

**ESTUDIO DE LA INESTABILIDAD
DE REPETICIONES DE
TRINUCLEÓTIDOS ASOCIADAS A
ENFERMEDADES GENÉTICAS
HUMANAS**

**TESIS DOCTORAL
Laura Fernández López
2004**



Universitat Autònoma de Barcelona

Departament de Genètica i de Microbiologia



**ESTUDIO DE LA INESTABILIDAD DE
REPETICIONES DE TRINUCLEÓTIPOS
ASOCIADAS A ENFERMEDADES
GENÉTICAS HUMANAS**

TESIS DOCTORAL

Laura Fernàndez López

2004

ESTUDIO DE LA INESTABILIDAD DE REPETICIONES DE TRINUCLEÓTIDOS ASOCIADAS A ENFERMEDADES GENÉTICAS HUMANAS

Memoria presentada para optar al grado
de Doctora en Ciencias Biológicas por
Laura Fernàndez López

Bellaterra (Cerdanyola del Vallès),
Abril 2004

Laura Fernàndez López

VºBº

Los Directores de la Tesis

Dr. Jordi Surrallés i Calonge

Dra. Antonia Velázquez Henar

A Eli, perquè esto es tan tuyo como mío

A MAC, encara que no l'hagi pres, gràcies per oferir-me el món.

Els anys han anat passant i finalment he arribat a allò que en un principi semblava una fita llunyana i força inaccessible. Encara em recordo dient: una tesis, jo? Apa home, què dius!!

Arribar fins aquí no ha estat gens fàcil, i sou tants a qui he d'agrair l'ajuda rebuda...

Per començar, voldria agrair al Dr. Jordi Surrallés i a la Dra. Antonia Velázquez haver-me donat l'oportunitat de realitzar aquesta tesis, permetent-me així endinsar-me en el món de la investigació. També, voldria agrair-los tot l'esforç que han dedicat tant a mi com a aquest treball.

Voldria agrair també a la resta dels professors del departament l'ajuda que m'han proporcionat sempre que ho he necessitat.

Gràcies a tots el companys que m'han ajudat en totes les qüestions de laboratori, en especial al "mestre" de tots els temes relacionats amb gels i radioactivitat.

Gràcies als companys i amics MUTAS, tots a la vostra manera heu fet que el llarg camí fos més planer: per estar sempre disposats a donar un cop de mà i a fer de mestres en les tècniques de laboratori, per tenir tanta paciència, per animar-me en els moments baixos, fent-me riure, deixant-me fer el pallaso, "cotillejant", per aconsellar-me, per escoltar la meva xerrera i les meves queixes, per amenitzar els àpats, per les converses cafeteres a les escales, ... Gràcies per tantes bones estones.

A les noves generacions (a les que incloc dins dels companys i amics MUTAS) ,gràcies per tota l'energia i optimisme que heu aportat, tan necessari en alguns moments.

Una menció especial a les meves quatre nenes, perquè sense vosaltres això no hagués sigut el mateix, gràcies per la vostra amistat.

A la família, gràcies per aguantar-me, i a tots els amics que han compartit aquesta etapa de la meva vida, gràcies per ser al meu costat i ser com sou.

I sobretot, moltíssimes gràcies Eli, per aguantar-me tot aquest temps, i per compartir tantes i tantes coses...

A més, em costa tant imaginar-me un demà sense vosaltres...

MOLTES GRÀCIES.

“Y aunque a veces me asuste volar, lejos del cobijo de un hogar,
se que habrá un viento cálido más para dejarme llevar.” Presuntos implicados
Ícaro

“Seguiré caminando, seguiré esperando...” Amparanoia La fiesta

“Comenzar es difícil
pero vamos dando los pasos
por un futuro, que los hijos puedan celebrar
somos el viento, que baila que canta
si estamos juntos, somos huracán
No estamos de paso
Nos somos fracaso” Amparanoia Somos viento

“No por caer voy a dejar de caminar” Amparanoia Tu ausencia

“Lo pequeño sabe crecer
y se hace grande sin querer
tu llévame a un lugar
donde pueda olvidar el dolor
tu llévame a un lugar
donde vea salir el sol” Amparanoia En algún lugar

“Veo al fin que hay una razón un porqué para continuar,
para levantarme y caminar,
para confiarme y regresar, un motivo para amar (o dar)” Presuntos implicados
Mi única razón

“Y aunque a veces me asuste volar, lejos del cobijo de un hogar,
se que habrá un viento cálido más para dejarme llevar.” Presuntos implicados
Ícaro

ÍNDICE

I. INTRODUCCIÓN.....	1
1. Microsatélites.....	2
1.1. Distribución genómica.....	2
1.2. Funciones.....	3
2. Inestabilidad de trinucleótidos.....	5
2.1. Enfermedades humanas asociadas a repeticiones de trinucleótidos.....	6
2.1.1. Distrofia Miotónica.....	9
2.1.1.1. Características fenotípicas.....	9
2.1.1.2. Base molecular.....	11
2.1.1.3. Efecto patológico de la expansión.....	12
2.1.2. Síndrome del X frágil.....	13
2.1.2.1. Características fenotípicas.....	14
2.1.2.2. Base molecular.....	15
2.1.2.3. Efecto patológico de la expansión.....	17
2.2. Generación de los alelos patogénicos con ERT.....	17
2.2.1. Origen de los alelos DM.....	18
2.2.2. Origen de los alelos FXS.....	19
3. Mecanismos de inestabilidad de las TNR.....	20
3.1. Factores que actúan en cis.....	22
3.2. Metabolismo del DNA.....	25
3.2.1. Replicación del DNA.....	25
3.2.1.1. Modelo clásico de deslizamiento entre las cadenas de DNA.....	25
3.2.1.2. Procesamiento aberrante de los fragmentos de Okazaki.....	27
3.2.1.3. Bloqueo de la horquilla de replicación.....	29
3.2.2. Transcripción.....	29
3.2.3. Reparación.....	30

3.2.3.1. Reparación de apareamientos erróneos.....	30
3.2.3.2. Reparación de las roturas de doble cadena.....	33
3.2.3.3. Reparación por escisión de nucleótidos.....	35
3.2.4. Recombinación.....	36
3.3. Mutagénesis.....	37
3.3.1. MMC.....	38
4. Detección de la inestabilidad de trinucleótidos.....	39
II.OBJETIVOS.....	41
III.ARTÍCULOS.....	43
1. Artículo 1:	
<u>Fernàndez-López L.</u> , Piñeiro E., Marcos R., Velázquez A. and Surrallés J. Induction of instability of normal length trinucleotide repeats within human disease genes. J. Med. Genet., 41:e3, 2004.....	44
2. Artículo 2:	
Piñeiro E., <u>Fernàndez-López L.</u> , Gámez J., Marcos R., Surrallés J. and Velásquez A. Mutagenic stress modulates the dynamics of CTG repeats instability associated with myotonic dystrophy type 1. Nucleic Acids Res., 31:6733-6740, 2003.....	51
IV.DISCUSIÓN.....	61
V.CONCLUSIONES.....	73
VI.BIBLIOGRAFÍA.....	75

I. INTRODUCCIÓN



I. INTRODUCCIÓN

1 Microsatélites

Las secuencias de DNA repetidas en tándem se encuentran ampliamente distribuidas en los organismos eucariotas. Todos los genomas eucarióticos estudiados hasta el momento presentan estas secuencias, que pueden llegar a constituir hasta un 10% del genoma humano. Los microsatélites forman una clase de secuencias repetitivas, caracterizadas por repeticiones directas en tándem de un motivo corto de DNA. Históricamente, el término microsatélite se aplicaba sólo a las repeticiones del dinucleótido CA (Litt y Luty, 1989; Weber y May, 1989), pero, a pesar de que la definición de este término sigue siendo arbitraria, su aplicación se ha extendido a repeticiones directas en tándem de motivos cortos de DNA de hasta 6 bp.

La tasa de mutación del genoma humano es muy variable entre las distintas secuencias que lo forman. Las secuencias génicas únicas varían muy lentamente, con unas tasas de mutación puntual del orden de 10^{-9} a 10^{-10} . Las secuencias repetidas de DNA, en cambio, son muy inestables con unas tasas de mutación de alrededor de 10^{-2} eventos por locus por replicación en *E.coli* (Levinson y Gutman, 1987) y alrededor de 10^{-4} a 10^{-5} en levaduras (Henderson y Petes, 1992; Strand *et al.*, 1993). Estimaciones realizadas a partir de análisis de pedigríes en humanos sugieren tasas de alrededor de 10^{-3} eventos por locus por generación (Weber y Wong, 1993), mientras estudios con ratones dan unas estimaciones de entre 10^{-3} y 10^{-4} (Dallas, 1992). En *Drosophila* las tasas de mutación de microsatélites parecen ser relativamente bajas, alrededor de 6×10^{-6} (Schug *et al.*, 1997) (revisado en Goldstein y Schlötterer, 1999).

1.1 Distribución genómica

Las repeticiones de mononucleótidos más comunes en el genoma humano son las secuencias poli(A) (Beckmann y Weber, 1992; Stallings, 1992; Hancock, 1996). Dentro de los dinucleótidos, Beckmann y Weber (1992) mostraron que los más abundantes son las repeticiones CA, con una frecuencia dos veces mayor que las repeticiones AT, y tres veces más que las AG. Entre las repeticiones de trinucleótidos (TNR, *trinucleotide repeats*), las más comunes son las repeticiones

CAG y AAT (Stallings, 1992). Estas proporciones se mantienen con algunas pequeñas diferencias en el resto de mamíferos, sin embargo, en especies más divergentes, las proporciones varían. Así, por ejemplo, respecto a los dinucleótidos, las levaduras son más ricas en las repeticiones AT que el genoma humano (Valle, 1993; Hanckoc, 1995), mientras en plantas además de las secuencias AT también son muy frecuentes las repeticiones GA (Stallings, 1992; Lagercrantz *et al.*, 1993), y en *Drosophila melanogaster* las repeticiones CA son las más comunes. Los organismos procariotas contienen muy pocas secuencias repetidas en tándem, pero las más comunes son las poli(A) (Hanckoc, 1995; Field y Wills, 1996) (revisado en Goldstein y Schlötterer, 1999).

Existen numerosas evidencias que indican que la distribución genómica de los microsatélites no es al azar. Así, una parte de los loci microsatélites seleccionan las expansiones o contracciones al azar en contra, presumiblemente por su posible efecto en la organización de la cromatina, regulación de la actividad génica, recombinación, replicación del DNA, ciclo celular, reparación, etc. (revisado en Goldstein y Schlötterer, 1999).

Li y colaboradores (2002) revisaron la literatura que estudiaba la distribución de los microsatélites en el genoma de muchas especies, desde procariotas hasta humanos. En estos estudios se vio que todos los tipos de microsatélites (desde mono- hasta hexanucleótidos) se encuentran mayormente en regiones no codificantes del genoma, a excepción de los tri- y hexanucleótidos que pueden encontrarse tanto en regiones codificantes como en regiones no codificantes. Estas diferencias en la distribución de microsatélites entre regiones codificantes y no codificantes podrían ser debidas a una selección específica contra las mutaciones de cambio de pauta de lectura en las regiones codificantes, resultantes de cambios en el tamaño en repeticiones no formadas por trinucleótidos o hexanucleótidos (Metzgar *et al.*, 2000).

1.2 Funciones

Durante mucho tiempo se ha creído que la evolución de los microsatélites sigue el modelo de la teoría neutra, es decir, al azar y sin constricciones de ningún tipo, considerándose únicamente como marcadores de DNA evolutivamente neutros. En los últimos tiempos han aparecido algunos estudios que han puesto en entredicho estos modelos, creando cierta controversia en cuanto a la evolución de estas secuencias. Estos estudios han demostrado que en algunos loci microsatélites

sí existen constricciones de tamaño, y que estas secuencias pueden tener un significado funcional como elementos codificantes o reguladores.

Kashi y colaboradores (1997) sugirieron que los microsatélites, considerados como elementos funcionales del genoma y debido a su elevada tasa de mutación, podrían ser una fuente importante de variación genética cuantitativa y de adaptación evolutiva, capacitando a la población a reponer la variación genética perdida por la deriva o selección.

Parece entonces que las secuencias microsatélites sí tienen alguna función en los organismos superiores. Li y colaboradores (2002) realizaron una revisión de todos los estudios publicados sobre el tema, y agruparon las funciones o efectos de los microsatélites en tres áreas: organización de la cromatina, regulación de procesos metabólicos del DNA y regulación de la actividad génica (ver figura 1).

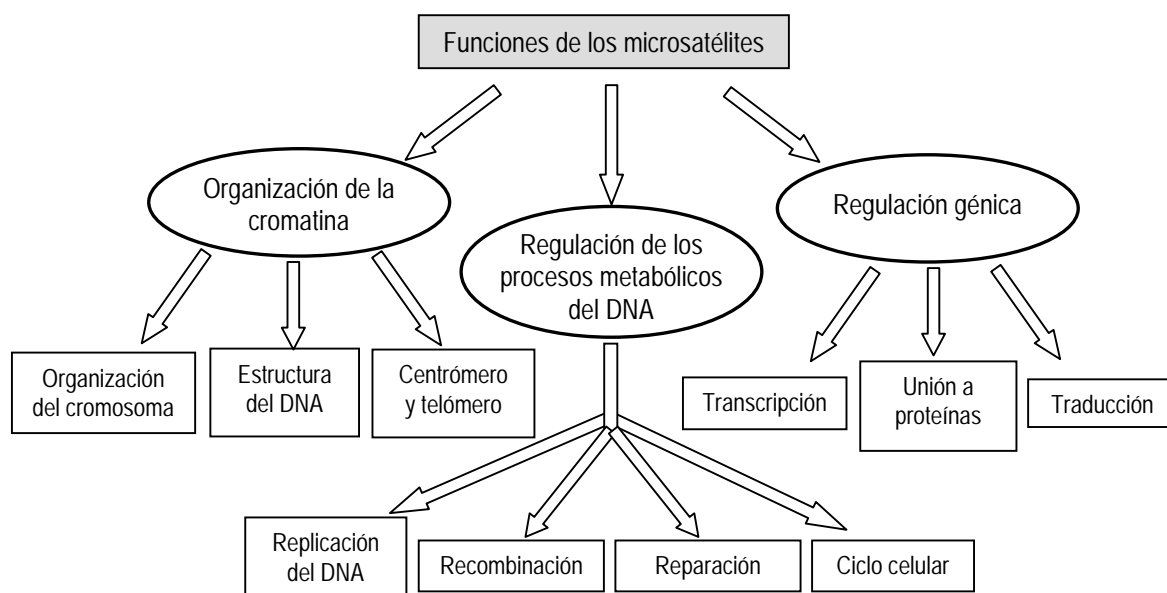


Figura 1. Clasificación de las funciones de los microsatélites (Adaptado de Li *et al.*, 2002)

En cuanto a la participación de los microsatélites en la organización de la cromatina y teniendo en cuenta la distribución de estas secuencias en el genoma, se sugiere que puedan tener un papel en la estructura del cromosoma. Algunas secuencias microsatélites se encuentran flanqueando los centrómeros, sugiriendo una función en la cohesión de las cromátidas hermanas y/o en la formación y función del cinetocoro.

Los microsatélites también están implicados en la regulación de los procesos metabólicos del DNA, tales como la replicación, la recombinación, el control del ciclo celular y el sistema de reparación de apareamientos erróneos (MMR, *mismatch*

repair). Está demostrado que los microsatélites pueden interferir en el proceso de replicación del DNA (Freudenreich *et al.*, 1997; Samadashwiliy *et al.*, 1997; Pelletier *et al.*, 2003). Jeffreys y colaboradores (1998) propusieron que los microsatélites y minisatélites podrían ser puntos calientes (*hot spots*) de recombinación. Así, se ha demostrado que el tipo de secuencia repetida y el número de repeticiones pueden afectar a la recombinación. Los microsatélites también pueden afectar a los enzimas responsables del control del ciclo celular, ya que algunos de estos genes, como *hMSH3*, *hMSH6*, *BAX*, *IGFIIR*, *TGFbetaIIIR*, *E2F4* y *BRCA2*, contienen pequeñas secuencias repetidas. Además, algunos de los genes del sistema MMR también contienen secuencias poli(A) en sus regiones codificantes, cuya alteración puede afectar a la función de la proteína que codifican. Dado que las proteínas del sistema de reparación MMR corrigen errores de la replicación e inhiben activamente la recombinación entre secuencias divergentes (Chen y Jinks-Robertson, 1998; Kolodner y Marsischky, 1999), Chang y colaboradores (2001) sugirieron que la elevada densidad de microsatélites en genes del sistema MMR permite modular la tasa de mutación adaptativa a lo largo del tiempo evolutivo.

El efecto de los microsatélites en la regulación génica es el mejor documentado, ya sea a través de su efecto en la transcripción, la traducción o la unión a proteínas. Hay numerosas evidencias de que los microsatélites localizados en las regiones promotoras y los localizados en intrones podrían afectar a la transcripción, actuando como elementos reguladores. También se conoce la importancia del número de repeticiones de los microsatélites en la regulación de la expresión génica. Algunos microsatélites localizados en regiones promotoras sirven como sitios de unión para una variedad de proteínas reguladoras (revisado en Csink y Henikoff, 1998), de modo que variaciones en el número de repeticiones de la secuencia podría afectar a la unión de estas proteínas.

2 Inestabilidad de trinucleótidos

El incremento en el número de repeticiones de secuencias de DNA es la base molecular de una lista creciente de enfermedades genéticas humanas. Estas secuencias son muy inestables y su tasa de mutación está relacionada con el número de repeticiones, así, la mutabilidad de la secuencia después de un cambio en el número de repeticiones difiere de la de su predecesor. Por este motivo, a este incremento de frecuencia de mutación se denomina **mutación dinámica** (Richards

y Sutherland, 1992), para diferenciarlo de la tasa de mutación constante que caracteriza a la mayoría de las secuencias del genoma.

La mutación dinámica no es exclusiva de los microsatélites ya que también se puede encontrar en minisatélites. Así, Richards y Sutherland en 1997 establecieron que las secuencias con mutaciones dinámicas pueden tener unidades repetidas de entre 3 y 33 bp. Se ha demostrado, por ejemplo, que una expansión masiva del número de repeticiones de un minisatélite rico en AT de 33 bp, es la base molecular de la fragilidad cromosómica FRA16B (Yu *et al.*, 1997). La mutación más común en el locus de la epilepsia EPM1 es una expansión de una repetición rica en GC de 12 bp (Virtaneva *et al.*, 1997; Lalioti *et al.*, 1997). En 1998 Hewett y colaboradores demostraron que la fragilidad cromosómica FRA10B está relacionada con la expansión de un minisatélite de 42 bp, rico en AT. A pesar de ello, las expansiones de repeticiones de trinucleótidos (ERT) son las más frecuentes dentro de los desórdenes genéticos relacionados con inestabilidad de secuencias repetidas, y se han convertido, por lo tanto, en objeto de particular interés en la comunidad científica en los últimos años.

2.1 Enfermedades humanas asociadas a repeticiones de trinucleótidos

En 1991 se descubrió la ERT como un nuevo mecanismo causante de enfermedad genética, describiéndose la expansión del trinucleótido $(CGG)_n$ como la base molecular del síndrome del X frágil (FXS, *Fragile X Syndrome*) (Fu *et al.*, 1991; Oberlé *et al.*, 1991; Verkerk *et al.*, 1991; Yu *et al.*, 1991). Desde entonces, se han asociado 14 enfermedades genéticas neurodegenerativas y 3 fragilidades cromosómicas con la expansión de los siguientes trinucleótidos: $(CTG)_n$, $(CGG)_n$ o $(GAA)_n$.

Las TNR asociadas a alguna patología tienen características comunes (Richards y Sutherland, 1997). La población normal presenta un grado elevado de polimorfismo en estas secuencias, que siempre mantienen un número de repeticiones bajo de generación en generación. A partir de un determinado número de repeticiones la secuencia se vuelve inestable, tanto en células somáticas como germinales, observándose un mosaicismo somático y una variación intergeneracional típicos de la patología. Además, todos estos desórdenes genéticos presentan el fenómeno de la anticipación, que consiste en un incremento de la severidad de los síntomas y una disminución de la edad de aparición de la enfermedad con el paso de las generaciones. Este fenómeno está relacionado con la

magnitud de la expansión, cuanto mayor sea el número de repeticiones más graves serán los síntomas y antes aparecerán. La magnitud de las expansiones también se ha relacionado con la inestabilidad de la secuencia, de modo que cuantas más repeticiones tenga la secuencia más inestable será. Así, las secuencias con un número de repeticiones que provoquen un fenotipo alterado son secuencias que sufren mutación dinámica. Todos estos desórdenes genéticos, no obstante, difieren en la secuencia de la unidad repetida y en la localización de la repetición respecto al gen afectado. Las repeticiones se pueden encontrar tanto en el extremo 5' del gen, dentro de la región codificante, dentro de un intrón o en el extremo 3' del gen (ver figura 2). Dada esta diferencia en la localización, el papel del trinucleótido en la etiología molecular de la enfermedad será distinto en los diferentes desórdenes. También dependiendo del desorden, la expansión es inestable preferentemente en la transmisión paterna o materna.

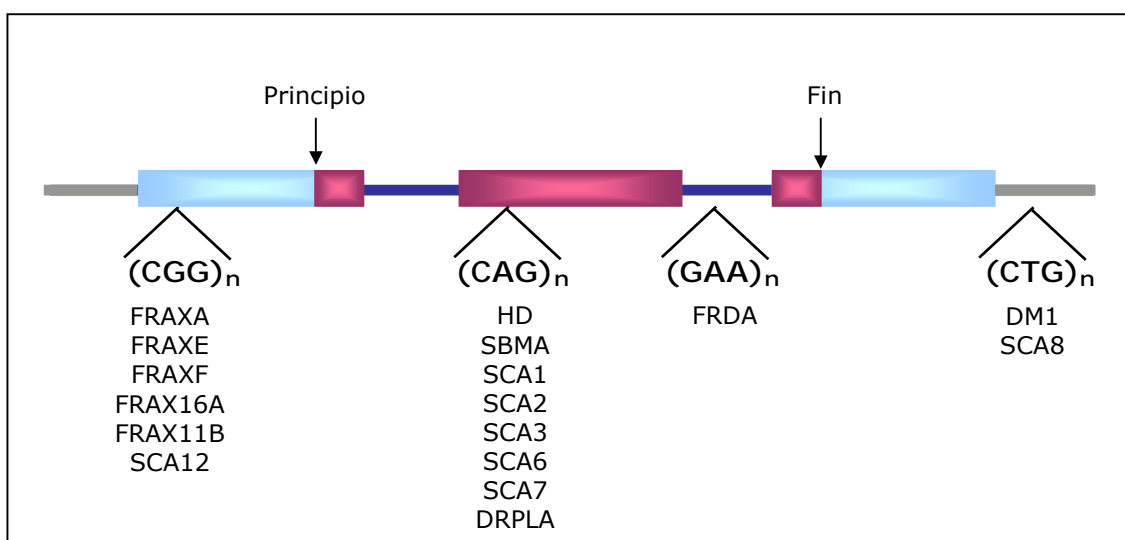


Figura 2. Esquema de la localización de las TNR en los diferentes desórdenes genéticos relacionados (Adaptado de Sinden *et al.*, 2002).

Desde el punto de vista patológico, los desórdenes producidos por TNR se dividen en dos grupos: Tipo I y Tipo II (Reddy y Housman, 1997). Las características más relevantes de los trinucleótidos asociados a estos tipos de patologías se muestran en la tabla 1. Los desórdenes de Tipo I se caracterizan por una pérdida neuronal progresiva, que depende de la edad de aparición y del tamaño de la repetición. El triplete repetido es CAG y se encuentra siempre en una zona codificante, dando lugar a un segmento poliglutamínico en la proteína. Debido a que el trinucleótido se encuentra en una zona codificante del gen, el número de

repeticiones en los individuos afectados es relativamente pequeño, en general entre 40-70 repeticiones. El incremento de tamaño del segmento poliglutamínico se asocia con un cambio significativo en las propiedades físicas de la proteína, provocando toxicidad en las neuronas. En este grupo se incluyen desórdenes como la enfermedad de Huntington, la atrofia muscular espinobulbar, la ataxia espinocerebelar y la enfermedad de Machado-Joseph.

Tabla 1. Desórdenes genéticos relacionados con ERT (Adaptado de Sinden, 1999 y Sinden *et al.*, 2002)

	Enfermedad	Locus		Número de repeticiones		
				TNR	Normal	Premutación
Tipo I	Síndrome de Kennedy (SMBA)	Xq13-21		9-36	?	38-62
	Enfermedad de Huntington (HD)	4p16.3		10-34	36-39	40-121
	Ataxia espinocerebelar 1 (SCA1)	6p23		6-39	...	40-82
	Ataxia espinocerebelar 2 (SCA2)	12q24.1	CAG	13-33	...	32-100
	Enfermedad de Machado Joseph (MJD) (SCA3)	14q32.1		13-44	?	55-84
	Ataxia espinocerebelar 6 (SCA6)	19p13		4-18	?	20-29
	Ataxia espinocerebelar 7 (SCA7)	13p12-13		4-35	?	37-306
	Síndrome de Haw-River (HRS) / DRPLA	12p13.31		7-25	?	49-75
Tipo II	Frágil XA (FRAXA)	Xq37.3		6-52	59-230	230-2000
	Frágil XF (FRAXF)	-	CGG	7-40	?	306-1008
	Síndrome de Jacobsen (FRA11B)	-		11	80	100-1000
	Frágil XE (FRAXE)	Xq28	CCG	4-39	(31-61)?	200-900
	FRA16A	-		16-49	?	1000-1900
	Ataxia espinocerebelar 8 (SCA8)	13q21	CTG	16-92	?	100-127
	Ataxia espinocerebelar 12 (SCA12)	5q31-33	CAG	7-45	?	55-78
	Distrofia miotónica (DM)	19q13	CTG	5-37	50-80	80-3000
	Ataxia de Friedreich (FRDA)	9q13-21.1	GAA	6-32	33-60	200-1700

Los desórdenes de Tipo II presentan una patología más multisistémica y más variada. La expansión de repeticiones se encuentra en regiones no codificantes, y por lo tanto, el intervalo de repeticiones es más amplio, desde 50 a miles de repeticiones. Dentro de este grupo encontramos el FXS, la ataxia de Friedreich y la distrofia miotónica tipo 1 (DM1). En el caso del FXS y de la DM1, el triplete repetido está localizado en regiones no traducidas, y en la ataxia de Friedreich se encuentra en un intrón. Por lo tanto, se considera que la relación entre expansión de las repeticiones y la patología en estos desórdenes reside en el efecto de la repetición

en la expresión del gen o en la función del mRNA. En los desórdenes asociados con las repeticiones (CGG)_n la metilación del DNA que contiene las repeticiones CGG, provocaría el silenciamiento del gen. En la DM1 y en las ataxias espinocerebelares 8 y 12, la localización de la expansión de repeticiones en la región 3' del gen puede llevar a un procesamiento aberrante del mRNA o a una expresión aberrante de los genes flanqueantes (ver tabla 1).

En esta tesis, vamos a centrarnos en la TNR asociadas a los desórdenes genéticos DM1 y FXS.

2.1.1 Distrofia miotónica

La DM1, o enfermedad de Steinert, fue descrita por primera vez en 1909 por Steinert, pero se cree que su existencia es muy antigua, remontándose incluso al antiguo Egipto. Se sospecha, a partir de las características físicas que se deducen de los dibujos y figuras, que Amenhotep (IV), más conocido como Akhenaton, uno de los últimos faraones de la decimoctava dinastía real y otros miembros de la dinastía, padecían esta enfermedad (Cattaino y Vicario, 1999), pudiendo ser ésta la causa del fin de la dinastía.

2.1.1.1 Características fenotípicas

La DM1 es la forma más común de distrofia muscular en el adulto, con una elevada incidencia en la población caucásica (1 de cada 8.000 individuos), baja frecuencia en orientales (1 de cada 18.000), y muy rara en africanos (Deka *et al.*, 1996). Es un desorden autosómico dominante, pero presenta un patrón de herencia no mendeliana debido al fenómeno de la anticipación.

La DM1 es una enfermedad multisistémica, que se caracteriza por distrofia muscular, miotonía y anomalías en varios sistemas. La distrofia muscular consiste en una degeneración muscular lenta y progresiva, con creciente debilidad y atrofia muscular. Los músculos más afectados son los más distales de las extremidades (antebrazos, manos, parte baja de las piernas y pies) y los músculos del cuello y la cara. La debilidad y pérdida de volumen de estos últimos confieren al paciente una apariencia característica: cara delgada y alargada con sienes hundidas y párpados caídos y calvicie frontal temprana en los hombres. La miotonía consiste en el retraso de la relajación muscular tras una contracción, ya sea voluntaria o provocada, que afecta tanto a la musculatura lisa como a la estriada. Normalmente

afecta a las manos y los dedos de tal forma que el paciente tiene dificultad de abrir la mano después de cerrarla. A pesar de que la miotonía es el síntoma más característico de la DM1, no suele representar un gran problema para el paciente que lo suele atribuir a una rigidez algo incómoda (revisado en Rittler *et al.*, 1997 y en Mason, 2000).

La miotonía y la debilidad muscular pueden afectar también a los músculos de órganos internos como el tracto digestivo y el útero, provocando complicaciones gastrointestinales y, en mujeres afectadas, problemas durante el embarazo e irregularidades menstruales. También puede afectar a los músculos respiratorios y de la garganta, alterando las funciones respiratorias y de deglución. Otro órgano que se ve afectado en la mayoría de los pacientes es el corazón provocando alteraciones del ritmo cardíaco o palpitaciones, siendo la segunda causa de muerte en pacientes DM1. La afectación más frecuente es la catarata, que aparece prácticamente en el 100% de los casos, y a veces es el único síntoma de la enfermedad. La mayoría de pacientes pueden desarrollar un tipo leve de diabetes llamada resistencia a la insulina con niveles elevados de azúcar en sangre (hiperglucemia insulinoresistente). Algunos pacientes presentan también algunas anormalidades del sistema nervioso central: retraso mental, personalidad evasiva característica e hipersomnolencia (revisado en Rittler *et al.*, 1997 y en Mason, 2000).

Una de las características más notables de este desorden es la gran variabilidad fenotípica que se observa entre individuos, incluso dentro de una misma familia. Según la sintomatología clínica y la edad de aparición de la enfermedad, los pacientes son divididos en tres grupos (figura 3). La forma intermedia, que a veces es difícil de diagnosticar, se observa en individuos de mediana o avanzada edad y se caracteriza por cataratas y, a veces, pequeños problemas musculares. La forma clásica se observa más frecuentemente en la adolescencia y principios de la edad adulta, presentándose los síntomas característicos de la enfermedad. La forma más severa de la enfermedad es la forma congénita, que está asociada con una hipoplasia muscular generalizada, retraso mental y elevada mortalidad neonatal.



Figura 3. Fotografía de tres generaciones de portadores del gen DM1, donde se observa la variación intergeneracional de la apariencia facial: (i) la abuela, a la izquierda, está ligeramente afectada y no presenta debilidad facial; (ii) su hija, a la derecha, presenta una cara estrecha y alargada característica de la DM1, con debilidad facial; (iii) el nieto, está severamente afectado, presentando la forma congénita de la enfermedad, con las características faciales típicas, como la boca abierta y de forma triangular (Shaw, 1995).

2.1.1.2 Base molecular

La DM1 es la forma clásica, más común de distrofia miotónica, pero en 1994 se describió una nueva forma de la enfermedad con los mismos síntomas, que tiene una base molecular distinta (Thornton *et al.*, 1994). Este nuevo locus (DM2) se encuentra en el cromosoma 3, y contiene un tetranucleótido CCUG, cuya expansión es característica de la DM2 (Thornton *et al.*, 1994).

La base molecular del desorden DM1 es una expansión de repeticiones del trinucleótido CTG en la región 3' no traducida del gen de la proteína kinasa DMPK, localizado en el cromosoma 19, concretamente en la región 19q13.3 (Brook *et al.*, 1992). Esta expansión se encuentra en una región rica en genes, por lo que es posible que provoque la disfunción de los genes próximos (ver figura 4).

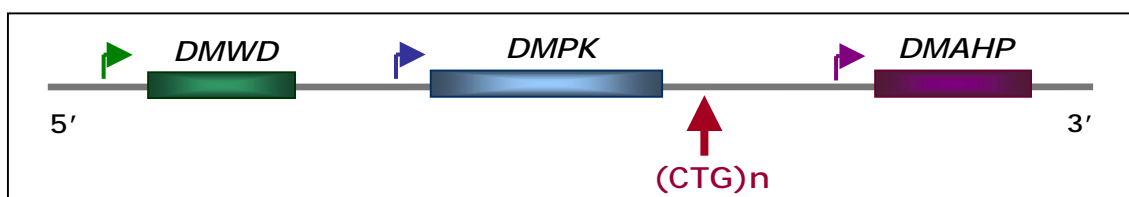


Figura 4. Representación esquemática del locus DM1. La ERT se localiza en la región 3' no traducida del gen DMPK, pero puede afectar a la expresión de los genes vecinos (adaptado de Groenen y Wieringa, 1998).

El intervalo de la expansión del trinucleótido CTG es muy amplio, entre 80 y miles de repeticiones, y muestra una correlación con la severidad y edad de aparición de la enfermedad (Harley *et al.*, 1992). En la población normal el número de repeticiones es polimórfico, y puede variar entre 5 y 40 repeticiones. Entre 50 y 80 repeticiones el individuo estaría medianamente afectado, pudiendo presentar sólo algunos síntomas leves de la enfermedad. A partir de unas 100 repeticiones, el individuo presenta el fenotipo clásico de la DM con todos los síntomas característicos de la enfermedad, observándose incluso la forma congénita al llegar a miles de repeticiones (Longshore y Tarleton, 1996). La longitud de las expansiones también se ha relacionado con la inestabilidad de estas secuencias tanto en células somáticas como germinales, de forma que a mayor número de repeticiones mayor inestabilidad (Monckton *et al.*, 1995; Martorell *et al.*, 1998). La inestabilidad meiótica se observa como una variación intergeneracional del número de repeticiones, y es la explicación molecular del fenómeno de la anticipación que se observa en las familias afectadas. La inestabilidad mitótica se refleja con un elevado mosaicismo somático y una heterogeneidad en el número de repeticiones entre y dentro de los tejidos del mismo individuo, que irá incrementando a lo largo de la vida del individuo. La secuencia presenta una tendencia a expandirse, característica que se puede asociar con la naturaleza progresiva de la enfermedad (Martorell *et al.*, 1998). Existe un umbral, alrededor de las 50 repeticiones, a partir del cual la secuencia se vuelve inestable. Así, el intervalo de 50 a 80 repeticiones se considera de premutación porque la secuencia es ya bastante inestable y puede sufrir grandes expansiones, pasando al intervalo considerado de mutación patológica.

2.1.1.3 Efecto patológico de la expansión

Los desórdenes genéticos dominantes son causados normalmente por la expresión de un gen alterado que produce una proteína con una función alterada. En el caso de la DM1 esto no se cumple, pues la expansión de las repeticiones CTG se halla en una región que no se traduce. La afectación multisistémica de la enfermedad dificulta conocer exactamente cuál es el efecto patológico de la expansión. La expansión CTG asociada a la DM1 se encuentra en el extremo 3' no-traducido del gen *DMPK*, de modo que el primer candidato responsable de la patología de la DM1 fue la reducción de los niveles de DMPK como consecuencia directa de la expansión del trinucleótido CTG. Los modelos de ratón *knockout* para el gen *DMPK* no muestran el fenotipo completo de la DM1 (revisado en Wansink y

Wieringa, 2003). Por lo tanto, se ha sugerido que la expansión puede afectar también a otros genes vecinos, debido a que las repeticiones CTG aumentan el ensamblaje de los nucleosomas, provocando una reestructuración de la cromatina (Otten y Tapscott, 1995) y afectando, así, a la accesibilidad y a la formación de complejos transcripcionales. De esta forma, la expresión de los genes cercanos como el *DMAHP (SIX5)* y el *DMWD*, que parece que pueden estar involucrados en la patología de la DM1, se verían afectados (Klessert *et al.*, 1997; Thornton *et al.*, 1997; Alwazzan *et al.*, 1999). Otro mecanismo de implicación del trinucleótido CTG en la patogénesis de la DM1, es a nivel del mRNA. Los mRNAs de *DMPK* con expansiones CUG, se acumulan como foci nucleares, y se propone que de esta forma se altera la regulación de proteínas de unión a CUG, alterando la función normal del mRNA (Revisado en Larkin y Fardaei, 2001 y en Ranum y Day, 2002a).

Todos los mecanismos propuestos podrían contribuir a la patogénesis de la DM1, y en este caso se trataría de un desorden génico regional. Sin embargo, la descripción de un nuevo locus relacionado con una segunda forma de distrofia miotónica, la DM2, localizado en otro cromosoma, era inconsistente con el modelo aditivo de patogénesis de la DM1 propuesto. Así, aunque la expansión de un tetranucleótido en el locus *DM2* afecte la regulación de varios genes en la región, es improbable que alteraciones en la regulación de conjuntos de proteínas diferentes en los dos loci resulte en enfermedades con características clínicas multisistémicas tan parecidas. Además, en el caso de la DM2 también se han encontrado foci nucleares de mRNA con expansiones CCUG, similares a los encontrados en el caso de la DM1, indicando que los efectos patogénicos comunes a las dos enfermedades son debidos a un efecto patogénico del mRNA con expansiones que, al formar foci nucleares, secuestran proteínas de unión a CUG y alteran así la función normal del mRNA. Asimismo, la alteración en la expresión de los genes de la región DM1 podrían contribuir a la patogénesis de aquellas características fenotípicas que sean diferentes de la DM2 (revisado en Ranum y Day, 2002b).

2.1.2 Síndrome del X frágil

La primera enfermedad genética que se asoció a una mutación dinámica de un microsatélite fue FXS (Yu *et al.*, 1991; Verkerk *et al.*, 1991; Oberlé *et al.*, 1991), la forma más común de retraso mental hereditario. No se estableció el origen genético de esta enfermedad hasta 1969, cuando Lubs observó la relación entre el síndrome clínico que ya habían descrito Martin y Bell en 1943 con una alteración citogenética en el cromosoma X (Lubs, 1969). Tal como indica el nombre de la

enfermedad, esa alteración es una fragilidad del cromosoma X localizada en la parte distal del brazo largo del cromosoma, concretamente en la banda q27.3, y denominada FRAXA. Como demostró Sutherland en 1977 en sus estudios sobre fragilidades cromosómicas, esta fragilidad es detectable citogenéticamente en cultivos de linfocitos utilizando medio deficiente en ácido fólico. Se trata, por lo tanto, de una enfermedad genética ligada al cromosoma X.

2.1.2.1 Características fenotípicas

El FXS es la forma más común de retraso mental hereditario, y la segunda cromosomopatía más frecuente después del Síndrome de Down, con una incidencia en la población caucásica de 1 de cada 4000 hombres y 1 de cada 8000 mujeres (Turner *et al.*, 1996). El desorden se da en todas las etnias, aunque la incidencia puede variar de un grupo a otro (Crawford *et al.*, 2001).

Las manifestaciones clínicas del síndrome son muy variables y dependen en gran medida de la edad y el sexo. Las principales características físicas del FXS en el varón afectado, asociadas al retraso mental típico, incluyen orejas prominentes acompañadas de una cara alargada y estrecha con un mentón prominente (ver figura 5). Además, el 80% de los hombres adultos afectados presentan macroorquidismo, sin efectos en la fertilidad. Estos rasgos físicos normalmente están muy atenuados en los niños prepuberales, pero se van acentuando con la edad. En los niños, la primera manifestación clínica suele ser un retraso en la aparición del lenguaje, acompañado de hiperactividad y falta de atención. Los pacientes pueden presentar también problemas oftalmológicos, como el estrabismo, problemas de otitis y sinusitis recurrentes y problemas ortopédicos, como pies planos. La piel de estos niños es fina, suave y aterciopelada, y presentan hiperlaxitud de las articulaciones, sobretodo de los dedos de las manos. El sistema nervioso central también se ve afectado de forma general, manifestándose como hipotonía y bajo tono muscular. El comportamiento de estos individuos presenta unas pautas características que a veces resultan más útiles para el diagnóstico que el fenotipo físico, como la hiperactividad y falta de atención, timidez y ansiedad social, contacto visual pobre, aleteo con los brazos, hablar reiterativamente y, en algunos casos, comportamiento similar al autismo (revisado en Hagerman y Cronister, 1996).

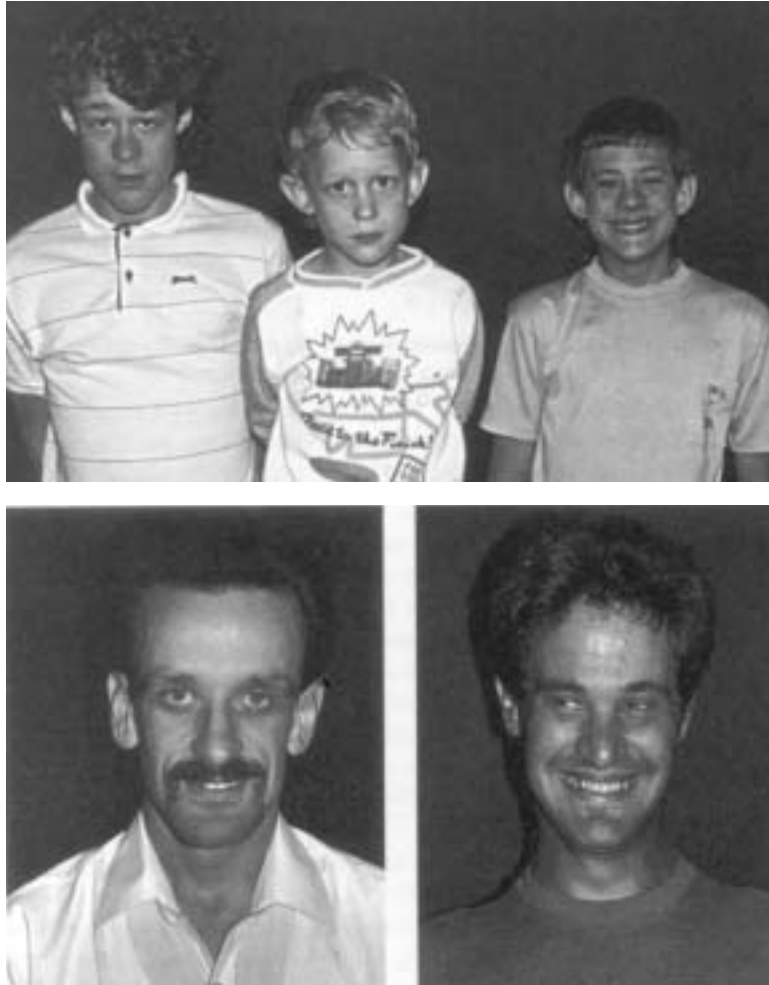


Figura 5. Varones con FXS que muestran algunos de los rasgos faciales típicos, como orejas y mentón prominente y/o una cara estrecha y alargada (Hagerman y Cronister, 1996).

El síndrome puede manifestarse también en mujeres, pero con una gran variabilidad. Existen desde casos con retraso mental y fenotipo parecido a los varones hasta casos en los que no se presenta ningún rasgo físico característico, que es lo más frecuente, y se observan tan sólo ligeros problemas de aprendizaje y dificultades emocionales.

2.1.2.2 Base molecular

El FXS está causado por la expansión del trinucleótido CGG en la región 5' no traducida del gen *FMR1*, concretamente en la región promotora del gen (Fu *et al.*, 1991; Verkerk *et al.*, 1991). Este gen codifica para una proteína llamada FMRP que se encuentra ausente o a nivel muy bajo en los individuos afectados. Cuando el

número de repeticiones CGG sobrepasa un umbral la isla CpG que se encuentra cerca de la región de las repeticiones (a 250 pb) se hipermetila, provocando la inactivación del gen y, por lo tanto, la falta de producción de la proteína.

El trinucleótido CGG asociado al FXS es muy variable, y en la población se encuentran esencialmente tres intervalos de longitud referidos como normal, premutación y mutación plena. En la población normal el número de repeticiones varía entre 6 y unas 50 repeticiones, entre las cuales el gen se expresa a nivel normal. El tamaño más común del trinucleótido son 30 repeticiones CGG con 1-3 secuencias simples AGG intercaladas (Fu *et al.*, 1991; Yu *et al.*, 1991) y esta secuencia es estable en la transmisión de padres a hijos. Cuando el número de repeticiones se sitúa entre 50 y 200, la secuencia adquiere una gran inestabilidad tanto en células somáticas como en células germinales, de manera que puede expandirse en tamaño al pasar de padres a hijos. Estos individuos, llamados también portadores, son sanos, ya que la isla CpG adyacente al trinucleótido no sufre hipermetilación y existe expresión del gen *FMR1*, sin embargo, estos individuos pueden tener descendencia afectada con el síndrome. La forma de mutación plena consiste en un tamaño del trinucleótido mayor de 200 repeticiones. A partir de este número de repeticiones, la isla CpG adyacente se encuentra anormalmente hipermetilada, de modo que la inactivación del gen impide la producción de proteína FMRP, manifestándose la enfermedad (revisado en Hagerman y Cronister, 1996).

La inestabilidad del trinucleótido CGG asociado al FXS está directamente relacionada con el número de repeticiones y con la edad del individuo. Así mismo, la pureza de la secuencia, es decir, si tienen o no secuencias simples intercaladas, afectan a la estabilidad de la secuencia repetida. Como en el caso de la DM1, la inestabilidad de las repeticiones se observa en células somáticas, dando lugar a un mosaicismo celular y una elevada heterogeneidad del número de repeticiones entre células y tejidos del mismo individuo. También se observa inestabilidad en células germinales, dando el fenómeno de la anticipación. En el caso del FXS el fenómeno de la anticipación fue denominado como paradoja de Sherman, pues cuando aún no se conocía la explicación molecular para el fenómeno, Sherman (1984; 1985) observó que en pedigrís afectados por el FXS, al avanzar en las generaciones mayor era la probabilidad de padecer la enfermedad.

2.1.2.3 Efecto patológico de la expansión

La causa última de la afectación clínica del FXS es la ausencia de la proteína FMRP, debido a una metilación anormal de la isla CpG del promotor del gen que la codifica. Esta proteína es fundamental en las primeras etapas del desarrollo, pero se sintetiza en todos los tejidos durante toda la vida del individuo, aunque de forma más abundante en cerebro y testículo. La FMRP es, pues, una proteína ubicua, que, dentro de la célula, se localiza principalmente en el citoplasma. Se trata de una proteína que se une al mRNA y participa en su transporte desde el núcleo de la célula hasta el citoplasma (Khandjian, 1999), de modo que su ausencia afecta a la producción de otras proteínas y, por lo tanto, las manifestaciones clínicas de la enfermedad son heterogéneas.

Este modelo presenta ciertas limitaciones ya que no explica porqué en individuos con un número de repeticiones CGG en los márgenes de la premutación (< de 200 repeticiones), que presentan una transcripción normal del gen *FMR1*, a veces se presenta un cuadro clínico leve. Se ha observado que en estos individuos, a pesar de que el gen se transcribe a nivel normal, los niveles de proteína son muy bajos (Tassone *et al.*, 2000 a,b,c). Esto indica que la traducción del mRNA y/o el transporte del mismo al citoplasma está afectado, dando lugar a niveles bajos de la proteína FMRP (revisado en Hagerman y Hagerman, 2001).

2.2 Origen de los alelos patogénicos con ERT

En un principio, existe la idea de que el mecanismo por el que un alelo normal pasa a ser un alelo premutado con predisposición a futuras expansiones, es distinto del mecanismo que genera diversidad de alelos dentro del intervalo normal. Se cree que en el rango de alelos cortos las variaciones siguen el modelo paso a paso (*step-wise*) que se observa en el resto de microsatélites, según el cual, el incremento o disminución en el número de repeticiones se da de uno en uno (Valdés *et al.*, 1993). No obstante, no está claro si las nuevas mutaciones que permiten el paso de un alelo normal a un alelo premutado son frecuentes o si se trata de sucesos raros.

Existen varias evidencias que indican que los sucesos de generación de alelos permutados son raros y, por lo tanto, sugieren un origen común de la mutación que causa la enfermedad. Esto significa que se darían pocas mutaciones que deriven en alelos permutados e incluso podría tratarse de un único suceso ancestral. Todo esto contrasta con la característica dominante de estas

enfermedades que conllevan una reducción de la eficacia biológica, por lo tanto, para mantener el alelo patogénico en la población, se necesita un nivel elevado de nuevas mutaciones. Dado que en las enfermedades asociadas con ERT los individuos más severamente afectados representan el punto final genético, debe haber una cantidad importante de individuos en la población que lleven y transmitan la premutación sin sufrir efectos fenotípicos graves.

2.2.1 Origen de los alelos DM1

En el caso de la DM1, hay evidencias claras de que en distintas poblaciones existe un desequilibrio de ligamiento entre el locus de la DM1 y polimorfismos en loci cercanos (Harley *et al.*, 1991; 1992). En uno de los primeros estudios realizados en pacientes europeos y japoneses, se encontró un desequilibrio total de ligamiento entre el alelo con la mutación plena y un polimorfismo de inserción de un elemento Alu (Alu(+)), situado a 5 kb de las repeticiones CTG (Harley *et al.*, 1992; Yamagata *et al.*, 1992; Mahadevan *et al.*, 1993). Posteriormente, se estudió la distribución de los alelos normales en la población con relación al polimorfismo Alu(+). Este estudio muestra que las poblaciones caucásicas presentaban una distribución tri-modal para los alelos normales. En esta distribución, el alelo de 5 repeticiones es el más frecuente, los alelos de 11-13 repeticiones representan el segundo modo en frecuencia, y existe un último modo heterogéneo, sin un pico claro (Zerylnick *et al.*, 1995). Además, se encontró una asociación del 100% entre los alelos con 5 repeticiones y Alu(+), del 98% entre los alelos (CTG)₁₁₋₁₃ y Alu(-), y del 100% en los alelos ≥ 18 repeticiones con Alu(+) (Yamagata *et al.*, 1996). En este estudio se sugiere que los 3 modos observados en la distribución podrían provenir de un número muy pequeño de sucesos iniciales. Asimismo, se propuso que los sucesos mutacionales que generan un nuevo modo son muy raros, mientras que los polimorfismos dentro de un modo tendrían su origen en el deslizamiento de la DNA polimerasa durante la replicación de estas secuencias (Yamagata *et al.*, 1996). A partir de todos estos datos se propuso que los alelos patogénicos provienen del cromosoma Alu(+) que lleva el alelo (CTG)₅, a partir del cual, por una o pocas mutaciones ancestrales, se llegó al alelo de ≥ 18 repeticiones, y estos alelos actuarían como reservorio para generar los alelos DM1 mediante mutaciones recurrentes. El otro cromosoma ancestral sería Alu(-) con un alelo (CTG)_{11, 12 o 13}. Un estudio más amplio llevado a cabo por Tishkoff y colaboradores (1998) permitió establecer que el haplotipo ancestral es el alelo (CTG)₉₋₁₇ con Alu(+), a partir del cual se generaron el resto de haplotipos. El alelo de 5 repeticiones tendría un origen más reciente, generado por pequeñas variaciones en el número de repeticiones, por

lo tanto, su elevada frecuencia se podría explicar por deriva genética o efecto fundador, por una mayor estabilidad del alelo de 5 repeticiones, y/o por una selección en contra de mutaciones que generen alelos <5 repeticiones. El resto de haplotipos se formarían por mutación y/o recombinación, llegando el haplotipo (CTG)₁₁₋₁₃ - Alu(-) a una frecuencia mayor por deriva genética. A partir del haplotipo ancestral se formarían los alelos >18 repeticiones, que componen el reservorio para la generación de alelos DM1 largos.

2.2.2 Origen de los alelos FXS

Los primeros estudios llevados a cabo con FXS también muestran que las nuevas mutaciones en el trinucleótido eran sucesos raros, lo cual era difícil de conciliar con la característica dominante de la enfermedad ligada al cromosoma X, en la que los varones afectados raramente se reproducen. Analizando los haplotipos de los cromosomas normales y mutantes, se demostró la existencia de un efecto fundador para los cromosomas X frágil, ya que se observó un desequilibrio de ligamiento entre el locus del FXS y determinados haplotipos (Richards *et al.*, 1992; Chakravarti, 1992). Todo esto sugirió que los alelos con premutaciones podrían ser llevados silenciosamente a través de pedigríes humanos durante muchas generaciones antes de su hiperexpansión e hipermetilación asociada a la enfermedad (Chakravarti, 1992; Oudet *et al.*, 1993). Morton y Macpherson (1992) propusieron un modelo multialélico para explicar el origen de la enfermedad, según el cual, la transición de un alelo normal estable a uno completamente expandido (>200 repeticiones) se da a través de una progresión entre 4 estados alélicos definidos. Un alelo normal estable (N) de forma rara podría sufrir una modesta inestabilidad, pasando a ser un inserto estable (S) de alrededor de las 40 repeticiones. Estos alelos con una pequeña predisposición a futuras expansiones pueden mantenerse en la población durante unas 90 generaciones antes de progresar a un alelo mayor más inestable (Z) o alelo premutado. La conversión de un alelo inestable (Z) a un alelo con la mutación completa (L) se da muy frecuentemente a través de la línea germinal femenina. Implícito a este modelo es el incremento gradual en el tamaño de las repeticiones CGG a medida que produce la progresión de N a S a Z. Según el modelo, debido a la longevidad del alelo S en las poblaciones humanas, se esperaría encontrar un enriquecimiento de alelos S largos, en aquellos haplotipos que presentan asociación al FXS, sin embargo, sólo se ha visto en algunos haplotipos. Esto sugiere que a parte de las influencias haplotipo-específicas, existe una influencia de las secuencias flanqueantes a las repeticiones o intrínsecas a las mismas, que pueden ser importantes a la hora de

determinar la estabilidad de un alelo, como el número y distribución de las interrupciones AGG (revisado en Lenzmeier y Freudenreich, 2003).

Eichler y colaboradores (1996), propusieron dos vías mutacionales distintas para explicar el origen de los alelos patológicos FXS. Por un lado, existirían unos haplotipos asociados a la mutación X frágil que mantendrían el patrón de interrupciones AGG y que, de manera gradual, progresarían hacia el umbral de inestabilidad, formando un reservorio de cromosomas con predisposición a sufrir grandes expansiones. Por otro lado, habría haplotipos también asociados a la mutación, pero en estas condiciones se darían pérdidas frecuentes de las interrupciones AGG, que llevarían a una progresión rápida hacia el umbral de inestabilidad. Por último, habría haplotipos no asociados o asociados negativamente con la mutación, que de forma esporádica podrían perder las interrupciones AGG, llevando también a una rápida progresión hacia el umbral de inestabilidad.

Según todo lo indicado anteriormente, parece claro que los alelos con la mutación completa se originan a partir del reservorio que suponen los alelos premutados, pero los eventos que llevan a la formación de estos alelos premutados parece que son más raros de lo que se esperaría debido a las características de estas enfermedades, y siguen sin estar del todo claros.

De todas formas hay que señalar que la mayoría de estudios se han centrado en los intervalos de premutación y mutación patológica, ya que son los alelos más inestables y donde encontramos patología. Por lo tanto, es interesante estudiar lo que ocurre en el intervalo normal, a nivel de pocas repeticiones, ver si también hay inestabilidad y cómo puede darse este paso a premutación.

3 Mecanismos de inestabilidad de las TNR

En los últimos años se han propuesto numerosos modelos para explicar la ERT, pero por el momento, poco se conoce realmente sobre la dinámica de los mecanismos de inestabilidad de estas repeticiones. Cada vez parece más claro que para generar el patrón de inestabilidad característico de las TNR asociadas a desórdenes humanos deben actuar múltiples mecanismos (McMurray, 1999). Varios estudios en organismos modelo, como en bacterias y levaduras, han permitido empezar a entrever los procesos celulares involucrados en las ERT.

Los desórdenes neurodegenerativos relacionados con las ERT muestran un umbral definido para el tamaño de la región de las repeticiones a partir del cual la

secuencia se vuelve altamente inestable, llevando a grandes expansiones. Aunque no se sabe exactamente qué es lo que lleva a las secuencias de repeticiones cortas a llegar a ese tamaño umbral, una posible explicación consiste en la acumulación lenta de pequeños incrementos en el número de repeticiones (Parniewski y Staczek, 2002). Este mecanismo de cambios en un número pequeño de repeticiones (*step-wise*) puede dar lugar a aumentos o pérdidas en el número de repeticiones y es el mecanismo de inestabilidad que se da en los microsatélites en general. Este mecanismo (ver figura 6) consiste en el deslizamiento de las dos cadenas complementarias posterior a una desnaturalización en la región de las repeticiones, formándose bucles de pocas unidades. Estos bucles, si no son reparados, durante la replicación van a producir pequeñas expansiones o contracciones. Si los bucles son reparados, también pueden producirse expansiones o contracciones de pocas unidades (ver figura 6). Una vez se ha llegado al tamaño umbral, las secuencias largas de repeticiones se vuelven muy inestables, produciéndose grandes expansiones mediante un mecanismo necesariamente distinto del responsable de la inestabilidad de las secuencias cortas.

Una diferencia importante entre los modelos generados para estudiar la inestabilidad y los datos obtenidos a partir de pedigríes humanos es la tasa de expansiones respecto a las contracciones. En todos los modelos estudiados se ha obtenido una mayor frecuencia de contracciones. En cambio, tanto en los estudios realizados en células humanas, como en los estudios *in vivo* llevados a cabo en pacientes de las enfermedades relacionadas con las ERT, se observa un claro sesgo hacia las expansiones, que aumenta al incrementarse el número de repeticiones de la secuencia. Este sesgo, no obstante, puede variar dependiendo del tipo celular, del locus estudiado, e incluso del paciente (revisado en Lenzmeier y Freudenreich, 2003).

Khajavi y colaboradores (2001) en cultivos linfoblastoides derivados de pacientes DM1 y Gomes-Pereira y colaboradores (2001) en cultivos celulares de ratones transgénicos, observaron que las células con números de repeticiones mayores presentaban una ventaja selectiva de crecimiento, de manera que las células con mayor número de repeticiones acababan siendo las más representadas en la población celular. Propusieron que este fenómeno, al que llamaron *mitotic drive*, era el responsable de la tendencia hacia las expansiones observada en los pacientes.

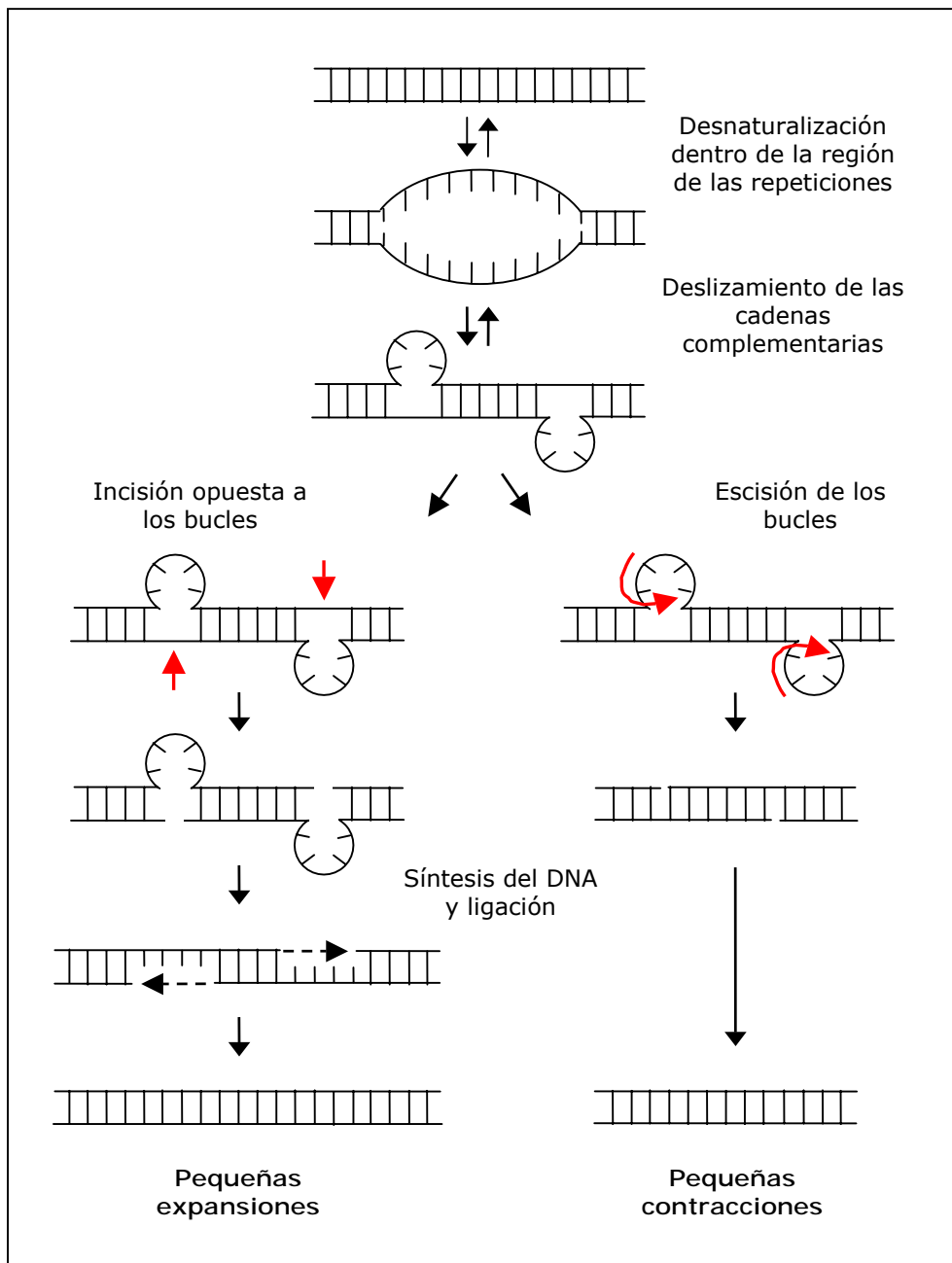


Figura 6. Mecanismo propuesto para explicar los pequeños cambios en el número de repeticiones (adaptado de Parniewski y Staczek, 2002).

3.1 Factores que actúan en cis

Los elementos en cis que afectan a la inestabilidad de las TNR, son todos aquellos que afectan a la formación y estabilidad de estas estructuras secundarias en la región de las repeticiones. Entre ellas se incluyen el tamaño, la pureza y el contexto cromosómico de la secuencia (como la dirección de la replicación, la proximidad al origen de replicación, la metilación del DNA, etc.).

Un factor importante para la inestabilidad de las TNR es la secuencia de repeticiones por sí misma. En varios experimentos con levaduras, en que se integró únicamente la región de las repeticiones en distintos loci, se observaron elevadas frecuencias de mutación (revisado en Lenzmeier y Freudenreich, 2003), sugiriendo que la inestabilidad es inherente a la secuencia, dependiendo de su composición de bases, número de repeticiones y la presencia de interrupciones en las repeticiones. Además, otros elementos en cis pueden contribuir a la inestabilidad, como las secuencias flanqueantes a la región de las repeticiones y las condiciones epigenéticas de alrededor, incluyendo la metilación del DNA, la estructura de la cromatina y la posición de los nucleosomas, ya que secuencias repetidas idénticas en distintos loci genéticos presentan niveles distintos de inestabilidad (revisado en Cleary y Pearson, 2003).

Otras evidencias que apuntan a la contribución de factores en cis en la inestabilidad de las TNR provienen de la comparación del patrón de inestabilidad de estas repeticiones con la inestabilidad genómica asociada a enfermedades causadas por una deficiencia en la reparación del DNA. A diferencia de la inestabilidad inducida por factores en trans típico de las células deficientes en reparación, la expansión de la secuencia de trinucleótidos en un individuo afectado se da únicamente en loci concretos, sin que se vean afectados otros loci con TNR. Así, en el caso del cáncer de colon hereditario no-poliposo (HNPCC, *hereditary nonpolyposic colorectal cancer*), con mutaciones en genes implicados en el sistema MMR, la inestabilidad se da en casi todos los microsatélites analizados (Goellner *et al.*, 1997; revisado en McMurray, 1999).

Todos los modelos propuestos para explicar la ERT se basan en la evidencia de que estas secuencias tienen una propensión a formar estructuras secundarias *in vitro* (revisado en Sinden, 1999; Sinden *et al.*, 2002). Así, la formación de estas estructuras *in vivo* se ha propuesto como primera fuente de inestabilidad genética de las repeticiones, ya que su presencia podría tener un impacto significativo en los procesos metabólicos del DNA que se den en la región de las repeticiones, como la replicación, la transcripción, la reparación y/o la recombinación. Las secuencias de TNR presentan unas características inherentes a la secuencia misma que permiten una mayor flexibilidad de la doble hélice y la formación de varias estructuras alternativas a la forma canónica B-DNA. Estas estructuras secundarias incluyen horquillas de cadena sencilla, cadenas triples o cuádruples de DNA, y deslizamientos entre las dos cadenas con la formación de pequeños bucles y/o horquillas (ver tabla 2 y figura 7).

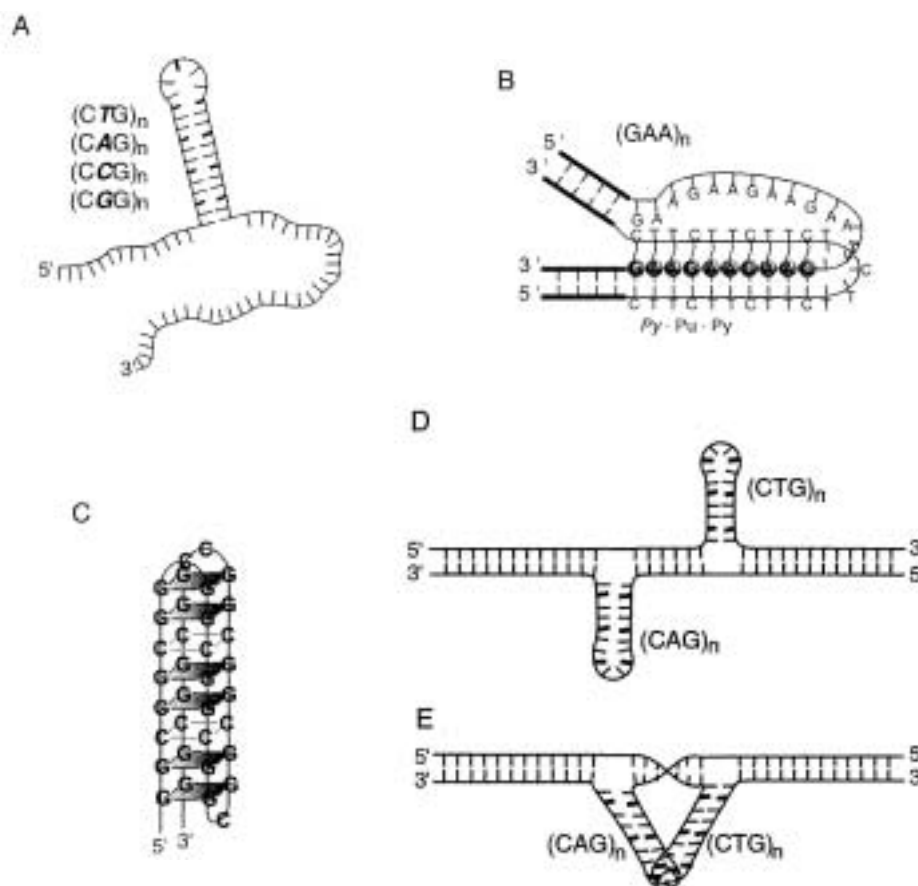


Figura 7. Estructuras alternativas de DNA formadas en las regiones con TNR: estructuras tipo horquilla (A); cadenas triples de DNA (B); cadenas cuádruples de DNA (C); deslizamientos entre las cadenas, formándose bucles y/o horquillas (D y E) (Sinden, 1999).

Tabla 2. Estructuras alternativas de DNA que pueden formar cada una de las TNR (adaptada de Sinden *et al.*, 2002).

Trinucleótido repetido	Estructuras alternativas de DNA
(CTG) _n •(CAG) _n	Hélice flexible, horquillas en cadena sencilla, deslizamientos entre cadenas.
(CGG) _n •(CCG) _n	Hélice flexible, horquillas en cadena sencilla, deslizamientos entre cadenas, cadena cuádruple de DNA.
(GAA) _n •(TTC) _n	Triple DNA intramolecular, deslizamientos entre cadenas?

Experimentos en sistemas modelo muestran que el avance de una horquilla de replicación por secuencias repetidas proporciona la oportunidad de formación de estructuras secundarias, y es igualmente probable que el proceso de reparación cerca de la región de las repeticiones pueda producir el mismo efecto.

También hay que destacar que en diferentes tejidos y en momentos diferentes del desarrollo, los elementos repetitivos de loci específicos, muestran niveles variables de inestabilidad, sugiriendo mecanismos diferentes entre los distintos loci y entre tejidos. Esto sugiere que elementos en cis específicos de tejido, desarrollo y locus contribuyen a la inestabilidad de las TNR. No obstante, dada la compleja naturaleza de las mutaciones dinámicas, es probable que tanto elementos en cis como factores en trans actúen de manera conjunta para producir una inestabilidad específica de tejido, desarrollo y locus, probablemente a través de los procesos de replicación del DNA, reparación y/o recombinación (revisado en Cleary y Pearson, 2003).

3.2 Metabolismo del DNA

Como se ha indicado anteriormente, la formación de estructuras secundarias complejas en las secuencias de TNR puede alterar el funcionamiento normal de los procesos metabólicos del DNA que se den en la región de las repeticiones, pudiendo generar inestabilidad en el número de repeticiones de la secuencia. Por otra parte, la resolución de estas estructuras secundarias proporciona oportunidades adicionales para que se den cambios en el número de repeticiones (revisado en Lenzmeier y Freudenreich, 2003).

3.2.1 Replicación del DNA

Parece claro que existe una conexión entre la inestabilidad de TNR y la replicación del DNA, ya que en desórdenes de TNR se observa inestabilidad en tejidos con una elevada proliferación celular. En el caso de la DM1, por ejemplo, se observa inestabilidad en células proliferativas, (Martorell *et al.*, 1995, 1998). Las grandes expansiones observadas en tumores de pacientes DM1 en comparación con sus tejidos no-neoplásicos sugieren que las expansiones pueden estar asociadas a la proliferación celular adquirida (Jinnai *et al.*, 1999).

3.2.1.1 Modelo clásico de deslizamiento entre las cadenas de DNA

Durante la replicación del DNA se pueden producir errores de inserción/delección de bases por deslizamientos entre la cadena naciente y la cadena molde. Por su naturaleza repetitiva, las secuencias de TNR son perfectos candidatos

para el deslizamiento de las cadenas, ya que pueden formar estructuras secundarias.

El modelo más aceptado de deslizamientos (figura 8) propone que las contracciones se producen si el bucle se forma en la cadena molde con síntesis discontinua (orientación II). Una región de repeticiones CTG o CGG de simple cadena dentro de la horquilla de replicación forma un bucle termodinámicamente estable, que no es replicado, produciéndose una deleción. A la inversa, las expansiones aparecen como consecuencia de la formación de estructuras secundarias en los fragmentos de Okazaki durante la síntesis de la cadena de síntesis discontinua (orientación I) (revisado en Parniewski y Staczek, 2002).

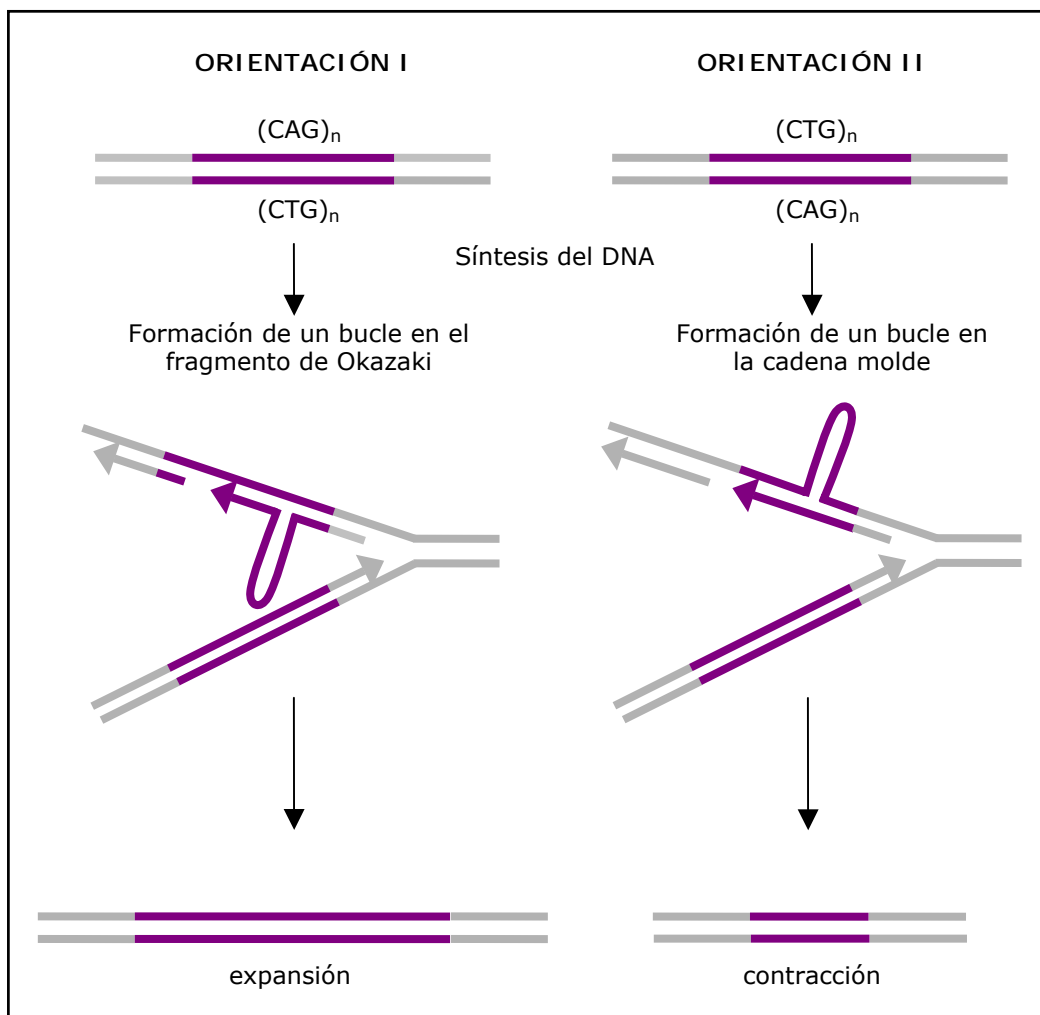


Figura 8. Modelo de inestabilidad por deslizamientos entre las cadenas durante la replicación (adaptado de Parniewski y Staczek, 2002).

Según este modelo, la dirección de la replicación a través de la región de las repeticiones afecta a su estabilidad, y depende de qué cadena contiene la secuencia de repeticiones más propensa a formar estructuras secundarias. Esta polaridad de las expansiones está de acuerdo con los estudios realizados *in vivo* en *E.coli* y levadura en los que se observa que la dirección de la replicación a través de las repeticiones CTG afecta a la inestabilidad de la secuencia (Kang *et al.*, 1995; Maurer *et al.*, 1996; Freudenreich *et al.*, 1997; Miret *et al.*, 1998). Lo mismo se ha encontrado en células de primate y en extractos de células humanas, demostrándose que la expansión es más probable cuando la secuencia CTG se encuentra en el fragmento de Okazaki (Cleary *et al.*, 2002; Panigrahi *et al.*, 2002). Asimismo, se ha encontrado que las repeticiones CGG en *E.coli* y levadura también muestran una inestabilidad dependiente de la dirección de la replicación (Hirst y White, 1998; Balakumaran *et al.*, 2000).

3.2.1.2 Procesamiento aberrante de los fragmentos de Okazaki

Algunos de los modelos propuestos para la inestabilidad de las TNR involucran al procesamiento aberrante de los fragmentos de Okazaki iniciados dentro de la región de las repeticiones (Gordenin *et al.*, 1997; Spiro *et al.*, 1999; Henricksen *et al.*, 2000). Estos modelos surgen a partir de los estudios realizados por Tishkoff y colaboradores (1997) en levaduras, donde mutantes deficientes en rad27/FEN-1, una proteína implicada en la eliminación de los cebadores de RNA que permiten la síntesis de los fragmentos de Okazaki (Lieber, 1997), mostraron un incremento de unas 1000 veces la tasa de duplicación de grandes secuencias aleatorias. Además, estos mutantes presentaban una tasa de recombinación elevada y sólo podían sobrevivir si eran capaces de reparar las roturas de doble cadena. Estos estudios sugieren que en la expansión de repeticiones pueden intervenir el procesamiento de los fragmentos de Okazaki, la recombinación y las roturas de doble cadena. La proteína RAD27 (igual que su homóloga en mamíferos, FEN-1) es la endonucleasa responsable de eliminar el fragmento de cadena sencilla que se genera en el extremo 5' de los fragmentos de Okazaki, durante la síntesis de la cadena discontinua cuando la DNA polimerasa encuentra el extremo 5' del siguiente fragmento de Okazaki y lo desplaza (Lieber, 1997) (figura 9A). Estudios adicionales en levadura, muestran una desestabilización de las secuencias CTG y CGG en mutantes Rad27⁻ (Freudenreich *et al.*, 1998; Schweitzer y Livingston, 1998; Spiro *et al.*, 1999; White *et al.*, 1999). A partir de todos estos estudios Gordenin y colaboradores (1997) propusieron un nuevo modelo para la ERT basado en el procesamiento de los fragmentos de Okazaki. Según este modelo, cuando en

el extremo 5' de los fragmentos de Okazaki se encuentran secuencias repetidas, éstas pueden formar estructuras secundarias resistentes a la endonucleasa FEN-1, proporcionando el paso inicial para la generación de grandes expansiones (figura 9B). Estudios *in vitro* sobre la especificidad de FEN-1 han demostrado que sólo puede actuar sobre DNA de cadena sencilla, de modo que las estructuras secundarias parcialmente de doble cadena bloquean la acción de la enzima (Bambara *et al.*, 1997; Lieber, 1997).

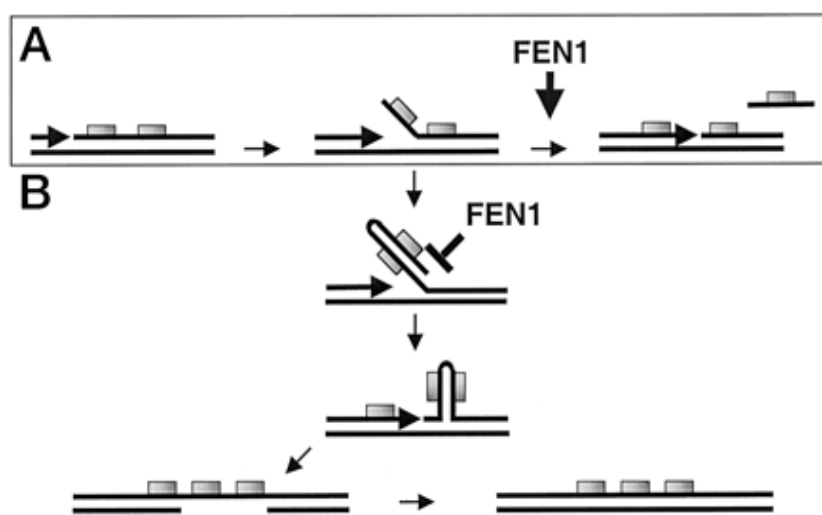


Figura 9. Modelo de la expansión de las repeticiones basado en el procesamiento de los fragmentos de Okazaki. (A) Acción de la endonucleasa FEN-1. (B) Vía propuesta para la ERT en el extremo 5' del fragmento de Okazaki (Henricksen *et al.*, 2000).

Hay varios caminos por los que un extremo 5' del fragmento de Okazaki no procesado puede producir una expansión. Por un lado, el bucle formado puede equilibrarse de nuevo con la cadena molde, permitiendo la ligación dando lugar a una expansión (figura 9B). Otra posibilidad, es que el extremo 5' del fragmento de Okazaki no escindido provoque un paro o retraso de la replicación, produciéndose una rotura de doble cadena que puede ser reparada vía recombinación homóloga (HR, *homologous recombination*) o vía unión de extremos no homólogos (NHEJ, *non-homologous end joining*), resultando en una expansión o delección (Gordenin *et al.*, 1997).

Lo importante que se puede inferir de estos estudios es que la expansión puede tener lugar en células donde FEN1/Rad27 es funcional, debido a la propensión de las secuencias de repeticiones de formar estructuras secundarias.

3.2.1.3 Bloqueo de la horquilla de replicación

Las estructuras secundarias formadas en las cadenas de DNA durante la replicación pueden provocar el bloqueo de la horquilla de replicación a su paso por la región de las repeticiones (Pelletier *et al.*, 2003; Freudenreich *et al.*, 1997; Samadashwiliy *et al.*, 1997). Estudios tanto *in vivo* como *in vitro* han mostrado una dependencia de tamaño y composición de las repeticiones en la pausa de la horquilla de replicación (revisado en Lenzmeier y Freudenreich, 2003 y en Cleary y Pearson, 2003). Así, las secuencias más largas de repeticiones provocan más pausas en la replicación que las secuencias más cortas, y esta pausa es dependiente de la secuencia de la cadena de síntesis discontinua molde, de tal forma que las secuencias CTG y CGG son más propensas a formar estructuras secundarias y provocar más pausas que las CAG y CCG. Asimismo, la presencia de interrupciones AGG en las repeticiones CGG produce una disminución de la pausa de la horquilla de replicación, facilitando así la progresión de la replicación a través de la región de las repeticiones (revisado en Lenzmeier y Freudenreich, 2003).

Se han propuesto distintos mecanismos que permiten restaurar la replicación dentro de la secuencia repetida después del bloqueo de la replicación. Estos mecanismos, entre los que se encuentran la inversión de la horquilla de replicación, la recuperación mediada por roturas de doble cadena o la recombinación, pueden facilitar la inestabilidad de las repeticiones, (revisado en Cleary y Pearson, 2003).

3.2.2 Transcripción

El proceso de transcripción de la región de las repeticiones es un factor potencial para influir en la inestabilidad de las repeticiones, pues supone la separación de las dos cadenas complementarias, permitiendo la formación de estructuras secundarias de las secuencias de repeticiones. En estudios realizados en bacterias se ha visto que la inducción de la transcripción incrementa la inestabilidad de las repeticiones, favoreciendo las contracciones (Bowater *et al.*, 1997; Schumacher *et al.*, 2001). Al igual que el efecto de la replicación en la estabilidad de las secuencias repetidas, el efecto de la transcripción también depende del tamaño y la composición de la secuencia. Las contracciones se producen cuando la cadena que se transcribe es la menos propensa a formar estructuras secundarias, de modo que la cadena complementaria queda en cadena sencilla, pudiéndose formar un bucle, que será saltado por el complejo de la DNA polimerasa durante la

replicación (Parniewski y Staczek, 2002). La transcripción puede afectar a la estabilidad de las repeticiones a través de la colisión con la horquilla de replicación o a través de una reparación aberrante acoplada a la transcripción (revisado en Cleary y Pearson, 2003).

3.2.3 Reparación

Como ya se ha comentado previamente, la inestabilidad de las TNR no está limitada a las células proliferativas, ya que también se ha observado en tejidos no proliferativos como el cerebro, en pacientes de la Enfermedad de Huntington, de Ataxia espinocerebelar 1 y de DM1 (revisado en Cleary y Pearson, 2003). Por lo tanto, se sugiere que procesos metabólicos independientes de la proliferación celular contribuyen a la inestabilidad de las repeticiones.

Las estructuras secundarias que se forman dentro de las TNR durante la síntesis del DNA, pueden alterar la estabilidad de las mismas si no son eliminadas mediante la reparación del DNA. No obstante, células con funciones normales de reparación también muestran inestabilidad de TNR.

Actualmente predominan dos visiones no-mutuamente excluyentes sobre el efecto de la reparación del DNA en la inestabilidad de las TNR (Lahue y Slater, 2003). Una de ellas supone que las estructuras secundarias formadas por las TNR bloquean la reparación, de modo que incluso las células con reparación normales no pueden evitar las expansiones o contracciones de estas secuencias. Otro modelo es que la reparación del DNA contribuye a la inestabilidad de las TNR, es decir, que la reparación tiene un efecto pro-mutagénico. Así, el daño en el DNA cerca o dentro de la región de las repeticiones o incluso la estructura aberrante formada por las repeticiones por sí misma, atraen a la maquinaria de reparación del DNA y la subsiguiente escisión de nucleótidos es seguida de síntesis tendiente a error. Esta síntesis, como en la replicación, puede generar expansiones o contracciones.

3.2.3.1 MMR

La vía MMR es un sistema fundamental involucrado en el mantenimiento de la integridad genómica, porque además de corregir los pares de bases mal apareados también repara algunas estructuras secundarias alternativas del DNA, como pequeños bucles y regiones no apareadas dentro del DNA. En algunos casos de cáncer hereditario, como el HNPCC, un defecto en el sistema MMR produce una

elevada inestabilidad de los microsatélites. Esto hizo pensar que el sistema MMR podía estar implicado en la inestabilidad de las TNR.

La función principal del sistema MMR consiste en eliminar los apareamientos entre bases incorrectos y los bucles de inserción/delección que se producen como consecuencia del deslizamiento de la DNA polimerasa durante la síntesis del DNA.

En *E. coli* este sistema de reparación está bien establecido y en él intervienen las proteínas MutS, MutL y MutH. La reparación es iniciada por el reconocimiento y la unión de MutS a las bases mal apareadas o al bucle, seguido de la unión de MutL. Este complejo activa a MutH, una endonucleasa que corta en los sitios hemimetilados GATC discriminando la cadena a reparar. Seguidamente, una exonucleasa 3'-5' ó 5'-3' elimina la secuencia que contiene el apareamiento erróneo y es reemplazada por la acción de la DNA polimerasa (Jiricny, 1998).

El sistema MMR está altamente conservado en la escala evolutiva. En humanos, el sistema es más complejo, pues se conocen varios homólogos para cada proteína MMR bacteriana. Para el reconocimiento del apareamiento erróneo, la proteína MSH2 forma un heterodímero con dos proteínas MMR adicionales, MSH6 o MSH3, dependiendo de si hay que reparar un mal apareamiento entre bases o bien un bucle, respectivamente. Otro heterodímero formado por MLH1 y PMS2 coordina la interacción entre el complejo que reconoce la alteración y el resto de proteínas necesarias para la reparación, que incluyen PCNA, exonucleasas, DNA polimerasas (δ y ϵ), factores de replicación y posiblemente helicasas (Peltomäki *et al.*, 2001). Mutaciones en los genes implicados en MMR producen un incremento de la tasa de mutación y una elevada inestabilidad en microsatélites, sugiriendo la posible implicación del sistema MMR en la inestabilidad de las TNR (Boyer *et al.*, 1995).

Sin embargo, el papel preciso del sistema MMR en la inestabilidad de las TNR es bastante controvertido. Las mutaciones de TNR pueden implicar tanto pequeños como grandes cambios de tamaño. Supuestamente, los cambios pequeños de tamaño se producen por el mecanismo *step-wise*, provocado por el deslizamiento de la DNA polimerasa durante la replicación, mientras que los cambios grandes se generan a partir de la formación de estructuras secundarias alternativas.

Al igual que en otros microsatélites, el sistema MMR juega un papel corrector en la rectificación de pequeños bucles de TNR que se producen durante la replicación, evitando las pequeñas mutaciones de TNR. Así, los defectos en el sistema MMR producen un incremento de pequeñas expansiones y contracciones en casi todos los microsatélites, incluyendo las TNR (Parniewski *et al.*, 2000; Wells *et al.*, 1998; Schmidt *et al.*, 2000; Hashem *et al.*, 2002b; Schweitzer y Livingston, 1997; revisado en Lahue y Slater, 2003).

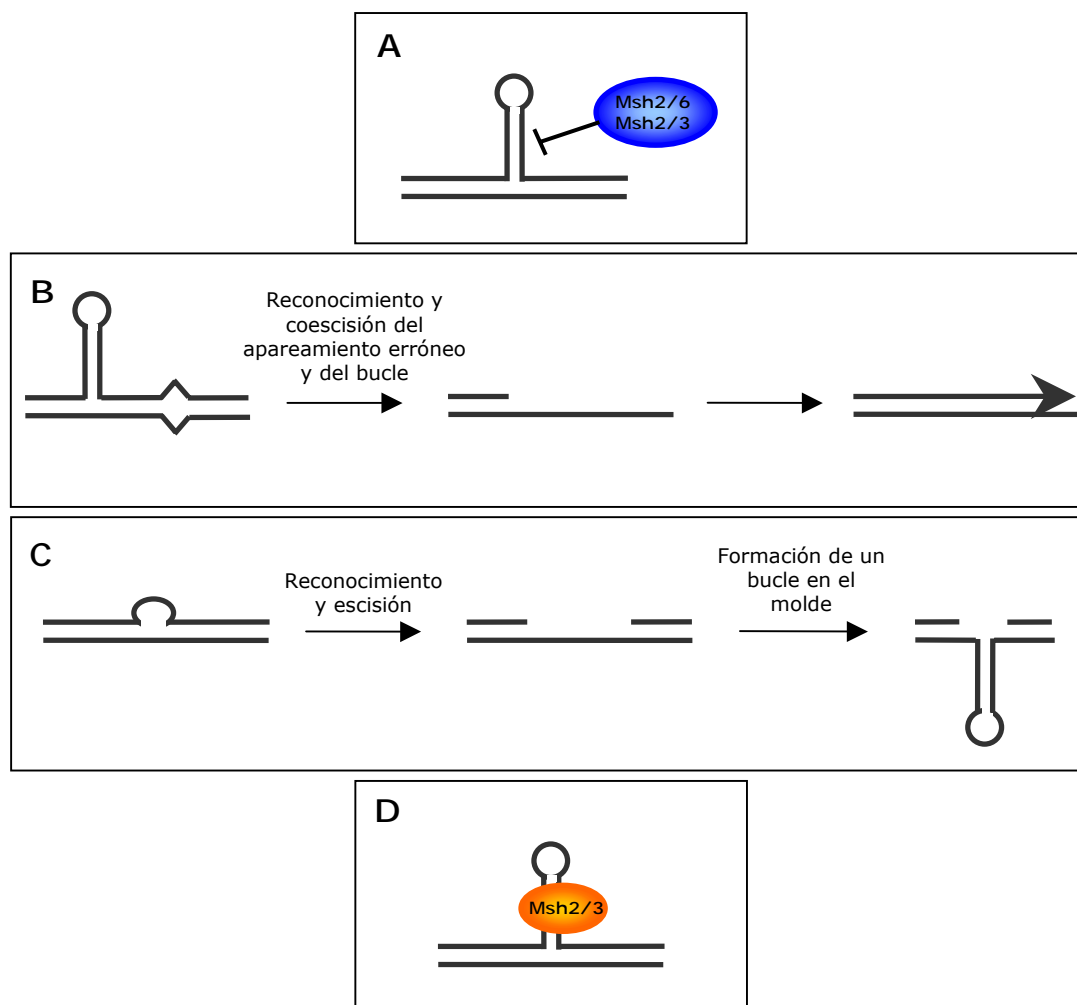


Figura 10. Modelos para explicar la implicación del sistema MMR en la inestabilidad de las TNR (adaptado de Lahue y Slater, 2003).

La influencia del sistema MMR en la generación de cambios grandes de tamaño en las TNR no está clara, ya que los resultados obtenidos en los distintos estudios son contradictorios. En levadura se ha observado que el sistema MMR no influye en la generación de grandes expansiones o contracciones (revisado en Lauhe y Slater, 2003). Una explicación de este hecho es que el sistema MMR no puede reparar los bucles formados por las TNR durante la síntesis del DNA (figura 10A) y estos precursores de expansiones y/o contracciones persisten en la célula. Otro modelo propone que los bucles formados en las TNR pueden ser reparados si el sistema MMR reconoce apareamientos erróneos localizados cerca y, al repararlos, también se repara el bucle (figura 10B). Estudios en bacterias muestran una mayor frecuencia de grandes contracciones en cepas salvajes, comparado con cepas deficientes en el sistema MMR, lo que sugiere un proceso en el que las contracciones se produzcan durante un funcionamiento normal de la reparación MMR (figura 10C).

Los estudios llevados a cabo con ratones deficientes (*knockouts*) en genes que intervienen en MMR permiten establecer un modelo adicional de la implicación de MMR en la inestabilidad de las TNR, indicado en la figura 10D (Kovtun y McMurray, 2001; Manley *et al.*, 1999; van den Broek *et al.*, 2002). Así, el complejo Msh2/Msh3, se une a la horquilla probablemente a través de los apareamientos erróneos que se dan dentro de la misma. Esta unión estabilizaría la estructura, permitiendo que la lesión persista y pueda ser usada como un molde expandido durante la subsiguiente síntesis del DNA. También se ha observado que el heterodímero Msh2/Msh6 tiene una función anti-expansión, quizás por la unión competitiva de Msh6 con Msh2, reduciendo así, la disponibilidad del complejo promutagénico Msh2/Msh3 (revisado en Lahue y Slater, 2003).

3.2.3.2 Reparación de las roturas de doble cadena

Estudios recientes en bacterias y levadura han implicado a la reparación de las roturas de doble cadena (DSBs, *double strand breaks*) como fuente de inestabilidad de TNR. En *E.coli* se ha visto que la pausa de la horquilla de replicación puede producir DSBs (Michel *et al.*, 1997), y en levadura se ha observado que las TNR pueden ser susceptibles a la formación de DSBs a través de procesos como el deslizamiento de la DNA polimerasa o el procesamiento endonucleolítico de las estructuras secundarias (Freudenreich *et al.*, 1998; Cohen *et al.*, 1999; Jankowski *et al.*, 2000; Richard *et al.*, 2000). Por lo tanto, se sugiere que la replicación deficiente de las TNR puede llevar a DSBs y la reparación de estas roturas puede contribuir a la inestabilidad de las repeticiones (revisado en Lenzmeier y Freudenreich, 2003 y en Cleary y Pearson, 2003).

Existen dos vías generales de reparación de las DSBs, por HR y por NHEJ. El sistema de HR es un tipo de reparación postreplicativa donde la cadena no dañada puede usarse como molde para la resíntesis y reparación de la cadena de DNA dañada. De este modo, la HR está restringida a la fase S del ciclo celular y permite que la maquinaria de la replicación continúe la síntesis, resolviendo las estructuras de DNA que bloquean el avance de la horquilla de replicación. La HR es un proceso muy preciso de reparación, en el que participan, entre otras, la proteína RAD51 y el complejo MRE11-RAD50-NBS1.

El proceso de NHEJ es mucho más robusto, pero con menos fidelidad de reparación de las DSBs, pues los extremos rotos simplemente se fusionan de nuevo sin tener en cuenta la secuencia del DNA. El sistema requiere al menos algunos pares de bases de homología entre los dos extremos, pero puede llevar a la pérdida de nucleótidos en la unión, por ello se le denomina también recombinación ilegítima

o recombinación tendiente a error. Es el mecanismo de reparación de DSBs predominante en G1 y las proteínas implicadas son principalmente Ku70, Ku80, DNA-Pkcs, DNA ligasa IV, XRCC4 (revisado en Haber, 2000 y en Barnes, 2001).

Estudios en levadura de inducción de DSBs y reparación por recombinación homóloga utilizando plásmidos con repeticiones CTG, mostraron que cuando el plásmido contiene secuencias cortas repetidas, al inducir los DSBs, se producen contracciones en el 20% de los casos (Richard *et al.*, 1999). Al aumentar el tamaño de la repetición del plásmido donador, un 30% de los cambios de tamaño observados fueron expansiones (Richard *et al.*, 1999). Estas expansiones se dieron por conversión génica durante la síntesis asociada a la reparación de los DSBs inducidos. Además, Richard y colaboradores (2000) encontraron que el complejo nucleasa Mre11-Rad50-Xrs2 en levadura previene las expansiones de las repeticiones CAG o CTG. Así, observaron que mutaciones del gen Mre11 o Rad50 provocaban más expansiones que contracciones, mientras que la sobreexpresión de Mre11 o de Rad50 no altera el ratio de expansiones/contracciones. Estos resultados juntos sugieren que el complejo Mre11-Rad50-Xrs2 permite eliminar las horquillas o bucles CAG o CTG que se dan durante la síntesis del DNA asociada con la reparación de las DSBs en estas repeticiones (Revisado en Lahue y Slater, 2003).

Por lo tanto, los estudios en levadura sugieren que la reparación de las DSBs vía HR puede llevar a expansiones o deleciones. En mamíferos, sin embargo, la vía preferencial de reparación de las DSBs es la NHEJ, y este mecanismo podría generar más expansiones que contracciones, explicando así lo observado en humanos. La baja tasa de NHEJ en bacterias y microorganismos podría explicar las frecuencias relativamente bajas de expansiones en estos organismos. No obstante, estudios realizados por Marcadier y Pearson (2003) en células de primate no mostraron esa tendencia hacia las expansiones observada en humanos. Sí observaron que la reparación, mediante NHEJ, de DSBs localizados dentro de la región de repeticiones del trinucleótido CTG producía alteraciones en el tamaño de la repetición con un sesgo hacia deleciones (Cleary *et al.*, 2002). Según estos resultados, es posible que las repeticiones CTG en los extremos libres formen estructuras secundarias que son eliminadas, produciendo deleciones. Estas estructuras formadas dentro de las TNR serían eliminadas por las proteínas que participan en la NHEJ (Marcadier y Pearson, 2003). No está claro qué proteínas específicas participan en el procesamiento de las estructuras repetidas deslizadas, pero lo que sí mostraron Marcadier y Pearson (2003) es que cuando estas estructuras se encuentran en una DSBs, su procesamiento produce deleciones (figura 10).

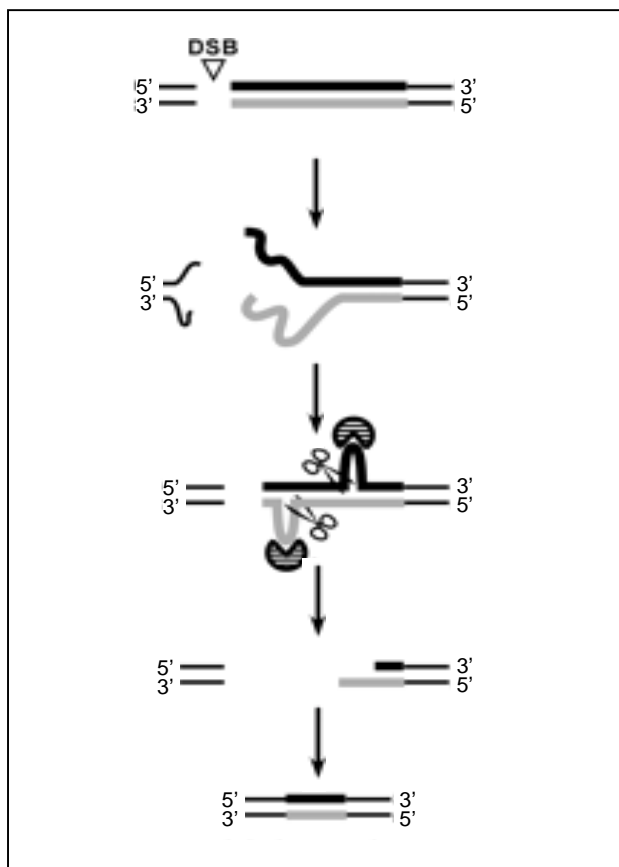


Figura 10. Modelo propuesto para explicar la delección de TNR a partir de una DSB (adaptado de Marcadier y Pearson, 2003).

3.2.3.3 Reparación por escisión de nucleótidos (NER)

Esta vía de reparación reconoce eficientemente y repara un gran número de lesiones, incluyendo grandes aductos en el DNA y enlaces cruzados que puedan causar una distorsión importante en la hélice, así como lesiones menos distorsionantes como bases metiladas. También se ha observado la participación de NER en la reparación de bucles de DNA *in vitro* (Ahn y Grossman, 1996; Kirkpatrick y Petes, 1997). Como las TNR pueden formar *in vivo* estructuras de DNA inusuales que provoquen la desestabilización de la doble hélice, es probable que estas estructuras atraigan a las proteínas de NER y que durante el proceso de reparación se incremente la inestabilidad de la secuencia repetida.

La transcripción a través de la región de las repeticiones puede desestabilizar adicionalmente las horquillas CTG, mediante la introducción de superenrollamientos negativos detrás del complejo de la RNA polimerasa. Es importante tener en cuenta que la vía NER está relacionada con la reparación de las cadenas que se transcriben (Surrallés et al., 2002). Cualquier clase de pausa de la

RNA polimerasa atrae al factor de transcripción asociado a la reparación. Esta proteína atrae a los componentes de NER a la región que se está transcribiendo, permitiendo la reparación rápida de lesiones en el DNA. Se podría asumir que la formación de estructuras en horquilla en la cadena molde así como en la cadena naciente de RNA podría llevar a la pausa de la RNA polimerasa (revisado en Parniewski y Staczek, 2002).

3.2.4 Recombinación

En humanos, las TNR asociadas a enfermedades exhiben una inestabilidad meiótica extrema. Así, se pensó que el entrecruzamiento desigual era un posible mecanismo para la ERT (McMurray 1995; Sia *et al.*, 1997). Una predicción de este modelo es que cada expansión en un alelo debería estar acompañada por una contracción en el otro alelo. No obstante, en las enfermedades asociadas a las ERT, el alelo normal siempre se mantiene. Además, los marcadores flanqueantes más cercanos no se ven afectados por el entrecruzamiento (McMurray, 1995).

En levadura, las repeticiones CAG inducen DSBs durante la meiosis, y la inestabilidad meiótica de las secuencias repetidas CAG es varias veces mayor que su inestabilidad mitótica (Cohen *et al.*, 1999; Jankowski *et al.*, 2000). Se ha demostrado que el mecanismo de reparación de DSBs es responsable de la elevada tasa de inestabilidad de las TNR durante la meiosis. Por lo tanto, en el fondo genético *spo11* (endonucleasa que inicia la recombinación) deficiente, la frecuencia de inestabilidad es similar al nivel mitótico, sugiriendo que la mayoría de alteraciones meióticas en el número de repeticiones de la secuencia se dan por el mecanismo de recombinación (Jankowski y Nag, 2002). Además, la reparación de las DSBs involucrando a las cromátidas hermanas contribuye significativamente a la inestabilidad de las secuencias de repeticiones. Experimentos con minisatélites también sugirieron que su inestabilidad es dependiente de recombinación, y que la conversión génica se produce utilizando los dos cromosomas homólogos y las cromátidas hermanas (Debrauwere *et al.*, 1999). En los estudios realizados en levadura, la reparación de DSBs resulta mayoritariamente en contracciones, con pocas expansiones. En humanos, las expansiones son más frecuentes que las contracciones, sugiriendo que los eventos que siguen a la formación de una DSB difieren en estos dos organismos.

Como hemos visto con anterioridad, la reparación de las DSBs puede ser una fuente importante de inestabilidad de TNR, y puesto que las repeticiones CAG inducen DSBs durante la meiosis y además la tasa de inestabilidad meiótica es mayor que la mitótica, es probable que la mayoría de alteraciones meióticas en el

número de repeticiones sean debidas a la reparación de DSBs inducidos por spo11, más que por el mecanismo de deslizamientos durante la replicación (Jankowski y Nag, 2002).

3.3 Mutagénesis

Hasta este momento no existe información acerca del alcance que pueden tener los factores ambientales en la inestabilidad de las TNR, a pesar de que se ha hipotetizado que estos factores podrían modular la inestabilidad de las repeticiones (Martorell *et al.*, 1998). Entre estos factores ambientales, los compuestos mutagénicos son buenos candidatos, pues pueden interferir con los mecanismos del metabolismo del DNA que se sabe están involucrados en la inestabilidad de TNR. Estos agentes mutagénicos pueden inducir lesiones en el DNA y su reparación, pueden producir el bloqueo de la transcripción y de la replicación y/o generar eventos recombinacionales para reparar las lesiones inducidas en el DNA. Como ya hemos visto, los procesos de replicación, recombinación y reparación del DNA están relacionadas con la inestabilidad de las TNR, de modo que podemos hipotetizar que el estrés mutagénico podría producir una inducción de la inestabilidad de las secuencias de TNR, llevando a un incremento en la frecuencia de las expansiones/contracciones en el número de repeticiones de la secuencia.

Muchos de los cánceres humanos están asociados con la exposición a agentes ambientales, por lo tanto, es probable que estos carcinógenos ambientales provoquen también inestabilidad genómica. Yamasaki y Mironov (2000) sugirieron que los carcinógenos podían inducir inestabilidad en los microsatélites, y que esta inducción sería mayor en células con inestabilidad de microsatélites. Estudios con *Drosophila* llevados a cabo en nuestro grupo (López *et al.*, 2002) demuestran la inducción de inestabilidad genómica por la bleomicina dependiente de MMR.

Jeffreys y colaboradores (1997) relacionaron la radiación ionizante con la inestabilidad de minisatélites. Analizaron la inestabilidad de minisatélites en poblaciones expuestas a radiación, usando una población de niños expuestos crónicamente al ^{137}Cs después del accidente de la central nuclear de Chernobyl en 1986. Encontraron que la tasa de mutación de las secuencias minisatélites estudiadas fue mayor en las familias expuestas que en las no expuestas, observándose además una relación dosis-respuesta. Sugirieron que tanto la radiación como posiblemente otros contaminantes ambientales, tales como los metales pesados, podían ser los inductores de la inestabilidad de minisatélites observada.

Ciertos agentes mutagénicos, incluyendo el EMS, ENU, luz UV y los agentes anticancerígenos mitomicina C (MMC), cisplatino y rayos-X, incrementaron la tasa de contracciones de las repeticiones CTG de un modo dependiente de tamaño y orientación en *E.coli* (Hashem y Sinden, 2002a). Asimismo, el daño oxidativo en el DNA también provoca contracciones en estas secuencias e inestabilidad de microsatélites (Jackson *et al.*, 1998; Zienolddiny *et al.*, 2000). A pesar de estos estudios, la experimentación es limitada con respecto a la inducción de inestabilidad en las TNR asociadas a enfermedades humanas.

Teniendo en cuenta las evidencias descritas anteriormente, en el presente estudio hemos analizado la inducción de la inestabilidad de trinucleótidos asociada a enfermedades humanas utilizando un agente inductor de enlaces cruzados, la MMC y bajo distintos fondos genéticos.

3.3.1 MMC

La MMC es un antibiótico con propiedades antitumorales que se usa en quimioterapia para el tratamiento de varios carcinomas. Su actividad biológica está relacionada fundamentalmente con su propensión a unirse covalentemente al DNA y formar un enlace cruzado entre las dos cadenas complementarias, uno de los tipos de lesiones más letales del DNA. A nivel molecular la MMC actúa como un agente alquilante que forma aductos en el N-2 de la guanina, enlaces intracatenarios y enlaces intercatenarios (Rink *et al.*, 1996). La reacción es específica de secuencia: la MMC alquila preferentemente guaninas en secuencias CpG, que pueden convertirse en un enlace cruzado intercatenario (Kumar *et al.*, 1992). Dentro de la célula, la MMC sufre un ciclo de oxidación y reducción, obteniéndose especies reactivas capaces de producir también radicales de oxígeno (Pritsos y Sartorelli, 1986).

Existen evidencias de que la citotoxicidad de la MMC es debida principalmente a la formación de aductos de DNA más que por los radicales de oxígeno. En particular, los enlaces cruzados intercatenarios, que resultan de la alquilación bifuncional del DNA por la MMC se han propuesto como la causa principal de muerte celular (Palom *et al.*, 2002). No obstante, mientras la atención se ha enfocado usualmente a los enlaces cruzados intercatenarios como la lesión más importante producida por la MMC, hay indicaciones de que algunos de los otros aductos también pueden ser citotóxicos (Rink *et al.*, 1996).

La MMC es un mutágeno S-dependiente, ya que los enlaces cruzados que genera producen un bloqueo de la horquilla de replicación al entrar en fase S. Este bloqueo induce DSBs que, en células de mamífero, son reparadas principalmente

por HR, generándose con frecuencia intercambios entre cromátidas hermanas (De Silva *et al.*, 2000; Jonhson *et al.*, 2000). Todo esto hace de este mutágeno un buen candidato para la inducción de inestabilidad de TNR, pues estos mecanismos están claramente relacionados con la de ERT.

4 Detección de la inestabilidad de trinucleótidos

Para el estudio de la inestabilidad de trinucleótidos se han usado varias aproximaciones distintas, que podemos dividir en estudios *in vivo* e *in vitro*. Los estudios *in vivo* incluyen los estudios con pacientes y pedigríes, y, más recientemente, con modelos de ratones transgénicos que contienen fragmentos de DNA genómico de pacientes, que incluyen TNR flanqueadas del DNA original humano (Monckton *et al.*, 1995; Lia *et al.*, 1998; Fortune *et al.*, 2000; Seznec *et al.*, 2000; Zhang *et al.*, 2002). Los estudios *in vitro* se han llevado a cabo en bacterias y levaduras, utilizando construcciones artificiales con las repeticiones insertadas en plásmidos, en el caso de las bacterias (Jaworski *et al.*, 1995; Kang *et al.*, 1995 y 1996), o en cromosomas, en el caso de levaduras (Schweitzer y Livingston, 1998). Además, una parte importante de estos estudios se llevan a cabo en líneas celulares con insertos que contienen las repeticiones o bien en líneas celulares derivadas de pacientes o de ratones transgénicos (Gomes-Pereira *et al.*, 2001).

Los estudios *in vivo* en pacientes presentan bastantes limitaciones, como la disponibilidad de las muestras apropiadas a lo largo de la vida del individuo, y están seriamente comprometidos por la variación genética inter-individual y ambiental (Fortune *et al.*, 2000). Además, los análisis de mutaciones en pedigríes están también limitados por el pequeño número de mutantes que se pueden detectar.

Tanto en los sistemas *in vivo* como *in vitro* se necesitan métodos de análisis sensitivos. Inicialmente los estudios *in vivo* en pacientes se basaban en análisis por Southern blot o PCR estándar usando grandes cantidades de DNA (típicamente $\geq 5\mu\text{g}$ para el Southern y $\geq 100\text{ng}$ para la PCR) (Brook *et al.*, 1992; Buxton *et al.*, 1992; Fu *et al.*, 1992; Lavedan *et al.*, 1993; Harley *et al.*, 1993; Ashizawa *et al.*, 1993). Estas dos aproximaciones sólo estiman el tamaño del alelo constitucional o más común pues analizan simultáneamente el tamaño de expansión en como mínimo 10^4 células, y los alelos expandidos se presentan como una mancha difusa, no como bandas discretas. En este tipo de análisis, se asume que la mancha está compuesta de múltiples fragmentos no resueltos derivados de los elevados niveles de inestabilidad de las repeticiones (Monckton *et al.*, 1995).

Jeffreys y colaboradores (1994) desarrollaron un método simple y sensible para detectar y cuantificar los alelos mutantes de un locus minisatélite en gametos individuales, mediante análisis por PCR de múltiples alícuotas diluidas de DNA (SP-PCR, *small pool PCR*). Al reducir la cantidad inicial de DNA, incrementa la probabilidad de amplificar un alelo distinto del constitucional, más raro, aunque para ello hay que realizar múltiples reacciones de la misma muestra. La SP-PCR puede detectar alelos mutantes *de novo* de cualquier muestra de DNA germinal o somático.

EL método de la SP-PCR se ha aplicado con éxito en el estudio de la estabilidad de las TNR. Esta técnica es el método de elección para examinar la inestabilidad de las repeticiones entre alelos con TNR usualmente estables, ya que con la SP-PCR no existe la limitación del tamaño de muestra que es necesario para observar cambios en el número de repeticiones utilizando Southern blot o PCR estándar (Crawford *et al.*, 2000). El método proporciona una aproximación robusta para el análisis preciso del nivel de variación de las repeticiones, proporcionando tanto datos cuantitativos como cualitativos.

En los estudios *in vitro* realizados con líneas celulares, debido al nivel de variabilidad inicial existente en las muestras, para poder detectar mutaciones *de novo* se han usado principalmente dos métodos: la SP-PCR y la clonación celular. El primer método, como ya se ha comentado, permite detectar nuevos alelos mutantes que, lógicamente, estarán menos representados que los alelos constitucionales y, comparando las distribuciones de los cambios detectados, permite determinar las tasas de mutación (Gomes-Pereira *et al.*, 2001; Ashizawa *et al.*, 1996). Otro método, muy usado también en los estudios de tumores con inestabilidad de microsatélites, es la clonación celular. Este método consiste en hacer cultivos clonales a partir de una única célula del cultivo inicial, de modo que se podrán detectar los nuevos alelos generados a lo largo de la progresión del cultivo (Yang *et al.*, 2003; Khajavi *et al.*, 2001).

II. OBJETIVOS



II. OBJETIVOS

Muchas enfermedades genéticas humanas están asociadas a la ERT. Entender los mecanismos responsables de estas expansiones es muy importante para conocer mejor la etiología y evolución de estas enfermedades. Las ERT se han relacionado con los procesos de replicación, reparación y recombinación, aunque también podrían estar implicados factores ambientales, como los agentes mutagénicos. Otro aspecto importante es conocer el mecanismo de generación de alelos patogénicos expandidos, así como la inestabilidad de alelos normales con pocas repeticiones.

En la presente tesis doctoral nos propusimos ampliar el conocimiento sobre los procesos genéticos involucrados en las ERT, e investigar si la inestabilidad de las TNR puede venir modulada por mutágenos ambientales. Así, los principales objetivos propuestos son los siguientes:

1. Comparar la inestabilidad en el número de repeticiones de los trinucleótidos CTG y CGG asociados a la DM1 y al FXS, respectivamente, en líneas celulares deficientes y no deficientes en el sistema MMR, con un número normal de repeticiones.
2. Comprobar si la MMC puede inducir cambios en el número de repeticiones de los trinucleótidos CTG y CGG asociados a la DM1 y al FXS, respectivamente, en alelos de tamaño normal, y si este fenómeno podría estar modulado por genes relacionados con el sistema MMR.
3. Estudiar la dinámica de la inestabilidad de las repeticiones CTG asociadas a la DM1 en alelos de tamaño normal y patogénico en células somáticas a lo largo del tiempo.
4. Evaluar el efecto de la MMC a lo largo del tiempo sobre la dinámica de la inestabilidad de las repeticiones CTG asociadas a la DM1 en células somáticas.

III. ARTÍCULOS



Most scientists regarded the new streamlined peer-review process as 'quite an improvement.'

ARTÍCULO 1

Induction of instability of normal length trinucleotide
repeats within human disease genes

J.Med.Genet., 41:e3, 2004

ELECTRONIC LETTER

Induction of instability of normal length trinucleotide repeats within human disease genes

L Fernández-López, E Piñeiro, R Marcos, A Velázquez, J Surrallés

J Med Genet 2004;41:e3 (<http://www.jmedgenet.com/cgi/content/full/41/1/e3>)

Mytotic dystrophy type 1 (DM1) is the most frequent cause of adult muscular dystrophy. It is a progressive multisystem disorder with an incidence of 1 in 8000 newborns. Its molecular basis is an expansion of a CTG repeat in the 3' untranslated region (UTR) of the *DMPK* gene on chromosome 19.¹ Fragile X syndrome is the most common inherited cause of mental retardation. It is caused by an expansion of a CGG repeat in the 5' UTR of the *FMR1* gene.²⁻⁴ The full mutation appears in approximately 1 in 4000 males and 1 in 9000 females. The number of trinucleotide repeats (TNRs) in most unaffected individuals is in the range of 5–35 repeats in DM1 and around 14–35 repeats in fragile X syndrome. However, pathogenic loci bear expanded alleles of up to several hundred or even thousand repeats in severe cases.¹⁻⁴

As with many other tandemly repeated microsatellite sequences, pathogenic TNRs are highly unstable in both the somatic and germ cell lineage, leading to expansions and contractions in the length of the repeats in successive cell or individual generations.⁵⁻⁸ Expansions in the number of TNRs are known to be a molecular basis of genetic anticipation,⁹ the progressively earlier age of onset of a disease with increasing severity in successive generations. An important consequence of genetic anticipation is that the genetic transmission is interrupted and, therefore, new pathogenic alleles have to be generated from unexpanded alleles. Little is known, however, about the mechanism(s) that cause a stable, non-expanded allele to become unstable and pathogenic, probably because normal length alleles are thought to be stable compared with expanded pathogenic alleles. In addition, much of the evidence for the mechanisms that are thought to produce repeat expansion in pathological TNRs are dismissed as irrelevant for normal alleles. Therefore, understanding the factors influencing the instability of normal unexpanded alleles is of critical importance.

A number of *trans* and *cis* acting factors have been proposed to play a modulating role in the instability of expanded, disease related TNRs.¹⁰ These factors are not mutually exclusive and include the purity and length of the repeated sequence,⁷⁻¹¹ age of individuals,^{4,7} mitotic drive,¹² the relative position of the TNR with respect to adjacent origins of DNA replication,¹³ and mutations in genes involved in DNA repair, replication, and recombination.¹⁴⁻¹⁶ Mutagenic agents are known to interfere with DNA metabolism by inducing DNA lesions and repair, blocking transcription and replication, and/or generating recombinational events to bypass or repair the induced DNA lesions.¹⁷ Therefore, we have hypothesised that mutagenic stress could result in an induction of instability of unexpanded TNR sequences, leading to an increase in the frequency of spontaneously occurring contractions/expansions in repeat number. To test this hypothesis, we analysed normal length TNR sequences at a single repeat resolution by radioactive PCR and denaturing polyacrylamide gels, using DNA from multiple single cell clones expanded in the presence or absence of mitomycin-C

Key points

- The objective of this study was to explore the effect of the mismatch repair genes *MLH1* and *MSH2/MSH6*, and mutagenic stress in the instability of normal length (CTG)_n and (CGG)_n repeats within the myotonic dystrophy and fragile X syndrome genes, respectively.
- Multiple single cell clones expanded in the presence or absence of mutagenic stress (mitomycin-C) show that normal length alleles can be destabilised with a clear bias towards repeat contractions. The data also suggest two mechanisms of mitomycin-C induced repeat instability: a major mechanism independent of mismatch repair and a minor mechanism dependent on *MLH1* deficiency.
- The major mechanism of mitomycin-C induced trinucleotide repeat instability acts *in trans*, as the clones presenting instability induced by mitomycin-C at the CTG locus also presented instability at the CGG locus, but not in a concurrently analysed mononucleotide repeat. Our data therefore indicate that (a) normal length trinucleotide repeats can be destabilised in their natural genomic context and (b) that mitomycin-C induces trinucleotide repeat contractions at normal length alleles, through a major mismatch repair independent *trans* acting mechanism.

(MMC), a mutagenic agent able to induce interstrand DNA cross links, strand breaks, and replication blockage. In addition, and in order to uncover a potential role of a functional mismatch repair (MMR) system in mutagen-induced TNR instability, all experiments were done in MMR proficient (SW480) or deficient cell lines (HCT116 and LoVo), as well as in the HCT116 cell line corrected for MMR deficiency by microcell mediated chromosome 3 transfer. Here we show that normal length alleles are not as stable as previously believed and that mutagenic stress induces TNR instability through an MMR independent *trans* acting mechanism.

MATERIALS AND METHODS

Cell lines

The human (male) colon cancer fibroblast cell lines SW480 (MMR proficient), HCT116 (hMLH1 deficient), and LoVo (hMSH2 and hMSH6 deficient) (kindly provided by Dr M. A. Peinado, Institut de Recerca Oncològica, Hospital Duran

Abbreviations: DM1, myotonic dystrophy type 1; MMC, mitomycin-C; MMR, mismatch repair; TNR, trinucleotide repeat; UTR, untranslated region

Table 1 Sequence and size of the trinucleotide repeats in the cell lines studied

Cell line (CTG)	Myotonic dystrophy		Fragile X*	
	Sequence (CTG)	Size (no. of repeats)	Sequence (CGG)	Size (no. of repeats)
SW480 Pure		12/13	10+AGG+8+AGG+9	29
HCT116 Pure		11/12	9+AGG+1+AGG+19	31
LoVo Pure		10/12	19+AGGAGG+8	29

HCT116+chr.3 was identical to HCT116 with the exception that the 11 CTG repeat allele was absent.
*All cells are male and therefore contain a single fragile X syndrome allele.

i Reynals, L'Hospitalet, Barcelona, Spain) were maintained in RPMI medium supplemented with 15% fetal bovine serum, 1% pyruvate, 1% l-glutamine, and (1%) penicillin-streptomycin as described previously.²² The HCT116 cell line corrected for MMR deficiency by microcell mediated chromosome 3 transfer (HCT116+chr.3 cell line; generously provided by Dr T Kunkel, National Institute on Environmental Health Sciences, Research Park, NC, USA) was maintained under G418 (400 µg/ml) selection in the same culture medium. The length and sequence of the repeats in the different cell lines are summarised in table 1.

Single cell cloning and culturing

Single cell clones were isolated from confluent cultures by limiting dilution into 96 well plates. The cells were diluted in culture medium at a concentration of 1 cell per 5 µl. Multiple 5 µl drops were deposited in Petri dishes and visualised using an inverted microscope. Only drops with one cell were individually plated into 96 well plates, in 100 µl of culture medium (with or without mutagen, see below). The clones were allowed to grow in the well and the medium was changed once a week. The expanded clones were transferred to 25 cm² flasks and maintained until confluence, when DNA was extracted by standard phenol-chloroform precipitation. Because of the limitations of the PCR to amplify rare alleles, only those contractions or expansions generated during the first population doublings can be detected; however, the clones were expanded until confluence to increase the yield of DNA.

Treatment of single cell clones

MMC was diluted to a concentration of 1 mg/ml and stored at 4°C. The concentration used in the microsatellite instability assays (0.005 µg/µl) induced a similar reduction of 30–50% of cell viability in all the cell lines analysed (data not shown). Mutagenic treatment was continuous in all cases. The medium with mutagen, at the indicated concentration, was replaced once a week. Parallel single cell cultures were left untreated to control for spontaneous instability.

Molecular analysis of microsatellite instability

A total of 32–57 (40 in most cases) single cell clones per treatment point and cell line were analysed for instability of the CTG repeats at the *DMPK* locus (chromosome 19), CGG repeats at the *FMR1* locus (X chromosome), and the polyA mononucleotide repeat BAT-25 (chromosome 1), which is known to be unstable in MMR deficient backgrounds and was included as an internal control of microsatellite instability. Microsatellite instability was analysed by radioactive PCR amplification and subsequent sequencing gel electrophoresis to detect changes in the length of the alleles at a single repeat resolution. The number of repeats in the amplified sequences was checked using a 30–330 bp ladder (GibcoBRL[®]) labelled with [γ -³²P]dATP. All PCR reactions

were performed in duplicate and run in parallel lanes of the gels to replicate the results. The reactions were carried out in a total volume of 25 µl with 1 µg of DNA using 3 µCi [α -³²P]dCTP or dATP. Amplification of the DM1-CTG repeats was performed as described by Cobo and co-workers²³ but using the primers 101 and 102 described previously by Brook *et al.*⁷ The amplification of the FMR1-CGG repeat was performed using the PCR conditions and primers (C and F) described by Fu *et al.*² The BAT-25 mononucleotide microsatellite was analysed basically as previously reported,²⁴ but using the primers described by Parsons and co-workers.²⁵ The reaction products were separated by denaturing electrophoresis in 6% polyacrylamide gels containing 8 mol/l urea and visualised by autoradiography, essentially as described previously.^{26–28}

Criteria for microsatellite instability quantification

Under the experimental design and conditions used, only those mutant alleles generated during the first divisions of the single cell clonal culture could be detected. Those mutations that occur in later divisions remained non-amplified because the normal pre-existing alleles were preferentially amplified in the competitive PCR reaction. However, those mutant alleles generated in the first divisions were predominant enough to be detected because they were clonally expanded. Each clone was derived from a single cell and, therefore, any clone presenting alleles with repeat sizes shorter (contracted) or longer (expanded) than the progenitor allele was counted as an unstable clone bearing mutant alleles. Thus, the mutation frequency can only be measured at the single clone level. All mutant clones presented new or additional alleles with a gain or loss of only one or two repeat units.

Statistics

Statistical analysis of mutagen induced instability of microsatellite sequences was performed using Fisher's exact test comparing treated and untreated cultures. The degree of coincidence in the profile of genomic instability of the three markers, in multiple single cell clones of each cell line, was performed either by analysing contingency tables and performing the χ^2 test, or by the p value of Fisher's exact test, depending on the number of unstable clones.

RESULTS

The instability of three microsatellites (BAT-25, DM1-CTGn, and FMR1-CGGn) has been studied in multiple single cell clones expanded under the effect of a biologically effective subtoxic concentration of MMC. A total of 32–57 (40 in most cases) single cells derived from MMR proficient (SW480 and HCT116+chr.3) and deficient (HCT116 and LoVo) cell lines were clonally expanded in the presence or absence of MMC. All single cell clones were analysed for instability of each of the three microsatellite loci by radioactive PCR, sequencing gels, and autoradiography. All single cell clones were PCR amplified twice and electrophoresed in consecutive lanes of the gels to confirm the results. PCR reactions and analyses were repeated when PCR products of replicated reactions gave non-coincident band patterns in consecutive lanes of the autoradiographs. Some examples of the banding patterns obtained in the present investigation are shown in fig 1.

Table 2 and fig 2 summarise the results on the spontaneous and MMC induced instability of the mononucleotide repeat BAT-25 and the DM1 and FMR1 TNRs in multiple single cell clones of MMR proficient (SW480 and HCT116+chr.3) and deficient (HCT116 and LoVo) cell lines. We found elevated levels of spontaneous instability of the microsatellite BAT-25 in HCT116 (11 clones with expansions or contractions in the repeat number out of 33 clones analysed, or 11/33 clones) and

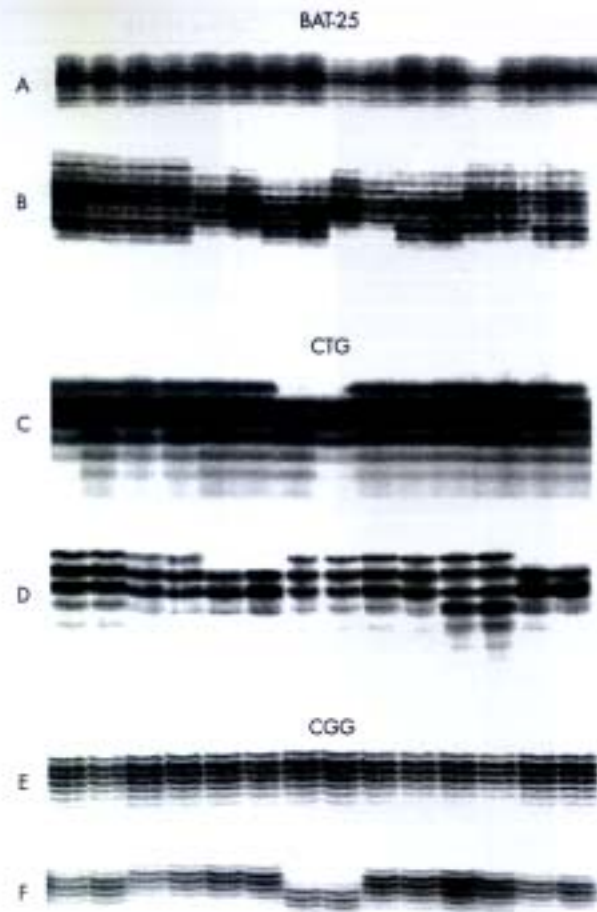


Figure 1 Representative images of the analysis of microsatellite instability by radioactive PCR and sequencing gels. Examples are shown on the microsatellite BAT-25 in untreated SW480 (A) and LoVo (B) cells, DM1 CTG repeat in untreated HCT116 cells (C) and MMC treated SW480 cells (D), and the fragile X syndrome CGG repeat in untreated (E) and MMC treated (F) SW480 cells. Eight (A,B) or seven (C-F) single cell clones are shown in each case, with two repeated PCR amplicons in each consecutive lane. The length of repeated sequences and the number of alleles in each cell line are shown in table 1.

especially in the LoVo cell line (25/35) compared with the MMR proficient cell lines SW480 (6/39) or HCT116+Chr.3 (0/40). This is consistent with the mutator phenotype

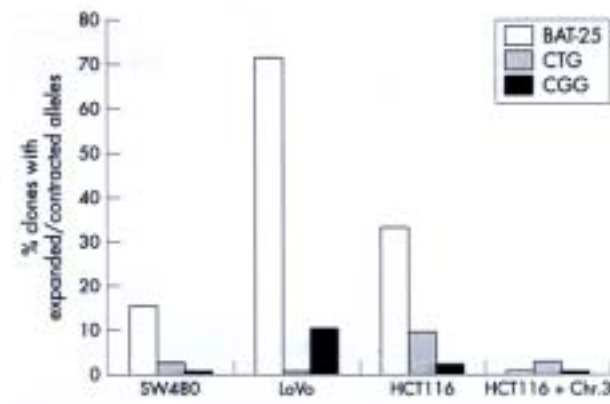


Figure 2 Spontaneous instability of the mononucleotide repeat BAT-25 and the trinucleotide repeats (CTG), and (CGG), in MMR proficient (SW480 and HCT116+chr.3) and deficient (LoVo and HCT116) cell lines.

characteristic of MMR deficiency and confirmed the suitability of our experimental design to detect microsatellite instability.

Although a statistically non-significant tendency towards higher instability of TNRs in MMR deficient cells (especially the CGG repeat in the LoVo cell line and the CTG repeat in the HCT116 cell line) was observed, both trinucleotides presented a low spontaneous instability compared with the mononucleotide repeat, irrespective of the MMR genetic background (table 2, fig 2). This result indicates that, in the absence of mutagenic stress, a functional MMR plays a secondary role in human cells to maintain the stability of unexpanded TNRs associated with human diseases in their natural genomic context.

MMC dramatically increased the frequency of unstable clones in the SW480 cell line and, to a lesser extent, in the HCT116 cell line (table 2, fig 3). DNA sequencing analysis of the mutants confirmed that changes in the size of the PCR products were to changes in the repeat number, rather than to deletions elsewhere in the flanking regions (data not shown). MMC induced an 18-fold increase in the frequency of CTG repeat mutations (from 2.6 to 45% in untreated and MMC treated cells, respectively; $p < 0.001$) and a >20-fold increase in the frequency of CGG repeat mutations (from 0.0 to 53.8%; $p < 0.001$) in the SW480 cell line. In addition, all MMC induced mutations had a contraction of the repeat sequence (fig 3). This increase was not observed in the MMR proficient HCT116+chr.3 cell line, indicating that the

Table 2 Spontaneous and mutagen induced instability of the mononucleotide repeat BAT-25 and the myotonic dystrophy type 1 (CTG) and fragile X syndrome (CGG) trinucleotide repeats in multiple single cell clones of MMR proficient (SW480 and HCT116+chr.3) and deficient (HCT116 and LoVo) cell lines

Treatment	Marker	Cell line			
		SW480	LoVo	HCT116	HCT116+chr3
Control	BAT-25	6/39 (15.4)	25/35 (71.4)	11/33 (33.3)	0/40 (0.0)
	CTG	1/39 (2.6)	0/36 (0.0)	3/32 (9.4)	1/40 (2.5)
	CGG	0/41 (0.0)	6/57 (10.5)	1/44 (2.3)	0/40 (0.0)
MMC	BAT-25	15/40 (37.5)	27/38 (71.0)	23/40 (57.5)	0/40 (0.0)
	CTG	18/40 (45.0)***	1/40 (2.5)	8/40 (20.0)	3/40 (7.5)
	CGG	21/39 (53.8)***	2/40 (5.0)	10/40 (25.0)**	0/40 (0.0)

MC/TC (%), mutant clones/total clones (percentage of mutant clones).

The difference between control and MMC treated cells is statistically significant with ** $p < 0.01$ or *** $p < 0.001$.

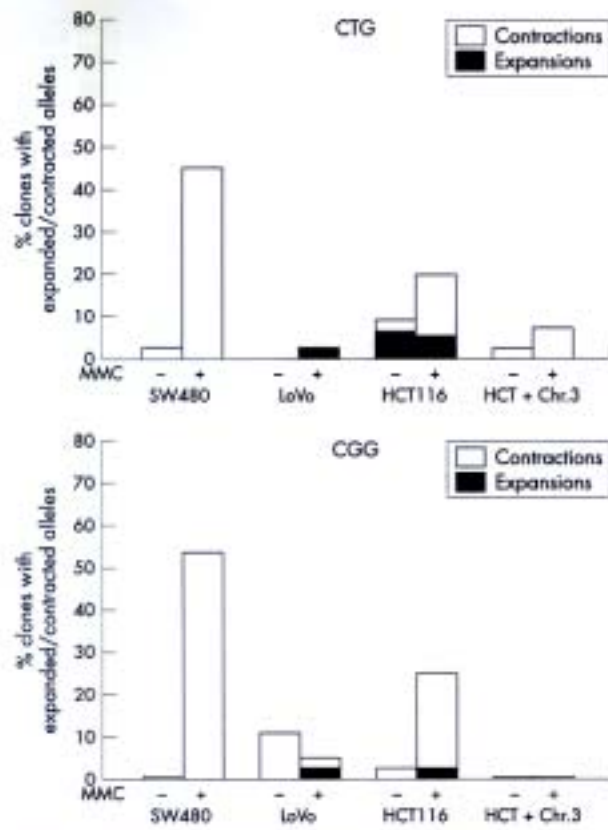


Figure 3 MMC induced expansions and contractions of the trinucleotide repeats (CTG), (DM1 locus) and (CGG), (fragile X locus) in MMR proficient (SW480 and HCT116+chr.3) and deficient (LoVo and HCT116) cell lines.

observed TNR instability induced by MMC in the SW480 cell line is not dependent on a functional MMR. However, the level of MMC induced TNR contractions is higher ($p < 0.001$) in the MLH1 deficient HCT116 cell line than in its genetically matched MLH1 proficient HCT116+chr.3 counterpart (fig 3), suggesting two mechanisms of MMC induced TNR instability: a major mechanism independent of MMR and a minor mechanism dependent on MLH1 deficiency. The negative results obtained in the LoVo cell lines indicate that MSH2/MSH6 deficiency plays no role in MMC induced instability.

To obtain further insights into the mechanisms of MMC induced TNR instability in the SW480 and the HCT116 cell lines, we next analysed the profile of MMC induced microsatellite instability in multiple single cell clones (fig 4). Surprisingly, we observed that most SW480 clones presenting instability in the CTG marker (black squares) also presented instability in the CGG marker but not necessarily in the mononucleotide marker which profile of instability is not coincidental. That was not the case in the HCT116 cell line after MMC treatment. The coincidence in the profile of MMC induced CTG and CGG contractions was highly significant ($p < 0.001$) and cell type (SW480) specific, suggesting a *trans* acting mechanism of MMC induced TNR instability in this cell line.

DISCUSSION

Expansions of TNRs are causally related to an increasing number of human genetic diseases. Most studies have been focused on expanded (pre-) pathogenic alleles that, unlike normal length alleles, are known to be highly unstable at both the somatic and germinal levels. Mouse models and

Cell line SW480				Cell line HT116			
N	Marker			N	Marker		
	BAT	CTG	CGG		BAT	CTG	CGG
1	Grey	Black	Black	1	Black	Grey	Grey
2	Grey	Black	Black	2	Black	Grey	Grey
3	Grey	Black	Black	3	Black	Grey	Grey
4	Grey	Black	Black	4	Black	Black	Grey
5	Grey	Black	Black	5	Black	Grey	Grey
6	Grey	Black	Black	6	Black	Grey	Grey
7	Grey	Black	Black	7	Black	Grey	Grey
8	Grey	Black	Black	8	Black	Black	Grey
9	Grey	Black	Black	9	Black	Grey	Grey
10	Grey	Black	Black	10	Black	Grey	Grey
11	Grey	Black	Black	11	Black	Grey	Grey
12	Grey	Black	Black	12	Black	Black	Black
13	Grey	Black	Black	13	Black	Grey	Grey
14	Grey	Black	Black	14	Black	Grey	Grey
15	Grey	Black	Black	15	Black	Black	Grey
16	Grey	Black	Black	16	Black	Grey	Grey
17	Grey	Black	Black	17	Black	Grey	Grey
18	Grey	Black	Black	18	Black	Grey	Grey
19	Grey	Black	Black	19	Black	Black	Black
20	Grey	Black	Black	20	Black	Black	Black
21	Grey	Black	Black	21	Black	Black	Black
22	Grey	Black	Black	22	Black	Grey	Black
23	Grey	Black	Black	23	Black	Grey	Grey
24	Grey	Black	Black	24	Black	Grey	Grey
25	Grey	Black	Black	25	Black	Black	Black
26	Grey	Black	Black	26	Black	Grey	Grey
27	Grey	Black	Black	27	Black	Grey	Grey
28	Grey	Black	Black	28	Black	Grey	Grey
29	Grey	Black	Black	29	Black	Grey	Grey
30	Grey	Black	Black	30	Black	Black	Black
31	Grey	Black	Black	31	Black	Grey	Grey
32	Grey	Black	Black	32	Black	Grey	Grey
33	Grey	Black	Black	33	Black	Black	Black
34	Grey	Black	Black	34	Black	Grey	Grey
35	Grey	Black	Black	35	Black	Black	Black
36	Grey	Black	Black	36	Black	Grey	Grey
37	Grey	Black	Black	37	Black	Black	Black
38	Grey	Black	Black	38	Black	Grey	Black
39	Grey	Black	Black	39	Black	Black	Black
40	Grey	Black	Black	40	Black	Black	Black

Figure 4 Profile of MMC induced BAT-25, (CTG), and (CGG), microsatellite instability in 40 independent single cell clones derived from the cell lines SW480 and HCT116. Grey and black boxes represent stable and unstable (with contracted or expanded alleles) clones, respectively, for each microsatellite analysed.

micro-organisms have been used to study such instability and the role of MMR deficiency, although in these genetic models the expanded pathogenic allele is out of its human natural genomic context. In addition, the results are conflicting, as MMR deficiency leads to TNR instability in

Escherichia coli and yeast,^{15, 21} whereas MSH2²⁴ and MSH3²² are required for TNR instability in mice. Other studies have predicted no role of MMR in TNR instability in yeast.²¹ All these studies were performed with expanded disease causing alleles. Studies on the instability of non-expanded alleles are very scarce. Here we show that non-expanded alleles are not as stable as previously believed and can be dramatically destabilised in their natural genomic context after mutagenic stress by a *trans* acting mechanism independent of MMR. A minor role for MLH1 was also observed in our study and, therefore we concluded that there are two mechanisms of MMC induced TNR instability: a major mechanism independent of MMR and a minor mechanism dependent on MLH1 deficiency. To reach this conclusion, we have used a novel approach based on the detection of instability of normal length alleles at single repeat resolution in cell clones expanded in the presence or absence of mutagenic MMC treatment. This approach led us to conclude that MMC was able to induce a >18-fold increase in the frequency of mutant clones bearing contractions in both TNRs in the SW480 cell line only. A minor increase in MMC induced TNR contractions was also observed in the MLH1 deficient cell line (HCT116); however, such inductions were not observed in the MSH2/MSH6 deficient cell line (LoVo) or in the HCT116 cell line corrected for MLH1 deficiency by microcell mediated chromosome 3 transfer. This is in agreement with previous data suggesting that MSH2 deficiency plays a minor role in disease associated TNR instability in human cells²⁴ or yeast.²³ A previous study in cell lines also indicated that mutations in *hMLH1* or *hMSH2/6* does not result in the instability of CTG or CGG tracts to the levels observed in individuals with DM1 or fragile X syndrome.²⁵ However, the method used in that study (Southern blot) did not allowed the detection of mutant alleles resulting from expansions or deletions of single repeats.

Interestingly, those single cell SW480 clones showing MMC induced contractions at the DM1 repeat also presented contractions at the FMR1 repeat. This finding is probably unrelated to the number of repeats and their sequence or purity in the different cell lines, as they all have a pure 10–12 CTG repeat sequence in the DM1 locus and 29–31 CGG repeats with two interruptions in the fragile X locus. The coincidence in the profile of TNR instability in the SW480 cells is highly indicative of a *trans* acting mechanism of TNR instability in this cell line instead of a direct *cis* action of MMC on the repeated sequence at both loci in each unstable clone.

The exact mechanism(s) explaining our observation is currently unknown. MMC is an S phase dependent mutagen that induces DNA cross links that block replication. The relationship between replication and TNR expansion is well documented, especially in yeast, where mutations in genes involved in the replication apparatus have been shown to increase TNR instability.^{26, 27} According to the literature, a possible candidate is FEN1, a protein involved in base excision repair and in the processing of the Okazaki fragments during replication of the lagging strand,²⁸ with an important role in TNR stability, at least in yeast.^{29, 30} However, human FEN1 immunoblotting experiments indicate that FEN1 is present at normal expression levels and molecular size in SW480 cell extracts (data not shown). Another candidate gene we are now extensively investigating is p53, as it is known that the SW480 cell line expresses a non-functional p53 protein.

The fact that the non-coincidental increase of CTG and CGG instability after MMC treatment is observed in the HCT116 cell line but not in the corrected HCT116+chr.3 cell line suggests a *cis* effect of MMC in TNR instability in an *hMLH1* deficient background. Our results also

indicate that *hMLH1* plays a role in the cellular response to MMC.

In the present study we have identified mutagenic stress as a new factor in disease associated normal length TNR instability, suggesting that normal length alleles are not as stable as previously believed. As all MMC induced altered clones had contractions, this effect is probably unrelated to the generation of (pre)pathogenic alleles from normal length alleles. Moreover, the process regulating repeat instability in somatic mitotic cells may be different to the situation in the germline *in vivo*. To clarify a potential role of MMC in the formation of pathogenic alleles, we are currently investigating by small pool PCR⁴ whether MMC modulates the repeat dynamics of DM1 alleles in successive cell generations.

ACKNOWLEDGEMENTS

We are greatly indebted to Professor Dr G Sutherland (Women's and Children's Hospital, Adelaide, Australia) for helpful comments and for reading the manuscript. We would also like to thank Dr L. Martorell (Sant Pau Hospital, Barcelona, Spain) for helpful comments and technical advice, and Drs M A Peinado and T Kunkel for sharing the cell lines. This work was partially funded by La Fundació La Marató de TV3 (project no. 1999-85), the Generalitat de Catalunya (SGR-00197-2002) and the Spanish Ministry of Science and Technology (MCYT) (project no. CICYT, PM98-0179). L Fernández-López and E Piñero are supported by predoctoral fellowships awarded by the MCYT. Dr J Surrallés is supported by a "Ramón y Cajal" project entitled "Genome stability and DNA repair", awarded by the MCYT and co-financed by the UAB.

Authors' affiliations

L Fernández-López, E Piñero, R Marcos, A Velázquez, J Surrallés, Group of Mutagenesis, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

The first two authors contributed equally to this work.

Correspondence to: Dr J Surrallés or Dr A Velázquez, Group of Mutagenesis, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain; jordi.surralles@uab.es or antonia.velazquez@uab.es

Received 6 May 2003

Accepted 2 July 2003

REFERENCES

- Brook JD, McCurach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 1992;68:799-808.
- Fu Y-F, Kuhl DPA, Fizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG Jr, Warren ST, Oostra BA, Nelson DL, Caskey CT. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991;67:1047-58.
- Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 1991;252:1711-14.
- Harley HG, Brook JD, Flynd J, Rundle SA, Crow S, Walsh KV, Thibout M-C, Harper PS, Shaw DJ. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992;355:545-6.
- Longshore JM, Torleton J. Dynamic mutations in human genes: a review of trinucleotide repeat diseases. *J Genet* 1996;75:193-217.
- Manckton DG, Wang LC, Ashizawa T, Caskey CT. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum Mol Genet* 1995;4:1-8.
- Martorell L, Manckton DG, Gomez J, Johnson KJ, Gich I, López de Munain A, Baiget M. Progression of somatic CTG repeats length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum Mol Genet* 1998;7:307-12.
- White PJ, Borts RH, Hirst MC. Stability of the human fragile X (CGG)_n triplet repeat array in *Saccharomyces cerevisiae* deficient in aspects of DNA metabolism. *Mol Cell Biol* 1999;19:5675-84.
- Richards RI, Sutherland GR. Dynamic mutation: A new class of mutations causing human disease. *Cell* 1992;70:709-12.
- Richards RI. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum Mol Genet* 2001;10:2187-94.

11. **Echler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeau SN, Richards CS, Ward PA, Nelson DL.** Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nature Genet* 1994;8:88–94.
12. **Khajavi M, Tori AM, Patel NB, Tsujii K, Sivak DR, Meistrich ML, Terry NH, Ashizawa T.** 'Mitotic drive' of expanded CTG repeats in myotonic dystrophy type 1 (DM1). *Hum Mol Genet* 2001;10:855–63.
13. **Cleary JD, Nichol K, Wang YH, Pearson CE.** Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nature Genet* 2002;31:37–46.
14. **Manley K, Shirley TL, Fishery L, Messer A.** Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nature Genet* 1999;23:471–3.
15. **Jaworski A, Rosche WA, Gellibolan R, Kang S, Shimizu M, Bowater RP, Sinden RR, Wells RD.** Mismatch repair in *Escherichia coli* enhances instability of (CTG)_n triplet repeats from human hereditary diseases. *Proc Natl Acad Sci USA* 1995;92:11019–23.
16. **Kang S, Jaworski A, Oshima K, Wells RD.** Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nature Genet* 1995;10:213–18.
17. **Kolodner RD, Tishkoff DX, Filosi N, Gaida GM.** A novel mutation avoidance mechanism dependent on *S. cerevisiae* RAD27 is distinct from DNA mismatch repair. *Cell* 1997;88:253–63.
18. **Jakupciak JP, Wells RD.** Genetic instabilities in (CTG/CAG) repeats occur by recombination. *J Biol Chem* 1999;274:23468–79.
19. **Moore H, Greenwell PW, Liu CP, Arnheim N, Petes TD.** Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc Natl Acad Sci USA* 1999;96:1504–9.
20. **Jankowski C, Nasr F, Nag DK.** Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc Natl Acad Sci USA* 2000;97:2134–9.
21. **Schweitzer JK, Livingston DM.** Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Hum Mol Genet* 1997;6:349–55.
22. **van den Broek WJAA, Nelen MR, Wasmink DG, Coenwinkel MM, te Riele H, Groenen PJ, Wieringa B.** Somatic expansion behaviour of the (CTG)_n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum Mol Genet* 2002;11:191–8.
23. **Niret JJ, Pessoa-Brandao L, Lahue RS.** Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1997;17:3382–7.
24. **Goellner GM, Tester D, Thibodeau S, Almquist E, Goldberg YP, Hayden MR, McMurray CT.** Different mechanisms underlie DNA instability in Huntington disease and colorectal cancer. *Am J Hum Genet* 1997;60:879–90.
25. **Kramer PR, Pearson CE, Sinden RR.** Stability of triplet repeats of myotonic dystrophy and fragile X loci in human mutator mismatch repair cell lines. *Hum Genet* 1996;98:151–7.
26. **Ireland MJ, Reinka SS, Livingston DM.** The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast. *Genetics* 2000;155:1657–65.
27. **Mirkin SM, Smirnova EV.** Positioned to expand. *Nature Genet* 2002;31:5–6.
28. **Gordenin DA, Kurikal TA, Resnick MA.** Repeat expansion—all in a flap? *Nature Genet* 1997;16:116–18.
29. **Freudenreich CH, Karrow SM, Zakian VA.** Expansion and length-dependent fragility of CTG repeats in yeast. *Science* 1998;279:853–6.
30. **Schweitzer JK, Livingston DM.** Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Hum Mol Genet* 1999;7:69–74.
31. **Sumalitis J, Ramirez MJ, Marcos R, Natarajan AT, Mullenders LHF.** Clusters of transcription coupled repair in the human genome. *Proc Natl Acad Sci USA* 2002;99:10571–4.
32. **Sumalitis J, Handels PH, Marcos R, Lanadarp P.** Accelerated telomere shortening in the human inactive X chromosome. *Am J Hum Genet* 1999;65:1616–22.
33. **Cobo A, Martinez JM, Martorell L, Baiget M, Johnson K.** Molecular diagnosis of homozygous myotonic dystrophy in two asymptomatic sisters. *Hum Mol Genet* 1993;2:711–15.
34. **Hoang J-M, Cotu PH, Thullie B, Salmon RJ, Thomas G, Hamelin R.** BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines. *Cancer Res* 1997;57:300–3.
35. **Parsons R, Guo-Min L, Longley M, Modrich P, Liu B, Berk T, Hamilton SR, Kinzler KW, Vogelstein B.** Mismatch repair deficiency in phenotypically normal human cells. *Science* 1995;268:738–40.
36. **Painado MA, Malkhosyan S, Velázquez A, Perucha M.** Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. *Proc Natl Acad Sci USA* 1992;89:10065–9.
37. **Yasuda J, Navarro JM, Malkhosyan S, Velázquez A, Arribas R, Sekiya T, Perucha M.** Chromosomal assignment of human DNA fingerprint sequences by simultaneous hybridization to arbitrarily primed PCR products from human/rodent monochromosome cell hybrids. *Genomics* 1996;34:1–8.
38. **López A, Xamena N, Marcos R, Velázquez A.** Germ cells microsatellite instability. The effect of different mutagens in a mismatch repair mutant of *Drosophila* [pel1]. *Mutat Res* 2002;514:87–94.

ARTÍCULO 2

Mutagenic stress modulates the dynamics of CTG
repeats instability associated with myotonic
dystrophy type 1

Nucleic Acids Res., 32:6733-6740, 2003

Mutagenic stress modulates the dynamics of CTG repeat instability associated with myotonic dystrophy type 1

Elisabeth Piñeiro, Laura Fernández-López, Josep Gamez¹, Ricard Marcos, Jordi Surrallés and Antonia Velázquez*

Grup de Mutagènesi, Unitat de Genètica, Departament de Genètica i de Microbiologia, Edifici Cn, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain and ¹Departamento de Neurología, Hospital de la Vall d'Hebron, Barcelona, Spain

Received September 23, 2003; Revised and Accepted October 13, 2003

ABSTRACT

The molecular basis of the myotonic dystrophy type 1 is the expansion of a CTG repeat at the *DMPK* locus. The expanded disease-associated repeats are unstable in both somatic and germ lines, with a high tendency towards expansion. The rate of expansion is directly related to the size of the pathogenic allele, increasing the size heterogeneity with age. It has also been suggested that additional factors, including as yet unidentified environmental factors, might affect the instability of the expanded CTG repeats to account for the observed CTG size dynamics over time. To investigate the effect of environmental factors in the CTG repeat instability, three lymphoblastoid cell lines were established from two myotonic dystrophy patients and one healthy individual, and parallel cultures were concurrently expanded in the presence or absence of the mutagenic chemical mitomycin C for a total of 12 population doublings. The new alleles arising along the passages were analysed by radioactive small pool PCR and sequencing gels. An expansion bias of the stepwise mutation was observed in a (CTG)₁₂₄ allele of a cell line harbouring two modal alleles of 28 and 124 CTG repeats. Interestingly, this expansion bias was clearly enhanced in the presence of mitomycin C. The effect of mitomycin C was also evident in the normal size alleles in two cell lines with alleles of 13/13 and 12/69 repeats, where treated cultures showed new longer alleles. In conclusion, our results indicate that mitomycin C modulates the dynamics of myotonic dystrophy-associated CTG repeats in LBCLs, enhancing the expansion bias

of long-pathogenic repeats and promoting the expansion of normal length repeats.

INTRODUCTION

Myotonic dystrophy type 1 (DM1) is an autosomal dominant neuromuscular disorder with a multisystemic phenotype. The molecular basis of DM1 is the expansion of an unstable CTG repeat in the 3'-untranslated region of the myotonic dystrophy protein kinase gene, on chromosome 19q13.3 (1). The size of the expanded allele directly correlates with the severity of the disease (2,3), and the number of CTG repeats progressively increases in successive generations of DM1 families, being the molecular explanation for anticipation (4).

Early Southern blot analysis of expanded CTG repeats revealed their somatic instability (5,6). Such instability was later demonstrated in both somatic and germ cells by using the small pool PCR (SP-PCR) technique (7). Somatic tissues of DM1 patients show mosaicism for CTG repeat size and, it has been found, a high tendency towards expansion throughout their lifetime. Thus the number of CTG repeats in peripheral blood leukocytes of young patients is rather stable whereas high levels of heterogeneity are observed in older patients (6,8). Germ cells from DM1 patients also show a high degree of variation in repeat size, including large contractions (9,10).

Analysis of tissue samples from DM1 subjects revealed that the major factor affecting the rate of expansion is the initial size of the expanded allele, with increased size heterogeneity over time (6,8). Although the model for the somatic instability proposed by Monckton *et al.* (7), based on the stepwise gain of a small number of repeats with age proceeding through expansion, agrees with these observations, a detailed analysis of the lifelong dynamics of somatic repeat size suggests that the expansion process is affected by other unknown transacting genetic factors (8). Additional *in vitro* studies using lymphoblastoid cell lines (LBCL) from DM1 patients have

*To whom correspondence should be addressed. Tel: +34 93 581 31 11; Fax: +34 93 581 23 87; Email: antonia.velazquez@uab.es
Correspondence may also be addressed to Jordi Surrallés. Email: jordi.surralles@uab.es

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

shown that two types of mutations arise in the pathogenic CTG repeats: (i) a frequent gain or loss of a small number of repeats, which is in concordance with the stepwise model described by Monckton *et al.* (7); and (ii) a less frequent large change of the repeat size (11), associated with a growth advantage of the cells bearing longer CTG repeats (12). This later phenomenon has been nominated mitotic drive (12).

The influence of environmental factors on the instability of CTG repeats associated with DM1 has not yet been properly addressed, although it has been hypothesized that such factors could modulate CTG repeat instability (8). Among these environmental factors, mutagenic compounds are good candidates since they can interfere with mechanisms known to be involved in trinucleotide repeat instability such as replication, recombination and DNA repair. In this study, we evaluate the effect of mutagenic stress on the dynamics of CTG repeat instability in mitotically dividing cells. To do so, we have established LBCLs from two DM1 patients and one normal individual to analyse the effect of mitomycin C (MMC) on the CTG repeat size dynamics of both normal length and long pathogenic alleles, along successive cell generations. The variant allele sizes in the heterogeneous population of transformed cells were resolved using the SP-PCR technique together with sequencing gel analysis. Using this procedure we have been able to study the dynamics of the CTG size and detect new alleles arising along the passages, in untreated and MMC-treated LBCL cultures.

MATERIALS AND METHODS

Lymphoblastoid cell lines and treatment

Peripheral blood samples were obtained from two symptomatic DM1 patients and one healthy individual by venipuncture, and lymphocytes were EBV-transformed into LBCLs. The cells were cultured as we previously reported (13,14). Transformed cells (2×10^6) were cultured in 75 cm² flasks at a concentration of 1×10^5 cells/ml and allowed to grow until the culture reached 2×10^7 cells (1×10^6 cells/ml). These expanded cultures were passed on as follows: 2×10^6 cells were transferred to new 75 cm² flasks and maintained until the culture reached 2×10^7 cells and then a new passage was performed. The remaining cells were split in two aliquots, one was kept frozen and the other was used for DNA extraction following a standard phenol-chloroform protocol. This procedure was repeated in each passage and the number of population doublings (PDs) in each passage was calculated.

MMC was diluted to a concentration of 1 mg/ml and stored at 4°C and treatments of the LBCL were performed by adding MMC to the medium to get a final concentration of 5 ng/ml. The mutagen treatment was continuous. The effect of MMC in cell proliferation and viability was evaluated by the Cell proliferation Kit II (XTT) (Boehringer Mannheim) and the fluorescein diacetate-mediated viability assay (15), respectively. We found that 5 ng/ml of MMC was not cytotoxic for any of the three LBCL but reduced the cell growth rate by 15% (data not shown). The effect of MMC on the culture growth dynamics is shown in the Supplementary Material available at NAR Online.

SP-PCR analysis

The concentration of the DNA extracted from the cell lines was measured spectrophotometrically and 5 µg of genomic DNA were digested with EcoRI. In order to estimate the true number of amplifiable genome equivalence (a.g.e.) in a volume of DNA sample, digested DNA was diluted in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 5 ng/µl carrier herring sperm DNA by serial 1:10 dilution steps to a final concentration of 5 pg/µl. Then the number of amplifiable molecules was determined by SP-PCR using the final dilution containing 5 pg/µl of digested genomic DNA. At this level of input DNA, some reactions will fail to contain an amplifiable copy of one or the other allele, and observing the number of reactions that amplified one allele or fail to amplify, the true number of a.g.e. in the input DNA was estimated (data not shown). Subsequently the CTG repeat size variations were analysed by performing 40 SP-PCRs using 200 a.g.e. of input DNA in each 25 µl reaction. The SP-PCR conditions previously described by Cobo *et al.* (16) were used, with some minor modifications. 36% of betaine (Sigma) was added to the reaction tube and the reactions were performed using the DM-93 and DM-103 primers described by Brook *et al.* (1). The amplified products were labelled by adding 3 µCi [α -³³P]dCTP to the SP-PCRs. Amplification conditions were 94°C 3 min for 1 cycle; 96°C 45 s, 68°C 45 s, 70°C 3 min for 28 cycles, with a final extension step at 68°C 1 min, 70°C 10 min, in a PT-100 Thermocycler (MJ Research, USA). Then, 7 µl of the SP-PCR products were resolved on 6% polyacrylamide sequencing gels and visualised by autoradiography. A 50 bp ladder (Amersham Biosciences) end-labelled with [γ -³³P]ATP was used as a molecular size marker. The length of the CTG repeats in the long alleles was estimated by densitometry using the Molecular Analyst 1.5 software (Bio-Rad).

Statistical analysis

Statistical comparisons between mutation frequencies were performed with the Fisher's exact test.

RESULTS

To investigate the effect of MMC on the CTG repeat instability of the DM1 locus, LBCLs were established from two DM1 patients and one normal individual. The sizes of the CTG repeats on these individuals were determined in peripheral blood samples by using standard radioactive PCR (Table 1). After Epstein-Barr virus (EBV) transformation, the lymphoblastoid cells were expanded until 1×10^6 cells/ml and the size of the CTG repeat was again analysed to establish the

Table 1. Size of the CTG repeat associated with DM1 in peripheral blood cells and their corresponding lymphoblastoid cell lines (LBCLs) of two DM1 patients and one normal individual

Individual	Age (years)	Genotype in blood ^a (CTG) _n	LBCL genotype at P ₀ ^a (CTG) _n
DM1-1	31	114/28	124/28
DM1-2	30	70/13	69/12
CM-1	27	13/13	13/13

^aFor the expanded allele, only the number of CTGs of the modal allele is indicated.

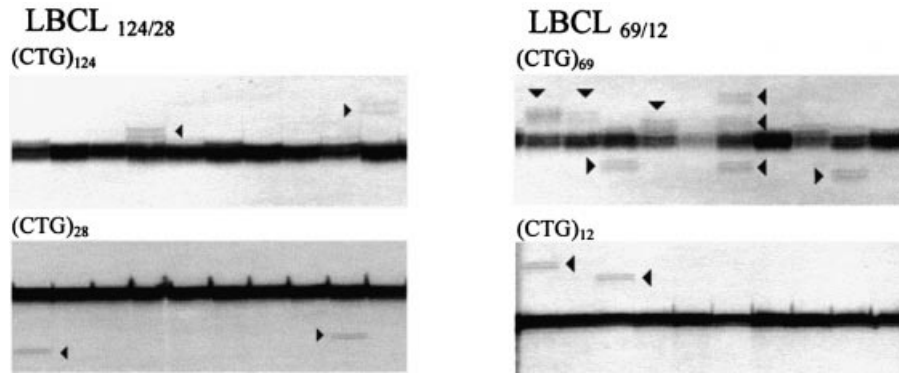


Figure 1. Representative examples of the SP-PCR analysis in the DM1 LBCLs using 200 a.g.e. of input DNA per reaction. Each line represents an SP-PCR. Variations of the normal and pathogenic alleles are indicated (arrowheads).

repeat length at the start of the experimental process. We have designated this stage as passage zero (P_0). Due to the heterogeneous size of the expanded CTG repeats in the peripheral blood cells of the two DM1 individuals, the LBCLs derived from their lymphocytes showed a modal pathogenic allele that differs from the modal size determined in the blood sample of the DM1 individuals. Normal size alleles, however, maintained their original size in the cell line (see Table 1). At P_0 , the cells were diluted to 1×10^5 cells/ml in medium with and without MMC and cultures were allowed to expand up to 1×10^6 cells/ml and then subpassed again. The same process was repeated four times. Following this approach, the molecular analysis of CTG instability was performed after 0, 6 and 12 PDs, both in the presence or absence of mutagenic treatment, by radioactive SP-PCR. To obtain the variant allele distribution in the cell population, at each point of the time-course study, we performed 40 identical SP-PCRs with an input of DNA of 200 a.g.e., which represents a total of 8000 analysed alleles in each PD. Under our SP-PCR conditions, and in the range of the allele size studied (12 to 124 CTGs), the amplification of the normal and the pathogenic alleles showed the same efficiency and, therefore, in the heterozygous LBCLs we considered a 1:1 ratio for normal and disease-causing alleles (4000 of each allele analysed at each point of study).

Mutations at the DM1-pathogenic CTG alleles

The SP-PCR approach, together with sequencing gel analysis, was applied to study the size-distribution of variant long alleles in two DM1 LBCLs, LBCL_{69/12} and LBCL_{124/28}. The cells were concurrently grown in the presence or absence of MMC. The SP-PCR is a technique that allows for the resolution of individual alleles by limiting the amount of input DNA in the PCR (7,17). However, the use of too little amplifiable DNA would limit the detection of infrequent mutant alleles, unless many PCRs are performed. On the other hand, due to the PCR stutter effect when large trinucleotide repeats are amplified (18), small variations in allele size would not be detected while larger amounts of input DNA are used. Therefore, in our experimental conditions where 200 a.g.e. of input DNA were used, small changes in the CTG repeat size of the two analysed large alleles (69 and 124 CTG repeats) were not resolved, whereas rare gross mutations were efficiently detected by performing 40 SP-PCRs (Fig. 1).

The SP-PCR analysis showed that the two LBCLs with long pathogenic alleles (69 and 124 CTG repeats), growing in standard conditions (without MMC), presented two types of spontaneous CTG repeat mutations: small changes in the number of CTG repeats and large size changes in the repeat number. The small size changes were manifested as a wide band containing a modal allele. The modal allele sizes and the upper and lower limits of these bands were determined as an indication of the allele size heterogeneity, being 69 ± 5 repeats for the (CTG)₆₉ allele and 124 ± 6 repeats for the (CTG)₁₂₄ allele at P_0 . This type of mutation was constant along the passages and was detected in all the SP-PCRs performed. It is unlikely that PCR stutter could account for this type of small change in repeat size, since this experimental artefact would have an effect on repeat size no larger than a three repeat shift, as recently reported (18). Nevertheless, whereas in the LBCL_{69/12} cell line the modal (CTG)₆₉ allele remains the same along the passages, in the LBCL_{124/28} cell line the modal (CTG)₁₂₄ allele at P_0 showed a gradual shift to slightly longer CTG tracks along the passages, being the new modal alleles 126 and 129 CTG tracks long after 6 and 12 PD, respectively (see Fig. 2). This shift was detected in all 40 SP-PCRs performed at each point of study, and it is clear evidence of a stepwise mutation mechanism with an expansion bias. In addition, new alleles with large size changes were also detected as discrete bands outside the wide major band (see Fig. 1). Based on the number of SP-PCRs, the relative frequency of such mutations with respect to the modal allele was determined, and their frequency distributions are represented in Figure 2. Large contractions and expansions were observed in both alleles and at all PD. Considering that 4000 expanded alleles per cell line were analysed at each point of study by using 200 a.g.e. of input DNA in each of the 40 SP-PCR performed, we have estimated the spontaneous frequency of large size changes at 0, 6 and 12 PD. As no differences were observed between PD, data were pooled and the overall spontaneous frequency of large mutations was 4.6×10^{-3} and 2.2×10^{-3} for the (CTG)₆₉ allele and for the (CTG)₁₂₄ allele, respectively.

MMC enhances the expansion bias of the (CTG)₁₂₄ allele

When the LBCL_{69/12} and the LBCL_{124/28} cultures were concurrently expanded in the presence of MMC, the same

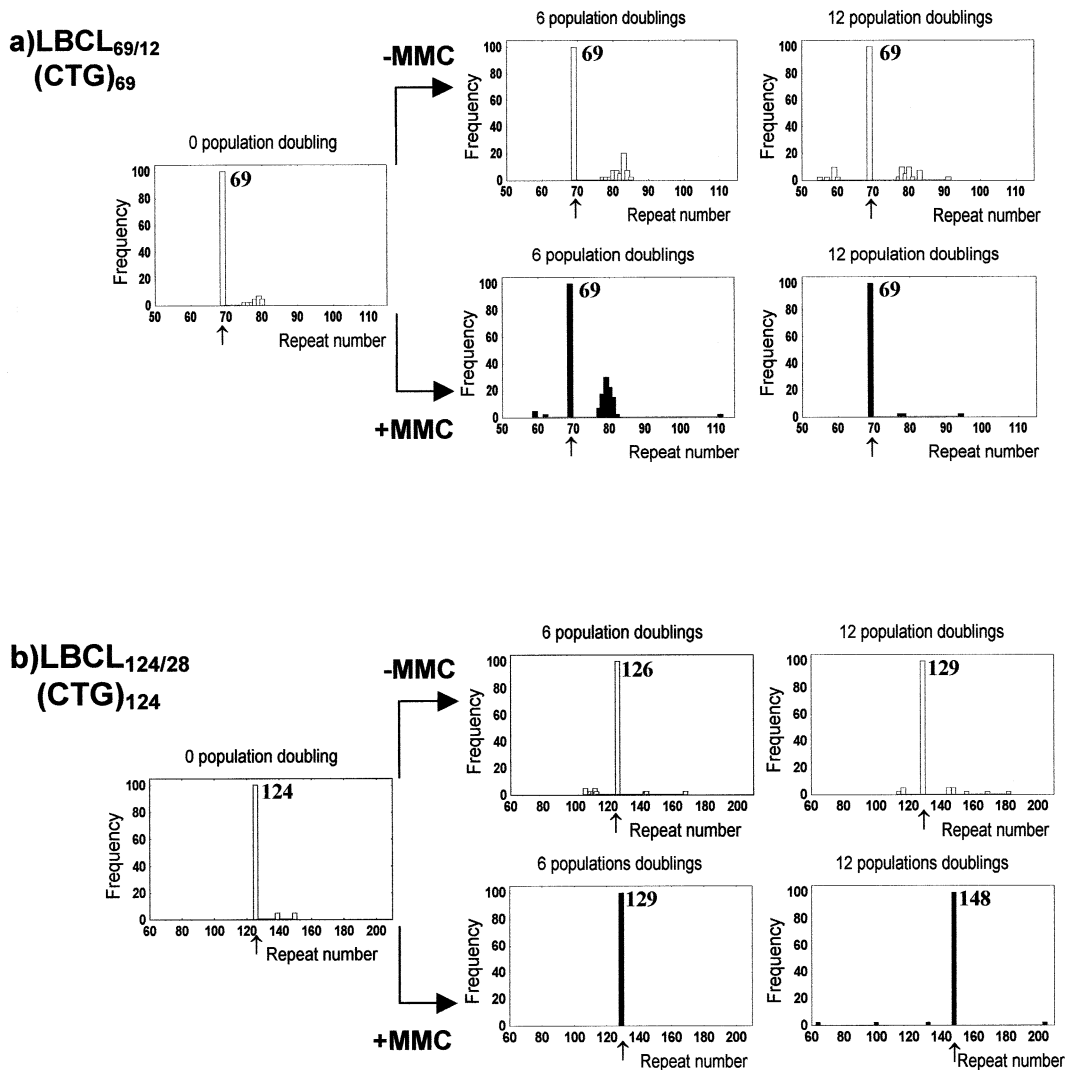


Figure 2. Distribution of the pathogenic DM1 alleles in the LBCL_{69/12} and LBCL_{124/28} cultures grown in the presence or absence of MMC up to 12 PD. Modal alleles and mutant alleles resulting from large contractions and expansions along the passages are shown. The frequency of mutant alleles is calculated considering the number of SP-PCRs that showed new bands with respect to modal allele bands. Modal alleles are indicated with an arrow. Number shows the modal CTG repeat size.

type of small changes observed in the long CTG repeats in the absence of MMC were found, i.e. along the passages a width SP-PCR band indicates the heterogeneity of the allele size around the modal allele. The heterogeneity for the (CTG)₆₉ and (CTG)₁₂₄ alleles was in the same range of the control experiment. As shown in Figures 2 and 3, the major effect of the MMC was an increase in the shift towards expansion of the (CTG)₁₂₄ modal allele along the passages, compared to the control culture. Thus, the modal allele at P₀ shifts from 124 to 129 and 148 CTG repeats after 6 and 12 PD, respectively. These results clearly suggest that MMC accelerated the expansion bias of the pathogenic allele in successive cell generations. In contrast, a statistically significant decrease in the number of large changes in the allele size was observed in the presence of MMC, with respect to control cultures. Thus, of 4000 alleles analysed after 12 PD, 22 and three alleles were large size variants of the (CTG)₆₉ allele in control and

MMC-treated cultures, respectively ($P < 0.001$). In the case of the (CTG)₁₂₄ allele, we found 10 and four of this type of variant allele in the absence or presence of MMC, respectively ($P < 0.01$). A feasible explanation for this fact is that a selection against those highly MMC-damaged cells is taking place along the passages, with a consequent elimination of the infrequent types of cells in the population.

Effect of MMC on the stability of normal length DM1 alleles

At the same time that the pathogenic alleles were analysed, the distribution of normal alleles in the DM1 LBCL_{69/12} and the LBCL_{124/28} cell lines was explored, together with the LBCL_{13/13} obtained from a healthy individual of the same age. Performing 40 SP-PCRs with an input of DNA of 200 a.g.e., we were able to detect rare alleles emerging along the passages, including those small variations in repeat number.

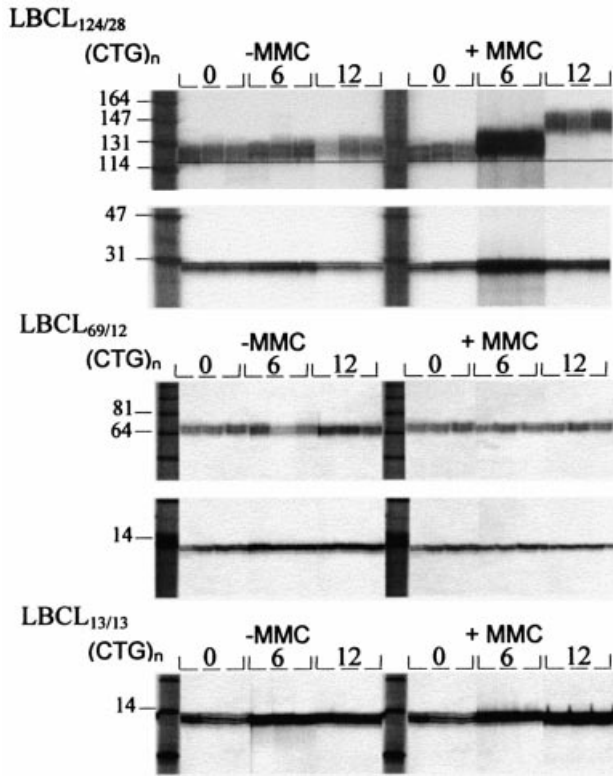


Figure 3. Evolution of the $(CTG)_{124}$ and $(CTG)_{28}$ allele sizes in the three lymphoblastoid cell lines along the culture passages in the presence or absence of MMC. The number of cell population doublings, 0, 6 and 12, is indicated. Each line represents an SP-PCR.

This was due to the fact that our sequencing gel conditions led to a higher resolution on the short CTG tracks range. The distribution of the normal alleles in the three cell lines expanded in the presence or absence of MMC is shown in Figure 4, including the relative frequencies of the rare alleles with respect to the modal normal alleles of 12, 13 and 28 CTG repeats. While both contractions and expansions of the $(CTG)_{12}$ and $(CTG)_{13}$ alleles were detected, for the $(CTG)_{28}$ allele only contractions were found. The same procedure used in the analysis of large changes in the pathogenic alleles was applied to estimate the frequency of rare alleles emerging from the normal DM1 alleles in the three cultures. Thus, at each experimental point, a total of 4000 normal alleles were analysed in the $LBCL_{69/12}$ and $LBCL_{124/28}$ cell lines and 8000 normal alleles in the $LBCL_{13/13}$ cell line. Because no differences between PD were observed, the estimated overall frequencies of spontaneous mutation (untreated cultures) of the normal alleles were 1.1×10^{-3} and 8.3×10^{-4} in the $LBCL_{69/12}$ and $LBCL_{124/28}$ cultures, respectively. These frequencies are significantly different and greater than the frequency of 2.9×10^{-4} , found in the $LBCL_{13/13}$ culture ($P < 0.01$ and $P < 0.05$, respectively).

Regarding the effect of MMC in the short alleles, we have to point out that the majority of alterations found were expansions into the normal range size. As shown in Figure 4a and b, a clear difference was observed in the length distributions of new alleles found in the MMC-treated $LBCL_{13/13}$ and

$LBCL_{69/12}$ cell lines; i.e. in these cultures, after 6 and 12 PDs, new long alleles to the normal size range were apparent. Most of these new alleles have a number of repeats that exceeds twice the number of repeats of the most represented allele (range 27–29 repeats). In addition, the frequency of new long alleles found in concurrent $LBCL_{13/13}$ cultures in the absence or presence of MMC, after 6 and 12 PD, is significantly higher (6.3×10^{-5} versus 7.5×10^{-4} , respectively; $P < 0.01$). In the $LBCL_{69/12}$ culture, a significant effect of MMC was also observed when we considered the frequency of expansion mutations in the normal size range (1.3×10^{-4} in control versus 1.0×10^{-3} in MMC-treated cells; $P < 0.05$). In contrast, only contractions were observed in the $(CTG)_{28}$ allele of the $LBCL_{124/28}$ cell line in all PD irrespective of MMC-treatment (see Fig. 4c).

DISCUSSION

Somatic CTG repeat instability associated with DM1 has a strong bias towards expansion that depends on the size of the repeat and the age of the affected individuals (6). In addition, previous studies on the repeat size dynamics hypothesized that environmental factors might affect the expansion process (8). In this study we have evaluated the effect of MMC on the instability of the CTG repeat of DM1 alleles of both normal and pathogenic length. The analysis of the CTG size dynamics was carried out in LBCLs from two DM1 patients and one normal individual. This analysis was performed along successive cell generations in a time-course study, using radioactive SP-PCR and sequencing gel analysis to resolve small variations in the CTG repeat size.

The instability of the expanded CTG alleles found in the untreated DM1 LBCLs cultures agrees with previously reported data on DM1 LBCLs (11,12). Thus, frequent small changes in the CTG repeat size (contractions or expansions of five to six repeats) were observed along the passages, in accordance with the stepwise mutation model proposed by Monckton *et al.* (7) and Khajavi *et al.* (12) to explain the *in vivo* and *in vitro* size heterogeneity of this repeat, respectively. Moreover, and as shown in Figure 3, the sensitive method of analysis used allowed us to detect a gradual increase of the $(CTG)_{124}$ allele size in successive cell generations, revealing the expansion bias of the stepwise mutation of this allele in short-term cultures. Such an effect was not observed in the $(CTG)_{69}$ allele. A possible explanation for the difference found between the two pathogenic alleles could be the extension time of the cultures, so that longer cultures would be necessary to detect a presumed expansion bias of the $(CTG)_{69}$ allele. Alternatively, a threshold in the number of CTG repeats might exist to display a bias towards expansion of the stepwise mutations occurring over time. Based on our results, this threshold would lie between 69–124 repeats, which includes the protomutation range (50–80 repeats), considered relatively stable *in vivo* somatic cells (19,20). In a similar study carried out in LBCLs (12) the authors reported expansion bias of alleles longer than 180 CTG repeats, which are not comparable to the 69 repeats allele of the present investigation. Martorell *et al.* (8) did not detect changes in the size of DM1 alleles of <200 repeats over time *in vivo*, although the resolution of their analysis procedure (Southern blot) would have limited the detection of small

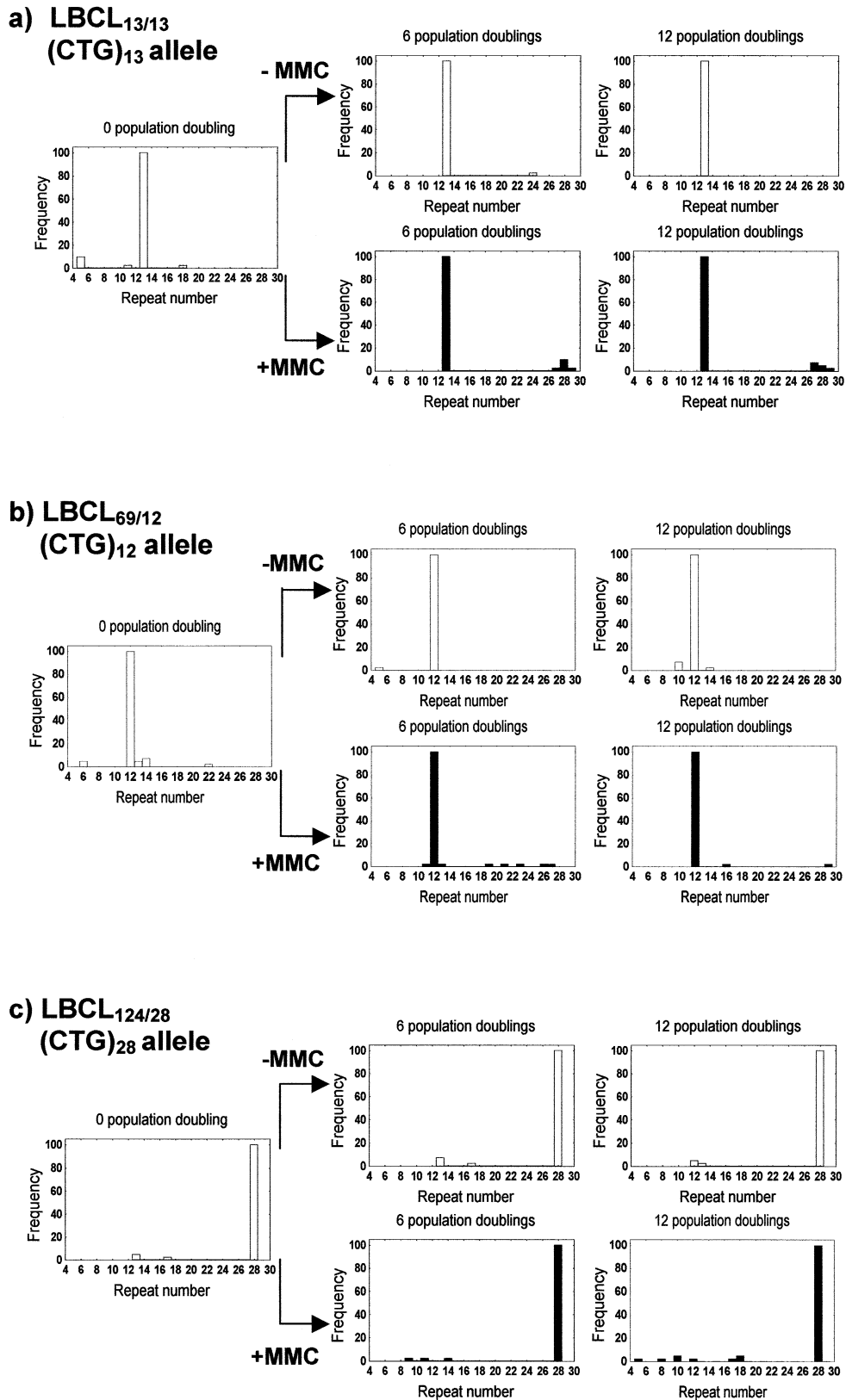


Figure 4. Distribution of the normal-length DM1 alleles in the LBCL_{13/13}, LBCL_{69/12} and LBCL_{124/28} cultures, grown in the presence or absence of MMC up to 12 PD. The frequency of mutant alleles is calculated considering the number of SP-PCRs that showed new bands with respect to modal allele band.

changes on repeat number. In addition, our data show no effect of MMC in the instability of the (CTG)₆₉ allele (see Fig. 2), supporting our hypothesis of the presence of a size threshold in the expansion bias of the stepwise mutations.

A less frequent type of change of the (CTG)₁₂₄ and the (CTG)₆₉ alleles, which corresponds to size changes exceeding five to six repeats, the modal allele size was also found along the passages, displaying discrete bands on the sequencing gels (see Fig. 1). Unlike previous studies that reported more contractions than expansions (11,12), our data show that this type of variant allele included a similar number of contractions and expansions.

The clear effect of MMC on the (CTG)₁₂₄ allele instability is intriguing. MMC increased the spontaneous expansion bias of this allele along the passages resulting in a shift of 24 CTG repeats in 12 PD (from 124 to 148 repeats). One may speculate that either the (CTG)₁₄₈ allele already existed at the starting point or that this mutation arose by chance in the culture. However, independently of MMC, and regardless of the possibility that mitotic drive could increase a possible likelihood of the cells with the (CTG)₁₄₈ allele to take over the rest of the cell population through the passages, the probability of this happening by chance would be very low, considering the short term cultures established. In addition, it should be considered that these rare cells would suffer a cell drift in successive cell generations and eventually disappear from the cell population, although a counter cell drift by mitotic drive is also expected. Therefore, this indicates that MMC plays a direct role in the observed repeat expansion. MMC produces DNA interstrand cross-links that are repaired by an error prone pathway at replication, consisting in the combination of nucleic excision repair and a lesion bypass process (21). The bypass process is responsible for the MMC S-phase dependent mutagenesis. Thus, new variant alleles could be induced by a replication-dependent stepwise mechanism in the presence of MMC. The MMC dose (5 ng/ml) used in our experiments had no effect on cell viability although it induced a 15% delay in the cell cycle and a consequent increase in the cell doubling time. As hypothesized before (11) this could be an important factor in DM1 repeat mutation. Accordingly, young patients with congenital DM1 do not manifest repeat heterogeneity regardless of the rapid cell divisions that take place in embryogenesis (22). Taking into account the above observations, we propose that MMC induces new variant alleles in culture cells and that the genetically heterogeneous cell population would be the substrate for the selection of those cells bearing larger alleles, ultimately displacing the original cell population in few passages. This model is supported by recent published data suggesting an association between high proliferating rate and large CTG repeats (12).

The spontaneous rates of instability of the non-pathogenic DM1 alleles in the LBCL_{69/12}, LBCL_{124/28} and LBCL_{13/13} cell lines were in the same order of magnitude that the overall rate of instability of microsatellite sequences reported in humans (23). However, the spontaneous mutation frequency of these normal alleles in the cell lines derived from DM1 patients was statistically higher than the mutation frequency of the allele in the cell line derived from a healthy individual. Although we do not have any feasible explanation for this finding, growth rate differences should be considered in interpreting our results.

The population doubling time of the DM1 cell lines LBCL_{69/12} and LBCL_{124/28} was 3 and 3.5 days, respectively, compared to 2 days for the LBCL_{13/13} cell line. Therefore the relative lower growth rate of the DM1 cell lines would have allowed additional time for spontaneous replication-independent mutations to occur. In addition, new alleles were observed in the presence of MMC, mainly with expansion of the CTG repeat, some of them reaching 29 repeats. The trinucleotide repeat instability induced by MMC is mainly independent of mismatch repair, as we have recently reported (24). Therefore these variant alleles could be generated by either recombination-dependent or -independent repair pathways involved in DNA interstrand cross-link repair (21,25). The MMC induced alleles lie within the long modal distribution of the normal size range and, therefore, could have significant importance in generating new premutational alleles to maintain the high incidence of DM1 in the population (15,26).

In conclusion, we have found that MMC modulates the dynamics of DM1-associated CTG repeats in mitotically dividing human cells, enhancing the expansion bias of long-pathogenic alleles and promoting the instability of normal length alleles. To our knowledge, this is the first description of an environmental factor affecting DM1-associated trinucleotide repeat instability. Although our *in vitro* results in somatic cells cannot be directly extrapolated to the *in vivo* situation, our findings suggest a novel mechanism of trinucleotide repeat instability with important clinical implications. Thus, mutagenic stress induced by environmental agents could accelerate the anticipation process and also increase the severity of the disease in symptomatic individuals due to an increased instability of the trinucleotide repeats in their somatic cells. It will be interesting to prove in mouse models whether mutagenic stress is involved in the anticipation process, by promoting the *in vivo* germ line expansion of trinucleotide repeats, and in the *in vivo* heterogeneity and mosaicism in the number of repeats.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

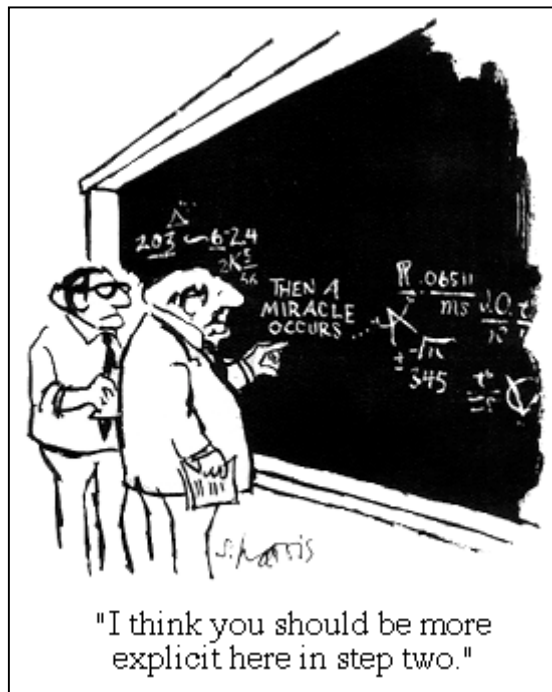
ACKNOWLEDGEMENTS

We would like to thank Dr L. Martorell (Sant Pau Hospital, Barcelona, Spain) for helpful comments and technical advice. This work was partially funded by La Fundació La Marató de TV3 (project 1999-85), the Generalitat de Catalunya (SGR-00197-2002), the Spanish Ministry of Health and Consumption (projects FIS PI020145, FIS PI020648 and FIS-Red G03/073), the Spanish Ministry of Science and Technology (projects PM 1999-0067, SAF2001-5138 and SAF2003-020328) and The European Union (project FIGH-CT-2002-00217). E.P. and L.F.-L. were supported by predoctoral fellowships awarded by the MCYT. J.S. is supported by a 'Ramón y Cajal' project entitled 'Genome stability and DNA repair' awarded by the MCYT and co-financed by the UAB.

REFERENCES

1. Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T. *et al.* (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*, **68**, 799–808.
2. Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., Harper, P.S., Housman, D.E. and Shaw, D.J. (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature*, **355**, 545–546.
3. Martorell, L., Martinez, J.M., Carey, N., Johnson, K. and Baiget, M. (1995) Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. *J. Med. Genet.*, **32**, 593–596.
4. Harley, H.G., Rundle, S.A., MacMillan, J.C., Myring, J., Brook, J.D., Crow, S., Reardon, W., Fenton, I., Shaw, D.J. and Harper, P.S. (1993) Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.*, **52**, 1164–1174.
5. Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., de Jong, P., Anvret, M., Riley, B. *et al.* (1992) Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature*, **355**, 547–548.
6. Wong, L.J., Ashizawa, T., Monckton, D.G., Caskey, C.T. and Richards, C.S. (1995) Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.*, **56**, 114–122.
7. Monckton, D.G., Wong, L.J., Ashizawa, T. and Caskey, C.T. (1995) Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.*, **4**, 1–8.
8. Martorell, L., Monckton, D.G., Gamez, J., Johnson, K.J., Gich, I., de Munain, A.L. and Baiget, M. (1998) Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum. Mol. Genet.*, **7**, 307–312.
9. Martorell, L., Monckton, D.G., Gamez, J. and Baiget, M. (2000) Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. *Eur. J. Hum. Genet.*, **8**, 423–430.
10. Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, C., Brunner, H. and Wieringa, B. (1994) Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)_n repeat variation and selection against extreme expansion in sperm. *Am. J. Hum. Genet.*, **54**, 575–585.
11. Ashizawa, T., Monckton, D.G., Vaishnav, S., Patel, B.J., Voskova, A. and Caskey, C.T. (1996) Instability of the expanded (CTG)_n repeats in the myotonin protein kinase gene in cultured lymphoblastoid cell lines from patients with myotonic dystrophy. *Genomics*, **36**, 47–53.
12. Khajavi, M., Tari, A.M., Patel, N.B., Tsuji, K., Siwak, D.R., Meistrich, M.L., Terry, N.H. and Ashizawa, T. (2001) Mitotic drive of expanded CTG repeats in myotonic dystrophy type 1 (DM1). *Hum. Mol. Genet.*, **10**, 855–863.
13. Callen, E., Samper, E., Ramirez, M.J., Creus, A., Marcos, R., Ortega, J.J., Olive, T., Badell, I., Blasco, M.A. and Surrallés, J. (2002) Breaks at telomeres and TRF2-independent end fusions in Fanconi anemia. *Hum. Mol. Genet.*, **11**, 439–444.
14. Surrallés, J., Ramirez, M.J., Marcos, R., Natarajan, A.T. and Mullenders, L.H. (2002) Clusters of transcription-coupled repair in the human genome. *Proc. Natl Acad. Sci. USA*, **99**, 10571–10574.
15. Straus, G. (1991) Non-random cell killing in cryopreservation: implications for performance of the battery of leukocyte tests (BLT), I. Toxic and immunotoxic effects. *Mutat. Res.*, **252**, 1–15.
16. Cobo, A., Martinez, J.M., Martorell, L., Baiget, M. and Johnson, K. (1993) Molecular diagnosis of homozygous myotonic dystrophy in two asymptomatic sisters. *Hum. Mol. Genet.*, **2**, 711–715.
17. Jeffreys, A.J., Tamaki, K., MacLeod, A., Monckton, D.G., Neil, D.L. and Armour, J.A. (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nature Genet.*, **6**, 136–145.
18. Zhang, Y., Monckton, D.G., Siciliano, M.J., Connor, T.H. and Meistrich, M.L. (2002) Age and insertion site dependence of repeat number instability of a human DM1 transgene in individual mouse sperm. *Hum. Mol. Genet.*, **11**, 791–798.
19. Barcelo, J.M., Mahadevan, M.S., Tsilfidis, C., MacKenzie, A.E. and Korneluk, R.G. (1993) Intergenerational stability of the myotonic dystrophy protomutation. *Hum. Mol. Genet.*, **2**, 705–709.
20. Martorell, L., Monckton, D.G., Sanchez, A., Lopez De Munain, A. and Baiget, M. (2001) Frequency and stability of the myotonic dystrophy type 1 premutation. *Neurology*, **56**, 328–335.
21. Zheng, H., Wang, X., Warren, A.J., Legerski, R.J., Nairn, R.S., Hamilton, J.W. and Li, L. (2003) Nucleotide excision repair- and polymerase eta-mediated error-prone removal of mitomycin C interstrand cross-links. *Mol. Cell. Biol.*, **23**, 754–761.
22. Mankodi, A. and Thornton, C.A. (2002) Myotonic syndromes. *Curr. Opin. Neurol.*, **15**, 545–552.
23. Weber, J.L. and Wong, C. (1993) Mutation of human short tandem repeats. *Hum. Mol. Genet.*, **2**, 1123–1128.
24. Fernández-López, L., Piñeiro, E., Marcos, R., Velásquez, A. and Surrallés, J. (2003) Induction of instability of normal length trinucleotide repeats within human disease genes. *J. Med. Genet.*, in press.
25. Cole, R.S., Levitan, D. and Sinden, R.R. (1976) Removal of psoralen interstrand cross-links from DNA of *Escherichia coli*: mechanism and genetic control. *J. Mol. Biol.*, **103**, 39–59.
26. Imbert, G., Kretz, C., Johnson, K. and Mandel, J.L. (1993) Origin of the expansion mutation in myotonic dystrophy. *Nature Genet.*, **4**, 72–76.

IV. DISCUSIÓN



IV. DISCUSIÓN

El incremento en el número de copias de repeticiones en ciertas secuencias microsátélites es la base molecular de una creciente lista de enfermedades genéticas humanas y, en concreto, la ERT es común en la mayoría de estos desórdenes, siendo objeto de particular interés para la comunidad científica, en los últimos años.

Las secuencias de TNR asociadas a desórdenes genéticos presentan un grado elevado de polimorfismo en la población normal, que siempre tiene un número de repeticiones bajo que se mantiene de generación en generación. A partir de un determinado número de repeticiones la secuencia se vuelve inestable, tanto a nivel somático como germinal, dando lugar a un mosaicismo somático y una variación intergeneracional, típico de estas enfermedades. Todos estos desórdenes genéticos, además, presentan el fenómeno de la anticipación, relacionado con la longitud de las expansiones, que consiste en un incremento de la severidad de los síntomas y una disminución de la edad de aparición de la enfermedad con el paso de las generaciones. Los estudios sobre la inestabilidad de los alelos no expandidos son muy escasos, ya que la mayoría de estudios se centran en los alelos patogénicos expandidos que presentan gran inestabilidad. No obstante, es importante estudiar la inestabilidad de las secuencias con un número normal de repeticiones, pues el carácter dominante de la enfermedad y el fenómeno de la anticipación implican que la transmisión genética de los alelos patogénicos se interrumpa, de modo que a partir de los alelos normales se tienen que generar nuevos alelos expandidos.

Se han propuesto varios factores que actuando en cis o en trans pueden jugar un papel modulador en la inestabilidad de las ERT asociadas a enfermedades humanas, como la DM1 y el FXS (Richards, 2001). Estos factores incluyen la pureza y tamaño de la secuencia repetida (Eichler *et al.*, 1994; Martorell *et al.*, 1998), la edad de los individuos (Monckton *et al.*, 1995; Martorell *et al.*, 1998), la deriva mitótica (Khajavi *et al.*, 2001), la posición relativa de la repetición respecto al origen de replicación (Cleary *et al.*, 2002) y mutaciones en genes involucrados en la reparación del DNA, en la replicación y en la recombinación (Manley *et al.*, 1999).

Uno de los procesos involucrados en el proceso de ERT podría ser el sistema MMR. Para estudiar el papel de la deficiencia en este sistema de reparación sobre la inestabilidad de las ERT, se han usado modelos de microorganismos y de ratón, con la limitación de que en estos modelos implica que los alelos estudiados se encuentran fuera de su contexto genómico natural. Además, los resultados

obtenidos en estos estudios son contradictorios, pues mientras que en *E.coli* y en levaduras la deficiencia en MMR lleva a la inestabilidad de las TNR (Jaworsky *et al.*, 1995; Schweitzer *et al.*, 1997), en ratón se requieren las proteínas MSH2 (Manley *et al.*, 1999) y MSH3 (van der Broek *et al.*, 2002) para la inestabilidad de estas repeticiones.

Dadas las contradicciones existentes sobre el efecto del sistema MMR sobre la ERT, decidimos estudiar la inducción de inestabilidad de TNR dependiente de MMR en alelos cortos y dentro de un contexto genómico natural. Con este objetivo, estudiamos el efecto de la MMC, un mutágeno químico inductor de enlaces cruzados, sobre la inestabilidad de las secuencias (CTG)_n y (CGG)_n endógenas, asociadas a la DM1 y al FXS, respectivamente, en distintas líneas celulares que contienen un número normal de repeticiones en un fondo genético deficiente en el sistema MMR (**Artículo 1**).

El análisis se llevó a cabo en cultivos celulares monoclonales expandidos bajo la presencia o ausencia de tratamiento mutagénico. Para la detección de la inestabilidad hemos usado una aproximación muy sensible que consiste en analizar la longitud de los alelos en geles de secuenciación, lo cual nos permite detectar pequeños cambios en el número de repeticiones (incluso de una sola repetición). En nuestro estudio hemos observado que los alelos no-expandidos no son tan estables como se creía y pueden ser inestabilizados en su contexto genómico natural por estrés mutagénico mediante dos mecanismos: un mecanismo en trans básicamente independiente del sistema MMR, y otro mecanismo menor dependiente de hMLH1.

Respecto a la inestabilidad espontánea, vemos que el comportamiento de los trinucleótidos asociados a la DM1 y al FXS es distinto que el del mononucleótido BAT-25, el cual se muestra inestable en las dos líneas celulares deficientes en el sistema MMR, HCT116 y LoVo (ver tabla 2 y figura 2, **Artículo 1**). Esto está de acuerdo con la bibliografía, ya que estas dos líneas celulares son derivadas de tumores y presentan una elevada tasa de inestabilidad de microsatélites (Boyer *et al.*, 1995). Actualmente se conoce que en algunos casos de cáncer hereditario, como el cáncer de colon hereditario no-poliposo (HNPCC), un defecto en enzimas del sistema MMR produce una elevada inestabilidad de microsatélites (Boyer *et al.*, 1995; Dietmaier *et al.*, 1997) y las líneas celulares derivadas de tumores deficientes en MMR presentan también este fenotipo mutador.

La mayor inestabilidad del mononucleótido BAT-25 observada en la línea LoVo respecto a la línea HCT116 está de acuerdo con estudios realizados en levadura donde las mutaciones en el gen *hMSH6* producen una mayor inestabilidad de mononucleótidos, por lo tanto, las líneas celulares derivadas de tumores deficientes en *hMSH6* presentaban también mayor inestabilidad de

mononucleótidos (Verma *et al.*, 1999). Dado que la línea LoVo tiene defectos tanto en el gen *hMSH2* como en el *hMSH6*, nuestros resultados muestran que la frecuencia de alteraciones en el mononucleótido BAT-25 en la línea celular LoVo es mayor que en la línea celular HCT116 que tiene mutado el gen *hMLH1*. Además, la elevada inestabilidad detectada en el mononucleótido BAT-25 en las líneas deficientes en MMR (LoVo y HCT116) en comparación con las líneas celulares normales para MMR (SW480 y HCT116+Chr.3), confirma que nuestro diseño experimental es adecuado y nos permitirá detectar y cuantificar la inestabilidad de de TNR dependiente de MMR.

Los trinucleótidos CTG y CGG analizados muestran niveles de inestabilidad muy bajos respecto al mononucleótido BAT-25. No obstante, se observa una tendencia no significativa hacia una mayor inestabilidad de estos trinucleótidos en las líneas deficientes en el sistema MMR (en especial el trinucleótido CGG en la línea LoVo) (ver tabla 2 y figura 2, **Artículo 1**). Se sabe que ciertos microsatélites son especialmente susceptibles a la inestabilidad. Así, en las células con MSI las repeticiones de mononucleótidos suelen ser las más inestables, seguidas de las repeticiones de dinucleótidos, siendo las repeticiones de tri-, tetra- y pentanucleótidos las que muestran la frecuencia de inestabilidad más baja (Dietmaier *et al.*, 1997). Esto explicaría la baja inestabilidad observada en los dos trinucleótidos analizados.

La baja inestabilidad espontánea de los dos trinucleótidos analizados tampoco es rara si tenemos en cuenta que la inestabilidad de estos dos microsatélites depende del número de repeticiones de la secuencia (Martorell *et al.*, 1998; Manley *et al.*, 1999). Según estos estudios, las secuencias con un número pequeño de repeticiones se mantienen estables. En nuestras líneas celulares, los dos trinucleótidos presentan un número de repeticiones dentro del intervalo normal (ver secuencias en Tabla 1, **Artículo 1**), lo que implica una tasa de mutación baja.

Kramer y colaboradores (1996) estudiaron la inestabilidad de las TNR CTG y CGG asociados a la DM1 y al FXS, respectivamente, en diversos tipos celulares, entre los que se incluían las líneas celulares HCT116 y LoVo. En este estudio, mutaciones en los genes *hMLH1* o *hMSH2/6* no mostraron efecto en la inestabilidad en estas secuencias, a los niveles observados en individuos con DM1 o FXS. Hay que destacar que el método usado en este estudio (Southern blot) no permite la detección de alelos mutantes raros resultantes de la expansión o contracción de pocas repeticiones y, por lo tanto, de este estudio no se puede descartar la posibilidad de que el sistema MMR pudiera estar implicado en la inestabilidad de los trinucleótidos CTG y CGG en el rango de longitud no patológica.

En nuestro estudio el hecho de detectar frecuencias mayores de clones inestables para TNR en la línea HCT116 comparada con la corregida HCT116+Chr.3 (ver tabla 2 y figura 2, **Artículo 1**), nos permite atribuir parte de esta inestabilidad espontánea a la deficiencia de hMLH1. A diferencia del estudio de Kramer y colaboradores (1996) nuestro análisis se llevó a cabo por un método altamente resolutivo que nos permite detectar frecuencias de mutación tan bajas como de 1 alelo mutante en 4000-20000 alelos pre-existentes. La resolución de la técnica se determinó a partir de estudios previos que nos permitieron establecer una sensibilidad de 1:100-1:500.

Como se ha indicado anteriormente, los resultados obtenidos después del tratamiento con MMC indican que existen dos mecanismos de inducción de inestabilidad de TNR: un mecanismo en trans independiente del sistema MMR y otro mecanismo secundario dependiente de la deficiencia en hMLH1.

Nuestros resultados, tras el tratamiento con MMC, muestran un incremento drástico (más de 18 veces) de la frecuencia de clones inestables para las TNR analizadas en la línea celular SW480 y, en menor medida, en la línea HCT116 (ver tabla 2 y figura 3, **Artículo 1**). En un principio, el hecho de que, al tratar con MMC, únicamente la línea celular SW480 presente esta inestabilidad tan elevada en los dos trinucleótidos analizados, podría indicar una inestabilidad dependiente de un sistema MMR funcional, lo que estaría de acuerdo con los estudios realizados en ratón que muestran que la inestabilidad de estas secuencias requiere las proteínas MSH2 (Manley *et al.*, 1999; Kovtun y McMurray, 2001) y MSH3 (van der Broek *et al.*, 2002) funcionales. Seguidamente analizamos la inducción de inestabilidad en los tres microsatélites estudiados en la línea HCT116 corregida para su deficiencia en MMR mediante la transferencia del cromosoma 3 (HCT116+Chr.3), donde no se observó ningún incremento de la inestabilidad (ver tabla 2 y figura 3, **Artículo 1**). Además, en la línea celular HCT116+Chr.3 los niveles de inestabilidad espontánea e inducida fueron menores que en la línea celular deficiente. Estos resultados nos llevan a concluir que la inducción de inestabilidad en los dos trinucleótidos analizados por la MMC observada en la línea celular SW480 no depende de un sistema MMR funcional, y que el responsable es un mecanismo independiente de este sistema de reparación.

Para profundizar más en el mecanismo responsable de la inestabilidad de trinucleótidos inducida por la MMC en la línea celular SW480, analizamos el perfil de inestabilidad de los tres loci microsatélites estudiados en las 4 líneas celulares utilizadas en el estudio. Sorprendentemente, encontramos una coincidencia en el perfil de inestabilidad de los dos trinucleótidos en la línea celular SW480, pero no con el mononucleótido BAT-25 (ver figura 4, **Artículo 1**). En esta línea celular, los

clones que presentaban inestabilidad inducida por la MMC en el locus *DMPK* también la presentaban en el locus *FMR1*. Esta coincidencia en el perfil de inestabilidad TNR en esta línea celular sugiere un mecanismo en trans de inestabilidad más que una acción directa de la MMC en la secuencia repetida. Además, la coincidencia en el perfil no estaría relacionada con el número de repeticiones, la secuencia repetida o la pureza de la misma, pues todas las líneas celulares analizadas tienen una secuencia pura de 10 a 12 repeticiones CTG en el locus de la DM1 y de 29 a 31 repeticiones CGG con dos interrupciones en el locus del FXS (ver tabla 1, **Artículo 1**).

Entender cual o cuales son los mecanismos en trans que puedan explicar nuestras observaciones es una cuestión importante para futuras investigaciones. Estos mecanismos en trans podrían ser activados/abolidos por el estrés mutagénico mediado por la MMC. Este mutágeno es dependiente de fase-S e induce enlaces cruzados en el DNA que bloquean la replicación. La relación entre replicación y ERT está bien documentada, especialmente en levaduras, donde mutaciones en los genes involucrados en la replicación incrementan la inestabilidad de las TNR (Ireland *et al.*, 2000; Mirkin y Smirnova, 2002). Yang y colaboradores (2003), en un estudio realizado en fibroblastos derivados de fetos DM1 utilizando inhibidores de la replicación, concluyeron que la perturbación de la dinámica de la horquilla de replicación podía tener algún papel en las expansiones CTG asociadas a la DM1.

Según la literatura un posible candidato para la inestabilidad de TNR es FEN-1, una proteína involucrada en la reparación por escisión de bases y en el procesamiento de los fragmentos de Okazaki durante la replicación de la cadena retardada (Gordenin *et al.*, 1997). Se ha visto que FEN-1 está implicada en la fidelidad de la replicación de las TNR y en la estabilidad cromosómica y genómica. Levaduras deficientes en Rad27 (homólogo de FEN-1) muestran incremento de la inestabilidad de las secuencias repetidas, sobretodo trinucleótidos (Freudenreich *et al.*, 1998; Schweitzer y Livingston, 1999). Tanto las células SW480 como las deficientes en Rad27 se caracterizan por inestabilidad genómica y cromosómica (Greene *et al.*, 1999; Melcher *et al.*, 2000), de modo que podría ser que la línea celular SW480 presentara una deficiencia en esta proteína, lo que explicaría la inestabilidad de trinucleótidos observada. Hay que tener presente no obstante, que si esta línea celular fuera deficiente para FEN-1, lo lógico sería observar también niveles elevados de inestabilidad espontánea, no solo bajo la inducción de la MMC. Nuestro análisis por western-blot con anticuerpos contra la FEN-1 en extractos celulares de las SW480 indica unos niveles de expresión normales y un tamaño molecular normal, indicando que FEN-1 no es responsable de la inestabilidad de trinucleótidos inducida por MMC observada en la línea celular SW480.

Sabemos que la línea celular SW480 es deficiente para la proteína p53 (Nigro *et al.*, 1989) y sobreexpresa una forma mutante de esta proteína. La expresión de p53 incrementa en respuesta al daño inducido por la MMC, y este incremento lleva a una disminución de la expresión de BRCA1 (Aritzi *et al.*, 2000; MacLachland *et al.*, 2000). La proteína BRCA1 está involucrada en la reparación de DSBs mediante recombinación homóloga, de modo que este efecto podría ser determinante para mantener la estabilidad de las TNR, así como del resto del genoma. Así, la proteína p53 es un buen candidato para explicar la inestabilidad de trinucleótidos inducida por MMC observada en la línea celular SW480.

Además de la inducción de inestabilidad de las TNR asociadas a la DM1 y al FXS en la línea celular SW480, también observamos un incremento de la inestabilidad de las repeticiones CTG y CGG (no coincidente) después del tratamiento con MMC en la línea HCT116 pero no en la corregida HCT116+Chr.3 (ver figura 3, **Artículo 1**), que sugiere un efecto en cis de la MMC en la inestabilidad de estas TNR en un fondo genético deficiente en hMLH1. Nuestros resultados también indican que la proteína hMLH1 juega un papel en la respuesta celular a la MMC. Esto está de acuerdo con estudios anteriores que sugieren que la deficiencia en el sistema MMR juega un papel menor en la inestabilidad de TNR asociadas a enfermedades en células humanas (Goellner *et al.*, 1997) o de levadura (Miret *et al.*, 1997).

Por lo tanto, nuestros resultados indican que la inestabilidad de TNR asociadas a enfermedades humanas es inducible. El hecho de que una secuencia corta de repeticiones pueda inestabilizarse por la acción de un mutágeno es importante, ya que estas secuencias se creían estables, aunque este efecto probablemente no está relacionado con la generación de alelos pre-patogénicos a partir de alelos de tamaño normal, pues todos los clones alterados por la inducción de la MMC presentan contracciones (ver figura 3, **Artículo 1**). Además, el proceso que regula la inestabilidad de repeticiones en células mitóticas somáticas puede ser diferente del que ocurre en la línea germinal in vivo.

Pero si la MMC es capaz de inestabilizar secuencias cortas de repeticiones teóricamente estables, se puede pensar que también tendrá efecto en las secuencias repetidas largas (pre)patogénicas, que son en sí mismas bastante inestables, afectando al proceso de expansión. Además, la inestabilidad de las TNR asociadas con desórdenes hereditarios presenta un fuerte sesgo hacia las expansiones que depende del número de repeticiones de la secuencia y de la edad de los individuos afectados (Wong *et al.*, 1995).

Así, quisimos evaluar el efecto de la MMC en la inestabilidad de las repeticiones CTG de los alelos DM1 de tamaño normal y patogénico (**Artículo 2**). El

análisis de la dinámica del tamaño de las repeticiones CTG se realizó en líneas celulares linfoblásticas (LBCLs) derivadas de dos pacientes DM1 y de un individuo normal. Este análisis se llevó a cabo a lo largo de sucesivas generaciones celulares (0-12 PD), usando SP-PCR radioactiva y análisis mediante geles de secuenciación para resolver pequeñas variaciones en el tamaño de las repeticiones CTG.

El análisis por SP-PCR nos permitió detectar dos tipos de mutación espontánea en las repeticiones CTG de los alelos patogénicos grandes ((CTG)₆₉ y (CTG)₁₂₄): (1) pequeños cambios en el número de repeticiones CTG y (2) grandes cambios en el número de repeticiones. Los cambios pequeños de tamaño se manifestaron como una banda ancha que contenía al alelo modal. El tamaño del alelo modal y los límites superior e inferior de estas bandas se determinaron como una indicación de la heterogeneidad del tamaño del alelo, siendo de 69±5 repeticiones para el alelo (CTG)₆₉ y de 124±6 repeticiones para el alelo (CTG)₁₂₄ al inicio del cultivo (punto P₀). Este tipo de mutación fue constante a lo largo de los pasajes y se detectó en todas las reacciones de SP-PCR realizadas. Esta inestabilidad de los alelos CTG expandidos observada en nuestros cultivos no tratados está de acuerdo con datos publicados previamente en LBCLs DM1 (Ashizawa, *et al.*, 1996; Khajavi *et al.*, 2001). Se ha propuesto un modelo de mutación paso a paso (*step-wise*) (Monckton *et al.*, 1995; Khajavi *et al.*, 2001) para explicar los cambios pequeños frecuentes en el tamaño de estas repeticiones *in vivo* e *in vitro*, respectivamente. Por lo contrario, es improbable que el tartamudeo de la PCR (*PCR stutter*) pueda explicar estos pequeños cambios en el número de repeticiones, pues como se ha publicado recientemente (Zhang *et al.*, 2002), este artefacto experimental solo explicaría los cambios en el tamaño de no más de tres repeticiones.

Sin embargo, mientras que en la línea celular LBCL_{69/12} el alelo modal (CTG)₆₉ permaneció estable a lo largo de los pasajes, en la línea celular LBCL_{124/28}, el alelo modal (CTG)₁₂₄ mostró un cambio gradual a secuencias ligeramente más largas a lo largo de los pasajes, siendo los nuevos alelos modales de 126 y 129 repeticiones CTG después de 6 y 12 duplicaciones de la población celular (PD), respectivamente (ver figura 2, **Artículo 2**). Este cambio se observó en las 40 reacciones de SP-PCR realizadas en cada punto del estudio, y es una clara evidencia de un mecanismo de mutación *step-wise* con un sesgo hacia las expansiones.

Como se muestra en la figura 3 (**Artículo 2**), la sensibilidad de nuestro método de análisis nos permitió detectar este incremento gradual del tamaño del alelo (CTG)₁₂₄ en las sucesivas generaciones celulares, demostrándose el sesgo hacia las expansiones de la mutación *step-wise* de este alelo en nuestros cultivos a

corto plazo. Este incremento en el tamaño del alelo, sin embargo, no se observó en el alelo (CTG)₆₉. El tiempo de extensión de los cultivos podría ser un factor importante para explicar las diferencias halladas entre los dos alelos patogénicos, pues tal vez serían necesarios cultivos más largos para poder detectar una tendencia hacia las expansiones de las mutaciones del alelo (CTG)₆₉. Otra explicación podría ser la diferencia de tamaño entre los dos alelos, es decir, que exista un umbral a partir del cual se manifieste el sesgo hacia las expansiones de las mutaciones *step-wise* que se den en el cultivo. Basándonos en nuestros resultados, este umbral podría estar entre 69-124 repeticiones, que incluye el intervalo de premutación (50-80 repeticiones), considerado relativamente estable en células somáticas *in vivo* (Barceló *et al.*, 1993; Martorell *et al.*, 2001). En un estudio similar realizado en LBCLs (Khajavi *et al.*, 2001), los autores mostraron una desviación hacia expansiones de alelos mayores de 180 repeticiones CTG, que no son comparables al alelo de 69 repeticiones de nuestro estudio. Martorell y colaboradores (Martorell *et al.*, 1998) no detectaron cambios en el tamaño de alelos de menos de 200 repeticiones a lo largo del tiempo *in vivo*, aunque la resolución de su método de análisis (Southern blot) podría haber limitado la detección de pequeños cambios en el número de repeticiones. Además, nuestros datos muestran que no hay un efecto de la MMC en la inestabilidad del alelo (CTG)₆₉ (ver figura 2, **Artículo 2**), lo que apoya nuestra hipótesis de la presencia de un umbral de tamaño en la desviación hacia expansiones de las mutaciones *step-wise*.

A parte de estos pequeños cambios en el número de repeticiones, a lo largo de los pasajes se detectó también otro tipo de cambio, menos frecuente, de los alelos (CTG)₁₂₄ y (CTG)₆₉. Estos cambios eran de más de 5-6 repeticiones respecto al tamaño del alelo modal, y en los geles de secuenciación se observaron como bandas discretas fuera de la banda ancha del alelo modal (ver figuras 1 y 2, **Artículo 2**). Nuestros datos muestran que este tipo de variantes alélicas incluyen un número similar de expansiones y contracciones.

A pesar de que se observa un claro efecto de la MMC sobre la inestabilidad del alelo (CTG)₁₂₄, no está claro como se da ese efecto. A lo largo de los pasajes, la MMC incrementó el sesgo hacia expansiones de este alelo, con un cambio de 24 repeticiones CTG en 12 PD (de 124 a 148 repeticiones), a diferencia del cambio gradual observado en los cultivos sin tratar, que fue de 124 a 126 y 129 repeticiones después de 6 y 12 PD, respectivamente (ver figura 2, **Artículo 2**). En los cultivos con MMC, se podría especular tanto que el alelo (CTG)₁₄₈ ya existía al inicio del cultivo, o que la mutación de 148 repeticiones se produjo por azar en el cultivo. Partiendo de cualquiera de las dos posibilidades, considerando nuestros cultivos a corto plazo, la probabilidad de que por azar las células con el alelo de 148

repeticiones reemplacen al resto de la población celular por la ventaja mitótica de las células que contienen este alelo (*mitotic drive*) sería muy baja. Además, se podría considerar que estas células raras podrían sufrir una deriva celular en las sucesivas generaciones celulares y, eventualmente, desaparecer de la población celular, a pesar de que se espera un aumento en el número de estas células si tienen *mitotic drive*. Por lo tanto, nuestros resultados indican que la MMC juega un papel directo en la expansión de repeticiones observada. La reparación de las DSBs producidas por el bloqueo de la replicación causado por la MMC, podría darse vía recombinación desigual entre los dos alelos, generándose así las nuevas variantes alélicas. La MMC produce enlaces cruzados intercatenarios (entre las dos cadenas de DNA) que pueden ser reparados por una vía de replicación tendiente a error, en la que intervienen la NER y un mecanismo de *bypass* de la lesión (Zheng *et al.*, 2003). El proceso de *bypass* de la lesión es responsable de la mutagénesis de la MMC dependiente de fase-S. Así, las nuevas variantes alélicas podrían ser inducidas mediante un mecanismo *step-wise* dependiente de replicación en la presencia de MMC. En el estudio realizado por Yang y colaboradores (2003) en fibroblastos derivados de fetos DM1, observaron un incremento de la inestabilidad de las repeticiones CTG al perturbar la dinámica de la horquilla de replicación con inhibidores de la replicación. Nuestros resultados parecen estar de acuerdo con los obtenidos independientemente por estos autores, pues la MMC es un mutágeno S-dependiente, que provoca el bloqueo de la horquilla de replicación. Estudios realizados en nuestro laboratorio con la bleomicina, un mutágeno S-independiente, no muestran inducción de las TNR asociadas a DM1 y al FXS en las mismas líneas celulares usadas en esta tesis, lo que refuerza la hipótesis de un mecanismo de inestabilidad dependiente de replicación.

Por otro lado, se conoce que las células tratadas con MMC reducen su tasa de crecimiento celular debido a la activación del punto de control de la división celular y, en un menor grado, a la activación del punto de control de la replicación (Lambert *et al.*, 2003). La activación de estos puntos de control extiende el tiempo para reparar el daño inducido en el DNA. La dosis de MMC (5 ng/mL) usada en nuestros experimentos no tiene un efecto en la viabilidad celular, aunque induce un retraso del 15% en la velocidad del ciclo celular y, como consecuencia, un incremento en el tiempo de duplicación de la población celular (ver anexo, **Artículo 2**). Tal como se había sugerido anteriormente (Ashizawa *et al.*, 1996) la reducción de la tasa de crecimiento celular podría ser un factor importante para la generación de mutaciones en las repeticiones DM1. De acuerdo con esto, los pacientes jóvenes con DM1 congénita no manifiestan heterogeneidad en el tamaño de las

repeticiones, a pesar de las rápidas divisiones que se dan en la embriogénesis (Mankodi *et al.*, 2002).

Tomando en consideración todas estas observaciones, proponemos que la MMC induce nuevas variantes alélicas en cultivos celulares y que la población celular genéticamente heterogénea podría ser el sustrato para la selección de aquellas células que llevan alelos mayores, desplazando finalmente a la población celular original en pocos pasajes. Este modelo está secundado por datos publicados recientemente sugiriendo una asociación entre una tasa elevada de proliferación celular y grandes repeticiones CTG (Khajavi *et al.*, 2001).

Las tasas de inestabilidad espontánea de los alelos DM1 no-patogénicos observadas en las líneas celulares LBCL_{69/12}, LBCL_{124/28} y LBCL_{13/13} fueron del mismo orden de magnitud que la tasa general de inestabilidad de secuencias microsatélite reportadas en humanos (Weber and Wong, 1993). No obstante, la frecuencia de mutación espontánea de estos alelos normales en las líneas celulares derivadas de pacientes DM1 fue estadísticamente mayor que la frecuencia de mutación del alelo en la línea celular derivada de un individuo sano. Aunque no tenemos ninguna explicación factible para este hallazgo, las diferencias en las tasas de crecimiento de las líneas celulares podrían explicar nuestros resultados. El tiempo de PD de las líneas celulares LBCL_{69/12} y LBCL_{124/28} fue de 3 y 3,5 días, respectivamente, comparados con los 2 días para la línea LBCL_{13/13}. Así, una tasa relativa de crecimiento menor en las líneas celulares DM1 podría proporcionar más tiempo para que se den mutaciones espontáneas independientes de la replicación. Además, se observaron nuevos alelos en la presencia de MMC, principalmente con expansiones de las repeticiones CTG, algunas de ellas alcanzando las 29 repeticiones. Como hemos visto en el estudio anterior (**Artículo 1**), la inestabilidad de las repeticiones CTG inducida por la MMC es principalmente independiente del sistema MMR. Por lo tanto, estas variantes alélicas podrían ser generadas durante la reparación de los enlaces cruzados intercatenarios inducidos por la MMC en la que intervienen procesos dependientes o independientes de recombinación (Zheng *et al.*, 2003; Cole *et al.*, 1976). Los nuevos alelos generados por el efecto de la MMC se encuentran dentro del intervalo de tamaño normal, pero cerca del límite superior, de modo que el estrés mutagénico podría tener una importancia significativa en la generación de nuevos alelos premutacionales para mantener la elevada incidencia de la DM1 en la población (Straus, 1991; Imbert *et al.*, 1993).

En conclusión, hemos encontrado que la MMC puede inducir inestabilidad de las TNR asociadas a la DM1 y al FXS por dos vías distintas: principalmente por un mecanismo en trans independiente de MMR y por un mecanismo menor dependiente de la deficiencia en hMLH1. También hemos encontrado que la MMC

modula la dinámica de las repeticiones CTG asociadas a la DM1 en células humanas en división mitótica, incrementando la desviación hacia expansiones de alelos patogénicos y promoviendo la inestabilidad de alelos de tamaño normal. Que nosotros sepamos, esta es la primera descripción de un factor ambiental afectando la inestabilidad de las TNR asociadas a la DM1 o al FXS. A pesar de que nuestros resultados *in vitro* en células somáticas no pueden extrapolarse directamente a la situación *in vivo*, nuestros hallazgos sugieren un mecanismo nuevo de inestabilidad de trinucleótidos con potenciales implicaciones clínicas. De esta forma, el estrés mutagénico inducido por agentes ambientales podría acelerar el proceso de anticipación y podría incrementar la severidad de la enfermedad en individuos sintomáticos debido a una mayor inestabilidad de las TNR en sus células somáticas. Sería interesante probar en modelos de ratón si el estrés mutagénico está involucrado en el proceso de anticipación promoviendo la expansión *in vivo* de las TNR en la línea germinal, y en la heterogeneidad y mosaicismo en el número de repeticiones *in vivo*.

V. CONCLUSIONES



V. CONCLUSIONES

De los resultados obtenidos en esta tesis doctoral se pueden extraer las siguientes conclusiones:

1. Las líneas celulares deficientes para el sistema MMR presentan una tendencia no significativa hacia una mayor inestabilidad de las TNR asociadas a la DM1 y al FXS.
2. La MMC induce inestabilidad de los alelos con un número normal de repeticiones CTG y CGG asociadas a la DM1 y al FXS, respectivamente, en su contexto genómico natural.
3. La inducción de la inestabilidad de las TNR asociadas a la DM1 y al FXS por la MMC se produce por dos vías distintas: principalmente por un mecanismo en trans independiente del sistema MMR y por un mecanismo menor dependiente de la deficiencia en hMLH1.
4. En LBCLs derivadas de pacientes DM1, las repeticiones CTG asociadas a la DM1 en alelos patogénicos muestran dos tipos de inestabilidad: grandes cambios en el número de repeticiones y cambios pequeños.
5. En las LBCLs derivadas de pacientes DM1, los cambios pequeños en el número de repeticiones muestran un sesgo hacia las expansiones a lo largo de las sucesivas generaciones celulares, generado por un mecanismo de mutación *step-wise*.
6. La MMC modula la dinámica de la inestabilidad *step-wise* de las repeticiones CTG asociadas a la DM1 en células humanas en división mitótica, incrementando la desviación hacia expansiones de los alelos patogénicos.
7. La MMC promueve la inestabilidad de los alelos con un número normal de repeticiones CTG asociadas a la DM1 en células humanas en división mitótica, generando alelos expandidos cerca del límite superior del intervalo de tamaño normal.

VI. BIBLIOGRAFÍA



VI. BIBLIOGRAFÍA

- Ahn B. and Grossman L. The binding of UvrAB proteins to bubble and loop regions in duplex DNA. *J. Biol. Chem.*, 271:21462-21470, 1996.
- Alwazzan M., Newman E., Hamshere M.G., Brook J.D. Myotonic dystrophy is associated with a reduced level of RNA from the DMWD allele adjacent to the expanded repeat. *Hum. Mol. Genet.*, 8:1491-1497, 1999.
- Arizti P., Fang L., Park I., Yin Y., Solomon E., Ouchi T., Aaronson S.A. and Lee S.W. Tumor suppressor p53 is required to modulate BRCA1 expression. *Mol. Cell Biol.*, 20:7450-7459, 2000.
- Ashizawa T., Dubel J.R. and Harati Y. Somatic instability of CTG repeat in myotonic dystrophy. *Neurology*, 43:2674-2678, 1993.
- Ashizawa T., Monckton D.G., Vaishnav S., Patel B.J., Voskova A. and Caskey C.T. Instability of the expanded (CTG)_n repeats in the myotonin protein kinase gene in cultured lymphoblastoid cell lines from patients with myotonic dystrophy. *Genomics*, 36:47-53, 1996.
- Balakumaran B.S., Freudenreich C.H. and Zakian V.A. CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in *Saccharomyces cerevisiae*. *Hum. Mol. Genet.*, 9:93-100, 2000.
- Bambara R.A., Murante R.S. and Henricksen L.A. Enzymes and reactions at the eukaryotic DNA replication fork. *J. Biol. Chem.*, 272:4647-4650, 1997.
- Barceló J.M., Mahadevan M.S., Tsilfidis C., MacKenzie A.E. and Korneluk R.G. Intergenerational stability of the myotonic dystrophy protomutation. *Hum. Mol. Genet.*, 2:705-709, 1993.
- Barnes D.E. Non-homologous end joining as a mechanism of DNA repair. *Curr. Biol.*, 11:455-457, 2001.
- Beckman J.S. and Weber J.L. Survey of human and rat microsatellites. *Genomics*, 12:627-631, 1992.
- Bowater R.P., Jaworski A., Larson J.E., Parniewski P. and Wells R.D. Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in *Escherichia coli*. *Nucleic Acids Res.*, 25:2861-2868, 1997.
- Boyer J.C., Umar A., Risinger J.I., Lipford J.R., Kane M., Yin S., Barrett J.C., Kolodner R.D. and Kunkel T.A. Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res.*, 55:6063-6070, 1995.
- Brook J.D., McCurrach M.E., Harley H.G., Buckler A.J., Church D., Aburatani H., Hunter K., Stanton V.P., Thirion J.P., Hudson T., Sohn R., Rundle S.A., Crow

- S., Davies J., Shelbourne P., Buxton J., Jones C., Juvonen V., Johnson K., Harper P.S., Shaw D.J. and Housman D.E. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*, 68:799-808, 1992.
- Buxton J., Shelbourne P., Davies J., Jones C., Van Tongeren T., Aslanidis C., de Jong P., Jansen G., Anvret M., Riley B., Williamson R. and Johnson K. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature*, 355:547-548, 1992.
- Cattaino G. and Vicario L. Myotonic dystrophy in Ancient Egypt. *Eur. Neurol.*, 41:59-63, 1999.
- Chakravarti A. Fragile X founder effect? *Nat. Genet.*, 1:237-238, 1992.
- Chang D.K., Metzgar D., Wills C. and Boland C.R. Microsatellites in the eukaryotic DNA mismatch repair genes as modulators of evolutionary mutation rate. *Genome Res.*, 11:1145-1146, 2001.
- Chen W. and Jinks-Robertson S. Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. *Mol. Cell Biol.*, 18:6525-6537, 1998.
- Cleary J.D. and Pearson C.E. The contribution of CIS-elements to disease-associated repeat instability: clinical and experimental evidence. *Cytogenet. Genome Res.*, 100:25-55, 2003.
- Cleary J.D., Nichol K., Wang Y.H. and Pearson C.E. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat. Genet.*, 31:37-46, 2002.
- Cohen H., Sears D.D., Zenvirth D., Hieter P. and Simchen G. Increased instability of human CTG repeat tracts on yeast artificial chromosomes during gametogenesis. *Mol. Cell Biol.*, 19:4153-4158, 1999.
- Cole R.S., Levitan D. and Sinden R.R. Removal of psoralen interstrand cross-links from DNA of *Escherichia coli*: mechanism and genetic control. *J. Mol. Biol.*, 103:39-59, 1976.
- Crawford D.C., Acuna J.M. and Sherman S.L. FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet. Med.*, 3:359-371, 2001.
- Crawford D.C., Wilson B. and Sherman S.L. Factors involved in the initial mutation of the fragile X CGG repeat as determined by sperm small pool PCR. *Hum. Mol. Genet.*, 9:2909-2918, 2000.
- Csank A.K. and Henikoff S. Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.*, 14:200-204, 1998.
- Dallas J.F. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mamm. Genome*, 3:452-456, 1992.

- De Silva I.U., McHugh P.J., Clingen P.H. and Hartley J.A. Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol. Cell Biol.*, 20:7980-7990, 2000.
- Debrauwere H., Buard J., Tessier J., Aubert D., Vergnaud G. and Nicolas A. Meiotic instability of human minisatellite CEB1 in yeast requires DNA double-strand breaks. *Nat. Genet.*, 23:367-371, 1999.
- Deka R., Majumder P.P., Shriver M.D., Stivers D.N., Zhong Y., Yu L.M., Barrantes R., Yin S.J., Miki T., Hundrieser J., Bunker C.H., McGarvey S.T., Sakallah S., Ferrell R.E. and Chakraborty R. Distribution and evolution of CTG repeats at the myotonin protein kinase gene in human populations. *Genome Res.*, 6:142-154, 1996.
- Dietmaier W., Wallinger S., Bocker T., Kullmann F., Fishel R. and Ruschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res.*, 57:4749-4756, 1997.
- Eichler E.E., Holden J.J., Popovich B.W., Reiss A.L., Snow K., Thibodeau S.N., Richards C.S., Ward P.A. and Nelson D.L. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat. Genet.*, 8:88-94, 1994.
- Eichler E.E., Macpherson J.N., Murray A., Jacobs P.A., Chakravarti A. and Nelson D.L. Haplotype and interspersed analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum. Mol. Genet.*, 5:319-330, 1996.
- Field D. and Wills C. Long, polymorphic microsatellites in simple organisms. *Proc. R. Soc. Lond. B. Biol. Sci.*, 263:209-215, 1996.
- Fortune M.T., Vassilopoulos C., Coolbaugh M.I., Siciliano M.J. and Monckton D.G. Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. *Hum. Mol. Genet.*, 9:439-445, 2000.
- Freudenreich C.H., Kantrow S.M. and Zakian V.A. Expansion and length-dependent fragility of CTG repeats in yeast. *Science*, 279:853-856, 1998.
- Freudenreich C.H., Stavenhagen J.B. and Zakian V.A. Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Mol. Cell Biol.*, 17:2090-2098, 1997.
- Fu Y.H., Kuhl D.P., Pizzuti A., Pieretti M., Sutcliffe J.S., Richards S., Verkerk A.J., Holden J.J., Fenwick R.G. Jr., Warren S.T., Oostra B.A., Nelson D.L. and Caskey C.T. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, 67:1047-1058, 1991.

- Fu Y.H., Pizzuti A., Fenwick R.G. Jr., King J., Rajnarayan S., Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P., Wieringa B., Korneluk R., Perryman M.B., Epstein H.F. and Caskey C.T. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science*, 255:1256-1258, 1992.
- Goellner G.M., Tester D., Thibodeau S., Almqvist E., Goldberg Y.P., Hayden M.R. and McMurray C.T. Different mechanisms underlie DNA instability in Huntington disease and colorectal cancer. *Am. J. Hum. Genet.*, 60:879-890, 1997.
- Goldstein D. and Schlötterer C. (editores). *Microsatellites. Evolution and applications*. Oxford University Press, 1^a ed., Oxford, 1999.
- Gomes-Pereira M., Fortune M.T. and Monckton D.G. Mouse tissue culture models of unstable triplet repeats: in vitro selection for larger alleles, mutational expansion bias and tissue specificity, but no association with cell division rates. *Hum. Mol. Genet.*, 10:845-854, 2001.
- Gordenin D.A., Kunkel T.A. and Resnick M.A. Repeat expansion - all in a flap? *Nat. Genet.*, 16:116-118, 1997.
- Greene A.L., Snipe J.R., Gordenin D.A. and Resnick M.A. Functional analysis of human FEN1 in *Saccharomyces cerevisiae* and its role in genome stability. *Hum. Mol. Genet.*, 8:2263-2273, 1999.
- Groenen P. and Wieringa B. Expanding complexity in myotonic dystrophy. *Bioessays*, 20:901-912, 1998.
- Haber J.E. Partners and pathways repairing a double-strand break. *Trends Genet.* 16:259-264, 2000.
- Hagerman R.J. and Cronister A. (editores). *Fragile X syndrome. Diagnosis, treatment and research*. The Johns Hopkins University Press, 2^a ed., Baltimore, 1996.
- Hagerman R.J. and Hagerman P.J. Fragile X syndrome: a model of gene-brain-behavior relationships. *Mol. Genet. Metab.*, 74:89-97, 2001.
- Hancock J.M. The contribution of slippage-like processes to genome evolution. *J. Mol. Evol.*, 41:1038-1047, 1995.
- Hancock J.M. Simple sequences in a "minimal" genome. *Nat. Genet.*, 14:14-15, 1996.
- Harley H.G., Brook J.D., Floyd J., Rundle S.A., Crow S., Walsh K.V., Thibault M.C., Harper P.S. and Shaw D.J. Detection of linkage disequilibrium between the myotonic dystrophy locus and a new polymorphic DNA marker. *Am. J. Hum. Genet.*, 49:68-75, 1991.

- Harley H.G., Brook J.D., Rundle S.A., Crow S., Reardon W., Buckler A.J., Harper P.S., Housman D.E. and Shaw D.J. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature*, 355:545-546, 1992.
- Harley H.G., Rundle S.A., MacMillan J.C., Myring J., Brook J.D., Crow S., Reardon W., Fenton I., Shaw D.J. and Harper P.S. Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.*, 52:1164-1174, 1993.
- Hashem V.I. and Sinden R.R. Chemotherapeutically induced deletion of expanded triplet repeats. *Mutat. Res.*, 508:107-119, 2002 *a*.
- Hashem V.I., Rosche W.A. and Sinden R.R. Genetic assays for measuring rates of (CAG).(CTG) repeat instability in *Escherichia coli*. *Mutat. Res.*, 502:25-37, 2002 *b*.
- Henderson S.T. and Petes T.D. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, 12:2749-2757, 1992.
- Henricksen L.A., Tom S., Liu Y. and Bambara R.A. Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *J. Biol. Chem.*, 275:16420-16427, 2000.
- Hewett D.R., Handt O., Hobson L., Mangelsdorf M., Eyre H.J., Baker E., Sutherland G.R., Schuffenhauer S., Mao J.I. and Richards R.I. FRA10B structure reveals common elements in repeat expansion and chromosomal fragile site genesis. *Mol. Cell*, 1:773-81, 1998.
- Hirst M.C. and White P.J. Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure. *Nucleic Acids Res.*, 26:2353-2358, 1998.
- Imbert G., Kretz C., Johnson K. and Mandel J.L. Origin of the expansion mutation in myotonic dystrophy. *Nat. Genet.*, 4:72-76, 1993.
- Ireland M.J., Reinke S.S. and Livingston D.M. The impact of lagging strand replication mutations on the stability of CAG repeat tracts in yeast. *Genetics*, 155:1657-1665, 2000.
- Jackson A.L., Chen R. and Loeb L.A. Induction of microsatellite instability by oxidative DNA damage. *Proc. Natl. Acad. Sci. USA*, 95:12468-12473, 1998.
- Jankowski C. and Nag D.K. Most meiotic CAG repeat tract-length alterations in yeast are SPO11 dependent. *Mol. Genet. Genomics*, 267:64-70, 2002.
- Jankowski C., Nasar F. and Nag D.K. Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc. Natl. Acad. Sci. USA*, 97:2134-2139, 2000.
- Jaworski A., Rosche W.A., Gellibolian R., Kang S., Shimizu M., Bowater R.P., Sinden R.R. and Wells R.D. Mismatch repair in *Escherichia coli* enhances instability

- of (CTG)_n triplet repeats from human hereditary diseases. *Proc. Natl. Acad. Sci. USA.*, 92:11019-11023, 1995.
- Jeffreys A.J., Murray J. and Neumann R. High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell*, 2:267-273, 1998.
- Jeffreys A.J., Bois P., Buard J., Collick A., Dubrova Y., Hollies C.R., May C.A., Murray J., Neil D.L., Neumann R., Stead J.D., Tamaki K. and Yardley J. Spontaneous and induced minisatellite instability. *Electrophoresis*, 18:1501-1511, 1997.
- Jeffreys A.J., Tamaki K., MacLeod A., Monckton D.G., Neil D.L. and Armour J.A. Complex gene conversion events in germline mutation at human minisatellites. *Nat. Genet.*, 6:136-145, 1994.
- Jinnai K., Sugio T., Mitani M., Hashimoto K. and Takahashi K. Elongation of (CTG)_n repeats in myotonic dystrophy protein kinase gene in tumors associated with myotonic dystrophy patients. *Muscle Nerve.*, 22:1271-1274, 1999.
- Jiricny J. Eukaryotic mismatch repair: an update. *Mutat. Res.*, 409:107-121, 1998.
- Johnson R.D. and Jasin M. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J.*, 19:3398-3340, 2000.
- Kang S., Jaworski A., Ohshima K. and Wells R.D. Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat. Genet.*, 10:213-218, 1995.
- Kang S., Ohshima K., Jaworski A. and Wells R.D. CTG triplet repeats from the myotonic dystrophy gene are expanded in *Escherichia coli* distal to the replication origin as a single large event. *J. Mol. Biol.*, 258:543-547, 1996.
- Kashi Y., King D. and Soller M. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.*, 13:74-78, 1997.
- Khajavi M., Tari A.M., Patel N.B., Tsuji K., Siwak D.R., Meistrich M.L., Terry N.H. and Ashizawa T. "Mitotic drive" of expanded CTG repeats in myotonic dystrophy type 1 (DM1). *Hum. Mol. Genet.*, 10:855-863, 2001.
- Khandjian E.W. Biology of the fragile X mental retardation protein, an RNA-binding protein. *Biochem. Cell Biol.*, 77:331-342, 1999.
- Kirkpatrick D.T. and Petes T.D. Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature*, 387:929-931, 1997.
- Klesert T.R., Otten A.D., Bird T.D. and Tapscott S.J. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat. Genet.*, 16:402-406, 1997.

- Kolodner R.D. and Marsischky G.T. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.*, 9:89-96, 1999.
- Kovtun I.V. and McMurray C.T. Trinucleotide expansion in haploid germ cells by gap repair. *Nat. Genet.*, 27:407-411, 2001.
- Kramer P.R., Pearson C.E. and Sinden R.R. Stability of triplet repeats of myotonic dystrophy and fragile X loci in human mutator mismatch repair cell lines. *Hum. Genet.*, 98:151-157, 1996.
- Kumar S., Lipman R. and Tomasz M. Recognition of specific DNA sequences by mitomycin C for alkylation. *Biochemistry*, 31:1399-1407, 1992.
- Lagercrantz U., Ellegren H. and Andersson L. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.*, 21:1111-1115, 1993.
- Lahue R.S. and Slater D.L. DNA repair and trinucleotide repeat instability. *Front. Biosci.*, 8:653-665, 2003.
- Lalioti M.D., Scott H.S., Buresi C., Rossier C., Bottani A., Morris M.A., Malafosse A. and Antonarakis S.E. Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature*, 386:847-851, 1997.
- Lambert S., Mason S.J., Barber L.J., Hartley J.A., Pearce J.A., Carr A.M. and McHugh P.J. *Schizosaccharomyces pombe* checkpoint response to DNA interstrand cross-links. *Mol. Cell Biol.*, 23:4728-4737, 2003.
- Larkin K. and Fardaei M. Myotonic dystrophy - a multigene disorder. *Brain. Res. Bull.*, 56:389-395, 2001.
- Lavedan C., Hofmann-Radvanyi H., Shelbourne P., Rabes J.P., Duros C., Savoy D., Dehaupas I., Luce S., Johnson K. and Junien C. Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am. J. Hum. Genet.*, 52:875-883, 1993.
- Lenzmeier B.A. and Freudenreich C.H. Trinucleotide repeat instability: a hairpin curve at the crossroads of replication, recombination, and repair. *Cytogenet. Genome Res.*, 100:7-24, 2003.
- Levinson G. and Gutman G.A. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Res.*, 15:5323-5338, 1987.
- Li Y-C., Korol A.B., Fahima T., Beiles A. and Nevo E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.*, 11:2453-2465, 2002.
- Lia A.S., Seznec H., Hofmann-Radvanyi H., Radvanyi F., Duros C., Saquet C., Blanche M., Junien C. and Gourdon G. Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but

- not correlated to the relative intertissue transcription levels and proliferative capacities. *Hum. Mol. Genet.*, 7:1285-1291, 1998.
- Lieber M.R. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays*, 19:233-240, 1997.
- Litt M. and Luty J.A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.*, 44:397-401, 1989.
- Longshore J.M. and Tarleton J. Dynamic mutations in human genes: a review of trinucleotide repeat diseases. *J. Genet.*, 75: 193-217, 1996.
- López A., Baida A., Marcos R., Xamena N. and Velazquez A. Spontaneous and bleomycin-induced genomic alterations in the progeny of *Drosophila* treated males depends on the Msh2 status. DNA fingerprinting analysis. *DNA Repair (Amst)*, 1:941-954, 2002.
- Lubs H.A. A marker X chromosome. *Am. J. Hum. Genet.* 21:231-244, 1969.
- MacLachlan T.K., Dash B.C., Dicker D.T. and El-Deiry W.S. Repression of BRCA1 through a feedback loop involving p53. *J. Biol. Chem.*, 275:31869-31875, 2000.
- Mahadevan M.S., Foitzik M.A., Surh L.C. and Korneluk R.G. Characterization and polymerase chain reaction (PCR) detection of an Alu deletion polymorphism in total linkage disequilibrium with myotonic dystrophy. *Genomics*, 15:446-448, 1993.
- Mankodi A. and Thornton C.A. Myotonic syndromes. *Curr. Opin. Neurol.*, 15:545-552, 2002.
- Manley K., Shirley T.L., Flaherty L. and Messer A. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat. Genet.*, 23:471-473, 1999.
- Marcadier J.L. and Pearson C.E. Fidelity of primate cell repair of a double-strand break within a (CTG).(CAG) tract. Effect of slipped DNA structures. *J. Biol. Chem.*, 278:33848-33856, 2003.
- Martorell L., Martinez J.M., Carey N., Johnson K. and Baiget M. Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. *J. Med. Genet.*, 32:593-596, 1995.
- Martorell L., Monckton D.G., Gamez J., Johnson K.J., Gich I., de Munain A.L. and Baiget M. Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum. Mol. Genet.*, 7:307-312, 1998.

- Martorell L., Monckton D.G., Sanchez A., Lopez De Munain A. and Baiget M. Frequency and stability of the myotonic dystrophy type 1 premutation. *Neurology*, 56:328-335, 2001.
- Mason C. An introduction to Myotonic Dystrophy, Signs and Symptoms. 2000. <http://www.nottingham.ac.uk/~pdzmggh/myotonicdystrophy/>
- Maurer D.J., O'Callaghan B.L. and Livingston D.M. Orientation dependence of trinucleotide CAG repeat instability in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 16:6617-6622, 1996.
- McMurray C.T. Mechanisms of DNA expansion. *Chromosoma*, 104:2-13, 1995.
- McMurray C.T. DNA secondary structure: a common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. USA*, 96:1823-1825, 1999.
- Melcher R., Steinlein C., Feichtinger W., Muller C.R., Menzel T., Luhrs H., Scheppach W. and Schmid M. Spectral karyotyping of the human colon cancer cell lines SW480 and SW620. *Cytogenet. Cell Genet.*, 88:145-152, 2000.
- Metzgar D., Bytof J. and Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.*, 10:72-80, 2000.
- Michel B., Ehrlich S.D. and Uzest M. DNA double-strand breaks caused by replication arrest. *EMBO J.*, 16:430-438, 1997.
- Miret J.J., Pessoa-Brandao L. and Lahue R.S. Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, 17:3382-3387, 1997.
- Miret J.J., Pessoa-Brandao L. and Lahue R.S. Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, 95:12438-12443, 1998.
- Mirkin S.M. and Smirnova E.V. Positioned to expand. *Nat. Genet.*, 31:5-6, 2002.
- Monckton D.G., Wong L.J., Ashizawa T. and Caskey C.T. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: smallpool PCR analyses. *Hum. Mol. Genet.*, 4:1-8, 1995.
- Morton N.E. and Macpherson J.N. Population genetics of the fragile-X syndrome: multiallelic model for the FMR1 locus. *Proc. Natl. Acad. Sci. USA*, 89:4215-4217, 1992.
- Nigro J.M., Baker S.J., Preisinger A.C., Jessup J.M., Hostetter R., Cleary K., Bigner S.H., Davidson N., Baylin S., Devilee P., Glover T., Collins F.S., Welson A., Modali R., Harris C.C. and Vogelstein B. Mutations in the p53 gene occur in diverse human tumour types. *Nature*, 342:705-708, 1989.

- Oberlé I., Rousseau F., Heitz D., Kretz C., Devys D., Hanauer A., Boue J., Bertheas M.F. and Mandel J.L. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science*, 252:1097-1102, 1991.
- Otten A.D. and Tapscott S.J. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. Natl. Acad. Sci. USA*, 92:5465-5469, 1995.
- Oudet C., Mornet E., Serre J.L., Thomas F., Lentes-Zengerling S., Kretz C., Deluchat C., Tejada I., Boue J., Boue A. and Mandel J.L. Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggests that fragile X chromosomes are derived from a small number of founder chromosomes. *J. Hum. Genet.*, 52:297-304, 1993.
- Palom Y., Suresh Kumar G., Tang L.Q., Paz M.M., Musser S.M., Rockwell S. and Tomasz M. Relative toxicities of DNA cross-links and monoadducts: new insights from studies of decarbamoyl mitomycin C and mitomycin C. *Chem. Res. Toxicol.*, 15:1398-1406, 2002.
- Panigrahi G.B., Cleary J.D. and Pearson C.E. In vitro (CTG)ⁿ(CAG) expansions and deletions by human cell extracts. *J. Biol. Chem.*, 277:13926-13934, 2002.
- Parniewski P. and Staczek P. Molecular mechanisms of TRS instability. *Adv. Exp. Med. Biol.*, 516:1-25, 2002.
- Parniewski P., Jaworski A., Wells R.D. and Bowater R.P. Length of CTG.CAG repeats determines the influence of mismatch repair on genetic instability. *J. Mol. Biol.*, 299:865-874, 2000.
- Pelletier R., Krasilnikova M.M., Samadashwily G.M., Lahue R. and Mirkin S.M. Replication and expansion of trinucleotide repeats in yeast. *Mol. Cell Biol.*, 23:1349-1357, 2003.
- Peltomaki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum. Mol. Genet.*, 10:735-740, 2001.
- Pritsos C.A. and Sartorelli A.C. Generation of reactive oxygen radicals through bioactivation of mitomycin antibiotics. *Cancer Res.*, 46:3528-3532, 1986.
- Ranum L.P. and Day J.W. Dominantly inherited, non-coding microsatellite expansion disorders. *Curr. Opin. Genet. Dev.*, 12:266-271, 2002 *a*.
- Ranum L.P. and Day J.W. Myotonic dystrophy: clinical and molecular parallels between myotonic dystrophy type 1 and type 2. *Curr. Neurol. Neurosci. Rep.*, 2:465-470, 2002 *b*.
- Reddy P.S. and Housman D.E. The complex pathology of trinucleotide repeats. *Curr. Opin. Cell Biol.*, 9:364-372, 1997.

- Richard G.F., Dujon B. and Haber J.E. Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol. Gen. Genet.*, 261:871-882, 1999.
- Richard G.F., Goellner G.M., McMurray C.T. and Haber J.E. Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. *EMBO J.*, 19:2381-2390, 2000.
- Richards R.I. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum. Mol. Genet.*, 10:2187-2194, 2001.
- Richards R.I. and Sutherland G.R. Dynamic mutations: a new class of mutations causing human disease. *Cell*, 70:709-712, 1992.
- Richards R.I. and Sutherland G.R. Dynamic mutation: possible mechanisms and significance in human disease. *Trends Biochem. Sci.*, 22:432-436, 1997.
- Richards R.I., Holman K., Friend K., Kremer E., Hillen D., Staples A., Brown W.T., Goonewardena P., Tarleton J., Schwartz C. and Sutherland G.R. Evidence of founder chromosomes in fragile X syndrome. *Nat. Genet.*, 1:257-260, 1992.
- Rink S.M., Lipman R., Alley S.C., Hopkins P.B. and Tomasz M. Bending of DNA by the mitomycin C-induced, GpG intrastrand cross-link. *Chem. Res. Toxicol.*, 9:382-389, 1996.
- Rittler M., Feld V. and Montagno M. Distrofia miotónica congénita. *Rev. Hosp. Mat. Inf. Ramón Sardá*, 1997, XVI, N°1 <http://www.sarda.org.ar/Revista/Sard%C3%A1/97a/34-40.pdf>.
- Samadashwily G.M., Raca G. and Mirkin S.M. Trinucleotide repeats affect DNA replication in vivo. *Nat. Genet.*, 17:298-304, 1997.
- Schmidt K.H., Abbott C.M. and Leach D.R. Two opposing effects of mismatch repair on CTG repeat instability in *Escherichia coli*. *Mol. Microbiol.*, 35:463-471, 2000.
- Schug M.D., Mackay T.F. and Aquadro C.F. Low mutation rates of microsatellite loci in *Drosophila melanogaster*. *Nat. Genet.*, 15:99-102, 1997.
- Schumacher S., Pinet I. and Bichara M. Modulation of transcription reveals a new mechanism of triplet repeat instability in *Escherichia coli*. *J. Mol. Biol.*, 307:39-49, 2001.
- Schweitzer J.K. and Livingston D.M. Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Hum. Mol. Genet.*, 6:349-355, 1997.
- Schweitzer J.K. and Livingston D.M. Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Hum. Mol. Genet.*, 7:69-74, 1998.

- Schweitzer J.K. and Livingston D.M. The effect of DNA replication mutations on CAG tract stability in yeast. *Genetics*, 152:953-963, 1999.
- Seznec H., Lia-Baldini A.S., Duros C., Fouquet C., Lacroix C., Hofmann-Radvanyi H., Junien C. and Gourdon G. Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. *Hum. Mol. Genet.*, 9:1185-1194, 2000.
- Shaw D.J. (editor). *Molecular genetics of human inherited disease*. Molecular medical Science series, John Wiley & Sons Ltd., Chichester, 1995.
- Sherman S.L., Morton N.E., Jacobs P.A. and Turner G. The marker (X) syndrome: a cytogenetic and genetic analysis. *Ann. Hum. Genet.*, 48:21-37, 1984.
- Sherman S.L., Jacobs P.A., Morton N.E., Froster-Iskenius U., Howard-Peebles P.N., Nielsen K.B., Partington M.W., Sutherland G.R., Turner G. and Watson M. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. *Hum. Genet.*, 69:289-299, 1985.
- Sia E.A., Jinks-Robertson S. and Petes T.D. Genetic control of microsatellite stability. *Mutat. Res.*, 383:61-70, 1997.
- Sinden R.R. Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am. J. Hum. Genet.*, 64:346-353, 1999.
- Sinden R.R., Potaman V.N., Oussatcheva E.A., Pearson C.E., Lyubchenko Y.L. and Shlyakhtenko L.S. Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J. Biosci.*, 27:53-65, 2002.
- Spiro C., Pelletier R., Rolfsmeier M.L., Dixon M.J., Lahue R.S., Gupta G., Park M.S., Chen X., Mariappan S.V. and McMurray C.T. Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol. Cell*, 4:1079-1085, 1999.
- Stallings R.L. CpG suppression in vertebrate genomes does not account for the rarity of (CpG)_n microsatellite repeats. *Genomics*, 13:890-891, 1992.
- Strand M., Prolla T.A., Liskay R.M. and Petes T.D. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, 365:274-276, 1993.
- Strauss G.H. Non-random cell killing in cryopreservation: implications for performance of the battery of leukocyte tests (BLT), I. Toxic and immunotoxic effects. *Mutat. Res.*, 252:1-15, 1991.
- Surrallés J., Ramírez M.J., Marcos R., Natarajan A.T. and Mullenders L.H. Clusters of transcription-coupled repair in the human genome. *Proc. Natl. Acad. Sci. USA*, 99:10571-10574, 2002.

- Sutherland G.R. Fragile sites on human chromosomes: demonstration of their dependence on the type of tissue culture medium. *Science*, 197:265-266, 1977.
- Tassone F., Hagerman R.J., Chamberlain W.D. and Hagerman P.J. Transcription of the FMR1 gene in individuals with fragile X syndrome. *Am. J. Med. Genet.*, 97:195-203, 2000 *a*.
- Tassone F., Hagerman R.J., Taylor A.K., Gane L.W., Godfrey T.E. and Hagerman P.J. Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. *Am. J. Hum. Genet.*, 66:6-15, 2000 *b*.
- Tassone F., Hagerman R.J., Taylor A.K., Mills J.B., Harris S.W., Gane L.W. and Hagerman P.J. Clinical involvement and protein expression in individuals with the FMR1 premutation. *Am. J. Med. Genet.*, 91:144-152, 2000 *c*.
- Thornton C.A., Griggs R.C. and Moxley R.T.3rd. Myotonic dystrophy with no trinucleotide repeat expansion. *Ann. Neurol.*, 35:269-272, 1994.
- Thornton C.A., Wymer J.P., Simmons Z., McClain C. and Moxley R.T.3rd. Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat. Genet.*, 16:407-409, 1997.
- Tishkoff D.X., Filosi N., Gaida G.M. and Kolodner R.D. A novel mutation avoidance mechanism dependent on *S. cerevisiae* RAD27 is distinct from DNA mismatch repair. *Cell*, 88:253-263, 1997.
- Tishkoff S.A., Goldman A., Calafell F., Speed W.C., Deinard A.S., Bonne-Tamir B., Kidd J.R., Pakstis A.J., Jenkins T. and Kidd K.K. A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. *Am. J. Hum. Genet.*, 62:1389-1402, 1998.
- Turner G., Webb T., Wake S. and Robinson H. Prevalence of fragile X syndrome. *Am. J. Med. Genet.*, 64:196-197, 1996.
- Umar A., Koi M., Risinger J.I., Glaab W.E., Tindall K.R., Kolodner R.D., Boland C.R., Barrett J.C. and Kunkel T.A. Correction of hypermutability, N-methyl-N'-nitro-N-nitrosoguanidine resistance, and defective DNA mismatch repair by introducing chromosome 2 into human tumor cells with mutations in MSH2 and MSH6. *Cancer Res.*, 57:3949-3955, 1997.
- Valdés A.M., Slatkin M. and Freimer N.B. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, 133:737-749, 1993.
- Valle G. TA-repeat microsatellites are closely associated with ARS consensus sequences in yeast chromosome III. *Yeast*, 9:753-759, 1993.

- van den Broek W.J., Nelen M.R., Wansink D.G., Coerwinkel M.M., te Riele H., Groenen P.J. and Wieringa B. Somatic expansion behaviour of the (CTG)_n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum. Mol. Genet.*, 11:191-198, 2002.
- Verkerk A.J., Pieretti M., Sutcliffe J.S., Fu Y.H., Kuhl D.P., Pizzuti A., Reiner O., Richards S., Victoria M.F., Zhang F.P., Eussen B.E., van Ommen G.J.B., Blonden L.A.J., Riggins G.J., Chastain J.L., Kunst C.B., Galjaard H., Caskey C.T., Nelson D.L., Oostra B.A. and Warren S.T. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, 65:905-914, 1991.
- Verma L., Kane M.F., Brassett C., Schmeits J., Evans D.G.R., Kolodner R.D. and Maher E.R. Mononucleotide microsatellite instability and germline *MSH6* mutation analysis in early onset colorectal cancer. *J. Med. Genet.*, 36:678-682, 1999.
- Virtaneva K., D'Amato E., Miao J., Koskiniemi M., Norio R., Avanzini G., Franceschetti S., Michelucci R., Tassinari C.A., Omer S., Pennacchio L.A., Myers R.M., Dieguez-Lucena J.L., Krahe R., de la Chapelle A. and Lehesjoki A.E. Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. *Nat. Genet.*, 15:393-396, 1997.
- Wansink D.G. and Wieringa B. Transgenic mouse models for myotonic dystrophy type 1 (DM1). *Cytogenet. Genome Res.*, 100:230-242, 2003.
- Weber J.L. and May P.E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.*, 44:388-396, 1989.
- Weber J.L. and Wong C. Mutation of human short tandem repeats. *Hum. Mol. Genet.*, 2:1123-1128, 1993.
- Wells R.D., Parniewski P., Pluciennik A., Bacolla A., Gellibolian R. and Jaworski A. Small slipped register genetic instabilities in *Escherichia coli* in triplet repeat sequences associated with hereditary neurological diseases. *J. Biol. Chem.*, 273:19532-19541, 1998.
- White P.J., Borts R.H. and Hirst M.C. Stability of the human fragile X (CGG)_n triplet repeat array in *Saccharomyces cerevisiae* deficient in aspects of DNA metabolism. *Mol. Cell Biol.*, 19:5675-5684, 1999.
- Wong L.J., Ashizawa T., Monckton D.G., Caskey C.T. and Richards C.S. Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.*, 56:114-122, 1995.
- Yamagata H., Miki T., Nakagawa M., Johnson K., Deka R. and Ogihara T. Association of CTG repeats and the 1-kb Alu insertion/deletion polymorphism

- at the myotonin protein kinase gene in the Japanese population suggests a common Eurasian origin of the myotonic dystrophy mutation. *Hum. Genet.*, 97:145-147, 1996.
- Yamagata H., Miki T., Ogihara T., Nakagawa M., Higuchi I., Osame M., Shelbourne P., Davies J. and Johnson K. Expansion of unstable DNA region in Japanese myotonic dystrophy patients. *Lancet*, 339:692, 1992.
- Yamasaki H. and Mironov N. Genomic instability in multistage carcinogenesis. *Toxicol. Lett.*, 15:251-256, 2000.
- Yang Z., Lau R., Marcadier J.L., Chitayat D. and Pearson C.E. Replication inhibitors modulate instability of an expanded trinucleotide repeat at the myotonic dystrophy type 1 disease locus in human cells. *Am. J. Hum. Genet.*, 73:1092-1105, 2003.
- Yu S., Mangelsdorf M., Hewett D., Hobson L., Baker E., Eyre H.J., Lapsys N., Le Paslier D., Doggett N.A., Sutherland G.R. and Richards R.I. Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. *Cell*, 88:367-374, 1997.
- Yu S., Pritchard M., Kremer E., Lynch M., Nancarrow J., Baker E., Holman K., Mulley J.C., Warren S.T., Schlessinger D., Sutherland G.R. and Richards R.I. Fragile X genotype characterized by an unstable region of DNA. *Science*, 252:1179-1181, 1991.
- Zerylnick C., Torroni A., Sherman S.L. and Warren S.T. Normal variation at the myotonic dystrophy locus in global human populations. *Am. J. Hum. Genet.*, 56:123-130, 1995.
- Zhang Y., Monckton D.G., Siciliano M.J., Connor T.H. and Meistrich M.L. Age and insertion site dependence of repeat number instability of a human DM1 transgene in individual mouse sperm. *Hum. Mol. Genet.*, 11:791-798, 2002.
- Zheng H., Wang X., Warren A.J., Legerski R.J., Nairn R.S., Hamilton J.W. and Li L. Nucleotide excision repair- and polymerase eta-mediated error-prone removal of mitomycin C interstrand cross-links. *Mol. Cell Biol.*, 23:754-761, 2003.
- Zienolddiny S., Ryberg D. and Haugen A. Induction of microsatellite mutations by oxidative agents in human lung cancer cell lines. *Carcinogenesis*, 21:1521-1526, 2000.