

DEPARTAMENTO DE GENÉTICA

MUSCLEBLIND, RELEVANCIA CLÍNICA Y ANÁLISIS DE
LA FUNCIÓN MOLECULAR EN DROSOPHILA
MELANOGASTER

MAYA PASCUAL LUCAS

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2005

Aquesta Tesi Doctoral va ser presentada a València el dia 10 de Novembre de 2005 davant un tribunal format per:

- D. José Viña Ribes
- D. Jaime Font de Mora Sainz
- D. Patrick Morcillo
- D. Francesc Palau Martínez
- D. Sergi Puig Todolí

Va ser dirigida per:

D. Rubén D. Artero Allepuz

D. Manuel Pérez Alonso

©Copyright: Servei de Publicacions
Maya Pascual Lucas

Depòsit legal:

I.S.B.N.:84-370-6376-0

Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115



VNIVERSITAT Æ VALÈNCIA

***MUSCLEBLIND, RELEVANCIA CLÍNICA Y
ANÁLISIS DE LA FUNCIÓN MOLECULAR EN
DROSOPHILA MELANOGASTER***

MEMORIA PRESENTADA POR Dña MAYA PASCUAL LUCAS
PARA OPTAR AL TÍTULO DE DOCTORA EN CIENCIAS
BIOLÓGICAS POR LA UNIVERSIDAD DE VALENCIA

VALENCIA, JULIO DE 2005

Dr. D. RUBÉN ARTERO ALLEPUZ, Investigador del Programa Ramón y Cajal del Ministerio de Educación y Ciencia y el Dr. D. MANUEL PÉREZ ALONSO, Profesor titular del Departamento de Genética de la Facultad de Ciencias Biológicas de la Universidad de Valencia

CERFIFICAN: que Dña. MAYA PASCUAL LUCAS, licenciada en Ciencias Biológicas, ha realizado bajo nuestra dirección el trabajo de investigación recogido en esta memoria bajo el título “*muscleblind*, relevancia clínica y análisis de la función molecular en *Drosophila melanogaster*”, para optar al grado de Doctora en Ciencias Biológicas y

AUTORIZAN: su presentación en el Departamento de Genética de la Facultad de Ciencias Biológicas para tal fin.

Para que conste, y en cumplimiento de la legislación vigente, firman el presente certificado en Burjasot a 20 de julio de 2005.

Fdo. Dr. D. Rubén Artero Allepuz

Fdo. Dr. D. Manuel Pérez Alonso

A MI MADRE

A JUAN

"Lo que no me mata me hace más fuerte"

(Nietzsche)

Quisiera dar las gracias a todas las personas que de una forma u otra han contribuido a que la realización de esta tesis doctoral fuera posible:

A la Conselleria de Cultura por concederme una beca predoctoral que ha permitido desarrollar este trabajo. Espero que pronto lo entiendan como tal y conviertan sus becas en contratos.

A mis directores de tesis, Manuel y Rubén. A Manuel por darme la oportunidad de iniciarme en la carrera investigadora y por el tiempo dedicado durante los primeros años de esta tesis, quizás los más duros para un director. A Rubén por acogerme en el seno de su proyecto, por las interesantes conversaciones científicas que hemos tenido y sobre todo, por su esfuerzo en las últimas etapas de este trabajo, más allá de las obligaciones de un director, para que finalmente viera la luz.

A mis compañer@s de laboratorio, Lidón, Marta, Silvia, Verónica, y Use con las que a lo largo de todos estos años he compartido momentos divertidos y no tan divertidos. Especial mención a Lidón por la empatía que nos une desde su llegada, por las interesantes conversaciones de pasillo que hemos tenido y por su apoyo.

A los alumn@s colaboradores que me han echado una mano en los momentos de agobio, especialmente a JuanMa, Yaiza, Amparo y Fabrice. A todos ellos les deseo lo mejor para las tesis que van a comenzar.

A todas las compañer@ de los distintos laboratorios del departamento: los moleculares, los bioquímicos y los genómicos, por atender a mis preguntas y echarme una mano siempre que los he necesitado.

Al Departamento de Genética del Hospital de La Fe, especialmente a Chema, Paqui y Silves por ayudarme en algunos aspectos de este trabajo.

Al personal de administración y servicios del departamento, Carmen, Fede, Jose Carlos y Begoña, imprescindibles para que podamos realizar un buen trabajo.

Por último y muy especialmente, a mi madre y hermanas que me han apoyado y animado en los momentos más difíciles, y por supuesto, a Juan, por su apoyo incondicional durante todos estos años, por creer en mí más que yo misma y por su ayuda y su paciencia en los durísimos últimos seis meses que preceden a estas palabras.

ÍNDICE

ABEVIATURAS	15
INTRODUCCIÓN	17
1. El gen <i>muscleblind</i> en <i>Drosophila melanogaster</i>	19
1.1. <i>muscleblind</i> es un gen esencial para la diferenciación terminal de la musculatura	21
1.2. <i>muscleblind</i> participa en la diferenciación de tipos celulares neuronales. ...	23
1.3. Estructura genómica y procesado alternativo de <i>muscleblind</i>	24
2. Función molecular de las proteínas Muscleblind	27
2.1. Regulación del procesado alternativo.....	27
2.2. Localización de proteínas dependiente del mRNA.....	29
3. La Distrofia Miotónica	31
3.1. DM1 y DM2	32
3.2. Sintomatología.....	35
3.3. Base genética de la enfermedad.....	37
3.4. Mecanismo de patogenicidad.....	38
3.5. Defectos Moleculares en la DM	41
3.5.1. Defectos en el patrón de procesado alternativo	42
3.5.2. El papel de Muscleblind en la ruta de patogénesis de la DM	46
4. <i>Drosophila</i> como modelo.....	48
4.1. Modelos del efecto tóxico de las repeticiones.....	49
OBJETIVOS.....	53
MATERIAL Y MÉTODOS.....	56
1. MATERIAL.....	57
1.1. Pacientes.....	57
1.2. <i>Drosophila melanogaster</i>	58
1.3. <i>Escherichia coli</i>	58
1.4. Anticuerpos.....	58
1.5. Enzimas.....	59
1.6. Reactivos.....	59
1.7. kits.....	59
1.8. Cebadores.....	60
1.9. Equipos.....	62
1.10. Miscelánea.....	62
1.11. Soluciones.....	62

2. MÉTODOS.....	63
2.1. Elaboración lista patologías candidatas MBNL3.....	63
2.2. Extracción de RNA y RT-PCR.....	64
2.3. Secuenciación de DNA.....	64
2.4. Rastreo fenotipo sobreexpresión <i>mbIC</i> en ojo.....	65
2.5. Clonación de MbIC:GFP.....	65
2.6. Otención de Moscas transgénicas.....	66
2.7. Rescate de la letalidad embrionaria.....	67
2.8. Obtención de la proteína MbIA:GST.....	68
2.9. Producción del antipéptido α -Muscleblind.....	68
2.10. Western blotting.....	69
2.11. Obtención material para IP.....	70
2.12. Preparación de los extractos para IP.....	71
2.13. Unión del Anticuerpo a la proteínaG-agarosa.....	71
2.14. Preabsorción de los extractos.....	72
2.15. Coimmunoprecipitación de proteínas.....	73
RESULTADOS.....	74
1. ANÁLISIS ESTRUCTURAL Y RELEVANCIA CLÍNICA DE LOS GENES	
HUMANOS.....	75
1.1. Muscleblind, una nueva familia de genes humanos.....	75
1.1.1. Existen tres genes humanos: MBNL1, MBNL2 y MBNL3.....	75
1.1.2. La estructura genómica esta altamente conservada.....	77
1.1.3. Los tres genes sufren un complejo patrón de procesado alternativo....	82
1.1.4. Expresión diferencial de los genes <i>muscleblind</i> humanos.....	87
1.1.5. Características de la familia de proteínas Muscleblind.....	90
1.2. Relevancia clínica de los genes <i>muscleblind</i> humanos.....	97
1.2.1.MBNL1, búsqueda de mutaciones en pacientes con USHIII.....	97
1.2.2.MBNL1, búsqueda de mutaciones en pacientes con DM.....	98
1.2.3.MBNL3, búsqueda de mutaciones en pacientes con Ptosis congénita..	105
2. ANÁLISIS DE LA FUNCIÓN MOLECULAR EN <i>Drosophila melanogaster</i>.....	108
2.1. Aproximación genética para la identificación de dianas moleculares.....	109
2.1.1. Rastreo de genes que interaccionan con <i>muscleblind</i>	112
2.1.1.1. Análisis de los genes de las deficiencias Df(2L)TW137 y	
Df(2L)sc19-8.....	114
2.1.1.2. Análisis de genes candidatos de todas las deleciones.....	116

2.1.2. Análisis de la interacción con <i>Jumeaux</i>	125
2.2. Aproximación bioquímica para la identificación de dianas moleculares....	135
2.2.1. Obtención de un anticuerpo α -Mbl.....	135
2.2.2. Obtención de moscas transgénicas.....	137
2.2.3. La proteína MbIC:GFP es funcional.....	138
2.2.4. Coinmunoprecipitación de proteínas.....	140
2.2.5. Verificación de candidatos mediante coinmunoprecipitación.....	145
2.2.6. Apéndices.....	148
2.2.6.1. Sistema UAS-GAL4.....	148
2.2.6.2. Esquema experimento colP.....	149
2.2.6.3. Tampones de homogenización.....	150
DISCUSIÓN.....	151
1. ANÁLISIS ESTRUCTURAL Y RELEVANCIA CLÍNICA DE LOS GENES	
HUMANOS.....	152
1.1. Estructura y procesado de los genes <i>muscleblind</i> humanos.....	152
1.1.1. Presencia de elementos repetitivos en MBNL1.....	152
1.1.2. Conservación en <i>Drosophila</i>	153
1.2. Expresión de los genes <i>muscleblind</i> humanos.....	153
1.3. Los genes <i>muscleblind</i> pueden tener funciones distintas.....	156
1.4. Historia evolutiva de los genes <i>muscleblind</i>	157
1.5. Efecto del polimorfismo poli(dT) sobre MBNL1.....	158
2. ANÁLISIS DE LA FUNCIÓN MOLECULAR EN <i>Drosophila melanogaster</i>.....	162
2.1. <i>Muscleblind</i> está implicado en distintas funciones celulares.....	162
2.1.1. <i>Muscleblind</i> y el metabolismo del RNA.....	162
2.1.2. <i>Muscleblind</i> y la regulación de la Transcripción.....	167
2.1.3. <i>Muscleblind</i> y la Apoptosis.....	168
2.2. Las moscas MbIC:GFP son una herramienta útil para colP.....	172
CONCLUSIONES.....	174
BIBLIOGRAFÍA.....	178

A continuación se detallan las abreviaturas y anglicismos usados a los largo de este trabajo.

OrR	Oregon R
PM	peso molecular
IP	inmunoprecipitación
coIP	co-inmunoprecipitación
cDNA	ADN complementario
<i>mbl</i>	<i>muscleblind</i>
GFP	Green Fluorescent protein
<i>mbl</i> ^{E27}	alelo letal embrionario del gen <i>mbl</i> generado por la escisión imprecisa del elemento P en la línea <i>mm</i> ^{k05507b}
DMP	dimethyl pimelimidate
RT	Retrotranscripción
PCR	Reacción en cadena de la polimerasa
ml	mililitro
min	minuto
s	segundo
h	hora
c/u	cada uno
BSA	Bovine Serum Albumine
GB	número de acceso de la base de datos GenBank
UTR	regiones no traducidas de los mRNAs
<i>splicing</i>	corte y empalme de los transcritos primarios
WGS	<i>Whole Genome Shotgun</i>
PCP	<i>Planar Cell Polarity</i>
<i>beads</i>	partículas de agarosa, en este trabajo proteína G agarosa
NMD	Non Mediated Decay

INTRODUCCIÓN

1. El gen *muscleblind* de *Drosophila melanogaster*

El gen *muscleblind* (*mb1*) fue identificado originariamente en *Drosophila melanogaster*. La caracterización del gen se llevó a cabo en colaboración entre el laboratorio del Prof. Manuel Pérez-Alonso, de la Universidad de Valencia (UEV) y el laboratorio del Dr. Marek Mlodzik del European Molecular Biology Laboratory (EMBL) de Heidelberg (Alemania) entre los años 1993 y 1998.

En el laboratorio del Dr. Mlodzik, *mb1* fue aislado durante un rastreo de cepas con inserciones de elementos P letales, como un supresor dominante del fenotipo de ganancia de función de *sev-svp2*, esto es, de la expresión ectópica de la isoforma 2 del gen *seven-up* (*svp*) en las células cono, bajo el control del *enhancer* *sevenless* (*sev*) (Begemann, 1997). Esta sobreexpresión produce un fenotipo que consiste, principalmente, en la transformación de estas células en fotorreceptores R7 (Begemann, 1995). Durante este rastreo se identificaron otros genes modificadores del fenotipo *sev-svp2* como son distintos elementos de la ruta Ras. Uno de los elementos P identificados en este rastreo, denominado I(2)k005507, se localiza en la región 54A1-3 del cromosoma 2 de *Drosophila* y, dado que se trata de una inserción de *enhancer trap*, permite detectar el patrón de expresión del gen adyacente en el genoma. Esta cepa expresa el gen reportero *lacZ* del *enhancer trap* en las células posteriores al surco morfogenético del disco imaginal de ojo. Los discos imaginales son grupos de células embrionarias organizadas en epitelios a partir de las cuales se formarán la mayoría de las estructuras de la mosca adulta. Por otro lado, en el laboratorio del Dr. Pérez-Alonso, durante la realización de un rastreo de cepas con inserciones de *enhancer trap* para la identificación de genes con una expresión interesante en el mesodermo, se identificó otra inserción en la misma región citogenética, referida como PP2, con el mismo patrón de expresión en ojo, que además se expresaba en las células precursoras de la musculatura embrionaria y en los músculos en desarrollo (Artero, 1995). Por ser un gen con una función en ojo y en músculo se le nombró *muscleblind*.

El gen *mb1* se expresa durante el desarrollo embrionario con un patrón que incluye derivados del ectodermo y el mesodermo. En el ectodermo, la expresión se localiza en el órgano de Bolwig donde se encuentran los fotorreceptores larvarios, y en el sistema nervioso central (SNC), originando un complejo patrón de expresión repetitivo y simétrico respecto al eje antero-posterior. En el mesodermo, la expresión se inicia en

doce grupos simétricos respecto de la línea media ventral compuestos por células mesodérmicas y otro grupo de células alrededor de la apertura del estomodeo todo ello alrededor del estadio 11 del desarrollo. Tras la retracción de la banda germinal, se produce un aumento de la expresión de *mb1* en el mesodermo somático, de manera que hacia el estadio 16 del desarrollo queda restringida a una serie de derivados mesodérmicos, no detectándose en las gónadas, cuerpo graso, corazón, hemocitos ni en las células de la línea ventral. Así, la expresión más intensa durante el desarrollo tardío se localiza en la musculatura faríngea, visceral y somática así como en la musculatura que sirve de sostén para el corazón (*alary muscles*) (Artero, 1995).

En larva se ha detectado expresión en el SNC, concretamente en el cerebro y ganglio ventral, y en discos imaginales de ojo-antena de larva de tercer estadio tras el paso del surco morfogenético (Artero, 1995). García-Casado (2002) hace un estudio detallado de la expresión de *mb1* en los diferentes discos imaginales usando la línea de *enhancer trap* PP2 y tinción con anticuerpo anti-Mbl. En el disco imaginal de ojo-antena (larvas de tercer estadio) la expresión en la parte de antena se corresponde con la parte más distal de la antena adulta, esto es, la arista. En el disco imaginal de ala la expresión se localiza en la región correspondiente al margen del ala y en una franja que se corresponde con el *notum* del adulto, así como también se detecta en la región que formará la articulación del ala con el cuerpo. En disco de pata *mb1* tiene un patrón de expresión en anillos concéntricos siguiendo la organización característica del disco, que corresponden a los segmentos más distales de la pata adulta. En concreto la expresión se limita a lo que serán los cinco segmentos tarsales, la tibia y, de forma más débil, también incluye el fémur. Además, en discos de pupa temprana (de 4 a 8 horas), destaca la expresión en unos grupos específicos de células en el último segmento tarsal (T5) que corresponden a la estructura de la garra en pata adulta.

García-Casado (2002) describe la expresión de *mb1* en el SNC. En embriones, *mb1* se expresa, entre otras, en el par de neuronas aCC y pCC. Las neuronas aCC son motoneuronas implicadas en la inervación de la musculatura por lo que la falta de función de *mb1* en ellas podría estar relacionado con el papel de este gen en el desarrollo muscular. Las neuronas pCC son interneuronas asociadas a las aCC que se localizan posterior y ventralmente respecto a éstas. En larvas, la expresión de *mb1* se detecta en los hemisferios cerebrales y en un complejo patrón a lo largo del ganglio ventral, no apareciendo señal en las regiones correspondientes a los lóbulos ópticos.

1.1 *muscleblind* es un gen esencial para la diferenciación terminal de la musculatura de *Drosophila*

Los mutantes *mb1* mueren en estadio 17 del desarrollo embrionario con los segmentos contraídos, especialmente en la región abdominal. En algunos casos llegan a eclosionar y mueren como larvas de primer estadio con una movilidad muy reducida que, en la mayoría de los casos, les impide realizar los movimientos con suficiente fuerza para salir del huevo. Los embriones mutantes no presentan defectos evidentes en el patrón muscular somático y visceral, aunque ocasionalmente se puede detectar la ausencia de algún músculo. Aunque la función del gen *mb1* no es esencial durante la especificación y el desarrollo temprano de la musculatura, sí que se han detectado defectos a nivel ultraestructural (Artero, 1998) que explican la parálisis y la hipercontracción de la musculatura. Así, los embriones mutantes presentan en las uniones indirectas del músculo a la epidermis una reducción de la matriz del tendón TM (Figura.1). Esta reducción provoca que los músculos tengan que competir por la superficie epidérmica implicada en el anclaje, esto es, las células tendón y, como consecuencia, podría contribuir a la parálisis y contracción de los embriones mutantes.

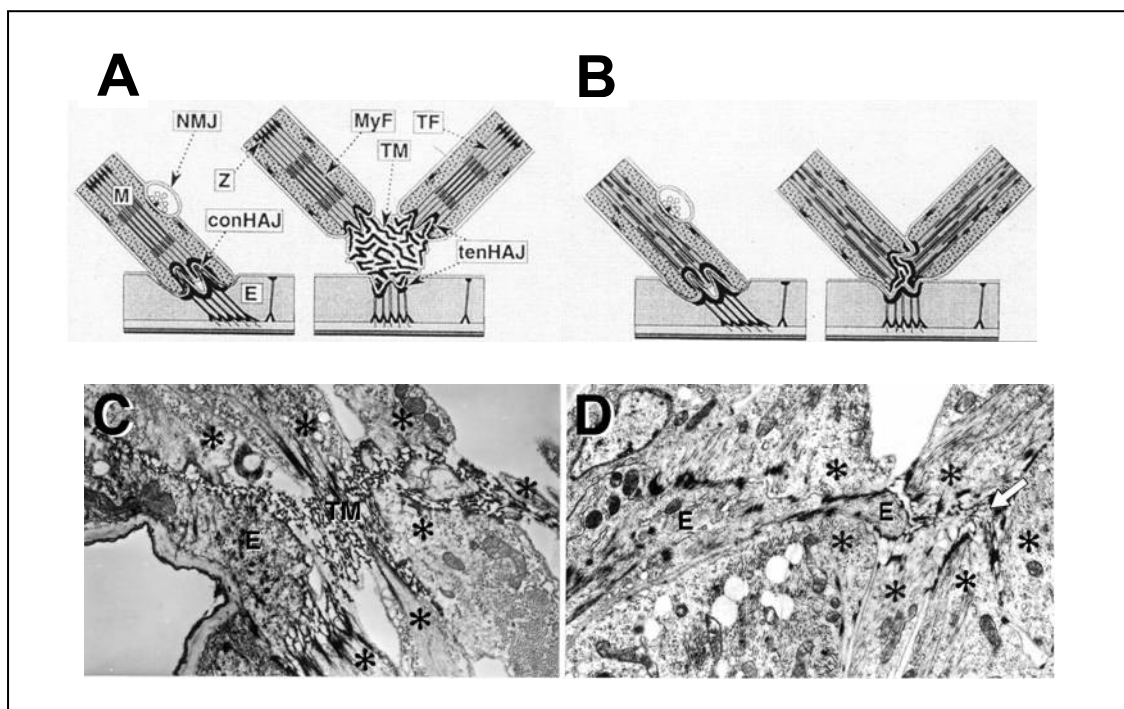


Figura 1. ver pie de figura en página siguiente

Figura 1. Alteraciones en la matriz extracelular de los mutantes *mb1*. (A) Las uniones del músculo a la epidermis pueden ser directas (representadas a la izquierda) o indirectas (derecha). En las uniones indirectas varios músculos (indicados con una M) establecen unión con una única célula epidérmica (señalada con una E) a través de una matriz extracelular (TM). También se indican las bandas Z de los sarcómeros (Z) y las uniones neuromusculares (NMJ). (B) Representación esquemática del fenotipo de embriones mutantes *mb1* en los que la matriz extracelular está prácticamente ausente. (C,D) Cortes ultrafinos de las uniones indirectas de los músculos (*) a la epidermis (E) en larvas de *Drosophila* silvestres (C) y mutantes *mb1* (D). En los mutantes *mb1* la matriz extracelular (TM) está muy reducida (flecha blanca) y las células epidérmicas están en contacto con las células musculares. Imágenes tomadas de Artero y colaboradores (1998).

Los componentes de la TM en las uniones indirectas del músculo a la epidermis se producen en los hemocitos y cuerpo graso (por ejemplo Colágeno IV, Tigrin), en las células epidérmicas (ej. Mascarade) y en los músculos (ej. M-spondin) (Fessler y Fessler, 1989; Le Parco, 1989; Fogerty, 1994; Murugasu-Oei, 1995 y Umemiya, 1997). Debido a que *mb1* sólo se expresa en la musculatura, y no en otros tipos celulares que expresan proteínas de la TM, los únicos componentes de la misma que podrían verse directamente afectados en los mutantes *mb1* serían los producidos en la musculatura somática, como la M-Spondin (Umemiya, 1997) y algunos colágenos tales como el colágeno II. De hecho, se ha comprobado que componentes no derivados de este tejido, como la proteína Tigrin, producida en los hemocitos y cuerpo graso (Fogerty, 1994) no están afectados en los embriones *mb1* (Artero, 1995; Artero 1998).

Por otro lado, el patrón típico de estriación sarcomérica también está afectado en mutantes *mb1* (Figura 2). Los filamentos de las bandas A se encuentran ordenados y empaquetados de manera menos densa que en la forma silvestre. Las bandas I se encuentran siempre ausentes, lo que es indicativo de la musculatura en estado hipercontraído. Además, las bandas Z aparecen desorganizadas, con un aspecto más difuso. Se ha comprobado que la expresión de uno de sus componentes, la proteína Kettin no está alterada en los mutantes *mb1* aunque no se organiza correctamente dentro de las bandas Z (Artero, 1998).

El fenotipo mutante de *muscleblind* es en parte similar al descrito para *Dmef2*, un factor de transcripción clave para el programa de diferenciación miogénica. Por este motivo, se ha estudiado la expresión de *mb1* en embriones mutantes nulos para *Dmef2* y en moscas que expresan dicho gen ectópicamente. Los resultados indican que en la musculatura faríngea, visceral y somática, *mb1* está bajo el control transcripcional de *Dmef2* y que debe existir otro mecanismo que regule su expresión en el resto de tejidos donde se expresa (Artero, 1998)

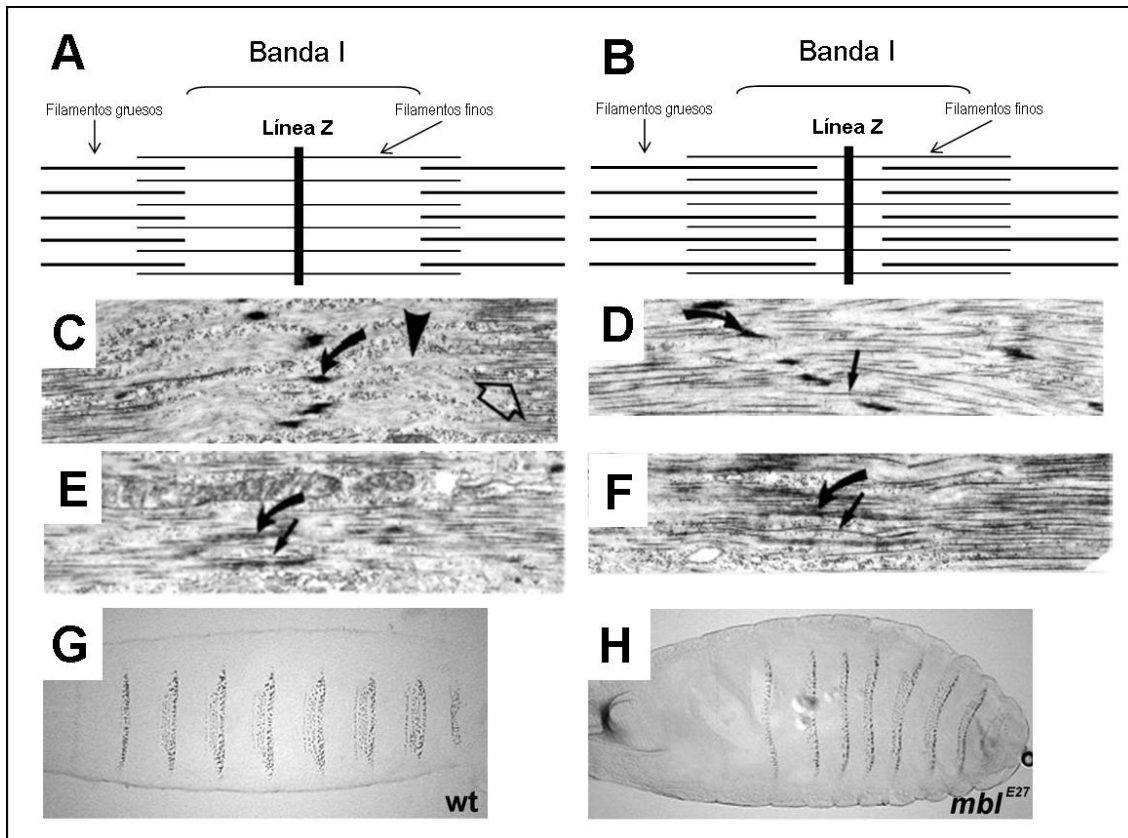


Figura 2. Alteraciones en el sarcómero de los embriones mutantes *muscleblind*. Representación esquemática de la ultraestructura del sarcómero en estado relajado (A) y contraído (B). La banda I viene dada por la región donde los filamentos finos, compuestos principalmente por actina, no solapan con los filamentos gruesos, formados principalmente por miosina. Las bandas I están biseccionadas por las bandas Z, el lugar de anclaje de los filamentos finos. Cuando el músculo esquelético se contrae se observa un acortamiento de las banda I de manera que los filamentos gruesos están muy cerca de la línea Z. (C-F) Cortes ultrafinos de músculos embrionarios silvestres de *Drosophila* en estado relajado (C) y contraído (D). Se señalan los filamentos finos (flecha) y las líneas Z (flecha curva). Las imágenes E y F se corresponden a músculos de mutantes de *mbi*. Se observa que el aspecto es similar al del músculo silvestre contraído (D). Además, las bandas Z que en los mutantes están desorganizadas apareciendo más difusas. (G, H) Preparaciones de cutículas de larvas de primer estadio silvestre (G) y mutantes *mbi* (H) en las que se aprecia el fenotipo de abdomen contraído de las larvas mutantes. Imágenes tomadas de Artero y colaboradores (1998).

1.2 *muscleblind* participa en la diferenciación de tipos celulares neuronales

Respecto al papel de *mbi* en el desarrollo de ojo de *Drosophila*, como ya se ha comentado anteriormente, el gen *mbi* se aisló como supresor del fenotipo por ganancia de función de *sev-svp2* aunque se comprobó que Muscleblind y Seven-up (*svp*) no interactúan físicamente a nivel proteico (Begemann, 1997).

El estudio de mosaicos genéticos generados mediante la técnica de recombinación mitótica ha permitido estudiar la falta de función del gen *mb1* en estadios post-embrionarios, específicamente durante la diferenciación de los fotorreceptores de la mosca. Estudios en clones mitóticos demostraron que *mb1* es necesario para la diferenciación tardía de todos los fotorreceptores. La apariencia externa de estos clones mutantes era normal. Sin embargo, el análisis de secciones transversales mostró defectos fenotípicos en los rabdómeros (estructuras en el interior de los fotorreceptores que captan la luz) que, frecuentemente, no llegaban a extenderse a las regiones más basales de la retina, mientras que las células pigmentarias y las células cono no se veían afectadas (Begemann, 1997). Aunque tanto el patrón de expresión como los defectos fenotípicos encontrados son similares a los descritos para el factor de transcripción *glass (gl)*, se ha comprobado que la expresión de *mb1* en ojo debe estar controlada por un mecanismo independiente de dicho factor de transcripción (Begemann, 1997).

El alelo *mindmelt* de *muscleblind* fue identificado a partir de una colección de 2000 líneas con inserciones de elementos P en el segundo cromosoma de *Drosophila* (Török, 1993) y posteriormente durante un rastreo genético en busca de defectos en el desarrollo del sistema nervioso periférico (SNP) embrionario (Kania, 1995). La mutación *mb1^{71/03}* se incluyó en el grupo de mutaciones identificadas causantes de defectos morfológicos. En concreto, las neuronas cordotonales laterales (Lch) aparecen asociadas más fuertemente que en la forma silvestre, además aparece una pérdida parcial de neuronas y morfologías anormales. Aunque estos defectos fenotípicos coinciden en parte con los descritos para el gen *foundue (fnd)*, se comprobó mediante ensayos de complementación que se trataba de genes diferentes. Por otro lado, la morfología de las neuronas Lch es similar a la de las neuronas mutantes para *sticky chordotonaals (stich1, stich2)* (Salzberg, 1994) por lo que se ha propuesto que todas estas mutaciones pueden estar afectando a la misma ruta del desarrollo (Kania, 1995).

1.3 Estructura genómica y procesado del gen *muscleblind*

El análisis de la unidad de transcripción *mb1* en *Drosophila* revela una organización genómica compleja (Artero 1995, Begemann 1997) (Figura. 3). Se han descrito 10 exones distribuidos a lo largo de 110 kb de DNA genómico. El tamaño de los exones oscila entre 42 y 2181 pb. Destaca la presencia de uno de los intrones más grandes

descritos en el genoma de *Drosophila* que abarca 75 kb (García-Casado, 2002). El transcrito *mbI* primario sigue un patrón de maduración complejo, dando lugar al menos a cuatro mRNAs por procesado alternativo. Estos cuatro mRNAs, denominados *mbIA*, *mbIB*, *mbIC* y *mbID* comparten la secuencia 5' no traducida (UTR) y los primeros dos exones (incluyendo el ATG y el extremo N-terminal de las proteínas codificadas) pero se diferencian en sus porciones 3' dando lugar a cuatro ORFs de diferentes longitudes que codifican para proteínas con distintos extremos C-terminales. Los transcritos denominados *mbIA*, *mbIB* y *mbIC* contienen tres ORFs de 203, 316 y 243 aminoácidos respectivamente, y comparten los primeros 179 residuos N-terminales (Figura. 4). El transcrito representado por *mbID* también contiene una ORF que comparte los primeros 63 aminoácidos con las otras isoformas de Mbl, pero debido al uso del exón 3, el cual contiene codones de parada, sólo codifica para un pequeño péptido de 84 aminoácidos. Las isoformas *mbIA*, *mbIB* y *mbIC* contienen dos copias del dedo de zinc tipo Cys₃His. Un motivo muy semejante se encuentra en las proteínas *Tristetrapolin* de ratón (TTP, también conocidas como TIS11 y NUP475) y *pharynx and intestine in excess protein 1* (PIE-1) en *C. elegans*. La isoforma corta, *MbID*, sólo contiene un dedo de zinc completo por lo que su función podría ser reguladora, actuando como una forma dominante negativa que es capaz de interactuar con las dianas normales de Mbl pero de una forma improductiva,

Un perfil de expresión mediante RT-PCR realizado para cada una de las cuatro isoformas de *splicing* descritas, indica que la expresión de las isoformas de *mbI* esta regulada a lo largo del desarrollo. Los transcritos de *mbI* tiene dos picos de expresión, uno alrededor del estadio 14-16 del desarrollo embrionario y otro en pupa temprana siendo este último el único donde se detectan transcritos para *mbIB*. Las isoformas *mbIA*, *mbIC* y *mbID* mantienen bajos niveles relativamente constantes entre estos dos picos. La isoforma *mbIC* se puede detectar en todos los estadios después de la pupa temprana (Houseley, 2003). Los transcritos *mbIC* y *mbID* se expresan en embriones de 0-4h . Como *mbIC* también se detecta en gran cantidad en hembras adultas se ha propuesto que los transcritos *mbIC* detectados en los primeros estadios del desarrollo embrionario puedan ser heredados de forma materna. El hecho de que la expresión de las diferentes isoformas proteicas de *muscleblind* esté regulada a lo largo del desarrollo sugiere que pueden tener funciones distintas (Monferrer, 2005)

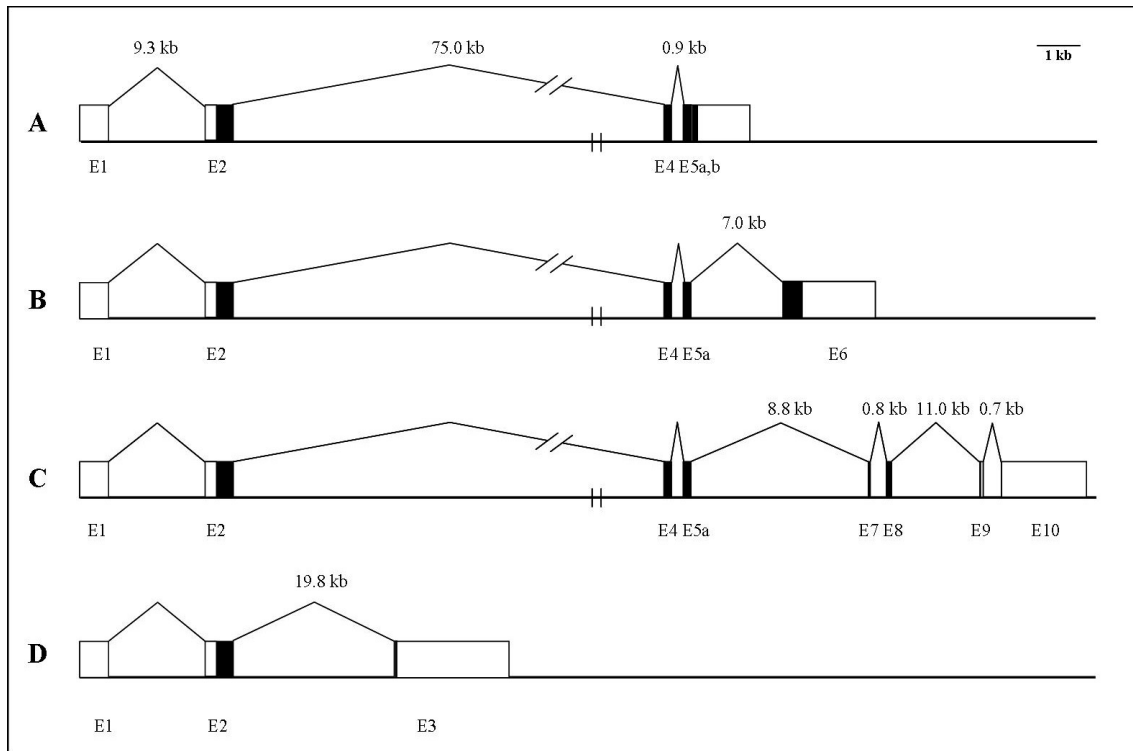


Figura 3. Estructura del gen *muscleblind* de *Drosophila melanogaster*. Se muestra la estructura de cada una de las variantes de procesado (*mbIA-D*) descritas por Begeman y colaboradores (1997) incorporando las modificaciones de García-Casado (2002). La línea horizontal indica la extensión del DNA genómico. Las cajas denotan exones y las líneas que las unen, intrones. En negro, región codificante. Figura dibujada a escala excepto los intrones (barritas verticales).

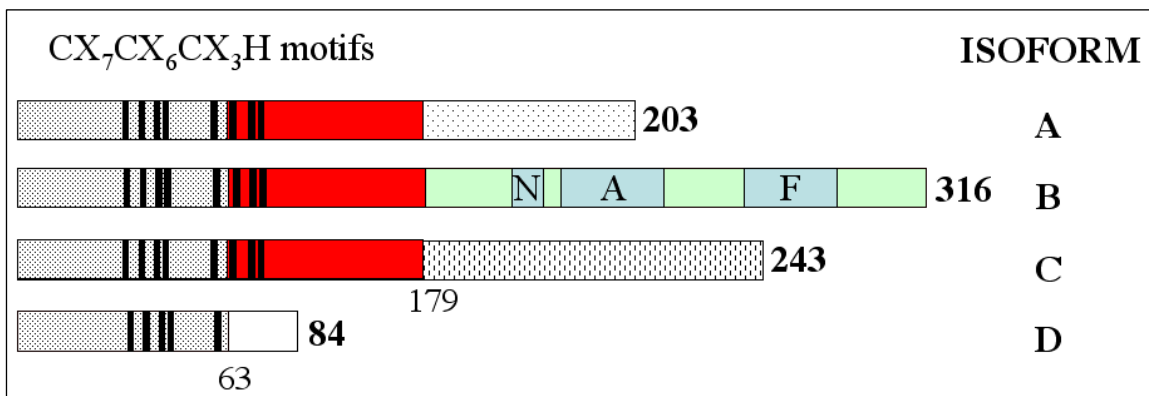


Figura 4. Representación esquemática de las cuatro isoformas proteicas generadas por corte y empalme alternativo del gen *muscleblind*. Las regiones comunes se indican con la misma trama y los dedos de zinc con cuatro líneas verticales (indicadoras de los residuos CCCH conservados). MbIB contiene tres regiones de baja complejidad ricas en glutamina (N), alanina (A) y fenilalanina (F).

2. Función molecular de las proteínas Muscleblind

En la descripción inicial del gen *mbl* de *Drosophila*, se indicó la existencia de genes homólogos en *C.elegans* y en vertebrados (Begemann, 1997). En humanos, se han descrito tres genes *muscleblind-like* (genéricamente *MBNL*). La característica estructural definitoria de las proteínas Muscleblind tanto en *Drosophila* como en humanos es la presencia de, al menos, un dedo de zinc del tipo Cys₃His. Este motivo proteico, también nombrado como CCCH, consiste en tres residuos de cisteína y uno de histidina con un espaciado entre los residuos característico y se identificó originariamente en la familia de proteínas TIS11/NUP475/TTP (Gomperts, 1990). Los miembros de esta familia juegan un papel relevante en distintos aspectos del metabolismo del RNA, como es la transcripción, el procesado y la estabilidad de los mensajeros. Por ejemplo, la proteína Tristetrapolin (TTP) se une a las secuencias ARE (*AU-rich elements*) dentro de la región 3'-UTR de los transcritos *tumour necrosis factor alpha* (*TNF*), promoviendo su deadenilación e inestabilidad (revisado en Blackshear, 2002). Otro ejemplo viene dado por la proteína *Pharynx and intestine in excess protein 1* (PIE-1) de *C.elegans* que contiene repeticiones similares a las del dominio C-terminal (CTD) de la RNA polimerasa II. La proteína PIE-1 mimetiza el dominio CTD y bloquea la actividad del factor de elongación de la transcripción P-TEFb inhibiendo, por tanto, el paso de elongación de la transcripción. Finalmente, otras proteínas pertenecientes a esta familia, como U2AF35 en mamíferos (Zhang, 1992) y *Suppressor of Sable* en *Drosophila* (Voelker, 1991), han sido implicadas en el procesado de los pre-mensajeros del RNA. Así pues, cabe esperar que la función molecular de la proteína Mbl de *Drosophila* esté relacionada con el metabolismo del RNA aunque no puede descartarse que participe o module ciertos aspectos de la transcripción.

2.1 Regulación del procesado alternativo

Actualmente no existen estudios funcionales sobre la proteína Mbl de *Drosophila*. Sin embargo, la implicación de las proteínas Muscleblind en la ruta de patogénesis de la Distrofia Miotónica (ver introducción 3.5) ha propiciado un rápido avance en el conocimiento de la función molecular de las proteínas MBNL humanas. Estas proteínas fueron aisladas a partir de extractos nucleares de células HeLa como proteínas que se unían específicamente a expansiones CUG largas (>50 repeticiones;

Miller, 2000). Experimentos *in vitro* demostraron que la unión de MBNL1 a este RNA era directa por lo que se evidenció que se trataba de proteínas de unión a RNA. Más tarde, la caracterización del ratón *knockout* de *Mbnl1* (Kanadia, 2003) mostró la posibilidad de que las proteínas Muscleblind tuvieran un papel directo sobre el procesado del pre-mRNA ya que se observaban defectos en la regulación del procesado alternativo de ciertos transcritos tales como *cardiac troponin T (cTNT)*, *muscle chloride channel (CIC-1)* y *fast skeletal troponin T (TNNT3)*. Una prueba directa de que las proteínas humanas MBNL regulan el *splicing* alternativo se obtuvo del estudio del patrón de *splicing* del exón 5 del gen de la *cTNT* y del exón 11 del *insulin receptor (IR)*. En estos experimentos, las tres proteínas humanas regulaban el *splicing* de los minigenes *cTNT* e *IR* en células transfectadas con la correspondiente proteína MBNL (Ho, 2004). Las proteínas MBNL actuaban como activadores o represores del *splicing* dependiendo del mensajero. La sobreexpresión de MBNL1, MBNL2 y MBNL3 promueve la exclusión del exón 5 en los minigenes *cTNT* pero promueven la inclusión del exón 11 en los minigenes *IR*. De acuerdo con esto, la falta de función de *MBNL1* en cultivos celulares generada por siRNA (*small interfering RNA*), muestra el efecto contrario sobre el *splicing* de dichos genes. Para llevar a cabo su función, MBNL1 se une a un motivo intrónico cerca del exón 5 del *cTNT* caracterizado por la secuencia consenso YGCU(U/G)Y (siendo Y cualquier pirimidina). En este estudio, la proteína *CUG binding protein 1 (CUGBP1)* y *embryonic lethal abnormal vision type RNA binding protein 3 (ETR-3)*, pertenecientes a la familia de proteínas CELF, antagonizan la función de MBNL en el control de la elección del exón sobre los pre-mRNA de *IR* y *cTNT*. Las proteínas CELF y MBNL se unen a distintos elementos en *cis* en los minigenes estudiados descartando, por tanto, un modelo de competición simple para la unión como base de su antagonismo. De hecho, minigenes portadores de mutaciones en los sitios de unión de las proteínas CELF o MBNL, son capaces de responder a la proteína antagonista lo que sugiere que ambas familias de factores de *splicing* regulan la elección de exones independientemente la una de la otra (Ho, 2004).

Trabajos realizados en cultivos de cardiomiocitos primarios también apoyan la actividad antagonista de las proteínas MBNL y CELF sobre el procesado alternativo del exón 5 del *cTNT* regulado durante el desarrollo (Ladd, 2005). El exón 5 se incluye en los transcritos *cTNT* embrionarios pero no en el corazón adulto. En estos estudios, los factores CUGBP1 y ETR-3 (CELF) promueven la exclusión del exón 5, mientras que las proteínas *polypyrimidine tract binding protein (PTB)* y MBNL1 reprimen su inclusión. El cambio durante el desarrollo depende de la expresión regulada de los

activadores de la inclusión del exón 5, CUGBP1 y ETR-3, y de sus represores, PTB y MBNL1. Mientras que la expresión de CUGBP1 y ETR-3 esta fuertemente disminuida en el corazón adulto, la expresión de PTB y MBNL1 se mantiene constante a lo largo de la diferenciación de este órgano. Por tanto, el balance entre los reguladores positivos y negativos de la inclusión del exón 5 es lo que determina el cambio en el *splicing* que se da a lo largo de la diferenciación del corazón.

MBNL1 también podría regular el procesado alternativo de otros mensajeros, como el *TNNT3* y el *CIC-1* ya que estos mensajeros muestran una regulación anormal del procesado en los ratones *knockout* para *Mbnl1*. De manera similar a lo que ocurre con el resto de mensajeros afectados, el defecto que se observa es la retención de las formas de *splicing* propias del tejido embrionario (Phillips, 1998; Charlet, 2002; Kanadia, 2003).

Así pues, las proteínas MBNL humanas están implicadas en el procesado alternativo de transcritos específicos de la musculatura esquelética promoviendo las formas de *splicing* características del tejido diferenciado. Esta función molecular es coherente con el fenotipo de los mutantes *muscleblind*, tanto en la mosca como en el ratón, donde encontramos una falta de diferenciación terminal en tejidos como el músculo o los fotorreceptores. De acuerdo con su papel como factor pro-miogénico, la expresión de los genes *muscleblind* se activa durante la diferenciación de los mioblastos en *Drosophila* (Artero, 1998) y en mamíferos (Miller, 2000; Squillace, 2002). Una excepción lo constituye el gen humano MBNL3 cuya expresión en músculo maduro no es detectable (Fardaei, 2002). Experimentos en cultivos celulares muestran que la expresión de MBNL3 en células C2C12 disminuye a lo largo del proceso de diferenciación. Además, la expresión constitutiva de esta proteína inhibe la inducción de las proteínas *sarcomeric myosin heavy chain* (MyHC) y la myogenin (un marcador de diferenciación temprano). Por otro lado, experimentos de siRNA contra *MBNL3* muestran una aceleración de la inducción de la MyHC durante la diferenciación de los mioblastos (Squillace, 2002). En su conjunto estos experimentos sugieren que *MBNL3* es un inhibidor de la miogénesis.

2.2 Localización de proteínas dependiente del mRNA

Un estudio reciente, ha identificado a MBNL2 como una proteína de unión a RNA que colocaliza con el mRNA de la *integrina $\alpha 3$* en el citoplasma lo que origina que la

proteína codificada se concentre en los complejos de adhesión focal (Adereth, 2005; en prensa). Experimentos de interferencia de RNA contra MBNL2 en células de carcinoma muestran que la proteína integrin $\alpha 3$ se localiza principalmente en el soma celular en vez de los complejos de adhesión donde normalmente colocaliza con la paxilina en los frentes de migración. Estos resultados implican a MBNL2 como una función esencial para la correcta localización de la proteína integrin $\alpha 3$. Las proteínas MBNL2 se localizan fundamentalmente en grandes placas citoplasmáticas que también contienen a la proteína *phospho focal adhesion kinase* (pFAK). Esta localización es dependiente de microtúbulos. Para controlar la localización subcelular de la integrin $\alpha 3$, MBNL2 se une físicamente con un motivo de secuencia denominado zipcode (ACACCC) dentro de la región 3'-UTR del mRNA de la *integrina* $\alpha 3$. Esta unión no es solo necesaria sino suficiente para dirigir la proteína GFP codificada por un mRNA quimérico, GFP-integrin $\alpha 3$, a los complejos de adhesión de una manera dependiente de MBNL2. En conjunto, se ha sugerido que MBNL2 sea un transportador del mRNA de la *integrina* $\alpha 3$ desde el núcleo hasta los lugares citoplasmáticos donde finalmente la proteína integrin $\alpha 3$ se sintetiza y se incorpora en los complejos de adhesión focal. En un cDNA array limitado a genes relacionados con la estructura y motilidad, además del mensajero de la *integrina* $\alpha 3$, se encontraron otros transcritos afectados por MBNL2 como son la *non-muscle β -actin* y el *macrophage colony stimulating factor 1 receptor* (c-fms). La regulación positiva de estos mensajeros producida por MBNL2 se produce a nivel posttranscripcional aunque los autores no han comprobado si MBNL2 está implicado en el *splicing* de la *integrina* $\alpha 3$ o alguno de estos mensajeros (Adereth, 2005).

Los resultados de Adereth y colaboradores identifican un nuevo mecanismo de localización proteica dependiente del mRNA mediado por MBNL2, cuya falta, podría explicar los cambios en la estructura celular y de adhesión, como la falta de rhabdomeros en clones mutantes *mbl* de la retina y la ausencia de matriz tendinosa en los sitios de unión de la musculatura. En *Drosophila*, los adultos dobles heterocigotos supervivientes para *mm*^{71/03} y *mbl*^{E27} (*scapers*) presentan un fenotipo característico de falta de función de las integrinas como es la presencia de burbujas en las alas (Prokopenko 1998; M. Vicente y R. Artero, datos no publicados).

En resumen, los genes *muscleblind* están implicados en la diferenciación terminal de tejidos como el músculo y el sistema nervioso. Tanto los mutantes *mbl* en *Drosophila*

como en ratón presentan un claro fenotipo muscular. En *Drosophila*, los mutantes nulos mueren en primer estadio larvario con los segmentos abdominales contraídos, posiblemente debido a la desorganización sarcomérica y la falta de matriz tendinosa que se observa a nivel ultraestructural (Artero, 1998). Los ratones mutantes *Mbnl1*, presentan hiperexcitabilidad del músculo esquelético, que se traduce en miotonía (retraso en la relajación muscular), y distintas características histológicas, algunas de las cuales, son propias de músculo no diferenciado (Kanadia, 2003). Además de la musculatura, *muscleblind* presentan fenotipos asociados al sistema nervioso. Los embriones mutantes *mbi* muestran alteraciones en las neuronas cordotoniales laterales así como defectos en la morfología, número y fasciculación neuronal, lo que sugiere que *muscleblind* es esencial durante las etapas finales de la diferenciación terminal y la morfogénesis de las neuronas (Kania, 1995; Prokopenko, 2000). De acuerdo con esto, *mbi* también es relevante en la diferenciación de los fotorreceptores del ojo de *Drosophila* (Begemann, 1997). De manera semejante a otras proteínas con los mismos dominios proteicos que Mbl, la función molecular mediante la cual llevan a cabo todas estas funciones biológicas parece ser el control postranscripcional de transcritos específicos implicados en la diferenciación terminal del tejido donde son necesarios. Para las proteínas MBNL humanas, se ha demostrado que son proteínas de unión a RNA que actúan a nivel del procesado alternativo (MBNL1, 2 y 3) y también regulando la localización subcelular de sus transcritos diana (MBNL2).

3. La Distrofia miotónica

La Distrofia miotónica (DM) forma parte del grupo de enfermedades hereditarias producidas por expansión de repeticiones inestables de microsatélites (entre tres y cinco nucleótidos). Estas repeticiones se encuentran en numerosos lugares del genoma y son polimórficas en la población general. En ocasiones estas repeticiones pueden volverse inestables durante la replicación del DNA, aumentando (expansiones) o disminuyendo (contracciones) el número de las repeticiones. Dependiendo de la secuencia de la repetición y de su localización genética, las expansiones pueden resultar patogénicas (Tabla 1). Se han descrito tres mecanismos por el cual las expansiones de repeticiones inestables desencadenan enfermedades: (1) Pérdida de función a nivel génico, como es el caso del síndrome del X frágil (FMR) en el que la expansión se produce en el extremo 5' del gen disminuyendo los niveles de proteína; (2) Ganancia de función a nivel de la proteína. En esta categoría se incluyen todas las

patologías en las que la expansión se localiza en la región codificante, por lo que es traducida a proteína. Casi todos los casos descritos corresponden a expansiones de trinucleótidos CAG, por lo que las proteínas afectadas incluyen largos tramos de poliglutaminas que alteran las interacciones normales con otras proteínas, un ejemplo arquetipo es la enfermedad de Huntington (3) Ganancia de función a nivel del RNA. Este es el mecanismo más recientemente descrito y explica los casos en los que las repeticiones se localizan en la región no traducida del gen. En este último caso se engloban enfermedades hereditarias como la Distrofia miotónica o la Ataxia espinocerebelar de tipo 8.

3.1 Distrofia miotónica tipo I y tipo II

Steinert (1909) y Batten y Gibb (1909) describieron de manera independiente la Distrofia Miotónica (DM, OMIM 160900 y OMIM 602668) como una enfermedad multisistémica que hoy en día se reconoce como una de las formas de distrofia muscular más común en adultos (1/13500 nacimientos vivos). Además de los síntomas de distrofia muscular (degeneración con falta de función) y miotonía (persistencia involuntaria de la contracción muscular), la DM provoca una gran variedad de síntomas que afectan a múltiples órganos y sistemas incluyendo defectos en la conducción cardíaca, cataratas y un conjunto característico de alteraciones del sistema endocrino (Harper, 2001). La DM tiene un patrón de herencia autosómico dominante y se han encontrado algunas características no mendelianas en esta patología como son penetrancia variable, anticipación genética (tendencia de una enfermedad a empeorar en las generaciones siguientes) y una transmisión materna preferente de las formas congénitas de la enfermedad.

La identificación de la primera mutación responsable de la DM se realizó en 1992. La causa genética de la DM1 consiste en una expansión del trinucleótido CTG en la región 3' no traducida (UTR) del gen *myotonia-dystrophica protein kinase* (*DMPK*) que

Tabla 1. Listado de patologías hereditarias asociadas a repeticiones inestables. Modificada de Cummings (2000) y actualizada según datos más recientes. UTR, región no traducida del mRNA; CDS indica región codificante; NH-term, extremo amino terminal de la proteína; COOH-term, extremo carboxi terminal de la proteína; p.f, pérdida de función; g.f, ganancia de función. Se indica a qué nivel se produce la ganancia de función (RNA o proteína). Cuando se indican ambos casos para una misma patología significa que los dos mecanismos contribuyen a la sintomatología (normalmente la falta de función se debe a una disminución en la actividad de la proteína que codifica el gen afectado por la expansión).

Patología	OMIM	Gen	Locus	Proteína	Tipo repetición	Número repeticiones		localización repetición	Mecanismo
						Normal	patológico		
Síndrome del X frágil	309550	<i>FMR1 (FRAXA)</i>	Xq27.3	FMR-1	CGG	6-53	>230 60-200*	5'-UTR	p.f
Síndrome del X frágil II	309548	<i>FMR2 (FRAXE)</i>	Xq28	FMR-2	GCC	6-35	>200 61-200	5'-UTR	p.f
Ataxia de Friederich I	229300	<i>FRDA</i>	9q13-21.1	Frataxina	GAA	7-34	>100 34-80	intrón	p.f
Distrofia Miotónica I	160900	<i>DMPK</i>	19q13.2-q13.3	DMPK	CTG	5-37	50- miles	3'-UTR	g.f (RNA) / p.f
Distrofia Miotónica II	602668	<i>ZNF9</i>	3q13.3-q24	ZNF9	CCTG	<30	75-11.000	intrón 1	g.f (RNA) / p.f
Atrofia muscular espinobulbar	313200	<i>AR</i>	Xq11-q12	AR	CAG	9-36	38-62	CDS (NH-ter)	g.f (proteína)
Distrofia Muscular Oculofaríngea	164300	<i>PABPN1</i>	14q11.2-q13	PABPN1	GCG	6	8-13	CDS (NH-ter)	g.f (proteína)
Enfermedad de Huntington	143100	<i>HD</i>	4p16.3	Huntingtina	CAG	6-35	36-121	CDS (NH-ter)	g.f (proteína)
Atrofia dentatorubral pallidolouisiana	125370	<i>DRPLA</i>	12p13.31	Atrofina-1	CAG	6-35	49-88	CDS (NH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo1	164400	<i>SCA1</i>	6p23	Ataxina-1	CAG	6-44	39-82	CDS (NH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo 2	183090	<i>SCA2</i>	12q24.1	Ataxina-2	CAG	15-31	36-63	CDS (NH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo 3	109150	<i>SCA3 (MJD1)</i>	14q32.1	Ataxina-3	CAG	12-40	55-84	CDS (COOH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo 6	183086	<i>SCA6</i>	19p13	CACNA1A	CAG	4-18	21-33	CDS (COOH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo 7	164500	<i>SCA7</i>	3p21.1-p12	Ataxina-7	CAG	4-35	37-306	CDS (NH-ter)	g.f (proteína)
Ataxia espinocerebelar tipo 8	608703	<i>SCA8</i>	13q21	no codificante	CTG	15-91	>100	3' RNA	g.f (RNA) / p.f ?
Ataxia espinocerebelar tipo 10	603516	<i>SCA10</i>	22q13	ATXN10	ATTCT	10-22	≤ 4500	intrón 9	g.f (RNA)
Ataxia espinocerebelar tipo 12	604326	<i>SCA12</i>	5q31-33	PPP2R2B	CAG	7-28	66-78	5'-UTR	pérdida función ?
Ataxia espinocerebelar tipo 17	607136	<i>SCA17(TBP)</i>	6q27	TBP	CAG	31-42	50-55	CDS (NH-ter)	g.f (proteína)
Enfermedad Huntington tipo 2	606438	<i>HDL</i>	16q24.3	JPH3	CTG	8-28	44-57	exón	g.f (RNA o proteína)

Tabla 1. Ver pie de figura página anterior

se localiza en 19q13.3 (Brook, 1992; Buxton, 1992; Fu, 1992; Harley, 1992 y Mahadevan, 1992). Poco después de que el test genético para la DM1 estuviera disponible, se describieron varias familias con DM que no tenían la expansión del trinucleótido en el cromosoma 19 (Ricker, 1994; Rowland, 1994; Thornton, 1994) lo que demostró la existencia de heterogeneidad genética en la DM. Debido a que los miembros afectados en estas familias solían acudir a la consulta debido a una debilidad muscular en la región pélvica que les dificultaba subir escaleras o incorporarse desde el suelo, Ricker y colaboradores nombraron a esta nueva forma de DM “*proximal myotonic myopathy*” (PROMM). En 1997, Udd y colaboradores describieron un fenotipo relacionado en el que la implicación de la musculatura proximal en la patología tenía un carácter más distrófico. Inicialmente pensaron que se trataba de una variante de PROMM y la llamaron “*proximal myotonic dystrophy*” (PDM). Finalmente, una familia de Minesota de gran tamaño con una segunda forma de distrofia miotónica (DM2), que presentaba una sintomatología muy similar a la de la forma adulta de la DM1, permitió mapear el gen responsable de la DM2 en 3q21 (Ranum, 1998; Day, 1999). Ricker y colaboradores pudieron demostrar que algunas familias PROMM también mapeaban en el locus DM2 (Ricker, 1999). El Consorcio Internacional de la Distrofia Miotónica (IDMC 2000) y la base de datos OMIM consensuaron la nomenclatura en base a los recientes descubrimientos clínicos y científicos de manera que el locus del cromosoma 19 se llamo DM1 (*myotonic dystrophy type I*) y al locus del cromosoma 3 se le llamó DM2. En 2001, Liquori y colaboradores demostraron que la causa genética de la DM2 era una expansión del tetranucleótido CCTG que se localiza en el intrón 1 del *zinc finger protein 9* (ZNF9).

Aunque la familia DM2 y la mayoría de las familias PROMM mapeaban en el locus DM2, otras permitieron excluir esta región (Kress, 2000; Wieser, 2000) por lo que se postuló la existencia de un tercer locus implicado en la DM (DM3). Recientemente, se ha comprobado que estas familias no poseen la expansión CCTG en el locus DM2 (Meola, 2004). Además, también se ha descrito una familia con características clínicas propias de la DM como son cataratas, miotonía y cambios histológicos miopáticos específicos cuyo gen responsable se ha mapeado en 13q21-24. Sin embargo, en esta familia, se observan también otras características clínicas que nunca se han descrito ni en DM1 ni en DM2 como son encefalopatía espongiiforme, degeneración de las motoneuronas y demencia (Le Ver, 2004). Por todo ello, actualmente, se debate la existencia de un tercer locus DM3.

3.2 Sintomatología

Músculo esquelético

La naturaleza multisistémica de la DM resulta en una gran variedad de síntomas (Tabla 2) aunque casi todos los pacientes que acuden a la consulta tienen la musculatura afectada cuando son diagnosticados. La disfunción muscular es la causa más común de dolencia tanto en la DM1 como en la DM2 (Moxley 1996; Day, 1999; Ricker 1999; Thornton 1999; Harper 2001) y se presenta en forma de debilidad muscular sintomática, dolor y miotonía. La miotonía se produce por el mantenimiento involuntario de la contracción muscular y el consecuente retraso en la relajación debido a la hiperexcitabilidad del músculo esquelético.

Existen similitudes y diferencias entre DM1 y DM2 a nivel de la musculatura. Ambas patologías tienen un patrón de debilidad muscular que afecta la musculatura facial, del cuello, de los dedos y de la cadera aunque la sintomatología inicial puede variar. Por ejemplo, en la DM2 la musculatura que suele afectarse en primer lugar es la de la cadera mientras que en la DM1 suele ser la de los dedos y se presenta junto con una atrofia muscular grave. La DM1 se asocia a una atrofia facial que confiere unos rasgos fisonómicos característicos en estos pacientes determinados por la presencia de ptosis bilateral (caída de los párpados) y debilidad de la musculatura facial. Además los pacientes con DM1 también sufren una atrofia grave de la musculatura de los antebrazos. La DM2 no se asocia ni con la atrofia de los antebrazos ni con la facial aunque si que se ha descrito la presencia de ptosis (Harper, 2001). Finalmente, aunque la miotonía se detecta en el electromiograma en casi todos los individuos DM1 y DM2, en general, los síntomas clínicos son más graves en pacientes con DM1. Respecto a las características histológicas del músculo en DM1 y DM2, ambas patologías son muy similares. Estudios de microscopía electrónica en biopsias de músculos DM1 que muestran cambios degenerativos en las miofibrillas de manera que las bandas Z e I aparecen especialmente afectadas (Ludatscher R.M, 1978).

Corazón

Las características más comunes de la DM1 y de la DM2 incluyen anomalías en la conducción atrioventricular e intraventricular, fibrilación atrial y arritmias ventriculares (Phillips, 1997). Los defectos cardíacos pueden llevar a arritmias potencialmente mortales que no están siempre precedidas por defectos en la conducción (Colleran, 1997).

Efectos multisistémicos

La DM1 y la DM2 tienen efectos similares aunque no idénticos en los sistemas no musculares a los que afectan. Las cataratas en estas dos enfermedades son idénticas y constituyen un tipo raro que se caracteriza por opacidades subcapsulares iridiscentes (Harper, 2001; Day, 2003).

Características clínicas		DM2	DM1
<i>Características músculo esquelético</i>			
Dolor muscular		++	ND
Miotonía	examen físico	+++	+++
	EMG	+++	+++
Debilidad	global	+++	+++
	facial	+	++
	cuello	+++	+++
	hombro	+	+
	dedos	++	+++
	cadera	+++	+
<i>Características multisistémicas</i>			
Cardiacas	arritmias/palpitaciones	+	++
	defectos conducción	+	++
Cataratas	examen o historial	+++	+++
<i>Características serológicas</i>			
	↑ CK	+++	+++
	↓ IgG	++	++
	↓ IgM	+	+/-
	↓ Testosterona	+	++
	↑ FSH	++	++
	insensibilidad insulina	+++	++

Tabla 2. Comparación de algunas de las características clínicas de la DM1 y la DM2. Las características clínicas comunes a ambas patologías incluyen miotonía, distrofia muscular, cataratas, arritmias cardiacas, insensibilidad a la insulina y diabetes, hipogammaglobulinemia y fallo testicular. Algunas de las diferencias entre la DM1 y la DM2 son la ausencia de una forma congénita en la DM2, con las consiguientes anomalías musculoesqueléticas y craneofaciales, y la no implicación del SNC que a veces ocurre en la DM1. En la Tabla se indican con símbolos (+) la frecuencia de los distintos síntomas para ambos tipos de DM. Para DM1 (+++) común; (++) fácilmente reconocible y/o común a lo largo del desarrollo de la enfermedad; (+) reconocible pero no común. Para el caso de la DM2 la simbología esta basada en un estudio sobre 234 pacientes (21-78 años) de manera que, en algún momento de la enfermedad, (+++) indica que >75% lo presentan; (++) >50% y (+) >20%. ND o (+/-) significa que no hay datos disponibles. ↑ y ↓ indican altos y bajos niveles, respectivamente. CK, creatina quinasa. Tabla traducida de Ranum y Day (2005).

El retraso mental aparece con frecuencia en las formas congénita y juvenil de la DM1 (Harper, 2001) mientras que no se ha asociado con certeza a la DM2. Por otro lado, tanto los pacientes DM1 como DM2 presentan una característica degenerativa del sistema nervioso central como es la aparición progresiva de depósitos de materia

blanca (Hund, 1997). Otro síntoma relacionado con el SNC es la hipersomnia que se observa en la edad adulta de algunos pacientes DM1 (Gibbs, 2002; Martínez-Rodríguez, 2003; Thorpy, 2003; Laberge, 2004).

Existen otros muchos síntomas comunes a las dos patologías tales como el fallo testicular, (tanto hipogonadismo como oligospermia), la hipogammaglobulinemia (niveles séricos reducidos de IgG y IgM) y la resistencia a la insulina. Una diferencia clara entre la DM1 y la DM2 es la no existencia de formas congénitas para la DM2. Los afectados con la forma congénita de la enfermedad nacen con graves alteraciones musculoesqueléticas, neuromusculares y neurológicas que incluyen anomalías craneofaciales, hipotonía, contractura de las articulaciones (artrogriposis) y retraso mental (Harper, 2001).

3.3 Base genética de las distrofias miotónicas

En 1992 se demostró que la causa genética de la DM1 consiste en una expansión del trinucleótido CTG en la región 3'-UTR del gen *DMPK*. La expansión de CTGs puede variar desde 80 hasta más de 4000 repeticiones en los individuos afectados mientras que en la población normal, el número de repeticiones oscila entre 5 y 37 CTGs. Existe una correlación positiva entre el número de repeticiones y la edad de inicio de la sintomatología para un número de CTGs < 400 repeticiones, pero una correlación más débil entre el número de repeticiones y la gravedad de la enfermedad para repeticiones mayores (Hamshere, 1999). Se ha descrito inestabilidad somática en algunos tejidos con un incremento de entre 50 y 80 repeticiones por año (Monckton, 1995; Martorell, 1998). También se ha observado inestabilidad intergeneracional que puede llegar a aumentar en varios miles el número de repeticiones. Esta inestabilidad de la línea germinal explica el fenómeno genético de la anticipación que se observa en las distrofias miotónicas.

Posteriormente, en 2001, Liquori y colaboradores demostraron que la causa genética de la DM2 era una expansión del tetranucleótido CCTG localizada en el intrón 1 del gen *ZNF9*. El fragmento afectado por la repetición es en realidad una secuencia consenso compleja (TG)_n(TCTG)_n(CCTG)_n; los tramos TG, TCTG y CCTG son polimórficos de manera independiente en la población en general pero solo el fragmento CCTG se expande en los individuos con DM2. En los alelos normales este fragmento aparece interrumpido por otras secuencias aunque, como ocurre en otras

enfermedades causadas por expansiones, estas interrupciones se pierden en los alelos mutados (y en los alelos de individuos no afectados que son portadores de un alelo premutado (Liquori, 2003). Aunque la DM2 es generalmente una enfermedad con una sintomatología más suave que la DM1, las expansiones CCTG pueden llegar a ser mucho más largas. En los pacientes de DM2 se encuentran alelos de entre 75 y 11.000 repeticiones (media de ~ 5000). Al igual que en la DM1 se ha comprobado la existencia de inestabilidad somática de la mutación DM2 así como de inestabilidad intergeneracional en un grado además, mucho mayor. Esta característica es la responsable de que más del 20% de las mutaciones DM2 no puedan ser detectadas mediante southern blot ya que existen tantos alelos que el DNA forma un *smear* en el que no se pueden reconocer bandas definidas. Por esta razón, el diagnóstico molecular de la DM2 se basa en un ensayo por PCR (Day, 2003).

3.4 Mecanismo de patogenicidad en las distrofias miotónicas

La DM1 fue la primera enfermedad con herencia dominante producida por la expansión de un trinucleótido que era transcrito pero no traducido y que, por tanto, no afectaba a la proteína codificada por este. Para explicar el efecto de esta mutación se han propuesto varias hipótesis:

1) Haploinsuficiencia de DMPK

Dada la localización de la mutación en la región 3' del gen *DMPK*, se propuso que las alteraciones en la expresión de este gen podrían explicar las características clínicas de la DM. Se demostró que los transcritos de *DMPK* mutantes, procesados y poliadenilados, se agregan en inclusiones nucleares. La mayoría de los estudios iniciales mostraron que los niveles de mRNA y proteína de *DMPK* estaban reducidos en pacientes y cultivos celulares (Fu, 1993; Novelli, 1993; Salvatori, 2005). Sin embargo, los ratones *knockout* para *Dmpk* generados para comprobar esta hipótesis no muestran las características multisistémicas típicas de la enfermedad. Estudios iniciales sobre estos ratones mostraron que padecían una tardía y miopatía muy suave que no es característica de la DM1 (Jansen, 1996; Reddy, 1996). Posteriormente, se demostró que tanto los ratones *Dmpk* hemicigóticos como homocigóticos presentan defectos en la conducción cardíaca (Berul, 2000). Estos resultados sugieren que *DMPK* podría contribuir a la sintomatología cardíaca de DM1 pero la haploinsuficiencia de este gen no es la causa principal del cuadro multisistémico que se observa en la

DM1. De acuerdo con esta idea, no se han encontrado mutaciones puntuales en el gen *DMPK* responsables de DM.

2) Haploinsuficiencia de *SIX5* y genes vecinos

Se ha propuesto que la expansión de trinucleótidos CTG en *DMPK* podría afectar a la expresión de genes localizados en las inmediaciones. Esta propuesta se basa en la observación de que la expansión CTG es un sitio de unión muy fuerte de nucleosomas por lo que podría estar alterando la estructura de la cromatina y tener efectos regionales sobre la expresión de diversos genes. En 1995, se determinó que la mutación DM1 afectaba a la región promotora de un gen adyacente, *SIX5* (Boucher, 1995) y que la expansión CTG disminuye los niveles del mRNA de dicho gen mediante silenciamiento transcripcional (Otten, 1995; Klesert, 1997; Thornton, 1997). Este gen es homólogo del gen *sine oculis* de *Drosophila melanogaster*, necesario para el desarrollo del ojo y de los testículos, y de una familia de genes murinos que regulan el desarrollo de la musculatura distal de las extremidades. Ya que las cataratas, la infertilidad y la atrofia de la musculatura distal son comunes en la DM1 se propuso que la haploinsuficiencia de *SIX5* podía contribuir a la enfermedad (Shaw, 1993; Boucher, 1995; Jansen, 1995). De acuerdo con esto, los ratones *knockout* para el gen *Six5* desarrollan cataratas y problemas en la conducción cardíaca (Klesert, 2000; Sarkar, 2000), así como una disminución de la masa testicular acompañada de esterilidad (Sarkar, 2004). Sin embargo, hay que resaltar que las cataratas del ratón *knockout* para *Six5* no poseen la localización subcapsular posterior y las opacidades iridiscentes características de las cataratas encontradas en los pacientes DM. Además, recientemente se ha comprobado que las propiedades contráctiles, el electromiograma y la histología del músculo de estos ratones *knockout* es normal (Personius, 2005).

Además de *SIX5*, se han propuesto otros genes cercanos a *DMPK* que pueden estar implicados en la ruta de patogénesis de la DM1 como son el gen *DMWD* (Alwazzan, 1999), con una fuerte expresión en tejidos afectados en pacientes DM1 tales como los testículos y el cerebro (Westerlaken, 2003) y el gen *FCGRT* que codifica para el receptor y transportador del fragmento Fc de la Inmunoglobulina G (Junghans, 2001).

En este modelo, las características multisistémicas de la DM1 se explicarían por la haploinsuficiencia de distintos genes localizados en las inmediaciones de la expansión de CTGs con un nivel de expresión, y por tanto, un grado de severidad, dependiente de la longitud de la misma.

3) Efecto tóxico del RNA

Una tercera hipótesis propone que las expansiones del trinucleótido CUG en el RNA de *DMPK*, interactúan con proteínas de unión a RNA de tal modo que el secuestro en las expansiones bloquea la función normal de estas proteínas (Taneja, 1995; Davis, 1997; Timchenko, 1997; Philips, 1998; Lu, 1999; Miller, 2000; Savkur, 2001). Una evidencia experimental que apoya esta hipótesis es un ratón transgénico en el cual la expansión CTG se inserta en la región 3' del gen *human skeletal actin (HSA)*, un gen no implicado directamente en la DM1 pero que se expresa en el músculo esquelético (Mankodi, 2000). Este modelo en ratón expresa 250 repeticiones de CTGs en músculo lo que provoca síntomas clínicos característicos de la DM1 como son miotonía y miopatía. Por tanto, los transcritos con repeticiones CUG son suficientes para producir un fenotipo similar a DM1 independientemente del gen *DMPK*. Además, los transcritos *DMPK* mutantes quedan retenidos en el núcleo formando inclusiones ribonucleares que se pueden detectar mediante hibridación *in situ*. Existen distintos tipos de inclusiones intracelulares que son marcadores característicos de toxicidad celular. Un ejemplo vendría dado por los cuerpos de Lewy, ligados a la enfermedad de Parkinson.

4) Reubicación de factores de transcripción

Ebralidze y colaboradores (2003) han descrito la reubicación de factores de transcripción desde la cromatina hacia complejos ribonucleoproteicos ("*RNA leaching*") en un modelo en cultivo celular de miocitos generados a partir de pacientes DM1. En este trabajo se describe que los RNAs con la repetición CUG expandida, unen y secuestran hasta el 90% de algunos factores de transcripción, como es la *Specificity protein 1 (Sp1)*, los miembros 1 y 3 de la familia *signal transducer and activator of transcription (STAT1, STAT3)* y el *retinoic acid receptor gamma (RAR γ)*. Como consecuencia del secuestro de estos factores de transcripción, diversos genes reducen su expresión. Este es el caso del *muscular chloride channel (CIC-1)*, una diana transcripcional de Sp1, cuya alteración ha sido implicada en la miotonía que padecen los paciente DM1 (Berg, 2005).

5) Por último, se ha postulado que el complejo y variable fenotipo de las DM es el resultado del efecto aditivo de varios sucesos como es, la disminución de la expresión del gen *DMPK*, la toxicidad de las repeticiones CUG, la expresión alterada de genes cercanos a la expansión de las repeticiones como *SIX5* (Tapscott, 2000) y finalmente, también debido a la disminución de la expresión de las dianas de los factores de transcripción secuestrados por las repeticiones, como son Sp1 o RAR γ . El problema más evidente del modelo aditivo es que los genes responsables de la DM1 y de la

DM2 no parecen tener funciones similares por lo que es difícil explicar que su haploinsuficiencia produzca los mismos síntomas clínicos. Además, estos genes se encuentran en entornos genómicos sin ninguna relación, esto es, los genes vecinos a las correspondientes expansiones tampoco presentan ningún tipo de similitud funcional. Por este motivo, es difícil explicar como la alteración de los genes adyacentes a las mutaciones DM1 y DM2 pueda ocasionar una sintomatología similar.

Sin embargo, existe un claro paralelismo clínico entre la DM1 y la DM2 que debe ser explicado por un mecanismo común de patogénesis. Como se ha comentado anteriormente tanto la expansión CCTG de la DM2 en el intrón 1 de la *ZNF9* como la expansión CTG de la DM1 en *DMPK* son transcritas a RNA pero no alteran la región codificante de las proteínas correspondientes por lo que no va a ser traducidas. Dada la naturaleza molecular de ambas mutaciones, expansiones que contienen la secuencia CUG, el mecanismo unificador que explicaría como dos mutaciones en genes y entornos genómicos diferentes pueden ocasionar el mismo cuadro clínico es el de la patogénesis a nivel del RNA.

3.5 Defectos moleculares en la DM

Los esfuerzos para tratar de entender cómo la expansión CUG en la DM1 expresada a nivel del RNA puede mediar un efecto dominante sobre otros genes no localizados en el locus DM1 se centraron en la identificar proteínas que pudieran unirse a las repeticiones CUG. Evidencias de que el RNA que contiene las repeticiones es responsable de la patogénesis de la DM son:

1. La expansión CTG en 3'-UTR del mensajero del gen *DMPK* inhibe la diferenciación de los mioblastos (Amack, 1999)
2. El modelo murino transgénico que expresa las repeticiones CTG a nivel de RNA padece miotonía y distrofia muscular (Mankodi, 2000; Seznec, 2001)
3. Los transcritos que contienen las expansiones CUG y CCUG se acumulan en foci nucleares (Liquori, 2001; Taneja, 1995; Mankodi, 2001; Fardaei, 2002)
4. Los transcritos que contienen las expansiones CUG y CCUG alteran la regulación o localización de proteínas de unión a RNA como CUG-BP (Timchenko, 1996) y los genes *muscleblind* humanos, MBNL1, MBNL2 y MBNL3 (Miller, 2000; Mankodi, 2001; Fardaei, 2002; Fardaei, 2001)

5. Las alteraciones en la actividad de las proteínas de unión a RNA causadas por las expansiones CUG y CCUG provocan una disfunción en el corte y empalme de distintos genes que están relacionados con la patofisiología de la DM. Actualmente, se conocen 11 genes cuyo patrón de procesado alternativo está alterado en pacientes (Tabla 3).

Gen	Nombre
TNNT2 (cTNT)	<i>cardiac troponin T</i>
IR	<i>insulin receptor</i>
CLCN1 (CIC-2)	<i>muscle-specific chloride channel</i>
MAPT (Tau)	<i>microtubule-associated protein Tay</i>
MTMR1	<i>miotubularin-related 1</i>
TNNT3	<i>fast skeletal muscle troponin T</i>
NMDAR1	<i>N-methyl-D-aspartate receptor channel</i>
APP	<i>amyloid beta A4 precursor protein</i>
RyR1	<i>skeletal muscle Ryanodine receptor type 1</i>
SERCA1	<i>sarcoplasmic/endoplasmic reticulum Ca²⁺ - ATPase protein 1</i>
SERCA2	<i>sarcoplasmic/endoplasmic reticulum Ca²⁺ - ATPase protein 2</i>

Tabla 3. Relación de genes cuyo patrón de procesado alternativo esta alterado en la Distrofia miotónica

3.5.1 Defectos en el patrón de splicing

Troponina T Cardíaca

En 1998, Philips y colaboradores descubrieron que un incremento en la actividad de la proteína CUGBP1, identificada anteriormente como una proteína de unión a repeticiones CUG (Timchenko, 1996a; Timchenko, 1996b), en el músculo de pacientes DM1 de manera que dicho incremento en la actividad alteraba el corte y empalme del mensajero de la cTNT ya que era capaz de unirse a las secuencias CUG intrónicas que sirven como señales de *splicing* en el mensajero inmaduro de la cTNT. La presencia de las repeticiones CUG en estudios de transfección y en cultivos de músculo esquelético y cardíaco de pacientes DM1 causa la inclusión del exón 5 en el mensajero cTNT lo que genera una isoforma producida durante el desarrollo temprano del corazón y músculo esquelético pero no en el tejido adulto. Posteriormente, se han

identificado nuevas proteínas que regulan el *splicing* del exón 5 de la cTNT favoreciendo su inclusión (CUGBP1 y ETR-3) o su exclusión (PTB y MBNL1-3) (Charlet-B, 2002a; Ho, 2004, Ladd, 2005, Ho, 2005). La proteína cTNT pertenece a la familia de proteínas de las troponinas cuya función esta implicada en la contracción del músculo estriado. cTNT se expresa principalmente en el corazón y también en el músculo esquelético inmaduro.

Receptor de la Insulina

Un síntoma característico tanto en la DM1 como en la DM2 es la resistencia a insulina (Moxley, 1980) que predispone a los pacientes a padecer diabetes. Se ha demostrado que el corte y empalme alternativo del mensajero del IR esta alterado en el músculo esquelético de DM1 de manera que el exón 11 es preferentemente excluido. Esto conduce a la expresión predominante de la forma resistente a la insulina, IR-A (Savkur, 2001; Seino 1989). Estos resultados se han reproducido para la DM2 (Savkur, 2004) demostrando que ambas enfermedades comparten la misma ruta de patogénesis. Se han identificado varias proteínas que se unen al mRNA de IR para regular el *splicing* específico del exón 11 tales como CUGBP1 (Savkur, 2001) y MBNL1-3 (Ho; 2004; Dansithong, 2005).

Canal de Cloro

Otra característica definitoria de la DM1 y la DM2 es la miotonía eléctrica. El músculo esquelético del modelo transgénico de la DM1 tiene reducido la conductancia transmembrana del cloro, un procesado aberrante del mensajero *CIC-1* y una pérdida de la proteína CIC-1 en la membrana celular (Mankodi, 2002). Esto también ocurre en los pacientes DM1 y DM2 (Mankodi, 2002) demostrando la relevancia de CIC-1 en el fenotipo de la DM. Se ha demostrado que el mRNA de *CIC-1* es una diana de la proteína CUGBP1, cuya actividad se encuentra aumentada en las células del músculo esquelético de DM1. Experimentos con minigenes (Mankodi, 2002) y ratones transgénicos que sobreexpresan CUGBP1 (Ho, 2005) muestran que un aumento en la actividad de CUGBP1 provoca la aparición de codones prematuros de parada debido a la retención del intrón 2 y a la inclusión del exón fetal 7a en el mensajero de *CIC-1*. Aunque no se ha comprobado que este mensajero sea una diana para las proteínas MBNL, en los ratones *knockout* para *Mbnl1* también se observan defectos de *splicing* similares (Kanadia, 2003). La proteína CIC-1 es el canal de cloro predominante en el músculo esquelético adulto y su función es esencial para mantener la estabilidad eléctrica de la membrana en dicho tejido.

Tau

Se han observado alteraciones en el corte y empalme en el pre-mRNA que codifica para la proteína Tau en el cerebro de pacientes DM1 (Sergeant, 2001) y en el ratón transgénico portados de repeticiones CTG (Seznec, 2001) lo que podría explicar algunos síntomas de la DM. Tau, pertenece a la familia de las proteínas de unión a microtúbulos. Estas proteínas se expresan fundamentalmente en neuronas y su función esencial es regular la red de microtúbulos celulares. Se ha descrito que los exones 2, 3 y 10 del gen que codifica para Tau, *MAPT*, sufren procesado alternativo lo que genera 6 isoformas distintas en el cerebro adulto (Goedert, 1989) Se ha demostrado que en cerebros de pacientes DM1, se da una reducción de las isoformas que contienen el exón 2 (Sergeant, 2001) así como también de las que contienen el exón 3 y 10 (Jiang H, 2004).

Proteína relacionada con la miotubularina 1

Buj-Bello y colaboradores (2002) identificaron tres nuevos exones codificantes en el intrón 2 del gen *MTMR1* que son procesados alternativamente y que dan lugar a 6 mRNAs distintos. Uno de estos transcritos es específico de músculo y esta inducido a lo largo de la miogénesis constituyendo la isoforma mayoritaria del gen *MTMR1* en el músculo adulto, denominada isoforma C. Los autores encuentran una disminución drástica de la isoforma C y la aparición de un transcrito anormal, la isoforma G, en cultivos diferenciados de mioblastos y en tejido muscular de pacientes con DM1 congénita. El *splicing* alternativo característico de la musculatura también esta alterado en la musculatura esquelética y el corazón de los ratones transgénicos que sobreexpresan CUGBP1, así como también, en tejido cardiaco procedente de pacientes DM (Ho, 2005). Se ha demostrado la actividad fosfatasa fosfoinositol de la proteína MTMR1 por lo que se cree que podría regular los niveles celulares de los lípidos inositol.

Troponina T tipo 3

Se ha comprobado un aumento de la inclusión del exón fetal 23 del gen *TNNT3* en tejido muscular adulto en músculo de pacientes DM1. Además, este defecto de *splicing* también ocurre en el ratón *knockout* para *Mbnl1* (Kanadia, 2003) por lo que se cree que este gen tiene un papel relevante en el procesado alternativo de este gen. La *TNNT3* es uno de los dos genes de la troponina que se expresan en el músculo esquelético siendo sus formas proteicas las de menor tamaño ("*fast form*")

NMDA receptor 1 y APP

Jiang y colaboradores (2004) realizaron un screening basado en RT-PCR en 45 exones pertenecientes a 32 genes diferentes comparando tejido cerebral normal y perteneciente a pacientes DM1. Como resultado de este trabajo encontraron diferencias en el procesado de los genes *NMDAR1* y *APP*.

Los mensajeros de *NMDAR1* sufren un aumento de la inclusión del exón 5. Este gen pertenece a la familia de los receptores de glutamato que son los receptores de neurotransmisores excitatorios más comunes en el cerebro de los mamíferos. La clase de los receptores NMDA se ha relacionado con la función del aprendizaje y el funcionamiento normal de los potenciales de larga duración en el hipocampo.

Por otro lado, los mensajeros de *APP* sufren una disminución en la inclusión del exón 7. Este gen codifica para una proteína de membrana que es el componente principal de las placas que se observan en los pacientes con la enfermedad de Alzheimer.

RyR1 y SERCA

Kimura y colaboradores (2005) investigan el *splicing* de los mRNAs de las dos proteínas principales del retículo sarcoplásmico (RS), *RyR1* y *SERCA1*. Para ambos genes, encuentran un aumento de las isoformas fetales en el músculo esquelético de DM1 y en los ratones transgénicos que sobreexpresan las CTGs para ambos genes. para una. Esto se traduce en una disminución de la inclusión del exón 70 de *RyR1* y del exón 22 de *SERCA1*. Además, identifican una nueva isoforma del gen *SERCA2* que se encuentra disminuida en el músculo esquelético de pacientes DM1 (*SERCA2b*). Las proteínas RyR1 y SERCA regulan la homeostasis intracelular del Calcio en el músculo esquelético. RyR1 libera Ca^{2+} del RS en respuesta a la despolarización de la membrana plasmática mientras que las proteínas SERCA transportan el Ca^{2+} al interior de la membrana del RS y el retículo endoplasmático (RE). La actividad de las proteínas SERCA mantiene la cantidad de calcio libre en el citosol 3-4X por debajo de la cantidad dentro del RS y el RE. Durante una contracción normal del músculo esquelético, el Ca^{2+} liberado del RS al citoplasma a través de RYR1 induce la contracción muscular. Este Ca^{2+} es posteriormente bombeado al interior del lumen del RS por SERCA para permitir la relajación.

3.5.2 El papel de muscleblind en la ruta de patogénesis de las Distrofias miotónicas

En el panorama actual de la ruta de patogénesis de la DM, dos proteínas de unión a RNA tienen un papel primordial: CUGBP1 y las proteínas *muscleblind* humanas, MBNL1-3. Aunque la actividad de CUGBP1 se aumenta en el músculo DM1 y esto resulta en un efecto dominante en *trans* sobre el *splicing* de otros genes, CUGBP1 no colocaliza con las inclusiones ribonucleares (foci) y el mecanismo por el cual su actividad aumenta se desconoce. Por el contrario, las proteínas de unión a RNA de la familia Muscleblind sí colocalizan con foci característicos de las DM aunque, inicialmente, no fueron asociadas con un efecto patogénico molecular específico. Sin embargo, recientemente se ha generado un ratón *knockout* para *Mbnl1* que presenta miotonía, miopatía, cataratas y alteraciones del *splicing* características de la DM1 y la DM2 (Kanadia, 2003). Además, experimentos recientes con células transfectadas han confirmado una implicación directa de MBNL1 en el *splicing* y se demuestra una relación entre los efectos de CUGBP1 y MBNL1. Parece que la sobreexpresión de CUGBP1 origina los mismos cambios en el patrón de corte y empalme que ocurren cuando se elimina MBNL1. En general, CUGBP1 parece promover las formas de *splicing* implicadas en el desarrollo fetal mientras que MBNL1 preferentemente dirige el *splicing* hacia formas de *splicing* que se encuentran en tejidos adultos (Ho, 2004). Estos datos predicen que la sobreexpresión de las expansiones CUG o CCUG, la sobreexpresión de la proteína CUGBP1 o la depleción de MBNL1 resultarían en el mismo conjunto de alteraciones del procesado, los cuales conducirían a las características moleculares y fisiológicas características de las distrofias miotónicas (figura 5).

Actualmente, se desconoce si existe una relación directa entre las funciones de CUGBP1 y MBNL1. En cultivos celulares se ha comprobado que CUBBP1 aumenta los niveles de proteína MEF2A durante la miogénesis (Timchenko, 2004). En *Drosophila* la expresión de *Dmef2* es necesaria para la expresión del gen *muscleblind*. Una posibilidad es que el aumento en la actividad de la proteína CUGBP1 debido a la presencia de las repeticiones CUG o CCUG produzca un alteración en los niveles del factor de transcripción MEF2A que a su vez repercute sobre la actividad de MBNL1.

Las diferencias clínicas entre DM1 y DM2 pueden ser debidas a las diferencias en el patrón de expresión temporal y espacial de los transcritos que contienen las

repeticiones CTG y CCTG respectivamente. Por otro lado, diferencias en la afinidad de las proteínas que interaccionan con estos dos tipos de repeticiones como CUGBP1 y MBNL1 podrían tener un papel en las características específicas de cada patología. Por último, las diferencias entre DM1 y DM2 podrían ser explicadas por cambios en la expresión de los genes que existen en cada uno de los loci como *DMPK*, *SIX5* o *DMWD* para DM1 y *ZNF9* para DM2.

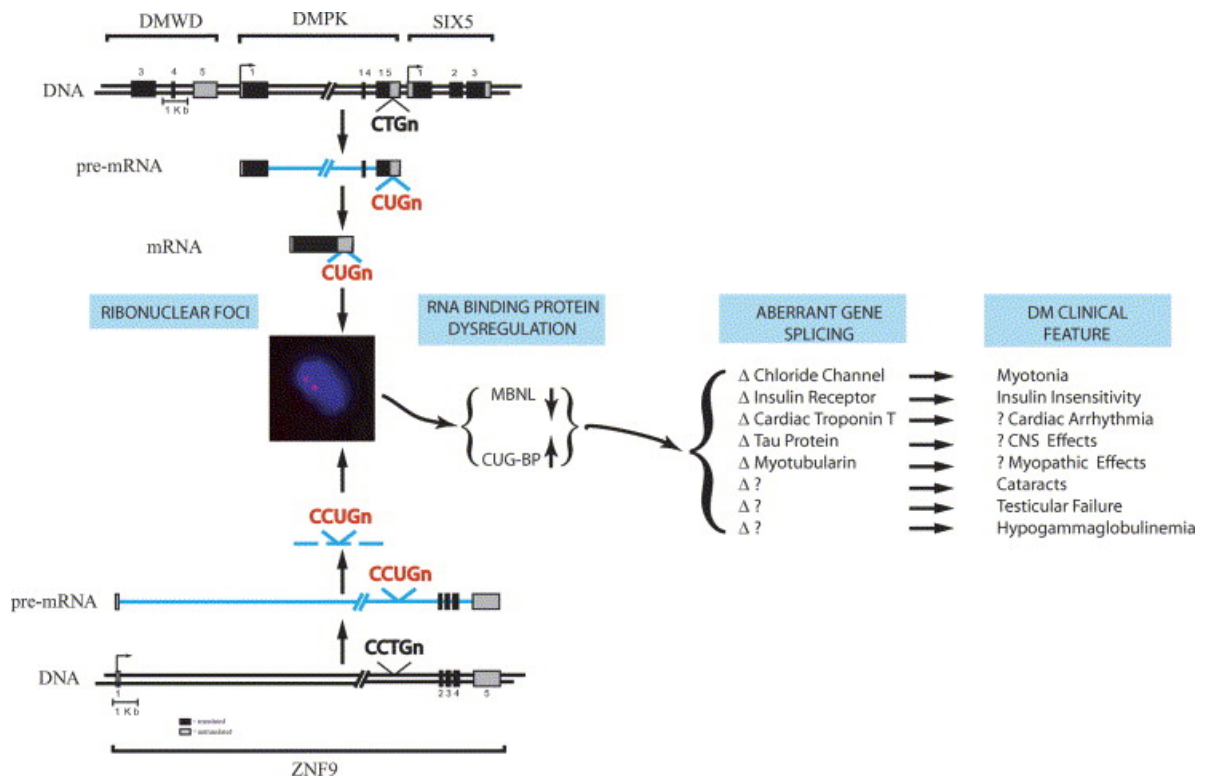


Figura 5. Papel de MBNL y CUGBP en las Distrofias miotónicas. La DM1 y la DM2 (genéricamente DM) están originadas por la expansión de los microsatélites CTG y CCTG en los genes *DMPK* y *ZNF9* respectivamente. Ambas expansiones tienen lugar en regiones transcritas pero no traducidas de los mencionados genes. Se trata de un modelo en el que la patogénesis de DM1 y DM2 está originada una ganancia de función a nivel del pre-mRNA. Los mensajeros mutados inmaduros de *DMPK* (DM1) y de *ZNF9* (DM2) forman inclusiones ribonucleoproteicas (foci) en el núcleo de las células afectadas. Se ha demostrado que las proteínas *muscleblind* (genéricamente MBNL) se unen específicamente a las expansiones patológicas de ambas patologías. Este secuestro proteico origina una falta de función de MBNL. Por otro lado, las expansiones originan un aumento en la actividad de CUGBP1 aunque el mecanismo es desconocido. La disminución de la actividad de MBNL y el aumento de la actividad CUGBP1 alteran el *splicing* de varios transcritos implicados en la patogénesis de la DM como son el canal de cloro y el receptor de la insulina. Sin embargo, los genes responsables de otras características clínicas de la DM no han sido identificados todavía como por ejemplo el fallo testicular o la hipogammaglobulinemia. Aunque para la DM1 existen otras hipótesis sobre el origen de la patogenicidad que explican estos síntomas, el hecho de que (a) estos síntomas ocurran también en la DM2 y (b) que las hipótesis no puedan ser aplicables a esta enfermedad, apoya que la causa por la que estos síntomas se producen sea el efecto tóxico de las expansiones en el RNA. Figura tomada de Ranum y Day (2005).

4. *Drosophila* como modelo

La aplicación del análisis genético en organismos modelo sencillos para descubrir los mecanismos moleculares y celulares básicos que acontecen en una célula normal proporciona una herramienta poderosa y económica para arrojar luz sobre la ruta de patogénesis de las enfermedades humanas. La mosca del vinagre, *Drosophila melanogaster*, es una excelente elección como organismo modelo. La comparación de los genomas completos de la mosca y el humano muestra una gran similitud entre las dos especies en términos de fisiología, desarrollo y comportamiento (Rubin, 2000). Como ejemplo, de una muestra de 287 genes humanos responsables de alguna patología y con equivalentes en el genoma de *Drosophila*, se encontró que 178 (62%) tenían un claro homólogo en *Drosophila* (Fortini, 2000). Además, cabe resaltar la gran cantidad de herramientas moleculares y genéticas disponibles para *Drosophila* (figura 6) y el tamaño relativamente pequeño de su genoma, lo que se traduce en que la mayoría de genes están representados en una sola copia. Esto permite que puedan ser analizados genéticamente sin problemas de redundancia genética.

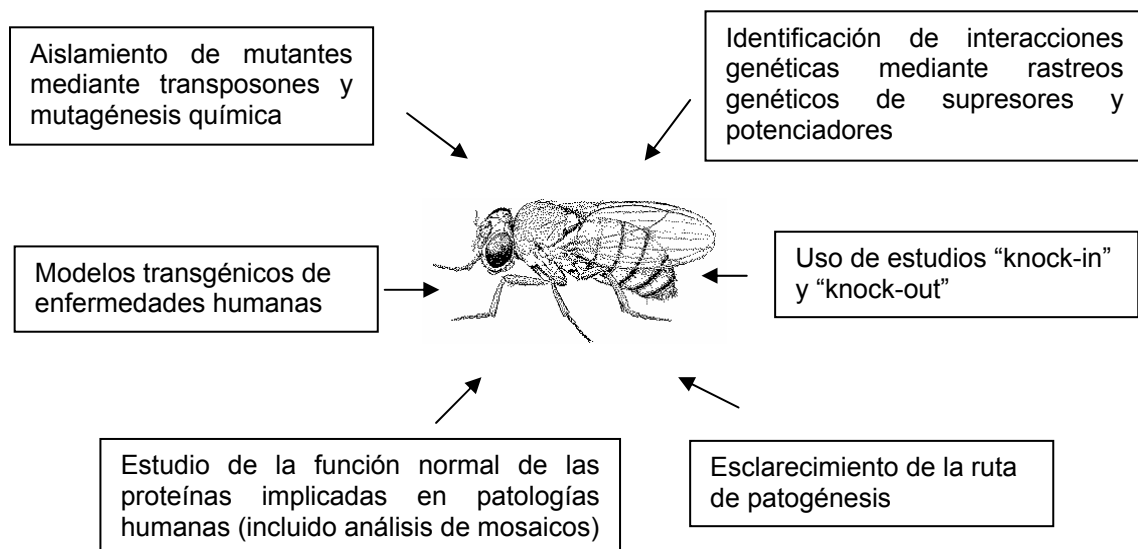


Figura 6. Distintas aproximaciones para determinar la función génica en *Drosophila*. La función de genes que son homólogos a genes implicados en enfermedades humanas pueden ser estudiadas a través de estudios clásicos de mutagénesis, haciendo uso de moscas transgénicas o de genética reversa. Los modelos transgénicos de enfermedades humanas aportan conocimiento sobre la ruta de patogénesis de la enfermedad haciendo uso de la gran variedad de técnicas desarrolladas en *Drosophila*. Los rastreos genéticos identifican los genes implicados en la ruta de patogénesis de enfermedades humanas a través de una búsqueda sistemática de mutaciones supresoras y potenciadoras. La identificación de estas mutaciones no solo ayudan a determinar la red de funciones proteicas necesarias para el proceso celular estudiado sino que también proporcionan potenciales dianas para el desarrollo de terapias, por ejemplo, la expresión de la chaperona humana hsp70 en la mosca suprime completamente el fenotipo de neurodegeneración que se produce cuando se expresa la proteína humana de la ataxia espinocerebelar 3 revelando que hsp70 tiene un papel clave en el evolución de esta enfermedad.

Drosophila se está utilizando como modelo para el estudio de numerosas enfermedades hereditarias humanas, es el caso de la Xeroderma pigmentosa (Mounkes, 1992), la corea de Huntington (Jackson, 1998; Lee, 2004), la enfermedad de Alzheimer (Ye, 1999; Crowther, 2004), las ataxias (Fernández-Fúnez, 2000), el cáncer (Potter, 2000; Busygina, 2004; Read, 2005) y la enfermedad de Parkinson (Feany, 2000; Auluck, 2002; Scherzer, 2003; Haywood, 2004; Chen, 2005). También se ha empleado en estudios de adicción a alcohol (Guarnieri, 2003) y a cocaína (Wolf, 1999). Gracias a la variedad de técnicas actualmente disponibles, se pueden identificar y caracterizar los genes implicados en la ruta de patogénesis de estas enfermedades aportando datos sobre los efectos fenotípicos de la falta y ganancia de función, el patrón espacio-temporal de expresión y, mediante estudios de epistasia, se pueden describir las relaciones genéticas entre los distintos genes llegando a establecer las rutas implicadas en el proceso estudiado en *Drosophila* para posteriormente identificar los genes ortólogos en humanos y validar los resultados obtenidos.

4.1 Modelos en *Drosophila* del efecto tóxico de las repeticiones CTG

Ataxia espinocerebelar tipo 8

La ataxia espinocerebelar tipo 8 (SCA8) es una ataxia espinocerebelar de progresión lenta, de herencia autosómica dominante y que se caracteriza por una baja penetrancia. La causa genética de la SCA8 es una expansión del trinucleótido CTG que se localiza en la región 3' de un transcrito no traducido (Koob, 1999). El extremo 5' del gen *SCA8* solapa con el extremo 5' de un gen transcrito en orientación inversa, *Kelch-like 1 (KLHL1)*. Este gen codifica para una proteína de unión a actina. Aunque no se ha demostrado una relación funcional entre los dos transcritos, la organización genómica de *SCA8* y *KLHL1* sugiere que la función normal del transcrito de *SCA8* pueda ser como un regulador antisentido del transcrito *KLHL1* (Nemes, 2000).

Las conocidas propiedades tóxicas de los transcritos con expansiones CUG y la similitud con las mutaciones DM1 y DM2 sugieren la posibilidad de que SCA8 pueda estar producida por el mismo mecanismo de ganancia de función del RNA descrito para las distrofias miotónicas. En este sentido, de acuerdo con las características multisistémicas de DM1 y DM2, la expresión de los transcritos *DMPK* y *ZNF9* se da en varios tejidos diferentes mientras que los transcritos *SCA8* se expresan casi exclusivamente en el cerebro de acuerdo con la implicación del SNC en esta

enfermedad. Aunque es posible que la expansión de CTG en SCA8 altere la regulación del gen *KLHL1* a través de un mecanismo antisentido, no hay datos experimentales que demuestren que este mecanismo exista o que tenga un papel en el desarrollo de la enfermedad. Sin embargo, si que existen evidencias experimentales que apoyan el efecto tóxico de las repeticiones CTG en la SCA8. Moseley y colaboradores (2002) generaron ratones transgénicos que sobreexpresan, según el promotor endógeno humano, el transcrito *SCA8* con las repeticiones CTG normales y expandidas. Todas las líneas de ratones con las CTG expandidas, y ninguna de las líneas control, desarrollaron un fenotipo neurológico progresivo que resultó letal para los animales que portaban mayor número de copias. Este modelo murino es coherente con el modelo de la ganancia de función del RNA en la patogénesis de SCA8

Recientemente, Mutsudi y colaboradores (2004) han desarrollado un modelo en *Drosophila* en el cual los transcritos humanos SCA8 son expresados en la retina. Tanto los transcritos controles, alelo (CTG)₉, como los expandidos, alelo (CTG)₁₁₂, inducen neurodegeneración. Esta se observa externamente como un fenotipo de ojo rugoso. El fenotipo generado por el alelo expandido se utilizó para rastrear una colección de 957 elementos P y 22 líneas con mutaciones conocidas en proteínas de unión a RNA con el fin de identificar potenciadores o supresores del fenotipo de ojos rugosos. Se identificaron 4 genes modificadores, *staufen*, *muscleblind (mbl)*, *split ends* y *CG3249*. Los tres primeros con un efecto potenciador y el último con un efecto supresor del fenotipo neurodegenerativo inducido por la expresión de los mensajeros SCA8. Aunque tanto los transcritos controles como los expandidos inducen neurodegeneración, se ha comprobado que las mutaciones en *mbl* potencian en mayor medida el fenotipo del alelo expandido que el alelo normal. Esto es coherente con el conocimiento que se tiene de las proteínas Muscleblind humanas ya que se ha demostrado que se unen a las repeticiones CUG en la región 3'-UTR del mensajero DM1 y que la fuerza de la interacción se incrementa con la longitud de las repeticiones (Miller, 2000).

Distrofia Miotónica

Para aportar nuevos datos sobre la ruta de patogénesis de la distrofia miotónica y dilucidar la función que *muscleblind* desempeña en ella, Houseley y colaboradores (2005) han desarrollado un Modelo en *Drosophila* en el que se expresan hasta 162 repeticiones CTG insertadas en un fragmento de ~100pb del extremo 3'-UTR del mRNA humano *DMPK* (modelo CTG₁₆₂). Este modelo reproduce los foci ribonucleares característicos de la DM, lo que demuestra que el proceso que subyace a su formación

está conservado en *Drosophila*. Además, demuestran que estos foci colocalizan con la proteína Muscleblind igual que ocurre con las proteínas ortólogas humanas MBNL-1, MBNL2 y MBNL3. Sin embargo, en el modelo CTG₁₆₂ no detectan ningún efecto patológico asociado a la expresión de las repeticiones, lo que demuestra que ni las repeticiones CTG expandidas ni los foci ribonucleares son tóxicos para la célula *per se*, al menos, en *Drosophila*. Estos datos contrastan con el fenotipo de ojos rugosos que Mutsuddi y colaboradores encuentran para la expresión de repeticiones CTG asociadas a SCA8 y los obtenidos en nuestro laboratorio por Monferrer y colaboradores (2004) en los que la expresión dirigida de 480 repeticiones CTG (modelo CTG₄₈₀) genera fenotipos en distintos tejidos como por ejemplo, ojo rugoso o alas extendidas. Una explicación posible para la falta de fenotipos asociados a la expresión de las 162 CTGs en el modelo de Houseley y colaboradores es que en sus experimentos utilizaron repeticiones puras CTG mientras que en los modelos de Mutsuddi y Monferrer emplearon expansiones CTG microsatélite interrumpidas regularmente por otros tripletes. Sin embargo, las expansiones interrumpidas empleadas en el modelo CTG₄₈₀ han sido utilizadas en diversos experimentos de cotransfección en cultivos celulares, de manera que reproducen correctamente aspectos de la patología (Philips, 1998; Savkur, 2001). Por tanto, la causa más plausible de que el modelo CTG₁₆₂ no genere fenotipos visibles es que el número de repeticiones con las que se ha trabajado no es suficiente.

El modelo de Houseley y colaboradores muestra que la formación de foci ribonucleares no es una manifestación obligada de la presencia de las repeticiones CTG₁₆₂, ya que los foci no aparecen en todos los tejidos donde se expresan. Por otro lado, determinan que Muscleblind no es ni necesaria ni suficiente para la formación de los foci ribonucleares ya que en algunos tejidos en los que se generan foci no se detecta proteína Mbl y en algunos tejidos en los que se detecta Mbl, como discos imaginales y músculos indirectos de vuelo, no se generan foci pese a la expresión de las repeticiones CTG₁₆₂. Estas observaciones indican que la formación de foci depende de un conjunto de proteínas de unión a RNA específicas para cada tipo celular. Además, estos foci no colocalizan con proteínas del proteosoma lo que indica que es improbable que los foci contengan proteínas mal plegadas como es el caso de los modelos en los que se sobreexpresan repeticiones en la región codificante (Warrick, 1998; Jackson 1998). Aunque no se ha encontrado colocalización con proteínas relacionadas con el espliceosoma, tales como la proteína SC-35, si que se ha demostrado la colocalización de la proteína non-A, una proteína cuyos ortólogos humanos, PSP1, PSF y p54^{nrb} están implicados en números procesos que incluyen

splicing y unión de los RNA de doble cadena (dsRNA) a la matriz nuclear. Houseley y colaboradores (2005) proponen un modelo en el que, en una célula normal, al menos parte de las proteínas Muscleblind están unidas a la matriz nuclear mediante su interacción con las proteínas PSF/p54nrb y que los transcritos que contienen tramos CUG se unen a una proteína Muscleblind que es liberada de la matriz y que puede entonces participar en la exportación, estabilidad o control traduccional del mensajero. Cuando ocurre la expansión del trinucleótido CUG se produce la unión de más de una molécula Muscleblind, lo que conduce a un retraso en la disociación de la matriz extracelular ya que muchas moléculas Muscleblind deben disociarse simultáneamente para que se produzca la liberación del complejo ribonucleoproteico, lo que conduciría a la formación de los foci.

El modelo de Houseley muestra que la sobreexpresión conjunta de las proteínas Muscleblind junto con las repeticiones CTG produce un aumento en la cantidad de estos mensajeros, lo que sugiere que Muscleblind podría estar implicado en la regulación de la estabilidad de los mensajeros.

A pesar del gran avance que se ha producido en el entendimiento de las funciones moleculares de las proteínas Muscleblind, aún quedan muchas cuestiones por resolver. Por ejemplo, cómo las proteínas MBNL pueden promover la inclusión de un exón en un sustrato y reprimirla en otro, si existe una interacción física con las proteínas de efecto antagónico CELF, cual es el mecanismo molecular por el que MBNL3 inhibe la diferenciación muscular, si MBNL2 está implicada en el *splicing* de su mRNA diana, *integrina $\alpha 3$* , o si MBNL1 y 3 también tienen algún papel en el mecanismo de localización proteica dependiente de RNA. Actualmente, no se conoce ninguna proteína que interaccione físicamente con Muscleblind, ni en mamíferos ni en *Drosophila*. La identificación de las dianas moleculares de las proteínas Muscleblind ayudará a entender cómo las proteínas MBNL llevan a cabo su función como factor de *splicing*, así como, permitirá conocer en qué otras funciones moleculares están implicadas. Finalmente, dada la relevancia de las proteínas MBNL en la ruta de patogénesis de la DM, el conocimiento de sus dianas moleculares aportará información clave para entender como la falta de función de MBNL contribuye al desarrollo de la enfermedad

OBJETIVOS

El trabajo de esta tesis doctoral se enmarca en el proyecto de investigación “Análisis genético y molecular del efecto tóxico de las repeticiones CTG en *Drosophila*” del laboratorio de Modelos Biomédicos del Departamento de Genética de la Universidad de Valencia. Dentro de este proyecto se están desarrollando otras dos tesis doctorales en las que se estudia el efecto tóxico de las repeticiones CTG en *Drosophila* (Lda. Lidón Monferrer) y se caracterizan los dominios funcionales de la proteína Muscleblind en *Drosophila* (Lda. Marta Vicente).

El objetivo general de este trabajo es aportar datos sobre la función molecular de la proteína Muscleblind con el fin de generar información de interés biomédico para entender, y potencialmente tratar, las enfermedades causadas por expansiones del trinucleótido CTG. Para la consecución de este objetivo general nos proponemos la realización de los siguientes objetivos específicos:

1. Comparar el gen *muscleblind* de *Drosophila* con los genes humanos a nivel de la estructura genómica y el corte y empalme alternativo.
2. Identificar las secuencias conservadas evolutivamente en la familia de proteínas Muscleblind para identificar regiones esenciales para la función de la proteína.
3. Estudiar la posible implicación de los genes *muscleblind* humanos en distintas patologías hereditarias.
4. Determinar los procesos biológicos en los que participa muscleblind mediante un rastreo genético de modificadores dominantes en *Drosophila*
5. Identificar proteínas que colaboren con Muscleblind en su función molecular mediante la técnica de coinmunoprecipitación

MATERIAL Y MÉTODOS

1. MATERIAL

1.1 Pacientes

Síndrome de Usher tipo III

El diagnóstico de los pacientes USH fue hecho conjuntamente por la Unidad de Neurofisiología y la Unidad de Otorrinolaringología del Hospital Universitario La Fe (Valencia), así como también por la Unidad de Oftalmología del Hospital Clínico (Valencia). En este trabajo se han analizado cuatro pacientes. Se trata de dos parejas de hermanos (608-609 y 934-935) afectados de USH1 pertenecientes a dos familias no relacionadas entre sí que presentaban un claro ligamiento a la región cromosómica 3q donde ha sido asignado previamente el locus USH3.

Distrofía Miotónica

El diagnóstico de los pacientes con DM (muestras 1-9, 13) fue realizado por distintos facultativos de la comunidad Valenciana y remitidos a la Unidad de Genética del Hospital Clínico Universitario La Fe para la detección molecular de la mutación responsable. En ninguno de los diez pacientes DM analizados en este trabajo se encontró expansión del trinucleótido CTG en el locus DMPK. Uno de los pacientes PROMM (muestra 10) fue diagnosticado por el Servicio de Neurología de la Fe y remitido a la Unidad de Genética para descartar la existencia de expansión en DMPK. El otro paciente PROMM (muestra 14) era el probando de una familia diagnosticada por el grupo de la Dra Tranebjaerg del University Hospital of Tromso (Noruega) que mostraba ligamiento al locus DM2, en 3q25. Los dos pacientes asintomáticos (muestras 11 y 12) tenían expansiones de 300 y 500 repeticiones y una edad de 30 y 32 años respectivamente. El motivo por el que acudieron a la consulta fue por fatigabilidad (muestra 11) y por dar a luz una niña con distrofia miotónica congénita (muestra 12).

Ptosis congénita ligada al cromosoma X

El diagnóstico de los pacientes fue realizado por McMullan y colaboradores y los detalles sobre el establecimiento de esta nueva condición oftalmológica pueden encontrarse en McMullan y col (2000) y McMullan y Tyers (2001)

1.2 *Drosophila melanogaster*

Cepa <i>Drosophila</i>	Abreviatura	Fuente
Oregon R	OrR	-
$y^1 w^{1118}; mbf^{E27}/CyOy^+; P\{w^{+mC}=UAS-mblC:GFP\}T15.3$	T15.3	Creada como parte de este trabajo
$y^1 w^{1118}; mbf^{E27}/CyOy^+$	-	Creada como parte de este trabajo
$y^1 w^{1118}; mbf^{E27}/CyOy^+; P\{w^{+mw.hs}=GAL4-da.G32\}UH1$	-	Creada por L. Monferrer
$w^{1118}; +; P\{w^{+mw.hs}=GAL4-da.G32\}UH1$	da-GAL4	-
$y^1 w^{1118}; mbf^{E16}/CyO$	-	Creada como parte de este trabajo
$y^1 w^{1118}; If/CyOy^+$	-	Creada como parte de este trabajo
$y^1 w^{1118}; Gla/CyOy^+$	-	Creada como parte de este trabajo
$y^1 w^{1118}$	-	-

1.3 *Escherichia coli*

La cepa de *Escherichia coli* que hemos empleado para todas las aplicaciones ha sido XL1-Blue. La principal ventaja de utilizar esta cepa con los plásmidos de la serie pUC es que se produce el fenómeno de la α -complementación cuando el plásmido no es recombinante y permite, por tanto, la selección de colonias recombinantes en un medio indicador con IPTG y X-gal, por su color blanco. Además, contiene resistencia a tetraciclina lo que disminuye el riesgo de contaminación bacteriana. Su genotipo completo es *supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac⁻ F'⁺{proAB⁺ lac^f lacZ Δ M15 Tn10(tet^r)}*.

1.4 Anticuerpos

Anticuerpo	creado	Proveedor	Descrito	Concentración de uso
anti-GFP	ratón	Roche		1/1000
anti-Jumu	conejo	Dr. K. Günter	<i>Strodicke et al, 2000</i>	1/1000
anti-Mbl (2º sangrado)	oveja	Dr. D. Monckton	<i>Houseley et al, 2003</i>	1/5000
anti-nonA (Bj6)	ratón	Prof. H. Saumweber	<i>Frasch et al, 1989</i>	1/500
anti-Mbl _{CUK333}	conejo	parte de este trabajo		1/500
anti-Aly (DmREF1)	conejo	Dr E. Izaurrealde	<i>Gatfield et al, 2002</i>	1/500
anti-Amos	conejo	Dr. A. Jarman	<i>Zur lage et al, 2003</i>	1/500
anti- α -tubulina	ratón	Sigma		1/5000

anti-guinea pig-HRP	conejo	Sigma	1/5000
anti-HA-HRP	rata	Roche	1/500
anti-mouse-HRP	cabra	Sigma	1/5000
anti-mouse-bio	-	Sigma	1/1000
anti-rabbit-HRP	cabra	Sigma	1/5000
anti-rabbit-bio	-	Pierce	1/1000
anti-sheep-HRP	burro	Sigma	1/5000
strepTactin-HRP	-	Bio-Rad	1/10.000

1.5 Enzimas

Enzima	Proveedor
T4 DNA ligasa	Roche
Netzyme [®] DNA polimerasa	N.E.E.D

1.6. Reactivos

Reactivo	Proveedor
Tri-Reagent	Sigma
EZview [®] Red Protein G affinity gel	Sigma
Dimetil Pimelidato (DMP)	Sigma
Complete [™] mini EDTA-free	Roche
Azida	Sigma
ProteinG-agarose beads	Roche
Proteogel	Pronadisa
Halocarbon 700	Sigma
Aceite mineral libre de nucleasas	Sigma
SYPRO-RUBY	Invitrogene
BSA	Roche
Azul Brillante Coomassie R50	Panreac
Glutation reducido	

1.7. kits

kit	Proveedor
QIAEXII agarose Gel Extraction	Quiagen

High Pure Plasmid Isolation kit	Roche
GenoPure Plasmid Midi Kit	Roche
Coomasie® Protein Assay Reagent kit	Pierce
Lumi-light Western Blotting substrate	Roche
First Strand cDNA synthesis for RT-PCR (AMV) kit	Roche
Vectastain® Elite ABC kit	Vector Lab
Silver Staining kit Protein	Amersham

1.8. Cebadores

Nombre	Secuencia 5'-3'	Diana
<i>XhoI-mbl</i>	CCG CTC GAG ATG GCC AAC GTT GTC AAT ATG	<i>muscleblind</i>
<i>mblC-EcoRIGC</i>	CGG AAT TCC GTC TTG GCA CAC CGG GAG GG	<i>muscleblind isoforma C</i>
actE5d	TTA TCC ATC GCC ATC GT	<i>α-actinina</i> (exón 5)
actE10r	TTG AAG TTG GTC TCC AGC	<i>α-actinina</i> (exón 10)
RP49d	ATG ACC ATC CGC CCA GCAT A	<i>Proteína ribosomal 49</i>
RP49r	CAT GTG GCG GGT GCG CTT GTT C	<i>Proteína ribosomal 49</i>
MBNLF	GGT TTC ACT GAA ACA TTT AAC TAC C	ORF de MBNL1 en 5'
MBNLR	AGC ACA CTG TCC TTT CCT CTT TAG C	ORF de MBNL1 en 3'
MBNLF1	TGC CCC ATT ACA ACC CGT GC	interno ORF MBNL1
MBNLF2	CCC TGC ACA TTT GCA AGC CA	linterno ORF MBNL1
MBNLR1	ACG GGT TGT AAT GGG GCA CC	interno ORF MBNL1
MBNLR2	TGG TTG ACC TGG TAT TGG GC	interno ORF MBNL1
E1F	GCA GAG AAC TGT ATG TCC	exon 1 de MBNL1
E1R	GGT ATC ACA TGG CAA GGT TG	exon 1 de MBNL1
E2F	GCC TGT GTT CAT TTG ATC CG	exon 2 de MBNL1
E2R	CAT TCA AGT GGG TGG TTT GC	exon 2 de MBNL1
E3F	GTC CAA CCT CTT ATC CAC	exon 3 de MBNL1
E3R	TAA AGG GAA GGA GGG AAG	exon 3 de MBNL1
E4F	CTT CTG TCA CCA GTC TTG C	exon 4 de MBNL1
E4R	GAG TAG GAG CAA TCT GAA CC	exon 4 de MBNL1

E5F	TGT CAG GAA ACC ACA CTC AG	exon 5 de MBNL1
E5R	CAA AGA GGA GTT ATC CTC CC	exon 5 de MBNL1
E6F	TTG GGG TGT GGT TTA TCT GC	exon 6 de MBNL1
E6R	TGC TTG CTT GCT CAT GCT TC	exon 6 de MBNL1
E7F	AAA GGG AAA TGG ACT TGC CC	exon 7 de MBNL1
E7R	AGT TAC ATG CCA CTG TTC CC	exon 7 de MBNL1
E8F	TCA ATG GGA AAT CCA CGT GC	exon 8 de MBNL1
E8R	CCT GTC ACA TGG CTT ATT GC	exon 8 de MBNL1
E9F	ACT ACA ATC CAC AGT CTG GC	exon 9 de MBNL1
E9R	CTG GAA ATA GCT GAT GCC AC	exon 9 de MBNL1
E10F	GTC TGA TAA AAT TCC TCA CAC CC	exon 10 de MBNL1 (solo ORF)
E10R	AAG GAC TAT CAC CTC ACA GCT G	exon 10 de MBNL1 (solo ORF)
EFIRSTF	GCA ACC TGT TAT CTC TTC GC	exón 1 de MBNL3
EFIRSTR	GTC GTT AAG ACA TTC TGA C	exón 1 de MBNL3
E165F	CCC CAG ATA ATG GAG TCC	exon 2 de MBNL3
E165R	ATT TCC CCA AGT CCC TGC	exon 2 de MBNL3
E192F	ACC TCT TTG TGT AAT GAC AGG	exon 3 de MBNL3
E192R	AAA CAA CAC ACT GGG CTC	exon 3 de MBNL3
E237F	AAA AGC CAG TCA GCT TCC	exon 4 de MBNL3
E237R	GAC CAT CCA TGA GTA TGG	exon 4 de MBNL3
E151F	TGG TGC TTG CTT GGT TTG	exon 6 de MBNL3
E151R	CTC CTC CTA TAC AAG CTC	exon 6 de MBNL3
E36F2	GGA GAT TGT TGG ACT TGC	exón 7 de MBNL3
E36R2	CAC ATC TTT CAG TCA TTC	exón 7 de MBNL3
E95F	TAC CAC ACC TCA TTC CAT CG	exón 8 de MBNL3
E95R	CTC CTC TGA TTC ATC TTG GG	exón 8 de MBNL3
E75F2	GTG GGA ACC ATG CTT ATA G	exón 10 de MBNL3
E75R2	CTA CAT CAC AAG CAA TGT C	exón 10 de MBNL3

1.9 Equipos

Equipo	Proveedor	Uso
Mini-Protean [®] 3 Cell	Bio-Rad	Electroforesis de proteínas
Hoefer [™] SE600 Ruby [™]	Amersham biosciences	Electroforesis de proteínas
Mini Trans-blot [®] Electroforetic Transfer Cell	Bio-Rad	Transferencia de proteínas en tanque
Nutator [®]	Clay Adams [®] Brand	Agitador nutacional
Rotator Drive STR4	Stuart Scientific	Agitador Noria
Sonicador	Sonics material VibraCell	Disrupción células

1.10 Miscelánea

Material	Proveedor
Placas Petri cuadradas Bibby Sterilin	Afora
Homogenizadores Potter-Elvehjem con capuchón de PTFE	Afora
PVDF Western blotting membrane	Roche
Precision Plus Protein Standard Dual color prestained	Bio-Rad

1.11 Soluciones

Tampón de carga para SDS-PAGE 2x

50mM Tris-HCl pH 6.8, 100mM DTT, 2% (p/v) SDS, 0.1% azul de bromofenol y 10% glicerol

Tampón de electroforesis SDS-PAGE

25mM Tris base, 250mM glicina pH8.3, 0.1% (p/v) SDS

Tampón de transferencia SDS-PAGE

48mM Tris base, 39mM glicina, 20% metanol y 0.0375% SDS

Tampón de Microinyección

10mM tampón fosfato pH7.8 y 5mM KCl

Tampón borato

200mM tetraborato sodico decahidratado pH9.0

Tampón DMP

20mM de Dimetil Pimelidate (DMP) en tampón borato,

Solución fijadora

50% metanol y 10% acético en agua

TBST

50mM Tris base, 150mM NaCl y 0.05% Tween 20

Solución de color

0.3 mg/ml DAB, 0.2 mg/ml NiCl₂ y 0.003% H₂O₂ en TBS

PBS

10mM fosfato sódico pH7.5 y 0.9% NaCl

Tampón de Elución

50mM TrisHCl pH8 y 5mM glutatión reducido

Solución de Coomassie

Azul brillante Coomassie R50 0.03M, 45% agua, 45% metanol y 10 ácido acético glacial.

2. MÉTODOS

Todos los reactivos y kits han sido usados según instrucciones del fabricante excepto cuando se especifique lo contrario. Los métodos moleculares han sido extraídos de Sambrook (1989, 2001), los métodos generales de proteínas de Ausubel (2002) y los protocolos generales de *Drosophila* de Sullivan (2000). El protocolo de producción de proteína GST se ha basado en las recomendaciones de Pharmacia Biotech (1997).

2.1. Lista de enfermedades hereditarias candidatas para secuenciación MBNL3

Conocida la localización cromosómica del gen MBNL3 en Xq25-q26.3 (clon genómico RP5-842K24) búscamos todas las patologías que mapean en la misma región utilizando la base de datos Online Mendelian Inheritance in Man (OMIM) en el NCBI (<http://www.ncbi.nlm.nih.gov>). Después, se realizó un estudio del documento OMIM o de los artículos originales de cada una de las patologías con el fin de conocer los dos marcadores moleculares que delimitan el intervalo crítico donde debe estar el gen responsable de esa enfermedad. DXS1047 y DXS998 son dos marcadores moleculares dentro de la región genómica que abarca MBNL3 por lo que comprobamos estos marcadores estaban dentro del intervalo crítico definido para cada enfermedad. Los mapas físicos del cromosoma X consultados durante la realización del trabajo fueron el Mapa de STS del NCBI (<http://www.ncbi.nlm.nih.gov/mapview>) y el Mapa físico del Xace junto con el Map Status Sanger Institute (<http://www.sanger.ac.uk/HGP/ChrX/mapping.shtml>).

Sobre el conjunto de patologías cuya región crítica incluía el gen MBNL3 seleccionamos aquellas cuya sintomatología afecta a los tejidos que suponemos más

relevantes para la función de este gen, esto es, musculatura, sistema nervioso y ojos. Esta selección en base al fenotipo de cada una de las patologías se hace en relación al conocimiento que se tiene del gen homólogo en *Drosophila melanogaster*, *muscleblind*, y de la conocida implicación de las proteínas MBNL en la ruta de patogénesis de la Distrofía miotónica.

2.2 Extracción de RNA y RT-PCR

Las muestras de sangre fueron facilitadas por los pacientes en la Unidad de Genética del Hospital Universitario “La Fe”. Las muestras se mezclaron con 1 volumen de suero fisiológico (aprox. 5ml) y se congelaron a -20°C hasta el momento de su procesamiento. La obtención de células mononucleares se hizo siguiendo el protocolo de LymphoprepTM, el pellet resultante se resuspendió en 700 μl del tampón RLT (Quiagen). La extracción de RNA se llevo a cabo con el RNeasy Mini Kit (Quiagen, VLC) a partir de 350 μl de la suspensión

Los cebadores utilizados para amplificar la ORF completa de MBNL fueron MBNLF y MBNLR. Las RT-PCR se hicieron con la Titan One Tube RT-PCR System (Roche Diagnostics, VLC) según las intrucciones del fabricante. Se utilizó 1 μg de RNA aproximadamente. La temperatura de annealing utilizada fue de 62°C y la concentración de iones Mg^{+} de 3mM. El tamaño del amplificado es de 1200 pb. Para poder secuenciar el producto de PCR completamente se utilizaron los siguientes cebadores internos: MBNLF1, MBNLF2, MBNLR1 y MBNLR2.

2.3 Secuenciación de DNA

Las reacciones de secuenciación se realizaron con el ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit en un termociclador modelo 9600 de PE. Los productos de las reacciones fueron analizados en un secuenciador automático modelo ABI PRISM 377.

2.4 Rastreo fenotipo de sobreexpresión de *mbIC* en ojo

La sobreexpresión de *mbIC* en el ojo de *Drosophila* con el promotor *sevenless* (*sev-Gal4*) produce un fenotipo de pérdida de fotorreceptores y problemas en la polaridad epitelial plana (PCP). El fenotipo es dependiente del grado de expresión lo que sugiere que puede estar sensibilizado a cambios en las dosis de genes relacionados funcionalmente con *mbIC* (Garcia-Casado, 2002). Hemos elegido este fenotipo para rastrear y validar *in vivo* diferentes genes candidatos a interactuar físicamente con la proteína Mbl. Hembras del recombinante UAS:*mbIC-sevGAL4* (Garcia-Casado, 2002) se cruzaron con machos de cada uno de los stocks de deficiencias testados (figura 1). Los ojos de las moscas de la primera generación filial (F1) portadoras del cromosoma UAS:*mbIC-sevGAL4* y de la deleción, se compararon con los ojos de las moscas que portaban el cromosoma UAS:*mbIC-sevGAL4* únicamente, obtenidos a partir de un cruce control paralelo.

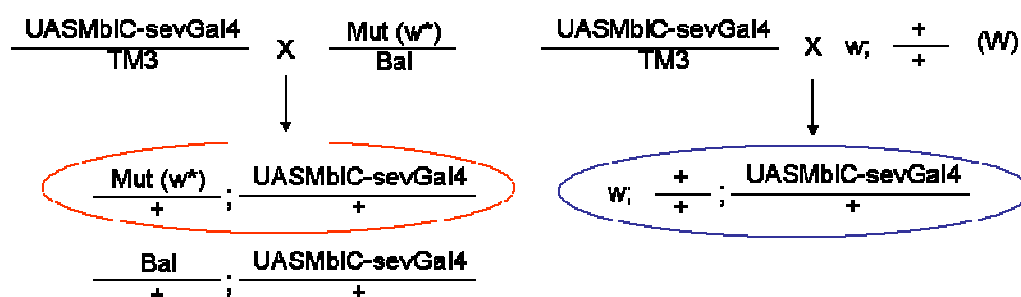


Figura 1. Esquema de los cruces para la búsqueda de interacciones genéticas con la sobreexpresión de *muscleblind* en el ojo (UAS*mbIC-sevGal4*). La interacción de la mutación introducida (Mut) se testaba comparando el fenotipo de ojos rugosos de los individuos que sobreexpresan la proteína MblC y que contienen la mutación que se quiere testar (círculo rojo) con el fenotipo de ojos rugosos de los individuos control, que sobreexpresan MblC sin la presencia de esa mutación (círculo azul). Si el fenotipo es similar concluimos que no hay interacción, si el fenotipo de ojos rugosos está potenciado o atenuado respecto al control determinamos que la mutación potencia o suprime, respectivamente. MblC, isoforma C de *muscleblind*. *sev*, *sevenless*. Mut, mutación. Bal, balanceador.

2.5 Clonación de MblC:GFP

La región codificante del cDNA *mbIC* (AF004536, GenBank) se amplificó mediante PCR usando cebadores que incluían dianas para los enzimas de restricción *XhoI* (*XhoI-mbl*) y *EcoRI* (*mbIC-EcoRIGC*) en la región 5' y 3', respectivamente. Este último cebador añadía, además, dos nucleótidos que posibilitan la clonación en pauta de este fragmento en el plásmido *peGFP-N3*. El producto de PCR resultante y el plásmido

peGFP-N3 se sometieron a digestión doble *XhoI-EcoRI* y se ligaron. El plásmido peGFP-N3 con la inserción mbIC fue digerido con *BglII* y *NotI* y el fragmento de restricción *mbIC*:GFP secuenciado para verificar la pauta de lectura de la proteína de fusión codificada. La clonación en pUAST se realizó ligando dicho fragmento con el vector digerido con *BglII-NotI*. Una vez comprobada la clonación se confirmó la identidad del inserto mediante digestión con enzimas de restricción.

Para introducir los plásmidos recombinantes en *Escherichia coli* hemos empleado el método de choque térmico, 42°C durante 1 min, a células competentes XL1-Blue. Para hacer células competentes para la transformación hemos seguido las instrucciones de Nishimura (1990). Las células competentes se almacenaron a -80°C hasta su utilización.

La selección de clones recombinantes ha dependido del vector de clonación utilizado y se ha basado en la resistencia a antibióticos que confiere la presencia del plásmido. Para comprobar la clonación del inserto, las colonias crecidas se sembraron en 4 ml de cultivo con el antibiótico seleccionador, se crecieron toda la noche en agitación a 37°C y se procedió a una extracción de DNA plasmídico a pequeña escala (miniprep) empleando el método de lisis alcalina (Ausubel et al, 1988) con las modificaciones señaladas en Pérez-Alonso (1989). Cuando fue necesario obtener mayores cantidades de DNA plasmídico se utilizó el kit "Geno Pure Plasmid Midi kit".

2.6. Obtención de Moscas Transgénicas

Microinyección de embriones de Drosophila

La microinyección de los embriones de *Drosophila* se realizó siguiendo las adaptaciones de García-Casado (2002) basadas en Rubin y Spradling (1982). El plásmido pUAST con el DNA de interés clonado, *mbIC*:GFP, y el plásmido "helper" fueron diluidos a una concentración de 0.2 µg/µl y 0.05 µg/µl, respectivamente, en tampón de microinyección y centrifugado a 16.000 g durante 15 min para eliminar cualquier partícula en suspensión. Se recogieron embriones *yw;mbI^{E27}/CyOy+* en placas petri de 35 milímetros con agar y se decoriorizaron a mano antes de montarse sobre pegamento obtenido a partir de celo Scotch de doble cara y heptano. Después de aproximadamente 5 minutos de deshidratación los embriones, sobre gel de sílice, se cubrieron con aceite Halocarbon 700 y fueron microinyectados en la región del

plasma polar usando un micromanipulador hidráulico (Leitz) cargado con aceite mineral libre de nucleasas y un microscopio de contraste de fases (Nikon Labophot).

Mapeo de la inserción

Los adultos supervivientes de la microinyección se cruzaron por moscas *yw;mb^{E16}/CyO* [cruce 1] con dos objetivos, a) comprobar si alguno de los gametos de estos individuos era portador de una inserción y b) cambiar el fondo genético sobre el que estaba la mutación *mb^{E27}* ya que distintas evidencias, como una disminución del número de embriones *mb^{E27}* homocigotos supervivientes, sugerían la existencia de segundas mutaciones en este cromosoma. De este modo, podíamos cambiar parte los cromosomas sin perder la condición de heterocigotos para *mb^{E27}*.

Los individuos de ojos coloreados descendientes del cruce 1 fueron cruzados por una cepa marcadora del cromosoma 2, *y w,lf / CyOy+* o *y w, Gla / CyOy+*, [cruce 2] con el fin de crear un stock para cada inserción. La localización cromosómica de cada inserción se determinó por el patrón de segregación de los marcadores según los resultados obtenidos en los cruces 1 y 2. En algunos casos, fue necesario un cruce adicional que consistió en cruzar los individuos de ojos naranjas del cruce 2 por una cepa *y w*.

2.7. Rescate de la letalidad embrionaria

Aproximadamente 40 hembras vírgenes de la cepa *y w, mb^{E27} / CyOy+;UAS-mb/C:GFP* se cruzaron con 30 machos *y w, mb^{E27}/CyOy+;da-GAL4* en un sistema de recogida de huevos. Tras un período de habituamiento de 24 h se procedió a un cambio de placa cada 20 h. Los huevos de cada placa retirada (de edad comprendida entre 0 y 20 horas) eran dispuestos en una nueva placa en pequeños grupos de entre 4-6 huevos para facilitar posteriormente el recuento. Pasadas ~24 horas desde la colocación de una placa se iniciaba el recuento de las larvas nacidas. La identificación del genotipo correspondiente al mutante homocigoto para *mb^{E27}* fue posible gracias al uso del marcador fenotípico *yellow* (*y*) que, en estadio larvario, provoca una falta de pigmentación en las mandíbulas que pasan de tener un color negruzco a amarronado. El número de larvas emergidas se anotó diariamente y se clasificaron como normales si poseían mandíbulas negras o como rescatadas si poseían mandíbulas amarillas. Las larvas normales se sacaban de la placa y se sacrificaban y las rescatadas se pasaban a una placa nueva para hacer un seguimiento del grado de rescate. El

recuento de las larvas nacidas (aquellas que conseguían salir del huevo) se realizó para dos temperaturas, 25°C y 19°C, y se calculó el porcentaje de embriones rescatados como el número de larvas nacidas sobre el número de larvas mutantes esperadas respecto del total de larvas nacidas.

2.8 Obtención de proteína MblA:GST

El plásmido pGEX-*mblA*:GST (Begemann, 1998) fue introducido en células competentes XL-Blue para inducir su expresión a gran escala. Para ello, se creció una colonia en 200 ml de LB con ampicilina y tetraciclina a 30°C durante toda la noche. A la mañana siguiente se realizó una dilución 1/10 de este cultivo en 100 ml y se indujo la expresión de la proteína de fusión con IPTG 0.1M dejándolo crecer durante 4 horas más a 30°C (D.O ~0.6). Las células se recogieron por centrifugación a ~6000 g durante 10 min., el pellet que se obtuvo se resuspendió en 10 ml de PBS con inhibidores de proteasas (Aprotinina y Leupeptina 10 µg/ml) y se sonicó en hielo con una punta de 8 milímetros de diámetro, 3 veces durante 30 s cada vez con intervalos de 1 min entre ellas. Después se añadió tritónX-100 al 1% y se mezcló por inversión evitando la producción de espuma. El sonicado se centrifugó a 10.000 g durante 10 min y el sobrenadante se mezcló con 1 ml de glutation sefarosa 4B al 50% en agitación suave durante 1 h a temperatura ambiente. La mezcla se centrifugó a 500g durante 5 min y el pellet de sefarosa se lavó 3 veces con 10 ml de PBS frío. La proteína MblA:GST unida se eluyó realizando 4 lavados de 15 min c/u en agitación suave con 600 µl de Tampón de Elución. La cantidad de proteína eluída se cuantificó mediante electroforesis en geles SDS-PAGE, tinción con solución de Coomasie y comparación con cantidades conocidas de proteína BSA.

2.9. Producción del anticuerpo anti-Muscleblind_{CUK333}

La elección del péptido para la producción del anticuerpo anti-muscleblind (α -Mbl_{CUK333}) se hizo en función de los distintos índices de antigenicidad aplicados (Hopp/Woods, Parker, Protrusion y Welling) y de su posición dentro de la proteína. De este modo, se eligió el péptido NH₂-VAQQKIPRSRLE-COOH ya que presentaba buenos valores de antigenicidad en todos los índices y no incluía ninguna secuencia con homología a otras proteínas que pudiera posteriormente darnos reacción cruzada.

El anticuerpo fue producido en conejo por la empresa CovalAb-Uk. Se realizaron cuatro inyecciones a lo largo de 53 días. El suero del último sangrado se inmunopurificó mediante una columna de afinidad del péptido inyectado. El anticuerpo recibido se guardó en alícuotas a -20°C en glicerol al 50% con azida al 0.02%.

2.10. Western blotting

La extracción de proteínas totales de embriones, larvas o adultos para análisis en western blot se realizó homogenizando en eppendorf el equivalente a ~100µl del material en 200µl de tampón de extracción (Tris-Cl 50 mM pH8) usando una mano de plástico para tubos de 1.5 ml tipo eppendorf. El homogenado se centrifugó a 14.000 rpm durante 5 minutos a 4°C para eliminar el material no homogenizado y los restos celulares. El sobrenadante se recuperó con cuidado y se repartió en alícuotas que se guardaron a -20°C hasta su utilización. Todo el proceso se realizó a 4°C o en hielo. La cantidad de proteínas totales de cada extracción se cuantificó mediante el método de Bradford.

Las muestras (30-50 µg) fueron desnaturalizadas en tampón de carga SDS 2X durante 3-7 minutos a 100°C (baño agua) y cargadas en geles SDS-PAGE al 10-12%. Las muestras se separaron mediante electroforesis durante 90 min a 12 mA en tampón de electroforesis (sistema Laemmli) y se transfirieron a una membrana PVDF en tampón de transferencia durante 1.5 h a 250 mA en sistema tipo tanque.

Para la inmunodetección, la membrana se equilibró en TBST y se bloqueó durante 1-2 h en solución de bloqueo que, dependiendo de la combinación de anticuerpos empleado consistió en leche desnatada al 5% o Caseína al 1% en TBST. La membrana fue incubada en anticuerpo primario diluido en solución de bloqueo durante una hora y lavada cuatro veces durante 10 min c/u con TBST. El anticuerpo secundario se diluyó en solución de bloqueo y se incubó entre 30-60 min seguido de 3 lavados de 10 min c/u. Todas las incubaciones se hicieron en agitación suave y a temperatura ambiente, excepto las incubaciones de >16h que se hicieron en cámara de 4°C La detección quimioluminiscente se hizo con el sustrato Lumi-light siguiendo las instrucciones del fabricante .

Para la detección cromogénica de la señal se sustituyeron los lavados con TBST por lavados con Caseína 1% en TBST, se incluyó un paso de amplificación de la señal

mediante la incubación en la solución ABC-HRP de Vector y se incubó con la solución de color hasta que la señal adquirió la intensidad deseada.

2.11. Obtención de material para IP

Para obtener grandes cantidades (miligramos) de embriones que expresasen la proteína de fusión MbIC:GFP se construyeron, en metacrilato, cajas de recogida de huevos de 11 cm de base por 20 cm de altura de modo que encajaran perfectamente con placas petri cuadradas (figura 2). Para cada cruce utilizamos entre 1000-1400 hembras vírgenes del genotipo *yw*, *mbI^{E27}/CyOy+*, *UAS-mbIC:GFP* y entre 700-1000 machos *w*; *da-GAL4* (inserción en el cromosoma III). Para recoger embriones con una edad comprendida entre 6 y 24 h se dejó que las moscas pusieran huevos en la placa durante 18 h a 25°C, después se retiró la placa y se mantuvo durante 6 h a la misma temperatura. Los huevos de varias placas se recogieron y lavaron en una cestilla. Se trataron con lejía comercial al 50% durante 3 min para decorionizar, se enjuagaron abundantemente con agua del grifo para eliminar cualquier resto de lejía, se transfirieron a un tubo eppendorf con ayuda de un pincel y se congelaron en nitrógeno líquido antes de almacenarlos a -80°C. Se recogió una media de 200 mg de embriones diarios por placa. Para disminuir el número de huevos no fecundados se dejó un período de habituamiento de al menos 24 h, y se comprobó sistemáticamente, bajo la lupa de fluorescencia, el porcentaje de embriones fecundados antes de cada recogida de embriones.



Figura 2. Cajas de recogida de huevos de *Drosophila melanogaster* a gran escala.

2.12. Preparación de los extractos para IP

Los extractos se prepararon homogenizando los embriones almacenados a -80°C en una proporción 1:2 (relación volumen de embriones y volumen de tampón) con tampón de homogenización al que se añadían inhibidores de proteasas. La rotura del material se realizó mecánicamente acoplado la mano de un homogenizador Potter-Elvehjem con capuchón de PTFE a un rotor mecánico. La homogenización se llevó a cabo hasta la completa desaparición de embriones enteros (~ 5 pulsos de 30 s. con descansos de 1 min entre ellos). Después se repartió el homogenado en tubos eppendorfs (~ 1 ml), y se centrifugó a 12.000 g durante 10 min a 4°C . Se recuperó el sobrenadante (~ 800 μl) y se pasó a un tubo eppendorf limpio que se almacenó inmediatamente a -80°C hasta su uso. Todo el proceso se realizó en cámara de 4°C o en hielo para minimizar la degradación de proteínas y RNA. Una pequeña alícuota de homogenado se guardaba para determinar la concentración de proteínas totales mediante el método de Bradford.

2.13. Unión del anticuerpo a la proteínaG-agarosa

Para una unión no covalente se incubó en 500 μl de PBS, 1 μg de anticuerpo anti-GFP (Roche, cat.1814460) con 30 μl de EZview proteínaG-agarosa al 50% previamente lavadas según instrucciones del fabricante. La incubación se llevó a cabo durante toda la noche (~ 16 h) a 4°C en un agitador tipo noria. Se añadió TritónX-100 al 0.01% para evitar que las *beads* quedasen pegadas a las paredes del tubo y mejorar así el mezclado. Posteriormente, se centrifugó a 16.000 g durante 2 segundos para recoger las *beads* y eliminar el sobrenadante que contenía el anticuerpo no unido. El pellet se lavó 3 veces con 1 ml de tampón de homogenización mezclando por inversión. Se centrifugó a 16.000 g 2 segundos y se retiró el sobrenadante. El pellet, que contiene las *beads* con el anti-GFP unido a la proteínaG-agarosa se mantuvo en hielo hasta su utilización.

Para unir covalentemente el anti-GFP a la proteínaG-agarosa se incubaron durante 30 min, a temperatura ambiente y en un volumen final de 300 μl , 250 μl de proteínaG-agarosa lavadas 6 veces con 1ml de PBS 0.1% tritónX-100 (volumen cama 125 μl) con 50 μg de anticuerpo anti-GFP. Se centrifugó y se retiró el sobrenadante que contenía el anticuerpo no unido. El pellet, que contenía la proteínaG-agarosa con el anticuerpo unido, se lavó dos veces con 1 ml de tampón borato y se incubó con 1ml de tampón

DMP preparado justo antes de ser usado. El pellet se enjuagó 3 veces con 1,5 ml de Etanolamina 0.2M pH8 y se mezcló durante dos h a temperatura ambiente. Se centrifugó y se retiró el sobrenadante. Para eliminar el anticuerpo unido no covalentemente, el pellet se preeluyó con 3 lavados de Glicina 0.1M pH 2.5 seguidos de 3 lavados del tampón de homogenización. Finalmente, la proteínaG-agarosa unida covalentemente al anticuerpo, se lavó con 1ml de PBS y se guardó al 50% en PBS con azida al 0.02% a 4°C.

2.14. Preabsorción de los extractos

Los extractos se descongelaron sobre hielo y se añadieron los inhibidores de proteasas. Se centrifugó a 16.000g durante 15 minutos y se recuperó el sobrenadante evitando tocar el pellet. Después se transfirió el homogenado a un eppendorf con un pellet de 30µl de proteínaG-agarosa al 50% previamente lavadas (según instrucciones fabricante). Se mezcló durante 30 minutos a 4°C en el agitador tipo noria y se centrifugó a 16.000g durante 2 segundos para eliminar todas las bolas. Se recuperó el sobrenadante dejando un pequeño volumen por encima para evitar tocar el pellet y se transfirió a un eppendorf que contenía un pellet de bolas proteínaG-agarosa previamente conjugadas con anti-GFP.

2.15 Coimmunoprecipitación de proteínas

Para cada inmunoprecipitación se incubaron entre 500 y 700 µl de homogenado, con 30 µl de EZView proteínaG-agarosa unida al anticuerpo (covalentemente o no) durante una hora y media a 4°C en la noria (velocidad mínima). Después de la incubación se centrifugó a 16.000 g durante 5 s a 4°C y se descartó el sobrenadante que contenía las proteínas no unidas. El pellet se lavó cuatro veces con 1 ml de tampón de homogenización frío con inhibidores durante un tiempo total de 30 min. Se mezcló por inversión y se mantuvo la muestra en hielo entre los lavados. Después se centrifugó a 16.000 g durante 2 segundos a 4°C y se lavó una última vez con PBS frío para eliminar los restos de TritónX-100 y otros componentes del tampón de homogenización que puedan interferir en la electroforesis. El pellet, que contiene las

proteínas inmunoprecipitadas, se guardó a -20°C o resuspendido en $20\ \mu\text{l}$ de tampón de carga SDS 2X.

Para una elución de las proteínas inmunoprecipitadas más suave, se usó un gradiente de glicina. Después de los lavados con tampón de homogenización, se hicieron dos lavados de 3-5 min c/u con $150\ \mu\text{l}$ de glicina 100mM $\text{pH}2.5$ seguido de otro lavado con $200\ \mu\text{l}$. Tras neutralizar el eluido con $200\ \mu\text{l}$ de Hepes 0.5M $\text{pH}7.6$ se centrifugó a 10.000g durante 5 minutos para eliminar *beads* que hubiera quedado en suspensión. El sobrenadante se transfirió a un tubo limpio y se precipitó con un volumen 1/10 de TCA al 80% en presencia de $20\ \mu\text{g}$ de aprotinina como *carrier*. El pellet de proteínas precipitadas se resuspendió en tampón de carga. Las proteínas inmunoprecipitadas separadas electroforéticamente, se visualizaron con tinción con plata.

En todos los casos el pellet de beads con el inmunoprecipitado se volvía a eluir con tampón de carga 2X. Este eluido era utilizado para la detección de la proteína MbIC:GFP mediante el anticuerpo $\alpha\text{-GFP}$.

Para la microsecuenciación de proteínas inmunoprecipitadas específicamente por la proteína MbIC:GFP se corrieron geles SDS-PAGE de tamaño estándar ($14\times 16\ \text{cm}$) que, además de las carreras de interés, contenían carreras con cantidades conocidas de marcador. El gel se selló dentro de una bolsa de plástico sumergido en solución fijadora y se envió al servicio de proteómica del Centro de Biología Molecular (CBM) Severo Ochoa (CSIC-Universidad Autónoma de Madrid). Las bandas de interés teñidas con SYPRO-Ruby fueron cortadas e identificadas mediante análisis MALDI-TOF.

RESULTADOS

1. ANÁLISIS ESTRUCTURAL Y RELEVANCIA CLÍNICA DE LOS GENES MUSCLEBLIND HUMANOS

La secuenciación de genomas completos ha puesto de manifiesto la gran conservación a nivel de secuencia de la mayoría de los genes. Además, un gran número de estudios funcionales demuestran la conservación de las distintas rutas y procesos en los que se organizan estos genes para llevar a término tareas celulares concretas, tales como señalización, regulación génica o respuesta a infecciones. Esta conservación de genes y rutas permite validar los resultados obtenidos en un organismo en otros diferentes, así como extrapolar datos de organismos modelo al estudio de, por ejemplo, enfermedades humanas. Como continuación del trabajo realizado en nuestro laboratorio en el marco de la caracterización del gen *muscleblind* en *Drosophila*, nos propusimos identificar y caracterizar los homólogos humanos de *muscleblind* así como comprobar su posible implicación en la patogénesis de enfermedades genéticas humanas.

1.1 *muscleblind*, UNA NUEVA FAMILIA DE GENES HUMANOS

1.1.1 Existen tres genes *muscleblind* en humanos

El primer cDNA *muscleblind* humano, Y13829, se identificó y depositó en las bases de datos gracias a la colaboración con el Dr. Giovanni Borsani perteneciente al TIGEM Telethon Institute of Genetics & Medicine (Borsani y Pérez-Alonso, no publicado). El rápido avance del Proyecto Genoma Humano y una búsqueda sistemática de nuevas secuencias relacionadas con el cDNA Y13829 permitió identificar los elementos necesarios para la caracterización de todos los genes humanos.

Hemos identificado tres genes homólogos a *muscleblind* en el genoma humano que, de acuerdo con la nomenclatura actual, nombraremos como *MBNL1*, *MBNL2* y *MBNL3* (Figura 1). El gen *MBNL1*, el primero que identificamos, abarca más de 160 kb de DNA genómico. Se localiza en el cromosoma 3, concretamente en 3q25, entre los marcadores D3S1279 y D3S1280. El gen *MBNL2* se extiende a lo largo de 120 kb aproximadamente y se localiza en 13q31, entre los marcadores D13S1494 y SGC38389. Por último, el menor de los tres genes, *MBNL3* cubre 110 kb aproximadamente y mapea en Xq25 entre los marcadores DXS1047 y DXS998.

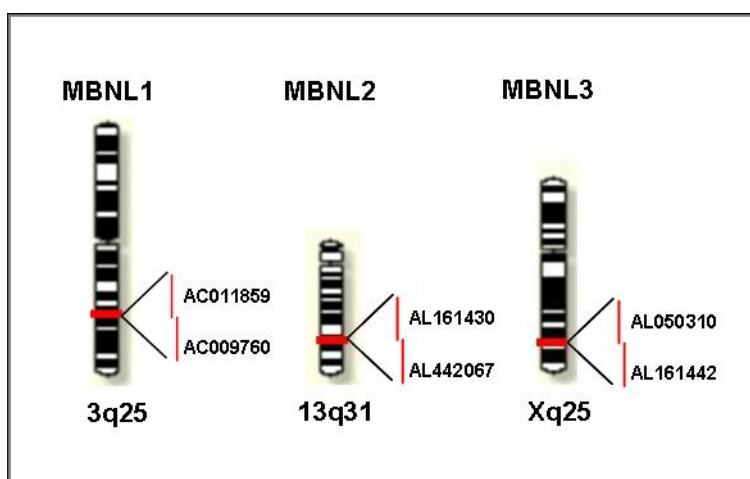


Figura 1. Existen tres genes *muscleblind* en el genoma humano *MBNL1*, *MBNL2* y *MBNL3*. Se muestra la localización cromosómica de *MBNL1*, *MBNL2* y *MBNL3* así como los números de acceso (GenBank) de los DNA genómicos que contienen el conjunto de los exones codificantes de cada uno de los genes

El mayor porcentaje de identidad nucleotídica se da entre los cDNAs de los genes *MBNL1* y *MBNL3* con un 57% siendo el menor entre *MBNL2* y *MBNL3* con un 41% (mejor alineamiento global entre las secuencia de nucleótidos de las regiones codificantes). Respecto a los cDNAs de *muscleblind* en *Drosophila*, la mayor identidad nucleotídica se da entre la región codificante de la isoforma *mbIc* y el gen humano *MBNL2* con un 50%. Los genes *MBNL1-3* codifican proteínas con cuatro dedos de zinc del tipo Cys₃H. La mayor identidad de secuencia entre las proteínas Muscleblind humanas se da entre la proteína *MBNL1* y *MBNL2* con un 77% siendo la menor entre la proteína *MBNL1* y *MBNL3* con un 68%. Respecto a *Drosophila*, todas las proteínas humanas tienen porcentajes de identidad similares; el mayor de ellos es entre *MBNL1* y la isoforma *Mblb* con un 33% de identidad y un 42% de similaridad (Figura 3)

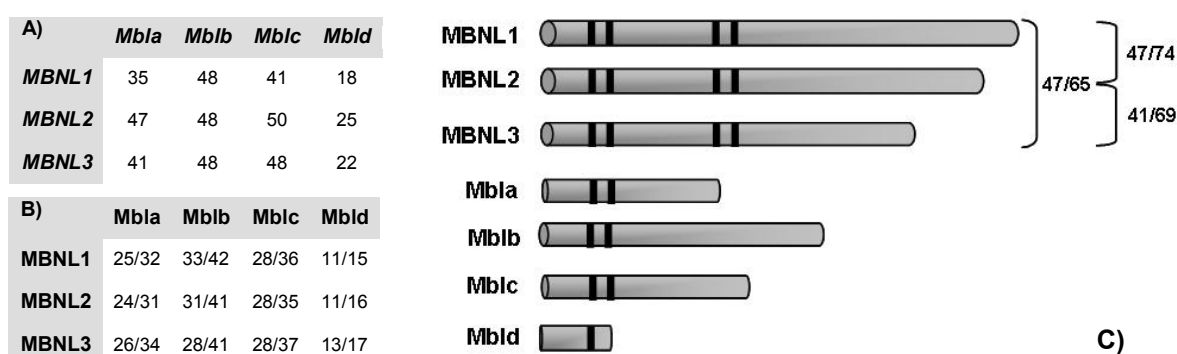


Figura 2. Comparaciones de secuencia entre *muscleblind* y *MBNL1-3*. (A) Se indica el porcentaje de identidad nucleotídica entre las regiones codificantes de los cDNAs de *MBNL1* (AJ308400), *MBNL2* (AF061261) y *MBNL3* (AK002178). (B) Se indica el porcentaje de identidad /similaridad entre las proteínas *MBNL1-3* y las isoformas de *Muscleblind*. *MBNL1-3* tienen, respecto a las distintas isoformas de *muscleblind*, porcentajes de identidad similares tanto a nivel de nucleótidos como de aminoácidos. La isoforma de *mbl* que presenta una identidad tanto nucleotídica como aminoacídica más alta con los tres genes humanos es *mblb*. (C) Las proteínas *MBNL1-3* poseen cuatro dedos de zinc (barras negras) mientras que *Mbl* de *Drosophila* contiene sólo dos. Se indica el porcentaje de identidad de secuencia entre los cDNA y las respectivas proteínas de *MBNL1-3*.

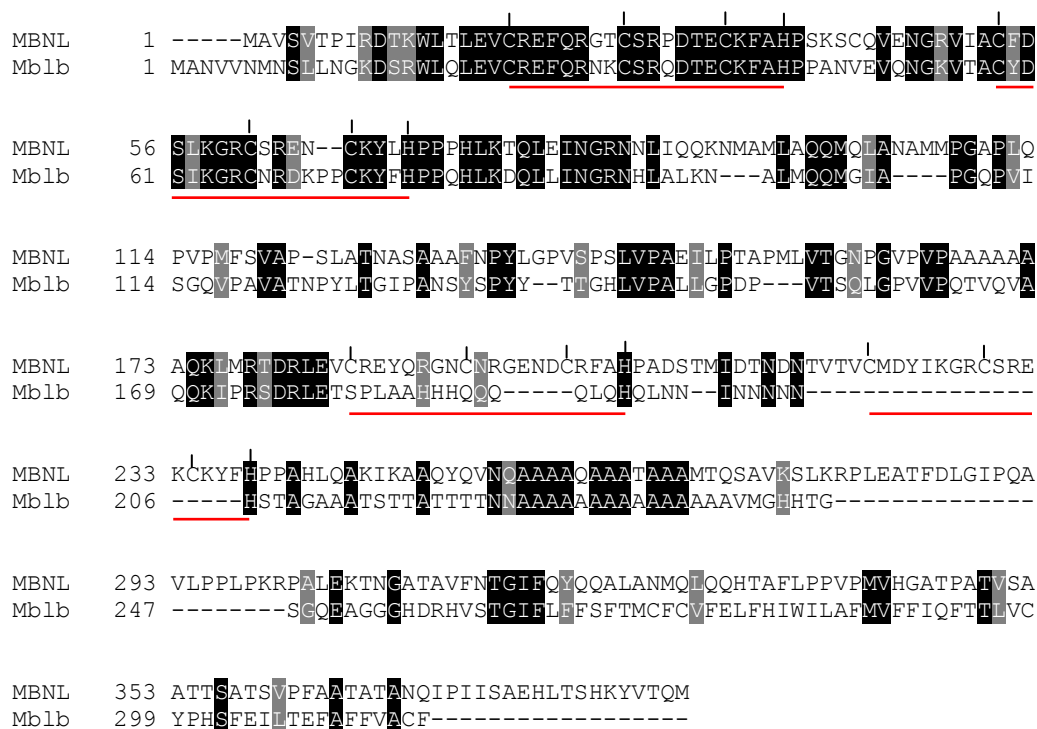


Figura 3. Alineamiento de las proteínas MBNL1 y Mblb realizado con el programa GAP del GCG *package program* y el programa Boxshade en <http://www.ch.embnet.org>. El alineamiento de la secuencia completa de las proteínas MBNL1 (BAA24858) y Mblb (AAC47758) muestra una identidad de secuencia de un 33% y una similaridad del 42% . Ambas proteínas se caracterizan por poseer dedos de zinc del tipo Cys₃H señalados con barras horizontales. MBNL1 posee cuatro dedos de zinc mientras que Mblb posee sólo los dos primeros que aparecen totalmente conservados.

1.1.2 La estructura genómica está altamente conservada en los tres genes humanos

Conocer la estructura exón-intrón de un gen es un paso previo y necesario a la búsqueda de mutaciones. Además, la comparación de la estructura génica entre distintos genes homólogos ayuda a determinar la relación filogenética que existe entre ellos. Por todo esto, determinamos con exactitud la estructura exón-intrón de los tres genes *muscleblind* humanos.

La estructura genómica de los genes MBNL1-3 pone de manifiesto la gran conservación que existe en cuanto a la composición exónica de los mismos. En los tres genes se pueden identificar fácilmente la equivalencia entre los exones codificantes de cada uno (Figura 4A). La mayoría coinciden exactamente en tamaño siendo esto especialmente evidente en los exones 5, 6, 7 y 8 de la región 3'. Cuando

el tamaño no es el mismo la diferencia no es nunca superior a 12 nt, excepto para el exón 4 de *MBNL3* que con 27 nt presenta la mayor diferencia en tamaño con el respectivo exón de *MBNL2*. Cabe resaltar la no existencia de ciertos exones equivalentes a los hallados para *MBNL1* en los otros dos genes humanos como es el caso de los exones 9 y 10 para *MBNL2* y los exones 5 y 9 para *MBNL3*. Este hecho se ha comprobado realizando búsquedas, mediante tblastn, del DNA genómico de cada uno de los genes que codifican las secuencias peptídicas correspondientes a los exones opcionales. Además, hemos descartado la existencia de ESTs específicos de cada uno de los genes que contuvieran dichos exones.

Respecto al loci *muscleblind* de *Drosophila*, la estructura exón-intrón de los genes humanos no está conservada ya que no se puede establecer una correlación evidente entre los exones de los genes humanos y los de la mosca.

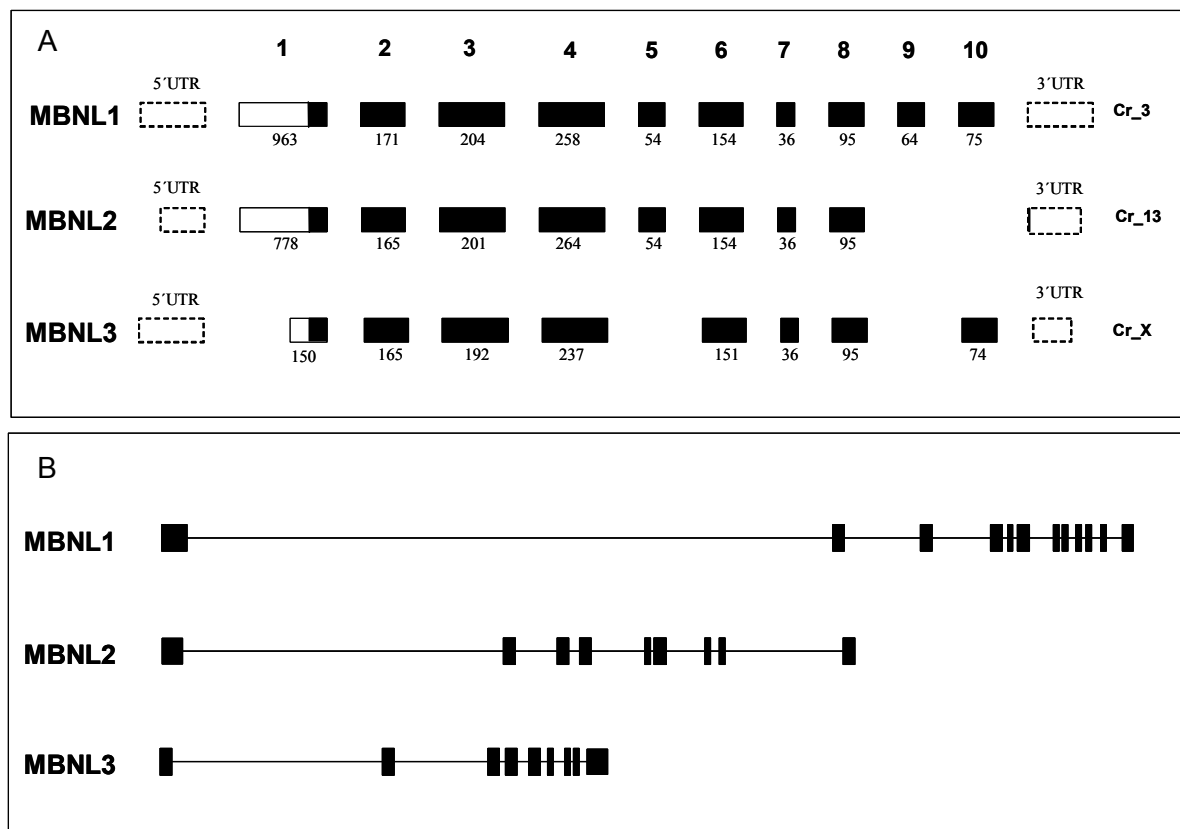


Figura 4. ver pie de figura página siguiente

Figura 4. Organización genómica de los genes *muscleblind* humanos. A) La composición exónica de todos los genes está altamente conservada excepto por la ausencia de los exones 9 y 10 en *MBNL2* y los exones 5 y 9 en *MBNL3*. Los exones se numeran a partir del primer exón codificante tomando *MBNL1* como referente. El tamaño de los exones está representado a escala pero no el de los intrones. El sombreado negro indica región codificante. Las cajas con borde punteado representan las regiones 5' y 3' no traducida (UTR) para cada gen. B) Visión genómica de los loci *muscleblind* humanos. No existe una clara conservación a nivel del tamaño de los intrones excepto por la presencia de un gran intrón entre el exón 1 y el 2. Las representaciones se han hecho proporcionales a las distancias en pb aunque debido a las grandes diferencias en el tamaño de los distintos intrones dentro de un mismo gen no se ha podido dibujar a escala.

Los tres loci *muscleblind* humanos se extienden varias kilobases sobrepasando las 100 kb en el caso de *MBNL1* y *MBNL2*. Aunque el tamaño de los intrones no está conservado entre los tres genes, cabe resaltar que en todos ellos existe un intrón 1-2 de gran tamaño que cubre más de la mitad de la extensión total del gen (Figura 4B), esto también ocurre en el locus del gen *muscleblind* de *Drosophila*.

El análisis *in silico* de las regiones 5'-UTR (UTR "untranslated regions") puso de manifiesto que los tres genes humanos contienen el ATG de inicio en un exón que es codificante sólo en sus últimos 174 pb (muy similar a *Drosophila* con 189 pb). Este primer exón codificante (exón 1, Figura 4A) es de gran tamaño y contiene gran cantidad de nucleótidos no codificantes. Esto es especialmente cierto para *MBNL1* y *MBNL2* donde el exón 1 tiene un tamaño de 963 y 778 pb, respectivamente. En *MBNL3* el exón queda reducido a 256 pb. Para los tres genes humanos, igual que para el de *Drosophila*, la región 5'-UTR está separada en 2 exones hecho que sólo ocurre en un 28% de los genes humanos (Pesole, 2000).

Por otro lado, los tres genes *muscleblind* tienen regiones 3'-UTR de varias kilobases resultando un tamaño superior a la media de los genes humanos (1027 pb). La región 3'-UTR más larga es la de *MBNL1* con 3673 pb seguida de *MBNL2* con 2798 pb y, por último, *MBNL3* con 1526 pb. Cabe resaltar que todas las regiones 3'-UTR de las isoformas del locus *muscleblind* en *Drosophila* tienen tamaños del mismo orden que los genes humanos siendo esto una característica poco común en invertebrados ya que el tamaño medio de sus regiones 3'-UTR suele ser de 444 pb. Las regiones *trailer* de gran longitud se relacionan con genes que son regulados a nivel de la estabilidad de su mRNA (Mignone, 2002).

Se pueden definir tres tipos de genes homólogos en función de su origen evolutivo: Ortólogos, Parálogos y Xenólogos. Los genes parálogos son aquellos homólogos producidos por duplicación génica y representan genes derivados a partir de un gen ancestral común. Nosotros proponemos que *MBNL1*, *MBNL2* y *MBNL3* son genes parálogos porque se parecen más entre sí que con respecto a *muscleblind*. Además, la conservación de la estructura genómica indica que los sucesos de duplicación génica son relativamente recientes o que existe una presión de selección para mantenerla. Una razón que determina el mantenimiento de la estructura exón-intrón es el mantenimiento de los patrones de *splicing* alternativo. Los genes parálogos tienden a tener funciones diferentes aunque esto va a depender, entre otras cosas, del tiempo transcurrido desde la duplicación que los originó.

Tabla 1. Información detallada sobre la estructura genómica de los genes *muscleblind* humanos

MBNL1

Exón	Tamaño Exón	cDNA ^a	Sitio aceptor	Sitio dador	Intrón	Tamaño Intrón
0	1053 [†]	1-626		AATGTGAGAG gtgagaggta	0-1	26.5
1	963	627-1588	ttttetacag GTTGGTACTA	TCATTGAAAG gtgagtaact	1-2	114.3
2	171	1589-1759	gtttetctag GGCCGTTGCT	ACAACCCGTG gtaagcatgt	2-3	17.6
3	204	1960-1963	tatttttcag CCAATGTTTT	CAGACTTGAG gtaggaatga	3-4	12.3
4	258	1964-2221	gacctcatag GTATGTCGAG	AGCTGCCATG gttgagtaga	4-5	1.1
5	54 ^{&}	1319-1373	gtaattaaag ACTCAGTCGG	CTTTGACCTG gtaactatgac	5-6	1.0
6	154	2222-2375	cactaaacag GGAATTCCTC	CTCCCACCAG gtaggggggg	6-7	7.8
7	36 [§]	991-1026	tgatgggcag GCTCAATATT	ACAAGTGTG gtagggtcct	7-8	0.7
8	95	2376-2470	cetcccaaag TTCCCATGGT	AGCCAACCAG gtttgctaat	8-9	1.8
9	64 [§]	213-277	gtttcgacag CCCCATCCTT	AGGAAAAATG gtaggagaag	9-10	1.0
10	75	2471-2545	tatttgctag ATACCCATAA	TCACTAAACA gtaagttcat	10-11	3.0
11	3673*	2546-2940	tecctttag ATCATGCTAA			

^acoordenadas del mensajero AB007888 excepto [&] que se corresponde con Y13829, [§] que se corresponde con AJ308400 y [§] que se corresponde con AF255334.

[†]máxima longitud de la 5'UTR según NM_021038.3

*máxima longitud de la 3'UTR según NM_207293.1

MBNL2

Exón	Tamaño Exón	cDNA ^a	Sitio aceptor	Sitio dador	Intrón	Tamaño Intrón
0	212 [†]	1-178		GCTAAACAAG gtaggagaat	0-1	53.0
1	778	179-956	tyttttgcag GTTTATCATT	TTCCCTAAAG gtaagagaat	1-2	57.9
2	165	957-1121	gtttttacag GGCCGTTGTT	TCATCCAGTG gtgagtacat	2-3	8.6
3	201	1122-1323	cetttaacag CCCACTTTCC	CAAACCTGGAG gtaactteaat	3-4	3.6
4	264	1323-1549	cctgtcttag GTATGCAGGG	CACAGTCATG gtaagtgcgg	4-5	9.7
5	54	1550-1603	ttattaaaag ACTCAGTCGA	TGTAGACCTG gecttteccc	5-6	0.7
6	154	1604-1757	teatttctag GCCTTCCCC	ATTCCAACAG gtagtgccec	6-7	7.5
7	36	1758-1793	atgacatcag GGTCAATTTT	ACCAGTATTG gtaggtttca	7-8	1.3
8	95	1794-1888	ctcctgaaag TACCCATGAT	AGCCAATCAG gtttgcctct	8-11	24.8
9	-	-	-	-	-	-
10	-	-	-	-	-	-
11	2798*	1889-2577	tctttaacag ATAATTCTGA			

^acoordenadas del mensajero AF061261

[†]máxima longitud de la 5'UTR según AF491866.1

*máxima longitud de la 3'UTR según CR749802.1

MBNL3

Exón	Tamaño Exón	cDNA ^a	Sitio aceptor	Sitio dador	Intrón	Tamaño Intrón
1	256	1-256		TTCTCTAAAG gtgagtcceca	1-2	33.0
2	165	257-421	tcattgtag GGTGGTGTA	GTCACTACTT gtaagtcgca	2-3	13.9
3	192	422-614	aacatacag GGTCTTTTTC	TAAACTGGAG gtatttgtga	3-4	1.0
4	237	615-850	gtgttotaag GTTTGCCGAG	CTCTGCCATG gtaagaaaag	4-6	4.0
5	-	-	-	-	-	-
6	151 [†]	596-746	cccccaacag GCCCTGCAGC	ATCCCTGCAG gtgagaacac	6-7	1.9
7	36	897-932	ggtgtacag GGCCAATACT	TCAAATATTG gtaggttttc	7-8	2.4
8	95	933-1027	caactgaaag TGCCCATGAT	AGGCAATCAG gtttggccat	8-10	0.7
9	-	-	-	-	-	-
10	74	1028-1101	actactatag ATACCCCAAT	CAGTAGAACA gtaagtttct	10-11	1.7
11	1527*	1102-1575	tcattttag CTGAAATTCT			

^acoordenadas del mensajero NM_133486.1 excepto [†] que se corresponde con AL133625

*máxima longitud de la 3'UTR según NM_0183882

1.1.3 Los tres genes *muscleblind* humanos sufren un complejo patrón de corte y empalme alternativo

El análisis de los ESTs y cDNAs disponibles en las bases públicas muestra que los tres genes humanos están sujetos a un complejo patrón de *splicing* alternativo que afecta principalmente a los exones del extremo 3' de cada uno de los loci. Actualmente, existen nueve mensajeros completos secuenciados para *MBNL1*, tres para *MBNL2* y cuatro para *MBNL3* que representan diferentes isoformas proteicas (Figura 5A). De manera análoga a lo que ocurre con el locus *muscleblind* de *Drosophila*, prácticamente todos los sucesos de procesamiento alternativo implican el extremo carboxi terminal de las proteínas humanas afectando a los exones 5, 7, 8, 9 y 10.

En *MBNL1* se encuentra una forma de procesamiento que afecta a la parte N-terminal de la proteína de manera que excluye el exón 3. Este exón constituye la región espaciadora entre los dos pares de dedos de zinc que codifican las proteínas Muscleblind humanas. Estudios de deleciones en experimentos de tres híbridos en levaduras demuestran que la región "linker" entre los dos pares de dedos de zinc es necesaria para mantener las propiedades de unión a RNA de la proteína MBNL1 (Kino, 2004). Esta variante de procesamiento está conservada en ratón y podría representar el equivalente funcional humano a la isoforma D de *Drosophila*. La isoforma *mbID* utiliza un exón opcional en el intrón 1-2 del gen lo que produce una proteína de reducido tamaño al que le falta el segundo dedo de zinc. Ambas isoformas generan proteínas que se espera que tengan actividades distintas al resto de las isoformas de sus respectivos genes.

El procesamiento alternativo que se observa para los tres genes permite extraer algunas observaciones generales. Para todos ellos se observa que los exones 2 y 4 son constitutivos ya que están siempre presentes. Estos exones contienen la segunda mitad del segundo dedo de zinc y el último par de dedos de zinc, respectivamente. El exón 5 no existe en *MBNL3* y es relativamente poco frecuente en *MBNL1*. El exón 7 suele ser constitutivo en los transcritos *MBNL3* y muy frecuente en *MBNL2*. El exón 8 es muy frecuente en *MBNL1* y *MBNL3*. El exón 9 sólo existe en *MBNL1* y normalmente no se incluye en los transcritos maduros. El exón 10, que contiene el codón de parada, sólo existe en los genes *MBNL1* y *MBNL3* y está presente en casi todos los transcritos de ambos genes. En los ESTs donde se ha comprobado la no inclusión del exón 10 y en los transcritos del gen *MBNL3*, donde no existe equivalente,

el codón de parada aparece al principio del exón 3'UTR por lo que, en estos casos, este exón codifica entre tres y cuatro aminoácidos. La presencia/ausencia de los exones 5 y 7 no altera la pauta de lectura mientras que el procesado de cualquier otro exón si que lo hace. En general, se observa que aquellas variantes que resultan en una alteración de la pauta de lectura son menos frecuentes que las variantes "en pauta". Por ejemplo, la falta del exón 6 provoca un corrimiento de la pauta de lectura que es posteriormente corregida por la incorporación del exón 9 de manera que el mRNA es portador de una proteína con 54 aminoácidos diferentes (aunque similares a los que le falta) pero con un extremo C-terminal idéntico al resto. Entre los mRNAs completos del gen *MBNL3* podemos observar que existen algunas variantes que usan un sitio aceptor alternativo 3' en el exón 6. Esto provoca la pérdida de 34 aminoácidos en estas proteínas sin que se produzca un corrimiento de la pauta en el resto de la secuencia codificante.

Aparte de los mRNA completos, también se analizaron todos los ESTs existentes para cada uno de los genes encontrándose formas de *splicing* que no están representadas en ninguno de los mensajeros completos conocidos hasta la fecha (Figura 5B). Además, para *MBNL1*, se identificaron y secuenciaron tres clones cuyos ESTs indicaban que podían contener nuevas variantes de procesado (id_IMAGE: 511793, 509472, 953096). La secuenciación de estos tres clones dió como resultado la identificación de una variante ya conocida, representada por el mensajero completo AB007888 (clon IMAGE: 509472), y dos nuevas variantes que eran visiones parciales del mensajero por lo que no fueron depositadas en las bases de datos públicas. Una de ellas (clon IMAGE: 953096) coincide con el mRNA completo que fue posteriormente depositado en las bases de datos como NM_207296. La otra variante (clon IMAGE: 511793) contiene un exón 6 mas largo por uso de un sitio dador alternativo, de manera que se extiende, más allá de las 154 pb habituales, hasta 2481 pb donde finaliza el transcrito.

A)

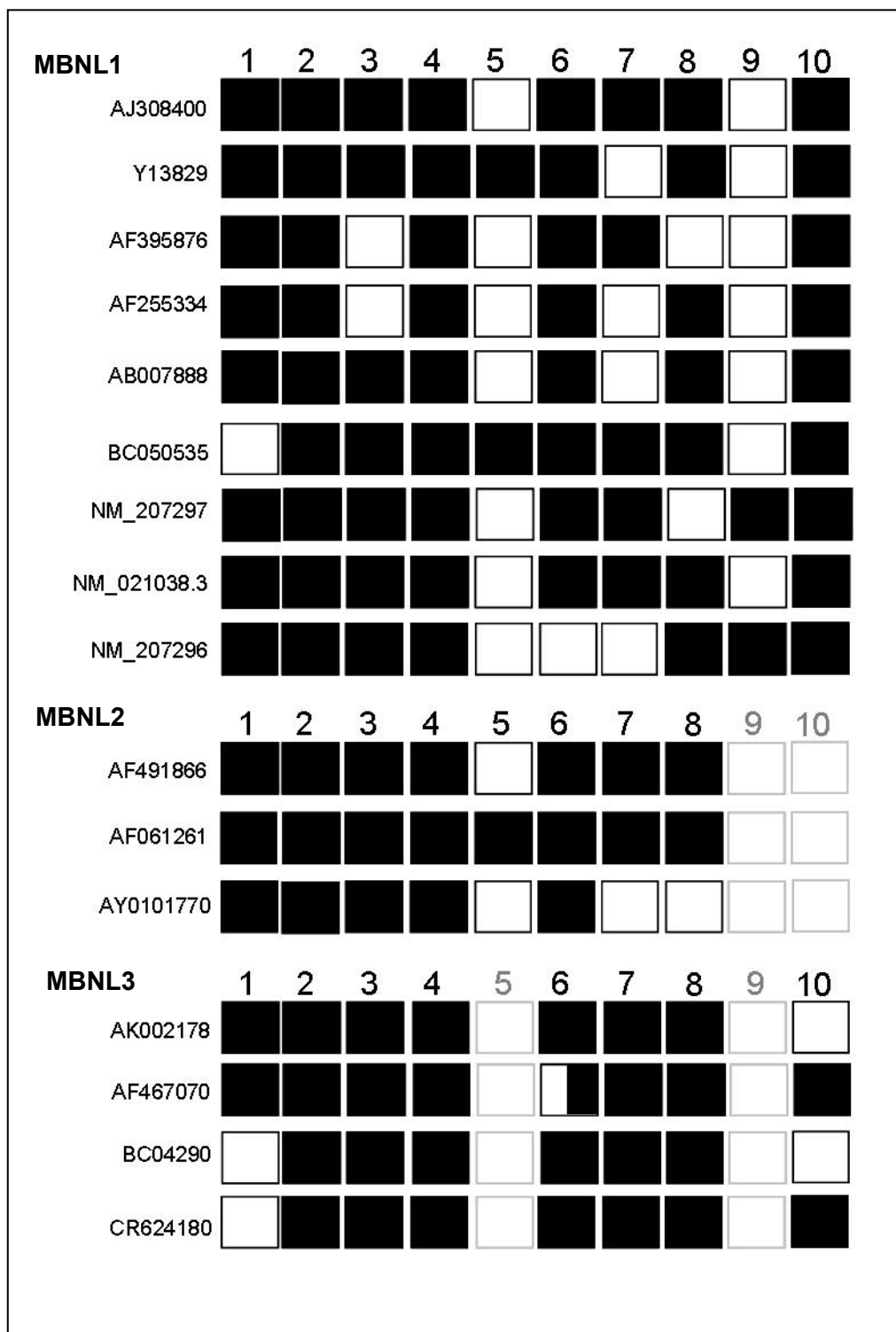


Figura 5A. Ver pie de figura en página siguiente

B)

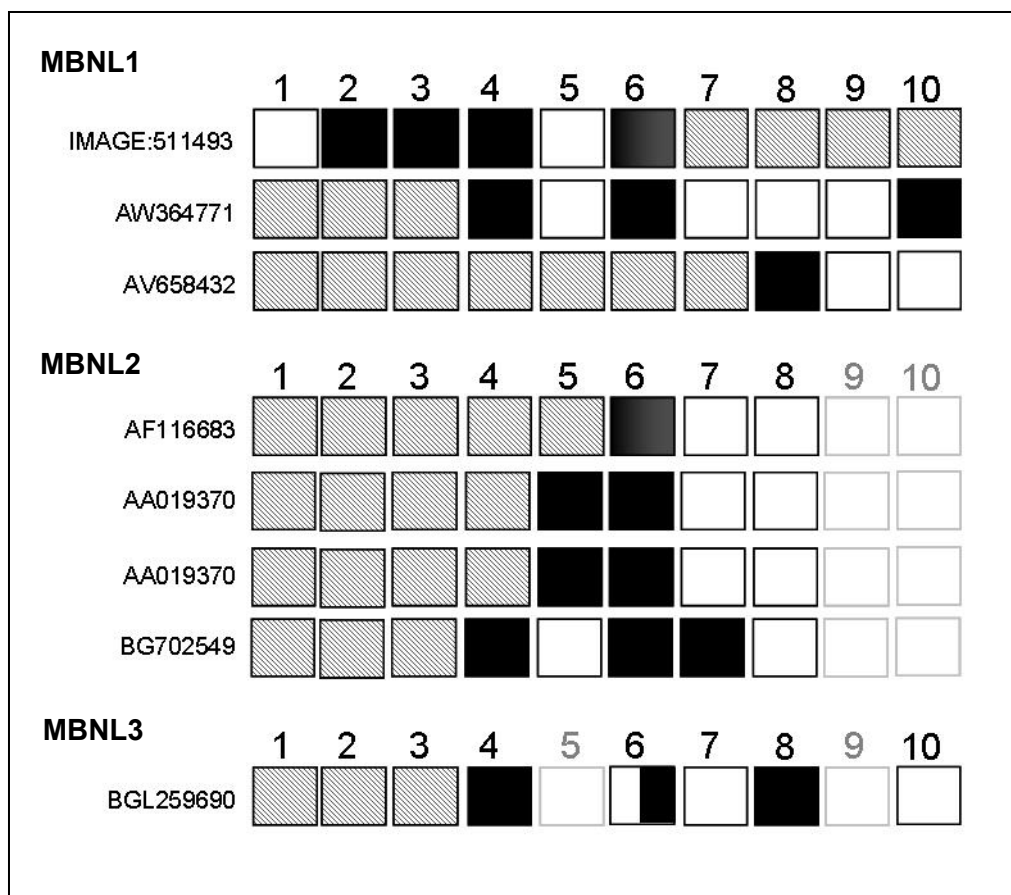


Figura 5. Procesado alternativo de los genes *muscleblind* humanos. Se representa la composición de exones codificantes de los mensajeros para el gen MBNL1, MBNL2 y MBNL3. A) Representación de los transcritos completos para cada uno de los genes. Se puede observar que el procesado alternativo afecta principalmente al exón 5 y 7 en todos los genes. Los mensajeros que sufren procesado alternativo del exón 1 pierden del ATG de inicio y el primer dedo de zinc por lo que las proteínas presentan extremos N-terminal diferentes. B) Representación de los transcritos que no cubren totalmente la ORF de los genes que aportan evidencia de nuevas formas de procesado alternativo. La numeración de los exones se corresponde con el de la Figura 4B. Las cajas representan exones. Caja negra, presencia del exón; caja blanca, ausencia del exón; caja a rayas, no hay información sobre su presencia o ausencia; caja sombreado degradado, exón con un sitio aceptor alternativo que lo extiende más allá. Cuando en alguno de los genes no existe la secuencia exónica equivalente a alguno de los exones numerados para MBNL1 se indica con los números y las cajas en gris.

Para valorar si se trataba de una forma de *splicing* funcional o si por el contrario se trataba de una forma aberrante de procesado, se realizaron búsquedas de otros ESTs que alinearan con las coordenadas de DNA genómico que cubrían el exón 6 “extendido”. Sólo se obtuvieron secuencias en el extremo 3’ final y en gran parte de estos ESTs se indicaba la existencia de secuencias *Alu*. Por otro lado, la traducción de este mensajero origina 57 nuevos aminoácidos a partir del aminoácido 282 de la proteína MBNL1 humana (AAH50535) y la búsqueda de homología con otras proteínas dió resultados negativos. Por todo esto, y pese a tratarse de un transcrito poliadenilado

(19 adeninas al final de la secuencia que no alinean con el DNA genómico) concluimos que debe tratarse de un forma aberrante de procesado producida por la presencia de un sitio aceptor en el elemento *Alu* próximo (revisado en Kreaehling, 2004 y Krull, 2005).

Nuevos exones en MBNL1

Diferentes ESTs denotan la existencia de tres nuevos exones: 1a, 1b y 1c, dentro del intrón 1 de MBNL1 (Figura 6). El exón 1a se localiza a 1.5 kb del exón 1 y es revelado por el EST AL562860. Tiene una extensión de 530 pb pero que se extiende hasta al menos 965 pb por otros ESTs solapantes como CR603447 y BF126161. Esta variante de procesado generaría una proteína de 83 aminoácidos (MBNL1_1a) que compartiría sólo los 58 primeros aminoácidos con las proteínas habituales de MBNL1 y un extremo C-terminal diferente con una región 3'-UTR diferente a la del resto de las variantes. Para comprobar la relevancia de este nuevo exón comprobamos si estaba conservado en otros organismos. Encontramos ESTs de ratón y vaca que alineaban en las mismas coordenadas de DNA genómico y que generan isoformas proteicas similares aunque las secuencias C-terminal no están conservadas. Esta nueva variante de procesado en humanos tiene un extraordinario paralelismo con la isoforma Mbld de *Drosophila* ya que ambas contienen tan solo el primer dedo de zinc (Figura 7).

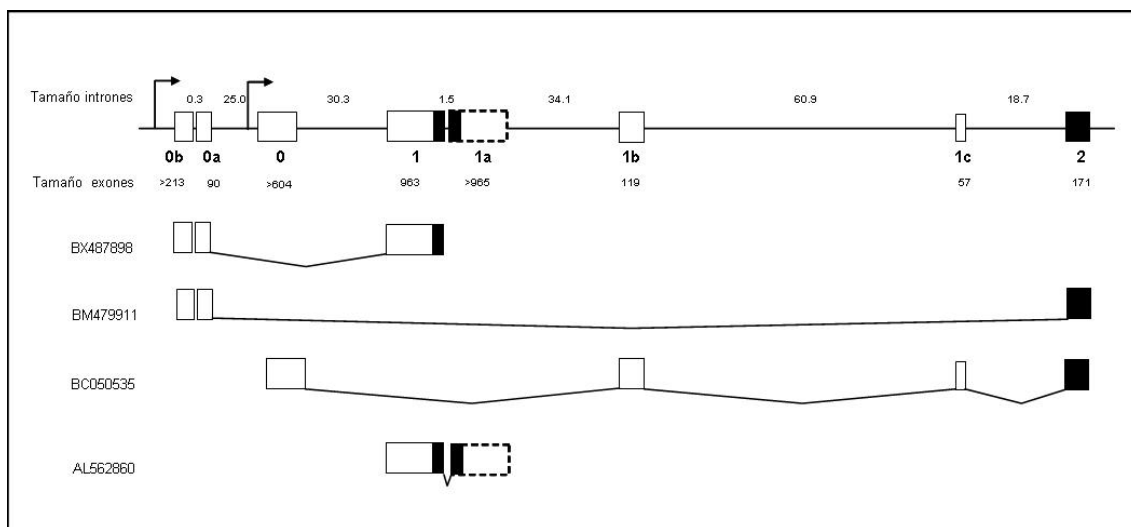


Figura 6. Nuevos exones para *MBNL1*. Además de los exones habituales, existen otros exones en *MBNL1* que son menos frecuentes. Dentro del intrón 1-2 se localizan los exones 1a, 1b y 1c que son evidenciados por los mensajeros AL562860 (1a que es extendido por CR603447 y BF126161) y BC0505335 (1b y 1c). Además existen dos exones 5'-UTR opcionales al ya conocido (exón 0) que se encuentran separados por 300 pb y que siempre se procesan juntos. Cabe resaltar que el exón 0b contiene DNA repetitivo. Las cifras que aparecen en el dibujo indican, de línea superior a línea inferior, exón, tamaño del exón en pb y tamaño del intrón. El sombreado en negro indica región codificante.

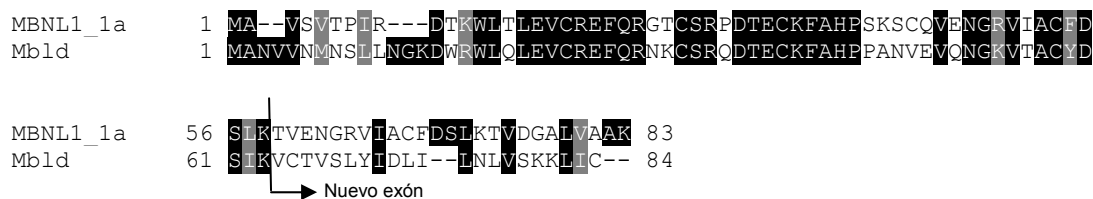


Figura 7. Alineamiento entre las isoformas MBNL1_1a humana y Mbl d de *Drosophila*. El uso de un exón alternativo en el intrón 1-2 del gen *MBNL1* y *mbi* produce proteínas de menor tamaño que las habituales, 83 y 84 aminoácidos, respectivamente. Ambas proteínas sólo contienen el primer dedo de zinc y tienen un extremo carboxi terminal de 25 (MBNL1_1a) y 21 (Mbl d) aminoácidos. Pese a tener una longitud similar, no existe conservación a nivel de secuencia entre los extremos carboxi terminal de la isoforma humana y de la mosca.

Por otro lado, el mensajero completo BC050335, que no contiene el exón 1, pone de manifiesto los exones 1b (119 pb) y 1c (57 pb). Esta variante genera una proteína Muscleblind a la que le falta la región N-terminal y que empieza a nivel del segundo dedo de zinc de la proteína. Estos nuevos exones 1b y 1c están representados por otras secuencias, como BQ2163772b y F00133, aunque su uso es minoritario. En este sentido, cabe resaltar que el exón 1c contiene secuencias de DNA repetitivas, en concreto, el elemento MER1. La presencia de este exón en transcritos procesados podría ser minoritaria debido al corte y empalme alternativo al que selectivamente se encuentra sometido este exón dada la presencia del elemento repetitivo. Este suceso está bien documentado para los elementos *Alu* (Sorek, 2002).

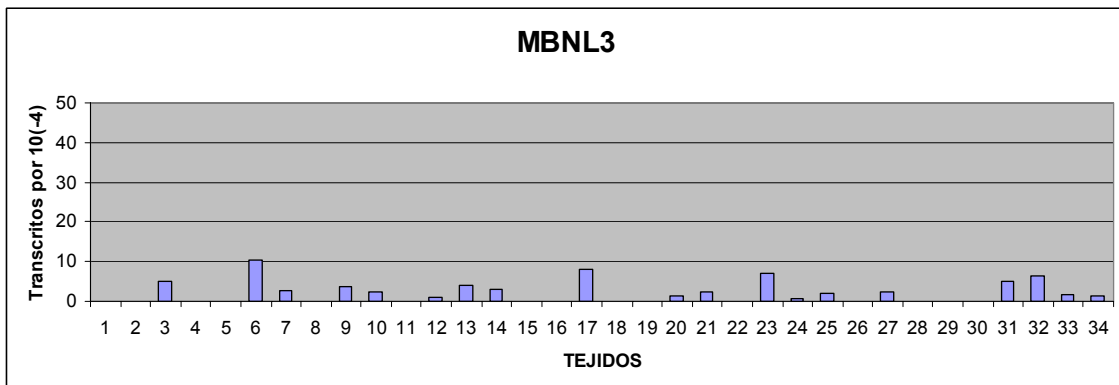
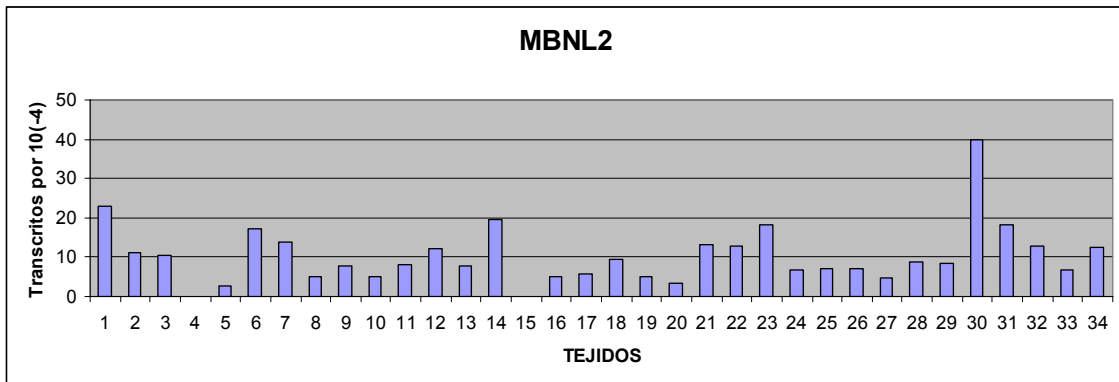
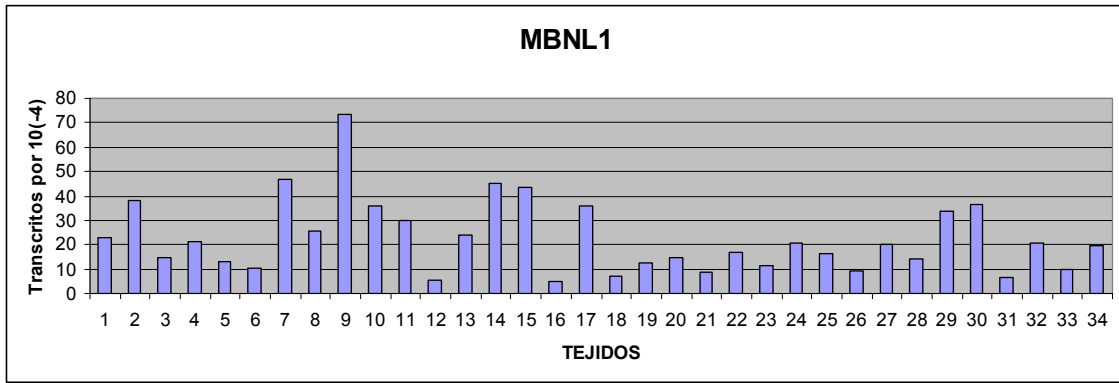
Además de los exones 1a, 1b y 1c dentro del intrón 1-2 de *MBNL1* hemos identificado dos exones adicionales en la región 5' del gen que llamamos 0a y 0b. Algunos de los mensajeros que lo evidencian son BX48798 y BM479911. La presencia de los exones 0a y 0b en un transcrito es incompatible con la presencia del exón 0 ya que no se encuentra ninguna secuencia que porte ambos exones a la vez. Proponemos por tanto, que se trata de dos sitios de inicio de la transcripción alternativos para el locus *MBNL1*.

1.1.4. Expresión diferencial de los tres genes humanos *muscleblind*

Durante el análisis de los cDNAs y ESTs de los genes *muscleblind* humanos, se observó que procedían de un gran número de tejidos diferentes. Con el fin de obtener información “*in silico*” acerca del patrón espacial de expresión de estos genes, examinamos el tejido de origen de todas las secuencias de expresión para el gen

muscleblind correspondiente, así como su número respecto al total de secuencias existentes para cada tejido. De esta manera podemos describir en qué tejidos se expresan los genes y con qué frecuencia relativa lo hacen. De acuerdo con el número y origen de los ESTs de cada uno de los genes podemos establecer que *MBNL1* es el que se expresa más abundantemente, tanto en cuanto al número de tejidos diferentes como en cuanto a la cantidad de transcritos identificados (Figura 8). Según la frecuencia relativa de los ESTs de los diferentes tejidos, *MBNL1* se expresa abundantemente en células endoteliales (73,3%) y en la musculatura esquelética (músculo 46,7%) y lisa (útero 40,0%, vejiga 23%) así como en tejidos de origen linfático (nódulo linfático 45,2%, sangre 35,7%, médula ósea 30%). El cerebro (5,3%), y el bazo (5,1%) serían tejidos donde cabría esperar una menor expresión de *MBNL1*. La expresión de *MBNL2* comparte gran cantidad de tejidos con *MBNL1* aunque en relación a *MBNL1* está más reducida. Los tejidos más representados para *MBNL2* son la laringe (39,8%) y la vejiga (23%), ambos con un gran contenido en musculatura lisa. En contraste, *MBNL3* sería el gen con la expresión más limitada de los tres ya que se expresa en un número muy reducido de tejidos y en una proporción mucho menor que los otros dos genes. Según las frecuencias relativas de sus ESTs, *MBNL3* se expresaría principalmente en músculo (lengua, 10,3%), testículos (8%) y placenta (7,1%).

Para determinar si la expresión de los genes *muscleblind* humanos se producía en algún momento preferente del desarrollo, esto es, el patrón temporal de expresión, se observó el momento del desarrollo de los tejidos de los que procedían los transcritos disponibles para cada gen (Figura 8, 32-34). Clasificados como embrionario, juvenil o adulto se observa que para *MBNL1* y 2, el estadio menos representado es el juvenil mientras que para *MBNL3* ocurre que más del 50% de las secuencias pertenecen a la categoría de embrionario por lo que sería en ese estadio en el que cabría esperar una mayor expresión de este gen. Cabe resaltar que como resultado del análisis del patrón espacial del gen *MBNL3* se obtuvo que la placenta era uno de los tejidos más representados para este gen lo que concuerda bien con el hecho de que su mayor expresión se da en estadios embrionarios.



- | | | | | |
|----------------------|----------------------|--------------------|--------------|-------------------|
| 1. vejiga | 8. corazón | 15. timo | 22. hígado | 29. tejido blando |
| 2. útero | 9. cél. endoteliales | 16. bazo | 23. placenta | 30. laringe |
| 3. colon | 10. sangre | 17. testículo | 24. próstata | 31. ojo |
| 4. intestino delgado | 11. médula ósea | 18. cervix | 25. riñón | 32. embrionario |
| 5. estómago | 12. cerebro | 19. ovario | 26. pulmón | 33. juvenil |
| 6. lengua | 13. S. Nerv. Perif. | 20. glánd. Mamaria | 27. piel | 34. adulto |
| 7. músculo | 14. nódulo linfático | 21. páncreas | 28. hueso | |

(Figura 8. ver pie de Figura en próxima página)

Figura 8. Expresión diferencial de los genes *muscleblind* humanos. Según el análisis del origen tisular y la cantidad de todos los transcritos para cada uno, *MBNL1* es el gen que más se expresa tanto por el número de transcritos por tejido como en la cantidad de tejidos en los que se detectan sus transcritos. Por el contrario, se observa que *MBNL3* tiene una expresión reducida encontrándose en un número limitado de tejidos, y comparativamente a los otros dos genes, en muy baja cantidad (tengase en cuenta el cambio de escala para los genes *MBNL2* y *MBNL3*). En relación con el estadio del desarrollo donde más frecuentemente se encuentran transcritos de estos genes, *MBNL3* parece expresarse mayoritariamente en estadio embrionario, ya que a esta categoría pertenecen más de la mitad del total de sus transcritos mientras que no se encuentra una desviación tan evidente en los otros dos genes. Ver texto para más comentarios sobre los gráficos. Se ha representado el número de transcritos de cada gen que se ha identificado para cada tejido en relación al número total de transcritos en dicho tejido. Datos extraídos de los documentos Unigene: *MBNL1* Hs.478000; *MBNL2* Hs.125715; *MBNL3* Hs.105134 (junio 2005). Sólo se han incluido EST provenientes de librerías de tejidos no cancerosos.

1.1.5 Muscleblind, una nueva familia de proteínas con dedos de zinc tipo CCCH

La identificación y caracterización *in silico* de los genes *muscleblind* humanos abre la puerta al estudio de las características de esta nueva familia proteica. Los objetivos que nos planteamos son determinar en qué organismos podemos encontrar proteínas *muscleblind* y cuales son las características comunes a todas estas proteínas.

La familia de dedos de zinc tipo CCCH es extensa y diversa, no sólo por el número y espaciado de los dedos que presentan las proteínas sino también por el espaciado de los residuos de cisteína dentro de los dedos. La proteína Muscleblind de *Drosophila* es la fundadora de una nueva familia de proteínas CCCH que se caracterizan, principalmente, por contener dedos del tipo CX₇CX₆CX₃H, donde C indica Cisteína, X cualquier aminoácido y H indica Histidina.

La identificación de proteínas Muscleblind se realizó usando la región de la proteína de *Drosophila* más conservada respecto a las proteínas humanas que se corresponde con los aminoácidos 9 a 84 (coordinadas sobre la secuencia AAC01949). Como resultado de las búsquedas en las bases de datos de proteínas (blastp) y de nucleótidos (tblastn) se identificaron más de 90 proteínas Muscleblind cuyo análisis nos permitió extraer información acerca de los organismos en los que encontramos el gen *muscleblind*, el número de copias que existe en cada uno de ellos y las características estructurales de esta nueva familia de proteínas.

El gen *muscleblind* sólo está presente en el grupo de los Metazoos ya que no hemos encontrado ningún representante en bacterias, hongos ni plantas. No se pudo comprobar si este gen está presente en los grupos Cnidaria o Porífera por no existir un

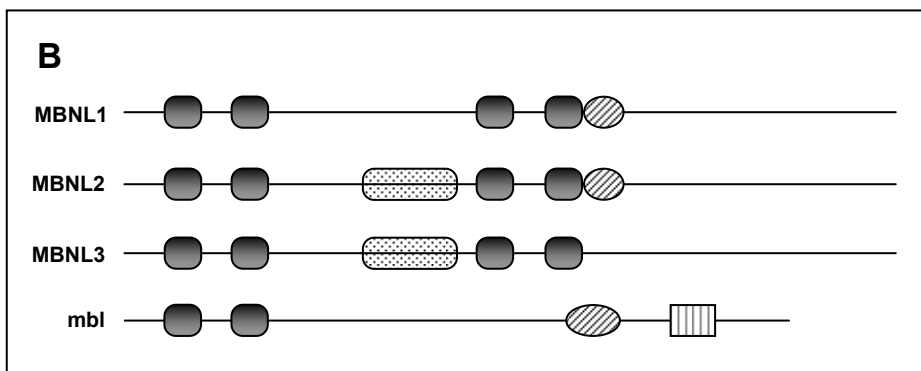
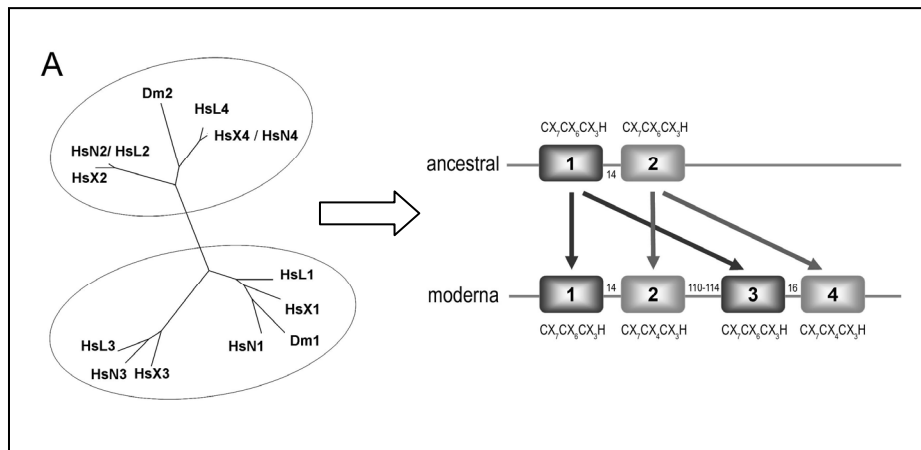
proyecto genoma de ningún organismo representativo. Dentro de los Bilateria, los Nematodos son el grupo más antiguo en el que la proteína Muscleblind puede ser identificada mediante conservación de secuencia (~35% de identidad).

Tomando como referencia las proteínas Muscleblind identificadas en los organismos cuyo genoma está completamente secuenciado podemos concluir que existe un solo gen *muscleblind* en todos los organismos invertebrados y tres genes en los vertebrados, a excepción de los peces donde hemos identificado hasta cinco genes *muscleblind* distintos en *Takifugu rugripes*. La Tabla 2 recoge los números de acceso de las proteínas utilizadas en este estudio.

Un análisis comparativo de los diferentes organismos en los que está presente el gen *muscleblind* indica que todos los invertebrados no cordados poseen un solo par de dedos de zinc mientras que todos los vertebrados poseen dos pares de dedos del mismo tipo. Cuando se hace un alineamiento múltiple entre la secuencia correspondiente a los dedos de zinc de invertebrados (ej. *Drosophila melanogaster*: Dm1-2) y vertebrados (ej. *Homo sapiens*: Hs1-4) se evidencia una mayor similitud entre el primer dedo de invertebrados y los dedos primero y tercero de vertebrados, y del segundo dedo de invertebrados con el tercer y cuarto dedo (Figura 9A). Esto sugiere la aparición del segundo par de dedos en vertebrados por duplicación de los dos dedos presentes inicialmente en invertebrados. Por otro lado, en invertebrados, el espaciado característico entre las cisteínas del dedos de zinc es siempre de CX₇CX₆CX₃H mientras que en vertebrados, el segundo dedo de zinc de cada par ha perdido dos aminoácidos entre las dos últimas cisteínas de manera que el espaciado es CX₇CX₄CX₃H.

Organismo	Genoma	Secuencia	Proteína
Nematoda			
<i>Caenorhabditis elegans</i>	completo	NP_510746.1	Muscleblind
Arthropoda			
<i>Anopheles gambiae</i>	WGS	reconstruida	Muscleblind C
<i>Apis mellifera</i>	WGS	reconstruida	Muscleblind C
<i>Bombyx mori</i>	WGS	reconstruida	Muscleblind C
<i>Drosophila melanogaster</i>	completo	AAC01949.1	Muscleblind C
Tunicata			
<i>Ciona intestinalis</i>	completo	reconstruida	Muscleblind
Vertebrata			
<i>Gallus gallus</i> (pollo)	WGS	CAG31624.1 XP_416979.1 NP_001012591.1	MBNL1 MBNL2 MBNL3
<i>Homo sapiens</i> (humano)	completo	BAA24858 NP_997187.1 AAH74776	MBNL1 MBNL2 MBNL3
<i>Mus musculus</i> (ratón)	completo	NP_064391.2 NP_997398.1 NP_598924.1	MBNL1 MBNL2 MBNL3
<i>Rattus norvegicus</i> (rata)	incompleto	XP_342253.2 XP_214253.2 XP_228685.2	MBNL1 MBNL2 MBNL3
<i>Takifugu rubripes</i> (pez)	WGS	SINFRUP00000169845 SINFRUT00000165687 AA584613 SINFRUP00000176586 SINFRUP00000133529	MBNL1 MBNL1a MBNL2 MBNL2a MBNL3
<i>Xenopus tropicalis</i> (rana)	incompleto	ENSXETP00000046798 ENSXETP00000050703 ENSXETP00000011905	MBNL1 MBNL2 MBNL3

Tabla 2. Secuencias analizadas en el estudio de las características de las proteínas Muscleblind. La etiqueta “completo” indica que el genoma del organismo está totalmente secuenciado y ensamblado. “WGS” indica que el genoma está completamente secuenciado pero no ensamblado. Se indican los números de acceso de la base de datos GenBank excepto los de *Takifugu* y *Xenopus* que se corresponden con la base de datos de Ensembl (<http://www.ensembl.org/>). En los organismos en los que no existe una secuencia proteica anotada, esta se reconstruyó a partir de búsquedas tblastn contra DNA genómico usando la proteína de *Drosophila* como “query”. En la columna “proteína” para invertebrados se indica a qué isoforma proteica de *Drosophila* se corresponde la proteína. Para vertebrados se indica el gen humano con el que tiene mayor homología de secuencia. *Takifugu* posee dos genes mbl adicionales que nombramos según el gen humano con el que tienen una homología de secuencia mayor.



C

MBNL1	1	MAVSVTPIRDTKWLTLLEVCREFORGTCSRPTTECKFAHPSKSCQVNGRVIACFDSLKG
MBNL2	1	MANVAVPRDTKWLTLLEVCRQFORGTCSRSDTECKFAHPPKSCQVNGRVIACFDSLKG
MBNL3	1	MTAVNVALIRDTKWLTLLEVCREFORGTCSRADADECKFAHPPRVCHVNGRVIACFDSLKG
MBNL1	60	RCSRENCKYLHPPPHLKTQLEINGRNNLIQOK-NMAMLAQQMQIANAMMPGAPIQPVPMF
MBNL2	60	RCSRENCKYLHPPPHLKTQLEINGRNNLIQOKTAAAMLAQQMCF---MFPPTPLHPVPIF
MBNL3	61	RCIRENCKYLHPPPHLKTQLEINGRNNLIQOKTAAAMFAQQMQL---MLQNAQMSISLGSF
MBNL1	119	SVAPSLATNASAAAFNPYLGVSF--SLVPAEILPTAPMLVLTGNPGVVPFAAAAAAOKL
MBNL2	117	PVGPATICTN-TAISFAPYLAPVTPGVGLVPTTEILPTTPVIVPGSP--PVTVPGSTATQKL
MBNL3	118	PMTPSIPAN-PPMAFNPLYT-P-HPGMGLVPAELVNPVLPVLPVGNP--PLAMPGAVG-PKL
MBNL1	177	MRTDRLEVCREVQRGNCNRGENDCRFAHPADSTMIDTNDNTVTVCMDYIKGRCSREKCKY
MBNL2	174	LRTDKLEVCREFORGNCRGENDCRFAHPADSTMIDTSDNTVTVCMDYIKGRCMREKCKY
MBNL3	172	MRSDKLEVCREFORGNCRGENDCRYAHPDASMIETASDNTVTTCMDYIKGRCSREKCKY
MBNL1	237	FHPPAHLQAKIKAAQYQVNOAAAAQAAATAAA--MTQSAVKSILKRPLEATFDLGIPOAVL
MBNL2	234	FHPPAHLQAKIKAAHQANQAAVAAQAAAAATVMTQSTAKAMKRPLEATVDLAFPPGAL
MBNL3	232	FHPPAHLQARIKAAHHQMNHSAASA-----M-----ALQPGTL
MBNL1	295	PPLPKRPALEKINGATAVNTGTFQYQQALANMQLQQHTAFIPEGSILCMTPATSVVPMV
MBNL2	294	HPLPKRQALEKSNGTSAVFNPSVLYHYQQALTSALQQQHAAFIPTGSVLCMTPATSIIVPMM
MBNL3	265	QLLIPKRSALEKINGATVFNPTVFHCCQALTNLQLPQP-AFIPIAGPILCMAPASNIIVPMM
MBNL1	355	HGATPATVSAATTSAATVPFAATATANQIPIISAEHLTSHKVIQOM
MBNL2	354	HSATSAATVSAATTPATSVFPFAATATANQIILK-----
MBNL3	324	HGATPTTVSAATTPATSVFPFAAPPTTGNQIQLSIDELNS-MEIVSQM

Figura 9. ver pie de figura en página siguiente

Figura 9. Consideraciones estructurales de las proteínas Muscleblind. **(A)** Muscleblind en invertebrados y vertebrados. Los dedos de zinc se nombran según el orden de aparición a lo largo de la proteína. Los invertebrados tienen dos dedos de zinc consecutivos del tipo CX₇CX₆CX₃H mientras que los vertebrados tienen cuatro dedos de zinc agrupados de dos en dos del tipo CX₇CX₆CX₃H (primero y tercero) y CX₇CX₄CX₃H (segundo y cuarto). Se muestra un árbol de similaridad entre las secuencias de los dedos de zinc entre humanos y mosca que muestra como el primer dedo de la mosca (Dm1) se agrupa junto a los dedos primero y tercero de los tres genes muscleblind humanos mientras que el segundo dedo (Dm2) lo hace junto con el segundo y cuarto dedo de zinc en humanos. El primer dedo de *Drosophila* es más parecido al primer dedo de humanos mientras que el segundo, actualmente, muestra una mayor similitud con el cuarto. Estas observaciones sugieren que durante la evolución de los vertebrados ocurrió una duplicación en tandem que cubría los dos dedos de zinc. Posteriormente, el segundo dedo de zinc de cada par en vertebrados perdió dos aminoácidos entre la segunda y tercera cisteína. Hs, *Homo sapiens*; Dm, *Drosophila melanogaster*; N, L y X para MBNL1, MBNL2 y MBNL3, respectivamente. **(B)** Esquema de las proteínas Muscleblind en vertebrados (MBNL1-3) e invertebrados (se representa la isoforma b de la proteína de *Drosophila*). Las proteínas muscleblind se caracterizan por sus motivos de dedos de zinc. Además, existen distintas regiones de “baja complejidad de secuencia” dentro de cada una de las distintas proteínas muscleblind. En la Figura sólo se han representado aquellas regiones de baja complejidad que comparten los tres genes muscleblind de vertebrados y todos los organismos analizados para cada uno de ellos. Punteado, rico en prolina; rayas oblicuas, rico en alanina y rayas verticales, dominio transmembrana predicho. **(C)** Alineamiento de las proteínas Muscleblind de Vertebrados. Para los tres genes, las proteínas Muscleblind de los diferentes vertebrados analizados (ver tabla 2) tienen una identidad global de secuencia mayor del 92%. Además de los dedos de zinc, existen otras secuencias altamente conservadas en todas las proteínas analizadas (cuadros negros) y una región de baja complejidad de secuencia que también está conservada (cuadro punteado). Las secuencias de MBNL1, MBNL2 y MBNL3 utilizadas para realizar este alineamiento se corresponden con las proteínas humanas y se han completado con todos los exones alternativos de cada gen que no alteran la pauta de lectura, con el fin de poder comparar los péptidos codificados por todos los exones conservados. Las proteínas utilizadas en este alineamiento se han reconstruido a partir de los siguientes transcritos: para MBNL1, Y13829 + exón 7; para MBNL2, AAM09798 y para MBNL3, AAH74776 + exón 10.

Los miembros de la familia *muscleblind* también se caracterizan por un espaciado definido entre los dos dedos de zinc de cada par siendo de 14 y 16 aminoácidos para el primer (el único en invertebrados) y el segundo par, respectivamente. Además, en las proteínas *muscleblind* de vertebrados, la longitud de la secuencia que une los dos pares de dedos de zinc o *linker* está conservada variando entre 110-112 aminoácidos según la especie. La región linker queda localizada en el exón 3 de los genes *MBNL1-3*. Existen variantes de *splicing* en las proteínas humanas en las que se excluye el exón 3. Estas isoformas proteicas también aparecen en otros organismos como el ratón. La eliminación de la secuencia que separa ambos pares de dedos de zinc en las proteínas Mbl humanas afecta a la función de la proteína tal y como demuestran los experimentos de tres híbridos realizados recientemente por Kino y col (2004).

Aparte de los dedos de zinc, las proteínas Muscleblind no presentan ningún otro dominio proteico relevante, si bien, se caracterizan por la presencia de diferentes regiones de “baja complejidad de secuencia” que se localizan, generalmente, después

del último dedo de zinc o entre los dos pares de dedos de zinc de vertebrados. Algunas de estas regiones son ricas en alanina o prolina (Figura 9B) las cuales han sido implicadas en la interacción proteína-proteína teniendo un efecto represor o activador de la transcripción respectivamente. De acuerdo con esta idea, se han publicado experimentos basados en tres híbridos que muestran que la eliminación de la región rica en alanina de la proteína MBNL1 humana aumenta su capacidad de unión a RNAs portadores de repeticiones CUGs (Kino, 2004). La región rica en alanina se encuentra en el exón 4 de los genes MBNL, justo después del segundo par de dedos de zinc por lo que su presencia no está determinada por ninguna variante de *splicing*.

Entre los vertebrados las proteínas Muscleblind tienen un porcentaje de identidad global con respecto al correspondiente gen humano superior al 88% en todos los casos. Esta homología de secuencia es similar a la que se encuentra entre las proteínas Muscleblind de invertebrados.

Aunque la identidad global entre las proteínas de vertebrados e invertebrados esta alrededor del 35%, se encuentra un porcentaje mucho mayor si nos fijamos solo en las secuencias que cubren los dedos de zinc (>50% respecto al primer par de vertebrados). Además, una comparación entre secuencias de proteínas Mbl pertenecientes a organismos alejados evolutivamente revelan dos motivos proteicos a parte de los dedos de zinc (Fig. 10). El primer motivo, al que hemos nombrado caja LEV, viene dado por la secuencia consenso WLXLEV, donde X indica cualquier aminoácido, y esta localizado inmediatamente antes del primer dedo de zinc. El segundo motivo consiste en una secuencia NGR que está precedida por un residuo de valina o asparagina, con dos residuos apolares antes y despues. Hemos nombrado a este motivo caja NGR y su secuencia consenso es (V/I)X₀₋₂NGR(V/N)X(A/L). Dos cajas NGR flanquean el segundo dedo de zinc de todas las proteínas Mbl estudiadas (Tabla 2). La conservación de estos residuos en especies tan distantes evolutivamente puede indicar una relevancia funcional en la actividad básica de la proteína Muscleblind.

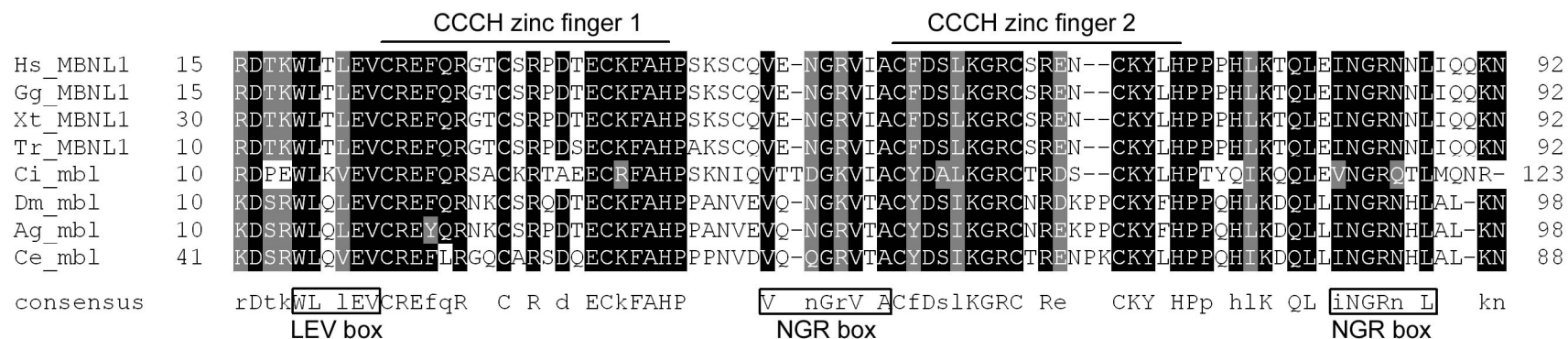


Figura 10. Conservación de la secuencia de aminoácidos entre distintos miembros de la familia Muscleblind. Se muestra el alineamiento dentro del primer par de dedos de zinc para distintos organismos: Dm, *Drosophila melanogaster*; Ag, *Anopheles gambiae*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*; Gg, *Gallus gallus*; Xt, *Xenopus tropicalis*; Tr, *Takifugu rubripes*; Ci, *Ciona intestinalis*. Para los vertebrados se ha utilizado la proteína MBNL1. Según la región mostrada, *C. elegans* es la proteína Mbl más distante compartiendo el 67% de los residuos aminoacídicos mientras que las proteínas de los vertebrados son prácticamente idénticas (99-100%). Como se aprecia en este alineamiento, la homología de secuencia se extiende más allá de los dominios de dedos de zinc, mostrando regiones altamente conservadas a lo largo de la evolución que hemos definido como las cajas LEV y NGR. Las barras horizontales señalan la extensión de los dos dedos de zinc tipo CCCH. Consensus, secuencia consenso establecida para el primer par de dedos de zinc de las proteínas Mbl, las mayúsculas indican un 100% de identidad y las minúsculas conservación de la similaridad.

1.2 RELEVANCIA CLÍNICA DE LOS GENES *muscleblind* HUMANOS

En *Drosophila*, el gen *muscleblind* es necesario para la diferenciación terminal de la musculatura y los fotorreceptores, resultando un gen esencial para el desarrollo de este organismo ya que sus mutaciones resultan letales en homocigosis. La identificación de los genes humanos homólogos plantea la cuestión de si mutaciones en algunos de estos genes pudieran ser responsables de la aparición de alguna patología humana o de algún síntoma clínico en particular. Por esta razón, uno de los objetivos de este trabajo de investigación fue la secuenciación de los genes *muscleblind* humanos en pacientes para cuyas patologías alguno de estos genes resultara candidato.

1.2.1 MBNL1, búsqueda de mutaciones en pacientes Usher tipo III

El síndrome de Usher (USH) es una enfermedad autosómica recesiva que se caracteriza por una pérdida auditiva sensoneural y una pérdida progresiva de la visión debido a Retinitis Pigmentosa. Se trata de una enfermedad clínica y genéticamente heterogénea. Se distinguen tres tipos de USH en base a la gravedad y la edad de inicio de las disfunciones auditivas y vestibulares: tipo I, II y III. El síndrome de Usher tipo III (USH3) se caracteriza por una pérdida progresiva de audición y una hipoactividad vestibular (OMIM: 276902). Aunque se ha estimado que USH3 representa el 2% del total de casos del síndrome de Usher, basado en criterios clínicos, en Finlandia el porcentaje se estima en un 40%.

En 1995 se localizó el locus USH3 en la región 3q21-q25 del genoma humano (Sankila, 1995). La localización del gen responsable de USH3 fue posteriormente delimitada en un intervalo genético de 1cM entre los marcadores D3S1299 y D3S3625. (Joensuu, 2000). El gen humano *MBNL1* mapea en la posición cromosómica 3q25, entre los marcadores D3S1279 y D3S1280 por tanto resultaba un gen candidato para ser USH3 porque se localizaba dentro de la región crítica (GenMap'99). Además, aunque actualmente la función de este gen humanos se desconoce, se asume que puede estar relacionada con la función visual, ya que esto es lo que ocurre con su homólogo en *Drosophila melanogaster*.

En este trabajo se han analizado cuatro pacientes. Se trata de dos parejas de hermanos (608-609 y 934-935) pertenecientes a dos familias no relacionadas entre sí, diagnosticadas inicialmente como USH I pero que presentaban un claro ligamiento a la región cromosómica 3q donde ha sido asignado previamente el locus USH3

La estrategia utilizada para la secuenciación del gen *MBNL1* en estos pacientes fue la amplificación de la ORF completa mediante RT-PCR a partir de RNA total extraído a partir de linfocitos de los cuatro pacientes y de un sujeto control. No se encontraron cambios nucleotídicos en la región codificante de *MBNL1* en ninguno de los pacientes analizados. Lo que sí revelaron es una nueva variante de procesado que depositamos en las bases de datos públicas (GenBank Accesion: AJ308400). La existencia de posibles alteraciones en los sitios de procesado del gen y en la región promotora así como el estudio de la relevancia de que en estos cuatro individuos se hubiera detectado una nueva variante de procesado, quedó pendiente de comprobación debido a la aparición de una publicación (Miller, 2000) en la que se implicaba la proteína MBNL1 en la ruta de patogénesis de la Distrofia Miotónica tipo I.

1.2.2 MBNL1, búsqueda de mutaciones en pacientes con Distrofia Miotónica

Durante el desarrollo de este trabajo de investigación fueron publicados los hallazgos de Miller y colaboradores (2000) que apoyan la hipótesis de que la proteína MBNL1 es “secuestrada” por el mensajero mutante del gen *DMPK*. Este gen es el responsable de la Distrofia Miotónica tipo I (DM1). El defecto genético que origina esta patología consiste en una expansión de trinucleótidos CTG en la región 3'UTR del gen *DMPK*.

Existen dos enfermedades de características clínicas difícilmente diferenciables a DM1 pero que, sin embargo, no presentan la expansión del trinucleótido en *DMPK*. Se trata de la Distrofia Miotónica tipo II (DM2) y la Miopatía Miotónica Proximal (PROMM) (OMIM: 602668 y 600109 respectivamente). El locus DM2 había sido mapeado en 3q25 (Ranum, 1998). Por otro lado, varios estudios con familias alemanas diagnosticadas como PROMM mostraron evidencia de heterogenidad genética en esta patología (Ricker, 1999). La mayoría de las familias mostraban ligamiento al locus DM2 por lo que se hablaba de enfermedades alélicas (DM2/PROMM), sin embargo, en un número limitado de familias quedaba excluido este locus por lo que se propuso la existencia de, al menos, un tercer locus, DM3.

De acuerdo con el modelo de patogénesis de la Distrofía Miotónica en el momento en el que llevabamos a cabo esta investigación, nos planteamos la siguiente hipótesis de trabajo. Ya que el secuestro de la proteína MBNL1 por el mensajero mutado de *DMPK* parecía ser, al menos, una de las causas de la aparición de la DM1 y dado que el secuestro de la proteína se traduce en una falta de función de este gen, propusimos que mutaciones en el propio gen *MBNL1* podrían tener un efecto similar y ser, por tanto, responsables directos de una enfermedad con síntomas muy similares como pueden ser DM2 o PROMM.

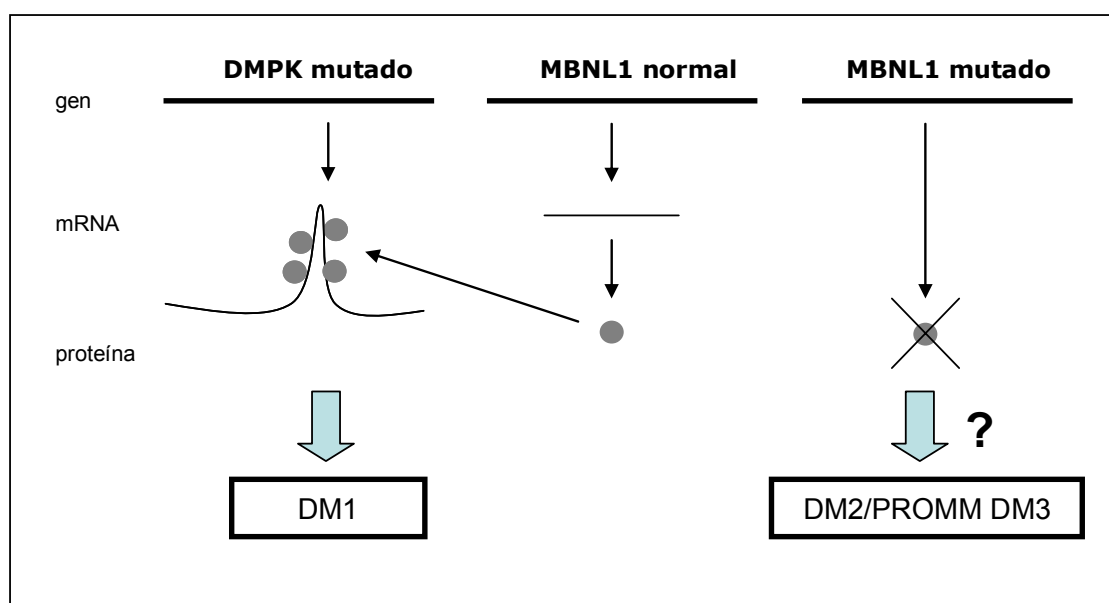


Figura 1. Implicación de la proteína MBNL1 en la ruta de patogénesis de la Distrofia miotónica tipo I y la hipótesis sobre las consecuencias de mutaciones en el gen *MBNL1* que originen una falta de función de este gen. Todas las patologías DM tienen una herencia autosómica dominante por lo que las mutaciones responsables en *MBNL1* deben generar dominantes negativos de esta función que puedan explicar que una sola copia mutante desencadene la patología. Otra posibilidad es que mutaciones en *MBNL1* sean haploinsuficientes.

Para comprobar esta hipótesis buscamos mutaciones en el gen *MBNL1* en un grupo de pacientes con sintomatología DM1 pero que no presentaban la expansión en el gen *DMPK*, es decir, se secuenció *MBNL1* en pacientes con DM cuyo gen responsable todavía no había sido identificado por si una mutación en *MBNL1* pudiera ser *per se* un desencadenante primario la enfermedad. Esto incluía pacientes diagnosticados como DM1 y también pacientes diagnosticados como PROMM. Por otro lado, incluimos muestras de personas que presentaban la expansión pero eran asintomáticas (Tabla 1). La secuenciación de *MBNL1* en estos pacientes asintomáticos podría revelar mutaciones o polimorfismos en el gen que tuvieran un efecto “protector” frente a la

expansión (Anderson, 2000; Houy, 2004; Medlej-Hashim, 2004), por ejemplo, una mutación en *MBNL1* que generase una mayor cantidad de proteína podría paliar el efecto de “falta de función” debido al secuestro por parte de las expansiones.

Nuestra Ref	Origen	Referencia	Fenotipo Clínico	Comentarios	MBNL alelo poli(dT)
1	HLF	DM-143 fDM-32	DM	sin expansión en DMPK	11
2	HLF	DM-317 fDM-73	DM	sin expansión en DMPK	11
3	HLF	ng96-221 fDM-85	DM	sin expansión en DMPK	11
4	HLF	ng97-88 fDM-99	DM	sin expansión en DMPK	11
5	HLF	ng97-590 fDM-107	DM	sin expansión en DMPK	11
6	HLF	ng97-591 fDM-108	DM	sin expansión en DMPK	11
7	HLF	ng99-46 fDM-133	DM	sin expansión en DMPK	11
8	HLF	ng99-499 fDM-145	DM	sin expansión en DMPK	11
9	HLF	00-374	DM	sin expansión en DMPK	11
10	HLF	ng99-523	DM	sin expansión en DMPK	11
11	HLF	DM 99-1	-	Expansión	11
12	HLF	DM-7	-	Expansión	11
13	HLF	ng98-712 fDM-127	PROMM	sin expansión en DMPK	10 [§]
14	HUT	LT/Ig NAMG96.1137	PROMM	sin expansión en DMPK	10 [§]

Tabla 1. Muestras para la secuenciación de *MBNL1*. Se incluyen tres tipos de pacientes. 1-10, pacientes diagnosticados como DM1 que no presentan la mutación en *DMPK*; 11-12, personas que presentan la mutación en *DMPK* pero son asintomáticas y 13-14, pacientes diagnosticados como PROMM. HLF, Hospital universitario La Fe, España. HUT, Hospital universitario de Tromsø, Finlandia.

Así pues, se diseñaron cebadores para amplificar, a partir de DNA genómico, todos los exones codificantes de *MBNL1* (material y métodos 1.8) en las distintas muestras, incluidos los límites intrón-exón, así como gran parte de la región 5' no traducida.

Se secuenciaron más de 126 productos de PCR y se buscaron mutaciones en más de 268 secuencias lo que supone un total de 64,5 kb de DNA analizado. Sin embargo, los resultados de la secuenciación de *MBNL1* descartaron la posibilidad de que existiera alguna mutación en la región codificante o en secuencias importantes para la maduración de los transcritos *MBNL1* en dichos pacientes.

Durante la secuenciación de estos pacientes, identificamos una secuencia en la región 5'-UTR constituida por varias timinas (secuencia poli(dT)) que era polimórfica en algunos pacientes con DM, en concreto en aquellos diagnosticados como DM2/PROMM. Estos individuos presentaban una secuencia de 10 timinas en la región 5'-UTR de *MBNL1* (alelo indicado como 10T) mientras que el resto presentaba una

secuencia de 11 timinas (alelo 11T). En todos los casos el alelo 10T se encontraba en heterocigosis (Figura 2).

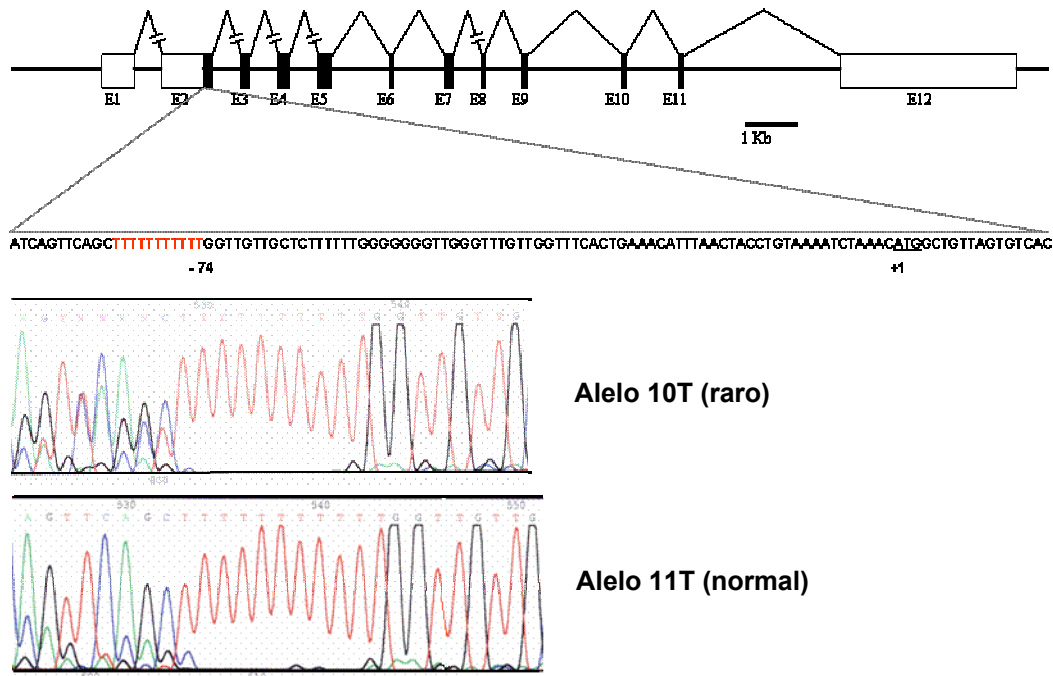


Figura 2. Polimorfismo poli(dT) en el gen humano *MBNL1*. La secuencia polimórfica encontrada consiste en una secuencia de timinas de longitud variable que se encuentra en el exón 1 del gen *MBNL1*. Esta secuencia, llamada poli(dT), se localiza en la región 5'-UTR del mRNA *MBNL1*, exactamente 74 pb aguas arriba desde el ATG de inicio de la traducción. Hemos encontrado dos alelos para la secuencia poli(dT), el más frecuente contiene 11 timinas (11T) y el más raro 10 (10T).

El descubrimiento de este polimorfismo nos planteó una nueva hipótesis de trabajo ya que, en las muestras analizadas había una correlación directa entre la presencia de este polimorfismo y la patología PROMM. Una de las muestras pertenecía a un caso no familiar de una mujer española con síntomas de hipotonía y miotonía en la parte proximal de las extremidades, que fue diagnosticada como PROMM y remitida a la Unidad de Genética del Hospital Clínico Universitario "La Fe" para excluir la DM1. La otra muestra pertenecía a un individuo de una familia noruega con PROMM publicada por Sun y colaboradores (1999). Para comprobar la relevancia de esta correlación procedimos a realizar un estudio poblacional que nos permitiera conocer la frecuencia de este polimorfismo en las dos poblaciones implicadas, la española y la noruega.

Para la amplificación, a partir de DNA genómico, de la región que incluía el polimorfismo poli(dT) diseñamos dos nuevos cebadores 5UTRpA y 5UTRpT.

Analizamos la presencia del polimorfismo en 100 individuos control españoles y 50 individuos control noruegos. Los resultados indicaron que se trata de un polimorfismo muy raro. En la población española obtuvimos una frecuencia de 0.5% (sólo un cromosoma de los 200 analizados presentaba el polimorfismo). Para la población noruega la frecuencia alélica del polimorfismo fue de un 0% ya que ningún cromosoma de los 100 analizados presentó el alelo raro. El hecho de que el polimorfismo poli(dT) fuera raro daba mayor relevancia a la asociación encontrada con los dos pacientes PROMM.

Para sustentar esta correlación determinamos la presencia del polimorfismo poli(dT) en el resto de individuos afectados del pedigree de la familia noruega con PROMM. Analizamos las muestras de seis de los componentes de este pedigree (Figura 3; II:2, III:1, III:2, IV:1; IV:2; IV:3). El genotipado del polimorfismo poli(dT) en los miembros de esta familia mostró que además de la muestra que ya conocíamos (III:2 se corresponde con la muestra 14 Tabla 1), solo un descendiente de esa persona lo poseía también, IV:2. En esta familia existen dos individuos, III:2 y IV:2, cuyo examen clínico muestra una afectación clara. Además, existe un tercer individuo, IV:3, cuyo diagnóstico clínico se debate en Sun y colaboradores (1999). Los autores consideran que el individuo IV:3 es afecto en base a un patrón típico de rigidez muscular que aparecería en la infancia, una aparición temprana de cataratas, calvicie frontal y elevados valores para la creatin kinasa (CK). De las tres personas afectadas en el pedigree, las dos con una sintomatología PROMM más clara (atrofia muscular evidente y otros síntomas característicos de la enfermedad) poseían el alelo raro 10T para el polimorfismo poli(dT) mientras que una tercera persona, afectada pero de una manera menos evidente, no lo poseía. Esta situación, descartaba que el polimorfismo pudiera ser la causa genética directa de la aparición de la patología PROMM. Sin embargo, no era incompatible con la idea de que este cambio nucleotídico en *MBNL1* pudiera estar actuando como un modificador del fenotipo PROMM de manera que el alelo raro actuase como potenciador de la mutación PROMM mientras que el alelo normal suavizase sus efectos.

Durante la realización de este trabajo, el grupo de la Dra. Laura Ranum del Instituto de Genética Humana la Universidad de Minesota (Estados Unidos) descubrió la mutación responsable del fenotipo DM2/PROMM en las familias que mapeaban en 3q25 (Liquori, 2001). Se trata de una expansión del tetranucleótido CCTG en el intrón 1 del gen *Zinc Finger protein 9* (ZNF9). En colaboración con el Dr. Ralf Krahe del *Anderson Cancer Center* de la Universidad de Texas (Estados Unidos) y el Dr. Bjarne Udd del

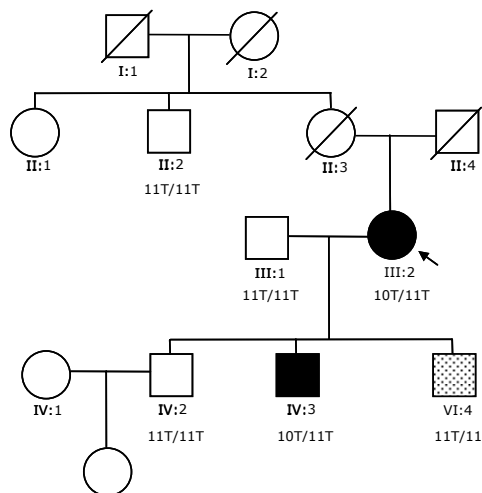


Figura 3. Pedigree de la familia PROMM noruega publicada por Sun y col. (1999). Mujeres y hombres se representan con círculos y cuadrados respectivamente. En negro se indican las personas afectadas. El paciente IV:4 muestra unas manifestaciones clínicas diferentes indicándose con un moteado. Una línea diagonal indica que la persona ha fallecido. El probando de la familia esta señalado con una flecha. Se indica el genotipo para la secuencia poliT de *MBNL1* en los individuos donde se ha determinado.

Departamento de Neurología del Hospital Central Vaasa (Finlandia) pudimos determinar si este era el gen mutado en la familia PROMM noruega con la que habíamos trabajado. Mediante amplificación del marcador CL3N58 descrito en Liquori (2001) comprobamos que la familia noruega PROMM publicada por Sun y col. (1999) tenía la expansión en ZNF9. Sin embargo, no existía mutación en el individuo PR6 cuyo estado clínico había sido debatido. Por tanto, este individuo debía considerarse normal respecto a PROMM. Estos resultados rebatieron la hipótesis del efecto modificador del polimorfismo poliT en *MBNL1* sobre la mutación PROMM ya que la situación con la que nos encontramos en estos momentos era de una asociación sencilla entre la presencia de un polimorfismo raro y tres casos de PROMM.

Para comprobar la relevancia de la asociación entre la presencia del alelo 10T y la enfermedad PROMM secuenciamos la región poli(dT) de *MBNL1* en más familias PROMM. Gracias a la colaboración con los Drs. Ralf Krahe y Bjarne Udd tuvimos acceso a tres familias finlandesas diagnosticadas como PROMM y positivas para la mutación en ZNF9: FIN001 con un fenotipo muy grave, FIN002 con un fenotipo PROMM estándar y FIN004, con un fenotipo más leve caracterizado por una aparición muy tardía de la debilidad muscular. El genotipado de la secuencia poli(dT) de, al menos, 2 miembros afectados y uno normal para cada familia, indicó que ninguno de ellos presentaba el alelo raro 10T.

En base a los resultados obtenidos tras la secuenciación de la región de *MBNL1* poli(dT) en diferentes familias PROMM concluimos que la presencia del alelo 10T no tiene un papel relevante en la determinación del fenotipo PROMM, por lo menos en cuanto a la gravedad de los síntomas más característicos de la enfermedad como son

la miotonía y la debilidad muscular. Sin embargo, la conservación de la región poli(dT) en otros organismos como el ratón y la rata (Figura 4) sugiere fuertemente que pueda ser relevante para algún aspecto de la expresión del gen. Dada la relevancia del gen

Homo sapiens

```
TGGGAGATCTTTTCCTTGATATTGACGACACAATTTCCATGTACTTTTA
AAGCAGGGAGTGGGGAAAAGTATTTTGAGGGGACATTTTCATCATCAGTT
CAGCTTTTTTTTTTTGGTTGTTGCTCTTTTTTGGGGGGTGGGTTTGT
GGTTTCAC TGAAACATTTAACTACCTGTAAAATCTAAACATGGCTGTTAG
TGTCACACCAATTCGGGACACAAAATGGCTAACACTGGAAGTATGTAGAG
AGTTCCAGAGGGGGACTTGCTCACGCCAGACACGGAATGTAAATTTGCA
```

Mus musculus

```
TTTTATCAGTTCCTATTTTGTAGTTTTTGTCTTTTGTGGTTGCTG
ATTTTTTTTTCTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTTCTTTTT
GTTTTTCCTTTTTTTTTTGGAGAGGGGTTGGGTTTGTGGTTTCATT
GAACATTTAACTACCTGTAAAAATAAACATGGCTGTTAGTGTACACCAA
TTCCGGGACACAAAATGGCTAACACTGGAAGTTGTAGAGAGTTTCAAAGGG
GGACTTGCTCACGACCAGACACGGAATGTAAATTTGCACACCTTCGAAAA
```

Rattus norvegicus

```
AATTTCTAGCGGAGGTCTTTCCCTTGATATTGAAGTCACACTTTTCCATGT
GCCGTTAAAGCGGGGACTGGGGAAGCAGCCTTTCGGACATTTTCACAGTCA
TCTCATCTTTTCTTTTCCTTTTTTTTTTTTTTTTTTGAGAGGGATTGGGT
TTGTTGGTTTCATTGAACATTTAACTACCTGTAAAATATAAACATGGCTGT
TAGTGTACACCAATTCGGGACACAAAATGGCTAACACTGGAAGTATGTAG
AGAGTTTCAAAGGGGGACTTGCTCACGACCAGACACGGAATGTAAATTTGC
```

Figura 4. Conservación de la región poli(dT) de *MBNL1* en distintos organismos. La secuencia poli(dT) se localiza en el exón 1 en la región 5'UTR (sombreado). La distancia al ATG de inicio en humanos es de 74 pb mientras que en los roedores es de 57 pb. La longitud de la secuencia poli(dT) es de 11 timinas en humanos, 12 en ratón y 17 en rata. La secuencia de humanos se corresponde con las coordenadas 323 a 622 de Y13829. La secuencia de ratón se corresponde con las posiciones 307 a 611 de AF231110. La secuencia de rata se corresponde con las posiciones 585 a 891 de XM_342252.2. Tanto en rata como en ratón la secuencia poli(dT) se encuentra embebida en una región muy repetitiva rica en timinas.

donde se encuentra este polimorfismo y su carácter de extraordinaria rareza, resultaría de gran interés, primero, determinar si el alelo 10T tiene algún efecto sobre la expresión de *MBNL1* y segundo, realizar un estudio clínico detallado de los tres pacientes PROMM portadores de este alelo raro para intentar encontrar algún característica común que pueda indicar qué otro tipo de pacientes pueden ser investigados (ver discusión 1.5).

1.2.3 *MBNL3*, búsqueda de mutaciones en pacientes con Ptosis Familiar

El clon genómico que contiene a *MBNL3*, AL050310, definió la localización cromosómica de este gen en el intervalo Xq25-26, entre los marcadores DXS1047 y DXS998. Por tanto, *MBNL3* es un gen candidato para todas las patologías cuya región crítica, determinada por estudios de ligamiento, incluya a los marcadores DXS1047 y DXS998. De acuerdo con la información existente en su momento elaboramos una lista de enfermedades genéticas ligadas al cromosoma X dentro de cuya región candidata estuviera incluido *MBNL3* (Tabla 2, ver material y métodos 2.1).

Nombre de la Patología	Siglas	Localización	OMIM	Intervalo	Máx. Z
<i>Cowchock syndrome</i>	NAMDS	Xq24-26.1	310490	DXS425-HPRT	DXS737
<i>Ptosis Hereditary congenital, 2</i>	PTOSX	Xq24-q26	300245	DXS1001-DXS1205	DXS984
<i>MR X-linked with isolated GH deficiency</i>	MRGH	Xq24-27.1	300123	DXS178-DXS292	DXS294
<i>Pettgrew syndrome</i>	PGS	Xq25-27	304340	DXS425-F9	HPRT
<i>Gustavson syndrome</i>	GUST	Xq26	309555	DXS424-DXS297	DXS297
<i>Shashi X-linked MR syndrome</i>	MRXS11	Xq26-q27	300238	GATA59C05-GATA31E08	DXS1047
<i>Borjeson-Ferssman-Lehmann syndrome</i>	BFLS	Xq26.3	301900	DXS155-DXS730	DXS294
<i>Albinis-deafness syndrome</i>	ADFN	Xq26.3-27.1	300700	SG34535-SG83750	-
<i>Nystagmus 1 congenital X-linked</i>	NYS1	Xq26-27	310700	GATA172D05-DXS1192	DXS1047
<i>Retinitis pigmentosa-24</i>	RP24	Xq26-27	300155	DXS8094-DXS8043	DXS8106

Tabla 2. Listado de enfermedades genéticas candidatas para la secuenciación del gen *MBNL3*. De entre todas las patologías cuya localización cromosómica solapaba con la conocida para *MBNL3* (Xq25-q26.3), se seleccionaron aquellas cuya región candidata contuviera los marcadores DXS1047 y DXS998 ya que incluyen a *MBNL3*. De estas, se seleccionaron aquellas cuya sintomatología incluyese a ojos y/o músculos o en general a cualquier órgano y/o tejido afectado en pacientes con Distrofia miotónica. Las enfermedades están ordenadas según su localización cromosómica. MR, “mental retardation”. La columna de “intervalo” indica los marcadores que delimitan la región candidata para cada enfermedad. En la columna “Máx. Z” se indica el marcador para el que se obtenía el máximo *lod score* según los estudios de ligamiento más recientes en cada patología.

Entre las enfermedades genéticas candidatas figuraba la “Ptosis congénita aislada ligada al cromosoma X” (OMIM 300245). Esta condición, de herencia dominante ligada al sexo, ha sido descrita únicamente en una familia caucasiana de gran tamaño del sureste de Inglaterra. La palabra “ptosis” es una abreviatura de *blepharoptosis*, término que se usa para describir la caída del párpado superior. Los miembros afectados de esta familia nacieron con ptosis bilateral con una función elevadora mínima en ambos párpados lo que casi les impedía la visión en una posición relajada normal. La Ptosis congénita se produce por una distrofia o una falta de desarrollo de

los músculos responsables del movimiento del párpado (Yilmaz N, 2004). Las proteínas Muscleblind han sido implicadas en la ruta de patogénesis de la Distrofia miotónica tipo I y tipo II. La Ptosis es característica de la DM y se considera uno de los síntomas diagnósticos más importantes (Harper, 2001). Dada la función de las proteínas Muscleblind en la diferenciación terminal del músculo en *Drosophila* y la implicación de las proteínas humanas en una patología donde la ptosis es un síntoma característico, nos planteamos si mutaciones en *MBNL3* podrían originar ptosis congénita.

Para ser un buen candidato, *MBNL3* debería expresarse en los tejidos afectados, idealmente en la musculatura de la zona ocular. Para confirmar experimentalmente los resultados del análisis *in silico* de la expresión de *MBNL3* (resultados 1.1.4) realizamos un perfil de expresión usando un conjunto de librerías de cDNA proporcionadas por la Dra. Jackeline M. Bye del *Sanger Institute* (Inglaterra). Dos de las tres librerías que dieron positivo con las condiciones probadas provenían de órganos con un alto contenido en tejido muscular como el corazón y el intestino delgado (Figura 5).

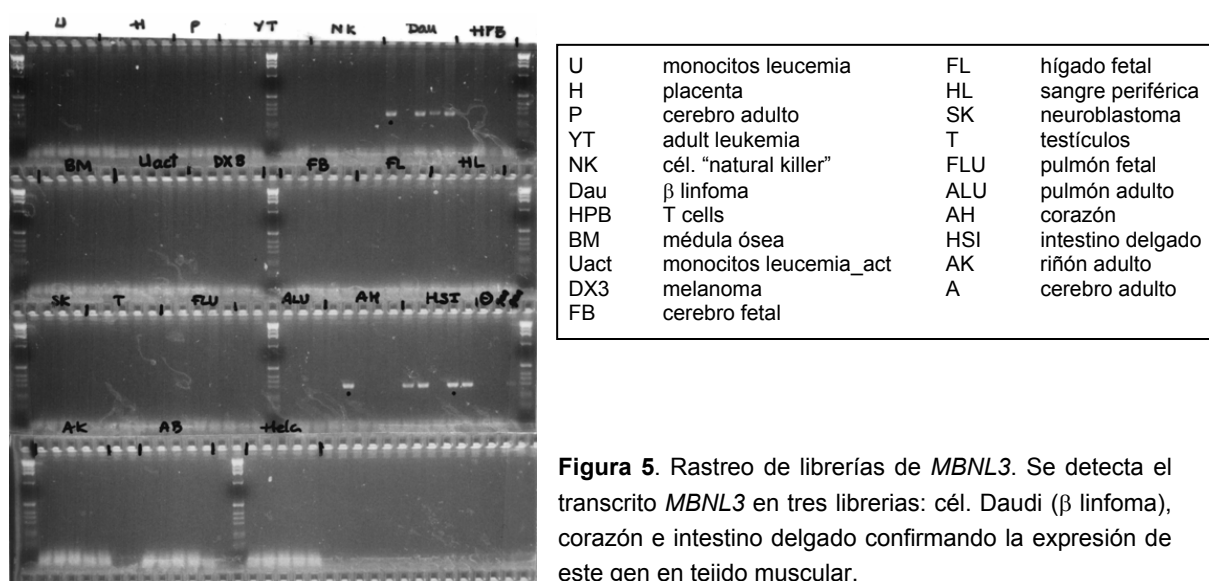


Figura 5. Rastreo de librerías de *MBNL3*. Se detecta el transcrito *MBNL3* en tres librerías: cél. Daudi (β linfoma), corazón e intestino delgado confirmando la expresión de este gen en tejido muscular.

Así pues, el gen *MBNL3* es un buen gen candidato para la Ptosis Congénita porque cae dentro de la región candidata y se expresa en la clase de tejido que aparece alterado en la enfermedad, esto es, tejido muscular. Además, este gen está implicado en la ruta de patogénesis de la Distrofia miotónica donde uno de los síntomas es la ptosis y por otro lado, el gen homólogo en *Drosophila* tiene una función en la diferenciación terminal de la musculatura. Por todo ello, secuenciamos el gen *MBNL3* en uno de los individuos del pedigree publicado por McMullan y col. (2000; Figura 3).

Para ello se diseñaron cebadores para la amplificación, a partir de DNA genómico, de todos los exones codificantes del gen incluyendo los límites intrón-exón (ver material y métodos 1.8)

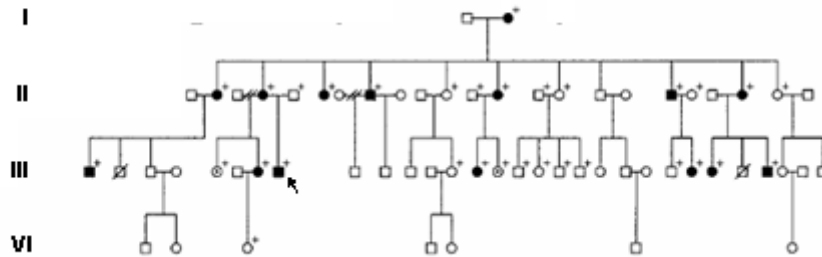


Figura 6. Familia con Ptosis congénita ligada al sexo estudiada por McMullan y col (2000). (+) Individuos para los que se tienen muestra de DNA. (●,■) mujer u hombre afectado, respectivamente. (†) probando y muestra elegida para la secuenciación completa del gen *MBNL3*.

La secuenciación completa de la región codificante del gen *MBNL3* en el paciente indicado no mostró ninguna alteración nucleotídica respecto de la secuencia del DNA genómico depositado en las bases de datos. Por ello, no se procedió a la secuenciación de ningún otro individuo del pedigree descartándose la implicación de mutaciones, en la región codificante o en sitios reguladores del splicing, en *MBNL3* como agente causante directo de la Ptosis congénita en esta familia.

2. ANÁLISIS DE LA FUNCIÓN MOLECULAR DE *muscleblind* EN *Drosophila melanogaster*

En nuestro laboratorio hemos desarrollado un modelo del efecto tóxico de las repeticiones (CTG)_n asociadas a la Distrofia miotónica que consiste en la sobreexpresión de 480 repeticiones del trinucleótido (Monferrer, 2004). Este modelo, reproduce diversos aspectos de la patología en los que se ha demostrado la relevancia de Muscleblind, como son la formación de foci nucleares en el músculo o una alteración del procesado alternativo de la α -actinina. La identificación de genes y proteínas que interaccionen con *muscleblind* permite conocer y entender no solo como *muscleblind* lleva a cabo su función molecular, sino también ayuda a explicar los acontecimientos moleculares y celulares que originan los defectos fenotípicos observados en nuestro modelo. La comprensión de la ruta de patogénesis en *Drosophila* ofrecerá diferentes dianas para el desarrollo de terapias que posteriormente serán validadas en otros organismos modelo como el ratón.

La proteína Muscleblind humana, MBNL1, ha sido descrita como un factor de *splicing* que actúa sobre un conjunto de pre-mRNAs diana (Ho, 2004; Jiang, 2004; Dansithong, 2005). Experimentos llevados a cabo en nuestro laboratorio muestran que la proteína MBNL1 humana es capaz de rescatar el fenotipo mutante en *Drosophila* asociado a la falta de función de *mbi*, mejorando aspectos como la viabilidad y el fenotipo de abdomen contraído (Monferrer, 2005b). Esto demuestra que existe una conservación funcional entre la proteína MBNL1 humana y la proteína Mbl de *Drosophila*, lo que sugiere que esta última también podría actuar como factor de *splicing*. Por otro lado, experimentos de cotransfección en células humanas que sobreexpresan repeticiones CUG demuestran que las proteínas MblA-C de *Drosophila* son capaces de unirse a las repeticiones al igual que su homólogas humanas (Monferrer, 2005). Además, en nuestro laboratorio hemos detectado cambios en el *splicing* del mensajero de la α -actinina asociados a la falta de función de *mbi*, y utilizando un ensayo de tres híbridos en levadura (SenGupta, 1996), hemos demostrado la interacción de la proteína MblC con un fragmento intrónico de dicho pre-mRNA (M Vicente, datos no publicados). Todas estas evidencias experimentales apoyan la idea de que, de manera análoga a su homóloga humana, la proteína Muscleblind de *Drosophila* actúa también como factor de *splicing*. En el ejercicio de esta función, Mbl requerirá del concurso de

otras proteínas y se unirá a RNAs. Además de factor de *splicing*, es posible que Mbl tenga otras funciones moleculares.

Para identificar dianas moleculares de *muscleblind* en *Drosophila*, desarrollamos dos aproximaciones diferentes pero complementarias entre sí, una aproximación genética y una aproximación bioquímica. En la aproximación genética realizamos un rastreo genético de modificadores dominantes de un fenotipo de sobreexpresión de *mbIC* en el que esperamos identificar genes cuya función esté relacionada con la de *mbI*. En la aproximación bioquímica, llevamos a cabo experimentos de co-inmunoprecipitación para identificar proteínas que interaccionen con la proteína Muscleblind a la vez que comprobamos la presencia de productos génicos candidatos identificados como resultado del rastreo genético.

2.1. APROXIMACIÓN GENÉTICA PARA EL ESTUDIO DE LA FUNCIÓN MOLECULAR DE LA PROTEÍNA Muscleblind DE *Drosophila*

En este trabajo utilizamos *Drosophila* para identificar genes relacionados con la función de *muscleblind* y nos proponemos realizar un rastreo de modificadores dominantes utilizando un fenotipo de sobreexpresión. El fenotipo empleado en un rastreo genético debe estar sensibilizado a la dosis de otros genes, esto es, debe ser modificable. También debe permitir una clasificación rápida y fácil de los resultados permitiendo determinar, con el menor número de observaciones posibles, si existe interacción génica o no. Por último, este fenotipo debe estar relacionado con la función del gen que se está estudiando con el fin de que las interacciones sean lo más relevantes posibles. Un ejemplo en el cual el fenotipo no estaría relacionado con la función del gen es aquel en el que su sobreexpresión desencadena apoptosis debido a un efecto tóxico para la célula.

En nuestro laboratorio se han descrito varios fenotipos generados por la sobreexpresión de *muscleblind* en ojo y en ala (García-Casado, 2004). La sobreexpresión de la isoforma *mbIC* en el ojo de *Drosophila* usando la línea *sev-Gal4* (*sev-Gal4*>UAS-*mbIC*) produce un fenotipo de ojos rugosos. El análisis de secciones transversales de estos ojos muestra defectos típicos de la polaridad epitelial plana que se manifiestan por la pérdida de la quiralidad de los omatidios y por una rotación

errónea de los mismos. Además, también se observa la ausencia de fotorreceptores en el ojo adulto. Esto puede ser debido a (1) un problema en la diferenciación del fotorreceptor de manera semejante a lo que ocurre en los mutantes por falta de función de *mb1* (Begemann, 1997) o (2) los fotorreceptores se diferencian pero posteriormente mueren. Para comprobar si la falta de fotorreceptores debida a la sobreexpresión de *mb1C* se debe a un fenotipo de apoptosis realizamos inmunodetección α -caspasa 3 en discos de ojo de larvas de tercer estadio. Los resultados no muestran un aumento de la apoptosis en comparación con los discos imaginales obtenidos a partir de larvas OrR. Esto descarta que el fenotipo de sobreexpresión de *mb1C* en ojo sea producido por la activación inespecífica de apoptosis y, por otro lado, sugiere que el fenotipo de pérdida de fotorreceptores debido a la sobreexpresión puede estar originado por la misma causa que el fenotipo por pérdida de función.

Por otro lado, la sobreexpresión de *mb1C* en el compartimento posterior del ala usando la línea *engrailed-GAL4* (*en-GAL4>UAS-mb1C*) muestra un fenotipo complejo que incluye defectos morfológicos en el crecimiento, ya que las alas presentan un tamaño inferior, y una falta de tejido laminar en el compartimento posterior, así como también, defectos en el patrón de venas y patrón de pelos epiteliales. Además, encontramos una gran variabilidad en este fenotipo de sobreexpresión. No se han descrito las modificaciones histológicas que subyacen a los defectos que se observan. En este sentido, nos propusimos investigar las causas que originaran al menos alguno de los defectos fenotípicos más relevantes de la sobreexpresión de *mb1C* en ala como es el caso de la falta de tejido laminar en el compartimento posterior (Figura 1).

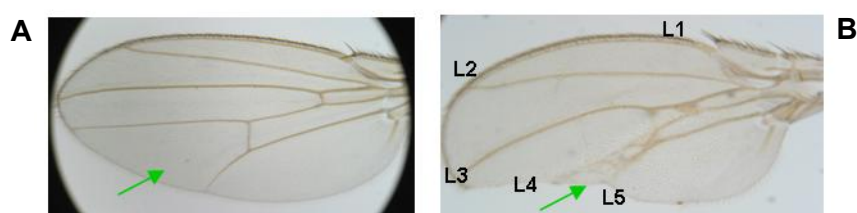


Figura 1. Fenotipo de sobreexpresión de *mb1C* en el compartimento posterior del ala con la línea *en-Gal4*. (A) ala salvaje (OrR) y (B) ala *en-GAL4>UAS-mb1C* en la que se puede apreciar la falta de tejido laminar en el compartimento segundo posterior, entre las venas L4 y L5.

El crecimiento de los órganos y tejidos es un equilibrio entre la proliferación y la apoptosis celular. La ausencia de material tisular puede deberse o bien a una falta de proliferación celular (ej. reciben alguna señal que inhibe la división celular) o bien a un exceso de muerte celular (ej. un número de células superior al normal entran en apoptosis) durante la formación del patrón básico del ala, esto es, en el disco imaginal

de ala. Para comprobar si la falta de tejido se debe a apoptosis realizamos tinciones con α -caspasa 3 en discos imaginales de ala de larvas de tercer estadio. Los resultados obtenidos muestran que la sobreexpresión del trangén *UASmbIC* con el driver *en-GAL4* activa la apoptosis en un región concreta dentro del compartimento donde se expresa *engrailed* en los discos imaginales de ala (fig 2). La activación de la apoptosis provoca pérdida de células. La pérdida anormal de células en esa región del disco imaginal correlaciona bien con la falta de tejido en el compartimento posterior del ala adulta.

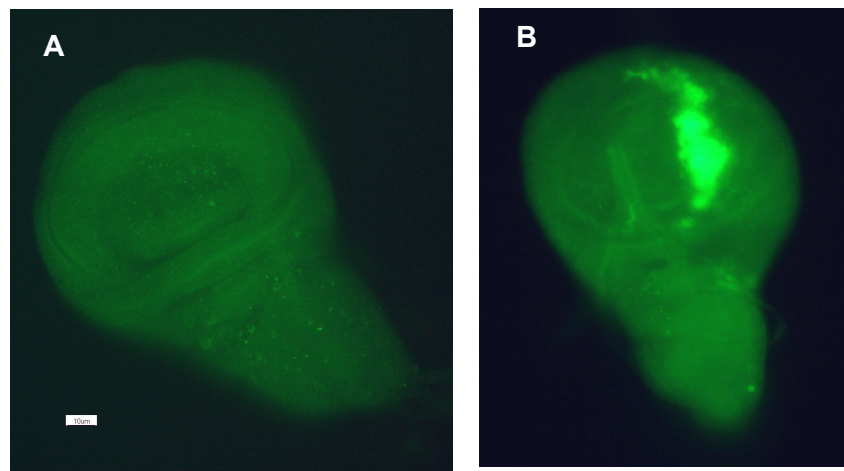


Figura 2. La sobreexpresión de *mbIC* en ala induce apoptosis. Los discos fueron extraídos de larvas de tercer estadio salvajes OrR (A) y de moscas transgénicas *engrailed-GAL4>UAS-mbIC* (B). Las células apoptóticas se marcan con un anticuerpo específico para la caspasa 3. La señal muestra una activación de la apoptosis en una región del disco que correlaciona bien la pérdida de tejido laminar en el ala adulta.

Para comprobar si también existe apoptosis en los discos de ojo que sobreexpresan *mbIC* realizamos inmunodetección α -caspasa 3 sobre discos de ojo de larvas de tercer estadio. Como se ha indicado anteriormente, los resultados de este experimento no muestran un aumento de la apoptosis en comparación con los discos imaginales obtenidos a partir de larvas OrR.

La gran variabilidad del fenotipo de sobreexpresión de *mbIC* en ala y el hecho de que existe un fuerte activación de la apoptosis hacen este fenotipo poco recomendable para los rastreos ya que, (1) la expresividad variable de un fenotipo hace más difícil identificar a los modificadores de ese fenotipo ya que obliga a observar un mayor número de muestras para valorar si hay modificación del fenotipo o no, y (2) la fuerte activación de la apoptosis como resultado de la sobreexpresión podría limitar o sesgar

la obtención de modificadores a solamente, aquellos relacionados indirectamente con la función de *mb1*, por ejemplo los implicados en apoptosis.

Por estas razones, decidimos utilizar el fenotipo de sobreexpresión en ojo para realizar el rastreo genético de modificadores dominantes pudiendo utilizar el ala para validar las interacciones encontradas así como también el fenotipo de ojos rugosos generado por la sobreexpresión de las repeticiones (CTG)₄₈₀ (Monferrer, 2004).

2.1.1 Rastreo de genes que interaccionen genéticamente con *muscleblind*

Para identificar modificadores dominantes del fenotipo de sobreexpresión de *mb1C* en ojo (Figura 3), realizamos un rastreo piloto usando un kit de deleciones que cubría aproximadamente el 55% del genoma de *Drosophila* (kit de deleciones del Bloomington *Drosophila* Research Center). De las 109 deleciones ensayadas, 13 (12%) modificaban visiblemente el fenotipo de ojos rugosos actuando, la mayoría de ellas, como supresores dominantes (Tabla 1). El porcentaje de modificadores hallados está acorde con el porcentaje esperable para un fenotipo sensibilizado (~10%).

DELECIÓN	Potencia	Suprime	REGIÓN DELECIONADA
Df(2R)017		++	56F5; F15
Df(2L)C144		++	23A1-2; 23C3-5
Df(2L)sc19-8		+++	24C2-8; 25C8-9
Df(2L)TW137		+++	36C2-4; 37B9-10
Df(3R)Tpl10		++	83C1-2; 84B1-2
Df(3L)66C-G28		++	66B8-9; 66C9-10
Df(3L)vin5		++	68A2-3; 69A1-3
Df(3L)BSC12		++	69F6-70A1; 70A1-2
Df(3L)st-f13		++	72C1-D1; 73A3-4
Df(3L)XS-533		++	76B4; 77B
Df(2R) w45-30n	+++		45A6-7; 45E2-3
Df(2R)CX1	++		49C1-4; 50C23-D2
Df(2L)r10	++		35D1; 36A6-7

Tabla 1. Resumen del rastreo piloto de modificadores dominantes del fenotipo de ojo rugoso de la sobreexpresión de *mb1C*. Se cuantifica el efecto modificador como (++) débil-moderado y (+++) fuerte.

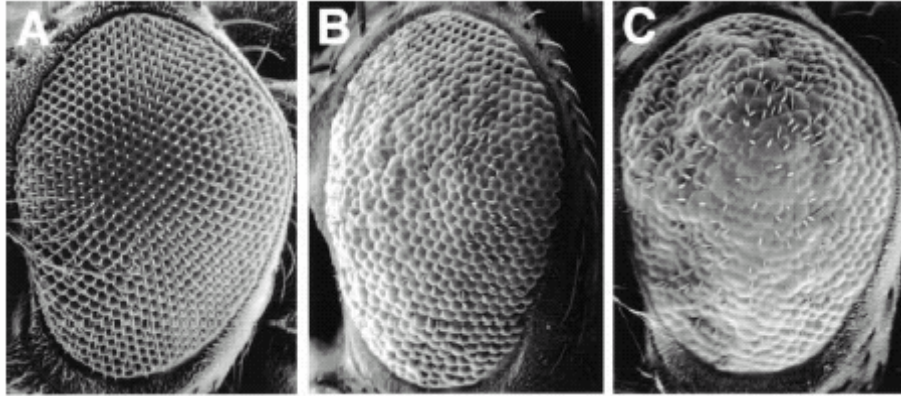


Figura 3. Fenotipo de sobreexpresión de *mbIC* en ojo y efecto de los modificadores. Imagen al microscopio electrónico de barrido de un fenotipo de ojo rugoso debido a la sobreexpresión del transgén *enGAL4>UASmbIC* (B). A la izquierda, ejemplo del efecto de un modificador dominante con un efecto supresor (A, delección *Df(2L)TW137*) y a la derecha un ejemplo de un modificador con un efecto potenciador (C, delección *Df(2L)w45-30n*). Se observa que el fenotipo es sensible a cambios en la dosis génica de otros genes.

Los modificadores con un efecto supresor tienen, en principio, mayor relevancia frente a las que tienen un efecto potenciador ya que, sin existir relación entre los dos elementos genéticos que interactúan, es más fácil que el fenotipo empeore (potenciador) que mejore (supresor). Por esto, centramos nuestra atención en dos de las delecciones con un efecto supresor que mostraban una interacción más fuerte, *Df(2L)sc19-8* y *Df(2L)TW137*.

Las delecciones *Df(2L)sc19-8* y *Df(2L)TW137* se estudiaron con más detalle para identificar los genes responsables del efecto modificador observado. Con el fin de reducir el número de genes candidatos a estudiar acotamos la región candidata haciendo uso de delecciones solapantes. Como resultado de este estudio se consiguió reducir las regiones candidatas a 350 kb y 700 kb respectivamente (Figura 4).

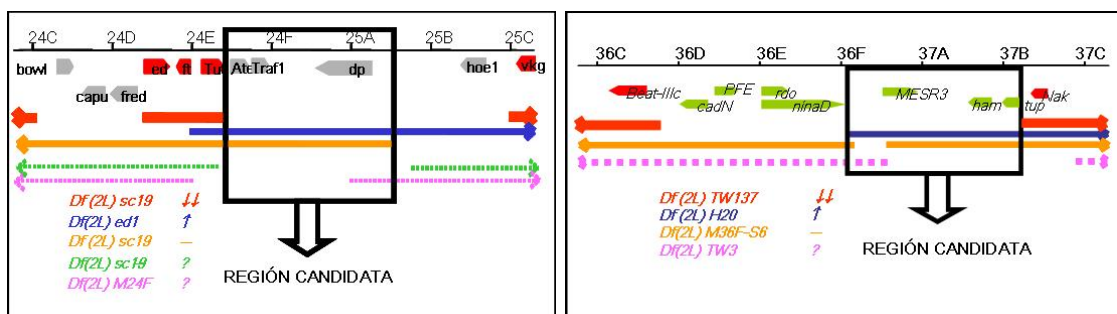


Figura 4. Representación de las regiones genómicas para las delecciones *Df(2L)sc19-8* y *Df(2L)TW137* y uso de delecciones solapantes para acotar la región candidata. ↑ potencia, ↓ suprime, - no modifica, ? no hay datos.

Con el objeto de ensayar individualmente cada uno de los genes dentro de la región candidata acotada buscamos todas las inserciones disponibles en el *Drosophila Genome Resource Center* “Kyoto fly” (<http://kyotofly.kit.jp/stocks/>). Este centro mantiene líneas de moscas portadoras de inserciones de elementos P de diferentes tipos.

2.1.1.1 Análisis de los genes de las deleciones Df(2L)TW137 y Df(2L)sc19-8

Delección Df(2L)sc19-8

Se identificaron 14 inserciones en la región candidata acotada (24D5; 25A2). De las 12 estudiadas cuatro interaccionaron: *CG15435*^{GS11154}, *tutI*^{GS13843}; *Traf1*^{GS2154} y *CG15433*^{NP434}, todas ellas con un efecto supresor del fenotipo de ojos rugosos de acuerdo con la interacción observada para la delección que las contiene.

Las cuatro inserciones que interaccionan pertenecen a dos tipos: la colección “*Gene Search*” (GS; Toba, 1999) y la colección del “*NP consortium*” (National Institute of Genetics, Japan). Ambas colecciones están basadas en una aproximación de mutagénesis por ganancia de función mediante el uso del sistema UAS/Gal4. La colección GS está constituida por elementos P con dos secuencias UAS que inducen transcripción de los genes situados a ambos lados de la inserción mientras que la colección NP está constituida por elementos *enhancer trap* del tipo P[Gal4].

El fenotipo que observamos al cruzar moscas de la colección GS con nuestro recombinante *en-GAL4 UAS-mbIC*, aunque también puede deberse a una reducción en la dosis génica del gen analizado debido a un efecto mutagénico de la inserción P[UAS], estará influenciado, en cualquier caso, por la sobreexpresión de dicho gen en el ojo debido al efecto del *driver engrailed* de nuestro recombinante. Para el caso de la colección NP, el fenotipo que observamos igualmente puede deberse a una reducción de la dosis génica debida a un efecto mutagénico de la inserción P[GAL4], pero de manera análoga, también estará determinado por la sobreexpresión del gen *muscleblind* bajo el patrón de expresión propio del gen analizado. En las líneas GAL4, o de *enhancer trap*, se sintetiza la proteína GAL4 con el patrón de expresión propio del gen de estudio. De esta manera la proteína GAL4 determinará que el transgén *UAS-mbIC* se exprese a su vez con ese mismo patrón de expresión.

Dada la naturaleza de estos mutantes, comprobamos si la interacción que estábamos observando se debía, en efecto, a los genes en los que se localizaba cada inserción. Para ello, se buscaron mutantes por falta de función para los genes correspondientes en el Bloomington *Drosophila* stock Center (<http://flystocks.bio.indiana.edu/>). Para dos de las cuatro inserciones interaccionantes, *CG15435*^{GS11154} y *CG15433*^{NP434}, conseguimos las correspondientes líneas letales *CG154435*^{KG05006} y *CG15433*^{KG02386}. Estos mutantes por falta de función se comportan como modificadores dominantes supresores verificando las interacciones anteriormente observadas en las líneas GS y NP, respectivamente.

CG15433 es un gen localizado en 24F cuya proteína predicha pertenece a la familia de las histonas acetiltransferasas de la familia GNAT (*GCN5-related N-acetyltransferase*) tipo GCN5. La supuesta actividad acetiltransferasa relaciona a esta proteína con el ensamblaje/desensamblaje de la cromatina (Neuwald, 1997). Por otro lado, este gen se ha identificado como un potenciador del fenotipo de apoptosis de *grim-reaper* (Wing, 2002). Los transcritos de *CG15433* son de origen materno y se detectan a lo largo de todo el desarrollo embrionario. Cabe resaltar su presencia en la musculatura somática de embriones y larvas, además de en las glándulas salivares (BDGP <http://www.fruitfly.org/cgi-bin/ex/insitu.pl>).

El gen *CG15435* se localiza en 24F3. La proteína codificada por este gen tiene dominios de dedos de zinc del tipo C2H2 lo que le otorga una supuesta función como factor de transcripción. Este gen fue identificado en un rastreo de genes precursores neurales a partir de una librería específica de cDNA de cabeza embrionaria (Brody, 2000) por lo que tiene una supuesta función como gen proneural. El gen *CG15435* tiene un máximo de expresión al principio de la embriogénesis. Los momentos del ciclo de vida de la mosca donde menos expresión ocurre son al final del estadio embrionario y principio del larvario, así como, al final de la prepupa (Arbeitman, 2002; base de datos on-line : <http://genome.med.yale.edu/Lifecycle/>)

Delección *Df(2L)TW137*

Se identificaron 16 inserciones en la región candidata acotada (36B10; 37B3). De las 13 estudiadas, las siguientes cuatro interaccionaron: *CG5790*^{NP1233}, *CG17424*^{NP7067}, *CG17323*^{GS9130} y *CG31751*^{NP1565}. Tres de ellas tienen un efecto potenciador del fenotipo de ojos rugosos, al contrario que la interacción observada para la delección que las contiene, mientras que una de ellas, *CG5790*^{NP1233}, resulta letal. Nuevamente

debido a la naturaleza de las líneas interaccionantes, se procedió a la búsqueda de líneas P letales para verificar la interacción. Solo se consiguió obtener la línea letal *CG17323*^{KG03187}. La falta de función también interacciona con el fenotipo de sobreexpresión de *mb1C* en ojo, en este caso actúa como supresor dominante.

El gen *CG17323* se localiza en 37B1. Se trata de uno de los 30 genes del genoma de *Drosophila* que codifica un producto con actividad UDP glucosil transferasa (Luque, 2002). *CG17323* es un gen cuya transcripción se induce como respuesta a apoptosis producida por la exposición de ecdisona y radiaciones (Lee, 2003). Además, ha sido identificado en un rastreo de genes que se expresan diferencialmente durante la miogénesis en concreto durante la diferenciación terminal (Artero, 2003). De acuerdo con esto, la expresión detectada en un microarray “*developmental time course*” indica que se expresa preferentemente en los últimos estadios del desarrollo embrionario, concretamente en los estadios 13 a 16 (<http://genome.med.yale.edu/Lifecycle/>).

2.1.1.2 Análisis de genes candidatos de todas las deleciones

Como disponíamos de datos funcionales de *muscleblind*, decidimos acotar el número de genes candidatos dentro de cada deleción centrándonos en aquellos con funciones relacionadas con una de las siguientes categorías: metabolismo del RNA y factores de transcripción.

Las proteínas Muscleblind humanas regulan el *splicing* alternativo de mRNAs como la troponina cardíaca T (cTNT) y el receptor de la insulina (IR) (Ho, 2004). Además, experimentos de tres híbridos demuestran que tienen capacidad de unión directa al RNA (Kino, 2004). De manera análoga, en nuestro laboratorio hemos encontrado que la proteína Muscleblind de *Drosophila* regula el *splicing* del mRNA de la α -actinina y , mediante experimentos de tres híbridos en levadura, que existe una unión directa de Mbl con un fragmento intrónico de dicho mRNA (M Vicente, datos no publicados). Por estas razones consideramos que Mbl es una proteína que puede estar implicada en el *splicing* de los pre-mRNAs. Por otro lado, no podemos descartar que, además del *splicing*, Mbl pueda estar relacionada con otros aspectos del metabolismo del RNA. Proteínas con dedos de zinc CCCH muy similares a los de las proteínas Muscleblind, han sido implicados en diversas funciones relacionadas con la unión a RNA. Por ejemplo, la proteína TTP, prototipo de la familia de proteínas con dedos de zinc del tipo CX₈CX₅CX₃H, se une directamente a los elementos ARE (*AU-rich elements*) en el

mRNA provocando la deadenilación y desestabilización de los transcritos a los que se unen (Lai, 2001). Por otro lado, Adhereth y colaboradores (comunicación personal) han demostrado muy recientemente una nueva función para la proteína MBNL2 humana relacionada con la localización del mensajero de la integrina $\alpha 3$ y su traducción en las adhesiones focales. Por tanto consideramos que los genes que codifican proteínas implicadas en cualquier aspecto del metabolismo del RNA, son candidatos evidentes para interactuar con la función de *muscleblind*.

Las razones para considerar candidatos a aquellos genes cuyos productos se unan a DNA, son principalmente tres. Primera, existen ejemplos de proteínas con dedos de zinc donde se ha demostrado su capacidad para unir tanto RNA como DNA. Un ejemplo lo constituye el factor de transcripción TFIIIA. Este es necesario para la transcripción de los genes que codifican el RNA ribosomal 5S pero también se ha encontrado asociado con mensajero ribosomal RNA 5S en las partículas de almacenamiento citoplasmático 7S en los ovocitos de anfibios y peces teleostios (Brown, 2005). Otro ejemplo viene dado por la proteína supresora del tumor de Wilm (WT1). El gen *WT1* codifica una proteína con dedos de zinc del tipo Krüppel que regula la expresión de distintos genes relacionados con el crecimiento por unión a secuencias específicas de DNA localizadas en la región promotora o 5'-UTR del gen diana. Además, se ha visto que WT1 puede encontrarse asociada a la maquinaria de *splicing* (Davies, 1998) y que se une a determinadas secuencias exónicas del mRNA *insulin-like growth factor 2* (Igf-2) siendo dependiente esta unión de los dedos de zinc (Caricasole, 1996).

La segunda razón por la que consideramos candidatos los genes que codifican productos que unen DNA es por la relevancia genérica que los factores de transcripción parecen tener en la ruta de patogénesis de la Distrofia miotónica. En 2003, Ebralidze y colaboradores demostraron que el mensajero mutante de *DMPK* une y secuestra factores de transcripción nucleares, lo que conduce a una redistribución de los mismos desde los sitios activos de la cromatina hacia la fracción ribonucleoproteica de la matriz nuclear, esto es, hacia los foci formados por el mensajero mutante *DMPK*. Dado el secuestro que sufren las proteínas *Muscleblind* por este mismo mensajero, nuestra hipótesis es que puede existir alguna relación funcional entre ambos tipos de proteínas, bien porque *muscleblind* participa de algún modo en el proceso de transcripción, bien por el conocido acoplamiento entre los procesos de transcripción y procesado de los mensajeros.

Finalmente, en un mapa de interacciones proteicas a escala genómica donde realiza un ensayo de dos híbridos entre más de 4500 proteínas de *Drosophila*, *muscleblind* es uno de los cDNAs ensayados y se encuentran 38 interacciones diferentes (Giot, 2003). Entre las 19 proteínas identificadas con una función conocida cabe resaltar que, al menos cinco tienen una actividad como factor de transcripción. Esto sugiere que *mbI* podría estar relacionado en el proceso de transcripción.

La identificación de los genes con funciones de unión a RNA o DNA se realizó haciendo búsquedas de líneas P letales en el *Bloomington Drosophila Stock Center* acotando la búsqueda a las regiones citogenéticas dentro de las deleciones y restringiéndola a aquellos genes cuyos productos proteicos estuvieran propuestos como factores de transcripción o proteínas implicadas en el metabolismo del RNA.

Por otro lado, comprobamos la interacción con distintos genes que resultaban buenos candidatos para la interacción con *muscleblind* según la información disponible. Las fuentes proceden de cuatro estudios: (1) genes que interaccionan con el modelo en *Drosophila* para la SCA8 (Mutsuddi, 2004; ver introducción); (2) inducción de la apoptosis por la sobreexpresión de *mbIC* en ala; 3) genes relevantes en la ruta de patogénesis de la DM en humanos; y 4) interacción física con la proteína Mbl según análisis de dos híbridos (Giot, 2003)

Metabolismo del RNA

Dentro de las deleciones que habían interaccionado se identificaron y estudiaron las inserciones letales que codificaban para las proteínas de unión a RNA: *Aly*⁰²²⁶⁷, *tsu*^{EP567} y *Rnp*^{EP1082}. Solo las dos primeras mutaciones interaccionan genéticamente con la sobreexpresión de *mbIC* en ojo actuando en la misma dirección que la deleción que las incluye, esto es, *Aly*⁰²²⁶⁷ como supresor y *tsu*^{EP567} como potenciador

El gen *Aly*, también conocido como DmREF1 se localiza en 84B2. Es una proteína con dominios de unión a RNA del tipo RRM (RNA Recognition Motif). Se trata de un componente del *Exón Junction Complex* (EJC) en *Drosophila*. El EJC es un complejo multiproteico que se une al mRNA en los puntos donde ha ocurrido el corte y empalme inmediatamente después de actuar el espliceosoma. Este complejo contiene factores implicados en el transporte al citoplasma de los mensajeros, localización citoplasmática y la ruta de degradación de mensajeros “erróneos” conocida como *Non-sense Mediated Decay* (NMD). La proteína *Aly* podría estar implicada en el transporte

al citoplasma de mensajeros específicos. Además, algunas evidencias apuntan a un papel en la regulación de la traducción de los mRNA exportados (Gatfield, 2002). Hibridaciones *in situ* en embriones de *Drosophila* muestran expresión de *Aly* desde los primeros estadios embrionarios (transcritos de origen materno). Se observa una expresión ubicua bastante fuerte a lo largo de todo el desarrollo embrionario aunque difusa (BDGP <http://www.fruitfly.org/cgi-bin/ex/insitu.pl>)

El gen *tsunagi* (*tsu*), también conocido como *DmY14*, se localiza en 45A7. La proteína posee dominios de unión a RNA del tipo RRM y, al igual que *Aly*, es un componente del ECJ (Gatfield, 2002). Interacciona físicamente con la proteína Mago Nashi formando un heterodímero funcional (Mohr, 2001;) y es necesario para la localización del mRNA de *oskar* en el ovocito (Hachet, 2001).

Factores de Transcripción

Se estudiaron seis inserciones letales: *amos*^{Tft-RM11}, *su(var)2-10*⁰³⁶⁹⁷, *Dp*^{KG00660}, *slp1*², *bow1*¹ y *esg*^{35Ce-1}. Solo los tres primeros interactúan; *amos*^{Tft-RM11} y *su(var)2-10*⁰³⁶⁹⁷, con un efecto contrario al de la delección que los incluye, mientras *Dp*^{KG00660} lo hace en la misma dirección. Así pues, *amos*^{Tft-RM11} y *Dp*^{KG00660} actúan como potenciadores del fenotipo de sobreexpresión de *mbIC* en ojo mientras *su(var)2-10*⁰³⁶⁹⁷ lo hace como un supresor.

El gen *absent MD neurons and olfactory sensilla* abreviadamente *amos*, también conocido como *Tufted* o *Roi*, está localizado en 36F6. Codifica una proteína con dominios de unión a DNA del tipo *Helix-Loop-Helix* (HLH) con una supuesta función como factor de transcripción. Se trata de un gen proneural implicado en la determinación de los órganos sensoriales y la diferenciación de la retina. Además del alelo de falta de función de *amos*^{Tft-RM11}, también comprobamos la interacción con el alelo de ganancia de función *amos*^{Roi-1}. Ambos se mostraron como potenciadores dominantes del fenotipo de sobreexpresión de *mbIC*. Además la sobreexpresión conjunta en el ojo de *mbIC* y *amos* resulta letal para la mosca (alelo *amosUAS*)

El gen *suppression of variegation 2-10* abreviadamente *Su(var)2-10*, también conocido como *pias* o *zimp*, se localiza en 45A8. *Su(var)2-10* codifica un miembro de la familia de proteínas PIAS (*Protein Inhibitor of Activated STAT*), un grupo muy conservado de proteínas que regulan negativamente los procesos mediados por las proteínas STAT (*signal transducers and activators of transcription*). Además, contiene dominios SAP

(SAFA/B, Acinus and PIAS) de unión a DNA. *Su(var)2-10* controla diversos aspectos de la estructura y función de los cromosomas estableciendo y/o manteniendo la organización de los cromosomas en el núcleo interfásico (Hari, 2001). También contiene dominios DEAD/H-box *helicase* que se relacionan con una actividad ATPasa que media reordenamientos estructurales en el RNA en procesos como splicing y degradación de RNA, transporte al citoplasma o ensamblaje de los ribosomas (InterPro_001410). Este gen ha sido relacionado con diversos procesos biológicos como son el desarrollo del ojo y la proliferación de las células sanguíneas (Betz, 2001)

El gen *Dp* se localiza en 49F10-11. Codifica para la proteína DP que forma parte del factor de transcripción E2F/DP. Se ha comprobado la implicación de este heterodímero en varios procesos biológicos como son la regulación del ciclo celular, la regulación de apoptosis inducida por daños en el DNA y la determinación del eje dorso ventral del ovocito de *Drosophila*, entre otros. Se han aislado mutantes por falta de función de *Dp* que tienen afectados numerosos tejidos, entre ellos, el ojo (tamaño y forma, desorganización pérdida de omatidios y pelos interomatidio) y el abdomen. El gen *Dp* se expresa en el SNC embrionario, en el disco ojo-antena, en la musculatura embrionaria y larvaria y en el cerebro larvario.

Interacción con el Modelo *Sca8*

La base genética de la Ataxia espinocerebelar 8 (SCA8) consiste en una repetición del trinucleótido CTG en el mensajero no codificante del gen *SCA8*. Se ha propuesto un efecto tóxico de las repeticiones CTG en el contexto del transcrito *SCA8* semejante al que explica la distrofia miotónica (ver introducción). De acuerdo con esta hipótesis, estudios genéticos en un modelo de *SCA8* en *Drosophila melanogaster* han demostrado que *muscleblind* actúa como un modificador dominante del fenotipo de ojos rugosos originado por la expresión de las repeticiones CTG (Mutsuddi, 2004). Además, se han descrito mutaciones modificadoras en otros genes. Los genes cuyas mutaciones interaccionen con el modelo de *SCA8* interaccionan potencialmente con una falta de función de *muscleblind*. Por este motivo, comprobamos la interacción de algunos de estos genes con el fenotipo de sobreexpresión de *mbIC* en ojo.

Se estudiaron cinco inserciones: *lolal*^{k02512}, *Atel*^{k10809}, *I(2)k10217*, *stau*^{ly9} y CG3249^{EP1440}, de los cuales sólo la mutación *I(2)k10217* interaccionó con el fenotipo de ojos rugosos generado por la sobreexpresión de *mbIC*. La interacción de esta línea es potenciadora al igual que la interacción con el modelo de *SCA8*.

La inserción *I(2)k10217* queda localizada entre los genes *CG15625* y *CG3036*. No hay estudios funcionales para ninguno de los dos genes. *CG15625* codifica para una proteína que posee dominios SAM (*Sterile Alpha Motifs*). Estos dominios aparecen en una gran variedad de proteínas implicadas en muy diversos procesos biológicos. Los dominios SAM suelen estar asociados a una interacción proteína-proteína. Sin embargo, se ha comprobado su capacidad de unión a RNA en otras proteínas. Por ejemplo, la proteína Smaug ayuda a establecer el gradiente morfogénico en los embriones de *Drosophila* reprimiendo la traducción de *nanos* mediante unión a la región 3-UTR de su mensajero (Kim, 2003). Por otro lado, el gen *CG3036* codifica una proteína transmembrana con actividad simportador sodio:fosfato. Datos de expresión a partir de un microarray “*developmental time course*” indican una baja expresión de transcritos *CG3036* en los primeros estadios del desarrollo embrionario y un pico de expresión en prepupa (<http://genome.med.yale.edu/Lifecycle/>).

d) Apoptosis

Como parte de este trabajo, hemos mostrado que la sobreexpresión de *mbIC* en el ala (*en-GAL4*) presenta un fenotipo de pérdida de tejido intervena en alas, debido a una inducción de la apoptosis en discos de tercer estadio. Para comprobar si *muscleblind* está relacionado con esta función biológica o si, por el contrario, la inducción de la apoptosis se produce tan solo como consecuencia del efecto tóxico de la sobreexpresión del *mbI*, comprobamos la interacción de distintos genes implicados en apoptosis con el fenotipo de ojos rugosos.

Los líneas estudiadas son: la delección *Df(3L)4W*, que abarca los genes proapoptóticos *head involution defective (Hid)*, *reaper (rpr)* y *grim*, la línea de sobreexpresión de *reaper GMR>rpr*, los alelos por falta de función del gen *thread Diap^{6-3s}*, *th⁴* y *th⁵* y, por último, dos alelos del gen *DiapII*, *DiapII^{K02205}* de falta de función y *GMR>DiapII* de sobreexpresión.

Así, hemos encontrado que la falta de función de los tres genes proapoptóticos *Hid*, *rpr* y *grim* suprime el fenotipo mientras que la sobreexpresión del gen *reaper* en el ojo (*GMR-GAL4*) junto con la sobreexpresión en el ojo de *mbIC* (*sev-GAL4*) resulta letal. Por otro lado, los dos alelos del gen antiapoptótico *thread*, *th⁴* y *th⁵*, resultan potenciadores del fenotipo de ojos rugosos. El hecho de que hayamos obtenido

interacción con *thread*, también llamado *Diapl*, y no con *Diapll*, un gen antiapoptótico muy relacionado, apoya la especificidad de esta interacción.

Ruta de patogénesis de la DM

Deducciones a partir de la información publicada nos han llevado a testar varios genes por su potencial para interactuar con *muscleblind* tales como los homólogos de CUG-BP en *Drosophila*, *aret*, *bruno-2* y *bruno-3* (Ho, 2004) y proteínas que colocalizan con *mbI* en los foci ribonucleoproteicos que contienen las expansiones CTG, tales como *nonA* y *Hsp70A* (Houseley, 2005; Jiang, 2004).

Las líneas por pérdida de función *aret*^{GS12289} y *nonA*^{Df(1)}, así como la línea de sobreexpresión de *Hsp70A*^{EY01148} mostraron una clara interacción con el fenotipo de ojos rugosos producido por la sobreexpresión de *mbIC* de manera que todos ellos actúan como potenciadores.

Interacción en un sistema de dos híbridos

En un ensayo de dos híbridos a escala genómica, Giot y colaboradores encontraron una interacción entre las proteínas Muscleblind y Jumu. Por este motivo, comprobamos si existía una interacción genética entre los dos genes encontrando que, el alelo mutante *Jumu*^{L70} actúa claramente como supresor del fenotipo de sobreexpresión de *mbIC*.

La Tabla 2 resume el resultado de la búsqueda de genes que interactúan genéticamente con la sobreexpresión de *mbIC* en ojo identificados en este trabajo

Tabla 2. Genes que interactúan genéticamente con la sobreexpresión de *muscleblind*. Para cada gen se indica su nombre, líneas empleadas y tipo de interacción con cada una de ellas. La función de cada gen se ha anotado en base a la bibliografía disponible y/o las anotaciones del Flybase. Abreviaturas empleadas son PEV “variegación debido al efecto de posición”, FKH/WH, “Forkhead/winged-helix”

Gen	Función	Sobreexpresión	Pérdida de función	Interacción
<i>CG15435</i>	contiene dedos de zinc C2H2	<i>CG15435</i> ^{GS11154}	<i>CG154435</i> ^{KG05006}	Supresor Supresor
<i>turtle</i>	contiene dominios inmunoglobulina; proteína con actividad kinasa implicada en adhesión célula-célula y control motor	<i>tutl</i> ^{GS13843}		Supresor
<i>TNF-receptor-associated factor 1</i>	contiene RING fingers; implicado en la ruta JNF; induce apoptosis	<i>Traf1</i> ^{GS2154}		Supresor
<i>CG15433</i>	actividad acetiltransferasa	<i>CG15433</i> ^{NP434}	<i>CG15433</i> ^{KG02386}	Supresor Supresor
<i>CG5790</i>	proteína kinasa serina/treonina; implicada en la forma celular	<i>CG5790</i> ^{NP1233}		Letal
<i>CG17424</i>	contiene motivos <i>stomatin</i> , proteína de membrana implicada en la regulación de cationes monovalente y organización del citoesqueleto	<i>CG17424</i> ^{NP7067}		Potenciador
<i>CG17323</i>	actividad UDP glucosil transferasa	<i>CG17323</i> ^{GS9130}	<i>CG17323</i> ^{KG03187}	Potenciador Supresor
<i>CG31751</i>	contiene motivo aminoglucósido fosfotransferasa	<i>CG31751</i> ^{NP1565}		Potenciador
<i>Aly</i>	componente del <i>exón junction complex</i> (EJC); actúa sobre mensajeros concretos; transporte núcleo-citoplasma; regulación de la traducción		<i>Aly</i> ⁰²²⁶⁷	Supresor
<i>DP transcription factor</i>	factor de transcripción E2F/DP; regulación ciclo celular y apoptosis		<i>Dp</i> ^{KG00660}	Potenciador
<i>absent MD neurons and olfactory sensilla</i>	contiene dominios de unión a DNA Helix-Loop-Helix (HLX); factor de transcripción implicado en neurogénesis	<i>amos</i> ^{Roi1} <i>amosUAS</i>	<i>amos</i> ^{Tr-RM11}	Potenciador Potenciador Letal
<i>suppression of variegation 2-10</i>	contiene motivos de unión a DNA; supresor del PEV; implicado en la determinación y mantenimiento de la estructura de la cromatina		<i>su(var)2-10</i> ⁰³⁶⁹⁷	Supresor
<i>CG15625</i>	actividad metiltransferasa dependiente de S-adenosylmethionine que afecta DNA, RNA y proteínas con propósitos reguladores		<i>CG15625</i> ^{k10217}	Potenciador
<i>reaper</i>	contiene dominios Proteinase inhibitor I32 (IAP); inductor de la apoptosis; implicado en la diferenciación sistema nervioso central	<i>GMRGal4<UASrpr</i>		Letal
<i>thread</i>	ubiquitina ligasa que actúa como inhibidor de la apoptosis a través de la inhibición de las caspasas		<i>th</i> ⁴ y <i>th</i> ⁵	potenciadores
<i>aret</i>	contiene motivo de unión a RNA RNP-1; factor de splicing		<i>aret</i> ^{GS12289}	Potenciador
<i>no on or off transient A</i>	contiene motivo de unión a RNA RNP-1; factor de splicing, implicado en función visual y sonidos de cortejo		<i>nonA</i> ^{Df(1)}	Potenciador
<i>jumeaux</i>	dominios de unión a DNA FKH/WH, supresor del PEV; determinación y mantenimiento de la estructura cromatina; implicado en neurogénesis y cierre dorsal		<i>Jumu</i> ^{L70}	Supresor
<i>Heat shock protein 70Ab</i>	proteína chaperona; respuesta al <i>stress</i>	<i>Hsp70Ab</i> ^{EY01148}		Potenciador

2.1.2 Análisis de la interacción con *jumeaux*

Existen varias razones por las que consideramos que *jumeaux* es un gen interesante para un estudio más detallado en relación a *muscleblind*. Una de las razones viene dada por el experimento de Giot y colaboradores (2003). En este estudio se presenta un mapa de interacciones proteicas a escala genómica y se realiza un ensayo de dos híbridos entre más de 4500 proteínas de *Drosophila*. *muscleblind* es uno de los cDNAs ensayados y se encuentran 38 interacciones diferentes. Entre las 19 proteínas identificadas con una función conocida, al menos cinco tienen una actividad como factor de transcripción. Una de estas proteínas está codificada por el gen *jumeaux*. Aunque las evidencias publicadas hasta la fecha apuntan hacia una función de *mbf* sobre el metabolismo del RNA, este dato experimental, junto con el hecho de encontrar varios genes relacionados con la función transcripcional que interactúan genéticamente (*CG15435*, *CG15433*, *Dp*, *amos*), nos hizo pensar en la función de Mbl como proteína implicada en los procesos de la transcripción y remodelación de la cromatina.

Otra de las razones por las que *jumeaux* resulta interesante es por su patrón de expresión y el fenotipo de sus mutantes. Hibridaciones *in situ* e inmunodetecciones muestran una fuerte señal en embriones tempranos (estadio 1 a 4) lo que sugiere su contribución materna. A partir del estadio 7 la señal disminuye y los transcritos empiezan a localizarse en regiones concretas que se corresponden con regiones neurogénicas, células cardíacas y precursores gonadales. Al final de la embriogénesis los transcritos de *jumeaux* se encuentran en el sistema nervioso central (SNC), periférico (SNP), células maxilares, y células precursoras de los discos imaginales (Cheah, 2000). Experimentos de hibridación e inmunodetección en discos imaginales de larvas de tercer estadio muestran fuertes señales especialmente en células neurogénicas. En disco de ojo, la señal aparece detrás del surco morfogenético y en los ommatidios en formación (Stroedicke, 2000). Los mutantes heterocigotos *jumeaux* no tienen ningún fenotipo visible. Sin embargo, mutaciones en este gen son letales en homocigosis. Un pequeño porcentaje de mutantes homocigotos para *jumeaux* consiguen llegar a la edad adulta. En ellas se puede observar un fuerte fenotipo de ojos rugosos donde no existe una estructura externa debido a una falta de células pigmentarias (que normalmente organizan a los ommatidios). El fenotipo de quetas mutantes indica una función de *jumu* en el SNP y consiste en un defecto de tipo numérico, no de localización. Por último, las alas de los homocigotos para *jumu* presentan un tamaño reducido, una disposición irregular de los pelos, muescas en el

margen posterior y a veces, la vena L5 acortada (Strodicke, 2000). Por tanto, el patrón de expresión de *jumeaux* es solapante al de *muscleblind*, al menos en tejidos como ojo, SNC y SNP. Además, los fenotipos de ambos mutantes resultan similares en cuanto a las estructuras afectadas como son los ojos, alas y quetas lo que puede indicar que están implicados en los mismos procesos biológicos.

Así pues, comprobamos si la mutación por falta de función de *jumeaux*, *jumu*^{L70}, interacciona genéticamente con el gen *muscleblind* mediante la modificación del fenotipo de sobreexpresión de *mbIC* en ojo. Además, comprobamos si *jumu*^{L70} interacciona con un “alelo” por falta de función de *muscleblind* como es la sobreexpresión de CTG₄₈₀. Según la hipótesis de trabajo de nuestro laboratorio, la sobreexpresión de las repeticiones CUG en el ojo de *Drosophila* provoca una falta de función de la proteína Mbl endógena debido a su secuestro por parte de dichas repeticiones. De acuerdo con esta hipótesis, la interacción genética que pudiera observarse en este cruce se debe, principalmente, a la falta de función de los genes *mbI* y *jumeaux*.

Los resultados indican que *jumu*^{L70} interacciona genéticamente con la sobreexpresión de *mbIC* pero no con la sobreexpresión de CTG₄₈₀ en ojo. Sin embargo, se comprobó la interacción con otro alelo de este gen, *jumu*⁴³¹², sí que interacciona, tanto con *mbIC* como con CTG₄₈₀. En ambos casos, la falta de función de *jumeaux* actúa como un supresor dominante del fenotipo de ojos rugosos.

El gen *jumeaux* codifica una proteína que tiene dominios de unión a DNA del tipo FKH/WH normalmente asociados a proteínas que actúan como factor de transcripción. Además, actúa como supresor dominante del fenotipo de variegación por efecto de posición (PEV). Su efecto sobre el PEV junto con la localización en los cromosomas politénicos sugiere que *jumu* tiene una función como remodelador de la cromatina en el establecimiento y/o mantenimiento de su estructura (Strodicke, 2000).

Para determinar la relevancia de la interacción génica detectada entre *jumeaux* y *muscleblind*, y en colaboración con laboratorio del Dr. Günter Korge (Universidad de Berlin, Alemania), nos propusimos comprobar (a) si *jumeaux* interacciona con otros fenotipos mutantes de *muscleblind*, (b) si *Muscleblind* puede localizarse en los cromosomas politénicos y si lo hace en los mismos sitios que *jumu* y c) si existe una interacción física entre las dos proteínas.

a) Interacción génica de *jumeaux* con otros fenotipos de *muscleblind*

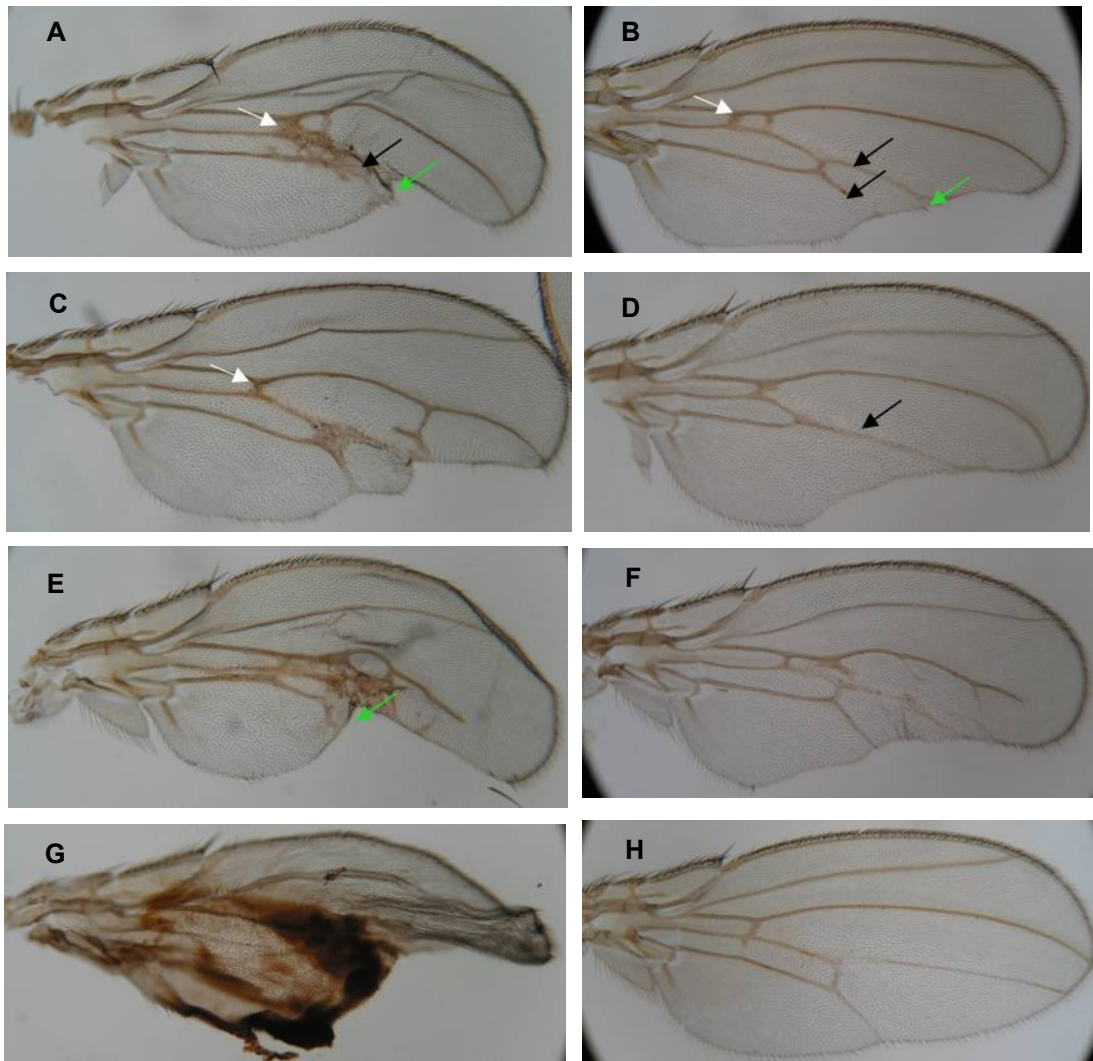
En primer lugar, comprobamos si existía una interacción entre *jumeaux* y *muscleblind* en ala ya que una interacción génica en otro sistema diferente al ojo refuerza la idea de la participación de estos dos genes en un mismo proceso celular. Para ello, comprobamos si la mutación *jumu*^{L70} era capaz de modificar el fenotipo característico de la sobreexpresión de *mbIC* en ala (*en-GAL4>UAS-mbIC*). Los resultados indican claramente que *jumeaux* y *muscleblind* interaccionan también en este órgano. En ala, la presencia de la mutación *jumu*^{L70} actúa en el mismo sentido que en ojo, esto es, como supresor dominante (Figura 5 - I).

Figura 5. *Jumeaux* interacciona genéticamente con la sobreexpresión de *mbIC* en ala. Cuando se estudia el efecto de la reducción a la mitad de la dosis génica de *jumeaux* sobre el fenotipo de sobreexpresión de *mbIC* en ala se observa una disminución en la gravedad del fenotipo. I) La columna de la izquierda presenta preparaciones de alas de la sobreexpresión de *mbIC* en disco de ala con la línea *engrailed* (*w; enGal4/TM3; UASmbIC/+*) que constituyen las alas control. La columna de la derecha presenta preparaciones de alas con el mismo fondo genético que las anteriores pero con el alelo mutante *jumu*^{L70} en heterocigosis (*Jumu*^{L70} *enGal4/TM3; UASmbIC/+*) a las que abreviadamente llamaremos alas *jumu*. II) Tabla en la que se muestran las observaciones recogidas sobre 12 alas control y 17 alas *jumu* y que demuestran que *jumu* es un supresor dominante del fenotipo de sobreexpresión de *mbI* en ala. IVA, interviena anterior

I)

*w; enGal4/TM3>UAS**MbIC***

*Jumu^{L70/+}; enGal4/TM3>UAS**MbIC***



II)

FENOTIPO	CONTROL	JUMU
presencia de mordisco extremo	50%	0%
presencia de mordisco suave	50%	94%
total de moscas con defectos de fusión L4 y L5	92%	71%
L4 y L5 se fusionan y desaparecen ambas	36%	0%
L4 y L5 se fusionan y desaparecen una de ellas	18%	33%
L4 y L5 se fusionan y se separan, no desaparece ninguna	46%	67%
IVA afectada (tamaño, definición...)	67%	41%

(Figura 5. ver pie de Figura en hoja anterior)

Las observaciones que demuestran que *Jumu*^{L70} es un supresor del fenotipo de sobreexpresión de *mb1C* con con el patrón de expresión de *engrailed* se resumen a continuación (Figura 5):

1. El fenotipo más extremo de ala “quemada” sólo se observa en la muestra control (8%) mientras que los fenotipos más suaves, esto es, que más se parecen a un ala normal, sólo se observan en la muestra con la mutación en *jumeaux* (12%).
2. Si dividimos el fenotipo de falta de tejido en el margen posterior del ala como (a) mordisco suave o (b) mordisco extremo el fenotipo más extremo sólo se observa en la muestra control (50%) y no en la muestra jummu (0%).
3. En algunas alas control se observan alteraciones de la vena L2, que consisten básicamente en una alteración de su forma normal apareciendo un arco o pico en el primer tercio de su longitud (33%). Esto no se observa en ninguna ala jummu.
4. El porcentaje de alas donde se observa que la intervena anterior (IVA) esta afectada, bien en cuanto a su presencia/ausencia, tamaño o definición, se reduce considerablemente en las alas jummu (67% → 41%), mostrando nuevamente una reducción en la gravedad de los efectos deletéreos de la sobreexpresión de *mb1C* en la ala.
5. Finalmente, las proporciones relativas de los fenotipos de fusión de las venas L4 y L5 tienden a ser menores para aquellas categorías más extremas. El fenotipo más extremo que se puede observar en la desaparición completa de ambas venas tras su fusión. Una vez más, esto solo se observa en las alas control.

El defecto más evidente que se produce cuando sobreexpresamos *mb1C* en ala con el patrón de expresión de *engrailed* es la pérdida de tejido laminar intervena que identificamos como un “mordisco” en el compartimento posterior. Tinciones α -caspasa 3 realizadas en discos imaginales muestran activación de la apoptosis en una región concreta del disco, que podría explicar la falta de tejido en el ala adulta (ver resultados 2.1). Este defecto fenotípico queda suprimido considerablemente cuando reducimos a la mitad la dosis del gen *jumeaux*. Estos resultados podrían indicar una disminución en la inducción de la apoptosis, por lo que ambos genes podrían participar en el proceso de apoptosis.

b) Localización de Mbl en cromosomas politénicos

Dada la función de *jumeaux* como factor de transcripción y la potencial función de *muscleblind* (dedos de zinc) en este sentido podemos comprobar, primero, si podemos localizar proteína Mbl en los cromosomas politénicos y, segundo, si la proteína Mbl y la proteína Jumu presentan un patrón de localización cromosómica solapante.

Usando las moscas transgénicas UAS-*mbIC*:GFP (ver material y métodos 2.6), dirigimos la sobreexpresión de la proteína de fusión a las glándulas salivares utilizando la línea *saga48E*-GAL4 y comprobando que MblC:GFP tenía una localización nuclear. Hicimos preparaciones de cromosomas politénicos y detectamos la proteína MblC:GFP usando un anticuerpo anti-GFP que permitió localizar la proteína de fusión en diferentes posiciones cromosómicas (Figura 6).

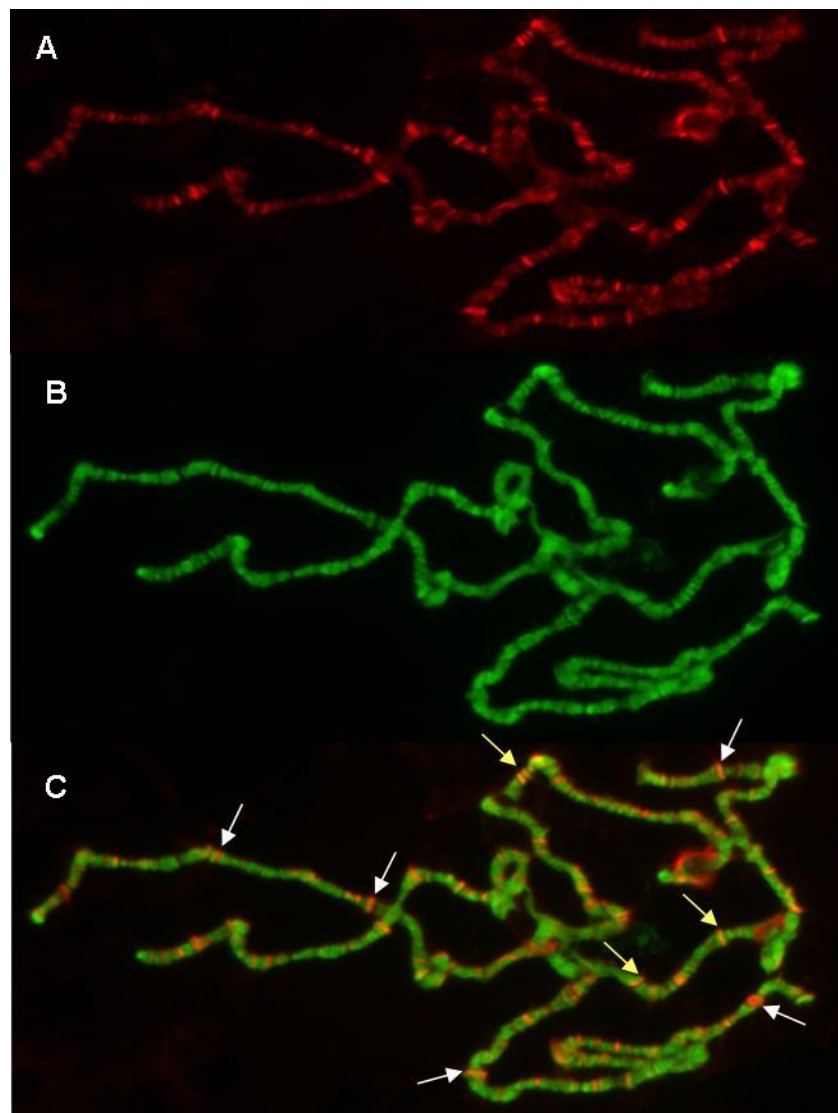


Figura 6. ver pie figura página siguiente

Figura 6. Localización de la proteína Muscleblind en los cromosomas politénicos de *Drosophila melanogaster*. Utilizando moscas transgénicas UAS-MblC:GFP hemos sobreexpresado la proteína en las glándulas salivares usando la línea saga48E-GAL4 y realizado preparaciones de cromosomas politénicos a partir de glándulas salivares pertenecientes a larvas de tercer estadio. (A) inmunodetección de MblC:GFP mediante anti-GFP y secundario acoplado al fluoróforo Cys3 (rojo) (B) visualización de la heterocromatina mediante tinción con colorante Hoechst (verde). (C) Imagen solapada de la tinción Hoechst y de la señal que identifica la proteína MblC:GFP. Se puede observar que Muscleblind se localiza en un número limitado de loci cromosómicos (~11/brazo). La señal MblC:GFP se localiza principalmente en la eucromatina y se puede ver en loci transcripcionalmente muy activos (*puffs*, flechas blancas) pero también en regiones heterocromáticas (flechas amarillas)

La proteína Muscleblind etiquetada (Mbl:GFP) se observa en lugares específicos de los cromosomas politénicos por lo que concluimos que Mbl ejerce su función sobre un grupo determinado de aproximadamente 60 genes. La presencia de Mbl en estos loci cromosómicos es congruente con una actividad como factor de transcripción, apoyado por la señal en sitios con una alta tasa de transcripción como son los *puffs*, o como factor de *splicing*.

Una inmunotinción doble con anticuerpo anti-GFP y anti-Jumu permitió observar los puntos de solapamiento entre las dos proteínas. Las inmunotinciones sobre cromosomas politénicos muestran que Mbl y Jumu tienen un patrón de localización parcialmente solapante. La intensidad de las señales para Jumu y Mbl varía pero ambas se localizan predominantemente en las interbandas, el cromocentro y algunos telómeros. Una gran diferencia entre ellas es que Jumu tiñe el nucleolo mientras que Mbl no lo hace (Figura 7).

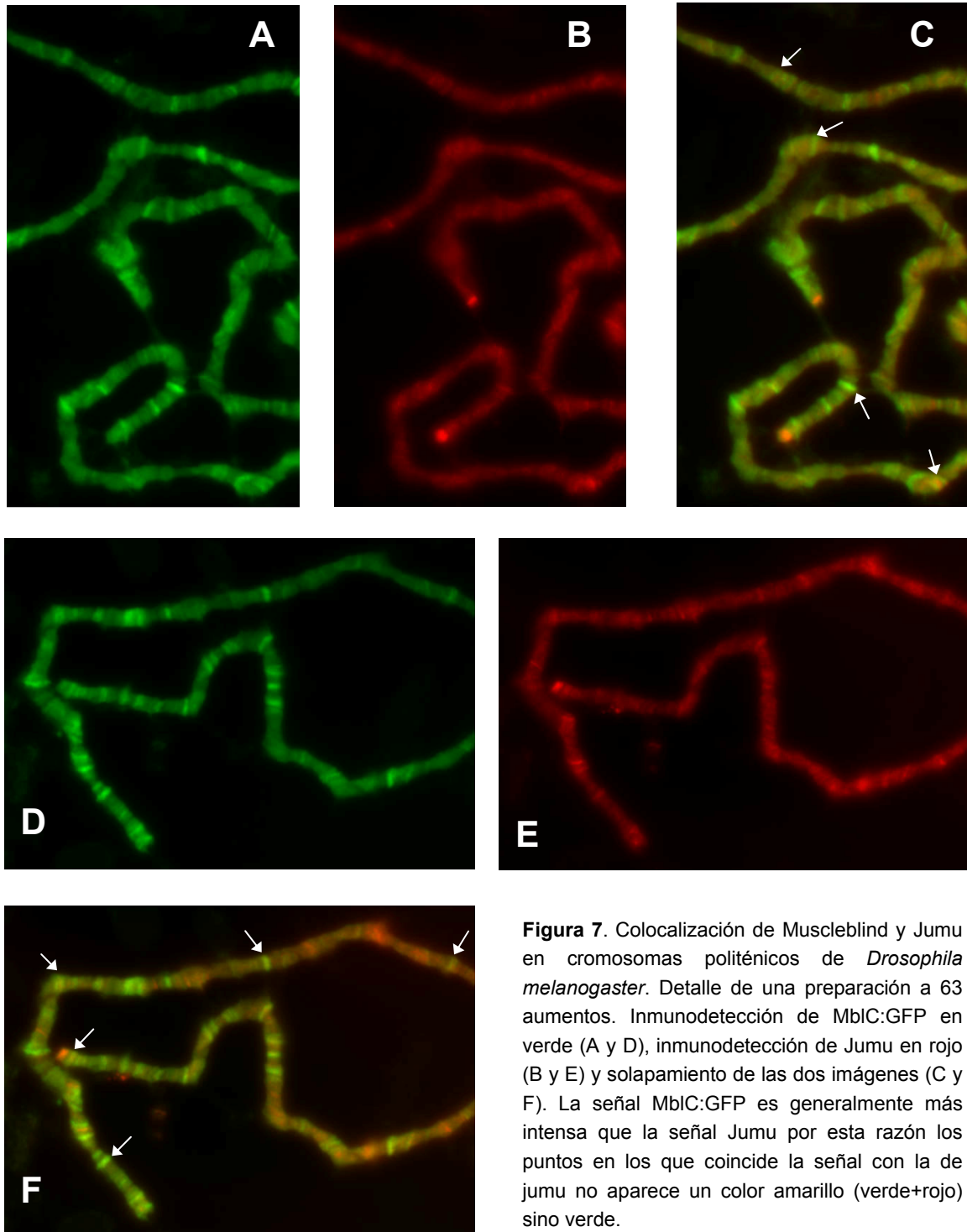


Figura 7. Colocalización de Muscleblind y Jumu en cromosomas politénicos de *Drosophila melanogaster*. Detalle de una preparación a 63 aumentos. Inmunodetección de MbIC:GFP en verde (A y D), inmunodetección de Jumu en rojo (B y E) y solapamiento de las dos imágenes (C y F). La señal MbIC:GFP es generalmente más intensa que la señal Jumu por esta razón los puntos en los que coincide la señal con la de jumu no aparece un color amarillo (verde+rojo) sino verde.

c) Interacción física entre Jumu y Muscleblind

La inmunoprecipitación (IP) es una técnica mediante la cual “precipitas” una proteína mediante el uso de un anticuerpo específico. El fundamento de la coinmunoprecipitación (coIP) se basa en que si dos proteínas interactúan, bien directamente o a través de intermediarios, cuando IP una de ellas en las condiciones que permitan mantener las uniones, la otra proteína también va a ser inmunoprecipitada.

Hemos puesto a punto un protocolo de coIP para precipitar la proteína de fusión MbIC:GFP a partir de homogenados de moscas transgénicas (ver resultados 2.2.4) de manera que, mediante Western blot, podemos comprobar si la proteína Jumu está presente en ese inmunoprecipitado o no. Aunque pudimos detectar correctamente la proteína jumu en las carreras control, no se pudo detectar la proteína jumu en los coinmunoprecipitados de MbIC:GFP (ver resultados 2.5). Este resultado nos permite concluir que Jumu no interactúa físicamente con Muscleblind en las condiciones utilizadas. Los resultados obtenidos en el laboratorio del Dr. Günter Korge con experimentos de dos híbridos y pull-down entre Jumu y Muscleblind confirma estos resultados (A. Schwendemann, comunicación personal)

En resumen, la sobreexpresión de la isoforma C del gen *muscleblind* en ojo genera un fenotipo de ojo rugoso que hemos utilizado en un rastreo de modificadores dominantes. Como resultado de este trabajo hemos identificado 19 genes que interaccionan genéticamente con *muscleblind* (Tabla 2). De ellos, más de la mitad (10) son genes sobre los que no existen estudios funcionales. Sin embargo, la función inferida a partir de los motivos que presentan sus proteínas o de los rastreos en los que han sido identificados permite tener una visión global de los procesos celulares en los que *muscleblind* podría estar implicado. Podemos distinguir cuatro funciones generales en el grupo de genes que interaccionan con Muscleblind (Figura 8) como son: (1) Metabolismo del RNA, en el que se incluyen componentes del *exon junction complex* como Aly o factores de *splicing* como aret y nonA; (2) Factores de Transcripción, que incluye proteínas de unión a DNA que activan la transcripción como Dp o amos y proteínas que regulan la estructura de la cromatina como jumeaux o su(var)2-10; (3) Regulación de la apoptosis, con proteínas como thread, reaper o Traf; y (4) otros, entre los que se incluirían distintas funciones enzimáticas implicadas en la adhesión célula-célula como tutl, determinación de la forma celular, CG5790, y organización del citoesqueleto, CG17424.

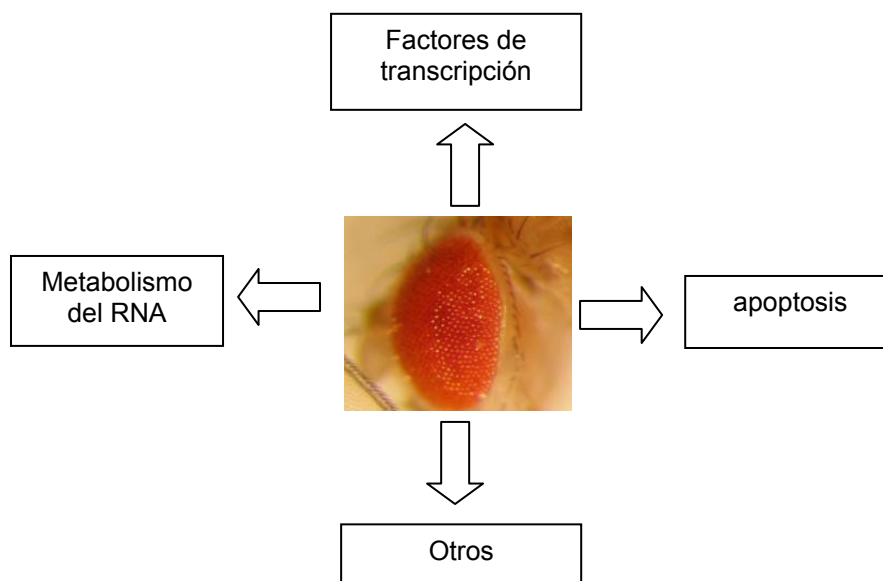


Figura 8. Categorías de genes que actúan como modificadores dominantes del fenotipo de sobreexpresión de *mbIC* en ojo (*sevenless*-GAL4). Estos resultados sugieren nuevas funciones para *mbI* como son la regulación de la transcripción o la apoptosis.

2.2 APROXIMACIÓN BIOQUÍMICA PARA LA IDENTIFICACIÓN DE DIANAS MOLECULARES DE MUSCLEBLIND

2.2.1 Obtención de un anticuerpo α - Muscleblind

Al inicio de este trabajo no existía ningún anticuerpo contra la proteína Muscleblind ya que el anticuerpo producido por Begemann y colaboradores (1998) se había terminado. Por ello, generamos un anticuerpo inmunopurificado, α -Mbl_{CUK331}, contra el péptido CUK-333 que cubre los últimos 13 aminoácidos compartidos por las isoformas A, B y C de Muscleblind (material y métodos 2.9). El test Elisa indicó una buena inmunoreactividad del anticuerpo purificado contra el péptido CUK-333 (Figura 1A). Para comprobar la inmunoreactividad del anticuerpo α -Mbl_{CUK331} contra la proteína Muscleblind inmovilizada en membrana PVDF realizamos un dot blot que mostró que el anticuerpo reconocía la proteína Muscleblind aunque débilmente (Figura 1B). Sin embargo, no se consiguió detectar proteína Muscleblind mediante Western blot a pesar de modificar gran cantidad de parámetros como la concentración de anticuerpo, la cantidad de muestra cargada, la solución bloqueante empleada, o el método de transferencia a membrana.

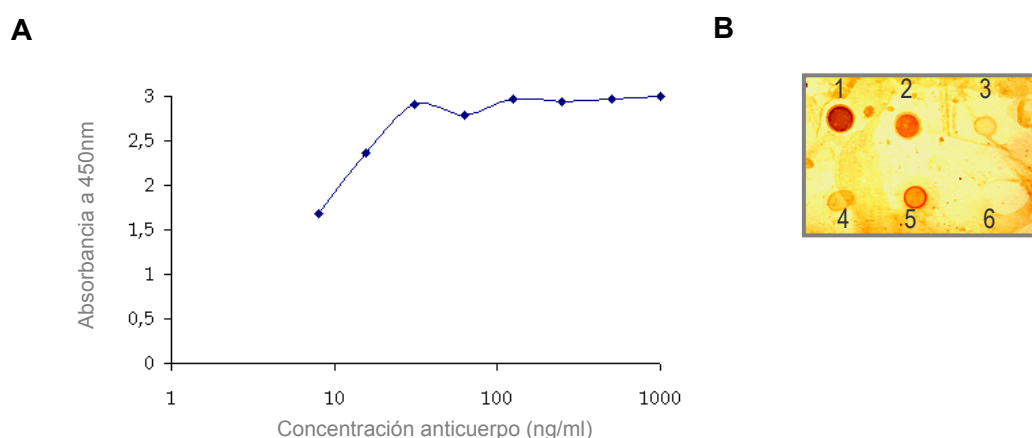


Figura 1. Inmunoreactividad del anticuerpo α -Mbl_{CUK333}. A) Test ELISA, la mínima concentración de anticuerpo que tiene una mayor reactividad es de aproximadamente 55 ng/ml. B) dot blot, hay señal en todas las carreras excepto en el control negativo, BSA. Las muestras y la cantidad de proteína total utilizada son: 1, mosca OrR adulta (42 μ g); 2, embriones OrR 0-24h (40 μ g); 3, MblA:GST purificada (2.5 μ g); 4, péptido CUK333 (1 μ g); 5, péptido CUK333 (10 μ g) y 6, proteína BSA (10 μ g). La concentración de anticuerpo empleada fue de 104 ng/ml (dilución 1/500).

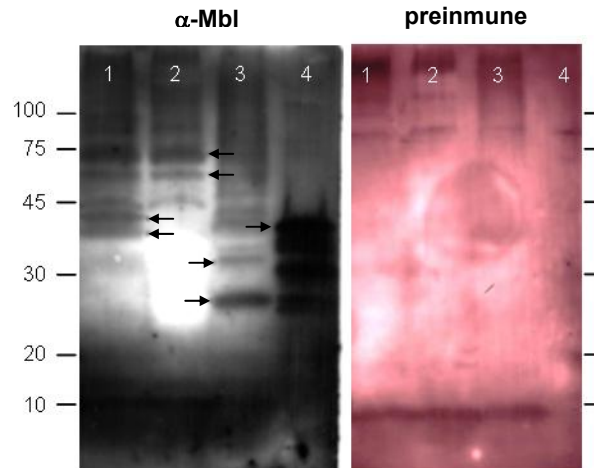


Figura 2. Western blot α -Mbl. Las muestras utilizadas se corresponden con 1, extracto de moscas silvestres adultas; 2, Extracto de mutantes hipomorfos *mbI* (*mbI^{E27}/mbI^{E16}*); 3, extracto de embriones 0-24h silvestres; 4, proteína purificada MbIA-GST. El suero contra Muscleblind reconoce las proteínas Mbl (flechas) ya que no se detecta señal con el suero preimmune y hay una fuerte señal en la carrera que contiene proteína MbIA-GST purificada que coincide con el tamaño esperado de la proteína. Sin embargo, la mayoría de las bandas tienen un tamaño que no se corresponde con ninguna isoforma de *muscleblind* conocida. Además, el suero α -Mbl genera un gran ruido de fondo y no permite obtener resultados reproducibles con facilidad. La cantidad de proteína total cargada es de 40 μ g excepto para la proteína purificada MbIA:GST que es de 1 μ g. Los sueros se han usado a 1:5000, al igual que el anticuerpo secundario (anti-oveja conjugado a HRP)

Por otro lado, en el laboratorio del Dr. Monckton de la Universidad de Glasgow (Reino Unido), se obtuvo un antisuero contra la proteína Muscleblind completa. Este suero, α -Mbl, se generó inyectando en oveja las isoformas proteicas A, B y C de Mbl fusionadas a la etiqueta GST. Gracias a una colaboración con este grupo, tuvimos acceso al suero α -Mbl y probamos su funcionamiento en Western blot (Figura 2). El suero reconoce las proteínas Muscleblind ya que detecta MbIA:GST y distintas isoformas proteicas de Mbl (flechas, Fig.2). Sin embargo, también detecta otras proteínas inespecíficamente ya que no se corresponden con ninguna isoforma conocida de Mbl. En Western blot, el suero α -Mbl genera un gran ruido de fondo y los resultados no son reproducibles por lo que no resulta adecuado para aplicaciones bioquímicas

A pesar de los esfuerzos realizados para conseguir un anticuerpo contra la proteína Muscleblind que sirviera para aplicaciones bioquímicas no hemos obtenido resultados satisfactorios.

2.2.2 Obtención de moscas transgénicas UAS*mbIC:GFP*

Para inmunoprecipitar la proteína Muscleblind (Mbl) de *Drosophila*, y debido a la inexistencia de un anticuerpo adecuado para esta aplicación, procedimos a generar moscas transgénicas capaces de expresar la proteína de fusión MblC:GFP gracias al sistema Gal4 (apéndice 2.2.6.1). La adición de la etiqueta GFP a la proteína Muscleblind tiene varias ventajas. En primer lugar, permite detectarla fácilmente mediante un anticuerpo específico anti-GFP. Además, su inserción en una proteína no suele afectar seriamente la función de la misma, dado el plegamiento en paralelo de sus extremos amino y carboxilo, que le confiere una estructura en farol (Marshall, 1995). Por último, el uso de la etiqueta GFP nos permite visualizar la proteína de fusión *in vivo* pudiendo comprobar de una manera inmediata que el transgén se expresa correctamente. Anteriormente, la proteína GFP ha sido utilizada con éxito como proteína de fusión para generar moscas transgénicas en aproximaciones similares (e.j Wilhelm, 2000)

Para reducir el número de cruces genéticos necesarios para llevar a cabo el experimento de rescate (resultados 2.2.3), se utilizaron embriones heterocigotos para el alelo *mb^{E27}* en la transformación de la línea germinal. De aproximadamente 800 embriones microinyectados obtuvimos 5 líneas transformantes procedentes todas ellas de un único progenitor. Los cruces genéticos realizados para mapear cada una de las inserciones indicó que en una de las líneas transformantes, el transgén se situaba en el cromosoma X mientras que en las otras cuatro lo hacía en el cromosoma 3 (Tabla 1) por tanto, obtuvimos, al menos, dos líneas transgénicas que representan sucesos independientes de inserción en el genoma. Todas las líneas se hicieron homocigotas para la inserción. La línea T15.3 fue amplificada y usada para todos los experimentos que se detallan en este trabajo. De aquí en adelante, haremos referencia a esta línea como UAS-*mbIC:GFP*.

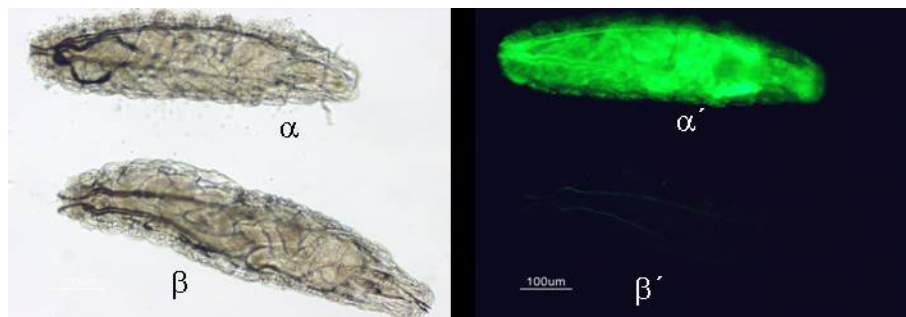
Línea transgénica	Cromosoma	Alelo
T15.1	X	<i>mb^{E16}</i>
T15.2	3	<i>mb^{E27}</i>
T15.3	3	<i>mb^{E27}</i>
T15.4	3	<i>mb^{E16}</i>
T15.5	3	<i>mb^{E16}</i>

Tabla 1. Líneas establecidas con el transgén UAS-*mbIC:GFP*. Localización cromosómica y fondo mutante para *muscleblind* (material y métodos 2.6)

2.2.3 La proteína MbIC:GFP es funcional

Para comprobar que la proteína de fusión MbIC:GFP era capaz de expresarse correctamente cruzamos la línea *UAS-mbIC:GFP* por una línea *da-GAL4*. Haciendo uso de la cualidad de emitir fluorescencia de la proteína MbIC:GFP, se observó, bajo la lupa de fluorescencia, que todos los embriones fecundados y larvas resultantes de este cruce emitían una fuerte fluorescencia verde (Figura 3A). Esto confirmó que el DNA de la construcción insertada era capaz de transcribirse y traducirse correctamente. El tamaño esperado de la proteína de fusión se comprobó analizando extractos de estos embriones mediante la técnica de Western blot (Figura 3B).

A



B

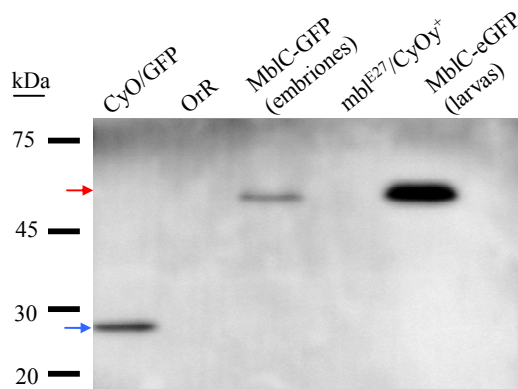


Figura 3. Expresión de la proteína MbIC:GFP. A) Imagen del microscopio de fluorescencia donde se puede apreciar el color verde característico de la proteína GFP en una larva que expresa la construcción *UAS-mbIC:GFP* con el patrón general de expresión ($\alpha \rightarrow \alpha'$) en contraste con una larva silvestre ($\beta \rightarrow \beta'$). B) El anticuerpo anti-GFP detecta la proteína MbIC:GFP como una banda de ~52k Da (flecha roja). La proteína GFP tiene un tamaño de ~29k Da.

Para comprobar que la proteína de fusión MbIC:GFP es funcional *in vivo*, realizamos un experimento de rescate. En este tipo de experimentos se realizan los cruces genéticos necesarios para construir un genotipo en el que, por un lado, se da la falta de función que se quiere estudiar (mutantes con un fenotipo concreto) y por otro, se expresa la proteína de interés. Si el fenotipo mutante cambia hacia el fenotipo silvestre, esto constituye una prueba de que dicha proteína es capaz de llevar a cabo esa función (Monferrer, 2005). El sistema GAL4 permite llevar a cabo esta expresión dirigida a un tejido concreto o de manera ubicua.

La mutación *mbI*^{E27} ha sido descrita como letal en fase embrionaria tardía (Begemann, 1997; Artero, 1998). En nuestro caso, queríamos saber si los mutantes nulos *mbI*^{E27} eran capaces de continuar su desarrollo al proporcionarles función *mbI* en forma de proteína de fusión MbIC:GFP. Puesto que no se conocen las regiones reguladoras del gen *mbI*, hicimos uso de una línea GAL4 que dirige la expresión del transgén diana de forma ubicua en el embrión de *Drosophila* (*da*-GAL4), solapando su expresión con la de *mbI* endógeno y por tanto con la de sus dianas moleculares.

En un primer experimento, confirmamos la letalidad embrionaria de la mutación *mbI*^{E27} y establecimos el porcentaje de mutantes *mbI* que superaban dicha letalidad, según nuestras condiciones experimentales. El porcentaje de larvas *mbI*^{E27} nacidas fue de 1.3% (n= 1334, 25°C). Posteriormente, comprobamos el grado de rescate que era capaz de alcanzar la proteína MbIC:GFP de *Drosophila* y lo comparamos con los datos del rescate obtenidos para la proteína MbIC sola. Los resultados, a 25°C, indican que la expresión de la proteína de fusión MbIC:GFP es igual de efectiva que la proteína MbIC sola en el rescate del fenotipo de letalidad embrionaria de los mutantes nulos para *mbI* aunque la diferencia no es significativa (fig. 4). Además, la cantidad de rescate es sensible a la cantidad de proteína MbIC:GFP expresada, ya que los valores de larvas nacidas son mayores para 25°C que para 19°C.

Para evaluar el período de supervivencia de las larvas rescatadas, estas se mantuvieron en placas con comida y se observaron varias veces al día. Tanto para las larvas rescatadas por la proteína MbIC:GFP, como por la proteína MbIC sola se observó una supervivencia de hasta tres días (48-72h). Además, y de manera similar a lo que ocurre con la sobreexpresión de *mbI*, se determinó la letalidad generada por la proteína MbIC:GFP al observarse que las larvas que sobreexpresan dicha proteína de fusión, aun poseyendo un fondo silvestre, no resultaban viables (supervivencia inferior a cuatro días).

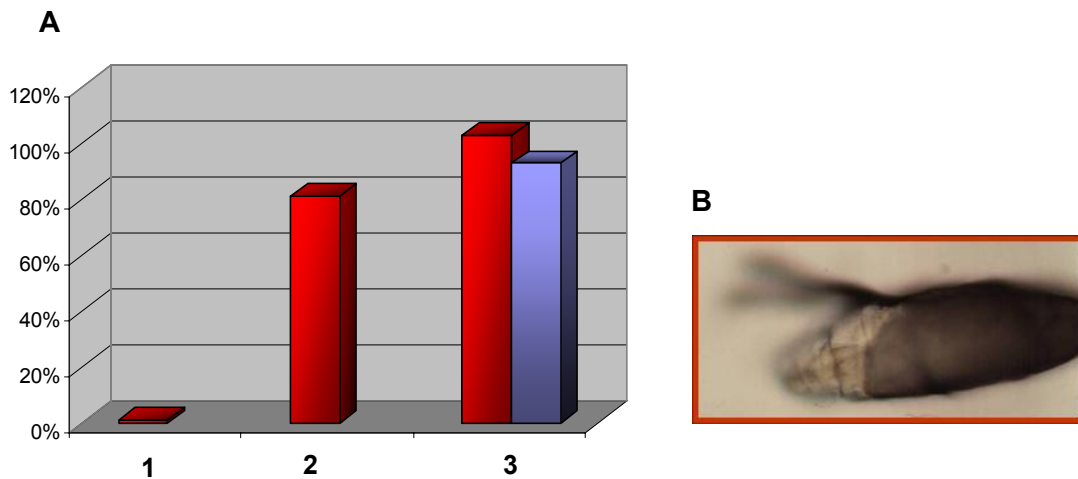


Figura 4. Experimento de rescate de la letalidad embrionaria de los mutantes *muscleblind*. A) Resultado de un experimento en el cual comparamos la capacidad de la proteína MbIC (2) y MbIC:GFP (3) de rescatar la letalidad embrionaria asociada con la falta de función de *muscleblind* (*mbf^{E27}*) (1). El porcentaje de larvas rescatadas con la proteína MbIC:GFP a 25°C fue del 103% (n=1029) y a 19°C del 93% (n=353), en contraste con el porcentaje de larvas rescatadas con la proteína MbIC sola que, a 25°C, fue del 81% (n=676). En rojo se muestran los datos a 25°C y en azul a 19°C. B) larva mutante rescatada naciendo

Por tanto, la proteína de fusión MbIC:GFP es, cuanto menos, igual de efectiva que la proteína MbIC en un experimento de rescate, por lo que concluimos que la etiqueta GFP no interfiere con la función normal de la proteína MbIC durante el desarrollo tardío del embrión.

2.2.4 Co-inmunoprecipitación de proteínas

Con el fin de dilucidar el mecanismo molecular por el cual la proteína Mbl realiza su función, nos propusimos llevar a cabo un experimento de coimmunoprecipitación para identificar proteínas que interaccionen *in vivo* con la proteína Mbl de *Drosophila*.

Así pues, diseñamos un experimento en el que, a partir de homogenados de embriones de edad comprendida entre 6 y 24 horas que expresan la proteína de fusión MbIC:GFP (embriones_GFP), procedimos a la IP de la proteína MbIC:GFP haciendo uso de un anticuerpo anti-GFP y al posterior análisis del eluido en electroforesis SDS-PAGE y Western blot (apéndice 2.2.6.2). Se eligieron embriones de 6 a 24h porque incluye los estadios del desarrollo donde *mbf* se expresa normalmente y es, por tanto,

el momento en el que están presentes las proteínas que interaccionan con Muscleblind cuando esta lleva a cabo su función.

Tampón de homogenización

Cuando el objetivo de una inmunoprecipitación es la coimmunoprecipitación de proteínas que interaccionen físicamente con el antígeno que se quiere inmunoprecipitar, las condiciones del tampón de homogenización han de favorecer al máximo la unión anticuerpo-antígeno para poder “bajarse” la mayor cantidad de antígeno posible, pero por otro lado, han de permitir el mantenimiento de las interacciones entre dicho antígeno y las proteínas asociadas.

Para optimizar las condiciones de inmunoprecipitación de la proteína MbIC:GFP probamos distintos tampones de homogenización (apéndice 2.2.6.3) en los que variamos la concentración salina y la fuerza iónica. Estos dos factores afectan a la unión antígeno-anticuerpo y por tanto a la eficiencia de la inmunoprecipitación. Los resultados obtenidos en geles teñidos con plata y en Western blots mostraban que (Figura 5):

- a) la proteína de fusión MbIC:GFP estaba siendo inmunoprecipitada ya que se detectaba en el inmunoprecipitado de los embriones GFP con el anticuerpo anti-GFP (Figura 5A). Sin embargo, no podía visualizarse en los geles de acrilamida, posiblemente, debido al solapamiento con la cadena pesada de mayor tamaño del anticuerpo anti-GFP empleado (Figura 5B)
- b) se estaba produciendo degradación de la proteína de fusión ya que se detectaban distintas bandas en el western blot (Figura 5A). De hecho, en algunas IP la banda de menor tamaño en el Western blot (~28k Da) si que se podía apreciar en los geles de acrilamida (punto rojo Figura 5B)
- c) las proteínas coimmunoprecipitadas en estas condiciones eran inespecíficas, ya que también estaban presentes en la IP control (Figura 5B).

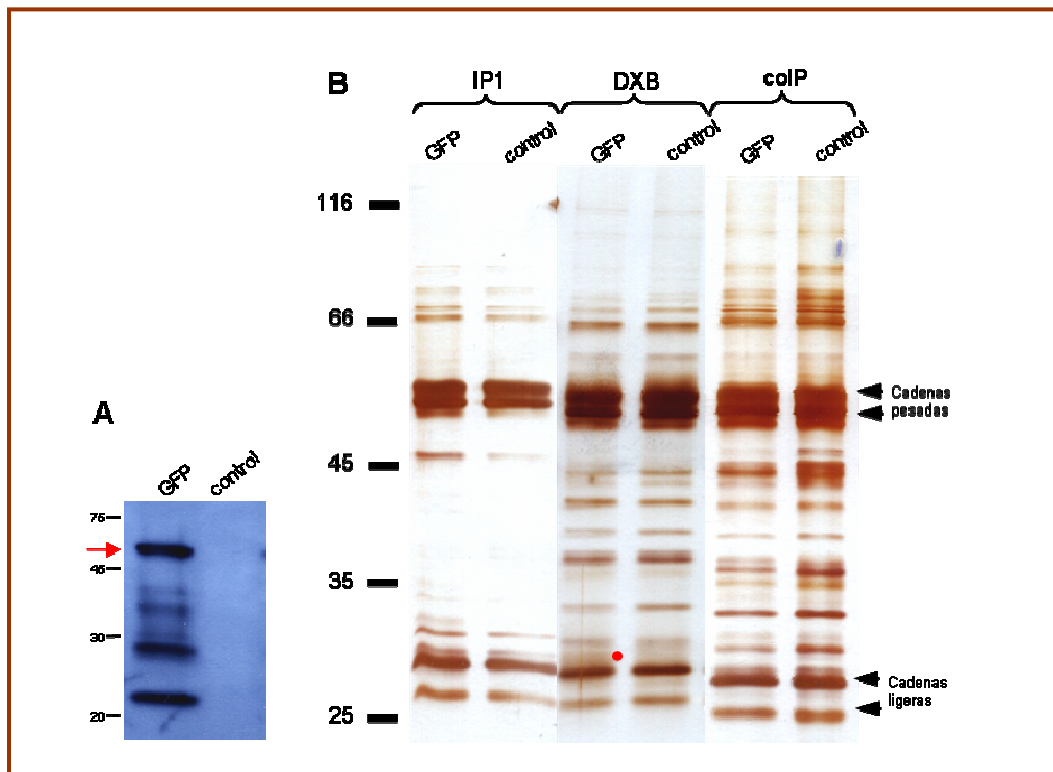


Figura 5. Inmunoprecipitación de MbIC:GFP. (A) Detección mediante Western blot de la proteína MbIC:GFP inmunoprecipitada con un anticuerpo α -GFP. La flecha roja señala la proteína de fusión íntegra con un tamaño de 52kDa. Las bandas de menor tamaño se corresponden con distintos fragmentos de degradación de la MbIC:GFP. (B) Inmunoprecipitación de MbIC:GFP con distintos tampones de homogenización. Se observa que a menor astringencia del tampón (IP1>DXB>colP) mayor es el número de proteínas coimmunoprecipitada (carrera GFP). Sin embargo, en todos los casos se trata de proteínas coimmunoprecipitadas inespecíficamente, ya que también aparecen en la colP control (carrera control). En ningún caso podemos observar la proteína MbIC:GFP íntegra, posiblemente por solapamiento con las cadenas pesadas del anticuerpo. El punto rojo señala un producto de la proteína MbIC:GFP degradada. El anticuerpo α -GFP empleado es una mezcla de dos anticuerpos monoclonales lo que explica la presencia de dos cadenas pesadas y dos cadenas ligeras.

De los seis tampones utilizados para la colP, obtuvimos los mejores resultados con el tampón DXB ya que fue el único en el que pudimos ver proteína MbIC:GFP inmunoprecipitada (punto rojo sobre Figura 5). El tampón DXB (25 mM Hepes 6.8, 50 mM KCl, 250 mM sucrosa y 1 mM DTT) es de astringencia media ya que contiene una cantidad moderada de iones y no posee detergentes.

Unión covalente del anticuerpo anti-GFP

Ya que el tamaño de la proteína de fusión MbIC:GFP (52 kDa) es similar al de las cadenas pesadas del anticuerpo anti-GFP empleado unimos covalentemente el anticuerpo a la proteínaG-agarosa de tal modo que no se liberasen durante la elución

y no se visualizasen posteriormente en los geles de acrilamida. Para ello se empleó el agente entrecruzador dimetil pimelidato (DMP) (ver material y métodos 2.13)

Degradación de MbC:GFP

Durante la homogenización de los embriones se produce la disgregación y rotura de sus células. Cuando las células y los orgánulos celulares se rompen, los lisosomas liberan las proteasas celulares cuya actividad produce la degradación de las proteínas. La adición de inhibidores de proteasas al tampón de homogenización resulta imprescindible para evitar la degradación de las proteínas en las muestras. Por ello, utilizamos un cóctel de inhibidores de proteasas: Aprotinina, Leupeptina, Pepstatina y PMSF, que añadíamos regularmente durante la homogenización y a lo largo de todo el protocolo de IP. Sin embargo, en los Western blot de los inmunoprecipitados observábamos que la proteína MbC:GFP continuaba degradándose.

Comprobamos la efectividad del uso y concentración de los inhibidores de proteasas durante la homogenización de los embriones, corriendo una pequeña muestra de homogenado en gel SDS-PAGE y visualizando las proteínas mediante tinción con Coomassie Blue. Observamos muestras con bandas bien definidas y de todos los tamaños, lo que indicó que no se está produciendo una degradación generalizada de las proteínas en la muestra. Para ver de qué factores dependía la degradación de la proteína MbC:GFP, tomamos muestras de homogenados de embriones recién preparados y homogenados almacenados a -80°C durante varios meses, así como también de muestras de un homogenado a lo largo del protocolo de IP. Los resultados indicaron que la degradación de la proteína MbC:GFP es directamente proporcional al tiempo de almacenamiento y al tiempo que transcurre desde la descongelación del homogenado.

Para reducir la degradación de la proteína de fusión MbIC:GFP se aumentó el número y la concentración de los inhibidores utilizados: Aprotinina, Leupeptina, pepstatina, PMSF y preparado *cOmplete* de Roche, y se redujeron los tiempos de incubación de la inmunoprecipitación de 2h a 1h. Con estas medidas se consiguió reducir significativamente la degradación de la proteína MbIC:GFP aunque continuó ocurriendo en una medida considerable.

Concentración de los homogenados

Para aumentar la cantidad de proteína MbIC:GFP inmunoprecipitada, y por tanto, de las proteínas coinmunoprecipitadas, realizamos IPs usando homogenados de diferentes concentraciones. En concreto, utilizamos homogenados a una concentración de 1:4, 1:2 y 1:1 (relación volumen de embriones y volumen de tampón de homogenización). Observamos que, a medida que aumentaba la concentración de los homogenados aumentaba ligeramente la cantidad de proteínas coIP, sin embargo, aumentaba en mayor medida la degradación de la proteína de fusión MbIC:GFP. Por este motivo, elegimos una concentración 1:2 (~10 µg/µl), ya que es la opción que mejor equilibra los dos parámetros: mayor cantidad de proteína IP y menor degradación.

Coinmunoprecipitación de proteínas

Aproximadamente 700 µl de extractos de embriones preabsorbidos se inmunoprecipitaron durante una hora con 50 µl de proteínaG-agarosa unida covalentemente a anticuerpo anti-GFP. El pellet de *beads* se lavó con varios volúmenes de tampón de homogenización DXB con inhibidores de proteasas recién añadidos y se eluyó con un gradiente de glicina. Las proteínas eluidas se precipitaron con TCA y se resuspendieron en tampón de carga SDS-PAGE. Las proteínas coIP se separaron en un gel de acrilamida SDS-PAGE al 10% durante 16 h y se visualizaron mediante tinción con plata. En paralelo, se procesaron muestras de los inmunoprecipitados de ambos tipos de embriones para detectar la proteína MbC:GFP mediante Western blot y poder identificar las bandas teñidas con plata correspondientes a dicha proteína. La Figura 6 muestra los resultados de la coinmunoprecipitación y el Western blot.

Se pueden identificar, al menos, 12 bandas que aparecen únicamente en la carrera GFP y no en la control. De ellas, seis se corresponden con la proteína MbIC:GFP (estrella, Fig.6) o fragmentos de degradación de la misma (puntos, Fig.6), tal y como muestra la detección con anti-GFP en los inmunoprecipitados. La microsecuenciación de las bandas "m" (fig.6) confirmó que efectivamente se trataba de proteína MbIC:GFP. Las otras 6 bandas se corresponden con proteínas coIP específicamente por la proteína MbIC:GFP ya que no aparecen en la muestra control (flechas verdes, Fig.6). Para conocer la identidad de las proteínas coIP específicamente, repetimos el experimento y enviamos el gel al servicio de proteómica del CBM Severo Ochoa de la Universidad Autónoma de Madrid. Sin embargo, la microsecuenciación de estas

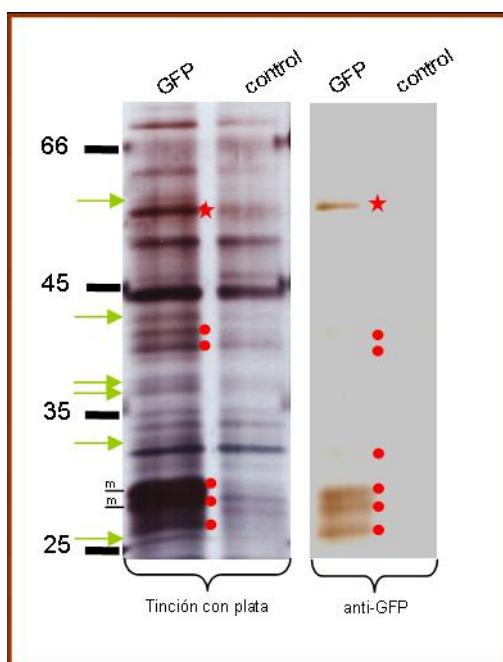


Figura 6. Coimmunoprecipitación de proteínas que interaccionan con MblC:GFP. Las proteínas colP específicamente, esto es, aquellas que no están presentes en la carrera control, se señalan con flechas verdes. El western blot permitió detectar cual de las bandas de la carrera GFP se correspondía con la proteína de fusión MblC:GFP. La proteína de fusión Íntegra se señala con una estrella mientras que las bandas correspondientes a fragmentos de degradación de la proteína se señalan con un punto (*). La colP se ha hecho en paralelo a partir de homogenizados de embriones que sobreexpresan la proteína de fusión MblC:GFP (carrera GFP) o de embriones que no lo hacen (carrera control). m, indica que las bandas han sido microsecuenciadas. * problemas de escaneo y resolución de la imagen impiden ver las dos bandas centrales

bandas no fue posible debido a que la cantidad de proteína que contiene cada una de ellas no es suficiente.

Actualmente trabajamos en la obtención de una mayor cantidad de proteínas inmunoprecipitadas, escalando el protocolo desarrollado de colP.

2.2.5 Verificación de candidatos mediante coimmunoprecipitación

Las proteínas colP, tras separarlas en un gel de acrilamida SDS-PAGE, pueden teñirse para ser cortadas e identificadas mediante microsecuenciación, o pueden transferirse a una membrana para comprobar mediante inmunodetección la presencia de una proteína dada. La colP de proteínas para microsecuenciación requiere que la proteína sea visible como una banda individual en el gel SDS-PAGE y que esté en cantidad suficiente para su identificación mediante técnicas de espectrometría de masas o similares. La detección mediante Western blot de proteínas coimmunoprecipitadas permite detectar una proteína que este en cantidades muy pequeñas, ya que la sensibilidad del método es varios órdenes de magnitud superior.

Una vez establecidas las condiciones de IP de la proteína de fusión MblC:GFP y como estrategia paralela a la búsqueda mediante escrutinio en geles SDS-PAGE de

proteínas interaccionadoras, procedimos a la comprobación directa mediante Western blot de la presencia de proteínas candidatas.

aly, *amos*, *nonA* y *jumu* son algunos de los genes que interaccionan genéticamente con la sobreexpresión de *mbIC* en ojo (resultados 2.1.1). La interacción génica que existe entre *muscleblind* y cada uno de estos genes puede reflejar una interacción física entre sus respectivas proteínas. Para comprobar esta posibilidad, detectamos la presencia de las proteínas Aly, Amos, NonA y Jumu en el colP de embriones GFP y control (Figura. 7).

Ninguno de los anticuerpos estaba caracterizado para la detección de la proteína en extractos crudos de embriones de *Drosophila*, por lo que el tamaño esperado de las bandas para cada anticuerpo se infiere a partir de otro tipo de experimentos (ej. extractos celulares) o del peso molecular predicho. Para cada caso, determinamos la concentración del anticuerpo y los homogenados de *Drosophila* que permitían detectar bandas de forma reproducible y con un tamaño similar al esperado de modo que nos sirvieran como control positivo del funcionamiento de cada anticuerpo en las colP.

Los anticuerpos anti-NonA y anti-Jumu no reconocen ninguna proteína en el colP de MbIC:GFP. Para Aly y Amos observamos bandas en el colP a partir de embriones GFP (carrera GFP) pero ninguna de ellas es específicamente colP por la proteína MbIC:GFP puesto que también aparece en el colP a partir de embriones que no sobreexpresan dicha proteína de fusión (carrera CONT). Por tanto, ninguna de las proteínas analizadas, Aly, Amos, NonA y Jumu, coinmunoprecipita específicamente con MbIC:GFP por lo que concluimos que ninguna de ellas interacciona físicamente con MbIC:GFP en las condiciones experimentales analizadas.

Un requisito indispensable para que se coinmunoprecipiten proteínas es que la proteína MbIC:GFP haya sido correctamente IP. Para cada experimento, comprobamos la presencia de la proteína de fusión en el colP mediante detección con α -GFP. En todos los casos detectamos señal (Figura. 7 anti-GFP, se muestra un solo ejemplo)

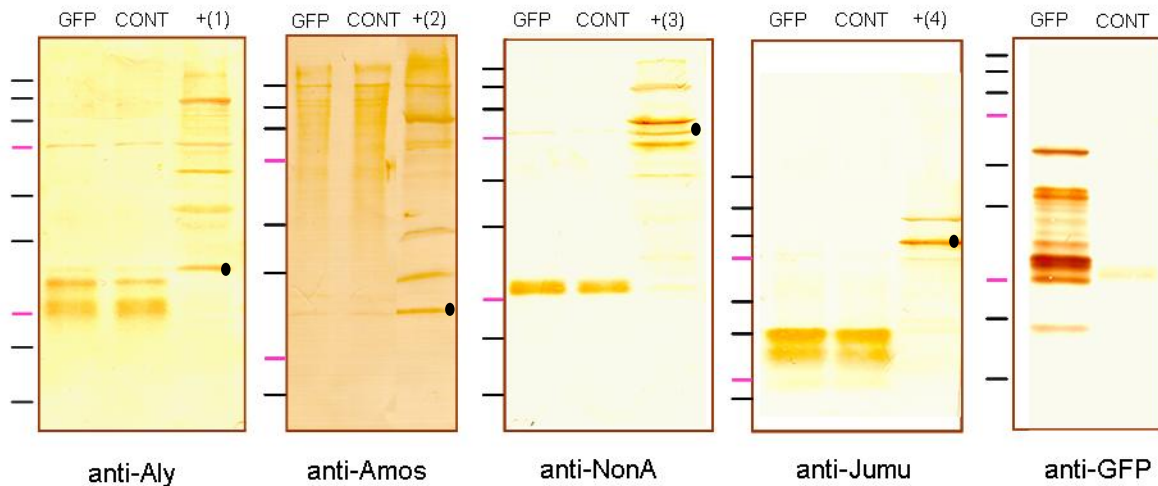


Figura 7. Validación de proteínas candidatas mediante coIP. Tras eluir las proteínas coIP en tampón de carga SDS-PAGE se separaron en gel de acrilamida al 10% o 12% y se transfirieron a membrana. Para la visualización de las proteínas candidatas en la membrana se utilizó un método de detección colorimétrico (sustrato DAB-H₂O₂) con un paso de amplificación de la señal (sistema avidina/biotina, material y métodos 2.10). No detectamos proteína Aly, Amos, NonA ni Jumu en el coIP de MbIC:GFP. La detección con α -GFP confirma que ha habido IP de MbIC:GFP. Además, para cada una de las proteínas en el coIP de MbIC:GFP, se incluye un control positivo (+) que permite ver el tamaño de las bandas que reconoce el anticuerpo empleado para cada una de las proteínas analizadas. Las muestras empleadas para los controles positivos (1) y (3) son extractos de embriones OrR 0-24h; (2), Mosca adulta; (4), glándulas salivares de larvas de 3er estadio que sobreexpresan *jumu* (*Saga49EGal4>UASjumu*). La cantidad de proteína total cargada es de 50 μ g por carrera para Aly y Amos, 150 μ g para NonA y 20 parejas de glándulas salivares para Jumu. Marcador de PM preteñido *Dual* de Bio-Rad: 250, 150, 100, 75, 50, 37, 25, 20, 15. 1.4. círculo (●), banda del tamaño esperado. Para información sobre los anticuerpos utilizados, ver material y métodos 1.4.

En resumen, con el fin de coIP proteínas que interaccionen con Muscleblind *in vivo* hemos generado moscas transgénicas que expresan la proteína de fusión MbIC:GFP. Hemos comprobado que MbIC:GFP es funcional *in vivo*, ya que rescata el fenotipo de letalidad embrionaria asociada a los alelos de *muscleblind* y por tanto resulta adecuada para coIP proteínas que interaccionen con Mbl para llevar a cabo su función. Una vez establecida la línea de moscas UASmbIC:GFP desarrollamos el protocolo de coIP que incluye la obtención de material a gran escala y la puesta a punto de las condiciones de coIP. Como resultado de este experimento hemos identificado, en geles de acrilamida teñidos con plata, seis proteínas de pesos moleculares entre 25 y 55 kDa. Para identificar mediante espectrometría de masas estas proteínas es necesario obtener una mayor cantidad de las mismas, por lo que, actualmente estamos escalando el protocolo de coIP. Por otro lado, hemos descartado que la identidad de alguna de esas proteínas sea la de Amos, nonA, Aly o Jumu ya que ninguna de ellas se detecta mediante Western blot en el inmunoprecipitado.

2.2.6. Apéndices

2.2.6.1 Sistema UAS-GAL4

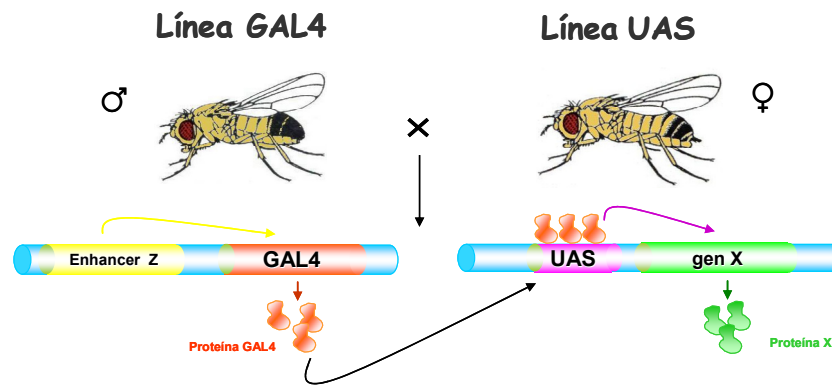


Figura 1. El sistema UAS/GAL4 de expresión ectópica (Brand y Perrimon, 1993) utiliza el factor de transcripción de levadura GAL4 y su secuencia diana UAS (*Upstream Activation Sequence*) a las cuales se une para activar la transcripción génica. GAL4 puede expresarse bajo diversos patrones de expresión colocándolo bajo el control de distintas secuencias promotoras de *Drosophila*. Ya que las secuencias UAS no existen en el genoma de *Drosophila melanogaster*, la expresión de GAL4 bajo un patrón de expresión determinado por las secuencias promotoras que utilizemos provocará la sobreexpresión del transgén colocado aguas abajo de las secuencias UAS con ese mismo patrón de expresión. El sistema UAS/GAL4 es un sistema de expresión bipartito en el que los dos componentes del sistema, la secuencia que determina el patrón de expresión o *driver* (enhancer Z) y el gen que queremos expresar bajo ese patrón de expresión o *responder* (gen X), se mantienen como líneas parentales separadas constituyendo la línea GAL4, que contiene secuencias *driver*, y la línea UAS que contiene el gen *responder*. Los descendientes del cruce entre machos y hembras de estas dos líneas poseen en su genoma ambos componentes a la vez. Cuando la región reguladora “enhancer Z” se activa provoca, la expresión de la proteína GAL4, esta se unirá a las secuencias reguladoras UAS que a su vez provocarán la expresión del gen X.

2.2.6.2. Esquema del experimento de colP

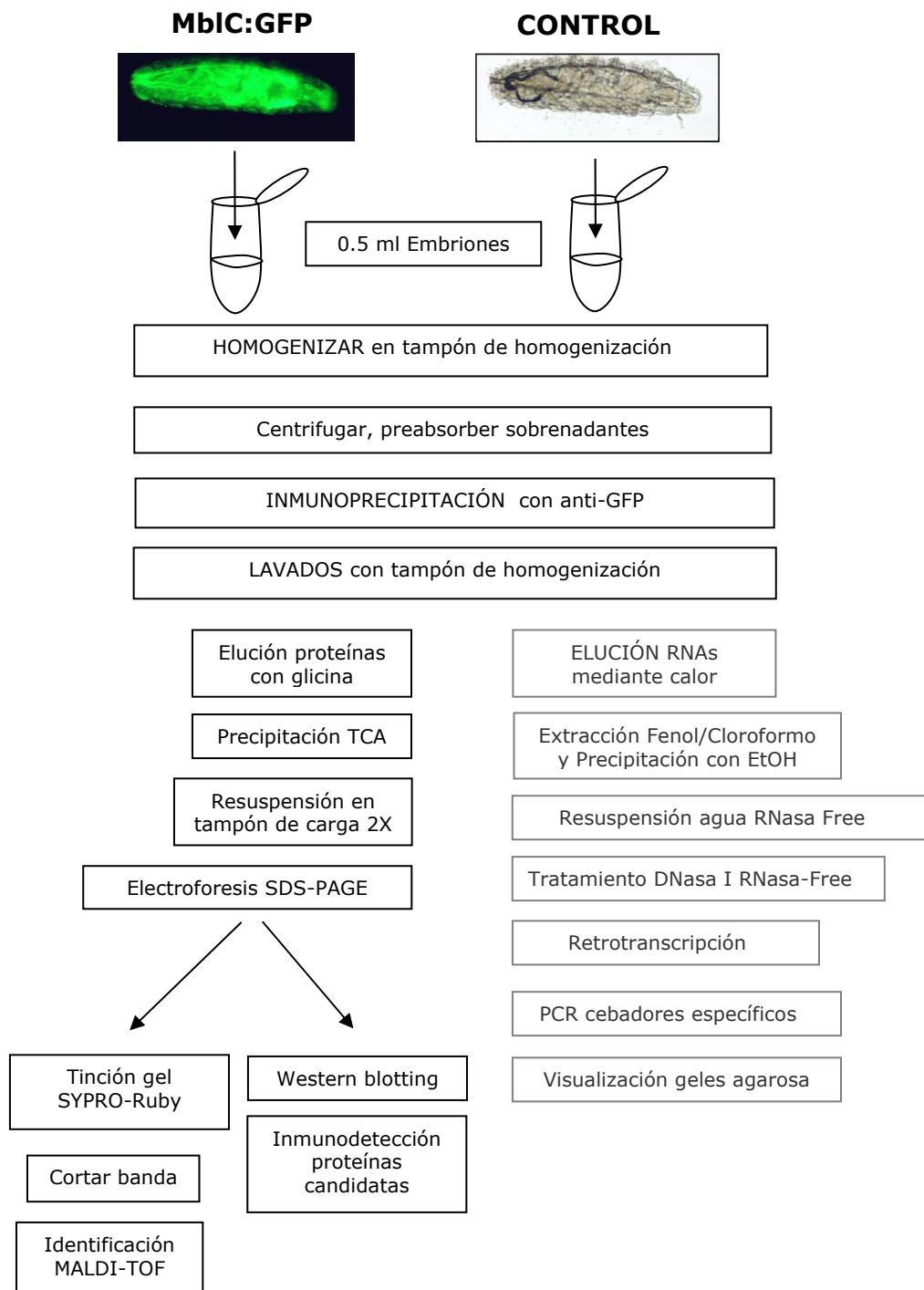


Figura 2. Esquema experimento de colP. El protocolo para la identificación de RNAs coimmunoprecipitados por la proteína de fusión MbIC:GFP no se ha desarrollado a lo largo de este trabajo aunque las moscas transgénicas MbIC:GFP han sido creadas para ambas utilidades.

2.2.6.3 Tampones de homogenización

Tampón	Composición	Fuente
DXB	25mM HEPES pH 6.8 50mM KCl 250mM sucrose 1mM DTT	Wilhelm <i>et al</i> 2000
IP-1	50mM Tris-Cl pH 7.5 300mM NaCl 5mM EDTA 1% Tritón X-100	Current protocols
coIP-1	50mM Tris-Cl pH 7.5 100mM NaCl 1mM MgCl ₂ 1mM DTT 0.1% Tritón X-100	Current protocols
IP-2	50mM Tris-Cl pH 7.5 500mM NaCl 5mM EDTA 1% Tritón X-100	Current protocols
coIP-2	50mM Tris-Cl pH 7.5 150mM NaCl 1mM DTT 1mM EDTA 1mM MgCl ₂ 0.1% Tritón X-100 10% glicerol	Current Protocols
Núcleos	15mM HEPES pH 7.6 10mM KCl 50mM MgCl ₂ 0.5mM EGTA 0.1mM EDTA 1mM DTT 0.35mM sacarosa 20mM HEPES pH 7.6 420mM NaCl 1.5mM MgCl ₂ 0.2mM EDTA 0.2mM DTT 25% glicerol	Instituto Ramón y Cajal, Madrid

Tabla 1. Tampones de homogenización testados para la coimmunoprecipitación de proteínas junto a MbIC:eGFP. IP-2 tiene una aumento en la fuerza iónica. coIP-2 también ligeramente y también de nueve el EDTA

DISCUSIÓN

1. ANÁLISIS ESTRUCTURAL Y RELEVANCIA CLÍNICA DE LOS GENES HUMANOS

1.1 Estructura y Procesado de los genes *muscleblind* humanos

El genoma humano contiene tres homólogos del gen *muscleblind* de *Drosophila*, *MBNL1*, *MBNL2* y *MBNL3*. Los tres genes humanos tienen una estructura intrón-exón muy conservada y sus transcritos presentan una gran variedad de procesado alternativo, especialmente *MBNL1*. Los resultados presentados en este trabajo coinciden con los publicados por Fardaei y colaboradores (2002) en cuanto a la estructura génica y el procesado alternativo de *MBNL1-3* excepto en dos puntos concretos. En el trabajo de Fardaei se describe la presencia del exón 10 para el gen *MBNL2* (también conocido como *MBLL*) y del exón 9 para el gen *MBNL3* (también conocido como *MBXL*). Nuestros resultados indican que estos exones no están contenidos en el DNA genómico que abarca los respectivos genes. Para llegar a esta conclusión, usamos los péptidos codificados por los exones 9 y 10 del gen *MBNL1* como *query* para realizar dos tipos de búsquedas (1) contra los ESTs específicos de cada uno de los genes *MBNL1* y *MBNL2* y (2) contra las bases de datos de DNA genómico humano. En ningún caso obtuvimos resultados positivos confirmado nuestros resultados e indicando la falta del exón 10 y 9 en los genes humanos *MBNL2* y *MBNL3*, respectivamente. Además, hemos encontrado cuatro nuevos exones de *MBNL1* no descritos hasta la fecha, que se localizan en el intrón 1-2 y en la región 5' del gen.

1.1.1 Presencia de elementos repetitivos en *MBNL1*

Hemos observado que algunos sucesos de *splicing* en *MBNL1* incluyen elementos repetitivos como secuencias *Alu* o secuencias MER1. Un ejemplo viene dado por la “extensión” que sufre el exón 6 cuando se utiliza un sitio dador opcional que se encuentra a 2227 pb aguas abajo del sitio 5' habitual de procesado. Las últimas 300 pb de este exón presentan una identidad del 89% con la secuencia complementaria del elemento *Alu-Sx*. Los elementos *Alu* son la familia más abundante de secuencias repetitivas de DNA en primates representado el 10% de la secuencia total del genoma humano. Se trata de pequeños elementos transponibles de aproximadamente 300 pb

que consisten en la repetición de dos unidades monoméricas unidas por una secuencia rica en adeninas. Los elementos *Alu* pueden ser ocasionalmente usados como exones alternativos y se calcula que aproximadamente el 5% de todos los exones internos del genoma humano derivan de elementos *Alu* (Gorek, 2002). Se han descrito varios ejemplos en los que una nueva variante de *splicing* incluye una parte o la totalidad de un elemento *Alu* (Mihovilovic, 1993; Mullersman, 1995; Vervoort, 1998; Svineng, 1998). Este fenómeno se describe como exonización de elementos transponibles (revisado en Krehling, 2004). La secuencia consenso *Alu* tiene nueve sitios 5' donadores potenciales y catorce sitios 3' dadores que, aunque no se corresponden con consensos "fuertes", eventualmente pueden ser usados por la maquinaria de *splicing*. Proponemos que esto es lo que ha ocurrido en el mensajero que se corresponde con el clon IMAGE identificado (ID 511793). Se ha comprobado que la mayoría de transcritos que incluyen un elemento *Alu* son marcados para su degradación lo que explicaría la rareza de este mensajeros en *MBNL1* (<0.1%). La inclusión de este elemento *Alu* causa la adición de 57 nuevos aminoácidos y un codón de parada prematuro. Aunque esta proteína contiene los cuatro dedos de zinc intactos, la relevancia funcional de esta isoforma se desconoce.

1.1.2 Conservación en *Drosophila melanogaster*

La estructura exón-intrón no está conservada entre los genes humanos y el gen de *Drosophila* ya que no se puede establecer una correlación evidente entre los exones de los loci humanos y de la mosca. Sin embargo, existen algunas similitudes entre los loci como son la presencia de un intrón 1-2 muy grande, regiones no traducidas (UTR) muy largas, y formas de *splicing* que generan proteínas con propiedades de unión diferentes al resto, como la isoforma *mbID* en muscleblind y la isoforma *MBNL1_1a* humana.

1.2 Expresión de los genes muscleblind humanos

El análisis *in silico* del patrón espacial y temporal predice que los genes *MBNL1* y *MBNL2* son los que se expresan en un mayor número de tejidos y de manera más abundante, especialmente *MBNL1*, siendo la expresión del gen *MBNL3* la más limitada, tanto por el número de tejidos como por la intensidad relativa de la transcripción. Además, más de la mitad de los ESTs del gen *MBNL3* (69,5%)

pertenecen a tejidos embrionarios, lo que indica que *MBNL3* se expresa preferentemente en este estadio del desarrollo. Esto concuerda bien con el hecho de que la placenta sea uno de los tejidos donde más ESTs encontramos.

Durante la realización de este trabajo se han publicado datos experimentales sobre la expresión de los genes *muscleblind* humanos. Así, Fardaei y colaboradores (2002) analizan la expresión de los genes *MBNL1*, *MBNL2* y *MBNL3* mediante Northern blot. El tamaño de los transcritos detectados es de 6.5 y 6.0 kb para *MBNL1*, 5.5 kb para *MBNL2* y de 10.0 kb para *MBNL3*. Para los mismos tejidos analizados, Miller y col (2000) detectan también dos transcritos de *MBNL1* cuyo tamaño estiman en 6.5 y 5.3 kb. El *splicing* alternativo de los exones en la región 3' produce variaciones en tamaño de entre 50-200 pb por lo que no puede explicar la aparición de estos dos especies de mensajeros en el Northern blot. En este sentido, hemos identificado varias señales de poliadenilación activas en la región 3'-UTR de *MBNL1*. La distancia entre la primera y la última señal es de 1200 pb lo que podría explicar *per se* las dos clases de mensajeros que se observan en los Northern blot.

MBNL1 y 2 se expresan en todos los tejidos analizados (Miller, 2000; Fardaei, 2002): páncreas, riñón, músculo esquelético, hígado, pulmón, placenta, cerebro y corazón. *MBNL1* lo hace predominantemente en músculo esquelético mientras *MBNL2* lo hace más homogéneamente, aunque ligeramente más en riñón, músculo, placenta y cerebro. La expresión de *MBNL3* se detecta principalmente en placenta y más débilmente en páncreas, riñón, hígado y corazón. Miller y col (2000) detectan proteína *MBNL1* (anticuerpo policlonal contra EXP42 que reconoce todas las isoformas) específicamente en corazón, ojo y de forma más abundante en músculo.

Los datos de expresión basados en el análisis de ESTs coinciden con los obtenidos mediante Northern y Western blot en dos aspectos fundamentales: (1) reflejan el hecho de que el gen *MBNL1* es el que se expresa en mayor cantidad y *MBNL3* en el que menos en base al número de ESTs por tejidos identificados en cada uno de los casos; y (2) muestran expresión de cada uno de los genes en los tejidos analizados: páncreas, riñón, músculo esquelético, hígado, pulmón, placenta, cerebro y corazón, con la excepción de *MBNL3* para el que no se han encontrado ESTs pertenecientes a librerías de corazón. Esto se debe, probablemente, a que, al tratarse de transcritos raros, no es fácil que estén representados en librerías que no estén hechas específicamente para captar este tipo de mensajeros.

Además de los tejidos mencionados, los datos *in silico* demuestran la expresión de estos genes en otros tejidos que no se han descrito hasta la fecha, tales como la lengua y los intestinos, para los tres genes humanos, la vejiga para *MBNL1* y *MBNL2* y el timo para *MBNL1*. Las hibridaciones *in situ* realizadas por Kanadia y colaboradores (2003b) en embriones de ratón son una evidencia indirecta de la expresión de *MBNL1*, *MBNL2* y *MBNL3* en estos tejidos tal y como predice la existencia de ESTs humanos.

Expresión en las células endoteliales

Otro tejido en el que cabe esperar la expresión de estos genes es el tejido vascular (células endoteliales). Para *MBNL1* el grupo de ESTs pertenecientes a librerías de células endoteliales es el más numeroso. También se encuentran ESTs de *MBNL2* y 3 en estas librerías. Las células endoteliales tienen diferentes funciones críticas en el control de la función vascular (revisado en Michiels, 2003). Una de ellas es la extravasación de los linfocitos que define la migración de estas células desde el sistema vascular hacia los tejidos que han sufrido un daño mecánico o están expuestos a patógenos. El proceso de migración de los linfocitos es posible gracias a una serie de moléculas de adhesión celular que se expresan en las células endoteliales en respuesta a las citoquinas producidas por los tejidos afectados. Se conocen tres tipos de moléculas de adhesión celular implicadas en este proceso: Selectinas, Inmunoglobulinas e Integrinas. Una de las integrinas expresadas por las células endoteliales y que median la unión de los leucocitos es la integrina alfa-3. Un estudio reciente llevada a cabo por Adereth y colaboradores ha demostrado que *MBNL2* regula la localización del mensajero de la integrina alfa-3 lo que conduce a que la proteína codificada por dicho mensajero, se localice correctamente en los *focal adhesion complexes* (Tien Hsu, comunicación personal; Adereth, 2005). Por tanto, la abundante expresión en las células endoteliales de los genes *MBNL1-3* que muestran los datos de expresión *in silico*, podría deberse a la función de los genes *MBNL* en el proceso de extravasación de linfocitos que se produce en gran medida en las células endoteliales.

Por otro lado, el perfil de expresión en base a los ESTs pone de manifiesto la expresión de *MBNL1* y 2 en la médula ósea donde se producen los componentes celulares de la sangre: eritrocitos, plaquetas y leucocitos. Un análisis de la expresión génica a escala genómica para determinar el programa genético de las células madre hematopoyéticas en ratón, ha identificado a *MBNL3* como un gen expresado en las células madre más indiferenciadas indicando que este gen podría tener alguna función en la biología de la hematopoyesis. Quizás en la diferenciación de los linfocitos ya que

también encontramos gran expresión en otros tejidos ricos en este tipo celular, como el timo y el nódulo linfático.

Expresión en células tumorales

Durante el análisis de los ESTs para elaborar el patrón de expresión de los genes *MBNL1*, *MBNL2* y *MBNL3* comprobamos la existencia de transcritos procedentes de librerías hechas a partir de tejidos cancerosos en los tres genes. Hemos comprobado experimentalmente la presencia de transcritos *MBNL3* pertenecientes a tejidos cancerosos mediante un rastreo de librerías de cDNA que incluían librerías hechas a partir de leucemia promonocítica, leucemia promonocítica activada, melanoma y β -linfoma obteniendo amplificado para este último caso. Por otro lado, *MBNL1* ha sido identificado como un loci implicado en las translocaciones de *BCL6*. Los reordenamientos génicos en este gen, se asocian a una transformación agresiva del linfoma folicular, por lo que *MBNL1* podría tener alguna relevancia en este hecho (Akasaka, 2003). Finalmente, Adereth y colaboradores identificaron *MBNL2* como un gen que se expresa diferencialmente en muchos tumores (Adereth, 2005). Dada la relación entre *MBNL2* y la función de la integrina alfa-3, este hecho podría estar relacionado con los cambios en la adhesión celular que sufren las células durante el proceso de transformación tumoral.

1.3 Los genes muscleblind humanos pueden tener funciones distintas

Los tres genes *muscleblind* humanos tienen una homología de secuencia del ~70% a nivel aminoacídico y presentan una estructura genómica muy conservada. Se pueden definir tres tipos de genes homólogos en función de su origen evolutivo: Ortólogos, Parálogos y Xenólogos. Los genes parálogos son aquellos homólogos producidos por duplicación génica y representan genes derivados a partir de un gen ancestral común. Nosotros proponemos que *MBNL1*, 2 y 3 son genes parálogos porque se parecen más entre sí que con respecto a el *muscleblind* de *Drosophila*. Además, la conservación de la estructura genómica indica que los sucesos de duplicación génica son relativamente recientes o que existe una presión de selección para mantenerla.

Los genes parálogos tienden a tener funciones diferentes ya que después de la duplicación los nuevos genes están libres de presión selectiva y el organismo puede explorar nuevas funciones para ellos. Los genes *muscleblind* humanos tienen un patrón de expresión distinto por lo que cabe esperar que desarrollen funciones

distintas. Además existen ciertos datos sobre la función de los genes humanos que apoyan esta hipótesis. Por ejemplo, se sabe que *MBNL1* es necesario en la diferenciación terminal de los mioblastos mientras que *MBNL3* es un inhibidor del programa miogénico (Squillace, 2002)

A favor de una diversificación en la función, un estudio sobre el patrón de expresión de los tres genes en ratón (99.9% identidad a nivel proteico) mediante hibridaciones *in situ* de preparaciones de embriones completos, muestra que los tres genes tienen una expresión coordinada con picos de expresión máxima diferentes a lo largo del desarrollo embrionario (Kanadia, 2003)

1.4 Historia evolutiva de los genes *muscleblind*

El gen *muscleblind* sólo existe dentro del grupo de los Metazoos. Dentro de los Bilateria, los Nematoda es el grupo más antiguo donde el gen *muscleblind* es reconocible por homología de secuencia. Las proteínas Muscleblind constituyen una nueva familia de proteínas CCCH. En invertebrados, existe un solo gen *muscleblind* y éste se caracteriza por poseer dos dedos de zinc del tipo $CX_7CX_6CX_3H$. En vertebrados existen tres genes *muscleblind* con una alta homología entre ellos (~70% identidad aminoacídica) y se caracterizan por tener cuatro dedos de zinc siendo el primero y tercero idénticos al de invertebrados y el segundo y cuarto del tipo $CX_7CX_4CX_3H$.

La similitud aminoacídica entre las proteínas Muscleblind de invertebrados y vertebrados es máxima en la región de los dedos de zinc CCCH (primer par de vertebrados) pero se extiende más allá de los mismos. Como parte de este trabajo hemos definido 2 motivos conservados evolutivamente que nombramos como caja LEV y caja NGR, la conservación de estos residuos puede poner de manifiesto su importancia para la función básica de *muscleblind*

Nuestro análisis de las proteínas Muscleblind revela que en invertebrados existen sólo dos dedos de zinc mientras en vertebrados existen cuatro. Por tanto, a lo largo de la evolución, la proteína Muscleblind ha sufrido una duplicación interna que incluyó, al menos, los dos dedos de zinc. Otra posibilidad es que la proteína ancestral tuviera cuatro dedos CCCH que han sido conservados a lo largo de la evolución de los vertebrados, mientras que los invertebrados han sufrido una pérdida.

Como hemos comentado anteriormente, en invertebrados existen un solo gen *muscleblind* mientras que en vertebrados existen tres. La evolución de dos genes *mbi* adicionales en vertebrados puede reflejar la necesidad de regular y modular la función de Mbl. La genómica comparativa ha permitido determinar que durante la especiación de los vertebrados se han producido dos duplicaciones a nivel genómico que explican por qué los vertebrados suelen tener de 2-4 loci por cada gen en invertebrados (Ohno, 1999). En peces hemos encontrado cuatro genes *muscleblind* en *Tetraodon nigroviridis* y cinco en *Takifugu rubripes*. El análisis del genoma completo de *Tetraodon nigroviridis* sugiere que se produjo una duplicación a escala genómica en el linaje de los peces teleosteos después de su separación de los mamíferos, lo que podría explicar la presencia de más genes *mbi* en peces que en el resto de los vertebrados (Jaillon, 2004).

Cabe resaltar que dos de los dedos de zinc de vertebrados no tienen el mismo espaciado que los dedos de zinc de invertebrados. Los dedos segundo y cuarto han perdido dos aminoácidos entre la segunda y tercera cisteína. O, del mismo modo que en lo referente al número de dedos de zinc en las proteínas, estos dos aminoácidos pueden haberse ganado a lo largo de la evolución, si la proteína ancestral no los poseía. En caso de que se trate de una pérdida de residuos y una ganancia de dedos CCCH, este hecho debe haberse producido muy tempranamente en la especiación de los vertebrados concretamente al principio de la aparición de los cordados ya que la proteína de *Ciona intestinalis* tiene un solo gen *muscleblind* pero con cuatro dedos de zinc iguales que los vertebrados superiores.

1.5 Efecto del polimorfismo poli(dT) sobre MBNL1

Durante la búsqueda de mutaciones en el gen *MBNL1* en pacientes con distrofia miotónica (DM) encontramos un polimorfismo muy raro (<0.5%) en dos pacientes PROMM/DM2 no relacionados. Este polimorfismo consiste en la delección de un residuo de timina embebida en una secuencia de 11 timinas. Esta secuencia, a la que llamamos poli(dT), presenta por tanto dos alelos, el alelo 11T y el alelo raro 10T. La DM es una de las enfermedades humanas más variables, tanto por el abanico de síntomas que presenta como por el grado de afectación que puede existir para cada uno de ellos. Hoy en día se conoce la base genética del 99% de los casos de DM. Aunque la gran variabilidad fenotípica de la enfermedad queda explicada, en parte, por

la naturaleza de las mutaciones dinámicas que la originan, no se puede descartar que el background genético donde se produce la mutación module algún aspecto de la enfermedad.

En este sentido, investigamos la asociación del polimorfismo poli(T) en cuanto a la gravedad de los síntomas más característicos de la enfermedad como son la miotonía o la debilidad muscular. El genotipado de este polimorfismo en familias PROMM/DM2 con rasgos clínicos diferenciables no nos permitió establecer dicha asociación. No obstante, un estudio clínico detallado de los pacientes PROMM/DM2 en los que se ha encontrado este polimorfismo, podría indicar el/los rasgos clínicos comunes que indiquen qué grupo de pacientes habría que investigar.

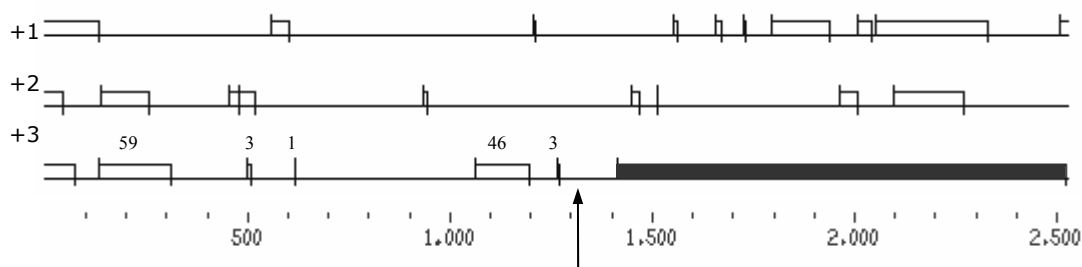
Un requisito imprescindible para que el alelo raro 10T pueda tener relevancia clínica es que tenga algún efecto sobre la expresividad del gen *MBNL1*. Por este motivo, discutimos a continuación los mecanismos por los que podría ocurrir esto. Dada la localización del polimorfismo en la región 5' UTR, la forma más plausible en la que puede afectar a la expresión del gen es regulando la tasa de traducción del mensajero. La eficiencia de la traducción puede ser modulada por características estructurales y de secuencia de las UTR del mensajero (Wilkie, 2003).

Efecto sobre las uORFs y uAUGs

Existen varias líneas de evidencia que indican que las pautas abiertas de lectura y los codones de inicio “aguas arriba” (*upstream*, respectivamente uORFs y uAUGs) del ATG de inicio principal regulan la expresión génica modulando la estabilidad del mensajero y la traducción (Iacono, 2005). Una uORF viene dada por un codón de inicio y un codón de parada en pauta que determinan una ORF anterior a la ORF principal. Un uAUG es un codón de inicio que no tiene ningún codón de parada en pauta antes del codón de inicio de la ORF principal por lo que puede competir por la maquinaria de traducción con este último y disminuir así los niveles de proteína. Para investigar el efecto que el alelo 10T puede tener sobre estas características de las UTRs de los mensajeros, analizamos la región 5'-UTR del gen *MBNL1* con el programa FRAME (GCG *package programs*). Como muestra la Figura 1, existen varias uORFs, algunas de ellas de un tamaño muy superior a la media (20 aa) aunque, pese a ser una UTR tan larga, no existe ningún uAUG. Esto indica que existe una fuerte presión de selección contra los uAUG en la región 5'-UTR de del gen *MBNL1*. La presencia del alelo 10T (delección de una timina respecto al alelo 11T) no afecta a ninguna uORF ni uATG (Figura 1). El único efecto que tiene el alelo 10T es la de

cambiar la pauta en la que coinciden las uORF respecto a la ORF principal. No parece que este hecho tenga ninguna relevancia sobre la expresión de la ORF principal según las recientes observaciones de Icono y col. (2005). Por este motivo, descartamos que la presencia del alelo 10T module la traducción alterando las uORFs que existen en el gen *MBNL1*.

a) Alelo normal (11T)



b) Alelo raro (10T)

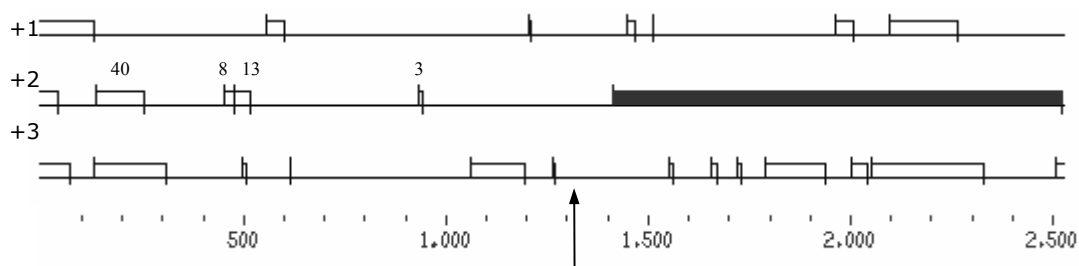


Figura 1. Efecto del alelo 10T en la región 5'-UTR de *MBNL1*. Se muestra la salida gráfica del programa Frames para la región 5'-UTR del cDNA AB007888 del gen *MBNL1* (1414 pb de 5'-UTR). Los codones de inicio y de parada potenciales se dibujan mediante líneas verticales por encima o por debajo de la horizontal respectivamente. La ORF de *MBNL1* aparece en gris. Los números por encima de las ORF anteriores corresponden al número de aminoácidos codificados por cada una de ellas. La flecha marca la posición del polimorfismo (delección de una timina). La ORF inmediatamente aguas arriba está a 215 pb de la ORF principal. El alelo raro (10T) provoca un cambio en la pauta de lectura en la que están ciertas uORF respecto a la ORF principal. Existen varios ejemplos de uORF con un claro efecto sobre la traducción de la ORF principal, sin embargo, no se ha visto que este efecto dependa de si están en pauta o no respecto de ella (revisado en Icono, 2005).

Efecto sobre la estructura secundaria

Para comprobar el efecto de la presencia del alelo 10T en la región 5'-UTR de *MBNL1*, en cuanto a las estructuras secundarias que pueden formarse en esta región, realizamos predicciones con el programa Mfold. Encontramos que es posible un cambio significativo de la estructura cuando está presente el alelo 10T ya que observamos que la energía libre de estas estructuras es mayor que para aquellas determinadas por el alelo 11T (Figura 2). Mayor energía libre implica mayor estabilidad. Una estructura secundaria más estable en la región 5' no traducida del

mensajero implicaría una tasa de traducción menor debido a una mayor dificultad en el escaneado que efectúan los ribosomas en busca del ATG de inicio. Menor tasa de traducción implicaría menor cantidad de proteína.

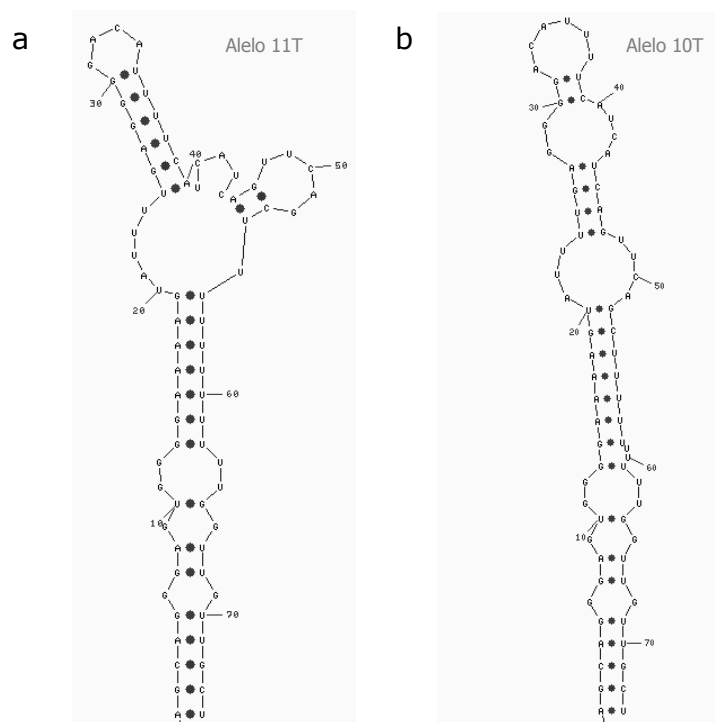


Figura 2. Efecto del polimorfismo 10T en la estructura secundaria de la región 5'-UTR. Se muestra la salida gráfica del programa de predicción de estructura secundaria Mfold (www.bioinfo.rpi.edu) para el alelo normal 11T (a) y el alelo raro 10T (b). Las energías libres asociadas a estas estructuras indican que la formada por el alelo raro ($\Delta G=15.50$ kcal/mol) es más estable que la formada por el alelo normal ($\Delta G=13.03$ kcal/mol).

Por tanto, nuestra hipótesis es que el polimorfismo poli(dT) afecta a la expresión del gen *MBNL1* modulando la tasa de traducción al alterar la estructura secundaria en su región 5'-UTR. La comprobación experimental de esta hipótesis puede realizarse mediante experimentos de traducción *in vitro* o mediante ensayos de transfección en cultivos celulares. De este modo, se compararían la cantidad de proteína que se obtiene a partir de mensajeros portadores del alelo 10T y del alelo 11T. En caso de confirmarse que la presencia del alelo 10T tiene un efecto sobre la tasa de traducción, y dada la relevancia de *MBNL1* en la ruta de patogénesis de la DM sería un buen aliciente para continuar los estudios de asociación que se han iniciado en este trabajo.

2. ANÁLISIS DE LA FUNCIÓN MOLECULAR DE *muscleblind* EN *Drosophila melanogaster*

Uno de los objetivos de este trabajo ha sido el estudio de la función molecular de *muscleblind* en *Drosophila*. Con este fin hemos desarrollado dos aproximaciones diferentes pero complementarias entre sí: una aproximación genética y una aproximación bioquímica. En la aproximación genética realizamos un rastreo genético de modificadores dominantes de un fenotipo de sobreexpresión de *mbi*. En la aproximación bioquímica, llevamos a cabo experimentos de co-inmunoprecipitación para identificar proteínas que interaccionen con la proteína Muscleblind a la vez que comprobamos la presencia de productos génicos candidatos identificados como resultado del rastreo genético.

2.1 Muscleblind está implicado en distintas funciones celulares

Entre los genes que han interactuado con *muscleblind* cabe destacar el hecho de que más de la mitad pertenecen a una de estas tres categorías: (1) proteínas relacionadas con el metabolismo del RNA; (2) Factores de transcripción; o (3) proteínas relacionadas con la regulación de la apoptosis.

2.1.1 *Muscleblind* y el metabolismo del RNA

Aspectos importantes del metabolismo del RNA como son, el transporte, la estabilidad, la localización en el citoplasma de los mensajero y el control de la traducción están acoplados e interconectados entre sí, ya que una misma proteína de unión a RNA puede afectar a muchas de estas funciones (revisado en Farina, 2002). Uno de los muchos ejemplos lo constituye la proteína hnRNAP A2. Esta proteína de unión a RNA está implicada en la localización del mRNA de la proteína básica mielina (MBP) en los oligodendrocitos de mamíferos pero también se ha relacionado con la función de *splicing* y la regulación de la traducción de ese mismo mensajero. De acuerdo con la implicación en todos estos procesos, y como ocurre con muchas otras proteínas de unión a RNA, la proteína hnRNAP A2 se localiza tanto en el núcleo como en el

citoplasma normalmente en gránulos ribonucleoproteicos o ribonucleocitoplasmáticos que contienen los mensajeros y el conjunto de proteínas necesarias para procesarlos.

Nuestros resultados indican que *muscleblind* podría estar implicado en el metabolismo del RNA ya que distintos genes relacionados con esta función son modificadores dominantes del fenotipo de sobreexpresión de *mbl* en ojo. Estos resultados son congruentes con la información funcional que existe acerca de las proteínas Muscleblind humanas y con los datos obtenidos en nuestro laboratorio sobre la función de la proteína en *Drosophila*.

Corte y empalme alternativo

Las proteínas Muscleblind están implicadas en el mecanismo de corte y empalme alternativo de transcritos. La proteína humana MBNL1 es una proteína de unión a RNA (Miller, 2000; Kino, 2004) que regula el *splicing* alternativo de mRNAs como la troponina cardiaca T (cTNT) y el receptor de la insulina (IR) (Ho, 2004). En nuestro laboratorio hemos encontrado que, de manera análoga, la proteína Muscleblind de *Drosophila* regula el *splicing* del mRNA de la α -actinina y, mediante experimentos de tres híbridos en levadura, que existe una unión *in vivo* de Mbl con un fragmento intrónico de dicho mRNA (M. Vicente, datos no publicados).

Los genes relacionados con el *splicing* de pre-mRNAs que hemos identificado son *arrest* (*aret*) y *no on or off transient A* (*nonA*). El gen *aret*, también conocido como *bruno*, se localiza en 33D3-E3, al lado de uno de sus homólogos *bruno-2*. Un microarray *developmental gene expression time course* indica que los momentos del desarrollo donde más se expresa es en la hembra adulta y en los primeros estadios embrionarios (Arbeitman, 2002; base de datos on-line : <http://genome.med.yale.edu/Lifecycle/>). Además, hibridaciones *in situ* en embriones evidencian que *aret* es un gen de efecto materno y que a lo largo del desarrollo embrionario se expresa casi exclusivamente en las células precursoras de la gónada (BDGP <http://www.fruitfly.org/cgi-bin/ex/insitu.pl>). Congruentemente, el gen *aret* fue identificado en un rastreo de mutaciones que causan infertilidad en hembras (Schupbach, 1991). *aret* codifica una proteína de unión a RNA que pertenece a la familia de las proteínas CELF, una familia muy conservada cuyos miembros están implicados en distintos aspectos del metabolismo del RNA. La proteína Aret es un regulador de la traducción de varios mRNAs de efecto materno (Webster, 1997). Un ejemplo es el mRNA de *oskar*. Aret se une a la región 3'-UTR de este mensajero reprimiendo su traducción hasta que éste alcanza la localización citoplasmática

adecuada dentro del ovocito (Castagnetti, 2000). *aret* es uno de los tres genes homólogos en *Drosophila* del gen CUGBP1 humano relacionado con el *splicing* de mensajeros implicados en la ruta de patogénesis de la DM. El efecto potenciador de mutaciones en *aret* sobre el fenotipo de ojos rugosos producido por la sobreexpresión de *mbi*, puede reflejar que ambos genes trabajan sobre los mismos mRNAs, pero con un efecto contrario, de manera que el efecto de la sobreexpresión de *mbi* sobre los mensajeros se ve acrecentado por la falta de función de *aret*, cuya presencia compensaría sus efectos. De acuerdo con esta idea, se ha demostrado que las proteínas humanas CUGBP1 y MBNL1 tienen efectos antagonistas sobre el *splicing* de ciertos mensajeros en el SNC (Jiang, 2004).

nonA se localiza en 14B18-C1. Se trata de un gen cuyas mutaciones afectan a la viabilidad, el comportamiento visual y el canto de cortejo en los macho de *Drosophila*. Codifica una proteína con motivos de unión a RNA del tipo RRM-1 (*RNA recognition motif*). Se ha propuesto como factor de *splicing* en base a su localización en loci concretos de los cromosomas politénicos y su gran homología con la proteína de mamíferos PSF/p54^{nrb}, que tiene una función reconocida como factor de *splicing* (Rendahl, 1996). Aunque *nonA* tiene una expresión homogénea y ubicua a lo largo del desarrollo, la relación con la función de *muscleblind* podría estar relacionada con los efectos de *nonA* en el sistema nervioso central donde podrían actuar sobre los mismos mRNA diana. La proteína NonA colocaliza bien con los focos formados por la expresión de 160 CTG en *Drosophila*, en los que también se localiza Muscleblind. Aunque la relevancia de este hecho no está clara, la proteína humana homóloga a NonA, PSF/p54^{nrb}, está implicada en la unión de ciertos mRNAs a la matriz nuclear. Muscleblind podría estar asociado a la matriz nuclear gracias a su interacción con *nonA* y desde allí ejercer su función como factor de *splicing*. A favor de esta hipótesis está la observación de la distribución subcelular de la proteína Mbl (Houseley, 2005):

Transporte, traducción y localización de mensajeros

Nuestros datos sugieren que Muscleblind podría estar implicado en otros aspectos del metabolismo del RNA aparte del corte y empalme de los mensajeros. Se han identificado dos genes modificadores cuyas proteínas forman parte del *Exon Junction Complex* (EJC), *Aly*, también conocido como REF, y *tsunagi*, también conocido como Y14. El EJC es depositado por el espliceosoma sobre los puntos en los que ha producido el corte y empalme de los exones de manera independiente a la secuencia del RNA. Este complejo está implicado en los procesos de transporte al citoplasma, degradación de mensajeros defectuosos mediante la ruta NMD (*non mediated decay*)

y localización de mensajeros. Recientemente, se ha comprobado el acoplamiento entre el *splicing* del mensajero de *oskar* y su localización citoplasmática (Hachet, 2004) aportando una nueva evidencia de la interconexión entre los distintos procesos del metabolismo del RNA.

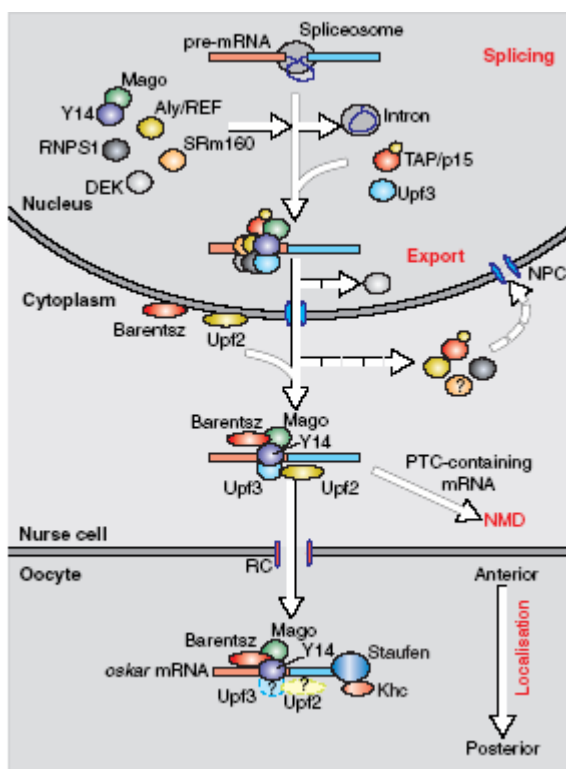


Figura 1. Modelo para el ensamblaje y la función del complejo del *Exon Junction Complex* (ECJ). Los seis componentes conocidos del ECJ (Y14, Mago, RNPS1, DEK, SRm160 y Aly/REF) son depositados sobre el mRNA de manera dependiente del espliceosoma, unos 6-8 pb aguas arriba del punto de unión entre dos exones. En el nucleoplasma, el ECJ recluta a Upf3 y el receptor de exportación heterodímero TAP/p15. Después de la translocación por el poro nuclear, varios componentes del complejo ribonucleoproteico exportado se disocian y vuelven al núcleo. Por el contrario Y14, Mago y Upf3 permanecen unidos al mRNA. Upf2, un factor de la degradación de mRNAs mediada por codones sin sentido (NMD), y Barentsz, una proteína requerida para la localización citoplasmática de oscar, son reclutados después de la translocación. Aquellos mensajeros que contienen un codón de parada prematuro (PTC) son degradados mediante la ruta del NMD.

El mRNA de *oskar* se transporta desde el polo anterior al posterior del ovocito gracias a la interacción de componentes específicos de la maquinaria de localización de mensajeros, entre los que se han identificado Staufen y la cadena pesada kinesina (Khc). El reciclado de SRm160 y el papel de Upf2 y Upf3 en la localización del mRNA de *oskar* no se conoce todavía (indicado con ?)

Además de formar parte del EJC, *Aly* y *tsu* están implicados en otros procesos del metabolismo del RNA, tales como la regulación de la traducción o la localización de los mensajeros a los que se unen, respectivamente. Cabe resaltar que una de las interacciones encontradas para Mbl en el ensayo de dos híbridos a nivel genómico de Giot y col. (2003) es con *Ef2*, que es un factor de elongación de la traducción. Mbl podría tener alguna función relacionada con la regulación de la traducción de sus mensajeros diana. Por otro lado, un trabajo reciente muestra que el gen humano *MBNL2* interviene en la localización citoplasmática y el control de la traducción del mensajero de la *integrina* $\alpha 3$ (Adhereth, 2005) por lo que no se puede descartar que esta función pueda estar conservada en la proteína Muscleblind de *Drosophila*. De hecho, mutantes *muscleblind* hipomorfos, presentan un fenotipo que recuerda, en

algunos aspectos, al de la falta de función de las integrinas, como es la aparición de burbujas en las alas (las integrinas mantienen unidas las dos capas del ala de *Drosophila*, de manera que una falta de función provoca que no se unan y por tanto la aparición de esas burbujas).

Aly actúa como un supresor dominante del fenotipo de ojos rugosos producidos por la sobreexpresión de la isoforma *mbIC*. Una posibilidad es que la supresión de *Aly* sea inespecífica, por ejemplo, si la función de *Aly* afecta a la exportación de los mensajeros de *mbIC* de manera que reduce la cantidad de mensajeros en el citoplasma, tendrá como consecuencia una menor cantidad de traducción de los mensajeros *mbIC* y por tanto una menor cantidad de proteína MblC. De esta manera, el fenotipo de ojo rugoso mejorará debido a una disminución en la expresión del transgén *UAS-mbIC*. En contra de un efecto inespecífico de *Aly*, observamos que la misma mutación en este gen, también modifica el fenotipo de ojos rugosos producido por la sobreexpresión de las repeticiones (CTG)₄₈₀. Ya que los dos mRNA no están relacionados, hace más probable que esté ocurriendo una interacción con la función de *mbIC*, bien por exceso (sobreexpresión *mbIC*), lo que ocasiona una mejora del fenotipo (supresión) o por defecto (sobreexpresión (CTG)₄₈₀) lo que produce un empeoramiento (potenciación).

Dada la función de Mbl como factor de corte y empalme, cabe esperar que la interacción génica entre *muscleblind* y otros genes implicados en el metabolismo del RNA refleje una relación a nivel de proteínas. Sin embargo, no hemos encontrado proteína *Aly* ni NonA en los inmunoprecipitados de MblC:GFP. Esto significa que, o bien no trabajan sobre las mismas moléculas diana o bien que aún ejerciendo su función sobre el mismo RNA diana, las proteínas no trabajan en el mismo complejo proteico, esto es, que no existe una relación física, directa ni indirecta entre ellas, por ejemplo que se unen a distintas secuencias del mensajero.

La identificación de los sucesos de *splicing* que regula Mbl, así como definir otras funciones de Mbl en el metabolismo del RNA, por ejemplo en la traducción, proporcionará pistas para explicar los fenotipos mutantes asociados a *mbI* y una base sobre la que reconstruir, al menos una parte, la ruta de patogénesis de la Distrofia miotónica. En este sentido, las moscas transgénicas MblC:GFP que hemos generado en este trabajo (ver material y métodos 2.6), representan una herramienta valiosa ya que pueden emplearse para la colP de mRNA asociados a Mbl y para visualizar la proteína asociada a los mismos, en tejidos o células.

2.1.2 Muscleblind y la regulación de la transcripción

Una de la categoría de genes más representada entre los genes modificadores del fenotipo de sobreexpresión de *mbl* es el de los factores de transcripción. Esta categoría incluye a proteínas con función de factor de transcripción en sentido amplio, esto es, tanto proteínas con motivos de unión a DNA, que aumentan la actividad de la RNA polimerasa, como proteínas que regulan la transcripción de los genes controlando la estructura de la cromatina. De igual modo, una de las categorías más representadas en el rastreo de dos híbridos en levadura de Giot y colaboradores (2003) es el de los factores de transcripción de manera que, de entre los 19 genes de función conocida cuyas proteínas interactúan con Mbl, seis de ellos están catalogados como factores de transcripción entre los que se incluyen dos proteínas que aumentan la actividad de la RNA polimerasa II.

Los genes identificados en el rastreo de modificadores llevado a cabo en este trabajo que están relacionados con la regulación de la transcripción son: *amos*, *CG15435*, *DP transcription factor*, *CG15433* y *jumeaux*.

Se ha estudiado con más detalle la posible relación de Mbl con la proteína codificada por *jumeaux*, Jumu, debido a un solapamiento de su patrón de expresión en varios tejidos como disco imaginal de ojo, SNC y SNP. Además los fenotipos de ambos mutantes resultan similares en cuanto a las estructuras afectadas como son ojo, alas y quetas. Y finalmente porque los resultados de Giot y col (2003) muestran una interacción física entre Mbl y Jumu en experimentos de dos híbridos en levaduras. Todo esto nos hizo pensar que Mbl y Jumu podían interactuar genética y físicamente.

Nuestro estudio genético sobre la relación funcional entre *mbl* y *jumeaux* demuestra que existe una interacción genética tanto en ojo como en ala. En ojo, hemos confirmado la interacción usando tanto la sobreexpresión de la isoforma *mbIC* como la sobreexpresión de las repeticiones (CTG)₄₈₀. En ala, hemos comprobado que *jumeaux* es un supresor dominante del fenotipo de sobreexpresión de *mbIC*. Hemos descartado que el efecto de reducir la dosis de *jumeaux* sobre los fenotipos de sobreexpresión se produzca indirectamente debido un requerimiento de Jumu como factor de transcripción para expresar el transgén *UAS-mbIC*, ya que *jumeaux* no interactúa con otros fenotipos sensibilizados (S. Muñoz-Descalzo, comunicación personal)

Jumu es una proteína con motivos del tipo FH/WH que actúa como factor de transcripción uniéndose al DNA para determinar la estructura de la cromatina. La proteína Jumu se localiza en loci concretos de cromosomas politénicos, de glándulas salivares de larvas de tercer estadio. Este gen participa en el proceso de neurogénesis y el desarrollo de ojo y ala. En este sentido, hemos comprobado que la proteína Mbl también se localiza en loci concretos de los cromosomas politénicos lo que indica que ejerce su función sobre un grupo limitado de genes. Hemos detectado aproximadamente 60 señales correspondientes a la proteína MblC:GFP sobre los cromosomas politénicos de las glándulas salivares de larvas de tercer estadio de *Drosophila*. Aproximadamente un 30% de estas señales coinciden con las de Jumu,. Esto podría indicar que Mbl y Jumu ejercen su función sobre un grupo común de genes diana, aunque la identidad de los mismos no ha sido establecida todavía. La señal de Mbl en los cromosomas politénicos puede reflejar la asociación de Mbl a complejos DNA/proteína que afecten a la transcripción, aunque también podría reflejar su tampoco asociación con los pre-mRNAs.

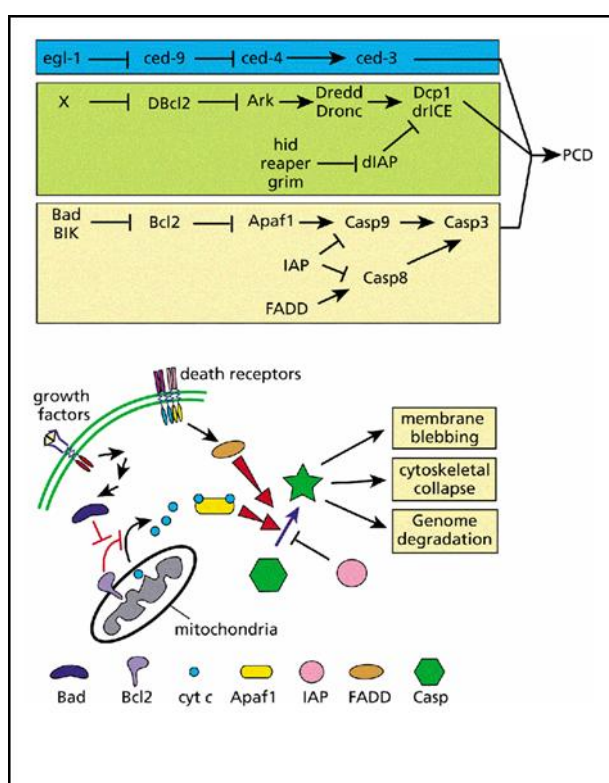
El gen *jumeaux* fue identificado como candidato para la interacción con *muscleblind* por los resultados del estudio genómico de dos híbridos de Giot y col. (2003), en los cuales se obtuvo una interacción débil entre estas dos proteínas. Con el objetivo de confirmar este resultado, realizamos experimentos de coIP *in vivo* con la proteína de fusión MblC:GFP, experimentos de dos híbridos con la porción C-terminal y N-terminal de la proteína Jumu, así como experimentos de *pull-down* con la porción C-terminal. Los resultados obtenidos indican que *muscleblind* no interacciona físicamente con los fragmento analizados de Jumu, al menos bajo las condiciones experimentales probadas (datos de dos híbridos y *pull-down*, A. Schwendemann, comunicación personal).

Hemos estudiado la interacción física entre Mbl y dos proteínas implicadas en la regulación de la transcripción, Jumu y Amos. En ninguno de los dos casos, hemos detectado proteína coimmunoprecipitada.

2.1.3 Muscleblind y la apoptosis

La apoptosis, o muerte celular programada, es un proceso del desarrollo fundamental para dar forma a los órganos, eliminar estructuras transitorias, ajustar el tamaño de las poblaciones de células que están interaccionando o para eliminar células que son

potencialmente dañinas o no funcionales (revisado en McCarthy, 2003). Existe una clara conservación evolutiva de las rutas por las que los diferentes organismos regulan el número de células. Esta conservación ha permitido generar un modelo general de inducción de la apoptosis en el cual múltiples rutas de señalización conducen finalmente a la activación de las Caspasas. Una vez activadas, estas proteasas actúan sobre sustratos celulares claves que desencadenan el ordenado proceso de muerte celular que caracteriza a la apoptosis. En *Drosophila*, la apoptosis también está regulada por una familia muy conservada de proteínas, las Caspasas. Además se ha descubierto el papel regulador clave de tres genes, *reaper*, *head involution defective (hid)* y *grim*, en la activación de la apoptosis. Estos genes se requieren para prácticamente todos los procesos apoptóticos que ocurren en el embrión. Los tres genes comparten una región de 14 aminoácidos en su región NH₂ terminal que



constituye el dominio RHG (Reaper Hid Grim) que les confiere propiedades específicas en la inducción de la apoptosis (Wing, 2001). Este dominio interacciona con las proteínas de la familia de Inhibidores de la Apoptosis (IAP) que incluye a las proteínas Diap1 y Diap2 (Goyal, 2000).

En el rastreo de interacciones genéticas que hemos llevado a cabo con el fenotipo de ojo rugoso por la sobreexpresión de *mbIC* en ojo, hemos identificado tres genes modificadores relacionados con la apoptosis *reaper*, *Traf1* y *diap-1*.

El gen *reaper*, abreviadamente *rpr*, se localiza en 75C6 y codifica una proteína proapoptótica que se caracteriza por poseer el motivo RHG que comparte con los genes proapoptóticos Hid y grim. La acción proapoptótica de *rpr* se produce por la inactivación de las proteínas DIAP que son inhibidores de la apoptosis. El gen *rpr* se expresa durante todo el desarrollo embrionario, con una señal específica en el SNC que se hace más evidente alrededor del estadio 11 (Lohmann, 2003). En los mutantes

rpr los neuroblastos del SNC en desarrollo donde se expresa sobreviven y proliferan lo que causa una hiperplasia neural. De acuerdo con esto, se observa que las moscas adultas presentan un SNC más largo de lo normal (Perterson, 2002). Los embriones mutantes *rpr* no presentan el patrón apoptótico alterado, por lo que se cree que *rpr* es requerido para la apoptosis de células que se produce tras la eclosión de la larva. De acuerdo con esto, la expresión de *rpr* es inducida por la ecdisona que se produce durante la metamorfosis. Recientemente se ha demostrado que *reaper* induce apoptosis vía la ruta JNK a través de DIAP1 (Kuranaga, 2002). Curiosamente, cuando se sobreexpresa *rpr* en el sistema nervioso central se observa la desaparición de las motoneuronas aCC y pCC (Landgraf, 2003) donde Mbl se expresa específicamente (García-Casado, 2002)

El gen *TNF-receptor Associated Factor1*, abreviadamente *Traf1*, se localiza en la posición 24E4. Es uno de los dos genes de *Drosophila* que pertenecen a la familia de proteínas TRAF. Esta familia incluye proteínas adaptadoras que actúan como transductores de la señal de la superfamilia del receptor *tumor necrosis factor* (TNF). Esto conduce a la activación de la ruta de las JNK entre otras, lo que, en algunos casos, se ha asociado con la activación de la apoptosis. La sobreexpresión ectópica de *Traf1* en el ojo induce apoptosis y ésta es dependiente de la ruta JNK (Cha, 2003).

El gen *thread* (*th*), también conocido como *diap-1*, se localiza en la posición 72D1. Codifica una proteína con motivos *baculovirus IAP repeat* (BIR) que caracteriza a la familia de proteínas IAPs (*Inhibition of Apoptosis protein family*). Además contiene un dominio RING finger en su extremo C-terminal que le confiere actividad ubiquitina ligasa E3 (Salvense, 2002). La sobreexpresión de este gen protege de la muerte celular inducida por *reaper* y otros genes proapoptóticos (Lisi, 1998). Las propiedades antiapoptóticas de Thread y otras proteínas IAP está mediada por la inhibición que ejercen sobre las Caspasas.

Hemos determinado que la causa molecular que explica el fenotipo de sobreexpresión de *mbIC* en ala es la apoptosis. La sobreexpresión de *mbIC* activa la apoptosis en disco imaginal de ala, en concreto en la parte superior de lo que será el segundo compartimento posterior del ala adulta, lo que correlaciona bien con la falta de tejido laminar. Esto significa que existen proteínas específicas de células que interacciona con el exceso de proteína Muscleblind y cuya consecuencia es la activación de la apoptosis.

En resumen, como resultado del rastreo genético, hemos encontrado tres proteínas que tienen un papel relevante en el proceso de la apoptosis. Es difícil determinar si estas interacciones son específicas o se deben a una respuesta celular protectora frente a la expresión en grandes cantidades de una proteína cualquiera como por ejemplo como MblC. A favor de que esta interacción sea específica está la observación de que no existe apoptosis en los discos imaginales de ojo de moscas que sobreexpresan *mbIC*, lo que indica que las células no responden a la presencia *per se* de la proteína MblC. Además, el hecho de que el fenotipo de ojo rugoso interaccione con diap-I y no con diap-II apoya la especificidad de la interacción. Diap1 y Diap2 responden a la apoptosis generada por los genes proapoptóticos *rpr*, *him* y *grim*. Sin embargo, la forma en que lo hacen puede ser diferente. Por ejemplo, diap1 parece interactuar físicamente con muchas Caspasas. Sin embargo, esto no ocurre con Diap2, lo que hace pensar que actúan por vías diferentes (Hay, 2000). Por último, cuando la sobreexpresión de *mbIC* activa la apoptosis, como por ejemplo en ala (resultados 2.1), no lo hace en toda la superficie donde esta presente si no que lo hace en regiones concretas, lo que determina que en esos puntos, la proteína Mbl interacciona con proteínas específicas para llevar a cabo su función, de manera que un desequilibrio en estas interacciones induce apoptosis por una vía que, de momento, no hemos identificado.

Cabe resaltar que tanto la sobreexpresión de *reaper*, un gen que induce apoptosis, como la reducción de la dosis génica de *thread*, un gen con un efecto inhibitor de la misma, potencian el fenotipo de ojos rugosos. Ambas situaciones van a favor de un aumento de la apoptosis. Curiosamente, la sobreexpresión conjunta de *Traf1* (línea GS2154) con *mbIC* suprime el fenotipo de ojos rugosos. Esto puede indicar un efecto antiapoptótico de Mbl ya que los ojos que sobreexpresan *Traf1* son rugosos debido a apoptosis. También puede indicar que la inserción de la línea *Traf1* empleada (resultados 2.1.1.1) genera una falta de función del gen, de manera que el efecto observado resulta de la sobreexpresión de ambos genes en un fondo mutante *Traf1*.

Un enlace posibles entre la función molecular conocida de *muscleblind* y la apoptosis podría ser que Mbl interacciona con ciertos mRNAs que codifican proteínas cuya falta o exceso desencadena apoptosis. Como ya se ha comentado, un descubrimiento reciente demuestra que la proteína humana *MBNL2* está implicada en la localización del mensajero de la integrina $\alpha 3$ en la adhesiones focales (Adhereth, 2005). La ausencia de integrinas en las adhesiones focales impide que las células se unan a la membrana basal y esto es un desencadenante de la apoptosis en estas células

(Giancotti, 1999). De igual manera, en el ojo podría ocurrir que una localización o regulación del *splicing* errónea para determinado mensajero, pueda desencadenar apoptosis.

2.2 Las moscas transgénicas MbC:GFP son una herramienta útil para colP

Con el fin de conseguir la colP de proteínas que interaccionen con Muscleblind *in vivo*, hemos generado moscas transgénicas que sobreexpresan la proteína de fusión MbC:GFP. Para ello, hemos comprobado que la proteína MbC:GFP es funcional en un ensayo *in vivo*, que consiste en rescatar el fenotipo de letalidad embrionaria asociada a los alelos de *muscleblind*. Las mutaciones en *mbI* provocan la muerte de los embriones en los últimos estadios del desarrollo embrionario impidiendo que las larvas eclosionen. La expresión de MbC:GFP en un fondo mutante para *muscleblind* permite la supervivencia de prácticamente todos los embriones mutantes al igual que la sobreexpresión de *mbI* por lo que se concluye que la etiqueta GFP no interfiere en la función normal de la proteína.

En nuestros experimentos, encontramos un rescate de la letalidad embrionaria. Este rescate es parcial en cuanto a la "calidad" del rescate, ya que los embriones rescatados no son capaces de completar el ciclo de vida. Este rescate parcial puede deberse a varias causas. En primer lugar, cada gen posee un patrón espacio-temporal de expresión que determina en qué tejidos y cuándo debe estar presente la proteína. Este patrón de expresión viene determinado principalmente por las regiones reguladoras de la transcripción del gen. La expresión de la proteína MbC:GFP no se ha dirigido con el mismo patrón que lo hace la proteína endógena, ya que no se conocen las regiones reguladoras del gen *muscleblind*. Por tanto, no podemos esperar un rescate total del fenotipo de las moscas mutantes *mbI*. Además, existe otro nivel de complejidad en la expresión de *mbI* que tampoco se ha reproducido. MbI tiene distintas isoformas proteicas que ejercen funciones específicas a lo largo del desarrollo (García-Casado, 2002; Monferrer, 2005). Los embriones mutantes carecen de todas estas isoformas y en este experimento de rescate solo estamos aportando una de ellas, la isoforma C. Finalmente, no podemos descartar que la fusión de la GFP pueda limitar alguna función particular de la proteína *mbI* que sea necesaria en un tejido o momento dado del desarrollo.

Usando las moscas transgénicas descritas, hemos identificado seis bandas coimmunoprecipitadas específicamente por la proteína MblC:GFP. La determinación de la identidad de las proteínas que se corresponden con estas bandas ayudará a entender la función de Mbl. Dado el diseño del experimento, cabe la posibilidad de que algunas de estas proteínas pueden haber sido colP únicamente por la fracción GFP de la proteína de fusión en este caso no aportarían información alguna sobre la función de *muscleblind*.

Por otro lado, hemos comprobado la presencia de proteínas en los inmunoprecipitados de MblC:GFP. De este modo, hemos descartado la presencia de las proteínas Aly, Amos, NonA y Jumu, bajo las condiciones de coimmunoprecipitación utilizadas.

En este trabajo hemos generado moscas transgénicas que sobreexpresan una proteína de fusión MblC:GFP, que es funcional *in vivo*, con el fin de IP proteínas que interaccionen con Mbl. Estas moscas constituyen, además, una herramienta de interés para otras aplicaciones como es la inmunolocalización de Mbl en los cromosomas politénicos, la colocación de Mbl en los focos formados por las repeticiones (CTG)₄₈₀ (L. Monferrer, no publicado) y la colP de los mensajeros diana asociados Mbl.

CONCLUSIONES

1. La estructura genómica de los tres genes humanos homólogos a *muscleblind*, MBNL1-3, está altamente conservada. Sin embargo, no se detecta conservación respecto al gen *muscleblind* de *Drosophila*
2. Los genes MBNL1-3 humanos, al igual que *muscleblind* en *Drosophila*, sufren un complejo patrón de corte y empalme alternativo que no afecta a los dedos de zinc de la proteína. MBNL1 es el gen que posee un mayor número de exones opcionales. Hemos descrito nueve variantes de *splicing* en MBNL1, tres en MBNL2 y cuatro en MBNL3.
3. Hemos identificado exones de uso poco frecuente en el procesado de los mensajeros MBNL1. El uso del exón opcional 1a de MBNL1 produce una variante de *splicing* cuya isoforma proteica es análoga a la isoforma *mbID* de *Drosophila*.
4. Los genes humanos MBNL1-3 tienen un patrón de expresión diferencial siendo MBNL1 es el gen que más se expresa tanto por el número de transcritos por tejido como por la cantidad de tejidos en los que se detecta expresión.
5. Hemos analizado los genes *muscleblind* en 12 especies representativas de la escala filogenética. Las proteínas Muscleblind constituyen una nueva familia proteica con motivos de unión a RNA del tipo CX₇CX₆CX₃H, que presentan otras regiones conservadas tanto en vertebrados como en invertebrados. Hemos definido las cajas LEV, GVA y INGRN como regiones de aminoácidos altamente conservadas.
6. Los nematodos son el *Phylum* más antiguo en el que se reconoce la proteína Muscleblind. Existe un sólo gen *muscleblind* en invertebrados mientras que en vertebrados encontramos entre tres (ej. *Mus musculus*) y cinco (ej. *Takifugu rugrupes*) genes.
7. Las proteínas Muscleblind de invertebrados poseen dos dedos de zinc mientras que las proteínas de vertebrados poseen cuatro. La homología entre los dedos de zinc de vertebrados e invertebrados sugiere que el origen de los dedos tercero y cuatro, exclusivos de vertebrados, se ha producido por duplicación de los dos primeros. La presencia en *Ciona intestinalis* de cuatro dedos de zinc sugiere que dicha duplicación se produjo en las primeras etapas de la especiación de vertebrados

8. Hemos descartado la presencia de mutaciones en la región codificante o en los sitios aceptores/donadores de *splicing* en el gen *MBNL1* en pacientes con Distrofía miotónica y en el gen *MBNL3*, en pacientes con Ptoxis congénita asociada al cromosoma X. También se ha descartado la presencia de mutaciones en la región codificante del gen *MBNL1* en pacientes con síndrome de Usher tipo III.
9. Hemos encontrado un polimorfismo muy raro (<0.5%) en la región 5'-UTR del gen *MBNL1* en dos pacientes DM2/PROMM, no relacionados entre sí, que consiste en la delección de una timina en una región rica en timinas, conservada evolutivamente, a 72 pb aguas arriba del punto de inicio de la traducción. La presencia de este polimorfismo no se ha podido correlacionar con la gravedad de síntomas relevantes de la enfermedad
10. Hemos identificado 20 genes que interactúan genéticamente con *muscleblind* que sugieren que las proteínas Muscleblind están implicadas en procesos celulares como apoptosis, control de la expresión génica, transporte, localización de mensajeros en el citoplasma, adhesión celular y *splicing* de pre-mRNAs.
11. Hemos generado moscas transgénicas que expresan la proteína de fusión MbIC:GFP y hemos comprobado que esta proteína es funcional *in vivo*. Esta herramienta resulta útil para coimmunoprecipitar proteínas y RNAs que interactúen físicamente con Muscleblind.
12. Hemos puesto a punto la técnica de coimmunoprecipitación de proteínas y hemos identificado seis bandas que coimmunoprecipitan específicamente con MbIC:GFP. La identificación de proteínas en gel mediante espectrometría de masas requiere escalar el protocolo para obtener mayor cantidad de proteínas.
13. Las proteínas Aly, Amos, NonA y Jumu no se detectan mediante inmunodetección en los coimmunoprecipitados de MbIC:GFP
14. La sobreexpresión induce apoptosis en una región específica del disco de ala que se correlaciona bien con la falta de tejido intervena en las alas adultas.
15. Muscleblind se localiza aproximadamente en 60 loci diferentes sobre cromosomas politénicos

BIBLIOGRAFÍA

- A -

Adereth Y, Dammai V, Kose N, Li R, Hsu T. (2005). RNA-dependent integrin $\alpha 3$ protein localisation regulated by the Muscleblind-like protein MLP1. *Nature Cell biology* (en prensa)

Akasaka T, Lossos IS, Levy R. (2003). *BCL6* gene translocation in follicular lymphoma: a harbinger of eventual transformation to diffuse aggressive lymphoma. *Blood*. Aug 15;102(4):1443-8.

Alfonso Martinez-Arias y Alison Steward. (2002). Molecular principles of animal development. May. Pág.51. *Oxford University Press*

Alwazzan M, Newman E, Hamshere MG, Brook JD. (1999). Myotonic dystrophy is associated with a reduced level of RNA from the *DMWD* allele adjacent to the expanded repeat. *Hum Mol Genet*. Aug;8(8):1491-7.

Amack JD, Paguio AP, Mahadevan MS. (1999). *Cis* and *trans* effects of the myotonic dystrophy (DM) mutation in a cell culture model. *Hum Mol Genet*. Oct;8(11):1975-84.

Anderson JL, Habashi J, Carlquist JF, Muhlestein JB, Horne BD, Bair TL, Pearson RR, Hart N. (2000). A common variant of the *AMPD1* gene predicts improved cardiovascular survival in patients with coronary artery disease. *J Am Coll Cardiol*. Oct;36(4):1248-52.

Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP. (2002). Gene expression during the life cycle of *Drosophila melanogaster*. *Science*. Sep 27;297(5590):2270-5.

Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, Mlodzik M, Perez-Alonso M, Baylies MK. (1998). The *muscleblind* gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by *Dmef2*. *Dev Biol*. Mar 15;195(2):131-43.

Artero R. (1995). Caracterización molecular de la región 54A de *Drosophila melanogaster*. Tesis Doctoral. Universidad de Valencia.

Auluck PK, Chan HY, Trojanowski JQ, Lee VM, Bonini NM. (2002). Chaperone suppression of alpha-synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science*. Feb 1;295(5556):865-8.

- B -

Begemann G, Paricio N, Artero R, Kiss I, Perez-Alonso M, Mlodzik M. (1997) *muscleblind*, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development*. Nov;124(21):4321-31.

Berg J, Jiang H, Thornton CA, Cannon SC. (2004). Truncated CIC-1 mRNA in myotonic dystrophy exerts a dominant-negative effect on the Cl current. *Neurology*. Dec 28;63(12):2371-5.

Berul CI, Maguire CT, Gehrman J, Reddy S. (2000). Progressive atrioventricular conduction block in a mouse myotonic dystrophy model. *J Interv Card Electrophysiol*. Jun;4(2):351-8.

Betz A, Lampen N, Martinek S, Young MW, Darnell JE Jr. (2001). A *Drosophila* PIAS homologue negatively regulates *stat92E*. *Proc Natl Acad Sci U S A*. Aug 14;98(17):9563-8.

Blackshear PJ. (2002). Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover. *Biochem Soc Trans*. Nov;30(Pt 6):945-52.

Boucher CA, King SK, Carey N, Krahe R, Winchester CL, Rahman S, Creavin T, Meghji P, Bailey ME, Chartier FL, et al. (1995). A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)_n repeat. *Hum Mol Genet*. Oct;4(10):1919-25.

Boucher CA, King SK, Carey N, Krahe R, Winchester CL, Rahman S, Creavin T, Meghji P, Bailey ME, Chartier FL, et al. (1995). A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)_n repeat. *Hum Mol Genet*. Oct;4(10):1919-25.

Brody T, Stivers C, Nagle J, Odenwald WF. (2002). Identification of novel *Drosophila* neural precursor genes using a differential embryonic head cDNA screen. *Mech Dev*. Apr;113(1):41-59.

Brown RS. (2005). Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol*. Feb;15(1):94-8.

Busygina V, Suphapeetiporn K, Marek LR, Stowers RS, Xu T, Bale AE. (2004). Hypermutability in a *Drosophila* model for multiple endocrine neoplasia type 1. *Hum Mol Genet*. Oct 15;13(20):2399-408.

- C -

Caricasole A, Duarte A, Larsson SH, Hastie ND, Little M, Holmes G, Todorov I, Ward A. (1996). RNA binding by the Wilms tumor suppressor zinc finger proteins. *Proc Natl Acad Sci U S A*. Jul 23;93(15):7562-6.

Castagnetti S, Hentze MW, Ephrussi A, Gebauer F. (2000). Control of oskar mRNA translation by Bruno in a novel cell-free system from *Drosophila* ovaries. *Development*. Mar;127(5):1063-8.

Colleran JA, Hawley RJ, Pinnow EE, Kokkinos PF, Fletcher RD. (1997). Value of the electrocardiogram in determining cardiac events and mortality in myotonic dystrophy. *Am J Cardiol*. Dec 1;80(11):1494-7.

Crowther DC, Kinghorn KJ, Page R, Lomas DA. (2004). Therapeutic targets from a *Drosophila* model of Alzheimer's disease. *Curr Opin Pharmacol*. Oct;4(5):513-6.

Cha GH, Cho KS, Lee JH, Kim M, Kim E, Park J, Lee SB, Chung J. (2003). Discrete functions of TRAF1 and TRAF2 in *Drosophila melanogaster* mediated by c-Jun N-terminal kinase and NF-kappaB-dependent signaling pathways. *Mol Cell Biol*. Nov;23(22):7982-91.

Charlet-B N, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA. (2002). Loss of the *muscle-specific chloride channel* in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol Cell*. Jul;10(1):45-53.

Cheah PY, Chia W, Yang X. (2000). *Jumeaux*, a novel *Drosophila* winged-helix family protein, is required for generating asymmetric sibling neuronal cell fates. *Development*. Aug;127(15):3325-35.

Chen L, Feany MB. (2005). Alpha-synuclein phosphorylation controls neurotoxicity and inclusion formation in a *Drosophila* model of Parkinson disease. *Nat Neurosci*. May;8(5):657-63.

- D -

Davies RC, Calvio C, Bratt E, Larsson SH, Lamond AI, Hastie ND. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. *Genes Dev*. Oct 15;12(20):3217-25.

Davis BM, McCurrach ME, Taneja KL, Singer RH, Housman DE. (1997). Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc Natl Acad Sci U S A*. Jul 8;94(14):7388-93.

Day JW, Ricker K, Jacobsen JF, Rasmussen LJ, Dick KA, Kress W, Schneider C, Koch MC, Beilman GJ, Harrison AR, Dalton JC, Ranum LP. (1999). Myotonic dystrophy type 2: molecular, diagnostic and clinical spectrum. *Neurology*. Feb 25;60(4):657-64.

Day JW, Roelofs R, Leroy B, Pech I, Benzow K, Ranum LP. (1999). Clinical and genetic characteristics of a five-generation family with a novel form of myotonic dystrophy (DM2). *Neuromuscul Disord*. Jan;9(1):19-27.

- E -

Ebralidze A, Wang Y, Petkova V, Ebralidse K, Junghans RP. (2004). RNA leaching of transcription factors disrupts transcription in myotonic dystrophy. *Science*. Jan 16;303(5656):383-7

- F -

Fardaei M, Larkin K, Brook JD, Hamshere MG. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res*. Jul 1;29(13):2766-71.

Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, Harper PS, Brook JD. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet*. Apr 1;11(7):805-14.

Farina KL, Singer RH. (2002). The nuclear connection in RNA transport and localization. *Trends Cell Biol*. Oct;12(10):466-72.

Fessler JH, Fessler LI. (1989). *Drosophila* extracellular matrix. *Annu Rev Cell Biol*. ;5:309-39.

Fogerty FJ, Fessler LI, Bunch TA, Yaron Y, Parker CG, Nelson RE, Brower DL, Gullberg D, Fessler JH. (1994). *Tiggrin*, a novel *Drosophila* extracellular matrix protein that functions as a ligand for *Drosophila* alpha PS2 beta PS integrins. *Development*. Jul;120(7):1747-58.

Fortini ME, Skupski MP, Boguski MS, Hariharan IK. (2000). A survey of human disease gene counterparts in the *Drosophila* genome. *J Cell Biol.* Jul 24;150(2):F23-30.

Frasch M, Saumweber H. (1989). Two proteins from *Drosophila* nuclei are bound to chromatin and are detected in a series of puffs on polytene chromosomes. *Chromosoma.* Jan;97(4):272-81.

Fu YH, Friedman DL, Richards S, Pearlman JA, Gibbs RA, Pizzuti A, Ashizawa T, Perryman MB, Scarlato G, Fenwick RG Jr, et al. (1993). Decreased expression of *myotonin-protein kinase* messenger RNA and protein in adult form of myotonic dystrophy. *Science.* Apr 9;260(5105):235-8.

- G -

Garcia-Casado MZ, Artero RD, Paricio N, Terol J, Perez-Alonso M. (2002). Generation of GAL4-responsive *muscleblind* constructs. *Genesis.* Sep-Oct;34(1-2):111-4.

Garcia-Casado Z . (2002). Caracterización funcional del gen *muscleblind*: valoración de *Drosophila* como modelo para el estudio de la Distrofia Miotónica. *Tesis Doctoral.* Universidad de Valencia.

Gatfield D, Izaurralde E. (2002). REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *J Cell Biol.* Nov 25;159(4):579-88.

Giancotti FG, Ruoslahti E. (1999). Integrin signaling. *Science.* Aug 13:1028-1032

Gibbs JW 3rd, Ciafaloni E, Radtke RA. (2002). Excessive daytime somnolence and increased rapid eye movement pressure in myotonic dystrophy. *Sleep.* Sep 15;25(6):662-5.

Giot L, Bader JS, Brouwer y otros. (2003). A protein interaction map of *Drosophila melanogaster*. *Science.* Dec 5;302(5651):1727-36.

Goedert M, Spillantini MG, Jakes R, Rutherford D, Crowther RA. (1989). Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron.* Oct;3(4):519-26.

Gomperts M, Pascall JC, Brown KD. (1990). The nucleotide sequence of a cDNA encoding an EGF-inducible gene indicates the existence of a new family of mitogen-induced genes. *Oncogene.* Jul;5(7):1081-3.

Goyal L, McCall K, Agapite J, Hartweg E, Steller H. (2000). Induction of apoptosis by *Drosophila reaper*, *hid* and *grim* through inhibition of IAP function. *EMBO J.* Feb 15;19(4):589-97.

Guarnieri DJ, Heberlein U. (2003). *Drosophila melanogaster*, a genetic model system for alcohol research. *Int Rev Neurobiol.* ;54:199-228.

- H -

Hachet O, Ephrussi A. (2001). *Drosophila* Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Curr Biol.* Oct 30;11(21):1666-74.

Hachet O, Ephrussi A. (2004). Splicing of *oskar* RNA in the nucleus is coupled to its

cytoplasmic localization. *Nature*. Apr 29;428(6986):959-63.

Hamshere MG, Harley H, Harper P, Brook JD, Brookfield JF. (1999). Myotonic dystrophy: the correlation of (CTG) repeat length in leucocytes with age at onset is significant only for patients with small expansions. *J Med Genet*. Jan;36(1):59-61.

Hari KL, Cook KR, Karpen GH. (2001). The *Drosophila Su(var)2-10* locus regulates chromosome structure and function and encodes a member of the PIAS protein family. *Genes Dev*. Jun 1;15(11):1334-48.

Harper PS. (2001). Myotonic dystrophy, vol 37. London: *W.B. Saunders*;

Hay BA. (2000). Understanding IAP function and regulation: a view from *Drosophila*. *Cell Death Differ*. Nov;7(11):1045-56.

Haywood AF, Staveley BE. (2004). Parkin counteracts symptoms in a *Drosophila* model of Parkinson's disease. *BMC Neurosci*. Apr 16;5(1):14.

Ho TH, Charlet-B N, Poulos MG, Singh G, Swanson MS, Cooper TA. (2004). Muscleblind proteins regulate alternative splicing. *EMBO J*. Aug 4;23(15):3103-12.

Hofmann-Radvanyi H, Lavedan C, Rabes JP, Savoy D, Duros C, Johnson K, Junien C. (1993). Myotonic dystrophy: absence of CTG enlarged transcript in congenital forms, and low expression of the normal allele. *Hum Mol Genet*. Aug;2(8):1263-6.

Houseley JM, Wang Z, Brock GJ, Soloway J, Artero R, Perez-Alonso M, O'Dell KM, Monckton DG. (2005). Myotonic dystrophy associated expanded CUG repeat *muscleblind* positive ribonuclear foci are not toxic to *Drosophila*. *Hum Mol Genet*. Mar 15;14(6):873-83.

Houy E, Raux G, Thibaut F, Belmont A, Demily C, Allio G, Haouzir S, Fouldrin G, Petit M, Frebourg T, Campion D. (2004). The promoter -194 C polymorphism of the nicotinic alpha 7 receptor gene has a protective effect against the P50 sensory gating deficit. *Mol Psychiatry*. Mar;9(3):320-2.

Hund E, Jansen O, Koch MC, Ricker K, Fogel W, Niedermaier N, Otto M, Kuhn E, Meinck HM. (1997). Proximal myotonic myopathy with MRI white matter abnormalities of the brain. *Neurology*. Jan;48(1):33-7.

- I -

Iacono M, Mignone F, Pesole G. (2005). uAUG and uORFs in human and rodent 5'untranslated mRNAs. *Gene*. Apr 11;349:97-105.

- J -

Jackson GR, Salecker I, Dong X, Yao X, Arnheim N, Faber PW, MacDonald ME, Zipursky SL. (1998). Polyglutamine-expanded human huntingtin transgenes induce degeneration of *Drosophila* photoreceptor neurons. *Neuron*. Sep;21(3):633-42.

Jaillon O, Aury JM, Brunet F, et al. (2004). Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature*. Oct 21;431(7011):946-57.

Jansen G, Bachner D, Coerwinkel M, Wormskamp N, Hameister H, Wieringa B. (1995).

Structural organization and developmental expression pattern of the mouse *WD-repeat gene DMR-N9* immediately upstream of the myotonic dystrophy locus. *Hum Mol Genet.* May;4(5):843-52.

Jansen G, Groenen PJ, Bachner D, Jap PH, Coerwinkel M, Oerlemans F, van den Broek W, Gohlsch B, Pette D, Plomp JJ, Molenaar PC, Nederhoff MG, van Echteld CJ, Dekker M, Berns A, Hameister H, Wieringa B. (1996). Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet.* Jul;13(3):316-24.

Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. (2004). Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum Mol Genet.* Dec 15;13(24):3079-88.

Junghans RP, Ebralidze A, Tiwari B. (2001). Does (CUG)_n repeat in DMPK mRNA 'paint' chromosome 19 to suppress distant genes to create the diverse phenotype of myotonic dystrophy?: A new hypothesis of long-range cis autosomal inactivation. *Neurogenetics.* Mar;3(2):59-67.

- K -

Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, Timmers AM, Hauswirth WW, Swanson MS. (2003). A *muscleblind* knockout model for myotonic dystrophy. *Science.* Dec 12;302(5652):1978-80.

Kanadia RN, Urbinati CR, Crusselle VJ, Luo D, Lee YJ, Harrison JK, Oh SP, Swanson MS. (2003). Developmental expression of mouse *muscleblind* genes Mbnl1, Mbnl2 and Mbnl3. *Gene Expr Patterns.* Aug;3(4):459-62.

Kania A, Salzberg A, Bhat M, D'Evelyn D, He Y, Kiss I, Bellen HJ. (1995). P-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. *Genetics.* Apr;139(4):1663-78.

Kim C.A. , Bowie J.U. (2003). SAM domains: uniform structure, diversity of function. *Trends Biochem Sci.* 28: 625-628

Kimura T, Takahashi MP, Okuda Y, Kaido M, Fujimura H, Yanagihara T, Sakoda S. (2000). The expression of ion channel mRNAs in skeletal muscles from patients with myotonic muscular dystrophy. *Neurosci Lett.* Dec 8;295(3):93-6.

Kino Y, Mori D, Oma Y, Takeshita Y, Sasagawa N, Ishiura S. (2004). Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum Mol Genet.* Mar 1;13(5):495-507

Klesert TR, Cho DH, Clark JI, Maylie J, Adelman J, Snider L, Yuen EC, Soriano P, Tapscott SJ. (2000). Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat Genet.* May;25(1):105-9.

Klesert TR, Otten AD, Bird TD, Tapscott SJ. (1997). Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat Genet.* Aug;16(4):402-6.

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LP. (1999). An untranslated CTG expansion causes a novel form of *spinocerebellar ataxia (SCA8)* *Nat Genet.* Apr;21(4):379-84.

Kreahling J, Graveley BR. (2004). The origins and implications of Aluternative splicing. *Trends Genet.* Jan;20(1):1-4.

Kress W, Mueller-Myhsok B, Ricker K, Schneider C, Koch MC, Toyka KV, Mueller CR, Grimm T. (2000). Proof of genetic heterogeneity in the proximal myotonic myopathy syndrome (PROMM) and its relationship to myotonic dystrophy type 2 (DM2). *Neuromuscul Disord.* Oct;10(7):478-80.

Krull M, Brosius J, Schmitz J. (2005). Alu-SINE Exonization: En Route to Protein Coding Function. *Mol Biol Evol.* May 18;

Kuan YS, Brewer-Jensen P, Searles LL. (2004). *Suppressor of sable*, a putative RNA-processing protein, functions at the level of transcription. *Mol Cell Biol.* May;24(9):3734-46.

Kuranaga E, Kanuka H, Igaki T, Sawamoto K, Ichijo H, Okano H, Miura M. (2002). Reaper-mediated inhibition of DIAP1-induced DTRAF1 degradation results in activation of JNK in *Drosophila*. *Nat Cell Biol.* Sep;4(9):705-10.

- L -

Laberge L, Begin P, Montplaisir J, Mathieu J. (2004). Sleep complaints in patients with myotonic dystrophy. *J Sleep Res.* Mar;13(1):95-100.

Landgraf M, Jeffrey V, Fujioka M, Jaynes JB, Bate M. (2003). Embryonic origins of a motor system: motor dendrites form a myotopic map in *Drosophila*. *PLoS Biol.* Nov;1(2):E41.

Lazarus A, Varin J, Ounnoughene Z, Radvanyi H, Junien C, Coste J, Laforet P, Eymard B, Becane HM, Weber S, Duboc D. (1999). Relationships among electrophysiological findings and clinical status, heart function, and extent of DNA mutation in myotonic dystrophy. *Circulation.* Mar 2;99(8):1041-6.

Le Parco Y, Le Bivic A, Knibiehler B, Mirre C y Cecchini JP. (1989). DCgl & IV collagen chain in *Drosophila melanogaster* is synthesized during embryonic organogenesis by mesenchymal cells and is deposited in muscle basement membranes. *Insect Biochem.* 19:789-802.

Lee CY, Clough EA, Yellon P, Teslovich TM, Stephan DA, Baehrecke EH. (2003) Genome-wide analyses of steroid- and radiation-triggered programmed cell death in *Drosophila*. *Curr Biol.* Feb 18;13(4):350-7

Lee WC, Yoshihara M, Littleton JT. (2004). Cytoplasmic aggregates trap polyglutamine-containing proteins and block axonal transport in a *Drosophila* model of Huntington's disease. *Proc Natl Acad Sci U S A.* Mar 2;101(9):3224-9.

Liquori CL, Ikeda Y, Weatherspoon M, Ricker K, Schoser BG, Dalton JC, Day JW, Ranum LP. (2003). Myotonic dystrophy type 2: human founder haplotype and evolutionary conservation of the repeat tract. *Am J Hum Genet.* Oct;73(4):849-62.

Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of *ZNF9*. *Science.* Aug 3;293(5531):864-7.

Lohmann I. (2003). Dissecting the regulation of the *Drosophila* cell death activator reaper. *Gene Expr Patterns*. May;3(2):159-63.

Lu X, Timchenko NA, Timchenko LT. (1999). Cardiac elav-type RNA-binding protein (ETR-3) binds to RNA CUG repeats expanded in myotonic dystrophy. *Hum Mol Genet*. Jan;8(1):53-60.

Ludatscher, R. M., Kerner, H., Amikam, S. and Gellei, B. (1978). Myotonia dystrophica with heart involvement: an electron microscopic study of skeletal, cardiac, and smooth muscle. *Journal of Clinical Pathology*. 31, 1057-1064.

Luque T, O'Reilly DR. (2002). Functional and phylogenetic analyses of a putative *Drosophila melanogaster* UDP-glycosyltransferase gene. *Insect Biochem Mol Biol*. Dec;32(12):1597-604.

- M -

Makalowski W, Mitchell GA, Labuda D. (1994). Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends Genet*. Jun;10(6):188-93

Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, Thornton CA. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*. Sep 8;289(5485):1769-73.

Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA. (2002). Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell*. Jul;10(1):35-44.

Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, Henderson D, Schalling M, Swanson MS, Thornton CA. (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet*. Sep 15;10(19):2165-70.

Marshall J, Molloy R, Moss GW, Howe JR, Hughes TE. (1995). The jellyfish green fluorescent protein: a new tool for studying ion channel expression and function. *Neuron*. Feb;14(2):211-5.

Martinez-Rodriguez JE, Lin L, Iranzo A, Genis D, Marti MJ, Santamaria J, Mignot E. (2003). Decreased hypocretin-1 (Orexin-A) levels in the cerebrospinal fluid of patients with myotonic dystrophy and excessive daytime sleepiness. *Sleep*. May 1;26(3):287-90.

Martorell L, Monckton DG, Gamez J, Johnson KJ, Gich I, Lopez de Munain A, Baiget M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum Mol Genet*. Feb;7(2):307-12.

McCarthy JV. (2003). Apoptosis and development. *Essays Biochem*. ;39:11-24

Medlej-Hashim M, Delague V, Chouery E, Salem N, Rawashdeh M, Lefranc G, Loiselet J, Megarbane A. (2004). Amyloidosis in familial Mediterranean fever patients: correlation with MEFV genotype and SAA1 and MICA polymorphisms effects. *BMC Med Genet*. Feb 10;5(1):4.

Michiels C. (2003). Endothelial cell functions. *J Cell Physiol*. Sep;196(3):430-43

Mihovilovic M, Mai Y, Herbstreith M, Rubboli F, Tarroni P, Clementi F, Roses AD. (1993). Splicing of an anti-sense Alu sequence generates a coding sequence variant for the alpha-3 subunit of a neuronal acetylcholine receptor. *Biochem Biophys Res Commun.* Nov 30;197(1):137-44.

Miller JW, Urbinati CR, Teng-Umuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS. (2000). Recruitment of human *muscleblind* proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* Sep 1;19(17):4439-48.

Miller JW, Urbinati CR, Teng-Umuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS. (2000). Recruitment of human *muscleblind* proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* Sep 1;19(17):4439-48.

Mohr SE, Dillon ST, Boswell RE. (2001). The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during *Drosophila* oogenesis. *Genes Dev.* Nov 1;15(21):2886-99.

Monckton DG, Wong LJ, Ashizawa T, Caskey CT. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum Mol Genet.* Jan;4(1):1-8.

Monferrer L, Artero R. (2005). An interspecific functional complementation test in *Drosophila* for introductory genetics laboratory courses. *J Hered.* (en prensa).

Monferrer L, Vicente M, Garcia S, Pérez-Alonso M, Artero R. (2004). A *Drosophila* model of human diseases caused by CTG microsatellite expansions. 4th International conference of unstable microsatellites and human disease. Banff, Alberta (Canada).

Monferrer L, Vicente M, Pascual M, Miranda M, Houseley J, Belacortu Y, Poulos M, O'Dell K, Monckton D, Swanson M, Artero R. (2005). 46th Annual *Drosophila* Research Conference. March 30-April 3. San Diego, California (EEUU).

Moseley ML et al. (2002). SCA8 BAC transgenic mice have a progressive and lethal neurological phenotype demonstrating pathogenicity of the CTG expansion. *Am J Hum Genet.* 71, A176.

Moxley RT, Griggs RC, Goldblatt D. (1980). Muscle insulin resistance in myotonic dystrophy: effect of supraphysiologic insulinization. *Neurology.* Oct;30(10):1077-83.

Mullersman JE, Pfeffer LM. (1995). An Alu cassette in the cytoplasmic domain of an interferon receptor subunit. *J Interferon Cytokine Res.* Sep;15(9):815-7.

Murugasu-Oei B, Rodrigues V, Yang X, Chia W. (1995). Masquerade: a novel secreted serine protease-like molecule is required for somatic muscle attachment in the *Drosophila* embryo. *Genes Dev.* Jan 15;9(2):139-54.

Mutsuddi M, Marshall CM, Benzow KA, Koob MD, Rebay I. (2004). The *spinocerebellar ataxia 8* noncoding RNA causes neurodegeneration and associates with *staufen* in *Drosophila*. *Curr Biol.* Feb 17;14(4):302-8.

- N -

Nemes JP, Benzow KA, Moseley ML, Ranum LP, Koob MD. (2000). The SCA8

transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). *Hum Mol Genet.* Jun 12;9(10):1543-51.

Neuwald AF, Landsman D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem Sci.* May;22(5):154-5.

Novelli G, Gennarelli M, Zelano G, Pizzuti A, Fattorini C, Caskey CT, Dallapiccola B. (1993). Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. *Biochem Mol Biol Int.* Feb;29(2):291-7.

- O -

Ohno S. (1999). Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999. *Semin Cell Dev Biol.* Oct;10(5):517-22

Otten AD, Tapscott SJ. (1995). Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc Natl Acad Sci U S A.* Jun 6;92(12):5465-9.

- P -

Peterson C, Carney GE, Taylor BJ, White K. (2002). Reaper is required for neuroblast apoptosis during *Drosophila* development. *Development.* Mar;129(6):1467-76.

Personius KE, Nautiyal J, Reddy S. (2005). Myotonia and muscle contractile properties in mice with *SIX5* deficiency. *Muscle Nerve.* Apr;31(4):503-5.

Philips AV, Timchenko LT, Cooper TA. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science.* May 1;280(5364):737-41.

Phillips MF, Harper PS. (1997). Cardiac disease in myotonic dystrophy. *Cardiovasc Res.* Jan;33(1):13-22.

Prokopenko SN y Bellen HJ. (1998). Comunicación personal a Flybase FBrf0102071.

Prokopenko SN, He Y, Lu Y, Bellen HJ. (2000). Mutations affecting the development of the peripheral nervous system in *Drosophila*: a molecular screen for novel proteins. *Genetics.* Dec;156(4):1691-715.

Proudfoot NJ, Furger A, Dye MJ. (2002). Integrating mRNA processing with transcription. *Cell.* Feb 22;108(4):501-12.

- R -

Read RD, Goodfellow PJ, Mardis ER, Novak N, Armstrong JR, Cagan RL. (2005). A *Drosophila* Model of Multiple Endocrine Neoplasia type 2. *Genetics.* Jun 18;

Reddy S, Smith DB, Rich MM, Leferovich JM, Reilly P, Davis BM, Tran K, Rayburn H, Bronson R, Cros D, Balice-Gordon RJ, Housman D. (1996). Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet.* Jul;13(3):325-35.

Rendahl KG, Gaukhshteyn N, Wheeler DA, Fry TA, Hall JC. (1996). Defects in courtship and vision caused by amino acid substitutions in a putative RNA-binding protein encoded by the no-on-transient A (nonA) gene of *Drosophila*. *J Neurosci.* Feb

15;16(4):1511-22.

Rubin GM y otros. (2000). Comparative genomics of the eukaryotes. *Science*. Mar 24;287(5461):2204-15.

- S -

Salzberg A, D'Evelyn D, Schulze KL, Lee JK, Strumpf D, Tsai L, Bellen HJ. (1994). Mutations affecting the pattern of the PNS in *Drosophila* reveal novel aspects of neuronal development. *Neuron*. Aug;13(2):269-87.

Salvatori S, Fanin M, Trevisan CP, Furlan S, Reddy S, Nagy JI, Angelini C. (2005) Decreased expression of *DMPK*: correlation with CTG repeat expansion and fibre type composition in myotonic dystrophy type 1. *Neurol Sci*. Oct;26(4):235-42

Sarkar PS, Appukuttan B, Han J, Ito Y, Ai C, Tsai W, Chai Y, Stout JT, Reddy S. (2000). Heterozygous loss of *Six5* in mice is sufficient to cause ocular cataracts. *Nat Genet*. May;25(1):110-4.

Sarkar PS, Paul S, Han J, Reddy S. (2004) *Six5* is required for spermatogenic cell survival and spermiogenesis. *Hum Mol Genet*. Jul 15;13(14):1421-31..

Savkur RS, Philips AV, Cooper TA, Dalton JC, Moseley ML, Ranum LP, Day JW. (2004). Insulin receptor splicing alteration in myotonic dystrophy type 2. *Am J Hum Genet*. Jun;74(6):1309-13. Epub Apr 26.

Savkur RS, Philips AV, Cooper TA. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet*. Sep;29(1):40-7.

Scherzer CR, Jensen RV, Gullans SR, Feany MB. (2003). Gene expression changes presage neurodegeneration in a *Drosophila* model of Parkinson's disease. *Hum Mol Genet*. Oct 1;12(19):2457-66.

Schoaser BG, Kress W, Walter MC, Halliger-Keller B, Lochmuller H, Ricker K. (2004). Homozygosity for CCTG mutation in myotonic dystrophy type 2. *Brain*. Aug;127(Pt 8):1868-77.

Schoaser BG, Ricker K, Schneider-Gold C, Hengstenberg C, Durre J, Bultmann B, Kress W, Day JW, Ranum LP. (2004). Sudden cardiac death in myotonic dystrophy type 2. *Neurology*. Dec 28;63(12):2402-4.

Schupbach T, Wieschaus E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics*. Dec;129(4):1119-36.

Seino S, Bell GI. (1989). Alternative splicing of human insulin receptor messenger RNA. *Biochem Biophys Res Commun*. Feb 28;159(1):312-6.

SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S, Wickens M. (1996). A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci U S A*. Aug 6;93(16):8496-501.

Sergeant N, Sablonniere B, Schraen-Maschke S, Ghestem A, Maurage CA, Watzet A, Vermersch P, Delacourte A. (2001). Dysregulation of human brain microtubule-

associated *tau* mRNA maturation in myotonic dystrophy type 1. *Hum Mol Genet.* Sep 15;10(19):2143-55.

Seznec H, Agbulut O, Sergeant N, Savouret C, Ghestem A, Tabti N, Willer JC, Ourth L, Duros C, Brisson E, Fouquet C, Butler-Browne G, Delacourte A, Junien C, Gourdon G. (2001). Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum Mol Genet.* Nov 1;10(23):2717-26.

Shaw DJ, McCurrach M, Rundle SA, Harley HG, Crow SR, Sohn R, Thirion JP, Hamshere MG, Buckler AJ, Harper PS, et al. (1993). Genomic organization and transcriptional units at the myotonic dystrophy locus. *Genomics.* Dec;18(3):673-9.

Sorek R, Ast G, Graur D. (2002). Alu-containing exons are alternatively spliced. *Genome Res.* Jul;12(7):1060-7.

Squillace RM, Chenault DM, Wang EH. (2002). Inhibition of muscle differentiation by the novel *muscleblind*-related protein CHCR. *Dev Biol.* Oct 1;250(1):218-30.

Strodicke M, Karberg S, Korge G. (2000). *Domina (Dom)*, a new *Drosophila* member of the FKH/WH gene family, affects morphogenesis and is a suppressor of position-effect variegation. *Mech Dev.* Aug;96(1):67-78.

Sun C, Henriksen OA, Tranebjaerg L. (1999). Proximal myotonic myopathy: clinical and molecular investigation of a Norwegian family with PROMM. *Clin Genet.* Dec;56(6):457-61.

Svineng G, Fassler R, Johansson S. (1998). Identification of beta1C-2, a novel variant of the integrin beta1 subunit generated by utilization of an alternative splice acceptor site in exon C. *Biochem J.* Mar 15;330 (Pt 3):1255-63.

- T -

Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol.* Mar;128(6):995-1002.

Tapscott SJ. (2000). Deconstructing myotonic dystrophy. *Science.* Sep 8;289(5485):1701-2.

Tenenhaus C, Subramaniam K, Dunn MA, Seydoux G. (2001) PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of *Caenorhabditis elegans*. *Genes Dev.* Apr 15;15(8):1031-40.

Thornton CA, Wymer JP, Simmons Z, McClain C, Moxley RT 3rd. (1997). Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking *DMAHP* gene. *Nat Genet.* Aug;16(4):407-9.

Thorpy M. (2003). Excessive daytime somnolence and increased rapid eye movement pressure in myotonic dystrophy. *Curr Neurol Neurosci Rep.* Mar;3(2):165-6.

Timchenko LT, Miller JW, Timchenko NA, DeVore DR, Datar KV, Lin L, Roberts R, Caskey CT, Swanson MS. (1996). Identification of a (CUG)_n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.* Nov

15;24(22):4407-14.

Timchenko LT, Timchenko NA, Caskey CT, Roberts R. (1996). Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum Mol Genet.* Jan;5(1):115-21.

Timchenko LT, Timchenko NA, Caskey CT, Roberts R. (1996). Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum Mol Genet.* Jan;5(1):115-21.

Timchenko NA, Patel R, Iakova P, Cai ZJ, Quan L, Timchenko LT. (2004). Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. *J Biol Chem.* Mar 26;279(13):13129-39.

Toba G, Ohsako T, Miyata N, Ohtsuka T, Seong KH, Aigaki T. (1999). The gene search system. A method for efficient detection and rapid molecular identification of genes in *Drosophila melanogaster*. *Genetics.* Feb;151(2):725-37.

Török T, Tick G, Alvarado M, Kiss I. (1993). P-lacW insertional mutagenesis on the second chromosome of *Drosophila melanogaster*: isolation of lethals with different overgrowth phenotypes. *Genetics.* Sep;135(1):71-80.

- U -

Umemiya T, Takeichi M, Nose A. (1997). M-spondin, a novel ECM protein highly homologous to vertebrate *F-spondin*, is localized at the muscle attachment sites in the *Drosophila* embryo. *Dev Biol.* Jun 15;186(2):165-76.

- V -

Vervoort R, Gitzelmann R, Lissens W, Liebaers I. (1998). A mutation (IVS8+0.6kbpdelTC) creating a new donor splice site activates a cryptic exon in an Alu-element in intron 8 of the human *beta-glucuronidase* gene. *Hum Genet.* Dec;103(6):686-93.

Vicente M, Pascual M, Garcia A, Artero R. (2005). Muscleblind protein shows differential RNA binding properties in a yeast three hybrid assay. *5th International Myotonic Dystrophy Consortium Meeting.* Québec, Canada

Voelker RA, Gibson W, Graves JP, Sterling JF, Eisenberg MT. (1991). The *Drosophila* suppressor of sable gene encodes a polypeptide with regions similar to those of RNA-binding proteins. *Mol Cell Biol.* Feb;11(2):894-905.

- W -

Warrick JM, Paulson HL, Gray-Board GL, Bui QT, Fischbeck KH, Pittman RN, Bonini NM. (1998). Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*. *Cell.* Jun 12;93(6):939-49.

Webster PJ, Liang L, Berg CA, Lasko P, Macdonald PM. (1997). Translational repressor *bruno* plays multiple roles in development and is widely conserved. *Genes Dev.* Oct 1;11(19):2510-21.

Westerlaken JH, Van der Zee CE, Peters W, Wieringa B. (1993). The DMWD protein

from the myotonic dystrophy (DM1) gene region is developmentally regulated and is present most prominently in synapse-dense brain areas. *Brain Res.* May 2;971(1):116-27.

Wieser T, Bonsch D, Eger K, Schulte-Mattler W, Zierz S. (2000). A family with PROMM not linked to the recently mapped PROMM locus DM2. *Neuromuscul Disord.* Feb;10(2):141-3.

Wilhelm JE, Mansfield J, Hom-Booher N, Wang S, Turck CW, Hazelrigg T, Vale RD. (2000). Isolation of a ribonucleoprotein complex involved in mRNA localization in *Drosophila* oocytes. *J Cell Biol.* Feb 7;148(3):427-40.

Wilkie GS, Dickson KS, Gray NK. (2003). Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem Sci.* Apr;28(4):182-8.

Wing JP, Schreader BA, Yokokura T, Wang Y, Andrews PS, Huseinovic N, Dong CK, Ogdahl JL, Schwartz LM, White K, Nambu JR. (2002). *Drosophila* Morgue is an F box/ubiquitin conjugase domain protein important for grim-reaper mediated apoptosis. *Nat Cell Biol.* Jun;4(6):451-6.

Wing JP, Schwartz LM, Nambu JR. (2001). The RHG motifs of *Drosophila* Reaper and Grim are important for their distinct cell death-inducing abilities. *Mech Dev.* Apr;102(1-2):193-203.

- X Y Z -

Xu B, Koenig RJ. (2004). An RNA-binding domain in the thyroid hormone receptor enhances transcriptional activation. *J Biol Chem.* Aug 6;279(32):33051-6.

Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR. (1992). Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc Natl Acad Sci U S A.* Sep 15;89(18):8769-73.

Zur Lage PI, Prentice DR, Holohan EE, Jarman AP. (2003). The *Drosophila* proneural gene amos promotes olfactory sensillum formation and suppresses bristle formation. *Development.* Oct;130(19):4683-93.