



Universitat Autònoma de Barcelona

Facultat de Biociències

Departament de Genètica i de Microbiologia

Unitat de Genètica

**VARIACIÓN EPIGENÉTICA Y EXPRESIÓN DEL PAR DE
GENES *w* Y *CG32795* EN DISTINTAS CEPAS MUTANTES
*zeste*¹ DE *Drosophila melanogaster*.**

Memoria presentada por ANNA PORTELA MESTRES, en el marco del programa de tercer ciclo de Genética de la Universitat Autònoma de Barcelona, para optar al grado de Doctor en Genética por la Universitat Autònoma de Barcelona.

El trabajo se ha realizado en los laboratorios del Grup de Mutagènesi bajo la codirección de los Doctores ORIOL CABRÉ FABRÉ y NOEL XAMENA LÓPEZ.

Director de Tesis

Director de Tesis

Autor

Dr. Oriol Cabré Fabré

Dr. Noel Xamena López

Anna Portela Mestres

*S'ha d'intentar comprendre fent treballar la fantasia,
i s'ha d'oblidar el que se sap,
de manera que la imaginació pugui rondar lliurement,
pugui córrer lluny cap a l'interior de les coses...*

Oceà. Alessandro Baricco.

Últimament quan em pregunten l'adreça, dubto entre l'Autònoma i casa. No hauré passat més hores al laboratori que a casa aquests darrers anys? Últimament quan penso en setembre m'adono de tot el que deixaré enrere, i és per això que vull aprofitar ara per donar les gràcies a tots aquells que m'han acompanyat i recolzat en aquesta aventura que és la Tesi.

Per començar a la meva família adoptiva: els meus pares "científics" Oriol i Noel i el germà gran Martí. Per ser-hi quan calia, per discutir allò que amb d'altres no podia ni parlar, per l'experiència prèvia, per moments d'histèria i alegria col·lectiva, per inventar la "migració conceptual" i per posar-m'ho tot tan fàcil.

En segundo lugar, a ti Víctor. Por hacer que tres meses en tu laboratorio parecieran siete u ocho. No porque se me hiciera largo! Al contrario... Por todo lo que aprendí y gané con la experiencia, contigo y con toda tu gente.

Perquè una Tesi no només és ciència. Per tots aquells que han fet que així sigui, els que encara hi són i els que ja han marxat, els que he tingut menys temps per a conèixer i els que fa anys que m'aguanten. Pels cafès a mig matí i els tes a la tarda, pels dinars a l'horrible self i per la millora qualitativa que ha suposat la sortida al campus, per cantar-me pel meu aniversari, pels "voy" o "vengo", pels LOST fòrums, per tantes converses a les escales, a les SEMAs, als Ferrocarrils o a l'avió, per fer-me riure quan més ho necessitava. Per haver-me deixat participar en moments decisius de les vostres vides. Per ensenyar-me castellà, paciència, tolerància, optimisme i per haver-me fet més lliure.

I ja us deia jo que una Tesi és una aventura, i com a tal, modifica el curs del teu dia a dia. I com fa poc em vareu dir: per tots aquells que fora del laboratori han patit la meva Tesi en les meves absències!

I finalment a la meva família de veritat: F+GOT. Per les visites, els mails, el messenger i la webcam allà on anés. Per preocupar-vos amb motiu o sense. Perquè encara que us ho he posat difícil heu intentat entendre'm i fer-me costat. Gràcies als quatre per la paciència.

I parlant de paciència només em quedes tu. Gràcies Pau. Pel teu optimisme i alegria contagiosos. Per fer-me veure tot allò que hi ha més enllà de la ciència. Per suportar els meus horaris intempestius. Per, amb una abraçada, fer-me superar les meves crisis existencials. Perquè sense tu no ho hauria fet.

ÍNDICE

ÍNDICE	7
RESUMEN	9
SÍMBOLOS, ACRÓNIMOS Y ABREVIATURAS	11
I. INTRODUCCIÓN	15
1. Epigenética	15
1.1. La metilación del DNA	15
1.1.1. Efectos de la metilación del DNA	17
1.1.2. Regulación de la metilación del DNA	19
1.1.3. DNA metilasas (DNMTasas)	21
1.1.4. La metilación del DNA en <i>D. melanogaster</i>	22
1.2. Código histónico	23
1.3. Remodelación de la estructura de la cromatina	25
1.3.1. El complejo remodelador de la cromatina SWI/SNF o Brahma	31
2. Los grupos <i>Polycomb</i> y <i>trithorax</i>	32
2.1. El grupo <i>Polycomb</i>	33
2.2. El grupo <i>trithorax</i>	37
2.2.1. El gen <i>zeste</i>	40
2.2.1.1. La mutación <i>zeste</i> ¹	43
3. El gen <i>white</i>	44
3.1. La interacción <i>zeste-white</i>	47
4. El gen <i>decapentaplegic</i>	48
4.1. Interacción <i>zeste-dpp</i>	50
5. El gen <i>CG3279</i>	51
6. Los elementos transponibles	53
6.1. El elemento transponible <i>FB-NOF</i>	57
6.2. Efectos de la inserción de un elemento transponible	60
7. Las cepas M115 y RM115	61
II. OBJETIVOS	65
III. RESULTADOS Y DISCUSIÓN	67
1. Estudio del efecto de la inserción de un elemento transponible <i>FB-NOF</i> en la expresión de los genes circundantes	67
1.1. Estudio del gen <i>CG32795</i>	67
1.1.1. Patrones de <i>splicing</i> alternativo	67
1.1.1.1. Clonación del extremo 5'	69
1.1.1.2. Clonación del extremo 3'	72
1.1.2. Predicciones sobre su posible función	73
1.2. Cuantificación de la expresión de los genes <i>white</i> y <i>CG32795</i>	79
1.2.1. El gen <i>w</i>	80

1.2.2. El gen <i>CG32795</i>	87
1.3. Discusión general de los resultados	92
2. Estudio epigenético del efecto de la mutación <i>zeste</i>¹	97
2.1. Cuantificación de la expresión de los genes <i>w</i> , <i>CG32795</i> y <i>dpp</i>	97
2.2. Posicionamiento de nucleosomas	100
2.3. Patrón de metilación del DNA en hembras CS y <i>z</i> ¹	104
2.4. Discusión general de los resultados	113
IV. CONCLUSIONES	119
V. MATERIALES Y MÉTODOS	121
1. Cepas de <i>Drosophila melanogaster</i>	121
2. Extracción de DNA genómico.....	122
3. PCR a tiempo real	122
3.1. <i>LightCycler</i>	122
3.1.1. PCR a tiempo real cuantitativa.....	127
3.1.2. Cebadores y reacción	130
3.2. Extracción de RNA.....	135
4. Cuantificación de pigmentos oculares	135
5. Ensayo de la nucleasa micrococcal	135
5.1. Aislamiento de núcleos.....	136
5.2. Digestión con nucleasa micrococcal.....	137
5.3. Purificación del DNA y digestión con enzimas de restricción.....	137
5.4. <i>Southern blot</i>	137
5.4.1. Sondas.....	138
6. Ensayo de modificación del DNA por bisulfito sódico.....	139
6.2. Digestión y tratamiento con bisulfito sódico	141
6.3. PCR.....	141
6.4. Clonación.....	143
6.5. Secuenciación y análisis de las secuencias	144
7. Análisis del gen <i>CG32795</i>	145
7.1. RACE	145
7.1.1. RACE 5'	145
7.1.2. RACE 3'	148
7.2. RT-PCR	149
7.3. Análisis bioinformático	151
VI. BIBLIOGRAFÍA	153

RESUMEN

El presente trabajo de tesis se centra en el estudio de diferentes mutantes *zeste*¹ y en el efecto que esta mutación tiene sobre la regulación de la transcripción en diversos genes.

En primer lugar se estudiaron los machos de las cepas M115 y RM115 cuya característica principal, además de ser portadores de la mutación *z*¹, es la inserción de un elemento transponible *FB-NOF* en el tercer intrón de *CG32795*, que forma con *w* un par de genes *head-to-tail*. Para poder estudiar cual es el efecto de la inserción de *FB-NOF* en un entorno *z*¹, fue necesario caracterizar en primer lugar el gen *CG32795*, a nivel de secuencia del mRNA y de predicción de la posible función de la proteína codificada. Se encontraron diversas variantes de *splicing*, tanto para su extremo 5' como para el 3', parecidas pero diferentes de las variantes predichas anteriormente. Conociendo más en profundidad este gen, nos propusimos estudiar los posibles efectos de la inserción de *FB-NOF* respecto a la expresión de los genes *w* y *CG32795*. Los resultados nos llevaron a un estudio más detallado de la expresión de *w* en diferentes partes del cuerpo, teniendo en cuenta la especificidad de tejido que la interacción *zeste-white* presenta. Así, vimos que el gen *w* no se ve afectado por la inserción de *FB-NOF* en su extremo 3' y que las diferencias de expresión observadas son debidas a la duplicación del *Zeste Binding Site* de *w* en un entorno *z*¹. Sin embargo, la inserción de *FB-NOF* en el tercer intrón de *CG32795* sí que modifica la expresión de este gen. No sólo la inserción es responsable de la alteración de la expresión de *CG32795* sino también la reordenación que se produce en la cepa RM115 eliminando la copia de *w* original y dejando la que se duplicó en M115.

La segunda parte de esta tesis consiste en el estudio de las hembras mutantes *z*¹. Analizamos la expresión de dos genes con un ZBS (*w* y *dpp*) y observamos como se reduce la expresión de dichos genes en las hembras *z*¹. Además, estudiando la expresión del gen *CG32795* constatamos que el efecto de la interacción *zeste-white* es local, sin afectar a este gen aunque su extremo 5' se encuentra a tan solo 700bp del extremo 3' del gen *w*. La reducción de la expresión nos hizo pensar que podía ser debida a alteraciones en la estructura de la cromatina, puesto que los agregados de Zeste en los ZBS son los

responsables de reclutar el complejo remodelador de la cromatina BRM, facilitando así la transcripción. Mediante un ensayo de nucleasa micrococcal detectamos algunas modificaciones en el posicionamiento de nucleosomas en los ZBS de *w* y *dpp*, siendo el posicionamiento en las hembras z^1 más estricto. Si el complejo BRM es el responsable de estas diferencias, pensamos que los patrones de metilación también podrían estar alterados, pues muchos factores asociados a los complejos SWI/SNF se han relacionado con actividades de regulación de la expresión mediante modificación de los mismos. Mediante el ensayo de modificación del DNA por bisulfito sódico estudiamos los patrones de metilación de cinco regiones. Sorprendentemente, el ZBS de *w* y de *dpp*, así como las regiones próximas a ellos, se encontraban hipometilados en las hembras z^1 . El desconocimiento general sobre la metilación en *D. melanogaster* nos hizo plantear cuales pueden ser las secuencias diana de la metilación en esta especie. Así realizamos un estudio estadístico sobre cuales son las dos bases anteriores y posteriores a las Cs metiladas en nuestras secuencias. Se obtuvieron diferencias en las frecuencias de cada posición y se estableció como secuencia más frecuente para ser metilada ApDp5mCpDpD. Aún así, es una secuencia muy degenerada, y se hacen necesarios estudios más profundos y amplios para confirmarla.

SÍMBOLOS, ACRÓNIMOS Y ABREVIATURAS

5AzaC - *5-AzaCytidine* - 5-Azacitidina

5mC - *5-methylcytosine* - 5 - metilcitosina

ash - *absent, small or homeotic discs*

BAF - *BRM Associated Factor* - Factor asociado a BRM

BAP - *BRM Associated Protein* - Proteína asociada a BRM

BRG - *BRM Related Gene* - Gen relacionado con BRM

BRM - *BRahMa*

BSM - *BiSulfite Modification* - Modificación por bisulfito

Caf - *Chromatin Assembly Factor* - Factor de empaquetamiento de la cromatina

ChrAC - *Chromatin Accessibility Complex* - Complejo de acceso a la cromatina

CHD - *Chromo-Helicase/ATPase DNA binding* - Cromo-helicasa/ATPasa de unión a DNA

CMT - *ChromoMeThylase* - Cromometilasa

C_p - *Crossing point* - Punto de cruce

CS - Canton S

CSB - *Cockayne Syndrome Group B* - Grupo B del síndrome de Cockayne

C_t - *Threshold cycle* - Ciclo umbral

cyc - *cyclin* - ciclina

dCBP - *Drosophila homolog of CREB Binding Protein* - Homólogo de *Drosophila* de la proteína de unión a CREB

DCC - *Dosage Compensation Complex* - Complejo de compensación de dosis

dpp - *decapentaplegic*

DDM1 - *Decreased DNA Methylation* - Metilación disminuida del DNA

DNMTasa - *DNA MethylTransferase* - DNA metiltransferasa

DRE - *Distant Regulatory Elements* - Elementos Reguladores Distantes

DRM - *Domains Rearranged Methyltransferases* - Metiltransferasa de dominios reorganizados

EN - *EndoNuclease* - Endonucleasa

Esc - *Extra sex combs*

ET - *Transposable Element* - Elemento Transponible

eya - *eyes absent*

E(bx) - *Enhancer of bithorax*

E(z) - *Enhancer of zeste*

FB - *Foldback*

H - *Histone* - Histona

HDAC - *Histone DesAcetilase* - Histona desacetilasa

HH - *Head-to-Head* - Cabeza-con-cabeza

Hin - *HaploINSuficiencia* - Haploinsuficiencia

HMMTOP - *Hidden Markov Model TOpology Prediction* - Modelo oculto de Markov para la predicción topológica

HMT - *Histone MethyTransferase* - Histona metiltransferasa

HT - *Head-to-Tail* - Cabeza-con-cola

HTH - *Helix Turn Helix* - Hélice-Giro-Hélice

ICF - *Immunodeficiency, Centromeric region instability, Facial anomalies* - Inmunodeficiencia, inestabilidad de la región centromérica y anomalías faciales

IN - *Integrase* - Integrasa

INO - *INOsitol requiring* - Requiere inositol

IR - *Inverted Repeat* - Repetición invertida

ISWI - *Imitation SWI* - Imitador de SWI

K - *Lysine* - Lisina

kDa - *kiloDalton* - kilodalton

LB - Luria-Bertani

LHP1 - *Like Heterochromatin Protein 1* - Proteína Parecida a la Heterocromatina 1

LINE - *Long Interspersed Nuclear Element* - Elemento nuclear intercalado largo

Lsh - *Lymphoid Specific Helicase* - Helicasa específica limfoidea

LTR - *Long Terminal Repeats* - Repeticiones terminales largas

MeCp - *Methyl C binding protein* - Proteína de unión a Cs metiladas

MBD - *Methyl Binding Domain* - Dominio de unión a grupos metilo

MBT - *Malignant Brain Tumor* - Tumor cerebral maligno

MITE - *Miniature Inverted-repeat Transposable Element* - Elemento transponible en miniature con repeticiones invertidas

MNasa - *Micrococcal Nuclease* - Nucleasa micrococcal

mod(mdg4) - *Modifier mdg4*

Mor - *moira*

ncRNA - *Non Coding RNA* - RNA no codificante

NER – *Nucleotid Excision Repair* - Reparación por escisión de nucleótidos

NOF - *Nofertette* - Nefertiti

NuRD - *Nucleosome Remodeling and Deacetylase* - Desacetilasa y remodelador de nucleosomas

NuRF - *Nucleosome Remodeling Factor* - Factor de remodelación de nucleosomas

ORF - *Open Reading Frame* - Marco de lectura abierto

Pc - *PolyComb*

PcG - *Pc Group* - Grupo *Pc*

PCR - *Polymerase Chain Reaction* - Reacción en cadena de la polimerasa

PEV - *Position Effect Variegation* - Variegación por efecto de posición

Ph – *Polyhomeotic*

Pho - *Pleiohomeotic*

PhoL - *Pho Like*

PRC - *Pc Repressive Complex* - Complejo represor *Pc*

PRE - *Pc Response Element* - Elemento de respuesta a *Pc*

Psc - *Posterior sex combs*

R - *Purine* - Purina

RACE - *Rapid Amplification of cDNA Ends* - Amplificación rápida de los extremos de cDNA

Rbf - *Retinoblastoma family* - Familia del retinoblastoma

RNAi - *RNA interference* - RNA de interferencia

rpm - Revoluciones por minuto

RSC - *Chromatin Remodeling Complex* - Complejo Remodelador de la Cromatina

RT - *RetroTranscriptase* - Retrotranscriptasa

RT-PCR - *Retrotranscriptive PCR* - PCR retrotranscriptiva

SAH - *S-adenosil-L-homocysteine* - S-adenosil-L-homocisteina

SAM - *S-adenosil-L-methionine* - S-adenosil-L-metionina

SAM domain - *Sterile Alpha Motif domain* - Dominio con motivo esteril alfa

SANT - SWI3, ADA2, N-CoR, TFIIB

Sbfl - *SET Binding Factor* - Factor de unión de SET

Scm - *Sex comb on midleg*

SET - *Su(var), E(z), Trx*

shv - *shortvein*

SINE - *Short Interspersed Nuclear Element* - Elemento nuclear intercalado corto

siRNA - *small interfering RNA* - RNA pequeño de interferencia

SNF - *Sucrose Non Fermenting gene*

SRP - *Signal Particle Recognition* - Partícula de reconocimiento de señal

SUMO - *Small Ubiquitin MOdifier* - Modificador de la ubiquitina pequeño

Su(z) - *Suppressor of zeste*

SWI - *SWItch gene*

TAC - *Trithorax Acetylation Complex* - Complejo acetilador de *trithorax*

TAF - *TBP Associated Factor* - Factor asociado a TBP

TBP - *TATA Binding Protein* - Proteína de unión a TATA

TCR - *Transcription Coupled Repair* - Reparación ligada a la transcripción

TF - *Transcription Factor* - Factor de transcripción

TGF- β - *Transforming Growth Factor β* - Factor de crecimiento transformante tipo β

T_m - *Melting Temperature* - Temperatura de fusión

TMHMM - *TransMembrane Hidden Markov Model* - Modelo oculto de Markov para transmembranas

TMPIT - *TransMembrane Protein Induced by Tumor necrosis factor alpha* - Proteína de membrana inducida por el factor de necrosis tumoral alfa

trl - *trx like* - Parecido a *trx*

trx - *trithorax*

trxG - *trx Group* - Grupo *trx*

Tth - *Thermus thermophilus*

Ubx - *Ultrabithorax*

UR - *Unstable Red* - Rojo inestable

UTR - *UnTranslated Region* - Región no traducida

UZ - *Unstable Zeste* - Zeste inestable

VEFS - VRN2-EMF2-FIS2-Su(z)12

VLP - *Virus Like Particle* - Partícula parecida a virus

w - *white*

y - *yellow*

Y - *Pyrimidine* - Pirimidina

z - *zeste*

ZBM - *Zeste Binding Motif* - Motivo de unión de *Zeste*

ZBS - *Zeste Binding Site* - Sitio de unión de *Zeste*

I. INTRODUCCIÓN

1. Epigenética

“La epigenética siempre ha sido todas las cosas extrañas y maravillosas que no pueden ser explicadas por la genética.” Denise Barlow.

“El genoma son palabras y el epigenoma es lo que determina cuales van en negrita, cuales van en mayúscula, donde va el punto y coma, etc. Es, por tanto, aquello que finalmente da un sentido a las palabras.” Manel Esteller.

La epigenética (del griego *epi*, en o sobre) estudia los cambios reversibles en el DNA que no afectan la secuencia nucleotídica de un gen pero producen cambios en su expresión. Dichos cambios son heredables de una célula a sus descendientes. Algunos indicios apuntan que esta información podría mantenerse durante la meiosis y por tanto, la herencia epigenética podría transmitirse incluso a los descendientes de organismos pluricelulares.

La epigenética estudia todos aquellos grupos químicos (metil, acetil o fosfato) o proteínas (ubiquitina o SUMO) que se unen al DNA o a las proteínas asociadas con éste, dando lugar a metilaciones, acetilaciones, fosforilaciones, ubiquitinaciones y sumoilaciones. Estas modificaciones epigenéticas intervienen en procesos tan importantes como: el destino celular (*cell fate*), el desarrollo embrionario, la inestabilidad genómica y el cáncer. Los principales mecanismos de regulación epigenética son: la metilación del DNA, la modificación de las colas de las histonas (código histónico) y la remodelación de la estructura de la cromatina.

1.1. La metilación del DNA

La metilación del DNA es una modificación covalente del DNA que no modifica su secuencia nucleotídica pero que sin embargo puede tener una gran influencia sobre la regulación de la expresión génica. Además de las diferentes metilaciones a las bases nitrogenadas del DNA inducidas por la acción aleatoria de los agentes metilantes, tanto

endógenos como exógenos, el DNA de la mayor parte de los organismos sufre metilaciones postreplicativas, dando lugar a tres tipos de bases metiladas:

- N4-metilcitosina → Se encuentra principalmente en procariotas.
- N6-metiladenina → Se encuentra principalmente en procariotas.
- C5-metilcitosina (5mC) → Es la principal base metilada, tanto en eucariotas como en procariotas. La formación de la 5mC se produce por acción de las DNA metiltransferasas (DNMTasas), que transfieren un grupo metilo de la S-adenosil-L-metionina (SAM) al carbono 5 del anillo pirimidínico de la citosina. Durante este proceso, el tiol de la cisteína del enzima ataca el carbono 6 de la citosina y forma un intermediario covalente DNA-proteína. La adición del grupo tiol de la cisteína activa el carbono 5 permitiendo la transferencia del grupo metilo de la SAM y liberando S-adenosil-L-homocisteína (SAH). Este mecanismo de reacción se encuentra conservado a lo largo de toda la evolución (Wu y Santi, 1987).

El mecanismo de reclutamiento de las DNMTasas en regiones específicas del genoma no se conoce en profundidad. Se supone que implica interacciones con proteínas cromatínicas, factores de transcripción y RNAs. La síntesis de RNAs antisentido cubriendo islas CpG parece inducir la metilación de dichas islas (Tufarelli *et al.*, 2003).

En procariotas, la metilación del DNA interviene en el sistema de restricción-modificación (R/M), en la reparación de falsos apareamientos y en el control de la replicación, mientras que en eucariotas, desempeña un papel principal en la regulación de la expresión génica.

En mamíferos, se encuentran 5mCs a lo largo de todo el genoma, principalmente en los dinucleótidos CpG, dando lugar a metilaciones simétricas en ambas cadenas del DNA (Bird, 2002). En las células somáticas humanas, la 5mC constituye casi un 1% del total de las bases del DNA, metilándose así un 70-80% de todos los dinucleótidos CpG (Ehrlich y Wang, 1981). En el promotor de aproximadamente el 40% de los genes, se encuentran islas CpG (regiones de más de 500bp con un porcentaje de G+C superior al

55%) no metiladas. Algunas de estas islas CpG se metilan durante el desarrollo embrionario, estableciendo así el silenciamiento génico de esos genes (Cardoso y Leonhardt, 1999). Por este motivo, se asocia la metilación del promotor de un gen con el silenciamiento de dicho gen, aunque no siempre sea cierto. Por ejemplo, la isla CpG del gen de la α -globina humana no está metilado ni en tejidos eritroides ni en tejidos no eritroides (Bird *et al.*, 1987).

1.1.1. Efectos de la metilación del DNA

La metilación del DNA juega papeles importantes en el desarrollo, la diferenciación, el envejecimiento, la inactivación del cromosoma X, el *imprinting* y la tumorigénesis, entre otros. Los mecanismos a través de los cuales la metilación participa en estos procesos se pueden clasificar en:

I) INHIBICIÓN DE LA TRANSCRIPCIÓN → Se ha observado una correlación entre metilación e inhibición de la transcripción. Por ejemplo, las islas CpG se encuentran altamente metiladas en el cromosoma X inactivo, mientras que el activo tiene esas mismas islas no metiladas (Bird, 2002). La metilación del DNA puede inhibir la transcripción por cuatro vías (Klose y Bird, 2006):

1 – La metilación del DNA puede bloquear directamente la unión de un factor de transcripción (ver Figura I.1. a) (ej. AP-2, Comb y Goodman, 1990). Sin embargo, otros factores de transcripción son insensibles a la metilación del DNA (ej. Sp1, Hantusch *et al.*, 2007).

2 – La metilación del DNA puede inhibir la actividad del promotor indirectamente (ver Figura I.1. b). Algunos factores reconocen específicamente las citosinas metiladas y se unen a ellas bloqueando el acceso de los factores de transcripción a los promotores. El reconocimiento de las citosinas metiladas se hace mediante el dominio de unión a metilo (MBD – *Methyl Binding Domain*).

3 – La metilación del DNA puede inhibir la transcripción remodelando la estructura cromatínica (ver Figura I.1. c). Algunos miembros de la familia proteica MBD, son subunidades de grandes complejos remodeladores de la cromatina (ej. el complejo MeCP1 está formado por MBD2 y Mi-2/NuRD). La presencia de

desacetilasas de histonas y de actividad remodeladora de la estructura de la cromatina en dichos complejos proteicos, establece una relación funcional entre la metilación del DNA y otros mecanismos epigenéticos (código histónico y estructura de la cromatina) (Bird y Wolffe, 1999). En *Drosophila* también se ha descrito una relación bidireccional entre la metilación del DNA y la metilación de las colas de las histonas.

4 – La metilación del DNA puede dificultar la elongación de la transcripción (ver Figura I.1. d). Los mecanismos anteriores de silenciamiento permitirían explicar el papel de la metilación en promotores. Sin embargo, gran parte del DNA metilado se encuentra en intrones y exones. Algunos experimentos demuestran que la metilación del DNA en secuencias génicas reduce la accesibilidad de la RNA polimerasa II a dicha secuencia, dificultando así la elongación de la transcripción (Lorincz *et al.*, 2004).

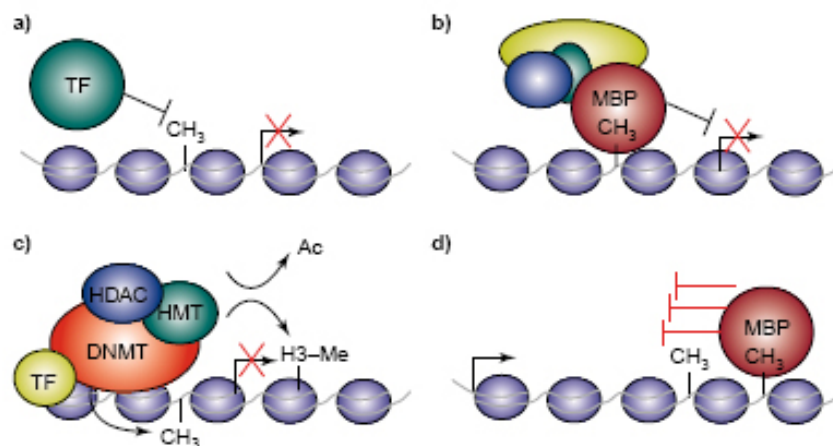


Figura I.1. – **Mecanismos de inhibición de la transcripción por metilación del DNA.** La metilación del DNA puede inhibir la transcripción: a) Bloqueando directamente la unión de un factor de transcripción. b) Inhibiendo la actividad del promotor indirectamente, reclutando proteínas que imposibilitarán la posterior unión de factores de transcripción. c) Remodelando la estructura de la cromatina. d) Dificultando la elongación de la transcripción. (Figura de Klose y Bird, 2006).

II) EFECTO EN LA ESTRUCTURA DE LA CROMATINA → La metilación del DNA tiene una gran influencia en la estructura del DNA (Leonhardt y Cardoso, 2000). Por ejemplo, la mutación del gen que codifica la DNMTasa3B está ligada al síndrome ICF (Inmunodeficiencia, inestabilidad de la región Centromérica y anomalías Faciales). Las células de los pacientes con ICF padecen deleciones y duplicaciones de brazos enteros de cromosomas y roturas de centrómero, además de una hipometilación

de los satélites II y III, principales componentes de la heterocromatina constitutiva. Esas regiones se encuentran normalmente muy metiladas. Así, la metilación de dichas regiones sería esencial para la correcta estructura y estabilidad del centrómero (Robertson y Wolffe, 2000b).

1.1.2. Regulación de la metilación del DNA

El patrón de metilación no se mantiene constante a lo largo de toda la vida de un individuo, sino que sufre sucesivas modificaciones. En mamíferos, los cambios en el patrón de metilación más importantes se producen durante el desarrollo embrionario. Después de la fecundación, se eliminan los patrones de metilación previamente establecidos (excepto en algunas zonas con *loci* con impronta génica) y se establece un nuevo patrón de metilación necesario para la expresión específica de tejido de algunos genes, para la inactivación del cromosoma X en hembras y para el *imprinting* genómico. Una vez establecidos, los nuevos patrones de metilación deben ser heredados por las células hijas. La herencia anormal de dichos patrones lleva a la formación de tumores y a procesos de envejecimiento. Los cambios en el patrón de metilación, tanto normalmente como en caso de enfermedad, se producen a través de tres mecanismos básicos:

I) METILACIÓN *DE NOVO* → Es la metilación de secuencias no metiladas, estableciendo así un nuevo patrón de metilación. Los niveles más altos de actividad DNMTasa *de novo* se han encontrado en carcinomas embrionarios y en células madre embrionarias (Stewart *et al.*, 1982; Lei *et al.*, 1996). La mayor parte de metilación *de novo* se produce durante el desarrollo embrionario, aunque también se da en los puntos de integración de DNA viral en el genoma (Toth *et al.*, 1990) y durante el envejecimiento (ej. hipermetilación del gen *c-myc* en el hígado de ratón (Ono *et al.*, 1989), entre otros. Así, aunque la metilación *de novo* ha sido demostrada en diferentes procesos biológicos y especies, se desconoce cómo se seleccionan las secuencias que van a ser metiladas *de novo*. Actualmente se propone que la metilación *de novo* de una secuencia es precedida por la metilación, en los nucleosomas próximos a dicha secuencia, de la Lys 9 de la histona H3 (Bachman *et al.*, 2003). Esta teoría sugiere que la metilación del DNA jugaría un papel importante en la estabilización de las secuencias silenciadas a través de las modificaciones de las colas de las histonas. Además de la selección de las secuencias a metilar *de novo*, es necesario el reclutamiento de las

DNMTasas en dichas secuencias. Se han propuesto tres vías de reclutamiento (Klose y Bird, 2006): a) Por interacción de dominios específicos de las propias DNMTasas con el DNA o la cromatina. b) Por interacción proteína-proteína, a través de factores de transcripción u otros factores. c) Por un sistema de RNAi.

II) METILACIÓN DE MANTENIMIENTO → Es el proceso mediante el cual se mantiene el patrón de metilación después de la replicación. Cada ronda de replicación genera una cadena de DNA de nueva síntesis no metilada, frente a la cadena complementaria vieja que sí está metilada. Esta reacción de metilación se lleva a cabo gracias a DNMTasas que prefieren un substrato hemimetilado (Goll y Bestor, 2005). El enzima clave para el mantenimiento de la metilación es la DNMT1.

III) DESMETILACIÓN → Es el proceso que permite eliminar los patrones de metilación previos. Ocurre principalmente durante el desarrollo preimplantacional, y en general, durante todo el desarrollo. Se puede producir mediante dos procesos:

1 – DESMETILACIÓN PASIVA → Se produce una reducción gradual de los niveles de metilación debida a la ausencia de metilación de mantenimiento a lo largo de sucesivas rondas de replicación. La desmetilación pasiva se produce de forma natural a lo largo del proceso de envejecimiento humano (Liu *et al.*, 2003). Sin embargo, este tipo de desmetilación también se puede favorecer con el tratamiento con drogas inhibitoras de las DNMTasas, como por ejemplo la 5AzaC (Jones y Taylor, 1980).

2 – DESMETILACIÓN ACTIVA → Se produce de forma independiente de la replicación del DNA, gracias a la intervención de enzimas específicos. Se han propuesto tres mecanismos bioquímicos para explicar la desmetilación del DNA: a) Escisión de las bases metiladas por una glicosilasa. b) Escisión del nucleótido metilado. c) Sustitución directa del grupo metil por un átomo de hidrógeno (Kress *et al.*, 2001). Algunos casos de desmetilación activa son la desmetilación global del genoma paterno en el cigoto (Mayer *et al.*, 2000) y la desmetilación global del genoma durante la diferenciación de los mioblastos (Jost y Jost, 1994).

1.1.3. DNA metilasas (DNMTasas)

Las DNMTasas han sido identificadas y clonadas en diversos organismos, tanto procariotas como eucariotas. Todas ellas comparten un dominio catalítico y se clasifican en cinco familias: DNMT1, DNMT2, DNMT3, Msc (DNMTasas de los hongos), y CMT (cromometilasa, exclusiva de plantas) (Colot y Rossignol, 1999). La existencia de diferentes DNMTasas sugiere que cada una de ellas tenga funciones específicas.

I) DNMT1 → Es la familia responsable de la metilación de mantenimiento, puesto que metila preferencialmente secuencias hemimetiladas. En cada ronda de replicación se genera una doble hélice de DNA que presenta la cadena parental metilada y la cadena de nueva síntesis sin metilar. La preferencia de la DNMT1 por las secuencias hemimetiladas permite copiar el patrón de metilación en la cadena recién sintetizada. En humanos, la DNMT1 está formada por dos dominios principales: el N-terminal regulador y el C-terminal catalítico. El dominio N-terminal puede participar en la inhibición transcripcional gracias a su interacción con desacetilasas de histonas (HDAC1 y HDAC2) y con represores de la transcripción (DMP1) (Fuks *et al.*, 2000; Robertson *et al.*, 2000a; Robertson *et al.*, 2000b). El dominio N-terminal también participa en la represión de la transcripción directamente, mediante una región relacionada con la proteína *trithorax* HRX (Fuks *et al.*, 2000). Se cree que DNMT1 no interviene solamente en el mantenimiento del patrón de metilación, sino que también juega un papel importante en el mantenimiento de la estructura de la heterocromatina. La pérdida de DNMT1 causa, entre otros efectos, desmetilación y alta expresión de los retrotransposones IAP. *Dnmt1* es el único gen conocido hasta el momento necesario para la represión de los transposones en células somáticas humanas. Se han encontrado homólogos de la DNMT1 en muchos organismos incluyendo: hongos, plantas (MET1), erizo de mar, anfibios, peces, pájaros y mamíferos (Goll y Bestor, 2005).

II) DNMT 2 → Las DNMTasas de la familia DNMT2 son las más altamente conservadas a lo largo de la evolución y las más ampliamente distribuidas entre las diferentes especies. Las DNMTasas de la familia DNMT2 son las únicas presentes en algunos organismos, como por ejemplo *D. melanogaster*. La familia de las DNMT2 se caracteriza por la ausencia de cola N-terminal, lo que hace recordar las citosin-metiltransferasas bacterianas. La función de DNMT2 sigue estando poco clara. Experimentos con células madre embrionarias demuestran que mutantes homocigotos

para *Dnmt2* no presentan ninguna anomalía en el patrón de metilación y su fenotipo es completamente normal.

III) DNMT3 → Son las metilasas consideradas *de novo*. En mamíferos, la familia DNMT3 está compuesta por tres miembros: DNMT3A, DNMT3B y DNMT3L. Las dos primeras se expresan en la mayoría de tejidos adultos y metilan indiferentemente secuencias no metiladas y secuencias hemimetiladas en los dinucleótidos CpG (ambas) y CpA (exclusivamente DNMT3A). DNMT3L (*DNMT3 - Like*) es la única DNMTasa que se expresa específicamente en células germinales. No posee actividad metiltransferasa y sin embargo es esencial para el establecimiento del *imprinting* materno en el oocito y la proespermatogonia (Buryanov y Shevchuk, 2005).

Se han descrito otros miembros de la familia DNMT3 en *Arabidopsis thaliana* (familia DRM (*Domains Rearranged Methyltransferases*), responsables de la metilación en secuencias asimétricas) y en *Apis mellifera*. No obstante, *D. melanogaster* y *Anopheles gambiae* no presentan ninguna DNMTasa de la familia DNMT3.

IV) Masc → Familia de DNMTasas específica de hongos. En estas especies la metilación actúa como mecanismo de protección ante la actividad de los transposones y como inhibidor de la transcripción. La metilación no se limita a secuencias simétricas CpG o CpNpG (Buryanov y Shevchuk, 2005).

V) Cromometilasas → Familia de DNMTasas exclusiva de plantas. Son responsables de la metilación de las secuencias CpNpG, algo poco frecuente en vertebrados (Buryanov y Shevchuk, 2005).

1.1.4. La metilación del DNA en *D. melanogaster*

La descripción de secuencias metiladas en *D. melanogaster* es muy reciente, pues anteriormente se creía que no metilaba el DNA. Según los datos bibliográficos, la metilación se produce mayoritariamente de forma asimétrica en los dinucleótidos CpT y CpA, gracias a la única DNMTasa descrita hasta el momento en esta especie, la dDNMT2. Los mutantes para *dDnmt2* pierden totalmente la metilación del DNA y sin embargo son fértiles y presentan un desarrollo y un fenotipo normal (Kunert *et al.*,

2003). Recientemente se ha demostrado que dDNMT2 participa también en la metilación de tRNAs (Goll *et al.*, 2006) y se ha propuesto una actividad dual DNA-RNA metiltransferasa para dicha enzima (Jeltsch *et al.*, 2006).

El papel de la metilación en *D. melanogaster* no está claro. Sin embargo, la existencia de la proteína MBD2/3 con un dominio MBD sugiere que la metilación del DNA desempeña un papel biológicamente importante en esta especie. Los mutantes nulos para el gen *MBD2/3* son viables, fértiles y supresores dominantes de la variegación por efecto de posición (PEV) (ver I.3. El gen *white*). Además presentan una tasa de defectos en la segregación de cromosomas muy elevada durante las primeras fases del desarrollo (Marhold *et al.*, 2004). Otros estudios en *D. melanogaster* han descrito una relación bidireccional entre la metilación del DNA y la metilación de las colas de las histonas (Kunert *et al.*, 2003). Hasta el momento, todos los datos apuntan a que la metilación del DNA en *D. melanogaster* esté asociada con estructuras y conformaciones de la cromatina específicas (Marhold *et al.*, 2004).

Se ha propuesto la posible existencia de islas CpG, que actuarían regulando la transcripción del mismo modo que ocurre en vertebrados. Se ha descrito la represión del gen *Rbf* debido a la hipermetilación de las islas CpG de su promotor (Ferres-Marco *et al.*, 2006). Sin embargo, la metilación preferencial de dinucleótidos CpA y CpT pone en entredicho esta propuesta.

1.2. Código histónico

Las modificaciones post-traduccionales de las histonas están estrechamente relacionadas con la regulación de la transcripción. Se ha propuesto la existencia de un código histónico, según el cual diferentes combinaciones de modificaciones de las colas amino-terminales de las histonas afectarían de forma específica, sinérgica o antagónicamente, las interacciones de dichas histonas con otras proteínas asociadas a la cromatina. De este modo, se vería también modificada la accesibilidad al DNA subyacente. Este “código” se transmitiría de cada célula a sus descendientes en forma de “memoria epigenética” de los programas transcripcionales (Jenuwein y Allis, 2001). Así, el código histónico no se vería implicado exclusivamente en la regulación de la transcripción sino también en otros procesos celulares como el control del ciclo celular y la replicación y reparación del DNA. Por ejemplo, la fosforilación de la serina (Ser)

10 de la histona H3 es esencial para la condensación de los cromosomas y por tanto, para la progresión del ciclo celular, tanto en mitosis como en meiosis (Nowak y Corces, 2004).

Las modificaciones de las colas de las histonas son reversibles. Por ejemplo, las fosforilaciones pueden ser revertidas por fosfatasas, las acetilaciones por desacetilasas y las metilaciones por desmetilasas. La posibilidad de revertir las modificaciones dota a la célula de mecanismos para adaptar rápidamente sus patrones de expresión génica a los cambios que puedan producirse (Shi y Whetstine, 2007).

Las modificaciones interactúan unas con otras, es decir son interdependientes. Por ejemplo, la fosforilación de la Ser 10 de la histona H3 inhibe la metilación de la Lys 9 y actúa sinérgicamente con las acetilaciones de las Lys 9 y 14 y la metilación de la Lys 4. La regulación de la transcripción se basa en las interacciones establecidas entre las distintas modificaciones (Cheung *et al.*, 2000; Rea *et al.*, 2000).

Cada uno de los cuatro tipos de histonas que forman el octámero del nucleosoma, presenta más de 30 residuos que pueden ser modificados post-traduccionamente. Los **acetilaciones** de las lisinas neutralizan su carga, debilitando así la interacción de los extremos N-terminales con el DNA y permitiendo una estructura de la cromatina más abierta y accesible (Strahl y Allis, 2000). La acetilación de las histonas se asocia con la activación de la transcripción, la regulación del ciclo celular, la reparación del DNA, la recombinación y la apoptosis (Sterner y Berger, 2000; Kurdistani y Grunstein, 2003). La **metilación** de las histonas *in vivo* se produce principalmente en seis Lys (H3K4, 9, 27, 36, 79 y H4K20) y en algunas Arg (ver Figura I.2.), regulando tanto la transcripción como la respuesta al daño génico y la formación de la heterocromatina. Las Lys pueden ser mono-, di- o trimetiladas, mientras que las Arg sólo pueden ser mono- o dimetiladas simétrica o asimétricamente (Shi y Whetstine, 2007). Los cinco tipos de histonas son susceptibles a la **fosforilación** en los residuos de Ser y Tre (ver Figura I.2.). Sin embargo, *in vivo* la fosforilación se produce principalmente en la Ser 10 de la histona H3 durante la interfase activando transcripcionalmente una serie de genes (ej. *Fos*, *Jun* y *heat-shock*), la mitosis (participando en la condensación de los cromosomas metafásicos) y la meiosis (participando generalmente en la condensación de los cromosomas) (Prigent y Dimitrov, 2003; Nowak y Corces, 2004). Además, la

fosforilación de la histona H2A participa en la reparación del DNA (Peterson y Laniel, 2004). La **ubiquitinación** de las histonas se produce generalmente en los residuos de Lys de las histonas H2A y H2B, aunque las Lys de las histonas H3 y H1 también pueden ser ubiquitinadas (ver Figura I.2.). En *Drosophila* y vertebrados las histonas H2A y H2B ubiquitinadas se encuentran en la región 5' no codificante de genes con niveles de expresión altos, y por tanto se asocian con la activación transcripcional. La **sumoilación** es la adición de un grupo SUMO (*Small Ubiquitin MOdifier*) y es un proceso antagonista a la acetilación. En levaduras las cuatro histonas que forman el octámero pueden ser sumoiladas y los principales residuos afectados por esta modificación son H2BK6, 7, 16, 17, H2AK126 y las cinco lisinas del extremo N-terminal de H4. Las principales modificaciones en las histonas de mamífero se resumen en la siguiente figura:

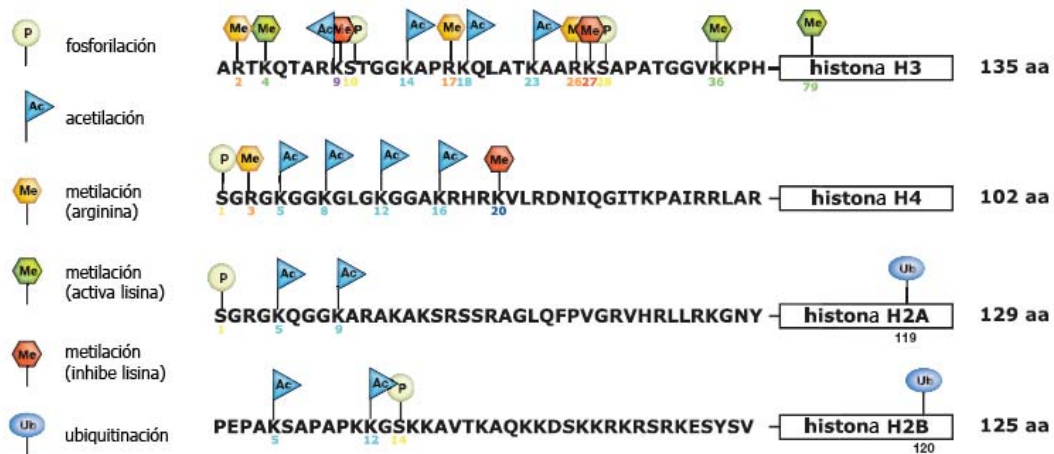


Figura I.2 - Principales modificaciones post-transcripcionales de las histonas en mamíferos. (Adaptada de Allis *et al.*, 2007).

La existencia de tantos tipos distintos de modificación y de tantos residuos modificables, hacen que el código histónico tenga una gran potencialidad y una elevada versatilidad para regular todos aquellos procesos en los que interviene.

1.3. Remodelación de la estructura de la cromatina

El alto nivel de compactación de la cromatina, hace que el DNA no sea accesible a proteínas no histonas importantes, como por ejemplo diversas enzimas que participan en los procesos de replicación, reparación, recombinación y transcripción del DNA. Para que estas proteínas puedan interactuar con el DNA es necesario que se produzcan

cambios en la estructura de la cromatina, proceso que se conoce con el nombre de remodelación de la cromatina. Así, la cromatina relaja su estructura y expone determinadas regiones del DNA a las proteínas que lo requieren. Pero ese proceso debe ser capaz de revertirse, es decir, de compactar las fibras de cromatina, en los periodos de inactividad.

Los complejos remodeladores de la cromatina pueden actuar de distintas formas: catalizando la movilización y reposicionamiento de nucleosomas, transfiriendo los octámeros a posiciones distintas del DNA o generando estructuras similares a dinucleosomas a partir de mononucleosomas para generar una torsión superhelicoidal en el DNA. Se trata de proteínas o complejos multiproteicos que utilizan la energía de la hidrólisis del ATP para modificar la estructura de la cromatina, y por tanto, todos ellos poseen una subunidad ATPasa que pertenece a la familia de ATPasas SNF2. Los miembros de la familia SNF2 pueden dividirse en subfamilias en función de los motivos proteicos situados al entorno del dominio ATPasa (ver Figura I.3) (Lusser y Kadonaga, 2003):

- Subfamilia SWI/SNF2 – Sus miembros contienen un bromodominio en el extremo C-terminal (ver Figura I.3.), con afinidad para la unión a Lys acetiladas y otros dos dominios conservados de función desconocida llamados dominio 1 y 2 respectivamente (de la Cruz *et al.*, 2005). Actúan sobre las interacciones DNA-histonas. Algunos de sus miembros más destacados son: la subunidad Swi2/Snf2 del complejo SWI/SNF y la subunidad Sth1 del complejo RSC en levaduras, la subunidad Brahma del complejo SWI/SNF en *Drosophila*, la subunidad BRG-1 y hBRM del complejo SWI/SNF en humanos.

- Subfamilia ISWI (*Imitation SWI*) – Sus miembros contienen un dominio SANT (SWI3, ADA2, N-CoR, IFIIB) y un dominio SLIDE (ver Figura I.3.). El dominio SANT es el responsable de dirigir o estabilizar la interacción cromatina-complejo remodelador. Además el dominio SANT también participa en el reclutamiento de otros complejos remodeladores y en la interacción entre histonas y enzimas modificadoras de éstas (de la Cruz *et al.*, 2005). El dominio SLIDE presenta carga total positiva y gran similitud con el dominio de unión al DNA de c-Myc, lo que sugiere que podría estar implicado en la unión al DNA (Grune *et al.*, 2003).

- Subfamilia CHD1 – Sus miembros contienen un cromodominio, capaz de reconocer Lys metiladas e involucrado en el control de las interacciones proteína-proteína necesarias para regular la estructura de la cromatina y controlar la expresión génica (de la Cruz *et al.*, 2005); y un dominio de unión al DNA (ver Figura I.3). Algunos de sus miembros se han encontrado en complejos con actividad desacetiltransferasa de histonas (HDAC).

- Subfamilia INO80 – Se caracteriza por la presencia de dominios ATPasa partidos (ver Figura I.3.) y por formar parte de complejos que contienen dos proteínas del tipo RuvB (Rvb1 y Rvb2).

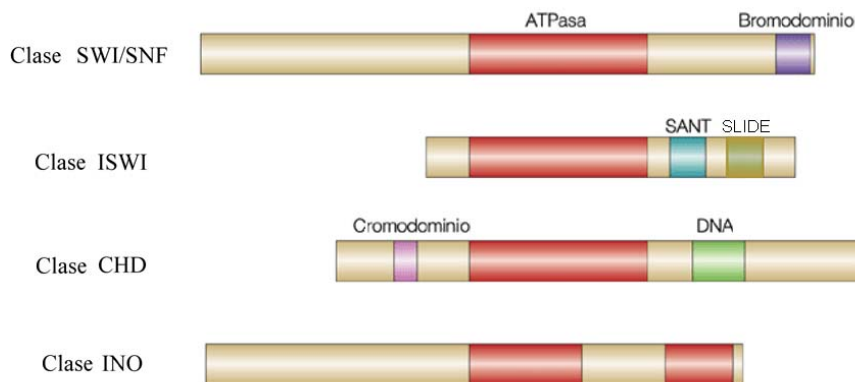


Figura I.3. – **Las cuatro principales subfamilias de ATPasas SNF2:** SWI/SNF2, ISWI, CHD e INO. El rectángulo rojo representa en todos los casos el dominio ATPasa. El rectángulo verde marcado como DNA representa un dominio de unión al DNA.

- Subfamilia CSB – Se caracteriza por un dominio ácido, una región rica en glicinas y señales putativas de localización nuclear. Es capaz de enrollar y desenrollar activamente el DNA, creando una torsión en éste al hacer rotar la hélice. Esta rotación afectaría a la interacción DNA-histonas y permitiría participar en procesos como la regulación de la transcripción, la reparación del DNA asociada a transcripción (TCR – *Transcription Coupled Repair*) y la reparación del DNA por escisión de nucleótidos (NER – *Nucleotid Excision Repair*) (Beerens *et al.*, 2005).

- Subfamilia RAD54 – Los miembros más destacados de la subfamilia RAD54 son: Rad54 (involucrado en la recombinación homóloga), ATRX y ARIP4 (interacciona con el receptor de andrógeno).

- Subfamilia DDM1 – Algunos miembros destacados de la subfamilia DDM1 son: DDM1 (*Decreased DNA Methylation* - Se cree que actúa modificando la estructura de la cromatina para hacerla accesible a las DNMTasas, participando así en el mantenimiento de la metilación del DNA y la estabilidad genómica) (Brzeski y Jerzmanowski, 2003) y Lsh (*Lymphoid Specific Helicase* – Participa en la metilación del DNA y es necesaria para la formación de la heterocromatina) (Yan *et al.*, 2003).

Todos los miembros de la subfamilia SNF2 participan en la remodelación de la cromatina. Sin embargo, dicha remodelación puede llevarse a cabo a través de distintos mecanismos, específicos de cada subfamilia:

1 – Interrupción de las interacciones DNA-histonas por parte de los miembros de la subfamilia SNF2.

2 – Movilización o deslizamiento de los nucleosomas. En determinados casos los nucleosomas pueden desplazarse en *cis* por el DNA, dejando expuestas secuencias que antes se encontraban entorno al octámero de histonas (ver Figura I.4.) (Becker, 2002).

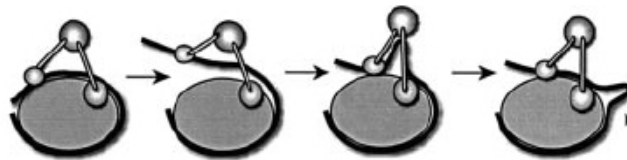


Figura I.4. – **Deslizamiento de nucleosomas.** Los óvalos grises representan el octámero de histonas. El complejo remodelador de la cromatina está representado por las esferas unidas con palos. El complejo se anclaría al DNA y al nucleosoma, haciendo que éste último se desplazara a lo largo de la hebra de DNA (Becker, 2002).

3 – Transferencia de nucleosomas de una cadena de DNA a otra por parte de los miembros de la subfamilia SNF2.

4 – Desensamblaje de nucleosomas, eliminando así el superenrollamiento de la estructura de la cromatina, gracias a los miembros de la subfamilia SNF2; o ensamblaje de nucleosomas gracias a los miembros de la subfamilia ISWI.

5 – Formación de dinucleosomas a partir de mononucleosomas, alterando así las interacciones DNA-histonas, mediante los miembros de la subfamilia SNF2 (Schnitzler *et al.*, 1998).

6 – Creación de una torsión superhelicoidal en la estructura del DNA. Dicha torsión puede causar una rotación del DNA en la superficie del octámero de histonas o incluso puede alterar formas de empaquetamiento superiores de la cromatina (Havas *et al.*, 2000).

7 – Translocación a lo largo del DNA. Los complejos remodeladores formados por miembros de la subfamilia ISWI pueden translocarse a lo largo de la hebra de DNA interrumpiendo las interacciones DNA-histonas. Una vez ha pasado el complejo, las interacciones se restablecen, dejando al octámero en una posición distinta a la inicial (ver Figura I.5.).

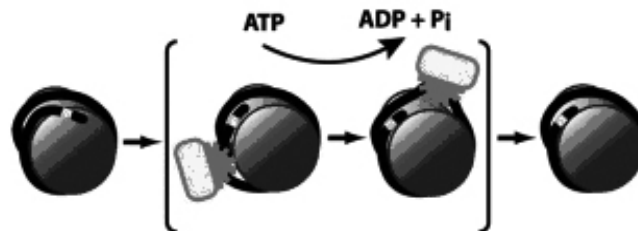


Figura I.5. – **Modelo de remodelación de la cromatina por translocación a lo largo del DNA.** El complejo remodelador se transloca a lo largo de la hebra de DNA interrumpiendo las interacciones DNA-histonas. Posteriormente las interacciones se restablecen y el octámero queda en una posición distinta a la inicial (Lusser y Kadonaga, 2003).

Los miembros de la superfamilia SNF2 no actúan de forma independiente, sino que siempre se encuentran formando parte de los llamados complejos remodeladores de la cromatina. Dichos complejos contienen proteínas sin actividad ATPasa cuya función es: regular la actividad ATPasa (ej. En el complejo remodelador hSWI/SNF, la actividad de la ATPasa hBRM se ve potenciada por la acción de las subunidades INI1, BAF155 y BAF170 de ese mismo complejo), interactuar con factores de transcripción (ej. La subunidad NuRF301 del complejo NuRF crea una interacción específica de secuencia con factores de transcripción), interactuar con el DNA (ej. La subunidad Acf1 del complejo ACF posee un dominio de interacción con el DNA), etc. Los principales complejos remodeladores descritos son (Fry y Peterson, 2001): *y*SWI/SNF, *y*RSC, ISW1 y ISW2 en levaduras; dSWI/SNF (Brahma), dACF, dNuRF (*Nucleosome-Remodeling Factor*) y dChrAC (*Chromatin Accessibility Complex*) en *Drosophila*; y hSWI/SNF, hNuRD y hRSF en humanos. Los principales complejos remodeladores de la cromatina descritos en levaduras, *Drosophila* y humanos, así como las homología presentes entre éstos, se muestran en la Figura I.6..

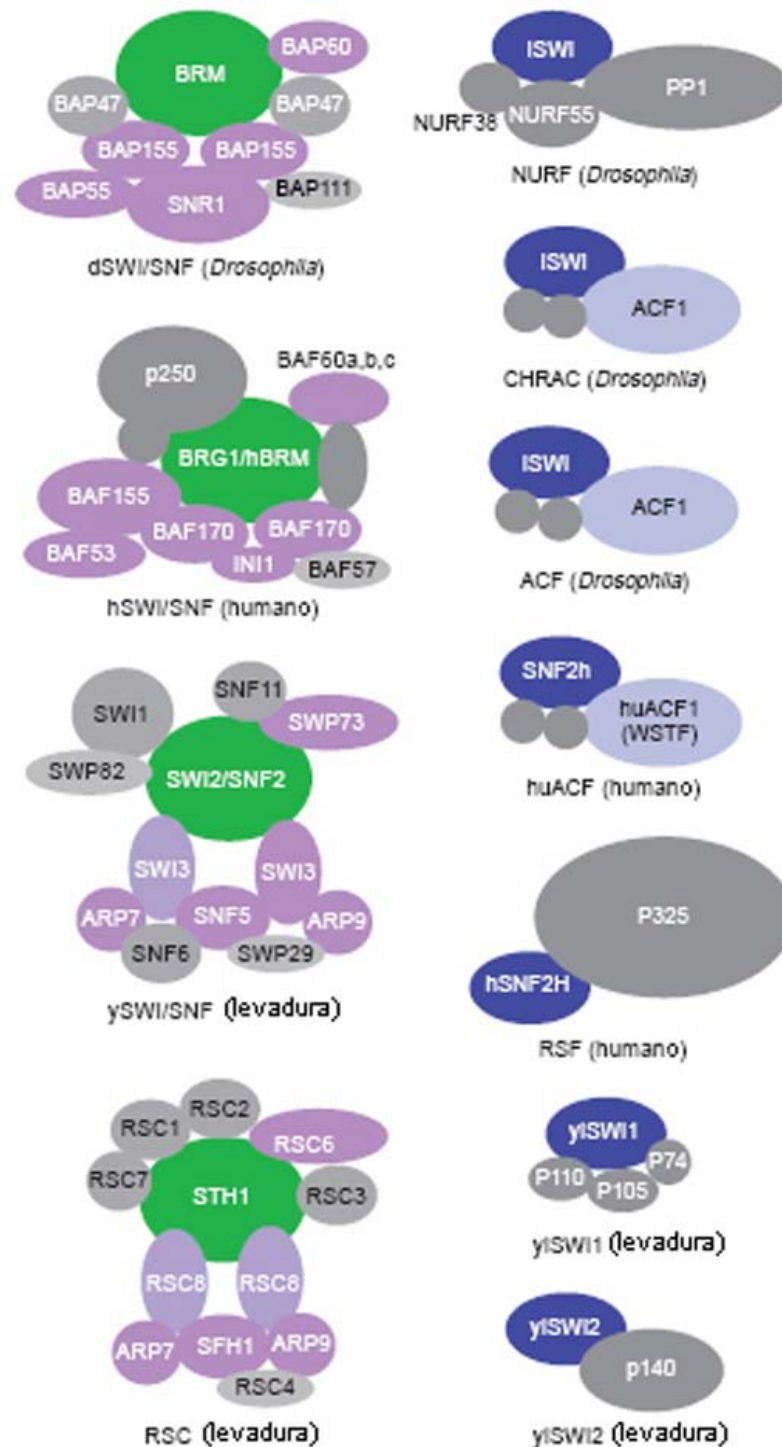


Figura I.6. – Principales complejos remodeladores de la cromatina en levaduras, *Drosophila* y humanos. Los óvalos verdes representan las subunidades ATPasa de los complejos SWI/SNF, mientras que los óvalos azules representan las subunidades ATPasa de los complejos ISWI. Los óvalos violetas representan subunidades que se encuentran conservadas en todos los complejos SWI/SNF y los azul claro, en algunos complejos ISWI. Los óvalos grises representan subunidades no conservadas en los complejos remodeladores descritos hasta el momento (Adaptada de Aalfs y Kingston, 2000). ATENCIÓN: No se representan todas las subunidades de cada complejo.

1.3.1. El complejo remodelador de la cromatina SWI/SNF o Brahma

Los genes SWI/SNF fueron identificados inicialmente en *Saccharomyces cerevisiae* como reguladores positivos de los genes *HO* y *SUC2*. Posteriormente se demostró que tanto *Drosophila* como los humanos poseían complejos remodeladores de la cromatina evolutivamente relacionados con el complejo SWI/SNF de levaduras. En *A. thaliana*, *Caenorhabditis elegans*, *Xenopus laevis*, pollo y ratón también se han descrito complejos similares, demostrándose así una gran conservación de su actividad a lo largo de la evolución. En general, los complejos remodeladores de la cromatina SWI/SNF contienen alrededor de 10 subunidades y su masa es de unos 2MDa.

En *Drosophila* el complejo remodelador de la cromatina SWI/SNF recibe el nombre de Brahma (BRM) y fue identificado en un rastreo (*screening*) para encontrar supresores dominantes de las mutaciones *Polycomb* (*Pc*) (Tamkun *et al.*, 1992). Es un complejo multimérico, formado por 10 subunidades y otros factores asociados más débilmente (Kal *et al.*, 2000). Entre las principales subunidades cabe destacar (ver Figuras I.6. y I.12.):

- *brahma* → Es la subunidad ATPasa del complejo. Es muy abundante en algunas fases del desarrollo y sus mutantes presentan defectos homeóticos pero también defectos en el sistema nervioso periférico y viabilidad reducida de los discos imaginales. Los individuos homocigotos para mutaciones en *brm* mueren en etapas embrionarias avanzadas sin defectos en el patrón corporal. Pertenece al grupo de genes *trithorax* (*trx*) (ver I.2.2. El grupo *trithorax*).

- *moira* → O BAP155. Esta subunidad es la principal diana para *zeste*, a la cual se une mediante su cremallera de leucina (Leu) (ver I.2.2.1. El gen *zeste*). Es homóloga a la subunidad SWI3 en levaduras y a BAF170 y BAF155 en humanos. Pertenece al grupo de genes *trx* (ver I.2.2. El grupo *trithorax*).

- *osa* → O p300. Contiene un dominio ARID de unión al DNA y se propone que podría reclutar el complejo BRM en los genes *antennapedia* entre otros (Vázquez *et al.*, 1999). Es homóloga a la subunidad SWI1 en levaduras y a BAF250/p270 en humanos. Pertenece al grupo de genes *trx* (ver I.2.2. El grupo *trithorax*).

- *snr1* → O BAP45. Esta subunidad interacciona con la ciclina E (*DmcyceE*) *in vivo*, lo que sugiere que podría estar implicada en regular la actividad del complejo BRM sobre el control del ciclo celular (Muchardt y Yaniv, 1999; Brumby *et al.*, 2002). Es homóloga a la subunidad SNF5 en levaduras y a INI1 en humanos. Pertenece al grupo de genes *trx* (ver I.2.2. El grupo *trithorax*).

- BAP60 → No es necesaria para remodelar la cromatina *in vivo*, sin embargo, se ha conservado a lo largo de la evolución desde las levaduras (SWP73) hasta los humanos (BAFa/b/c). Participa tanto en la activación como en la inactivación de la transcripción y es la subunidad encargada de reclutar el complejo BRM a determinados promotores específicos a través de la interacción con factores de transcripción (ej. *sisA* y *sc*) (Moller *et al.*, 2005).

2. Los grupos *Polycomb* y *trithorax*

Los genes de los grupos *Pc* (*PcG*) y *trx* (*trxG*) se encuentran conservados a lo largo de toda la escala evolutiva, desde las levaduras hasta los humanos. Algunos experimentos incluso han demostrado que cDNAs de ratón o humano son capaces de rescatar parcialmente el fenotipo provocado por mutantes del gen homólogo en *Drosophila*, demostrando que existe una conservación de la funcionalidad entre las proteínas (Muyrers-Chen *et al.*, 2004).

Los genes del *PcG* fueron descubiertos como represores y activadores de la expresión de los genes homeóticos en fases determinadas del desarrollo. Estudios posteriores han demostrado que las proteínas de los *PcG* y *trxG* forman complejos multiméricos no necesarios para iniciar la regulación de los genes homeóticos si no para mantener los patrones de expresión establecidos una vez desaparecen del embrión los factores reguladores iniciales. Así, los genes del *PcG* codifican proteínas capaces de generar estructuras parecidas a la heterocromatina, impidiendo a los factores de transcripción u otras proteínas acceder a los genes regulados. La represión creada por los genes del *PcG* puede ser contrarestanda por los genes del *trxG*.

Además de participar en la regulación de la expresión génica, las proteínas de los *PcG* y *trxG* intervienen en la regulación de la proliferación celular, en la diferenciación de las células madre, en procesos cancerígenos, en fenómenos de *imprinting* y en la inactivación del cromosoma X (Schuettengruber *et al.*, 2007).

Algunos miembros de los *PcG* y *trxG* poseen actividad metiltransferasa específica para las Lys de la histona H3, mientras que otros miembros son capaces de interpretar estas marcas histónicas.

2.1. El grupo *Polycomb*

En *Drosophila*, las proteínas del *PcG* se clasifican en tres grandes grupos según el complejo del que forman parte. Los dos principales (PRC1 y PRC2) se muestran en la siguiente figura:

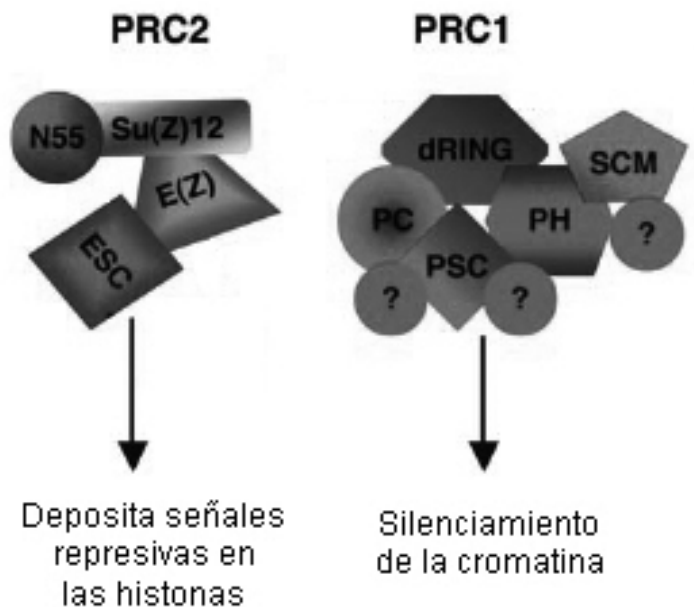


Figura I.7. – **Complejos represores *Pc***. Se muestran las principales subunidades de los complejos PRC2 y PRC1. Los interrogantes representan posibles TAFs. (Adaptada de Grimaud *et al.*, 2006).

Complejo <i>Pc</i>	Proteína <i>Pc</i> en <i>Drosophila</i>	Características a destacar	Proteína homóloga en humanos
PhoRC	dSfmbt	Dominios MBT que se unen a las lisinas mono- y dimetiladas H3K9 y H4K20.	--
	Pho	Dedo de zinc que se une secuencias específicas de DNA. Interacciona con Pc.	--
PRC2	E(z)	Metiltransferasa de histonas. Su dominio SET metila H3K27.	EZH2
	Esc	Repeticiones WD40.	EED
	Su(z)12	Dedos de zinc. Caja VEFS.	SUZ12
	Nurf55	Repeticiones WD40. Interacciona con Su(z)12.	RpAp48/RpAp46
PRC1	dRing	Dedo de zinc del tipo RING. Interacciona con Pc.	RING1A
	Pc	Cromodominio. Se une a H3K27Me3 y H3K9Me3.	HPC1-3
	Ph	Dedos de zinc. Dominio SAM que permite la interacción con Scm.	HPH1-3
	Psc	Dedos de zinc. Dominio HTH. Interacción con Pc y Ph.	BMI1
	Scm	Dominio SAM que permite la interacción con Ph.	SCMH1-2
	TAF	Interaccionan con Pc, Ph y Psc.	--

Tabla I.1 – **Componentes centrales de los tres complejos Pc descritos en *D. melanogaster*.** (Adaptada de Schuettengruber *et al.*, 2007). PRC – *Polycomb Repressive Complex*. dSfmbt – Gen relacionado con *Scm* con dominios MBT (*Malignant Brian Tumor*). Pho – *Pleiohomeotic*. E(z) – *Enhancer of zeste*. Esc – *Extra sex combs*. Repeticiones WD – Dominios de interacción proteína-proteína. Su(z)12 - *Suppressor of zeste 12*. Caja VEFS – Caja VRN2-EMF2-FIS2-Su(z)12. Dominio RING – Dominio de interacción proteína-proteína. Pc – *Polycomb*. Ph – *Polyhomeotic*. Dominio SAM – Dominio *Sterile Alpha Motif*. Dominio de interacción proteína-proteína. Psc – *Posterior sex combs*. Dominio HTH – Dominio *Helix Turn Helix*. Scm – *Sex comb on midleg*. TAF – *TBP Associated Factors*. TBP – *TATA Binding Proteins*.

Los complejos PRC1 y PRC2 no poseen dominios de unión al DNA. Sin embargo, Pho y PhoL (*Pho Like*) se unen a unidades de PRC2 e inducen su reclutamiento en los elementos de respuesta a Pc (PRE – *Pc Response Element*) del gen *ultrabithorax* (*Ubx*). Hasta hoy los PREs sólo se han caracterizado en *Drosophila*. Se definen como elementos de DNA necesarios y suficientes para el reclutamiento de los complejos *Pc* y para el silenciamiento de los promotores asociados a ellos. El modelo de reclutamiento de los complejos *Pc* propuesto sugiere que el proceso se iniciaría con el reconocimiento de secuencias específicas por parte de Pho y PhoL. Éstas a su vez reclutarían a PRC2 que trimetilaría la H3K27, dejando una marca que sería reconocida por PRC1. Sin embargo, Pho no sólo es capaz de interactuar con PRC2, si no que *in vitro* interactúa con Pc y Ph (subunidades de PRC1). Además, la presencia de secuencias de unión de Pho sola es incapaz de reclutar los complejos *Pc*. Pho también interactúa con el complejo remodelador de la cromatina INO80, lo que sugiere que puede participar en procesos independientes de *Pc*. Los mutantes de *Drosophila* para el gen *pho* y *phoL* son letales en estadios tardíos del desarrollo, pero en sus cromosomas politénicos los sitios *PcG* tiñen normalmente, demostrando que los complejos *Pc* pueden ser reclutados por otros medios (Schuettengruber *et al.*, 2007).

Se cree que los genes *trl* (*trx like* – Codifica el factor GAGA) y Pipsqueak (*psq*) podrían estar involucrados también en el reclutamiento de los complejos *Pc*, aunque no se encuentran conservados en humanos. GAF interactúa con *Pc* y potencia la capacidad represora de PRC1 en moldes que contengan secuencias GAGA (Mulholland *et al.*, 2003). *Psq* forma parte de una variante del complejo PRC1 llamada CHRASCH aislada en células S2 de *Drosophila* (Huang y Chang, 2004).

También se piensa que los smRNAs (*small RNA*) y la maquinaria del RNAi pueden estar involucrados en el reclutamiento de los complejos del *PcG*. Recientemente se ha demostrado que en humanos, el gen *Ago1* consigue silenciar la transcripción de algunos genes a través de siRNAs (*small interfering RNAs*) mediante el reclutamiento de EZH2 (Kim *et al.*, 2006).

Se han propuesto diferentes mecanismos a través de los cuales los complejos *Pc* pueden reprimir la transcripción:

- Sin intervención de PRC1 → PRC1 podría ser sustituido por otros factores (ej. LHP1 - *Like Heterochromatin Protein 1*). Sin embargo también podría ser que la acción de PRC2 al trimetilar H3K27 fuera suficiente, e inhibiera algún paso necesario para la activación transcripcional (ej. ubiquitinación de H2B o trimetilación de H3K4). La transcripción de RNAs no codificantes (ncRNAs – *Non Coding RNAs*) también podría inhibir la unión de algunos TAFs y reprimir así la transcripción sin necesidad de PRC1 (ver Figura I.8. A).

- Con intervención de PRC1 → PRC1 puede reprimir la actividad remodeladora de la cromatina de los complejos SWI/SNF *in vitro*. También es capaz de condensar la cromatina en ausencia de modificaciones de histonas. Los TAFs de PRC1 sugieren que PRC1 está en contacto con los promotores de algunos genes. Así, PRC1 aunque no necesariamente inhiba el reclutamiento de la RNA polimerasa II, podría interferir en la separación de las hebras de DNA durante el inicio de transcripción (ver Figura I.8. B).

- Formación de un bucle → La posición de los PREs respecto a los promotores es variable: pueden sobreponerse o pueden encontrarse separados por varias Kb. Un modelo para la inhibición de la transcripción a través de PREs distantes propone que E(z) se uniría a los PREs creando grandes dominios de H3K27 trimetiladas. Esta marca silenciaría todos los promotores que se encontraran dentro del bucle. En humanos EZH2 es capaz de reclutar DNMTasas. Estudios recientes han demostrado que los genes con proteínas del *PcG* asociadas son unas de las principales dianas para la metilación *de novo*. Además las proteínas del *PcG* que se encontraran unidas a un PRE podrían establecer contactos con la maquinaria de transcripción unida en otro PRE, evitando así que se produjera la transcripción (ver Figura I.8. C).

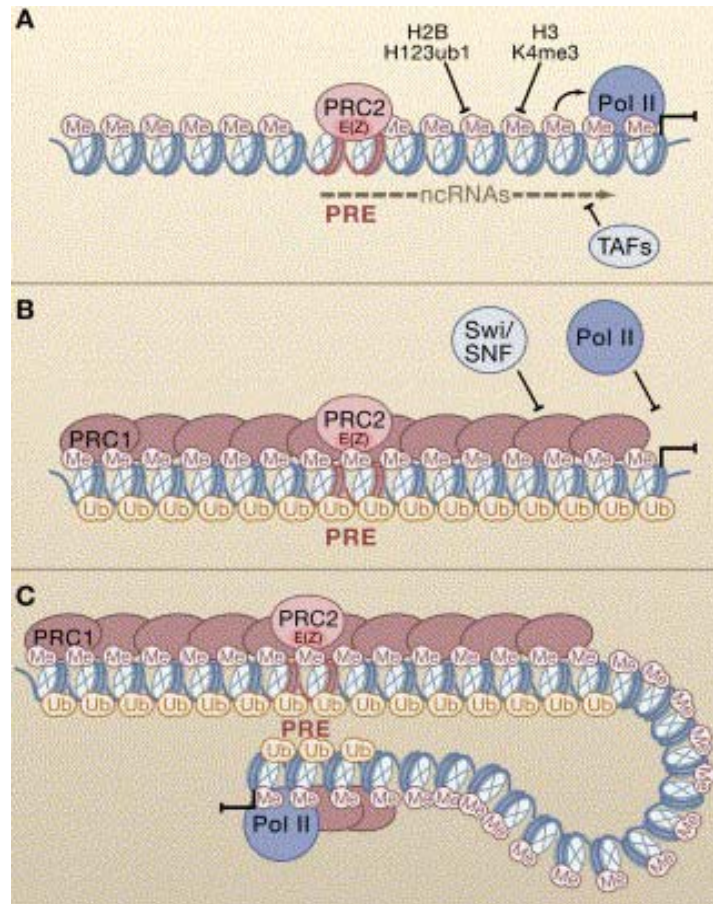


Figura I.8 – Principales mecanismos de inhibición de la transcripción por complejos *Pc*. A) Inhibición de la transcripción sin intervención de PRC1. B) Inhibición de la transcripción mediada por PRC1. C) Inhibición de la transcripción por formación de un bucle a través de E(z). (Schuettengruber *et al.*, 2007).

2.2. El grupo *trithorax*

Los primeros genes del *trxG* fueron identificados como supresores del fenotipo causado por el gen del *PcG* *esc*. Sin embargo, aunque todos los genes del *trxG* suprimen los fenotipos causados por las mutaciones de *PcG*, sólo los alelos de *trithorax* (*trx*), *absent*, *small or homeotic discs 1* (*ash1*) y *absent, small or homeotic discs 2* (*ash2*) provocan transformaciones homeóticas por sí mismos (Shearn, 1989). Estos tres genes serían antagonistas específicos de las funciones de los genes del *PcG* y el resto de genes del *trxG* serían activadores o co-activadores transcripcionales generales.

Igual que las proteínas del *PcG*, las proteínas del *trxG* también se agrupan formando complejos. Los principales complejos *trx* son (ver Tabla I.2):

Complejo <i>trx</i>	Proteína <i>trx</i> en <i>Drosophila</i>	Características a destacar	Proteína homóloga en humanos
BRM	Brm	Ver I.1.3.1. El complejo remodelador de la cromatina SWI/SNF o Brahma.	BRM
	Osa		BAF250
	Moirá		BAF170
	Snr1		BAF47
NuRF	ISWI	Actividad ATPasa. Dominio SANT y dominio SLIDE. (Ver I.1.3. Remodelación de la estructura de la cromatina).	SNF2L
	NuRF38	Actividad pirofosfatasa inorgánica.	--
	E(bx)	Interacción específica de secuencia con factores de transcripción.	BPTF
	NuRF55 (Caf1)	Repeticiones WD40.	RpAp46/RpAp48
TAC1	Trx	Dominios SET y PHD. Actividad histona metiltransferasa H3K4.	--
	dCBP	Histona acetiltransferasa. Interacciona con Ash1 y Trx.	--
	Sbfl	Fosfatasa de proteínas.	--
Ash1	Ash1	Dominio SET. Actividad histona metiltransferasa H3K9, H3K4 y H4K20.	--
	dCBP	Histona acetiltransferasa. Interacciona con Ash1 y Trx.	--

Tabla I.2. – **Componentes centrales de los principales complejos *trx* descritos en *D. melanogaster*.** (Adaptada de Schuettengruber *et al.*, 2007). BRM – *Brahma*. BAF – *Brahma Associated Factor*. NuRF – *Nucleosome Remodeling Factor*. ISWI – *Imitation Switch Protein*. SANT - SWI3, ADA2, N-CoR, TFIIB. E(bx) – *Enhancer of bithorax*. Caf1 – *Chromatin Assembly Factor 1*. Repeticiones WD – Dominios de interacción proteína-proteína. TAC – *Trithorax Acetylation Complex*. Trx – *Trithorax*. SET - *Su(var)*, *E(z)*, *Trx*. dCBP – *Drosophila homolog of the CREB Binding Protein*. Sbfl – *SET Binding Factor*. Ash1 - *absent, small or homeotic discs 1*.

Como se puede ver en la Tabla I.2., las proteínas del *trxG* son bastante heterogéneas. Aún así, se podrían clasificar en dos grandes grupos: las que contienen un dominio SET (Trx y Ash1 en *Drosophila* y MLL en vertebrados) y las que forman parte de los complejos remodeladores de la cromatina ATP dependientes (SWI/SNF y NuRF).

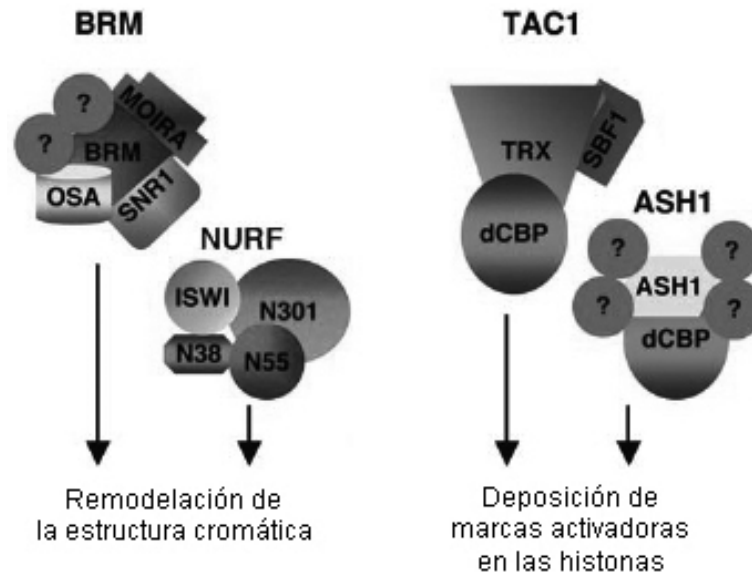


Figura I.9. – **Complejos activadores *trithorax***. Se muestran las principales subunidades de los complejos BRM, NuRF, TAC1 y ASH1. Los interrogantes representan subunidades no caracterizadas. (Adaptada de Grimaud *et al.*, 2006).

El reclutamiento de los complejos *trx* sigue siendo un tema bastante confuso y se conocen distintas vías para llevarla a cabo. En *Drosophila*, Trx se une al elemento de DNA *Fab-7* de las glándulas salivales en ausencia de activación transcripcional. Otros estudios sugieren que un segundo elemento de DNA sería el responsable del mantenimiento de la activación de *Ubx* por Trx. Otros miembros del *trxG* son reclutados directamente en regiones de la cromatina transcripcionalmente activas. Por ejemplo, inmediatamente después de activar la transcripción de *Ubx*, Ash1 es reclutado en 3' del punto de inicio de la transcripción. Otra vía es el reclutamiento de BRM llevado a cabo por Zeste en sus puntos de unión (ZBS – *Zeste Binding Sites*) (Schuettengruber *et al.*, 2007).

El mecanismo preciso de actuación de todos estos complejos no está del todo definido. Lo que si es obvio es que existe una relación e interacción entre los complejos

encargados de depositar marcas activadoras en las histonas y los complejos encargados de remodelar la estructura de la cromatina. La existencia de diferentes fenotipos mutantes dependiendo del complejo afectado sugiere un alto nivel de especialización de dichos complejos. Algunos de ellos (ej. BRM, NuRF) jugarían papeles más generales en el mantenimiento y remodelación de la estructura de la cromatina, mientras que otros (ej. TAC1) tendrían una función más estrictamente antagonista a la de los complejos Pc (Grimaud *et al.*, 2006).

2.2.1. El gen *zeste*

El gen *zeste* (*z*) de *D. melanogaster* pertenece al *trxG* aunque no forma parte de ninguno de los complejos *trx* y no se encuentra conservado en humanos. Se localiza citológicamente en la posición 3A3 del cromosoma 1, según el mapa de los cromosomas politénicos (Lindsley y Zimm, 1992).

La región codificante de *z* está compuesta por 3 exones, separados por dos intrones de 121 y 62bp (ver Figura I.10.). El transcrito principal (*z*-RA) tiene 2398 bases y es muy poco abundante, no superando el 0,01% del total de mRNAs poliA en individuos adultos. Su expresión presenta especificidad temporal, llegando a valores máximos en huevos no fertilizados y durante el desarrollo pupal, y a valores mínimos durante los estadios larvarios (Mansukhani *et al.*, 1988; Pirrotta, 1991).

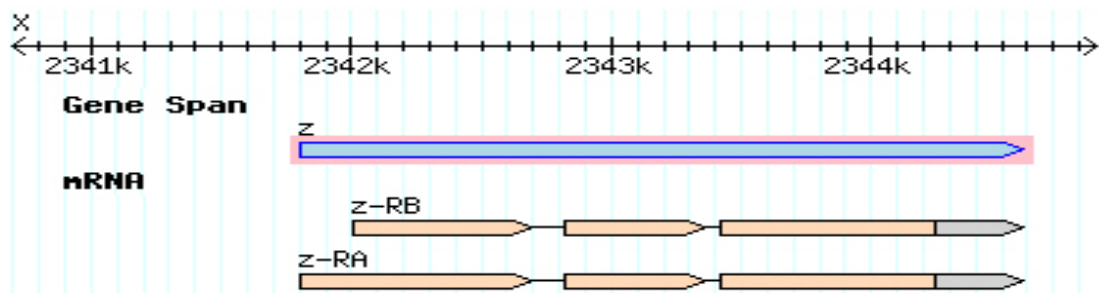


Figura I.10. – Esquema del gen *zeste* de *Drosophila melanogaster*. La barra superior indica el tamaño del gen. La barra azul indica la extensión completa del gen. La representación inferior, indica mediante cajas, la posición de los exones, en los dos transcritos descritos (Adaptada de *FlyBase*).

El gen *z* codifica Zeste (*Z*) una proteína de unión al DNA de 574 aminoácidos y ~62 kDa. Los sitios de unión de *Z* reciben el nombre de *Zeste Binding Sites* (ZBS). Cada ZBS está formado por 4 o más *Zeste Binding Motives* (ZBM) en una región

aproximada de menos de 300bp. La secuencia consenso de los ZBM es YGAGYG (Y = C o T, es decir, una pirimidina) (Chen y Pirrotta, 1993a).

La estructura de la proteína puede dividirse básicamente en tres partes (Pirrotta, 1991; Chen y Pirrotta, 1993a):

- Región N-terminal → Posee un dominio hélice-giro-hélice (*Helix-Turn-Helix* – HTH) de 90 aminoácidos que le permite interactuar con el DNA. Es una región rica en glicinas (Gly).

- Región central → Es la región necesaria para la activación transcripcional. Es una zona rica en tiradas de glutamina (Gln), alanina (Ala) o ambos aminoácidos intercalados. Estos motivos aminoacídicos se asocian a las llamadas repeticiones OPA presentes en muchas proteínas de regulación de *Drosophila* (ej. *Notch*) y en factores de transcripción de vertebrados.

- Región C-terminal → Es la región necesaria para la unión de Z a sus ZBS y para la interacción con otras proteínas. Se divide en dos partes: una región de 42 aminoácidos con seis residuos de Leu, capaces de formar una cremallera de Leu (*leucine zipper*) (Landschulz *et al.*, 1988) y; el dominio 11G3, que contiene la región hidrofóbica FFLLI. Ambas regiones son necesarias para la multimerización de la proteína. La cremallera de Leu participa en la formación de dímeros de Z, mientras que el dominio 11G3 agrega estos dímeros para formar multímeros. Este proceso requiere una secuencia ordenada de eventos, así la formación de dímeros es necesaria para la posterior formación de los multímeros.

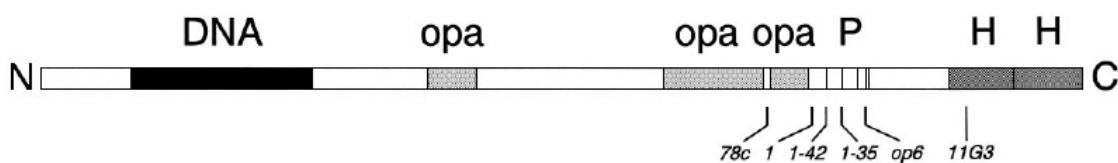


Figura I.11. – **Estructura de la proteína Zeste.** En el extremo N-ter se muestra el dominio de unión al DNA. En la región central se observan las repeticiones opa, motivos conservados en diversas proteínas de regulación de *Drosophila* y en factores de transcripción de vertebrados. En la región C-ter se muestran dos regiones hidrofóbicas (H) involucradas en el proceso de agregación de Zeste. (Adaptada de Rosen *et al.*, 1998).

En los cromosomas politénicos, *Z* se encuentra asociada a más de 60 puntos distintos. Participa en el apareamiento de copias homólogas y en los efectos de transvección de *Ultrabithorax (Ubx)*, *yellow (y)*, *eyes absent (eya)*, *decapentaplegic (dpp)* y *white (w)* (Pirrota, 1991; Rosen *et al.*, 1998).

Zeste es un regulador de la transcripción específico de secuencia, considerado del grupo *Trithorax (trxG)*, puesto que es capaz de activar la transcripción, actuando tanto en *cis* como en *trans*. Su capacidad de transactivar (actuar en *trans*) fue demostrada a través de la PEV (ver I.3. El gen *white*), ya que los mutantes de *z*, especialmente sus alelos nulos, son *enhancers* recesivos de la PEV. De estos experimentos se desprende el actual modelo para explicar la función de *zeste*. Éste propone que la proteína salvaje *Z* actúa conjuntamente con otras proteínas para establecer y estabilizar dominios cromatínicos. Las funciones de dichas proteínas serían redundantes y por eso la ausencia de *Z* no tiene efectos drásticos sobre el individuo (Goldberg *et al.*, 1989). Las hembras sin el gen *z* son normales y fértiles y no presentan transformaciones homeóticas ni anomalías en el desarrollo. La única característica distintiva es un color de ojos marrón, que indica que la expresión del gen *white* no es del todo normal. Cabe resaltar, que los individuos sin el gen *z* aparecen en una frecuencia mucho menor de la esperada. Así, puede que la viabilidad de estos individuos sea más baja de lo normal o bien que nazcan con malformaciones en alas y/o patas y por eso mueran atrapados en el medio. Además los individuos sin el gen *z* también presentan unas altas tasas de segregación de cromosomas anormal. Los individuos homocigotos y heterocigotos para la delección de *z* presentan una clara tendencia a la segregación conjunta de los cromosomas X maternos, obteniéndose así una proporción elevada de individuos XXY (Pirrota, 1991; Judd, 1995).

Zeste activa la transcripción gracias al reclutamiento del complejo BRM en los ZBS. El complejo BRM es un remodelador de la cromatina del tipo SWI/SNF que actúa una vez unido al promotor de los genes (ver I.1.3.1. El complejo remodelador de la cromatina dSWI/SNF o Brahma). La proteína *Z* recluta el complejo BRM a través de interacciones directas con varias de sus subunidades, como las proteínas del grupo *trx* Moira y Osa (ver Figura I.12.). Así, la cremallera de Leu de *Z* es la responsable de interactuar específicamente con la subunidad Moira. La activación transcripcional

mediada por Z es un proceso ordenado. En primer lugar es necesario que Z se una al DNA en los ZBS. Para este primer paso, la presencia del complejo BRM es innecesaria. Una vez Z se encuentra unida a los ZBS se procede al reclutamiento del complejo BRM. Este segundo paso es indispensable para que se produzca la activación transcripcional mediada por Z. En este caso no es necesario el reclutamiento de los remodeladores ISWI, que son totalmente prescindibles (Kal *et al.*, 2000; Dejardin y Cavalli, 2004).

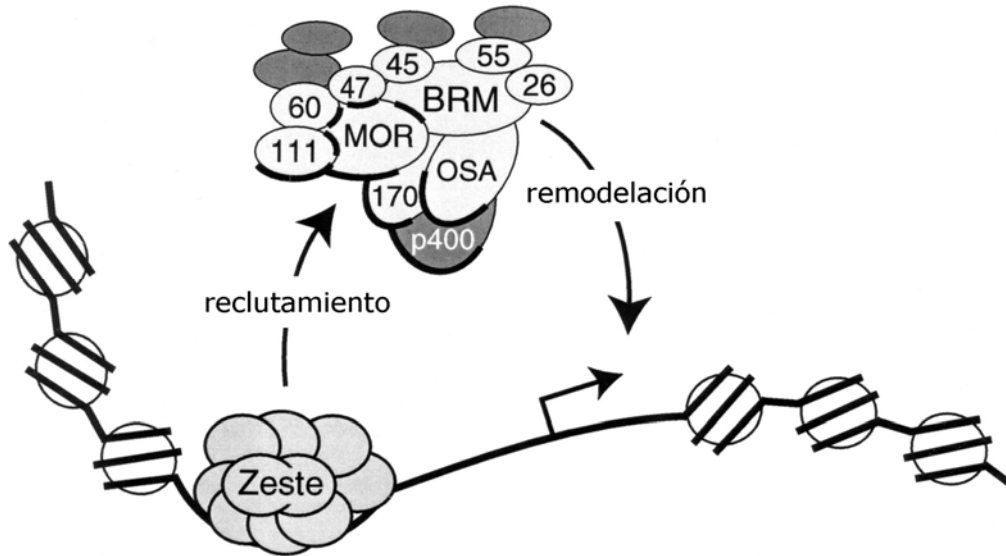


Figura I.12. – **Esquema del reclutamiento del complejo BRM por parte de Z.** El núcleo del complejo BRM está formado por 10 subunidades (en blanco). Además de este núcleo, BRM contiene otros factores (en gris) asociados más débilmente. Zeste se une al DNA e interactúa con BRM a través de diferentes regiones de las subunidades (marcadas en negrita). Una vez unido a Z, el complejo BRM crea una conformación abierta de la cromatina que permite el anclaje de otros factores de transcripción (Adaptada de Kal *et al.*, 2000).

El reclutamiento del complejo BRM en los ZBS genera una reestructuración local de la cromatina. Así, la adquisición de una nueva conformación cromatínica más abierta permite el anclaje de otros coactivadores y también de la maquinaria de transcripción. La relación molar BRM:nucleosoma para que se produzca la activación transcripcional es menor que 1:50 (Kal *et al.*, 2000).

2.2.1.1. La mutación *zeste*¹

El primer mutante descrito para el gen *z* fue el mutante neomórfico *zeste*¹ (*z*¹), que presenta un fenotipo de ojos amarillos siempre que tenga copias apareadas del gen *w* (ver I.3.1. Interacción *zeste-white*) (Jack y Judd, 1979). La secuencia de Z en los mutantes *z*¹ diverge de la salvaje en dos aminoácidos: una sustitución de Lys a

metionina (Met) en la posición 425 (K425M) y una, de Ser a Ala en la posición 496 (S496A). Este último cambio se considera un polimorfismo, ya que también se encuentra presente en algunos individuos salvajes. La sustitución K425M hace que Z pierda el único aminoácido cargado en un dominio central de la proteína de 166 aminoácidos y crea un nuevo dominio hidrofóbico que da propiedades hiperagregantes a la proteína Zeste¹ (Z¹). Gracias a estas nuevas propiedades Z¹ forma agregados más grandes y más estables que Z (Chen y Pirrotta, 1993a).

3. El gen *white*

El gen *white* (*w*) ha sido utilizado como modelo en infinidad de estudios en *D. melanogaster*. Se localiza citogenéticamente en la posición 3C2 del cromosoma X, según el mapa de los cromosomas politénicos (Lindsley y Zimm, 1992).

La región codificante del gen *w* está compuesta por seis exones, el primero de los cuales se encuentra separado de los demás por un gran intrón de 2,9 Kb. Se han descrito tres transcritos distintos para el gen *w*: 1- Transcrito de 2,6 Kb poli-(A)⁺. Transcrito mayoritario presente en embriones, larvas, pupas y adultos (ver Figura I.13.). 2- Transcrito de 2 Kb presente en larvas de tercer estadio y pupas. 3- Transcrito de 2,2 Kb presente en machos. Se considera que podría ser el responsable de la pigmentación gonadal (Fjose *et al.*, 1984; Pirrotta y Brockl, 1984).

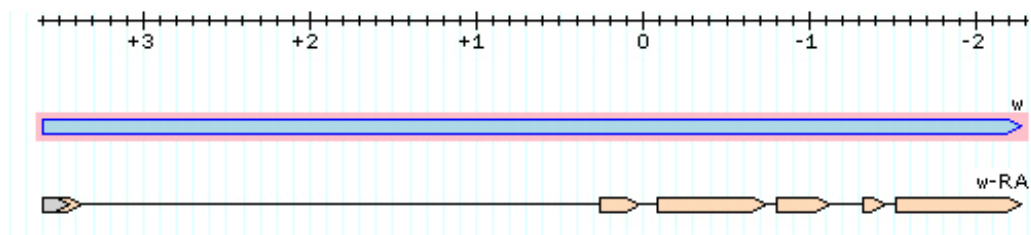


Figura I.13. – **Esquema del gen *white* de *D. melanogaster*.** La barra superior muestra las coordenadas de Levis, en las que el 0 corresponde al punto de inserción del elemento *copia* en el alelo *white apricot* (*w^a*) (Levis *et al.*, 1982). La barra azul indica la extensión completa del gen. La representación inferior, indica mediante cajas, la posición de los exones. El transcrito representado es el mayoritario, de 2,6 KB (Adaptada de *FlyBase*).

Los transcritos de *w* son escasos, aunque se encuentran, en menor o mayor cantidad, a lo largo de todas las fases del desarrollo de *D. melanogaster*. La expresión

del gen *w* muestra especificidad temporal. El mRNA de *w* no supera el 0,0005% del total de mRNAs poli-(A)⁺ en pupas e individuos adultos. Las larvas de tercer estadio son las que presentan una mayor proporción del transcrito de *w*, llegando a valores de 0,003% del total de mRNAs poli-(A)⁺. Es en esta fase del desarrollo cuando se produce un incremento en la tasa de deposición de pigmentos en los túbulos de Malpighi y los discos imaginales anteno-oculares (O'Hare *et al.*, 1983 Hazelrigg, 1987).

El gen *w* codifica una proteína de membrana de aproximadamente 78kDa, cuya función es la deposición de pigmentos en el ojo, los testículos y los túbulos de Malpighi. Su función es ATP dependiente y aún no está claro si transporta los precursores de las pteridinas (pigmentos rojos) y ommocromos (pigmentos marrones), deposita los precursores en los gránulos de pigmento o mantiene la integridad de los gránulos. Los mutantes de *w* no ven afectada su viabilidad ni su fertilidad, pero presentan alteraciones en la cantidad y la distribución de los pigmentos en los ojos, los testículos y los túbulos de Malpighi. La delección del locus resulta en una ausencia total de pigmento. Cabe destacar que la relación entre la actividad de *w* y la coloración detectable no es lineal (O'Hare *et al.*, 1984; Hazelrigg, 1987; Pirrotta y Rastelli, 1994).

El promotor de *w* es relativamente débil y carece de caja TATA. Una serie de cajas reguladoras (*enhancers*) situadas en 5' del gen *w* son las encargadas de modular, de forma bastante compleja, la expresión. Los *enhancers* son los responsables de la especificidad tisular. Algunos de ellos, pueden interaccionar con proteínas para modificar la expresión del gen *w* (Hazelrigg, 1987).

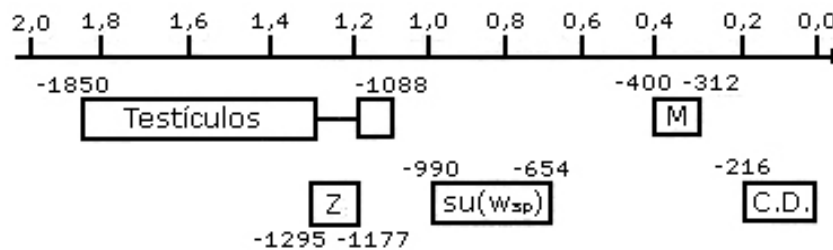


Figura I.14. – **Los principales enhancers de white.** La barra superior indica la distancia desde el punto de inicio de la transcripción, en Kb. Se muestran los *enhancers* responsables de la activación de *w* en testículos, túbulos de Malpighi (M), la región necesaria para la interacción *zeste-white* (Z), la zona donde interactúan los productos del gen *su(w_{sp})* y la caja responsable de la compensación de dosis (C.D.) (Adaptada de Hazelrigg, 1987).

En la región 3' del gen *w* no se ha descrito ninguna secuencia reguladora. Aún así, los individuos con deleciones en esa zona presentan un fenotipo de ojos marrón oscuro. Puesto que generalmente la transcripción se suele alargar unos cien pares de bases a partir de la señal de poliadenilación, se propuso que dicha deleción afecta de alguna manera el procesado y maduración del mRNA. De todas formas, los individuos mutantes con una deleción en 3' no ven afectada su interacción con z^1 ni tampoco la compensación de dosis (Levis *et al.*, 1985).

La regulación de la expresión del gen *w*, como la del resto de genes localizados en el cromosoma X, presenta compensación de dosis. En *Drosophila* los dos cromosomas X de las hembras permanecen activos y los machos consiguen la compensación de dosis gracias a un incremento de la expresión del cromosoma X. Para la compensación de dosis del gen *w* es necesaria su localización en el cromosoma X además de la presencia de múltiples determinantes presentes en la región 5' del gen y en su misma secuencia codificante. Parte de la secuencia requerida para la compensación de dosis se localiza en los 216bp anteriores al inicio de transcripción. Así se consigue que hembras y machos presenten un color de ojos idéntico, aún teniendo las primeras el doble de copias del gen *w* (Hazelrigg *et al.*, 1984; Hazelrigg, 1987; Qian y Pirrotta, 1995).

La expresión del gen *w* también puede verse modificada debido a la PEV. Cuando un gen abandona su posición eucromática original en el genoma y se sitúa en un entorno heterocromático, se puede producir una inactivación parcial del gen. En los casos de PEV, el gen *w* se inactiva en algunas células en un momento concreto del desarrollo tisular, y ese estado es heredado por clones de células descendientes. De ese modo se obtiene un mosaico que presenta los ojos blancos con manchas rojas. Así pues, la inactivación por PEV es un fenómeno arbitrario que puede producirse en algunas células y en otras no. La inactivación es debida a la formación, en zonas heterocromáticas, de complejos proteicos (productos de los genes *Su(var)*). Una vez establecidos, los complejos se pueden expandir a lo largo de la cromatina y llegar a invadir centenares de kilobases eucromáticas (Hazelrigg, 1987; Pirrotta y Rastelli, 1994). La teoría de compartimentalización nuclear sugiere que esta inactivación también podría verse favorecida por un cambio de posición nuclear, que facilitaría la acción de factores silenciadores (Wallrath, 1998).

En general cualquier alteración genética puede causar una desregulación del gen *w*. Se han descrito mutantes fenotípicos de *w* debidos a reordenaciones cromosómicas (ej. *white-mottled*, Tartof *et al.*, 1989), inserciones de elementos transponibles (ej. *white apricot*, Zachar y Bingham, 1982), deleciones (ej. *white-spotted2*, Zachar y Bingham, 1982) y, como en nuestro trabajo, duplicaciones (ej. M115, Badal *et al.*, 2006a). El efecto observado de variación en la expresión del gen *w* en los mutantes M115 es debido a la duplicación del gen *w* (explicado en detalle en el apartado I.6.), afectado así la llamada interacción *zeste-white*.

3.1. La interacción *zeste-white*

La interacción *zeste-white* es la represión de la expresión del gen *w* en los ojos de *D. melanogaster*, debida a un efecto en *trans* producido por la presencia de la proteína mutante Z^1 . La interacción *zeste-white* afecta exclusivamente al *enhancer* de los ojos, por lo que la expresión de *w* en testículos y túbulos de Malpighi no se ve afectada. La región necesaria para la interacción *zeste-white* es el ZBS, compuesto por cinco ZBM localizados entre -1295 y -1177, según las coordenadas de Levis (ver Figura I.14.). El fenotipo z^1 se observa siempre que los individuos presenten dos o más copias apareadas de *w* y por eso se da habitualmente en hembras y no en machos, aunque se han descrito machos con fenotipo z^1 , portadores de duplicaciones del gen *w* o de la región de los ZBS. Del mismo modo se han descrito también hembras z^1 de ojos rojos, en las que una copia del ZBS de *w* ha sido delecionada (Jack y Judd, 1979; Pirrotta, 1991). La producción de mRNA de *w* parece no verse alterada al analizarse muestras de RNA total de individuos enteros (O'Hare *et al.*, 1983; Pirrotta y Brockl, 1984), pero se aprecia una ligera disminución de éste al analizar muestras de RNA total de cabezas (Bingham y Zachar, 1985).

El efecto de z^1 sobre la expresión del gen *w* presenta similitudes con la PEV, puesto que los ojos amarillos típicos del fenotipo z^1 pueden tomar, dependiendo de la temperatura, distintas tonalidades e incluso un fenotipo punteado, con omatídios más oscuros en los que la represión de la expresión de *w* no haya sido tan eficaz. Aún teniendo en cuenta las similitudes fenotípicas, cabe destacar que los dos fenómenos (PEV y interacción *zeste-white*) tienen orígenes moleculares distintos (Rasmuson-Lestander *et al.*, 1993).

El modelo molecular actual explica la interacción *zeste-white* gracias a la formación de hiperagregados de proteína mutante Z^1 en el promotor de *w*. Cuando se encuentran dos copias apareadas del gen *w*, el agregado de Z^1 que se forma es mayor que en el caso de contar con una sola copia de *w*. Así, la represión de la expresión del gen *w* sólo puede ser observada fenotípicamente si se cuenta con copias apareadas de *w* (Chen y Pirrotta, 1993a). Sin embargo, cabe destacar que si se sobreexpresa Z^1 , formando hiperagregados suficientemente grandes aún en posesión de una única copia de *w*, el fenotipo ojos *zeste* puede observarse incluso en machos (Bickel y Pirrotta, 1990). El modelo también propone que los hiperagregados podrían crear interacciones anormales con otras proteínas (ej. BRM) implicadas en la estabilización del dominio cromatínico proximal de *w*, en la región 3C de los cromosomas politénicos o unirse anormalmente a los ZBS (Judd, 1995). Algunas teorías también proponen la participación de genes *Pc* en la represión de *w*, puesto que tres mutantes con pérdida de función de *Psc*, *Scm* y *E(z)* respectivamente, son capaces de suprimir el fenotipo *z*. Aún así, el papel de los genes *Pc* en la interacción *zeste-white* sigue sin estar claramente definido (Wu *et al.*, 1989).

4. El gen *decapentaplegic*

El gen *decapentaplegic* (*dpp*) participa en numerosos procesos del desarrollo de *Drosophila*. Se localiza citogenéticamente en la posición 22F1 del cromosoma 2, en el extremo distal de su brazo izquierdo (Lindsley y Zimm, 1992).

Las mutaciones en *dpp* pueden clasificarse en tres grandes grupos, que mapean en zonas distintas del gen: 1 – Mutaciones *shortvein* (*shv*). Todas ellas muestran defectos recesivos en la formación de la vena longitudinal del ala. 2 – Mutaciones de haploinsuficiencia (*Hin*). Los mutantes presentan un patrón dorso-ventral deficiente. 3 – Mutaciones de disco imaginal (*disk*). Presentan deleciones en el patrón epidérmico de los derivados de los discos imaginales. Dentro de cada tipo de mutaciones se han descrito distintos alelos (Gelbart, 1982). Los mutantes de *dpp* adultos son viables y presentan múltiples defectos epidérmicos en las estructuras derivadas de 15 de los 19 discos imaginales. Por ese motivo el gen recibe el nombre de *decapentaplegic* (Spencer *et al.*, 1982).

El gen *dpp* es una gran unidad genética, de unas 50Kb, formada en su mayor parte por elementos de regulación en *cis*. La región codificante del gen *dpp* mide alrededor de 30Kb y está compuesta por tres exones. Se han descrito cinco transcritos distintos para *dpp*, que miden entre 3,5 y 5Kb y se diferencian en su extremo 5' gracias a usos alternativos de exones (ver Figura I.15.). Aún así, todos los transcritos poseen los dos últimos exones en 3', localizados en la región *Hin*, que son los que contienen el ORF entero (Padgett *et al.*, 1987).

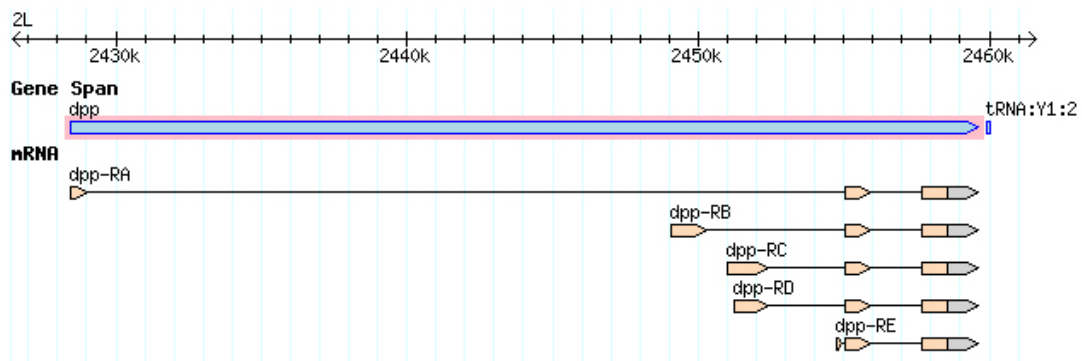


Figura I.15. – Esquema del gen *dpp* de *Drosophila melanogaster*. La barra superior indica el tamaño del gen. La barra azul indica la extensión completa del gen. La representación inferior, indica mediante cajas, la posición de los exones, en los cinco transcritos descritos (Adaptada de *FlyBase*).

Los diferentes transcritos codifican una única proteína de 588 aminoácidos, perteneciente a la familia del factor de crecimiento transformante β (TGF- β – *Transforming Growth Factor* β) (Padgett *et al.*, 1987). En una primera fase, Dpp es responsable de establecer la polaridad dorso-ventral. En una segunda fase, su actividad define los límites de los segmentos embrionarios (*segmental compartments*). Como parte de este proceso, junto con *wingless* y *hedgehog*, define la posición de los futuros apéndices, incluyendo las alas, las patas y las antenas. Dpp también posee una función independiente en la estructuración del mesodermo. En la parte final del proceso de desarrollo de los apéndices Dpp, junto con los productos de *engrailed* y *hedgehog*, define los límites de los compartimentos donde se localizan los apéndices (*appendage compartments*), asegurando una correcta polaridad antero-posterior. Además Dpp presenta una función análoga en el desarrollo del ojo, donde se encarga de la progresión del surco morfogénico, el sitio de inducción de la retina de *Drosophila* (Ferguson y Anderson, 1992; Campbell *et al.*, 1993; Heberlein *et al.*, 1993).

4.1. Interacción *zeste-dpp*

El gen *dpp* se encuentra sujeto a los efectos de transvección, para los que requiere la intervención del gen *z*. La transvección es un tipo inusual de complementación interalélica dependiente de la presencia de copias apareadas del gen en cuestión, en nuestro caso *dpp*. Es decir, dos alelos de un gen pueden complementarse parcialmente gracias a su proximidad física dada por el apareamiento de cromosomas. La transvección en *dpp* se debe a la acción de unos *enhancers* que mediante *z* actúan en *trans* favoreciendo la expresión de *dpp* (ver Figura I.16.).

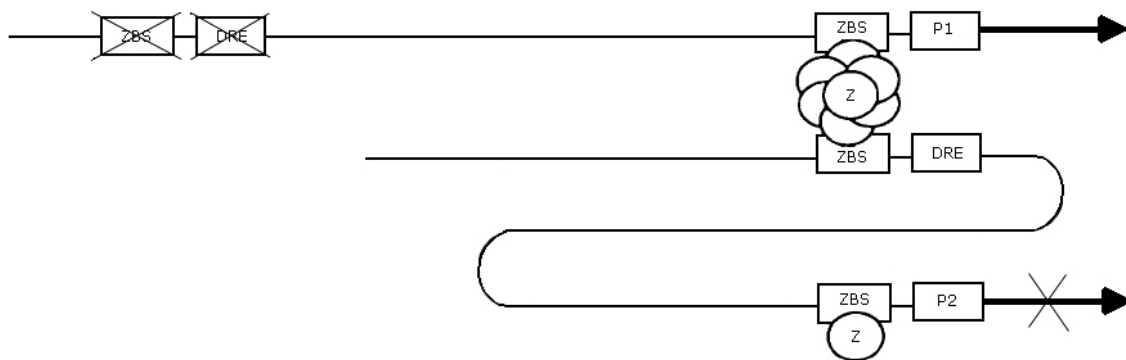


Figura I.16. – **Modelo del efecto de transvección mediado por *zeste***. El elemento regulador distante (DRE - *Distant Regulatory Element*) de la copia 1 del gen está deletado, mientras que la copia 2 presenta una mutación en su secuencia que imposibilita su expresión. El complejo formado por la proteína Z se une a los ZBS y acerca el DRE de la copia 2 del gen al promotor de la copia 1, permitiendo que se expresen ambas copias. Así gracias al fenómeno de transvección se consigue una complementación que permite la correcta expresión del gen (Figura adaptada de Benson y Pirrotta, 1988).

La transvección en *dpp* mediada por *z* ha sido demostrada en los mutantes *held out* (dpp^{d-ho}) y dpp^{hr4} . dpp^{d-ho} es una mutación *disk* consistente en una delección de 3Kb a ~23Kb del extremo 3' de la unidad de transcripción de *dpp*, que en homocigosis causa variaciones en el ángulo que forman las alas. dpp^{hr4} es una mutación sin sentido en la secuencia codificante de *dpp*, que en homocigosis causa una reducción drástica del tamaño de las alas. Con cromosomas estructuralmente normales, los heterocigotos dpp^{d-ho}/dpp^{hr4} se complementan dando lugar a individuos normales (Gelbart, 1982).

El gen *dpp* posee varios ZBSs. El ZBS que elegimos para nuestro estudio se encuentra en la región próxima al inicio de transcripción y está formado por cuatro ZBM localizados en menos de 130bp. El resto de ZBSs se encuentran en 3' de la región codificante, esparcidos a lo largo de más de 20 Kb. La región necesaria para que se

produzca el efecto de transvección mediado por z se extiende hasta la región 35E, según el mapa de los cromosomas politénicos. En presencia del alelo z^1 , dicha región se extiende hasta el centrómero, llegando hasta la posición 40, según el mapa de los cromosomas politénicos (Smolik-Utlaut y Gelbart, 1987; Benson y Pirrotta, 1988; Pirrotta, 1991).

Actualmente se desconoce si la unión de Z es necesaria para la correcta expresión de dpp . Se propone que Z podría unirse cerca del promotor e interactuar con otras proteínas Z unidas cerca de elementos reguladores distantes (DRE). De esta forma, se crearía un bucle de DNA que acercaría los DREs del propio alelo a su promotor, permitiendo la correcta expresión del gen. En este caso no sería necesaria la presencia de copias apareadas (Benson y Pirrotta, 1988) (ver Figura I.17.).

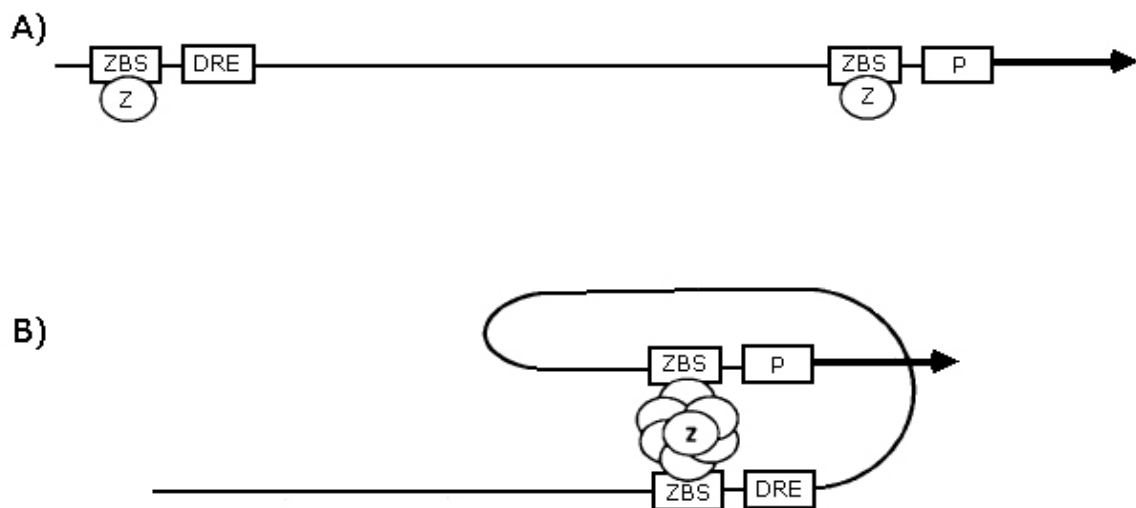


Figura I.17. - **Modelo de bucle mediado por zeste.** El complejo formado por la proteína Z se une al ZBS y acerca el DRE al promotor del gen, permitiendo la expresión del gen. En A) se muestra la cadena de DNA antes de la formación del bucle. En B) se muestra la estructura que adquiere la hebra de DNA para acercar, mediante Z , el DRE al promotor (Benson y Pirrotta, 1988).

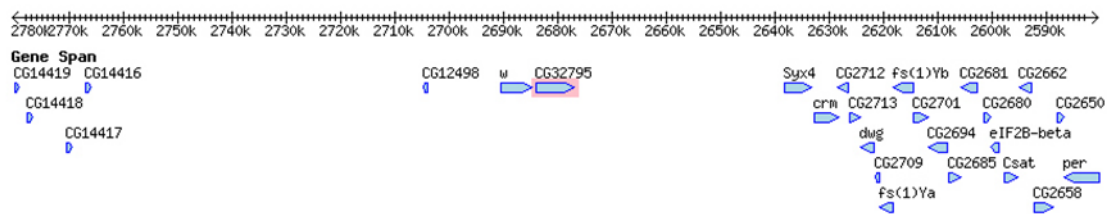
5. El gen *CG3279*

El gen *CG32795* se localiza citogenéticamente en la posición 3B6 del cromosoma X, según la predicción bioinformática. La unidad transcripcional de *CG32795* se localiza sólo a 700bp del final de transcripción de *w*, en un entorno casi vacío de genes (ver Figura I.18. A). Así, *w* y *CG32795* forman una pareja de genes (*gene pair*) del tipo cola-cabeza (TH - *tail-to-head*). Las parejas más abundantes en *Drosophila* son, igual

que en humanos, las de tipo cabeza-cabeza (HH - *head-to-head*) (Koyanagi *et al.*, 2005).

La región codificante del gen *CG32795* está compuesta por siete (CG32795-RB) u ocho (CG32795-RA, CG32795-RC) exones, según el transcrito (ver Figura I.18. B). Se han descrito tres transcritos distintos de entre 1,8 y 1,9 Kb, aunque todos ellos codifican la misma proteína de 387 aminoácidos (*FlyBase*).

A)



B)

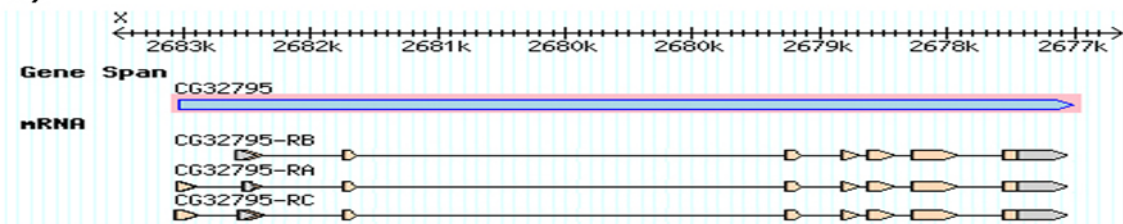


Figura I.18. – **Esquema del gen *CG32795* y su entorno.** A) El gen *CG32795* en su entorno, evidenciándose que forma parte de una pareja de genes del tipo HT. B) Esquema del gen *CG32795* de *D. melanogaster*. La barra superior indica el tamaño del gen. La barra azul indica la extensión completa del gen. La representación inferior, indica mediante cajas, la posición de los exones, en los tres transcritos descritos (Adaptado de *FlyBase*).

CG32795 es un gen predicho bioinformáticamente por el equipo de anotadores de *FlyBase* y su función es desconocida. Dispone de una línea EP (EP3617^{EP3617}; Rorth *et al.*, 1998; Dominski *et al.*, 2002), una cepa con un elemento P insertado a unos 350bp 5' del inicio del gen *CG32795* (*FlyBase*). Cuando se utiliza en esta cepa el controlador *ey-GAL4* (Hauck *et al.*, 1999), para sobrexpresar *CG32795* en los discos imaginales que darán lugar a los ojos, se obtiene un fenotipo de ojos reducidos. Este fenotipo es rescatado mediante la sobrexpresión de la ciclina E (*cycE*). Cuando se utilizan los controladores *sev-GAL4* y *en-GAL4* (Neufeld *et al.*, 1998), se obtienen individuos de

fenotipo salvaje, demostrando así que la sobreexpresión de *CG32795* no tiene ningún efecto en células postmitóticas, ni en otros tejidos que no sean los discos imaginales que dan lugar a los ojos (Tseng y Hariharan, 2002).

6. Los elementos transponibles

Los primeros elementos transponibles (ETs) fueron descritos en los años 50 por la Dr. Barbara McClintock en *Zea mays* (McClintock, 1956). Desde entonces hasta la actualidad, se han descrito ETs en organismos tan diferentes como bacterias (ej. *Tn10* en *E. coli*, Foster, 1977), nemátodos (ej. *Tc1* en *C. elegans*, Emmons *et al.*, 1983), levaduras (ej. *Ty1* en *S. cerevisiae*, Gafner y Philippsen, 1980), vegetales (ej. *Tnt1* en *N. tabacum*, Grandbastien *et al.*, 1989), peces (ej. *Tdr1* en *D. rerio*, Izsvak *et al.*, 1995) y mamíferos (ej. *THE* en *H. sapiens*, Paulson *et al.*, 1985).

Los ETs son secuencias de DNA que tienen la capacidad de moverse y/o multiplicarse dentro del genoma, creando copias de ellos mismos repartidas a lo largo del DNA genómico. Pueden transponerse a altas frecuencias, a razón de 10^{-3} - 10^{-5} por elemento y por generación, dependiendo del tipo de elemento. Algunos, pueden perder su movilidad y/o su capacidad replicativa al integrarse en el genoma, debido a mutaciones o a truncamientos de su secuencia. En un principio, fueron tratados como DNA “basura” o “egoísta” (Ohno, 1972), pero actualmente se consideran uno de los principales componentes del genoma (ver Tabla I.3.) (Biémont y Vieira, 2006).

Espece	Tamaño del genoma (picogramos)	% TEs
<i>Zea mays</i>	5,0	60
<i>Homo sapiens</i>	3,5	45
<i>Mus musculus</i>	3,4	40
<i>Drosophila melanogaster</i>	0,18	15-22
<i>Caenorhabditis elegans</i>	0,1	12
<i>Saccharomyces cerevisiae</i>	0,012	3-5
<i>Escherichia coli</i>	0,0046	0,3

Tabla I.3. – **Tamaño de genomas y porcentaje en ETs** (Biémont y Vieira, 2006).

Los ETs pueden integrarse prácticamente en cualquier punto del genoma. Aún así, los análisis realizados a nivel genómico demuestran que algunos ETs (ej. elemento *P*) tienen tendencia a integrarse en zonas activas del genoma, siendo los inicios de transcripción de los genes del genoma huésped las principales zonas de inserción. Este fenómeno puede explicarse gracias a la estructura abierta de la cromatina en dichas regiones que facilitaría la inserción, o bien gracias a la presencia de factores celulares que atraerían al ET. Otros ETs (ej. transposones *Tc*) parecen insertarse al azar, influenciados en algún caso por la propia secuencia del DNA genómico huésped. Finalmente, algunos ETs (ej. elementos *Ty*) se integran en una diana específica, interaccionando claramente con factores celulares del organismo huésped (Wu y Burgess, 2004).

Según su mecanismo de transposición, los ETs se clasifican en tres grandes grupos (Finnegan, 1989):

- ELEMENTOS de CLASE I → Son elementos que utilizan un intermediario de RNA en sus procesos de transposición. Este tipo de movilización recibe el nombre de transposición replicativa, puesto que el número de copias del ET aumenta en cada fenómeno de transposición. Dentro de esta clase se pueden establecer dos subgrupos:

RETROTRANSPOSONES → Son elementos de estructura similar a los retrovirus, dotados de largas repeticiones terminales directas (LTR). La región entre las LTRs contiene dos ORFs que mantienen cierta similitud con los genes víricos *gag* y *pol*, (ver Figura I.19.). Ejemplos de ETs de la clase de los retrotransposones de *Drosophila* son *copia* (Levis *et al.*, 1980) y *B104* (Tschudi *et al.*, 1982). Algunos retrotransposones pueden tener un tercer ORF que recuerda el gen viral *env*, responsable de la formación de la cápside (ver Figura I.19.) y se consideran retrovirus. Un ejemplo de retrovirus es *gypsy* (Song *et al.*, 1994).

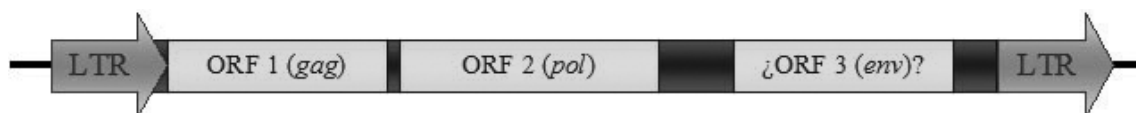


Figura I.19. – **Retrotransposón.** Los retrotransposones se caracterizan por sus LTR. Entre los LTR siempre se encuentran dos ORF que codifican proteínas implicadas en la formación de VLP (*gag*) y en la transcripción inversa del intermediario de RNA a DNA durante la transposición (*pol*). Algunos retrotransposones, considerados actualmente retrovirus, poseen un tercer ORF que recuerda el gen viral *env*, responsable de la formación de la cápside.

RETROPOSONES → Su principal característica es la ausencia de LTR y la presencia de una larga cola poliA en el extremo 3'. Los retroposones pueden a su vez clasificarse en dos subgrupos, según su capacidad codificante: LINEs (*Long Interspersed Nuclear Element*) y SINEs (*Short Interspersed Nuclear Element*). Los LINEs codifican tres proteínas: una proteína de unión al RNA, una endonucleasa (EN) y una transcriptasa reversa (RT) (ver Figura I.20.). Algunos ejemplos de LINEs son los elementos *L1* en ratón (Burton *et al.*, 1985), y el elemento *I* en *D. melanogaster* (Chambeyron y Bucheton, 2005).



Figura I.20. – **LINE**. Los LINEs no presentan LTR pero sí una cola poliA en el extremo 3'. Codifican tres proteínas: una de unión al RNA, una endonucleasa (EN) y una transcriptasa inversa (RT). Poseen un promotor interno para la RNA pol II que se encuentra separado de los ORF por una región no traducida (*UnTranslated Region* – UTR).

Los SINEs a diferencia de los LINEs, no tienen capacidad codificante y por lo tanto no capaces de transponer por sí solos, es decir, no son elementos autónomos. El extremo 3' de los SINEs suele presentar homología con los LINEs y por eso se propone que los LINEs podrían participar en la transposición de los SINEs (Dewannieux *et al.*, 2003). Algunos ejemplos de SINEs son los elementos *B1* y *B2* de ratón y las secuencias *Alu* en humanos.



Figura I.21. – **SINE**. Los SINEs se caracterizan por la ausencia de LTR, la presencia de una cola poliA y de un promotor interno para la RNA pol III. Suelen medir entre 100 y 300bp.

ELEMENTOS de CLASE II → Son elementos que no utilizan intermediario de RNA. Se movilizan por escisión directa del transposón, que al reinsertarse crea duplicaciones directas en el punto de inserción. Los ETs de clase II tienen extremos terminales invertidos (IRs - *inverted repeats*). Generalmente contienen un único ORF que codifica una transposasa encargada de escindir el elemento del genoma, gracias al reconocimiento de la secuencia de los IRs. Algunos ejemplos de elementos de clase II son los elementos *P* (Bingham *et al.*, 1982) y *hobo* (Streck *et al.*, 1986) de *D. melanogaster*.

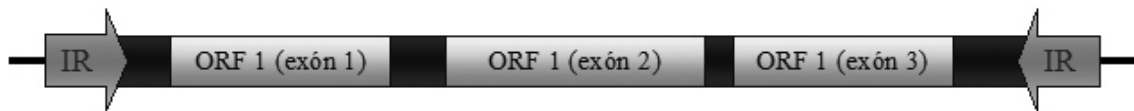


Figura I.22. – **ET de clase II.** Se caracterizan por la presencia de IR. Generalmente tienen un único ORF que codifica la transposasa.

Otros ETs de clase II a resaltar son los MITEs (*Miniature Inverted-repeat Transposable Elements*), ETs pequeños (~100 - 500bp) con IR cortos (10-15bp) y una secuencia interna rica en A y T. Se propone que los MITEs son ETs de clase II truncados, es decir, no autónomos. Algunos ejemplos de MITEs son los elementos de la familia *Emigrant* en *A. thaliana* (Casacuberta *et al.*, 1998) y los de la familia *mPing* en *O. sativa* (Kikuchi *et al.*, 2003).



Figura I.23. – **MITE.** Los MITEs son ETs de clase II no autónomos. Se caracterizan por ser pequeños, con IR cortos y una región interna no codificante rica en AT.

Finalmente, dentro de los ETs de clase II hay que nombrar el grupo de los helitrones. Dichos elementos no siguen el mecanismo de transposición típico de los ETs de clase II, sino que transponen mediante replicación por círculo rodante. Los helitrones se insertan entre nucleótidos A y T, poseen unos extremos terminales TCT en 5' y CTAG en 3' y tienen una secuencia palindrómica de 16 - 20bp capaz de formar una horquilla (*hairpin loop*) a unos 10 - 12bp del extremo 3' terminal. No presentan IRs ni duplicaciones del punto de inserción. Los helitrones autónomos codifican una DNA helicasa y una o dos copias de una nucleasa/ligasa (ver Figura I.24). Hasta el momento, la mayor parte de los helitrones han sido clasificados como no autónomos (Kapitonov y Jurka, 2001). Algunos ejemplos de helitrones son el elemento *Hel-It1* en *Ipomoea tricolor* (Choi *et al.*, 2007) y el elemento Mo17 en *Z. mays* (Brunner *et al.*, 2005).



Figura I.24. – **Helitron autónomo.** Los helitrones poseen extremos terminales TCT en 5' y CTAG en 3'. Poseen una secuencia palindrómica en 3' (caja rayada) capaz de crear una horquilla. Codifican una DNA helicasa y una nucleasa/ligasa.

ELEMENTOS de CLASE III → En esta categoría se agrupan todos aquellos ETs difíciles de clasificar porque presentan características inusuales y un mecanismo de transposición desconocido. La mayor familia de ETs de clase III es la superfamilia *Foldback (FB)* (ver I.6.1. El elemento transponible *FB-NOF*). Los elementos *FB* se encontraron inicialmente en *D. melanogaster* y se caracterizan por sus largos IR, que pueden llegar a medir varias Kb. Además los dos IRs de un mismo elemento pueden tener tamaños distintos. Aunque se pueden encontrar ORFs entre sus IR, las proteínas codificadas no presentan ninguna similitud con las codificadas en los ORFs de los ETs de clase I y de clase II (Truett *et al.*, 1981; Potter, 1982; Templeton y Potter, 1989). Algunos elementos de la familia *FB* son *FARE* en *A. thaliana* (Windsor y Waddell, 2000) y *Galileo* en *Drosophila buzzatii* (Casals *et al.*, 2003).

Dentro de los ETs de clase III, también se encuentra la familia de los elementos *TU* descritos en *Strongylocentrotus purpuratus*. Su mecanismo de transposición no está claro y por eso se les considera ETs de clase III (Liebermann *et al.*, 1986).

6.1. El elemento transponible *FB-NOF*

FB fue descrito como una nueva familia de transposones con IRs (Potter *et al.*, 1980) y actualmente se considera la superfamilia más numerosa dentro de los TEs de clase III. *FB* ha sido clasificado como ET de clase III puesto que no se han encontrado en él secuencias codificantes y además se desconoce su mecanismo de transposición. Sin embargo, en los puntos de inserción de *FB* siempre se encuentra una duplicación directa de 9bp, lo que sugiere que *FB* debería ser transpuesto mediante una transposasa. No hay una secuencia consenso para los puntos de inserción de *FB*, pero las duplicaciones directas suelen ser ricas en A/T excepto en la primera posición que suele ser rica en C/G. *FB* tiende a insertarse en intrones o en regiones 5' de genes, igual que otros ETs y retrovirus (ej. elemento *P* y *HIV-1*) (Wu y Burgess, 2004; Badal *et al.*, 2006b).

Los miembros de la familia *FB* presentan diferencias notables, pero mantienen una estructura común: tienen largos IRs formados por repeticiones directas en tándem de secuencias de menor longitud (Harden y Ashburner, 1990). Recientemente, se ha descrito una secuencia consenso para los IRs de *FB* (ver Figura I.25. A). Ésta empieza con 470bp de secuencia no repetitiva (a excepción de algunas repeticiones directas

cortas). A continuación le siguen una serie de bloques de 154bp que empiezan con el trinucleótido CTC. Cada bloque está formado por cinco repeticiones, casi idénticas, de 32bp. Las repeticiones siguen el mismo orden en todos los bloques. El primer bloque suele presentar una primera repetición un poco más larga, pero el resto de los bloques suelen ser idénticos, aunque se pueden encontrar bloques irregulares (ver Figura I.25. C) (Badal *et al.*, 2006b).

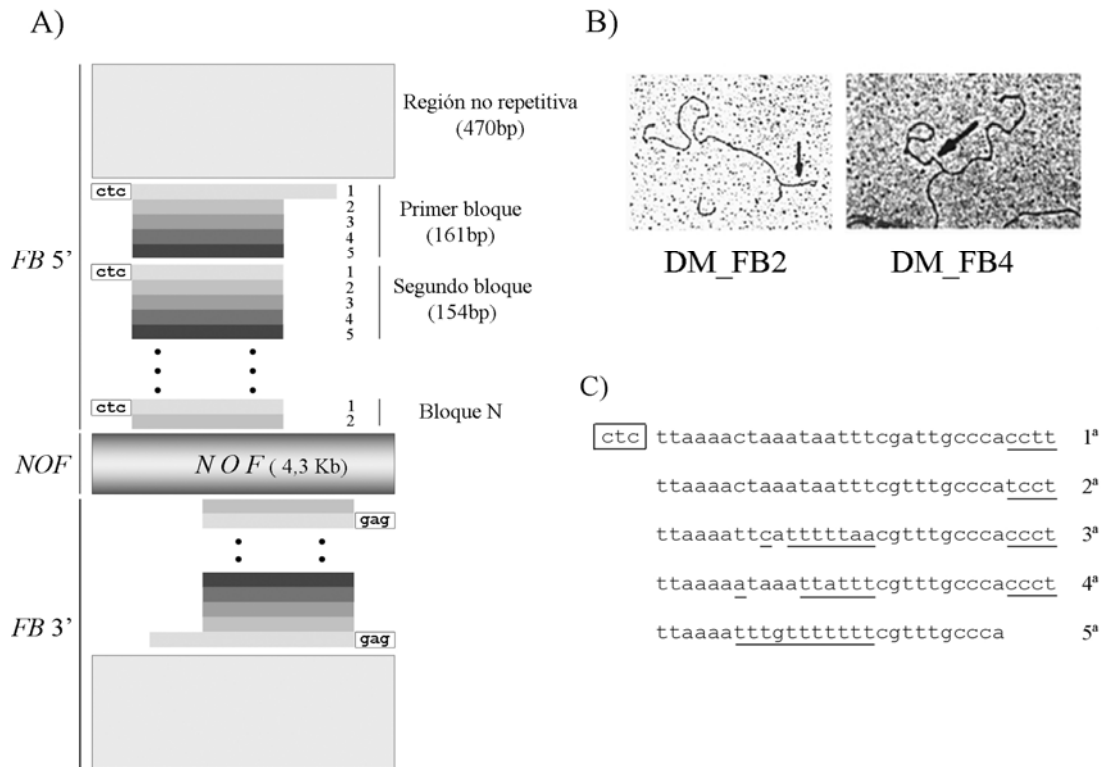


Figura I.25. – **El elemento *FB-NOF***. A) Esquema de la estructura de los IRs de *FB*. B) Imágenes de microscopio electrónico en las que se observa la formación de horquillas con bucle correspondientes a los *FB* en ssDNA. C) Esquema de las cinco repeticiones que conforman cada uno de los bloques de los IRs de *FB* (Adaptada de Potter *et al.*, 1980; Badal *et al.*, 2006b).

La observación de ssDNA en el microscopio electrónico permite identificar los elementos *FB* gracias a las horquillas (*hairpins*) que éstos forman (ver Figura I.25. B) (Potter *et al.*, 1980; Truett *et al.*, 1981). La presencia de bucles (*loops*) en las horquillas se explicó como la presencia de secuencias no-*FB* entre sus IRs. Actualmente pero, se tiende a considerar que la presencia de dichos bucles es producida principalmente por un apareamiento imperfecto de los IRs, puesto que éstos suelen tener diferente longitud (Brierley y Potter, 1985). Aún así, en algunos casos se pueden encontrar secuencias no-*FB* entre los IRs de *FB* (Paro *et al.*, 1983).

Una de estas secuencias no-*FB* es *NOF*. Éste está siempre asociado a *FB*, y no se ha encontrado nunca fuera de los IRs de *FB*. Los dos extremos de *NOF* se insertan siempre en un bloque de *FB* detrás de la secuencia TCCT de la segunda repetición de un bloque (ver Figura I.25. B). Por ese motivo se considera que *FB-NOF* es el elemento completo (Templeton y Potter, 1989; Harden y Ashburner, 1990; Badal *et al.*, 2006b).

NOF mide 4,3Kb y posee unos IRs idénticos de 308bp y una región interna codificante, con tres ORFs (Templeton y Potter, 1989):

- ORF 1 → Codifica una proteína de ~71 kDa capaz de unirse al DNA. Se ha propuesto que podría tratarse de la transposasa de *NOF*.

- ORF2 → Codifica una proteína predicha bioinformáticamente capaz de unirse al DNA. Se ha logrado amplificar su mRNA mediante RT-PCR. Se ha propuesto que los ORF1 y 2 se encuentran formando un único transcrito primario (Badal, 2007).

- ORF 3 → Es un ORF predicho bioinformáticamente que se encontraría en la cadena contraria a la de los ORFs 1 y 2. No se ha logrado amplificar su mRNA mediante RT-PCR (Badal, 2007).

El hecho de que *NOF* tenga sus propios IRs y ORFs sugiere que *NOF* pudo ser en su momento un ET autónomo. Así, tradicionalmente *FB-NOF* ha sido considerado como dos ETs independientes. Sin embargo, los últimos estudios apuntan que podría tratarse de un único elemento que podría presentarse de forma entera (*FB-NOF*) o de forma truncada (*FB*) (Badal, 2007).

La cantidad de elementos *FB-NOF* en el genoma de *D. melanogaster* es muy variable. En general se pueden llegar a encontrar entre 25 y 50 copias de *FB* y no más de 10 copias de *NOF*. Cabe destacar la relación existente entre el número de *FB-NOFs* y el número de *FBs* de un genoma: a mayor número de elementos completos, mayor número de elementos truncados.

Finalmente, resaltar la contribución de los elementos *FB-NOF* a la plasticidad genómica. *FB* es capaz de promover cualquier tipo de reordenación genómica: inversiones (Cáceres *et al.*, 1999), translocaciones (Moschetti *et al.*, 2004) e incluso duplicaciones (Badal *et al.*, 2006a).

6.2. Efectos de la inserción de un elemento transponible

Gracias a su capacidad de transposición, los ETs pueden invadir un genoma causando mutaciones por inserción y reordenaciones cromosómicas. Un claro ejemplo lo encontramos en *D. melanogaster*, el 50-80% de cuyas mutaciones han sido atribuidas a la inserción de ETs. Los ETs son elementos activos del genoma que pueden fomentar la recombinación (Chance *et al.*, 1994), aportar elementos reguladores de la transcripción (van de Lagemaat *et al.*, 2003), señales de poliadenilación (Levis *et al.*, 1984) o de *splicing* (Goldin y Papaioannou, 2003) y hasta codificar una proteína (Gotea y Makalowski, 2006). Así mismo, la inserción de ETs puede desregular o incluso llegar a inhibir la expresión de un determinado gen. Por ese motivo, se les considera responsables de 0,5-1% de las enfermedades humanas (ej. inserciones de SINEs y LINEs se señalan como posibles causas de la hemofilia, distrofia de Duchenne y algunos tumores) (Deininger y Batzer, 1999).

A lo largo de la evolución, los organismos han desarrollado estrategias de defensa para limitar el número de ETs en su genoma y minimizar las transposiciones (Rizzon *et al.*, 2002). El balance entre las fuerzas que controlan los ETs en poblaciones naturales, sigue siendo hoy en día un tema controvertido (Nuzhdin, 1999). Se ha propuesto que el control del número de ETs implique una selección negativa contra las reorganizaciones causadas por recombinación ectópica entre ETs (Gray, 2000) o bien una selección negativa contra las mutaciones inducidas por ETs (Biémont *et al.*, 1997).

La inserción de ETs en un genoma no tiene por que ser siempre negativa para el genoma huésped, ya que su actividad mutagénica puede significar una contribución a la diversidad de la especie. Algunos ETs han sido “domesticados” por el genoma huésped y en la actualidad actúan como genes o elementos reguladores, representando una fuente de innovación génica (Biémont y Vieira, 2006).

La expresión de los ETs, es decir de los RNAs que transcribe, presenta especificidad tisular y temporal, y en algunos casos es dependiente del sexo. La presencia de patrones de expresión tan específicos ha sido explicada por la interacción de las secuencias reguladoras del ET con numerosos genes del desarrollo del genoma huésped, o bien por la interacción con genes cercanos altamente expresados en el genoma huésped (según el mecanismo llamado *readthrough*). Ambos mecanismos implican la inversión de una gran cantidad de energía por parte de los organismos implicados. Así, ese patrón de expresión determinado, permitiría al organismo ganar una función (Biémont y Vieira, 2006). Se ha demostrado que algunos retrotransposones pueden regular determinados genes jugando un papel determinante en el desarrollo de los oocitos y embriones preimplantacionales de ratón (Peaston *et al.*, 2004). Pero el papel de los ETs va más allá del control de ciertos procesos de desarrollo. Se ha demostrado que los retrotransposones *LINE-1* se hallan insertados preferentemente en las regiones reguladoras de determinados genes neuronales de ratón. De esta forma, alteran la expresión génica creando diferentes poblaciones de células neuronales (Muotri *et al.*, 2005).

Los ETs pueden verse sujetos a un control epigenético, a través de metilaciones en el DNA y metilaciones y acetilaciones de histonas. Estas marcas epigenéticas tienen un efecto en los niveles de expresión de los ETs y de los genes cercanos al punto de inserción. Los ETs pueden tener efectos sobre las marcas epigenéticas bien induciendo ellos mismos la metilación del DNA próximo al punto de inserción o bien desestructurando el patrón epigenético de la zona cercana al punto de inserción (Lippman *et al.*, 2004). Un buen ejemplo de la relación gen-ET, lo encontramos en el control del color del pelaje de los ratones *agouti*. La sobreexpresión del gen *agouti* da un pelaje amarillo a los ratones. Si el retrotransposón *IAP* se inserta justo en 5' del gen *agouti*, la expresión del gen depende del estado de metilación del elemento *IAP* insertado. Si hay diferencias en el estado de metilación en distintas células, se obtiene un pelaje manchado (Morgan *et al.*, 1999).

7. Las cepas M115 y RM115

Las cepas M115 y RM115 fueron halladas por la Dr. Eva Baldrich (Baldrich, 1998) y descritas por nuestro grupo (Badal *et al.*, 2006a). Ambas cepas son portadoras

del alelo z^1 y presentan alteraciones del mapa de restricción en la región del extremo 3' del gen w . Dicha alteración recordaba al alelo w^{+UR} , en el que la inserción de un elemento $FB-NOF$ en 3' del gen w daba lugar a machos de ojos amarillos en una cepa z^1 . La pérdida de NOF y la presencia de el elemento truncado FB generaba el alelo w^{+UZ} que en cepas z^1 se observaba fenotípicamente como machos de ojos rojos (Rasmuson-Lestander y Ekstrom, 1996). Así, se podría establecer un paralelismo entre las cepas w^{+UZ} y M115 y las cepas w^{+UR} y RM115. El punto de inserción de FB o $FB-NOF$ coincide y se sitúa en la posición -5655 según las coordenadas de Levis, dentro del tercer intrón del gen $CG32795$. Sin embargo, el tamaño de los elementos insertados es distinto, demostrando así que las cepas (w^{+UZ} y M115) y sus revertientes (w^{+UR} y RM115) tienen orígenes distintos.

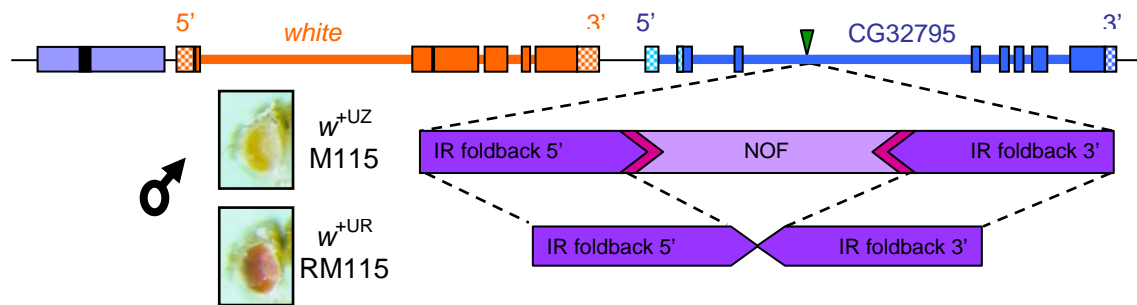


Figura I.26. – **Esquema de la inserción de FB y $FB-NOF$ en las cepas M115 y RM115.** El elemento transponible se inserta en la posición -5655 según las coordenadas de Levis, en 3' del gen w , dentro del tercer exón del gen $CG32795$. Los machos de las cepas w^{+UZ} y M115 tienen ojos amarillos, mientras que los correspondientes machos de las cepas revertientes (w^{+UR} y RM115) presentan un fenotipo de ojos rojos.

Se creía que la inserción de $FB-NOF$ en una región tan próxima al extremo 3' del gen w influenciaba de algún modo la interacción *zeste-white* (Rasmuson-Lestander y Ekstrom, 1996). Sin embargo, posteriormente se demostró que el fenotipo de ojos amarillos presente en los machos de la cepa M115 no era debido a ningún efecto de $FB-NOF$ sobre la interacción *zeste-white*, sino a una duplicación que se extendía desde más allá del extremo 5' del gen w (incluyendo la región de los *enhancers* y el ZBS) hasta el punto de inserción de $FB-NOF$. Así, se propone que el papel FB más que el de modificador de la interacción *zeste-white*, sería el de mediador de reordenaciones cromosómicas que, en este caso concreto, llevaron a la duplicación de la región circundante del gen w (Badal *et al.*, 2006a).

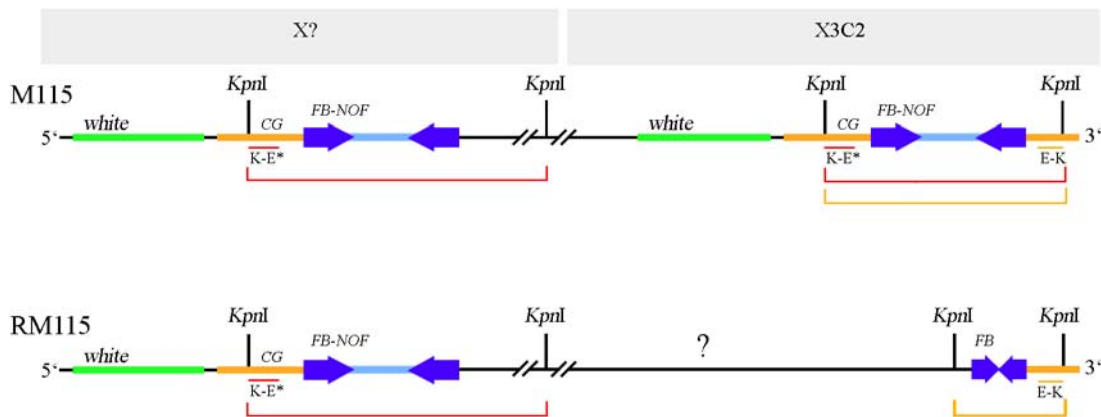


Figura I.27. – Modelo propuesto para explicar la duplicación en la cepa M115 y la reversión de ésta en la cepa RM115. La sonda K-E* en rojo, y la sonda E-K en amarillo, permiten el estudio del punto de inserción desde 5' y desde 3' respectivamente. La barra gris superior indica la posición cromosómica. (Adaptada de Badal, 2007).

La cepa M115 es, por tanto, portadora de una duplicación que incluye el gen *w* y su promotor y el extremo 5' del gen *CG32795* hasta el punto de inserción del elemento *FB-NOF* (-5655 según las coordenadas de Levis). Así pues, se conoce el extremo 3' de la duplicación, pero se desconoce cual es su extremo 5' (ver Figura I.27.). En la cepa RM115, se produce una reordenación que elimina una de las regiones duplicadas (Badal *et al.*, 2006a). Los indicios apuntan que la copia delecionada es la original, puesto que estudiando por *Southern blot* el entorno del punto de inserción en la cepa RM115, se observa que no se mantiene el mapa de restricción que correspondería a la secuencia salvaje esperada en ambos lados del punto de inserción. Estudiando el mapa de restricción obtenido acercándose al punto de inserción del ET por el extremo 5' (sonda K-E*), se encuentra la secuencia del gen *CG32795* y en -5655 se entra dentro de un elemento *FB-NOF*. Sin embargo, el mapa de restricción obtenido acercándose al punto de inserción del ET (sonda E-K), muestra la secuencia 3' del gen *CG32795* y en -5655 se entra dentro de un elemento *FB* (ver Figura I.27.). Así, ambos fragmentos parecen no encontrarse situados en continuidad lineal en el genoma de la cepa RM115.

II. OBJETIVOS

El presente trabajo de Tesis Doctoral se enmarca en una línea de investigación del Grup de Mutagènesis de la Universitat Autònoma de Barcelona que busca explicar a nivel molecular la inserción de un elemento *FB-NOF* en el tercer intrón del gen *CG32795*. El origen de esta línea se encuentra en los mutantes M115 y RM115, aislados por la Dr. Baldrich a lo largo de su Tesis Doctoral (Baldrich, 1998) y en el parecido de éstos con los mutantes w^{+UR} y w^{+UZ} descritos por la Dr. Rasmuson-Lestander (Rasmuson-Lestander y Ekstrom, 1996). Posteriormente, los mutantes M115 y RM115 fueron caracterizados por nuestro grupo (Badal *et al.*, 2006a). Sin embargo, faltaba explorar los efectos que la inserción de *FB-NOF* en el tercer intrón de *CG32795* tiene sobre la expresión del par de genes formado por *w* y *CG32795*. Tanto los individuos mutantes M115 como los RM115 son portadores del alelo z^{1} . Puesto que la mutación z^{1} afecta a la expresión de *w*, consideramos que sería interesante conocer con más detalle cuál era el papel que z^{1} desempeñaba en la regulación de la expresión de diferentes genes con ZBS. Así los objetivos propuestos son los siguientes:

- 1 - Estudiar el efecto de la inserción de un elemento *FB-NOF* (entero o truncado) en el tercer intrón del gen *CG32795*, sobre la expresión del par de genes *w* y *CG32795*.
 - 1A - Dada la poca información disponible acerca del gen *CG32795*, nos propusimos caracterizarlo a nivel de mRNA y de predicción de la estructura proteica.
 - 1B - Estudiar la expresión del par de genes *w* y *CG32795* en machos de las cepas M115 y RM115, portadores de un elemento *FB-NOF* (entero o truncado) en el tercer intrón del gen *CG32795*.
- 2 - Ampliar los conocimientos acerca de los efectos del alelo mutante z^{1} sobre la expresión de genes con un ZBS y la posible extensión de dicho efecto a los genes próximos, mediante el estudio de:

- 2A - La expresión de los genes *w*, *dpp* y *CG32795* en hembras z^1 .
- 2B - El posicionamiento de nucleosomas en los ZBS de *w* y *dpp* y en el extremo 5' del gen *CG32795* en hembras z^1 .
- 2C - Los patrones de metilación de los genes *w*, *dpp* y *CG32795* en hembras z^1 .

III. RESULTADOS Y DISCUSIÓN

1. Estudio del efecto de la inserción de un elemento transponible *FB-NOF* en la expresión de los genes circundantes

Las cepas M115 y RM115 son portadoras del alelo z^1 y poseen un elemento *FB-NOF* completo (*FB-NOF*) o incompleto (*FB*) insertado en -5655 según las coordenadas de Levis, dentro del tercer intrón del gen *CG32795*. El gran potencial recombinogénico de *FB-NOF* medió una reordenación en la cepa M115, duplicando desde al menos el ZBS de *w* hasta -5655. Dicha recombinación fue revertida espontáneamente en la cepa RM115, dando lugar a individuos con una única copia de la región previamente duplicada en M115 (Badal *et al.*, 2006a).

Los ETs pueden alterar la expresión de los genes localizados en regiones próximas a su punto de inserción (Biémont y Vieira, 2006). Por este motivo, nos propusimos estudiar la expresión y el posicionamiento de nucleosomas de los genes más cercanos al punto de inserción de este elemento *FB-NOF*: el gen *w* y el gen predicho *CG32795*. Para poder llevar a cabo este estudio fue necesario en primer lugar caracterizar el gen *CG32795*.

1.1. Estudio del gen *CG32795*

El gen *CG32795* estaba bioinformáticamente predicho (ver Figura I.18. en el apartado I.5. El gen *CG32795*). Sin embargo, consideramos imprescindible comprobar dicha predicción y, con la información obtenida, tratar de conocer su posible función y conservación.

1.1.1. Patrones de *splicing* alternativo

El estudio del gen *CG32795* se inició con la obtención del cDNA completo para dicho gen en machos de tres cepas distintas: la cepa control CS, la cepa M115 y la RM115. En una primera aproximación, decidimos comprobar si *CG32795* se expresaba tanto en su región 5' (antes del punto de inserción en M115 y RM115) como en su región 3' (después del punto de inserción en M115 y RM115), en las tres cepas a estudiar.

Las RT-PCRs llevadas a cabo con las parejas de cebadores: BACNex2b con BACNex3b y BACNex7b con BACNex8b, demostraron no sólo que dichas amplificaciones se producían y que por tanto existía mRNA de aquellas regiones, sino que el mRNA encontrado tenía el tamaño esperado según la predicción bioinformática; 185bp en el caso de la pareja BACNex2b/BACNex3b y 493bp en el de la pareja BACNex7b/ BACNex8b (ver Figura III.1. A).

Posteriormente decidimos comprobar que los dos fragmentos del mRNA correspondían a un único mRNA y con este fin realizamos una amplificación por RT-PCR de toda la región central del gen, desde el segundo hasta el octavo exón, usando la pareja de cebadores BACNex2b/ BACNex8b y obteniendo una banda de 1209bp (ver Figura III.1. B). Esta amplificación fue clonada para su posterior secuenciación, dando lugar a los clones pCR_2v8CS, pCR_2v8M115 y pCR_2v8RM115.

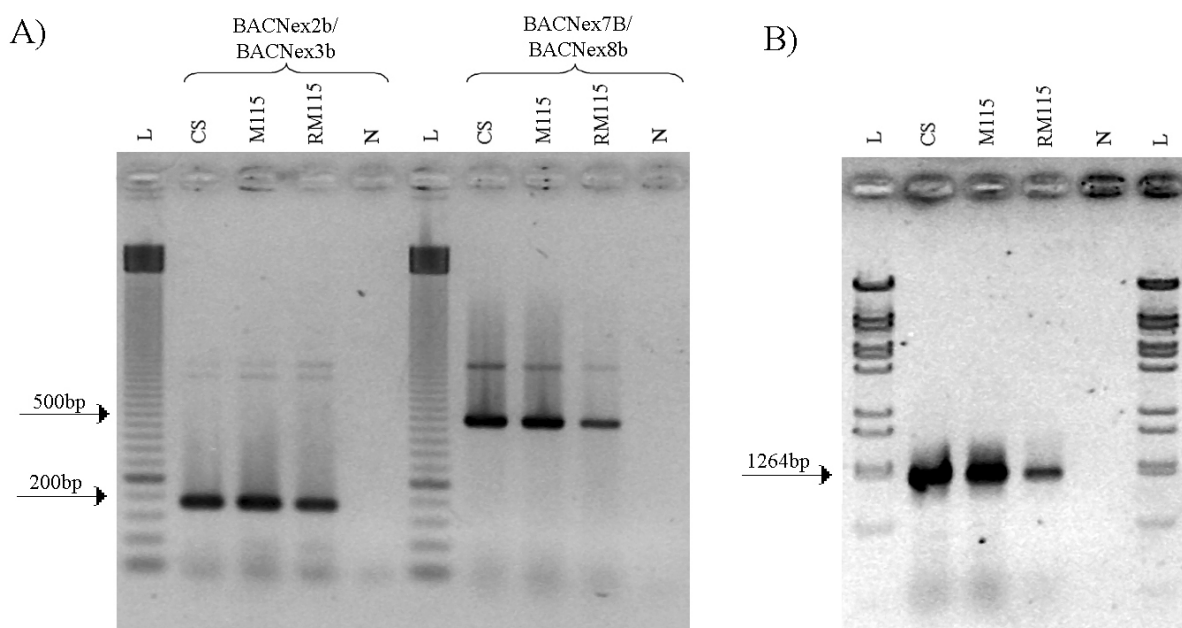


Figura III.1. – **RT-PCRs del gen CG32795.** Figura A) Se muestran las amplificaciones sobre RNA total del exón 2 al 3 (185bp) y del exón 7 al 8 (493bp). L – Marcador de 50bp. N – Control negativo de la reacción de RT-PCR. Figura B) Se muestra la amplificación sobre RNA total del exón 2 al exón 8 (1209bp). L – Marcador fago λ digerido con *BstEII*. N – Control negativo de la reacción de RT-PCR.

Las secuencias de los clones pCR_2v8CS, pCR_2v8M115 y pCR_2v8RM115 desde el exón 3 hasta el exón 7 son idénticas en los tres casos y se ajustan perfectamente a las predicciones bioinformáticas de *FlyBase*.

Habiendo clonado la parte central del cDNA del gen *CG32795*, sólo nos quedaba amplificar los extremos 5' y 3' de dicho transcrito para poder obtener la secuencia completa del cDNA en las tres cepas estudiadas. Con ese fin realizamos un RACE, tanto del extremo 5' como del 3'.

1.1.1.1. Clonación del extremo 5'

El extremo 5' del transcrito de *CG32795* se obtuvo con un RACE 5', usando para el paso final de PCR la pareja de cebadores BACNex3b y Oligo-dT Ancla que daba lugar a una amplificación de 457bp, 345bp o 518bp según se tratara de la forma predicha *CG32795-A*, *CG32795-B* o *CG32795-C*. No se consiguió una única banda en ninguna de las tres cepas estudiadas, aún usando condiciones de PCR más restrictivas. Se obtuvo un doblete de bandas con tamaños alrededor de los 400bp (ver Figura III.2. A).

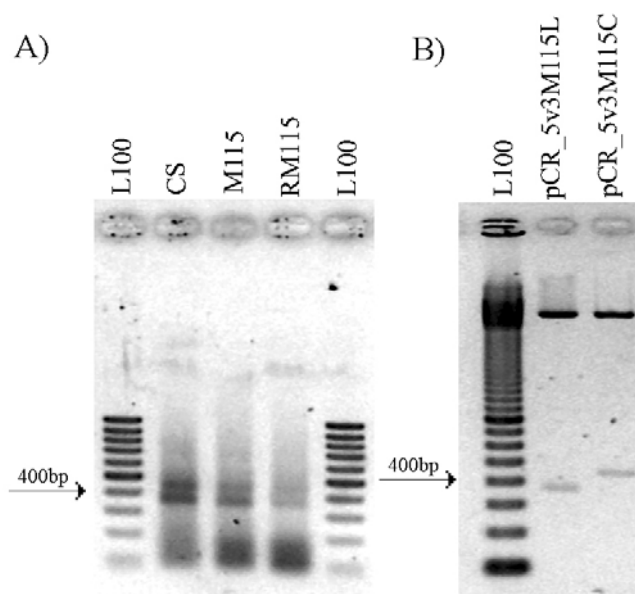


Figura III.2. – **PCR final del RACE 5' y clonación de los fragmentos.** A) La RT-PCR usando la pareja de cebadores BACNex3b y Oligo-dT Ancla, no dio en ninguna de las tres cepas estudiadas una única banda, sino un doblete. B) Liberación de los productos clonados por digestión con *EcoRI*. Las dos bandas de cada cepa fueron clonadas para su posterior secuenciación, dando lugar a los clones pCR_5v3XC (en el caso de la banda más corta) y pCR_5v3XL (en el caso de la banda más larga), en los que X es el nombre de la cepa. En la imagen superior se muestran concretamente los clones correspondientes a la cepa M115. L100 – Marcador de 100bp.

El doblete obtenido se clonó para su posterior secuenciación, dando lugar a los clones pCR_5v3XC (en el caso de clonar la banda más corta) y pCR_5v3XL (en el caso

de clonar la banda más larga), en los que X es el nombre de la cepa (CS, M115 o RM115) (ver Figura III.2. B).

Una vez obtenida la colección entera de clones (pCR_5v3XC, pCR5v3XL, pCR_2v8X, pCR_7vAXC, pCR_7vAXL siendo X = CS, M115 y RM115) estos se mandaron a secuenciar. Puesto que los clones eran solapantes se usó para la secuenciación el mismo cebador que se había usado para la última PCR de cada proceso: BACNex3b para los clones pCR_5v3XC, pCR5v3XL; BACNex2b y BACNex8b para los clones pCR_2v8X; y BACNex7b para los clones pCR_7vAXC, pCR_7vAXL. Sin embargo, nuestra sorpresa fue al descubrir que tanto los clones pCR_5v3XC, pCR5v3XL como los clones pCR_7vAXC, pCR_7vAXL eran diferentes variantes de *splicing* del mismo transcrito.

En el extremo 5' del transcrito de *CG32795* se encontraron cuatro variantes de *splicing*, aunque la presencia de sólo dos tipos de banda al liberar los productos clonados, mediante digestión *EcoRI*, hacía pensar en la presencia exclusiva de dos variantes (ver Figura III.2. B). Tres de las variantes encontradas eran parecidas a las variantes previamente predichas en *FlyBase*: *CG32795-SA* se parece a *CG32795-RA* con un primer exón más corto, *CG32795-SB* se parece a *CG32795-RB* con un primer exón más largo y *CG32795-SC* se parece a *CG32795-RC* con un primer exón más corto (ver Figura III.3.).

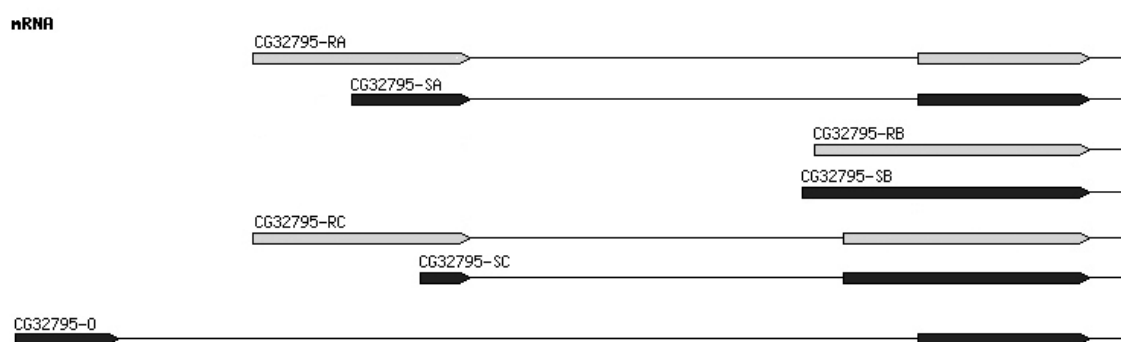


Figura III.3. – **Variantes de *splicing* del extremo 5' del gen *CG32795*.** Las variantes previamente predichas se representan en gris claro. Las nuevas variantes encontradas al realizar un RACE5' sobre RNA total de machos adultos de *D. melanogaster*, se representan en gris oscuro. La representación de cada variante predicha junto a la variante encontrada permite una comparación más cómoda.

Sin embargo se encontró una cuarta variante (*CG32795-0*) que no recuerda ninguna de las predichas con anterioridad. Dicha variante posee un primer exón localizado en el extremo 5' de lo que se había considerado hasta ahora el inicio de transcripción del gen. El que según la predicción génica era el primer exón, no se transcribe en *CG32795-0* y el segundo exón de la variante *CG32795-RA*, es idéntico al segundo exón de *CG32795-0*.

La similitud entre algunas de las variantes encontradas y las variantes predichas podría hacer sospechar que se cometió algún error de secuenciación. Cabe remarcar que nuestras cuatro variantes fueron clonadas y secuenciadas en las tres cepas estudiadas, obteniendo secuencias idénticas para todas ellas. Sin embargo, en ninguna de las cepas se encontró ninguno de los transcritos predichos por *FlyBase*. Puede que las diferencias entre las variantes predichas y las variantes que nosotros obtuvimos sean debidas a la utilización de tejidos distintos. Nuestro estudio se realizó en machos adultos de *D. melanogaster* de las cepas CS, M115 y RM115. Los cDNAs y ESTs de *Flybase* se obtuvieron a partir de embriones, larvas y cabezas de individuos adultos. La existencia de variantes de *splicing* específicas de un sexo o un tejido o una etapa del desarrollo no es rara en *D. melanogaster*. Sin embargo, queda por demostrar si la existencia de variantes específicas de tejido es real o si se produjo algún error en la predicción.

La existencia de las cuatro variantes de *splicing* descritas por nosotros (*CG32795-SA*, *CG32795-SB*, *CG32795-SC* y *CG32795-0*) queda confirmada al haberse encontrado en tres cepas distintas con una secuencia idéntica. Las diferentes variantes podrían presentar especificidad tisular, pero no sexual o temporal, puesto que todas se encontraron en machos adultos. Sin embargo, la proteína codificada en todas las variantes sería la misma, puesto que el origen de traducción no cambiaría. Así, el único cambio entre las variantes sería el 5' UTR (*UnTranslated Region*). La presencia de variantes de *splicing* en el 5' UTR ya ha sido descrita en otros genes de *D. melanogaster* (ej. *Dstpk61*, Clyde y Bownes, 2000). Se considera que dichas variantes podría ser debida al uso de distintos promotores, que se encontrarían regulados en *trans* por diferentes factores. La actividad de un promotor determinado dependería de la presencia o ausencia de las proteínas reguladoras necesarias, que podrían ser específicas de tejido o de etapa del desarrollo. El 5' UTR también podría contener sitios de unión

para factores en *trans* que podrían regular tanto la traducción como la localización subcelular del transcrito (Clyde y Bownes, 2000).

1.1.1.2. Clonación del extremo 3'

El extremo 3' del transcrito de *CG32795* se obtuvo con un RACE 3', usando para el paso final de PCR la pareja de cebadores BACNex7b y PCR Ancla que daba lugar a una amplificación de 862bp (ver Figura III.4. A).

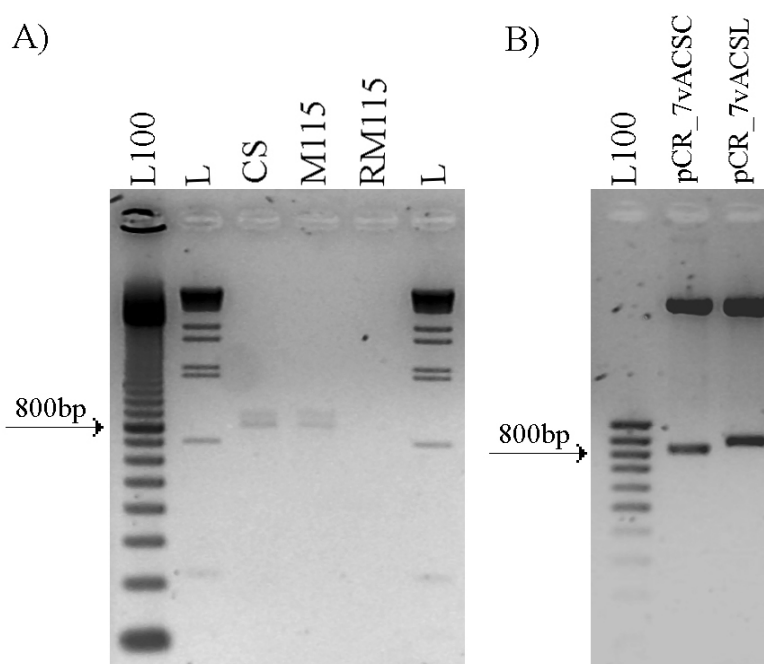


Figura III.4. –**PCR final del RACE 3' y clonación de los fragmentos.** A) La RT-PCR usando la pareja de cebadores BACNex7b y PCR ANCLA, no dio en ninguna de las tres cepas estudiadas una única banda, sino un doblete. Aunque en el carril correspondiente a la cepa RM115 no se observa claramente la amplificación, había material, puesto que posteriormente se pudo clonar sin dificultad. B) Liberación de los productos clonados por digestión con *EcoRI*. Las dos bandas de cada cepa fueron clonadas para su posterior secuenciación, dando lugar a los clones pCR_7vAXC (en el caso de la banda más corta) y pCR_7vAXL (en el caso de la banda más larga), en los que X es el nombre de la cepa. En la imagen superior se muestran concretamente los clones correspondientes a la cepa CS. L100 – Marcador de 100bp. L – Marcador fago λ digerido con *BstEII*.

En el paso final de PCR del RACE 3' no se consiguió una única banda en ninguna de las tres cepas estudiadas (ver Figura III.4. A). En todos los casos, y aunque la PCR se repitió en condiciones más restrictivas, se obtuvo un doblete. Dicho doblete se clonó para su posterior secuenciación, dando lugar a los clones pCR_7vAXC (en el caso de clonar la banda más corta) y pCR_7vAXL (en el caso de clonar la banda más larga), en los que X es el nombre de la cepa (CS, M115 o RM115) (ver Figura III.4. B).

En el extremo 3' se encontraron dos variantes, para el 3' UTR: una de ellas la predicha y otra 88bp más corta. La señal de poliadenilación presente en aproximadamente un 50% de los mRNAs de *D. melanogaster* es una secuencia AAUAAA unos 30 nucleótidos antes de el sitio de poliadenilación. Sin embargo, ésta no es la única señal que se puede encontrar. En la nueva variante para el 3' UTR descrita por nosotros, hay una secuencia AUUAAA, que podría ser usada dando lugar a un transcrito más corto. La señal de poliadenilación AUUAAA es usada en aproximadamente un 10% de los mRNAs de *D. melanogaster* (Retelska *et al.*, 2006). Podría ser que las dos variantes del 3' UTR contuvieran, como en el caso de las variantes del 5' UTR, diferentes sitios de unión para factores en *trans* que podrían regular tanto la traducción como la localización subcelular del transcrito (Clyde y Bownes, 2000).

1.1.2. Predicciones sobre su posible función

El gen *CG32795* así como la proteína que codifica (*CG32795*) están predichos bioinformáticamente. Se desconoce su función, aunque se cuenta con la línea EP3617^{EP3617} (Rorth *et al.*, 1998; Dominski *et al.*, 2002) que sobreexpresando *CG32795* bajo el controlador *ey-GAL4* (Hauck *et al.*, 1999), da un fenotipo de ojos reducidos que puede ser rescatado mediante la sobreexpresión de la ciclina E (*cycE*).

Con la intención de aportar información acerca de la posible función de *CG32795*, decidimos generar una nueva cepa de *D. melanogaster* que tuviera el gen *CG32795* deletado. Mediante una serie de cruces, el Dr. Oriol Cabré consiguió una cepa con una deleción en el cromosoma X que abarcaba el gen *w* y el gen *CG32795* enteros y con una copia del gen *w* (incluidos sus *enhancers*) en el cromosoma 2 (ver V.1. Cepas de *Drosophila melanogaster*). Dicha cepa (CGnulw3), no presenta ninguna alteración del fenotipo visible a ojo desnudo y sus individuos son viables y fértiles. Así podemos decir que el gen *CG32795* no es vital.

En segundo lugar, nos decantamos por una aproximación bioinformática. Las secuencias nucleotídicas que obtuvimos al secuenciar los diferentes clones de las tres cepas estudiadas, se ensamblaron dando lugar a distintos cDNAs completos. Sin embargo, al traducirlos, todos ellos daban lugar a una única secuencia de 387 aminoácidos, idéntica a la secuencia predicha en *FlyBase* para las tres isoformas de la

proteína (CG32795-A, CG32795-B y CG32795-C). Así, decidimos usar diferentes herramientas bioinformáticas para intentar predecir la estructura y posible función de dicha proteína. Los resultados obtenidos con ellas son los siguientes:

- SOUSI → Es un sistema de predicción de la estructura proteica que permite, a partir de la secuencia aminoacídica, discriminar entre proteínas solubles y proteínas transmembranales. Además si la proteína es transmembranal predice las hélices que se pueden formar. El sistema SOUSI asume tres supuestos básicos: 1 – Las proteínas membranales se caracterizan por al menos un dominio altamente hidrofóbico, la hélice transmembranal principal. 2 – Pueden existir hélices transmembranales secundarias, pero su hidrofobicidad será parecida a la de los dominios hidrofóbicos de las proteínas solubles. 3 – La hélice transmembranal primaria se estabiliza gracias a una combinación de cadenas laterales anfifílicas (moléculas que contienen un grupo polar ligado a una cadena no polar) y una región central altamente hidrofóbica. Los cuatro parámetros que toma en cuenta el programa son: el índice de hidrofobicidad, el índice de anfifilicidad, el índice de cargas aminoacídicas y la longitud de la secuencia. Asegura una precisión >99% para la predicción de proteínas membranales y de ~97% para proteínas solubles. El programa SOSUI predijo que la proteína CG32795 era transmembranal y que poseía seis hélices transmembranales, dos de ellas primarias (ver Figura III.5.).

La secuencia aminoacídica de las hélices así como el aminoácido N-terminal y C-terminal, el tipo de hélice y la longitud, se especifican en la siguiente tabla (ver Tabla III.1.):

Hélice	N-ter.	Zona transmembranal	C-ter.	Tipo	Long.
1	131	KLILNVIGLIMAFFNLI FN YRAL	153	Primaria	23
2	155	LAFIFLLVWYYCTLTIRE SILKV	177	Primaria	23
3	186	WRAH HFISTVAAGVLLVWPQGE	207	Secundaria	22
4	217	MYFYISIVQYLQFGYQKGLLYR	238	Secundaria	22
5	262	GLSFLLPFLFIGYGYQAYNAWTL	284	Secundaria	23
6	298	VSVMSGLFLLL FVGNMATT LWV	320	Secundaria	23

Tabla III.1. – **Detalle de las hélices transmembranales predichas por el programa SOSUI.** Las columnas N-terminal y C-terminal indican el aminoácido en el que se inicia y termina la hélice predicha.

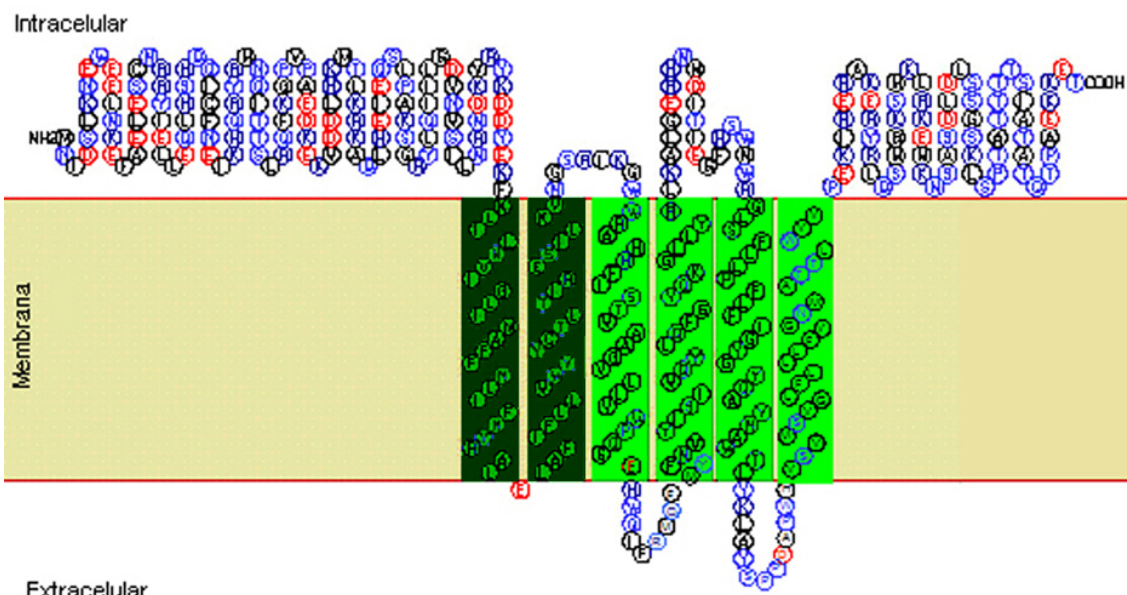


Figura III.5. – Esquema gráfico de la predicción de hélices transmembranales para la proteína CG32795 hecha por SOSUI. Las dos primeras hélices (en verde más oscuro) serían primarias, mientras que las otras cuatro (verde claro) serían secundarias.

- TMHMM (*TransMembrane Hidden Markov Model*)→ El modelo oculto de Markov se compone de una serie de submodelos que se utilizan para modelar las distintas regiones de una proteína de membrana. El programa permite la transición de un submodelo a otro, de forma que la longitud de las distintas regiones se puede ajustar. Además aunque se trate de un programa para determinar la topología de proteínas transmembranales, también permite discriminar entre proteínas transmembranales y solubles. El programa TMHMM confirmó los resultados obtenidos anteriormente con el programa SOSUI, y predijo que CG32795 es una proteína con seis hélices transmembranales (ver Figura III.6.).

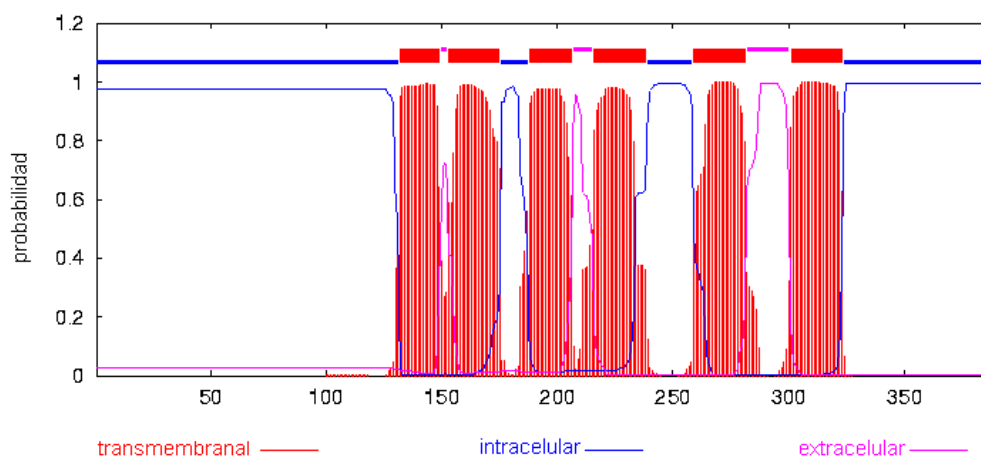


Figura III.6. – Esquema gráfico de la predicción de hélices transmembranales para la proteína CG32795 hecha por TMHMM. Las líneas azules indican regiones intracelulares, mientras que las rosas indican las regiones extracelulares. Las zonas rojas representan las hélices transmembranales.


```

seq  DITIEGFHSW MWRGLSFLLP FLFIGYGYQA YNAWTLYKLA YSPDPAPWHV 300
pred  iiiiiiiiii iiiHHHHHHH HHHHHHHHHH Hooooooooo ooooooooooH

seq  SVMSGLFLLL FVGNMATTW VVPEKIRERA KERYRLQSMG KSMKLRKEMK 350
pred  HHHHHHHHHH HHHHHHHHHH HHiiiiiii iiiiiiiiii IIIIIIIIII

seq  NSASDLDLSS GSKLSPTATT TTSIATATQT PAEKKET 387
pred  IIIIIIIIII IIIIIIIIII IIIIIIIIII IIIIIIII

```

Siendo I – Intracelular, H – Hélice, O – Extracelular. Las mayúsculas representan una mayor probabilidad de que el aminoácido pertenezca a la categoría asignada, mientras que las minúsculas representan regiones de transición entre dos categorías.

La secuencia aminoacídica de las hélices así como el aminoácido N-terminal y C-terminal y la longitud de la hélice, se especifican en la siguiente tabla (ver Tabla III.3.):

Hélice	N-ter.	Zona transmembranal	C-ter.	Longitud
1	132	LILNVIGLIMAFFNLIF	148	17
2	153	LELAFIFLLVWYYCTLTI	170	18
3	185	WWRAHHFISTVAAGVLLV	202	18
4	217	MYFYISIVQYLQFGY	233	15
5	264	GLSFLLPFLFIGYGYQAY	281	18
6	300	SVMGSLFLLL FVGNMATTWV	322	22

Tabla III.3. – **Detalle de las hélices transmembranales predichas por el programa HMMTOP.** Las columnas N-terminal y C-terminal indican el aminoácido en el que se inicia y termina la hélice predicha.

La predicción de seis hélices transmembranales y de extremos C-terminal y N-terminal intracelulares para la proteína CG32795, hecha por el programa HMTOP, coincide perfectamente con los resultados obtenidos con los programas SOSUI y TMHMM. Sin embargo, los tres modelos propuestos presentan ligeras variaciones respecto al extremo C-terminal y N-terminal de cada hélice.

Es importante recordar en todo momento que, aunque los tres programas nos den una estructura proteica similar, se trata sólo de modelos que deberían ser comprobados y refinados mediante metodologías experimentales. Por ejemplo, en el primer modelo

(SOSUI) nos encontramos con una incompatibilidad, puesto que entre las hélices 1 y 2 existe un único aminoácido. Por motivos estéricos es imposible la formación de un lazo extracelular que permita el empaquetamiento de dos hélices alfa volviendo a entrar en dirección antiparalela en la membrana, lo cual requiere un mínimo de 3-6 aminoácidos.

La topología predicha para CG32795 nos recordó la estructura de los canales de sodio y potasio dependientes de voltaje, no sólo por las seis hélices transmembranales sino también por la presencia de un gran lazo externo entre las hélices cinco y seis, que recuerda la región P de dichos canales. Sin embargo, la cuarta hélice carece del “sensor de voltaje” típico los canales dependientes de voltaje (dominio de 4 a 7 aminoácidos positivos (R o K) separados por dos o tres aminoácidos, al menos dos de ellos hidrofóbicos).

Se puede descartar también que se trate de un canal de potasio, pues no posee la secuencia canónica GYGD en la denominada región H5 o región del poro, entre las hélices 5 y 6. Esta secuencia canónica constituye la “firma” para la permeación selectiva del potasio. En algunos canales de potasio (ej. h-ERG) esta firma varía a GFGN, pero en cualquier caso las dos G son necesarias, pues sino de la cadena principal de esa zona (que constituye la pared del poro selectivo) sobresaldrían las cadenas laterales y el poro estaría taponado. Cabe destacar la presencia de la secuencia GYGY en medio de la quinta hélice. No se encuentra en la posición correcta para hacer de región del poro, puesto que está en medio de una de las hélices. Sin embargo, se podría hipotetizar que las hélices 4 y 6 sean realmente las hélices 5 y 6. Nos encontraríamos entonces ante un canal de tipo rectificador anómalo que abre a voltajes negativos y no al despolarizar y que contiene dos únicas hélices transmembranales, la quinta y la sexta (ej. canal Kir, Nichols y Lopatin, 1997). No obstante, esta hipótesis es un poco arriesgada, teniendo en cuenta, no sólo las predicciones bioinformáticas para la topología proteica de CG32795, sino el hecho de que al hacer un BLAST con CG32795 no se obtiene ninguna similitud con ningún canal Kir de dos hélices ni nada parecido.

No se puede descartar que se trate de un canal de sodio o de otro ion (ej. Inespecífico de cationes). Sin embargo, si se trata de un canal de sodio, sería un canal “ancestral”, puesto que los canales más evolucionados presentan las cuatro repeticiones de seis hélices en un solo polipéptido (ej. los canales de los axones).

Funcionalmente, parece improbable que se trate de ningún canal convencional de seis hélices, porque en todos ellos el poro está siempre entre las hélices 5 y 6, en una región llamada H5. En nuestras predicciones, la región H5 es bastante más corta de lo habitual. Además la presencia de tres prolinas en una secuencia de cinco aminoácidos (PPDAP) hace improbable que la región H5 se pliegue a modo de los poros convencionales.

En relación a posibles dominios funcionales se encontró con el programa NCBI CD, una similitud importante (valor $E = 2 \times 10^{-160}$) desde el aminoácido 5 hasta el 333 de CG32795, con la familia proteica TMPIT (*TransMembrane Protein Induced by Tumor necrosis factor alpha*). El dominio TMPIT se conserva en metazoos a lo largo de toda la evolución. Sin embargo, no se tiene ninguna información a cerca de su función así como tampoco de las rutas metabólicas en las que participa.

1.2. Cuantificación de la expresión de los genes *white* y *CG32795*

La inserción de un elemento *FB-NOF* cercana al extremo 3' del gen *w* se había propuesto como principal motivo para explicar el fenotipo de ojos zeste en los machos de la cepa w^{+UZ} (Rasmuson-Lestander y Ekstrom, 1996). Sin embargo, posteriormente nuestro grupo describió una duplicación del ZBS de *w* en la cepa M115. Dicha cepa poseía también machos de ojos zeste y una inserción de un elemento *FB-NOF* en el mismo punto que la cepa w^{+UZ} . La región duplicada no afectaba exclusivamente al ZBS del gen *w* sino a todo el gen *w* y al extremo 5' del gen *CG32795* (Badal *et al.*, 2006a). Por ese motivo, y conociendo de múltiples casos en los que la inserción de un ET altera la expresión de los genes circundantes al punto de inserción, decidimos estudiar que efecto tiene la inserción de *FB-NOF* sobre la expresión de los genes localizados alrededor del punto de inserción: *w* y *CG32795*.

El estudio de la expresión se hizo con PCR a tiempo real, analizando tres regiones distintas: 1 – Región *w* → Localizada entre el segundo y tercer exón del gen *w* (ver en naranja Figura III.7.). 2 – Región *CG5* → Localizada entre el segundo y el tercer exón del gen *CG32795*, en 5' del punto de inserción del elemento *FB-NOF* en la cepa M115 (ver en azul claro Figura III.7.). 3 – Región *CG3* → Localizada entre el séptimo y el

octavo exón del gen *CG32795*, en 3' del punto de inserción del elemento *FB-NOF* en la cepa M115 (ver en azul oscuro Figura III.7.).

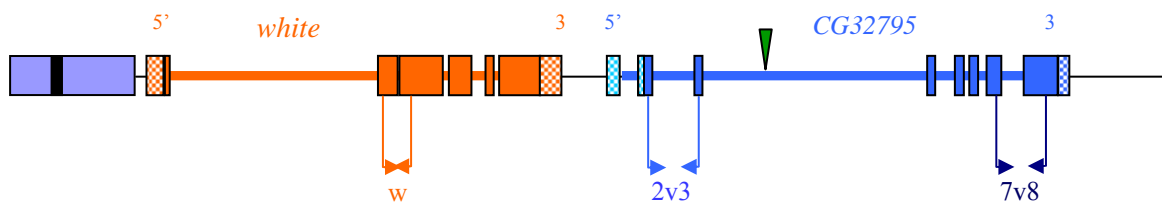


Figura III.7. – **Regiones amplificadas para el estudio de la expresión de los genes *w* y *CG32795*.** Las flechas representan los cebadores usados para cada reacción de PCR (descritos en V.3.1.2. Cebadores y reacción). El triángulo verde en el tercer intrón del gen *CG32795* señala el punto de inserción del elemento *FB-NOF* en la cepa M115.

1.2.1. El gen *w*

Originalmente el estudio se realizó a partir de mRNA total de machos adultos enteros de las cepas CS, z^1 , M115, RM115 y Transw3. Se utilizó el gen de la actina como control, puesto que previamente demostramos que se expresaba igual en las cinco cepas de interés. La expresión de las regiones problema se normalizó respecto a la expresión de la actina en cada cepa, y posteriormente se compararon los valores obtenidos para cada región problema entre las cinco cepas. La expresión en la cepa CS se consideró siempre del 100%. Así se obtuvo una cuantificación relativa de la expresión del gen *w*, permitiéndonos establecer relaciones entre las cinco cepas estudiadas (ver Figura III.8.).

Considerando la cantidad de mRNA de *w* en la cepa CS como el 100%, se observa que las cepas z^1 con dos copias de *w* (M115 y Transw3) tienen aproximadamente el doble de mRNA de *w*, y sin embargo poseen un fenotipo de ojos amarillos. Dicho fenotipo se explica por la presencia de dos ZBS de *w* en una cepa z^1 , permitiendo así el apareamiento de los ZBS de *w* y causando la interacción *zeste-white* que sólo afecta a la expresión de *w* en ojos. Nuestra propuesta es que el gen *w* vea reducida su expresión en los ojos, pero no en el resto de tejidos, donde se expresaría el doble, dando un valor de mRNA de *w* en todo el cuerpo de aproximadamente el doble de la cepa salvaje (ver Figura III.8.).

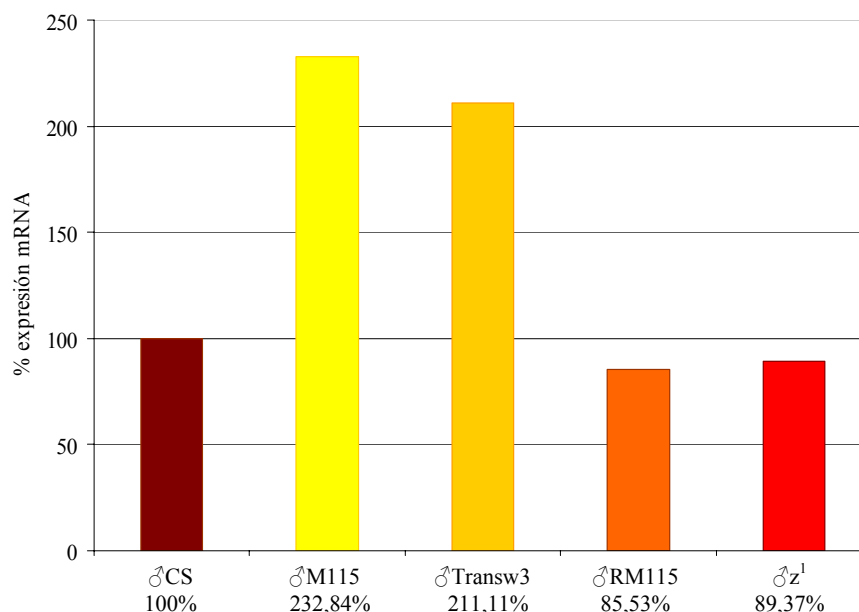


Figura III.8. – **Cuantificación de la expresión del gen *w*.** El mRNA total se obtuvo a partir de machos adultos de *D. melanogaster*. Considerando la expresión del gen *w* en CS como el 100%, se observa que las cepas z^1 con dos copias de *w* expresan el doble respecto a la cepa salvaje, y sin embargo poseen un fenotipo de ojos amarillos. Las cepas z^1 con una única copia del gen *w*, expresan aproximadamente lo mismo que la cepa salvaje CS, aunque se observa una ligera disminución de la expresión.

Las cepas z^1 con una única copia del gen *w* (RM115 y z^1), poseen aproximadamente la misma cantidad de mRNA de *w* que la cepa salvaje CS (ver Figura III.8.). Sin embargo, al tratarse de cepas z^1 también se observa una ligera disminución de la expresión de *w*, aunque ésta no pueda ser percibida *de visu* en los ojos de los individuos. La reducción de la expresión de *w* en los machos z^1 es un fenómeno que concuerda con las observaciones anteriores que describen un fenotipo de ojos zeste, tanto para machos como para hembras, en individuos con la mutación K425L en Zeste. Este cambio aminoacídico permite que se formen hiperagregados más grandes que los de Z^1 , dando lugar a un fenotipo visible incluso sin la presencia de ZBS de *w* apareados. La sustitución K425V, da lugar a una forma de Zeste que produce hiperagregados menores que los de Z^1 . En este caso, los machos presentan ojos rojos y las hembras ojos marrones (Chen y Pirrotta, 1993b). Así, la interacción *zeste-white* no depende exclusivamente de la presencia de copias apareadas del ZBS de *w*, sino que se puede ver modificada por los diferentes alelos de *z*. Nuestros resultados sugieren que los machos z^1 presentan una reducción de la expresión de *w*, que no puede ser detectada de vista en el fenotipo ocular, pero sí a nivel de mRNA. Dicha reducción se explica por la formación de hiperagregados de Z^1 en la única copia del ZBS de *w* presente. Estos

agregados serían menores que los que se forman en caso de encontrar copias apareadas del ZBS de w , pero aún así dificultarían el inicio de transcripción y por tanto causarían una reducción de la expresión del gen w .

En los individuos adultos de *D. melanogaster* los tejidos de la cabeza representan un 10% de la masa corporal y los del tórax-abdomen un 85% (Bingham y Zachar, 1985). Así, los efectos detectados en estudios con individuos enteros pueden no reflejar lo que está ocurriendo en tejidos específicos. Por este motivo, habiendo estudiado la expresión del gen w en individuos enteros, pensamos que se podría obtener una información más precisa realizando el mismo experimento en cabezas, tóraxs y abdómenes por separado. Además, esta nueva aproximación nos permitiría confirmar nuestra hipótesis a cerca de los valores de $\sim 200\%$ de expresión de w en las cepas con dos copias del gen.

Empezamos estudiando la expresión del gen w en cabezas de las cinco cepas estudiadas anteriormente. Los principales tejidos de la cabeza donde se expresa el gen w son los ojos y los ocelos. Los resultados obtenidos fueron los siguientes:

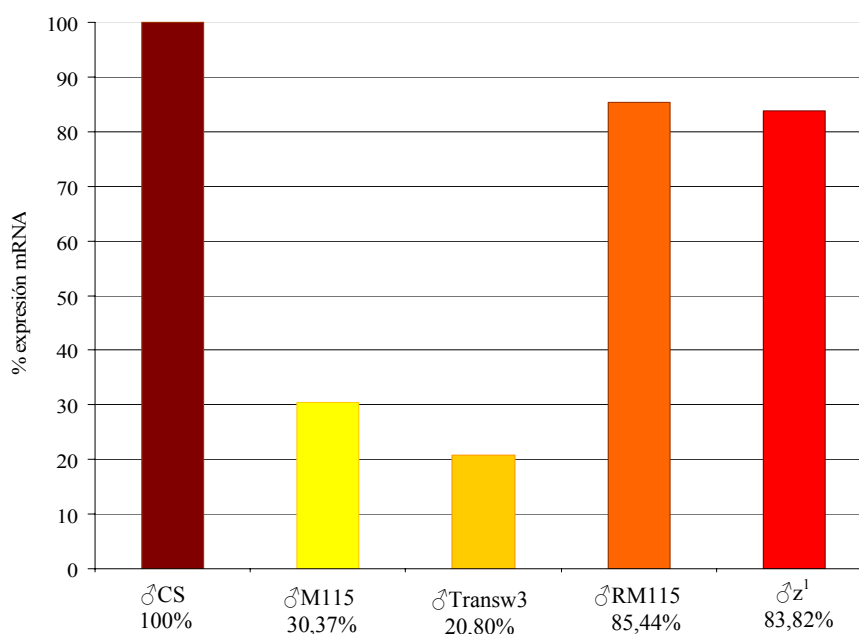


Figura III.9. – **Cuantificación de la expresión del gen w en cabezas.** El mRNA total se obtuvo a partir de cabezas de machos adultos de *D. melanogaster*. Considerando la expresión del gen w en CS como el 100%, se observa que las cepas z^1 con dos copias de w expresan $\sim 25\%$ respecto a la salvaje, debido al efecto de la interacción *zeste-white*. Las cepas z^1 con una única copia del gen w , expresan aproximadamente lo mismo que la cepa salvaje CS.

Se observa que las cepas z^1 con dos copias del gen w (M115 y Transw3) presentan una reducción drástica de la expresión de dicho gen en la cabeza, como era de esperar por su fenotipo de ojos amarillos. Estos resultados confirman parcialmente nuestra hipótesis sobre la expresión del gen w en mutantes z^1 con dos copias para dicho gen, demostrando que aunque la expresión global del gen w parece ser el doble que en presencia de una única copia, la expresión se ve reducida en ojos y ocelos. La reducción de la expresión de w en estos tejidos es debida a la interacción *zeste-white*. Las cepas z^1 con una única copia de w (RM115 y z^1) presentan una ligera disminución de la expresión de w . Tal y como sugerimos para los resultados en individuos enteros, dicha reducción podría ser debida al efecto de Z^1 . *Zeste* actúa como factor de transcripción, reclutando el complejo BRM y permitiendo la abertura de la estructura de la cromatina para la posterior transcripción del gen. Puede que Z^1 no sea tan eficaz como Z y por tanto, la transcripción, aunque se da y en altos niveles, presente sus dificultades.

Aunque no existe una relación lineal entre la expresión del gen w y la cantidad de pigmento depositada en los ojos, decidimos comprobar si la reducción de la expresión de w detectada en cabezas se percibía también en la cantidad de pigmentos oculares, aunque no fuéramos capaces de detectarlo *de visu*. Con este propósito realizamos una cuantificación de pigmentos oculares. Los resultados obtenidos fueron los siguientes:

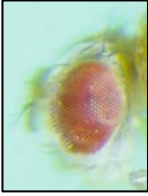
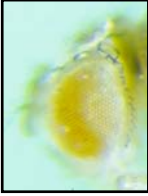
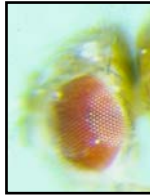
	♂CS	♂M115	♂RM115
Fenotipo ocular			
Pigmentos oculares	100%	4,36%	90,22%
Expresión del gen w	100%	30,37%	85,44%

Tabla III.4. – **Resultados de la cuantificación de los pigmentos oculares.** Se muestran los fenotipos oculares de las tres cepas estudiadas así como la relación entre las absorbancias de los pigmentos oculares y los resultados de cuantificación por PCR a tiempo real, considerando CS como el 100%.

Claramente no existe una relación lineal entre el fenotipo ocular que se observa y la cantidad de pigmento medida. Sin embargo, si que se aprecia una ligera disminución de la cantidad de pigmento en los machos RM115, que, aunque no puede ser observada a ojo desnudo, si se corresponde con la disminución de la expresión del gen *w* observada por PCR a tiempo real.

En segundo lugar, decidimos estudiar la expresión del gen *w* en tórax. Teóricamente, el tórax no contiene ningún tejido que pueda expresar el gen *w*. Así pues, el estudio de expresión del gen *w* en tórax era para nosotros un control negativo que nos mostraría los errores intrínsecos de la propia técnica. Los resultados obtenidos fueron los siguientes:

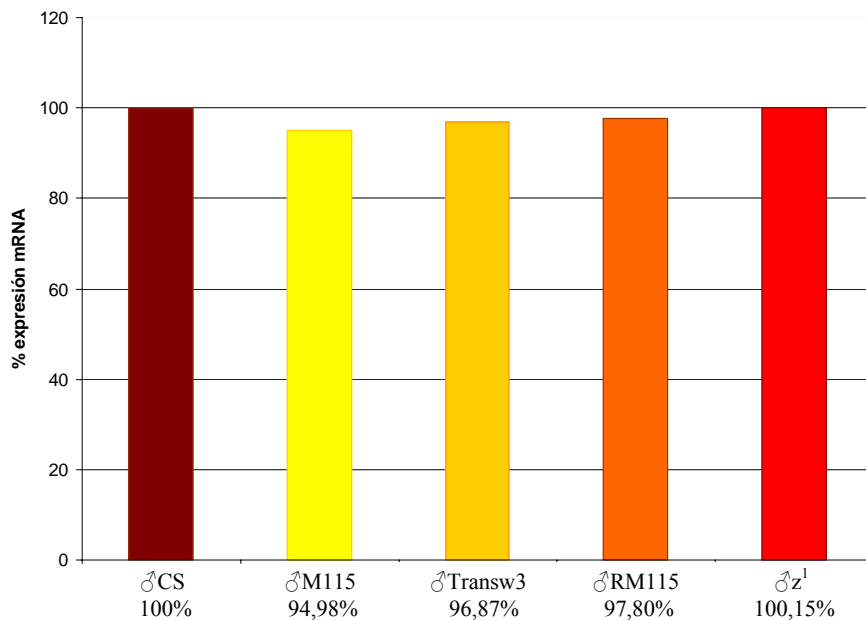


Figura III.10. – **Cuantificación de la expresión del gen *w* en tórax.** El mRNA total se obtuvo a partir de tórax de machos adultos de *D. melanogaster*. Considerando la expresión del gen *w* en CS como el 100%, se observa que el resto de cepas presentan una expresión similar. La expresión de *w* detectada en tórax es residual, mínima, puesto que el tórax no presenta ningún tejido que exprese *w*.

La expresión del gen *w* en tórax es la misma en las cinco cepas estudiadas aunque, como ya mencionamos anteriormente, en esta región del cuerpo no se encuentra ningún tejido que exprese *w*. La expresión detectada puede ser debida probablemente a restos de tejidos (mayoritariamente testículos) que se hayan podido arrastrar a la hora de cortar los fragmentos. El error cometido es aproximadamente el mismo en todas las cepas.

Además, al ser un error tan pequeño, la PCR a tiempo real no nos permite diferenciar entre las cepas con una o dos copias de *w*. Sin embargo, es imposible descartar que se produzca una expresión basal de *w* en algún tejido del tórax que no haya sido descrita con anterioridad.

En tercer lugar, estudiamos la expresión del gen *w* en abdómenes. El abdomen de los machos de *D. melanogaster* contiene dos tejidos que expresan el gen *w*: los testículos y los túbulos de Malpighi. Los resultados obtenidos fueron:

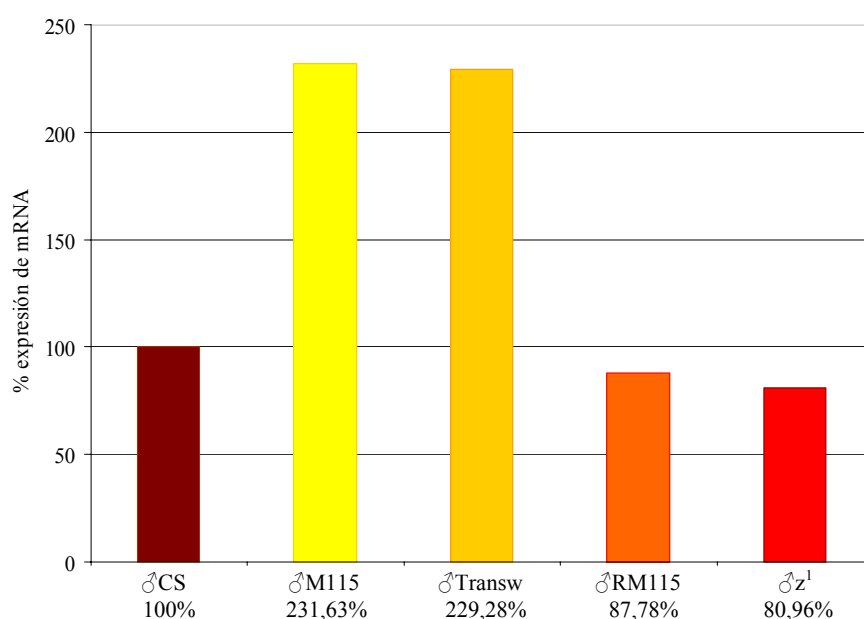


Figura III.11. – **Cuantificación de la expresión del gen *w* en abdómenes.** El mRNA total se obtuvo a partir de abdómenes de machos adultos de *D. melanogaster*. Considerando la expresión del gen *w* en CS como el 100%, se aprecia que las cepas *z*¹ con dos copias del gen *w* expresan el doble de mRNA de *w*. Las cepas *z*¹ con una única copia de *w* presentan una ligera reducción de la expresión de *w*.

Se observa que las cepas *z*¹ con dos copias del gen *w* (M115 y Transw3) presentan una expresión de más del doble que la cepa salvaje. Así se demuestra, una vez más, que la interacción *zeste-white* afecta exclusivamente a la expresión del gen *w* en ojos y ocelos, sin modificar la expresión del gen *w* en otros tejidos (testículos y túbulos de Malpighi). Las cepas M115 y Transw3 al presentar dos copias del gen *w* y no verse sujetas al efecto de la interacción *zeste-white* en el abdomen, poseen el doble de mRNA del gen *w*. Las cepas *z*¹ con una única copia del gen *w* (RM115 y *z*¹) presentan una expresión del gen *w* similar a la de la cepa salvaje CS. Sin embargo, se observa una

ligera disminución de la expresión, que no puede ser justificada por la mutación z^1 , puesto que como hemos dicho anteriormente, la interacción *zeste-white* sólo afecta la expresión de *w* en ojos y ocelos.

Igual como sucede en los ojos, tampoco existe una relación lineal entre la expresión del gen *w* y la cantidad de pigmento depositada en los testículos. Aún así, decidimos comprobar si la reducción de la expresión de *w* detectada en abdómenes se percibía también en los pigmentos testiculares. En este caso, tuvimos que realizar una observación a ojo desnudo, pues el reducido tamaño de los testículos dificulta la cuantificación de pigmentos con el mismo protocolo empleado en las cabezas. Los resultados obtenidos fueron los siguientes:




	♂CS	♂M115	♂RM115
Fenotipo testicular			
Expresión del gen <i>w</i>	100%	235,54%	87,78%

Tabla III.5. – **Observación de los pigmentos testiculares.** Se muestran los fenotipos testiculares de las tres cepas estudiadas así como los resultados de cuantificación de la expresión de *w* por PCR a tiempo real, considerando CS como el 100%.

Se observa un incremento de la pigmentación testicular en los individuos M115, de acuerdo con la expresión del 235,54%, respecto la salvaje en CS, del gen *w*. Sin embargo, se hace difícil observar diferencias entre la pigmentación de los testículos de CS y los de RM115, puesto que las diferencias de expresión de *w* no son lo suficientemente grandes.

Finalmente, decidimos estudiar cual era la contribución de cada fragmento estudiado (cabeza, tórax y abdomen) a la expresión global del gen *w*, obtenida en primera instancia. Así, realizamos un último experimento de PCR a tiempo real midiendo la expresión de *w* en cabeza, tórax y abdomen, en la cepa salvaje CS.

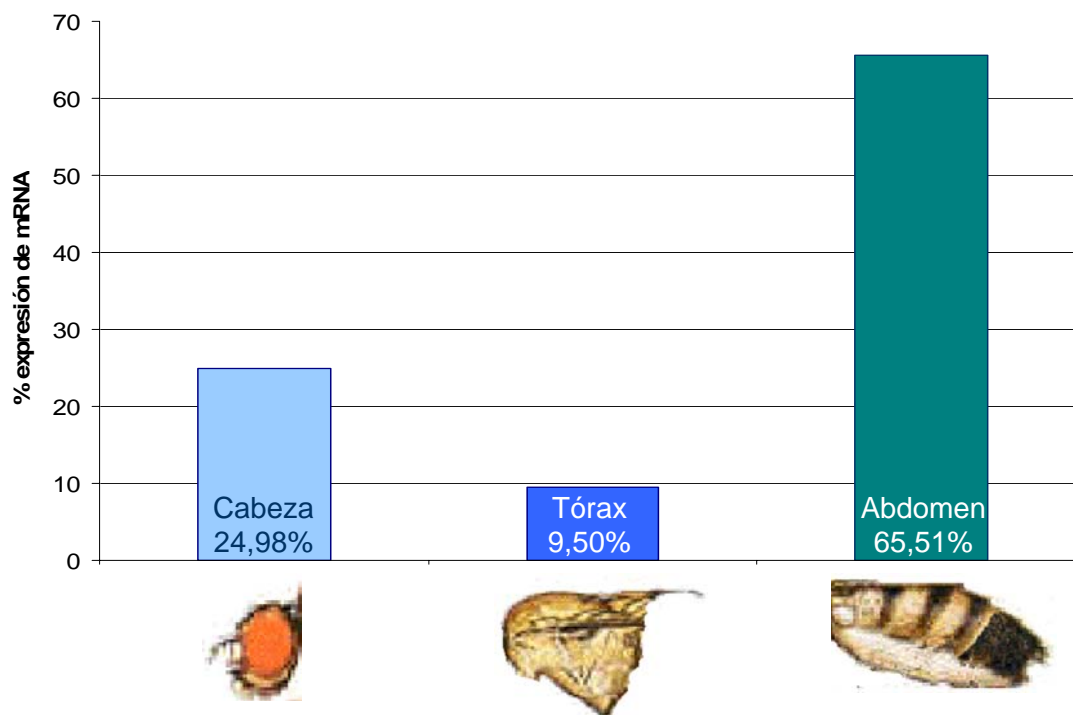


Figura III.12. – **Cuantificación de la expresión del gen *w* en CS.** La contribución de la cabeza a la expresión del gen *w* recae sobre los ojos y ocelos, que son responsables de transcribir un 25% del total de mRNA de *w* del individuo. Los tejidos que expresan *w* en el abdomen de los machos son los túbulos de Malpighi y los testículos. Éstos son responsables de un 65% de la expresión de *w*. El tórax es un control negativo, puesto que no posee ningún tejido que exprese el gen *w*.

Los resultados obtenidos indican que en machos CS de *D. melanogaster* la expresión del gen *w* se produce principalmente en testículos y túbulos de Malpighi, siendo estos tejidos responsables de transcribir el 65% de los mRNAs de *w*. Los ojos y ocelos, localizados en la cabeza, producen el 25% de los mRNAs de *w*. En tórax se supone que no hay ningún tejido que exprese el gen *w* y por ese motivo, se esperaría encontrar 0 transcritos de *w*. Sin embargo, errores a la hora de separar los diferentes fragmentos corporales nos dan como resultado una expresión del 10% de los tejidos, que seguramente no corresponde con la realidad. De todas maneras, no se puede descartar que en el tórax exista una expresión basal de *w*, que aún siendo residual permite la amplificación de su mRNA mediante PCR a tiempo real.

1.2.2. El gen *CG32795*

El estudio de la expresión del gen *CG32795* se realizó en machos de las cepas CS, M115 y RM115. Puesto que anteriormente habíamos demostrado que la mutación *z*¹ no

afectaba la expresión del gen *CG32795* (ver III.2.1. Cuantificación de la expresión de los genes *w*, *CG32795* y *dpp*) consideramos que no era necesario incluir las cepas *z¹* y *Transw3* en nuestro estudio. Los datos de expresión del gen *CG32795* son muy pocos, y se desconoce cuales son los tejidos en los que se expresa. Por este motivo, realizamos nuestro estudio sobre individuos adultos enteros.

Se estudiaron dos regiones del gen *CG32795*: la *CG5* (en 5' más arriba del punto de inserción del ET *FB-NOF* en la cepa M115) y la *CG3* (en 3' más abajo del punto de inserción del ET *FB-NOF* en la cepa RM115). Los resultados obtenidos fueron los siguientes:

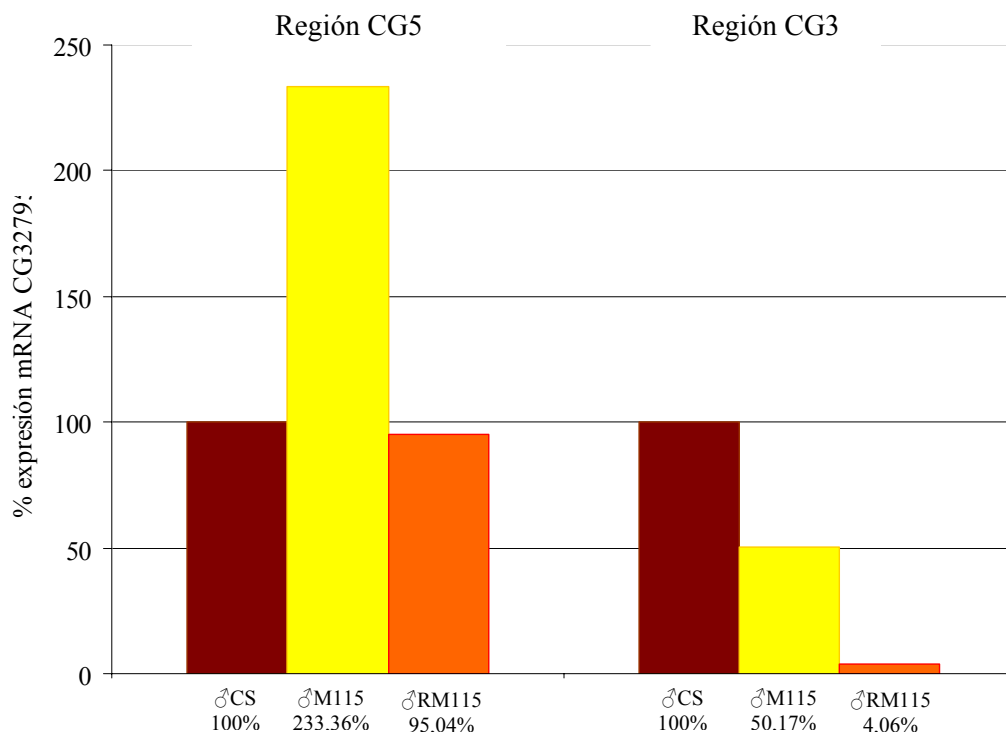


Figura III.13. – **Cuantificación de la expresión del gen *CG32795***. La cuantificación de la expresión se realizó en mRNA total de machos adultos enteros de *D. melanogaster* de las cepas CS, M115 y RM115. La región CG5 corresponde a la amplificación de los exones 2 y 3 localizados en 5' de la inserción de *FB-NOF*. La región CG3 corresponde a la amplificación de los exones 7 y 8 localizados en 3' del punto de inserción de *FB-NOF*.

Los resultados obtenidos muestran claras diferencias entre lo que ocurre en el extremo 5' y en el 3' del gen *CG32795*. En el extremo 5' y tomando la expresión en la cepa salvaje CS como el 100%, la cepa M115 expresa aproximadamente el doble, dato que concuerda con la presencia de dos copias del extremo 5' de *CG32795* (Badal *et al.*,

2006a). La cepa RM115, que posee una única copia del gen *CG32795* expresa aproximadamente lo mismo que la cepa salvaje CS.

Sin embargo, los resultados obtenidos para el extremo 3' del transcrito *CG32795* son completamente distintos. Tomando la expresión de *CG32795* en la cepa CS como el 100%, se observa una reducción radical de la expresión en el resto de cepas estudiadas, siendo la expresión del 50% en la cepa M115 y del 4% en la RM115. Ambas cepas poseen una única copia del extremo 3' del gen *CG32795*, puesto que la duplicación en la cepa M115 sólo se extiende hasta el punto de inserción de *FB-NOF*, en el tercer intrón del gen *CG32795* (ver Figura I.27.) (Badal *et al.*, 2006a). Diversas hipótesis podrían explicar tan drástica disminución de la cantidad de transcritos del extremo 3' de *CG32795* en las cepas M115 y RM115.

La inserción de un ET podría dificultar la obtención del transcrito entero, dando lugar a menos mRNAs que contengan el extremo 3' de *CG32795*. Un caso similar se da en la cepa *w^a* de *D. melanogaster*, en la que un ET *copia* se inserta en el segundo intrón del gen *w* (en el punto 0 según las coordenadas de Levis). Dicha cepa presenta un fenotipo de ojos anaranjados, puesto que la inserción de *copia* aporta nuevos sitios de poliadenilación, permitiendo la producción de RNA aberrantes de 5,7, 3,0 y 1,25Kb y disminuyendo la producción del mRNA de *w* con *splicing* correcto (2,6Kb) (Levis *et al.*, 1984; Pirrotta y Brockl, 1984). En las cepas M115 y RM115 podría ocurrir un fenómeno similar con el gen *CG32795*. La inserción de un elemento *FB-NOF* en el tercer intrón del gen *CG32795* (ver Figura I.27.) aporta un par de nuevas señales de poliadenilación (presentes en la secuencia de *NOF*), que podrían ser utilizados dando lugar a transcritos aberrantes. Estos transcritos aberrantes contendrían el extremo 5' del gen *CG32795* pero se truncarían al encontrar las secuencias de poliadenilación de *NOF* y por tanto, no contendrían el extremo 3' del gen *CG32795*. Esta hipótesis explicaría porque se encuentran los niveles de expresión esperados al comparar con la cepa salvaje en 5', pero se observa una drástica reducción de la expresión en 3'.

Otra hipótesis para explicar el mismo fenómeno, sería la reducción de la procesividad de la RNA polimerasa al encontrarse con la secuencia del elemento *FB-NOF*. Un caso similar fue descrito en mamíferos para los retrotransposones L1, cuyo ORF2 produce truncamiento prematuro del transcrito (cuando se encuentra insertado en

sentido inverso al de la transcripción del gen) o defectos en la elongación transcripcional (cuando se encuentra insertado en sentido directo al de la transcripción del gen) (Han *et al.*, 2004). En los casos de inserción directa, el fenómeno sería parecido al descrito para el elemento *copia* en w^a . En los casos de inserción inversa, se produciría un efecto llamado “reóstato molecular”, según el cual a mayor longitud de las secuencias insertadas, mayor grado de pausas o disociaciones de la RNA polimerasa. Se desconoce cuales son las características del ORF2 de L1 que le permiten actuar como reóstato. En ambos casos, se aprecia una disminución de la producción de pre-mRNAs enteros para el gen en el que se encuentran insertados. Esta hipótesis también concuerda con los resultados obtenidos para la expresión del gen *CG32795* en las cepas M115 y RM115. Así, la inserción de un elemento *FB-NOF* en el tercer intrón de *CG32795*, produciría un efecto similar a la inserción de un elemento L1.

Finalmente, en la cepa RM115, una tercera hipótesis permitiría explicar el efecto, aún más drástico, observado en el extremo 3' del gen *CG32795*: el trans *splicing*. Cabe recordar que en la cepa RM115, el gen *CG32795* no posee continuidad lineal en la secuencia genómica, puesto que las reordenaciones mediadas por *FB-NOF* han situado el extremo 5' del gen en una posición distinta a la original (ver Figura I.27.). Sin embargo, dicha cepa es capaz de producir transcritos enteros de *CG32795*, tal como demostramos al conseguir amplificar del exón 2 al exón 8 (ver III.1.1.1. Patrones de *splicing* alternativo). Dichos mRNAs no pueden ser transcritos de forma continua y por eso proponemos un modelo en el que la transcripción de los extremos 5' y 3' es independiente. Posteriormente un proceso de trans *splicing* permitiría a ambos pre-mRNAs unirse para formar un único mRNA maduro. Un fenómeno similar se observa en los experimentos de Dorn *et al.*, 2001, en los que diferentes exones del gen *mod(mdg4)* clonados en vectores se insertan en diferentes cromosomas y sin embargo son capaces de dar lugar a un mRNA maduro completo. Para ser posible el trans *splicing*, es necesario que todos los fragmentos posean un promotor que permita el inicio de transcripción. En el caso del gen *CG32795*, usando el programa de predicción de promotores *Neural Network Promoter Prediction* (Reese, 2001) (http://www.fruitfly.org/seq_tools/promoter.html), se obtienen tres secuencias promotoras con puntuaciones superiores al 0,8 para la secuencia localizada entre el punto de inserción de *FB-NOF* y el inicio del cuarto exón (ver Tabla III.6.).

Secuencia promotora	Puntuación
atacaaatcaaataaatagcaggctgCGcaatgatgaaacatcagagatt	0,83
tgacgagtcatatagatgtgCaccgaaggTggaacatacgTttatggtaa	0,82
acgcaaagtgtatatattggcccacccgatccttccttgaatGttatgtaca	0,93

Tabla III.6. – **Promotores predichos para el inicio de transcripción de CG32795 en 3' del punto de inserción de FB-NOF.** La puntuación indica la probabilidad que la secuencia actúe como promotor. Cuando más cercano se encuentre el valor a 1, más probable es que dicha secuencia sea un promotor. La base más grande, indica el punto en el que se produciría el inicio de transcripción (+1).

Cabe recordar que las tres secuencias promotoras propuestas también se encuentran presentes en las cepas CS y M115. Puede que en estas cepas dicho promotor se encuentre silenciado y que la reordenación producida en la cepa RM115 lo active. Sin embargo, se trata de una hipótesis que debería ser probada. Otra alternativa es que el trans *splicing* se diera en todas las cepas, pero éste no parece necesario en la cepa M115 y menos en la cepa salvaje CS. Los casos de trans *splicing* descritos con anterioridad en *D. melanogaster*, eran siempre casos en los que dicho proceso era necesario: bien porque los exones se encontraban codificados en cadenas complementarias del DNA (gen *mod(mdg4)*, Dorn *et al.*, 2001; Labrador *et al.*, 2001) o bien porque el gran número de variantes de *splicing* para un solo gen lo hacía la forma más simple de conseguir las (gen *lola*, Ohsako *et al.*, 2003).

Una hipótesis alternativa es la presencia de una secuencia promotora en 5' del punto de inserción de *FB*. En la cepa RM115, la región que se encuentra en 5' del punto de inserción de *FB* no coincide con la secuencia intrónica de *CG32795* (ver Figura I.27.). Se desconoce cual es dicha secuencia, pero sin embargo, no se puede descartar que contenga un promotor que permita la transcripción del extremo 3' de *CG32795* haciendo así posible el trans *splicing*. En caso de que esta hipótesis fuera la cierta, el trans *splicing* se produciría exclusivamente en la cepa RM115, pues tanto la cepa M115 como la cepa CS, no poseerían dicha secuencia promotora en el lugar adecuado para permitir la transcripción del extremo 3' de *CG32795*. La presencia de un promotor en 5' de la inserción de *FB*, implicaría que los transcritos generados contuvieran el elemento *FB*. La presencia de la secuencia *FB* dotaría a los pre-mRNAs de una región de homología con los pre-mRNAs del extremo 5' de *CG32795*. Las repeticiones invertidas de *FB* presentes tanto en los pre-mRNAs del extremo 5' como en los pre-mRNAs del

extremo 3', generarían secuencias complementarias a través de las cuales los dos tipos de pre-mRNAs podrían unirse para favorecer el trans *splicing*, tal y como fue propuesto para el gen *mog(mdg4)* (Labrador *et al.*, 2001). Además el uso, para la transcripción del extremo 3' del gen *CG32795*, de un promotor en 5' de la inserción del elemento *FB* junto con la teoría del “reóstato molecular” (Han *et al.*, 2004), explicaría aún mejor el porqué de una reducción tan drástica para la presencia de transcritos del extremo 3' de *CG32795*.

1.3. Discusión general de los resultados

El estudio molecular de los mutantes M115 y RM115 ha permitido esclarecer muchas de las incógnitas a nivel de expresión que éstos presentaban. Sin embargo, quedan algunas preguntas abiertas que deberían ser estudiadas con profundidad para poder descartar o confirmar las hipótesis propuestas.

En primer lugar, demostramos que la inserción de un elemento *FB-NOF* en la región próxima al extremo 3' del gen *w*, no afecta a la expresión de dicho gen. Los efectos observados en el fenotipo ocular de la cepa M115 (con fondo genético z^1) son debidos a la duplicación de la región que contiene el gen *w*. En la cepa RM115, la reversión al fenotipo ocular salvaje es debida a una reordenación espontánea, que elimina una de las copias de la región de *w* (Badal *et al.*, 2006a). Considerando la expresión del gen *w* en CS como 100%, los resultados obtenidos en las cepas M115 y RM115 se resumen en la siguiente tabla:

♂ cepa	Región	% expresión mRNA de <i>w</i>
M115	Cabeza	30,37
	Tórax	94,98
	Abdomen	235,54
	TOTAL	232,84
RM115	Cabeza	85,44
	Tórax	97,80
	Abdomen	87,78
	TOTAL	85,53

Tabla III.7. – **Resumen de la cuantificación de la expresión del gen *w* por PCR a tiempo real.** Se toma la expresión en la cepa salvaje CS como el 100%.

La cepa M115, con dos copias del gen *w* y de su ZBS y con el alelo mutante z^1 , presenta un fenotipo de ojos amarillos. Sin embargo, no es sorprendente que la expresión de *w* a nivel global sea de alrededor del 230%, puesto que la interacción *zeste-white* que reduce la expresión del gen *w* siempre que se halle un entorno z^1 y se encuentren copias apareadas del ZBS de *w*, sólo se da en los ojos. Así, aunque la expresión de *w* en los ojos de M115 se ve reducida (~30% de expresión del mRNA de *w*), la expresión global del gen se ve altamente incrementada (~230%) gracias a la expresión de las dos copias del gen en testículos y túbulos de Malpighi (~235%).

La cepa RM115, con una única copia del gen *w* y de su ZBS y con el alelo mutante z^1 , presenta un fenotipo de ojos rojos, no distinguible del salvaje a ojo desnudo. Sin embargo, estudios de los niveles de pigmentación ocular nos muestran una reducción de la cantidad de pigmentos en dicha cepa (~90% respecto a la cepa salvaje CS). Cabe recordar que la relación entre la cantidad de pigmento ocular y la expresión del gen *w* no es lineal. Así, la disminución de pigmentos observada, concuerda con la reducción de la expresión de *w* en la cabeza (~85%). Al haber una única copia de *w* y de su ZBS, el efecto de la interacción *zeste-white* no es tan grande como en el caso de tener dos copias, pero aún así, se produce un ligero efecto. Llama la atención el hecho de que dicha reducción de la expresión no se produzca exclusivamente en la cabeza, sino también en el abdomen (~88% de expresión del mRNA del gen *w*), puesto que la interacción *zeste-white* se da en el *enhancer* de los ojos y por tanto, afecta únicamente a la expresión de *w* en ese tejido. Sin embargo, un fenómeno similar ocurre en la cepa z^1 , en la que la expresión de *w* en abdomen también se ve reducida (~89%). Por tanto, y gracias a la presencia de réplicas prácticamente idénticas en el mismo experimento y en experimentos distintos, podemos descartar que se trate de un artefacto.

A diferencia del gen *w*, el gen *CG32795* ve modificada su expresión debido a la inserción del elemento *FB-NOF* en la cepa M115 y a la reordenación producida por el ET en la cepa RM115. Para poder estudiar la expresión del gen *CG32795*, en primer lugar fue necesario caracterizarlo. Así, descubrimos que se trata de un gen no vital que presenta cuatro variantes de *splicing* para el extremo 5' y dos señales de poliadenilación distintas para el extremo 3'. La presencia de tantas variantes no afecta a la proteína final, puesto que en todas ellas el inicio de traducción y el codón de stop son los mismos. El gen *CG32795* codifica una proteína que según las predicciones

bioinformáticas, está compuesta por seis hélices transmembranales. Su estructura recuerda la de un canal de sodio y potasio dependiente de voltaje. Sin embargo, un análisis más minucioso indica que es poco probable que se trate de un canal de potasio o de un canal dependiente de voltaje. Podría tratarse de un canal “ancestral” de sodio, aunque obviamente, es necesario hacer estudios *in vitro* e *in vivo* para poder aportar más información acerca de su estructura y función.

La expresión del gen *CG32795* se cuantificó en el extremo 5’ (antes del punto de inserción de *FB-NOF*) y en el 3’ (después del punto de inserción de *FB-NOF*). Los resultados obtenidos considerando la expresión en la cepa CS como el 100% se resumen en la siguiente tabla:

Cepa	% expresión mRNA CG32795 5’	% expresión mRNA CG32795 3’
♂ M115	233,36	50,17
♂ RM115	95,04	4,06

Tabla III.8. – **Resumen de los resultados de cuantificación por PCR a tiempo real de la expresión del gen *CG32795*.** Se considera la expresión en la cepa CS como 100%.

La cantidad de mRNA del gen *CG32795* en su extremo 5’ es parecida a lo que cabía esperar. En la cepa M115 se encuentran dos copias de dicha región y su expresión es de alrededor del 230%. Dicho valor es muy parecido al valor obtenido en la cuantificación de la expresión de *w* en individuos enteros de las cepas M115 y Transw3. Sorprende que el valor se aproxime a 230% y no a 200%. Los experimentos se repitieron y los resultados obtenidos fueron muy parecidos en todas las ocasiones. Tanto el gen *w* como el gen *CG32795* se encuentran originalmente en el cromosoma X. En *D. melanogaster* las dos copias del cromosoma X de las hembras se encuentran activas. Los machos consiguen la compensación de dosis por hipertranscripción de su único cromosoma X, gracias a la intervención de complejos compensadores de dosis (DCC – *Dosage Compensation Complex*). Los DCCs se encuentran preferencialmente en el cromosoma X y producen una remodelación de la cromatina que permite la sobreexpresión de los genes que allí se encuentran. El mecanismo mediante el cual los DCCs reconocen específicamente el cromosoma X, se desconoce. Sin embargo, se han descrito “sitios de entrada” de los DCCs, formados por secuencias con motivos

complejos y degenerados a los que se unen los DCCs (Straub *et al.*, 2005; Gilfillan *et al.*, 2006). La presencia de dos copias de *w* en el cromosoma X de un macho podría dar lugar a una expresión superior (~230%) a la esperada (200%) justamente por la hiperexpresión del cromosoma X. Sin embargo, el mismo fenómeno se observa en la cepa Transw3, que contiene el gen *w* localizado en el cromosoma 2. La secuencia transgénica insertada en el cromosoma 2 es lo suficientemente amplia como para contener un “sitio de entrada” para DCC. Por este motivo, la hiperactivación de *w* podría darse aunque *w* no se encuentre en el cromosoma X. De todas formas, se trata de una hipótesis preliminar que debería ser probada mediante una serie de experimentos concluyentes.

La cantidad de mRNA del extremo 5' del gen *CG32795* en la cepa RM115, también se aproxima (~95%) a lo que cabría esperar (100%). Dicha cepa posee una única copia del gen *CG32795*, cuya expresión no se ve afectada por el efecto de z^1 .

Los resultados de la cuantificación del mRNA del extremo 3' del gen *CG32795* resultaron mucho más sorprendentes, tanto en la cepa M115 como en la RM115 (ver Tabla III.8.). Ambas cepas presentan una única copia de dicha región. En la cepa M115 dicha copia se encuentra en su posición original, precedida en 5' por la secuencia correspondiente al gen *CG32795*, con la inserción de un elemento *FB-NOF* en el tercer intrón. En la cepa RM115 debido a la reordenación que da lugar a la reversión del fenotipo *zeste*, la copia del extremo 3' del gen *CG32795* no se encuentra precedida en 5' por la secuencia original de *CG32795* sino por una secuencia desconocida dentro de la cual encontramos un elemento *FB* (ver Figura I.27.). La drástica reducción de la expresión del extremo 3' de *CG32795* en M115 (~50% de expresión) y en RM115 (~4% de expresión) puede ser explicada mediante distintas hipótesis: 1 – El correcto *splicing* del transcrito se ve dificultado por la presencia de un ET insertado en el tercer intrón, de forma similar a lo que ocurre en la cepa w^a (Levis *et al.*, 1984; Pirrotta y Brockl, 1984). 2 – Reducción de la procesividad de la RNA polimerasa en presencia de un ET, produciéndose el efecto llamado “reóstato molecular” (Han *et al.*, 2004). 3 – Efecto del trans *splicing*. Este fenómeno debería darse en la cepa RM115, pues el extremo 5' y el 3' de *CG32795* no se encuentran de forma contigua en la secuencia genómica. Así, un fenómeno de trans *splicing* es necesario para que ambos extremos del gen den lugar a un único transcrito del tamaño predicho. El trans *splicing* no sería tan

efectivo como el *splicing*. De esta manera, los transcritos 3' de *CG32795* que no lograran unirse correctamente al transcrito del extremo 5', serían degradados. Este fenómeno de trans *splicing* también podría darse en la cepa M115, pues aunque ésta posee una copia del gen *CG32795* no partida, sino simplemente interrumpida por el elemento *FB-NOF*, también posee otra copia del extremo 5' del gen que no se encuentra contigua a la secuencia del extremo 3'. Así, esta segunda copia del extremo 5' de *CG32795*, podría participar en procesos de trans *splicing*. Cabe destacar que las tres hipótesis no son excluyentes, y que lo más probable es que la menor cantidad de mRNAs del extremo 3' de *CG32795* en las cepas M115 y RM115, sea debida a una mezcla de los mecanismos propuestos en las tres hipótesis.

2. Estudio epigenético del efecto de la mutación *zeste*¹

Los mutantes *z*¹ de *D. melanogaster* se caracterizan por su fenotipo de ojos amarillos siempre que se encuentren copias apareadas del ZBS de *w*, es decir, generalmente en hembras (ver I.3.1. La interacción *zeste-white*). Puesto que *z* pertenece al *trxG* y dicha familia génica es la encargada de establecer y mantener dominios de la cromatina transcripcionalmente activos, pensamos que la mutación *z*¹, podía de algún modo modificar la función de *z* más allá del fenotipo ocular. La proteína Zeste se une al ZBS de *w* formando agregados que cuya función es reclutar el complejo remodelador de la cromatina BRM. Debido a una mutación puntual que elimina el único aminoácido cargado de una región y crea un nuevo dominio hidrofóbico, la proteína mutante Zeste¹ se une al ZBS de *w* formando hiperagregados. Consideramos interesante estudiar en algunos genes con ZBS si los niveles de expresión, el posicionamiento de nucleosomas y/o la metilación del DNA, se veían afectados en los mutantes *z*¹.

2.1. Cuantificación de la expresión de los genes *w*, *CG32795* y *dpp*

Las hembras mutantes *z*¹ presentan un fenotipo de ojos amarillos. De dicho fenotipo se desprende que la expresión del gen *w* en dichas hembras debe de estar reducida. Decidimos comprobar mediante un experimento de PCR a tiempo real, si era cierto. A sólo 700bp del extremo 3' del gen *w* se encuentra el gen *CG32795*, cuya función se desconoce haciendo imposible detectar a nivel fenotípico observar si dicho gen también ve reducida su expresión. Consideramos que sería interesante incluir el gen *CG32795* en el estudio de PCR a tiempo real, comprobando así si el efecto de la interacción *zeste-white* se extiende más allá del gen *w*. Finalmente decidimos incorporar un último gen en el estudio de la expresión: el gen *dpp*, puesto que posee un ZBS en su primer exón. El gen *dpp* debía indicarnos si los posibles efectos observados en *w* eran exclusivos de dicho gen o también podían darse en otros genes con ZBS.

El estudio de la expresión se realizó mediante un experimento de PCR a tiempo real, analizando cuatro regiones distintas: 1 – Región *w* → Localizada entre el segundo y tercer exón del gen *w* (ver en naranja Figura III.7.). 2 – Región *CG5* → Localizada entre el segundo y el tercer exón del gen *CG32795* (ver en azul claro Figura III.7.). 3 – Región *CG3* → Localizada entre el séptimo y el octavo exón del gen *CG32795* (ver en

azul oscuro Figura III.7.). 4 – Región *dpp* → Localizada entre el primer y el segundo exón del gen *dpp* (ver Figura III.14.).



Figura III.14. –Región amplificada para el estudio de la expresión del gen *dpp*. Las flechas representan los cebadores usados para la reacción de PCR (descritos en V.3.1.2. Cebadores y reacción).

El estudio se realizó a partir de mRNA total de hembras adultas de las cepas CS y z^1 . Se utilizó el gen de la actina como control, puesto que previamente demostramos que se expresaba igual en las dos cepas de interés. La expresión de las regiones problema se normalizó respecto a la expresión de la actina en cada cepa, y posteriormente se compararon los valores obtenidos para cada región problema entre las dos cepas. La expresión en la cepa CS se consideró siempre del 100%. Así se obtuvo una cuantificación relativa de la expresión del gen *w*, *CG32795* y *dpp* permitiéndonos establecer relaciones entre las dos cepas estudiadas (ver Figura III.15).

Respecto al gen *w* se observa que las hembras z^1 mantienen sólo un 17,74% del mRNA de dicho gen (ver Figura III.15.). En hembras el gen *w* se expresa en los ojos y en los túmulos de Malpighi, aunque el tamaño de éstos hace que la expresión de *w* en ese tejido pueda ser menospreciada en comparación con la cantidad de mRNA de *w* que se encuentra en los ojos. Así, viendo el fenotipo de ojos amarillos de las hembras z^1 , era de esperar una reducción drástica de la expresión de *w*. Dicho fenotipo se explica por la presencia de dos ZBS de *w* en una cepa z^1 , permitiendo así el apareamiento de los ZBS de *w* y causando la interacción *zeste-white* que afecta a la expresión de *w* en ojos.

La expresión del gen *CG32795* fue estudiada en dos puntos: uno en la región 5' del gen y otro en la región 3'. Los resultados obtenidos, como cabía esperar, fueron similares para los dos puntos: la región 5' presentaba un valor de expresión de 92,84%, mientras que los resultados de la región 3' indicaban que la expresión era del 95,41% (ver Figura III.15.). Los genes *CG32795* y *w* forman un par de genes *head-to-tail*, puesto que su extremo 5' se encuentra a menos de 700bp del extremo 3' del gen *w*, lo

que puede hacer pensar que su expresión quizá se vea afectada, en las hembras z^1 , por la interacción *zeste-white*. Sin embargo, los resultados obtenidos demuestran claramente que eso no ocurre, y que la expresión del gen *CG32795* es completamente independiente de la del gen *w* y de la interacción *zeste-white*.

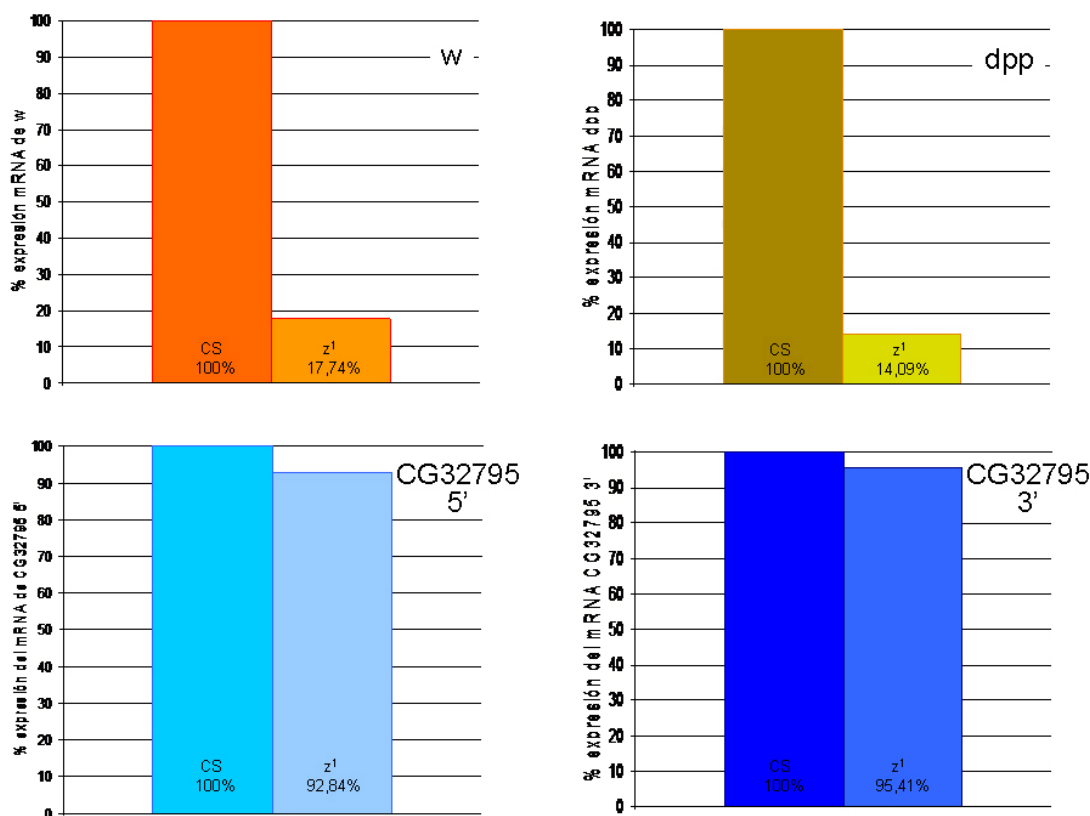


Figura III.15. – **Resultados de la cuantificación de la expresión de los genes *w*, *CG32795* y *dpp* en hembras de las cepas CS y z^1 de *D. melanogaster*.** El mRNA total se obtuvo a partir de hembras adultas de *D. melanogaster*. Considerando la expresión en CS como el 100%, se observa que la cepa z^1 presenta una expresión reducida de los dos genes con un ZBS (*w* y *dpp*). Sin embargo, el gen *CG32795*, aunque se encuentra localizado muy cercano a *w*, no ve afectada su expresión. Cabe destacar que los cuatro gráficos son independientes y no se pueden establecer comparaciones entre ellos. La única comparación posible es entre las dos cepas, para una misma región.

El estudio de la expresión del gen *dpp* dio unos resultados muy parecidos a los obtenidos en *w*: las hembras z^1 expresan un 14,09% del mRNA de *dpp*, tomando la expresión en las hembras CS como el 100%. Como *w*, el gen *dpp* posee un ZBS, pero a diferencia del ZBS de *w* que se encuentra en la zona reguladora del gen, el ZBS de *dpp* se encuentra en el primer exón de este gen. Sin embargo, los resultados obtenidos para la expresión de ambos genes en las hembras z^1 son parecidos. Así, la mutación z^1 no afecta única y exclusivamente la expresión del gen *w*, sino que también modifica la

expresión de otro gen con un ZBS (*dpp*), aunque este hecho no se pueda observar a ojo desnudo en el fenotipo de las hembras mutantes.

2.2. Posicionamiento de nucleosomas

La reducción de la expresión de los genes con un ZBS en las hembras z^1 y el hecho que la proteína Z sea la encargada de reclutar el complejo remodelador de la cromatina BRM, nos indujeron a pensar que quizás la expresión de los genes con un ZBS se veía alterada en los mutantes z^1 por problemas en la remodelación de la cromatina, causados por un reclutamiento alterado del complejo BRM a través de los hiperagregados de la proteína mutante Z^1 . Por ese motivo, decidimos estudiar el posicionamiento de nucleosomas en los ZBS de *w* y *dpp*, así como también en el gen *CG32795*.

El estudio del posicionamiento de los nucleosomas se realizó mediante el ensayo de la nucleasa micrococcal (ver V.5. Ensayo de la nucleasa micrococcal) utilizando tres sondas distintas: a) la sonda P4362, para estudiar la región del ZBS de *w*, b) la sonda SP, para estudiar el extremo 5' del gen *CG32795* y el extremo 3' del gen *w* y c) la sonda *dpp*, para estudiar el ZBS del gen *dpp* (ver Figura III.16. B). El mismo filtro fue hibridado con cada una de las sondas y posteriormente deshibridado para prepararlo para el siguiente experimento. Cada hibridación se realizó al menos dos veces. Todas las autoradiografías obtenidas fueron analizadas con el programa TotalLab y los resultados obtenidos se muestran en la Figura III.16. A.

Cuando la cromatina es digerida con nucleasa micrococcal (MNasa), las regiones de DNA (unos 140-150bp) en las que se encuentra un nucleosoma posicionado quedan protegidos y flanqueados por sitios susceptibles de ser digeridos. Sin embargo, cuando se expresan, la mayor parte de los genes sufren cambios en la estructura de la cromatina (Kingston *et al.*, 1996). Los tres genes estudiados se están expresando en los individuos adultos y además dos de ellos (*w* y *dpp*) lo hacen a diferente nivel, dificultando así la interpretación de los resultados obtenidos y la comparación entre ambas cepas (Wu *et al.*, 1979). Los patrones de posicionamiento de nucleosomas obtenidos no fueron muy estrictos en ninguna de las tres regiones estudiadas, pues los tres genes estaban siendo transcritos; pero aún así, los patrones obtenidos en las cepas CS y z^1 fueron muy parecidos (ver Figura III.16. A).

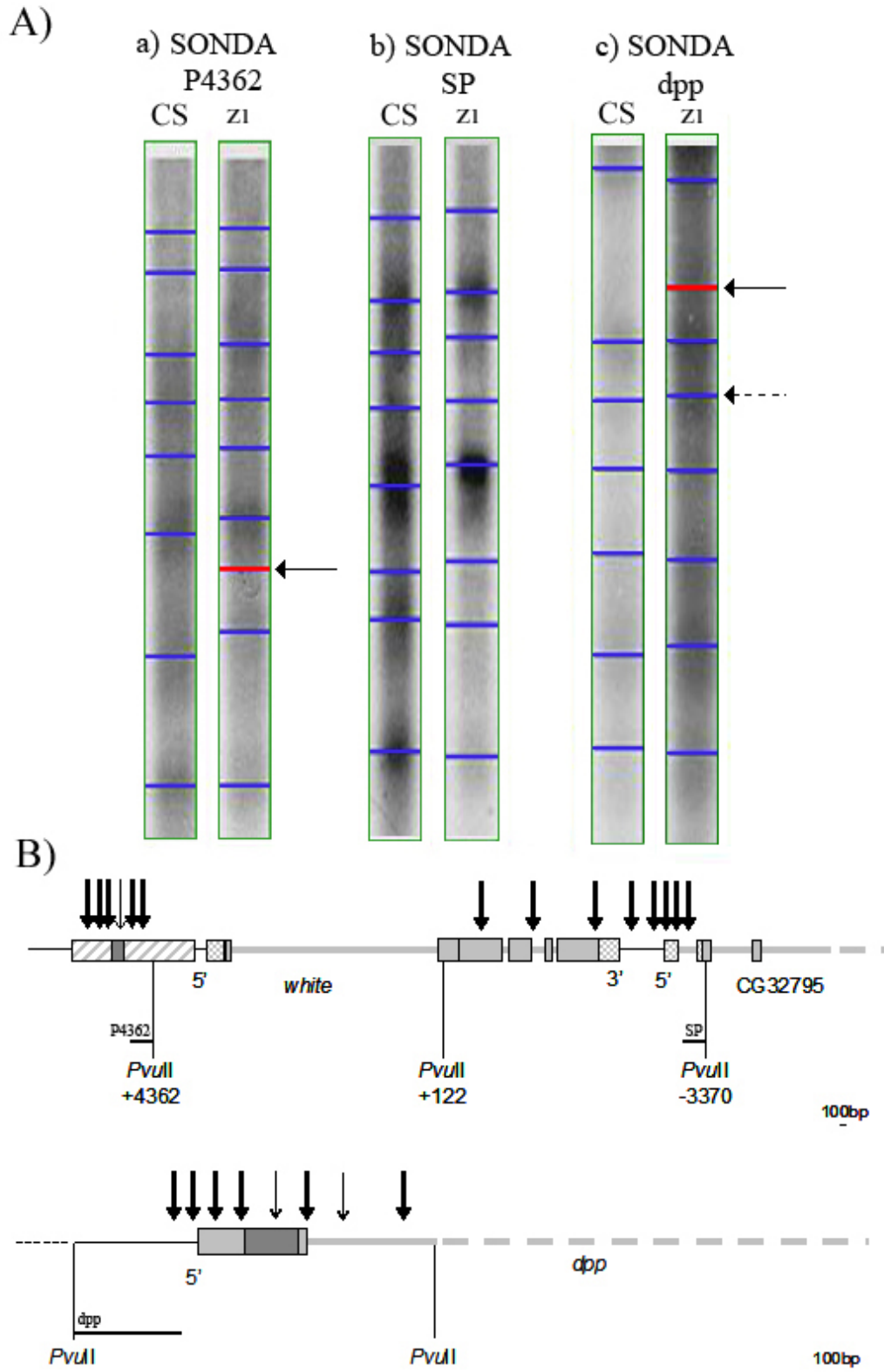


Figura III.16. – Resultados del ensayo de la nucleasa micrococcal y esquema del posicionamiento de nucleosomas. A) Se muestran los carriles del ensayo de la MNasa que fueron analizados con el programa TotalLab. Las líneas azules indican bandas que se encuentran compartidas en ambas cepas (CS y z¹), mientras que las líneas rojas indican bandas presentes exclusivamente en la cepa z¹. Las flechas señalan las líneas rojas. La flecha punteada indica una banda presente en ambas cepas pero mucho más intensa en la cepa z¹. B) Esquema del posicionamiento de nucleosomas. Las flechas más gruesas indican puntos de corte de la MNasa en ambas cepas, mientras que las flechas más delgadas indican puntos de corte de la MNasa exclusivos de la cepa z¹.

La hibridación con la sonda P4362 permitió analizar el posicionamiento de nucleosomas en el promotor del gen *w*, donde se encuentra su ZBS. En esta región, los patrones de posicionamiento eran idénticos para la cepa CS y z^1 , a excepción de una única banda que se encontraba exclusivamente en las hembras z^1 . Dicha banda corresponde a un lugar sensible a MNasa que localiza justo en el ZBS de *w* para la cepa z^1 . La presencia de esta banda, indica que el posicionamiento de este nucleosoma es mucho más estricto en las hembras z^1 que en las hembras salvajes CS, en las que la banda no aparece (ver Figura III.16.).

La hibridación con la sonda SP permitió analizar el extremo 5' del gen *CG32595* y cubrió también el extremo 3' del gen *w*. En esta región se observan las mismas bandas en los carriles de las hembras CS y de las hembras z^1 , indicando que no existe diferencia alguna en el posicionamiento de nucleosomas en ambas cepas (ver Figura III.16.). No es para nada un resultado sorprendente, puesto que esta región no se ve afectada por la interacción *zeste-white*, tal y como demostramos anteriormente (ver Figura III.15.) al comprobar que el gen *CG32795* tiene niveles de expresión parecidos en las dos cepas.

Finalmente, la hibridación con la sonda *dpp* permitió analizar el posicionamiento de nucleosomas en el primer exón del gen *dpp*, donde se encuentra su ZBS. En este caso, el patrón de bandas obtenido también era muy similar para ambas cepas (ver Figura III.16.). Sin embargo, cabe destacar una banda que aparece en las hembras de la cepa z^1 y no en las hembras de la cepa CS, mostrando otra vez un posicionamiento más estricto en z^1 que en CS. En este caso, el nucleosoma posicionado más estrictamente no corresponde al ZBS sino que se localiza en el primer intrón. El nucleosoma posicionado en el ZBS de *dpp* presenta una banda mucho más intensa en las hembras z^1 que en las CS, aunque la banda también es detectable en la cepa salvaje. De todas maneras, la diferencia de intensidad, una vez normalizada respecto a la intensidad global del carril, es tan grande que la consideramos una diferencia en el patrón de posicionamiento (ver Figura III.16. B).

En el núcleo celular, el DNA se encuentra enrollado alrededor de los octámeros de histonas, dificultando así el acceso de la polimerasas y los complejos de reparación o recombinación a dichas secuencias. Para que un gen pueda ser transcrito es necesario

que la cromatina sea remodelada dejando accesibles las secuencias en las que se formará el complejo de inicio de la transcripción. Las tres regiones analizadas en nuestro estudio pertenecen a genes que están siendo transcritos en el individuo adulto. Sin embargo, el estudio de cuantificación de la expresión realizado no permite establecer una relación entre los niveles de expresión de los distintos genes, es decir, no es posible saber cuál de los tres genes es el que más se transcribe en adultos ni cual el que menos. Así, resulta inútil intentar establecer relaciones entre los niveles de expresión y la rigidez del patrón de posicionamiento de nucleosomas. Lo que sí resulta revelador es la relación entre los niveles de expresión y el posicionamiento de nucleosomas para un mismo gen en las dos cepas estudiadas. Así, llama la atención el hecho que el patrón de posicionamiento sea totalmente idéntico sólo en la región en la que no se detectaron diferencias de expresión entre las cepas, es decir, en el primer exón del gen *CG32795*. Sin embargo, los dos genes que presentaban diferencias de expresión entre las hembras CS y las z^1 , también presentan diferencias en el patrón de posicionamiento de los nucleosomas. El gen *w* presenta diferencias de posicionamiento en la región del promotor, concretamente en el ZBS, pero no en el extremo 3'. El gen *dpp* presenta diferencias de posicionamiento en el ZBS de su primer exón y en el primer intrón. Cabe destacar que los cambios observados han sido siempre en la misma dirección: posicionamiento de nucleosomas más estricto en la cepa z^1 .

Resulta difícil decir si los cambios observados en el patrón de posicionamiento de los nucleosomas de los genes *w* y *dpp* son la causa o la consecuencia de los cambios asociados a los niveles de expresión de dichos genes, puesto que ambas opciones son posibles:

1 – El posicionamiento de nucleosomas más estricto en los genes *w* y *dpp* de las hembras z^1 es la CAUSA de la reducción de la expresión dichos genes. La presencia de nucleosomas posicionados estrictamente en el promotor o en el inicio de transcripción de un gen puede reducir la accesibilidad de la maquinaria de transcripción a dichas secuencias. Los agregados de Z en los ZBS son los encargados de reclutar el complejo remodelador de la cromatina BRM, encargado a su vez de abrir la estructura de la cromatina permitiendo la transcripción del gen. La proteína mutante Z^1 no forma agregados sino hiperagregados en los ZBS. Puede que dichos hiperagregados no tengan la misma capacidad que los agregados de Z para reclutar el complejo BRM. En ese

caso, la cantidad de BRM reclutado en los ZBS de la cepa Z^1 sería menor, dificultando la remodelación de la cromatina y por tanto el acceso de la maquinaria de transcripción, reduciendo de esta manera los niveles de expresión de aquellos genes con un ZBS, como es el caso de w y dpp . Sin embargo, aunque Z^1 sea más ineficiente a la hora de reclutar BRM, no significa que su capacidad para hacerlo sea nula, puesto que aunque en menor grado, los genes con ZBS siguen siendo transcritos en las hembras z^1 .

2 – El posicionamiento de nucleosomas más estricto en los genes w y dpp de las hembras z^1 es la CONSECUENCIA de la reducción de la expresión dichos genes. Los niveles de transcripción de los genes con un ZBS son menores en las hembras z^1 que en las hembras salvajes CS. Así, las hembras z^1 presentarán un patrón de posicionamiento de nucleosomas más estricto, pues a menor tasa de transcripción más estricto tiende a ser el posicionamiento (Wu *et al.*, 1979). Si la tasa transcripción se reduce, las secuencias de DNA enrolladas en torno al octámero de histonas no deben ser accesibles a la maquinaria de transcripción tan a menudo, produciéndose un menor desplazamiento de los nucleosomas y distorsionando en menor medida el patrón de nucleosomas posicionados.

Los resultados obtenidos hasta el momento no permiten descartar ninguna de las dos hipótesis. Recordar además que los estudios fueron realizados en hembras adultas enteras. La expresión de los genes w y dpp es específica de tejido y en el caso del gen $CG32795$ se desconoce su patrón de expresión. Puede que el uso de individuos enteros dificulte aún más la interpretación de los resultados obtenidos.

2.3. Patrón de metilación del DNA en hembras CS y z^1

Siguiendo la hipótesis que el posicionamiento más estricto de los nucleosomas en el promotor del gen w y el primer exón del gen dpp en la cepa z^1 sea la causa de la reducción de la expresión de dichos genes debido a un menor reclutamiento de BRM en los ZBS, decidimos estudiar cual era el patrón de metilación del DNA en diferentes regiones de los genes w , $CG32795$ y dpp . Muchos factores relacionados con los complejos remodeladores de la cromatina del tipo SWI/SNF (al que BRM pertenece) han sido asociados con el silenciamiento o activación de la transcripción mediante regulación de los patrones de metilación del DNA (Banine *et al.*, 2005).

El estudio de los patrones de metilación del DNA se realizó mediante el ensayo de modificación del DNA por bisulfito sódico (ver V.6. Ensayo de modificación del DNA por bisulfito sódico) en hembras adultas de las cepas CS y z^1 . Se analizaron cinco regiones distintas:

a) El ZBS del gen w . Es la región del gen w en la que se forman los agregados de Z en la cepa CS o los hiperagregados de Z^1 en la cepa z^1 . Anteriormente habíamos demostrado que la expresión del gen w era reducida y el posicionamiento de nucleosomas en esta región era más estricto para la cepa z^1 . Si se producían cambios en el patrón de metilación, mediados por BRM, como mecanismo para activar/reprimir la transcripción, esperábamos que fuera en esta región.

b) Una región del primer exón de w . Región localizada a 670bp del extremo 3' del ZBS de w . En la cepa z^1 la expresión del gen w se encuentra reducida. Sin embargo, en mamíferos la regulación de la expresión por metilación suele darse en islas CpG a nivel de promotor, por lo que si lo mismo ocurría en *D. melanogaster* (Ferres-Marco *et al.*, 2006), esperábamos que esta región ya no se viera afectada por cambios en la metilación del DNA.

c) El ZBS del gen dpp . Es la región del gen dpp en la que se forman los agregados de Z en la cepa CS o los hiperagregados de Z^1 en la cepa z^1 . Anteriormente habíamos demostrado que la expresión del gen dpp era reducida y el posicionamiento de nucleosomas en esta región era más estricto para la cepa z^1 . Igual que en el caso del ZBS de w , esperábamos que si se producían cambios en el patrón de metilación, mediados por BRM, como mecanismo para activar/reprimir la transcripción, fuera en esta región.

d) Una región del primer intrón de dpp . Región localizada a 938bp del extremo 3' del ZBS de dpp . En la cepa z^1 la expresión del gen dpp se encuentra reducida. Como en el caso de la región localizada en 3' del ZBS de w , esperábamos que esta región ya no se viera afectada por cambios en la metilación del DNA.

e) Una región del tercer exón del gen *CG32795*. El gen *CG32795*, aún teniendo su extremo 5' a menos de 700bp del extremo 3' del gen *w*, no presentaba diferencias de expresión ni en el posicionamiento de nucleosomas entre las hembras de las cepas CS y z^1 . Por ese motivo consideramos que sería un buen control, en el que no esperábamos que las cepas presentaran diferencias en los patrones de metilación del DNA.

La extracción del DNA se realizó a partir de cabezas de hembras adultas de las cepas CS y z^1 , para el estudio de las dos regiones del gen *w* y de hembras adultas enteras para el estudio de las dos regiones del gen *dpp* y la región del gen *CG32795*. El DNA extraído fue digerido con *HindIII* y modificado según el protocolo del tratamiento con bisulfito sódico (ver V.6. Ensayo de modificación del DNA por bisulfito sódico). Se realizó una PCR para cada región a estudiar y los fragmentos obtenidos fueron clonados. Se secuenció un mínimo de 9 colonias y un máximo de 20 para cada región a estudiar. Las secuencias fueron analizadas como un todo, puesto que estábamos buscando patrones de metilación y no metilación de bases concretas. Los resultados obtenidos se resumen en la siguiente tabla:

Región	Cepa	%5mC	No met./Met.	p
ZBS de <i>w</i>	CS	3,8	1373/52	0,01
	z^1	2,1	1581/34	
Primer exón de <i>w</i>	CS	0,8	2229/18	0,02
	z^1	0,3	2240/7	
ZBS de <i>dpp</i>	CS	0,7	1667/13	0,05
	z^1	0,3	1675/5	
Primer intrón de <i>dpp</i>	CS	0,4	2605/12	0,04
	z^1	0,1	1918/3	
Tercer exón de <i>CG32795</i>	CS	5,0	1130/57	0,98
	z^1	5,0	968/49	

Tabla III.9. – Resultados del estudio de las cinco regiones analizadas mediante el ensayo de la modificación del DNA por bisulfito sódico.

Los resultados se analizaron mediante un test de chi cuadrado, comparando el número total de Cs metiladas y de Cs no metiladas para cada región estudiada en cada cepa y después entre cepas. Se consideró que la relación de Cs metiladas y no metiladas entre cepas era diferente de forma estadísticamente significativa en aquellos casos en los que se obtuvo una $p \leq 0,05$. Los resultados muestran que las hembras z^1 se encuentran significativamente hipometiladas en las regiones de los ZBSs de *w* y *dpp*, así como también en las regiones de *w* y *dpp* próximas a los ZBS. Contrariamente a lo que ocurre en estas regiones, nuestra región control en el gen *CG32795*, sin ZBS en su secuencia o en alguna secuencia cercana, aparece igualmente metilada en ambas cepas. Sorprendentemente, las regiones hipometiladas corresponden a las hembras z^1 , cepa en la que la expresión de los genes *w* y *dpp* se ve reducida. Así, pues queda descartado que la regulación de la transcripción de los genes *w* y *dpp* se deba a una hipermetilación de islas CpG, tal y como ha sido descrito en diversos genes de mamíferos y en el gen *Rbf* (*RetinoBlastoma-family*) de *D. melanogaster* (Ferres-Marco *et al.*, 2006).

Llama la atención el hecho que los dos ZBSs analizados y sus regiones próximas se encuentren hipometiladas en la cepa z^1 , mientras que la región control en el gen *CG32795* presenta niveles de metilación iguales a los de la cepa salvaje CS. Estos resultados podrían ser explicados gracias a la formación de hiperagregados de Z^1 en los ZBSs, que bloquearían por impedimento estérico la accesibilidad de la metilasa a la cadena de DNA. Se han descrito algunos casos similares en las que un agregado proteico bloquea el acceso de otras enzimas a la secuencia de DNA. Por ejemplo, el gen *agn43* codifica Ag43 una proteína de la membrana externa en *E. coli*. La regulación de los niveles de transcripción de *agn43* depende de la metilasa de mantenimiento Dam y de la proteína de unión al DNA de respuesta al estrés oxidativo OxyR. Dam es una metilasa de mantenimiento que metila los residuos de adenina de las secuencias GATC. OxyR reprime la expresión de *agn43* uniéndose a la región reguladora de *agn43* situada en 5' del inicio de transcripción de dicho gen. Si las tres secuencias GATC de la región reguladora de *agn43* se encuentran metiladas, la unión de OxyR resulta menos frecuente y la transcripción de *agn43* se activa. La unión de OxyR a la secuencia reguladora de *agn43* “secuestra” sus tres secuencias GATC haciendo que Dam no pueda acceder a ellas (Wallecha *et al.*, 2002). En nuestro caso, serían los hiperagregados de Z^1 los que secuestrarían la zona de los ZBSs y las regiones próximas, creando un impedimento estérico que dificultaría el acercamiento de la Dnmt2 y por tanto la metilación de dichas

secuencias. En las cepas salvajes para z , los agregados de Z en los ZBSs serían de menor tamaño que los de Z^1 y por ese motivo, el impedimento estérico que supondrían para las Dnmt2 sería menor. Esta hipótesis encajaría con el hecho que los niveles de metilación encontrados son bastante más altos (5,0%) en *CG32795* donde no habría ningún impedimento estérico, que en el resto de regiones analizadas (ver Tabla III.9.).

La metilación es un tema muy desconocido en *D. melanogaster* y hasta hace poco se consideraba que esta especie no poseía metilación en el DNA. Sin embargo, estudios recientes han descrito la presencia de secuencias metiladas e incluso de regulación de la expresión génica a través de hipermetilación de islas CpG (Field *et al.*, 2004; Ferres-Marco *et al.*, 2006). Debido a ese gran desconocimiento referente a la metilación en *D. melanogaster* consideramos que sería de interés realizar un análisis de las frecuencias de metilación y de las secuencias metiladas. El porcentaje de Cs metiladas fue de 0,1% a 5,0% según la región, un poco más alto de lo esperado en *D. melanogaster* (0-3%, Field *et al.*, 2004), pero dentro de unos rangos nada sorprendentes.

El análisis de las secuencias metiladas resultó ser un poco más complejo y desconcertante. Se desconoce cuales son las secuencias metiladas en *D. melanogaster*, y mientras algunos trabajos consideran que la metilación se produce en islas CpG como en mamíferos (Ferres-Marco *et al.*, 2006), otros aseguran que la metilación no se produce exclusivamente en dinucleótidos CpG sino también en CpT y CpA (Kunert *et al.*, 2003). Con la intención de aportar más información respecto a cuales son las secuencias metiladas en *D. melanogaster*, alineamos todas las C que aparecían metiladas en nuestras secuencias junto con las 5 bases anteriores y posteriores. El alineamiento de dichas secuencias dio lugar al weblogo (Crooks *et al.*, 2004) (<http://weblogo.berkeley.edu/>) siguiente:

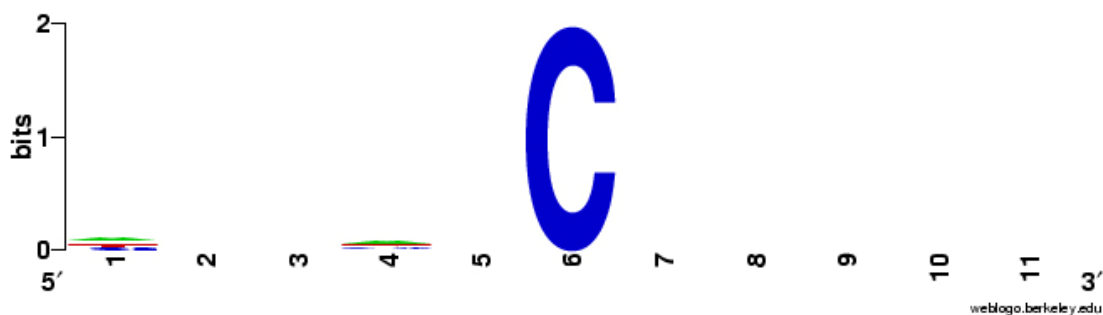


Figura III.17. – Weblogo obtenido mediante el alineamiento de 253 Cs metiladas con las 5 bases anteriores y posteriores a la 5-mC.

El weblogo no resulta informativo y no permite encontrar una secuencia consenso, pues ninguna de las posiciones estudiadas es lo bastante informativa. Por ese motivo, decidimos realizar una nueva aproximación al análisis de las secuencias, estudiando en ese caso las dos bases anteriores y posteriores a la 5mC independientemente, para ver si había una distribución normal de bases en dichas posiciones. Los análisis estadísticos se realizaron con el programa SPSS, utilizando las pruebas no paramétricas de Friedman para comparar más de dos grupos entre sí y la de los rangos de signo de Wilcoxon para comparar dos grupos entre sí. Los resultados obtenidos se resumen en las siguientes tablas:

a) Análisis de la primera base posterior a la 5mC

Región	CpC	CpT	CpA	CpG	CpY	CpR
ZBS de <i>w</i>	5	31	26	27	36	53
1r exon <i>w</i>	6	8	4	7	14	11
ZBS <i>dpp</i>	3	0	2	13	3	15
1r intrón <i>dpp</i>	3	9	4	0	12	4
3r exon <i>CG.</i>	22	29	28	26	51	54
TOTAL	39	77	64	73	116	137
% del total	15,4**	30,4	25,3	28,9	45,85	54,15

Tabla III.10. – Análisis estadístico de la primera base posterior a la 5mC. Y – Pirimidina. R – Purina. ** - $p=0,003$.

La prueba de Friedman para comparar los grupos CpC, CpT, CpA y CpG dio una significación de 0,003, lo que indica que la distribución de la primera base posterior a la 5mC no es normal. Comparando los grupos de dos en dos, se observa que el grupo CpC es el que muestra una frecuencia inferior a lo que cabría esperar si la distribución fuera normal. Comparando los grupos CpT, CpA y CpG entre sí, se obtiene una distribución normal ($p=0,537$). Así, según nuestros resultados, las secuencias metiladas en *D. melanogaster* serán mayoritariamente CpT, CpA y CpG, aunque eso no significa que no se produzca metilación en CpC. Agrupando los datos según CpY (C seguida de pirimidina) y CpR (C seguida de purina), la distribución de la base posterior a la 5mC sigue una distribución normal ($p=0,187$).

b) Análisis de la segunda base posterior a la 5mC

Región	CpNpC	CpNpT	CpNpA	CpNpG	CpNpY	CpNpR
ZBS de <i>w</i>	10	33	33	13	43	46
1r exon <i>w</i>	6	9	4	6	15	10
ZBS <i>dpp</i>	3	3	11	1	6	12
1r intrón <i>dpp</i>	3	1	5	7	4	12
3r exon <i>CG.</i>	13	17	33	42	30	75
TOTAL	35	63	86	69	98	155
% del total	13,8***	24,9	34,0	27,3	38,7***	61,3

Tabla III.11. – Análisis estadístico de la segunda base posterior a la 5mC. Y – Pirimidina. R – Purina. *** - $p < 0,001$.

La capacidad de *A. thaliana* para metilar en secuencias asimétricas (CpNpG, CpHpH; H – no G) (Chan *et al.*, 2005) hizo que nos preguntáramos si la metilación en *D. melanogaster* también podría ser asimétrica. Así, decidimos estudiar las frecuencias de la segunda base posterior a la 5mC. La prueba de Friedman para comparar los grupos CpNpC, CpNpT, CpNpA y CpNpG dio una significación de $p < 0,001$, lo que indica que la distribución de la segunda base posterior a la 5mC no es normal. Comparando los grupos de dos en dos, se observa que el grupo CpNpC es el que muestra una frecuencia inferior a lo que cabría esperar si la distribución fuera normal. Comparando los grupos CpT, CpA y CpG entre sí, se obtiene una distribución normal ($p = 0,141$). Así, según nuestros resultados, las secuencias metiladas en *D. melanogaster* serán mayoritariamente CpNpT, CpNpA y CpNpG, aunque eso no significa que no se produzca metilación en CpNpC. Agrupando los datos según CpNpY y CpNpR, la distribución de la segunda base posterior a la 5mC tampoco sigue una distribución normal, siendo significativamente ($p < 0,001$) inferior la frecuencia de CpNpY. Sin embargo, esta menor frecuencia de pirimidinas en la segunda base posterior a la 5-mC se ve claramente influenciada por la baja frecuencia de C en dicha posición. Cabe destacar pero, que la frecuencia de Ts en dicha posición es la segunda más baja, hecho que favorece la distribución no homogénea cuando se agrupan los datos según pirimidinas y purinas.

c) Análisis de la primera base anterior a la 5mC

Región	CpC	TpC	ApC	GpC	YpC	RpC
ZBS de <i>w</i>	4	27	41	17	31	58
1r exon <i>w</i>	3	9	4	9	12	13
ZBS <i>dpp</i>	6	3	2	7	9	9
1r intrón <i>dpp</i>	3	2	4	7	5	11
3r exon <i>CG.</i>	22	23	27	33	45	60
TOTAL	38	64	78	73	102	151
% del total	15,0**	25,3	30,8	28,9	40,3**	59,7

Tabla III.12. – Análisis estadístico de la primera base anterior a la 5mC. Y – Pirimidina. R – Purina. ** - $p=0,002$.

La prueba de Friedman para comparar los grupos CpC, TpC, ApC y GpC dio una significación de 0,002, lo que indica que la distribución de la primera base anterior a la 5mC no es normal. Comparando los grupos de dos en dos, se observa que el grupo CpC es el que muestra una frecuencia inferior a lo que cabría esperar si la distribución fuera normal. Comparando los grupos TpC, ApC y GpC entre sí, se obtiene una distribución normal ($p=0,495$). Así, según nuestros resultados, las secuencias metiladas en *D. melanogaster* serán mayoritariamente DpC (D – no C), aunque eso no significa que no se produzca metilación en CpC. Agrupando los datos según YpC y RpC, la distribución de la primera base anterior a la 5mC tampoco sigue una distribución normal, siendo significativamente inferior ($p=0,002$) la frecuencia de YpC. Como en el caso de la segunda base posterior a la 5-mC, no podemos asegurar que esta menor frecuencia de pirimidinas no sea debida a la baja frecuencia de C en dicha posición. Destacar también, como en el caso anterior, que la frecuencia de Ts en dicha posición es la segunda más baja, hecho que favorece la distribución no homogénea cuando se agrupan los datos según pirimidinas y purinas.

d) Análisis de la segunda base anterior a la 5mC

Región	CpNpC	TpNpC	ApNpC	GpNpC	YpNpC	RpNpC
ZBS de <i>w</i>	17	34	32	6	51	38
1r exon <i>w</i>	10	8	5	2	18	7
ZBS <i>dpp</i>	3	3	8	4	6	12
1r intrón <i>dpp</i>	4	1	7	4	5	11
3r exon <i>CG.</i>	18	24	50	13	42	63
TOTAL	52	70	102	29	122	131
% del total	20,6	27,7	40,3***	11,5***	48,2	51,8

Tabla III.13. - Análisis estadístico de la segunda base anterior a la 5mC. Y – Pirimidina. R – Purina. *** - $p < 0,001$.

La prueba de Friedman para comparar los grupos CpNpC, TpNpC, ApNpC y GpNpC dio una significación de $p < 0,001$, lo que indica que la distribución de la segunda base anterior a la 5mC no es normal. Comparando los grupos de dos en dos, se observa que el grupo GpNpC es el que muestra una frecuencia inferior a lo que cabría esperar si la distribución fuera normal. Comparando los grupos TpNpC, ApNpC y GpNpC entre sí, tampoco se obtiene una distribución normal ($p < 0,001$), evidenciando al comparar los grupos entre sí de dos en dos que el grupo ApNpC presenta una frecuencia más elevada de lo esperado según una distribución normal. Así, según nuestros resultados, las secuencias metiladas en *D. melanogaster* serán mayoritariamente ApNpC, aunque eso no significa que no se produzca metilación en BpNpC (B – no A). Agrupando los datos según YpNpC y RpNpC, la distribución de la segunda base anterior a la 5mC sigue una distribución normal ($p = 0,103$), puesto que la baja frecuencia de GpNpC se ve compensada por la alta frecuencia de ApNpC, dando lugar a una frecuencia de RpNpC parecida a la de YpNpC.

Así, el análisis estadístico de las secuencias con 5mCs en las regiones que estudiadas desvela una secuencia consenso que sería la más frecuentemente metilada en *D. melanogaster*: ApDpCpDpD (D – no C). Eso no significa que secuencias que no cumplan dichas condiciones no vayan a ser metiladas, pero lo serán en menor frecuencia. Así podríamos decir que la metilación en *D. melanogaster* tiende a

producirse en Cs localizadas en un entorno en el que las Cs sean poco frecuentes. Aún así, no se puede descartar que nuestra observación se deba a una distribución no homogénea de las bases a lo largo del genoma. Si las Cs aisladas fueran más abundantes que las Cs rodeadas de Cs, entonces la secuencia consenso propuesta sólo reflejaría la distribución de las bases en el genoma. Por ese motivo, para establecer una secuencia consenso más precisa en lo que a la metilción en *D. melanogaster* se refiere, sería necesario realizar más experimentos así como revisar lo poco publicado por otros grupos hasta el momento.

2.4. Discusión general de los resultados

El estudio molecular de las hembras de la cepa mutante z^1 ha permitido aportar nueva información respecto al papel que esta metilación juega en la regulación de la expresión de determinados genes. Sin embargo, quedan algunas preguntas abiertas que deberían ser estudiadas con profundidad para poder descartar o confirmar las hipótesis propuestas.

La primera parte del estudio nos permitió evaluar el efecto de la mutación z^1 sobre la expresión de dos genes con un ZBS: *w* y *dpp*. Se estudió además el gen *CG32795* que se encuentra a sólo 700bp del extremo 3' del gen *w*, formando una pareja de genes *head-to-tail* con éste. Los resultados obtenidos mediante el análisis de la expresión por PCR a tiempo real se resumen en la siguiente tabla:

Cepas	% expresión de mRNA			
	Región <i>w</i>	Región CG5	Región CG3	Región <i>dpp</i>
♀ CS	100	100	100	100
♀ z^1	17,74	92,84	95,41	14,09

Tabla III.14 – Resumen de los resultados del análisis de la expresión de los genes *w*, *CG32795* y *dpp*.

Los resultados indican que en las hembras z^1 se produce una clara reducción de la expresión de aquellos genes con un ZBS (*w* ~ 18% y *dpp* ~ 14% de la expresión, considerando la expresión en la cepa salvaje CS como el 100%), tanto si éste se encuentra en el promotor (el caso de *w*) como si se encuentra en el primer exón (el caso

de *dpp*). Sin embargo, la expresión del gen *CG32795* (~ 93% de la expresión en el extremo 5' y ~95% en el extremo 3') se mantiene a unos niveles parecidos a los de las hembras salvajes CS. De los resultados obtenidos se infiere que la formación de hiperagregados de Z^1 en los ZBSs puede dificultar la transcripción de los genes con ZBSs. Se desconoce si la reducción de la expresión es debida a la dificultad que los hiperagregados de Z^1 puedan representar para el avance de la maquinaria de transcripción o bien a la posible menor habilidad de Z^1 para reclutar el complejo remodelador de la cromatina BRM, encargado de abrir la estructura de la cromatina para facilitar la transcripción. Sea cual sea el mecanismo, los resultados también indican claramente que el efecto de la mutación z^1 es local y sólo se produce en aquellos genes con un ZBS, puesto que la expresión del gen *CG32795* aunque forma un par de genes con *w*, no ve reducida su expresión.

La posibilidad de que la segunda hipótesis fuera la correcta, es decir, que la reducción de la expresión de los genes con un ZBS en las hembras z^1 se produzca debido a una menor habilidad de la proteína mutante Z^1 para reclutar BRM, nos indujo a pensar que en los mutantes z^1 la expresión de los genes con un ZBS se podía ver alterada por problemas en la remodelación de la cromatina. Por ese motivo, decidimos realizar un análisis del posicionamiento de nucleosomas mediante el ensayo de la nucleasa micrococcal, rehibridando un mismo filtro con tres sondas diferentes, cubriendo el ZBS de los genes *w* y *dpp* y el extremo 5' del gen *CG32795*.

Los resultados indican que en las hembras z^1 el posicionamiento de los nucleosomas es más estricto en el ZBS de los genes *w* y *dpp* (ver Figura III.16.). Las hembras z^1 presentan un nucleosoma posicionado en el ZBS del gen *w*, que no se encuentra presente en el ZBS de las hembras CS. Lo mismo ocurre en el primer exón de *dpp*. En este caso, el nucleosoma no posicionado en las hembras z^1 no se encuentra justo en la secuencia del ZBS, sino unas bases más abajo en 3', en una secuencia intrónica. El nucleosoma que se encuentra claramente posicionado en las hembras z^1 sobre el ZBS de *dpp* aparece como una banda mucho menos intensa en las hembras CS, indicando que aún encontrarse posicionado en las hembras salvajes, su posicionamiento es mucho menos estricto. El posicionamiento de nucleosomas en el gen *CG32795*, tal y como era de esperar a la luz de los resultados de expresión, es prácticamente idéntico en las hembras CS y z^1 .

Las diferencias en el posicionamiento de los nucleosomas en los genes con ZBS, coincidiendo con la reducción de la expresión de dichos genes, reafirman la hipótesis de una posible menor capacidad de los hiperagregados de la proteína mutante Z^1 para reclutar el complejo remodelador de la cromatina BRM. Cabe destacar pero, al contrario que esperaríamos según nuestra hipótesis, que la mutación z^1 es puntual y produce un cambio aminoacídico en la región rica en prolinas (ver Figura I.11.), no afectando para nada la cremallera de leucina que es la encargada de interaccionar con BRM. Aún así, puede que la creación de un nuevo dominio hidrofóbico en la proteína mutante Z^1 , altere la posición de la cremallera respecto al resto de la proteína, dificultando la accesibilidad de BRM a ésta.

Sin embargo, cabe recordar que el análisis del posicionamiento de nucleosomas se realizó en hembras adultas enteras, pues aunque la expresión del gen w es específica de tejido y se podrían haber usado sólo cabezas para este fin, la expresión de dpp se produce en otras partes del cuerpo y la de $CG32795$ se desconoce su patrón. No sólo el hecho de trabajar con hembras enteras dificulta la interpretación de los resultados. La expresión a diferente nivel entre las distintas cepas de los genes w y dpp dificulta el análisis. A menor nivel de expresión del gen, menor es el reposicionamiento de nucleosomas necesario para la transcripción. Ese podría ser un factor de confusión importante a la hora de aceptar que el posicionamiento de nucleosomas en las hembras z^1 es más estricto que en las hembras CS. Por ese motivo, no podemos asegurar si el posicionamiento de nucleosomas más estricto en las hembras z^1 es causa o consecuencia de la menor expresión de los genes w y dpp .

Aún así, considerando que el posicionamiento más estricto de los nucleosomas en el promotor del gen w y el primer exón del gen dpp en la cepa z^1 es la causa de la reducción de la expresión de dichos genes debido a un menor reclutamiento de BRM en los ZBS, decidimos estudiar cual es el patrón de metilación del DNA en diferentes regiones de los genes w , $CG32795$ y dpp , puesto que diversos factores relacionados con los complejos remodeladores de la cromatina SWI/SNF (al que BRM pertenece) han sido asociados con la regulación de los patrones metilación del DNA (Banine *et al.*, 2005).

El análisis de los patrones de metilación se realizó mediante modificación del DNA por bisulfito sódico en el ZBS de los genes *w* y *dpp*, en dos regiones cercanas localizadas en 3' de éstos y en una región control del gen *CG32795*, que no había mostrado diferencias ni en los niveles de expresión, ni en el posicionamiento de los nucleosomas. Los resultados obtenidos (ver Tabla III.9.) indican que el ZBSs de los genes *w* y *dpp*, así como las regiones próximas a éstos, se encuentran hipometilados en las hembras z^1 de forma estadísticamente significativa ($p \geq 0,05$). Sin embargo, la región control en el gen *CG32795* presenta unos niveles de metilación similares en ambas cepas. Las regiones hipometiladas corresponden a las hembras z^1 , en las que la expresión de los genes *w* y *dpp* se ve reducida. Así, se puede descartar la posibilidad de que la regulación de la transcripción de los genes *w* y *dpp* se deba a una hipermetilación de islas CpG. Sin embargo, nuestros resultados podrían ser explicados por el “secuestro” de las secuencias de los ZBSs y de las regiones circundantes por arte de los hiperagregados de Z^1 . Éstos serían más grandes que los agregados de Z e impedirían el acceso de las metilasas Dnmt2 a las secuencias a metilar, de forma similar a lo que hace OxyR en la región reguladora del gen *agn43* en *E. coli*. Esta hipótesis también encajaría con el hecho que la región del gen *CG32795* es la que presenta niveles más altos de metilación (5,0%) puesto que en ésta no habría impedimento estérico de ningún tipo.

El estudio del patrón de metilación, también nos permitió realizar un análisis de cuáles son las secuencias metiladas en *D. melanogaster*. Así, aunque el weblogo no reveló ninguna secuencia consenso, analizamos las dos posiciones anteriores y posteriores a cada 5mC. Mediante análisis estadísticos llegamos a una secuencia consenso para C metiladas ApDpCpDpD (D – no C), aunque eso no significa que la metilación no se produzca en otro tipo de secuencias. El hecho de encontrar Cs metiladas en secuencias tan variables, y no exclusivamente en dinucleótidos CpG (como en el caso de mamíferos) puede hacer pensar que la modificación del DNA por bisulfito sódico fue incompleta. Sin embargo cabe resaltar que la metilación no siempre se produce en dinucleótidos CpG, si no que por ejemplo *A. thaliana* tiene metilación simétrica en CpG pero también metilación asimétrica en CpNpG y CpHpH (H – no G) (Chan *et al.*, 2005). Incluso en humanos, que supuestamente metilan exclusivamente en dinucleótidos CpG, se han encontrado Cs metiladas en dinucleótidos CpA y CpT en el exón 5 del gen *p53* de muestras de cáncer de pulmón (Kouidou *et al.*, 2005). Además el hecho de que las frecuencias de 5mC encontradas en nuestras muestras (0,1-5%) se

aproximen a las esperadas en *D. melanogaster* (0-3%) y que se haya descrito la presencia de 5mC en dinucleótidos CpT y CpA para dicha especie (Kunert *et al.*, 2003), confirma nuestros resultados. Otro hecho que reafirma la validez de nuestros resultados es el haber encontrado 5mC en la misma posición en clones diferentes de la misma cepa y de la otra cepa. Así, algunas posiciones CpA, CpC y CpT aparecen como metiladas en extracciones de DNA y modificaciones por bisulfito sódico distintas. Las reacciones de modificación fueron repetidas y algunas regiones vueltas a analizar y los resultados obtenidos fueron muy similares, encontrándose otra vez las mismas posiciones metiladas. Sin embargo, podría ser interesante comprobar la total modificación de nuestras muestras mediante ensayos de enzimas sensibles a metilación. Mucho queda por saber referente a la metilación en *D. melanogaster* y más estudios amplios, detallados y cuidadosos son necesarios, además de una revisión y compilación de los estudios publicados hasta el momento para la metilación en dicha especie.

IV. CONCLUSIONES

A partir de los objetivos propuestos y mediante el trabajo experimental, el presente trabajo de Tesis nos permite llegar a las siguientes conclusiones:

1 - El gen *CG32795* no es vital, puesto que la cepa *CGNulw3*, *knock-out* para dicho gen, es viable y fértil.

2 - Las predicciones bioinformáticas indican que la proteína *CG32795* posee seis hélices transmembranales y una alta similitud de secuencia con la familia proteica *TMPIT*.

3 - En machos adultos, el gen *CG32795* presenta cuatro formas de *splicing* alternativo en su extremo 5' y dos señales de poliadenilación distintos en 3'. Las diferentes variantes de *splicing* predichas anteriormente en *FlyBase* se parecen a las encontradas, pero no son idénticas.

4 - La inserción de un elemento transponible *FB-NOF*, entero o truncado, en el tercer intrón del gen *CG32795* no afecta a la expresión del gen *w* en las cepas *M115* y *RM115*.

5 - Las cepas z^1 y *RM115*, portadoras del alelo z^1 , presentan expresión reducida de *w* a nivel de mRNA. La cantidad de pigmentos en los ojos de los individuos de estas cepas también se ve reducida.

6 - En la cepa *M115*, el estudio de la expresión de *w* en machos enteros resulta engañoso. Al tratarse de una cepa z^1 con una duplicación que incluye el ZBS de *w*, la expresión de dicho gen se ve reducida en la cabeza, pero no así en el resto del cuerpo, en el que la expresión de *w* es aproximadamente el doble.

7 - La inserción de un elemento transponible *FB-NOF* en el tercer intrón del gen *CG32795*, que da lugar a la cepa *M115* y la posterior reordenación cromosómica, que da lugar a la cepa *RM115* reducen la expresión del gen *CG32795* en ambas cepas.

8 - En las hembras de la cepa z^1 la expresión de los genes w y dpp , ambos poseedores de un ZBS, se ve reducida respecto a la expresión en las hembras salvajes CS.

9 - El efecto de z^1 en la expresión génica es local. Por este motivo, la expresión del gen $CG32795$, vecino próximo del gen w con el que forma un par de genes *head-to-tail*, no se ve alterada en las hembras de la cepa z^1 .

10 - El posicionamiento de nucleosomas es más estricto en el ZBS de w y dpp en las hembras de la cepa z^1 que en las hembras salvajes CS.

11 - Las hembras de la cepa z^1 se encuentran hipometiladas en el ZBS de w y dpp y en las regiones circundantes a éstos.

12 - El análisis de las secuencias con 5mC en los genes w , dpp y $CG32795$ en hembras z^1 y CS, dio como secuencia consenso para la metilación en *D. melanogaster*: ApDpCpDpD, siendo las frecuencias para cada posición: ~40%, ~85%, 5mC, ~85%, ~86%.

13 - La frecuencia de 5mC en las regiones estudiadas de los genes w , dpp y $CG32795$ oscila entre el 0,1% y el 5,0%.

V. MATERIALES Y MÉTODOS

1. Cepas de *Drosophila melanogaster*

Las cepas de *D. melanogaster* usadas para la realización de esta tesis son las siguientes:

- **CS** (Canton S) → Cepa salvaje de referencia.
- **z^1** (*zeste*¹) → Mutante z^1 de referencia.
- **M115** → (*sc*, z^1 , *f*) Cepa mutante aislada por Baldrich (1998) y analizada por Badal (Badal *et al.*, 2006a). Los machos de esta cepa son portadores de una duplicación del gen *w* y de la región 5' del gen *CG32795*. Se caracterizan por un fenotipo de ojos color *zeste*.
- **RM115** → (*sc*, z^1 , *f*) Descrita por Badal (Badal *et al.*, 2006a). Cepa que proviene de la reversión de la cepa M115, dando un genotipo con una única copia del gen *w* y de la región 5' del gen *CG32795*. Los machos presentan un fenotipo de ojos salvajes.
- **Transw3** → (z^1) Cepa transgénica obtenida por microinyección por nuestro grupo (Portela, 2003) gracias a la colaboración de la Dr. Preiss (Hohenheim Universität, Stuttgart, Alemania). Proviene de una cepa w^{1118} a la que se insertó un alelo *w* completo (incluyendo sus *enhancers*) en el cromosoma 3R.
- **wIR8** → Cepa amablemente proporcionada por el Dr. A. Pélisson (Institut de Génétique Humaine, Montpellier, Francia). Cepa con fenotipo *white* debido a una deleción que abarca el extremo 3' del gen *w* y el extremo 5' del gen *CG32795*.
- **CGNulw3** → (z^1) Cepa con una deleción en el cromosoma X que comprende el gen *w* y el gen *CG32795*. Una copia entera del gen *w* (incluidos sus *enhancers*) se encuentra en el cromosoma 3, reestableciendo así la función de *w*, pero siendo aún *knock out* para *CG32795*. Cepa obtenida por el Dr. Oriol Cabré mediante varios cruces entre las cepas Transw3 y wIR8.

Las cepas han sido mantenidas en un medio de cultivo estándar para *D. melanogaster* a una temperatura de 25°C, una humedad relativa aproximada del 60% y un ciclo luz/oscuridad de 12 horas.

2. Extracción de DNA genómico

El DNA genómico se ha obtenido siguiendo principalmente el protocolo descrito por Piñol *et al.*, 1988.

Se homogenizan 0,20 – 0,25g de moscas en 10mL de tampón A (100mM NaCl, 10mM EDTA, 0,5% Triton X-100, 20mM TrisCl pH 7,5). Se filtra el homogenizado a través de una malla de nailon de 150 *mesh*, para separar los núcleos de los restos de tejido y de las cubiertas celulares. El homogenizado filtrado se centrifuga a 4000rpm durante 5 minutos y se lava dos veces en tampón B (100mM NaCl, 10mM EDTA, 20mM TrisCl pH 7,5). Se añaden, al sedimento, 0,75mL de perclorato sódico 4M y 4% SDS con el fin de disociar las proteínas unidas al DNA. A continuación se desproteíniza con un lavado con fenol y dos lavados adicionales de cloroformo-álcohol isoamílico (24:1 v/v). El DNA se precipita con etanol absoluto y se lava con etanol 70%. Finalmente, se deja secar el *pellet* al aire y se disuelve en un volumen de agua o tampón TE apropiado según el rendimiento de la extracción (es decir, según el tamaño de los *pellets*).

3. PCR a tiempo real

Desde la descripción de la PCR por Mullis en los años 80, ésta ha sido una de las técnicas que más ha contribuido a la actual revolución de la biología molecular. Debido al impacto causado por la PCR, Mullis recibió el Premio Nobel en el año 1993. La PCR a tiempo real nos lleva un paso más allá, permitiendo monitorizar la amplificación de los ácidos nucleicos a lo largo de todo el proceso (Higuchi *et al.*, 1993).

3.1. *LightCycler*

El sistema *LightCycler* de Roche Molecular Biochemicals (Roche, Alemania) es uno de los sistemas disponibles en el mercado para realizar PCR a tiempo real, y ha sido

el utilizado en nuestro caso. Su diseño permite realizar la amplificación y el análisis del resultado en un mismo capilar, acelerando el proceso analítico y reduciendo el riesgo de contaminación.

En este sistema de PCR a tiempo real, los ciclos típicos de cualquier reacción de PCR (desnaturalización, hibridación de los cebadores y extensión) son más cortos, gracias al uso de capilares y de un sistema de transferencia de calor por aire altamente eficaz. El aire se calienta mediante una resistencia que permite adquirir la temperatura programada con un error de $\pm 0,3^{\circ}\text{C}$. El aire casi no tiene masa y por eso, el proceso es mucho más rápido que usando termocicladores con bloque térmico. El modo de transmitir el calor, junto la elevada relación superficie-volumen de los capilares, hacen que la transición de temperaturas puede llegar a los $20^{\circ}\text{C}/\text{segundo}$ (ver Figura V.1.).

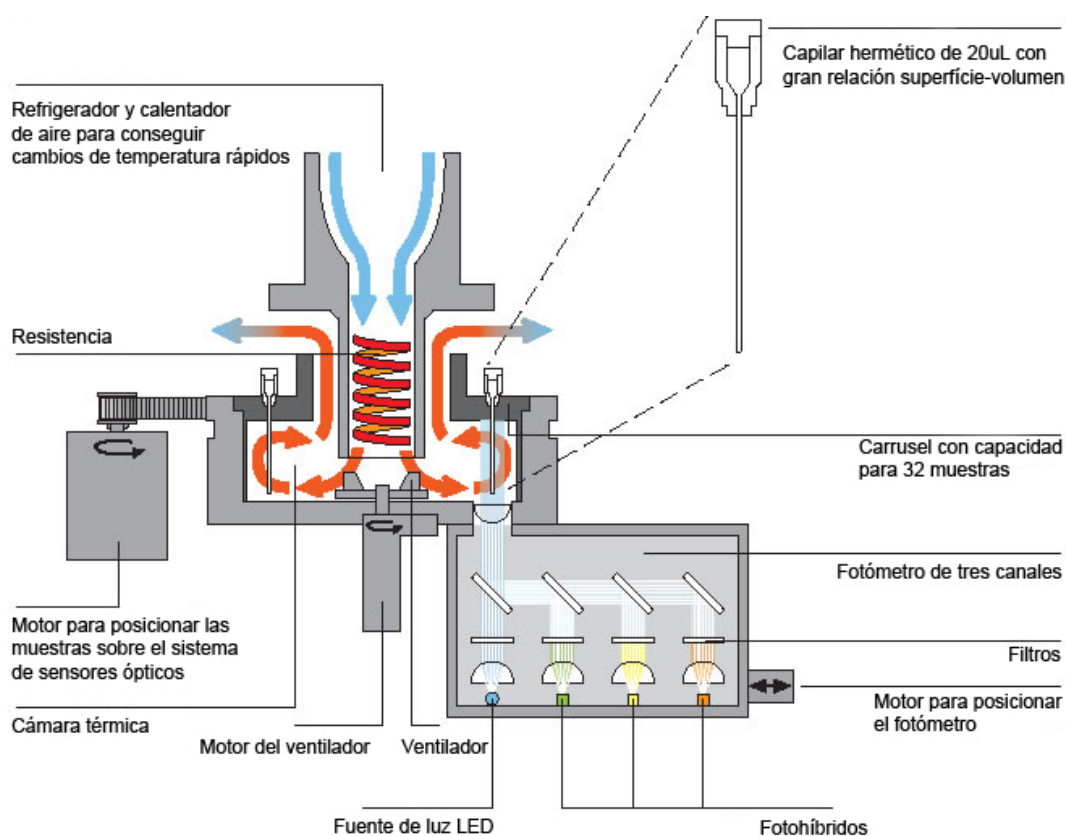


Figura V.1. – Esquema del funcionamiento del *LightCycler*.

Los capilares (Roche, Alemania) en los que se lleva a cabo la reacción de PCR tienen una capacidad de 20 μ L, son de borosilicato y están diseñados especialmente para el *LightCycler* (Roche, Alemania) (ver Figura V.2.). Las propiedades del cristal borosilicado son las óptimas para permitir la medición de la fluorescencia.

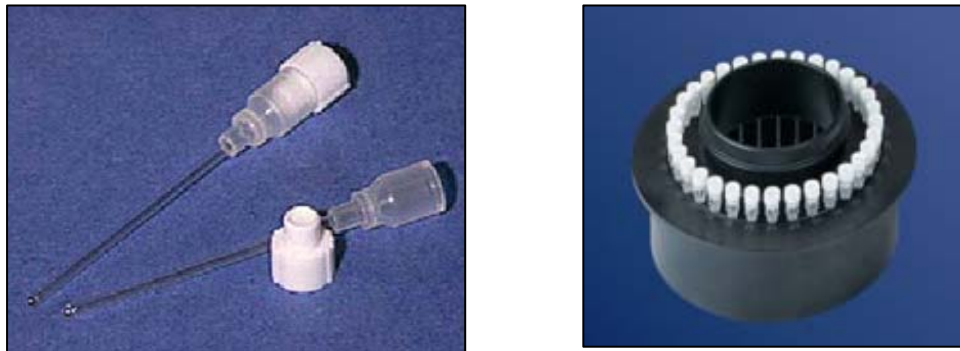


Figura V.2. – Capilares y rotor de 32 muestras del *LightCycler*.

El rotor donde se colocan los capilares tiene capacidad para 32 muestras (ver Figura V.2.). Durante el proceso de la PCR, el rotor rueda de manera que cada capilar se ajusta de forma perfecta al dispositivo óptico. La optimización de esta posición, con precisión de micrómetros, se lleva a cabo cada vez antes de empezar el programa, en un proceso llamado *seek*. Una vuelta completa de rotor dura 5 segundos, de manera que el tiempo de medición por capilar es de unos 20 milisegundos.

La fluorescencia se mide a intervalos pre-programados. La unidad óptica del *LightCycler* tiene tres canales de detección que miden la luz emitida a diferentes longitudes de onda (canal 1 \rightarrow 530nm, canal 2 \rightarrow 640nm y canal 3 \rightarrow 710nm). Además hay una fuente de luz azul llamada LED (*light-emitting diode*), que emite a 470nm. La luz de la LED es la encargada de excitar los fluoróforos. Una vez excitado el fluoróforo, la luz emitida es conducida a la unidad óptica, donde una serie de filtros y espejos separan la luz en las diferentes longitudes de onda que pueden ser detectadas por los tres canales. Los datos recogidos aparecen inmediatamente en pantalla, de manera que se puede hacer un seguimiento de la reacción a tiempo real (ver Figura V.1.).

En nuestros experimentos de PCR a tiempo real, el fluoróforo utilizado ha sido el *SYBR Green I*, una molécula de unión específica al surco pequeño del DNA de doble cadena (dsDNA). La elección del *SYBR Green I* como fluoróforo para la PCR a tiempo real, se basa principalmente en dos de sus cualidades: 1 - Su fluorescencia se ve incrementada cuando se encuentra unido a dsDNA. 2 - No inhibe la reacción de PCR. El *SYBR Green I* se une a todo el dsDNA que se generó durante la reacción de PCR, de modo que los niveles de fluorescencia captada se incrementan en cada ciclo.

La medida de la fluorescencia se realizó a través del canal 1, al final de cada fase de elongación, pues en este momento del ciclo se produce la máxima emisión fluorescente por parte del *SYBR Green I* unido al dsDNA.

La unión del *SYBR Green I* a cualquier dsDNA le hace polivalente, pudiéndolo utilizar para monitorizar cualquier producto amplificado. En su contra cabe decir que la versatilidad del *SYBR Green I* puede conducir a errores, si éste se une a productos de amplificaciones inespecíficas. Para evitar estos errores, hay que analizar la curva de fusión.

La curva de fusión se genera al terminar los ciclos de amplificación de la PCR. Es un paso del programa que consiste en un aumento progresivo y sostenido de la temperatura desde los 65°C hasta los 95°C, a una velocidad de 0,1°C/segundo. Durante este proceso se mide la fluorescencia constantemente. Al inicio de la curva todo el DNA amplificado se encuentra en forma de dsDNA y la fluorescencia emitida por el *SYBR Green I* es máxima. A medida que incrementa la temperatura, la fluorescencia va disminuyendo ya que también disminuye la estabilidad de las estructuras de dsDNA. Cuando se llega a la T_m del producto de PCR, la mayor parte del producto se encontrará en forma de ssDNA y se producirá una caída drástica de la fluorescencia. Los datos de las medidas de fluorescencia tomados a lo largo de este proceso se pueden representar gráficamente, dando lugar a la llamada curva de fusión (ver Figura V.3.).

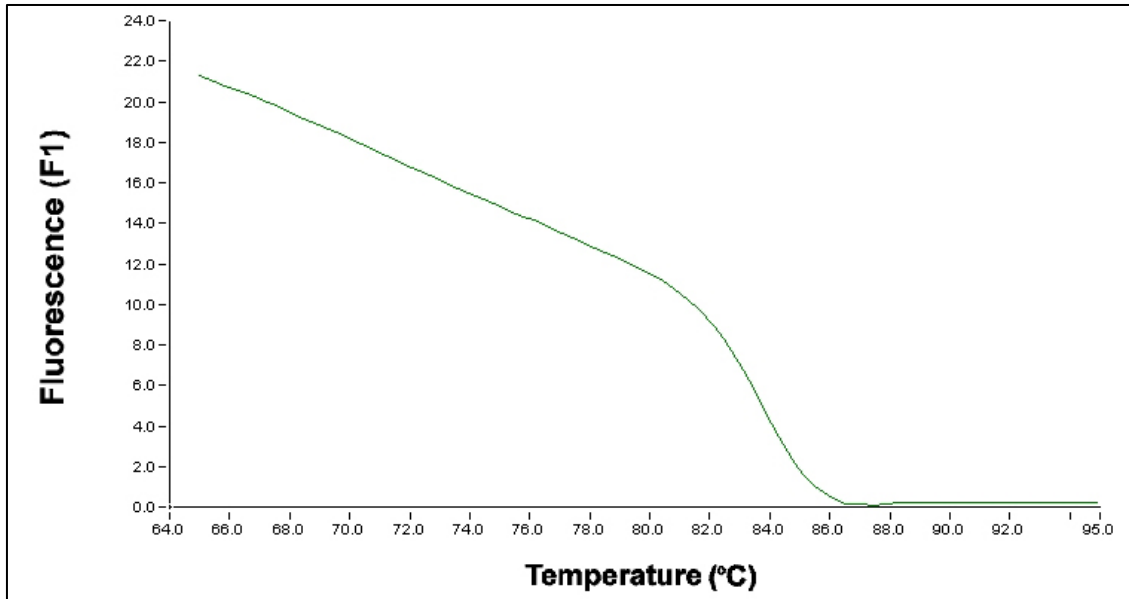


Figura V.3. – Ejemplo de curva de fusión de uno de nuestros productos.

Estudiando la curva de fusión se puede saber si el producto amplificado es específico o si ha habido amplificaciones inespecíficas. Si se realiza un gráfico de $-dF/dT$ versus temperatura, se obtiene una curva más sencilla de interpretar; puesto que los cambios de pendiente alrededor de los puntos de la T_m se muestran como a picos en los que el vértice del pico corresponde a la T_m . Si hay más de un producto de amplificación aparece más de un pico (ver Figura V.4.).

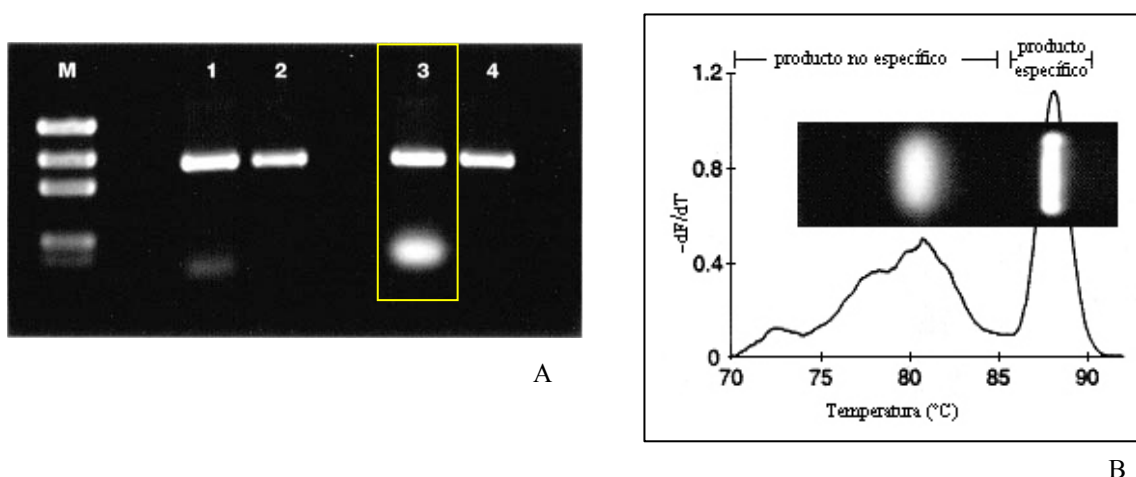


Figura V.4. – **Posible contaminación o inespecificidad en las amplificaciones.** A) Gel de agarosa correspondiente al análisis de los productos de una PCR. En el carril 3 del gel se observa una amplificación inespecífica que puede ser detectada en la curva de fusión (B) gracias a la presencia de un pico inesperado. En una amplificación específica, como la del carril 2 sólo se observa el segundo pico.

3.1.1. PCR a tiempo real cuantitativa

Una de las principales aplicaciones de la PCR a tiempo real es la cuantificación de la expresión de un gen, a través de la estimación del mRNA transcrito. Esta cuantificación puede ser absoluta (obteniendo el número aproximado de copias de un determinado mRNA que se encuentran por célula) o relativa (obteniendo la relación entre el número de copias de un determinado mRNA en dos muestras distintas). En nuestro caso hemos realizado una cuantificación relativa que nos ha permitido comparar la expresión de diversos genes en diferentes cepas de *D. melanogaster*.

En la cuantificación relativa, la expresión del gen de interés se normaliza respecto a la expresión de un gen de referencia (cuya expresión se considera constante entre las distintas cepas analizadas). El resultado se expresa como cociente entre los niveles de expresión de ambos genes y no como valor absoluto (Roche Applied Science, 2002). Esta normalización permite corregir errores causados por: variaciones en la cantidad inicial de mRNA total, posible degradación inicial de una muestra de RNA, diferencias en la calidad de las muestras, errores de pipeteo, variaciones en la eficiencia de síntesis del cDNA...

Como genes de referencia suelen usarse genes de expresión constitutiva. Estos genes codifican proteínas esenciales para el mantenimiento de las funciones celulares, como por ejemplo: componentes del citoesqueleto (β -actina, α -tubulina), componentes del complejo mayor de histocompatibilidad (β 2-microglobulina), enzimas de la ruta glicolítica (GAPDH), subunidades ribosomales... Aún así, la expresión de los genes constitutivos puede verse afectada por el diseño experimental (Thellin *et al.*, 1999). Por eso es necesario validar la expresión del gen de referencia escogido para cada experimento.

Como ya se ha dicho anteriormente, la PCR a tiempo real permite monitorizar la amplificación de los ácidos nucleicos a lo largo de todo el proceso. Así, realizando un gráfico del incremento de la fluorescencia *versus* el ciclo de la PCR (ver Figura V.6.), se obtiene una curva de crecimiento con tres zonas diferenciadas:

- Zona inicial de *lag* → Fluorescencia por debajo del umbral de sensibilidad del fluorómetro. Dura hasta que la fluorescencia del producto de PCR es superior a la fluorescencia de fondo.
- Zona de crecimiento exponencial → Empieza cuando se ha acumulado suficiente producto de PCR, de modo que su señal fluorescente es superior a la del fondo. Se termina cuando cae la eficiencia de la reacción.
- Zona final de *plateau*.

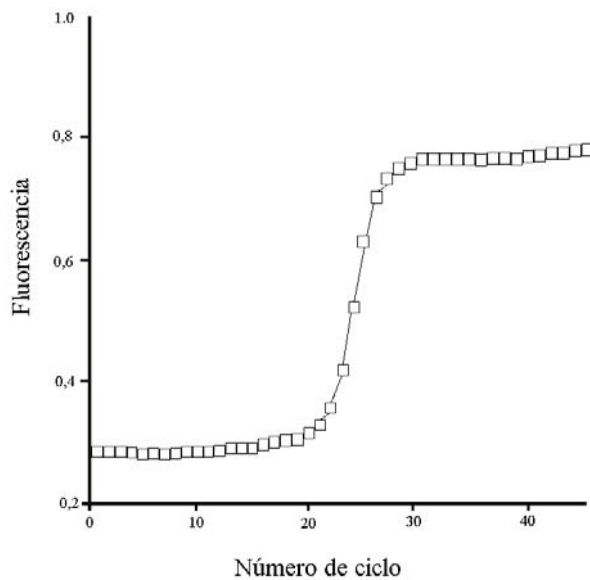


Figura V.5. – **Representación de las medidas de fluorescencia obtenidas a lo largo de los ciclos de amplificación en una PCR a tiempo real.**

La PCR a tiempo real permite cuantificar el número inicial de copias molde a partir de la cinética de cada reacción. Durante la fase de crecimiento exponencial, la cinética de la amplificación se describe según la ecuación:

$$N_n = N_0 \times (1+E)^n \quad (1)$$

En la que:

N_n → La cantidad de producto acumulada en el ciclo n.

N_0 → La cantidad de molde inicial.

E → Eficiencia de la amplificación. La eficiencia máxima es 1 (en el caso que el producto de PCR se duplique a cada ciclo) y la mínima 0 (no hay amplificación).

n → Número de ciclo.

El ciclo en que cada cinética se eleva por encima del ruido de fondo depende de la cantidad inicial de molde al inicio de la reacción (N_0). El ciclo en que eso sucede recibe el nombre de ciclo umbral (C_t – *threshold cycle*) o punto de cruce (C_p – *crossing point*).

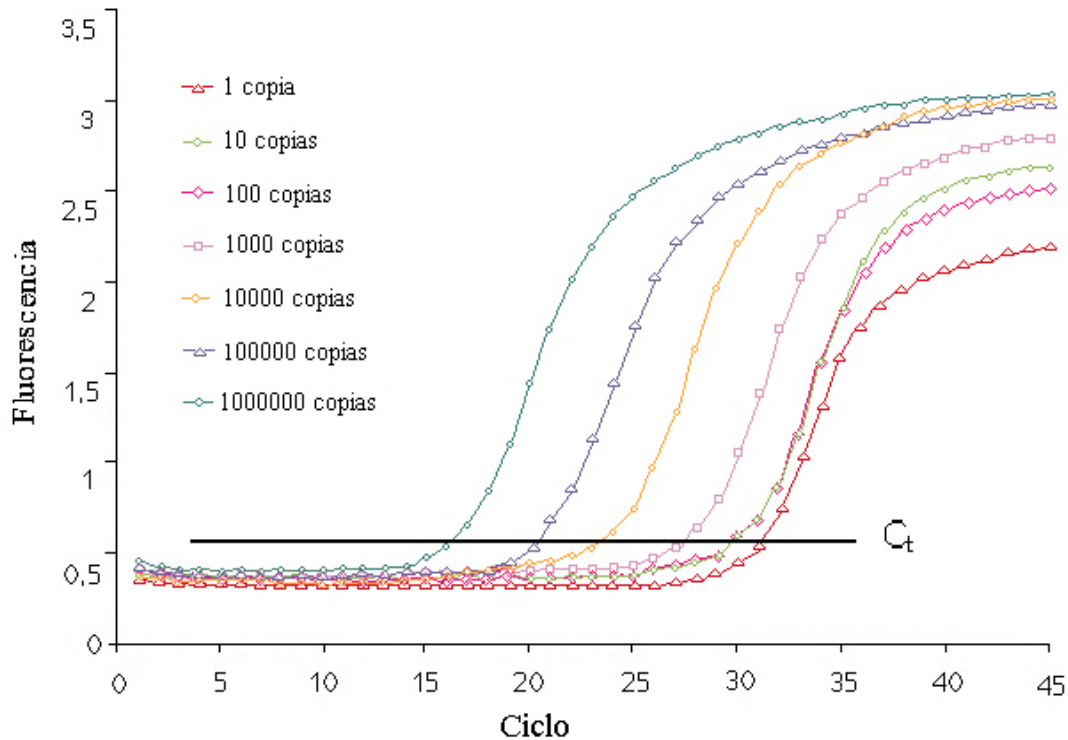


Figura V.6. – Representación gráfica de los resultados de una PCR a tiempo real.

De ese modo se puede cuantificar la expresión de un gen a partir de RNA total, realizando una RT-PCR específica de dicho gen y amplificando posteriormente el cDNA obtenido con los mismos cebadores. Podremos determinar la cantidad inicial de un mRNA concreto, monitorizando la reacción de amplificación del cDNA. A menor C_p , mayor cantidad inicial del mRNA de dicho gen. El único problema es que al cuantificar un gen de interés, no podremos detectar si la cantidad de mRNA total inicial es igual en todas las muestras. Para saber si las diferencias de expresión detectadas son debidas a diferencias reales o a diferencias en las concentraciones iniciales, usaremos un gen de referencia, que se cuantifica paralelamente al gen de interés.

A partir de diluciones seriadas del mRNA total y amplificando el gen de referencia, se construye la recta patrón. Ésta se usará para interpolar el resto de los valores obtenidos: tanto las amplificaciones del gen de referencia en las demás muestras, como las amplificaciones del gen de interés. Para obtener la recta patrón, llamamos K a la cantidad de producto de PCR acumulado en el C_p y substituimos en la ecuación 1 (Rasmussen, 2001):

$$K = N_0 \times (1+E)^{C_t} \tag{2}$$

Si se convierte a una recta se obtiene la siguiente ecuación:

$$\log K = \log N_0 + C_t \times \log(1+E) \tag{3}$$

Reordenando la fórmula en la estructura de una recta Y = mx + b (Figura XXX):

$$C_t = -[1/\log(1+E)] \times \log N_0 + [\log K/\log(1+E)] \tag{4}$$

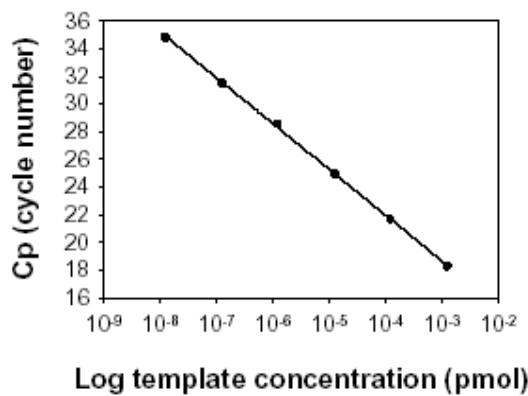


Figura V.7. – Ejemplo de una recta patrón. En el eje de abscisas se representa el logaritmo de la cantidad de ácido nucleico inicial (N₀) y en el eje de ordenadas el punto de cruce (C_p).

En esa recta se interpolan todos los valores C_p obtenidos, estimando así la cantidad inicial de molde, y por tanto la expresión de cada gen. Una vez obtenidos todos los valores se realizan los cocientes gen de interés/gen de referencia y se establecen las relaciones de expresión entre las distintas cepas.

3.1.2. Cebadores y reacción

Los cebadores usados para las diferentes amplificaciones en el *LightCycler* se resumen en la siguiente tabla:

Cebador	Secuencia	Región amplificada	Longitud
RTactinF	5'-GACGATGAAGTTGCTGCTCT-3'	Gen de la actina	153bp
RTactinR	5'-ATACCGACCATCACACCTTG-3'		
RTwhiteF	5'-CGTCATCCTGACCATTTCATC-3'	Gen <i>white</i>	295bp
RTwhiteR	5'-GCCAACAACTGCTCCATATC-3'		
BACNex2	5'-AACCACCAGAATGAATATCG-3'	Región 5' del gen <i>CG32795</i>	185bp
BACNex3	5'-GAGACGTCGTTATCTGGTTC-3'		
BACNex7	5'-GCCTGCTCTATCGACTGAAG-3'	Región 3' del gen <i>CG32795</i>	288bp
BACNex8	5'-ATGCCAGGCTAATCAGAATC-3'		
RTdppF	5'-TCCGAGTCGAACCAGTGTAT-3'	Gen <i>dpp</i>	294bp
RTdppR	5'-TCAGCATCAGGATCAGGATC-3'		

Tabla V.1. – Cebadores utilizados en los diferentes experimentos de PCR a tiempo real.

Es importante resaltar que los cebadores para la PCR a tiempo real fueron diseñados de modo que se encontraran todos en exones diferentes. Así, conociendo el tamaño de la amplificación era posible descartar contaminaciones por amplificación directa en DNA que no hubiera sido eliminado durante la extracción de RNA.

Las reacciones de PCR a tiempo real se han realizado con el *kit RNA Master SYBR Green I* (Roche, Alemania). Dicho *kit* incluye una polimerasa *Hot start*, que sólo será activa a altas temperaturas, cuando los aptámeros que se unen a su centro activo sean liberados. Es la polimerasa *Tth* (*Thermus thermophilus*) que posee actividad DNA polimerasa DNA dependiente y actividad transcriptasa reversa en presencia de RNA e iones manganeso. De esta forma es posible realizar la retrotranscripción y la posterior PCR sobre el cDNA amplificado en un mismo capilar y en un único paso.

Las reacciones de la RT-PCR se han realizado según las condiciones recomendadas por el fabricante:

Componentes	Volumen (por tubo)	Concentración final
H ₂ O	9µL	-
Mn (OAc) ₂ (50mM)	1,3µL	3,25mM
Cebador D	0,6µL	0,3µM
Cebador R	0,6µL	0,3µM
RNA Master SYBR Green I	7,5µL	1x
RNA	2µL	-

Tabla V.2. – **Concentraciones de los diferentes componentes necesarios para realizar una reacción de PCR a tiempo real.**

El programa usado se divide en cinco partes:

Parte 1 – **TRANSCRIPCIÓN REVERSA**. Parte del programa en la que se sintetiza el cDNA a partir del RNA total.

Datos	Valor
Ciclos	1
Modo de análisis	Ninguno
Características de la temperatura	Segmento 1
Temperatura (°C)	61
Tiempo (h:min:seg)	00:20:00
Velocidad de transición de la temperatura (°C/seg)	20
Temperatura secundaria (°C)	0
Intervalo de paso (°C)	0
Retraso del paso (Ciclos)	0
Modo de adquisición	Ninguno

Parte 2 – **DESNATURALIZACIÓN**. El cDNA generado es desnaturalizado de manera que todo el DNA presente queda en forma de ssDNA.

Datos	Valor
Ciclos	1
Modo de análisis	Ninguno
Características de la temperatura	Segmento 1
Temperatura (°C)	95
Tiempo (h:min:seg)	00:00:30
Velocidad de transición de la temperatura (°C/seg)	20
Temperatura secundaria (°C)	0
Intervalo de paso (°C)	0
Retraso del paso (Ciclos)	0
Modo de adquisición	Ninguno

Parte 3 – **AMPLIFICACIÓN**. El cDNA generado durante la retrotranscripción es amplificado mediante ciclos sucesivos de desnaturalización, templado de los cebadores y síntesis de DNA. Se añadió un último paso en el que se medía la fluorescencia. Al incrementar la temperatura en la que se medía la fluorescencia, evitábamos la posible fluorescencia que viniera del *SYBR Green I* unido a los dímeros de cebadores.

Datos	Valor			
Ciclos	45			
Modo de análisis	Cuantificación			
Características de la temperatura	Seg. 1	Seg. 2	Seg. 3	Seg. 4
Temperatura (°C)	95	63	72	80
Tiempo (h:min:seg)	00:00:01	00:00:10	00:00:14	00:00:05
Velocidad de transición de la temperatura (°C/seg)	20	20	20	20
Temperatura secundaria (°C)	0	0	0	0
Intervalo de paso (°C)	0	0	0	0
Retraso del paso (Ciclos)	0	0	0	0
Modo de adquisición	Ninguno	Ninguno	Ninguno	Puntual

Parte 4 – **ANÁLISIS DE LA CURVA DE FUSIÓN**. En esta fase, todo el DNA sintetizado es desnaturalizado y renaturalizado. Partiendo de dsDNA se sube lentamente

la temperatura tomando medidas continuas de fluorescencia. De este modo se puede conocer la T_m del producto sintetizado.

Datos	Valor		
Ciclos	1		
Modo de análisis	Curva de Fusión		
Características de la temperatura	Seg. 1	Seg. 2	Seg. 3
Temperatura (°C)	95	65	95
Tiempo (h:min:seg)	00:00:05	00:00:15	00:00:00
Velocidad de transición de la temperatura (°C/seg)	20	20	0,1
Temperatura secundaria (°C)	0	0	0
Intervalo de paso (°C)	0	0	0
Retraso del paso (Ciclos)	0	0	0
Modo de adquisición	Ninguno	Ninguno	Continuo

Parte 5 – **ENFRIAMIENTO**. Permite enfriar los tubos para poder sacarlos del *LightCycler*.

Datos	Valor
Ciclos	1
Modo de análisis	Ninguno
Características de la temperatura	Segmento 1
Temperatura (°C)	40
Tiempo (h:min:seg)	00:00:30
Velocidad de transición de la temperatura (°C/seg)	20
Temperatura secundaria (°C)	0
Intervalo de paso (°C)	0
Retraso del paso (Ciclos)	0
Modo de adquisición	Ninguno

3.2. Extracción de RNA

La extracción de RNA total se hizo utilizando *TRIzol Reagent* (Invitrogen, Estados Unidos), siguiendo las instrucciones y recomendaciones dadas por el fabricante. Para cada extracción se parte de unos 0,16 gramos de *Drosophila* guardados a -80°C . Después de una desproteinización con cloroformo, el RNA se precipita, primero en isopropanol y finalmente en etanol 70%. El *pellet* se resuspende en 25-50 μL de agua libre de RNasas. Unos 11 μg del RNA total obtenido se tratan durante 30 minutos a 37°C con la DNasa I del *kit* DNA-free™ (Ambion, Reino Unido). Para comprobar la ausencia de DNA contaminante en las muestras se corre una PCR usando los cebadores RTactinF y RTactinR (ver Tabla V.1.). Cada experimento de PCR incluye un control negativo y uno positivo, obteniendo exclusivamente amplificación en el control positivo y en aquellas muestras contaminadas por DNA. Para terminar, las muestras se corren en un gel de agarosa al 1% para comprobar la integridad del RNA.

La concentración del RNA se mide en Tris-EDTA (TE, pH 7,5), asegurándonos siempre que el cociente A_{260}/A_{280} esté entre 1,8 y 2,1.

4. Cuantificación de pigmentos oculares

Con la intención de relacionar el fenotipo ocular con los resultados obtenidos por PCR a tiempo real, realizamos a lo largo de cinco semanas, la cuantificación de los pigmentos oculares en cinco cabezas de individuos adultos de las cepas CS, M115 y RM115.

Se siguió el protocolo descrito por Gubb *et al.*, 1997. Las cabezas se homogenizan con 0,8 mL de etanol 30% (pH=2) y posteriormente se centrifuga. El sobrenadante se recupera y se mide su absorbancia a 480 nm.

5. Ensayo de la nucleasa micrococcal

El ensayo de la nucleasa micrococcal se utilizó con el fin de conocer cual es la disposición de los nucleosomas en regiones concretas del genoma de *D. melanogaster*.

El ensayo requiere aislar núcleos celulares de los organismos a estudiar. La cromatina de esos núcleos es digerida con nucleasa micrococcal, que corta las dos hebras de DNA entre nucleosomas. La digestión es desproteinizada y el DNA obtenido se vuelve a digerir con un enzima de restricción a elegir (*PvuII* en nuestro caso). Finalmente el DNA se corre en un gel de agarosa a partir del cual se realiza un *Southern blot*. El resultado que se obtiene es un escalado de bandas que nos permite posicionar los nucleosomas a lo largo de la secuencia estudiada.

5.1. Aislamiento de núcleos

Todo el proceso debe realizarse en una cámara de 4°C. Para cada extracción se parte de un gramo de individuos de *D. melanogaster*, almacenado previamente a -80°C. Los individuos se pulverizan con un mortero en nitrógeno líquido. Después de que el nitrógeno se evapora, el homogenado se transfiere a un homogenizador de 15mL y se le añaden 3mL de *buffer A*⁺ (60mM KCl, 15 mM NaCl, 13mM EDTA, 0,1mM EGTA, 15mM Tris-HCl (pH 7,4), 0,15mM espermina, 0,5mM espermidina, 0,5mM DTT, 0,5% Nonidet P-40). El homogenado se desagrega con diez emboladas de pistón a 2000 rpm, y se filtra a través de dos capas de malla de nailon. Se usan 2mL de *buffer A*⁺ para limpiar el pistón y la suspensión se filtra a través de las mismas dos capas de malla de nailon. El filtrado se traspa a un tubo de centrifuga y se le añade 1mL de *buffer AS* (60mM KCl, 15mM NaCl, 1mM EDTA, 0,1mM EGTA, 15mM Tris-HCl (pH 7,4), 0,15mM espermina, 0,5mM espermidina, 0,5mM DTT, 0,3M sacarosa). Se centrifuga el tubo a 4°C, 3000rpm durante 5 minutos en una centrifuga de cabezales oscilantes. El sobrenadante se elimina y el *pellet* obtenido se resuspende en 3mL de *buffer A*⁺ y se transfiere al homogenizador, donde se desagrega con cinco emboladas de pistón a 2000 rpm. La suspensión se coloca en un tubo de centrifuga y se le añade 1mL de *buffer AS*. Se vuelve a centrifugar en las mismas condiciones que la vez anterior. El sobrenadante se elimina y el *pellet* obtenido se resuspende en 3mL de *buffer C* (60mM KCl, 15 mM NaCl, 1mM EDTA, 0,1mM EGTA, 15mM Tris-HCl (pH 7,4), 0,15mM espermina, 0,5mM espermidina, 0,5mM DTT). Se transfiere la mezcla a un homegenizador y el *pellet* se desagrega con cinco emboladas de pistón. Se centrifuga a 4°C, 2000 rpm durante 5 minutos. El *pellet* obtenido ya son los núcleos, y se resuspenden en 1mL de *buffer* de la nucleasa micrococcal (60mM KCl, 15mM NaCl, 15mM Tris-HCl (pH 7,4), 0,5mM DTT, 0,25 sacarosa, 1mM CaCl₂).

5.2. Digestión con nucleasa micrococcal

Los núcleos obtenidos se digieren con nucleasa micrococcal (MNasa). Se divide la suspensión de núcleos en cinco tubos que se incuban con 0,0U, 1,0U, 1,5U, 2,0U y 2,5U de MNasa (Roche Applied Science, Alemania) a temperatura ambiente, con agitación, durante 3 minutos. La digestión se finaliza añadiendo 5 μ L de 0,5M EDTA.

5.3. Purificación del DNA y digestión con enzimas de restricción

Terminada la digestión con MNasa, se añaden 10 μ L 10% SDS y 1 μ L proteinasa K (10mg/mL) (Roche Applied Science, Alemania) y se incuban los núcleos toda la noche a 37°C. A la mañana siguiente se desproteína con fenol una vez y dos veces con cloroformo-isoamilalcohol (24:1). Se agrega un volumen de H₂O y 2 μ L de RNasa A (10 mg/mL) (Roche Applied Science, Alemania) y se incuba la mezcla a 40°C durante 2 horas. Finalmente se precipita el DNA con 1/10 volúmenes de acetato sódico 3M y 2 volúmenes de etanol. El *pellet* obtenido se resuspende en 16 μ L de H₂O.

El DNA obtenido se digiere con un enzima de restricción a elegir. Para nuestros experimentos usamos el enzima *PvuII* (Roche Diagnostics GmbH, Alemania), según las instrucciones dadas por el fabricante. Dicha digestión nos permite llevar a cabo un marcaje indirecto del extremo durante el *Southern blot*, puesto que nos da un punto de referencia conocido en cada uno de los fragmentos que se detectaran en el mismo.

5.4. Southern blot

Se corren unos 10 μ g del DNA digerido, durante unas 5 horas a 100V en un gel de agarosa al 1,5%. Una vez terminada la electroforesis se procede a transferir el DNA a un filtro de nailon de carga positiva (Roche Diagnostics GmbH, Alemania). En primer lugar se desnatura el DNA sumergiendo el gel en agitación durante 1 hora en solución de desnaturalización (1,5M NaCl, 0,5M NaOH). A continuación se sumerge el gel durante 1 hora con agitación en solución de neutralización (1M Tris-HCl (pH 8), 1,5M NaCl). Finalmente se monta el dispositivo de transferencia tradicional en 10x SSC (75mM NaCl, 7,5mM citrato sódico 2H₂O) y se deja transfiriendo toda la noche. El DNA se fija en el filtro irradiándolo con luz UV durante 3 minutos.

La hibridación del DNA fijado en el filtro con las sondas se realizó según Sambrook *et al.*, 1989. El filtro se lava de restos del gel de agarosa sumergiéndolo 2

minutos en 6x SSC. Posteriormente se prehibrida durante dos horas a 68°C con solución de prehibridación (6x SSC, 5x reactivo Denhardt, 0,5% SDS, 100 µg/mL DNA de esperma de salmón fragmentado). A continuación se añade la sonda desnaturalizada y marcada radiactivamente con [α -³²P]dATP (Amersham Biosciences, Reino Unido) (ver apartado 5.4.1. Sondas) y se deja hibridando a 68°C toda la noche. A la mañana siguiente se elimina la solución de hibridación y se lava el filtro a temperatura ambiente con 2x SSC y 0,5% SDS. Pasados cinco minutos se retira dicha solución y se deja el filtro durante 15 minutos a temperatura ambiente con 2x SSC y 0,1% SDS. Eliminamos la solución y dejamos el filtro agitando 1 hora a 37°C en 0,1x SSC y 0,5% SDS. Sustituimos la solución con 0,1x SSC y 0,5% SDS nuevo e incrementamos la temperatura a 68°C. Finalmente lavamos 2 minutos el filtro en 0,1x SSC. Se envuelve el filtro en plástico y se deja exponiendo en un casete con autoradiografías Kodak Scientific Imaging Film (Eastman Kodak Company, Estados Unidos) durante una semana.

Los filtros obtenidos fueron reutilizados en distintas ocasiones. Para eliminar la sonda del filtro, se hierve 5 minutos con *buffer* D (10mM Tris-HCl (pH 7,4), 1mM EDTA, 1% SDS). Una vez deshibridados los filtros pueden volver a ser hibridados siguiendo el mismo protocolo, a partir del paso de prehibridación.

5.4.1. Sondas

Todas las sondas para el ensayo de la nucleasa micrococcal se obtuvieron a partir de reacciones de PCR. Los cebadores que se usaron son los que se muestran en la siguiente tabla:

Sonda	Cebador	Secuencia	Tamaño
P4362	4362F	5'-AGGAAGTCTCTTACGAATCG-3'	438bp
	4362R	5'-GGGCATTATCAATCTTCATC-3'	
SP	SPF	5'-ATATTAGTCGACGGTAGCGG-3'	343bp
	SPR	5'-AGAAGAAGCAGAAGATGTTG-3'	
dpp	MNdppF	5'-ACTGAGGACCGAAGACTCTG-3'	390bp
	MNdppR	5'-TTGATGTTGCTGTTGCTGCT-3'	

Tabla V.3. – Cebadores utilizados para generar las sondas del ensayo de la MNasa.

Las reacciones de PCR se llevaron a cabo en un termociclador iCycler™ Thermal Cycler (BIO-RAD Laboratories Inc., Estados Unidos), de efecto Peltier con tapa termoregulable y tubos de 0,2 mL.

Los productos de PCR se purificaron con el *kit* NucleoSpin® Extract II (Macherey-Nagel, Alemania), con el fin de eliminar nucleótidos no incorporados y dímeros de cebadores.

Posteriormente, los productos de PCR se marcaron con [α -³²P]dATP por *random primer* con el *kit* Prime-a-Gene® (Promega, Estados Unidos). Dicho *kit* utiliza una mezcla de hexadeoxiribonucleótidos que actúan como cebador de cualquier dsDNA. Se hierven durante 2 minutos 25ng del DNA a marcar. A continuación se mezcla el DNA desnaturalizado con 10 μ L tampón de marcado 5x, 2 μ L dNTP no marcados (a una concentración de 20 μ M cada uno), 2 μ L BSA, 5 μ L [α -³²P]dATP (50 μ Ci), 5U Klenow y H₂O hasta alcanzar un volumen de 50 μ L. La reacción se deja durante una hora a temperatura ambiente. Finalmente se termina el proceso, hirviendo la sonda durante 2 minutos y añadiendo EDTA hasta una concentración de 20mM. Para eliminar los nucleótidos radioactivos no incorporados y evitar ruido de fondo en el *Southern blot*, las sondas ya marcadas se purifican con el *kit* NucleoSpin® Extract II.

6. Ensayo de modificación del DNA por bisulfito sódico

El ensayo de modificación del DNA por bisulfito sódico (BSM) se usó para determinar cual era el estado de metilación de cinco regiones distintas del genoma de *D. melanogaster*. Dicho ensayo consiste en convertir, a través de una serie de reacciones químicas, aquellas citosinas no metiladas (C) en uracilos (U) (ver Figura V.9.). Las 5mC no serán sustrato de la reacción y por consiguiente permanecerán sin modificar. El DNA tratado con bisulfito sódico, es amplificado por PCR en las regiones de interés. El producto de PCR resultante se clona y posteriormente se analizan 10-20 clones independientes. En las secuencias obtenidas las C originales aparecerán como timinas (T) mientras que las 5mC originales se mantendrán como C.

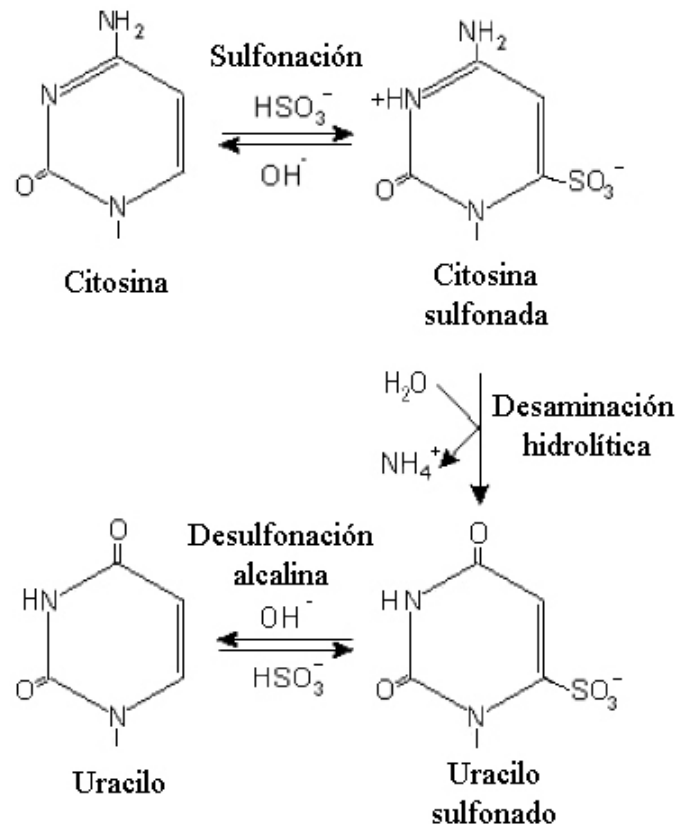


Figura V.9. – Esquema de las reacciones químicas que ocurren durante la modificación del DNA por bisulfito sódico. Permite la transformación de las citosinas (C) no metiladas a uracilos (U).

6.1. Extracción de DNA

La obtención de DNA para el experimento de BSM se realizó a partir de cabezas de hembras de diferentes cepas de *D. melanogaster*. Para conseguir suficiente cantidad de DNA partiendo de tan poco material, se realizó un protocolo de extracción diferente del descrito en el apartado V.2 – Extracción del DNA genómico. El protocolo descrito a continuación es útil también en el caso de querer obtener material a partir de un único individuo.

Se separan las cabezas de los individuos de *D. melanogaster* y se congelan durante al menos 1 minuto a -80°C . Se añaden $100\mu\text{L}$ de solución 1 (60mM NaCl, 50mM EDTA, 10mM Tris-HCl (pH 7,5), 5% sacarosa) y se homogeniza la muestra. Se añaden $100\mu\text{L}$ de solución 2 (100mM EDTA, 300mM Tris-HCl (pH 9), 5% sacarosa, 1,25% SDS) y se mezcla por inversión repetidamente. La muestra se incuba durante 30-40 minutos a 60°C para asegurar la lisis celular y la liberación de su DNA. Se añaden $30\mu\text{L}$

de acetato potásico 3M, se agita y se incuba 45 minutos en hielo. Se centrifuga 1 minuto a 13000 rpm a 4°C. Se transfiere el sobrenadante a un tubo nuevo, se añaden 2 volúmenes de etanol 100% y se deja a temperatura ambiente durante 10 minutos. Se centrifuga 10 minutos a 13000 rpm. Se lava el *pellet* en 500µL de etanol 70%. Se centrifuga 1 minuto a 13000 rpm, se retira el etanol y se deja secar al aire durante 5 minutos. Finalmente se disuelve el *pellet* en 20 µL de H₂O.

6.2. Digestión y tratamiento con bisulfito sódico

Para hacer el DNA más accesible se digirió con el enzima de restricción *HindIII* (Roche, Alemania), siguiendo las instrucciones del fabricante. El enzima elegido depende de las zonas a amplificar, puesto que no se puede elegir un enzima que digiera en medio de una de las regiones que se quieran estudiar.

Una vez digerido el DNA se hizo el tratamiento de bisulfito sódico con el *kit* CpGenomeTM Fast DNA Modification (Chemicon International, Estados Unidos). En primer lugar se mezcla 1µg de DNA y 7µL de NaOH 3M en 100µL de H₂O y se incuba durante 10 minutos a 37°C. Este proceso permite la desnaturalización del DNA a una temperatura suave, gracias al medio alcalino. A continuación se añaden 550µL de reactivo de modificación del DNA (reactivo que incluye el *kit* y que contiene el ion bisulfito (HSO₃⁻) y se incuba durante 20 horas a 50°C, permitiendo así la sulfonación y desaminación hidrolítica de las C no metiladas. De ese modo se obtiene un intermediario de la reacción, los uracilos sulfonados. El DNA tratado se incuba en hielo y pasados 5 minutos se le añaden 750µL de tampón de retención. La muestra se pasa a través de una columna de afinidad, permitiendo eliminar las sales acumuladas. La conversión final a uracilo se completa con una desulfonación, gracias a los 50µL de NaOH 20mM, etanol 90% que se añaden a la muestra. Así, finalmente las C no metiladas habrán sido convertidas a U (ver Figura V.9.).

6.3. PCR

Mediante modificación por bisulfito sódico se estudiaron cinco regiones del genoma de *D. melanogaster*:

- El ZBS del gen *w*.
- Una región del primer exón de *w*, localizada 670bp 3' del ZBS de *w*.

- El ZBS del gen *dpp*.
- Una región del primer intrón de *dpp*, localizada 938bp 3' del ZBS de *dpp*.
- Una región del tercer exón del gen *CG32795*.

Los cebadores que se usaron para amplificar las diferentes regiones tuvieron que ser modificados para que fueran complementarios a la secuencia modificada por la reacción del bisulfito sódico. Así pues, dichos cebadores directos no contenían ninguna C y en su lugar había T, y los cebadores reversos no contenían ninguna G y en su lugar había A. En la siguiente tabla se detallan los cebadores usados:

Cebador	Secuencia	Región amplificada	Tamaño
wZBS_BMPF	5'-GGTTTTAAGTATTTTTGTG-3'	ZBS del gen <i>w</i>	522bp
wZBS_BMPR	5'-ATTTAACTCAAATAATAACTTTT-3'		
w_BSMF	5'-GATATTATTGAATTTATTTGATTA-3'	Primer exón del gen <i>w</i>	490bp
w_BSMR	5'-ACACATATACTACTCACATTAT-3'		
CG_BSMF	5'-AAATAGTTGTTTGTTTTAAAAA-3'	Tercer exón del gen <i>CG32795</i>	601bp
CG_BSMR	5'-CAATAACCTCCAATACTCACT-3'		
dZBS_BSMF	5'-GGTATTAGATAATATATTTGTAA-3'	ZBS del gen <i>dpp</i>	448bp
dZBS_BSMR	5'-TAATTCAATTAATCTCAAATT-3'		
dpp_BSMF	5'-TTTTGAAATGTGTAAAGGTATA-3'	Primer intrón del gen <i>dpp</i>	464bp
dpp_BSMR	5'-TACTAATACCACCACTTTTAA-3'		

Tabla V.4. – Cebadores utilizados en los diferentes experimentos de modificación del DNA por bisulfito sódico.

Las reacciones de PCR se llevaron a cabo en un termociclador iCycler™ Thermal Cycler (BIO-RAD Laboratories Inc., Estados Unidos), de efecto Peltier con tapa termoregurable y tubos de 0,2mL.

El programa usado en cada reacción fue distinto. En la siguiente tabla se detallan los mismos:

Paso	wZBS	w	CG	dZBS	dpp	Ciclos
Desnaturalización	3 min 94°C	3 min 94°C	3 min 94°C	3 min 94°C	3 min 94°C	x1
Desnaturalización	1 min 94°C	30 seg 94°C	30 seg 94°C	30 seg 94°C	1 min 94°C	x10
Templado	30 seg 54°C	1 min 50°C	1 min 54°C	1 min 50°C	30 seg 54°C	
Extensión	1 min 72°C	30 seg 72°C	30 seg 72°C	30 seg 72°C	1 min 72°C	
Desnaturalización	1 min 94°C	30 seg 94°C	30 seg 94°C	30 seg 94°C	1 min 94°C	x20
Templado	30 seg 51°C	1 min 47°C	1 min 51°C	1 min 47°C	30 seg 51°C	
Extensión	1 min 72°C	30 seg 72°C	30 seg 72°C	30 seg 72°C	1 min 72°C	
Extensión	3 min 72°C	3 min 72°C	3 min 72°C	3 min 72°C	3 min 72°C	x1
Enfriamiento	∞ 4°C	∞ 4°C	∞ 4°C	∞ 4°C	∞ 4°C	x1

Tabla V.5. – **Programas de PCR para amplificar las distintas regiones del análisis del patrón de metilación por BSM.**

6.4. Clonación

Los productos amplificados gracias a la PCR fueron purificados con el *kit* NucleoSpin[®] Extract II (Macherey-Nagel, Alemania) para eliminar nucleótidos no incorporados y dímeros de cebadores. Posteriormente se clonaron con el sistema TOPO TA Cloning[®] Kit (Invitrogen, Estados Unidos).

El sistema TOPO TA Cloning[®] Kit facilita la clonación de productos de PCR gracias a las deoxitimas presentes en los extremos 3' del vector. Dichas timinas son complementarias a las deoxiadeninas presentes en los extremos 3' del producto de PCR, que son añadidas debido a la actividad transferasa terminal independiente de molde de la *taq* polimerasa. Además, los dos extremos del vector TOPO[®] poseen un enlace fosfo-tirosil al que se une la topoisomerasa I del virus *Vaccinia*. El enlace fosfo-tirosil puede ser atacado por un 5'-OH del producto de PCR, liberando la topoisomerasa y dejando el producto de PCR unido al vector (ver Figura V.10.).

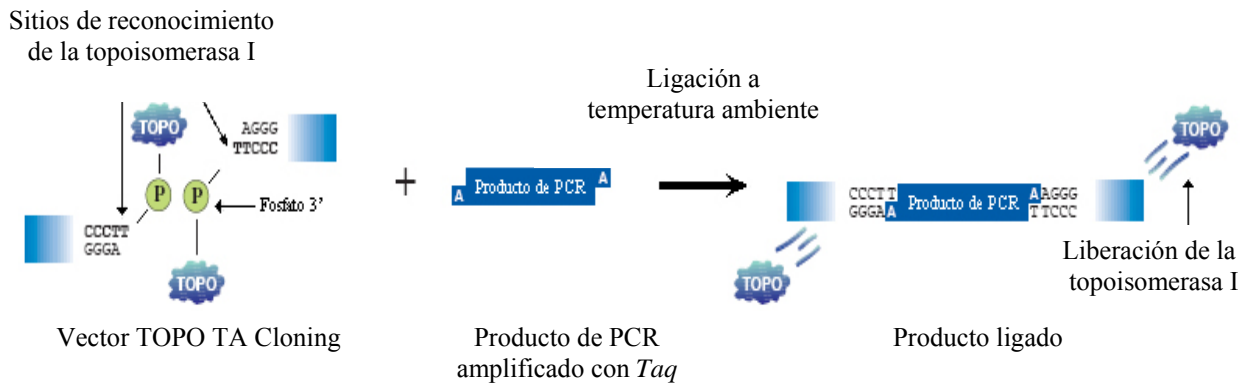


Figura V.10. – Esquema de ligación de un producto de PCR con el vector TOPO TA Cloning®.

Para realizar la reacción de clonación, se mezclan 4µL de producto de PCR, 1µL de solución salina (1,2M NaCl, 60mM MgCl₂) y 1µL de vector TOPO®. La mezcla se deja a temperatura ambiente durante 30 minutos. Este tiempo se aprovecha para descongelar las células TOP10 (Invitrogen, Estados Unidos) que son las que se usarán en la transformación. Pasados los 30 minutos, 2µL de la reacción de ligado se introducen en el vial de las células y se deja reposando en hielo durante 30 minutos. A continuación se somete a las células a un choque térmico de 45 segundos a 42°C. Inmediatamente se les añaden 250µL de medio de cultivo SOC (2% triptona, 0,5% extracto de levadura, 10mM NaCl, 2,5mM KCl, 10mM MgCl₂, 10mM MgSO₄, 20mM glucosa) a temperatura ambiente. El cultivo se deja en agitación (200 rpm) a 37°C durante 1 hora. Finalmente se siembran 50-100µL en placas de LB con ampicilina. Las placas se incuban a 37°C toda la noche.

A la mañana siguiente han crecido ya las colonias, a partir de las cuales se realizan cultivos de noche en medio LB con ampicilina. La extracción del DNA plasmídico (*miniprep*) se llevó a cabo con el *kit* NucleoSpin® Plasmid QuickPure (Macherey-Nagel, Alemania).

6.5. Secuenciación y análisis de las secuencias

El DNA plasmídico obtenido fue secuenciado por la empresa Macrogen (Corea) con el sistema ABI3730XL.

Las secuencias de cada región estudiada fueron consideradas como un conjunto. No buscábamos sitios específico de metilación sino un patrón global de la zona. Así considerando todas las secuencias de una misma región como un todo, la potencia estadística conseguida era mucho mayor. Evidentemente, las secuencias se clasificaron no sólo por regiones sino también por cepas.

Se estableció una relación entre las bases no metiladas y las metiladas de cada región y cepa. El programa usado para analizar nuestros datos fue el G-Stat con un test Chi-cuadrado de datos agrupados. Se consideraron estadísticamente significativos aquellos valores de $p \leq 0,05$.

7. Análisis del gen *CG32795*

La única información disponible a cerca del gen *CG32795* era la predicción bioinformática hecha por *FlyBase*. Decidimos estudiarlo más detenidamente, con el fin de conocer más detalles sobre su secuencia, *splicing*, posible función, etc.

7.1. RACE

La técnica RACE (Amplificación Rápida de los Extremos de cDNA - *Rapid Amplification of cDNA Ends*) nos permite, como su nombre indica, amplificar los dos extremos de un determinado cDNA. En nuestro trabajo conseguimos capturar los extremos 5' y 3' del gen *CG32795*.

Para realizar el RACE, primero fue necesario obtener RNA total, siguiendo exactamente el mismo protocolo detallado en el apartado V.3.2.

A partir del RNA total se usó el 5'/3' RACE Kit, 2nd Generation (Roche Applied Science, Alemania) para la obtención del cDNA. La estrategia para capturar el extremo 5' de un cDNA es distinta de la usada para capturar el extremo 3'. Así pues, vamos a detallar los protocolos de forma independiente.

7.1.1. RACE 5'

El primer paso consiste en realizar, partiendo de RNA total (obtenido según se detalla en el apartado V.3.2.), una PCR retrotranscriptiva (RT-PCR) con un cebador

específico, en nuestro caso el BACNex7 (5' – CTTCAGTCGATAGAGCAGGC – 3') (ver Figura V.11.). La reacción contiene: 4μL tampón de síntesis del cDNA, 2μL dNTP, 1μL cebador BACNex7, 2μg RNA total, 1μL transcriptasa reversa y H₂O hasta conseguir un volumen de 20μL. La mezcla se incuba a 55°C durante 1 hora y después a 85°C durante 5 minutos.

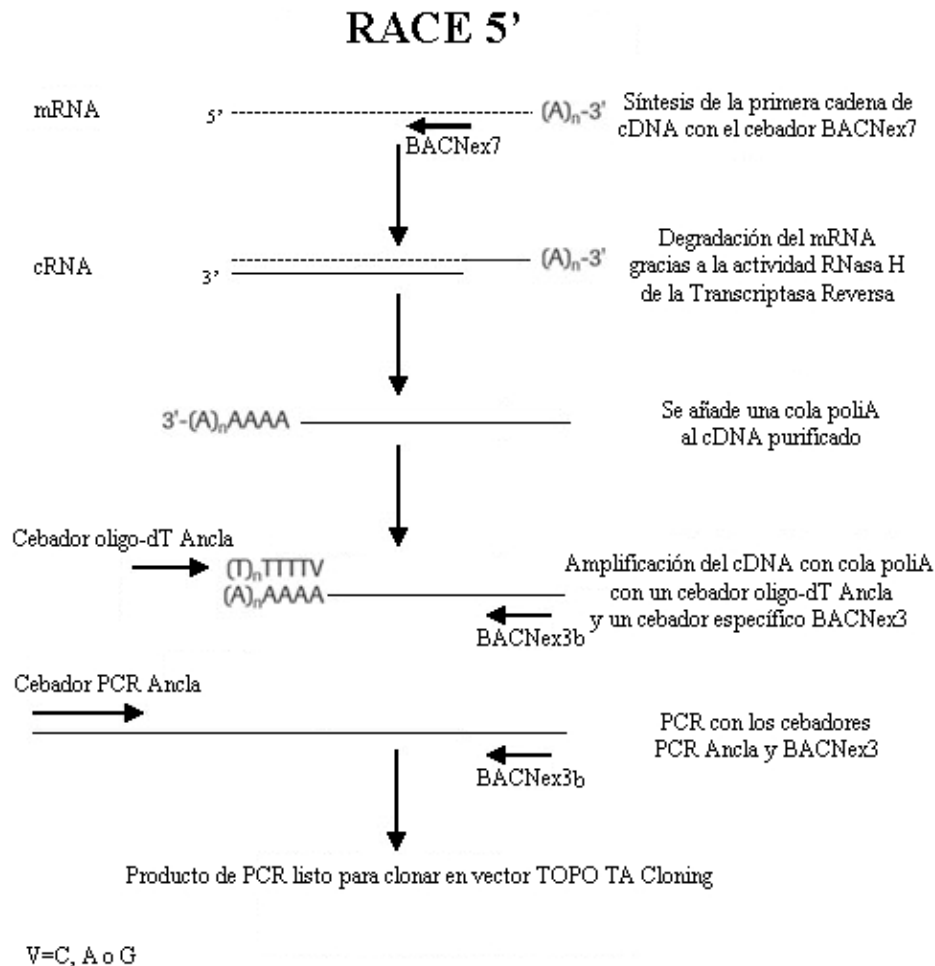


Figura V.11. – Esquema del protocolo completo de RACE 5'.

En segundo lugar hay que unir una cola poliA al extremo 5' del cDNA sintetizado (ver Figura V.11.). Para eso es necesario purificar el cDNA obtenido. Utilizamos con este propósito el *kit* NucleoSpin[®] Extract II (Macherey-Nagel, Alemania). Se mezclan 19μL cDNA purificado, 2,5μL tampón de reacción (10x) y 2,5μL dATP (2mM) y se incuban a 94°C durante 3 minutos. Se deja enfriar en hielo, se añade a la mezcla 1μL de transferasa terminal (80U/μL) y se incuba a 37°C durante 30 minutos. Finalmente se incuba a 70°C durante 10 minutos para inactivar la transferasa terminal.

En tercer lugar se realiza una PCR sobre el cDNA con la cola poliA (ver Figura V.11.). En este caso, los cebadores utilizados son: Oligo-dT Ancla (5' - GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTTTT - 3') (V=A, C o G) y un cebador específico, en nuestro caso BACNex3b (5' – GAGACGTCGTTATCTGGTTC – 3'). Para la reacción se mezclan: 5µL cDNA con cola poliA, 1µL cebador Oligo-dT Ancla, 1µL BACNex3b (12,5µM), 1µL dNTP, 0,5µL *Taq* polimerasa (5U/µL), 5µL tampón de reacción, 36,5µL H₂O. La reacción de PCR se llevó a cabo en un termociclador iCycler™ Thermal Cycler (BIO-RAD Laboratories Inc., Estados Unidos), de efecto Peltier con tapa termoregurable y tubos de 0,2mL y el programa usado fue el descrito en la siguiente tabla:

Paso	Temperatura y tiempo	Ciclos
Desnaturalización	5 min 94°C	x1
Desnaturalización	45 seg 94°C	x10
Templado	30 seg 54°C	
Extensión	40 seg 72°C	
Desnaturalización	45 seg 94°C	x25 Incrementando 20 segundos la extensión a cada ciclo.
Templado	30 seg 54°C	
Extensión	40 seg 72°C	
Extensión	7 min 72°C	x1
Enfriamiento	∞ 4°C	x1

Tabla V.6. – Programa de PCR para amplificar el extremo 5' del gen *CG32795*.

El producto resultante de esta PCR vuelve a ser amplificado usando en esta ocasión los cebadores PCR Ancla (5' – GACCACGCGTATCGATGTCGAC – 3') y el cebador específico BACNex3b (ver Figura V.11.). Esta reacción de PCR se realizó con las mismas concentraciones de reactivo y condiciones de tiempo y temperatura que la anterior.

Finalmente, el producto obtenido fue clonado con el sistema TOPO TA Cloning[®] descrito previamente en el apartado V.5.4. Los clones obtenidos fueron secuenciados y los resultados se analizaron con el programa Chromas 1.45 (Conor McCarthy School of Health and Science, Australia).

7.1.2. RACE 3'

El RACE 3' permite obtener el extremo 3' del cDNA de interés. En este caso, el primer paso consiste en realizar, partiendo de RNA total (obtenido según se detalla en el apartado V.3.2.), una PCR retrotranscriptiva (RT-PCR) con el cebador oligo-dT Ancla (ver Figura V.12.). La reacción contiene: 4µL tampón de síntesis del cDNA, 2µL dNTP, 1µL cebador oligo-dT Ancla, 2µg RNA total, 1µL transcriptasa reversa y H₂O hasta conseguir un volumen de 20µL. La mezcla se incuba a 55°C durante 1 hora y después a 85°C durante 5 minutos.

En segundo lugar se realiza una PCR sobre el cDNA sintetizado, usando como cebadores: PCR Ancla y un cebador específico, en nuestro caso BACNex7b (5' – GCCTGCTCTATCGACTGAAG – 3'). Para la reacción de PCR se mezclan: 1µL cDNA, 1µL cebador PCR Ancla, 1µL BACNex7b (12,5µM), 1µL dNTP, 0,5µL *Taq* polimerasa (5U/µL), 5µL tampón de reacción, 36,5µL H₂O. La reacción de PCR se llevó a cabo en el termociclador iCycler[™] Thermal Cycler (BIO-RAD Laboratories Inc., Estados Unidos), y el programa usado fue el descrito en la siguiente tabla:

Paso	Temperatura y tiempo	Ciclos
Desnaturalización	3 min 94°C	x1
Desnaturalización	45 seg 94°C	x30
Templado	45 seg 60°C	
Extensión	1 min 72°C	
Extensión	10 min 72°C	x1
Enfriamiento	∞ 4°C	x1

Tabla V.7. – Programa de PCR para amplificar el extremo 3' del gen CG32795.

El producto resultante de esta PCR fue clonado con el sistema TOPO TA Cloning[®] descrito previamente en el apartado V.5.4. Los clones obtenidos fueron secuenciados y los resultados se analizaron con el programa Chromas 1.45 (Conor McCarthy School of Health and Science, Australia).

RACE 3'

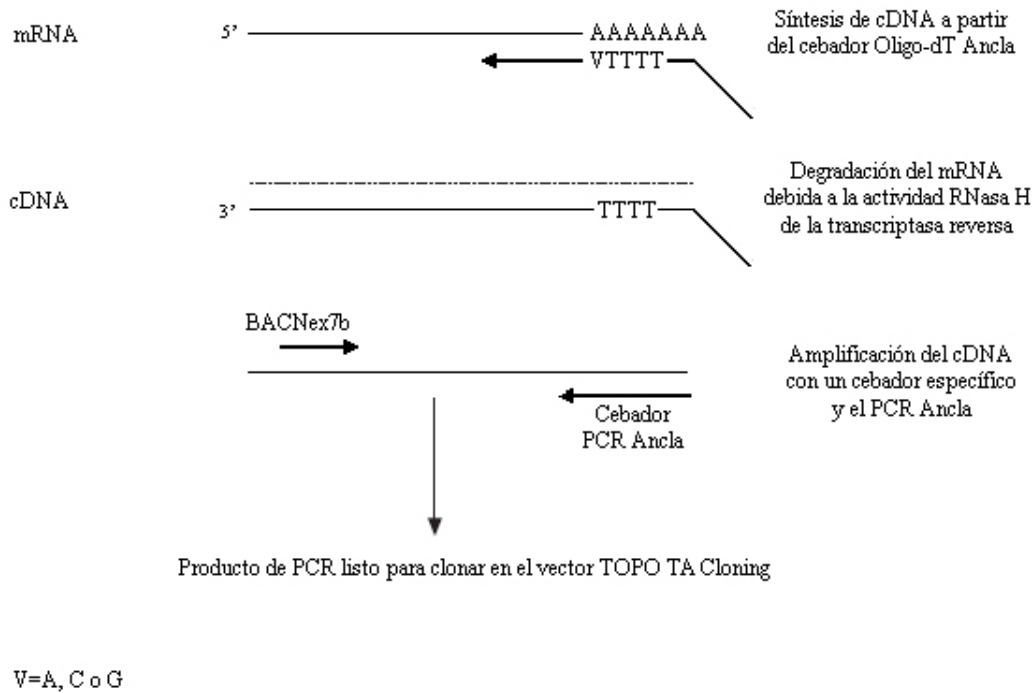


Figura V.12. – Esquema del protocolo completo de RACE 3'.

7.2. RT-PCR

Habiendo obtenido los extremos 5' y 3' del cDNA de interés, sólo quedaba obtener la región central de dicho transcrito. Con ese fin se realizó una RT-PCR (PCR retrotranscriptiva).

La RT-PCR se realizó con la transcriptasa reversa SuperScript[™]II RNasa H (Invitrogen, Estados Unidos), partiendo de RNA total (obtenido según se detalla en el apartado V.3.2.). Para la reacción se mezcla 5ng de RNA total, 2pmoles de un cebador específico (en nuestro caso BACNex8b (5' - ATGCCAGGCTAATCAGAATC - 3')), 1µL dNTP (10mM de cada nucleótido) y H₂O hasta alcanzar un volumen total de 12µL. La mezcla se incubaba a 65°C y pasados 5 minutos se coloca en hielo. Se añaden 4µL 5x

tampón de síntesis de la primera cadena y 2µL 0,1M DDT y se incubaba a 42°C durante 2 minutos. Se añade 1µL de retrotranscriptasa SuperScript™II y se incubaba a 42°C durante 50 minutos. Finalmente se inactiva la retrotranscriptasa incubando a 72°C durante 15 minutos.

El cDNA obtenido en esta RT-PCR se amplifica mediante una PCR. En nuestro experimento los cebadores utilizados fueron: BACNex2b (5' - AACCACCAGAATGAATATCG - 3') y BACNex7b. Para la reacción de PCR se mezclan: 1µL cDNA, 0,25µM de ambos cebadores, 1x tampón de la *Taq* polimerasa, 0,25mM MgCl₂ y 1,5U *Taq* polimerasa y H₂O hasta llegar a un volumen de reacción de 25µL. La reacción de PCR se llevó a cabo en el termociclador iCycler™ Thermal Cycler (BIO-RAD Laboratories Inc., Estados Unidos), y el programa usado fue el descrito en la siguiente tabla:

Paso	Temperatura y tiempo	Ciclos
Desnaturalización	5 min 94°C	x1
Desnaturalización	30 seg 94°C	x30
Templado	30 seg 55°C	
Extensión	1 min 72°C	
Extensión	7 min 72°C	x1
Enfriamiento	∞ 4°C	x1

Tabla V.8. – Programa de RT-PCR para amplificar la región central del gen CG32795.

El producto resultante de esta PCR fue clonado con el sistema TOPO TA Cloning® descrito previamente en el apartado V.5.4. Los clones obtenidos fueron secuenciados y los resultados se analizaron con el programa Chromas 1.45 (Conor McCarthy School of Health and Science, Australia).

La RT-PCR también se utilizó en una aproximación previa al RACE para comprobar que efectivamente el gen *CG32795* se expresaba en todas las cepas que se querían estudiar. En este caso se usaron las parejas de cebadores: BACNex2b con BACNex3b y BACNex7b con BACNex8b, siguiendo el protocolo descrito sin el último paso de clonación.

7.3. Análisis bioinformático

Las secuencias obtenidas del transcrito completo, fueron traducidas, mediante el programa EMBOSS Transeq (Rice *et al.*, 2000) (<http://www.ebi.ac.uk/translate/>), para conseguir la secuencia aminoacídica de la proteína resultante.

Las secuencias aminoacídicas se analizaron con los programas: SOSUI (Hirokawa *et al.*, 1998) (<http://bp.nuap.nagoya-u.ac.jp/sosui/>), TMHMM (Krogh *et al.*, 2001) (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) y HMMTOP (Tusnady y Simon, 2001) (<http://www.enzim.hu/hmmtop/>) con el fin de determinar la estructura secundaria de la proteína estudiada. También se usó el programa NCBI CD (*Conserved Domain Search*) (Marchler-Bauer *et al.*, 2005) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) para descubrir la presencia de posibles dominios funcionales descritos con anterioridad. El programa Pfam (Finn *et al.*, 2006) (<http://www.sanger.ac.uk/Software/Pfam/>) permitió evaluar la conservación del dominio proteico encontrado, a lo largo de la evolución.

Tanto las secuencias nucleotídicas como las aminoacídicas fueron analizadas con el programa BLAST (Altschul *et al.*, 1997) (<http://www.ncbi.nlm.nih.gov/BLAST/>), con la intención de encontrar similitudes locales con secuencias de otras especies. Las secuencias obtenidas mediante el BLAST fueron alineadas mediante el programa ClustalW (Thompson *et al.*, 1994) (<http://www.ebi.ac.uk/clustalw/>), en busca de regiones o motivos conservados a lo largo de la evolución.

VI. BIBLIOGRAFÍA

- Aalfs, J. D. y R. E. Kingston (2000). "What does 'chromatin remodeling' mean?" Trends Biochem Sci **25**(11): 548-55.
- Allis, C. D., T. Jenuwein y D. Reinberg (2007). Epigenetics, Cold Spring Harbor Laboratory Press.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller y D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.
- Bachman, K. E., B. H. Park, I. Rhee, H. Rajagopalan, J. G. Herman, S. B. Baylin, K. W. Kinzler y B. Vogelstein (2003). "Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene." Cancer Cell **3**(1): 89-95.
- Badal, M. (2007). "Origen del fenotipo *zeste* en machos de la cepa M115 de *D. melanogaster* y causas de su reversión: el elemento transponible *FB-NOF*." Tesis doctoral. Universitat Autònoma de Barcelona.
- Badal, M., A. Portela, E. Baldrich, R. Marcos, O. Cabré y N. Xamena (2006a). "An *FB-NOF* mediated duplication of the *white* gene is responsible for the *zeste*¹ phenotype in some *Drosophila melanogaster* unstable strains." Mol Genet Genomics **275**(1): 35-43.
- Badal, M., A. Portela, N. Xamena y O. Cabré (2006b). "Molecular and bioinformatic analysis of the *FB-NOF* transposable element." Gene **371**(1): 130-5.
- Baldrich, E. (1998). "Análisis genético y molecular de los mutantes de *Drosophila melanogaster* obtenidos a partir de individuos *white-apricot* sometidos a choque térmico y a tratamiento N-etil-N-nitrosurea." Tesis doctoral. Universitat Autònoma de Barcelona.
- Banine, F., C. Bartlett, R. Gunawardena, C. Muchardt, M. Yaniv, E. S. Knudsen, B. E. Weissman y L. S. Sherman (2005). "SWI/SNF chromatin-remodeling factors induce changes in DNA methylation to promote transcriptional activation." Cancer Res **65**(9): 3542-7.
- Becker, P. B. (2002). "Nucleosome sliding: facts and fiction." Embo J **21**(18): 4749-53.
- Beerens, N., J. H. Hoeijmakers, R. Kanaar, W. Vermeulen y C. Wyman (2005). "The CSB protein actively wraps DNA." J Biol Chem **280**(6): 4722-9.

- Benson, M. y 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**(12): 3907-15.
- Bickel, S. y V. Pirrotta (1990). "Self-association of the *Drosophila* zeste protein is responsible for transvection effects." EMBO J **9**(9): 2959-67.
- Biémont, C., A. Tsitrone, C. Vieira y C. Hoogland (1997). "Transposable element distribution in *Drosophila*." Genetics **147**(4): 1997-9.
- Biémont, C. y C. Vieira (2006). "Junk DNA as an evolutionary force." Nature **443**(7111): 521-4.
- Bingham, P. M., M. G. Kidwell y G. M. Rubin (1982). "The molecular basis of P-M hybrid dysgenesis: the role of the *P* element, a *P*-strain-specific transposon family." Cell **29**(3): 995-1004.
- Bingham, P. M. y Z. Zachar (1985). "Evidence that two mutations, w^{DZL} and z^1 , affecting synapsis-dependent genetic behavior of *white* are transcriptional regulatory mutations." Cell **40**(4): 819-25.
- Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev **16**(1): 6-21.
- Bird, A. P., M. H. Taggart, R. D. Nicholls y D. R. Higgs (1987). "Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene." EMBO J **6**(4): 999-1004.
- Bird, A. P. y A. P. Wolffe (1999). "Methylation-induced repression--belts, braces, and chromatin." Cell **99**(5): 451-4.
- Brierley, H. L. y S. S. Potter (1985). "Distinct characteristics of loop sequences of two *Drosophila foldback* transposable elements." Nucleic Acids Res **13**(2): 485-500.
- Brumby, A. M., C. B. Zraly, J. A. Horsfield, J. Secombe, R. Saint, A. K. Dingwall y H. Richardson (2002). "*Drosophila* cyclin E interacts with components of the Brahma complex." EMBO J **21**(13): 3377-89.
- Brunner, S., G. Pea y A. Rafalski (2005). "Origins, genetic organization and transcription of a family of non-autonomous *helitron* elements in maize." Plant J **43**(6): 799-810.
- Brzeski, J. y A. Jerzmanowski (2003). "Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors." J Biol Chem **278**(2): 823-8.

- Burton, F. H., D. D. Loeb, S. F. Chao, C. A. Hutchison, 3rd y M. H. Edgell (1985). "Transposition of a long member of the L1 major interspersed DNA family into the mouse *beta globin* gene locus." Nucleic Acids Res **13**(14): 5071-84.
- Buryanov, Y. I. y T. V. Shevchuk (2005). "DNA methyltransferases and structural-functional specificity of eukaryotic DNA modification." Biochemistry (Mosc) **70**(7): 730-42.
- Cáceres, M., J. M. Ranz, A. Barbadilla, M. Long y A. Ruiz (1999). "Generation of a widespread *Drosophila* inversion by a transposable element." Science **285**(5426): 415-8.
- Campbell, G., T. Weaver y A. Tomlinson (1993). "Axis specification in the developing *Drosophila* appendage: the role of *wingless*, *decapentaplegic*, and the homeobox gene *aristales*." Cell **74**(6): 1113-23.
- Cardoso, M. C. y H. Leonhardt (1999). "DNA methyltransferase is actively retained in the cytoplasm during early development." J Cell Biol **147**(1): 25-32.
- Casacuberta, E., J. M. Casacuberta, P. Puigdomenech y A. Monfort (1998). "Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the *Emigrant* family of elements." Plant J **16**(1): 79-85.
- Casals, F., M. Cáceres y A. Ruíz (2003). "The *foldback-like* transposon *Galileo* is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*." Mol Biol Evol **20**(5): 674-85.
- Chambeyron, S. y A. Bucheton (2005). "*I* elements in *Drosophila*: in vivo retrotransposition and regulation." Cytogenet Genome Res **110**(1-4): 215-22.
- Chan, S. W., I. R. Henderson y S. E. Jacobsen (2005). "Gardening the genome: DNA methylation in *Arabidopsis thaliana*." Nat Rev Genet **6**(5): 351-60.
- Chance, P. F., N. Abbas, M. W. Lensch, L. Pentao, B. B. Roa, P. I. Patel y J. R. Lupski (1994). "Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17." Hum Mol Genet **3**(2): 223-8.
- Chen, J. D. y V. Pirrotta (1993a). "Multimerization of the *Drosophila* *zeste* protein is required for efficient DNA binding." EMBO J **12**(5): 2075-83.
- Chen, J. D. y V. Pirrotta (1993b). "Stepwise assembly of hyperaggregated forms of *Drosophila* *zeste* mutant protein suppresses *white* gene expression in vivo." EMBO J **12**(5): 2061-73.

- Cheung, P., K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu y C. D. Allis (2000). "Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation." Mol Cell **5**(6): 905-15.
- Choi, J. D., A. Hoshino, K. I. Park, I. S. Park y S. Iida (2007). "Spontaneous mutations caused by a *Helitron* transposon, *Hel-It1*, in morning glory, *Ipomoea tricolor*." Plant J **49**(5): 924-34.
- Clyde, D. y M. Bownes (2000). "The *Dstpk61* locus of *Drosophila* produces multiple transcripts and protein isoforms, suggesting it is involved in multiple signalling pathways." J Endocrinol **167**(3): 391-401.
- Colot, V. y J. L. Rossignol (1999). "Eukaryotic DNA methylation as an evolutionary device." Bioessays **21**(5): 402-11.
- Comb, M. y H. M. Goodman (1990). "CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2." Nucleic Acids Res **18**(13): 3975-82.
- Crooks, G. E., G. Hon, J. M. Chandonia y S. E. Brenner (2004). "WebLogo: a sequence logo generator." Genome Res **14**(6): 1188-90.
- de la Cruz, X., S. Lois, S. Sanchez-Molina y M. A. Martinez-Balbas (2005). "Do protein motifs read the histone code?" Bioessays **27**(2): 164-75.
- Deininger, P. L. y M. A. Batzer (1999). "Alu repeats and human disease." Mol Genet Metab **67**(3): 183-93.
- Dejardin, J. y G. Cavalli (2004). "Chromatin inheritance upon Zeste-mediated Brahma recruitment at a minimal cellular memory module." EMBO J **23**(4): 857-68.
- Dewannieux, M., C. Esnault y T. Heidmann (2003). "LINE-mediated retrotransposition of marked Alu sequences." Nat Genet **35**(1): 41-8.
- Dominski, Z., X. C. Yang, C. S. Raska, C. Santiago, C. H. Borchers, R. J. Duronio y W. F. Marzluff (2002). "3' end processing of *Drosophila melanogaster* histone pre-mRNAs: requirement for phosphorylated *Drosophila* stem-loop binding protein and coevolution of the histone pre-mRNA processing system." Mol Cell Biol **22**(18): 6648-60.
- Dorn, R., G. Reuter y A. Loewendorf (2001). "Transgene analysis proves mRNA trans-splicing at the complex *mod(mdg4)* locus in *Drosophila*." Proc Natl Acad Sci U S A **98**(17): 9724-9.
- Ehrlich, M. y R. Y. Wang (1981). "5-Methylcytosine in eukaryotic DNA." Science **212**(4501): 1350-7.

- Emmons, S. W., L. Yesner, K. S. Ruan y D. Katzenberg (1983). "Evidence for a transposon in *Caenorhabditis elegans*." Cell **32**(1): 55-65.
- Ferguson, E. L. y K. V. Anderson (1992). "*Decapentaplegic* acts as a morphogen to organize dorsal-ventral pattern in the *Drosophila* embryo." Cell **71**(3): 451-61.
- Ferres-Marco, D., I. Gutiérrez-García, D. M. Vallejo, J. Bolívar, F. J. Gutiérrez-Aviño y M. Domínguez (2006). "Epigenetic silencers and Notch collaborate to promote malignant tumours by *Rb* silencing." Nature **439**(7075): 430-6.
- Field, L. M., F. Lyko, M. Mandrioli y G. Prantera (2004). "DNA methylation in insects." Insect Mol Biol **13**(2): 109-15.
- Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer y A. Bateman (2006). "Pfam: clans, web tools and services." Nucleic Acids Res **34**(Database issue): D247-51.
- Finnegan, D. J. (1989). "Eukaryotic transposable elements and genome evolution." Trends Genet **5**(4): 103-7.
- Fjose, A., L. C. Polito, U. Weber y W. J. Gehring (1984). "Developmental expression of the *white* locus of *Drosophila melanogaster*." EMBO J **3**(9): 2087-2094.
- Foster, T. J. (1977). "Insertion of the tetracycline resistance translocation unit *Tn10* in the *lac* operon of *Escherichia coli* K12." Mol Gen Genet **154**(3): 305-9.
- Fry, C. J. y C. L. Peterson (2001). "Chromatin remodeling enzymes: who's on first?" Curr Biol **11**(5): R185-97.
- Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies y T. Kouzarides (2000). "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." Nat Genet **24**(1): 88-91.
- Gafner, J. y P. Philippsen (1980). "The yeast transposon *Ty1* generates duplications of target DNA on insertion." Nature **286**(5771): 414-8.
- Gelbart, W. M. (1982). "Synapsis-dependent allelic complementation at the *decapentaplegic* gene complex in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **79**(8): 2636-40.
- Gilfillan, G. D., T. Straub, E. de Wit, F. Greil, R. Lamm, B. van Steensel y P. B. Becker (2006). "Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex." Genes Dev **20**(7): 858-70.
- Goldberg, M. L., R. A. Colvin y A. F. Mellin (1989). "The *Drosophila zeste* locus is nonessential." Genetics **123**(1): 145-55.

- Goldin, S. N. y V. E. Papaioannou (2003). "Unusual misregulation of RNA splicing caused by insertion of a transposable element into the T (*Brachyury*) locus." BMC Genomics **4**(1): 14.
- Goll, M. G. y T. H. Bestor (2005). "Eukaryotic cytosine methyltransferases." Annu Rev Biochem **74**: 481-514.
- Goll, M. G., F. Kirpekar, K. A. Maggert, J. A. Yoder, C. L. Hsieh, X. Zhang, K. G. Golic, S. E. Jacobsen y T. H. Bestor (2006). "Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2." Science **311**(5759): 395-8.
- Gotea, V. y W. Makalowski (2006). "Do transposable elements really contribute to proteomes?" Trends Genet **22**(5): 260-7.
- Grandbastien, M. A., A. Spielmann y M. Caboche (1989). "*Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics." Nature **337**(6205): 376-80.
- Gray, Y. H. (2000). "It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements." Trends Genet **16**(10): 461-8.
- Grimaud, C., N. Negre y G. Cavalli (2006). "From genetics to epigenetics: the tale of *Polycomb* group and *trithorax* group genes." Chromosome Res **14**(4): 363-75.
- Grune, T., J. Brzeski, A. Eberharter, C. R. Clapier, D. F. Corona, P. B. Becker y C. W. Muller (2003). "Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI." Mol Cell **12**(2): 449-60.
- Gubb, D., J. Roote, J. Trenear, D. Coulson y M. Ashburner (1997). "Topological constraints on transvection between *white* genes within the transposing element TE35B in *Drosophila melanogaster*." Genetics **146**(3): 919-37.
- Han, J. S., S. T. Szak y J. D. Boeke (2004). "Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes." Nature **429**(6989): 268-74.
- Hantusch, B., R. Kalt, S. Krieger, C. Puri y D. Kerjaschki (2007). "Sp1/Sp3 and DNA-methylation contribute to basal transcriptional activation of human podoplanin in MG63 versus Saos-2 osteoblastic cells." BMC Mol Biol **8**(1): 20.
- Harden, N. y M. Ashburner (1990). "Characterization of the *FB-NOF* transposable element of *Drosophila melanogaster*." Genetics **126**(2): 387-400.
- Hauck, B., W. J. Gehring y U. Walldorf (1999). "Functional analysis of an eye specific enhancer of the *eyeless* gene in *Drosophila*." Proc Natl Acad Sci U S A **96**(2): 564-9.

- Havas, K., A. Flaus, M. Phelan, R. Kingston, P. A. Wade, D. M. Lilley y T. Owen-Hughes (2000). "Generation of superhelical torsion by ATP-dependent chromatin remodeling activities." Cell **103**(7): 1133-42.
- Hazelrigg, T. (1987). "The *Drosophila white* gene: a molecular update." Trends in Genetics **3**(2): 43-47.
- Hazelrigg, T., R. Levis y G. M. Rubin (1984). "Transformation of *white* locus DNA in *drosophila*: dosage compensation, *zeste* interaction, and position effects." Cell **36**(2): 469-81.
- Heberlein, U., T. Wolff y G. M. Rubin (1993). "The TGF beta homolog *dpp* and the segment polarity gene *hedgehog* are required for propagation of a morphogenetic wave in the *Drosophila* retina." Cell **75**(5): 913-26.
- Higuchi, R., C. Fockler, G. Dollinger y R. Watson (1993). "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions." Biotechnology (N Y) **11**(9): 1026-30.
- Hirokawa, T., S. Boon-Chieng y S. Mitaku (1998). "SOSUI: classification and secondary structure prediction system for membrane proteins." Bioinformatics **14**(4): 378-9.
- Huang, D. H. y Y. L. Chang (2004). "Isolation and characterization of CHRASCH, a polycomb-containing silencing complex." Methods Enzymol **377**: 267-82.
- Izsvak, Z., Z. Ivics y P. B. Hackett (1995). "Characterization of a *Tc1*-like transposable element in zebrafish (*Danio rerio*)." Mol Gen Genet **247**(3): 312-22.
- Jack, J. W. y B. H. Judd (1979). "Allelic pairing and gene regulation: A model for the *zeste-white* interaction in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **76**(3): 1368-1372.
- Jeltsch, A., W. Nellen y F. Lyko (2006). "Two substrates are better than one: dual specificities for Dnmt2 methyltransferases." Trends Biochem Sci **31**(6): 306-8.
- Jenuwein, T. y C. D. Allis (2001). "Translating the histone code." Science **293**(5532): 1074-80.
- Jones, P. A. y S. M. Taylor (1980). "Cellular differentiation, cytidine analogs and DNA methylation." Cell **20**(1): 85-93.
- Jost, J. P. y Y. C. Jost (1994). "Transient DNA demethylation in differentiating mouse myoblasts correlates with higher activity of 5-methyldeoxycytidine excision repair." J Biol Chem **269**(13): 10040-3.

- Judd, B. H. (1995). "Mutations of *zeste* that mediate transvection are recessive enhancers of position-effect variegation in *Drosophila melanogaster*." Genetics **141**(1): 245-53.
- Kal, A. J., T. Mahmoudi, N. B. Zak y C. P. Verrijzer (2000). "The *Drosophila* Brahma complex is an essential coactivator for the *trithorax* group protein *zeste*." Genes Dev **14**(9): 1058-71.
- Kapitonov, V. V. y J. Jurka (2001). "Rolling-circle transposons in eukaryotes." Proc Natl Acad Sci U S A **98**(15): 8714-9.
- Kikuchi, K., K. Terauchi, M. Wada y H. Y. Hirano (2003). "The plant MITE *mPing* is mobilized in anther culture." Nature **421**(6919): 167-70.
- Kim, D. H., L. M. Villeneuve, K. V. Morris y J. J. Rossi (2006). "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells." Nat Struct Mol Biol **13**(9): 793-7.
- Kingston, R. E., C. A. Bunker y A. N. Imbalzano (1996). "Repression and activation by multiprotein complexes that alter chromatin structure." Genes Dev **10**(8): 905-20.
- Klose, R. J. y A. P. Bird (2006). "Genomic DNA methylation: the mark and its mediators." Trends Biochem Sci **31**(2): 89-97.
- Kouidou, S., T. Agidou, A. Kyrkou, A. Andreou, T. Katopodi, E. Georgiou, D. Krikelis, A. Dimitriadou, P. Spanos, C. Tsilikas, H. Destouni y G. Tzimagiorgis (2005). "Non-CpG cytosine methylation of *p53* exon 5 in non-small cell lung carcinoma." Lung Cancer **50**(3): 299-307.
- Koyanagi, K. O., M. Hagiwara, T. Itoh, T. Gojobori y T. Imanishi (2005). "Comparative genomics of bidirectional gene pairs and its implications for the evolution of a transcriptional regulation system." Gene **353**(2): 169-76.
- Kress, C., H. Thomassin y T. Grange (2001). "Local DNA demethylation in vertebrates: how could it be performed and targeted?" FEBS Lett **494**(3): 135-40.
- Krogh, A., B. Larsson, G. von Heijne y E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." J Mol Biol **305**(3): 567-80.
- Kunert, N., J. Marhold, J. Stanke, D. Stach y F. Lyko (2003). "A Dnmt2-like protein mediates DNA methylation in *Drosophila*." Development **130**(21): 5083-90.
- Kurdistani, S. K. y M. Grunstein (2003). "Histone acetylation and deacetylation in yeast." Nat Rev Mol Cell Biol **4**(4): 276-84.

- Labrador, M., F. Mongelard, P. Plata-Rengifo, E. M. Baxter, V. G. Corces y T. I. Gerasimova (2001). "Protein encoding by both DNA strands." Nature **409**(6823): 1000.
- Landschulz, W. H., P. F. Johnson y S. L. McKnight (1988). "The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins." Science **240**(4860): 1759-64.
- Lei, H., S. P. Oh, M. Okano, R. Juttermann, K. A. Goss, R. Jaenisch y E. Li (1996). "*De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells." Development **122**(10): 3195-205.
- Leonhardt, H. y M. C. Cardoso (2000). "DNA methylation, nuclear structure, gene expression and cancer." J Cell Biochem Suppl **Suppl 35**: 78-83.
- Levis, R., P. M. Bingham y G. M. Rubin (1982). "Physical map of the *white* locus of *Drosophila melanogaster*." Proc Natl Acad Sci U S A **79**(2): 564-8.
- Levis, R., P. Dunsmuir y G. M. Rubin (1980). "Terminal repeats of the *Drosophila* transposable element *copia*: nucleotide sequence and genomic organization." Cell **21**(2): 581-8.
- Levis, R., T. Hazelrigg y G. M. Rubin (1985). "Separable cis-acting control elements for expression of the *white* gene of *Drosophila*." EMBO J **4**(13A): 3489-99.
- Levis, R., K. O'Hare y G. M. Rubin (1984). "Effects of transposable element insertions on RNA encoded by the *white* gene of *Drosophila*." Cell **38**(2): 471-81.
- Liebermann, D., B. Hoffman-Liebermann, A. B. Troutt, L. Kedes y S. N. Cohen (1986). "Sequences from sea urchin *TU* transposons are conserved among multiple eucaryotic species, including humans." Mol Cell Biol **6**(1): 218-26.
- Lindsley, D. L. y G. G. Zimm (1992). The genome of *Drosophila melanogaster*, Academic Press, Inc.
- Lippman, Z., A. V. Gendrel, M. Black, M. W. Vaughn, N. Dedhia, W. R. McCombie, K. Lavine, V. Mittal, B. May, K. D. Kasschau, J. C. Carrington, R. W. Doerge, V. Colot y R. Martienssen (2004). "Role of transposable elements in heterochromatin and epigenetic control." Nature **430**(6998): 471-6.
- Liu, L., R. C. Wylie, L. G. Andrews y T. O. Tollefsbol (2003). "Aging, cancer and nutrition: the DNA methylation connection." Mech Ageing Dev **124**(10-12): 989-98.

- Lorincz, M. C., D. R. Dickerson, M. Schmitt y M. Groudine (2004). "Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells." Nat Struct Mol Biol **11**(11): 1068-75.
- Lusser, A. y J. T. Kadonaga (2003). "Chromatin remodeling by ATP-dependent molecular machines." Bioessays **25**(12): 1192-200.
- Mansukhani, A., P. H. Gunaratne, P. W. Sherwood, B. J. Sneath y M. L. Goldberg (1988). "Nucleotide sequence and structural analysis of the *zeste* locus of *Drosophila melanogaster*." Mol Gen Genet **211**(1): 121-8.
- Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWeese-Scott, L. Y. Geer, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. Zhang y S. H. Bryant (2005). "CDD: a Conserved Domain Database for protein classification." Nucleic Acids Res **33**(Database issue): D192-6.
- Marhold, J., K. Kramer, E. Kremmer y F. Lyko (2004). "The *Drosophila* MBD2/3 protein mediates interactions between the MI-2 chromatin complex and CpT/A-methylated DNA." Development **131**(24): 6033-9.
- Mayer, W., A. Niveleau, J. Walter, R. Fundele y T. Haaf (2000). "Demethylation of the zygotic paternal genome." Nature **403**(6769): 501-2.
- McClintock, B. (1956). Controlling elements and the gene.
- Moller, A., F. W. Avila, J. W. Erickson y H. Jackle (2005). "*Drosophila* BAP60 is an essential component of the Brahma complex, required for gene activation and repression." J Mol Biol **352**(2): 329-37.
- Morgan, H. D., H. G. Sutherland, D. I. Martin y E. Whitelaw (1999). "Epigenetic inheritance at the *agouti* locus in the mouse." Nat Genet **23**(3): 314-8.
- Morrison, T. B., J. J. Weis y C. T. Wittwer (1998). "Quantification of low-copy transcripts by continuous *SYBR Green I* monitoring during amplification." Biotechniques **24**(6): 954-8, 960, 962.
- Moschetti, R., R. M. Marsano, P. Barsanti, C. Caggese y R. Caizzi (2004). "*FB* elements can promote exon shuffling: a promoter-less *white* allele can be reactivated by *FB* mediated transposition in *Drosophila melanogaster*." Mol Genet Genomics **271**(4): 394-401.
- Muchardt, C. y M. Yaniv (1999). "ATP-dependent chromatin remodelling: SWI/SNF and Co. are on the job." J Mol Biol **293**(2): 187-98.

- Mulholland, N. M., I. F. King y R. E. Kingston (2003). "Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA." Genes Dev **17**(22): 2741-6.
- Muotri, A. R., V. T. Chu, M. C. Marchetto, W. Deng, J. V. Moran y F. H. Gage (2005). "Somatic mosaicism in neuronal precursor cells mediated by *L1* retrotransposition." Nature **435**(7044): 903-10.
- Muyrers-Chen, I., T. Rozovskaia, N. Lee, J. H. Kersey, T. Nakamura, E. Canaani y R. Paro (2004). "Expression of leukemic MLL fusion proteins in *Drosophila* affects cell cycle control and chromosome morphology." Oncogene **23**(53): 8639-48.
- Neufeld, T. P., A. F. de la Cruz, L. A. Johnston y B. A. Edgar (1998). "Coordination of growth and cell division in the *Drosophila* wing." Cell **93**(7): 1183-93.
- Nichols, C. G. y A. N. Lopatin (1997). "Inward rectifier potassium channels." Annu Rev Physiol **59**: 171-91.
- Nowak, S. J. y V. G. Corces (2004). "Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation." Trends Genet **20**(4): 214-20.
- Nuzhdin, S. V. (1999). "Sure facts, speculations, and open questions about the evolution of transposable element copy number." Genetica **107**(1-3): 129-37.
- O'Hare, K., R. Levis y G. M. Rubin (1983). "Transcription of the *white* locus in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **80**(22): 6917-6921.
- O'Hare, K., C. Murphy, R. Levis y G. M. Rubin (1984). "DNA sequence of the *white* locus of *Drosophila melanogaster*." J Mol Biol **180**(3): 437-55.
- Ohno, S. (1972). "So much "junk" DNA in our genome." Brookhaven Symp Biol **23**: 366-70.
- Ohsako, T., T. Horiuchi, T. Matsuo, S. Komaya y T. Aigaki (2003). "*Drosophila lola* encodes a family of BTB-transcription regulators with highly variable C-terminal domains containing zinc finger motifs." Gene **311**: 59-69.
- Ono, T., N. Takahashi y S. Okada (1989). "Age-associated changes in DNA methylation and mRNA level of the *c-myc* gene in spleen and liver of mice." Mutat Res **219**(1): 39-50.
- Padgett, R. W., R. D. St Johnston y W. M. Gelbart (1987). "A transcript from a *Drosophila* pattern gene predicts a protein homologous to the transforming growth factor-beta family." Nature **325**(6099): 81-4.

- Paro, R., M. L. Goldberg y W. J. Gehring (1983). "Molecular analysis of large transposable elements carrying the *white* locus of *Drosophila melanogaster*." EMBO J **2**(6): 853-860.
- Paulson, K. E., N. Deka, C. W. Schmid, R. Misra, C. W. Schindler, M. G. Rush, L. Kadyk y L. Leinwand (1985). "A transposon-like element in human DNA." Nature **316**(6026): 359-61.
- Peaston, A. E., A. V. Evsikov, J. H. Graber, W. N. de Vries, A. E. Holbrook, D. Solter y B. B. Knowles (2004). "Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos." Dev Cell **7**(4): 597-606.
- Peterson, C. L. y M. A. Laniel (2004). "Histones and histone modifications." Curr Biol **14**(14): R546-51.
- Piñol, J., O. Francino, A. Fontdevila y O. Cabré (1988). "Rapid isolation of *Drosophila* high molecular weight DNA to obtain genomic libraries." Nucleic Acids Res **16**(6): 2736.
- Pirrotta, V. (1991). "The genetics and molecular biology of *zeste* in *Drosophila melanogaster*." Adv Genet **29**: 301-48.
- Pirrotta, V. y C. Brockl (1984). "Transcription of the *Drosophila white* locus and some of its mutants." EMBO J **3**(3): 563-8.
- Pirrotta, V. y L. Rastelli (1994). "*white* gene expression, repressive chromatin domains and homeotic gene regulation in *Drosophila*." Bioessays **16**(8): 549-56.
- Portela, A. (2003). "Anàlisi per PCR a temps real de l'efecte d'un element *FB-NOF* sobre la interacció *zeste-white* i avaluació experimental de mètodes de transgènia per aprofundir aquest estudi." Trabajo de Máster. Universitat Autònoma de Barcelona.
- Potter, S., M. Truett, M. Phillips y A. Maher (1980). "Eucaryotic transposable genetic elements with inverted terminal repeats." Cell **20**(3): 639-47.
- Potter, S. S. (1982). "DNA sequence analysis of a *Drosophila foldback* transposable element rearrangement." Mol Gen Genet **188**(1): 107-10.
- Prigent, C. y S. Dimitrov (2003). "Phosphorylation of serine 10 in histone H3, what for?" J Cell Sci **116**(Pt 18): 3677-85.
- Qian, S. y V. Pirrotta (1995). "Dosage compensation of the *Drosophila white* gene requires both the X chromosome environment and multiple intragenic elements." Genetics **139**(2): 733-44.

- Rasmuson-Lestander, A. y K. Ekstrom (1996). "Genetic and molecular analysis of a set of unstable *white* mutants in *Drosophila melanogaster*." Genetica **98**(2): 179-92.
- Rasmuson-Lestander, A., J. Larsson y B. Rasmuson (1993). "Position-effect variegation and z^1 mediated *white* repression in the $In(1)w^{is}$ system in *Drosophila melanogaster*." Hereditas **119**(3): 209-18.
- Rasmussen, R. (2001). Quantification on the LightCycler, Springer.
- Rea, S., F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis y T. Jenuwein (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-9.
- Reese, M. G. (2001). "Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome." Comput Chem **26**(1): 51-6.
- Retelska, D., C. Iseli, P. Bucher, C. V. Jongeneel y F. Naef (2006). "Similarities and differences of polyadenylation signals in human and fly." BMC Genomics **7**: 176.
- Rice, P., I. Longden y A. Bleasby (2000). "EMBOSS: the European Molecular Biology Open Software Suite." Trends Genet **16**(6): 276-7.
- Rizzon, C., G. Marais, M. Gouy y C. Biémont (2002). "Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome." Genome Res **12**(3): 400-7.
- Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones y A. P. Wolffe (2000a). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet **25**(3): 338-42.
- Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones y A. P. Wolffe (2000b). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet **25**(3): 338-42.
- Robertson, K. D. y A. P. Wolffe (2000b). "DNA methylation in health and disease." Nat Rev Genet **1**(1): 11-9.
- Roche Applied Science (2002). "Selection of housekeeping genes." Technical Note No. LC 15/2002.
- Rorth, P., K. Szabo, A. Bailey, T. Laverty, J. Rehm, G. M. Rubin, K. Weigmann, M. Milan, V. Benes, W. Ansorge y S. M. Cohen (1998). "Systematic gain-of-function genetics in *Drosophila*." Development **125**(6): 1049-57.

- Rosen, C., D. Dorsett y J. Jack (1998). "A proline-rich region in the Zeste protein essential for transvection and *white* repression by Zeste." Genetics **148**(4): 1865-74.
- Sambrook, J., O. Fisher y T. Maniatis (1989). Molecular Cloning.
- Schnitzler, G., S. Sif y R. E. Kingston (1998). "Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state." Cell **94**(1): 17-27.
- Schuettengruber, B., D. Chourrout, M. Vervoort, B. Leblanc y G. Cavalli (2007). "Genome regulation by polycomb and trithorax proteins." Cell **128**(4): 735-45.
- Shearn, A. (1989). "The *ash-1*, *ash-2* and trithorax genes of *Drosophila melanogaster* are functionally related." Genetics **121**(3): 517-25.
- Shi, Y. y J. R. Whetstine (2007). "Dynamic regulation of histone lysine methylation by demethylases." Mol Cell **25**(1): 1-14.
- Smolik-Utlaut, S. M. y W. M. Gelbart (1987). "The effects of chromosomal rearrangements on the *zeste-white* interaction in *Drosophila melanogaster*." Genetics **116**(2): 285-98.
- Song, S. U., T. Gerasimova, M. Kurkulos, J. D. Boeke y V. G. Corces (1994). "An *env*-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus." Genes Dev **8**(17): 2046-57.
- Spencer, F. A., F. M. Hoffmann y W. M. Gelbart (1982). "*Decapentaplegic*: a gene complex affecting morphogenesis in *Drosophila melanogaster*." Cell **28**(3): 451-61.
- Sterner, D. E. y S. L. Berger (2000). "Acetylation of histones and transcription-related factors." Microbiol Mol Biol Rev **64**(2): 435-59.
- Stewart, C. L., H. Stuhlmann, D. Jahner y R. Jaenisch (1982). "*De novo* methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells." Proc Natl Acad Sci U S A **79**(13): 4098-102.
- Strahl, B. D. y C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.
- Straub, T., I. K. Dahlsveen y P. B. Becker (2005). "Dosage compensation in flies: mechanism, models, mystery." FEBS Lett **579**(15): 3258-63.
- Streck, R. D., J. E. Macgaffey y S. K. Beckendorf (1986). "The structure of *hobo* transposable elements and their insertion sites." EMBO J **5**(13): 3615-3623.

- Tamkun, J. W., R. Deuring, M. P. Scott, M. Kissinger, A. M. Pattatucci, T. C. Kaufman y J. A. Kennison (1992). "*brahma*: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2." Cell **68**(3): 561-72.
- Tartof, K. D., C. Bishop, M. Jones, C. A. Hobbs y J. Locke (1989). "Towards an understanding of position effect variegation." Dev Genet **10**(3): 162-76.
- Templeton, N. S. y S. S. Potter (1989). "Complete *foldback* transposable elements encode a novel protein found in *Drosophila melanogaster*." EMBO J **8**(6): 1887-94.
- Thellin, O., W. Zorzi, B. Lakaye, B. De Borman, B. Coumans, G. Hennen, T. Grisar, A. Igout y E. Heinen (1999). "Housekeeping genes as internal standards: use and limits." J Biotechnol **75**(2-3): 291-5.
- Thompson, J. D., D. G. Higgins y T. J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-80.
- Toth, M., U. Muller y W. Doerfler (1990). "Establishment of *de novo* DNA methylation patterns. Transcription factor binding and deoxycytidine methylation at CpG and non-CpG sequences in an integrated adenovirus promoter." J Mol Biol **214**(3): 673-83.
- Truett, M. A., R. S. Jones y S. S. Potter (1981). "Unusual structure of the *FB* family of transposable elements in *Drosophila*." Cell **24**(3): 753-63.
- Tschudi, C., V. Pirrotta y N. Junakovic (1982). "Rearrangements of the 5S RNA gene cluster of *Drosophila melanogaster* associated with the insertion of a B104 element." EMBO J **1**(8): 977-985.
- Tseng, A. S. y I. K. Hariharan (2002). "An overexpression screen in *Drosophila* for genes that restrict growth or cell-cycle progression in the developing eye." Genetics **162**(1): 229-43.
- Tufarelli, C., J. A. Stanley, D. Garrick, J. A. Sharpe, H. Ayyub, W. G. Wood y D. R. Higgs (2003). "Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease." Nat Genet **34**(2): 157-65.
- Tusnady, G. E. y I. Simon (2001). "The HMMTOP transmembrane topology prediction server." Bioinformatics **17**(9): 849-50.

- van de Lagemaat, L. N., J. R. Landry, D. L. Mager y P. Medstrand (2003). "Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions." Trends Genet **19**(10): 530-6.
- Vázquez, M., L. Moore y J. A. Kennison (1999). "The *trithorax* group gene *osa* encodes an ARID-domain protein that genetically interacts with the brahma chromatin-remodeling factor to regulate transcription." Development **126**(4): 733-42.
- Wallecha, A., V. Munster, J. Correnti, T. Chan y M. van der Woude (2002). "Dam- and OxyR-dependent phase variation of *agn43*: essential elements and evidence for a new role of DNA methylation." J Bacteriol **184**(12): 3338-47.
- Wallrath, L. L. (1998). "Unfolding the mysteries of heterochromatin." Curr Opin Genet Dev **8**(2): 147-53.
- Windsor, A. J. y C. S. Waddell (2000). "*FARE*, a new family of foldback transposons in *Arabidopsis*." Genetics **156**(4): 1983-95.
- Wu, C., Y. C. Wong y S. C. Elgin (1979). "The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity." Cell **16**(4): 807-14.
- Wu, C. T., R. S. Jones, P. F. Lasko y W. M. Gelbart (1989). "Homeosis and the interaction of *zeste* and *white* in *Drosophila*." Mol Gen Genet **218**(3): 559-64.
- Wu, J. C. y D. V. Santi (1987). "Kinetic and catalytic mechanism of *Hha*I methyltransferase." J Biol Chem **262**(10): 4778-86.
- Wu, X. y S. M. Burgess (2004). "Integration target site selection for retroviruses and transposable elements." Cell Mol Life Sci **61**(19-20): 2588-96.
- Yan, Q., J. Huang, T. Fan, H. Zhu y K. Muegge (2003). "*Lsh*, a modulator of CpG methylation, is crucial for normal histone methylation." EMBO J **22**(19): 5154-62.
- Zachar, Z. y P. M. Bingham (1982). "Regulation of *white* locus expression: the structure of mutant alleles at the *white* locus of *Drosophila melanogaster*." Cell **30**(2): 529-41.