricción, resultaron ser idénticos (a todos ellos les hemos denominado clon DsU2). Dicho clon fue mapeado por análisis de restricción; las regiones de solapamiento con el clon lambda DsC3 han sido determinadas mediante hibridaciones cruzadas de DNAs de ambos clones. Este solapamiento con el clon lambda DsC3 ha sido deducido de los mapas de restricción, así como de las hibridaciones cruzadas entre ambos clones. La extensión de este clon sobre el DNA genómico puede observarse en la figura 7.

4. MAPA DE RESTRICCIÓN DE LA REGIÓN CLONADA

En total, todos los clones que hemos caracterizado hasta la fecha abarcan 31 kb de DNA genómico de la región 3' del gen *Antennapedia* de *Drosophila subobscura*. Esto supone, aproximadamente, un tercio de la longitud total del gen, si asumimos que el tamaño de este gen en las dos especies de *Drosophila* es similar. Esta región incluye los exones 4, 5, 6, 7 y 8 cuya disposición, a grandes rasgos, describimos en la figura 7. En esta misma figura hemos representado también la extensión de cada uno de estos exones sobre el mapa de restricción del genoma. Algunos de los detalles descritos en

esta figura los conocemos a través del posterior análisis de las secuencias, descrito más adelante. Sin embargo, hemos preferido agrupar información procedente de distintos experimentos en esta figura.

La polaridad de las coordenadas utilizadas en esta figura es la misma que la usada por SCOTT y otros (1983) para el Complejo Antennapedia de *D.melanogaster*, y opuesta a la utilizada por SCHNEUWLY y otros (1986). El punto de referencia (100 kb) que hemos tomado es el centro del fragmento EcoRI que contiene la secuencia homeobox. Las regiones del genoma que dan señal positiva al hibridar DNAs de los clones con una sonda de cDNA de *D.melanogaster* las hemos representado con una línea negra bajo el mapa de restricción.

A grandes rasgos, la estructura del gen en nuestra especie (fig. 7) no muestra grandes diferencias respecto de *D.melanogaster*: el tamaño del intrón que separa los exones 4-5-6-7 del exón 8 es, dentro del error de la medida, el mismo: unas 10-11 kb; el intrón 3, que precede a los exones 4-5-6-7 es superior a 15 kb, dado que no hemos detectado la presencia de ningún exón entre las coordenadas 128 kb y 113 kb del genoma de *D.subobscura* (a juzgar por la no hibridación de una sonda de DNA genómico que incluye el exón 3 de *D.melanogaster* frente al DNA de dicha región). Todavía desconocemos el tamaño de este intrón 3, ya que aún no disponemos de clones que incluyan el exón 3.

Esta conservación de la estructura general no significa que no se hayan producido variaciones a menor escala en la

organización del gen, tal como veremos en los siguientes apartados.

5. SECUENCIACION DE LOS EXONES DE LA REGION

Los datos parciales de la secuencia del clon lambda DsC3, junto con la identificación del fragmento EcoRI de 1.8 kb como la posible región de máxima conservación entre el genoma de *D.subobscura* y la sonda de cDNA de *D.melanogaster*, hacen de dicho fragmento el objeto de nuestro interés para el inicio de una secuenciación de los exones. Con el propósito de obtener la secuencia completa de este fragmento, hemos combinado varias aproximaciones en nuestra estrategia de secuenciación. También hemos secuenciado el fragmento de 1.0 kb del clon lambda DsA1. Este fragmento corresponde, como hemos dicho anteriormente, a una parte del fragmento de 1.8 kb del clon DsC3.

Además de la secuencia de esta región, presentaremos datos preliminares de la secuencia del exón 8 (fragmento EcoRI de 2.2 kb del clon lambda DsA1).

5.1 Estrategia de secuenciación

Todos los fragmentos EcoRI de los clones de lambda

originales, aislados de la genoteca, y que han dado señal de hibridación con la sonda de cDNA, han sido estudiados mediante secuenciación. Los subclones plasmídicos de estos fragmentos han sido secuenciados siguiendo las adaptaciones del método de SANGER para la secuenciación de DNA dúplex.

Nuestro trabajo se ha iniciado, en un principio, en los fragmentos que han dado las señales más fuertes: el fragmento de 1.8 kb del clon lambda DsC3 y el de 1.0 kb del lambda DsA1. Los correspondientes subclones plasmídicos son, respectivamente, pDsC.07 y pDsA.14. El clon pDsC.07 ha sido secuenciado en su práctica totalidad y el pDsA.14 lo suficiente para comprobar su solapamiento con la secuencia determinada en el clon anterior.

La estrategia de secuenciación que hemos seguido está descrita en la parte inferior de la figura 8. Hemos combinado los resultados de los primeros experimentos de secuenciación, llevados a cabo sobre clones del vector M13, con la secuencia obtenida a partir de clones plasmídicos.

La secuencia de los extremos del inserto del clon pDsC.07 se ha obtenido utilizando el primer universal directo y el primer reverso. Además, hemos obtenido un total de 5 puntos de acceso a la secuencia valiéndonos de un sitio de restricción HindIII, dos sitios PstI y un sitio SmaI. Las secuencias obtenidas a partir de estos sitios de restricción están representadas con una flecha que, en su origen, tiene una línea vertical. Estos sitios de restricción vienen señalados, en la figura 8, sobre el esquema de *D.subobscura*.

Para obtener la secuencia a partir de estos sitios de restricción buscamos, en primer lugar, sitios de restricción que se encontrasen a la vez en el inserto y en el polilinker del vector plasmídico utilizado (pUC18). A continuación, efectuamos una digestión del plásmido con el enzima de restricción correspondiente y, tras verificar la digestión completa, el enzima fué inactivado. Posteriormente, efectuamos una religación del producto de cada digestión, consiguiendo así una reordenación y/o una delección parcial del inserto. Tras la transformación de una cepa apropiada de E.coli con cada una de estas ligaciones, obtuvimos tres "librerías", una por cada uno de los enzimas SmaI, HindIII y PstI. Decidimos preparar DNA de 5 clones de cada librería y llevar a cabo las reacciones de secuenciación. El resultado fué la obtención de 5 puntos de entrada diferentes, a partir del total de 15 clones secuenciados.

Las secuencias obtenidas a partir de clones del vector M13 (ver apartado 2.1) se representan por una flecha sin ninguna marca en su origen.

Con el fin de completar la secuencia del clon pDsC.07, y aprovechando la información obtenida a partir de nuestras secuencias, hemos diseñado dos oligonucleótidos, que nos han servido para completar la mayor parte de la secuencia del plásmido pDsC.07. Las secuencias obtenidas con estos oligos vienen representadas con sendas flechas en cuyo origen hay dibujada una estrella de ocho puntas (fig. 8).

La secuencia determinada con el oligonucleótido No. 2

(apartado 15 de MATERIALES Y METODOS) ha sido necesaria para completar la elaboración de una secuencia continua (contig) con el ordenador. Aunque en la estrategia de secuenciación presentada en la figura 8 no parece necesario el uso del oligo, su uso se justifica por el hecho de que la región donde todas las demás secuencias se solapan es una región con un gran número de repeticiones del codón CAG. Por este motivo, el solapamiento entre secuencias que encuentra el ordenador no tiene validez para considerar que es posible generar un contig a ese nivel.

El clon pDsC.07 se extiende a lo largo de un segmento de DNA situado entre los dos sitios EcoRI de la figura 8. Corresponde a un fragmento EcoRI (1.8 kb) del genoma de *D.subobscura*, que habíamos detectado en un Southern con DNA total (apartado 1). No es este el caso del clon pDsA.14: se trata de un fragmento EcoRI (1.0 kb) que está situado en un extremo del clon lambda DsA1 (ver fig. 7). Por este motivo, sólo uno de los sitios EcoRI del fragmento clonado corresponde realmente a un sitio EcoRI del genoma de *D.subobscura*; el otro sitio EcoRI procede del polilinker del vector lambda EMBL 4. El extremo izquierdo del inserto del clon pDsA.14 corresponde, entonces, a un sitio de restricción para el enzima MboI (perteneciente al genoma de *D.subobscura*) situado en la posición 298 (exón 5) de nuestra secuencia (fig. 10). Este punto está marcado en la estrategia de secuenciación con una estrella de cinco puntas (fig. 8). El extremo derecho del clon pDsA.14 es el mismo sitio EcoRI del extremo derecho de

pDsC.07.

El hecho de que uno de los sitios EcoRI del inserto del clon pDsA.14 sea artificial explica el hecho de que este mismo fragmento, cuando se observa en la figura 6B del Southern de lambda, aún dando una señal de hibridación muy fuerte, no sea observable en el Southern genómico (fig. 6A).

Hemos determinado la secuencia de los extremos de dicho clon (utilizando los primers directo y reverso). La secuencia obtenida es suficiente para demostrar que el fragmento EcoRI que contiene es, en efecto, parte del fragmento EcoRI del clon pDsC.07.

Por último, hemos determinado la secuencia de los extremos del otro fragmento EcoRI del clon lambda DsA1 que también dió señal positiva al hibridar con la sonda de cDNA de *D.melanogaster*. Se trata del fragmento de 2.2 kb, cuyo subclón plasmídico es pDsA.02. En la última parte del trabajo demostramos que se trata del exón 8 del gen, que contiene la secuencia homeobox y la mayor parte del segmento final no traducido de los transcritos (trailer). Dado que no hemos secuenciado la totalidad de este fragmento, la secuencia protéica que presentaremos más adelante no incluye la correspondiente a los 70-80 aminoácidos del extremo carboxiterminal de la proteína Antp de *D.subobscura*.

5.2. Asignación de intrones y exones: descripción de la ORF

Siguiendo la estrategia que hemos descrito, obtuvimos la secuencia completa del fragmento EcoRI de 1.8 kb (pDsC.07), que incluye los exones 4-5-6-7, los intrones 5 y 6, así como unos 100 pb del intrón 4. El alineamiento por ordenador de nuestra secuencia con la de *D.melanogaster*, ha permitido determinar los límites entre los intrones y los exones del gen *Antp* en nuestra especie. Esto nos ha permitido deducir la pauta de lectura y efectuar la traducción conceptual.

El primer análisis que hemos llevado a cabo ha sido una comparación con la secuencia de *Drosophila melanogaster* mediante una matriz de puntos, calculada con uno de los programas ofrecidos por DNASTAR. En la figura 9 presentamos el resultado obtenido aplicando las opciones de partida del programa. Se observa una diagonal, que representa una homología buena, lo que ofrece la posibilidad de efectuar el alineamiento de ambas secuencias. La diagonal está interrumpida por secuencias que no demuestran homología. En la sección 5.4 discutiremos ampliamente el resultado obtenido con este análisis. En esta sección sólo nos referiremos a la información que esta gráfica ha proporcionado para la determinación de las posiciones de los intrones y de los exones, así como para la posterior aplicación de los programas de alineamiento de secuencias.

Como hipótesis de partida, proponemos que las regiones

conservadas (que forman una diagonal en la matriz de puntos) son los exones de la región; las regiones no conservadas son, en principio, los intrones. A continuación presentaremos las evidencias que hemos utilizado para asignar las posiciones de los intrones y exones, así como para determinar cual es la región traducida y el producto protéico.

En la figura 10 presentamos la secuencia obtenida, alineada frente a la secuencia de *D.melanogaster* con la ayuda de los programas de alineamiento de DNASTAR. La secuencia de *D.melanogaster*, que ha sido extraída de la base de datos EMBL Data Library, es un fragmento genómico que contiene los exones 4-5-6-7, así como los intrones que los separan. Sus referencias son idénticas a las dadas en el artículo original que la describe (SCHNEWULY y otros, 1986). Sobre la secuencia de *D.melanogaster* está la secuencia de *D.subobscura* que hemos determinado nosotros. La identidad entre nucleótidos está señalada con ":" y los huecos introducidos por el ordenador durante el alineamiento se representan con "-". Téngase presente que, debido a un vacío de unos 100 pb de nuestra estrategia de secuenciación, la secuencia que presentamos tiene una discontinuidad a nivel del intrón 4 (ver fig. 8). Esta discontinuidad, representada por una línea horizontal, hace que presentemos dos secuencias: una corta, de 325 pb, que corresponde al extremo izquierdo de nuestro clon; y otra larga, de 1414 pb, que incluye la casi totalidad del resto del clon.

Las regiones con repeticiones CAG están marcadas con una

línea de trazo grueso (por ejemplo, posiciones 535 a 582). Los sitios de restricción utilizados para la clonación o como puntos de partida en la estrategia de secuenciación vienen señalados sobre la secuencia. El significado de los demás signos se especificará conforme discutamos aspectos concretos de las secuencias.

La secuencia protéica que hemos deducido para *D. subobscura* es un dato que se obtiene después de los razonamientos que presentamos en los párrafos siguientes. Sin embargo, hemos preferido representar la secuencia protéica sobre este dato "primario" de alineamiento, con el fin de facilitar el examen de las conclusiones que presentamos.

Para determinar cuál es la región traducida, un requisito previo es comprobar si la posición de los intrones se ha conservado en las dos especies. Los puntos de transición entre intrones y exones están señalados en la secuencia (fig. 10) con flechas verticales. En la Tabla 4 describimos la secuencia de los sitios dadores y aceptores de los intrones de *D. subobscura*, en comparación con los de *D. melanogaster*. A continuación mostraremos las pruebas de que las posiciones de los intrones son, en efecto, las que indicamos en la figura 10 y en la Tabla 4.

Para asignar los límites de los intrones que no interrumpen la ORF, el criterio que hemos seguido es la coincidencia de dos hechos: (1) que sea una región en la que el ordenador haya presentado alineamiento entre las dos secuencias; (2) que esta región coincida con el límite

intrón/exón de *D.melanogaster*, aunque este punto no coincide con las secuencias consenso.

Para asignar los límites de los intrones que interrumpen la ORF hemos hecho uso de una evidencia adicional: la gran homología entre estas dos especies, a nivel de los exones traducidos. Los límites que presentamos son verosímiles, según las evidencias que presentamos a continuación (fig. 10). Por un lado, la homología entre nuestra secuencia y la de *D.melanogaster* es muy desigual: hay segmentos que comparten un grado muy alto de similitud, interrumpidos por segmentos sin similaridad aparente, tal como se observa en el alineamiento de secuencias de la figura 10 y en la matriz de puntos de la figura 9 (en la matriz se observan interrupciones en la diagonal de homología). Los puntos de transición entre las regiones similares y las no similares coinciden perfectamente con los límites intrón/exón descritos en *D.melanogaster*. De ahí nuestra interpretación de que los segmentos conservados son los exones y de que los segmentos no conservados son los intrones.

En la Tabla 4 presentamos una lista con todos los sitios dadores y aceptores que hemos deducido. Las discrepancias de nuestros sitios dadores y aceptores con las secuencias consenso son discrepancias en común con *D.melanogaster*.

Los tres intrones que siguen a los tres exones traducidos (exones 5, 6 y 7) contienen codones de parada en fase con la pauta de lectura que proponemos, como es habitual en los intrones (WATSON y otros, 1987). Así, el intrón 5 tiene un

codón TAA en la posición 874 y un codón TAG en la posición 880; el intrón 6 contiene un codón TAA en posición 1104; el intrón 7 tiene un codón TGA en la posición 1405. Con esto concluimos que los intrones se sitúan en posiciones homólogas en las dos especies.

Una evidencia más de que la secuencia que presentamos como la ORF de *D.subobscura* es la correcta radica en que las regiones de homología presentan un esquema general de conservación que puede describirse como ++-/++-/++-. Dicho de otra forma, las diferencias nucleotídicas tienden a aparecer cada tres pares de bases, dentro de las regiones que presentamos como pertenecientes a la ORF. Esto sugiere muy fuertemente que se trata de sustituciones sinónimas en tercera posición de los codones, que no afectan al producto protéico codificado.

Las pruebas que hemos presentado hasta ahora permiten deducir cuál es la pauta de lectura en *D.subobscura*. El último dato necesario para determinar la ORF es conocer el sitio de inicio de la síntesis protéica. En principio, la coincidencia entre, por un lado, el punto donde comienza la homología significativa en el exón 5 y, por otro lado, el sitio de inicio de la síntesis protéica en *D.melanogaster* hace pensar que en *D.subobscura* la síntesis de la proteína se produce en la región homóloga (posición 193 del exón 5). Sin embargo, no puede descartarse a priori que el sitio de inicio de la traducción se encuentre en un punto desplazado en dirección 5' (dando lugar a una proteína Antp con una "extensión" de su

extremo aminoterminal). Esta posibilidad, la descartamos por el hecho de que no hay ningún codón ATG ni en el exón 5 ni en el exón 4.

El sitio que proponemos como el de inicio de la traducción para *D.subobscura* (CCGCAATG) se ajusta bien a la secuencia consenso para los codones de iniciación eucarióticos (KOZAK, citado en SCHNEUWLY y otros, 1986): CC(A/G)CCATG. El sitio de inicio de la traducción de *D.melanogaster* (CCACGATG) difiere en la misma medida en que lo hace el de nuestra especie. En ambas especies, sólo el nucleótido que precede al codón ATG está fuera del consenso.

Con este razonamiento, hemos representado la traducción conceptual de la ORF de *D.subobscura* sobre la secuencia nucleotídica de la figura 10. En aquellas posiciones en las que hay diferencias entre las dos especies, hemos escrito la secuencia protéica de *D.melanogaster*. Todas aquellas posiciones donde sólo señalamos la secuencia protéica de *D.subobscura* corresponden a lugares donde no ha habido sustituciones de aminoácidos.

Concluimos que la ORF de *D.subobscura* se compone de los siguientes segmentos:

Exón 5 (parte traducida): Posiciones 193 a 855.

Exón 6 (completo): Posiciones 1002 a 1040.

Exón 7 (completo): Posiciones 1168 a 1387/1399 (en función del sitio dador que se utilice).

A esta ORF (de 934 pb; 311.3 aminoácidos) hay que añadir el exón 8, que contiene la secuencia homeobox, cuyos límites desconocemos por no haber secuenciado este exón. No obstante, en otros estudios de comparación interespecífica de genes del desarrollo se ha observado una extrema conservación a nivel del dominio homeótico. Por este motivo, asumimos que la región traducida del exón 8 (que contiene el dominio homeótico) de *D.subobscura* debe tener una longitud muy similar a la de *D.melanogaster* (248 pb; 82.6 aminoácidos).

5.3. Secuencia del producto protéico: dominios constantes y dominios variables

La secuencia de la proteína Antennapedia de *D.subobscura* (sin incluir los aminoácidos codificados en el exón 8) está representada en la figura 11. Dicha secuencia la hemos obtenido por traducción conceptual de la ORF. El peso molecular de este producto protéico es de 34.450 Dalton, en comparación con los 32.660 Dalton del producto de *D.melanogaster* (en ambos casos, excluyendo el dominio carboxiterminal, estimado en 10.140 Dalton). Se trata de una proteína predominantemente hidrofílica, con una carga neta negativa (-2.8) a pH 7.0 y un punto isoeléctrico de 6.6.

En esta figura hemos señalado los aminoácidos cargados (cuadros negros sobre los aminoácidos ácidos y cuadros blancos sobre los básicos), así como los residuos de cisteína

(incluidos en un recuadro). Las regiones ricas en glutamina, que corresponden a repeticiones del codón CAG, están subrayadas (estas regiones repetidas también se indican en la secuencia de la fig. 10). Puede observarse el elevado contenido en este aminoácido (23%), todavía superior al de la proteína Antp de *D.melanogaster* (22%). Las tres regiones codificadas por diferentes exones están separadas por flechas verticales: los aminoácidos de las posiciones 1 a 221 corresponden al exón 5; las posiciones 222 a 234 corresponden al exón 6; y las posiciones 235 a 307/311 del exón 7.

Ninguno de los aminoácidos cargados ha sido sustituido a lo largo de la evolución; tampoco han variado las posiciones del aminoácido cisteína. Muy diferente es la situación de las regiones con repeticiones del aminoácido glutamina, donde más adelante demostraremos que ha habido una gran variación.

El alineamiento por ordenador entre las dos secuencias protéicas (fig. 12) revela que la distribución de la variación dista mucho de ser homogénea. Al contrario, observamos regiones que prácticamente no han variado durante la evolución, frente a otras que han divergido en gran medida. Esto se observa mejor en una representación gráfica (fig. 13, parte inferior), que fué utilizada por primera vez por KASSIS y otros (1986). Esta gráfica representa la proteína Antp de *D.melanogaster* con un rectángulo.

La variación encontrada en nuestra especie se representa en la figura 13 de la siguiente forma: las inserciones de uno o dos codones se representan con flechas negras pequeñas y

grandes, respectivamente; las inserciones de más de dos codones se representan con un triángulo cuya base es proporcional al tamaño de la inserción; la única "delección" se representa con una flecha invertida. Nos referimos a inserciones y delecciones como una forma arbitraria de describir las diferencias en el número de aminoácidos en posiciones concretas. Hablamos de inserción cuando *D.subobscura* posee más aminoácidos que *D.melanogaster*; al hecho inverso le denominamos delección. Ninguno de estos términos se refiere, necesariamente, a inserciones o delecciones como procesos evolutivos.

Por último, la posición de los aminoácidos que han sufrido una sustitución no conservativa se representa con una línea vertical negra.

Debajo de esta representación (fig. 13) hemos marcado las regiones ricas en el aminoácido glutamina (repeticiones CAG) con recuadros vacíos. Los recuadros negros señalan las posiciones de dos péptidos conservados, que son comunes a otros genes homeóticos (estos dos péptidos están enmarcados en un recuadro en la figura 12). Uno de ellos es el péptido Tyr-Pro-Trp-Met, situado justo antes de los sitios dadores del exón 7. Este péptido no sólo es idéntico en el gen *Antp* de las dos especies, sino que también presenta la particularidad de que se encuentra en los genes *Ultrabithorax*, *Deformed*, *abdominal-A* (WILDE y AKAM, 1987; REGULSKI y otros, 1987). También hemos encontrado este péptido en el gen *labial*, aunque difiere en un aminoácido (Tyr-Lys-Trp-Met) (DIEDERICH y otros,

1989).

El otro péptido conservado tiene la secuencia Met-X-Ser-Tyr-Phe, y se sitúa cerca del extremo aminoterminal de la proteína Antp (X=Thr). El gen Ubx contiene esta misma secuencia (siendo X=Asn) justo en sus 5 aminoácidos del extremo aminoterminal.

En una comparación del gen Ubx entre tres especies de *Drosophila* y *Musca domestica*, se ha observado que ambos péptidos son idénticos en las cuatro especies. Además, ciertos genes con homeobox en el ratón contienen estas dos secuencias (WILDE y AKAM, 1987).

Comparando, por una parte, cómo se distribuye la variación de la proteína y, por otra parte, la localización de los segmentos repetidos se pone de manifiesto que la mayor parte de la variación del producto protéico de *D. subobscura* se ha producido dentro de la región de poliglutamina (repeticiones CAG). Esto podría estar de acuerdo con la idea de que la función de esta región rica en glutamina es tal vez sólo la de unir los dominios realmente funcionales.

La presencia de repeticiones de un aminoácido es una característica común en muchos genes que contienen el dominio homeótico, lo que sugiere que dichas repeticiones poseen una función común. Así, se ha propuesto que podrían actuar en forma de bisagra, uniendo otros dominios de la proteína (BEACHY y otros, 1985; LAUGHON y otros, 1985).

Los exones 5 y 7 son de mayor tamaño en nuestra especie que en *D. melanogaster* debido a la acumulación de 7 inserciones

de codones en *D.subobscura* (figuras 10 y 12). El exón 5 ha acumulado dos inserciones de 1 codón, una inserción de 2 codones, una de 5 codones, y una de 6. El exón 7 ha acumulado dos inserciones de 1 codón. La única "delección" es un codón del exón 5 de *D.melanogaster* que no aparece en la secuencia de *D.subobscura*. El resultado de estos cambios es la presencia de 16 aminoácidos más en la proteína Antp de *D.subobscura*, en comparación con *D.melanogaster*.

El total de sustituciones de aminoácidos detectadas en la proteína Antennapedia es de 14. De estas sustituciones, sólo 3 han sido de tipo conservativo (cambio de Ser a Thr, o viceversa). Hemos considerado como cambio conservativo cualquier variación de un aminoácido en otro dentro de los siguientes grupos: [Ser-Thr]; [Glu-Asp]; [Lys-Arg]; [Tyr-Phe]; [Asn-Gln]; y [Leu-Val-Ile-Ala]. Sin embargo, como hemos dicho, sólo tres sustituciones de aminoácidos son de este tipo. Las otras 11 sustituciones son de las consideradas como no conservativas.

Las sustituciones de aminoácidos se han producido en los dos exones de mayor tamaño. En el exón 5 se han producido 10 sustituciones no conservativas y 2 sustituciones conservativas (Ser-Thr); en el exón 7 se ha producido 1 sustitución conservativa de aminoácidos y 1 no conservativa (figuras 10, 12 y 13). El porcentaje de conservación calculado para la secuencia proteica de cada uno de los exones viene presentado en la Tabla 5.

Ninguna de estas sustituciones no conservativas ha

afectado a los aminoácidos ácidos ni a los básicos, ni tampoco a los residuos de cisteína. Las sustituciones que observamos vienen señaladas en la representación de KASSIS (figura 13, parte inferior) en forma de líneas verticales. Sobre esta representación, hemos situado el perfil de hidropatía de HOPP y WOODS de las proteínas Antp de las dos especies. Dicho perfil predice los dominios hidrofóbicos e hidrofílicos de una proteína en función de los valores de hidropatía de sus aminoácidos. Los dominios hidrofílicos aparecen en la mitad superior de la escala; y los dominios hidrofóbicos en la mitad inferior. En esta figura puede verse que el producto de Antp es una proteína predominantemente hidrofílica, cuya organización se ha mantenido, a pesar de las divergencias.

Hemos representado el eje horizontal del perfil de hidropatía a la misma escala que la representación de KASSIS, de tal forma que la variación de la proteína Antp puede localizarse sobre el perfil de hidropatía de *D.melanogaster*. Todas las inserciones (así como la única delección) se han producido en dominios hidrofílicos de la proteína; las sustituciones no conservativas de aminoácidos han tenido lugar tanto en los hidrofóbicos como en los hidrofílicos, con una cierta tendencia a estos últimos (4 a 6).

Como ya hemos mencionado, el alineamiento de las secuencias de DNA denota que la mayoría de las sustituciones nucleotídicas producidas dentro de la ORF se han producido en lugares silenciosos de los codones, sin afectar a la secuencia protéica. Las sustituciones nucleotídicas que han producido el

cambio de un codón a otro que especifique un aminoácido diferente son sólo una pequeña fracción.

La comparación de las frecuencias de uso de codones (Tabla 6) para el gen Antp de las dos especies analizadas denota una sutil (aunque distinguible) tendencia en *Drosophila subobscura* al uso de codones sinónimos que presenten un menor contenido en Guanina y Citosina (en comparación con *D.melanogaster*). Esta tendencia la observamos para los aminoácidos: Ala, Asn, Asp, Cys, Gln, Glu, His, Leu, Pro, Ser, y Thr; el fenómeno inverso (preferencia de codones en *D.subobscura* que impliquen un mayor contenido en Guanina y Citosina) sólo se observa para los aminoácidos : Gly, Phe, y Val. En total, hay un cambio neto de 12 G ó C por A ó T, lo que da lugar a una ligera diferencia en el porcentaje de A+T (39.7%), en comparación con *D.melanogaster*, donde es del 38.3%

No es suficiente con la comparación de estas frecuencias de uso de codones en un sólo gen para concluir que hay realmente una tendencia en *D.subobscura* a utilizar codones que impliquen un menor contenido en G y C. Sin embargo, hemos comparado las frecuencias de uso de codones en otro gen de esta especie: el gen rp49 (AGUADE, 1988) y hemos encontrado la misma tendencia. En este caso, sólo los aminoácidos Gly y Ala se "resisten" a la tendencia general, que implica, en este gen, una sustitución neta de 14 G ó C por A ó T.

5.4. Segmentos conservados en el líder y en el trailer del

transcrito

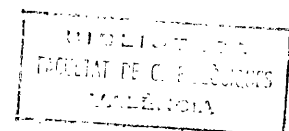
El alineamiento de las secuencias denota, además de la conservación de los exones traducidos, un nivel de conservación muy elevado en el exón 4, que forma parte del líder no traducido. Al observar las secuencias conservadas en esta región, se encuentra que contienen repeticiones que se pueden agrupar en tres secuencias consenso: TAGAGGAA, TACTTAAG y TCAA. Estas repeticiones están marcadas con asteriscos en la secuencia de la figura 10; en la Tabla 8 presentamos estas repeticiones, agrupadas por tipos para cada especie. En esta tabla, cada repetición de *D.melanogaster* tiene su homólogo en *D.subobscura* a la derecha. Los nucleótidos que difieren de la secuencia consenso están representados en letra minúscula.

Además de las repeticiones, las secuencias conservadas del exón 4 contienen el palíndromo imperfecto TTTGGCCAAGA. El significado de estas secuencias conservadas es oscuro. A juzgar por las predicciones de la estructura secundaria del transcrito Antp de *D.melanogaster* (datos no presentados), estas secuencias no parecen participar en la formación de segmentos de estructura secundaria en el mRNA.

A pesar de no poder determinar cuál es la función de estas secuencias conservadas, es evidente que no se han conservado por puro azar. Las secuencias afuncionales (intrones) presentan una conservación mucho menor que este exón 4 (ver Tabla 7). Este exón presenta un nivel de conservación casi tan elevado como el de los exones traducidos.

Además de la conservación del exón 4, el líder de *D.subobscura* presenta una característica muy peculiar, aunque esta vez no está conservada entre las dos especies: al comparar las secuencias nucleotídicas de las dos especies mediante una matriz de puntos (fig. 9) se demuestra que hay una región del líder de *D.subobscura* que comparte una elevada homología con una parte del producto protéico. Esta matriz de puntos muestra una gran nube de puntos en el centro de la diagonal de homología (posiciones 530 a 700 del eje de *D.subobscura* frente a las posiciones 900-1050 de *D.melanogaster*). Dicha nube se produce por la acumulación de una gran cantidad de codones CAG y CAA en esos segmentos de las dos secuencias. La otra gran mancha que se observa en la matriz de puntos, fuera de la diagonal, corresponde a las posiciones 110-154 de nuestra secuencia, que presenta una gran similitud con la misma secuencia de *D.melanogaster* (posiciones 900-1050). El examen de esta región no traducida del exón 5 evidencia que hay 10 repeticiones del codón CAG, todas ellas en la misma fase, en una región del líder (45 pb) muy cercana al sitio esperado de inicio de síntesis de la proteína (fig. 10, exón 5).

En principio, descartamos la posibilidad de que el inicio de la síntesis protéica se produzca, en *D.subobscura* en un lugar desplazado en dirección 5'. Para afirmar esto, nos basamos en dos hechos: (1) que la pauta de lectura de estas repeticiones CAG no está en fase con la común con *melanogaster*; y (2) que hay evidencias muy fuertes de que la



ORF que presentamos en la sección 5.2 es la correcta.

Por otra parte, un análisis preliminar de la región no traducida del exón 8 (trailer) demuestra la presencia de segmentos extraordinariamente conservados entre la posición de las dos señales de poliadenilación. Estos segmentos conservados están embebidos en secuencias que han divergido totalmente y que no tienen ninguna homología aparente entre las dos especies. En la figura 14 presentamos la secuencia del extremo derecho del fragmento EcoRI que contiene el exón 8 (ver fig. 7). Esta secuencia, alineada frente a *D.melanogaster* muestra una homología muy elevada en varios segmentos situados entre los dos sitios de poliadenilación. El segmento de mayor similitud presenta 46 nucleótidos idénticos, en una secuencia de 48 pb.

En la bibliografía, sólo encontramos un precedente de conservación en la región 3' no traducida: se trata del trailer del gen *hunchback* (*hb*), que presenta regiones de conservación entre dos especies de *Drosophila*, inmersas en un fondo no conservado (TREIER y otros, 1989). El caso normal es la ausencia de conservación: en las comparaciones interespecíficas de los genes *engrailed*, *Adh* y *rp49* no se describe ningún tipo de conservación en la región 3' del gen. Estos tres genes tienen en común el hecho de que sus secuencias trailer son cortas (hasta unos 150 pb). El gen *Antp*, sin embargo, comparte con *hunchback* la característica de poseer un trailer anormalmente largo. El de *hb* tiene una longitud de 660 nucleótidos y el de *Antp* es de 600 o 2000

nucleótidos (dependiendo de cual de los dos sitios de poliadenilación alternativos se utilice; ver figura 5). Esto sugiere que estas secuencias trailer largas contienen una serie de señales que afectan a la función del gen, ya sea a través de la estabilidad del transcrito, o bien por su localización intracelular.

Cualquier conservación observable dentro de las regiones no traducidas (bien sea en 5' o en 3') puede considerarse, en principio, significativa, si se tiene en cuenta el bajo nivel general de conservación que se observa en dichas regiones.

5.5. ¿Conservación en los intrones?

En los intrones que hemos secuenciado no aparece ninguna secuencia conservada. El único segmento que tiene una ligera similitud es el final del intrón 5 (posiciones 960 a 985 de *D.subobscura*). No atribuimos, sin embargo, ninguna función a esta corta región conservada: estimamos que su posible conservación se debe únicamente al azar.

El gen *Antp* de *D.melanogaster* comprende una región de 100 kb de DNA, donde sólo unas 2 kb corresponden a exones. Las 98 kb restantes son secuencias intrónicas, donde destacan por su tamaño los intrones 1, 2, 3 y 7 (SCHNEWULY y otros, 1986; LAUGHON y otros, 1986). El intrón 7 es el que precede al exón del extremo 3'. Este intrón 7 tiene un tamaño de 10-11 kb en *D.melanogaster*. Nuestros datos muestran que su tamaño en

D. subobscura es el mismo, por lo que se puede interpretar que al menos el tamaño del intrón es importante para la función del gen.

Queda por determinar si los otros intrones del gen, de longitud todavía mayor en *D. melanogaster* han conservado su tamaño en *D. subobscura*. Analizar este aspecto sobre la estructura de la unidad de transcripción en las dos especies es uno de los objetivos de nuestro trabajo en curso. Con los datos que ya tenemos, podemos adelantar que el intrón 3 tiene una longitud superior a 15-16 kb, por lo que su tamaño podría ser similar al de *D. melanogaster* (apartado 4 de este capítulo).

También especulamos sobre la posibilidad de que alguno de los intrones gigantes contenga alguna información. Nos referimos a la posibilidad de que haya alguna unidad de transcripción totalmente embebida en alguno de los intrones. Esta posibilidad no sería un caso excepcional ya que hay, al menos, dos casos conocidos de este fenómeno (HENIKOFF y otros, 1986; CHEN y otros, 1987). Por otra parte esto no sería de extrañar si consideramos que los intrones del gen *Antennapedia* de *D. melanogaster* representan una fracción nada despreciable del total del genoma no repetitivo: el 0.1%.

La organización general de la región que hemos clonado es, a grandes rasgos, similar a la de la misma región en *Drosophila melanogaster*. El mayor de los intrones que hemos clonado totalmente es el intrón 7, que separa el exón 3' terminal del resto de los exones que codifican el producto

protéico, tiene el mismo tamaño que en *D.melanogaster*, 10-11 kb. El tamaño del intrón 6 es también comparable a la especie mencionado, mientras que el intrón 5 es considerablemente menor, 145 pb *D.subobscura* frente a los 378 pb en *D.melanogaster* (fig. 8, parte superior). A pesar de no haber secuenciado el intrón 4, sabemos que su tamaño es igual al de *D.melanogaster*.

6. ASPECTOS CONSERVADOS Y CARACTERISTICAS DIFERENCIALES EN LA ESTRUCTURA DEL GEN

En la figura 8 representamos un esquema de la organización del gen en el principal segmento que hemos secuenciado, en comparación con el segmento homólogo de *D.melanogaster*. Con una línea representamos el DNA y con recuadros sobre la línea la localización de los exones. Las secuencias que corresponden a exones (o parte de un exón) no traducidos son de color blanco; en negro representamos las secuencias incluidas en la ORF. Sobre la secuencia de *D.subobscura* hemos representado los sitios de restricción utilizados para la subclonación de fragmentos durante la estrategia de secuenciación. El codón de iniciación lo hemos marcado con AUG y la posición de las regiones conservadas del exón 4 están marcadas con puntas de flecha. Los dos sitios dadores alternativos para el splicing del exón 7 están marcados con dos flechas, al final de dicho

exón.

Las diferencias nucleotídicas entre el segmento de *D.subobscura* que hemos secuenciado y su homólogo en *D.melanogaster* se reflejan en un 64% de conservación general. Esta cifra, sin embargo, enmascara las grandes diferencias de conservación entre distintas regiones de la unidad de transcripción. En la Tabla 7 presentamos los porcentajes de conservación de cada región, donde puede verse la gran diferencia de conservación entre intrones y exones. Dentro de los exones, los que codifican para el producto protéico presentan el mayor nivel de conservación. El nivel de conservación de la ORF es del 87% a nivel de DNA; a nivel de la proteína la conservación es del 95%, superior a la conservación de DNA debido a que en esta región muchas de las sustituciones nucleotídicas se han producido en posiciones silenciosas de los codones.

Solamente una fracción de las posiciones nucleotídicas de los genes estructurales pueden cambiar sin alterar la secuencia de aminoácidos del producto protéico: son las posiciones silenciosas. Al comparar la conservación de nuestra secuencia con *D.melanogaster* a nivel de DNA (en la ORF) y a nivel de proteína, puede verse que una gran mayoría de las sustituciones nucleotídicas se han producido en lugares silenciosos, lo que indica que existe una fuerte presión de selección que elimina las sustituciones deletéreas (que producen cambios no permisibles en la secuencia de aminoácidos).

Hay muy pocos datos sobre genes secuenciados en especies de *Drosophila* del grupo obscura (al que pertenecen *D.subobscura* y *D.pseudoobscura*). Se ha estimado que el grupo obscura y el grupo melanogaster han divergido durante el Eoceno Medio (hace unos 45 millones de años). Es esperable, en principio, que los niveles de conservación entre distintos genes de los dos grupos sean comparables (asumiendo una misma tasa de variación entre distintas regiones del genoma). En un estudio en el que se determinó la secuencia de la región aminoterminal de la proteína Ultrabithorax en varias especies (WILDE y AKAM, 1987), la secuencia proteica de *D.pseudoobscura*, muestra un 90% de conservación (calculado de la misma forma que en la Tabla 5).

Un porcentaje de conservación similar al que observamos ha sido también demostrado por AGUADE (1988) para el producto protéico del gen *rp49* de *D.subobscura*, que codifica para una proteína ribosómica.

Estos números, sin embargo, no reflejan las variaciones de los niveles de conservación en las distintas partes de la proteína Antp. Como hemos puesto de manifiesto en los apartados anteriores, hay regiones que son completamente idénticas en las dos especies, frente a regiones que han divergido en gran medida. Dicho de otra manera, las sustituciones de aminoácidos se distribuyen de forma muy diferente a lo largo de la proteína Antp, lo que parece indicar que ciertas partes de la proteína Antp sufren una presión de selección superior a otros dominios. Así, en las figuras 12 y

13 se aprecia que los primeros 104 aminoácidos del producto Antp de *D.melanogaster* se han conservado casi totalmente en *D.subobscura*, con sólo tres diferencias de aminoácidos (dos sustituciones y una inserción). En estas figuras también puede encontrarse una región extremadamente variable (aminoácidos 105 a 155 de *D.melanogaster*), que coincide con la región con gran contenido en glutamina. Los siguientes 77 aminoácidos (posiciones 156 a 232), que incluyen un exón ó idéntico en las dos especies sólo han variado en una posición (168), siendo sinónimas las dos sustituciones producidas (Ser-Thr). Por último, una región (32 aminoácidos) que ha acumulado tres cambios significativos, viene seguida de los 35 aminoácidos del extremo carboxiterminal, previo al dominio homeótico. Esta región ha permanecido invariable en la evolución.

Encontramos que este patrón de variación a lo largo del producto protéico del gen *Antennapedia* tiene una similaridad muy llamativa con el esquema de conservación del producto protéico del gen *Ultrabithorax* (WILDE y AKAM, 1987, fig.5 de esta referencia). Esto sugiere, a nuestro juicio, que los dominios funcionales de ambos genes homeóticos se organizan de forma muy similar.

La diferente conservación entre distintos dominios protéicos coincide con los resultados de la comparación del gen *Adh* entre distintas especies (SCHAEFFER y AQUADRO, 1987). No parece, sin embargo, que pueda generalizarse que hay distintas presiones de selección para cada uno de los dominios de cualquier proteína. Los resultados de AGUADE (1988) al

comparar la secuencia del gen *rp49* de *D.subobscura* con la de *D.melanogaster* muestran que las sustituciones de aminoácidos se distribuyen al azar entre los dos exones de este gen. Pensamos, no obstante, que esta ausencia de restricciones en distintas regiones de las proteínas no se da normalmente en los genes que controlan el desarrollo, muchos de los cuales poseen dominios de unión al DNA extremadamente conservados.

En la comparación del gen *engrailed* entre *D.virilis* y *D.melanogaster* (KASSIS y otros, 1986) se ha observado una identidad absoluta a nivel de los aminoácidos del exón 3', a pesar de la notable variación en otros dominios de la proteína. Este exón 3' contiene el dominio homeótico (implicado en la unión a secuencias específicas de DNA). De forma similar, la comparación del gen *hunchback* en estas dos mismas especies (TREIER y otros, 1989) demuestra también un 100% de conservación en los dominios de unión al DNA (fingers) de la proteína. *D.virilis* es una especie más distante de *D.melanogaster* que nuestra especie de estudio. Pensamos que si los dominios de unión al DNA se han conservado totalmente en las dos especies (para los dos genes antes citados), es razonable pensar que la situación será similar para los dominios de unión al DNA en el gen *Antp* entre *D.subobscura* y *D.melanogaster*.

Dada la extrema conservación esperable a nivel de los dominios de unión al DNA, en el trabajo que aquí presentamos no hemos centrado nuestra atención en la secuenciación del exón 8 del gen *Antp* de *D.subobscura*, que contiene la secuencia

homeobox.

Vemos la necesidad de remarcar que, en nuestra comparación de Antp, todas las variaciones de la región traducida que afectan al número de aminoácidos han sido en el sentido de incrementar el número de aminoácidos en *D.subobscura*; sólo 1 codón de la ORF de *D.melanogaster* está ausente de la ORF de nuestra especie. Esto da lugar a un producto protéico que, en *D.subobscura*, contiene 16 aminoácidos más que en *D.melanogaster*. Será necesario conocer la secuencia de otros genes de *D.subobscura* para determinar si esto es una tendencia general en todo el genoma de esta especie. Lo que sí parece una tendencia es que estas "inserciones" y "delecciones" tienden a producirse precisamente en las regiones con repeticiones de un aminoácido, no sólo al comparar un gen entre varias especies (nuestros datos; WILDE y AKAM, 1987; KASSIS y otros, 1986; TREIER y otros, 1989), sino que también se han encontrado variaciones entre distintas cepas de *D.melanogaster*, al menos para el gen Deformed (REGULSKI y otros, 1987: figura 6 de esta referencia). Más adelante nos referiremos al posible significado de estas variaciones.

Es de señalar la extremada conservación nucleotídica entre las dos especies que se ha producido a nivel de los dos sitios dadores alternativos para el procesado del intrón 7 (fig. 10 y Tabla 4), a pesar de que el sitio dador D2 difiere del consenso en que el intrón comienza en GA, en vez de GT, que es la situación normal en la casi totalidad de los intrones eucarióticos. Debido a la conservación de los dos sitios

dadores alternativos, deducimos que *D.subobscura* es también susceptible a un procesado diferencial que, al igual que en *D.melanogaster*, de lugar a pequeñas variaciones en el producto protéico de Antp en las inmediaciones del siguiente dominio protéico: el dominio homeótico (BERMINGHAM y SCOTT, 1988).

Esto sugiere que, en *D.subobscura* también se produce un procesado diferencial nivel de los dos sitios alternativos dadores para el splicing, situados en el exón 7 (ver figs. 10 y 11). Estos dos sitios se han mantenido totalmente idénticos en las dos especies. Deducimos, pues, que las microvariaciones resultantes en el producto protéico son del mismo tipo. En otro gen homeótico, *Ultrabithorax*, ocurre una situación similar: el exón que codifica la mayor parte del producto protéico posee también dos sitios dadores para el splicing, que se utilizan de forma diferencial a lo largo del desarrollo.

En una comparación del gen *Ubx* de *D.melanogaster* con otras dos especies del mismo género y con *Musca domestica* (WILDE y AKAM, 1987) se ha demostrado que estos dos sitios se han conservado en *D.funnebris* y en *D.pseudoobscura*, aunque no en *Musca*, donde sólo se encuentra el primer sitio dador. En cuanto al péptido codificado (9 aminoácidos) entre los dos sitios dadores, en las tres especies de *Drosophila*, éste se ha conservado entre *D.melanogaster* y *D.pseudoobscura*; en *D.funnebris* tres aminoácidos han divergido.

La función de estas variaciones en el producto protéico no se conoce, si bien su cercanía al dominio de unión al DNA

(dominio homeótico) sugiere que, de alguna forma, podría modular la especificidad o afinidad de la proteína por el DNA. Si esto es así, las variaciones que se han descrito en Ubx entre distintas especies de *Drosophila* podrían conducir a variaciones en la actividad de este gen en estas especies. La conservación que nosotros observamos entre *D.subobscura* y *D.melanogaster* a nivel del péptido alternativo de Antp indica que no hay tal posibilidad de variación entre las dos especies.

También se ha propuesto que estas variaciones podrían afectar a las posibilidades de interacción con otras proteínas reguladoras (O'CONNOR y otros, 1988; BERMINGHAM y SCOTT, 1988). Lo que sí parece evidente es que estas variaciones se producen entre regiones muy conservadas: el dominio homeótico y el péptido Tyr-Pro-Trp-Met (ver fig. 12).

Este fenómeno podría ser común a casi todos los genes con homeobox. Así, en el gen homeótico labial y en el gen de efecto materno bicoid también se han descrito variaciones del producto protéico, producidas por la presencia de dos sitios aceptores alternativos, también en las cercanías del dominio homeótico (MLODZIK y otros, 1988; BERLETH y otros, 1988).

La parte no traducida del exón 5 del gen *Antennapedia* de *D.subobscura* muestra un porcentaje de conservación similar al de los intrones; mientras que el exón 4 (no traducido) presenta una conservación extraordinaria, comparable a la de los exones traducidos. Un examen de las regiones conservadas en el exón 4 ha demostrado la presencia de varias secuencias

consevadas, así como un palíndrome casi perfecto.

El alineamiento de los intrones es prácticamente imposible, dada la gran divergencia que se produce; el que presentamos en la figura 10 es el que proporciona el ordenador, intentando maximizar los aciertos. Sin embargo, cualquier variación de los parámetros en la aplicación del programa da lugar a cambios en el alineamiento, lo que significa a nuestro juicio que los alineamientos de los intrones son datos poco consistentes. Esto difiere del alineamiento de los exones traducidos, donde el resultado es siempre el mismo, independientemente de las opciones que utilicemos en la aplicación de los programas de alineamiento. A pesar de los cambios de longitud, la posición de los intrones dentro de la ORF no ha variado entre las dos especies.

7. OTRAS CONSIDERACIONES

A lo largo de estas secciones, hemos descrito una serie de clones de lambda (cuatro en total) correspondientes al gen Antp de *D.subobscura*, así como la determinación de la secuencia nucleotídica de los fragmentos de interés. Sin embargo, el número de clones que hemos analizado es mayor, aunque nos hemos limitado a describir aquellos cuya pertenencia a este gen hemos podido demostrar.

De todos los clones originales del screening de la genoteca, sólo hemos llegado a determinar la identidad de tres. Hemos discutido ampliamente la inevitabilidad de obtener, junto con los clones de Antp, otros clones que compartan con este gen una cierta similitud, bien sea a través de la homeobox o bien por medio de repeticiones del codón CAG. En efecto, otro de los clones obtenidos, lambda DsJ3, que no tenía ninguna región de solapamiento con el resto de los clones que hemos descrito, resultó contener un fragmento EcoRI que proporcionaba una buena señal de hibridación frente a la sonda de cDNA. Al determinar la secuencia de parte de este fragmento pudimos comprobar que no contenía, como esperábamos, uno de los exones todavía no aislados, sino una ORF de, al menos, 309 nucleótidos. La secuencia protéica correspondiente contiene, de los 103 aminoácidos codificados 30 glutaminas, codificadas mayoritariamente por el codón CAG. Una búsqueda en la base de datos EMBL Data Library con esta secuencia encuentra homología con genes que contienen segmentos de poliglutamina (repeticiones M), tales como engrailed, hairy, Deformed, fushi-tarazu y Notch. Sin embargo, esta homología es únicamente a nivel de las repeticiones CAG, por lo que no es significativa.

Consideramos la posibilidad de que el clon lambda DsJ3 pueda contener un gen todavía no descrito en *D.melanogaster*, ya que ninguna de las homologías que hemos detectado llegan más allá de las repeticiones M (CAG).

La mayor parte de las sustituciones en la secuencia

protéica de Antp se han producido, como hemos demostrado, a nivel de las regiones con repeticiones de glutamina. También ocurre lo mismo con los sucesos de inserción/delección, lo que apoya la idea de que las regiones con codones repetidos son propensas a un "desplazamiento de codones", o codon slippage (DOVER, 1989). El mecanismo molecular de este proceso se desconoce, aunque se sospecha que se produce durante la replicación del DNA. Lo que es un hecho es que cualquier gen, de los hasta ahora estudiados en varias especies, que contenga segmentos de poliaminoácidos presenta una acumulación de la variación protéica precisamente en estos segmentos. Es el caso de la comparación de los genes Ultrabithorax, engrailed, hunchback (WILDE y AKAM, 1987; KASSIS y otros, 1986; TREIER y otros, 1989). En el gen Deformed se ha descrito una variación entre las cepas Canton-S y Oregon-R. Esta variación se sitúa exactamente en las regiones con repeticiones de los codones CAG (Gln) y AAC (Asn) (REGULSKI y otros, 1987).

En el estudio comparativo del gen period se ha llegado más lejos: se ha demostrado que las variaciones producidas por desplazamiento de codones en *D.pseudoobscura* son responsables de las diferencias de comportamiento (ritmos circadianos y comportamiento durante el cortejo) entre esta especie y *D.melanogaster* (COLOT y otros, 1988). Las variaciones en el gen huchback también parecen tener relación con las diferencias en el patrón de expresión en *D.melanogaster* y en *D.virilis*, aunque no se ha demostrado todavía una relación causal.

DOVER (1989) propone que estas variaciones entre especies en el número y posición de las repeticiones podrían ser un mecanismo de especiación tan eficaz como lo puedan ser las sustituciones de aminoácidos. El hecho de que se produzca una variación permisiva a este nivel incluso dentro de una misma especie (como en el gen Deformed) nos parece significativo. Probablemente, el fenómeno del desplazamiento de codones sea una característica inherente a todas las regiones repetidas, de tal forma que todos los genes que las posean estén sujetos a una tasa de variación elevada. Esto contrasta con lo que cabría esperar en los genes implicados en el control de desarrollo donde, precisamente se ha descrito la casi universal presencia de repeticiones (ver el capítulo INTRODUCCION). Pensamos que este fenómeno podría constituir una base importante de la evolución.

Este tipo de razonamientos podría invalidar la frecuente interpretación de que las regiones de la proteína que han divergido entre dos especies corresponden a regiones no críticas para su actividad (esto ha sido también nuestra hipótesis de partida). Hay que tener presente, sin embargo, que las regiones con divergencias podrían contener funciones características de cada especie (DOVER, 1989). No es posible concluir que las regiones con repeticiones de aminoácidos sean afuncionales, a pesar de la extrema variación que se ha descrito en todas las comparaciones inespecíficas. Si fueran afuncionales, se habría descrito algún caso de delección completa a nivel de estas regiones, cosa que no ha ocurrido.

Nuestra interpretación es que las regiones que contienen repeticiones, aunque pudieran contener instrucciones distintas para cada especie, no son las regiones esenciales para la función. Pensamos que las regiones invariables son las comprometidas en "qué hacer" (por ejemplo, en los genes homeóticos: unirse a ciertas secuencias de DNA; interaccionar con otras proteínas reguladoras; etc) y que las regiones variables especificarían el "cómo hacerlo" (por ejemplo, variaciones más o menos sutiles en los dominios de expresión espaciales o temporales).

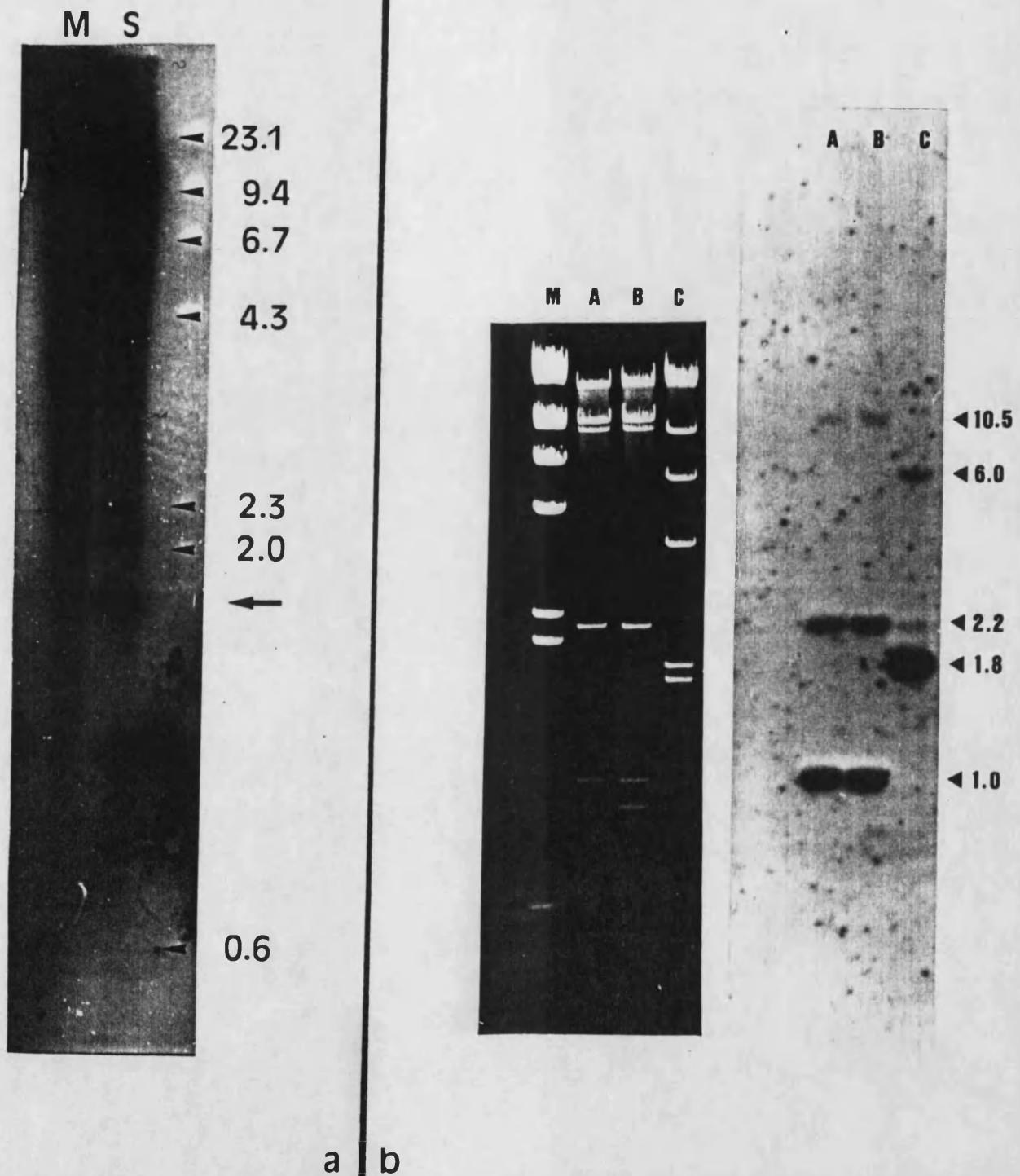


Figura 6. (A) Detección de fragmentos homólogos al gen *Antennapedia* de *D. melanogaster* en el DNA genómico de *Drosophila subobscura*. La flecha indica el fragmento que produce la señal de hibridación más fuerte (≈ 1.8 kb), M es la muestra de *D. melanogaster* y S es la muestra de *D. subobscura*, ambos DNAs están digeridos con EcoRI. Las puntas de flecha señalan las posiciones del marcador de pesos moleculares (λ HindIII). (B) Análisis de los clones recombinantes de λ aislados de la genoteca. En la izquierda se muestran los DNAs digeridos con EcoRI y en la derecha una hibridación, de ese mismo gel, con la sonda de cDNA de *D. melanogaster* (la misma que en la figura 4,A). Hemos señalado con puntas de flecha los fragmentos a los que nos referimos en el texto. A, B y C son los clones λ DsA1, λ DsB1 y λ DsC3.

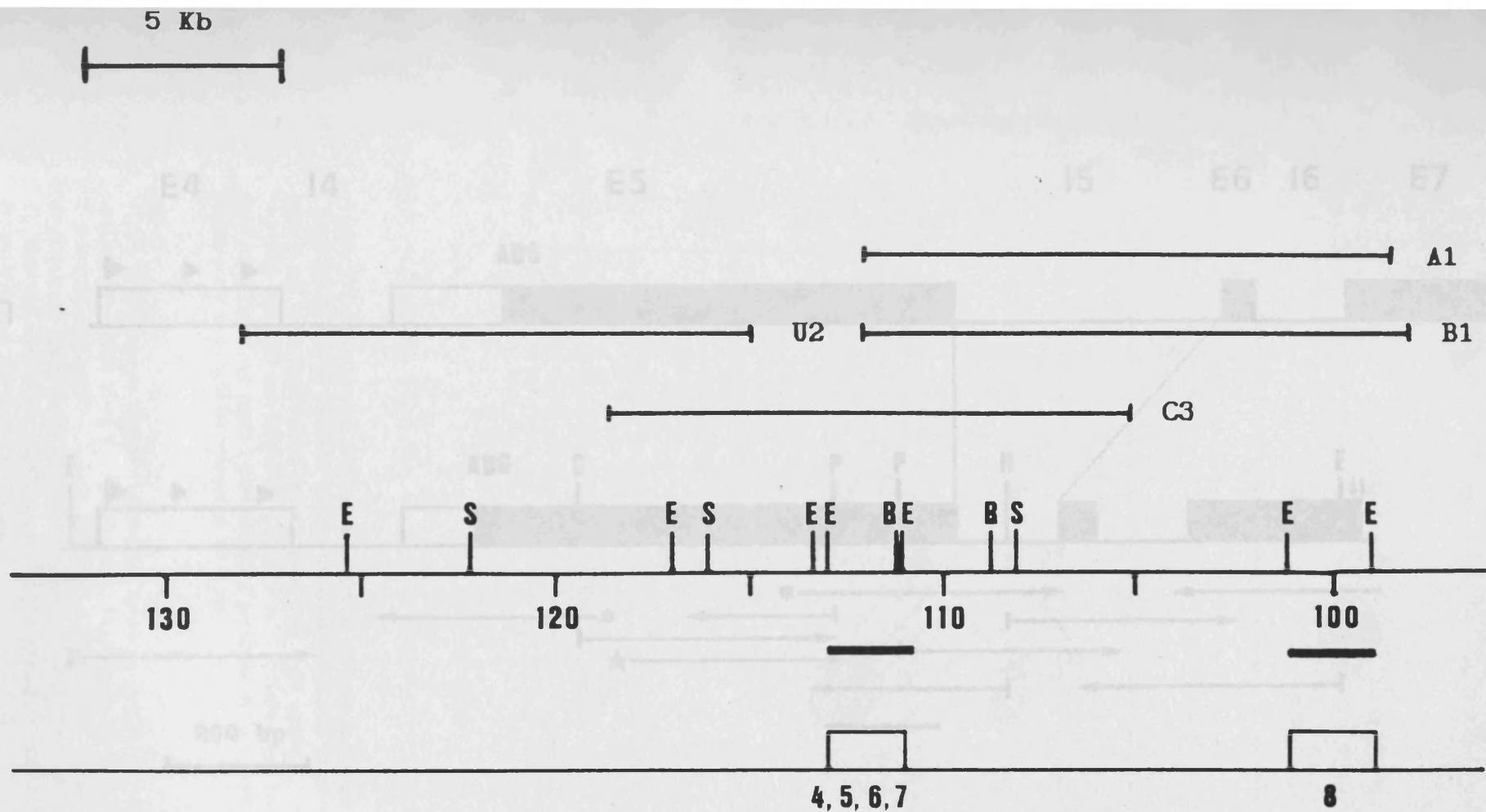


Figura 7. Mapa de restricción de la región 3' del gen *Antp* de *D. subobscura*. Esta región incluye 31 kb de DNA genómico que incluye los exones 4-5-6-7 y 8. La extensión de los clones del fago λ que hemos obtenido la hemos indicado sobre el mapa. Las unidades del mapa están medidas en kb; hemos considerado como punto 100 el centro del fragmento que contiene la secuencia homeobox (exón 8). La polaridad que hemos seguido para las referencias es la utilizada por SCOTT y otros (1983), y opuesta a la utilizada por SCHNEUWLY y otros (1986). Las regiones que hibridan con la sonda de cDNA de *D. melanogaster* se señalan con una línea gruesa, debajo de la línea que representa el genoma. Los sitios de restricción analizados son: E, EcoRI; S, Sall; y B, BamHI.

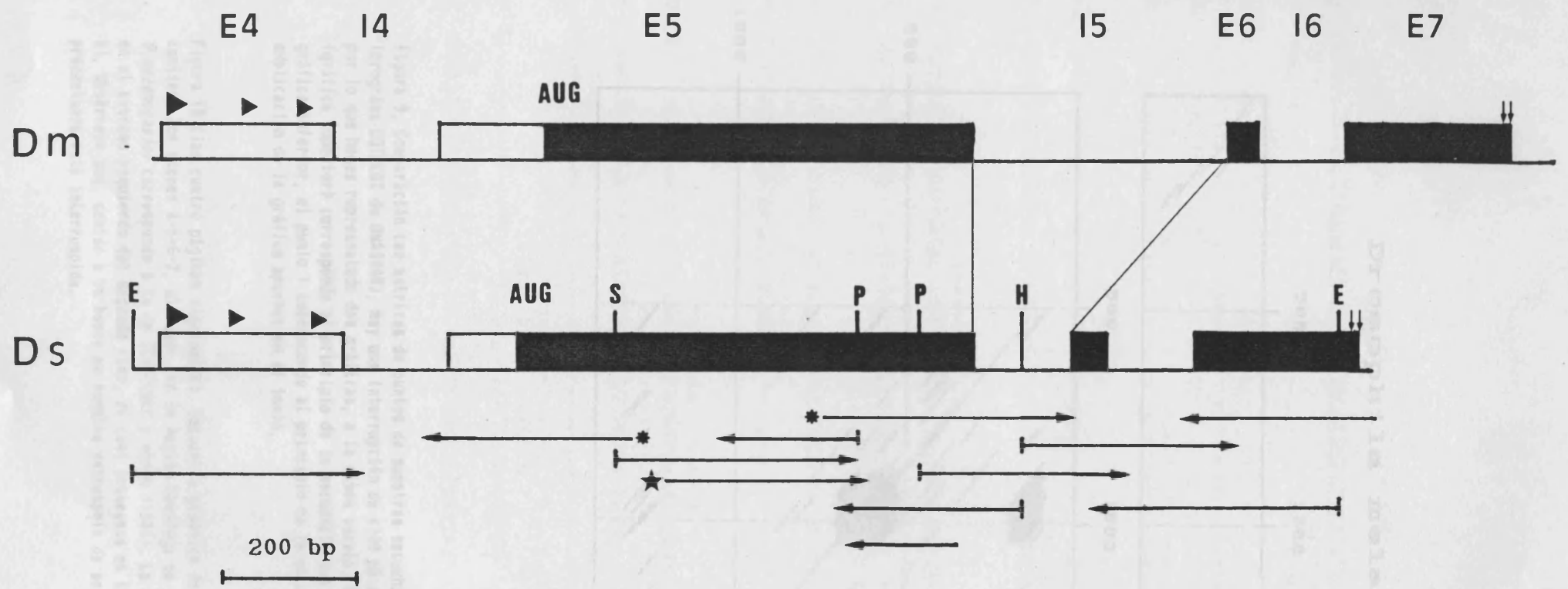


Figura 8. Organización de la región del gen *Antp* que incluye los exones 4-5-6-7. Los recuadros blancos son los exones (o parte de exones) no traducidos; los recuadros negros son los exones traducidos. La posición de las secuencias conservadas en el exón 4 se muestra con puntas de flecha. Con dos flechas pequeñas hemos señalado los dos sitios dadores alternativos para el procesado, los cuales se han conservado perfectamente en las dos especies. La estrategia de secuenciación utilizada se indica con las flechas horizontales. Las flechas con una línea vertical en su origen son las obtenidas a partir de un sitio de restricción; las flechas con una estrella en su origen son las obtenidas con oligonucleótidos sintéticos; las flechas sin ningún símbolo en su origen son las obtenidas a partir de clones en el vector M13 (el resto de las secuencias se han obtenido a partir de clones plásmidicos). La secuencia obtenida a partir del extremo del clon λ DsA1 es la señalada con una estrella de cinco puntas (explicación en el texto). Los sitios de restricción son: S, SmaI; E, EcoRI; H, HindIII; y P, PstI.

Drosophila melanogaster

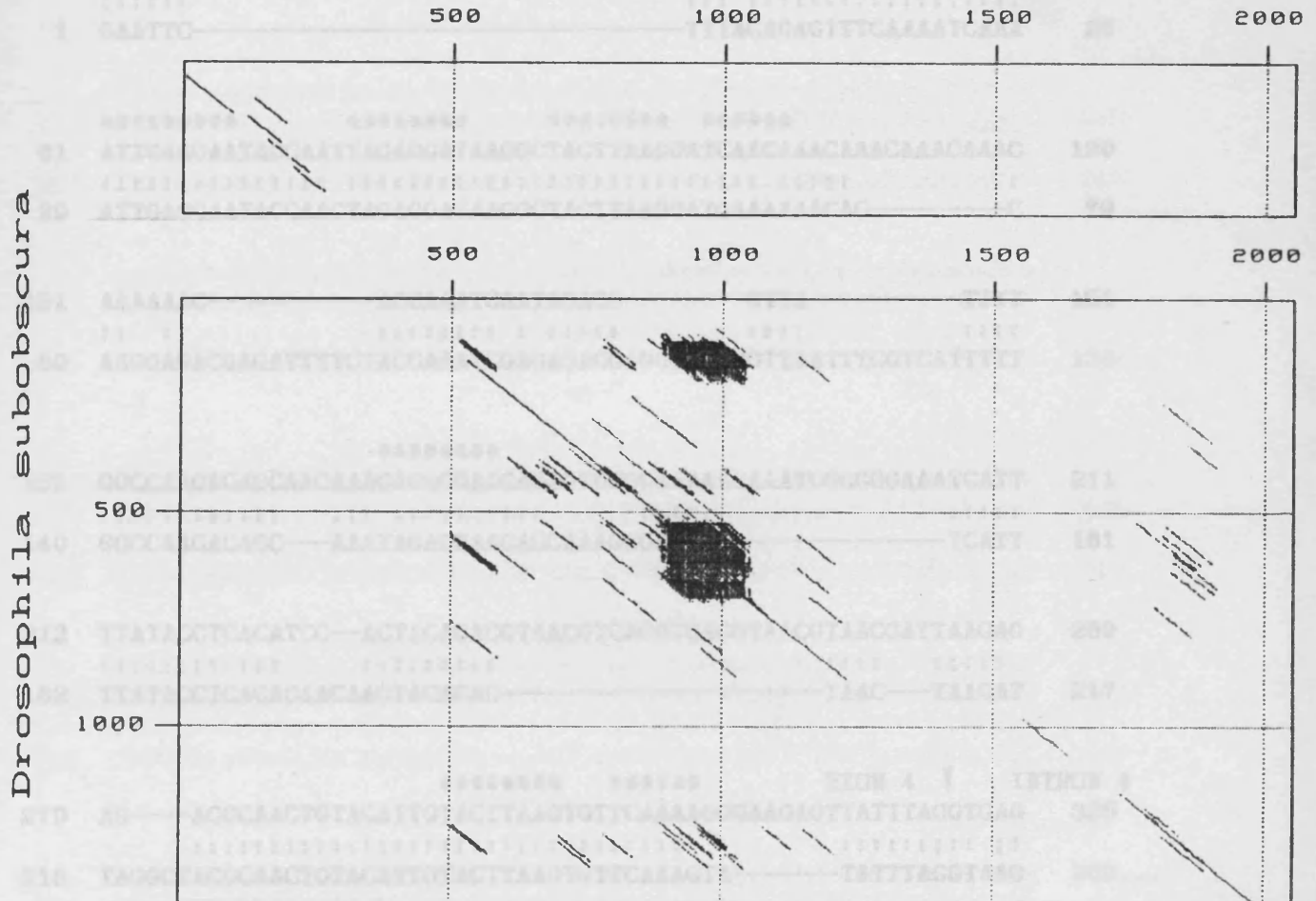
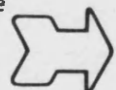


Figura 9. Comparación con matrices de puntos de nuestras secuencias con la de *D. melanogaster* (programa DOTPLOT de DNASTAR). Hay una interrupción de ≈ 100 pb en nuestra secuencia (ver fig. 10), por lo que hemos representado dos gráficas, a la misma escala. El punto 1 del eje de *D. subobscura* (gráfica superior) corresponde al principio de la secuencia que presentamos en la figura 10. En la gráfica inferior, el punto 1 corresponde al principio de la secuencia después de la interrupción. La explicación de la gráfica aparece en el texto.

Figura 10 (las cuatro páginas siguientes). Secuencia genómica de la región de *D. subobscura* que contiene los exones 4-5-6-7, alineada con la región homóloga de *D. melanogaster*. La numeración de *D. melanogaster* corresponde a la de SCHNEUWLY y otros (1986). La numeración de *D. subobscura* empieza en el extremo izquierdo del segundo clon, el cual mostramos en la estrategia de secuenciación (fig. 8). Obsérvese que, debido a un hueco en nuestra estrategia de secuenciación, la secuencia que presentamos está interrumpida.



INTRON 3 EXON 4

EcoRI ▼ *****

Ds	1	GAATTCGAACTTTAACGAAAACTCTACTTCATTCTCTTTTTTCAGAGTTTCAAAATCAAA	60
		:::::	::: ::::::::::::::::::::
Dm	1	GAATTC-----TTTACAGAGTTTCAAAATCAAA	28

Ds	61	ATTGAGGAATACCAATTAGAGGATAAGGCTACTTAAGGATCAACAAACAAACAAACAAAC	120
		::::::::::: :	
Dm	29	ATTGAGGAATACCAACTAGAGGATAAGGCTACTTAAGGATCAAAAAACAC-----C	79

Ds	121	AAAAAAC-----ACCAAATCAATAGACG-----GTTA-----TTTT	151
		:: :	
Dm	80	AAGGAGACGAGATTTTCTACCAAATCGAGAGACGAGGGGCAGGTTAATTCGTCATTTTT	139

Ds	152	GGCCAAGACAGCAACAAACAGAGGAACAGAGTGCGGAAAAACAAATCGCGGAAATCATT	211
		::::::::::: :	
Dm	140	GGCCAAGACAGC-----AAATAGAGGAACAGCAAAGCGAAAA-----TCATT	181

Ds	212	TTATACCTCACATCC--ACTACACACGTAACGTCACGTCACGTAACGTAACGATTAAGAG	269
		::::::::::: :	
Dm	182	TTATACCTCACACAACAACACTACACAC-----TAAC--TAAGAT	217

Ds	270	AG----ACGCAACTGTACATTGTAAGTGTTCAAAAGGGAAGAGTTATTTAGGTGAG	325
		::::::::::: :	
Dm	218	TAGGCTACGCAACTGTACATTGTAAGTGTTCAAAAGTA-----TATTTAGGTAAG	269...

INTRON 4 ▼ EXON 5

Ds	84	-----TTCGCAGTCACGTTTCACCACGG--AAGCAGCAGAAATA	120
		::::: :	
Dm	... 392	CTAATTTGTGTTTTTCTTATCATCCAGTTTACTTTGTATATAAGAAAAGTAGCTAAAAGC	451

Ds	121	G-----CAGCA--GCAGCAGCAG-----AAGCAGA-----	143
		::: :	
Dm	452	ACGCGGACAGGGAGGCAGGAGCACACCACAGTCACTAGCCACTAAGCAGAGTCACAGTCAGC	511

Ds	144	-----AGCAGCCGCAGTTGCCGCCGCC	165
		::::: :	
Dm	512	ATCACGTTCACTCCAGGATCAGGACTCGGGGCGGGATCAGCAGACGCTGAGGAAGCTGCC	571

	INTRON 5	▼	EXON 6	EXON 6	▼	INTRON 6	
			LeuGlyTyrThrAspValGlyValProAspValThrGlu				
Ds	992		ACCAAAACAGTTGGGTTACACGGACGTTGGAGTTC	CCCGATGTTACAGAGGTAAGA	-----		1046
Dm	1567		ACCTG—CAGTTGGGCTATACGGACGTTGGAGTTC	CCCGACGTGACAGAGGTAAGATCGAG			1624
Ds	1047		-----CGCCCGTCATCCCGCCCTCTCTCAAT	-----ACC	---	CATCCCTC	1084
Dm	1625		TCCACATTGCGCTTGTC	-----CCTACTTCAAATTCTGCTAAAACCTAACATCTTTTA			1678
Ds	1085		TCCCAATATGGTGGCTGGTTAATTGCACCCTCTAACTGTAGACAATAA	---	AATATTTACAT		1143
Dm	1679		CGGTTGATGGTTCAAGATAT	----ACCC-CTAA	-----	AATTACAATATCACTTA	1725
			INTRON 6	▼	EXON 7		
					ValHisGlnAsnHisHisAsnMetGlyMetTyrGly		
Ds	1144		TCTCTTTTTGAATCGTTCATTTAGGTCCATCAA	AACCATCACAACATGGGAATGTACGGA			1203
Dm	1726		TCTAAAAACTTACCCATCAC	--AGGTCCATCAGAACCATCACAACATGGGCATGTAC	---		1780
Ds	1204		GlnGlnGlnThrGlyValProProValGlyAlaProProGlnAlaMetMetHisGlnGly				
Dm	1781		CAACAGCAGACGGGTGTGCCACCTGTGGGGGCCACCACCACAGGCAATGATGCACCAGGGC				
			Ser Gly				
Ds	1264		GlnGlyProProGlnMetHisGlnGlyHisLeuProGlyGlnHisThrProProSerGln				
Dm	1841		CAGGGGCCGCGCAGATGCACCAGGGCCATTTACGGGCCAACATACGCCGCCATCCCAA				
			EcoRI				
Ds	1324		AsnProAsnSerGlnSerSerGlyMetProSerProLeuTyrProTrpMetArgSerGln				
Dm	1898		AACCCGAATTCGCAATCCTCGGGGATGCCGTCAACACTGTATCCCTGGATGCCGAAGTCAG				
			▼	EXON 7	▼	INTRON 7	
			PheGlyLysCysGln				
Ds	1384		TTTGGTAAGTGTCAAGGAAAGTGATCCTCTA			Drosophila subobscura	1414
Dm	1958		TTTGGTAAGTGTCAAGGAAAGTGATCGACAATTCC			Drosophila melanogaster	1992

Met Thr Met Ser Thr Asn Asn Cys Glu Ser Met Thr Ser Tyr Phe 15
 Thr Asn Ser Tyr Met Gly Ala Asp Met His His Gly His Tyr Pro 30
 Gly Asn Gly Val Thr Asp Leu Asp Ala Gln Gln Met His His Tyr 45
 Ser Gln Asn Pro Asn His Gln Gly Asn Met Pro Tyr Pro Arg Phe 60
 Pro Pro Tyr Asp Arg Met Pro Tyr Tyr Asn Gly Gln Gly Met Asp 75
Gln Gln Gln Gln Gln His Gln Gly Tyr Ser Arg Pro Asp Ser Pro 90
 Ser Ser Gln Val Gly Gly Val Met Pro Gln Ala Gln Thr Asn Gly 105
 Leu Pro Thr Gly Gly Gln Leu Val Ala Gln Gln Gln Gln Gln Gln 120
Pro Gln Gln Gln Ser Gln Thr Pro Gln Gln Gln Gln Ala Gln Gln 135
Pro Gln Gln Gln Gln Gln Gln Leu Gln Gln Leu Pro Gln Val Thr 150
Gln Gln Val Thr His Pro Gln Gln Gln Gln Gln Gln Gln Gln Gln 165
Gln Gln Gln Pro Val Val Tyr Ala Ser Cys Lys Leu Gln Ala Ala 180
 Val Gly Val Gly Leu Gly Met Val Pro Glu Gly Gly Ser Pro Pro 195
 Leu Val Asp Gln Met Thr Gly His His Met Asn Ala Gln Met Ser 210
 Leu Pro His His Met Gly His Pro Gln Ala Gln Leu Gly Tyr Thr 225
 Asp Val Gly Val Pro Asp Val Thr Glu Val His Gln Asn His His 240
 Asn Met Gly Met Tyr Gly Gln Gln Gln Thr Gly Val Pro Pro Val 255
 Gly Ala Pro Pro Gln Ala Met Met His Gln Gly Gln Gly Pro Pro 270
 Gln Met His Gln Gly His Leu Pro Gly Gln His Thr Pro Pro Ser 285
 Gln Asn Pro Asn Ser Gln Ser Ser Gly Met Pro Ser Pro Leu Tyr 300
 Pro Trp Met Arg Ser Gln Phe Gly Lys Cys Gln 311

Figura 11. Secuencia proteica codificada por la ORF de *Drosophila subobscura*. Los residuos de cisteína estan enmarcados y las regiones ricas en glutamina estan subrayadas. Los aminoácidos con carga se indican con cuadros, negros para los ácidos y blancos para los aminoácidos básicos. Ninguno de dichos aminoácidos ha sido sustituido (ver fig. 12). Las puntas de flecha separan los péptidos codificados por los distintos exones (5, 6 y 7). La parte carboxiterminal, que incluye la secuencia homeobox, no ha sido determinada.

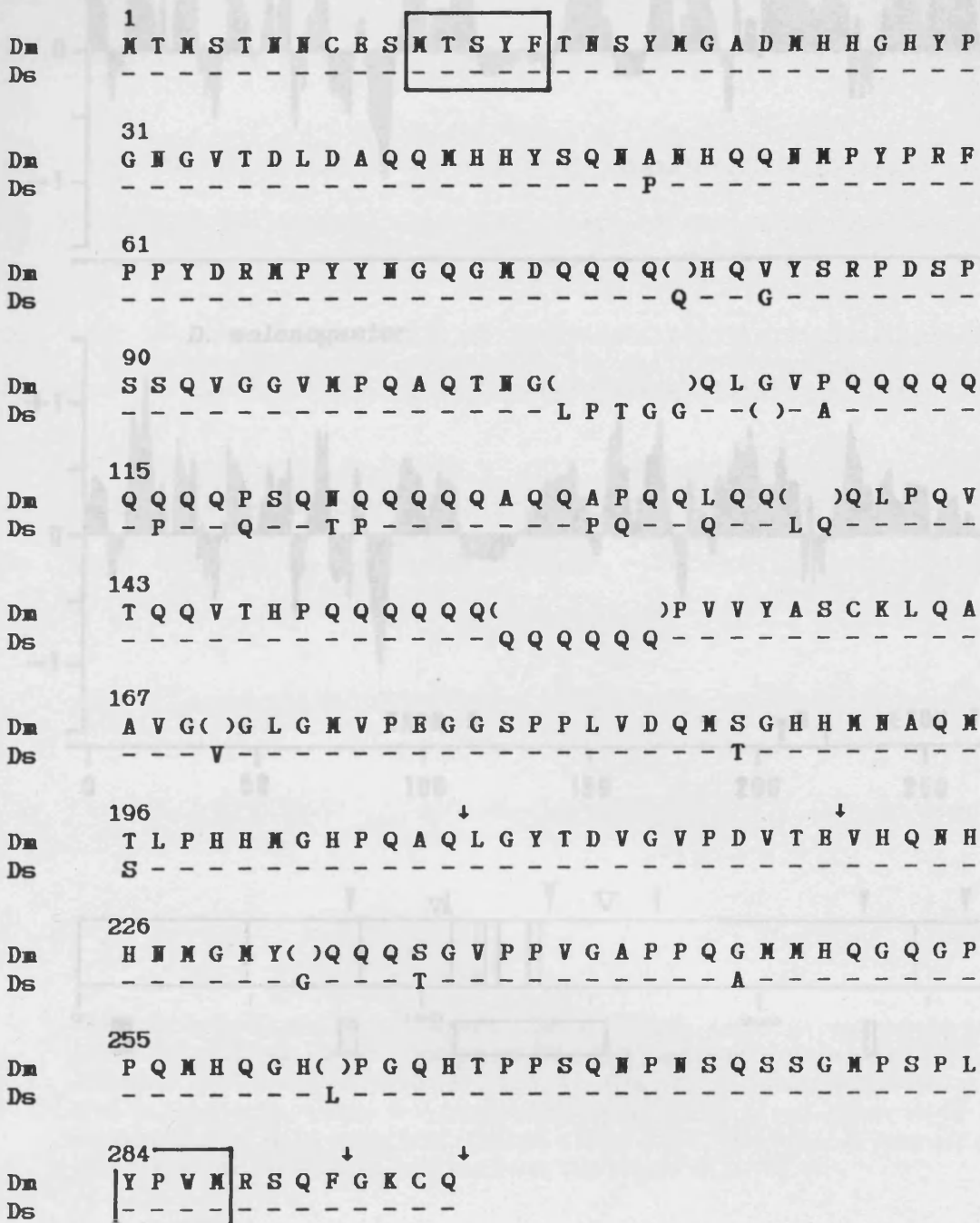


Figura 12. (Parte superior) Perfil de hidropatía de la proteína Antp de *D. subobscura* y de *D. melanogaster*, utilizando el algoritmo de Dayhoff. Los cuadrillos triangulares señalan en la parte superior de la gráfica. (Parte inferior) Representación de la variación de la proteína en las dos especies. El recuadrado representa al producto proteico Antp de *D. melanogaster* y, entre éste, la variación encontrada en *D. subobscura* (ver texto). Las flechas invertidas señalan las inserciones, la base de estas triángulos es proporcional a la cantidad de dichas inserciones. Las flechas indican las posiciones de unión entre las regiones codificadas en exones diferentes (las dos flechas de la parte C-terminal denotan las posiciones de los sitios alternativos para el procesamiento). Los dos péptidos comunes a otros genes del desarrollo están señalados con un recuadro. Hemos utilizado el código de una letra de Dayhoff.

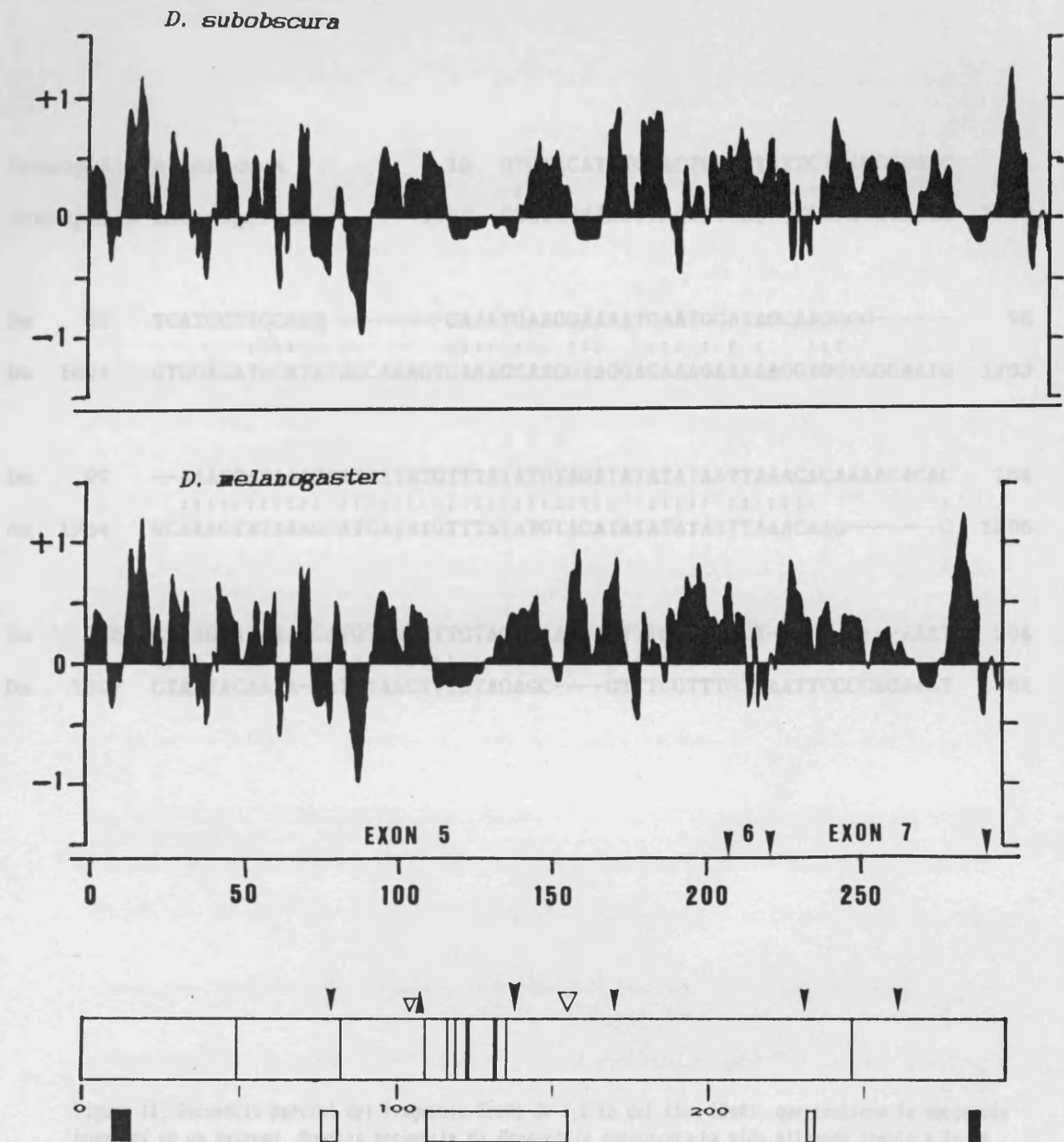


Figura 13. (Parte superior), Perfil de hidropatía de la proteína *Antp* de *D. subobscura* y de *D. melanogaster*, utilizando el algoritmo de Hopp y Woods. Los dominios hidrofílicos quedan en la parte superior de la gráfica. (Parte inferior), Representación de la variación de la proteína en las dos especies. El rectángulo representa el producto protéico *Antp* de *D. melanogaster* y, sobre éste, la variación encontrada en *D. subobscura* (método de KASSIS). Los triángulos invertidos señalan las inserciones; la base de estos triángulos es proporcional a la longitud de dichas inserciones. Las puntas de flecha señalan las inserciones o deleciones de uno o dos aminoácidos (ver fig. 12). Las sustituciones no conservativas de aminoácidos están marcadas con una línea vertical en el extremo. Bajo el esquema de la proteína, los recuadros blancos representan las regiones de la proteína cuyo contenido en Glutamina es muy elevado. Los recuadros negros señalan los dos péptidos comunes a otros genes del desarrollo.

```

Drosophila subobscura      18  GTGTACATCTCAACTCTCATATTCTCAGCTCGTG      51
      : : : : : : : : : : : : : : : : : :
Drosophila melanogaster   1660 GTGTAGTAACTTAAAAATCATTTCACGCCATCT      1693

Ds  52  TGATCCTTCCAAG-----CAAATGAAGGAAAAATGAATGGATAGCAAGGGG-----      96
      : : : : : : : : : : : : : : : : : :
Dm  1694 GTGGAAATCCATATACCAAAGTCAAAGGAACGAAGGAGAAAAGAAAAAGGAGGACGGAATG      1753

Ds  97  --AAACTATAAAGCATCATATGTTTATATGTAGATATATATAATTA AACACAAAAACACAC      154
      : : : : : : : : : : : : : : : : : :
Dm  1754 GCAAACATAAAGTATCATATGTTTATATGTAGATATATATATTTAAACAAG-----C      1806

Ds  155  CTAAGCATAAAGCATGTAAC TTTGTAGAGAAAAGAGTTTCGTTTTGT-----AAAT      204
      : : : : : : : : : : : : : : : : : :
Dm  1807 CTAATACAAAA-CATGTAAC TTTATAGAGC----GTTTCGTTTGTA AATTCCCCAGAAAT      1861

```

Figura 14. Secuencia parcial del fragmento EcoRI de 2,2 kb del clon λ DsA1, que contiene la secuencia homeobox en un extremo. Muestra secuencia de *Drosophila subobscura* ha sido alineada frente a la de *Drosophila melanogaster*. Esta secuencia corresponde al extremo no traducido, y muestra un elevado nivel de conservación respecto de *D. melanogaster*. Las referencias de esta especie patrón corresponden a las de la secuencia de SCHNEUWLY y otros (1986). Esta región de homología está situada entre los dos sitios de poliadenilación (ver esquema de la fig. 5).

```

AATDCAGAA
-----
VTAGGTCTT
-----
GAGTTCGTT
-----

```

Figura 15. Secuencia de un fragmento del clon λ DsA1, que coincide con la zona utilizada en nuestro screening. Su contenido en repetitones AA es considerablemente elevado. La traducción computacional de esta secuencia muestra que, de caso de ser traducida, el contenido en glutamato será superior incluso al del producto proteico del gen *Ubx*. Este sitio es homólogo al elemento homeobox. La homología de secuencias similares se una base de datos de genes con homología a nivel de las repetitones.

TCACATCAGCATCCACAGCAGGTCAGGTTTCAGGTTTCAGGTCAGCAACAGCAGCAGCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
AGTGTAGTCGTAGGTGTCTGTCACGTCACAGTCCAAAGTCCACGTCGTTGTCTGTCGTCGA

SerHisGlnHisProGlnGlnValGlnValGlnValGlnValGlnGlnGlnGlnGlnGlnPro

CATCCACACCCAGCAGCAGATTATTGCTCATACGCACGCCACCAACAACACCAACAGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
GTAGGTGTGGTGTCTCTAATAACGAGTATGCGTGCAGGTTGGTGTGTGGTGTCTGTC

HisProHisGlnGlnGlnIleIleAlaHisThrHisAlaHisGlnGlnHisGlnGlnGln

CAGCAGCAACAACAGTCTCCTCAGACCATTTCAGGCTCATCCCACAGCACTTGCAATAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
GTCGTCTGTTGTTCAGAGGAGTCTGGTAAGTCCGAGTAGGGCGTGTCTGTAACGTTATG

GlnGlnGlnGlnGlnSerProGlnThrIleGlnAlaHisProAlaGlnHisLeuGlnTyr

GGACATCCGCCTCCGCCGCCACAGGCACAGGGCGTGGCTAACTCCGCAGTGGTTGTAGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
CCTGTAGGCAGAGGCAGGCGGTGTCCGTGTCCCGCACCGATTGAGGCGTCACCAACATCCT

GlyHisProProProProProGlnAlaGlnGlyValAlaAsnSerAlaValValValGly

GGAGGAGGAGGAGGAGGCGCTGCCGTTGCCATTGCGGGTGTCTGCTGCTGAGAGCATCGAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
CCTCCTCCTCCTCCTCCGACGGCAACGGTAACGCCACGACGACGACTCTCGTAGCTA

GlyGlyGlyGlyGlyGlyAlaAlaValAlaIleAlaGlyAlaAlaAlaGluSerIleAsp

AATCCAGAA
-----+ 309
TTAGGTCTT

AsnProGlu

Figura 15. Secuencia de un fragmento del clon λ DsJ3, que hibrida con la sonda utilizada en nuestro screening. Su contenido en repeticiones CAG es extremadamente elevado. La traducción conceptual de esta secuencia muestra que, en caso de traducirse, el contenido en glutamina sera superior incluso al del producto proteico del gen *Anip*. Este clon no pertenece al complejo Antennapedia. La busqueda de secuencias similares en una base de datos solo proporciona homologas a nivel de las repeticiones.

Límite entre los intrones y los exones

	Sitio dador	Sitio aceptor
CONSENSO	CAG/GTAAGT A G	(T)nNCAG/G C T
Intrón 3		TTTCAG/A Ds :: : : : TTACAG/A Dm
Intrón 4	TAG/GTGAGT :: : : : TAG/GTAAGT	TCGCAG/T Ds :: : : : ATCCAG/T Dm
Intrón 5	CAG/GTGAGA :: : : : CAG/GTGAGC	AAACAG/T Ds :: : : : CTGCAG/T Dm
Intrón 6	GAG/GTAAGA :: : : : GAG/GTAAGA	ATTTAG/G Ds :: : : : TCACAG/G Dm
Intrón 7		
	D1 TTG/GTAAGT :: : : : TTG/GTAAGT	Ds Dm
	D2 AAG/GAAAGT :: : : : AAG/GAAAGT	Ds Dm

D1 y D2 corresponden a los dos sitios alternativos para el procesado del exón 7.

Tabla 4. Comparación de los sitios dadores y aceptores para el procesado del gen *Antennapedia* entre *Drosophila subobscura* y *D. melanogaster*. Estas posiciones están señaladas con puntas de flecha en la figura 10.

Conservación de aminoácidos en la proteína *Antp*

REGIÓN	LONGITUD		Nº Diferencias *	% Conservación **
	<i>D. sub</i>	<i>D. mel</i>		
Exón 5	221 aa	207 aa	13	94 %
Exón 6	13 aa	13 aa	0	100 %
Exón 7	77 aa	75 aa	2	97 %
TOTAL ***	311 aa	295 aa	15	95 %

* El Nº estimado de diferencias es el de sustituciones de aminoácidos más el nº de aminoácidos de *D. melanogaster* (sólo uno) que, en el alineamiento, no aparecen frente a una secuencia homóloga en *D. subobscura*.

** El porcentaje estimado de conservación es el nº de aminoácidos homólogos dividido por el número total de aminoácidos de *D. melanogaster* (x100).

*** El total se refiere al segmento de *D. subobscura* que presentamos alineado frente a *D. melanogaster* (región aminoterminal).

Tabla 5. Niveles de conservación en la secuencia del producto protéico de *Antp* en *D. subobscura*, en comparación con *D. melanogaster*. Las estimaciones están calculadas siguiendo la aproximación de KASSIS y otros (1986).

gca	Ala(A)	3	#	cag	Gln(Q)	56	#	uug	Leu(L)	1	#	uaa	Ter(.)	0
gcc	Ala(A)	5	#	---	Gln(Q)	73	#	---	Leu(L)	12	#	uag	Ter(.)	0
gcg	Ala(A)	4	#	gaa	Glu(E)	1	#	aaa	Lys(K)	0	#	uga	Ter(.)	0
gcu	Ala(A)	0	#	gag	Glu(E)	2	#	aag	Lys(K)	2	#	---	Ter(.)	0
---	Ala(A)	12	#	---	Glu(E)	3	#	---	Lys(K)	2	#	aca	Thr(T)	4
aga	Arg(R)	0	#	gga	Gly(G)	5	#	aug	Met(M)	22	#	acc	Thr(T)	4
agg	Arg(R)	0	#	ggc	Gly(G)	17	#	---	Met(M)	22	#	acg	Thr(T)	7
cga	Arg(R)	1	#	ggg	Gly(G)	3	#	uuc	Phe(F)	2	#	acu	Thr(T)	0
cgc	Arg(R)	3	#	ggu	Gly(G)	7	#	uuu	Phe(F)	1	#	---	Thr(T)	15
cgg	Arg(R)	0	#	---	Gly(G)	32	#	---	Phe(F)	3	#	ugg	Trp(W)	1
cgu	Arg(R)	0	#	cac	His(H)	9	#	cca	Pro(P)	11	#	---	Trp(W)	1
---	Arg(R)	4	#	cau	His(H)	11	#	ccc	Pro(P)	11	#	uac	Tyr(Y)	11
aac	Asn(N)	9	#	---	His(H)	20	#	ccg	Pro(P)	12	#	uau	Tyr(Y)	2
aau	Asn(N)	5	#	aua	Ilu(I)	0	#	ccu	Pro(P)	2	#	---	Tyr(Y)	13
---	Asn(N)	14	#	auc	Ilu(I)	0	#	---	Pro(P)	36	#	gua	Val(V)	0
gac	Asp(D)	5	#	auu	Ilu(I)	0	#	agc	Ser(S)	3	#	guc	Val(V)	4
gau	Asp(D)	4	#	---	Ilu(I)	0	#	agu	Ser(S)	4	#	gug	Val(V)	11
---	Asp(D)	9	#	cua	Leu(L)	0	#	uca	Ser(S)	2	#	guu	Val(V)	3
ugc	Cys(C)	1	#	cuc	Leu(L)	1	#	ucc	Ser(S)	4	#	---	Val(V)	18
ugu	Cys(C)	2	#	cug	Leu(L)	9	#	ucg	Ser(S)	5	#	nnn	???(X)	0
---	Cys(C)	3	#	cuu	Leu(L)	0	#	ucu	Ser(S)	1	#	TOTAL		311
caa	Gln(Q)	17	#	uua	Leu(L)	1	#	---	Ser(S)	19	#			

gca	Ala(A)	0	#	cag	Gln(Q)	52	#	uug	Leu(L)	3	#	uaa	Ter(.)	0
gcc	Ala(A)	6	#	---	Gln(Q)	64	#	---	Leu(L)	10	#	uag	Ter(.)	0
gcg	Ala(A)	6	#	gaa	Glu(E)	0	#	aaa	Lys(K)	0	#	uga	Ter(.)	0
gcu	Ala(A)	0	#	gag	Glu(E)	3	#	aag	Lys(K)	2	#	---	Ter(.)	0
---	Ala(A)	12	#	---	Glu(E)	3	#	---	Lys(K)	2	#	aca	Thr(T)	3
aga	Arg(R)	0	#	gga	Gly(G)	6	#	aug	Met(M)	22	#	acc	Thr(T)	4
agg	Arg(R)	0	#	ggc	Gly(G)	10	#	---	Met(M)	22	#	acg	Thr(T)	5
cga	Arg(R)	1	#	ggg	Gly(G)	6	#	uuc	Phe(F)	1	#	acu	Thr(T)	0
cgc	Arg(R)	3	#	ggu	Gly(G)	8	#	uuu	Phe(F)	2	#	---	Thr(T)	12
cgg	Arg(R)	0	#	---	Gly(G)	30	#	---	Phe(F)	3	#	ugg	Trp(W)	1
cgu	Arg(R)	0	#	cac	His(H)	12	#	cca	Pro(P)	5	#	---	Trp(W)	1
---	Arg(R)	4	#	cau	His(H)	8	#	ccc	Pro(P)	12	#	uac	Tyr(Y)	11
aac	Asn(N)	14	#	---	His(H)	20	#	ccg	Pro(P)	11	#	uau	Tyr(Y)	2
aau	Asn(N)	1	#	aua	Ilu(I)	0	#	ccu	Pro(P)	6	#	---	Tyr(Y)	13
---	Asn(N)	15	#	auc	Ilu(I)	0	#	---	Pro(P)	34	#	gua	Val(V)	0
gac	Asp(D)	8	#	auu	Ilu(I)	0	#	agc	Ser(S)	5	#	guc	Val(V)	6
gau	Asp(D)	1	#	---	Ilu(I)	0	#	agu	Ser(S)	2	#	gug	Val(V)	6
---	Asp(D)	9	#	cua	Leu(L)	0	#	uca	Ser(S)	1	#	guu	Val(V)	6
ugc	Cys(C)	2	#	cuc	Leu(L)	0	#	ucc	Ser(S)	5	#	---	Val(V)	18
ugu	Cys(C)	1	#	cug	Leu(L)	7	#	ucg	Ser(S)	6	#	nnn	???(X)	0
---	Cys(C)	3	#	cuu	Leu(L)	0	#	ucu	Ser(S)	1	#	TOTAL		295
caa	Gln(Q)	12	#	uua	Leu(L)	0	#	---	Ser(S)	20	#			

Tabla 6. Frecuencia de uso de codones en la ORF del gen *Antp* en *D.subobscura* (parte superior) y en *D.melanogaster* (parte inferior).

Conservación de nucleótidos en la unidad de transcripción *Antp*

REGIÓN	LONGITUD		Nº Diferencias * <i>D. subobscura</i>	% Conservación **
	<i>D. sub</i>	<i>D. mel</i>		
Intrón 3	>15 kb	≈25 kb	N.D.	?
Exón 4	275 pb	251 pb	60	76 %
Intrón 4	≈150 pb	154 pb	N.D.	N.D.
Exón 5				
Región no traducida	102 pb	156 pb	108	31 %
Región traducida	663 pb	621 pb	83	87 %
Intrón 5	146 pb	379 pb	293	23 %
Exón 6	39 pb	39 pb	4	90 %
Intrón 6	127 pb	134 pb	69	49 %
Exón 7	232 pb	226 pb	25	89 %
Intrón 7	≈11 kb	≈11 kb	N.D.	?
TOTAL ***	1584 pb	1806 pb	642	64 %

* El Nº estimado de diferencias es el de sustituciones nucleotídicas más el nº de nucleótidos de *D. melanogaster* que, en el alineamiento, no aparecen frente a una secuencia homóloga en *D. subobscura*.

** El porcentaje estimado de conservación es el nº de nucleótidos homólogos dividido por el número total de bases de *D. melanogaster* (x100).

*** El total se refiere los segmentos de *D. subobscura* que presentamos alineados frente a *D. melanogaster*.

Tabla 7. Niveles de conservación nucleotídica en la secuencia del gen *Antp* de *Drosophila subobscura* en comparación con *D. melanogaster*, a lo largo de las distintas regiones. Las estimaciones están calculadas siguiendo la aproximación de KASSIS y otros (1986).

Secuencias conservadas en el exón 4 de *Antp*

<i>D. melanogaster</i>	CONSENSO	<i>D. subobscura</i>
<hr/>		
TAGAGGAA		
30 TtGAGGAA 37		62 TtGAGGAA 69
155 TAGAGGAA 162		170 cAGAGGAA 177
45 TAGAGGAt 52		77 TAGAGGAt 84
<hr/>		
TACTTAAG		
58 TACTTAAG 65		90 TACTTAAG 97
240 TACTTAAG 247		288 TACTTAAG 295
<hr/>		
TCAAAA		
18 TCAAAA 23		50 TCAAAA 55
24 TCAAAA 29		56 TCAAAA 61
68 TCAAAA 73		100 TCAAcA 105
251 TCAAAg 256		299 TCAAAA 305

▼		
136	T T T T G G C C A A G A	149 <i>D. melanogaster</i>
148	T T T T G G C C A A G A	159 <i>D. subobscura</i>
▲		

Los tres grupos superiores corresponden a secuencias repetidas: los nucleótidos indicados en letra minúscula son los que se salen del consenso que se propone. La secuencia inferior es un palíndromo imperfecto.

Tabla 8. Regiones conservadas en el líder del gen *Antp*. Estas regiones están señaladas con asteriscos en la figura 10.

CONCLUSIONES

CONCLUSIONES

1. Las 30 kb de la región 3' del gen Antp de *Drosophila subobscura* presentan una estructura casi idéntica a la de la región homóloga de *D.melanogaster*, con la excepción de un pequeño intrón (I-6) cuyo tamaño es un tercio del de la especie patrón.
2. Encontramos un segmento muy conservado que se sitúa en la región del líder 5' no traducido del transcrito. Este segmento consiste en repeticiones de las secuencias TAGAGGAA, TACTTAAG, TCAAAA. Además este segmento conservado contiene el palíndromo imperfecto TTTTGCCAAGA.
3. Los primeros datos que tenemos sobre el otro extremo del transcrito, el trailer 3' no traducido, indican que hay una baja homología general, donde destaca la presencia de, al menos, un segmento de 50 pb que presenta un elevado porcentaje de conservación.
4. La conservación en los intrones pequeños es mínima, al igual que en los exones no traducidos (salvo la conservación citada en la conclusión 2).
5. El producto protéico, analizado en su mayor parte, ha acumulado la variación de forma muy desigual: casi toda la variación (inserciones y sustituciones de aminoácidos) se ha acumulado en la región de poliglutamina, codificada por repeticiones múltiples del codón CAG.
6. Especulamos que las regiones ricas en codones repetidos

tienen una función relacionada con la captación de
variabilidad en genes críticos para el desarrollo.

APENDICE :
MATERIALES Y METODOS

MATERIALES Y METODOS

1. ENZIMAS

1.1. DNA ligasa

Para la unión de fragmentos de DNA hemos utilizado el enzima DNA ligasa del fago T4 (E.C. 6.5.1.1), del proveedor Boehringer Mannheim (B.M.) (Cat.No. 799 009). El tampón utilizado difiere del recomendado por MANIATIS y otros en el pH del tampón. La composición de nuestro tampón es: 50 mM Tris-HCl, pH 8.0; 10 mM MgCl₂; 10 mM DTT; 1 mM espermidina; 2 mM ATP; 100 µg/ml BSA.

Siempre que ha sido posible, las ligaciones se han realizado en un volumen mínimo, normalmente entre 5 y 10 µl. La cantidad de ligasa utilizada ha sido siempre de 1-2 U. La relación molar de vector a inserto ha variado según el experimento, aunque se ha intentado mantener una relación del orden de 1:1 en la mayoría de los casos.

Para la ligación de extremos cohesivos, la temperatura de incubación ha sido de 15-16°C y el tiempo de incubación de 12-16 h. Tras esta incubación, la mezcla se ha mantenido a 4°C hasta el momento de la transformación o transfección.

Para la ligación de extremos romos, la incubación se ha llevado a cabo a temperatura ambiente (en presencia de un 10% de PEG), manteniéndose a esta temperatura hasta la transformación.

1.2. Polimerasas

Para la preparación de sondas radioactivas, hemos utilizado indistintamente la DNA polimerasa I de E.coli de B.M. (Cat.No. 104 493) en una reacción de "nick translation", o bien el fragmento Klenow en una reacción de "oligolabelling". El tampón de incubación utilizado para la DNA pol I ha sido el que recomiendan MANIATIS y otros (tampón 10x: 0.5 M Tris-HCl, pH 7.5; 0.1 M MgSO₄; 1 mM DTT; 500 µg/ml BSA) y la incubación se ha realizado a 15-16°C durante 1-2 h.

Para el "oligolabelling", el tampón empleado ha sido el recomendado por FEINBERG y VOLGESTEIN (1983). El enzima utilizado fué el fragmento Klenow de la DNA polimerasa I (Labelling Grade; 2U/µl) de B.M. (Cat.No. 1008 412). Normalmente, esta reacción la hemos dejado durante 12-16 h a temperatura ambiente. Las condiciones exactas para ambos tipos de reacciones de marcaje están descritas en el apartado 9.

Las primeras reacciones de secuenciación las hemos efectuado con el fragmento Klenow de la DNA pol I (Sequencing Grade; 1 U/µl) de B.M. (Cat.No. 997 463), en las condiciones detalladas en el apartado 16. No obstante, la mayor parte de las reacciones de secuenciación las hemos llevado a cabo con la DNA polimerasa del fago T7 del proveedor Pharmacia-LKB, siguiendo las condiciones recomendadas por el fabricante

(apartado 16).

1.3. Enzimas de restricción

Las digestiones de DNA con endonucleasas de restricción han sido realizadas utilizando enzimas de B.M., con los tampones recomendados por el fabricante. Los cinco tipos de tampones utilizados son los denominados A, B, L, M y H, cuyas composiciones de trabajo se describen a continuación. Tampón A: 33 mM Tris-acetato, pH 7.9; 10 mM Mg.acetato; 66 mM K.acetato; 0.5 mM DTT. Tampón B: 10 mM Tris-HCl, pH 8.0; 5 mM MgCl₂; 100 mM NaCl; 1 mM β-mercaptoetanol. Tampón L: 10 mM Tris-HCl, pH 7.5; 10 mM MgCl₂; 1mM ditioeritrol (DTE). Tampón M: 10 mM Tris-HCl, pH 7.5; 10 mM MgCl₂; 50 mM NaCl; 1 mM DTE. Tampón H: 50 mM Tris-HCl, pH 7.5; 10 mM MgCl₂; 100 mM NaCl; 1 mM DTE.

Para las digestiones dobles, se ha utilizado el tampón que permitiera la máxima actividad de los enzimas usados, según la hoja de instrucciones del fabricante. En caso de incompatibilidad, hemos procedido a precipitar el DNA digerido con uno de los enzimas y, una vez resuspendido, hemos digerido con el segundo enzima.

La cantidad de enzima (número de unidades) utilizado para las digestiones ha sido siempre la suficiente para digerir todo el DNA en 1 h, si bien la incubación ha sido normalmente de 2-4 h a 37°C (salvo las digestiones de DNA genómico que se llevaron a cabo durante 12-16 h). Para las digestiones de DNA genómico de *Drosophila*, el tiempo de incubación ha sido de 12-16 h. El volumen añadido de solución enzimática ha sido, en todos los casos, inferior al 10% del volumen final, con el fin de evitar la relajación de la especificidad que puede producirse en muchos enzimas de restricción.

En todas las incubaciones con estos enzimas, se añadió a la mezcla de digestión BSA libre de nucleasas (BRL), a una concentración final de 100 µg/ml.

Antes de interrumpir la reacción, siempre se ha comprobado que ésta es completa (en un minigel). Cuando se ha alcanzado este punto, los tubos de reacción se han pasado a 4°C, sin ningún otro tratamiento. Cuando ha sido necesario inactivar los enzimas (por ejemplo, durante la preparación de vectores), la mezcla de digestión se ha calentado a 65-68°C durante 15-20 min.

1.4. Desoxirribonucleasa

En las reacciones de "nick translation" hemos usado la DNasa I (E.C. 3.1.21.1) de B.M. (Cat.No. 104 132). Se preparó una solución stock a una concentración de 1 mg/ml en un tampón de nick translation que incluye un 50% de glicerol, y se ha mantenido a -20°C hasta la actualidad. Para cada reacción de marcaje, se preparó una dilución 10⁻⁴, añadiendo 1 µl de la dilución a la mezcla de reacción según se describe en el punto

9.1.

Para la preparación de DNA del fago lambda a gran escala (apartado 6.3.2) también se utilizó este producto de B.M., aunque añadiéndolo directamente en polvo.

1.5. Ribonucleasa

Para la eliminación del RNA durante la preparación de DNA, hemos utilizado RNasa A de B.M. (Cat.No. 109 169). Se preparó una solución stock con una concentración de 10 mg/ml en un tampón (10 mM Tris-HCl, pH 8.0, 15 mM NaCl). Esta solución se calentó a 100°C durante 15-20 min, y se dejó enfriar lentamente para facilitar la renaturalización del enzima. El stock se mantiene a -20°C indefinidamente. Para la preparación de DNA del fago lambda, hemos utilizado RNasa cruda de B.M. (Cat.No. 109 126).

1.6. Fosfatasa alcalina

Para la defosforilación de los extremos 5' del DNA, hemos usado siempre la fosfatasa alcalina de intestino de ternera (E.C. 3.1.3.1) suministrada por B.M. a una concentración de 1 U/ μ l (Cat.No.713 023) y el tampón de incubación es el recomendado por MANIATIS (tampón 10x: 0.5 M Tris-HCl, pH 9.0; 10 mM MgCl₂; 1 mM ZnCl₂; 10 mM espermidina). El protocolo seguido fue, al principio, el descrito en dicho manual. Posteriormente, hemos seguido una alternativa mucho más simple (ver apartado 12.1).

1.7. Polinucleótido kinasa

Las reacciones de marcaje de extremos 5' se ha llevado a cabo con el enzima polinucleótido kinasa de B.M. (Cat.No. 633 542) sobre DNA previamente defosforilado. El tampón se preparó a una concentración 10x (0.5 M Tris-HCl, pH 7.6; 0.1 M MgCl₂; 50 mM DTT; 1 mM espermidina; 1 mM EDTA) y se ha mantenido a -20°C).

La incubación se llevó a cabo durante 45 min a 37°C, siguiendo el protocolo del apartado 9.3.

1.8. Proteínasa K

Para la digestión de las proteínas contaminantes durante la preparación de DNA genómico de Drosophila y de DNA del bacteriófago lambda hemos utilizado Proteínasa K del proveedor B.M. (Cat.No. 1000 144). Para el DNA genómico hemos utilizado una solución stock a una concentración de 20 mg/ml (en agua), mantenida a -20°C; para la extracción a gran escala de DNA del fago lambda se añadió este enzima en polvo.

1.9. Lisozima

Para producir la lisis bacteriana durante la preparación de DNA plasmídico, el enzima utilizado ha sido la lisozima (E.C. 3.2.1.17) (de clara de huevo, cristalizada) del proveedor B.M. (Cat.No. 837 059). El enzima fué resuspendido justo antes de su uso.

2. TAMPONES Y REACTIVOS

Para preparar los tampones descritos en este trabajo, hemos partido de soluciones concentradas, preparadas según se describe en MANIATIS y otros. Las sales utilizadas, los alcoholes y los disolventes orgánicos son de calidad "Para Análisis" del fabricante MERCK, excepto el CsCl que es de B.M. (Cat.No. 757 292). El fenol de MERCK (Ref. 206) se ha utilizado tal como viene de la casa comercial, sin ningún otro paso de purificación.

El isopropil- β -D-tiogalactopiranosido (IPTG) y el 5-bromo-4-cloro-3-indolil- β -D-galactopiranosido (X-gal) fueron adquiridos del proveedor B.M. (Cat.No. 724 815 y Cat.No. 703729, respectivamente), al igual que el Tris (Cat.No. 604 205).

Aquellos tampones y soluciones cuya composición no describimos han sido preparados según el manual mencionado. La seroalbúmina bovina (BSA) de las mezclas de hibridación es del proveedor SERVA (Ref. 11930). En el resto de soluciones se ha utilizado BSA libre de nucleasas (acetilada) del proveedor BRL (Cat.No. 540-5561UB).

3. CEPAS UTILIZADAS

3.1. E.coli

NM538 : F- supF hsdR mcrB (r-m+) lamda-.

Q358 : supF hsdR mcrB (r-m+) attphi80

DH5 α : F- SupE44 hsdR17 mcrA mcrB (r-m+) gyrA96 recA1 relA1 endA1 thi-1 attfi80 lacZdeltaM15 lambda-.

JM105 : [F' traD36 proAB+ lacIqZdeltaM15] delta(pro-lac) hsdR4 (r-m+) rpsL sbcB15 thi endA1.

3.2. Drosophila

La cepa a partir de cuyo DNA se preparó la librería genómica fué la H-271 de *Drosophila subobscura*. Para las

hibridaciones con DNA genómico se utilizó esta misma cepa, así como un control con DNA de la cepa Oregon de *Drosophila melanogaster*.

4. VECTORES DE CLONACION

4.1. Derivados del bacteriófago lambda

La librería genómica utilizada en este trabajo tenía como vector un derivado del fago lambda: lambda EMBL4, un vector de reemplazamiento diseñado por FRISCHAUF y otros (1983).

En esquema, la genoteca se preparó digiriendo parcialmente DNA genómico de la cepa H-271 de *Drosophila subobscura* con el enzima de restricción Mbo I e insertándolo en el sitio BamH I del vector EMBL4. Una alícuota de esta librería, una vez amplificada, nos fue facilitada por G. Marfany y R. Gonzalez-Duarte (Universidad de Barcelona) y ha sido nuestro punto de partida para la clonación.

Así pues, todos los clones del fago lambda de este trabajo son derivados del EMBL4. Este vector carece del gen de la proteína represora cI, así como de los genes para el sistema de integración, de tal forma que estos virus tienen necesariamente un crecimiento lítico.

4.2. Derivados del virus M13

Parte de la secuencia que hemos obtenido en este trabajo ha sido obtenida a partir de fragmentos clonados en este virus filamentososo de cadena simple, aprovechando su fase replicativa de doble cadena (RF) para la digestión con un enzima de restricción y tratamientos análogos a los de los plásmidos.

El virus M13 no produce lisis, sino que su infección provoca un crecimiento retardado de las bacterias. Así, cuando se lleva a cabo una infección o una transfección, el resultado es la aparición de calvas que no son placas de lisis, sino más bien calvas turbias debidas al crecimiento enlentecido de las bacterias.

De las distintas variantes de este vector, las utilizadas han sido el M13mp10 y el M13mp11.

Estos vectores contienen un segmento del gen que codifica la β -galactosidasa (lacZ). concretamente el extremo 5', que corresponde a la parte aminoterminal del enzima. Ambos vectores contienen un sitio de clonación múltiple (insertado en dicho segmento, pero manteniendo la pauta de lectura) que incluye dianas para los mismos enzimas de restricción (EcoRI, SacI, SmaI, XmaI, BamHI, XbaI, SalI, AccI, HindII, PstI y HindIII) aunque en direcciones opuestas.

Cuando se cultivan estos vectores, en su estado intacto, en un hospedador bacteriano mutante que sintetiza la enzima β -galactosidasa carente de los aminoácidos 11 a 41 (mutación lacZdeltaM15), la síntesis de ese fragmento aminoterminal

codificado por el vector es capaz de complementar dicha mutación y dar un fenotipo β -galactosidasa positivo (α -complementación). La inserción de un fragmento de DNA (inserto) en el sitio de clonación múltiple (MCS) provoca, en la mayoría de los casos, la inactivación del fragmento codificado por el vector, por lo que ya no se produce esa complementación. Esto es la base de la selección de los clones recombinantes frente a los parentales: utilizando un medio indicador IPTG/X-gal, los clones recombinantes son de color blanco, mientras que los resultantes de la recircularización del vector producen un halo azul.

La adsorción de este virus a las células de E.coli requiere que éstas posean el pelo F, cuya síntesis está codificada en el factor de fertilidad F (episoma F). Por ello, las cepas hospedadoras para este virus deben ser F'. Con el fin de mantener esta condición F' las cepas hospedadoras llevan una delección de los genes responsables de la biosíntesis de prolina, los cuales han sido insertados en el episoma F. El cultivo de la cepa hospedadora en un medio mínimo, permite evitar la pérdida de dicho factor.

Además, el DNA cromosómico del hospedador lleva también una delección completa del gen lacZ, mientras que el episoma lleva una inserción de la variante mutante del gen lacZdeltaM15. Además, el episoma lleva la mutación lacIq, que conlleva la superproducción del represor lac, permitiéndolo así regular la expresión del operón lac mediante el inductor IPTG.

4.3. Plásmidos

Para la subclonación de fragmentos de restricción de los clones aislados de la genoteca, hemos utilizado los siguientes vectores.

pUC18 y pUC19. Estos vectores contienen un sitio de clonación múltiple, con dianas para los siguientes enzimas: EcoRI, SacI, KpnI, SmaI, XmaI, BamHI, XbaI, SalI, AccI, HindII, PstI, SphI y HindIII, dispuestos en ordenación opuesta. Este sitio de clonación, al igual que en los vectores del apartado anterior, está situado en el interior de un segmento del gen lacZ capaz de manifestar α -complementación frente a la mutación lacZdeltaM15.

Por este motivo los vectores de esta serie permiten la selección de los clones recombinantes por su coloración blanca, por el mismo mecanismo descrito en el apartado anterior.

5. METODOS MICROBIOLÓGICOS

El mantenimiento y propagación de las cepas bacterianas y víricas utilizadas a lo largo de este trabajo se ha llevado a cabo siguiendo los protocolos de MANIATIS y otros.

5.1 Medios de cultivo

El medio de cultivo empleado habitualmente en este trabajo es el LB, aunque con distintos aditivos. Su composición es:

Triptona	10 g
Extracto de levadura	5 g
NaCl	10 g
Agua destilada	hasta 1 l

(pH ajustado a 7.5)

Para la preparación de medio sólido se añadió agar al 15%; el agar blando para la obtención de placas de lisis (fago lambda y fago M13) se preparó añadiendo agar hasta una concentración del 7%. Este medio se ha utilizado incluso para el screening de la genoteca y no encontramos la necesidad de utilizar agarosa como solidificante.

La ampicilina para la selección de transformantes (B.M.; Cat.No. 835 242) se utilizó en la dosis máxima recomendada por MANIATIS y otros (50 µg/ml).

Para la detección de actividad β-galactosidasa, el medio indicador LB/Amp/X-gal/IPTG se preparó de la siguiente forma:

I. Preparar una disolución de X-gal en dimetilformamida con una concentración de 20 mg/ml.

II. Preparar una disolución de IPTG en agua, con una concentración de 20 mg/ml (84 mM) y esterilizarla por filtración (ambas disoluciones pueden almacenarse a -20°C).

III. Por cada placa que se desee preparar, mezclar 40 µl de cada una de las dos disoluciones con 120 µl de agua estéril. Extender esta mezcla sobre una placa de medio sólido LB con ampicilina.

IV. Dejar secar en la cámara estéril. Estas placas pueden mantenerse, al menos, durante dos semanas a 4°C, protegidas de la luz.

Para la selección del factor F' en las cepas hospedadoras del virus M13, se ha utilizado un medio mínimo preparado de la siguiente forma:

I. Preparar un stock de sales M9 (10x) con la siguiente composición:

Na ₂ HPO ₄	60 g
KH ₂ PO ₄	30 g
NaCl	5 g
NH ₄ Cl	10 g
Agua destilada	hasta 1 l

II. Preparar un stock 1 M de MgSO₄ y un stock 1 M de CaCl₂.

III. Esterilizar en el autoclave las tres soluciones.

IV. Preparar un stock al 20% de glucosa y una solución de tiamina (10 mg/ml) y esterilizarlas por filtración.

V. Para 1 l de medio mínimo, disolver 15 g de agar en 900

ml de agua y añadir 100 ml de sales M9 (10x). Esterilizar en el autoclave.

VI. Cuando la temperatura de la mezcla anterior sea alrededor de 55°C, añadir 1 ml del stock MgSO₄, 100 µl del stock de CaCl₂, 1 ml de tiamina y 10 ml de glucosa.

Estas placas pueden mantenerse a 4°C durante tres meses.

5.2. Preparación de células competentes

El protocolo utilizado está basado en MANIATIS y otros, siendo la principal diferencia que, en nuestro protocolo, las células se resuspenden en un volumen mucho menor que el recomendado por ese manual. Las células competentes preparadas con este protocolo tienen, mantenidas a -80°C, la misma eficiencia que las preparadas en el momento. Las eficiencias de transformación obtenidas con la cepa DH5α son del orden de 10⁷ y 10⁸ transformantes/µg plásmido intacto. El protocolo seguido es el siguiente.

I. Cultivar 200 ml de medio LB con la cepa apropiada hasta una OD₆₀₀ = 0.2 (o bien OD₅₅₀ = 0.5), realizando el inóculo a partir de un cultivo de noche.

II. Enfriar en hielo durante 10 min.

III. Centrifugar en el rotor JA-20 a 3.000 rpm durante 15 min a 4°C.

IV. Resuspender las células en un total de 100 ml de 0.1 M CaCl₂ frío y estéril.

V. Mantener en hielo durante 20-30 min.

VI. Centrifugar en el mismo rotor a 3.000 rpm durante 15 min a 4°C,

VII. Resuspender en 2 ml de 0.1 M CaCl₂.

VIII. Añadir 300 µl de glicerol estéril y distribuir en alícuotas de 100 µl.

IX. Congelar en nitrógeno líquido y mantener a -80°C.

Las células preparadas con este sistema mantienen la misma eficiencia de transformación al cabo de 1 año.

5.3. Transformación

El protocolo seguido para la introducción de DNA en las células competentes ha sido:

I. Descongelar sobre hielo un tubo de células competentes preparadas según el protocolo de la sección anterior. Dejar sobre el hielo hasta que se descongelen (20-30 min).

II. Añadir el DNA o mezcla de ligación en cantidades apropiadas y mezclar suavemente.

III. Mantener en hielo durante 30-60 min (mínimo 10 min), mezclando ocasionalmente.

IV. Transferir a un baño de 42°C durante exactamente 2 min.

V. Durante el instante que transcurre en poder pasar a la etapa vi, mantener el tubo sobre hielo.

VI. Añadir esta mezcla a un tubo de 10 ml estéril, que contenga 1 ml de medio LB sin antibiótico, precalentado a 37°C y previamente aireado mediante agitación intensa en el vórtex.

VII. Incubar sin agitación durante 1 h a 37°C, con el fin de permitir la expresión del gen que codifica la resistencia a la ampicilina .

VIII. Extender 300-400 µl de este cultivo sobre placas de medio LB con ampicilina (distribuir el cultivo de la etapa anterior en 3 placas), que contengan, si procede, la mezcla indicadora IPTG/Xgal.

IX. Invertir e incubar durante 12-16 h a 37°C.

X. Si las colonias resultantes no muestran diferencias apreciables de coloración, dejar las placas durante 1-2 h a 4°C y volver a incubarlas a 37°C durante 2-4 h.

XI. Inocular las colonias recombinantes (de color blanco) en un tubo con 5 ml de medio LB con ampicilina.

5.4. Preparación de células para la infección con el bacteriófago lambda (plating bacteria)

Este protocolo produce células aptas para la infección, que pueden mantenerse a 4°C durante 3-4 semanas. Está basado en MANIATIS.

I. Inocular 5 ml de medio LB con una colonia de la cepa NM 538 de E.coli. Cultivar a 37°C durante 12-16 h.

II. Añadir 1 ml del cultivo anterior a 100 ml de medio LB que contiene maltosa al 0.2%. Cultivar hasta que la OD₅₄₀ sea 0.8 aprox. (si se utiliza una cepa rec⁻, esperar hasta una OD₅₄₀ = 2).

III. Centrifugar en el rotor JA-20 a 7.000 rpm durante 10 min a 15°C.

IV. Eliminar el sobrenadante y resuspender el sedimento con 40 ml de 10mM MgSO₄ estéril.

V. Mantener a 4°C. Para una placa de cultivo de 90 mm de diámetro, utilizar 100 µl de esta suspensión celular.

También es posible inocular la colonia directamente en los 100 ml de medio (etapa ii) y cultivar durante 12-16 h. Las células obtenidas de esta manera son también utilizables, aunque su eficiencia de infección es algo menor.

5.5. Obtención de placas de lisis del fago lambda

Para obtener placas de lisis, hemos seguido el siguiente protocolo:

I. Mezclar 100 µl de plating bacteria con una cantidad de virus dada.

II. Dejar a temperatura ambiente durante 20 min para permitir

la adsorción de los virus a las bacterias.

III. Añadir 3 ml de agar blando, fundido a 47°C y verter inmediatamente sobre una placa de medio LB sólido (calentada previamente a 37°C).

IV. Dejar solidificar e incubar durante 12 h, hasta que aparezcan las placas de lisis.

Efectuando diluciones seriadas del stock vírico en tampón SM se ha obtenido el título (pfu/μl) de los clones aislados a lo largo de este trabajo, así como de la librería genómica de *Drosophila subobscura*. La composición (por litro) del tampón SM es:

NaCl	5.8 g
MgSO ₄ .7H ₂ O	2 g
Gelatina al 2%	5 ml
1 M Tris-HCl, pH 8.0	50 ml
Agua destilada	hasta 1 l

(esterilizar en el autoclave)

5.6. Preparación de lisados líquidos

Este es uno de los "cuellos de botella" que hemos encontrado a lo largo del trabajo. No conocemos ningún método infalible para la obtención (al primer intento) de lisados líquidos a partir de un clon dado. El protocolo más reproducible que hemos conseguido es una adaptación del recomendado por AUSUBEL y otros (1988).

I. Mezclar 100 μl de plating bacteria (sección anterior), 100 μl de tampón de adsorción (10 mM MgCl₂; 10 mM CaCl₂), y 100 μl de un stock preparado por difusión de una placa de lisis en 0.5 ml de tampón SM (o bien de 10⁵ a 10⁸ pfu de cualquier otro tipo de stock).

II. Incubar durante 15 min a 37°C.

III. Con estas células infectadas, inocular 50 ml de medio LB con 10 mM MgCl₂ y cultivar con la máxima agitación durante 8-12 h a 37°C.

IV. En caso de no aparecer la lisis, aumentar el volumen de cultivo hasta 100 ml con el mismo tipo de medio.

V. Cuando aparezca la lisis, añadir unas gotas de cloroformo y agitar durante 10 min más.

VI. Centrifugar en el rotor JA-20 A 10.000 rpm durante 10 min a 4°C, con el fin de eliminar los restos celulares.

VII. Recuperar el sobrenadante, añadir unas gotas de cloroformo y guardar a 4°C.

Cuando se requiera una gran cantidad de lisado líquido (apartado 6.3.2.), preparar un cultivo de 400 ml de LB + Mg²⁺ de la cepa hospedadora hasta llegar a una OD₆₅₀ = 0.3-0.4 (o bien OD₅₄₀ = 0.15-0.2) y añadir 5 ml del lisado obtenido en la etapa v. Incubar entonces hasta que aparezca la lisis (unas 8 h) y añadir entonces 1 ml de cloroformo.

5.7. Transfección con DNA del virus M13

Para preparar las células competentes, se ha seguido el mismo método que describimos en el apartado 5.3., salvo que se usaron directamente las células de la etapa vii, sin añadir glicerol.

El protocolo de transfección que hemos utilizado es:

I. Al iniciar la preparación de células competentes, inocular igualmente un tubo con 10 ml de medio LB, anadiendo una gota de un cultivo de noche.

II. Preparar las células competentes (apartado 5.3, hasta la etapa vii).

III. Seguir las etapas ii a v del protocolo de transformación con DNA plasmidico (apartado 5.4.).

IV. Fundir medio de agar blando y mantenerlo en un baño a 42°C. Distribuir 3 ml de este medio en un tubo estéril y mantenerlo en el baño.

V. Añadir a este tubo 40 µl de solución de IPTG y 40 µl de solución de X-gal (su preparación se describe en el apartado 5.1).

VI. Añadir al tubo 300 µl del cultivo de la etapa i (debe estar en fase exponencial)

VII. Inmediatamente, añadir las células transformadas, mezclar evitando formar burbujas y verter sobre una placa previamente precalentada a 37°C.

VIII. Dejar solidificar e incubar a 37°C durante 12-16 h, hasta que aparezcan los halos de inhibición.

6. AISLAMIENTO DE DNA

6.1. Aislamiento de DNA genómico de Drosophila

Para la detección de secuencias en los genomas de *D. melanogaster* y de *D. subobscura*, el protocolo que hemos seguido es una adaptación del método recomendado por JOWETT (1986). Nuestra experiencia es que este método es válido tanto para la extracción de DNA a gran escala, como para la extracción de el DNA de una sola mosca. Las cantidades recomendadas pueden, por tanto variarse proporcionalmente a la cantidad de material de partida. El DNA que se obtiene está contaminado con RNA, el cual se puede eliminar añadiendo RNasa al tampón final.

I. Homogenizar unas 100 moscas adultas en 300 µl de tampón de homogenización (10 mM Tris-HCl, pH 7.5; 60 mM NaCl; 10 mM EDTA; 0.15 mM espermina; 0.15 mM espermidina; 5% sacarosa).

II. Añadir 300 µl del siguiente tampón: 1.25% SDS; 0.3 M Tris-HCl, pH 9.0; 0.1 M EDTA; 5% sacarosa; 200 µg/ml proteinasa K.

III. Incubar durante 30-40 min a 60-65°C, con el fin de liberar el DNA.

IV. Enfriar en hielo y añadir 90 µl de 8 M K.acetato (preparado sin ajustar el pH).

V. Incubar en hielo durante 45-60 min y centrifugar en la centrifuga de sobremesa durante 1-2 min.

VI. Recuperar el sobrenadante, evitando tomar la capa de lípidos de la superficie.

VII. Extraer con fenol, fenol-cloroformo y cloroformo.

VIII. Añadir 2 volúmenes de etanol a temperatura ambiente y centrifugar los ácidos nucleicos durante 5 min en la centrifuga de sobremesa.

IX. Lavar con etanol al 80% y secar en vacío.

X. Disolver en 60 µl de tampón TE que contenga 100 µg/ml de RNasa.

XI. Añadir 140 µl de etanol para precipitar nuevamente. Centrifugar y lavar como se describe en las etapas viii y ix.

XII. Disolver en 200 µl de tampón TE.

6.2. Aislamiento de DNA plasmídico

6.2.1. Minipreparaciones de plásmidos

Las preparaciones de DNA plasmídico a pequeña escala han sido realizadas siguiendo el método de la lisis alcalina, descrito en MANIATIS y otros. La principal novedad de nuestro protocolo consiste en que la adición de cada disolución va seguida de una agitación en vórtex. Hemos suprimido todas aquellas indicaciones que se refieren a mezclar suavemente, etc. Esta simple modificación aumenta sensiblemente el rendimiento del proceso.

I. Inocular la colonia que interese en 5 ml de medio LB con ampicilina y cultivar durante 12-16 h a 37°C.

II. En un tubo Eppendorf centrifugar 1.5 ml del cultivo en la centrifuga de sobremesa durante 3 min a temperatura ambiente.

III. Eliminar el sobrenadante y centrifugar de nuevo 1.5 ml de cultivo sobre el mismo tubo.

IV. Eliminar el sobrenadante y resuspender en 100 µl de tampón de lisis (50 mM glucosa; 25 mM Tris-HCl, pH 8.0; 10 mM EDTA; 5 mg/ml lisozima). El tampón sin lisozima puede almacenarse indefinidamente a 4°C; el tampón con lisozima no debe guardarse.

V. Mezclar fuertemente con el vórtex hasta resuspender totalmente el sedimento. Incubar en hielo durante 5 min.

VI. Añadir 200 µl de solución alcalina (0.2 N NaOH; 1% SDS) y mezclar con el vórtex. Incubar en hielo durante 5 min.

VII. Añadir 150 µl de 5 M K.acetato (pH 4.8) y mezclar con el vórtex (se formará un precipitado). Incubar en hielo durante 5 min.

VIII. Precipitar los restos celulares en la centrifuga de sobremesa durante 5 min a temperatura ambiente.

IX. Recuperar el sobrenadante, evitando cualquier contaminación con la capa de lípidos de la superficie.

X. Extraer con fenol, fenol-cloroformo y cloroformo.

XI. Añadir 1.1 volúmenes de isopropanol, mezclar por inversión y dejar a -20°C durante, al menos, 10 min.

XII. Precipitar los ácidos nucleicos en la centrifuga de sobremesa durante 5 min.

XIII. Eliminar el sobrenadante y lavar con 1 ml de etanol al 80%.

XIV. Centrifugar durante 2 min, eliminar el sobrenadante y secar al vacío.

XV. Resuspender el precipitado en 50 μl de tampón TE que contiene 20 $\mu\text{g/ml}$ de RNasa A y guardar a 4°C .

El rendimiento de este proceso es de más de 7 μg de DNA plasmídico por ml de cultivo bacteriano, suficiente para llevar a cabo al menos 25 análisis de restricción. El DNA obtenido de esta forma no presenta contaminación apreciable de nucleasas después de 1 año mantenido a 4°C . Su calidad es suficiente para utilizarlo como vector, aunque hemos preferido utilizar únicamente vectores purificados siguiendo el método del punto 6.2.2.

Todos los moldes para la secuenciación de plásmidos han sido preparados siguiendo este método a pequeña escala, sin ninguna purificación adicional.

6.2.2. Aislamiento de plásmidos a gran escala

A pesar de que el DNA obtenido con el procedimiento anterior es de una gran calidad y puede utilizarse como vector de clonación, en este trabajo hemos preferido preparar los vectores únicamente a partir de DNA extraído mediante ultracentrifugación en gradiente de CsCl .

El procedimiento seguido está basado en el protocolo de la lisis alcalina recomendado por MANIATIS, con modificaciones.

6.3. Aislamiento de DNA del fago lambda

6.3.1. Preparación a partir de cultivos en placa

Para obtener DNA por este método, se ha seguido exactamente el protocolo descrito en MANIATIS y otros.

6.3.2. Preparación a gran escala

Para obtener gran cantidad de DNA, hemos seguido un protocolo modificado de MANIATIS y otros. La principal modificación es la densidad de las disoluciones de CsCl utilizadas y la omisión de una diálisis al final de la purificación de DNA. El protocolo seguido es el siguiente:

I. Partiendo de 400 ml de lisado líquido, distribuirlo en dos botellas y centrifugar en el rotor JA-14 (o bien el GS3) a 8.000 rpm durante 10 min.

II. Recuperar el sobrenadante y, si todavía quedan restos celulares, efectuar una segunda centrifugación.

III. Añadir 4 mg de RNasa cruda y 4 mg de DNasa I (para el total de 400 ml de lisado). Incubar durante 1 h a 37°C.

IV. Añadir 4 g de NaCl y 16 g de polietilenglicol (PEG) por cada 200 ml y mantener durante 12-16 h a 4°C (concentraciones finales: 0.34 M NaCl; 8% PEG).

V. Centrifugar en el mismo tipo de rotor a 8.000 rpm durante 15 min a 4°C.

VI. Eliminar el sobrenadante y dejar las botellas invertidas en posición vertical sobre papel absorbente.

VII. Disolver los sedimentos en un total de 5 ml de tampón de fagos (10 mM Tris-HCl, pH 8.0; 20 mM NaCl; 10 mM MgCl₂). Dejar durante una hora a 4°C.

VIII. Centrifugar en el mismo rotor a 7.000 rpm durante 2 min, con el fin de eliminar el PEG.

IX. Recuperar el sobrenadante.

X. Preparar tres soluciones de CsCl en tampón de fagos, con las siguientes densidades: 1.7 g/ml (92.5 g de CsCl en 100 ml de tampón); 1.5 g/ml (67.5%); y 1.3 g/ml (38.6%).

XI. Preparar un gradiente escalonado de CsCl en dos tubos ULTRA-CLEAR de BECKMAN (Ref. 344060, para el rotor SW 40 Ti; Ref. 344059 para el rotor SW 41 Ti), de la siguiente forma: depositar en el tubo 1.5 ml de la solución de densidad 1.7; después añadir lentamente 2 ml de la solución de densidad 1.5, procurando que se forme una interfase; por último depositar 3 ml de la solución de densidad 1.3, procurando también mantener la interfase. Marcar con rotulador la interfase entre las densidades 1.5 y 1.7.

XII. Depositar el sobrenadante de la etapa ix. sobre el gradiente, procurando no destruir el gradiente.

XIII. Centrifugar en el rotor SW 40 Ti o en el SW 41 Ti) a 30.000 rpm durante 1 h a 4°C.

XIV. Utilizando una jeringuilla y una aguja de calibre grueso, pinchar la banda situada entre la fase de densidad 1.3 y la de 1.5. La banda tiene un aspecto verdeazulado. En caso de dificultades para ver la banda iluminar los tubos desde arriba y mirar sobre un fondo oscuro. Si aparece una banda blanca gruesa es porque el PEG no ha sido eliminado suficientemente durante la fase viii. En este caso, el gradiente ha quedado colapsado y no aparecerá la banda del fago.

XV. Poner el volumen recuperado de la banda en un tubo nuevo y añadir 5 ml de la solución de densidad 1.5. Rellenar el tubo con la solución de densidad 1.3, hasta que falten 7-8 mm para llegar al borde.

XVI. Centrifugar en el mismo rotor a 30.000 rpm durante 2 h a 4°C.

XVII. Recuperar la banda de la forma descrita en la etapa xiv y dializar frente a 1000 volúmenes de tampón de fagos durante 1 h a temperatura ambiente.

- XVIII. Cambiar el tampón y dializar de nuevo durante 1 h.
- XIX. Transferir la suspensión de fagos a un tubo y añadir 0.5 M EDTA para obtener una concentración final de 20 mM.
- XX. Añadir proteinasa K en polvo (o de un stock recién preparado) para obtener una concentración final de 50 µg/ml y añadir SDS al 10% para obtener una concentración del 0.5%.
- XXI. Mezclar e incubar durante 1 h a 65°C.
- XXII. Añadir un volumen igual de fenol y mezclar durante 10 min. Separar las fases por centrifugación y recuperar la fase acuosa.
- XXIII. Añadir un volumen de fenol-cloroformo, mezclar brevemente y separar las fases por centrifugación.
- XXIV. Repetir la etapa anterior, pero esta vez con cloroformo.
- XXV. Transferir a un tubo y añadir 1 volumen de isopropanol. Se formará una "medusa" de precipitado, que puede recogerse con una punta desechable de micropipeta.
- XXVI. Lavar con etanol al 80% , secar al vacío y disolver en tampón TE.

6.3.3. Minipreparaciones de lisados líquidos

En la última parte de este trabajo, hemos probado (con éxito) el protocolo recomendado por AUSUBEL y otros (1988) para pequeños volúmenes de lisados, con ligeras modificaciones. La principal modificación es la omisión de la etapa de ultracentrifugación a 30.000 rpm durante 90 min (en un rotor SW-27) que, en nuestro protocolo, hemos sustituido por el uso de una centrífuga de alta velocidad. El protocolo es el siguiente:

- I. Partiendo de 50 ml de lisado líquido, añadir 10 µl de DNasa (5 mg/ml) y 25 µl de RNasa (10 mg/ml). Incubar a 37°C durante 1 h.
- II. Sedimentar los fagos mediante centrifugación a 20.000 rpm durante 135 min en el rotor JA-20 (tubos Corex).
- III. Retirar el sobrenadante e invertir los tubos sobre papel absorbente, para eliminar el líquido al máximo. Resuspender los fagos en 50 mM Tris-HCl, pH 8.0.
- IV. Pasar esta solución a un tubo Eppendorf y añadir 200 µl de fenol. Dejar en agitación durante 15-20 min. Centrifugar durante 2 min y recuperar la fase acuosa.
- V. Repetir la extracción con fenol.
- VI. Añadir 200 µl de cloroformo y mezclar por agitación. Centrifugar brevemente y recuperar la fase acuosa.
- VII. Repetir la extracción con cloroformo.
- VIII. Añadir 20 µl de acetato sódico 3 M y precipitar el DNA con 2 volúmenes de etanol al 96% a temperatura ambiente.
- IX. Centrifugar durante 10 min y eliminar el sobrenadante.
- X. Lavar el precipitado con etanol al 80%, centrifugar 5 min y eliminar el sobrenadante.
- XI. Secar en vacío y resuspender el DNA en 100 µl de tampón TE. En caso de observar precipitados de distinto tamaño (en

diferentes preparaciones), resuspender en volúmenes de TE proporcionales al tamaño de los mismos.

Con este protocolo, hemos obtenido cantidades de DNA del orden de 50-100 μ l. La calidad de este DNA es suficiente para el análisis de restricción y para la subclonación de fragmentos en vectores plasmídicos. Si al digerir con enzimas de restricción se obtienen bandas difusas, esto se debe a la presencia de nucleasas contaminantes en el DNA (por una mala extracción con fenol). Si esto ocurre, se puede solucionar efectuando una nueva extracción con fenol (aunque las bandas aparezcan con degradación, probablemente el DNA no estará degradado, ya que la degradación se produce durante la incubación a 37°C para el análisis de restricción).

6.4. Aislamiento de ssDNA del fago M13

Con el fin de obtener moldes de simple cadena para los primeros experimentos de secuenciación, hemos seguido el método descrito en el Manual de Instrucciones "M13 Cloning/Dideoxy Sequencing" del proveedor GIBCO-BRL.

I. Inocular una colonia de la cepa JM105, u otra apropiada para el cultivo de M13 en 5 ml de medio LB y cultivar durante 12-16 h a 37°C.

II. Inocular 2 ml de medio LB con 2 gotas del cultivo anterior, junto con una placa de lisis. Incubar durante 5-6 h a 37°C con fuerte agitación. Evitar períodos más largos de incubación, pues se puede producir la lisis celular y el ssDNA obtenido estaría contaminado con DNA cromosómico.

III. Transferir 1.5 ml del cultivo a un tubo Eppendorf y centrifugar 5 min a temperatura ambiente.

IV. Transferir 1.2 ml del sobrenadante a un tubo nuevo, evitando tocar el sedimento celular. El sedimento puede conservarse, pues contiene el DNA del virus en su forma replicativa (RF). Para preparar el RF DNA, basta con procesar este precipitado del mismo modo que en la preparación de DNA plasmídico a pequeña escala. El sedimento también sirve como stock del clon.

V. Añadir al sobrenadante 300 μ l de una solución de precipitación (20% PEG; 2.5 M NaCl). Mezclar brevemente en el vórtex y dejar 15 min a temperatura ambiente.

VI. Sedimentar los virus por centrifugación durante 10 min en la centrífuga de sobremesa.

VII. Tirar el sobrenadante y volver a centrifugar durante 2 min.

VIII. Eliminar al máximo el sobrenadante, con la ayuda de una pipeta Pasteur con la punta estirada a la llama.

IX. Resuspender el precipitado de virus con 200 μ l de tampón TE que contiene RNasa (50 μ g/ml). Extraer con un volumen igual de fenol. Agitar violentamente con la ayuda del vórtex.

X. Transcurridos 10 min, agitar nuevamente en el vórtex y separar las fases por centrifugación durante 2-3 min.

- XI. Recuperar la fase acuosa y extraer con cloroformo.
- XII. Precipitar añadiendo 20 μ l de 3 M Na.acetato (pH 5.0) y 500 μ l de etanol al 96%. Mantener durante 12-16 h a -20°C o bien durante al menos 15 min a -70°C .
- XIII. Centrifugar en frío durante 10 min. Lavar con 1 ml de etanol al 80% frío y volver a centrifugar.
- XIV. Secar el precipitado al vacío y resuspenderlo en 20 μ l de tampón TE. Guardar a -20°C .

El rendimiento de este proceso es de 1-2 μ g de DNA por preparación. Para una reacción de secuenciación, suele ser suficiente con 5 μ l de este DNA.

7. DETERMINACION DE LA CONCENTRACION

Aunque en algunos casos hemos determinado la concentración midiendo la A_{260} , en la mayoría de los casos lo hemos hecho comparando la intensidad de fluorescencia de la muestra en un minigel con las diferentes intensidades de los fragmentos de un marcador de peso molecular.

El marcador utilizado es el lambda Hind III de Boehringer Mannheim (B.M.) (Cat.No. 236 250). Los tamaños de los fragmentos de DNA que contiene se enumeran en el apartado 13.1.

8. SONDAS UTILIZADAS

8.1 Derivadas de D.melanogaster

cDm903:

En los primeros rastreos (screenings) de la librería genómica, hemos utilizado como sonda el plásmido cDm903, que contiene un inserto de cDNA clonado en el sitio EcoRI del vector pBR322. El inserto, de 2.2 kb, contiene la combinación de exones 1-2-4-5-6-7-8, y corresponde a un RNA derivado del promotor distal (P1) (SCHNEUWLY y otros, 1986). Esta sonda nos ha sido facilitada por el Prof. Walter J. Gehring, de la Universidad de Basilea (Suiza).

Este inserto fue purificado y mantenido a -20°C hasta el momento de marcarlo con un isótopo radioactivo.

pDmA.2015:

Con el fin de identificar la región que contiene la secuencia homeobox, hemos preparado sondas del plásmido pDmA.2015, que es un clon de DNA genómico que incluye parte del exón 3' terminal, concretamente un fragmento EcoRI de 1.7 kb (donde se encuentra la secuencia homeobox) insertado en el sitio EcoRI del vector pUC 8. Este clon es un derivado del clon homónimo descrito en LAUGHON y otros (1986). Esta sonda nos ha sido facilitada por Sonsoles Campuzano, del Centro de

8.2 Derivadas de *D. subobscura*

Para las etapas de paseo cromosómico (chromosomal walk), hemos utilizado como sondas fragmentos de restricción subclonados en el vector pUC 18. El uso de cada uno de estos clones se describe en el apartado RESULTADOS.

Con el fin de determinar el grado de solapamiento entre dos clones dados del fago lambda, hemos utilizado una sonda que incluye todo el DNA de uno de los clones y ésta se ha hibridado frente a un filtro que contiene DNA del otro clon, digerido con varios enzimas de restricción.

9. MARCAJE DE DNA

9.1. Nick-translation

Para marcar sondas mediante este sistema, hemos seguido un protocolo basado en MANIATIS y otros. El tampón utilizado (NT) es el que recomienda este autor.

O. Tampones de partida.

Tampón NT (10x): 0.5 M Tris-HCl, pH 7.2; 0.1 M MgSO₄; 1 mM DTT; 500 µg/ml BSA.

Mezcla CGT: 2 mM dCTP; 2 mM dGTP; 2 mM dTTP (diluídos a partir de las disoluciones stock, preparadas según MANIATIS y otros).

I. Mezclar en un tubo Eppendorf: 2 µl de tampón NT; 1 µl de mezcla CGT; x µl de DNA (mínimo 50 ng); 1 µl de DNA pol I (1-3 U, según concentración suministrada); 1 µl de solución diluida de DNasa I (0.1 µg/ml, ver apartado 1.4) y agua hasta 17 µl.

II. Añadir 3 µl de [³²P]dATP reciente. Si el nucleótido radiactivo tiene más de dos semanas, aumentar la cantidad añadida y modificar la cantidad de agua.

III. Incubar durante 1-2 h a 15-16°C.

IV. Parar la reacción con 5 µl de 0.5 M EDTA.

V. Preparar una columna de Sephadex G-100 en el interior de una pipeta Pasteur. Determinar el volumen vacío (normalmente entre 450 y 500 µl) con una solución de azul dextrano y equilibrar la columna con tampón TE.

VI. Aplicar los 25 µl de muestra y dejar este volumen en la columna.

VII. Sin dejar que el gel se seque, aplicar lentamente un volumen de TE igual al 90% del volumen vacío. El avance del pico de DNA marcado puede seguirse con un monitor de radiactividad de tipo Geiger.

VIII. Añadir 450 µl de TE y recoger todo el volumen eluido de la columna.

Si el marcaje ha funcionado, debe observarse claramente un

pico de radiactividad que debe salir con el volumen vacío. Se puede obtener una buena estimación del porcentaje de radiactividad incorporado en el DNA, midiendo con el monitor las cpm de la muestra antes y después de pasarla por la columna (colocando el detector del monitor a una misma distancia, en ambos casos).

El porcentaje de incorporación obtenido con este método ha sido, normalmente, del orden del 40-50%. La actividad específica resultante ha sido del orden de 0.5×10^8 a 2×10^8 cpm/ μ g.

9.2. Oligolabelling

Este método permite obtener actividades específicas más elevadas que con el sistema anterior, aunque la principal ventaja es, a nuestro juicio, la mayor facilidad de control de este sistema, pues sólo es preciso controlar la actividad de un enzima.

Para la aplicación de este protocolo, hemos seguido exactamente el protocolo sugerido por FEINBERG y VOLGESTEIN (1983), excepto que la mezcla de nucleótidos a sido adaptada para el uso de [32 P]dATP marcado, en vez de [32 P]dCTP.

Después de incubaciones durante 12-16 h, la actividad específica del DNA marcado con este sistema ha oscilado entre 0.5×10^8 cpm/ μ g.

La eficiencia de incorporación ha oscilado normalmente entre el 60 y el 90% y la actividad específica obtenida ha sido de 2×10^8 a 6×10^8 cpm/ μ g.

9.3. Marcaje de extremos 5'

Cuando no se precisa una actividad específica muy elevada, es posible utilizar sondas marcadas por este método:

I. Partiendo de 0.5-1 μ g de DNA lineal defosforilado con fosfatasa alcalina intestinal, aumentar el volumen hasta 200 μ l.

II. Extraer con fenol, fenol-cloroformo y cloroformo.

III. Añadir 0.1 volúmenes de 3 M Na.acetato y 2.5 volúmenes de etanol al 96%.

IV. Incubar a -70°C durante, al menos, 20 min y centrifugar en la centrifuga de sobremesa durante 15 min.

V. Lavar en etanol al 80% y secar al vacío.

VI. Disolver en 18.5 μ l de agua. Añadir 2.5 μ l de tampón (10x) de kinasa (apartado 1.4) y 1 μ l de polinucleótido kinasa (unas 4 U).

VII. Añadir 3 μ l de [γ - 32 P]ATP e incubar durante 45 min a 37°C .

VIII. Separar del DNA marcado del nucleótido radioactivo tal como se describe en las etapas v a viii de la sección 9.1.

9.4. Marcaje no radioactivo

En una segunda fase de nuestro trabajo, hemos prescindido del uso de isótopos radioactivos, aprovechando la comercialización de un kit de marcaje y detección de DNA que incorpora la digoxigenina (un esteroide) como molécula marcadora del DNA (B.M. Cat.No. 1093 657). La detección de la señal se efectúa utilizando un anticuerpo conjugado al enzima fosfatasa alcalina y efectuando una reacción de color.

El método de marcaje es el de oligolabelling, descrito en el apartado 9.2. y el protocolo específico sigue las recomendaciones del proveedor:

I. Desnaturalizar entre 100 ng y 2 µg de DNA lineal, calentando a 100°C durante 10 min.

II. Enfriar rápidamente en una mezcla de etanol y hielo seco.

III. Añadir 2 µl de mezcla de hexanucleótidos (vial 5), 2 µl de mezcla de nucleótidos (1 mM dATP; 1 mM dCTP; 1 mM dGTP; 0.65 mM dTTP; 0.35 mM dUTP-digoxigenina; pH 6.5) y agua hasta 19 µl.

IV. Añadir 1 µl de fragmento Klenow (Labelling Grade, vial 7) e incubar a 37°C entre 60 min y 16 h.

V. Parar la reacción añadiendo 1 µl de 0.5 M EDTA y precipitar el DNA marcado añadiendo 2 µl de 4 M LiCl y 60 µl de etanol enfriado a -20°C. Dejar a -20°C durante 2 h o a -70°C durante 30 min.

VI. Centrifugar durante 10 min en la centrifuga de sobremesa. Eliminar el sobrenadante y secar en vacío.

VII. Disolver en 50 µl de TE.

10. TRANSFERENCIA A FILTROS

En la mayoría de los experimentos de hibridación, hemos utilizado filtros de nylon del proveedor Amersham (Refs. RPN.3050N y RPN.87N).

Hemos comprobado que la solución de neutralización recomendada por estos autores (0.5 M Tris-HCl, pH 8.0; 1.5 M NaCl) tiene el mismo resultado que la recomendada por la hoja de instrucciones del proveedor Pall (3M Na.acetato, pH 5.5). En este trabajo nos hemos inclinado por esta segunda alternativa.

A continuación describimos los protocolos específicos que hemos seguido.

10.1. Réplica de placas de lisis

Durante los screening de la genoteca, así como durante la purificación de los clones positivos, el protocolo de transferencia a filtros que hemos seguido ha sido una modificación del recomendado por KAISER y MURRAY (1985). En todos los casos, hemos utilizado placas de cultivo de 90 mm de

diámetro y membranas de 87 mm.

I. Una vez que las placas de lisis alcanzan un tamaño de alrededor de 0.2 mm, interrumpir la incubación y pasarlas a 4°C, con el fin de endurecer la capa de agar blando.

II. Marcar con lápiz un extremo del filtro de nylon con una clave que se corresponda con la placa a replicar. Colocar con cuidado la membrana (no es necesario humedecerla previamente) sobre la superficie del agar.

III. Con la ayuda de una aguja estéril, marcar 4 orificios asimétricos sobre la membrana, atravesando también el medio de cultivo. Esperar entre 30 s y 1 min.

IV. Si se precisan filtros duplicados, colocar un nuevo filtro sobre la placa (sin olvidar marcarlo antes con el lápiz), marcar los agujeros siguiendo el patrón establecido sobre el medio de cultivo y esperar 1-2 min.

En caso contrario, pasar a la etapa siguiente.

V. Con la ayuda de unas pinzas estériles, separar el filtro de la placa y colocarlo (con la superficie de contacto con la placa boca arriba) sobre una hoja doble de papel Whatman 3MM empapada en solución desnaturalizante (0.5 M NaOH; 1.5 M NaCl).

VI. Esperar 5 min y transferir el filtro (en la misma posición) a otra hoja de papel de filtro empapada en solución de neutralización (3 M Na.acetato, pH 5.5).

VII. Esperar 5 min y enjuagar el filtro brevemente en 2 x SSC.

VIII. Dejar secar el filtro sobre papel de filtro durante 30-60 min a temperatura ambiente (o bien durante 10 min a 80°C).

IX. Irradiar el filtro durante 3-5 min en el transiluminador UV (o bien durante 1 h a 80°C).

X. No es necesario seguir este protocolo en la cámara estéril. Con el fin de evitar el crecimiento de cualquier bacteria contaminante, al final de hacer todas las réplicas en filtro, colocar las placas en una campana de vidrio junto con un recipiente con 20-30 ml de cloroformo y dejarlas así durante 12-16 h. Tras este tratamiento, envolver las placas con plástico y mantenerlas a 4°C.

10.2. Réplica de colonias bacterianas

En algunos casos, para la selección de subclones determinados, ha sido preciso hibridar réplicas de colonias bacterianas, utilizando sondas específicas, con el fin de seleccionar los clones de interés.

El método seguido ha sido idéntico al anterior, salvo que, tras el contacto entre la membrana y la placa de cultivo, se sometió al filtro a un tratamiento con SDS al 10% (3 min sobre papel 3 MM empapado).

10.3. Transferencia de Southern

Con el fin de detectar secuencias específicas en el DNA genómico de *Drosophila* y en el DNA de los clones recombinantes, hemos seguido el protocolo de transferencia desde geles de agarosa a filtros, diseñado por Southern y descrito en MANIATIS y otros.

I. Después de la electroforesis (en condiciones normales, incluyendo bromuro de etidio en el gel y en el tampón), tomar una fotografía de referencia en el transiluminador.

II. Si el tamaño de los fragmentos de DNA a transferir es superior a 20 kb, sumergir el gel en una cubeta con 0.25 N HCl durante 15 min y mantener en agitación. En caso contrario, pasar a la etapa iii.

III. Lavar el gel con agua y transferir a una cubeta con solución de desnaturalización (idéntica a la de los apartados anteriores). Mantener durante 30-60 min en agitación.

IV. Lavar el gel con agua y sumergirlo en una cubeta con solución de neutralización (también idéntica a la de los apartados anteriores) y mantener durante 30-60 min en agitación.

V. Montar el sistema de transferencia por capilaridad de la forma descrita en MANIATIS y otros, y transferir el DNA a un filtro de nylon utilizando 20 x SSC como tampón de transferencia (tiempo de transferencia: 12-24 h).

VI. Después de la transferencia, procesar la membrana de la misma forma que en los apartados anteriores (etapa ix del apartado 10.1). La membrana con el DNA fijado se puede mantener a temperatura ambiente envuelta en papel de aluminio.

Para la transferencia bidireccional se ha seguido este mismo proceso, con el esquema descrito en MANIATIS y otros.

10.4. Dot blot

Cuando se ha precisado una información rápida para la simple determinación de presencia/ausencia de una secuencia en un clon concreto, hemos seguido el siguiente protocolo de dot blot:

I. Efectuar varias diluciones de los DNAs problema, hasta llegar a concentraciones entre 10 pg/ μ l y 0.1 pg/ μ l. Desnaturalizar en un baño a 100°C durante 10 min.

II. Enfriar las muestras en hielo y depositar entre 1 y 5 μ l de cada una sobre una membrana de nylon (sin humidificación previa), en la que previamente se ha señalado la posición de cada muestra.

III. Dejar secar y fijar el DNA mediante tratamiento con UV o bien calentando a 80°C durante 1 h.

IV. Prehibridar e hibridar siguiendo los protocolos específicos.

Cuando las muestras son de ssDNA del virus M13, no es

precisa la etapa i, y puede procederse a depositar las muestras como se describe en la etapa ii.

Hemos observado que los filtros de nylon proporcionan señales mucho más intensas que los de nitrocelulosa debido, al parecer, a que el DNA difunde muy poco durante el momento en que se deposita la gota sobre el filtro.

11. HIBRIDACIONES

Todas las hibridaciones han sido llevadas a cabo en el interior de tubos de vidrio, y la temperatura se ha mantenido en un horno giratorio del proveedor BACHOFER.

Las soluciones de hibridación utilizadas con las sondas marcadas radioactivamente son:

Mezcla con formamida: 5x SSC; 0.1% BSA; 0.1% Ficoll-400; 0.1% polivinilpirrolidona; 50 mM tampón fosfato, pH 7.0; 0.5% SDS; 250 µg/ml DNA de salmón sonificado y desnaturalizado; 46% formamida desionizada. Esta mezcla sólo ha sido utilizada para hibridaciones de estringencia moderada.

Mezcla sin formamida: contiene los mismos componentes en las mismas concentraciones excepto la formamida (que está ausente) y el SSC (que en esta mezcla se añadió en una concentración de 6x).

Para las hibridaciones con sondas marcadas con digoxigenina, se empleó la mezcla recomendada en la hoja de instrucciones del proveedor, incrementando la concentración de agente de bloqueo hasta el 1%, para su uso en filtros de nylon.

Mezcla para sondas no radioactivas: 5x SSC; 1% agente de bloqueo (componente 11 del kit); 0.1% N-laurilsarcosina (sal sódica); 0.02% SDS.

El protocolo seguido para las hibridaciones es:

I. Incubar los filtros, con el DNA ya fijado, en solución de hibridación (sin sonda) durante, al menos, 1 h a una temperatura apropiada:

Mezcla con formamida: 42°C.

Mezcla sin formamida:

Hibridaciones de estringencia moderada: 65°C.

Hibridaciones de alta estringencia: 68°C.

II. Añadir la mezcla de hibridación que contiene la sonda desnaturalizada e incubar a la misma temperatura entre 6 y 24 h, dependiendo de la complejidad del DNA que se pretende analizar.

III. Recuperar la sonda y mantenerla a -20°C.

IV. Lavar los filtros dos veces a temperatura ambiente en un volumen grande de 2x SSC; 0.1% SDS (5-10 min cada lavado, en agitación).

V. Efectuar los lavados de estringencia (15 min cada lavado, en agitación), a la temperatura apropiada:

Hibridaciones de estringencia moderada:

Temperatura: 50°C

Tampón: 2x SSC; 0.1% SDS.

Hibridaciones de alta estringencia:

Temperatura: 68°C.

Tampón: 0.1% SSC; 0.1% SDS.

VI. Dejar secar el filtro y proceder a la detección de la señal (autorradiografía en el caso de las sondas radioactivas).

Todas las autorradiografías de sondas marcadas con ^{32}P se han realizado con película radiográfica sometida a una breve preexposición (preflashing) en luz roja. Aunque en algunos manuales se afirma que este tratamiento no aumenta la sensibilidad para la radiación del ^{32}P , hemos comprobado que no es así.

Para la detección de la señal de las sondas marcadas con digoxigenina hemos seguido exactamente las condiciones recomendadas por el fabricante del kit (B.M.). Normalmente, la señal de los filtros que contienen DNA de un clon del fago lambda, digerido con un enzima de restricción, aparece en 1-5 min. La señal de el DNA contenido en las placas de lisis aparece entre 3 y 12 h. La señal de secuencias en el DNA genómico de *Drosophila* aparece entre 12 y 24 h de reacción enzimática.

En caso de volver a utilizar por segunda vez una misma mezcla de hibridación, siempre se ha procedido a desnaturalizar la sonda durante 10 min a 100°C (y después a enfriarla en hielo).

12. MANIPULACION DEL DNA

12.1. Preparación de vectores plasmidicos

Estos vectores han sido preparados partiendo de plásmidos purificados mediante ultracentrifugación en gradiente de CsCl. Unos 5 µg de DNA fueron linearizados con el enzima de interés, en todos los casos algún enzima de restricción de los que cortan en el sitio de clonación múltiple (multiple cloning site) del vector. Durante los primeros experimentos, los vectores fueron defosforilados posteriormente a la digestión, siguiendo esencialmente el protocolo recomendado por MANIATIS y otros. Posteriormente, hemos encontrado varios manuales que recomiendan defosforilar el vector al mismo tiempo que se realiza la digestión con el enzima de restricción. El protocolo definitivo que hemos elaborado es una combinación de varias fuentes.

I. Digerir entre 1 y 10 µg de DNA plasmidico (o bien de la forma replicativa del virus M13), purificado en CsCl, con el enzima de restricción apropiado, incluyendo en la mezcla de digestión 1 µl de fosfatasa alcalina de intestino de ternera (CIP, 1 U/µl).

II. Observar aproximadamente 0.1 µg de el DNA digerido en un

minigel de agarosa y comprobar si la digestión es completa.

III. Cuando no se observe ninguna traza de DNA circular, proceder a inactivar el enzima de restricción y la fosfatasa calentando a 68°C durante 15-20 min.

IV. Extraer con fenol, fenol-cloroformo y cloroformo.

V. Precipitar con etanol, secar en vacío y resuspender en TE, con una concentración de 0.1-0.5 µg/µl.

VI. Efectuar una ligación de 0.5 µg de vector y transformar, con el fin de determinar el nivel de ruido de fondo.

En nuestra experiencia, un vector del tipo del pUC 18 preparado con este protocolo produce, después de ligarlo en ausencia de inserto, menos de 200 colonias (todas azules) por µg de DNA (aplicando la prueba de la etapa vi).

12.2. Clonación y selección de recombinantes

Los diversos experimentos de clonación llevados a cabo se describen en el apartado RESULTADOS. En la medida de lo posible (plásmidos y M13) se ha preferido el uso de medio indicador IPTG/Xgal para la selección de los clones recombinantes por su color blanco (en las colonias plasmídicas o en los halos de inhibición de M13).

La identidad de los clones recombinantes ha sido determinada mediante hibridaciones de el DNA obtenido frente a sondas específicas y mediante análisis de sus secuencias.

13. ELECTROFORESIS

13.1. Geles de agarosa

La mayoría de las electroforesis se han llevado a cabo en geles de agarosa, dentro de un rango del 0.8 al 1.2% de concentración. Para la elección de la concentración, se ha tenido en cuenta el rango de separación descrito en MANIATIS y otros. Todos los geles se han preparado incluyendo bromuro de etidio en una concentración de 0.5 µg/ml; el tampón de migración ha incluido también este colorante en la misma concentración. Estas mismas condiciones se han utilizado para los geles que luego han sido transferidos por el procedimiento de SOUTHERN.

El tampón de electroforesis ha sido el Tris-borato (TBE), con una ligera variación recomendada por el Manual de secuenciación del proveedor BRL. Su composición es, para un stock 10x: disolver 121.1 g de Tris base, 55 g de ácido bórico, 7.4 g Na₂.EDTA en un volumen final de 1 l. El pH final es 8.3.

El marcador de pesos moleculares (y de concentración) que hemos empleado es una digestión del DNA del fago lambda con el enzima de restricción HindIII, que produce fragmentos de los siguientes tamaños (en pb): 23130, 9416, 6682, 4361, 2322,

2027 y 564.

El tampón utilizado para cargar las muestras ha sido el tipo II de MANIATIS, por la comodidad de poder mantenerlo a 4°C. Su composición es (tampón 10x): 0.25% azul bromofenol, 0.25% xilen-cianol, 25% Ficoll-400.

13.2. Geles de poliacrilamida

Los geles de acrilamida para la secuenciación se han preparado con gradiente de grosor de 0.2 a 0.4 mm, utilizando las placas de vidrio y el sistema de electroforesis MacroPhor del fabricante PHARMACIA-LKB.

Para la preparación de los geles, hemos seguido un protocolo que permite almacenar una mezcla (INSTAGEL) con todos los componentes del gel excepto el TEMED y el persulfato amónico:

I. Preparar un stock de acrilamida (19:1) al 40% disolviendo 380 g de acrilamida (B.M.; Cat.No. 100 137) y 20 g de bisacrilamida (SERVA; Ref. 29196) en un volumen final de 1 l. Filtrar en papel Whatmann 3MM y mantener a 4°C en la oscuridad (se mantiene indefinidamente).

II. Preparar una mezcla INSTAGEL al 8% mezclando 100 ml del stock de acrilamida al 40%, 250 g de urea (B.M.; Cat.No. 100 164), 50 ml de tampón TBE 10x y agua hasta 500 ml. Disolver en agitación, evitando calentar la mezcla (para compensar el enfriamiento producido por la disolución de la urea, es posible aplicar calor, pero evitando que la mezcla llegue a calentarse). Mantener a temperatura ambiente en la oscuridad (se mantiene durante 1 mes).

III. Preparar una disolución de persulfato amónico (PSA) al 10% y mantener a 4°C (estable durante 1 semana).

IV. En un matraz Kitasatos, desgasificar en vacío 75 ml de Instagel durante 5-10 min.

V. Añadir 750 µl de solución al 10% de PSA y 15 µl de TEMED.

VI. Verter el gel sobre el sistema de placas de vidrio que han sido tratadas de la forma recomendada por el fabricante (Pharmacia-LKB). Dejar polimerizar durante, al menos 1 h.

VII. Efectuar una preelectroforesis durante 30-60 min a 1000 V.

14. PURIFICACION DE FRAGMENTOS DE DNA

Para purificar fragmentos, hemos seguido un protocolo simplificado basado en MANIATIS y otros. La principal ventaja de esta alternativa es que la electroelución se lleva a cabo en la misma cubeta que la electroforesis, así como en el mismo tampón de electroforesis: no encontramos necesario llenar la cubeta de electroforesis con 0.5x TBE.

Alternativa A.

I. Escindir del gel el fragmento de interés utilizando una hoja de bisturí estéril. Si el DNA es para subclonación, minimizar la exposición del fragmento a la radiación UV.

II. Rellenar un tubo de diálisis (cerrado por un extremo) con tampón TBE 0.5x e introducir el fragmento de agarosa que contiene el fragmento de DNA. Sacar la mayor cantidad posible de tampón y cerrar el otro extremo de la bolsa.

III. Introducir en una cubeta de electroforesis que contiene TBE 1x y aplicar 100 voltios durante 30 min.

IV. Invertir la polaridad durante 30 s y comprobar en el transiluminador UV que el DNA ha salido de la agarosa.

V. Recuperar el tampón de la bolsa y lavar con algo más de tampón.

VI. Centrifugar durante 10 min en la microcentrífuga, con el fin de sedimentar los posibles restos de agarosa.

VII. Recuperar el sobrenadante, evitando tocar el precipitado de agarosa.

VIII. Extraer con fenol, fenol-cloroformo y cloroformo.

IX. A la fase acuosa final, añadir 2 µl de una solución de tRNA (10 mg/ml) y añadir 2 volúmenes de etanol.

X. Dejar precipitar durante, al menos, 1 h a -70°C. Centrifugar, secar a vacío y resuspender en el tampón adecuado.

Alternativa B.

Este protocolo se basa también en la electroelución del fragmento y la diferencia radica en el método de purificación del DNA.

I. LLenar una jeringuilla de 5 ml con 2 ml de tampón HIGH (1 NaCl; 20 mM Tris-HCl, pH 7.5; 1 mM EDTA).

II. Cortar la punta de una columna Elutip-D (Schleicher Schüll, Ref. NA 010/2) justo por debajo de la matriz y unirla a la jeringuilla. Forzar el tampón a través de la columna.

III. LLenar otra jeringuilla con 5 ml de tampón LOW (0.2 M NaCl; 20 mM Tris-HCl, pH 7.5; 1 mM EDTA).

IV. Separar la columna de la primera jeringuilla y unirla a la segunda. Hacer pasar el tampón.

V. Separar la columna de la segunda jeringuilla y cargar esta jeringuilla con la muestra de DNA (obtenida en las etapas I a V del apartado anterior). El DNA debe ajustarse a una concentración salina de 0.2 M NaCl.

VI. Unir un filtro desechable de acetato de celulosa (Schleicher Schüll, Ref. FP 030/20) a la jeringuilla ya cargada y unir la columna al extremo del soporte del filtro. Hacer pasar lentamente la muestra de DNA a través del filtro y de la columna.

VII. Separar la columna, cargar la misma jeringuilla con 3 ml de tampón LOW y volver a conectar la columna al soporte del filtro. Hacer pasar este tampón por el filtro.

VIII. Separar la columna. Cargar la primera jeringuilla con 0.4 ml de tampón HIGH y unirla directamente a la columna.

IX. Forzar la elución del DNA y precipitarlo con etanol si es preciso.

15. SINTESIS DE OLIGONUCLEOTIDOS

Para completar parte de la secuencia que hemos determinado, hemos utilizado dos oligonucleótidos sintéticos. La síntesis fué realizada por el Dr. Enrique Herrero y por el Dr. Julio Polaina, utilizando un sintetizador Gene Assembler Plus de Pharmacia-LKB y siguiendo las instrucciones del fabricante.

La secuencia de los oligonucleótidos sintetizados son, en dirección 5'-3', TTCTGGCTGTAGTGATGCA (oligo1) y TGGCGTGATGCCACAGGCACA (oligo2). El oligo1 se une a las posiciones 317-335 (ambas incluidas), de tal forma que la secuencia que hemos determinado con su uso es la de la cadena complementaria a la que presentamos en las figuras del capítulo Resultados. El oligo2 corresponde a las posiciones 477-497 de la secuencia que hemos obtenido (ver la secuencia proporcionada en el capítulo de Resultados; fig. 10), de tal forma que la secuencia determinada corresponde a la cadena codificante del gen Antp.

Su diseño se llevó a cabo partiendo de la secuencia parcial de *D. subobscura*, previamente determinada, evitando incluir repeticiones de la secuencia CAG (o su complementaria CTG), dado que se encuentra en numerosas posiciones a lo largo del gen *Antennapedia*.

Para su uso como primer en las reacciones de secuenciación, hemos utilizado los oligonucleótidos sin ninguna purificación de tamaños. La cantidad utilizada por reacción es de 20 ng.

16. SECUENCIACION

El método de secuenciación que hemos seguido es el de SANGER (1977). El isótopo utilizado en todos los casos ha sido el análogo [³²S]dATP (desoxiadenosina 5'-α[³²S]tiotriofosfato) de Amersham (Cat.No. SJ.1304; actividad específica superior a 1000 Ci/mmol). Hemos preferido el uso de este isótopo por la mayor facilidad de su manejo, vida media más larga, y por la buena resolución que proporciona durante la autorradiografía.

El primer que hemos utilizado para la secuenciación de los clones de M13 es el primer universal de 17 pb, cuya secuencia es GTAAAACGACGGCCAGT (B.M., Cat.No. 1010 077). Para la secuenciación de los clones plasmídicos hemos empleado tanto el primer universal de 17 pb como el primer reverso, también de 17 pb, cuya secuencia es CAGGAAACAGCTATGAC (B.M. Cat.No. 1010 093). Además, en dos ocasiones, hemos empleado oligonucleótidos de síntesis (ver apartado anterior).

16.1. Secuenciación de moldes de cadena simple

Para secuenciar los clones de M13 hemos seguido el siguiente protocolo:

I. Partiendo de las soluciones stock 10 mM de desoxirribonucleótidos (preparadas según se describe en MANIATIS y otros), preparar unas diluciones 1:20 en agua (concentración final de cada stock: 0.5 mM).

II. Con estas soluciones diluidas, preparar las siguientes mezclas de nucleótidos, que son estables a -20°C:

	A'	C'	G'	T'
0.5 mM dCTP	20 µl	1 µl	20 µl	20 µl
0.5 mM dGTP	20 µl	20 µl	1 µl	20 µl
0.5 mM dTTP	20 µl	20 µl	20 µl	1 µl
10x tampón de reacción	20 µl	20 µl	20 µl	20 µl

Tampón de reacción (10x): 100 mM Tris-HCl, pH 8.0; 100 mM MgCl₂; 300 mM NaCl.

III. Partiendo de los stock 10 mM de didesoxirribonucleótidos, preparar las siguientes diluciones: 0.5 mM ddATP; 0.5 mM ddCTP; 0.5 mM ddGTP; y 1 mM ddTTP.

IV. Preparar las mezclas de reacción de la siguiente forma:

Mezcla A: 20 µl A' + 20 µl 0.5 mM ddATP
Mezcla C: 20 µl C' + 20 µl 0.5 mM ddCTP
Mezcla G: 20 µl G' + 20 µl 0.5 mM ddGTP
Mezcla T: 20 µl T' + 20 µl 0.5 mM ddTTP

V. En un tubo Eppendorf, mezclar 0.5-1 µg de ssDNA del clon de M13; 10 ng del primer universal; 1 µl de tampón de reacción (10x) y agua hasta completar un volumen final de 12.4 µl.

VI. Calentar en un baño a 55-60°C durante 10 min y dejar enfriar lentamente (20-30 min). El DNA con el primer unido puede mantenerse a -20°C.

VII. Marcar 4 tubos con las letras A, C, G y T y poner en cada uno de ellos 2 µl de la mezcla correspondiente.

VIII. Al tubo con el DNA unido al primer, añadir 1 µl de fragmento Klenow (1-2 U); 1 µl 0.1 M DTT; y 3 µl [³⁵S]dATP. Mezclar bien y distribuir 3 µl de esta mezcla a cada uno de los cuatro tubos de la etapa anterior.

IX. Incubar a 50°C durante 5 min.

X. Añadir 1 µl de 0.5 mM dATP a cada tubo e incubar durante otros 5 min a 50°C.

XI. Interrumpir las reacciones añadiendo 5 µl de tampón de parada. En esta fase, las muestras pueden guardarse a -20°C.

XII. Para cargar el gel, desnaturalizar la muestra en un baño a 100°C durante 3-5 min. Durante el tiempo que transcurra

hasta el momento de cargar, mantener sobre hielo.
XIII. Cargar 3 μ l en cada carrera del gel de secuenciación. El resto de la muestra puede guardarse a -20°C .

Tampón de parada: mezclar 100 ml de formamida desionizada, 0.1 g de xilene cianol, 0.1 g de azul bromofenol, y 4 ml de 0.5 M EDTA. Mantener a 4°C .

16.2. Secuenciación de moldes de doble cadena

La mayoría de las secuencias analizadas corresponden a segmentos clonados en plásmidos. El método seguido es el recomendado por la hoja de instrucciones de B.M., con las siguientes modificaciones:

- (1) Hemos añadido 3 μ l de [^{35}S]dATP, en vez de 2 μ l;
- (2) las reacciones se han llevado a cabo a 50°C durante 5 min, en vez de a 37°C durante 20 min; y
- (3) hemos cargado 3 μ l de muestra, en vez de 1-2.

En las últimas partes de este trabajo, hemos llevado a cabo las reacciones de secuenciación utilizando la DNA polimerasa del fago T7 (Pharmacia; Ref. 27-1682-01), que aporta numerosas ventajas frente al fragmento Klenow. Estos experimentos se han realizado siguiendo el protocolo de la hoja de instrucciones, con la siguiente modificación:

(*) Para la desnaturalización, secar al vacío 2-3 μ g de DNA y resuspenderlo en 10 μ l de 0.2 N NaOH.

17. ANALISIS DE SECUENCIAS

Parte de las secuencias que hemos obtenido ha sido analizadas con el software del Genetics Computer Group (Universidad de Wisconsin), instalado en un ordenador de tipo VAX. El resto de los análisis, incluyendo la elaboración de la estrategia de clonación y las búsquedas en las bases de datos, se han llevado a cabo con los programas DNASTAR, instalados en un ordenador personal Olivetti M380, XP3.

REFERENCIAS

ABBOTT, M. K. and 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.

AGUADE, M (1988).

Nucleotide Sequence Comparison of the rp49 Gene Region between *Drosophila subobscura* and *D. melanogaster*.
Mol. Biol. Evol. 5(4): 433-441.

AKAM, M (1984).

A common segment in genes for segments in *Drosophila*.
Nature 308: 402-403.

AKAM, M (1989).

Hox and HOM: Homologous Gene Cluster in Insects and Vertebrates.
Cell 57: 347-349.

AKAM, M. E (1983).

The location of Ultrabithorax transcripts in *Drosophila* tissue sections.
EMBO J. 2: 2075-2084.

AKAM, M. E. and A. MARTINEZ-ARIAS (1985).

The distribution of Ultrabithorax transcripts in *Drosophila* embryos.
EMBO J. 4: 1689-1700.

AKAM, M. E., A. MARTINEZ-ARIAS, R. WEINZIERL and C. D. WILDE (1985).

Function and expression of Ultrabithorax in the *Drosophila* embryo.
Cold Spring Harbor Symp. Quant. Biol. 50: 195-200.

ANDERSON, K. V (1987).

Dorsal-ventral embryonic pattern genes of *Drosophila*.
Trends in Genetics 3: 91-97.

ANDERSON, K. V. and C. NÜSSLEIN-VOLHARD (1984).

Information for the dorsal-ventral pattern of the *Drosophila* embryo is stored as maternal mRNA.
Nature 311: 223-227.

ATKINSON, P. W., L. E. MILLS, W. T. STARMER and D. T. SULLIVAN (1988).

Structure and Evolution of the Adh Genes of *Drosophila mojavensis*.
Genetics 120: 713-723.

AUSUBEL, F. M., R. BRENT, R. E. KINGSTON, D. D. MOORE, J. G. SEIDMAN, J. A. SMITH and K. STRUHL (1988).

Current protocols in Molecular Biology.
Wiley-Interscience

AWGULEWITSCH, A., M. F. UTSET, C. P. HART, W. McGINNIS AND F. H. RUDDLE (1986).

Spatial restriction in expression of a mouse homeobox locus within the central nervous system.
Nature 320: 328-335.

BAKER, N. E (1988).

Localization of transcripts from the wingless gene in whole *Drosophila* embryos.
Development 103: 289-298.

BALCELLS, L., J. MODOLELL and M. RUIZ-GOMEZ (1988).

A unitary basis for different Hairy-wing mutations of *Drosophila melanogaster*.
EMBO J. 7: 3899-3906.

BANERJEE, U., P. J. RENFRANZ, D. R. HINTON, B. A. RABIN and S. BENZER (1987).

The sevenless⁺ protein is expressed apically in cell membranes of developing *Drosophila* retina; it is not restricted to cell R7.
Cell 51: 151-158.

BANERJEE, U., P. J. RENFRANZ, J. A. POLLOCK and S. BENZER (1987).

Molecular characterization and expression of sevenless, a gene involved in neuronal pattern formation in the *Drosophila* eye.
Cell 49: 281-291.

BARAD, M., T. JACK, R. CHADWICK and W. McGINNIS (1988).

A novel, tissue-specific, *Drosophila* homeobox gene.
EMBO J. 7: 2151-2161.

BARTON, S. C., M. A. H. SURANI and M. L. NORRIS (1984).

Role of paternal and maternal genomes in mouse development.
Nature 311: 374.

BEACHY, P. A., M. A. KRASNOW, E. R. GAVIS and D. S. HOGNESS (1988).

An Ultrabithorax protein binds sequences near its own and the Antennapedia P1 promoters.
Cell 55: 1069-1081

BEACHY, P. B., S. L. HELFAND and D. S. HOGNESS (1985).

Segmental distribution of bithorax complex proteins during *Drosophila* development.
Nature 313: 545-551.

BENDER, W., B. WEIFFENBACH, F. KARCH and M. PEIFER (1985).

Domains of cis-interaction in the *Drosophila* Bithorax Complex.
Cold Spring Harbor Symp. Quant. Biol. 50: 173-180.

BENDER, W., M. AKAM, F. KARCH, P. A. BEACHY, M. PFEIFER, P. SPIERER, E. B. LEWIS and D. S. HOGNESS (1983).

Molecular genetics of the Bithorax Complex in *Drosophila melanogaster*.
Science 221: 23-29.

BENSON, M. and V. PIRROTTA (1987).

The product of the *Drosophila* zeste gene binds to specific DNA sequences in white and Ubx.
EMBO J. 6: 1387-1392.

BENSON, M. and V. PIRROTTA (1988).

The *Drosophila* zeste protein binds cooperatively to sites in many gene regulatory regions: implications for transvection and gene regulation.
EMBO J. 7: 3907-3915.

BERLETH, T., M. BURRI, G. THOMA, D. BOPP, S. RICHSTEIN, G. FRIGERIO, M. NOLL and C. NÜSSLEIN-VOLHARD (1988).

The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo.
EMBO J. 7: 1749-1756.

BERMINGHAM, J. R. and M. P. SCOTT (1988).

Developmentally regulated alternative splicing of transcripts from the *Drosophila* homeotic gene Antennapedia can produce four different proteins.
EMBO J. 7: 3211-3222.

BIENZ, M. and G. TREMML (1988).

Domain of Ultrabithorax expression in *Drosophila* visceral mesoderm from autoregulation and exclusion.
Nature 333: 576-578.

BIENZ, M., G. SAARI, G. TREMML, J. MÜLLER, B. ZÜST and P. A. LAWRENCE (1988).

Differential Regulation of Ultrabithorax in Two Germ Layers of *Drosophila*.
Cell 53: 567-576.

BIGGIN, M. D. and R. TJIAN (1988).

Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts.
Cell 53: 699-711.

BIGGIN, M. D. and R. TJIAN (1989).

A Purified *Drosophila* Homeodomain Protein Represses Transcription In Vitro.
Cell 58: 433-440.

- BIGGIN, M. D., S. BICKEL, M. BENSON, V. PIRROTA and R. TJIAN (1988).
Zeste encodes a sequence-specific transcription factor that activates the Ultrabithorax promoter in vitro.
Cell 53: 713-722.
- BODNER, M., J. L. CASTRILLO, L. E. THEIL, T. DEERINCK, M. ELLISMAN and M. KARIN (1988).
The Pituitary-Specific Transcription Factor GHF-1 Is a Homeobox-Containing Protein.
Cell 55: 505-518.
- BONNER, J. T. (Ed.) (1982).
Evolution and Development
Life Sciences Research Report 22 (Dahle Konferenzen). Springer-Verlag.
- BOPP, D., M. BURRI, S. BAUMGARTNER, G. FRIGERIO and M. NOLL (1986).
Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila.
Cell 47: 1033-1040.
- BOSSY, B., L. M. C. HALL and P. SPIERER (1984).
Genetic activity along 315 kb of the Drosophila chromosome.
EMBO J. 3: 2537-2541.
- BROWER, D. J. (1984).
Posterior-to-anterior transformation in engrailed wing imaginal disks of Drosophila.
Nature 310: 496.
- BROWN, S. D. M. and A. J. GREENFIELD (1985).
Possible functions of the homeobox.
Nature 313: 185.
- BURGLIN, T. R. (1988).
The yeast regulatory gene PHO2 encodes a homeobox.
Cell 53: 339-340.
- BURGLIN, T. R., C. V. E. WRIGHT and E. M. DE ROBERTIS (1987).
Translational control in homeobox mRNAs?
Nature 330: 701-701.
- CABRERA, C. V., A. MARTINEZ-ARIAS and M. BATE (1987).
The expression of three members of the achaete-scute gene complex correlates with neuroblast segregation in Drosophila.
Cell 50: 425-433.
- CARRERA, C. V., J. BOTAS and A. GARCIA-BELLIDO (1985).
Distribution of Ultrabithorax proteins in mutants of Drosophila bithorax complex and its transregulatory genes.
Nature 318: 569-571.
- CABRERA, C. V., M. C. ALONSO, P. JOHNSTON, R. G. PHILLIPS and P. A. LAWRENCE (1987).
Phenocopies induced with antisense RNA identify the wingless gene.
Cell 50: 659-663.
- CAMPUZANO, S., L. BALCELLS, R. VILLARES, L. CARRAMOLINO, L. GARCIA-ALONSO and J. MODOLELL (1986).
Excess function Hairy-wing mutations caused by gypsy and copia insertions within structural genes of the achaete-scute locus of Drosophila.
Cell 44: 303-312.
- CARROL, S. B. and M. P. SCOTT (1986).
Zygotically active genes that affect the spatial expression of the fushi tarazu segmentation gene during early Drosophila embryogenesis.
Cell 45: 113-126.
- CARROL, S. B., G. M. WINSLOW, T. SCHUPBACH and M. P. SCOTT (1986).
Maternal control of Drosophila segmentation gene expression.
Nature 323: 278-280.

CARROLL, S. B. and M. P. SCOTT (1985).

Localization of the fushi tarazu protein during *Drosophila* embryogenesis.
Cell 43: 47-57.

CARROLL, S. B., R. A. LAYMON, M. A. McCUTCHEON, P. D. RILEY and M. P. SCOTT (1986).

The localization and regulation of Antennapedia protein expression in *Drosophila* embryos.
Cell 47: 113-122.

CASANOVA, J., E. SANCHEZ-HERRERO and G. MORATA (1985).

Prothoracic transformation and functional structure of the Ultrabithorax gene of *Drosophila*.
Cell 42: 663-669.

CASANOVA, J., E. SANCHEZ-HERRERO and G. MORATA (1986).

Identification and characterization of a parasegment specific regulatory element of the Abdominal-B gene of *Drosophila*.
Cell 47: 627-636.

CASTRILLO, J. L., M. BODNER and M. KARIN (1989).

Purification of Growth Hormone-Specific Transcription Factor GHF-1 Containing Homeobox.
Science 243: 814-817.

CELNIKER, S. E. and E. B. LEWIS (1987).

Transabdominal, a dominant mutant of the Bithorax Complex, produces a sexually dimorphic segmental transformation in *Drosophila*.
Genes & Development 1: 111-123.

CHADWICK, R. and W. McGINNIS (1987).

Temporal and spatial distribution of transcripts from the Deformed gene of *Drosophila*.
EMBO J. 6: 779-789.

CHEN, C., T. MALONE, S. K. BECKENDORF and R. L. DAVIS (1987).

At least two genes reside within a large intron of the dunce gene of *Drosophila*.
Nature 329: 721-724.

CHISHOLM, R. L. (1986).

Homeoboxes: what do they do?
Trends in Genetics 2: 224-225.

CHO, K. W. Y., J. GOETZ, C. V. E. WRIGHT, A. FRITZ, J. HARDWICKE and E. M. DE ROBERTIS (1988).

Differential utilization of the same reading frame in a *Xenopus* homeobox gene encodes two related proteins sharing the same DNA-binding specificity.
EMBO J. 7: 2139-2149.

CHOWDHURY, K., U. DEUSCH and P. GRUSS (1987).

A multigene family encoding several "finger" structures is present and differentially active in mammalian genomes.
Cell 48: 771-778.

CITRI, Y., H. V. COLOT, A. C. JACQUIER, Q. YU, J. C. HALL, D. BALTIMORE and M. ROSBASH (1987).

A family of unusually spliced biologically active transcripts is encoded by a *Drosophila* clock gene.
Nature 326: 42.

COLOT, H. V., J. C. HALL and M. ROSBASH (1988).

Interspecific comparison of the period gene of *Drosophila* reveals large blocks of non-conserved coding DNA.
EMBO J. 7: 3929-3937.

DE ROBERTIS, E. M., A. FRITZ, J. GOETZ, G. MARTIN, I. W. MATAJ, E. SALO, G. D. SMITH, C. WRIGHT and R. ZELLER (1985).

The *Xenopus* homeoboxes.
Cold Spring Harbor Symp. Quant. Biol. 50: 271-275.

DEAROLF, C. R., J. TOPOL and C. S. PARKER (1989).

The caudal gene product is a direct activator of fushi tarazu transcription during *Drosophila* embryogenesis.
Nature 341: 340-343.

- DELOTTO, R. and P. SPIERER (1986).
A gene required for the specification of dorsal-ventral pattern in *Drosophila* appears to encode a serine protease.
Nature 323: 688-692.
- DENELL, R. E., K. R. HUMMELS, B. T. WAKIMOTO and T. C. KAUFMAN (1981).
Developmental studies of lethality associated with the Antennapedia gene complex in *Drosophila melanogaster*.
Dev. Biol. 81: 43-50.
- DESPLAN, C., J. THEIS and P. H. O'FARRELL (1985).
The *Drosophila* developmental gene, engrailed, encodes a sequence-specific DNA binding activity.
Nature 318: 630-635.
- DESPLAN, C., J. THEIS and P. H. O'FARRELL (1988).
The sequence specificity of homeodomain-DNA interaction.
Cell 54: 1081-1090.
- 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 and Development 3: 399-414.
- DINARDO, S., E. SHER, J. HEEMSKERK-JONGENS, J. KASSIS and P. H. O'FARRELL (1988).
Two-tiered regulation of spatially patterned engrailed gene expression during *Drosophila* embryogenesis.
Nature 332: 604-609.
- DOE, C. Q., D. SMOUSE and C. S. GOODMAN (1988).
Control of neuronal fate by the *Drosophila* segmentation gene even-skipped.
Nature 333: 376-378.
- DOVER, G. A. (1989).
Slips, strings and species.
TIG 5: 100-102.
- DOYLE, H. J., K. HARDING, T. HOEY and M. LEVINE (1986).
Transcripts encoded by a homeobox gene are restricted to dorsal tissues of *Drosophila* embryos.
Nature 323: 76-79.
- DRIEVER, W. and C. NÜSSLEIN-VOLHARD (1988).
The bicoid Protein Determines Position in the *Drosophila* Embryo in a Concentration Dependent Manner.
Cell 54: 95-104.
- DRIEVER, W. and C. NÜSSLEIN-VOLHARD (1988).
A Gradient of bicoid Protein in *Drosophila* Embryos.
Cell 54: 83-93.
- DRIEVER, W. and C. NÜSSLEIN-VOLHARD (1989).
The bicoid protein is a positive regulator of hunchback transcription in the early *Drosophila* embryo.
Nature 337: 138-143.
- DRIEVER, W., G. THOMA and C. NÜSSLEIN-VOLHARD (1989).
Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the bicoid morphogen.
Nature 340: 363-367.
- DUNCAN, I (1986).
Control of Bithorax Complex functions by the segmentation gene fushi tarazu of *D. melanogaster*.
Cell 47: 297-309
- DUNCAN, I. W. and T. C. KAUFMAN (1975).
Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: mapping of the proximal portion of the right arm.
Genetics 80: 733-752.

DURA, J. M., N. B. RANDSHOLT, J. DEATRICK, I. ERK, P. SANTAMARIA, J. D. FREEMAN, S. J. FREEMAN, D. WEDDELL and H. W. BROCK (1987).
A Complex Genetic Locus, polyhomeotic, Is Required for Segmental Specification and Epidermal Development in *D. melanogaster*.
Cell 51: 829-839.

EDGAR, B. E., M. P. WEIR, G. SCHUBIGER and T. KORNBERG (1986).
Repression and turnover pattern of fushi tarazu RNA in the early *Drosophila* development.
Cell 47: 747-754.

FJOSE, A., W. J. MCGINNIS and W. J. GEHRING (1985).
Isolation of a homeobox containing gene from the engrailed region of *Drosophila* and the spatial distribution of its transcripts.
Nature 313: 284-289.

FRASCH, M. and M. LEVINE (1987).
Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in *Drosophila*.
Genes & Development 1: 981-995.

FREI, E., D. BOPP, M. BURRI, S. BAUMGARTNER, J.-E. EDSTRÖM and M. NOLL (1985).
Isolation and Structural Analysis of the extra sex combs Gene of *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 127-134.

FRENCH, V. (1988).
Gradients and insect segmentation.
Development 104 (Supp): 3-16.

FRIGERIO, G., M. BURRI, D. BOPP, S. BAUMGARTNER and M. NOLL (1986).
Structure of the segmentation gene paired and the *Drosophila* PRD gene set as part of a gene network.
Cell 47: 735-746.

FRISCHER, L. E., L. S. HAGEN and R. L. GARBER (1986).
An inversion that disrupts the Antennapedia gene causes abnormal structure and localization of RNAs.
Cell 37: 1017-1023.

FRÖHNHÖFER, H. G. and C. NÜSSLEIN-VOLHARD (1986).
Organization of anterior pattern in the *Drosophila* embryo by the maternal gene bicoid.
Nature 324: 120-125.

FRÖHNHÖFER, H. G. and C. NÜSSLEIN-VOLHARD (1987).
Maternal genes required for the anterior localization of bicoid activity in the embryo of *Drosophila*.
Genes and Development 1: 880-890.

GARBER, R. L., A. KUROIWA and W. J. GEHRING (1983).
Genomic and cDNA clones of the homeotic locus Antennapedia in *Drosophila*.
EMBO J. 2: 2027-2036.

GARCIA-BELLIDO, A., P. A. LAWRENCE y G. MORATA (1979).
Compartimentos en el desarrollo de los animales.
Investigación y Ciencia 36: 66-75.

GAUGER, A., R. G. FEHON and G. SCHUBIGER (1985).
Preferential binding of imaginal disk cells to embryonic segments of *Drosophila*.
Nature 313: 395.

GAUL, U. and H. JÄCKLE (1987).
Pole region-dependent repression of the *Drosophila* gap gene Krüppel by maternal gene products.
Cell 51: 549-555.

GAUL, U. and H. JÄCKLE (1987).
How to fill a gap in the *Drosophila* embryo.
TIG 3: 5 127-131.

- GAUL, U., E. SEIFERT, R. SCHUH and H. JÄCKLE (1987).
Analysis of Krüppel protein distribution during early Drosophila development reveals posttranscriptional regulation.
Cell 50: 639-647.
- GAY, N. J., S. FOOLE and T. KORNBERG (1988).
Association of the Drosophila melanogaster engrailed protein with specific soluble nuclear protein complexes.
EMBO J. 7: 4291-4297.
- GEHRING, W. J (1985).
Homeotic Genes, the Homeobox, and the Genetic Control of Development.
Cold Spring Harbor Symp. Quant. Biol. 50: 243-251.
- GEHRING, W. J (1985).
The Homeobox: A Key to the Understanding of Development?
Cell 40: 3-5.
- GEHRING, W. J (1987).
Homeoboxes in the study of Development.
Science 236: 1245-1252.
- GEHRING, W. J (1987).
The Homeobox: Structural and Evolutionary Aspects.
Molecular Approaches to Developmental Biology, 115-129.
- GEHRING, W. J (1987).
Homeotic Genes, The Homeobox, and the Spatial Organization of the Embryo.
The Harvey Lectures, 81: 153-172.
- GEHRING, W. J. and Y. HIROMI (1986).
Homeotic genes and the homeobox.
Ann. Rev. Genet. 20: 147-173.
- GERGEN, J. P. and E. WIESCHAUS (1986).
Dosage Requirements for runt in the Segmentation of Drosophila Embryos.
Cell 45: 289-299.
- GRAHAM, A., N. PAPALOPULU and ROBB KRUMLAUF (1989).
The Murine and Drosophila Homeobox Gene Complexes Have Common Features of Organization and Expression.
Cell 57: 367-378.
- HAFEN, E., A. KUROIWA and W. J. GEHRING (1984).
Spatial Distribution of Transcripts from the Segmentation Gene fushi tarazu during Drosophila Embryonic Development.
Cell 37: 833-841.
- HAFEN, E., M. LEVINE and W. J. GEHRING (1984).
Regulation of Antennapedia transcript distribution by the bithorax complex in Drosophila.
Nature 307: 287-289.
- HAFEN, E., M. LEVINE, R. L. GARBER and W. J. GEHRING (1983).
An improved in situ hybridization method for the detection of cellular RNAs in Drosophila tissue sections and its application for localizing transcription of the homeotic Antennapedia gene complex.
EMBO J. 2: 617-623.
- HARDING, K., C. RUSHLOW, H. J. DOYLE, T. HOEY and M. LEVINE (1986).
Cross-Regulatory Interactions Among Pair-Rule Genes in Drosophila.
Science 233: 953-959.
- HARDING, K., C. WEDEEN, W. MCGINNIS and M. LEVINE (1985).
Spatially Regulated Expression of Homeotic Genes in Drosophila.
Science 229: 1236-1242.

- HARDING, K., T. HOEY, R. WARRIOR and M. LEVINE (1989).
Autoregulatory and gap gene response elements of the even-skipped promoter of *Drosophila*.
EMBO J. 8: 1205-1212.
- HASHIMOTO, C., K. L. HUDSON and K. V. ANDERSON (1988).
The Toll Gene of *Drosophila*, Required for Dorsal-Ventral Embryonic Polarity, Appears to Encode a Transmembrane Protein.
Cell 52: 269-279.
- HAZELRIGG, T. and T. C. KAUFMAN (1983).
Revertants of dominant mutations associated with the Antennapedia gene complex of *Drosophila melanogaster*: cytology and genetics.
Genetics 105: 581-600.
- HEBERLEIN, U. and R. TJIAN (1988).
Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts.
Nature 331: 410-415.
- HENIKOFF, S., M. A. KEENE, K. FECHTEL and J. W. FRISTROM (1986).
Gene within a Gene: Nested *Drosophila* Genes Encode Unrelated Proteins on Opposite DNA Strands.
Cell 44: 33-42.
- HIROMI, Y. and W. J. GEHRING (1987).
Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*.
Cell 50: 963-974.
- HIROMI, Y., A. KUROIWA and W. J. GEHRING (1985).
Control Elements of the *Drosophila* Segmentation Gene *fushi tarazu*.
Cell 43: 603-613.
- HOEY, T. and M. LEVINE (1988).
Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*.
Nature 332: 858-861.
- HOEY, T., H. J. DOYLE, K. HARDING, C. WEDEEN and M. LEVINE (1986).
Homeobox gene expression in anterior and posterior regions of the *Drosophila* embryo.
Proc. Natl. Acad. Sci. USA 83: 4809-4813.
- HOGAN, B (1985).
Homeoboxes and strings for the packaging of genes?
Nature 314: 670-671.
- HOGNESS, D. S., H. D. LIPSHITZ, P. A. BEACHY, D. A. PEATTIE, R. B. SAINT, M. GOLDSCHMIDT-CLERMONT, P. J. HARTE, E. R. GAVIS, and S. L. HE LFAND (1985).
Regulation and Products of the Ubx Domain of the Bithorax Complex.
Cold Spring Harbor Symp. Quant. Biol. 50: 181-194.
- HOLLAND, P. W. H. and B. L. M. HOGAN (1986).
Phylogenetic distribution of Antennapedia-like homeoboxes.
Nature 321: 251.
- HOWARD, K. and P. INGHAM (1986).
Regulatory Interactions between the Segmentation Genes *fushi tarazu*, *hairy*, and *engrailed* in the *Drosophila* Blastoderm.
Cell 44: 949-957.
- INGHAM, P (1985).
Abdominal gene organization.
Nature 313: 98.
- INGHAM, P. W (1985).
Genetic Control of the Spatial Pattern of Selector Gene Expression in *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 201-208.

- INGHAM, P. W (1985).
Patterns of differentiation.
Nature 317: 202-203.
- INGHAM, P. W (1988).
The molecular genetics of embryonic pattern formation in *Drosophila*.
Nature 335: 25-34.
- INGHAM, P. W. and A. MARTINEZ-ARIAS (1986).
The correct activation of Antennapedia and bithorax complex genes requires the fushi tarazu gene.
Nature 324: 592-597.
- INGHAM, P. W., K. R. HOWARD and D. ISH-HOROWICZ (1985).
Transcription pattern of the *Drosophila* segmentation gene hairy.
Nature 318: 439.
- INGHAM, P. W., N. E. BAKER and A. MARTINEZ-ARIAS (1988).
Regulation of segment polarity genes in the *Drosophila* blastoderm by fushi tarazu and even skipped.
Nature 331: 73-75.
- INGHAM, P., A. MARTINEZ-ARIAS, P. A. LAWRENCE and K. HOWARD (1985).
Expression of engrailed in the parasegment of *Drosophila*.
Nature 317: 634-636.
- IRISH, V. F., A. MARTINEZ-ARIAS and M. AKAM (1989).
Spatial regulation of the Antennapedia and Ultrabithorax homeotic genes during *Drosophila* early development.
EMBO J. 8: 1527-1537.
- ISH-HOROWICZ, D. and S. M. PINCHIN (1987).
Pattern Abnormalities Induced by Ectopic Expression of the *Drosophila* Gene hairy Are Associated with Repression of ftz Transcription.
Cell 51: 405-415.
- ISH-HOROWICZ, D., K. R. HOWARD, S. M. PINCHIN and P. W. INGHAM (1985).
Molecular and Genetic Analysis of the hairy Locus in *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 135-144.
- JACK, T., M. REGULSKI and W. MCGINNIS (1988).
Pair-rule segmentation genes regulate the expression of the homeotic selector gene, Deformed.
Genes and Development 2: 635-651.
- JÄCKLE, H., D. TAUTZ, R. SCHUH, E. SEIFERT and R. LEHMANN (1986).
Cross-regulatory interactions among the gap genes of *Drosophila* genetics.
Nature 324: 668-670.
- JÄCKLE, H., U. B. ROSENBERG, A. PREISS, E. SEIFERT, D. C. KNIPPLE, A. KIENLIN, and R. LEHMANN (1985).
Molecular Analysis of Krüppel, a Segmentation Gene of *Drosophila melanogaster*.
Cold Spring Harbor 50: 465-473.
- JAYNES, J. B. and P. H. O'FARRELL (1988).
Activation and repression of transcription by homeodomain-containing proteins that bind a common site.
Nature 336: 744-749.
- JORGENSEN, E. M. and R. L. GARBER (1987).
Function and misfunction of the two promoters of the *Drosophila* Antennapedia gene.
Genes & Development 1: 544-555.
- JOWETT, T. (1986).
Preparation of nucleic acids.
Drosophila: a practical approach (Ed. D. B. Roberts). IRL Press: 275-286.

JÜRGENS, G (1985).

A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*.
Nature 316: 153-155.

JÜRGENS, G (1988).

Head and tail development of the *Drosophila* embryo involves spalt, a novel homeotic gene.
EMBO J. 7: 189-196.

KAISER, K. and N. E. MURRAY (1985).

The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries.
DNA cloning: a practical approach, vol. I (Ed. D. M. Glover). IRL Press: 1-47.

KARCH, F., B. WEIFFENBACH, M. PEIFER, W. BENDER, I. DUNCAN, S. CELNIKER, M. CROSBY, and E. B. LEWIS (1985).

The Abdominal Region of the Bithorax Complex.
Cell 43: 81-96.

KASSIS, J. A., M. L. WONG and P. H. O'FARRELL (1985).

Electron Microscopic Heteroduplex Mapping Identifies Regions of the engrailed Locus That Are Conserved between *Drosophila melanogaster* and *Drosophila virilis*.
Molecular and Cellular Biology 5: 3600-3609.

KASSIS, J. A., S. J. POOLE, D. K. WRIGHT and P. H. O'FARRELL (1986).

Sequence conservation in the protein coding and intron regions of the engrailed transcription unit.
EMBO J. 5: 3583-3589.

KAUFFMAN, S. A., R. M. SHYMKO and K. TRABERT (1978).

Control of Sequential Compartment Formation in *Drosophila*.
Science 199: 259-269.

KAUFMAN, T. C., R. LEWIS and B. WAKIMOTO (1980).

Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polytene chromosome interval 84A-B.
Genetics 94: 115-133.

KIDD, S., M. R. KELLEY and M. W. YOUNG (1986).

Sequence of the Notch Locus of *Drosophila melanogaster*: Relationship of the Encoded Protein to Mammalian Clotting and Growth Factors.
Molecular and Cellular Biology 6: 3094-3108.

KILCHHERR, F., S. BAUMGARTNER, D. BOPP, E. FREI and M. NOLL (1986).

Isolation of the paired gene of *Drosophila* and its spatial expression during early embryogenesis.
Nature 321: 493-499.

KINZLER, K. W., J. M. RUPPERT, S. H. BIGNER and B. VOGELSTEIN (1988).

The GLI gene is a member of the Kruppel family of zinc finger proteins.
Nature 332: 371-374.

KNIPPLE, D. C., E. SEIFERT, U. B. ROSENBERG, A. PREISS and H. JÄCKLE (1985).

Spatial and temporal patterns of Krüppel gene expression in early *Drosophila* embryos.
Nature 317: 40-44.

KO, H. S., P. FAST, W. McBRIDE and L. M. STAUDT (1988).

A Human Protein Specific for the Immunoglobulin Octamer DNA Motif Contains a Functional Homeobox Domain.
Cell 55: 135-144.

KRASNOW, M. A., E. E. SAFFMAN, K. KORNFELD and D. S. HOGNESS (1989).

Transcriptional Activation and Repression by Ultrabithorax Proteins in Cultured *Drosophila* Cells.
Cell 57: 1031-1043.

KRAUSE, H. M. and W. J. GEHRING (1989).

Stage-specific phosphorylation of the fushi tarazu protein during *Drosophila* development.
EMBO J. 8: 1197-1204.

KUROIWA, A., E. HAFEN and W. J. GEHRING (1984).

Cloning and Transcriptional Analysis of the Segmentation Gene *fushi tarazu* of *Drosophila*.
Cell 37: 825-831.

KUROIWA, A., U. KLOTER, P. BAUMGARTNER and W. J. GEHRING (1985).

Cloning of the homeotic *Sex combs reduced* gene in *Drosophila* and in situ localization of its transcripts.
EMBO J. 4: 3757-3764.

KUZIORA, M. A. and W. MCGINNIS (1988).

Autoregulation of a *Drosophila* Homeotic Selector Gene.
Cell 55: 477-485.

LASKO, P. F. and M. ASHBURNER (1988).

The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A.
Nature 335: 611-617.

LAUGHON, A. and M. P. SCOTT (1984).

Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins.
Nature 310: 25-31.

LAUGHON, A., A. M. BOULET, J. R. BERMINGHAM, R. A. LAYMON and M. P. SCOTT (1986).

Structure of transcripts from the homeotic *Antennapedia* gene: Two promoters control the major protein-coding region.
Mol. Cell. Biol. 6: 4676-4689.

LAUGHON, A., S. B. CARROLL, F. A. STORFER, P. D. RILEY and M. P. SCOTT (1985).

Common Properties of Proteins Encoded by the *Antennapedia* Complex Genes of *Drosophila melanogaster*.
Cold Spring Harbor Symp. Quant. Biol. 50: 253-262.

LAWRENCE, P. (1983).

A new homeotic gene.
Nature 306: 643.

LAWRENCE, P. A. (1985).

Pattern Formation in the Muscles of *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 113-117.

LAWRENCE, P. A. (1985).

Compartment genes in hand.
Nature 313: 268.

LAWRENCE, P. A. (1987).

Fair-Rule Genes: Do They Paint Stripes or Draw Lines?
Cell 51: 879-880.

LAWRENCE, P. A. (1988).

Background to *bicoid*.
Cell 54: 1-2.

LAWRENCE, P. A., F. R. S., and A. MARTINEZ-ARIAS (1985).

The cell lineage of segments and parasegments in *Drosophila*.
Phil. Trans. R. Soc. Lond. B 312: 83-90.

LAWRENCE, P. A., P. JOHNSTON, P. MACDONALD and G. STRUHL (1987).

Borders of parasegments in *Drosophila* embryos are delimited by the *fushi tarazu* and *even-skipped* genes.
Nature 328: 440-442.

LEHMANN, R. (1988).

Phenotypic comparison between maternal and zygotic genes controlling the segmental pattern of the *Drosophila* embryo.
Development 104 (Supp): 17-27.

- LEHMANN, R. and C. NUSSLEIN-VOLHARD (1986).
Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in *Drosophila*.
Cell 47: 141-152.
- LEHMANN, R. and C. NUSSLEIN-VOLHARD (1987).
Involvement of the nanos gene in the transport of an abdominal signal in the *Drosophila* embryo.
Nature 329: 167-170.
- LEVINE, M. (1988).
Molecular Analysis of Dorsal Ventral Polarity in *Drosophila*.
Cell 52: 785-786.
- LEVINE, M. and T. HOEY (1988).
Homeobox Proteins as Sequence-Specific Transcription Factors.
Cell 55: 537-540.
- LEVINE, M., E. HAFEN, R. L. GARBER and W. J. BEHRING (1983).
Spatial distribution of Antennapedia transcripts during *Drosophila* development.
EMBO J. 2: 2037-2046.
- LEVINE, M., K. HARDING, C. WEDEEN, H. DOYLE, T. HOEY and H. RADOMSKA (1985).
Expression of the Homeobox Gene Family in *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 209-222.
- LEWIS, E. B. (1985).
Regulation of the Genes of the Bithorax Complex in *Drosophila*.
Cold Spring Harbor Symp. Quant. Biol. 50: 155-164.
- LEWIS, R. A., B. T. WAKIMOTO, R. E. DENELL and T. C. KAUFMAN (1980).
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, R. A., T. C. KAUFMAN and R. E. DENELL (1980).
Genetic analysis of the Antennapedia gene complex (ANT-C): mutant screen of proximal 3R, section 84B-D.
DIS 55: 85.
- LEWIS, R. A., T. C. KAUFMAN, R. E. DENELL and P. TALLERICO (1980).
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.
- MACDONALD, P. M. and G. STRUHL (1986).
A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern.
Nature 324: 537-545.
- MACDONALD, P. M. and G. STRUHL (1988).
Cis-acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos.
Nature 336: 595-598.
- MACDONALD, P. M., P. INGHAM and G. STRUHL (1986).
Isolation, structure and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeobox.
Cell 47: 721-734.
- MAHAFFEY, J. W. and T. C. KAUFMAN (1987).
Distribution of the Sex combs reduced Gene Products in *Drosophila melanogaster*.
Genetics 117: 51-60.
- MAHOWALD, A. P. and P. A. HARDY (1985).
Genetics of *Drosophila* Embryogenesis.
Ann. Rev. Genet. 19: 149-77.

- MANIATIS, T., E. F. FRITSCH and J. SAMBROOK (1982).
Molecular Cloning.
CSHL Press
- MARTINEZ-ARIAS, A (1986).
The Antennapedia gene is required and expressed in parasegments 4 and 5 of the *Drosophila* embryo.
EMBO J. 5: 135-141.
- MARTINEZ-ARIAS, A (1987).
On the developmental and evolutionary role of some genes from the Ant-C.
Molecular Approaches to Developmental Biology, 131-145.
- MARTINEZ-ARIAS, A., and P. A. LAWRENCE (1985).
Parasegments and compartments in the *Drosophila* embryo.
Nature 313: 639-642.
- MARTINEZ-ARIAS, A., and P. INGHAM (1986).
Form and diffusion.
Nature 324: 510-511.
- MARTINEZ-ARIAS, A., and P. W. INGHAM (1985).
The origin of pattern duplications in segment polarity mutants of *Drosophila melanogaster*.
J. Embryol. Exp. Morph. 87: 129-135.
- MARTINEZ-ARIAS, A., and R. A. H. WHITE (1988).
Ultrabithorax and engrailed expression in *Drosophila* embryos mutant for segmentation genes of the pair-rule class.
Development 102: 325-338.
- MARTINEZ-ARIAS, A., N. E. BAKER and P. W. INGHAM (1988).
Role of segment polarity genes in the definition and maintenance of cell states in the *Drosophila* embryos.
Development 103: 157-170.
- MARTINEZ-ARIAS, A., P. W. INGHAM, M. P. SCOTT and M. E. AKAM (1987).
The spatial and temporal deployment of Dfd and Scr transcripts throughout development of *Drosophila*.
Development 100: 673-683.
- MARTINEZ-CRUZADO, J. C., C. SWIMMER, M. G. FENERJIAN and F. C. KAFATOS (1988).
Evolution of the Autosomal Chorion Locus in *Drosophila*. I. General Organization of the Locus and Sequence Comparisons of Genes sl5 and sl9 in Evolutionarily Distant Species.
Genetics 119: 663-677.
- MATSUO, Y. and T. YAMAZAKI (1989).
Nucleotide Variation and Divergence in the Histone Multigene Family in *Drosophila melanogaster*.
Genetics 122: 87-97.
- McGINNIS, W (1985).
Homeobox Sequences of the Antennapedia Class Are Conserved Only in Higher Animal Genomes.
Cold Spring Harbor Symp. Quant. Biol. 50: 263-270.
- McGINNIS, W., M. S. LEVINE, E. HAFEN, A. KUROIWA and W. J. GEHRING (1984).
A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and bithorax complexes.
Nature 308: 428-433.
- MEINHARDT, H (1986).
The threefold subdivision of segments and the initiation of legs and wings in insects.
TIG 2: 36-41.
- MLODZIK, M., A. FJOSE and W. J. GEHRING (1985).
Isolation of caudal, a *Drosophila* homeobox containing gene with maternal expression, whose transcripts form a concentration gradient at the preblastoderm stage.
EMBO J. 4: 2961-2969.

- 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.
- MLODZIK, M., and W. J. GEHRING (1987).
Expression of the caudal Gene in the Germ Line of Drosophila: Formation of an RNA and Protein Gradient during Early Embryogenesis.
Cell 48: 465-478.
- MLODZIK, M., C. M. DE MONTRION, Y. HIROMI, H. M. KRAUSE and W. J. GEHRING (1987).
The influence on the blastoderm fate map of maternal-effect genes that affect the antero-posterior pattern in Drosophila.
Genes & Development 1: 603-614.
- MÜLLER, M. M., S. RUPPERT, W. SCHAFFNER and P. MATTHIAS (1988).
A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells.
Nature 336: 544-551.
- MÜLLER, M., M. AFFOLTER, W. LEUPIN, G. OTTING, K. WÜTHRICH and W. J. GEHRING (1988).
Isolation and sequence-specific DNA binding of the Antennapedia homeodomain.
EMBO J. 7: 4299-4304.
- NUSSLEIN-VOLHARD, C. and E. WIESCHAUS (1980).
Mutations affecting segment number and polarity in Drosophila.
Nature 287: 795-801.
- NUSSLEIN-VOLHARD, C., H. KLUDING and G. JÜRGENS (1985).
Genes Affecting the Segmental Subdivision of the Drosophila Embryo.
Cold Spring Harbor Symp. Quant. Biol. 50: 145-154.
- O'CONNOR, M. B., R. BINARI, L. A. PERKINS and W. BENDER (1988).
Alternative RNA products from the Ultrabithorax domain of the bithorax complex.
EMBO J. 7: 435-445.
- O'FARRELL, P. H., C. DESPLAN, S. DI NARDO, J. A. KASSIS, J. M. KUNER, E. SHER, J. THEIS, and D. WRIGHT (1985).
Embryonic Pattern in Drosophila: The Spatial Distribution and Sequence-specific DNA Binding of engrailed Protein.
Cold Spring Harbor Symp. Quant. Biol. 50: 235-242.
- OTTING, G., Y. QIAN, M. MÜLLER, M. AFFOLTER, W. GEHRING and K. WÜTHRICH (1988).
Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif.
EMBO J. 7: 4305-4309.
- PANKRATZ, M. J., M. HOCH, E. SEIFERT and H. JÄCKLE (1989).
Krüppel requirement for knirps enhancement reflects overlapping gap gene activities in the Drosophila embryo.
Nature 341: 337-339.
- PETERSEN, G., J. C. HALL and M. ROSBASH (1988).
The period gene of Drosophila carries species-specific behavioral instructions.
EMBO J. 7: 3939-3947.
- PFEIFER, M., F. KARCH and W. BENDER (1987).
The bithorax complex: control of segmental identity.
Genes & Development 1: 891-898.
- POOLE, S. J., L. M. KAUVAR, B. DREES and T. KORNBERG (1985).
The engrailed Locus of Drosophila: Structural Analysis of an Embryonic Transcript.
Cell 40: 37-43.
- PREISS, A., D. A. HARTLEY and S. ARTAVANIS-TSAKONAS (1988).
The molecular genetics of Enhancer of split, a gene required for embryonic neural development in Drosophila.
EMBO J. 7: 3917-3927.

- PREISS, A., U. B. ROSENBERG, A. KIENLIN, E. SEIFERT and H. JÄCKLE (1985).
Molecular genetics of Krüppel, a gene required for segmentation of the *Drosophila* embryo.
Nature 313: 27-33.
- PTASHNE, M (1988).
How eukaryotic transcriptional activators work.
Nature 335: 683-689.
- PULTZ, M. A., R. J. DIEDERICH, D. L. CRIBBS and T. C. KAUFMAN (1988).
The proboscipedia locus of the Antennapedia Complex: a molecular and genetic analysis.
Genes and Development 2: 901-920.
- RAFF, E. C. and R. A. RAFF (1985).
Possible functions of the homeobox.
Nature 313: 185-186.
- REGULSKI, M., N. MCGINNIS, R. CHADWICK and W. MCGINNIS (1987).
Developmental and molecular analysis of Deformed: a homeotic gene controlling *Drosophila* head development.
EMBO J. 6: 767-777.
- RICKOLL, W. L. and S. J. COUNCE (1980).
Morphogenesis in the Embryo of *Drosophila melanogaster* - Germ Band Extension.
Roux's Archives of Developmental Biology 188: 163-177.
- RIJDEWIJK, F., M. SCHUERMAN, E. WAGENAAR, P. PARREN, D. WEIGEL and R. NUSSE (1987).
The *Drosophila* homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless.
Cell 50: 649-657.
- RILEY, P. D., S. B. CARROLL and M. P. SCOTT (1987).
The expression and regulation of Sex combs reduced protein in *Drosophila* embryos.
Genes & Development 1: 716-730.
- ROBERTSON, M (1987).
A genetic switch in *Drosophila* morphogenesis.
Nature 327: 556-557.
- ROBERTSON, M (1988).
Homeoboxes, POU proteins and the limits to promiscuity.
Nature 336: 522-524.
- ROMANI, S., S. CAMPUZANO and J. MODOLELL (1987).
The achaete-scute complex is expressed in neurogenic regions of *Drosophila* embryos.
EMBO J. 6: 2085-2092.
- ROSENBERG, U. B., A. PREISS, E. SEIFERT, H. JÄCKLE and D. C. KNIPPLE (1985).
Production of phenocopies by Krüppel antisense RNA injection into *Drosophila* embryos.
Nature 313: 703-706.
- ROSENBERG, U. B., C. SCHRÖDER, A. PREISS, A. KIENLIN, S. COTE, I. RIEDE and H. JÄCKLE (1986).
Structural homology of the product of the *Drosophila* Krüppel gene with *Xenopus* transcription factor IIIA.
Nature 319: 336-339.
- RUBIN, G. M (1988).
Drosophila melanogaster as an Experimental Organism.
Science 240: 1453-1459.
- RUBIN, M. R., L. E. TOTH, M. D. PATEL, P. D'EUSTACHIO, M. C. NGUYEN-HUU (1986).
A Mouse Homeobox Gene Is Expressed in Spermatocytes and Embryos.
Science 233: 663-667.

RUDDLE, F. H., C. P. HART, A. AWGULEWITSCH, A. FAINSDOD, M. UTSET, D. DALTON, N. KERK, M. RABIN, A. FERGUSON-SMITH, A. FIENBERG and W. McGINNIS (1985).

Mammalian Homeobox Genes.

Cold Spring Harbor Symp. Quant. Biol. 50: 277-284.

RUIZ-GOMEZ, M. and J. MODOLELL (1987).

Deletion analysis of the achaete-scute locus of *Drosophila melanogaster*.

Genes and Development 1: 1238-1246.

RUSHLOW, C., H. DOYLE, T. HOEY and M. LEVINE (1987).

Molecular characterization of the *zerknüllt* region of the Antennapedia gene complex in *Drosophila*.

Genes and Development 1: 1268-1279.

RUSHLOW, C., M. FRASCH, H. DOYLE and M. LEVINE (1987).

Maternal regulation of *zerknüllt*: a homeobox gene controlling differentiation of dorsal tissues in *Drosophila*.

Nature 330: 583-586.

SAMSON, M. L., L. JACKSON-BRUSBY and R. BRENT (1989).

Gene Activation and DNA Binding by *Drosophila* Ubx and abd-A Proteins.

Cell 57: 1045-1052.

SANCHEZ-HERRERO, E. and M. A. CROSBY (1988).

The Abdominal-B gene of *Drosophila melanogaster*: overlapping transcripts exhibit two different spatial distributions.

EMBO J. 7: 2163-2173.

SANCHEZ-HERRERO, E., I. VERNOS, R. MARCO and G. MORATA (1985).

Genetic organization of *Drosophila* bithorax complex.

Nature 313: 108-113.

SANCHEZ-HERRERO, E., J. CASANOVA, S. KERRIDGE and G. MORATA (1985).

Anatomy and Function of the Bithorax Complex of *Drosophila*.

Cold Spring Harbor Symp. Quant. Biol. 50: 165-172.

SANGER, F., S. NICKLEN and A. R. COULSON (1977).

DNA sequencing with chain-terminating inhibitors.

Proc. Natl. Acad. Sci. USA 74: 5463-5467.

SCHAEFFER, S. W. and C. F. AQUADRO (1987).

Nucleotide Sequence of the Adh Gene Region of *Drosophila pseudoobscura*: Evolutionary Change and Evidence for an Ancient Gene Duplication.

Genetics 117: 61-73.

SCHIEDEREIT, C., J. A. CROMLISH, T. GERSTER, K. KAWAKAMI, C. G. BALMACEDA, R. A. CURRIE and R. G. ROEDER (1988).

A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein.

Nature 336: 551-557.

SCHLEIF, R (1988).

DNA Binding by Proteins.

Science 241: 1182-1187.

SCHNEUWLY, S. and W. J. GEHRING (1985).

Homeotic transformation of thorax into head: Developmental analysis of a new Antennapedia allele in *Drosophila melanogaster*.

Dev. Biol. 108: 377-386.

SCHNEUWLY, S., A. KUROIWA and W. J. GEHRING (1987).

Molecular analysis of the dominant homeotic Antennapedia phenotype.

EMBO J. 6: 201-206.

SCHNEUWLY, S., A. KUROIWA, P. BAUMGARTNER and W. J. GEHRING (1986).

Structural organization and sequence of the homeotic gene Antennapedia of *Drosophila melanogaster*.

EMBO J. 5: 733-739.

SCHNEUWLY, S., R. KLEMENZ and W. J. GEHRING (1987).

Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene *Antennapedia*.
Nature 325: 816-818.

SCHUH, R., W. AICHER, U. GAUL, S. COTE, A. PREISS, D. MAIER, E. SEIFERT, U. NAUBER, C. SCHRÖDER, R. KEMLER and H. JÄCKLE (1986).
A Conserved Family of Nuclear Proteins Containing Structural Elements of the Finger Protein Encoded by *Krüppel*, a *Drosophila* Segmentation Gene.

Cell 47: 1025-1032.

SCOTT, M. P. (1985).

Molecules and puzzles from the *antennapedia* homeotic gene complex of *Drosophila*.
TIG 1: 74-79.

SCOTT, M. P. (1987).

Complex loci of *Drosophila*.
Ann. Rev. Biochem. 56: 195-227.

SCOTT, M. P. and 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. USA 81: 4115-4119.

SCOTT, M. P. and P. H. O'FARRELL (1986).

Spatial programming of gene expression in early *Drosophila* embryogenesis.
Ann. Rev. Cell Biol. 2: 49-80.

SCOTT, M. P. and S. B. CARROLL (1987).

The Segmentation and Homeotic Gene Network in Early *Drosophila* Development.
Cell 51: 689-698.

SCOTT, M. P., A. J. WEINER, T. I. HAZELRIGG, B. A. POLISKY, V. PIRROTTA, F. SCALENGHE and T. C. KAUFMAN (1983).

The molecular organization of the *Antennapedia* locus of *Drosophila*.
Cell 35: 763-776.

SEEGER, M. A. and T. C. KAUFMAN (1987).

Homeotic Genes of the *Antennapedia* Complex (*ANT-C*) and Their Molecular Variability in the Phylogeny of the *Drosophilidae*.
Development as an Evolutionary Process 179-202.

SPRENGER, F., L. M. STEVENS and C. NÜSSLEIN-VOLHARD (1989).

The *Drosophila* gene *torso* encodes a putative receptor tyrosine kinase.
Nature 338: 478-483.

STANOJEVIC, D., T. HOEY and M. LEVINE (1989).

Sequence-specific DNA-binding activities of the gap proteins encoded by *hunchback* and *Krüppel* in *Drosophila*.
Nature 341: 331-335.

STEWARD, R., F. J. McNALLY and P. SCHEDL (1984).

Isolation of the dorsal locus of *Drosophila*.
Nature 311: 262-264.

STEWARD, R., L. AMBROSIO and P. SCHEDL (1985).

Expression of the Dorsal Gene.
Cold Spring Harbor Symp. Quant. Biol. 50: 223-228.

STEWARD, R., S. B. ZUSMAN, L. H. HUANG and P. SCHEDL (1988).

The dorsal Protein Is Distributed in a Gradient in Early *Drosophila* Embryos.
Cell 55: 487-495.

STROEHER, V. L., E. M. JORGENSEN and R. L. GARBER (1986).

Multiple Transcripts from the *Antennapedia* Gene of *Drosophila melanogaster*.
Molecular and Cellular Biology 6: 4667-4675.

- STRUHL, G (1981).
A homeotic mutation transforming leg to antenna in *Drosophila*.
Nature 292: 635-638.
- STRUHL, G (1985).
Near-reciprocal phenotypes caused by inactivation or indiscriminate expression of the *Drosophila* segmentation gene *ftz*.
Nature 318: 677-680.
- STRUHL, G. and M. AKAM (1985).
Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of *Drosophila*.
EMBO J. 4: 3259-3264.
- STURM, R. A. and W. HERR (1988).
The POU domain is a bipartite DNA-binding structure.
Nature 336: 601-604.
- SULLIVAN, W (1987).
Independence of fushi tarazu expression with respect to cellular density in *Drosophila* embryos.
Nature 327: 164-167.
- TAUTZ, D (1988).
Regulation of the *Drosophila* segmentation gene hunchback by two maternal morphogenetic centres.
Nature 332: 281-284.
- TAUTZ, D., R. LEHMANN, H. SCHNÜRCH, R. SCHUH, E. SEIFERT, A. KIENLIN, K. JONES and H. JÄCKLE (1987).
Finger protein of novel structure encoded by hunchback, a second member of the gap class of *Drosophila* segmentation genes.
Nature 327: 383-389.
- TEUGELS, E. and A. GHYSEN (1985).
Domains of action of bithorax genes in *Drosophila* central nervous system.
Nature 314: 558.
- THALI, M., M. M. MÜLLER, M. DE LORENZI, P. MATTHIAS and M. BIENZ (1988).
Drosophila homeotic genes encode transcriptional activators similar to mammalian OTF-2.
Nature 336: 598-601.
- TOMLINSON, A., B. E. KIMMEL and G. M. RUBIN (1988).
rough, a *Drosophila* Homeobox Gene Required in Photoreceptors R2 and R5 for Inductive Interactions in the Developing Eye.
Cell 55: 771-784.
- TOMLINSON, A., D. D. L. BOWTELL, E. HAFEN and G. M. RUBIN (1987).
Localization of the sevenless protein, a putative receptor for positional information, in the eye imaginal disc of *Drosophila*.
Cell 51: 143-150.
- TREIER, M., C. PFEIFLE and D. TAUTZ (1989).
Comparison of the gap segmentation gene hunchback between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change.
EMBO J. 8: 1517-1525.
- TREISMAN, J. and C. DESPLAN (1989).
The products of the *Drosophila* gap genes hunchback and Krüppel bind to the hunchback promoters.
Nature 341: 335-337.
- UTSET, M. F., A. AWGULEWITSCH, F. H. RUDDLE and W. MCGINNIS (1987).
Region-Specific Expression of Two Mouse Homeobox Genes.
Science 235: 1379-1382.
- VÄSSIN, H. and J. A. CAMPOS-ORTEGA (1987).
Genetic Analysis of Delta, a Neurogenic Gene of *Drosophila melanogaster*.
Genetics 116: 433-445.

- VASSIN, H., K. A. BREMER, E. KNUST and J. A. CAMPOS-ORTEGA (1987).
The neurogenic gene Delta of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats.
EMBO J. 6: 3431-3440.
- WAKIMOTO, B. T., R. A. LEWIS and T. C. KAUFMAN (1980).
Genetic analysis of the Antennapedia gene complex: mutant screen of proximal 3R, bands 84A-84B1.
DIS 55: 140.
- WEDEEN, C., K. HARDING AND M. LEVINE (1986).
Spatial Regulation of Antennapedia and Bithorax Gene Expression by the Polycomb Locus in *Drosophila*.
Cell 44: 739-748.
- WEINZIERL, R., J. M. AXTON, A. GHYSEN and M. AKAM (1987).
Ultrabithorax mutations in constant and variable regions of the protein coding sequence.
Genes & Development 1: 386-397.
- WEIR, M. P. and T. KORNBERG (1985).
Patterns of engrailed and fushi tarazu transcripts reveal novel intermediate stages in *Drosophila* segmentation.
Nature 318: 433-438.
- WHITE, R. A. H. and M. E. AKAM (1985).
Contrabithorax mutations cause inappropriate expression of Ultrabithorax products in *Drosophila*.
Nature 318: 567.
- WHITE, R. A. H. and M. WILCOX (1984).
Protein Products of the Bithorax Complex in *Drosophila*.
Cell 39: 163-171.
- WHITE, R. A. H. and M. WILCOX (1985).
Regulation of the distribution of Ultrabithorax proteins in *Drosophila*.
Nature 318: 563-567.
- WHITE, R. A. H. and R. LEHMANN (1986).
A Gap Gene, hunchback, Regulates the Spatial Expression of Ultrabithorax.
Cell 47: 311-321.
- WIESCHAUS, E. and C. C. NÜSSLEIN-VOLHARD (1986).
Looking at embryos.
Drosophila: a practical approach (Ed. D. B. Roberts). IRL Press: 199-227.
- WIESCHAUS, E. and R. RIGGLEMAN (1987).
Autonomous requirements for the segment polarity gene armadillo during *Drosophila* development.
Cell 49: 177-184.
- WILDE, C. D. and M. AKAM (1987).
Conserved sequence elements in the 5' region of the Ultrabithorax transcription unit.
EMBO J. 6: 1393-1401.
- WILKINS, A. S. (1986).
Genetic analysis of animal development.
Wiley-Interscience.
- WINSLOW, G. M., S. HAYASHI, M. KRASNOW, D. S. HOGNESS and M. P. SCOTT (1989).
Transcriptional Activation by the Antennapedia and fushi tarazu Proteins in Cultured *Drosophila* Cells.
Cell 57: 1017-1030.
- WIRZ, J., L. I. FESSLER and W. J. BEHRING (1986).
Localization of the Antennapedia protein in *Drosophila* embryos and imaginal discs.
EMBO J. 5: 3327-3334.

WRIGHT, C. V. E., K. W. Y. CHO, A. FRITZ, T. R. BÜRLIN and E. M. DE ROBERTIS (1987).
A *Xenopus laevis* gene encodes both homeobox-containing and homeobox-less transcripts.
EMBO J. 6: 4083-4094.

ZINK, B. and R. FARO (1989).
In vivo binding pattern of a trans-regulator of homeotic genes in *Drosophila melanogaster*.
Nature 337: 468-471.

BIBLIOTECA
FACULTAT DE C. BIOLÒGUES
VALÈNCIA

UNIVERSIDAD DE VALENCIA

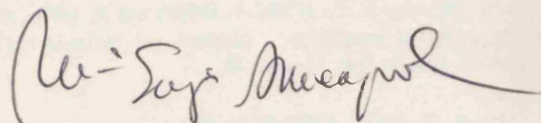
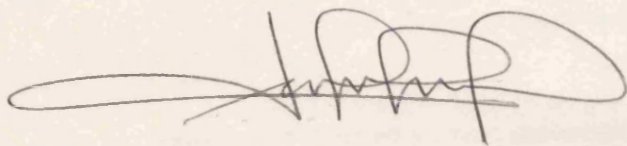
FACULTAD DE CIENCIAS BIOLÓGICAS

Reunido el Tribunal que suscribe, en el día de la fecha,
acordó otorgar, por unanimidad, a esta Tesis doctoral de
D. Manuel Pérez Alonso
la calificación de Apto "Cum Laude"

Valencia, a 22 de Diciembre de 1989

El Secretario,

El Presidente



DEPARTAMENT DE
GENÈTICA
ENTRADA
Nº 177
data 15.11.89